Original scientific paper

A ROLLING BEARING FAULT DIAGNOSIS METHOD BASED ON EXTREME LEARNING MACHINE OPTIMIZED BY IMPROVED WHALE OPTIMIZATION ALGORITHM

Xin Li[1], Shiliang Guo[2], Dejie Sun[1], Lijun Cao[1], Cong Li[1], Shuyao Tian[3], Peng Liu[2], Yadong Qi[2]

[1]School of Mathematics and Information Science & Technology,
Hebei Normal University of Science & Technology, Qinhuangdao, China
[2]School of Electrical Engineering, Yanshan University, Qinhuangdao, China
[3]College of Electronics and Control Engineering,
North China Institute of Aerospace Engineering, Langfang, China

Abstract. *Rolling bearing is one of the most commonly used components in rotating machinery, and researching fault diagnosis techniques for it has important practical significance. In this paper, a fault diagnosis method based on extreme learning machine optimized by improved whale optimization algorithm (IWOA-ELM) is proposed for rolling bearing vibration signals. Firstly, Variational Mode Decomposition (VMD) is used to decompose the vibration signal of the bearing, and the energy entropy is calculated to form the eigenvector. Secondly, based on the original whale optimization algorithm, a hybrid initialization population strategy is adopted to generate an initial population with a certain quality. Selecting convergence factors based on reinforcement learning to improve global search capability, and using adaptive weights and random jumps to update individual positions. In this process, the t-distribution-levy flight variation strategy is introduced to avoid being attracted by local extremum. Then, the improved whale optimization algorithm is used to optimize the input weights and hidden layer thresholds of the Extreme Learning Machine (ELM). Finally, the feature set is input into an improved ELM model for training and testing. Experiments on fault diagnosis of rolling bearings of different types and degrees have shown that the model proposed in this paper can effectively improve the accuracy of fault classification.*

Key words: *Fault diagnosis, Variational mode decomposition, Whale optimization algorithm, Reinforcement learning*

1. INTRODUCTION

Rotating machinery plays an important role in modern society, and rolling bearings are critical components that impact its performance. However, the working environment of most rolling bearings is relatively harsh, which makes them prone to failure, resulting in significant economic losses and even catastrophic events. The fault vibration signal of rolling bearings has nonlinear and non-stationary characteristics. How to extract fault information from vibration signals and accurately locate bearing faults are essential for fault diagnosis and normal machinery operation [1-4]

The application of machine learning methods in fault diagnosis can minimize manual involvement and greatly improve the accuracy of fault diagnosis, such as support vector machines (SVM) [5-7], backpropagation neural networks (BPNN) [8-10], and convolutional neural networks (CNN) [11-13]. However, these models face limitations: SVM needs to store large datasets, making it difficult to solve multi classification problems; BPNN and CNN have slow convergence speed and high computational cost, particularly with deeper layers.

Extreme Learning Machine (ELM) has the characteristics of fast training speed, few learning parameters, strong generalization ability, and is widely used in the field of fault diagnosis [14-16]. Wang et al. [17] proposed a bearing fault diagnosis method based on ELM for adaptive parameter optimization, and introduced spectral cross-correlation to help select the optimal penalty factor, improving classification accuracy. In order to further enhance the generalization ability of ELM, Zhang et al. [18] used the Multitask Beetle Antennae Swarm Algorithm (MBAS) to optimize the input weights and biases of ELM, which can simultaneously reduce the number of conditions and regression errors. From the above description, it can be seen that optimizing the parameters of ELM based on optimization algorithms can further improve the performance of the model. Whale Optimization Algorithm (WOA) is a meta-heuristic algorithm developed in recent years, which is used to handle various optimization problems in different fields due to its advantage of fewer adjustable parameters [19-21]. However, as the dimensionality of the search space increases, WOA faces problems such as easily falling into local optima and slow convergence speed.

Based on the above analysis, we propose a rolling bearing fault diagnosis method based on extreme learning machine optimized by improved whale optimization algorithm (IWOA-ELM), as shown in Fig. 1. Firstly, the variational mode decomposition (VMD) method is used to extract the fault signal features of rolling bearings, which are then divided into a training set and a testing set. Then, based on the WOA, a hybrid initialization population strategy, a convergence factor selection based on reinforcement learning, an adaptive weight and random jump update strategy, and a t-distribution-levy flight variation strategy are introduced to solve the problem of the algorithm falling into local optimum. Finally, the IWOA was used to optimize the parameter settings of the extreme learning machine, thereby improving the speed and accuracy of bearing fault diagnosis. The IWOA-ELM algorithm was compared with WOA-ELM, PSO-ELM, GWO-ELM, and the experimental results verified the effectiveness of the improved rolling bearing fault diagnosis method proposed in this paper.
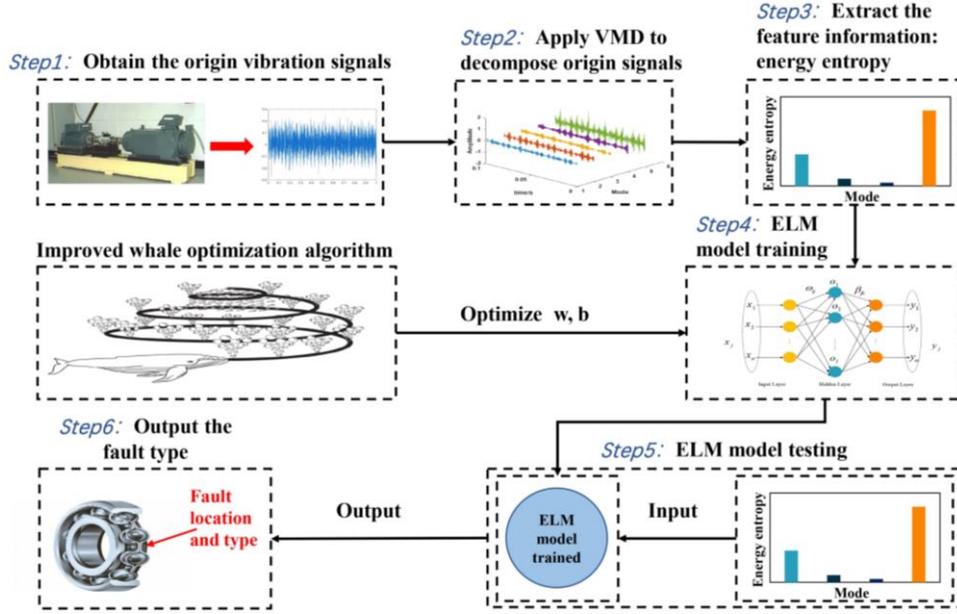
Fig. 1 Methodological framework of this study

2. SIGNAL DECOMPOSITION AND FEATURE EXTRACTION

2.1 VMD Theory

VMD is an innovative and adaptive signal processing method [22] that can solve the common endpoint effects and mode component mixing problems in Empirical Mode Decomposition (EMD). VMD decomposes the initial signal into k discrete sub-signals, ensuring that each decomposition represents a finite bandwidth modal component centered at a specific frequency. At the same time, the sum of estimated bandwidths for each modality is minimized, and the sum of all modes is equal to the original signal as a constraint. The constraint variational expression can be written as:

$$\left\{ \min_{\{u_k\},\{\omega_k\}} \left\{ \sum_k \left\| \partial_t \left[\left(\delta(t) + j/\pi t \right) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\} \atop \text{s.t.} \sum_{k=1}^{K} u_k = f \right\} \tag{1}$$

where, k is the number of decomposed modes. The symbols $\{u_k\}$ and $\{\omega_k\}$ refer to the kth mode component and center frequency after decomposing, respectively. $\delta(t)$ is the Dirac function where t represents time. * is the convolution operator.

Furthermore, we introduce the Lagrange multiplication operator λ and the quadratic penalty factor α to find the optimal solution for the constrained variational problem, and obtain the augmented function:

$$L\left(\{u_k\},\{\omega_k\},\lambda\right)=\alpha\sum_k\left\|\partial_t\left[\left(\delta(t)+\mathrm{j}/\pi t\right)*u_k(t)\right]\mathrm{e}^{-\mathrm{j}\omega_k t}\right\|_2^2$$
$$+\left\|f(t)-\sum_k u_k(t)\right\|_2^2+\left\langle\lambda(t),f(t)-\sum_k u_k(t)\right\rangle \tag{2}$$

where, α is a quadratic penalty factor used to reduce the interference of Gaussian noise. The alternate direction method of multipliers (ADMM) iterative algorithm combined with Parseval/Plancherel and Fourier isometry transform is used to optimize each eigenmode component and center frequency, and to search for saddle points of the extended Lagrangian function. The expressions for u_k, ω_k, and λ change after alternating optimization iteration is as follows:

$$\hat{u}_k^{n+1}(\omega)\leftarrow\frac{\hat{f}(\omega)-\sum_{i\neq k}\hat{u}_i(\omega)+\hat{\lambda}(\omega)/2}{1+2\alpha\left(\omega-\omega_k\right)^2} \tag{3}$$

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

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

where, γ is the noise margin, $\hat{u}_k^{n+1}(\omega)$, $\hat{u}_i(\omega)$, $\hat{f}(\omega)$ and $\hat{\lambda}(\omega)$ correspond to the Fourier transforms of $u_k^{n+1}(t)$, $u_i(t)$, $f(t)$ and $\lambda(t)$ respectively. The main iterative VMD solution process is as follows:

   (1) Initialize $\hat{u}_k^1,\omega_k^1,\hat{\lambda}^1$ and set the maximum number of iterations N, $n\leftarrow 0$;
   (2) Using Eqs. (3) and (4), update $\hat{u}_k$ and ω_k;
   (3) Update $\hat{\lambda}$ using Eq. (5);
   (4) Accuracy convergence criterion $\varepsilon>0$, if $n<N$ and $\sum_k\left\|\hat{u}_k^{n+1}-\hat{u}_k^n\right\|_2^2/\left\|\hat{u}_k^n\right\|_2^2<\varepsilon$ are not

satisfied, return to the second step. Otherwise, the iteration is completed.

