IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 18 February 2024, accepted 29 February 2024, date of publication 4 March 2024, date of current version 14 March 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3373554

RESEARCH ARTICLE

Intelligent Diagnosis of Rolling Bearing Based on ICEEMDAN-WTD of Noise Reduction and Multi-Strategy Fusion Optimization SCNs

KUN LI, HAO WU, XINMING LIU, JIANING XU, AND YING HAN
Faculty of Electrical and Control Engineering, Liaoning Technical University, Huludao 125105, China

Corresponding author: Ying Han (hyfengyan@163.com)

This work was supported in part by the Basic Scientific Research Project of Education Department of Liaoning Province (Key Research) under Grant JYTZD2023074, and in part by the National Natural Science Foundation of China under Grant 62203197.

ABSTRACT Aiming at the problem of noise interference leading to poor fault diagnosis effect of rolling bearing, a two-stage signal noise reduction method based on multi-strategy coati optimization algorithm (MFCOA) optimized ICEEMDAN combined with wavelet threshold denoising (ICEEMDAN-WTD) is proposed. The MFCOA-optimized stochastic configured networks (MFCOA-SCNs) is used for fault type identification. Firstly, ICEEMDAN is used to decompose noisy signals into several IMF signals, and then a wavelet threshold denoising method is used to process IMF signals whose arrangement entropy is lower than the pre-set value to obtain reconstructed signals. In this process, the ICEEMDAN decomposition is affected by the number of white noise additions and the degree of addition. Therefore, MFCOA is introduced to optimize the parameters to ensure the effect of signal noise reduction. Secondly, ICEEMDAN is applied to the reconstructed signal, and the entropy of IMF signal samples of each order is calculated to obtain the fault characteristics in different simultaneous frequency domains. Finally, the scale factor λ and regularization coefficient r in SCNs are optimized by MFCOA, and the sample entropy feature is used as the input of the MFCOA-SCNs model to achieve fault identification. Several simulation experiments have verified that the proposed method can effectively reduce the impact of noise in bearing signals, and the experimental results are better than other methods, and the diagnostic accuracy is better.

INDEX TERMS Coati optimization algorithm, ICCEMDAN, wavelet threshold denoising, stochastic configuration networks, parameter optimization.

NOMENCLATURE

EMD	Empirical Mode Decomposition.
ICEEMDAN	Improved Complete Ensemble EMD.
WTD	Wavelet Threshold Denoising.
VMD	Variational Mode Decomposition.
WPD	Wavelet Packet Decomposition.
SVM	Support Vector Machine.
LSSVM	Least Squares Support Vector Machine.
RF	Random Forest Algorithm.
DBN	Deep Belief Networks.
SCNs	Stochastic Configuration Networks.
COA	Coati Optimization Algorithm.
MFCOA	Multi-strategy Fusion Coati Optimization Algorithm.
GA	Genetic Algorithm.
PSO	Particle Swarm Optimization.
HBA	Honey Badger algorithm.
SSA	Sparrow Search Algorithm.
WOA	Whale Optimization Algorithm.
AO	AquilaOptimizer.
BWO	Beluga Whale Optimization.
GJO	Golden Jackal Optimization.
BOA	Butterfly Optimization Algorithm.
SE	Sample Entropy.
PE	Permutation Entropy.
FE	Fuzzy Entropy.
Acc	Accuracy.
Pre	Precision.

The associate editor coordinating the review of this manuscript and approving it for publication was Lei Shu.

I. INTRODUCTION

A. BACKGROUND RESEARCH

Rolling bearings are the basic components of rotating machinery and equipment such as motors, and their health status directly determines the safety of the equipment. When there are faults such as fractures, they are often accompanied by serious personal safety threats and cause huge economic losses. Therefore, it is crucial to judge the health status of bearings in time in the industrial production [1].

In the current research on bearing fault diagnosis, how to improve the accuracy and robustness of fault recognition is still an important topic for many scholars. Due to the complexity of industrial application scenarios of rolling bearings, the mixed noise in the vibration signal will lead to the submerged fault features, and it is difficult to obtain the ideal diagnosis effect. Therefore, the noise reduction and reconstruction process of the signal is particularly important in the fault diagnosis process. In addition, the selection of a diagnostic model is also an important part of the diagnosis process, and the feature learning ability of the model largely determines the accuracy of diagnosis. Based on the current research work, further works are carried out in this paper.

B. RELATED WORK

1) SIGNAL DECOMPOSITION METHOD

At present, the bearing fault diagnosis method based on signal decomposition is still the mainstream method. Common signal decomposition methods mainly include VMD [2], EMD [3], WPD [4], etc. The bearing early fault features are easy to be submerged by noise. The signal decomposition method can effectively separate high and low-frequency information to realize noise reduction reconstruction and time-frequency domain feature extraction. However, the above methods have problems such as mode aliasing, and the decomposition ability is limited by parameter settings, so it is difficult to achieve ideal results when dealing with strong noise problems [5].

To solve the above problems, many scholars have carried out related research. Liu et al. [6] determined the optimal number of decomposition modes according to the difference of correlation function wavelet packet coefficients, and obtained the optimal wavelet basis with logarithmic energy entropy as the cost function to avoid the performance limitation of wavelet packet decomposition due to manual parameter setting. Lu et al. [7] proposed a Mirror Extension Variational Modal Decomposition (MVMD) to suppress end effects. Marcelo et al. [8] proposed an Improved Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (ICEEMDAN). The IMF component was defined as the difference between the residual of the previous decomposition stage and the local mean, and the way of adding white noise was changed, which effectively reduced the pseudo mode and noise residual problems existing in CEEMDAN. However, its parameter selection still interferes with the decomposition effect to some extent.

2) COMBINED WITH SWARM INTELLIGENCE OPTIMIZATION MODELING METHOD

In the process of bearing fault diagnosis model establishment, parameter interference will lead to unsatisfactory calculation accuracy. To reduce the subjective error caused by manual setting, the modeling method based on swarm intelligence optimization has attracted attentions. Zhu et al. [9] considered the influence of weight and threshold in the BP neural network on its performance, determined the optimal parameters for experimental data by GA, and proved that this method could effectively improve diagnostic accuracy. Tang et al. [10] developed Particle Swarm optimization and Random forest classifier (PSO-RF) to optimize the parameters of random forest (RF) through particle swarm optimization to improve the accuracy of bearing fault diagnosis. Guo et al. [11] optimized the parameters of SVM by CS algorithm, obtained the SVM classification model with optimized parameters, and applied it to the bearing fault diagnosis problem to verify its effectiveness. Xu et al. [12] used QPSO to optimize SVM parameters, ensuring the classification performance of SVM. Despite the strong anti-interference ability of SCNs, the generalization ability of SCNs is still affected by the unsatisfactory parameter configuration. Therefore, Zhang et al. [13] proposed a random configuration network model based on chaotic sparrow optimization (CSSA-SCNs) to improve the performance of SCNs.

The drawback of such methods is that the selection of model parameters depends heavily on the optimization algorithm. If the algorithm has poor computational performance, it may lead to negative optimization of the model. Therefore, choosing the appropriate optimization algorithm is the core of this method.

3) DIAGNOSIS METHOD BASED ON NONLINEAR DYNAMICS

Feature extraction is an important part of the fault recognition process, and the fault diagnosis method based on nonlinear dynamics has been widely concerned by scholars.

Li et al. [14] proposed a nonlinear dynamic index combining variable step size multi-scale (RCVM) and multi-mapping dispersion entropy (MDE), and verified that this method can effectively extract depth information at different time scales and has strong robustness through a variety of signals. Aiming at the problem that the local fault pulse is too weak to be easily submerged by noise, Li et al. [15] proposed a method that combines smooth sparse low-rank matrix (SSLRM) with asymmetric and singular value decomposition (SVD) penalty regularization, which effectively reduces the noise interference and ensures the operation speed. Since SloEn can only reflect the information of signals at a single scale, Li et al. [16] proposed variable step size multi-scale single threshold SloEn (VSM-StSloEn) and introduced snake optimizer (SO) for parameter selection to avoid the influence of threshold values. Finally, the excellent engineering practicability of the method was verified through experiments. In view of the limitation that SLoEn is only suitable for

single-scale, single-channel time series analysis, Li et al. [17] developed multivariate and multi-scale SloEn, and optimized the threshold through the hippocampal optimizer. The data verified that this method has a better feature extraction effect compared with other methods.

4) IMPROVED VERSION OF SWARM INTELLIGENCE ALGORITHM

Bionic algorithms such as PSO [18], SSA [19], WOA [20], and HBA [21] have been widely used to solve engineering problems because of their simple principle and high calculation accuracy. However, the swarm intelligence optimization algorithm is affected by the search method and diversity consumption, which often leads to local stagnation and convergence rate decay in the later iteration. In this problem, many scholars have designed improved versions for each algorithm.

Liu et al. [22] introduced sinusoidal chaos theory and differential evolution strategy into WOA to enhance its global search ability and efficiency. Fei et al. [23] introduced the update strategy of the HBA mining stage to enhance the global exploration ability, and used the disturbance operator and Levy flight strategy to disturb the ability of population individuals to jump out of local areas, aiming at the problem of slow convergence speed and easy local optimization of Sparrow algorithm. Qu et al. [24] used Singer mapping to generate the initial population of PSO to enhance diversity, and introduced the Levy flight method to make it obtain the ability to jump out of local areas. Hou et al. [25] promoted the diversity of the grey wolf population through Tent mapping, and used the nonlinear convergence factor based on the Gaussian distribution change curve to balance the search stage. Finally, the dynamic proportional weighting strategy was used to enhance the convergence efficiency of GWO.

Dehghani et al. [26] proposed coati optimization algorithm (COA) based on the behavior of coati hunting and escaping from natural enemies. Compared with other popular algorithms, this algorithm has the advantages of strong robustness and fewer parameter settings, although its convergence efficiency has been verified. However, the problems of individual assimilation behavior and diversity loss also exist, so COA still has room for improvement.

C. OUR CONTRIBUTIONS

Considering the complexity of the working environment of the bearing, the collected vibration signal is usually unstable, which limits the ability of wavelet transform, and makes it difficult for a single wavelet threshold method to process the noisy signal overlapping with useful information, resulting in difficult to ensure the noise reduction effect. Secondly, the mode aliasing phenomenon may lead to the overlay of time-frequency domain features, which makes it difficult to extract the active components. In addition, there are some hyper-parameters in the diagnosis model, if they are set manually, it may cause subjective errors and affect diagnostic accuracy. To solve the above problems, this paper proposes a multi-task optimization bearing fault diagnosis method based on multi-strategy fusion of COA, aiming at achieving high-precision bearing fault diagnosis under noise interference.

Our contributions are summarized as follows:

(1) A multi-strategy fusion coati optimization algorithm (MFCOA) is proposed, which combines quasi-opposition-based learning, enclosing predation strategy, shock transfer factor, and survival selection strategy to improve the shortcomings of COA such as too high population assimilation in the later iteration and limited development caused by unbalanced search.

