Research Article

Composite Fault Diagnosis of Rolling Bearing Based on Optimized Wavelet Packet AR Spectrum Energy Entropy Combined with Adaptive No Velocity Term PSO-SOM-BPNN

Hongwei Fan [1,2] **Yang Yan,**[1] **Xuhui Zhang,**[1,2] **Xiangang Cao,**[1,2] **and Jiateng Ma**[1]

[1]*School of Mechanical Engineering, Xi'an University of Science and Technology, Xi'an 710054, China*
[2]*Shaanxi Key Laboratory of Mine Electromechanical Equipment Intelligent Monitoring, Xi'an 710054, China*

Correspondence should be addressed to Hongwei Fan; hw_fan@xust.edu.cn

Aiming at the problem of low diagnosis efficiency and accuracy, due to noise and cross aliasing among various faults when diagnosing composite faults of rolling bearing under actual working conditions, a composite fault diagnosis method of rolling bearing based on optimized wavelet packet autoregressive (AR) spectral energy entropy and adaptive no velocity term particle swarm optimization-self organizing map-back propagation neural network (ANVTPSO-SOM-BPNN) is proposed. The energy entropy feature is extracted from the bearing vibration signal through wavelet packet AR spectrum, and SOM and BPNN are combined to form a series network. For PSO, the velocity term is discarded and the inertia weight and learning factor are adaptively adjusted. Finally, the Dempster-Shafer (D-S) evidence fusion diagnosis is carried out. To get closer to the application condition, the data are collected near and far away from the fault point for the composite fault diagnosis, which verifies the effectiveness of the proposed method.

1. Introduction

Rolling bearing is one of the most important components in rotating machinery. It plays an important role in supporting rotating shaft and reducing friction. Its working state is of great significance to the normal operation of the whole rotating machine [1–4]. In practical engineering, the fault often does not appear alone, and the probability of composite fault of the same bearing is also large. Composite faults are two or more faults that are interrelated and cross influenced at the same time [5]. And the vibration signals caused by different faults will interfere with each other and produce the coupling phenomenon, which makes the signal more complex and difficult to accurately diagnose faults. Therefore, the fault diagnosis of rolling bearing has the important practical value [6].

For the fault diagnosis of bearing, the feature extraction, neural network, and multi-information fusion are the research focuses. For the feature extraction, the wavelet packet decom-

position is a typical processing method of unsteady signal, which can process the signal more finely [7]. Tang and Deng [8] proposed a composite bearing fault feature separation method based on the improved harmonic wavelet packet decomposition to decompose the signal of intermediate frequency part and extract more effective signals. He et al. [9] applied the adaptive redundant multiwavelet packet to composite fault diagnosis of rotating machinery, proposed the normalized multifractal entropy as the evaluation criterion, adaptively constructed multiwavelet, and determined the fault sensitive frequency band by the relative energy ratio of characteristic frequency. Ma et al. [10] decomposed the composite fault signal using multiwavelet packet, reconstructed the signal with permutation entropy as the evaluation index, and finally demodulated and extracted the fault features using energy operator. For the fault diagnosis of rolling bearing, Abbasion et al. [11] preprocessed the vibration signal through wavelet analysis and then used support vector machine (SVM) to

diagnose the faults. Janssens et al. [12] applied the convolutional neural network (CNN) to multi fault diagnosis. Lv and Yao [13] used wavelet packet decomposition combined with back propagation neural network (BPNN) to diagnose the faults. Among them, BPNN is widely used for the fault diagnosis of rolling bearing due to its strong nonlinear mapping ability and high self-learning and adaptive ability [14]. However, the standard BPNN is easy to fall into the local optimal solution and relies too much on samples. According to the defects of BPNN, Huang et al. [15] used the global search ability of genetic algorithm to optimize BPNN. Gong et al. [16] combined the self-organizing map (SOM) with BPNN to obtain the better classification results and improve the convergence speed. Ju et al. [17] optimized the weight and threshold of BPNN through particle swarm optimization (PSO) and extracted the feature energy through wavelet packet, which improved the diagnosis efficiency and accuracy. The standard PSO also has its disadvantages, such as low convergence accuracy and easy to fall into local extremum. Wang and Wang [18] introduced the decline index and iteration threshold to improve the linear decline weight of the standard PSO and verified the advantages of improved PSO in search accuracy, convergence speed, and stability. Zhu and Xue [19] adaptively modified the learning factor to better balance the local and global search ability in view of the problem that the fixed value of learning factor in PSO affects the algorithm performance. Aiming at the signal fuzziness and uncertainty of composite fault, the diagnosis result is further improved by information fusion. Khazaee et al. [20] fused vibration and sound signals through Dempster-Shafer (D-S) evidence theory for fault diagnosis of gearbox and achieved ideal results. Feng and Pereira[21] applied the wavelet neural network and evidence theory to fault diagnosis of rotating machine and verified the effectiveness.

This paper proposes a new diagnosis method based on optimized wavelet packet AR spectral energy entropy to adaptive no velocity term PSO-SOM-BPNN (ANVTPSO-SOM-BPNN). In order to be closer to the real working conditions on site, data are collected near the fault points and far away from the fault points, respectively. The energy entropy characteristics of bearing vibration signals are extracted through wavelet packet AR spectrum. The basis function and decomposition layers of wavelet packet decomposition are optimally selected. SOM and BPNN are combined to form a series network, and PSO discards the velocity term and adaptively adjusts the inertia weight and learning factor. Finally, the proposed method is used to fuse the diagnosis results at two measuring points at D-S evidence decision level to improve the efficiency and accuracy in the composite fault diagnosis of rolling bearing.

2. Methodology

During the operation of rolling bearing, due to the interaction of inner ring, outer ring, and rolling element, it is easy to form overlapping composite faults. Among them, the fault features with weak energy may be submerged by the features with other strong energy or noise, which affects the accuracy of fault diagnosis. Therefore, at first, vibration accelerators are installed at two different measuring points to collect the vibration signals.

Secondly, two kinds of collected signal are preprocessed and then extract the signal features. And two kinds of extracted fault feature are diagnosed in the new method, to obtain two kinds of basic probability distribution. Finally, two probability distributions are fused by D-S evidence theory to achieve the purpose of fault diagnosis using multi-information fusion. The overall research idea of composite fault diagnosis of rolling bearing is shown in Figure 1.

2.1. Wavelet Packet AR Spectral Entropy Feature Extraction Method

2.1.1. Principle of Wavelet Packet Decomposition. Wavelet packet decomposition overcomes the defect that wavelet analysis only decomposes the low-frequency part of signal. It also decomposes the high-frequency part and improves time-frequency resolution. The specific algorithm is as follows.

Given scaling function $\phi(t)$ and wavelet basis function $\psi(t)$, two-scale equations are satisfied between them:

$$\begin{cases} \phi(t) = \sqrt{2} \sum_{k \in Z} h_k \phi(2t - k), \\ \psi(t) = \sqrt{2} \sum_{k \in Z} g_k \phi(2t - k), \end{cases} \tag{1}$$

where k is the time translation factor; h_k is the low-pass filter coefficient; and g_k is the high-pass filter coefficient.

The wavelet packet decomposition algorithm is

$$\begin{cases} d_l^{j,2n} = \sum_k a_{k-2l} d_k^{j+1,n}, \\ d_l^{j,2n+1} = \sum_k b_{k-2l} d_k^{j+1,n}, \end{cases} \tag{2}$$

where j is the number of wavelet packet decomposition layers; $d_l^{j,2n}$ is the low-frequency coefficient decomposed by layer j; and $d_l^{j,2n+1}$ is the high-frequency coefficient decomposed by layer j.

The wavelet packet reconstruction algorithm is

$$d_l^{j+1,n} = \sum_k \left(h_{l-2k} d_k^{j,2n} + g_{l-2k} d_k^{j,2n+1} \right), \tag{3}$$

where h_{l-2k} is the low-frequency coefficient reconstructed by wavelet packet and g_{l-2k} is the high-frequency coefficient reconstructed by wavelet packet.

