Rolling Bearing Fault Diagnosis Based on ELM Optimized by Variational Quasi-Reflective Learning Arithmetic Optimization Algorithm

Yin-Yin Bao, Er-Chao Li*, Xiao-Qian Dai, Bao-Xin Zhang

Abstract—**Accurate fault diagnosis of rolling bearings is critical for ensuring the safety and efficiency of rotating machinery. This study introduces a novel Extreme Learning Machine (ELM) model optimized by the Modified Quasi-Reflective Arithmetic Optimization Algorithm (MQRAOA) to improve diagnostic performance. Initially, vibration signals are decomposed using Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN), enabling the extraction of intrinsic mode functions (IMFs) across multiple frequency bands. Subsequently, Composite Multiscale Approximate Entropy (CMAE) is utilized to capture the multiscale features of the decomposed signals. These features are then input into the ELM model, where MQRAOA optimizes the random initialization of weights and biases by enhancing global search capabilities and mitigating the risk of local optima. Experimental results reveal that the MQRAOA-optimized ELM outperforms models optimized using WOA, SCA, SSA, and TLBO regarding classification accuracy and robustness. This study establishes an effective framework for intelligent fault diagnosis and offers a promising approach for broader industrial applications.**

Index Terms—**fault diagnosis, extreme learning machines, approximate entropy, optimization algorithm**

I. Introduction

IN intelligent manufacturing, the manufacturing chain is undergoing unprecedented transformations. Tang et al. highlighted that smart manufacturing has significantly enhanced production efficiency and product quality by integrating digitalization, automation, and intelligent technologies[1]. In this paradigm, the convenience, accuracy, and reliability of intelligent troubleshooting methods for equipment are particularly critical. Technology-driven sensors and real-time data acquisition systems provide precise data that help rapidly detect and localize abnormalities. High diagnostic accuracy minimizes production downtime and reduces maintenance costs, driving the manufacturing industry toward greater efficiency and reliability.

Manuscript received October 7, 2024; revised March 5, 2025. This work was supported by the National Natural Science Foundation of China (Grant No. 62306131).

Yin-Yin Bao is a postgraduate student at the School of Electrical Engineering and Information Engineering, Lanzhou University of Science and Technology, Lanzhou, 730050, P. R. China (e-mail: bao17640667213@163.com).

Er-Chao Li is a professor at the School of Electrical Engineering and Information Engineering, Lanzhou University of Science and Technology, Lanzhou, 730050, P. R. China (Corresponding author: lecstarr@163.com).

Xiao-Qian Dai is a postgraduate student at the School of Electrical Engineering and Information Engineering, Lanzhou University of Science and Technology, Lanzhou, 730050, P. R. China (e-mail: daiiii07@163.com).

Bao-Xin Zhang is a postgraduate student at the School of Electrical Engineering and Information Engineering, Lanzhou University of Science and Technology, Lanzhou, 730050, P. R. China (e-mail: 3197522253@qq.com).

As a critical component of mechanical equipment, rolling bearings are essential in equipment performance and longevity. A failure in rolling bearings can drastically reduce mechanical efficiency and may even cause severe equipment damage. Studies indicate that approximately 40% of rotating machinery failures originate from bearing damage. Therefore, research into rolling bearing fault diagnosis is crucial for improving equipment reliability and extending its service life. Current research predominantly focuses on feature extraction from rolling bearing vibration signals to achieve fault identification. However, due to the influence of operating environments and various other factors, the vibration signals generated by rolling bearings exhibit high complexity. In failure scenarios, these signals are often nonlinear and non-stationary. Directly using these signals for fault identification may result in weak feature representations, compromising diagnostic accuracy. To address this issue, Variational Mode Decomposition (VMD) is widely employed to reduce noise in weak vibration signals caused by early bearing faults.

Entropy, a metric that reflects signal complexity, is often used to construct feature vectors for nonlinear and nonsmooth data. Common entropy metrics include sample entropy and permutation entropy (PE). PE has been widely applied in fault diagnosis for feature extraction. For example, Yang Jingzong et al. combined local mean decomposition with PE and an optimized Extreme Learning Machine (ELM) for fault diagnosis [2], while Yang Yun et al. utilized VMD and PE to identify and classify rolling bearing faults [3]. These approaches have effectively improved diagnostic accuracy. However, despite its success, PE evaluates signal complexity only on a single scale and fails to capture rolling bearing vibration signal characteristics across multiple scales.Compared to single-scale entropy, Composite Multiscale Approximate Entropy (CMAE) can analyze signals across multiple time scales, enabling it to capture the complexity of rolling bearing vibration signals more effectively. This is particularly advantageous in nonlinear and nonsmooth signal processing, as CMAE enhances the ability to detect weak fault characteristics and comprehensively characterizes the multiscale dynamics of the signals. Consequently, CMAE significantly improves fault diagnosis accuracy.

Fault diagnosis methods for rolling bearings commonly include decision trees, support vector machines (SVMs), extreme learning machines (ELMs), and neural networks. Yu et al. noted that decision trees perform well in small-scale fault scenarios but are less effective when dealing with high-dimensional, complex data [4]. Neural networks rely on backpropagation algorithms, but as the network depth increases, the training time grows substantially—this issue is particularly pronounced in rolling bearing fault diagnosis, where the complexity of vibration signal characteristics can hinder real-time diagnostic performance. To address these challenges, the Extreme Learning Machine (ELM) has gained popularity due to its fast training speed and strong classification capability. ELM, a feed-forward neural network algorithm, is characterized by strong global generalization and high training efficiency. Once fault features are extracted, ELM can effectively learn and recognize these features in rolling bearing fault diagnosis, significantly improving diagnostic efficiency and real-time performance [5-6].

Despite its advantages, ELM's random initialization of weights and biases in the input and hidden layers can result in instability and poor robustness during operation [7]. This randomness affects the model's classification performance, particularly when processing complex or noisy fault signals, and reduces its generalization capability. Therefore, improving ELM's initialization strategy or introducing more robust training methods is crucial to enhancing its practical application. Swarm intelligence algorithms have shown promise in optimizing the classification performance of ELMs. For example, Feng et al. optimized ELM weights and thresholds using the Sparrow Search Algorithm (SSA) and its improved version (ISSA), achieving better fault feature extraction and classification accuracy [8]. Du et al. introduced a Hybrid Kernel Extreme Learning Machine (HKELM) optimized by the Improved Northern Eagle Optimization Algorithm (INGO) to enhance fault diagnosis accuracy [9]. Yang et al. proposed a rolling bearing fault diagnosis method combining Variational Mode Decomposition (VMD), improved Artificial Fish Swarm Algorithm (AFSA), and multi-feature vector fusion in ELM [10]. While these swarm intelligence algorithms have improved ELM optimization, they still face challenges, including lengthy optimization times, susceptibility to local optima, and limited robustness, which must be addressed in practical fault diagnosis scenarios.

The Arithmetic Optimization Algorithm (AOA) is a metaheuristic optimization algorithm designed for global optimization based on the distributional properties of arithmetic operators. While it achieves better solution efficiency than other algorithms, it shares a common limitation of swarm intelligence methods: susceptibility to local optima [11]. This paper proposes an Arithmetic Optimization Algorithm based on Variational Quasi-Reflective Learning (MQRAOA) to overcome this issue. The MQRAOA aims to effectively avoid local optima and enhance the algorithm's convergence accuracy and speed.

In this study, MQRAOA is applied to optimize the weights and biases of the ELM network, leading to the development of an MQRAOA-ELM-based fault classification model. First, the MQRAOA-optimized ELM network is constructed to classify bearing faults. Then, extracted fault features are then input into the model for accurate identification of fault type. Comparative experimental results with four other fault classification models—SCA-ELM, SSA-ELM, TLBO-ELM, and WOA-ELM demonstrate that the MQRAOA-ELM model outperforms the alternatives in recognition accuracy and classification performance. These findings validate the superiority of MQRAOA in rolling bearing fault diagnosis.