(2) Aiming at the problem of noise interference in bearing signal, the two-stage noise reduction method of ICEEMDAN-WTD is introduced, and the parameters are optimized by MFCOA to ensure the decomposition effect and help to eliminate noise.

(3) In order to improve the generalization of SCNs, MFCOA is used to optimize the scale factor λ and the regularization coefficient r, and the MFCOA-SCNS model is established. The application ability of MFCOA-SCNS is verified by a bearing fault diagnosis example.

The rest of this paper is organized as follows. In Section II, the basic algorithm is introduced; Section III introduces the two-stage noise reduction methods of MFCOA and MFCOA-ICEEMDAN-WTD, and finally puts forward an overall fault diagnosis model; in Section IV, the performance test of benchmark functions and the Wilcoxon signed-rank test verify the superiority of MFCOA; in Section V, the superiority of the proposed fault diagnosis model is proved by different experiments; Section VI concludes the main work of this paper.

II. BASIC ALGORITHM

A. COA

COA simulates the group predation and natural enemy avoidance behavior of coatis, and the algorithm realizes the serial cooperation between populations, so it has the advantages of efficient convergence and high accuracy in solving optimization problems [26].

In the initialization phase of the algorithm, individual coati in the population is randomly assigned positions as candidate solutions. The mathematical model is expressed by

$$X_{i,j} = R_L + rand \times (R_H - R_L) \tag{1}$$

where $X_{i,j}$ represent the jth coatis individual of generation i; R_L and R_H denote the upper and lower boundaries of the solution space, respectively; *rand* represents a random number in [0, 1].

In the global development stage, half of the coatis will gather around the current best individual to intimidate the green iguana. This process is expressed by

$$X_{i,j}^{P1} = X_{i,j} + rand \times (X_{best} - I \times X_{i,j}), i = 1, 2, \ldots, N/2 \tag{2}$$

When the green iguana is scared, it randomly falls into the solution space, and the remaining raccoons hunt quickly. The mathematical model of this behavior is defined by

$$I_r = R_L + rand \times (R_H - R_L) \tag{3}$$

$$X_{i,j}^{P1} = \begin{cases} X_{i,j} + rand \times (I_r - I \times X_{i,j}), & F_{Ir} < F_i \\ X_{i,j} + rand \times (X_{i,j} - I_r), & else, \end{cases}$$
$$i = N/2 + 1, N/2 + 2, \ldots, N \tag{4}$$

where $X_{i,j}$ represents the updated individual position; X_{best} is the current best position; I represents a random number of 1 or 2; I_r indicates the random drop location of the green iguana; N is the number of cocoon individuals set.

If the updated position achieves fitness growth, it will be retained; otherwise, the updated position will be abandoned. This process can be expressed by

$$X_{i,j} = \begin{cases} X_{i,j}^{P1}, & F_{X_{i,j}^{P1}} < F_{X_{i,j}} \\ X_{i,j}, & else \end{cases} \tag{5}$$

After the hunting process of the coatis, the natural enemy may attack, so the coatis will quickly escape after feeling the crisis. This behavior can be expressed by

$$X_{i,j}^{p2} = X_{i,j} + (1 - 2 \times rand) \times (R_L/t + rand \times (R_H/t - R_L/t)) \tag{6}$$

where t represents the current number of iterations.

B. WAVELET THRESHOLD DENOISING

Wavelet threshold denoising is a common signal processing method, which uses wavelet transform to deconstructure the original sequence into signals with different frequencies, and uses the frequency characteristics of the original signal and the noise signal to remove the high-frequency noise according to the set threshold, to reduce the interference information in the original signal [27]. There are two functions in the traditional wavelet threshold method: hard threshold function and soft threshold function, which are given by

$$\hat{\omega}_{j,k} = \begin{cases} \omega_{j,k}, & |\omega_{j,k}| \geqslant TH \\ 0, & |\omega_{j,k}| < TH \end{cases} \tag{7}$$

$$\hat{\omega}_{j,k} = \begin{cases} sgn(\omega_{j,k})(|\omega_{j,k}| - TH), & |\omega_{j,k}| \geqslant TH \\ 0, & |\omega_{j,k}| < TH \end{cases} \tag{8}$$

where $\omega_{j,k}$ represents the wavelet coefficients of the signal to be processed after wavelet transform. TH is the set threshold; $\hat{\omega}_{j,k}$ is the estimated value of wavelet decomposition coefficient.

The difference between the two methods is that the hard threshold process retains the wavelet coefficients whose absolute value is not less than the set value, and the wavelet coefficients less than the threshold are set to 0. Soft threshold similarly sets the wavelet coefficients less than the threshold to 0 but performs sgn calculation on the wavelet coefficients greater than or equal to the threshold.

It is easy to see that the effect of the wavelet threshold denoising method largely depends on the choice of TH. When TH is too large, effective information may be filtered out, and when TH is too small, the denoising effect cannot be guaranteed.

$$TH = \sigma\sqrt{2\lg L_x} \tag{9}$$

where L_x represents the length of signal x,σ = median(ωj, k)/0.6745.

C. SCNs

Stochastic Configuration Networks (SCNs) was proposed by Wang et al. [28]. The basic structure of SCNs is shown in FIGURE 1.

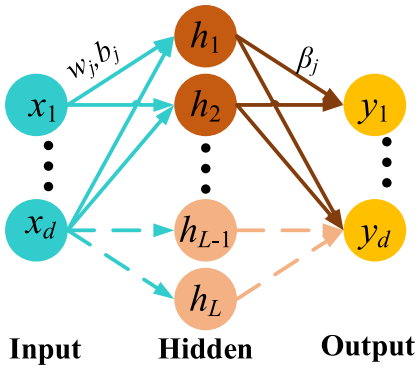


FIGURE 1. SCNs structure diagram.

Input features $X = [x_1, x_2, \ldots, x_n]$, where $x_i = [x_{i,1}, x_{i,2}, \ldots, x_{i,d}]$, $i = 1,2,\ldots, n$, $x_i \in R^d$; Setting labels $O = [o_1, o_2, \ldots, o_n]$, where $o_i = [o_{i,1}, o_{i,2}, \ldots, o_{i,m}]$, $o_i \in R^m$.

Step 1 Suppose that L-1 nodes have been created, and the current network model output is expressed by

$$f_{L-1}(X) = \sum_{j=1}^{L-1} \beta_j \varphi(w_j^{\mathrm{T}} X + b_j), f_0 = 0 \tag{10}$$

where β_j represents the output weight of the jth hidden layer node, $j = 1,2,\ldots, L_{max}$; φ denotes the activation function.

Step 2 Calculate the SCN residual:

$$E_{L-1} = Y - f_{L-1}(X) = [E_{L-1,1}(X), E_{L-1,2}(X), \ldots, E_{L-1,m}(X)] \tag{11}$$

Step 3 If $\|E_{L-1}\|$ does not reach the set tolerance tol or has not been built to L_{max} node, add node L, node parameters are randomly generated by Equations (12)-(13), and the supervised mechanism of Equation (14) retains w_j and b_j that maximize Equation (15) as node parameters:

$$w_j = \lambda \times (2 \times rand(d, L) - 1) \tag{12}$$

$$b_j = \lambda \times (2 \times rand(1, L) - 1) \tag{13}$$

$$h_L = [\varphi(w_L^{\mathrm{T}} x_1 + b_L), \ldots, \varphi(w_L^{\mathrm{T}} x_N + b_L)]^{\mathrm{T}} \in R^N \tag{14}$$

$$\xi_{L,j} = \frac{\left(E_{L-1,j}^T h_L\right)^2}{h_L^T h_L} - (1 - r - u_L) \parallel E_{L-1,j} \parallel_F^2 \tag{15}$$

$$\xi_L = \sum_{j=1}^{m} \xi_{L,j} \geq 0 \tag{16}$$

where T_{max} represents the maximum number of candidate nodes; Regularization coefficient $r \in [0, 1]$; The non-complex $u_L \leq 1-r$ and the u_L tends to 0 as L tends to infinity.

Step 4 Output the weight β^* according to

$$\beta^* = \arg\min_{\beta} \parallel H_L\beta - Y \parallel_F^2 = H_L^+ Y \tag{17}$$

where $H_L = [h_1, h_2, \ldots, h_L]$; H_L^+ denotes the Moore-Penrose generalized inverse matrix of H_L; $|| \cdot ||_2^F$ denotes the 2-norm.

Step 5 Output result f, there is:

$$f = H\beta^* \tag{18}$$

III. IMPROVED ALGORITHM AND PROPOSED METHOD

A. MFCOA

According to the basic COA, it can be known that the random generation of the coatis population cannot guarantee the quality of individuals, which affects the convergence efficiency of the algorithm. The division of labor and cooperation in the attack phase strengthened the information exchange between populations, but it is difficult to control the balanced search of the search space, which led to invalid computation and limited vision spot phenomenon. In the avoidance phase, the assimilation degree is too high due to the population following behavior, which makes the algorithm appear local stagnation. Based on the above problems, the COA is improved in this paper.

1) QOBL

To enhance the quality of the initial population and improve the diversity problem, this paper introduces the Quasi-opposition-based learning strategy (QOBL) into COA to improve the quality of the initial population. In the iterative process, the QOBL strategy generates mirror solutions far away from the local optimal solution according to Equations (19)-(20) to help the initial population achieve space traversal to find the optimal solution [29]:

$$x_{OBL} = R_H + R_L - x \tag{19}$$

$$x_{QOBL} = rand \times (\frac{R_H + R_L}{2}, x_{OBL}) \tag{20}$$

2) SHOCK TRANSFER FACTOR AND ENVELOPING PREDATION STRATEGY

After achieving the improvement of individual quality, high-quality coatis perform predation behavior through serial cooperation. However, the lack of regional convergence control behavior in the two-stage search process makes the algorithm unable to effectively traverse in the early stage and easy to skip the optimal position in the later stage. Therefore, this paper introduces the triangle walking strategy [30] and the shock transfer factor to modify the predator coati Equation (4) to ensure the balanced search of the algorithm and reduce the limited vision spots to enhance the development ability of the algorithm. The improved predation formula is given by

$$s = \begin{cases} 1\text{-}1.15 \times (\frac{t}{0.8t_{max}})^2 \times \cos(\pi(1 - \frac{1.4t}{0.8t_{max}})/2), \\ \quad t \leq 0.8t_{max} \\ 1\text{-}1.15 \times \cos(-0.2\pi) + 0.7 \times (\frac{t}{t_{max}})^{10} \\ \quad \times \cos(0.15\pi \times t), else \end{cases} \tag{21}$$

$$L_1 = X_{i,j} - I_r \tag{22}$$

$$\overrightarrow{L_2} = rand \times \overrightarrow{L_1} \tag{23}$$

$$L = L_1^2 + L_2^2 - 2 \times L_1 \times L_2 \cos(2 \times \pi \times rand) \tag{24}$$

$$X_{i,j}^{P1} = \begin{cases} X_{i,j} + rand \times (I_r - I \times X_{i,j}), & F_{Ir} < F_i \\ X_{i,j} + 2 \times s \times rand \times L, & else, \end{cases}$$
$$i = N/2 + 1, N/2 + 2, \ldots, N \tag{25}$$

where s is the proposed shock transfer factor and its curve is shown in FIGURE 2. In the early iteration stage of $t \leq 0.8t_{max}$, s is reduced generation by generation to ensure the limited vision spot development in the early stage and the rapid convergence in the middle and late stages. In the late iteration of $t > 0.8\ t_{max}$, the algorithm is perturbed by the transmission of vibration wave to avoid local stagnation.

