DIAGNOSTYKA, 2026, Vol. 27, No. 2

e-ISSN 2449-5220
DOI: 10.29354/diag/221643

INTELLIGENT DIAGNOSIS METHOD FOR ROLLING BEARINGS BASED ON ADAPTIVE SIGNAL PROCESSING AND PARAMETER OPTIMIZATION

Xinjun JIANG [1], Baohua WANG [1], Guoxing WU [1], Jun HU [1], Ge YIN [1], Han YAN [1], Xinrong HE [1], Jing ZHU [2,*]

[1] Guoneng Changzhou Second Power Generation Co., Ltd., Changzhou 213000, Jiangsu, China
[2] Henan University of Science and Technology, Luoyang 471003, Henan, China
[*] Corresponding author, email: 1907175086@qq.com

Abstract

To address the limitations of conventional deep-learning-based fault diagnosis methods in complex pattern recognition tasks, particularly insufficient feature extraction and suboptimal parameter configuration, this paper proposes a novel IDCS-SSA-BiTCN framework. The proposed method integrates adaptive signal processing with deep learning in a unified diagnosis architecture. In the signal enhancement stage, an improved singular spectrum analysis (SSA) method is employed, in which the embedding dimension is optimized using the IDCS algorithm and the decomposed components are reconstructed according to a correlation-coefficient matrix, thereby achieving adaptive feature enhancement and noise suppression. In the modeling stage, a bidirectional temporal convolutional network (BiTCN) is constructed, and its key hyperparameters are globally optimized by IDCS to improve fault recognition accuracy and generalization capability across different fault categories. Furthermore, to enhance the stability and convergence efficiency of high-dimensional optimization, the original differential creative search (DCS) algorithm is improved by incorporating a nonlinear parameter control strategy and a Lévy flight mechanism. Experiments conducted on the rolling bearing datasets from Xi'an Jiaotong University and Southeast University demonstrate that the proposed method achieves accuracies of 99.20% and 98.75%, respectively, indicating excellent fault-classification performance.

Keywords: rolling bearing fault diagnosis; singular spectrum analysis; intelligent optimization algorithm

1. INTRODUCTION

As a key component of rotating machinery, the operating status of rolling bearings is directly related to the stability and safety of the equipment. In critical domains such as aerospace, rail transportation, and intelligent manufacturing, bearing failures can lead to severe equipment damage, operational downtime, and catastrophic safety incidents. Therefore, efficient and reliable fault diagnosis has important engineering significance and application value [1, 2]. With the increasing complexity of industrial systems, bearing fault signals show strong nonlinearity, non-stationarity, and high noise, which puts higher requirements on the robustness and accuracy of diagnostic algorithms.

At present, research on rolling bearing fault diagnosis mainly focuses on two types of methods: one is the traditional method based on signal processing, such as time domain statistical analysis, frequency domain spectrum analysis, wavelet transform and singular spectrum analysis (SSA) [3-4]. This type of method focuses on extracting key statistical features from vibration signals for pattern recognition and fault classification. However, this type of method is insensitive to feature changes under strong noise or variable working conditions, and is highly dependent on artificial experience parameter settings, making it difficult to achieve adaptive extraction of fault information [5].

Another important category of bearing fault diagnosis methods is based on deep neural networks, such as convolutional neural networks (CNNs), recurrent neural networks (RNNs), and residual networks (ResNets), which have been widely applied in industrial signal classification and intelligent fault diagnosis of rotating machinery [6-8].These methods possess strong representation-learning capability and can automatically extract multi-level structural information from raw signals, thereby improving fault identification accuracy. However, deep models still face three major challenges when dealing with raw vibration data: (1) the quality of the input features is often degraded by redundant information and environmental noise, which interferes with deep feature learning; (2) the network architecture is usually complex, while the hyperparameter configuration is often determined

Received 2025-08-28; Accepted 2026-05-10; Available online 2026-05-11

empirically and may not match different fault patterns well; and (3) the training process may suffer from local optima, unstable convergence, and limited generalization capability. Under real industrial operating conditions, where the data distribution changes dynamically and frequently, these issues further reduce the reliability of conventional deep-learning-based diagnosis models [9].