2.2. Extraction of Feature: Energy Entropy

Yu [23] proposed the theory of energy entropy, and judged the fault information of rolling bearings based on the difference in energy by different frequency bands. The energy for any part of the waveform is defined as:

$$E_i=\int_{-\infty}^{+\infty}\left|c_i(t)\right|^2\mathrm{d}t=\sum_{k=1}^N\left|x_{ik}^2\right| \tag{6}$$

where, N is value of sample points of the ith IMF, and x_{ik} is the the amplitude.

Decompose the original vibration signal into n IMF vibration components, and the energy of each IMF component is E_i. The energy of the original vibration wave is

$E_{\text{total}} = \sum_{i=1}^{n} E_i$. P_i is the amount of energy in the i-th IMF part compared to the total energy of the original signal:

$$P_i = E_i / E_{total} \tag{7}$$

The corresponding energy entropy equation is:

$$H_{EN} = -\sum_{i=1}^{n} \left(P_i \log P_i \right) \tag{8}$$

3. IWOA-ELM MODEL

3.1 ELM theory

ELM is used to train single hidden layer feedforward neural networks (SLFN). Fig. 2 illustrates the network structure of ELM.
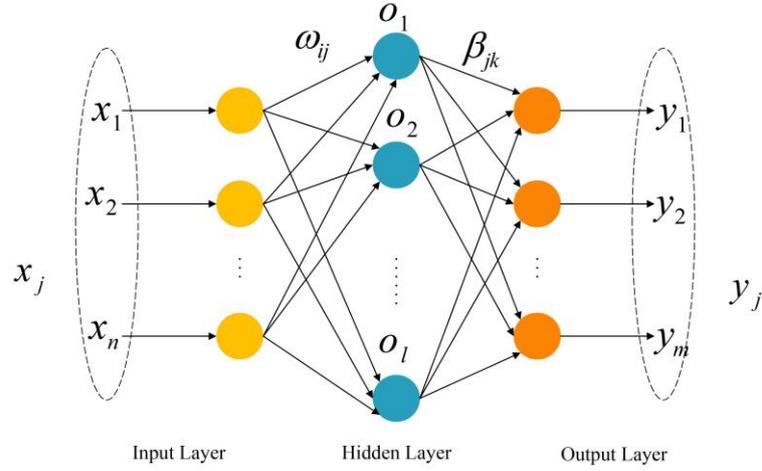

Fig. 2 The network structure of ELM

Assuming there are N random samples $(x_i,\ t_i)$, $i\ =\ 1,\ 2,…,\ N$, where $x_i = \left[x_{i1}, x_{i2}, \cdots, x_{in} \right]^T \in R^n$ and $t_i = \left[t_{i1}, t_{i2}, \cdots, t_{in} \right]^T \in R^m$. For a SLFN with L hidden layer nodes and an activation function $g(x)$ can be calculated:

$$\sum_{i=1}^{L} \beta_i g_i \left(x_j \right) = \sum_{i=1}^{L} \beta_i g_i \left(\omega_i \cdot x_j + b_i \right) = t_j, \mathrm{j} = 1, 2, \cdots, n \tag{9}$$

where, ω_i, b_i, and β_i represent the input weight, deviation, and output weight of the ith node in the hidden layer. We can use a matrix to represent it: $H\beta = T$, where H is the output of the hidden layer node, β is the output weight, and T is the output matrix.

$$H\left(\omega_1, \cdots, \omega_L, b_1, \cdots, b_L, x_1, \cdots, x_L\right) = \begin{bmatrix} g\left(\omega_1 \cdot x_1 + b_1\right) & \cdots & g\left(\omega_L \cdot x_1 + b_L\right) \\ \vdots & \cdots & \vdots \\ g\left(\omega_1 \cdot x_N + b_1\right) & \cdots & g\left(\omega_L \cdot x_N + b_L\right) \end{bmatrix}_{N \times L} \tag{10}$$

$$\beta = \begin{bmatrix} \beta_1^{\mathrm{T}} \\ \beta_2^{\mathrm{T}} \\ \vdots \\ \beta_L^{\mathrm{T}} \end{bmatrix}, \mathrm{T} = \begin{bmatrix} t_1^{\mathrm{T}} \\ t_2^{\mathrm{T}} \\ \vdots \\ t_N^{\mathrm{T}} \end{bmatrix}_{N \times m} \tag{11}$$

Thus, the output weight is:

$$\beta = H^+ T \tag{12}$$

where, H^+ is the Moore-Penrose generalized inverse of H.

3.2 WOA Theory

In WOA algorithm, each whale represents a set of solutions and can be optimized by its location and speed. When a whale detects its prey, it performs a spiral motion around it and releases a bubble net to trap the prey. The whale individuals can interact with each other in various ways, including "encircling prey," "bubble net attack," and "search for prey," which are described as follows:

3.2.1. Encircling Prey

During hunting, whales will continuously update their position and swim towards their prey after determining its location. In the solution space of the algorithm, assume that the position of the individual whale close to the prey is the optimal position. Other individual whales follow the best whale to update their position and gradually approach their prey. The above behavior is expressed as:

$$\overrightarrow{X}(t+1) = \overrightarrow{X}^*(t) - \overrightarrow{A} \cdot \overrightarrow{D} \tag{13}$$

$$\overrightarrow{D} = \left| \overrightarrow{C} \cdot \overrightarrow{X}^*(t) - \overrightarrow{X}(t) \right| \tag{14}$$

where, t is the current number of iterations, $\vec{X}^*$ is best position, $\vec{X}$ is current position of an individual whale, D stands for distance. $\vec{A}$ and $\vec{C}$ are coefficient vectors, which are calculated as follows:

$$\overrightarrow{A} = 2\overrightarrow{\alpha} \cdot \overrightarrow{r_1} - \overrightarrow{\alpha} \tag{15}$$

$$\overrightarrow{C} = 2 \cdot \overrightarrow{r_2} \tag{16}$$

$$a = 2 - \frac{2t}{M} \tag{17}$$

where, a is the convergence factor with its value decreasing linearly from 2 to 0 during the iteration, M is the maximum number of iterations. $\vec{r_1}$ and $\vec{r_2}$ are both random vectors in [0,1].

3.2.2. Bubble-Net Attacking Method (Exploitation Phase)

At this stage, the shrinking encircling mechanism of the WOA algorithm is shown in equation (13). In addition, after determining the position of the prey, the humpback whale will move along the spiral towards the prey and update its individual position. The mathematical model is as follows:

$$\overline{X}(t+1) = \overline{D}' \cdot e^{bl} \cdot \cos(2\pi l) + \overline{X}^*(t) \tag{18}$$

$$\overline{D}' = \left| \overrightarrow{X^*}(t) - \overrightarrow{X}(t) \right| \tag{19}$$

where, $\overrightarrow{D}'$ is the distance between the current whale's position and the best whale's position. b is a logarithmic spiral shape constant, and l is a random value between [-1,1].

The two hunting behaviors of shrinking encircling and spiral movement are performed simultaneously. These two actions are selected by setting the parameter p, where p is a random number between [0,1], and the probability of $p<0.5$ and $p \geq 0.5$ is 50%. The mathematical model is as follows:

$$\overrightarrow{X}(t+1) = \begin{cases} \overline{X}^*(t) - \overline{X} \cdot \overline{D} & p < 0.5 \\ \overline{D}' \cdot e^{bl} \cdot \cos(2\pi l) + \overline{X}^*(t) & p > 0.5 \end{cases} \tag{20}$$

3.2.3. Search for Prey (Exploration Phase)

At this stage, the behavior of whales is controlled by coefficient vectors $\vec{A}$ to expand the search range and improve search capabilities. When the value of A is not within the range of [-1,1], individual whales will randomly select other whales as guides to update their position. The mathematical model is as follows:

$$\overrightarrow{X}(t+1) = \overrightarrow{X}_{rand} - \overrightarrow{A} \cdot \overrightarrow{D} \tag{21}$$

$$\overrightarrow{D} = \left| \overrightarrow{C} \cdot \overrightarrow{X}_{rand} - \overrightarrow{X}(t) \right| \tag{22}$$

where, $\vec{X}_{rand}$ is a vector of random individual positions.

3.3 Improved Whale Optimization Algorithm

The WOA algorithm was briefly summarized in the previous section. In order to improve the algorithm's ability and efficiency in finding the global optimal solution, this section proposes IWOA. A hybrid initialization population scheme is designed based on the original WOA. Meanwhile, we select convergence factor based on reinforcement

learning and design adaptive weight. The random jump update strategy and t-distribution-levy flight variation strategy are implemented to improve the search ability of the algorithm.

3.3.1. Hybrid Initialization

In this section, we use improved Tent chaos mapping and opposition-based learning to initialize the population, and design a random exchange strategy for disadvantaged populations.