II. FUNDAMENTAL PRINCIPLE

A. Fully adaptive noise ensemble empirical modal decomposition

Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) is an improved algorithm for signal processing based on Ensemble Empirical Mode Decomposition (EEMD), designed to address the standard mode aliasing problem in traditional Empirical Mode Decomposition (EMD) [12]. As a data-driven, non-linear, and non-smooth signal processing method, EMD is widely used to analyze complex signals. However, its decomposition results are often unstable due to noise interference. EEMD reduces the effect of modal aliasing by adding white noise to the signal and performing multiple decompositions, but the problem is not completely resolved. CEEMDAN builds on this foundation by introducing an adaptive noise adjustment mechanism that gradually modifies the noise level in each decomposition step. This approach makes the decomposition results more stable and accurate, significantly improving feature extraction capabilities and noise suppression effects[13]. As a result, CEEMDAN has been widely applied and verified in the analysis of non-linear and non-smooth signals, particularly in fault diagnosis and medical signal processing.

The core idea of CEEMDAN is to introduce noise through recursive iterations and adaptively adjust the noise level to maintain signal characteristics while reducing modal aliasing during each decomposition step. First, multiple sets of white noise are added to the original signal, and the ensemble average of the first *IMF* layer is obtained through EMD decomposition. Subsequently, in each iteration, the *IMFs* extracted in the previous step are subtracted from the signal to produce a residual signal. White noise is then introduced to the residual signal to extract the next layer of *IMFs*. As the decomposition progresses, the noise amplitude is gradually adjusted to ensure the accuracy of subsequent *IMF* extractions. This iterative process continues until the residual signal can no longer be decomposed into additional *IMFs* or a preset stopping condition is reached. Finally, the original signal is decomposed into a series of *IMFs* and a residual term. Algorithm steps for CEEMDAN:

Assume that the original signal $X(t)$ is given:

1)White $\omega_i(t)$ noise (i-rd noise) is added to the signal and EMD decomposition is performed on it to obtain the first $IMF_1^{(i)}(t)$, and the decomposition results of all the noises are averaged to obtain the first layer of *IMF*:

$$IMF_1 = \sum_{i=1}^{N} IMF_1^{(i)}(t) \tag{1}$$

2)Remove the first layer of *IMF* from the signal to get the residual $r_1(t) = x(t) - IMF_1(t)$ and add white noise to the residual signal for the next decomposition.

3)When calculating the next layer of *IMF*, the noise amplitude is adaptively adjusted based on the residuals' characteristics to gradually reduce the noise's impact. The EMD decomposition is continued, and the second layer of *IMF* is derived:

$$IMF_2 = \sum_{i=1}^{N} IMF_2^{(i)}(t) \tag{2}$$

4)Repeat the above steps until the residual signal cannot decompose further or a preset condition is reached.

The original rolling bearing vibration signal is decomposed into multiple *IMFs* using CEEMDAN, as shown in Fig. 1. CEEMDAN effectively decomposes complex, nonlinear, and non-smooth signals by gradually introducing noise and adaptively adjusting the noise level, successfully avoiding the modal overlapping problem commonly found in traditional Empirical Mode Decomposition (EMD). With an increase in the *IMF* order, the frequencies in the modes are gradually reduced, enabling the extraction of signal features across different frequency ranges.

The Pearson correlation coefficients of each *IMF* component with the original signal are shown in Fig. 2. These coefficients reveal the degree of correlation between each *IMF* and the original signal. *IMFs* with higher correlation coefficients indicate that the component contains more critical information.

In comparison, *IMFs* with lower correlation coefficients primarily contain noise or minor components, providing a visual measure of their contribution to the signal's features. Additionally, the variance contribution ratio of each *IMF* component is shown in Fig. 3, which reflects the proportion of energy each *IMF* accounts for in the original signal. A higher variance contribution ratio indicates that the *IMF* plays a more significant role in interpreting the signal and represents a critical component for fault feature extraction. These analyses demonstrate that CEEMDAN effectively captures important features in rolling bearing signals and provides reliable data for fault diagnosis.

The core advantage of CEEMDAN lies in its mechanism of gradually adjusting the noise amplitude, significantly improving the decomposition accuracy by avoiding modal aliasing and preserving the local features of the signal during the decomposition process. After each decomposition, CEEMDAN adaptively adjusts the noise according to the results, gradually reducing its influence on subsequent *IMFs*. This mechanism enhances the stability of the decomposition, enabling CEEMDAN to perform effectively when dealing with complex, nonlinear, and non-smooth signals, particularly in high-noise fault diagnosis scenarios.

B. Composite multiscale approximate entropy

Compound Multiscale Approximate Entropy (CMAE) is a powerful tool for analyzing the complexity of time series. It combines the advantages of Multiscale Entropy (MSE) and Approximate Entropy (*ApEn*) to evaluate signal complexity across different time scales comprehensively. CMAE is especially suitable for processing nonlinear and non-smooth signals, with applications in rolling bearing fault diagnosis, physiological signal analysis, and other fields.

Approximate Entropy (*ApEn*) is a nonlinear method for measuring the irregularity and complexity of time series. While it performs well with short time series, it has limitations when applied to long time series or multiscale complex signals. In contrast, Multiscale Entropy (MSE) provides a better understanding of the global dynamic characteristics of signals by reconstructing them across multiple time scales. However, single-scale analysis methods often fail to reflect the complexity of signals at different scales fully.

CMAE overcomes these limitations by combining the strengths of MSE and *ApEn*, providing a comprehensive assessment of signal complexity by calculating approximate entropy at multiple scales. This multiscale approach enhances the robustness and accuracy of the analysis, particularly when dealing with complex signals. CMAE is capable of capturing essential signal features more accurately. Compared with single-scale approximate entropy, CMAE offers a more comprehensive and adaptable analysis across different time scales, demonstrating higher reliability in processing complex, nonlinear, and non-smooth signals. The steps for calculating the CMAE are as follows:

1)Signal reconstruction: for a given time series $X = \{x_1, x_2, \cdots, x_N\}$, reconstruct it by different scale factors τ to get a new series $X^{(\tau)}$.

2)Approximate entropy calculation:for the reconstructed signal at each scale factor τ, the approximate entropy value is calculated $ApEn^{(\tau)}$. The approximate entropy is used to quantify the complexity of the signal, with larger values indicating more complex signals.

3)Compound Multiscale Processing: The approximate entropy calculated at each time scale is compounded and averaged to obtain the final CMAE value, which measures the signal's multiscale complexity.

When applying Composite Multiscale Approximate Entropy (CMAE) to analyze rolling bearing fault signals, the entropy values at various time scales are plotted, as shown in Fig. 4. This plot illustrates the entropy variation of rolling bearing signals across four states: normal, typical inner ring failure, rolling element failure, and outer ring failure. The signal complexity in the normal state is higher, with smoother entropy values across all scales, indicating consistently high complexity at all time scales. In contrast, the entropy values for the inner ring and rolling element faults are higher at small scales but gradually decrease as the scale factor increases. This decrease is particularly prominent for inner ring faults, which exhibit a noticeable trend of declining complexity. Meanwhile, the entropy value for outer ring faults remains consistently lower across all scales, indicating lower signal complexity, making distinguishing from other fault types easier.

C. Principles of Extreme Learning Machine Algorithm

Extreme Learning Machine (ELM) is a machine learning algorithm similar to the single hidden layer feedforward neural network (SLFN) model. Unlike SLFN, ELM replaces gradient-based algorithms with random initialization of input layer weights and biases. It uses generalized inverse matrix theory to compute the output layer weights during testing. ELM's advantages, such as fast training and good generalization, make it well-suited for diagnosing rolling bearing faults under complex working conditions. The implementation process of the ELM algorithm is as follows: The ELM network model is depicted in Fig. 5. The model consists of three layers: the input layer, the hidden layer, and the output layer, arranged from left to right. All layers are fully connected, ensuring efficient information transfer between them.