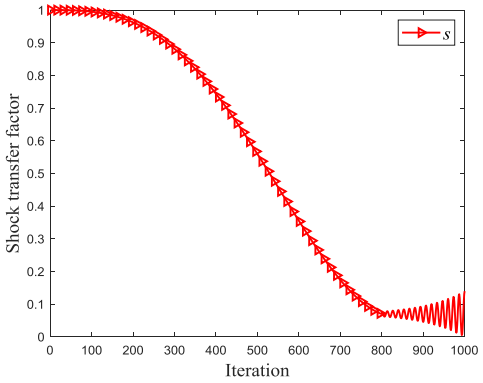


FIGURE 2. Shock transfer factor curve.

3) SURVIVAL SELECTION STRATEGY

In order to further avoid the problem of excessive individual assimilation in the later stage of iteration, this paper uses the Cauchy-Gaussian mutation operator to generate the mirror optimal solution X_{C-G} according to the current optimal solution X_{best} to promote the algorithm to get rid of the local obstacle. This process is as follows:

$$X_{C-G} = X_{best}(1 + (\alpha \times Cauchy(0, \sigma^2) + \beta \times Gauss(0, \sigma^2)) \tag{26}$$

$$\begin{cases} \alpha = (t/t_{max})^2 \\ \beta = 1 - \alpha \end{cases} \tag{27}$$

where $Cauchy(0,\sigma^2)$ and $\text{Gauss}(0,\sigma^2)$ represent random numbers satisfying the *Cauchy* and *Gaussian* distributions, respectively; σ^2 is the standard deviation.

However, after the introduction of mutation strategy, invalid operation may occur, resulting in increased time cost. Therefore, the survival rate *su* of DOA [31] is used for survival selection in this paper. If the survival rate is lower than the dynamic probability p, the individual mutation operation will be carried out. The survival rate *su* is calculated as follows:

$$\begin{cases} su = \dfrac{fit_{max} - fit(i)}{fit_{max} - fit_{min}}, & fit_{max} = fit_{min} \\ su = 1, & else \end{cases} \tag{28}$$

$$p = 0.4 + 0.6 \times \frac{2t}{t + t_{max}} \tag{29}$$

4) TIME COMPLEXITY ANALYSIS

Suppose that the number of populations is N, the maximum number of iterations is T, and the number of decision variables of the problem to be optimized is D.

For basic COA:

The time complexity of generating the initial population is $O_1(N \times D)$, the time complexity of attack and predation stages is $O_2(0.5N \times D \times T)$, and the time complexity of avoidance stage is $O_3(N \times D \times T)$, so the time complexity of COA is $O_1 + 2O_2 + O_3 = O(N \times D \times T)$.

For MFCOA:

The time complexity of generating the QOBL population is $O_4(\text{N}\times\text{D}\times\text{T})$, the time complexity of the attack and predation stage after introducing the shock transfer factor is $O_5(0.5\text{N}\times\text{T}\times\text{D})$, and the time complexity of the mutation update stage is $O_6(\text{N}\times\text{T})$, because the survival rate and dynamic probability need to be calculated. In summary, the time complexity of MFCOA is as follows: $O_1 + O_4 + 2O_5 + O_3 + O_6 = \text{O}(N \times D \times T)$, which is the same as COA.

The flow chart of MFCOA is shown in FIGURE 3, and the pseudocode of the algorithm is Algorithm 1. The initial population of the algorithm is generated according to Equation.(1) and Equation.(20), the update mode of the attack stage is referred to Equation.(2) and Equation.(25), and the individual variation is updated according to Equations.(26)-(29).

Algorithm 1 Pseudo-Code of MFCOA

Input: Population number N, Dimension D, Search space $[R_H, R_L]$, Maximum number of iterations $t_{\max}$
Output: The global optimal position X_{best}, Best fitness fit_{best}
(1) **Start**
(2) Parameter setting.
(3) The initial population position is obtained according to Eq.(1).
(4) **While** $t<t_{\max}$ **do**
(5) The reverse population is generated according to Eq.(20).
(6) Keep the current optimal position X_{best} and the best fitness fit_{best}.
(7) **If** $i \leq N/2$ **then**
(8) The individual position is updated according to Eq.(2).
(9) **Else if**
(10) The shock transfer factor is calculated according to Eq.(21).
(11) The individual position is updated according to Eq.(25).
(12) **End if**
(13) Keep the current optimal position.
(14) **For** $i = 1: N$
(15) The individual position is updated according to Eq.(6).
(16) The individual survival rate *su* and dynamic probability p are determined according to Eqs.(28)-(29).
(17) **If** $su< p$
(18) Generate mutation population according to Eq.(26).
(19) **End If**
(20) Save the best individual position and fitness value.
(21) **End For**
(22) **End While**
(23) Output the global optimal position X_{best} and best fitness fit_{best}.
(24) **End**

B. PROPOSED METHOD IN THIS PAPER

The basic contents of the proposed method are as follows:

Firstly, the two-stage noise reduction method of ICEEMDAN-WTD is used for signal processing. Considering that the performance of the ICEEMDAN algorithm is limited by the number of white noise addition and the degree of white noise addition, the proposed MFCOA is used to optimize its parameters to ensure the decomposition effect.

Permutation Entropy (PE) is solved for the IMF signal obtained after decomposition and the selection value h is set to judge. The IMF signal with PE value greater than h is processed by the 'db3' wavelet threshold denoising method, and the other signals are retained and participated in noise reduction reconstruction. The reason for choosing PE as the standard is that PE can more effectively reflect the degree of random variation and complexity of time series, and is more conducive to screening out IMF signals with noise. The key to wavelet threshold denoising is the selection of threshold function. However, the hard threshold function can preserve boundary information more effectively, but its discontinuity leads to too rough noise reduction effect. Although the soft threshold function improves the continuity, it will produce a constant deviation when the absolute value of the coefficient is too large, which will affect the approximation degree between the reconstructed signal and the original signal. Therefore, this paper chooses the improved threshold function in literature [32] for wavelet threshold noise reduction, and the regulatory factor a is set

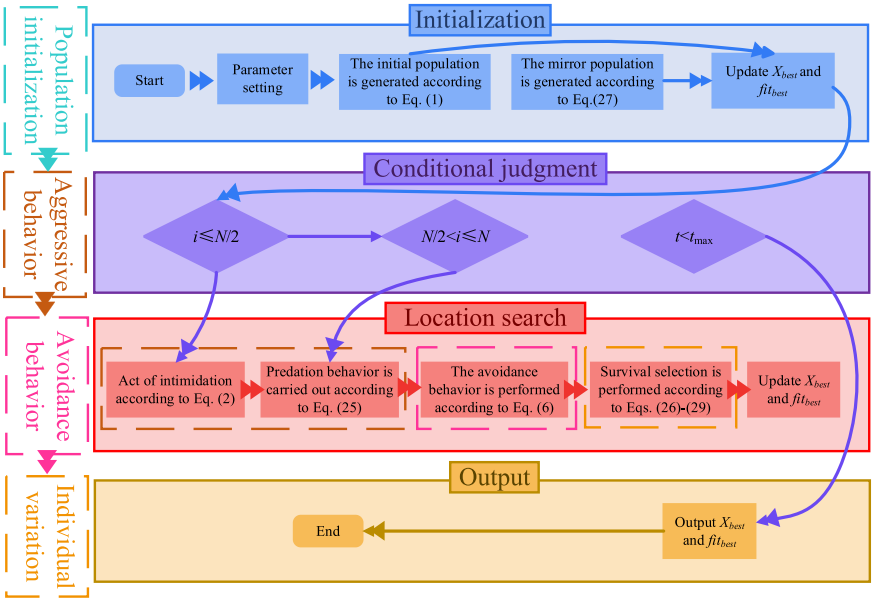


FIGURE 3. Flow chart of MFCOA.

to 200,

$$\hat{\omega}_{j,k} = \begin{cases} \omega_{j,k} \times \frac{2 \times \arctan([|\omega_{j,k}| - TH] \times a)}{\pi}, & |\omega_{j,k}| \geq TH \\ 0, & |\omega_{j,k}| < TH \end{cases} \tag{30}$$

Secondly, ICEEMDAN is used again to decompose the reconstructed signal to obtain the Sample Entropy (SE) of different time-frequency scales as fault features. Sample entropy is used to analyze the complexity of time series, which can effectively characterize the nonlinear characteristics of signals. Time domain features can effectively represent the state of signal changes over time, but they are one-sided because they cannot capture the information in the frequency domain. Moreover, time domain features are mainly obtained through quantization and other methods, which may cause information loss and interfere with subsequent diagnostic accuracy to a certain extent. A similar situation also exists in the frequency domain. In addition, there may be more interference information in the excessive time-frequency domain information, which requires feature screening, etc. Compared with it, sample entropy is more practical, can process discrete and continuous signals and effectively characterize the complexity of information contained in signals. Meanwhile, sample entropy has the advantages of high computation efficiency, good consistency, and short required data, etc. Therefore, sample entropy is selected as the fault feature in this paper.

Finally, in order to further improve the accuracy of model diagnosis and recognition and reduce the limited performance of SCNs caused by artificial experience, MFCOA is used to optimize the SCNs scale factor λ and regularization coefficient r, and the MFCOA-SCNS model is constructed to realize fault diagnosis.

In summary, the flow chart of the proposed method is shown in FIGURE 4.

IV. PERFORMANCE VERIFICATION OF IMPROVED ALGORITHM

In experiments, MATLAB R2021a is used as the software and running on the computer with Intel(R) Core (TM) i7-10750H CPU @ 2.60GHz 2.59GHz CPU and 16GB RAM.

A. TEST FUNCTION SELECTION

In this paper, 12 common test functions are selected for simulation verification, among which the number ratio of unimodal, multimodal and fixed-dimensional multimodal functions is 4:4:4, and the specific function information is shown in TABLE 1.

B. PARAMETER SETTING AND RESULT ANALYSIS

In order to verify the performance of MFCOA proposed in this paper, five intelligent algorithms with excellent performance are selected for testing and comparison: COA [26], AO [33], GJO [34], PSOGWO [35], and BWO [36]. The specific parameter settings of each algorithm are shown in TABLE 2.