(1) Improved Tent chaos mapping

Tent chaotic mapping is a piecewise linear mapping that has more uniform traversal and faster search speed than Logistic chaos mapping. The expression of Tent chaos mapping is:

$$x_{i+1} = \begin{cases} 2x_i & 0 \le x \le \dfrac{1}{2} \\ 2(1-x_i) & \dfrac{1}{2} < x \le 1 \end{cases} \tag{23}$$

The expression after Bernoulli shift transformation is:

$$x_{i+1} = (2x_i)\bmod 1 \tag{24}$$

To prevent the Tent chaotic sequence from getting stuck in unstable periodic points during the iteration, Zhang [24] introduced a random variable to the original Tent chaotic mapping equation as follows:

$$x_{i+1} = \begin{cases} 2x_i + rand(0,1)\times\dfrac{1}{N} & 0 \le x \le \dfrac{1}{2} \\ 2(1-x_i) + rand(0,1)\times\dfrac{1}{N} & \dfrac{1}{2} < x \le 1 \end{cases} \tag{25}$$

The expression after Bernoulli shift transformation is:

$$x_{i+1} = (2x_i)\bmod 1 + rand(0,1)\times\dfrac{1}{N} \tag{26}$$

where N is the number of particles within the sequence.

(2) Opposition-based learning

Opposition-based learning mechanism can effectively improve the quality of the population and is applied to the improvement of various optimization algorithms. The expression of OBL is as follows:

$$x_i = lb + rand(ub - lb) \tag{27}$$

$$x_i^{'} = lb + ub - x_i \tag{28}$$

where *rand* is a random value on the interval [0,1], *lb* and *ub* are the upper and lower bounds of the search space, respectively. Applying the OBL idea to the randomly generated

N initial solutions $X_N = \{x_1, x_2, \cdots, x_n\}$ to obtain N reverse solutions $X_N' = \{x_1', x_2', \cdots, x_n'\}$. In the process of searching for the optimal solution, the current solution and the reverse solution are simultaneously searched, and the better solution among them is taken as the problem solution, which can expand the search scope and improve the search efficiency of the algorithm. The initial solutions generated by chaotic mapping and opposition-based learning are mixed into a population with $2N$ initial solutions, and then the N initial solutions with smaller fitness values are selected according to the ranking of fitness values. This move can reduce the loss of good solutions.

(3) Random exchange of inferior populations

In order to further improve the quality of the initial population, two different individuals were randomly selected from the N initial solutions with poor fitness generated in the previous step to form a group (total $N/2$ groups). The algorithm randomly select K non repeating positions for each individual in each group, and replace the K positions of individual 1 with those of individual 2. Similarly, perform the same operation on individual 2 to generate N new initial solution individuals. Then the fitness value of each individual is calculated and compared with the previous N smaller initial solutions, and the N initial solutions with smaller fitness values are selected. The random exchange operation for the inferior groups is shown in Fig. 3.
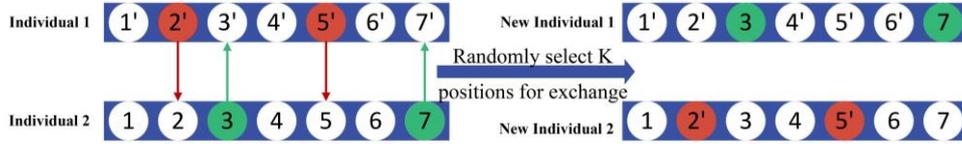

Fig. 3 Random exchange

3.3.2. Improved Hunting Patterns

(1) Adaptive weight

In order to improve the ability of traditional WOA algorithms to find optimal solutions, we introduce an exponentially changing adaptive weight during the algorithm's prey encirclement stage. The algorithm uses larger weights in the early stage to improve its global search ability, but when approaching the optimal solution, the weight values decrease exponentially to enhance the algorithm's local optimization ability. The weight expression can be written as:

$$\omega = e^{-10(t/T)} \tag{29}$$

Taking the above equation into Eq. (20), the individual position update model for the encircling prey stage is obtained:

$$\vec{X}(t+1) = \begin{cases} \omega \cdot \vec{X}^*(t) - \vec{A} \cdot \vec{D} & p < 0.5 \\ \vec{D}' \cdot e^{bl} \cdot \cos(2\pi l) + \omega \cdot \vec{X}^*(t) & p \geq 0.5 \end{cases} \tag{30}$$

where T is the highest number of iterations and t is the current number of iterations.

(2) Random jump update strategy

The superiority of WOA in the population iteration process depends on the search agent selected from the previous generation of the population. From Eqs. (21) and (22), it can be derived that the search agent position update during exploration is based on a randomly selected individual, and the new population convergence speed and accuracy are necessarily affected by the random individual. To address this phenomenon, a random jump update strategy is used to select the search agent more efficiently. The improved formula is as follows:

$$\vec{D} = \left| \vec{C} \cdot \vec{X}_{rand} - \vec{X}(t) \right| \tag{31}$$

$$\vec{X}(t+1) = \vec{C} \bullet \vec{D} + \vec{C} \bullet \vec{X}(t) \tag{32}$$

(3) Reinforcement learning based convergence factor selection

Whales act 'nonlinearly' in the real hunting process, and the linear descent strategy used by traditional algorithms cannot fully reflect the true situation of WOA. Therefore, this paper adopts different nonlinear control strategies to optimize the convergence factor a.

Ref [25] uses a nonlinear convergence factor based on sinusoidal function to improve whale optimization algorithm, and its mathematical model is:

$$a = a_{max} - \left(a_{max} - a_{min} \right) \sin \left(t\pi / 2T \right). \tag{33}$$

On this basis, several adjustment strategies for control parameter a were proposed in references [26, 27], such as:

Nonlinear decreasing strategy based on quadratic function:

$$a = a_{max} - \left(a_{max} - a_{min} \right) \left(t / T \right)^2 \tag{34}$$

Nonlinear decreasing strategy based on logarithmic form:

$$a = a_{max} - \left(a_{max} - a_{min} \right) \ln \left(1 + t(e-1) / T \right) \tag{35}$$

Through numerous experiments in references [26, 27], it can be seen that the convergence factor adjustment strategy has different effects on the utility of test functions. In order to investigate the experimental effect, the various adjustment strategies are considered as actions, and a suitable action is chosen to control the position of the whale at each step of the whale hunt. In addition, the possibility of making the right decision is greater if one can consider a few more steps forward before making a decision. Therefore, when whales choose actions, the effectiveness of multiple evolutionary steps can be judged by evaluating the advantages and disadvantages of different actions, which increases the whales' ability to find the best.

Reinforcement learning can gradually learn the optimal action policy by selecting the action that maximizes the agent's cumulative return with discount. Q-learning is a model independent reinforcement learning algorithm that obtains the optimal action strategy through multi-step learning of the agent to find the action with the maximum discounted return. If the adjustment strategy of the convergence factor is regarded as an action, the individual whale's choice of the optimal adjustment strategy is transformed into the agent's

choice of the optimal action. Based on this, the standard whale optimization algorithm, Eq. (33), Eq. (34), and Eq. (35) four convergence factor adjustment strategies are selected as the set of actions in this paper.

The basic format of Q-learning is

$$Q(s,a) = r(s,a) + \gamma \max_{a'} Q(s',a') \tag{36}$$

where, $r(s,a)$ is the immediate reward of the agent after taking action a in state s. s and s' represent the two states of the agent, while a and a' represent the two actions of the agent. $\gamma(0 \leq \gamma < 1)$ is the discount factor, and $\max Q(s',a')$ denotes the agent the maximum value of the reward that can be obtained by choosing a different action in the next state s'.

Assume that the set of operations available to the agent in any state is $A = \{a_1, a_2, \ldots, a_n\}$. If the first action taken by the agent is a, then the revenue when the agent chooses the action a after $m+1$ steps is calculated as:

$$Q(a) = r(a) + \gamma Q(a^{(1)}) + \gamma^2 Q(a^{(2)}) + \cdots + \gamma^m Q(a^{(m)}) \tag{37}$$

where, a, $a^{(i)} \in A$, and $1 \leq i \leq m$, the parameter m controls the number of steps to calculate the Q value looking forward.

With the above Q-learning algorithm, the mapping between WOA and Q-learning algorithm can be achieved by considering the whale as an agent and various distance control parameter adjustment strategies as a set of agent actions. For any individual whale in the WOA, looking forward $m+1$ steps, the gain $Q(a)$ when the whale chooses action a can be calculated by Eq. (37). Define the individual immediate gain, $r(a) = f_p(a) - f_o(a)$. Here, $f_p(a)$ is value of fitness function corresponding to parent whale individual and $f_o(a)$ is value of fitness function about offspring whale individual generated after selecting action a. If the number of optional actions of whales is n, then n^m whales will be generated after m times of evolution. To calculate the optimal evolutionary strategy for the whales, an exponential calculation is required, which is too computationally intensive. For this purpose, the probability of each offspring being retained is calculated using the Boltzmann distribution:

$$p(a_i) = e^{r(a_i)/T} / \sum_{i=1}^{n} e^{r(a_i)/T} \tag{38}$$

Furthermore, the highest probability of the n whale offspring generated will be retained. After this simplification, only one of the n offspring produced by a whale after one evolution is preserved during the evolution process, which greatly reduces the complexity of the algorithm. The convergence of the evolutionary planning algorithm based on Q-learning is proved in the [28]. The proof process of WOA is similar, and its convergence will not be proved in this paper due to the limitation of space.

3.3.3. T-Distribution-Levy Flight Variation Strategy

(1) T-distribution variation strategy

T-distribution has a degree of freedom parameter of n and its probability density function is:

$$p_t\left(x\right) = \frac{\Gamma\left(\dfrac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\dfrac{n}{2}\right)}\left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}} \quad -\infty < x < +\infty \tag{39}$$

For each whale, the algorithm generates a random number in the [0,1]. If the variation probability is greater than the generated random number, t-distribution perturbation is applied to it. The specific location update method is as follows:

$$x_{new}^{j} = x_{best}^{j} + t(iteration)\cdot x_{best}^{j} \tag{40}$$

where, $t(iteration)$ is the t-distribution with degree of freedom parameter as the number of iterations. x_{best}^{j} and x_{new}^{j} are the positions of the optimal whale individual in the jth dimension before and after variation perturbation, respectively. If the new solution after t-distribution perturbation has a better fitness value, the old solution is replaced; otherwise, the old solution is retained.