The core of ELM lies in the processing of the inputs and outputs of the hidden layer node $h_i(x)$,Suppose given the

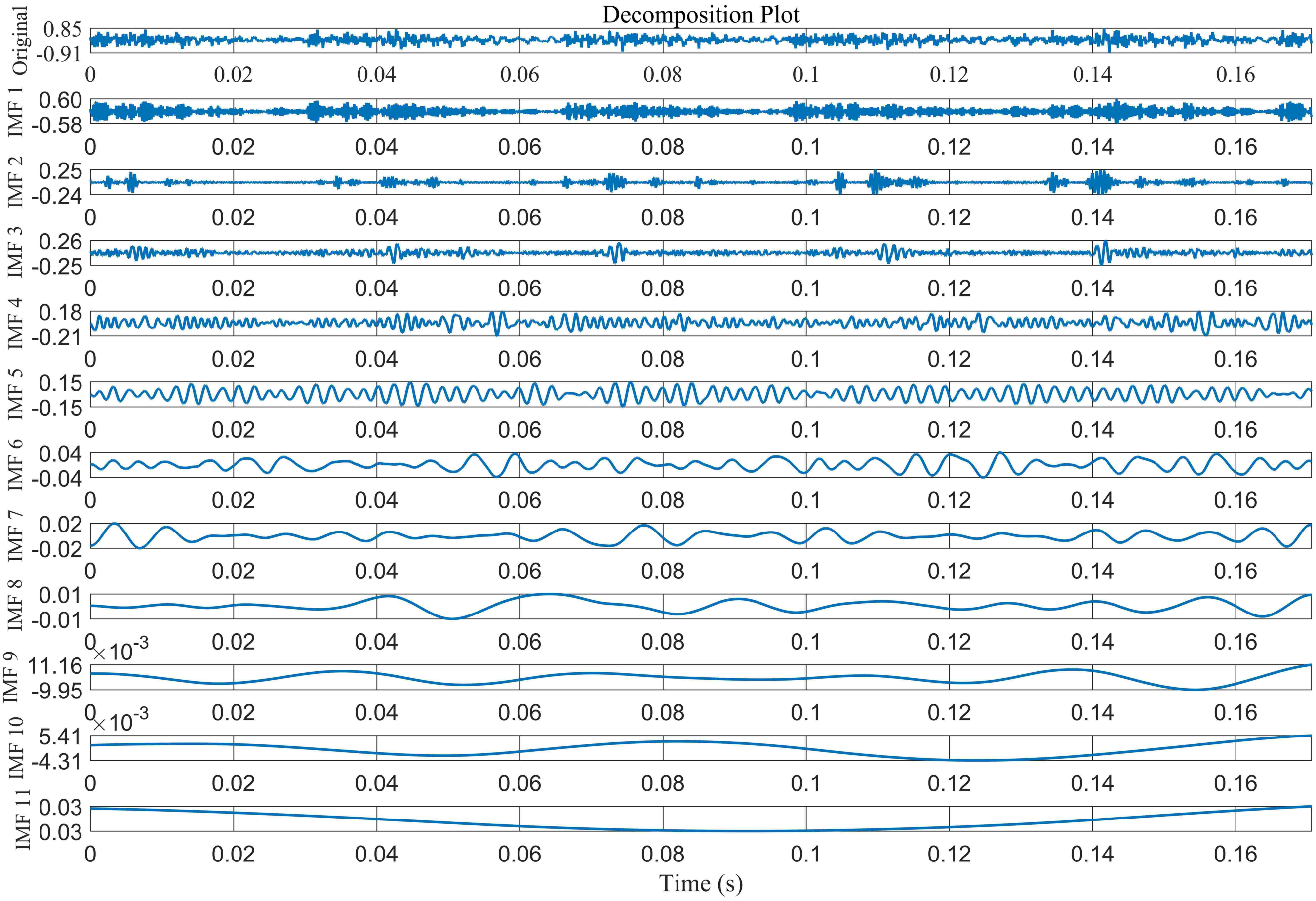

Fig. 1. Raw signals and their intrinsic modal functions (IMFs) obtained by CEEMDAN

training set $\{x_i, t_i \mid x_i \in R^D, t_i \in R^m, i = 1, 2, \cdots, N\}$; denotes the i-rd data, t_i denotes the label corresponding to x_i, according to the input relation, then the hidden layer node $h_i(x)$ after the weighting operation can be expressed as.

$$h_i(x) = \mathrm{g}\left(\omega_i, b_i, x\right) = \mathrm{g}\left(\omega_i x + b_i\right), \omega_i \in R^D, b_i \in R \quad (3)$$

Where, $\mathrm{g}\left(\omega_i, b_i, x\right)$ is the activation function, ω_i and b_i are the weights and biases of the hidden layer nodes, the activation function used in this paper is the sigmoid function.

The set of hidden layer nodes can be denoted as $H(x) = [h_1(x), h_2(x), \cdots, h_L(x)]$,so the output of the single hidden layer used for "generalization"after the hidden layer is:

$$f_L(x) = \sum_{i=1}^{L} \beta_i h_i(x) = H(x)\beta \mid \quad (4)$$

Where, $\beta = [\beta_1, \beta_2, \cdots, \beta_L]^T$ is the output weight between the L nodes of the hidden layer and the nodes of the output layer.

Up to this point, three types of unknown quantities appear in the model ω,b,β. In ELM, the hidden layer node parameter,at initialization, is randomly generated from any continuous probability distribution and kept constant, which can be solved by combining Eq. (3) and its set $H(x)$. β is determined by training on the data, and in order to minimize the training error, the minimum squared deviation of $H(x)\beta$ from the sample labels T, i.e.min $\|H\beta - T\|^L$, $\beta \in R^{L\times m}$, is taken as an objective of the solving function. The objective function contains the matrices H and T, which can be expressed as follows:

$$\begin{aligned} H &= \left[h\left(x_1\right), h\left(x_2\right), \cdots, h\left(x_N\right)\right]^T \\ &= \begin{pmatrix} h_1\left(x_1\right) & \cdots & h_L\left(x_1\right) \\ \vdots & \ddots & \vdots \\ h_L\left(x_1\right) & \cdots & h_L\left(x_N\right) \end{pmatrix} \\ T &= \left[t_1^T, t_2^T, \cdots, t_N^T\right]^T \end{aligned} \quad (5)$$

By matrix transformation, it can be deduced that the objective function yields an optimal solution:

$$\beta^* = H^* T \quad (6)$$

Where, H is the Moore-Penrose generalized inverse matrix of matrix H .

In the ELM model, the input and output parameters of the hidden layer are determined in advance. Unlike traditional neural networks, ELM does not require backpropagation to update its parameters, enabling faster model training.

III. Rolling Bearing Fault Diagnosis with ELM Optimized by Arithmetic Optimization Algorithm Based on Variational Quasi-Reflective Learning

A. Arithmetic Optimization Algorithm

Arithmetic operators—multiplication, division, subtraction, and addition—are traditional computational methods for studying numbers. Arithmetic Optimization Algorithms (AOA) leverage these operators as mathematical optimization