The initial population size $N = 50$ and the maximum number of iterations $t_{max} = 1000$ are set. In order to avoid one-sided results caused by experimental contingency, various algorithms independently execute the optimization 30 times to obtain the Best, Worst, Average, and standard deviation. The average convergence curve is drawn as shown in FIGURES 5-6, and the results of each index are shown in TABLE 3.

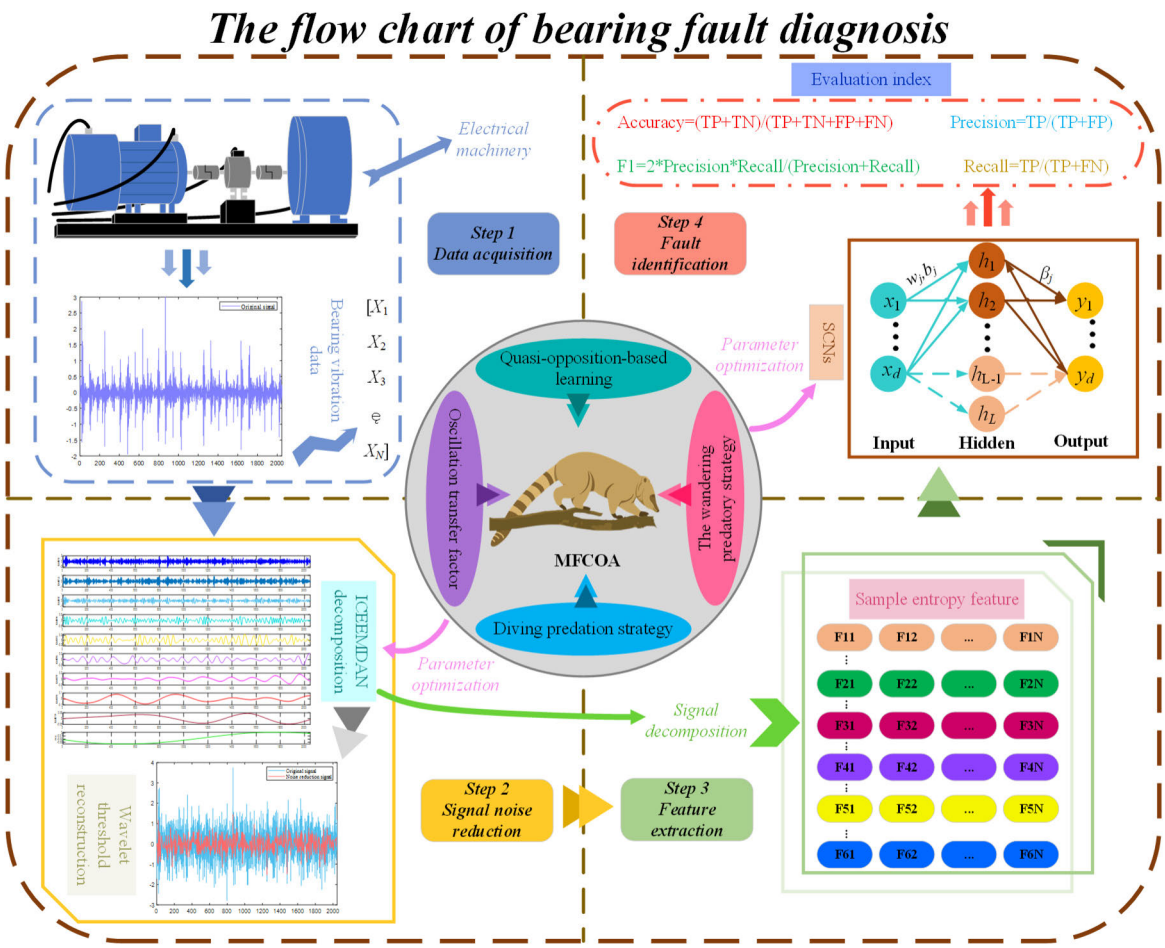


FIGURE 4. The flow chart of bearing fault diagnosis.

Analysis of TABLE 3 and FIGURES 5-6 shows that: For functions F2-F4 and F6-F7, MFCOA and COA are significantly better than other algorithms, and for other functions, MFCOA is almost better than other algorithms in convergence accuracy. This is due to the addition of mixed strategy, which makes the algorithm obtain better diversity and development ability, and effectively reduces the situation of local obstacles. Combined with the iterative convergence curve, ICOA can almost always search for the ideal result with the least number of iterations, which shows that it effectively balances the convergence efficiency while ensuring accuracy, and further reflects the effectiveness of the hybrid strategy.

C. WILCOXON SIGNED-RANK TEST

It is too one-sided to judge the algorithm by only the four indicators mentioned above. Therefore, the Wilcoxon signed-rank test is used to compare the matching degree between the algorithms to further verify the performance of MFCOA. With MFCOA as the standard, the p values of the output results of 30 independent detection experiments performed by each algorithm are shown in TABLE 4, where N/A represents the beyond evaluation accuracy, and the performance of the two algorithms is highly close. The boundary value of significant difference is set as 0.05, and "+", "−" and "=" are used to represent the degree of superiority between algorithms. For example, "+" means that MFCOA shows obvious advantages in the comparison.

According to the results in TABLE 4, the proposed MFCOA and COA show comparable performance in solving unimodal functions and F6-F7, which is attributed to the outstanding performance of the COA algorithm itself, which is easy to solve the optimal value. However, MFCOA is significantly superior to COA in the convergence efficiency of the solution, which reflects the advantages of MFCOA. In addition, the test p-values of MFCOA compared with other algorithms are almost all less than 0.05, which intuitively illustrates the superiority of MFCOA.

V. BEARING FAULT DIAGNOSIS EXPERIMENT

A. DATA PROVENANCE AND SIGNAL PROCESSING

The bearing composite fault data of this simulation experiment comes from the motor test bench of Liaoning Technical University (LNTU), China, as shown in FIGURE 7. In the

TABLE 1. The information of unimodal and multimodal benchmark functions.

Function	*Dim*	*Range*	F_{best}
$F_1(x)=\sum_{i=1}^{d}x_i^2$	20	[-100,100]	0
$F_2(x)=\max_i\{\lvert x_i\rvert,1\le i\le d\}$	20	[-100,100]	0
$F_3(x)=\sum_{i=1}^{d-1}\left[100\left(x_{i+1}-x_i^2\right)^2+\left(x_i-1\right)^2\right]$	20	[-30,30]	0
$F_4(x)=\sum_{i=1}^{d}\left(\lvert x_i+0.5\rvert\right)^2$	20	[-100,100]	0
$F_5(x)=\sum_{i=1}^{d}-x_i\sin\sqrt{\lvert x_i\rvert}$	20	[-500,500]	-12569.5
$F_6(x)=-20\exp\left(-0.2\sqrt{\frac{1}{n}\sum_{i=1}^{d}x_i^2}\right)-\exp\left(\frac{1}{d}\sum_{i=1}^{d}\cos\left(2\pi x_i\right)\right)+20+e$	20	[-32,32]	0
$F_7(x)=\frac{\pi}{d}\left\{10\sin\left(\pi y_1\right)+\sum_{i=1}^{d-1}\left(y_i-1\right)^2\left[1+10\sin^2\left(\pi y_{i+1}\right)\right]+\left(y_n-1\right)^2\right\}+\sum_{i=1}^{d}u\left(x_i,10,100,4\right)$ $y_i=1+\frac{x_i+1}{4}$ $u(x_i,a,k,m)=\begin{cases}k(x_i-a)^m, & x_i>a\\ 0, & -a<x_i<a\\ k(-x_i-a)^m, & x_i<a\end{cases}$	20	[-50,50]	0
$F_8(x)=\left[\frac{1}{500}+\sum_{j=1}^{25}\frac{1}{j+\sum_{i=1}^{2}\left(x_i-a_{ij}\right)^6}\right]^{-1}$ $a_{ij}=\begin{pmatrix}-32 & -16 & 0 & \text{L} & 32\\ -32 & -32 & -32 & \text{L} & 32\end{pmatrix}$	2	[-65.54,65.54]	1
$F_9(x)=\sum_{i=1}^{11}\left[a_i-\frac{x_1\left(b_i^2+b_ix_2\right)}{b_i^2+b_ix_3+x_4}\right]^2$ $a_i=0.1957,0.1947,0.1735,0.1600,0.0844,0.0627,$ $0.0456,0.0342,0.0235,0.246$ $b_i^{-1}=0.25,0.5,1,2,4,6,8,10,12,14,16$	4	[-5,5]	0.0003075
$F_{10}=4x_1^2-2.1x_1^4+\frac{1}{3}x_1^6+x_1x_2-4x_2^2+4x_2^4$	2	[-5,5]	-1.316285
$F_{11}(x)=\left(x_2-\frac{5\cdot1}{4\pi^2}x_1^2+\frac{5}{\pi}x_1-6\right)^2+10\left(1-\frac{1}{8\pi}\right)\cos x_1+10$	2	[-5,0] [10,15]	0.398
$F_{12}(x)=-\sum_{i=1}^{5}\left[\left(x-a_i\right)\left(x-a_i\right)^T+c_i\right]^{-1}$	4	[0,10]	−10.1532

TABLE 2. Setting of algorithm parameters.

Algorithm	parameter setting
AO	The development adjustment parameters a=δ =0.1, number of search cycles R_1=10.
GJO	Prey initial energy control coefficient c from 1.5 to 0.
PSOGWO	Control parameters a_{in} = 2 and a_{fin} = 0, Learning factors c_1 = 2 and c_2 = 2.
BWO	The whale fall probability W_F decays from 0.1 to 0, the fluctuation range of the initial balance factor B_F is (0,1).

experiment, the sampling frequency of the sensor is 10 kHz, the power frequency of the frequency regulator is 50 Hz, the rotor speed is 1453 r/min, and the fixed motor load is 1HP after the steady state of operation. A total of six states of bearing vibration signals are collected for analysis, and the bearing fault situation is shown in Table 5. Each 2048 sampling points are taken as a group of samples, and 100 groups of bearing vibration signal samples in different states are obtained. The training set and the test set are divided according to the ratio of 7:3.

Bearing signals of different states in Table 5 are collected by the experimental bench in FIGURE. 7, and a noise signal with a SNR of -6db is synthesized by awgn function to simulate the industrial application scenario of rolling bearings. Part of the time domain waveforms of the original signal and noise signal are shown in FIGURE. 8. The noise has almost covered the original signal.

For noisy signals, the MFCOA-ICEEMDAN method is first used for signal decomposition and screening. The parameters in the optimization are set as follows: the maximum number of iterations t_{max} = 5, the number of populations N = 18, the search lower bound R_L = [0.15 30], and the search upper bound R_H = [0.6 60]. Each order of IMF signals after decomposition is retained and the last order residual signal is eliminated. The IMF signals with PE>0.7 are selected for wavelet threshold denoising and

TABLE 3. Comarison of optimization experimental results of various algorithms.