(2) Levy flight strategy

Levy flight is a search method that obeys levy distribution [29]. This paper introduces the Levy flight strategy into the whale optimization algorithm after T-distribution perturbation. By conducting small-scale exploration around the optimal solution, the search area of the algorithm is further increased to avoid the population falling into local optima. The location update formula is

$$x_{new}^{k} = x_{best}^{k} + levy(\beta) \tag{41}$$

where, $levy(\beta)$ is a method to generate random numbers that obey the levy distribution. In this paper, we take advantage of the normal distribution to resolve random numbers [30]. The steps to generate random numbers submitted to levy distribution are as follows:

$$s \sim levy(\beta) = \frac{u}{\left|v\right|^{\frac{1}{\beta}}} \tag{42}$$

where, u obeys a normal distribution, i.e. $u \sim N\left(0,\sigma^2\right)$, and v obeys a standard normal distribution, i.e. $v \sim N\left(0,1\right)$. The value of σ is taken as follows:

$$\sigma = \left[\frac{\Gamma(1+\beta)\sin(\dfrac{\pi\beta}{2})}{2^{\frac{\beta-1}{2}}\beta\Gamma\left(\dfrac{1+\beta}{2}\right)}\right]^{\frac{1}{\beta}} \tag{43}$$

Generally speaking, the range of values for β is $1\leq\beta\leq3$, which is taken as 1.5 in this paper.

3.3.4. IWOA Algorithm Performance Test

In order to verify the feasibility of the algorithm, 8 benchmark test functions are chosen to evaluate the IWOA algorithm. Table 1 displays the test functions. For each benchmark function, IWOA is compared with WOA, OBCWOA [31], REMWOA [32], GWO [33] and GOA [34], SSA [35]. The population size is set to 30 and the maximum number of iterations is 500.The parameters of the OBCWOA method are configured as follows: $a=4$, $k=1$. The relevant parameters of GOA algorithm are $c_{max}=1$ and $c_{min}=0.00004$. The SSA algorithm is set to $ST=0.8$, the proportion of producers is 20% and the proportion of scouts is 10%.

Table 1 Functions for testing

Function	Dim	Range	f_{min}	Peak
$f_1(x) = \sum_{i=1}^{n} x_i^2$	30	[-100,100]	0	Single
$f_2(x) = \sum_{i=1}^{n} \|x_i\| + \prod_{i=1}^{n} \|x_i\|$	30	[-10,10]	0	Single
$f_3(x) = \sum_{i=1}^{n} \left(\sum_{j-1}^{i} x_j \right)^2$	30	[-100,100]	0	Single
$f_4(x) = \sum_{i=1}^{n} i x_i^4 + random[0,1)$	30	[-1.28,1.28]	0	Single
$f_5(x) = \sum_{i=1}^{n} \left[x_i^2 - 10\cos(2\pi x_i) + 10 \right]$	30	[-5.12,5.12]	0	Multiple
$f_6(x) = -20\exp\left(-0.2\sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2} \right)$ $-\exp\left(\frac{1}{n}\sum_{i=1}^{n}\cos(2\pi x_i) \right) + 20 + e$	30	[-32,32]	0	Multiple
$f_7(x) = \frac{1}{400}\sum_{i=1}^{n} x_i^2 - \prod_{i=1}^{n}\cos\frac{x_i}{\sqrt{i}} + 1$	30	[-600,600]	0	Multiple
$f_8(x) = \sum_{i=1}^{11} \left[a_i - \frac{x_1\left(b_i^2 + b_i x_2\right)}{b_i^2 + b_i x_3 + x_4} \right]^2$	4	[-5,5]	0.0003	Multiple

Fig. 4 illustrates the iterative optimization process of several algorithms. It can be seen that the original WOA has a slow convergence rate. Although the convergence speed of OBCWOA and REMWOA has improved, it is easy to fall into local optimum in different test functions. IWOA improves the original whale optimization algorithm by using multiple optimization strategies and improves the optimization capability of WOA. The IWOA algorithm can jump out of local optima and find the optimal solution of the function in a timely manner even when encountering local optima, which improves the optimization performance of the algorithm.
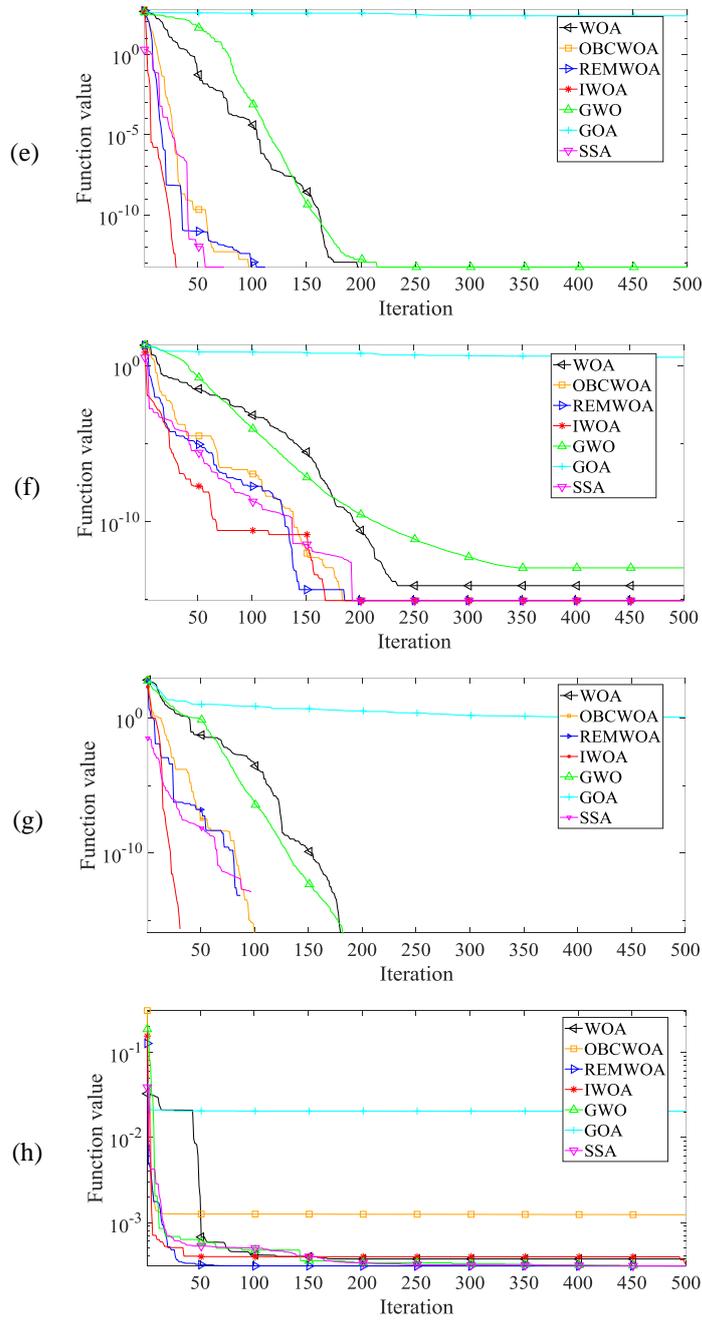
(a)

(b)

(c)

(d)
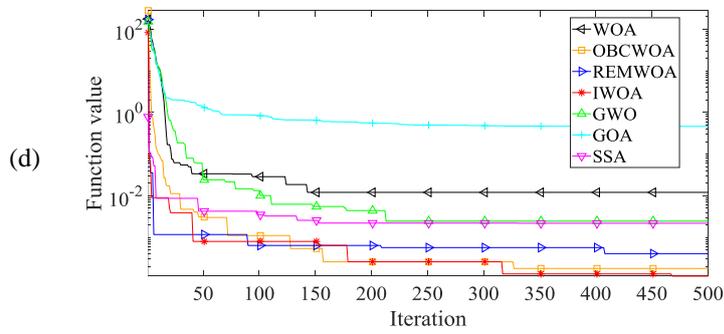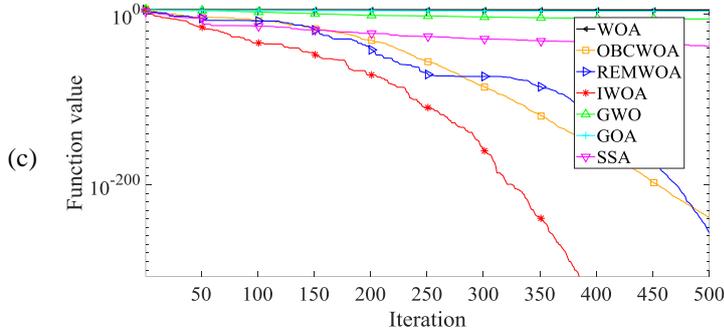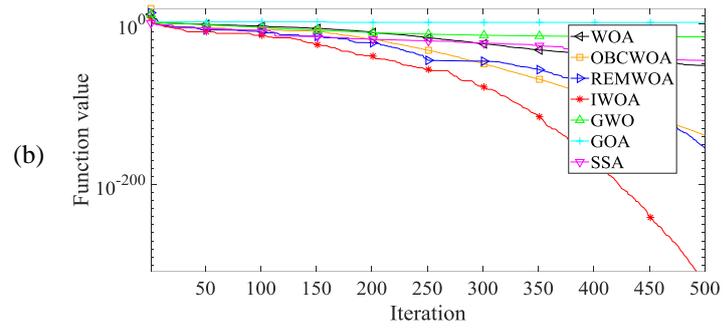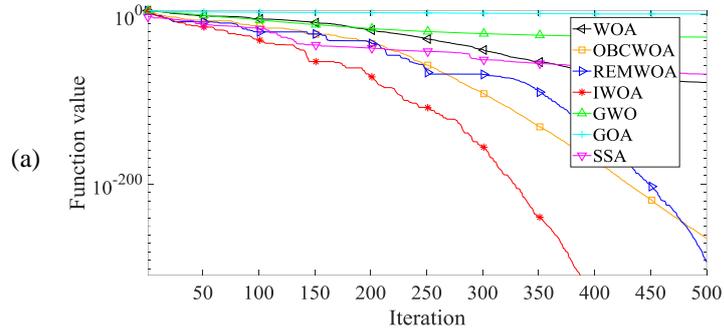
Fig. 4 Convergence curves of 8 test functions:
(a) $f_1(x)$; (b) $f_2(x)$; (c) $f_3(x)$; (d) $f_4(x)$; (e) $f_5(x)$; (f) $f_6(x)$; (g) $f_7(x)$; (h) $f_8(x)$

All algorithms were executed 30 times on the 8 benchmark functions, and the average values were computed. The results of experiment are given in Table 2.