In summary, standalone signal processing methods struggle to meet feature extraction demands under complex working conditions, while relying solely on deep learning renders models vulnerable to the dual constraints of input data quality and architectural sub-optimality. Therefore, building an end-to-end diagnostic model that integrates signal processing and intelligent optimization mechanisms is an effective way to improve the overall performance of the diagnostic system. In recent years, some studies have attempted to integrate methods such as wavelet transform and empirical mode decomposition (EMD) with neural networks to achieve efficient recognition of some fault types [10]. Yu et al. [11] used wavelet packet transform to construct multi-scale time-frequency feature maps and global statistical feature matrices of vibration signals, and then designed a deep feature extraction network combining ResNet and self-attention mechanism to achieve fusion extraction of local and global time-frequency features. Hamdaoui et al. [12] used variational mode decomposition and wavelet threshold to denoise vibration signals, and used time-frequency continuous wavelet transform to generate scale map images, and finally combined deep learning structures to classify and diagnose bearing faults. Zhang et al. [13] used synchronous compressed wavelet transform to convert to a higher resolution two-dimensional time-frequency image, and used an improved and optimized deep convolutional neural network model to extract high-level features of the time-frequency image, and further input the extracted high-level features into the extreme learning machine classifier for fault classification. Jiang et al. [14] combined continuous wavelet transform and deep residual network for fault diagnosis. Ren et al. [15] used wavelet packet transform to construct a time-frequency feature matrix (MWCM); used a shallow robustly optimized residual network to extract features from the MWCM, and used a visual converter to extract time-frequency global features from the MWSM. However, there are still problems of strong manual intervention and lack of adaptability at the two levels of feature extraction and network structure design. In addition, the hyperparameter tuning of deep networks often uses inefficient methods such as grid search and genetic algorithms, which makes it difficult to obtain a stable optimal solution in high-dimensional space [16].

To address the above issues, this paper proposes an intelligent fault diagnosis framework, termed IDCS-SSA-BiTCN, that integrates adaptive signal enhancement with deep model optimization. In the signal preprocessing stage, an improved SSA method is employed to construct a w-correlation map for component reconstruction, while the improved differential creative search algorithm (IDCS) is used to determine the optimal embedding dimension, thereby enhancing feature integrity and robustness. In the model construction stage, a BiTCN architecture is introduced to strengthen temporal feature modeling, and its key hyperparameters are further optimized by IDCS to improve the compatibility between the network structure and the processed input features. The main contributions of this paper are summarized as follows:

1) An intelligent diagnosis framework integrating signal enhancement and structural optimization is proposed, systematically resolving the bottlenecks of inadequate feature extraction and model fragility;
2) An improved difference creative search algorithm (IDCS) was designed, and nonlinear control strategy and Levy flight mechanism were introduced to enhance the convergence performance and global exploration ability of the algorithm in high-dimensional search space;
3) A SSA component aggregation method based on correlation graph was proposed, and the weighted periodic consistency index between components was used to guide the reconstruction of high-quality periodic features;
4) A dynamic structural optimization strategy for the BiTCN model is developed, significantly boosting the generalization performance of the diagnostic model across diverse fault modes and varying operating conditions.

The experimental results show that the method proposed in this paper achieves recognition accuracy rates of 99.20% and 98.75% respectively on the bearing datasets of Xi 'an Jiaotong University and Southeast University, significantly outperforming the existing mainstream methods. It demonstrates its outstanding performance and has high engineering application value and promotion potential.