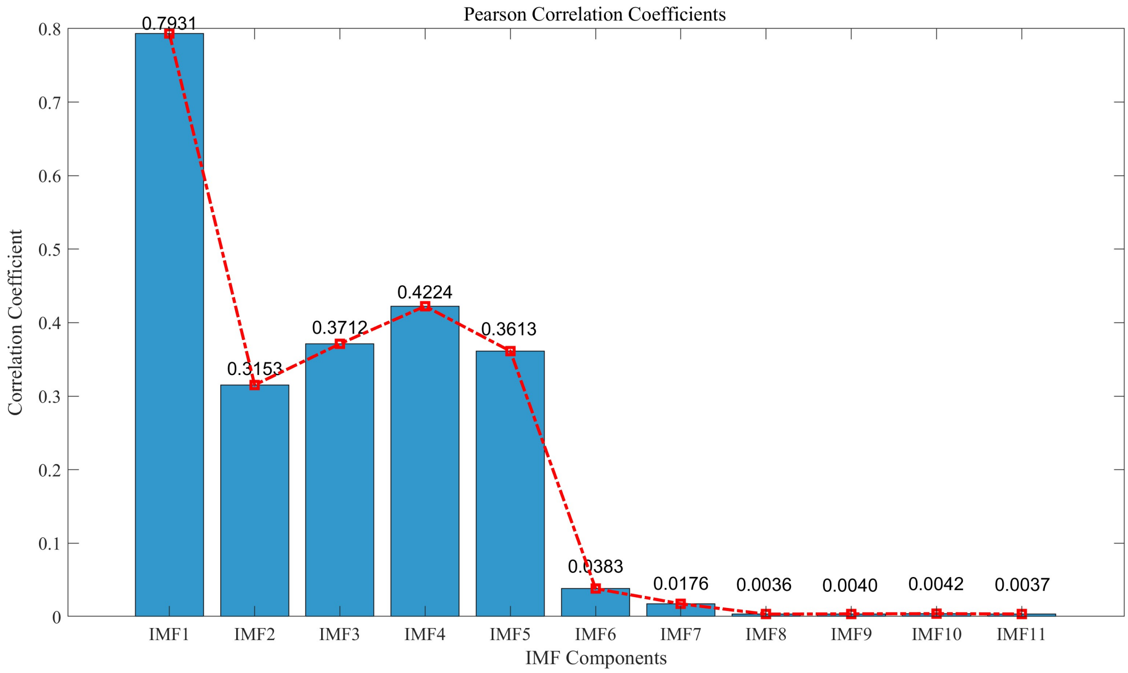

Fig. 2. Pearson's correlation coefficient between each IMF component and the original signal

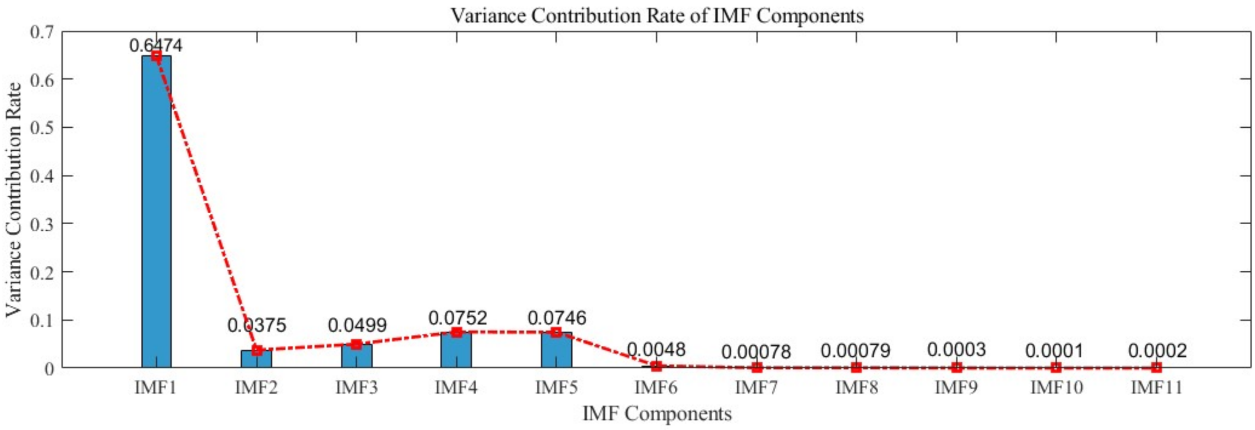

Fig. 3. Variance Contribution of IMF Componentsl

tools to identify the optimal solution from a set of candidate solutions that best meet predefined criteria. The AOA process is divided into initialization, exploration, and exploitation. Fig. 6 illustrates the hierarchical structure of arithmetic operators within the AOA framework.

(1) Initialization phase

In arithmetic optimization algorithms (AOA), initial candidate solutions are randomly generated, as shown in Eq. (7). The best candidate solution in each iteration is regarded as the best or near-optimal solution obtained so far [6].

$$X = \begin{bmatrix} x_{1,1} & \cdots & x_{1,n} \\ \vdots & \ddots & \vdots \\ x_{N,1} & \cdots & x_{N,n} \end{bmatrix} \tag{7}$$

Before the AOA begins operating, it must determine the appropriate search phase—exploration or exploitation. For this purpose, the Mathematical Optimization Acceleration (MOA) function, a coefficient used to guide the search phase, is calculated using the formula provided in Eq. (8).

$$\text{MOA}\ (C\_\text{Iter}\ ) = \text{Min}\ + C\_\text{Iter}\ \times (\ \text{Max}\ - \text{Min}\ )/M\_\ \text{Iter} \tag{8}$$

Where, $MOA\,(C\_Iter)$ is the value of *MOA* for the current iteration, which is calculated by Eq.(8); C Iter is the current iteration and M Iter is the maximum number of iterations;Min and Max denote the minimum and maximum values of *MOA*, respectively. Its change image with the number of iterations is shown in Fig. 7.

(2) Exploration phase

The exploration phase of AOA is performed using either the division or multiplication operator; the *MOA* controls this search phase. When the condition is $r_1 > MOA$ (r_1 is a random number), the algorithm carries out the exploration phase; when $r_1 < 0.5$ (r_2 is a random number), the division operator is used and the multiplication operator will be ignored, and the multiplication operator will be used to perform the current task instead of the division operator when $r_2 > 0.5$.The equation for updating the position of the exploration phase of the AOA is Eq. (9).

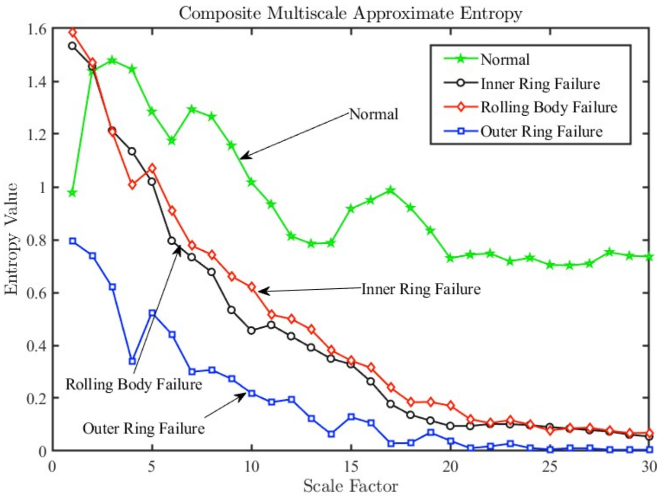

Fig. 4. Multi-scale approximate entropy analysis of rolling bearing signals for different fault types

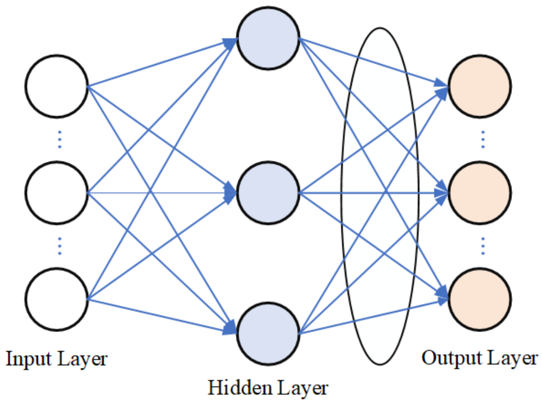

Fig. 5. ELM network model

$$x_{i,j}\left(C\_\text{Iter}\right)=\begin{cases}\text{best}\ \left(x_j\right)\div\left(MOP+\in\right)\times\left(\left(UB_j-LB_j\right)\times\mu+LB_j\right) & r_2<0.5\\ \text{best}\left(x_j\right)\times MOP\times\left(\left(UB_j-LB_j\right)\times\mu+LB_j\right) & \text{otherwise}\end{cases}\tag{9}$$

Where, $x_{i,j}\left(C\_\text{Iter}\right)$ denotes the *j*-rd position of the *i*-nd solution of the current iteration, while $\text{best}\left(x_j\right)$ is the *j*-th position of the best solution obtained so far. $\in$ is a small integer, UB_j and LB_j denote the upper and lower bound values for the *j*-th position, respectively. μ is a control parameter regulating the search process, $\mu=0.5$.

$$MOP\left(C\_\text{Iter}\right)=1-\left(C\_\text{Iter}^{\ 1/\alpha}\right)/\left(M\_\text{Iter}^{\ 1/\alpha}\right)\tag{10}$$

Where, *MOP* is a coefficient, $MOP\left(C\_\text{Iter}\right)$ denotes the value of the function of the first iteration. α is a sensitive parameter that defines the iteration's development accuracy, $\alpha=5$. The image of *MOP* variation with the number of iterations is shown in Fig. 8.