Table 2 Optimization performance of 8 benchmark functions

Function	WOA	REMWOA	OBCWOA	IWOA	GOA	SSA	GWO
$f_1(x)$	$5.97×10^{-72}$	$1.26×10^{-290}$	$2.06×10^{-262}$	0	93.63	$3.99×10^{-60}$	$1.00×10^{-27}$
$f_2(x)$	$6.11×10^{-52}$	$1.26×10^{-147}$	$6.55×10^{-135}$	0	77.88	$1.79×10^{-30}$	$7.39×10^{-17}$
$f_3(x)$	43722.24	$1.27×10^{-233}$	$6.4×10^{-225}$	0	2007.86	$2.63×10^{-24}$	$3.92×10^{-5}$
$f_4(x)$	0.003952	$8.29×10^{-5}$	$9.39×10^{-5}$	$7.25×10^{-5}$	0.463665	0.001386	0.002197
$f_5(x)$	0	0	0	0	189.4202	0	2.3886621
$f_6(x)$	$4.32×10^{-15}$	$8.88×10^{-16}$	$8.88×10^{-16}$	$8.88×10^{-16}$	8.60	$8.88×10^{-16}$	$1.05×10^{-13}$
$f_7(x)$	0.014453	0	0	0	2.535341	0	0.002539
$f_8(x)$	0.000584	0.000352	0.000714	0.000397	0.009826	0.000332	0.005038

Table 2 illustrates that the WOA is particularly obvious disadvantaged when solving for the optimal value of high dimensional functions. The reason for this result is that there are three ways to update it, but only one way of obtaining the exploration, and the probability of getting it is low. Therefore the WOA is prone to falling into local optima. Compared with WOA, the algorithm used in this paper improves the convergence accuracy of the algorithm in later iterations through the choice of convergence factors and the variational strategy. Meanwhile, compared with other improved WOA, IWOA can also effectively solve the problem of algorithm falling into local extremes. It shows significant advantages over several other algorithms about convergence accuracy and stability.

3.4 IWOA-ELM Fault Diagnosis Model

In this section, the IWOA algorithm is selected to optimize the input weights and biases of the ELM model, and use the training set error rate and validation set error rate of the extremum learning machine as fitness values, as described in equation (44), in order to improve the accuracy and generalization ability of the ELM model classification. The feature set is divided into training set, validation set and testing set as the input of ELM, and the minimum fitness value F_{min} and population extreme value G_{best} and individual extreme value P_{best} are obtained. The optimal solution, individual extreme value P_{best} and population extreme value G_{best} are updated iteratively until the best fitness value is obtained or the maximum number of iterations is reached, and the optimal hidden layer weights and bias are obtained.

$$fitness = 2 - Accuracy_{train} - Accuracy_{verify} \tag{44}$$

where, $Accuracy_{train}$ and $Accuracy_{verify}$ represent the correct classification ratio of training set and validation set, respectively. The procedure of IWOA-ELM model is shown in Fig. 5.
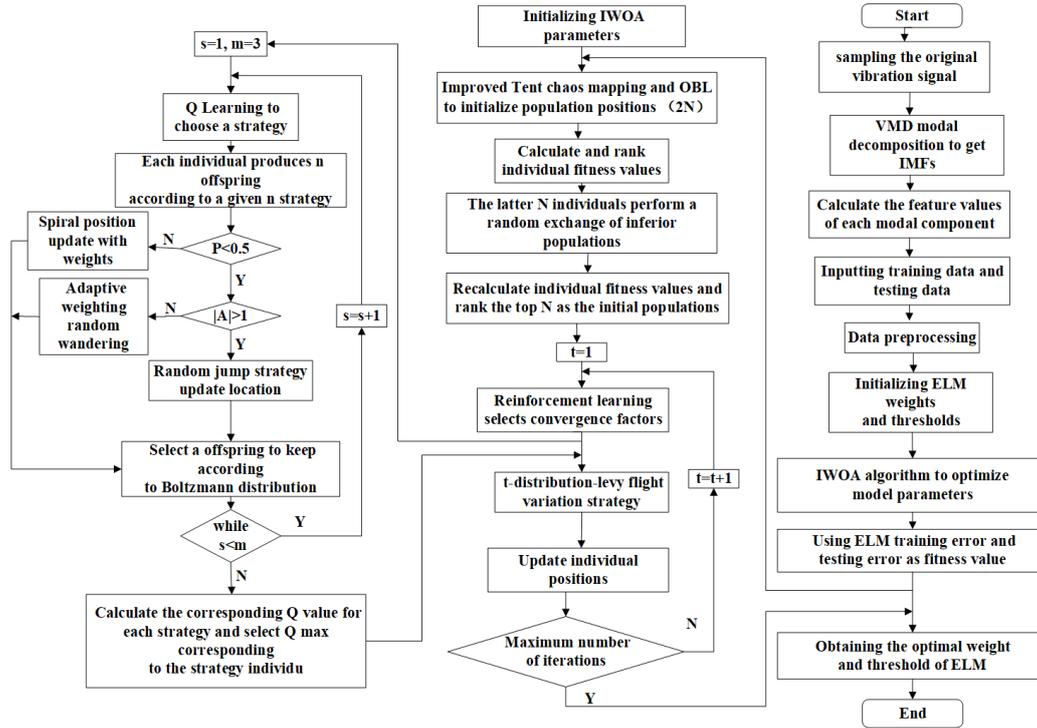
Fig. 5 Fault diagnosis flowchart

4. EXPERIMENTAL STUDY

4.1 Equipment and Data

In this paper, we used data from both normal and faulty rolling bearings, which were provided by Case Western Reserve University (CWRU). Fig. 6 illustrates the experimental simulation platform of CWRU. The experimental setup includes a 2 hp motor, a torque transducer, a dynamometer, and the corresponding control electronics. During the experiment, vibration signals were collected using an accelerometer with a sampling frequency of 12 kHz.

Case 1: Vibration signals for different fault conditions were diagnosed under varying loads, with the fault diameter in the rolling bearings kept constant. The data that collected from the 6 o'clock position at the 12K drive end serves as the primary source, including fault damage diameter of 0.36mm inner and outer rings, rolling elements, and normal bearing data. The specifics of the bearing data are provided in Table 3.

Case 2: Diagnose vibration signals under different fault diameters while keeping the load on the rolling bearing constant. The data source acquisition method remains unchanged, including the normal state of the bearing inner ring and the fault damage diameters of 0.18 mm, 0.36 mm, and 0.53 mm, respectively. The bearing data are presented in Table 4.
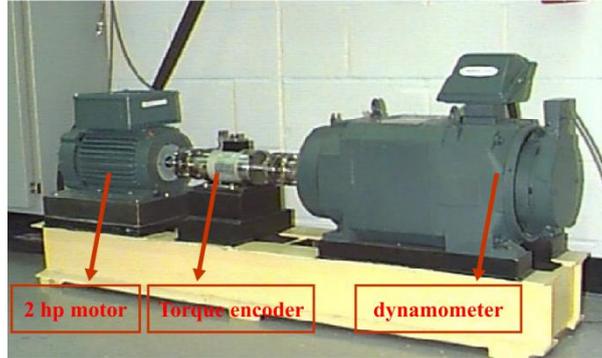
Fig. 6 Baring fault simulation test-bed

Table 3 Overview of the data sets for Case 1

Bearing status	Fault diameter(mm)	Acronym	Label
Normal	-	N	1
Ball fault	0.36	BF	2
Inner race fault	0.36	IRF	3
Outer race fault	0.36	ORF	4

Table 4 Overview of the data sets for Case 2

Bearing status	Fault diameter(mm)	Acronym	Label
	-	N	1
Normal	0.18	IRF-0.18	2
Inner race fault	0.36	IRF-0.36	3
	0.53	IRF-0.53	4

4.2 Proposed Model Evaluation

4.2.1. Experiment 1 Validation

Among the collected data, we selected 120,000 vibration signal samples and 10 seconds sampling time. The vibration signals were divided into 400 segments, each containing 300 points. The time-domain waveform of the initial signal is shown in Fig. 7. From Fig. 7, it can be seen that identifying the fault type from the time-domain waveform is challenging and can be time-consuming even when successful, highlighting the urgent need for the proposed diagnostic method. We use VMD to decompose the vibration signal of rolling bearings, with VMD decomposition K and penalty factor α values of 4 and 1989, respectively [36]. Fig. 8 illustrates the modal components along with their corresponding frequency domain diagrams for various fault types. It can be seen from Figure 8 that the energy in the low-frequency range is more pronounced for the normal state, while the inner and outer ring failures exhibit the highest energy in their respective frequency bands. The rolling element fault shows significant energy distribution in both low and high-frequency ranges. Short-duration low-frequency pulses in the vibration signal indicate shock events due to partial damage in the bearing. Additionally, the signal contains resonant components

generated by these shock events, which excite high-frequency intrinsic vibrations, modulated by their amplitude.