2. PROPOSED METHOD

In order to improve the denoising and feature extraction capabilities of rolling bearing fault diagnosis systems under complex working conditions, this paper constructs an intelligent diagnosis framework that integrates signal optimization and deep model parameter adjustment. The framework consists of three parts: improved singular spectrum analysis (SSA) is used to adaptively extract key periodic features in vibration signals; bidirectional temporal convolutional network (BiTCN) is used to capture time-dependent structures and realize fault classification; and the improved difference creative search algorithm (IDCS) runs through the above two modules, which are used for embedding dimension selection and

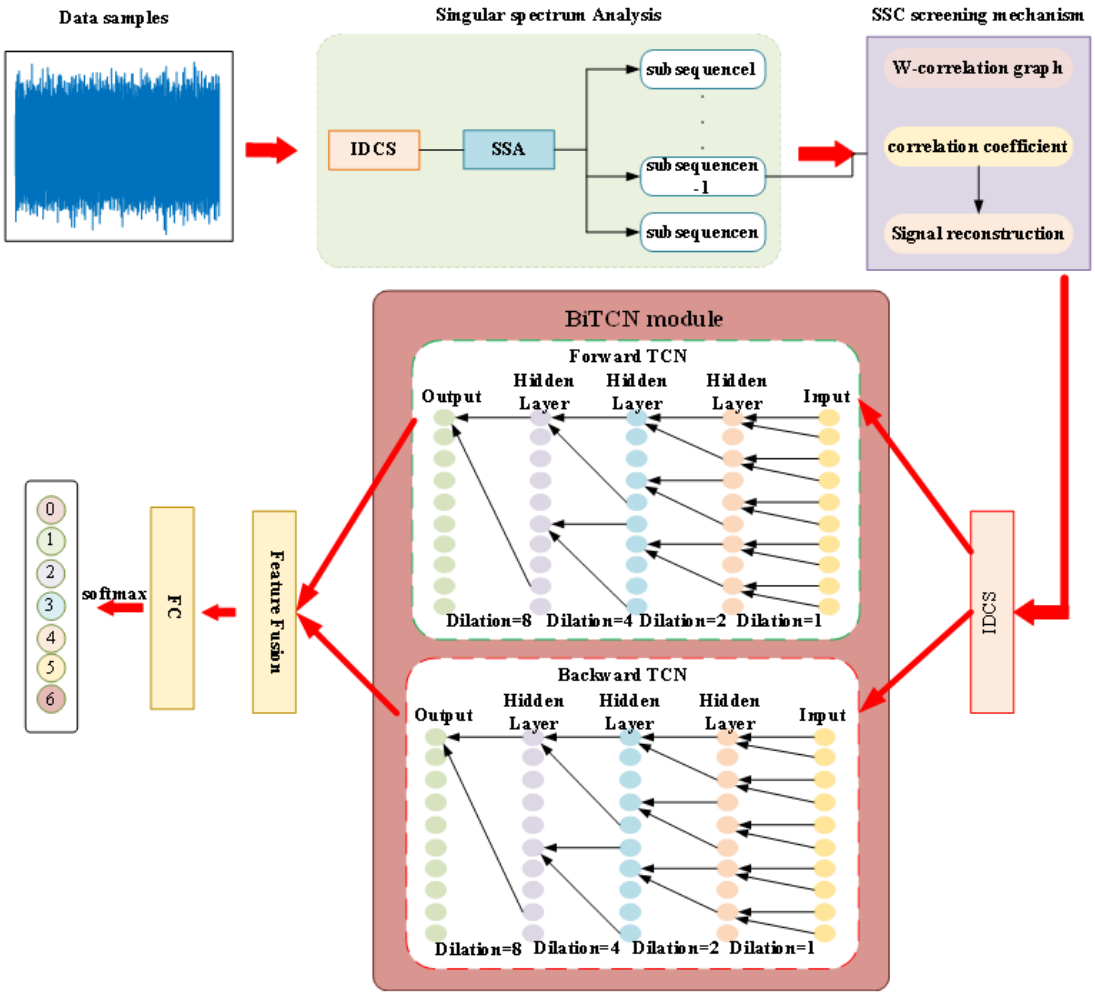


Fig. 1. Model structure diagram(Note: Inside the BiTCN module, the data stream flows from right to left, meaning the raw input enters from the right side and propagates to the left. Correspondingly, the dilation factor correctly increases exponentially as 1, 2, 4, and 8 from the input layer to the output layer

component reconstruction in SSA, and key hyperparameter search in BiTCN, respectively, to construct a parameter adaptive mechanism.