Function	Result	Algorithm					
		MFCOA	COA	AO	GJO	PSOGWO	BWO
F1	Best	0.00E+00	0.00E+00	2.27E-303	3.36E-133	0.00E+00	0.00E+00
	Worst	0.00E+00	0.00E+00	2.06E-294	6.83E-125	0.00E+00	0.00E+00
	Average	0.00E+00	0.00E+00	4.13E-295	1.37E-125	0.00E+00	0.00E+00
	Std.	0.00E+00	0.00E+00	0.00E+00	3.05E-125	0.00E+00	0.00E+00
F2	Best	0.00E+00	0.00E+00	2.09E-158	3.96E-41	1.08E-156	5.22E-261
	Worst	0.00E+00	0.00E+00	3.22E-130	1.17E-37	2.37E-152	1.37E-254
	Average	0.00E+00	0.00E+00	1.61E-131	1.58E-38	1.83E-153	1.33E-255
	Std.	0.00E+00	0.00E+00	7.19E-131	2.93E-38	5.39E-153	0.00E+00
F3	Best	0.00E+00	0.00E+00	7.43E-06	2.62E+01	2.75E+01	2.63E-21
	Worst	0.00E+00	0.00E+00	7.57E-03	2.86E+01	2.89E+01	1.33E-14
	Average	0.00E+00	0.00E+00	8.39E-04	2.73E+01	2.81E+01	1.49E-15
	Std.	0.00E+00	0.00E+00	1.71E-03	7.41E-01	2.98E-01	3.96E-15
F4	Best	0.00E+00	0.00E+00	1.45E-08	1.99E+00	4.16E+00	7.86E-30
	Worst	0.00E+00	0.00E+00	3.88E-07	2.75E+00	4.55E+00	1.25E-28
	Average	0.00E+00	0.00E+00	2.01E-07	2.37E+00	4.36E+00	6.62E-29
	Std.	0.00E+00	0.00E+00	2.64E-07	5.36E-01	2.77E-01	8.26E-29
F5	Best	-1.26E+04	-1.26E+04	-1.26E+04	-6.40E+03	-3.22E+03	-1.26E+04
	Worst	-1.26E+04	-1.26E+04	-5.04E+03	-2.94E+03	-2.28E+03	-1.26E+04
	Average	-1.26E+04	-1.26E+04	-8.82E+03	-4.40E+03	-2.74E+03	-1.26E+04
	Std.	8.50E-04	1.11E-03	3.34E+03	1.17E+03	2.75E+02	1.92E+00
F6	Best	8.88E-16	8.88E-16	8.88E-16	4.44E-15	4.44E-15	8.88E-16
	Worst	8.88E-16	8.88E-16	8.88E-16	4.44E-15	4.44E-15	8.88E-16
	Average	8.88E-16	8.88E-16	8.88E-16	4.44E-15	4.44E-15	8.88E-16
	Std.	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
F7	Best	1.57E-32	1.57E-32	3.49E-08	1.44E-01	4.49E-01	4.22E-27
	Worst	1.57E-32	1.57E-32	4.11E-08	2.16E-01	5.11E-01	5.88E-26
	Average	1.57E-32	1.57E-32	3.80E-08	1.80E-01	4.80E-01	3.15E-26
	Std.	0.00E+00	0.00E+00	4.38E-09	5.10E-02	4.44E-02	3.86E-26
F8	Best	9.98E-01	9.98E-01	9.98E-01	2.98E+00	9.98E-01	9.98E-01
	Worst	9.98E-01	9.98E-01	9.98E-01	2.98E+00	2.98E+00	9.98E-01
	Average	9.98E-01	9.98E-01	9.98E-01	2.98E+00	1.99E+00	9.98E-01
	Std.	1.05E-16	4.75E-12	1.80E-09	5.27E-12	1.40E+00	8.16E-15
F9	Best	3.08E-04	3.08E-04	3.86E-04	3.07E-04	3.38E-04	3.08E-04
	Worst	3.17E-04	6.05E-04	5.35E-04	1.42E-03	6.77E-04	3.27E-04
	Average	3.12E-04	3.55E-04	4.31E-04	4.19E-04	4.67E-04	3.13E-04
	Std.	2.62E-06	8.99E-05	4.68E-05	3.52E-04	9.74E-05	6.11E-06
F10	Best	-1.03E+00	-1.03E+00	-1.03E+00	-1.03E+00	-1.03E+00	-1.03E+00
	Worst	-1.03E+00	-1.03E+00	-1.03E+00	-1.03E+00	-1.03E+00	-1.03E+00
	Average	-1.03E+00	-1.03E+00	-1.03E+00	-1.03E+00	-1.03E+00	-1.03E+00
	Std.	1.35E-05	9.86E-06	1.08E-04	2.33E-06	2.21E-05	1.69E-05
F11	Best	3.98E-01	3.98E-01	3.98E-01	3.98E-01	3.98E-01	3.98E-01
	Worst	3.98E-01	3.98E-01	3.98E-01	3.98E-01	4.00E-01	4.06E-01
	Average	3.98E-01	3.98E-01	3.98E-01	3.98E-01	3.99E-01	4.00E-01
	Std.	2.55E-07	2.72E-05	6.15E-05	2.59E-06	7.95E-04	3.17E-03
F12	Best	-1.02E+01	-1.02E+01	-1.02E+01	-1.02E+01	-4.73E+00	-1.02E+01
	Worst	-1.02E+01	-1.01E+01	-1.01E+01	-5.10E+00	-3.94E+00	-1.02E+01
	Average	-1.02E+01	-1.01E+01	-1.02E+01	-9.65E+00	-4.29E+00	-1.02E+01
	Std.	1.51E-06	2.88E-02	4.31E-03	1.60E+00	2.46E-01	1.26E-07

signal reconstruction. Part of the time-domain waveform of the reconstructed signal is plotted as shown in FIGURE 9, and a large amount of noise has been removed from the reconstructed signal.

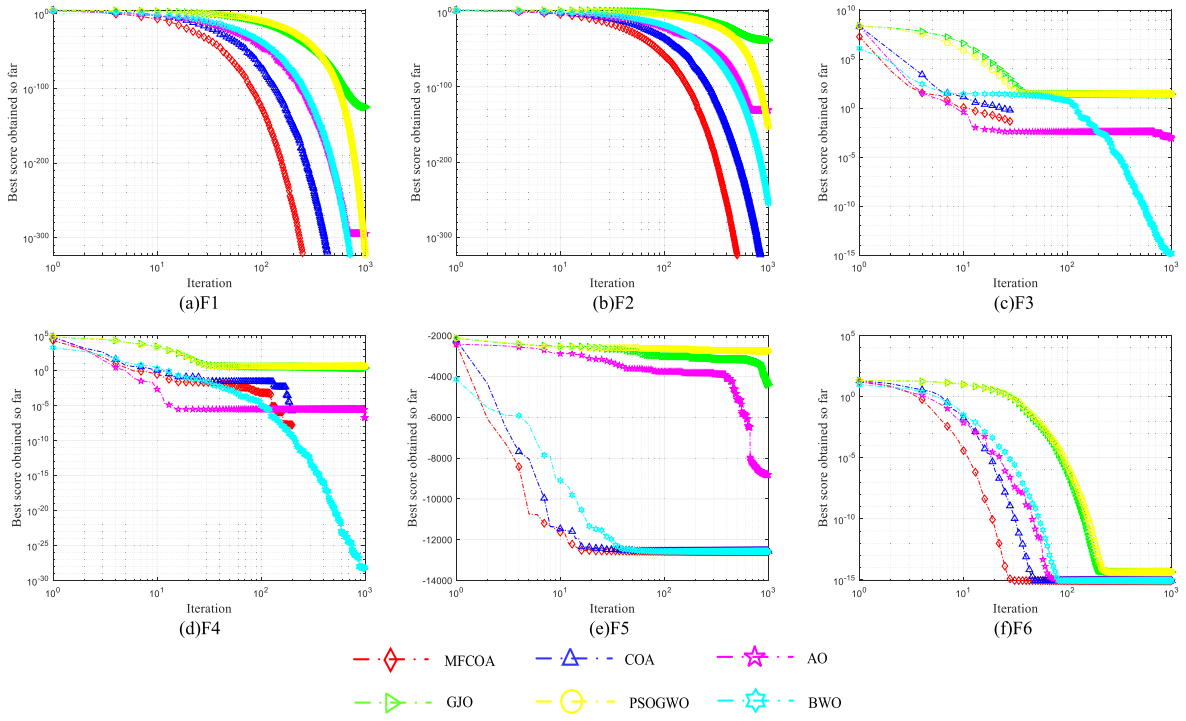


FIGURE 5. Comparison of convergence curves for F1–F6.

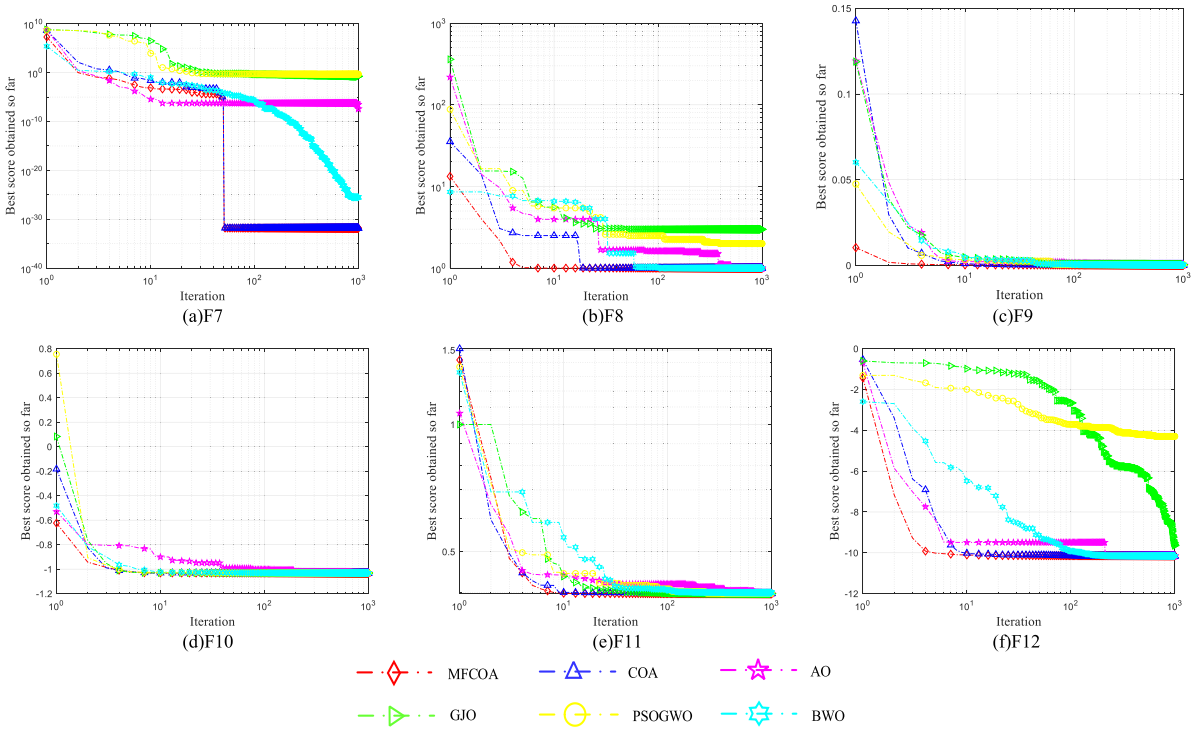


FIGURE 6. Comparison of convergence curves for F7–F12.