(3) Development phase

The development phase of the AOA algorithm is carried out using subtraction and addition operators. The *MOA* function also controls the development phase, and the algorithm carries out the development phase under the condition that $r_1<MOA$ (r_1 is a random number). When $r_3<0.5$ (r_3 is a random number), subtraction operator is used and addition operator will be ignored and when $r_3>0.5$ addition operator will be used to perform the current task instead of subtraction operator.The AOA development phase position update equation is Eq. (11).

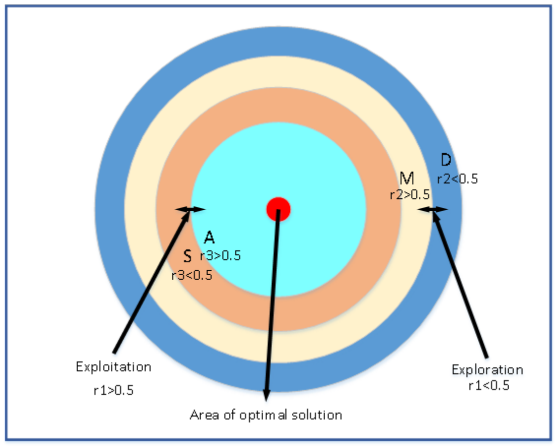

Fig. 6. Hierarchy of arithmetic operators

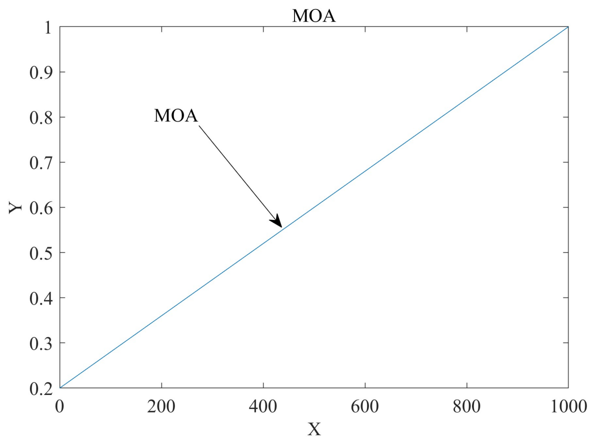

Fig. 7. Image of MOA

$$x_{i,j}\left(C\_\text{Iter}\right)=\begin{cases}\text{best}\left(x_j\right)-MOP\times\left(\left(UB_j-LB_j\right)\times\mu+LB_j\right) & r_3<0.5\\ \text{best}\left(x_j\right)+MOP\times\left(\left(UB_j-LB_j\right)\times\mu+LB_j\right) & \text{otherwise}\end{cases}\tag{11}$$

B. Arithmetic Optimization Algorithm Based on Mutation Quasi-Reflection Learning

To overcome the local optimization problem in AOA, mutation quasi-reflection learning (MQR) was introduced to improve the original algorithm using mutation and quasi-reflection operations[14]. Mutation quasi-reflective learning is a learning-theory-based optimization mechanism that aims to enhance the global search capability of the algorithm by combining random mutation and quasi-reflective operations. First, the mutation operation prevents the solution from falling into the local optimum in the local search by introducing random perturbations in the neighborhood of the solution; second, the quasi-reflection operation simulates the reflection phenomenon and maps the current suboptimal solution to another symmetric location, thus realizing the jump search in the global scope. With these two mechanisms, the algorithm remains versatile in later development stages, avoids local optimum traps, and significantly improves global search efficiency.

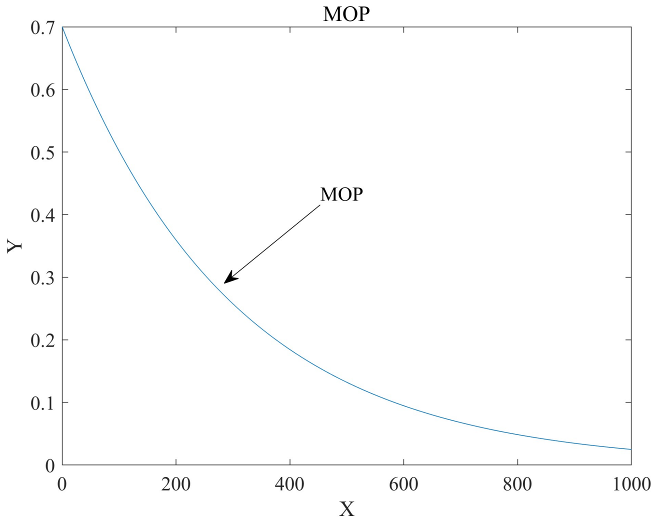

Fig. 8. Enter Caption

The mutation operation prevents the solution from falling into a local optimum by introducing a random perturbation in the neighborhood of the solution, which is updated using the following formula:

The random perturbation is generated using a normal distribution with a mean of 0 and a specified standard deviation. This ensures that most perturbation values are concentrated near the current solution while there is a certain probability of introducing larger perturbations. A stochastic coefficient controls the strength of the perturbation, determining how much the random perturbation influences the update of the current solution. The value of this coefficient ranges between [0, 1].

With the introduction of variational quasi-reflective learning, the updated formulation of the Arithmetic Optimization Algorithm (AOA) significantly improves the algorithm's global search capability and the efficiency of local exploitation by introducing random perturbations and control coefficients. These changes ensure that the algorithm can avoid local optima and quickly converge to a globally optimal solution when dealing with complex optimization problems. Specifically, random perturbations enhance the diversity of solutions, allowing the algorithm to escape local regions and maintain diversity in the solution space. The control coefficients enable the algorithm to adjust dynamically according to the search requirements at different stages, thus improving robustness and adaptability. During the exploitation phase, the introduced perturbations and jump searches accelerate the convergence process, rapidly approaching the optimal solution and improving optimization efficiency. The flowchart of the improved AOA algorithm is shown in Fig. 9.

C. ELM optimized based on variational quasi-reflective learning arithmetic optimization algorithm

Proposed Methodology for Rolling Bearing Fault Diagnosis:

(1)Signal Preprocessing: The vibration signals of rolling bearings are first preprocessed. This involves denoising the signals, decomposing them using the Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN), and extracting key Intrinsic Mode Functions (*IMFs*). This *IMFs* capture critical fault-related features for further analysis.

(2) Feature Extraction:The Complexity Measure Adaptive Entropy (CMAE) is employed to extract the complexity features of the signals across multiple scales. This multi-scale analysis allows for a comprehensive characterization of the rolling bearing fault signals, enriching the representation of fault characteristics and enhancing classification accuracy.

(3) ELM Model Construction:An Extreme Learning Machine (ELM) model is constructed. The input-to-hidden layer connection weights and bias parameters are initialized randomly, while the hidden layer output is calculated using the generalized inverse matrix theory, ensuring efficient computation.

(4) MQR-AOA Optimization of ELM Parameters: The Multi-Quasi-Reflective Arithmetic Optimization Algorithm (MQR-AOA) is introduced to optimize the ELM model's random weights and biases. MQR improves the algorithm's global search capability by incorporating random perturbations and quasi-reflective operations, effectively avoiding local optima.

(5) Optimization process:During each iteration, the weights and biases of the ELM model are updated using the MQR-AOA algorithm. Subtraction and addition operators are applied during the developmental phase to accelerate convergence toward the global optimal solution. The global search capability and exploration diversity are further enhanced through variational quasi-reflective learning (MQR). This mechanism improves the global jump potential of the search process, effectively avoiding local optima and significantly boosting optimization efficiency.

(6) **Classification and diagnosis:** The optimized ELM model is trained using the extracted fault features. This enables the accurate classification and diagnosis of rolling bearing faults, ensuring high reliability and performance.