This method not only integrates the advantages of traditional signal processing and deep learning models, but also realizes unified coordination and global performance drive between modules through intelligent optimization algorithms, effectively solving key problems such as insufficient feature extraction, model configuration dependence on experience, and unstable training. Its overall structure is shown in fig. 1.

2.1. Improved Differential Creativity Optimization Algorithm (IDCS)

In rolling bearing fault diagnosis, the choice of optimization algorithm is crucial for extracting representative fault features from raw vibration signals and for efficiently tuning the parameters of deep learning models. Mainstream meta-heuristic optimization methods, such as particle swarm optimization (PSO), genetic algorithm (GA), differential evolution (DE), and whale optimization algorithm (WOA), often suffer from premature convergence, limited search capability, or strong dependence on parameter tuning when handling nonlinear or high-dimensional problems. By contrast, the differential creative search (DCS) algorithm provides stronger population diversity, richer knowledge evolution behavior, and better adaptability to complex optimization tasks. Therefore, DCS is selected in this study as the baseline optimization framework.

Although the DCS algorithm has excellent performance in multidimensional optimization problems, it still has bottlenecks in convergence speed and jumping out of local optimality. In order to improve the performance of the Differential Creative Search (DCS) algorithm in high-dimensional fault diagnosis optimization tasks, this paper proposes an improved Differential Creative Search (IDCS) algorithm. By introducing nonlinear parameter control strategy and Levy flight mechanism, the convergence of the search process, the ability to jump out of local optimality and parameter adaptability are effectively enhanced. At

the same time, this paper integrates and reconstructs multiple strategies of the original DCS to build a more efficient and stable optimization framework. The IDCS algorithm proposed in this paper not only optimizes the search strategy in theory, but also serves as a core optimization component in the fault diagnosis system, and is applied to: 1) Signal processing stage: IDCS is used to optimize the embedding dimension and component reconstruction selection in SSA decomposition to achieve adaptive denoising and feature enhancement of vibration signals, and improve the input quality of subsequent models; 2) Model training stage: IDCS is used to optimize key hyperparameters such as convolution kernel size, number of convolution channels, and number of hidden layer neurons in the BiTCN network to avoid the subjectivity of empirical parameter adjustment and improve network generalization performance and accuracy. IDCS proposes the following three key improvements based on the original algorithm, which reflects the innovation of this paper in optimizer design:

1) Nonlinear parameter control strategy

The mutation intensity parameter in the original DCS algorithm is set to a fixed value or a simple linear change, lacking dynamic adaptive adjustment of the search process, which can easily lead to falling into the local optimum too early in the early stage of the search, or failing to effectively converge to the exact solution in the later stage. To solve this problem, this paper introduces a nonlinear parameter control strategy to dynamically adjust the parameters in the mutation and crossover operations so that they gradually decrease during the search process, thereby improving the convergence speed and global search ability of the algorithm. The parameter a changes with the number of iterations as follows:

$$a = cos\left(\frac{\pi}{2}\left(\frac{l}{Max_{iter}}\right)^4\right) \tag{1}$$

Among them, l is the current iteration number, and Maxiter is the maximum iteration number. This function shows a nonlinear decreasing trend, so that in the early stage of the search, the algorithm maintains a strong perturbation ability, which is conducive to exploring the potential optimal area in a large range; in the later stage, the perturbation intensity gradually weakens, which helps the algorithm focus on local development and improves the final convergence accuracy and stability.