Finally, the IMFs constituents obtained from the signals of the four fault types are decomposed according to the VMD and the energy entropy is calculated. The characteristic sample sets are obtained by VMD decomposition and calculation of energy entropy, with 300 sets per sample set and a total of 1200 sets. From each group, 250 groups are randomly selected, of which 600 groups are used as the training sample set and 400 groups are used as the test sample set. Set the population size of IWOA N=20, the maximum value of iterations M=50, a_{max}=2, a_{min}=0, Q-learning discount factor γ=0.5, m=3, the value of nodes of ELM hidden layer is set to 100. The parameters of the ELM network weights and thresholds are optimized by the IWOA. Table 5 shows the results of diagnosing four fault states using the IWOA-ELM method. The accuracy of fault diagnosis for the training samples is 98.83%, and for the test samples, it is 97.5%. These findings indicate that the method successfully identify all four fault states of rolling bearings.
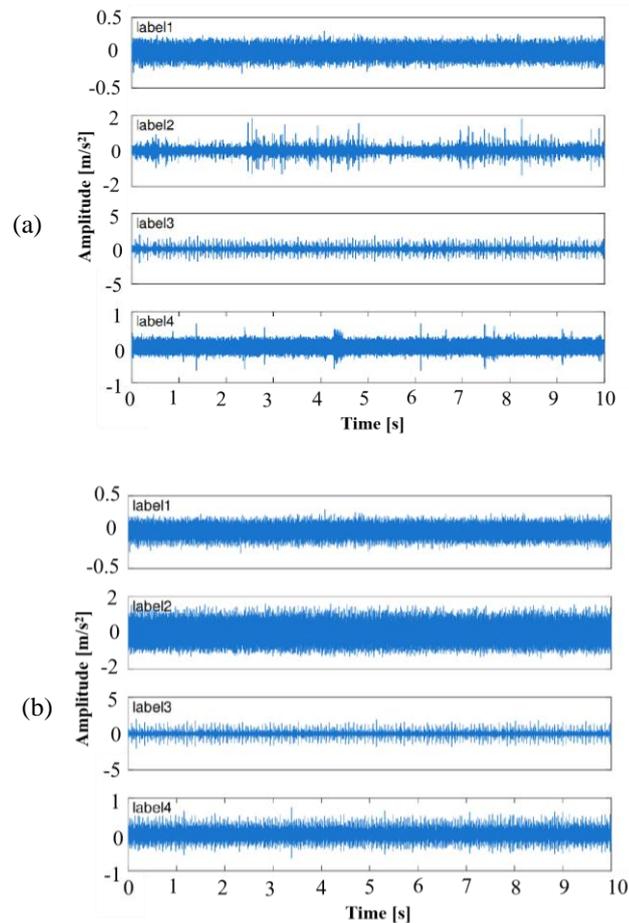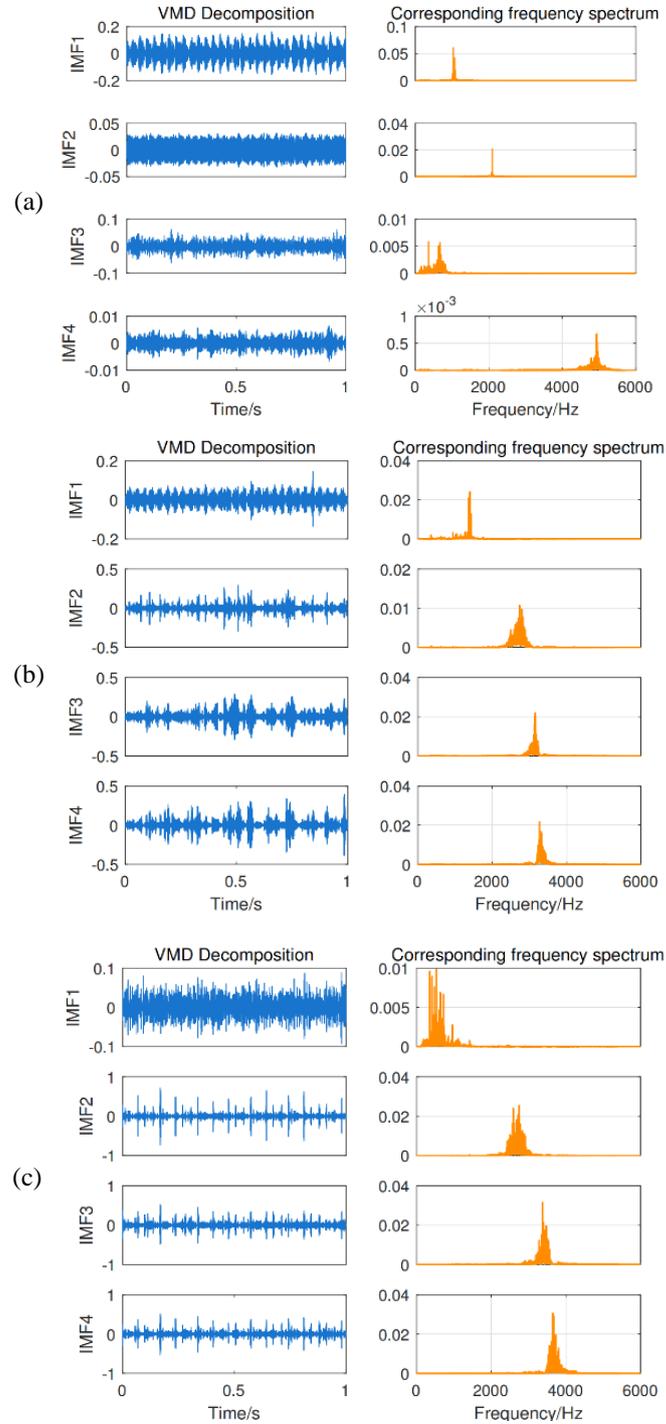

Fig. 7 The time-domain waveforms of rolling bearing fault types: (a) Case 1; (b) Case 2
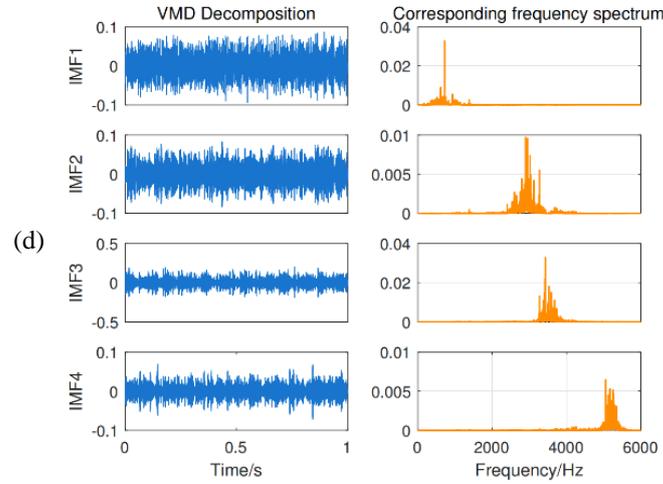
Fig. 8 Modal components and frequency distribution diagrams obtained after fault decomposition: (a) Normal; (b) Ball fault; (c) Inner race fault; (d) Outer race fault

Table 5 Diagnostic performance results for samples in diverse fault scenarios, divided into training and test groups.

Bearing status	Training Number/testing samples		BF	IRF	ORF	Training accuracy	Testing accuracy
N	150/100	149/99	1/1	0/0	0/0	99.33%	99%
BF	150/100	1/1	147/97	1/2	1/0	98%	97%
IRF	150/100	0/0	0/2	148/95	2/3	98.67%	95%
ORF	150/100	0/1	0/0	1/1	149/98	99.33%	98%