IV. Simulation experiment and result analysis

A. Data set

To verify the validity and superiority of the MQRAOA-ELM fault diagnosis model proposed in this study, the rolling bearing fault vibration signals publicly provided by the Bearing Data Center of Case Western Reserve University were selected as the research object. The experimental data consists of deep groove ball bearing fan-end data, where the fault types include inner ring, outer ring, and rolling element failures, each with a fault damage diameter of 0.007 mm. The vibration signals were collected at a sampling frequency of 12 kHz across four states: normal bearing, inner ring failure, rolling element failure, and outer ring failure. A total of 480 samples were collected for each state, with each sample containing 2048 data points.

B. Fault diagnosis and comparison verification

The experimental dataset was preprocessed by extracting features from the raw vibration signals and applying dimensionality reduction. The resulting dataset comprises four fault states, each representing 100 feature samples. The dataset was divided into training and testing sets to ensure model generalization: the first 70 samples were used for training, and the remaining 30 were used for testing. All input data

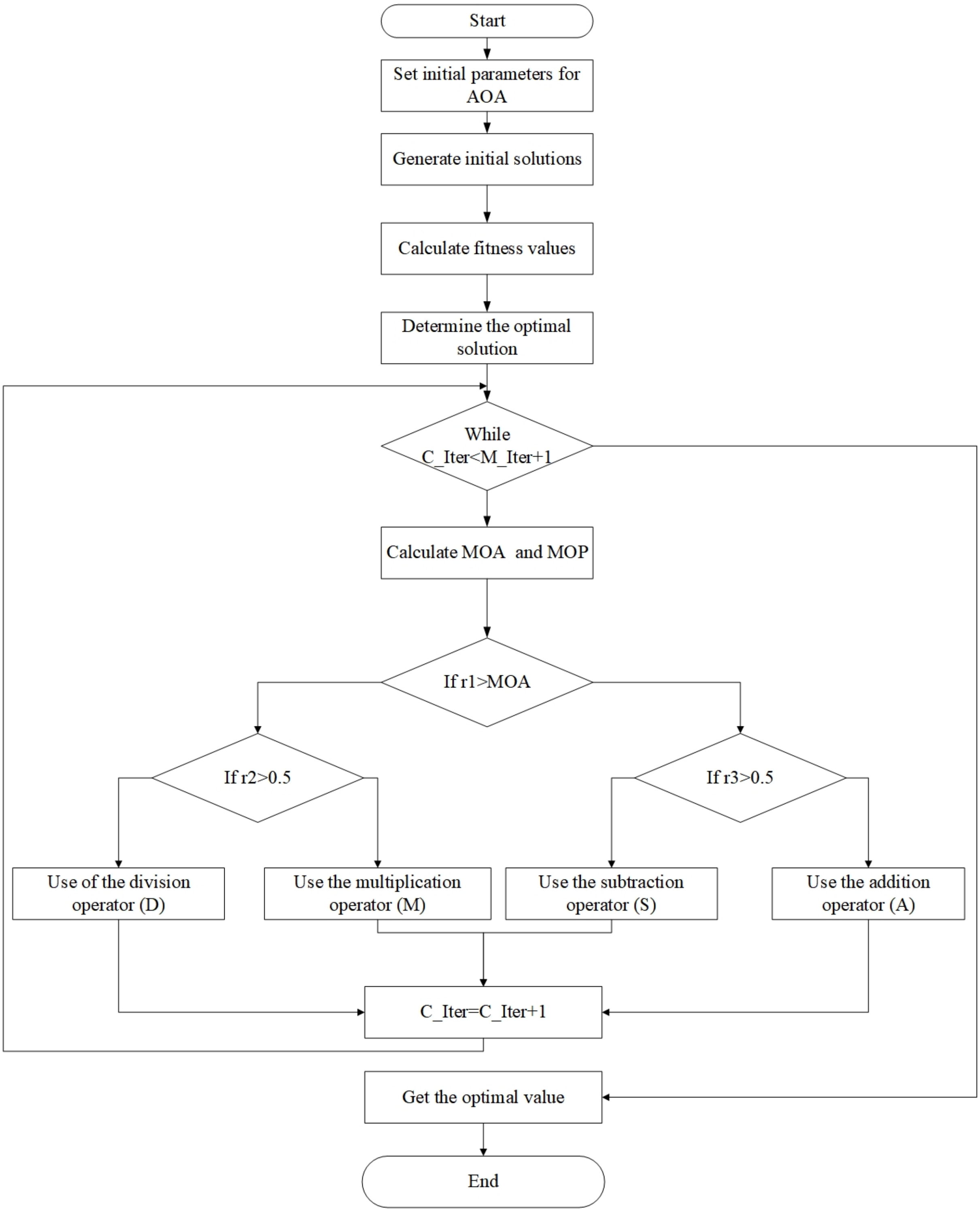

Fig. 9. MQRAOA Flowchart

were normalized to avoid the influence of feature value discrepancies on model training. Specifically, the mapminmax function was applied to linearly transform the data into the [0, 1] range, accelerating network convergence and improving model stability and robustness.

This study employs a hybrid model that combines the Extreme Learning Machine (ELM) and the Multi-Quasi-Reflective Arithmetic Optimization Algorithm (MQRAOA) for classification tasks. The ELM network consists of an input layer, a hidden layer, and an output layer. The input layer matches the input feature dimensions, the hidden layer contains 45 nodes with a Sigmoid activation function, and the output layer corresponds to the number of classification categories. In the standard ELM model, input weights and biases are randomly initialized, and the output weight matrix is determined using the least squares method.

To enhance classification accuracy, the MQRAOA algorithm is introduced to optimize ELM's input weights and biases. MQRAOA is a global optimization algorithm that combines multi-population search strategies with quantum mechanics principles to achieve efficient global optimization. Its main goal is to maximize classification accuracy. The algorithm uses a population size of 30 and up to 100 iterations. The optimized weights and biases obtained through this process are used to retrain and test the ELM model. Classification accuracy is then calculated on the test set to assess model performance.

To improve the robustness of the experimental results and reduce randomness, the optimization process was repeated 10 times, each starting from random initialization. Classification accuracy, fitness curves, and confusion matrices were recorded during each run. The average results of these 10 experiments were used to evaluate the optimization effect, reducing the risk of errors from a single experiment and ensuring the stability and reliability of the results.

Furthermore, the performance of the MQRAOA-ELM model was compared with several classic optimization algorithms, including WOA[15], SCA[16], SSA[17], and TLBO[18]. Each algorithm was executed for 10 iterations, and the average classification accuracy was computed for comparison. Fitness curves and confusion matrices were analyzed to demonstrate each algorithm's optimization effectiveness and classification performance, providing a comprehensive comparison of their advantages and disadvantages.

Figure 10 shows the fitness trends of various optimization algorithms in optimizing the ELM model for rolling bearing fault diagnosis. The results indicate that MQRAOA outperforms other algorithms, especially regarding convergence speed and optimization effectiveness. MQRAOA achieves rapid convergence, reaching near-optimal results in fewer than 50 iterations with a final fitness value close to 1.0. In contrast, WOA and SCA exhibit significant fluctuations in early iterations, especially in the first 30 iterations, indicating a tendency to get stuck in local optima during the global search phase. Although WOA and SCA stabilize later, with fitness values approaching 0.98, their convergence speed and overall optimization performance still lag behind MQRAOA.

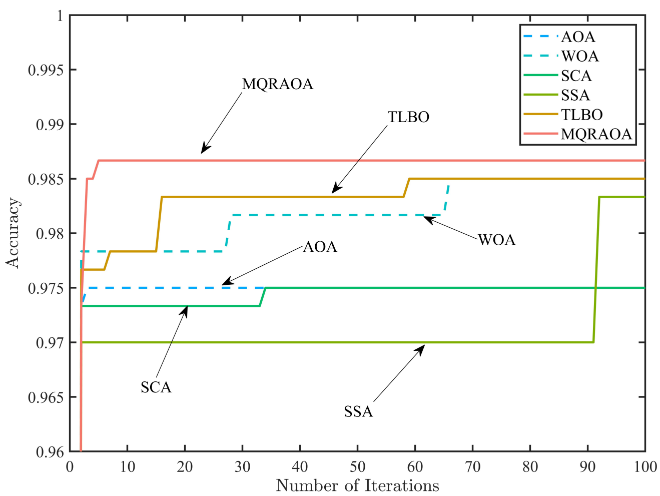

Fig. 10. Algorithm fitness plot

Similarly, SSA and TLBO show relatively smooth convergence patterns, with final fitness values approaching 0.99, slightly lower than MQRAOA. However, their fluctuations in the early iterations are less pronounced compared to WOA and SCA. The advantage of MQRAOA lies in its variation and quasi-reflection operations, which significantly enhance its global search ability, enabling it to more effectively locate near-optimal solutions while avoiding local optima. While other algorithms possess some global search capabilities, their convergence speed and final optimization results are insufficient when handling complex fault signals, highlighting the superior performance of MQRAOA in terms of convergence speed and optimization effectiveness.