2) Levy flight strategy

Although the DCS algorithm achieves a certain degree of diversity preservation by simulating team role differences, it is still easy to fall into local optimality in high-dimensional and complex search spaces, especially in the divergent thinking stage. Its position update mechanism relies on conventional Gaussian perturbations and has a limited search range. In order to improve the ability to jump out of local optimality, this paper introduces a divergent search strategy based on Levy flight. This strategy uses the random step length of the long-tail distribution to introduce sparse and long-distance jumping behavior in the individual generation stage. Its step length is generated by Levy distribution:

$$\Delta x = 0.01\frac{u}{|v|^{\frac{1}{\beta}}} \tag{2}$$

where u and v follow normal distributions N(0,σ_u^2) and N(0,σ_v^2), respectively, and

$$\sigma_u = \left(\frac{\Gamma(1+\beta)\,sin\left(\frac{\pi\beta}{2}\right)}{\Gamma\left(\frac{1+\beta}{2}\right)\beta 2^{\frac{\beta-1}{2}}}\right)^{\frac{1}{\beta}} \tag{3}$$

This mechanism enables individuals to have the dual capabilities of "short-distance meticulous search" and "occasional long jumps" near the current optimal solution, significantly improving the probability of IDCS jumping out of local extreme values. Especially in high-dimensional non-convex spaces, this strategy can effectively expand the search range and improve the coverage of the optimal solution.

3) Optimize the algorithm behavior logic and execution process

The original DCS has problems such as complex behavior triggering rules and fuzzy update paths at the implementation level, especially the lack of clear behavior judgment boundaries between "differentiated knowledge acquisition", "convergent thinking" and "divergent thinking", resulting in reliance on a large number of hyperparameter controls and unstable behavior in actual execution. This paper systematically reconstructs the algorithm execution process, based on the principle of "role drive + behavior mapping", clarifies the trigger conditions and update paths of three types of core behaviors, and coordinates nonlinear control functions and probabilistic control mechanisms for scheduling management: 1) Divergent search stage: high-performance individuals conduct remote exploration according to the Levy flight strategy; 2) Convergent search stage: medium and low-performance individuals build target guidance paths according to the current optimal individuals; 3) Diversity maintenance stage: low-performance individuals are replaced by newly generated individuals to ensure population vitality.

In addition, a boundary correction mechanism and a simplified retrospective evaluation mechanism (RA) are added to ensure that the population individuals are always in the feasible solution space, while achieving continuous tracking of the historical optimal solution. The overall execution logic is clearer, the operation process is more stable, and it is more suitable for deployment and reproduction in complex engineering tasks.

2.2. SSA optimization reconstruction method based on w-correlation

Singular spectrum analysis is a signal decomposition method that decomposes time series signals into trend terms, periodic terms, and noise components. It has good trend extraction and noise reduction capabilities in rolling bearing fault

diagnosis. However, the traditional SSA method still faces two key challenges in practical applications: the empirical selection of embedding dimensions and the lack of scientific criteria for subcomponent reconstruction, which greatly limits its diagnostic reliability under complex working conditions. To address these problems, this study proposes an improved SSA optimization framework to improve the accuracy and stability of feature extraction.

The embedding dimension determines the fineness of SSA decomposition and is a key control parameter for SSA performance. In order to avoid structural errors that may be caused by subjective settings, this paper models the selection of the embedding dimension L as an optimization problem and introduces an improved difference creative search algorithm (IDCS) for solution. The optimization objective function is as follows:

$$F(L) = \alpha \cdot MSE(L) + \beta \cdot H(L) \tag{4}$$

Where, MSE(L) is the reconstruction error when the embedding dimension is L; H(L) is the information quantity measure (such as entropy or singular value energy) when the embedding dimension is L; α and β are weight coefficients used to balance the error and information quantity. It is worth noting that relying solely on MSE could lead to signal over-smoothing, potentially eliminating subtle fault impulses. By incorporating the information measure H(L) into the multi-objective fitness function, the IDCS algorithm effectively balances noise reduction and feature preservation. Furthermore, the subsequent w-correlation clustering mechanism ensures that transient fault impulses are accurately identified and aggregated rather than being smoothed out.