2.1.2. AR Spectrum Estimation. Due to the complexity of composite fault signal of rolling bearing, it is difficult to obtain the accurate fault characteristics only by wavelet packet decomposition. Therefore, it needs to be further processed on the basis of wavelet packet decomposition. The basic idea of AR spectrum estimation is to establish an AR model for time series signal and then calculate the self-power spectrum of signal with model coefficients [22].
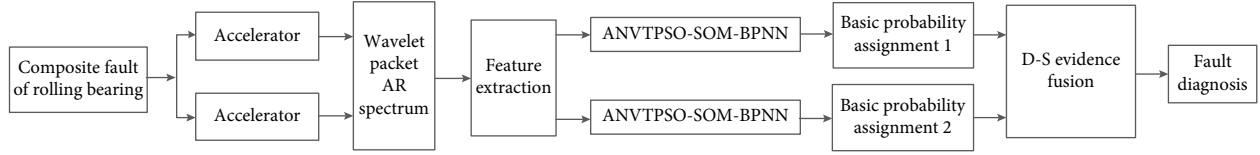
FIGURE 1: General research idea of rolling bearing composite fault diagnosis.

The general expression of AR model is

$$g(x) = s(x) - \sum_{i=1}^{R} a_i g(x - i), \tag{4}$$

where $g(x)$ is autoregressive time series; $s(x)$ is finite bandwidth white noise with normal distribution with mean value of 0 and variance of σ_s^2; a_i is regression coefficient; and R is model order.

If equation (4) is regarded as the input/output equation of a system, $s(x)$ can be regarded as the white noise input of the system, and $g(x)$ is the response output of the system under the excitation of limited bandwidth white noise.

According to the definition of self-power spectrum and transfer function, the unilateral spectrum of signal can be expressed by the following formula:

$$G_y(f) = \frac{2T_s \sigma_s^2}{\left| 1 + \sum_{i=1}^{R} a_i e^{-i2\pi k T_s} \right|^2}, \tag{5}$$

where $f \in [0, f_s/2.56]$; $T_s = 1/f_s$; and f_s is the sampling frequency.

2.1.3. Determination of Wavelet Packet Decomposition Levels. The selection of decomposition levels not only affects the fault feature extraction but also determines the dimension of feature vector. When the number of decomposition layers is too small, the information of each frequency band cannot be completely decomposed, and the bearing feature information is not accurately extracted, which affects the accuracy of fault diagnosis. Although increasing the number of wavelet packet decomposition layers can analyze the fault signal more finely, the number of signals after decomposition increases. When the number of decomposition layers is too many, the dimension of feature vector is too large, which affects the efficiency of fault identification. Therefore, the number of wavelet packet decomposition layers must consider the characteristics of the signal itself. In this paper, the optimal number of decomposition layers is calculated by the following equation [23]:

$$J = \left\{ \max \left| j < \log_2 \frac{f_s}{4 f_{sf}} \right. \right\}, \quad j \in Z, \tag{6}$$

where J is the maximum number of layers; f_s is the sampling frequency; and f_{sf} is the signal frequency.

For the vibration signal of rolling bearing, especially the fault state signal, the frequency of useful signal is divided into two types: (1) rotation frequency and (2) fault frequency [24]. The wavelet packet decomposition aims to find fault features, so the signal frequency f_{sf} can be replaced by fault feature frequency [25].

2.1.4. Selection of Wavelet Packet Basis Function

(1) Information Entropy Principle. Information entropy is the measure of information disorder in information theory. The greater the entropy, the greater the disorder of information and the smaller the contribution of information. On the contrary, the smaller the entropy, the smaller the disorder of information and the greater the contribution of information. The working state of rolling bearing is often expressed in the form of vibration state. When rolling bearing fails, the vibration signal will change accordingly. Therefore, extracting information entropy from vibration signal in the time-frequency domain can reflect the vibration state of rolling bearing.

(2) Wavelet Packet Energy Entropy. The construction steps of wavelet packet energy entropy are as follows.

Step 1. The composite fault signal of rolling bearing is decomposed by wavelet packet. After the signal is decomposed in j layers, 2^j sub-signals are generated. The energy of the node n of layer j, the $S_{j,n}$, is expressed as

$$E_{j,n} = \sum_{q=1}^{Q} \left(S_{j,n}(q) \right)^2, \tag{7}$$

where j is the number of wavelet packet decomposition layers; $n = 0, 1, \cdots, 2^j - 1$ is the node n of the layer j; and Q is the signal length.

Step 2. The total signal energy is expressed as

$$E_j = \sum_{n=0}^{2^j - 1} E_{j,n}. \tag{8}$$

Step 3. The proportion of energy of each node in the total energy is recorded as

$$p_{j,n} = \frac{E_{j,n}}{E_j}. \tag{9}$$

Step 4. $p_{j,0}, p_{j,1}, \cdots, p_{j,2^j-1}$ is the energy distribution of each frequency band in layer j after the signal is decomposed by wavelet packet. According to Shannon's theorem, the wavelet packet energy entropy corresponding to each node is defined as

$$H_{j,n} = -p_{j,n} * \log p_{j,n}. \tag{10}$$

Step 5. The total energy entropy of the signal is expressed as

$$H_j = -\sum_{n=0}^{2^j-1} p_{j,n} * \log p_{j,n}. \tag{11}$$

(3) Selection of Wavelet Packet Basis Functions. In wavelet packet decomposition, the parts where the signal waveform is similar to the waveform of the selected wavelet packet basis function are enhanced and the rest are suppressed [26], so the greater the wavelet packet energy after decomposition. In information theory, the more regular the signal is, the higher the contribution value of information will be, and the smaller the energy entropy of wavelet packet will be. According to the principle of maximum ratio of total energy and total energy entropy of wavelet packet, the larger the ratio is, the more similar the selected wavelet packet basis function is to the original signal [27]. The ratio formula of total energy and total energy entropy of wavelet packet is

$$\sigma = \frac{E_j}{H_j}. \tag{12}$$

2.1.5. Construction of Wavelet Packet AR Spectral Entropy Eigenvector. The construction steps of wavelet packet AR spectral entropy eigenvector are as follows.

Step 1. The optimal wavelet basis is selected to decompose the collected vibration signals by j level wavelet packet decomposition and generate 2^j wavelet packet coefficients.

Step 2. According to the wavelet packet filter selected in the decomposition process, its dual filter is selected for reconstruction. When reconstructing a certain frequency band signal, set the wavelet packet coefficients of other frequency bands to zero to make the reconstructed signal only contain the time-domain waveform of the frequency band signal.

Step 3. The AR spectrum of each reconstructed signal is estimated to obtain the AR spectrum containing only specific frequency information.

Step 4. Calculate the energy entropy of wavelet packet AR spectrum band.

Step 5. The energy entropy of wavelet packet AR spectrum band is normalized, and the feature vector is constructed.

2.2. Fault Diagnosis Model of ANVTPSO-SOM-BPNN

2.2.1. SOM-BPNN Algorithm. BPNN is a multilayer feedforward neural network trained according to the error back propagation. It is a supervised learning network, which is trained on the premise of known expected output. SOM is an unsupervised, self-organizing, and visual network composed of fully connected neuron arrays. The two are connected in series

to form a combined SOM-BPNN model, which has both the advantages of SOM and BPNN. After the sample data enters SOM, the preliminary classification of samples is realized. The essence of training the secondary network is to add a dimension to the training sample vector and use it as the input of the secondary network. The newly added dimension is used to mark the classification results of the primary network, which can promote the training of the secondary network. Theoretically, it can effectively reduce the training time of the secondary network and make the whole combined network converge faster. As the primary network training, SOM does not need a large sample set, so SOM-BPNN also has the same characteristics. Therefore, the combination of two neural networks can achieve the complementary advantages, so as to improve the accuracy of fault diagnosis. The essence of SOM-BPNN is to add a competition layer in front of the hidden layer of BPNN, and its structure is shown in Figure 2.

The implementation process of SOM-BPNN is as follows.

Step 1. Construct the training samples, and normalize the input samples.