Figures 11–13 and Table 1 clearly illustrate the performance of ELM models with different optimization algorithms in the rolling bearing fault classification task. These experimental results reveal the impact on classification accuracy and robustness, particularly when dealing with complex nonlinear fault signals. By analyzing the results of the standard ELM, AOA-ELM, and MQRAOA-ELM models, the importance of introducing optimization algorithms to improve classification performance becomes evident.

Figure 11 and Table 1 show that the standard ELM performs relatively well on simple categories (e.g., categories 1 and 4), with classification accuracies of 86.7% and 100%. However, its performance in categories 2 and 3 is poor, with only 46.7% and 53.3% accuracies. This highlights a major limitation in handling complex categories. Specifically, the left plot in Figure 11 shows that there is a significant deviation between predicted and true values for categories 2 and 3, leading to frequent misclassifications. This limitation stems from the standard ELM's weak ability to model nonlinear data, which makes it difficult to capture the complex distribution of such data, resulting in unstable overall classification performance.

In contrast, the AOA-optimized ELM classifier significantly improves performance, with classification accuracy increasing to 81.67%, as shown in Figure 12. According to Table 1, the classification accuracy for categories 1 and 4 remains stable at 86.7% and 100%, while for categories 2 and 3, accuracy improves to 70.0%. These improvements indicate that the AOA algorithm enhances the ELM model's ability to capture complex category features. In the left plot of Figure 12, the overlap between predicted and true values is significantly improved, and the number of misclassified samples is reduced, demonstrating the effectiveness of the AOA algorithm and indicating that the optimized model is more reliable when handling complex data.

The MQRAOA-optimized ELM classifier shows even greater enhancement, with outstanding performance, as illustrated in Figure 13 and Table 1. Its overall classification accuracy is 99.17%, the highest among all models. The classification performance for all categories is close to or reaches 100%, with categories 2 and 3 improving from 46.7% and 53.3% under the standard ELM to 100% and 96.7%, respectively. Furthermore, in the left plot of Figure 13, the predicted values almost perfectly align with the true values and misclassification points are nearly eliminated. These findings highlight the strong global optimization capability of the MQRAOA algorithm. Its efficient search mechanism and rapid convergence significantly enhance the parameter optimization of the ELM classifier, improving its ability to model and classify complex data.

A comprehensive analysis of Figures 11, 12, 13, and Table 1 leads to the conclusion that different optimization algorithms significantly affect the performance of the ELM classifier. While the standard ELM performs well on simple categories, its limited modeling ability restricts its perfor-