B. COMPARISION EXPERIMENT 1

After dividing the samples of the above reconstructed signals, the sample entropy features are extracted as the input of the model, and some of their eigenvalues are shown in TABLE 6, Where the embedding dimension $m = 2$ and the set value z is 0.2 times the standard deviation of the data. The sample labels 1-6 in TABLE 6 correspond to the fault states in TABLE 5, and SEn represents the

TABLE 4. Comparision of optimization experimental result of various algorithms.

Function	COA		AO		GJO		PSOGWO		BWO	
	p	Result	p	Result	p	Result	p	Result	p	Result
F1	N/A	=	8.01E-09	+	8.01E-09	+	N/A	=	N/A	=
F2	N/A	=	8.01E-09	+	8.01E-09	+	8.01E-09	+	8.01E-09	+
F3	N/A	=	6.39E-05	+	6.39E-05	+	6.39E-05	+	6.39E-05	+
F4	N/A	=	8.01E-09	+	8.01E-09	+	8.01E-09	+	8.01E-09	+
F5	2.55E-02	+	7.55E-04	+	1.79E-04	+	1.79E-04	+	2.21E-03	+
F6	N/A	=	N/A	=	1.59E-05	+	1.59E-05	+	N/A	=
F7	N/A	=	6.39E-05	+	6.39E-05	+	6.39E-05	+	6.39E-05	+
F8	1.70E-03	+	2.73E-01	+	7.57E-02	+	1.31E-03	+	1.83E-04	+
F9	6.40E-02	+	1.83E-04	+	2.83E-03	+	1.83E-04	+	7.91E-01	+
F10	3.45E-01	+	7.69E-04	+	3.07E-01	+	7.69E-04	+	9.11E-03	+
F11	2.73E-01	+	3.07E-01	+	9.10E-01	+	1.01E-03	+	2.83E-03	+
F12	1.83E-04	+	2.73E-01	+	9.70E-02	+	1.83E-04	+	1.83E-04	+

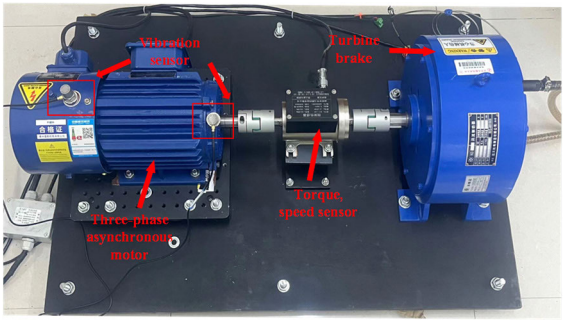


FIGURE 7. LNTU motor test bench.

TABLE 5. Fault conditions of bearings.

Fault type	Degree of fault
1.Normal	--
2.Inner ring fault	0.8mm deep, 0.05mm crack
3.Outer ring fault	0.8mm deep, 0.05mm crack
4.Rolling element fault	0.8mm deep,0.5mm cautery
5.Compound fault 1(1 and 2)	0.8mm deep, 0.05mmcrack
6.Compound fault 2(1,2 and 3)	0.8mm deep, 0.05mm crack,0.5mm cautery

TABLE 6. Characteristic values of samples.

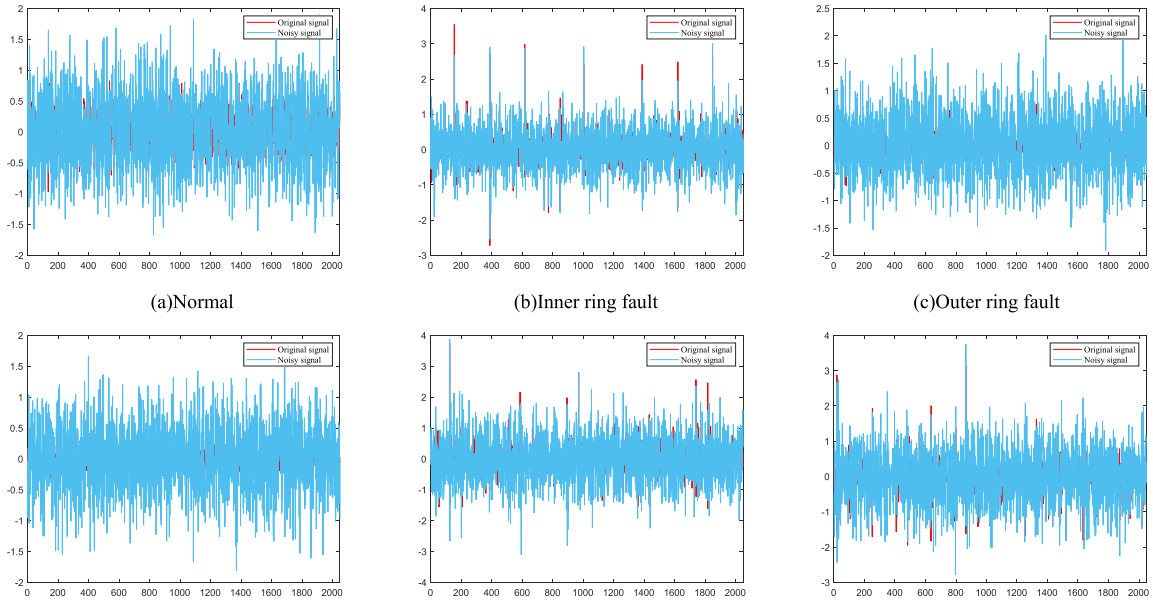
Sample label	SE_1	SE_2	SE_3	...	SE_7	SE_8
1	1.6999	1.2795	0.6847	...	0.1352	0.0550
2	1.6079	1.2341	0.6868	...	0.1084	0.0497
3	1.7176	1.2797	0.6755	...	0.1020	0.0249
4	1.5622	1.2001	0.6787	...	0.1141	0.0457
5	1.6537	1.2521	0.6680	...	0.0894	0.0429
6	1.4993	1.2162	0.7038	...	0.0943	0.0308

TABLE 7. Each method in experiment 1.

	Method
Method 1	ICEEMDAN decomposes the signal and extracts SE features,and classification by SCNs.
Method 2	MFCOA-ICEEMDAN-WPD processes the signal and extracts SE features, and classification by SCNs.
Method 3	MFCOA-ICEEMDAN-WPD processes the signal and extracts SE features, and classification by PSO-RF.
Method 4	MFCOA-ICEEMDAN-WPD processes the signal and extracts SE features, and classification by GWO-SVM.

sample entropy of the extracted IMF signals of each order.

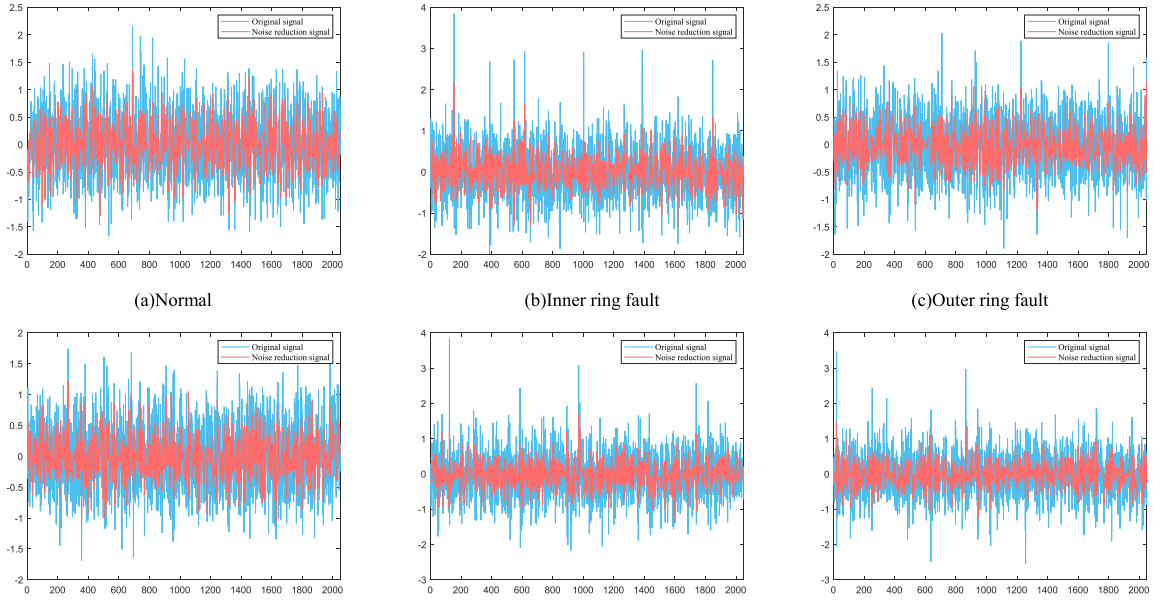
To verify the effectiveness of the proposed method, the following comparison experiments are designed. Method 1 does not perform noise reduction and directly inputs the SCNs model for diagnosis after calculating the sample entropy through ICEEMDAN; based on Method 1, Method 2 introduces the MFCOA-ICEEMDAN-WPD two-stage noise reduction method for diagnosis; Method 3 and Method 4 use the proposed method for noise reduction and feature extraction, and the difference was that PSO-RF [10] and GWO-SVM [37] are respectively used for diagnostic experiments. The detailed introduction of each method is shown in TABLE 7. This paper chooses accuracy (*Acc*), Precision (*Pre*), *Recall*, and F1 coefficient ($F1$) as the evaluation criteria to compare the proposed method with the other four methods. The calculation methods of each index are Equations (31)-(34). The experimental results are shown in TABLE 8 and FIGURE 10. It should be noted that the experimental results are the average results of 20 independent experiments. The SCNs set as follows: $L_{max} = 50$,



(a)Normal (b)Inner ring fault (c)Outer ring fault

(d)Rolling element fault (e)Compound fault 1 (f)Compound fault 2

FIGURE 8. Time domain waveform of a noisy signal.