4.2.2. Experiment 2 Validation

In the data collected, we selected 120,000 vibration signal samples from normal and three bearing inner rings with different degrees of failure (sampling time: 10 seconds). The vibration signals were divided into 400 segments, each containing 300 points. Fig. 7 shows the time-domain waveform of the raw vibration signal. Decompose the vibration signal of rolling bearings using VMD, with the decomposition number K and the penalty factor α value of 4 and 1992, respectively. Fig. 9 displays the modal components and their frequency plots for inner ring of the bearing after fault decomposition at damage levels of 0.18 mm and 0.53 mm. Since the normal state and the 0.36 mm fault diameter have been discussed, we will not elaborate further here. As shown in Fig. 9, the energy associated with the inner ring failure is primarily focused on in the mid-frequency range. Furthermore, for the same section of the bearing, the amplitude corresponding to specific frequencies increases with the degree of damage. This comparison shows that as a bearing transitions from a normal state to a failed state, the primary energy in its spectrum is gradually transferred from low frequency band to middle frequency band. Subsequently, we calculated the energy entropy of IMF based on VMD decomposition of normal state signals and three types of fault vibration signals in the inner ring.
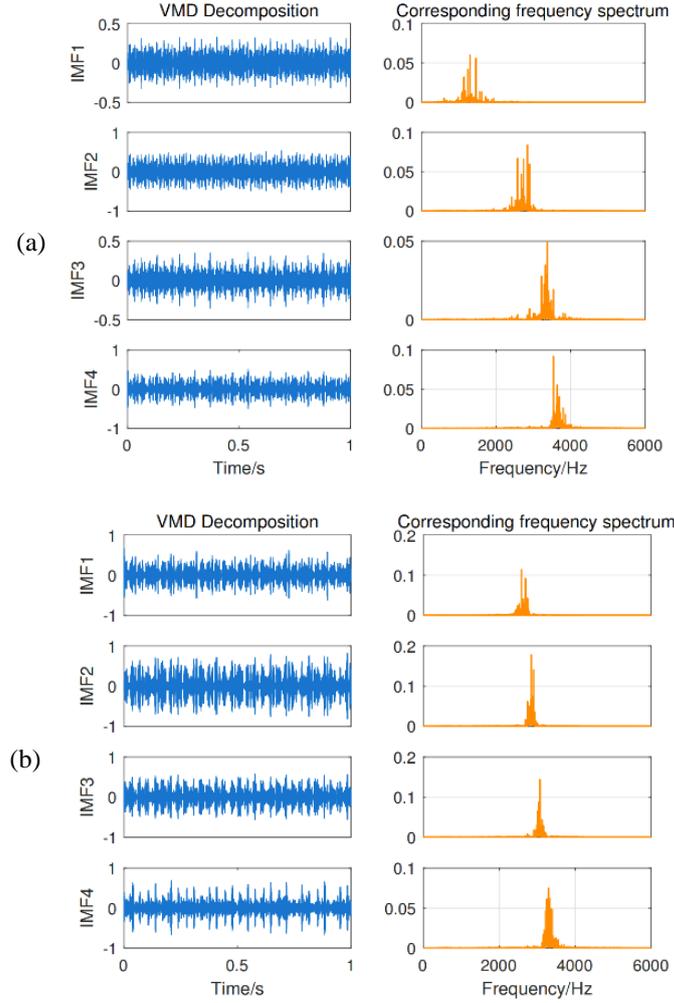
Fig. 9 Frequency domain diagrams show inner ring damage with diameters of:
(a) 0.18 mm and (b) 0.53 mm

Finally, 1200 sets of feature samples were obtained through VMD decomposition and energy entropy calculation, including 300 sets for each of the three different fault degrees and normal states. Randomly select 250 groups from each group, with 600 groups as training samples and 400 groups as testing samples. Set the population size of IWOA to N=20, the maximum iterations number M=50, a_{max}=2, a_{min}=0, Q-learning discount factor γ=0.5, m=3, the ELM activation function is a sigmoid function, and the number of hidden layer nodes is 100. The weights and thresholds of ELM are optimized through IWOA. The results of using IWOA-ELM to diagnose normal state and three different degrees of fault samples in the bearing inner race are shown in Table 6. The diagnosis accuracy reached 99% for the training samples and 96.75% for the test samples. This discrepancy may be attributed to certain failure feature parameters being less distinct at varying degrees of

failure damage, making it challenging to differentiate between states. However, with the increase of the degree of failure, salient characteristics of the characteristic parameters of the failure state become more obvious.

Table 6 Comparison diagnostic results on training and testing samples across multiple fault severity levels

Bearing status	Number of samples	N	IRF − 0.18	IRF − 0.36	IRF − 0.53	Training accuracy	Testing accuracy
N	150/100	149/99	0/0	1/0	0/1	99.33%	99%
IRF-0.18	150/100	0/1	145/90	5/9	0/0	96.67%	90%
IRF-0.36	150/100	0/0	0/0	150/100	0/0	100%	100%
IRF-0.53	150/100	0/1	0/0	0/1	150/98	100%	98%

4.3 Comparison of Different Fault Diagnosis Models

In order to verify the effectiveness of optimization algorithm in improving the accuracy of fault diagnosis, we compared it against several other methods, including SVM, ELM, PSO-ELM, GWO-ELM, and WOA-ELM. The population sizes for IWOA, WOA, and GWO were all set to 30, with a maximum of 100 iterations. The PSO algorithm was configured with a population size of 30 and executed for 100 iterations, setting the maximum particle velocity to 0.8 and the minimum particle velocity to -0.8, and the inertia weight to 0.9. For extreme learning machines, $g(x)$ was the sigmoid function and there were 100 hidden layer nodes. In the case of the SVM, the penalty coefficient was chosen as C=2 and the kernel width coefficient was set as g=0.2. The diagnostic results of different fault diagnosis methods for Case 1 and Case 2 are shown in Table 7 and Table 8, respectively. The convergence curves and comparison between predicted and actual results are shown in Fig. 10 and Fig. 11, respectively.

Comparing the results in Tables 7 and 8, the IWOA-ELM model achieved the highest fault identification accuracy for both different types and levels of bearing fault diagnosis

Table 7 Diagnosis results of different fault diagnosis methods for Case 1

Model	Training accuracy rate	Testing accuracy rate
IWOA-ELM	98.83%	97.5%
WOA-ELM	98.67%	95.5%
GWO-ELM	98.33%	95%
PSO-ELM	98.17%	95%
ELM	98.17%	90.5%
SVM	92%	91.25%

Table 8 Diagnosis results of different fault diagnosis methods for Case 2

Model	Training accuracy rate	Testing accuracy rate
IWOA-ELM	99%	96.75%
WOA-ELM	98.17%	96%
GWO-ELM	98%	94.75%
PSO-ELM	98.17%	94.75%
ELM	98%	92.75%
SVM	96.17%	94.5%

in the training set samples. This suggests that the model effectively leverages the training set, resulting in optimal training outcomes. The test set samples further validate this conclusion, as the IWOA-ELM model also achieves the highest diagnostic accuracy. Furthermore, comparing the fault diagnosis results with and without optimization shows that the optimized method consistently outperforms the unoptimized approach in terms of accuracy. In addition, while the ELM diagnostic model surpasses the SVM model on the training set, the SVM model performs better on the test set. This discrepancy highlights the differing applicability of these methods to various forms of fault diagnosis, leading to varying diagnostic results. However, this does not affect the diagnostic conclusions. The convergence curves of the different diagnostic models in Figs. 10 and 11 show that the IWOA-ELM diagnostic model gets better fitness values relative to the optimized ELM diagnostic model by other algorithms at the same number of iterations.
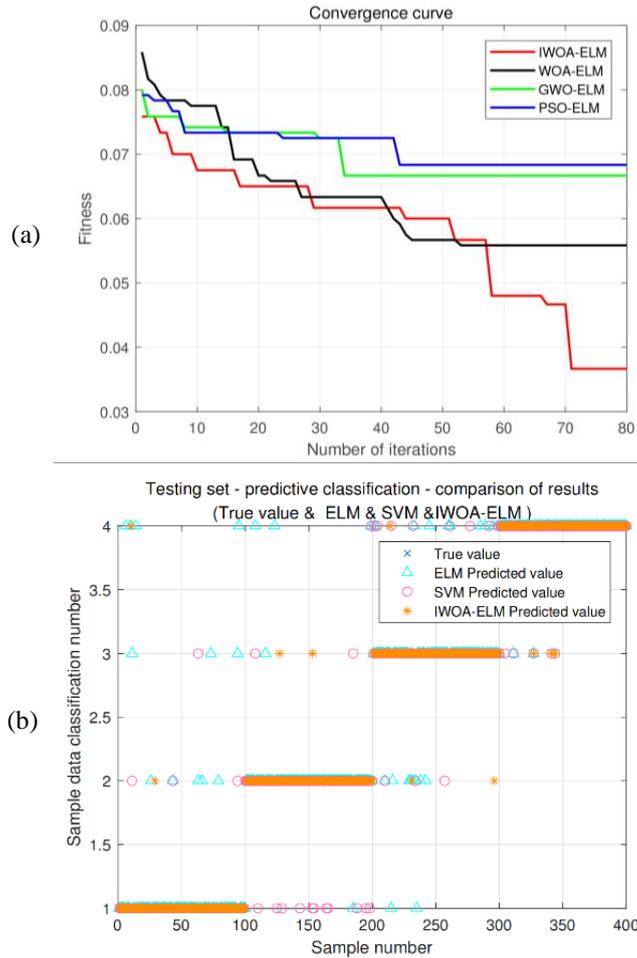
(a)

(b)

Fig. 10 Comparison of different fault diagnosis models under Case 1: (a) The convergence curves; (b) Comparison of test set prediction results
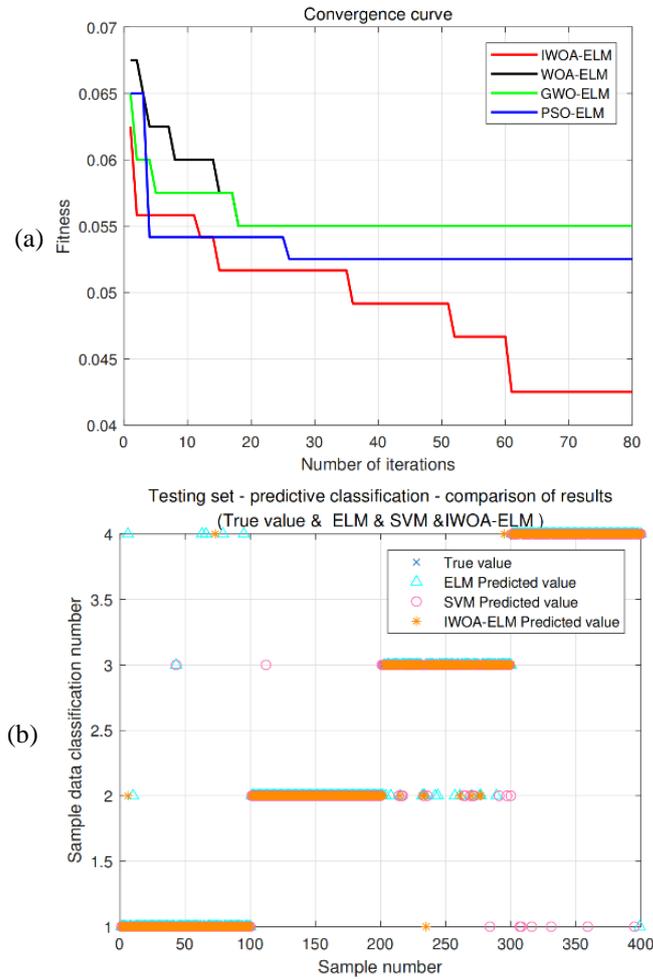
Fig. 11 Comparison of different fault diagnosis models under Case 2:
(a) The convergence curves; (b) Comparison of test set prediction results