Step 2. Determine the number of layers and nodes of SOM and BPNN, respectively.

Step 3. Classify the input samples preliminarily with SOM.

Step 4. Add a dimension to the training sample vector according to the preliminary classification results of SOM, and use the new vector as the input of the secondary BPNN.

Step 5. Start training after the BPNN input layer of the secondary network receiving the new sample vector, until the model reaches the convergence requirement.

The combined network is the SOM-BPNN model which can classify the input sample set more accurately. The classification of test samples is realized by inputting the test sample set into the model.

2.2.2. ANVTPSO Algorithm. PSO is the search for the optimal solution through the cooperation among individuals in the group. In practice, a group of random particles is initialized, and in each iterative search process, the particles continuously update through the extreme (P_{ia}, P_{ga}) until the optimal solution is found within the set number of iterative steps. Among them, P_{ia} is the optimal solution found so far by the particle itself, which is the individual extreme value, and P_{ga} is the optimal solution found so far by the whole population, which is the global extreme value.

$$V_{ia}(t+1) = w(t)V_{ia}(t) + c_1 r_1 (P_{ia}(t) - X_{ia}(t)) + c_2 r_2 (P_{ga}(t) - X_{ia}(t)),$$
$$X_{ia}(t+1) = X_{ia}(t) + V_{ia}(t+1), \tag{13}$$

where $V_{ia}(t+1)$ and $X_{ia}(t+1)$ are the particle velocity and position of the i particle in the a dimension in the $t+1$ iteration, respectively; $w(t)$ is inertia weight; t is the number of
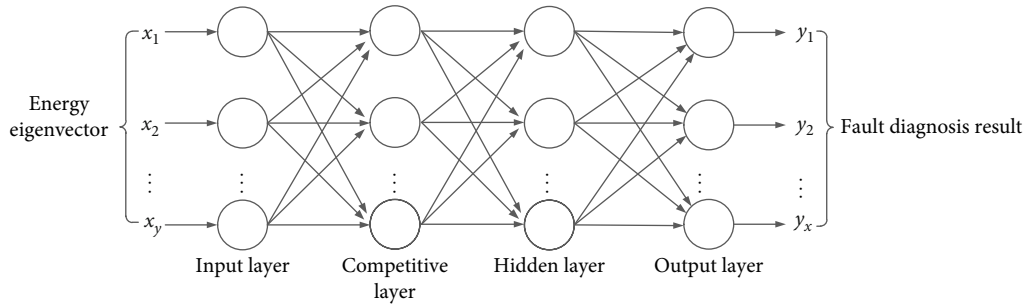
Figure 2: Topological structure of the SOM-BPNN model.

iterations; c_1 and c_2 are learning factors; and r_1 and r_2 are random numbers at $[0, 1]$.

In order to avoid the influence of randomly given initial velocity of particles on the convergence speed and accuracy, the velocity term of the standard PSO is abandoned [28], and the position is updated according to the following equation:

$$X_{ia}(t + 1) = w(t)X_{ia}(t) + c_1r_1(P_{ia}(t) - X_{ia}(t)) + c_2r_2(P_{ga}(t) - X_{ia}(t)). \tag{14}$$

PSO has the disadvantages of easy premature convergence, low convergence accuracy, and low later iteration efficiency [29]. Inertia weight w regulates the searching ability of particles in solution space, and its value affects the optimization level of the algorithm. Meanwhile, because PSO has the evolutionary stages, different learning factors should be set in different stages. Based on this, this paper uses an adaptive method to modify the inertia weight, which changes with the change of particle objective function value [30], expressed by equation (15). Asynchronous nonlinear adaptive adjustment learning factor is adopted [31], which is expressed by equation (16).

$$w = \begin{cases} w_{min} - \dfrac{(w_{max} - w_{min}) * (f - f_{min})}{f_{avg} - f_{min}}, & f \le f_{avg}, \\ w_{max}, & f > f_{avg}, \end{cases} \tag{15}$$

where w_{max} is the maximum inertia weight; w_{min} is the minimum inertia weight; f is the real-time objective function value of the particle; and f_{avg} and f_{min} are the average and minimum values of all current particles, respectively.

$$\begin{cases} c_1 = 2 + \dfrac{f - f_{avg}}{f_{avg} - f_{min}}, \\ c_2 = 2 - \dfrac{f - f_{avg}}{f_{avg} - f_{min}}, \end{cases} \tag{16}$$

where 2 is the initial value of learning factors c_1 and c_2.

2.2.3. ANVTPSO-SOM-BPNN Model. The preliminary classification of input samples is realized through SOM. According to the preliminary classification results, a dimension is added to the training sample vector, and the newly formed feature vector is used as the input of SOM-BPNN. However, the initial network connection weight and node threshold of

SOM-BPNN, like BPNN, are usually determined based on experience and are easy to fall into local optimal solution, which limits the convergence efficiency of the network. But PSO can search in a large space, and when it is used to optimize the threshold and weight of SOM-BPNN, it can avoid the above problems to a certain extent. Because the parameter setting of PSO has a great impact on the final result, this paper adopts an adaptive way to adjust the inertia weight and learning factor of PSO and round off its velocity term to avoid the influence of particle initial velocity on the convergence speed and solution accuracy, which is the new ANVTPSO algorithm, used for SOM-BPNN threshold and weight optimization, to improve the accuracy of fault diagnosis. ANVTPSO-SOM-BPNN diagnostic model is constructed, and the process is shown in Figure 3.

The process of ANVTPSO-SOM-BPNN algorithm is as follows.

Step 1. Set the input node, network competition layer, and other parameters in SOM, according to the characteristic data. Use the classification results obtained by SOM as the training sample vector, and add a dimension; then, form a new feature data set with the original feature data.

Step 2. Set the input node N, hidden layer node L, output node M, and other parameters according to the new feature data set. Clarify the structure of SOM-BPNN.

Step 3. Initialize PSO, calculate its search space dimension a, and set parameters such as population number and maximum iteration times T_{max}.

Step 4. Use the characteristic data as the input of SOM-BPNN to calculate the fitness value of each particle. Fitness function takes the mean square error function MSE between the actual training output and the expected output.

Step 5. Calculate the initial individual optimal position P_{ia} and global optimal position P_{ga} of PSO.