By searching for L* that satisfies the maximum fitness, the reconstruction structure with the largest amount of information and closest to the original signal can be obtained, thereby improving the accuracy and stability of subsequent diagnosis.

At the same time, in traditional singular spectrum analysis, a series of subcomponents are generated after the signal is decomposed, and each subcomponent corresponds to a different characteristic scale. However, in the context of multi-periodic signals or strong noise, these periodic components are often dispersed into different subcomponents, resulting in confusion of periodic information and unconcentrated features, which in turn affects the reconstruction quality and subsequent diagnostic performance. Therefore, accurately identifying and reconstructing subcomponents with similar periodic characteristics is a key link to improve the practicality and diagnostic reliability of SSA.

To solve the above problems, this paper introduces a correlation spectrum method (w-correlation) based on weighted correlation to measure the periodic similarity between each SSA subcomponent. This method systematically evaluates the similarity relationship between components by constructing a global correlation matrix, thereby guiding the identification and aggregation of periodic features. Specifically, w-correlation is defined as:

$$w_{ij} = \frac{\langle x_i, x_j \rangle}{\sqrt{\langle x_i, x_i \rangle \langle x_j, x_j \rangle}} \tag{5}$$

Among them, x_i and x_j represent the i-th and j-th subcomponents obtained by SSA decomposition, respectively, and the inner product operation ⟨x_i,x_j ⟩ represents the weighted inner product between the two:

$$\langle x_i, x_j \rangle = \sum_{t=1}^{N} \omega(t) x_i(t) x_j(t) \tag{6}$$

In the formula, ω(t) is a weighting function used to adjust the contribution weight of data at different time points. This weighting function ensures that the weight of the edge points of the time series is small and the weight of the center point is large, which effectively reduces the impact of the boundary effect on the component correlation evaluation.

By calculating ω_ij between all sub-component pairs, a symmetric correlation matrix can be constructed, which can quantitatively characterize the structural consistency between components. When the value of ω_ij exceeds the threshold, it indicates that the two components have significant periodic consistency, which may originate from the same potential periodic signal and should be considered for aggregation.

Although the improved SSA method significantly improves the effectiveness of signal processing, it is difficult to cope with the challenges of fault pattern recognition under complex working conditions by relying solely on signal processing technology. The operating environment of modern industrial equipment is complex and changeable, and fault characteristics are often highly nonlinear and non-stationary, which requires the diagnostic system to have strong adaptive learning capabilities. Therefore, this paper further introduces deep learning technology and constructs an intelligent fault recognition framework based on a bidirectional temporal convolutional network (BiTCN). The high-quality features optimized by SSA are used as input to achieve accurate classification of fault types.

2.3. BiTCN structure optimization strategy based on IDCS

As a deep learning architecture that combines convolutional neural networks and bidirectional information flow, Bidirectional Temporal Convolutional Network (BiTCN) has shown excellent performance in time series data processing. The network can effectively process the vibration signal features optimized by SSA, thereby realizing accurate fault diagnosis of rolling bearings. However, the effectiveness of the BiTCN model depends largely on the configuration of its key hyperparameters, which have a direct impact on the model's feature extraction ability and fault identification accuracy. Traditional parameter tuning methods such as grid search or random search are inefficient and prone to local optimality when

processing high-dimensional parameter spaces, making it difficult to meet the accuracy requirements of complex fault diagnosis of rolling bearings. To address this challenge, this paper proposes to use an improved differential creative search algorithm (IDCS) to globally optimize the key parameters of the BiTCN model and construct an adaptive parameter tuning framework.