5. CONCLUSION

The setting of parameters has a significant impact on the performance of recognition algorithms for rolling bearing fault diagnosis. In response to this issue, the paper introduces an improved whale optimization algorithm into rolling bearing fault detection and proposes a fault diagnosis model based on the IWOA-ELM algorithm. Firstly, the VMD method is used to decompose the vibration signals of rolling bearings, and the energy entropy values of IMF decomposed for different scenarios are calculated. Furthermore, multiple optimization strategies are used to improve the performance of the WOA algorithm, and it is combined with ELM to optimize the input layer weights and hidden layer thresholds. Comparing the improved algorithm with SVM, ELM, GWO-ELM, PSO-ELM, and WOA-

ELM algorithms, the diagnostic results show that the rolling bearing fault diagnosis method based on IWOA-ELM can effectively improve the convergence speed and have a higher recognition rate. The idea proposed in this paper of optimizing ELM related parameters through improved optimization algorithms to enhance the performance of rolling bearing fault diagnosis can provide some reference for future research on rolling bearing fault diagnosis.

REFERENCES

1. Rai, A., Upadhyay, S. H., 2016, *A review on signal processing techniques utilized in the fault diagnosis of rolling element bearings*, Tribology International, 96, pp. 289-306.
2. Singh, J., Azamfar, M., Li, F., Lee, J., 2020, *A systematic review of machine learning algorithms for prognostics and health management of rolling element bearings: fundamentals, concepts and applications*, Measurement Science and Technology, 32(1), 012001.
3. El-Thalji, I., Jantunen, E., 2015, *A summary of fault modelling and predictive health monitoring of rolling element bearings,* Mechanical Systems and Signal Processing, 60, pp. 252-272.
4. Kamat, P., 2024, *Fault precognition system for remaining useful life estimation in bearing systems using autoencoder-LSTM and clustering techniques*, Journal Européen des Systèmes Automatisés, 57(6), pp. 1721-1728.
5. Araste, Z., Sadighi, A., Jamimoghaddam, M., 2023, *Fault diagnosis of a centrifugal pump using electrical signature analysis and support vector machine*, Journal of Vibration Engineering & Technologies, 11(5), pp. 2057-2067.
6. Islam, M. M., Kim, J. M., 2024, *Time–frequency envelope analysis-based sub-band selection and probabilistic support vector machines for multi-fault diagnosis of low-speed bearings,* Journal of Ambient Intelligence and Humanized Computing, 15, pp. 1527–1542.
7. Shi, Q., Zhang, H., 2020, *Fault diagnosis of an autonomous vehicle with an improved SVM algorithm subject to unbalanced datasets*, IEEE Transactions on Industrial Electronics, 68(7), pp. 6248-6256.
8. Seghiour, A., Abbas, H. A., Chouder, A., Rabhi, A., 2023, *Deep learning method based on autoencoder neural network applied to faults detection and diagnosis of photovoltaic system*, Simulation Modelling Practice and Theory, 123, 102704.
9. Ghorvei, M., Kavianpour, M., Beheshti, M. T., Ramezani, A., 2023, *Spatial graph convolutional neural network via structured subdomain adaptation and domain adversarial learning for bearing fault diagnosis*, Neurocomputing, 517, pp. 44-61.
10. Lee, C. Y., Ou, H. Y., 2021, *Induction motor multiclass fault diagnosis based on mean impact value and PSO-BPNN*, Symmetry, 13(1), 104.
11. Thomas, J. B., Chaudhari, S. G., Shihabudheen, K. V., Verma, N. K., 2023, *CNN-based transformer model for fault detection in power system networks*, IEEE Transactions on Instrumentation and Measurement, 72, 2504210.
12. Huang, T., Zhang, Q., Tang, X., Zhao, S., Lu, X., 2022, *A novel fault diagnosis method based on CNN and LSTM and its application in fault diagnosis for complex systems*, Artificial Intelligence Review, 55(2), pp. 1289-1315.
13. Janssens, O., Slavkovikj, V., Vervisch, B., Stockman, K., Loccufier, M., Verstockt, S., Van Hoecke, S., 2016, *Convolutional neural network based fault detection for rotating machinery*, Journal of Sound and Vibration, 377, pp. 331-345.
14. Huang, G. B., Zhu, Q. Y., Siew, C. K., 2006, *Extreme learning machine: Theory and applications*, Neurocomputing, 70(1-3), pp. 489-501.

15. Zhao, Y. P., Chen, Y. B., 2022, *Extreme learning machine based transfer learning for aero engine fault diagnosis*, Aerospace Science and Technology, 121, 107311.
16. Jayapal, P. K., Muvva, V. R., Desanamukula, V. S., 2023, *Stacked Extreme Learning Machine with Horse Herd Optimization: A Methodology for Traffic Sign Recognition in Advanced Driver Assistance Systems*, Mechatronics and Intelligent Transportation Systems, 2(3), pp. 131-145.
17. Wang, J., Zhang, Y., Zhang, F., Li, W., Lv, S., Jiang, M., Jia, L., 2021, *Accuracy-improved bearing fault diagnosis method based on AVMD theory and AWPSO-ELM model*, Measurement, 181, 109666.
18. Zhang, X., Yang, Z., Cao, F., Cao, J., Wang, M., Cai, N., 2020, *Conditioning optimization of extreme learning machine by multitask beetle antennae swarm algorithm*, Memetic Computing, 12, pp. 151-164.
19. Wu, X., Zhang, S. E. N., Xiao, W., Yin, Y., 2019, *The exploration/exploitation tradeoff in whale optimization algorithm*, IEEE Access, 7, pp. 125919-125928.
20. Shim, E. S., Kim, H. D., Lee, S. 2024, *Optimal multi-impulse elliptic rendezvous using stagnation escaping whale optimization algorithm*, International Journal of Aeronautical and Space Sciences, 25, pp. 1092-1104.
21. Abdel-Basset, M., Manogaran, G., El-Shahat, D., Mirjalili, S., 2018, *A hybrid whale optimization algorithm based on local search strategy for the permutation flow shop scheduling problem*, Future Generation Computer Systems, 85, pp. 129-145.
22. Dragomiretskiy, K., Zosso, D., 2013, *Variational mode decomposition*, IEEE Transactions on Signal Processing, 62(3), pp. 531-544.
23. Yu, Y., Yu, D. J., Cheng, J. S., 2006, *A roller bearing fault diagnosis method based on EMD energy entropy and ANN*, Journal of sound and vibration, 294(1-2), pp. 269-277.
24. Zhang, N., Zhao, Z. D., Bao, X. A., Qian, J., Wu, B., 2020, *Gravitational search algorithm based on improved Tent chaos*, Control and Decision, 35(4), pp. 893-900.
25. Jiang, T., Zhang, C., Zhu, H., Gu, J., Deng, G., 2018, *Energy-efficient scheduling for a job shop using an improved whale optimization algorithm*, Mathematics, 6(11), 220.
26. Lv, H., Feng, Z., Wang, X., Zhou, W., Chen, B., 2021, *Structural damage identification based on hybrid whale annealing algorithm and sparse regularization*, J. Vib. Shock, 40, pp. 85-91.
27. Teng, X., Luo, Y., Zheng, T., Zhang, X., 2022, *An improved whale optimization algorithm based on aggregation potential energy for QoS-driven web service composition*, Wireless Communications and Mobile Computing, 2022(1), 9741278.
28. Zhang, H. X., Lu, J., 2008, *An adaptive evolutionary programming algorithm based on Q learning*, Acta Automatica Sinica, 34(7), pp. 819-822.
29. Motamarri, R., Bhookya, N., 2020, *JAYA algorithm based on Lévy flight for global MPPT under partial shading in photovoltaic system*, IEEE Journal of Emerging and Selected Topics in Power Electronics, 9(4), pp. 4979-4991.
30. Mantegna, R. N., 1994, *Fast, accurate algorithm for numerical simulation of Levy stable stochastic processes*, Physical Review E, 49(5), 4677.
31. Chen, H., Li, W., Yang, X., 2020, *A whale optimization algorithm with chaos mechanism based on quasi-opposition for global optimization problems*, Expert Systems with Applications, 158, 113612.
32. Liu, J., Shi, J., Hao, F., Dai, M., 2022, *A reinforced exploration mechanism whale optimization algorithm for continuous optimization problems*, Mathematics and Computers in Simulation, 201, pp. 23-48.
33. Mirjalili, S., Mirjalili, S. M., Lewis, A., 2014, *Grey wolf optimizer*, Advances in engineering software, 69, pp. 46-61.
34. Saremi, S., Mirjalili, S., Lewis, A., 2017, *Grasshopper optimisation algorithm: theory and application*, Advances in Engineering Software, 105, pp. 30-47.
35. Xue, J., Shen, B., 2020, *A novel swarm intelligence optimization approach: Sparrow search algorithm*, Systems science & Control Engineering, 8(1), pp. 22-34.
36. He, C., Wu, T., Gu, R., Jin, Z., Ma, R., Qu, H., 2021, *Rolling bearing fault diagnosis based on composite multiscale permutation entropy and reverse cognitive fruit fly optimization algorithm–extreme learning machine*, Measurement, 173, 108636.