Step 6. Discard the velocity term of PSO, update the position according to equation (14), update the inertia weight according to equation (15), and update the learning factor according to equation (16), so as to obtain the individual and global
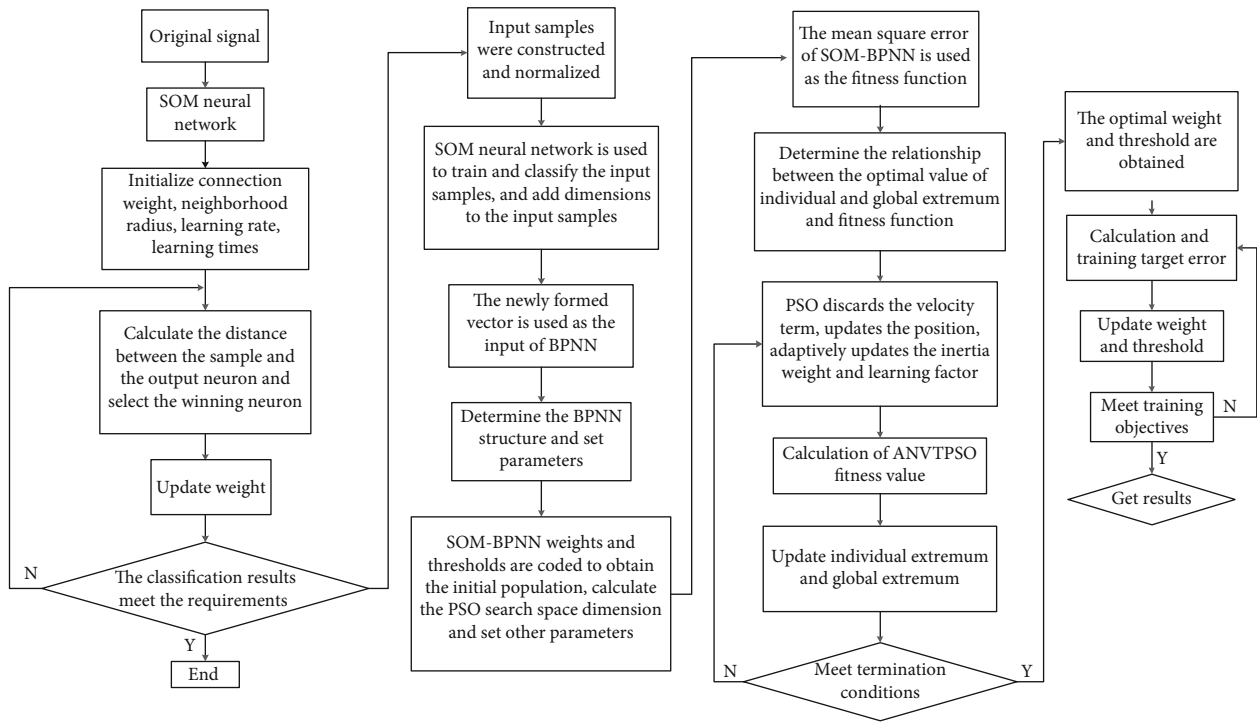
Figure 3: Flow chart of ANVTPSO-SOM-BPNN algorithm.
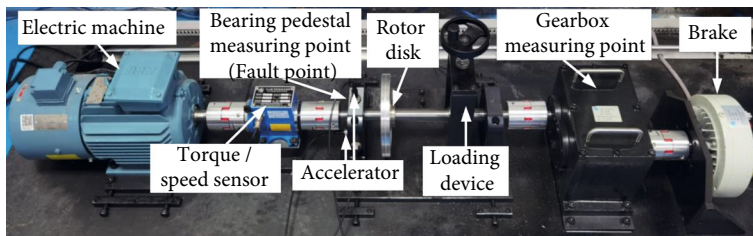

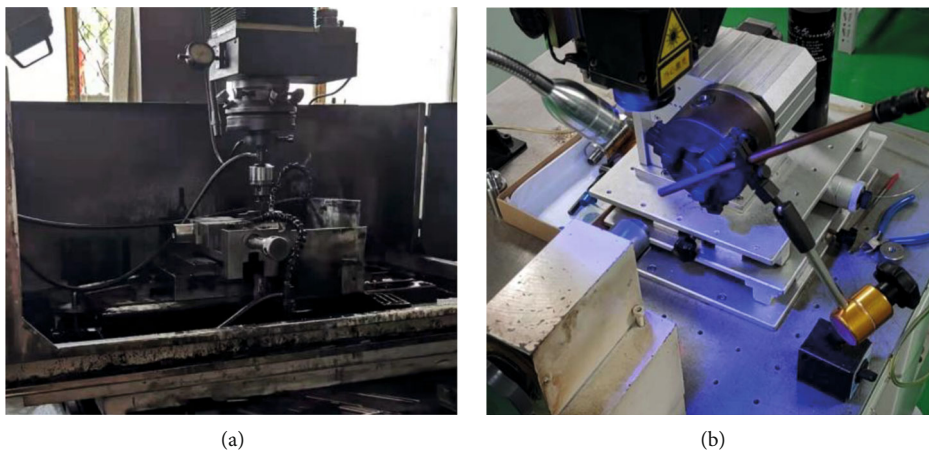Figure 4: Fault diagnosis comprehensive test platform.

(a)   (b)

Figure 5: Fault machining equipment: (a) EDM; (b) laser welding machine.

(a)

(b)

(c)

FIGURE 6: Finished fault bearings: (a) inner ring and outer ring; (b) inner ring and rolling element; (c) outer ring and rolling element.

TABLE 1: Fault size of rolling bearing.

Bearing status	Fault size (mm)
Normal	0
Inner ring crack and outer ring crack	$2 * 1.5 * 0.5 + 2 * 1.5 * 0.5$
Inner ring crack and rolling element pitting	$2 * 1.5 * 0.5 + 1s$ (pit corrosion)
Outer ring crack and rolling element pitting	$2 * 1.5 * 0.5 + 1s$ (pit corrosion)

TABLE 2: Fault characteristic frequency of rolling bearing.

Fault location	Frequency (Hz)
Inner ring	162.23
Outer ring	107.77
Rolling element	71.33

optimal extreme value. And then, the PSO position is mapped to obtain the optimal weight and threshold.

Step 7. Bring the optimized weight and threshold into SOM-BPNN, and continue tuning until the training objectives are met.

2.3. D-S Evidence Theory

2.3.1. Principle of D-S Evidence Theory. D-S evidence theory has good practicability, so it is widely used in the field of multisensor target recognition [32, 33]. Its main characteristics include the following: it satisfies the weaker conditions than Bayesian probability theory and has the ability to directly express "uncertainty" and "do not know" [34].

Definition 1. If Θ is a finite sample set and all propositions in the set are mutually exclusive, set Θ is called the identification framework. The power set of Θ is composed of all subsets and various combinations of identification framework Θ, noted as 2^{Θ}. Take Θ as the identification framework of

TABLE 3: Decomposition level of wavelet packet.

Fault location	Decomposition level
Inner ring	$J < 3.98$
Outer ring	$J < 4.57$
Rolling element	$J < 5.16$

proposition A. A assign the basic probability function m as $m(A)$ which is the mapping of set 2^Θ to $[0,1]$.

The function $m : 2^\Theta \longrightarrow [0, 1]$ satisfies the following conditions:

$$\begin{cases} m(\phi) = 0, \\ \sum_{A \subseteq \Theta} m(A) = 1, \end{cases} \tag{17}$$

where $0 \leq m(A) \leq 1$, A is called focal element, and $m(A)$ is the basic probability assignment of A, indicating the trust degree in A.

Definition 2. Mapping $Bel(A)$: $2^\Theta \longrightarrow [0, 1]$ is the confidence function defined on Θ, which reflects the exact trust degree of A. The expression is

$$Bel(A) = \sum_{B \subseteq A} m(B). \tag{18}$$

Mapping $PI(A)$: $2^\Theta \longrightarrow [0, 1]$ is a plausible function defined on Θ, which represents the degree of nonfalse trust in proposition A. It is also an uncertainty measure that seems to be possible for proposition A. The expression is

$$PI(A) = 1 - Bel(\bar{A}) \tag{19}$$

where $PI(A)$ and $Bel(A)$ represent the upper and lower limits of the function, respectively.

Definition 3. $[Bel(A), Pl(A)]$ is defined as the trust interval of proposition A. $[0, Bel(A)]$ represents the support evidence interval of proposition A. $[Pl(A), 1]$ represents the rejection evidence interval of proposition A.

Definition 4. D-S evidence theory synthesis rule: let m_1 and m_2 be the basic reliability distribution on the same identification framework Θ and meet the following conditions:

$$\sum_{A_i \cap A_j = \phi} m_1(A_i) m_2(A_j) < 1. \tag{20}$$

Then, the combined basic probability distribution function is

$$m(A) = \frac{\sum_{A_i \cap A_j = A} m_1(A_i) m_2(A_j)}{1 - \sum_{A_i \cap A_j = \phi} m_1(A_i) m_2(A_j)}. \tag{21}$$

2.3.2. Fault Diagnosis Based on D-S Evidence Fusion. The composite fault signals of rolling bearing obtained by multiple sensors are processed by wavelet packet AR spectral entropy, and the relevant eigenvalues are extracted. The composite fault diagnosis is carried out by using ANVTPSO-SOM-BPNN, and the output is used as evidence which is fused through D-S evidence theory to construct a new fault diagnosis model. The model makes full use of the advantages of D-S theory in dealing with uncertain problems and the powerful nonlinear processing ability of neural network and uses the self-learning ability of neural network to solve the problem that it is difficult to obtain the basic probability assignment in D-S theory. At the same time, if there is no noise, the target recognition will be easy, but in practice, the noise is inevitable. Therefore, using multiple sensors for recognition and fusing the recognition results of each sensor can improve the recognition rate.