(a)Normal (b)Inner ring fault (c)Outer ring fault

(d)Rolling element fault (e)Compound fault 1 (f)Compound fault 2

FIGURE 9. Time domain waveform of a reconstructed signal.

$T_{\max} = 100$, $\lambda = [100{:}2{:}200]$, $r = [0.9,\ldots,0.999]$. MFCOA sets the maximum number of iterations $t_{max} = 10$ and the initial number of populations $N = 20$.

$$Acc = (\mathrm{TP} + \mathrm{TN})/(\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}) \tag{31}$$

$$Pre = \mathrm{TP}/(\mathrm{TP} + \mathrm{FP}) \tag{32}$$

$$Recall = \mathrm{TP}/(\mathrm{TP} + \mathrm{FN}) \tag{33}$$

$$F1 = 2 \times Precision \times Recall/(Precision + Recall) \tag{34}$$

According to the information in TABLE 8 and FIGURE 10, it can be intuitively shown that the proposed method can effectively reduce the influence of noise, which is conducive to mining fault features containing different time-frequency domain information, thus improving the diagnosis accuracy. In addition, all indexes of methods 3 and 4 are lower than those of the proposed method, which verifies the performance superiority of MFCOA-SCNs and proves the engineering practicability of MFCOA.

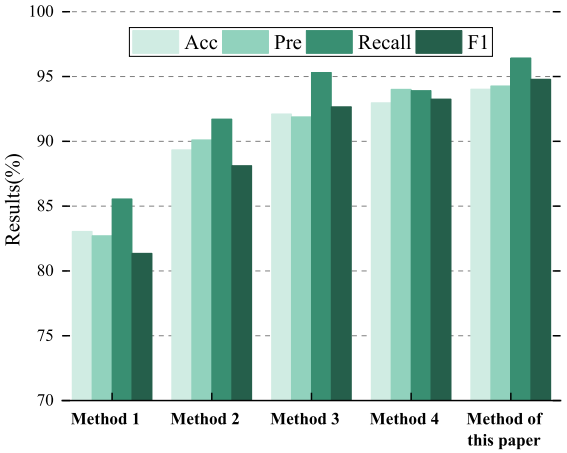


FIGURE 10. Results of experiment 1.

TABLE 8. Classification results of experiment 1.

Method	*Acc*	*Pre*	*Recall*	*F*1
Method 1	83.05%	82.72%	85.56%	81.36%
Method 2	89.34%	90.11%	91.72%	88.12%
Method 3	92.11%	91.88%	95.30%	92.67%
Method 4	92.97%	94.01%	93.91%	93.26%
Proposed method	94.02%	94.27%	96.43%	94.78%

TABLE 9. Classification results of experiment 2.

Method	*Acc*	*Pre*	*Recall*	*F*1
COA-SCNs	93.15%	93.44%	94.65%	92.93%
AO-SCNs	92.97%	92.89%	95.92%	91.54%
GJO-SCNs	93.22%	93.71%	94.31%	93.28%
PSOGWO-SCNs	91.95%	91.45%	91.49%	90.63%
BWO-SCNs	92.78%	94.02%	93.07%	92.19%
MFCOA-SCNs	94.02%	94.27%	96.43%	94.78%

C. COMPARISION EXPERIMENT 2

To further verify the performance of MFCOA-SCNs proposed in this paper, the five algorithms mentioned above are used to optimize the SCNs to construct a variety of SCNs optimized versions. The fault samples obtain by the proposed method are also used to compare the performance with MFCOA-SCNs. The parameter settings of each algorithm are the same as shown in TABLE 2, and the average results of 20 independent experiments are shown in TALBE 9 and FIGURE 11. The optimization time of each model is shown in FIGURE 12.

In the results shown in TALBE 9 and FIGURE 12, although MFCOA increases the time cost due to the addition of multiple strategies, the index value of MFCOA-SCNs are better than other methods, which indicates that for the data set used in this paper, MFCOA has higher convergence accuracy for the optimization solution of SCNs, which makes SCNs obtain stronger performance. It illustrates the strong competitiveness of MFCOA-SCNs in engineering applications.

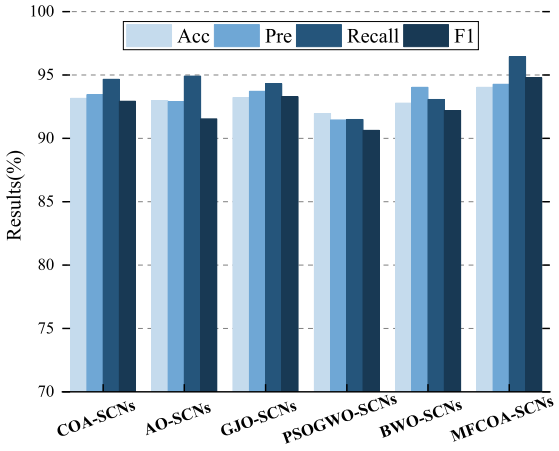


FIGURE 11. Results of experiment 2.

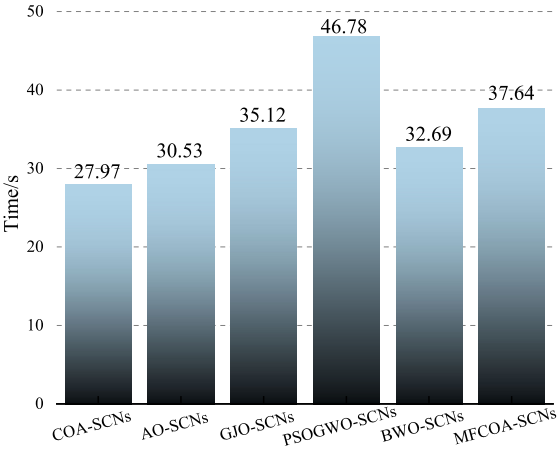


FIGURE 12. Optimization time.

TABLE 10. Each method in experiment 3.

	Method
Method 5	ICEEMDAN-WPD processes the signal and extracts SE features, and classification by SCNs.
Method 6	The WPD is used to reduce noise, the sample entropy of the decomposed signal is calculated as the feature, and the classification is carried out by SVM.
Method 7	Wavelet threshold noise reduction processing, using GA to optimize VMD and calculate sample entropy, classifying by WOA-LSSVM.
Method 8	The VMD decomposes the signal and extracts the fuzzy entropy feature, and realizes the fault classification by BOA-DBN.

D. COMPARISION EXPERIMENT 3

To comprehensively consider the practicability of the method proposed in this paper, it is compared with other four methods: Method 5, Method 6 [38], Method 7 [39], and

TABLE 11. Classification results of experiment 2.

Method	*Acc*	*Pre*	*Recall*	*F*1
Method 5	88.12%	89.12%	90.02%	88.65%
Method 6	86.89%	88.35%	86.71%	89.39%
Method 7	92.43%	93.23%	93.16%	92.69%
Method 8	91.76%	91.45%	93.26%	92.86%
Proposed method	94.02%	94.27%	96.43%	94.78%

Method 8 [40]. The contents of each method are explained in TABLE 10, and the experimental results are shown in TABLE 11 and FIGURE 13.

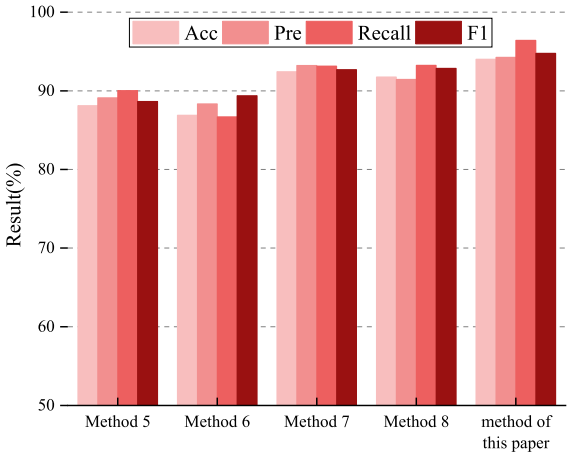


FIGURE 13. Results of experiment 3.

According to the experimental results, the proposed method is superior to other methods in various indexes, which is due to the fact that MFCOA-ICEEMDAN-WPD method effectively reduces noise interference and is conducive to extracting important feature information. Meanwhile, MFCOA ensures the performance of SCNs.

VI. CONCLUSION

In this paper, a bearing fault diagnosis method based on MFCOA multi-task optimization was proposed. A two-stage decomposition method of ICEEMDAN-WTD was used to reduce the effects of noise interference in bearing vibration signals. As some hyper-parameters of the model may affect the diagnosis accuracy, so a multi-strategy hybrid MFCOA algorithm was proposed to optimize them in ICEEMDAN and SCNs, in order to reduce subjective influence of the manual setting. The experimental results showed that the *Acc*, *Pre*, *Recall* and *F*1 of the proposed method can reach 94.02%, 94.27%, 96.43% and 94.78%, respectively, which was significantly better than other methods mentioned in this paper.

The method proposed in this paper showed ideal results in dealing with bearing fault diagnosis. However, the results obtained by only analyzing the vibration signal, which limited the diagnosis accuracy to a certain extent. Therefore, the fusion diagnosis method of vibration signals combined with other features such as images is the direction of our future work.

ACKNOWLEDGMENT

The authors would like to thank the Faculty of Electrical and Control Engineering, Liaoning Technical University.