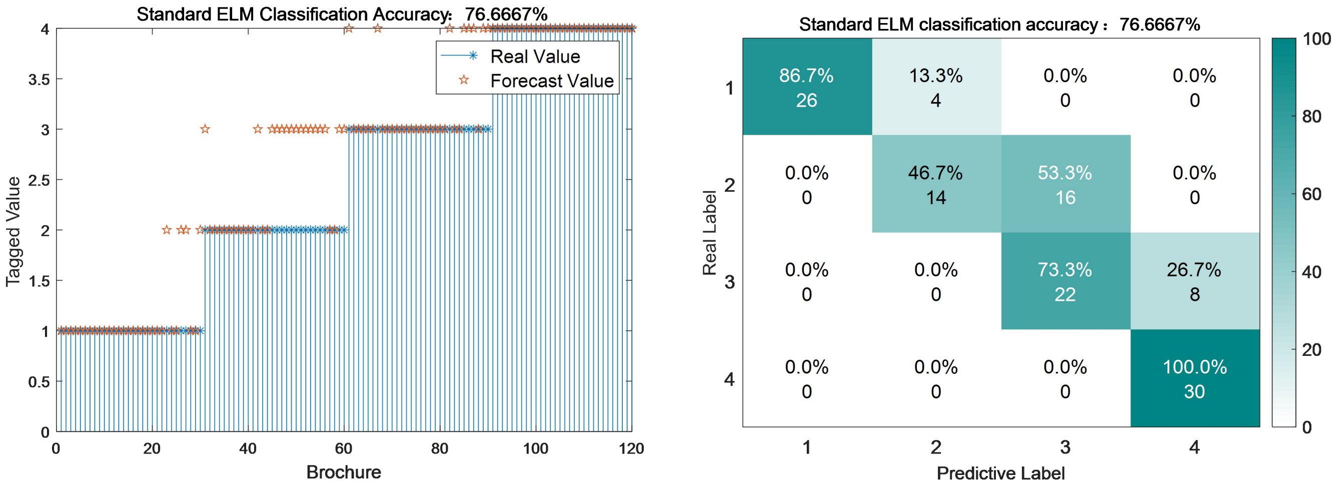

Fig. 11. Standard ELM classification results

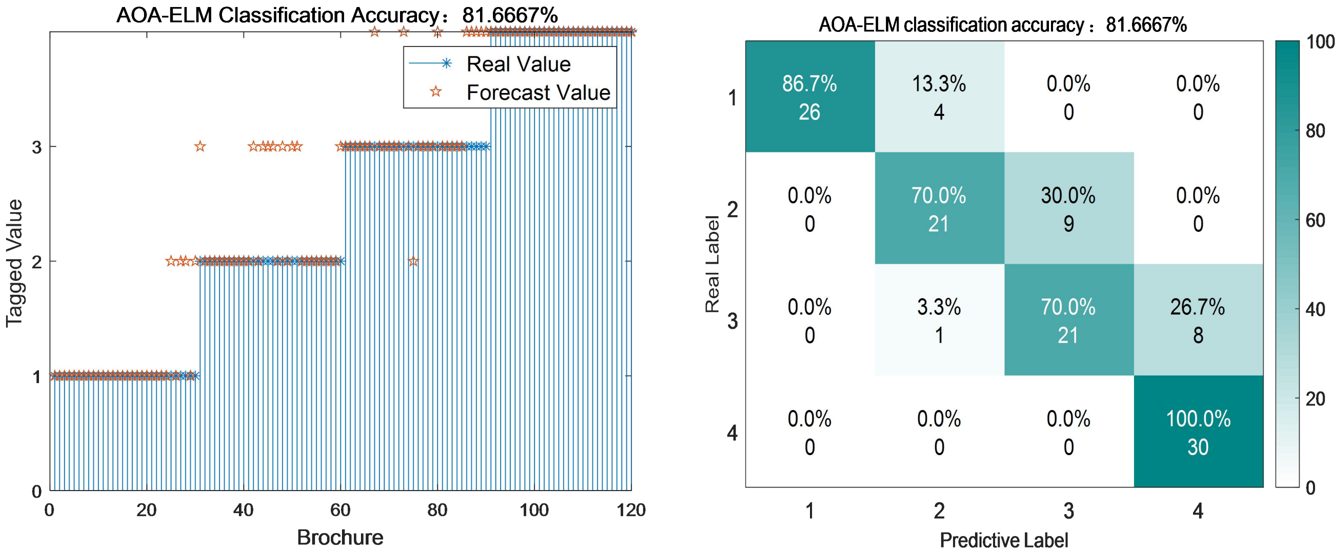

Fig. 12. AOA-ELM classification results

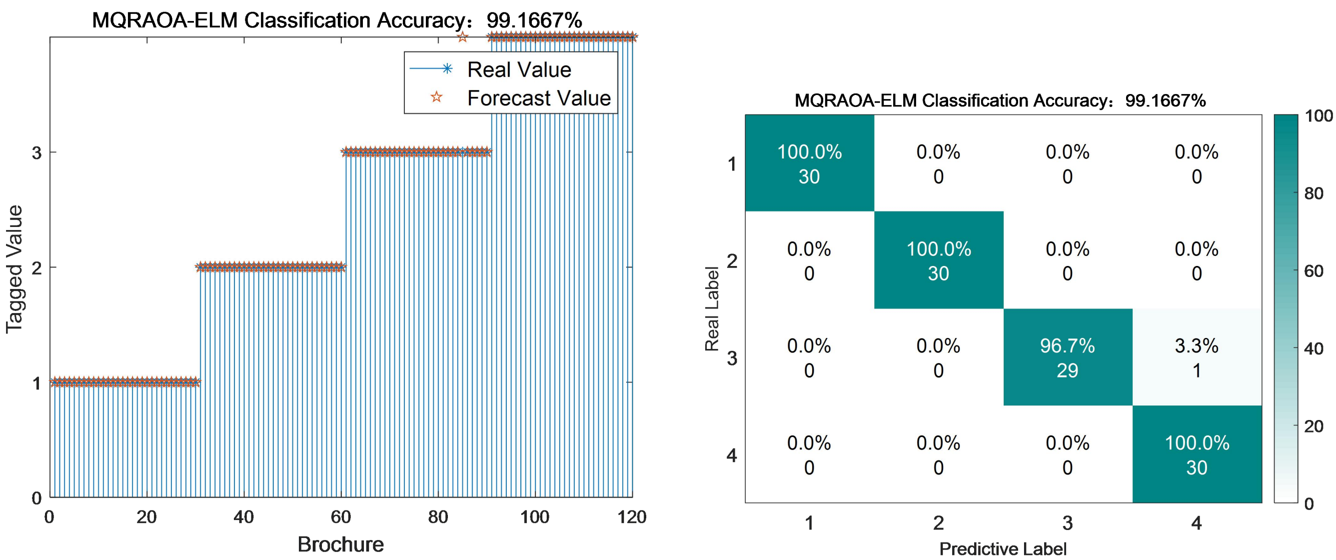

Fig. 13. MQRAOA-ELM classification results

TABLE I
CLASSIFICATION RESULTS FOR DIFFERENT CLASSIFIERS

Classifier	Category 1	Category 2	Category 3	Category 4	Average
Standard ELM	86.7%	46.7%	53.3%	100%	76.6667%
AOA-ELM	86.7%	70.0%	70.0%	100%	81.1667%
WOA-ELM	90.0%	63.3%	66.7%	100%	80.0000%
SCA-ELM	86.7%	70.0%	80.0%	100%	84.1667%
SSA-ELM	90.0%	66.7%	76.7%	100%	83.3333%
TLBO-ELM	86.7%	63.3%	80.0%	100%	82.5000%
MQRAOA-ELM	100.0%	100.0%	96.7%	100%	99.1667%

mance on complex categories. Optimization algorithms like AOA enhance the overall performance, particularly when dealing with complex categories. However, MQRAOA stands out, with its optimized ELM achieving nearly perfect classification accuracy across all categories, with an overall accuracy of 99.17%. These results demonstrate that MQRAOA, with its excellent optimization capability, provides a robust and efficient solution for complex fault signal classification tasks.

V. CONCLUSION

This study proposes an arithmetic optimization algorithm based on variational quasi-reflective learning ELM for rolling bearing fault diagnosis. The accuracy of fault diagnosis is significantly improved by extracting vibration signal features using CEEMDAN and CMAE.MQRAOA avoids the local optimum problem by optimizing the weights and biases of the ELM model and enhances the algorithm's global searching ability and convergence speed. The experimental results show that the MQRAOA-ELM model outperforms other algorithms' classification accuracy and robustness, verifying the method's effectiveness. The method provides a reliable solution to improve the efficiency and accuracy of rolling bearing fault diagnosis and has good application prospects.

REFERENCES

[1] H. Tang, S. Gao, L. Wang, X. Li, B. Li, and S. Pang, "A novel intelligent fault diagnosis method for rolling bearings based on Wasserstein generative adversarial network and convolutional neural network under unbalanced dataset," *Sensors*, vol. 21, no. 20, pp. 6754, 2021.

[2] J. Yang, C. Shi, T. Yang, and L. Wu, "A research on fault diagnosis method based on LMD and ABC optimized KELM," *Industrial Engineering Journal*, vol. 25, no. 3, pp. 124–131, 2022.

[3] H. Zhang, L. Ding, Y. Xue, and Y. Yang, "Research on fault diagnosis of rolling bearing based on VMD and permutation entropy," *Modular Machine Tool & Automatic Manufacturing Technique*, pp. 90–93, 2021.

[4] Z. Yu, B. Zhang, G. Hu, and Z. Chen, "Early fault diagnosis model design of reciprocating compressor valve based on multiclass support vector machine and decision tree," *Scientific Programming*, pp. 1–7, 2022.

[5] G. Xu, D. Hou, H. Qi, and L. Bo, "High-speed train wheel set bearing fault diagnosis and prognostics: A new prognostic model based on extendable useful life," *Mechanical Systems and Signal Processing*, vol. 146, pp. 107050, 2021.

[6] Z. Liu, Z. Lin, and C. Wang, "Kent-PSO optimized ELM fault diagnosis model in analog circuits," *Journal of Physics Conference Series*, vol. 1871, no. 1, pp. 012053, 2021.

[7] K. Yan, Z. Ji, H. Lu, J. Huang, W. Shen, and Y. Xue, "Fast and accurate classification of time series data using extended ELM: Application in fault diagnosis of air handling units," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 49, no. 7, pp. 1349–1356, 2019.

[8] C. Feng, Y. Li, and C. Feng, "Bearing failure diagnosis based on VMD sample entropy and improved extreme learning machine," *Journal of Gansu Agricultural University*, vol. 58, no. 2, 2023.

[9] D. Du, M. Wang, Z. Mao, et al., "Fault diagnosis of gearbox based on improved northern goshawk algorithm and hybrid kernel extreme learning machine," *Control Theory & Applications*, pp. 1–9, 2024.

[10] S. Yang, H. Wang, Y. Cui, C. Li, and Y. Tang, "Bearing fault diagnosis based on parameter optimized VMD with improved AFSA and ELM," *Journal of Vibration Engineering*, vol. 47, pp. 162–168, 2023.

[11] L. Abualigah, A. Diabat, S. Mirjalili, M. A. Elaziz, and A. H. Gandomi, "The arithmetic optimization algorithm," *Computer Methods in Applied Mechanics and Engineering*, vol. 376, pp. 113609, 2021.

[12] M. E. Torres, M. A. Colominas, G. Schlotthauer, and P. Flandrin, "A complete ensemble empirical mode decomposition with adaptive noise," *2011 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pp. 4144–4147, 2011.

[13] Z. Wu and N. E. Huang, "Ensemble empirical mode decomposition: A noise-assisted data analysis method," *Advances in Adaptive Data Analysis*, vol. 1, no. 01, pp. 1–41, 2009.

[14] W. N. Chen and J. Zhang, "A novel set-based particle swarm optimization method for discrete optimization problems," *IEEE Transactions on Evolutionary Computation*, vol. 17, no. 2, pp. 278–300, 2013.

[15] M. O. Okwu and L. K. Tartibu, "Whale Optimization Algorithm (WOA)," *Advances in Engineering Software*, vol. 53–60, 2020.

[16] S. Mirjalili, "Sca: A Sine Cosine Algorithm for Solving Optimization Problems," *Knowledge-Based Systems*, vol. 96, pp. 120–133, 2016.

[17] J. Xue and B. Shen, "A novel swarm intelligence optimization approach: Sparrow search algorithm," *Systems Science & Control Engineering*, vol. 8, no. 1, pp. 22–34, 2020.

[18] R. V. Rao, V. J. Savsani, and D. P. Vakharia, "Teaching-learning-based optimization: A novel method for constrained mechanical design optimization problems," *Computer-Aided Design*, vol. 43, no. 3, pp. 303–315, 2011.