The implementation process of the proposed diagnosis model based on D-S evidence fusion is as follows.

Step 1. Obtain the target feature vector. The collected composite fault signals of rolling bearing are extracted by wavelet packet AR spectral entropy.

Step 2. Input the target eigenvector into ANVTPSO-SOM-BPNN model.

Step 3. Normalize the diagnostic output of ANVTPSO-SOM-BPNN model, with a range of $[0, 1]$; calculate the error E_n between the actual output and the expected output of the diagnostic model, as shown in equation (22). The basic probability value of each focus element is shown in equation (23). The uncertainty degree $m(\theta)$ of the diagnostic model is shown in equation (24).

$$E_n = \frac{1}{2} * \sum (t_{ni} - y_{ni}), \tag{22}$$

where t_{ni} is the expected value of the output neuron and y_{ni} is the actual value of the output neuron.

$$m(A_i) = \frac{y(A_i)}{S_n}, \tag{23}$$

where $m(A_i)$ is the basic probability of each focal element; $y(A_i)$ is the diagnostic result; and $S_n = \sum_{i=1}^{n} y(A_i) + E_n$.

$$m(\theta) = 1 - \sum_{i=0}^{R} m(A_i). \tag{24}$$

Step 4. Obtain the final result by multi-information fusion with evidence combination rules.

3. Experiments

3.1. Experimental Data Collection. In order to verify the effect of the composite fault diagnosis method of rolling bearing based on ANVTPSO-SOM-BPNN combined with wavelet packet AR spectral entropy, the experimental test

TABLE 4: Ratio of total energy and total energy entropy of wavelet packet (radial mean value of direct fault point).

Type	sym8	db4	db5	db8	db10
Normal	0.7071	0.7075	0.7069	0.7072	0.7078
Inner ring and outer ring	0.8215	0.8218	0.8212	0.8216	0.8221
Inner ring and rolling element	0.8714	0.8710	0.8705	0.8715	0.8728
Outer ring and rolling element	1.0236	1.0220	1.0219	1.0237	1.0249

TABLE 5: Ratio of total energy and total energy entropy of wavelet packet (axial mean value of direct fault point).

Type	sym8	db4	db5	db8	db10
Normal	1.0148	1.0118	1.0122	1.0149	1.0156
Inner ring and outer ring	1.1893	1.1892	1.1886	1.1891	1.1896
Inner ring and rolling element	0.9888	0.9871	0.9872	0.9893	0.9903
Outer ring and rolling element	1.0543	1.0529	1.0528	1.0544	1.0553

TABLE 6: Ratio of total energy and total energy entropy of wavelet packet (radial mean value of gearbox).

Type	sym8	db4	db5	db8	db10
Normal	2.6329	2.6375	2.6346	2.6329	2.6332
Inner ring and outer ring	2.5379	2.5415	2.5385	2.5393	2.5401
Inner ring and rolling element	2.6280	2.6309	2.6279	2.6286	2.6282
Outer ring and rolling element	2.5099	2.5119	2.5096	2.5088	2.5097

TABLE 7: Ratio of total energy and total energy entropy of wavelet packet (axial mean value of gearbox).

Type	sym8	db4	db5	db8	db10
Normal	3.2479	3.2479	3.2456	3.2472	3.2497
Inner ring and outer ring	2.8983	2.8962	2.8951	2.8985	2.9004
Inner ring and rolling element	3.2146	3.2151	3.2128	3.2148	3.2167
Outer ring and rolling element	3.1967	3.1949	3.1932	3.1968	3.1991

TABLE 8: Parameters of ANSVTPSO-SOM-BPNN.

ANVTPSO parameters	Value	SOM parameters	Value	BPNN parameters	Value
Number of population particles x	20	Input node	9	Input node N	10
Spatial dimension a	199	Network competition layer	$6 * 6$	Hidden layer node L	13
Position X	[-1, 1]	Topological function	Hextop	Output node M	4
Learning factor c_1	Undetermined	Distance function	Linkdist	Maximum training times	1000
Learning factor c_2	Undetermined	Classification stage learning rate	0.9	Training objectives	10^{-8}
Inertia weight W_{max}	0.9	Classification stage learning step	1000	Learning rate	0.001
Inertia weight W_{min}	0.4	Learning rate in tuning phase	0.02	Hidden layer transfer function	Tansig
Maximum number of iterations T_{max}	20	Tuning phase domain distance	1	Output layer transfer function	Purelin

is carried out on the comprehensive experimental platform for fault diagnosis of mechanical transmission system, as shown in Figure 4. The test-bed consists of three-phase variable frequency motor, rotor bearing system, radial loading device, parallel shaft gearbox, and magnetic particle brake. The rolling bearing in the bearing pedestal on the left side of the rotor is selected as the tested object. The used bearing model is NSK6205, the number of rolling elements Z is 9, the diameter of rolling elements d is 7.94 mm, the pitch diameter D is 39.36 mm, and the contact angle α is $0°$.

Set the motor speed at 1800 r/min and no load; install the composite fault part at the bearing seat as the fault source. Set the bearing pedestal and gearbox as two measuring points for vibration signal acquisition, and the vertical

TABLE 9: Comparison of bearing fault diagnosis results (radial direction of fault point).

Diagnostic method	Iteration steps	Diagnostic accuracy (%)
BPNN	12	100

TABLE 10: Comparison of bearing fault diagnosis results (axial direction of fault point).

Diagnostic method	Iteration steps	Diagnostic accuracy (%)
BPNN	135	97
ANVTPSO-SOM-BPNN	81	97

radial and axial direction corresponding to the measuring points adopt the accelerators with a sensitivity of 103 mV/g (g is gravity acceleration). When collecting the vibration signal of composite fault of rolling bearing, the sampling time is set as 1 s and the sampling rate is set as 10.24 kHz. A total of 300 groups of vibration acceleration signals are collected, including the normal, inner ring crack and outer ring crack, inner ring crack and rolling element pitting, and outer ring crack and rolling element pitting, and each type has 75 groups. The signal samples are divided into a training set and test set in 2 : 1.

The inner ring and outer ring are machined by using electrical discharge machine (EDM), and the rolling element has the 1 second pit corrosion by using TH-RFT300 high-speed laser welding machine. The fault machining equipment is shown in Figure 5. The finished fault bearings are shown in Figure 6. The fault size is shown in Table 1, and the fault characteristic frequency is shown in Table 2.

3.2. Experimental Data Analysis

3.2.1. Determining the Optimal Number of Wavelet Packet Decomposition Layers.
The purpose of wavelet packet decomposition is to find fault characteristics. Therefore, the signal frequency can be replaced by fault characteristic frequency. The number of decomposition layers can be calculated by equation (6), as shown in Table 3.

Table 3 shows that according to the characteristic frequencies of different fault parts of the bearing, the best values of wavelet packet decomposition layers are 3 to 5. Because the composite fault signal is more complex than a single fault case, in order to retain the useful information of four types of bearing vibration signals to the greatest extent, in the selection of unified decomposition layers, if the number of decomposition layers exceeds 3, the inner ring signal may be over decomposed, resulting in the loss of useful information in the composite fault. After comprehensive consideration, the number of wavelet packet decomposition layers in this paper is 3.

3.2.2. Determining the Optimal Wavelet Basis Function.
75 groups of 4 types of bearing data are selected, and the 5 types of wavelet bases sym8, db4, db5, db8, and db10 are decomposed in 3 layers by wavelet packet and calculated according to formula (12) to obtain the ratio of wavelet packet total energy and total energy entropy corresponding to the 4 types of bearing data. In order to eliminate the uncertain influence caused by individual signals, the mean value of parameters under various states is calculated. The corresponding calculation results are shown in Tables 4–7.