REFERENCES

[1] Y. Liu, Y. Cheng, Z. Zhang, and J. Wu, "Acoustic fault diagnosis of rotor bearing system," *Shock Vib.*, vol. 2022, pp. 1–9, Apr. 2022.
[2] Y. Li, B. Tang, X. Jiang, and Y. Yi, "Bearing fault feature extraction method based on GA-VMD and center frequency," *Math. Problems Eng.*, vol. 2022, pp. 1–19, Jan. 2022.
[3] M. Wang, Y. Chen, X. Zhang, T. K. Chau, H. H. Ching Iu, T. Fernando, Z. Li, and M. Ma, "Roller bearing fault diagnosis based on integrated fault feature and SVM," *J. Vib. Eng. Technol.*, vol. 10, no. 3, pp. 853–862, Apr. 2022.
[4] H. Habbouche, T. Benkedjouh, and N. Zerhouni, "Intelligent prognostics of bearings based on bidirectional long short-term memory and wavelet packet decomposition," *Int. J. Adv. Manuf. Technol.*, vol. 114, nos. 1–2, pp. 145–157, Mar. 2021.
[5] Y. Wenhua, Y. Lei, W. Zhenhuan, Y. Jianhua, T. Tiejun, and L. Liansheng, "Geotechnical engineering blasting: A new modal aliasing cancellation methodology of vibration signal de-noising," *Earthq. Eng. Eng. Vib.*, vol. 21, no. 2, pp. 313–323, Apr. 2022.
[6] Y. Liu, X. Lu, G. Bei, and Z. Jiang, "Improved wavelet packet denoising algorithm using fuzzy threshold and correlation analysis for chaotic signals," *Trans. Inst. Meas. Control*, vol. 43, no. 6, pp. 1394–1403, Apr. 2021.
[7] Q. Lu, X. Shen, X. Wang, M. Li, J. Li, and M. Zhang, "Fault diagnosis of rolling bearing based on improved VMD and KNN," *Math. Problems Eng.*, vol. 2021, pp. 1–11, Oct. 2021.
[8] M. A. Colominas, G. Schlotthauer, and M. E. Torres, "Improved complete ensemble EMD: A suitable tool for biomedical signal processing," *Biomed. Signal Process. Control*, vol. 14, pp. 19–29, Nov. 2014.
[9] Z. Zhu, X. Xu, L. Li, Y. Dai, and Z. Meng, "A novel GA-BP neural network for wireless diagnosis of rolling bearing," *J. Circuits, Syst. Comput.*, vol. 31, no. 10, Jul. 2022, Art. no. 2250173.
[10] G. Tang, B. Pang, T. Tian, and C. Zhou, "Fault diagnosis of rolling bearings based on improved fast spectral correlation and optimized random forest," *Appl. Sci.*, vol. 8, no. 10, p. 1859, Oct. 2018.
[11] Z. Guo, M. Liu, Y. Wang, and H. Qin, "A new fault diagnosis classifier for rolling bearing united multi-scale permutation entropy optimize VMD and cuckoo search SVM," *IEEE Access*, vol. 8, pp. 153610–153629, 2020.
[12] M. Xu and H. Yao, "Fault diagnosis method of wheelset based on EEMD-MPE and support vector machine optimized by quantum-behaved particle swarm algorithm," *Measurement*, vol. 216, Jul. 2023, Art. no. 112923.
[13] C. Zhang and S. Ding S, "A stochastic configuration network based on chaotic sparrow search algorithm," *Knowl.-Based Syst.*, vol. 220, May 2021, Art. no. 106924.
[14] Y. Li, S. Jiao, S. Deng, B. Geng, and Y. Li, "Refined composite variable-step multiscale multimapping dispersion entropy: A nonlinear dynamical index," *Nonlinear Dyn.*, vol. 112, no. 3, pp. 2119–2137, Feb. 2024.
[15] Q. Li, "New sparse regularization approach for extracting transient impulses from fault vibration signal of rotating machinery," *Mech. Syst. Signal Process.*, vol. 209, Mar. 2024, Art. no. 111101.
[16] Y. Li, B. Tang, S. Jiao, and Y. Zhou, "Optimized multivariate multiscale slope entropy for nonlinear dynamic analysis of mechanical signals," *Chaos, Solitons Fractals*, vol. 179, Feb. 2024, Art. no. 114436.
[17] Y. Li, B. Tang, S. Jiao, and Q. Su, "Snake optimization-based variable-step multiscale single threshold slope entropy for complexity analysis of signals," *IEEE Trans. Instrum. Meas.*, vol. 72, pp. 1–13, 2023.
[18] D. Wang, D. Tan, and L. Liu, "Particle swarm optimization algorithm: An overview," *Soft Comput.*, vol. 22, no. 2, pp. 387–408, Jan. 2018.
[19] J. Xue and B. Shen, "A novel swarm intelligence optimization approach: Sparrow search algorithm," *Syst. Sci. Control Eng.*, vol. 8, no. 1, pp. 22–34, Jan. 2020.

[20] S. Mirjalili and A. Lewis, "The whale optimization algorithm," *Adv. Eng. Softw.*, vol. 95, pp. 51–67, May 2016.
[21] K. Li, X. Yan, and Y. Han, "Multi-mechanism swarm optimization for multi-UAV task assignment and path planning in transmission line inspection under multi-wind field," *Appl. Soft Comput.*, vol. 150, Jan. 2024, Art. no. 111033.
[22] L. Liu and R. Zhang, "Multistrategy improved whale optimization algorithm and its application," *Comput. Intell. Neurosci.*, vol. 2022, pp. 1–16, May 2022.
[23] T. Fei, H. Wang, L. Liu, L. Zhang, K. Wu, and J. Guo, "Research on multi-strategy improved sparrow search optimization algorithm," *Math. Biosci. Eng.*, vol. 20, no. 9, pp. 17220–17241, 2023.
[24] P. Qu and F. Du, "Improved particle swarm optimization for laser cutting path planning," *IEEE Access*, vol. 11, pp. 4574–4588, 2023.
[25] Y. Hou, H. Gao, Z. Wang, and C. Du, "Improved grey wolf optimization algorithm and application," *Sensors*, vol. 22, no. 10, p. 3810, May 2022.
[26] M. Dehghani, Z. Montazeri, E. Trojovská, and P. Trojovský, "Coati optimization algorithm: A new bio-inspired metaheuristic algorithm for solving optimization problems," *Knowl.-Based Syst.*, vol. 259, Jan. 2023, Art. no. 110011.
[27] J. Ma, H. Li, B. Tang, J. Wang, Z. Zou, and M. Zhang, "Rolling bearing fault diagnosis based on improved VMD-adaptive wavelet threshold joint noise reduction," *Adv. Mech. Eng.*, vol. 14, no. 10, Oct. 2022, Art. no. 168781322211283.
[28] D. Wang and M. Li, "Stochastic configuration networks: Fundamentals and algorithms," in *Proc. IEEE Trans. Cybern.*, Oct. 2017, vol. 47, no. 10, pp. 3466–3479.
[29] R. Liu, T. Wang, J. Zhou, X. Hao, Y. Xu, and J. Qiu, "Improved African vulture optimization algorithm based on quasi-oppositional differential evolution operator," *IEEE Access*, vol. 10, pp. 95197–95218, 2022.
[30] D. Wu, H. Rao, C. Wen, H. Jia, Q. Liu, and L. Abualigah, "Modified sand cat swarm optimization algorithm for solving constrained engineering optimization problems," *Mathematics*, vol. 10, no. 22, p. 4350, Nov. 2022.
[31] H. Peraza-Vázquez, A. F. Peña-Delgado, G. Echavarría-Castillo, A. B. Morales-Cepeda, J. Velasco-Álvarez, and F. Ruiz-Perez, "A bio-inspired method for engineering design optimization inspired by dingoes hunting strategies," *Math. Problems Eng.*, vol. 2021, pp. 1–19, Sep. 2021.
[32] D. Zhang, S. Wang, F. Li, S. Tian, J. Wang, X. Ding, and R. Gong, "An efficient ECG denoising method based on empirical mode decomposition, sample entropy, and improved threshold function," *Wireless Commun. Mobile Comput.*, vol. 2020, pp. 1–11, Dec. 2020.
[33] L. Abualigah, D. Yousri, M. A. Elaziz, A. A. Ewees, M. A. A. Al-Qaness, and A. H. Gandomi, "Aquila optimizer: A novel meta-heuristic optimization algorithm," *Comput. Ind. Eng.*, vol. 157, Jul. 2021, Art. no. 107250.
[34] N. Chopra and M. Mohsin Ansari, "Golden jackal optimization: A novel nature-inspired optimizer for engineering applications," *Expert Syst. Appl.*, vol. 198, Jul. 2022, Art. no. 116924.
[35] N. Singh and S. B. Singh, "Hybrid algorithm of particle swarm optimization and grey wolf optimizer for improving convergence performance," *J. Appl. Math.*, vol. 2017, pp. 1–15, Jan. 2017.
[36] C. Zhong, G. Li, and Z. Meng, "Beluga whale optimization: A novel nature-inspired mestaheuristic algorithm," *Knowl.-Based Syst.*, vol. 251, Sep. 2022, Art. no. 109215.
[37] X. Su, C. Cao, X. Zeng, Z. Feng, J. Shen, X. Yan, and Z. Wu, "Application of DBN and GWO-SVM in analog circuit fault diagnosis," *Sci. Rep.*, vol. 11, no. 1, pp. 1–14, Apr. 2021.
[38] M. Qiao, X. Tang, Y. Liu, and S. Yan, "Fault diagnosis method of rolling bearings based on VMD and MDSVM," *Multimedia Tools Appl.*, vol. 80, no. 10, pp. 14521–14544, Jan. 2021.
[39] J. Li, W. Chen, K. Han, and Q. Wang, "Fault diagnosis of rolling bearing based on GA-VMD and improved WOA-LSSVM," *IEEE Access*, vol. 8, pp. 166753–166767, 2020.
[40] Z. Jin and Y. Sun, "Bearing fault diagnosis based on VMD fuzzy entropy and improved deep belief networks," *J. Vib. Eng. Technol.*, vol. 11, no. 2, pp. 577–587, Feb. 2023.

KUN LI was born in Taian, Shandong, China, in 1983. He received the B.Sc. degree from Shandong University of Science and Technology, in 2005, the M.Sc. degree from Liaoning Technical University, China, in 2008, and the Ph.D. degree from Northeastern University of China, in 2013.

From September 2013 to January 2021, he was an Assistant Professor with the College of Control Science and Engineering, Bohai University, China. Since March 2021, he has been a Professor with the Faculty of Electrical and Control Engineering, Liaoning Technical University. He is currently the author of more than 30 articles and more than ten inventions. His research interests include industrial artificial intelligence, complex system modeling, intelligent computing, machine learning, multi-objective optimization, swarm intelligence optimization algorithm, and their applications in petroleum, coal mine, power, communication network, and UAV path planning.

HAO WU was born in Weifang, Shandong, China, in 1998. He received the B.Sc. degree from Taishan University, in 2021. He is currently pursuing the M.Sc. degree with Liaoning Technical University, Huludao, China. His current research interests include fault diagnosis method and intelligent optimization algorithm.

XINMING LIU was born in Chaoyang, Liaoning, China, in 1984. He received the B.Sc. and M.Sc. degrees from Liaoning Technical University, in 2002 and 2009, respectively, and the Ph.D. degree from Northeastern University, China, in 2015. He is currently a Lecturer with the Faculty of Electrical and Control Engineering, Liaoning Technical University. His research interests include fault diagnosis methods and applications and intelligent control theory and applications.

JIANING XU was born in Chaoyang, Liaoning, China, in 1991. She received the B.Sc. degree from Northeastern University, China, in 2013, and the M.Sc. degree from Northeast Electric Power University, in 2016. She is currently a Lecturer with the Faculty of Electrical and Control Engineering, Liaoning Technical University. Her research interests include intelligent mine inspection theory and technology and intelligent optimization algorithm.

YING HAN was born in Jinzhou, Liaoning, China, in 1982. She received the B.Sc. and M.Sc. degrees from Liaoning Technical University, in 2005 and 2008, respectively, and the Ph.D. degree from Northeastern University, China, in 2020. She is currently an Associate Professor with the Faculty of Electrical and Control Engineering, Liaoning Technical University. She is the author of more than 20 articles and more than ten inventions. Her research interests include prognostics and health management (PHM) in industrial applications, time series analysis and applications, and complex system prediction.