It can be seen from Tables 4 and 5 that the ratio of total energy and total energy entropy of wavelet packet in 4 types of bearings in the measuring points of bearing pedestal (direct fault point) is db10, which is the largest in the data of radial and axial measuring points. According to the principle that the greater the ratio of total energy and total energy entropy of wavelet packet, the better the decomposition effect of wavelet packet, db10 is regarded as the optimal wavelet basis function of wavelet packet decomposition of 4 kinds of bearing signals in radial measuring points and axial measuring points of bearing pedestal. It can be seen from Tables 6 and 7 that the ratio of total energy of wavelet packet to total energy entropy of four types of bearings in the gearbox measuring points is db4 and db10, respectively, in the radial and axial data. Similarly, db4 and db10 are taken as the optimal wavelet basis function for wavelet packet decomposition of 4 types of bearing signals in the gearbox radial direction measuring points and axial measuring points.

3.3. Determination of the ANVTPSO-SOM-BPNN Parameters.
The parameters of ANVTPSO-SOM-BPNN are shown in Table 8, where the spatial dimension of particles a [35] and the selection of the optimal number of nodes L in the hidden layer are shown in equations (25) and (26), respectively.

$$a = N * L + L * M + L + M, \tag{25}$$

$$L = \sqrt{(M + N)} + \varepsilon, \quad 0 < \varepsilon < 10. \tag{26}$$

3.4. Result Analysis.
BPNN, SOM-BPNN, PSO-SOM-BPNN, and ANVTPSO-SOM-BPNN were used for the composite fault diagnosis of rolling bearing, respectively. The expected output of the state categories of the tested rolling bearing are normal (1 0 0 0), inner ring fault (0 1 0 0), outer ring fault (0 0 1 0), and rolling element fault (0 0 0 1).

In order to verify the advantages of the method proposed in this paper, at first, optimize the wavelet basis function and decomposition levels of wavelet packet AR spectrum energy entropy, extract the characteristics of energy entropy, and compare SOM-BPNN with standard BPNN to verify that the series network has more advantages in convergence speed than a single network. Second, the PSO-SOM-BPNN is compared with SOM-BPNN to verify the optimization effect of PSO on SOM-BPNN. Then, compare the above 3 schemes BPNN with ANVTPSO-SOM-BPNN, study the series advantages of both unsupervised learning network and supervised learning network, and verify the impact of improved PSO on fault diagnosis results. Finally, the collected multisensor data are used for fault diagnosis through the ANVTPSO-SOM-BPNN constructed in this paper, and the results are fused at the decision level through D-S evidence theory, so as to improve the final fault diagnosis rate.
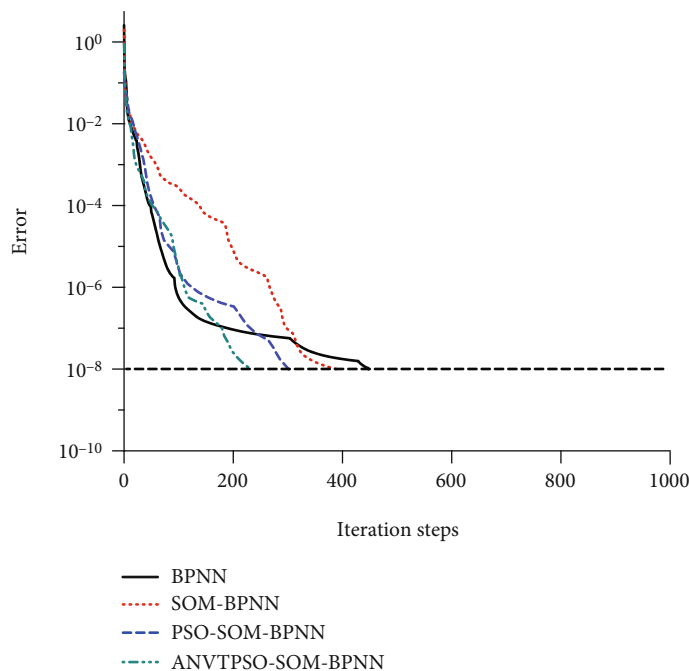
FIGURE 7: Error curves of four kinds of diagnosis models (gearbox radial direction).
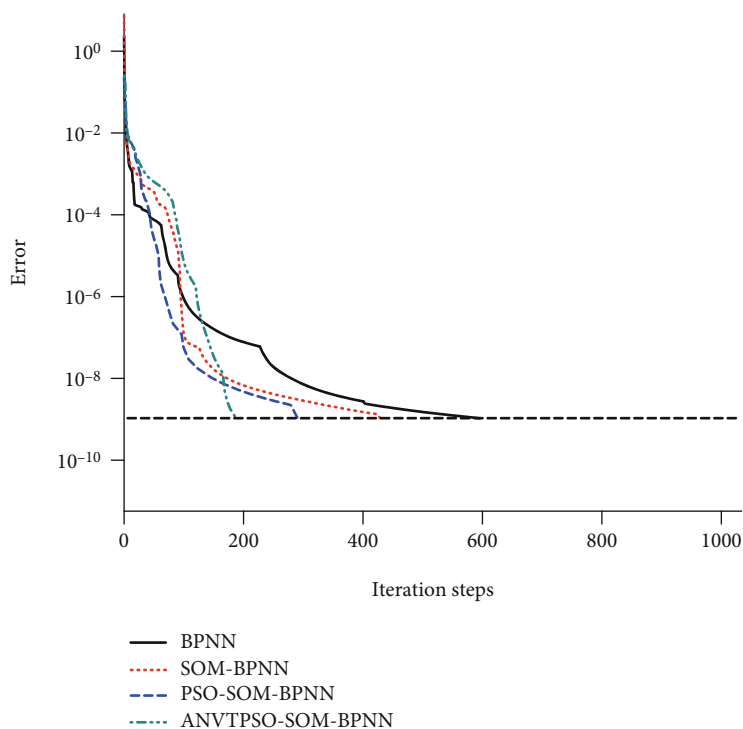

FIGURE 8: Error curves of four kinds of diagnosis models (gearbox axial direction).

Through fault diagnosis of vibration signals of radial measuring points and axial measuring points of bearing pedestal, the diagnosis results are shown in Tables 9 and 10, respectively.

As shown in Table 9, in the radial measuring points of the bearing pedestal, because they are close to the fault point,

there is less noise interference, and the fault characteristics of the collected vibration signals are obvious. Therefore, the fault diagnosis using standard BPNN can reach 100%, and the number of iterative steps is only 12. Table 10 presents that in the axial measuring point, the standard BPNN used for fault diagnosis reach 97%, and the number of

TABLE 11: Comparison of bearing fault diagnosis results (gearbox radial direction).

Diagnostic method	Iteration steps	Diagnostic accuracy (%)
BPNN	448	82
SOM-BPNN	389	85
PSO-SOM-BPNN	302	87
ANVTPSO-SOM-BPNN	240	92

TABLE 12: Comparison of bearing fault diagnosis results (gearbox axial direction).

Diagnostic method	Iteration steps	Diagnostic accuracy (%)
BPNN	576	91
SOM-BPNN	415	93
PSO-SOM-BPNN	282	95
ANVTPSO-SOM-BPNN	182	96

iterative steps is 135. The ANVTPSO-SOM-BPNN method proposed in this paper is used for diagnosis, with an accuracy of 97%, which is the same as the diagnosis result of the standard BPNN, but the number of iterative steps is 81; compared with the former, it reduces 54 steps.

To sum up, the data collected at the measuring point of the bearing pedestal has less interference and obvious fault characteristics, so the basic diagnosis algorithm used in both radial and axial direction data has a high accuracy. However, in real working conditions, due to the influence of various on-site factors, it is impossible to install sensors to collect vibration signals close to the direct fault point. Therefore, the indirect gearbox measuring point is more universal in line with the actual working conditions.

Figures 7 and 8, respectively, demonstrate the mean square error of the four diagnostic methods of the radial measuring point and axial measuring point of gearbox with the number of training times. The comparison of diagnostic results of the composite fault of rolling bearing is shown in Tables 11 and 12.

Figure 7 indicates that BPNN, PSO-SOM-BPNN, and ANVTPSO-SOM-BPNN intersect near 50 steps and the error of ANVTPSO-SOM-BPNN is the smallest before the intersection. After the intersection, the BPNN always keeps the minimum error until the second intersection with ANVTPSO-SOM-BPNN near 200 steps. After the second intersection, ANVTPSO-SOM-BPNN converges faster. Before SOM-BPNN intersects with BPNN, SOM-BPNN error is always the largest, and it converges faster after the bifurcation point. From the details of the iterative process, the four methods all have fallen into the local minimum for a short time, resulting in the increase of the total iterative steps, but ANVTPSO-SOM-BPNN performs better than the other methods. Figure 8 displays that BPNN and SOM-BPNN intersect near 100 steps.

The error of BPNN before intersection is the smallest, but it is easier to fall into the local minimum than SOM-BPNN, and the convergence speed becomes slower and the total number of iterative steps increases after intersection. Before the intersection of PSO-SOM-BPNN and ANVTPSO-SOM-BPNN, the error is the smallest among the four methods, and the number of relative falling into the local minimum is the least. After the intersection, ANVTPSO-SOM-BPNN converges faster and takes the least iterative steps to reach the training target.

Above all, ANVTPSO-SOM-BPNN has the advantages of series connection of unsupervised learning network and supervised learning network. Combined with ANVTPSO, at the gearbox measuring points with more interference, it can reach the training target faster for both radial and axial vibration signal diagnosis, which proves that the proposed new method has an obvious optimization effect.

Tables 11 and 12 show the quantitative data of the four methods in the radial and axial measuring points of gearbox. Table 11 reveals that the diagnostic accuracy of SOM-BPNN is 3% higher than that of BPNN. Compared with SOM-BPNN, the diagnostic accuracy of PSO-SOM-BPNN is improved by 2%, and the number of iterative steps is reduced by 87. Table 12 suggests that the diagnostic accuracy of SOM-BPNN is 2% higher than that of BPNN. Compared with SOM-BPNN, the diagnostic accuracy of PSO-SOM-BPNN is improved by 2%, and the number of iterative steps is reduced by 133. In both radial and axial directions, the PSO learning factor in PSO-SOM-BPNN is taken as $c_1 = c_2 = 1.49445$ according to experience, while the ANVTPSO-SOM-BPNN adaptively adjusts the inertia weight and learning factor, so that the inertia weight is taken as $w = 0.9$ in the initial stage and $w = 0.4$ in the later stage, the radial learning factor is $c_1 = 1.4175$ and $c_2 = 2.5825$ in the later stage, and the axial learning factor is $c_1 = 1.6984$ and $c_2 = 2.3016$ in the later stage. As mentioned above, the axial learning factor of bearing pedestal is $c_1 = 1.5736$ and $c_2 = 2.4264$ in the later stage. Both the inertia weight and learning factor meet the needs of different stages of the algorithm through adaptation. After many tests, the accuracy of ANVTPSO-SOM-BPNN is higher than other methods, and the number of iterative steps also has great advantages. Table 11 presents that in the radial direction, the diagnostic accuracy reaches 92% at step 240. Table 12 shows that in the axial direction, the diagnostic accuracy reached 96% at step 182.

Tables 13 and 14 display the partial basic probability distribution values of the radial and axial measuring points of gearbox after the output results of the ANVTPSO-SOM-BPNN model are processed according to formula (22) and formula (23) and the uncertainty degree of the diagnostic model according to equation (24). Table 15 presents the partial basic probability distribution values and uncertainty degree of the two measuring points after D-S evidence fusion. Table 16 indicates that the diagnosis results of the newly proposed ANVTPSO-SOM-BPNN method in this paper at the radial and axial measuring points of gearbox are fused at the decision level, and the accuracy of fault diagnosis reaches 100%. The diagnostic accuracy after fusion was improved by 8% and 4% compared with the radial and axial ANVTPSO-SOM-BPNN in the gearbox, respectively. When

TABLE 13: Basic probability distribution value of ANVTPSO-SOM-BPNN (gearbox radial direction).

Type	$m_1(A_1)$	$m_1(A_2)$	$m_1(A_3)$	$m_1(A_4)$	$m_1(\theta)$
Normal	0.9373	0.0416	0.0200	0	0.0011
Inner ring and outer ring	0.0764	0.8403	0.0764	0	0.0069
Inner ring and rolling element	0.0212	0.0422	0.9354	0	0.0012
Outer ring and rolling element	0.0839	0	0.0840	0.8235	0.0086

TABLE 14: Basic probability distribution value of ANVTPSO-SOM-BPNN (gearbox axial direction).

Type	$m_2(A_1)$	$m_2(A_2)$	$m_2(A_3)$	$m_2(A_4)$	$m_2(\theta)$
Normal	0.7670	0	0.2050	0.0006	0.0274
Inner ring and outer ring	0	0.9994	0.0004	0.0002	0
Inner ring and rolling element	0.1051	0	0.8887	0	0.0062
Outer ring and rolling element	0.0000	0.0059	0.0117	0.9823	0.0001

TABLE 15: Basic probability distribution value after D-S evidence fusion.

Type	$m(A_1)$	$m(A_2)$	$m(A_3)$	$m(A_4)$	$m(\theta)$
Normal	0.9920	0.0015	0.0065	0	0
Inner ring and outer ring	0	1.0000	0	0	0
Inner ring and rolling element	0.0030	0.0003	0.9967	0	0
Outer ring and rolling element	0	0.0001	0.0013	0.9986	0

TABLE 16: D-S evidence fusion diagnosis results.

Diagnostic method	Diagnostic accuracy (%)
ANVTPSO-SOM-BPNN (radial & axial)	100

compared with the diagnostic accuracy of BPNN in the radial direction of the fault point, it also reaches 100%. And compared with the diagnostic accuracy of ANVTPSO-SOM-BPNN in the axial direction of the fault point, it is improved by 3%. In conclusion, it is proved that the method proposed in this paper can achieve high diagnosis accuracy even at the gearbox measuring point far away from the fault point.

4. Conclusions

(1) The method of wavelet packet AR spectrum energy entropy can effectively extract the composite fault feature components in the vibration signal of rolling bearing and can better eliminate interference and noise. The optimal selection of wavelet packet decomposition layers and basis function in wavelet packet AR spectrum energy entropy can avoid the external interference caused by blind selection

(2) The adaptive inertia weight and learning factor are introduced into the standard PSO algorithm to meet the needs of the algorithm for parameters in different stages, and the velocity term is discarded to avoid the influence of the initial particle velocity on the convergence speed and solution accuracy of the algorithm, which significantly improves the search speed and convergence accuracy of the algorithm compared with the conventional method

(3) Build ANVTPSO-SOM-BPNN diagnostic model. SOM-BPNN avoids the influence of the limitations of a single algorithm on the diagnosis results, so that the primary network can promote the training of the secondary network. Then ANVTPSO is used to optimize the threshold and weight of SOM-BPNN to avoid falling into the local optimal solution, so as to improve the diagnostic accuracy

(4) In the actual working condition, it is common for the same rolling bearing to coexist multiple faults, and the installation scheme of sensors also have a great impact on the accuracy of diagnosis. In this paper, both the data of fault point and far away from fault point are collected by multiple acceleration sensors, and the proposed method based on optimal wavelet packet AR spectrum energy entropy combined with ANVTPSO-SOM-BPNN is used for multi-information fusion diagnosis. By comparing the diagnosis results of two measuring points, it is found that even at the gearbox measuring point far away from the direct fault point, the diagnosis results can achieve high accuracy and effectively diagnose the composite fault of rolling bearing under noise

Data Availability

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

References

[1] A. Gałęzia and K. Gryllias, "Application of the combined Teager-Kaiser envelope for bearing fault diagnosis," *Measurement*, vol. 182, article 109710, 2021.

[2] D. T. Hoang and H. J. Kang, "A survey on deep learning based bearing fault diagnosis," *Neurocomputing*, vol. 335, pp. 327–335, 2019.

[3] K. Yu, R. L. Tian, J. W. Tan, and H. Ma, "An adaptive sensitive frequency band selection method for empirical wavelet transform and its application in bearing fault diagnosis," *Measurement*, vol. 134, pp. 375–384, 2019.

[4] X. Li, Y. Yang, H. D. Shao, X. Zhong, J. Cheng, and J. Cheng, "Symplectic weighted sparse support matrix machine for gear fault diagnosis," *Measurement*, vol. 168, article 108392, 2021.

[5] D. D. Liu, W. D. Cheng, W. G. Wen, and G. T. Wan, "Composite fault diagnosis of rolling bearing based on envelope demodulation filter," *Journal of Central South University(natural science)*, vol. 49, no. 4, pp. 881–887, 2018.

[6] L. L. Cui, C. G. Wu, and N. Wu, "Composite fault diagnosis of rolling bearings based on EMD and ICA algorithm," *Journal of Beijing University of Technology*, vol. 40, no. 10, pp. 1459–1464, 2014.

[7] X. X. Zheng and M. N. Li, "Health state evaluation based on wavelet packet and pchmm for vulnerable components of wind turbines," *Acta Energiae Solaris Sinica*, vol. 40, no. 2, pp. 370–379, 2019.

[8] G. J. Tang and F. Y. Deng, "Compound fault features separation method of rolling element bearing based on improved harmonic wavelet packet decomposition," *Chinese Journal of Scientific Instrument*, vol. 36, no. 1, pp. 143–151, 2015.

[9] S. He, J. Chen, Z. Zhou, Y. Zi, Y. Wang, and X. Wang, "Multifractal entropy based adaptive multiwavelet construction and its application for mechanical compound-fault diagnosis," *Mechanical Systems and Signal Processing*, vol. 76-77, pp. 742–758, 2016.

[10] B. D. Ma, S. J. Hu, Y. F. Meng, B. Song, and T. Wu, "Rolling bearings compound fault diagnosis method based on multiwavelet packet and energy operator demodulating method," *Noise and Vibration Control*, vol. 38, no. 4, pp. 154–158, 2018.

[11] S. Abbasion, A. Rafsanjani, A. Farshidianfar, and N. Irani, "Rolling element bearings multi-fault classification based on the wavelet denoising and support vector machine," *Mechanical Systems & Signal Processing*, vol. 21, no. 7, pp. 2933–2945, 2007.

[12] O. Janssens, V. Slavkovikj, B. Vervisch et al., "Convolutional neural network based fault detection for rotating machinery," *Journal of Sound and Vibration*, vol. 377, pp. 331–345, 2016.

[13] N. Lv and P. X. Yao, "Fault diagnosis of rolling bearing based on BP neural network," *Coal Mine Machinery*, vol. 41, no. 8, pp. 172-173, 2020.

[14] J. M. Li, X. F. Yao, X. D. Wang, Q. W. Yu, and Y. G. Zhang, "Multiscale local features learning based on BP neural network for rolling bearing intelligent fault diagnosis," *Measurement*, vol. 153, 2020.

[15] W. C. Huang, L. J. Wang, J. W. Liu, and B. J. Chen, "Rolling bearings fault recognition based on multi feature fusion and GA-BP mode," *Machine Tool & Hydraulics*, vol. 49, no. 6, pp. 170–173, 2021.

[16] M. Gong, H. X. Pan, and H. L. Lan, "Diesel engine fault diagnosis based on combination of optimal wavelet packet and SOM-BP," *Coal Mine Machinery*, vol. 33, no. 10, pp. 278–280, 2012.

[17] C. Ju, C. Zhang, H. W. Fan, X. H. Zhang, Y. Q. Yang, and Y. Yan, "Rolling bearing fault diagnosis based on wavelet packet decomposition and PSO-BPNN," *Industry and Mine Automation*, vol. 46, no. 8, pp. 70–74, 2020.

[18] L. Wang and X. K. Wang, "Modified particle swarm optimizer using non-linear inertia weight," *Computer Engineering and Applications*, vol. 43, no. 4, pp. 47-48, 2007.

[19] Y. M. Zhu and P. X. Xue, "The research of PSO based on the adaptive changes of acceleration coefficients," *Journal of Shaanxi University of Science & Technology(Natural Science Edition)*, vol. 33, no. 4, pp. 172–177, 2015.

[20] M. Khazaee, H. Ahmadi, M. Omid, A. Moosavian, and M. Khazaee, "Classifier fusion of vibration and acoustic signals for fault diagnosis and classification of planetary gears based on Dempster-Shafer evidence theory," *Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering*, vol. 228, no. 1, pp. 21–32, 2014.

[21] D. C. Feng and J. M. D. Pereira, "Study on information fusion based on wavelet neural network and evidence theory in fault diagnosis," in *2007 8th International Conference on Electronic Measurement and Instruments*, Xi'an, China, 2007IEEE Press.

[22] L. L. Zhang, Y. G. Zhao, Y. K. Xiao, S. D. Luo, H. Y. Liao, and Q. X. Pan, "Automobile transmission bearing fault feature extraction based on wavelet packet-auto regressive model spectrum," *Journal of Vibration, Measurement & Diagnosis*, vol. 31, no. 4, pp. 492–495, 2011.

[23] W. Q. Wang and S. Yang, "A method for choosing the wavelet decomposition level in structural fault analysis," *Structure & Environment Engineering*, vol. 36, no. 4, pp. 34–41, 2009.

[24] X. L. Liu, H. G. Liu, J. H. Yang, G. Litak, G. Cheng, and S. Han, "Improving the bearing fault diagnosis efficiency by the adaptive stochastic resonance in a new nonlinear system," *Mechanical Systems and Signal Processing*, vol. 96, pp. 58–76, 2017.

[25] M. Y. Liu, *Research of motor bearing fault diagnosis methods based on vibration signal [M.S. thesis]*, Harbin Institute of Technology, Harbin, China, 2016.

[26] M. L. Tian and X. L. Wang, "Wavelet analysis method applied in fault diagnosis of motor and selection of wavelet base," *Coal Mine Machinery*, vol. 28, no. 5, pp. 176–178, 2007.

[27] P. K. Kankar, S. C. Sharma, and S. P. Harsha, "Fault diagnosis of rolling element bearing using cyclic autocorrelation and wavelet transform," *Neurocomputing*, vol. 110, no. 8, pp. 9–17, 2013.

[28] W. Hu and Z. S. Li, "A simpler and more effective particle swarm optimization algorithm," *Journal of Software*, vol. 18, no. 4, pp. 861–868, 2007.

[29] X. Wei, N. Q. Shu, P. C. Cui, and B. Wu, "Power transformer fault integrated diagnosis based on improved PSO-BP neural networks and D-S evidential reasoning," *Automation of Electric Power Systems*, vol. 30, no. 7, pp. 46–50, 2006.

[30] Z. Wen and H. K. Sun, *MATLAB intelligent algorithm*, Tsinghua University Press, Beijing, China, 2017.

[31] Q. J. Tong, Q. Zhao, and M. Li, "Particle swarm optimization algorithm based on adaptive dynamic change," *Microelectronics & Computer*, vol. 36, no. 2, pp. 6–10, 2019.

[32] Y. B. Li, J. Chen, F. Ye, and D. D. Liu, "The improvement of DS evidence theory and its application in IR/MMW target recognition," *Journal of Sensors*, vol. 2016, no. 6, 2016.

[33] P. Wang, C. X. Shang, and Z. Z. Han, "Target identification based on evidence theory of improved basic probability assignment," in *Electronics, Communications and Networks IV*, pp. 1171–1175, London, U.K., 2015.

[34] S. Li, J. W. Tan, and K. Yu, "Composite fault diagnosis research of rolling bearing based on combination of neural network and improved D-S evidence theory," *Machine Tool & Hydraulics*, vol. 46, no. 1, pp. 153–157, 2018.

[35] J. Dong, W. W. Qin, Y. J. Li, Q. Q. Li, and L. W. Deng, "Fast multi-objective antenna design based on improved back propagation neural network surrogate model," *Journal of Electronics & Information Technology*, vol. 40, no. 11, pp. 2712–2719, 2018.