The BiTCN model integrates the causal convolution characteristics of the temporal convolutional network with the bidirectional information processing capability. Its core structure includes components such as multi-layer causal convolution, residual connection, bidirectional information fusion, and attention mechanism. Among them, the convolution kernel size k determines the receptive field range of the model, which directly affects the ability to capture fault features of different time scales; the number of convolution channels c controls the dimensional richness of feature expression, which is closely related to the complexity of the fault mode; the expansion factor d adjusts the modeling ability of long-term dependencies, which is particularly important for periodic fault features. In addition, parameters such as the number of network layers l, learning rate (η), batch size (b) and dropout rate (dr) also jointly affect the training stability and generalization ability of the model. The selection of these parameters usually relies on expert experience and lacks a systematic optimization method, making it difficult to cope with the multi-class fault diagnosis needs of rolling bearings under complex working conditions.

In this study, the hyperparameter optimization of BiTCN is formulated as a multivariable constrained optimization problem with the decision vector $X = [k, c, d, l, \eta, b, dr]$. The objective of IDCS is to maximize the classification accuracy on the validation set, whereas the held-out test set is strictly excluded from the hyperparameter search and reserved only for the final unbiased performance evaluation. The candidate ranges of the hyperparameters are determined based on theoretical analysis and preliminary experiments as follows: convolution kernel size $k \in \{3, 5, 7, 9, 11\}$, number of convolution channels $c \in \{16, 32, 64, 128, 256\}$, dilation factor $d \in \{1, 2, 4, 8, 16\}$, number of layers $l \in \{2, 3, 4, 5, 6\}$, learning rate $\eta \in [0.0001, 0.01]$, batch size $b \in \{16, 32, 64, 128\}$, and dropout rate $dr \in [0.1, 0.5]$. These search ranges cover the characteristic scales of typical bearing fault features while remaining consistent with practical computational constraints.

2.4. Algorithm flow of this paper

In order to achieve efficient diagnosis from raw vibration signals to fault type identification, this paper constructs an intelligent fault diagnosis system that integrates signal enhancement and model parameter optimization. The whole process consists of four stages: data input, feature extraction, parameter optimization and fault classification. The overall process is shown in fig. 2, which specifically includes the following key steps:

First, the system receives the collected raw rolling bearing vibration signal as input. In order to improve the quality of subsequent modeling, the signal is first preprocessed by the improved singular spectrum analysis (SSA) method. In this stage, the improved difference creative search algorithm (IDCS) is used to adaptively search the embedding dimension in SSA to optimize the trajectory matrix construction method; then the w-correlation map is constructed to quantify the periodic consistency between SSA components, further screen and reconstruct the periodic components with high correlation, and output a noise reduction signal with a significant characteristic structure.

Secondly, the processed signal is sent to the bidirectional temporal convolutional network (BiTCN) for fault identification. The model can simultaneously model the forward and backward time-dependent structures of the signal and effectively capture multi-scale time series features. In order to improve the accuracy and stability of the BiTCN model, this paper introduces the IDCS algorithm to globally optimize its key structural parameters (including convolution kernel size, number of channels, network depth, number of neurons in the fully connected layer, etc.), and construct an adaptive mapping relationship between the deep structure and the input features.

Finally, the trained and optimized BiTCN model discriminates the input signal and outputs the corresponding fault category label to realize automatic diagnosis of the rolling bearing status. Through this process, the system realizes a complete closed-loop optimization from signal acquisition, feature extraction to classification and discrimination, effectively improving the accuracy, stability and engineering practicality of fault diagnosis.

To eliminate information leakage, each dataset was first divided chronologically into three contiguous and non-overlapping subsets: 60% for training, 10% for validation, and 30% for final testing. The split was performed on the original continuous vibration signal before sample generation, so that no overlapping windows from adjacent time intervals were assigned to different subsets. The training subset was used to fit the model parameters. The validation subset was used exclusively for IDCS-driven hyperparameter optimization, including the SSA embedding dimension, component-selection-related settings, and the BiTCN structural hyperparameters. The test subset was strictly held out during the entire model development process and was accessed only once for the final performance evaluation. Therefore, no testing samples were involved in parameter tuning, threshold selection, model selection, or early stopping.

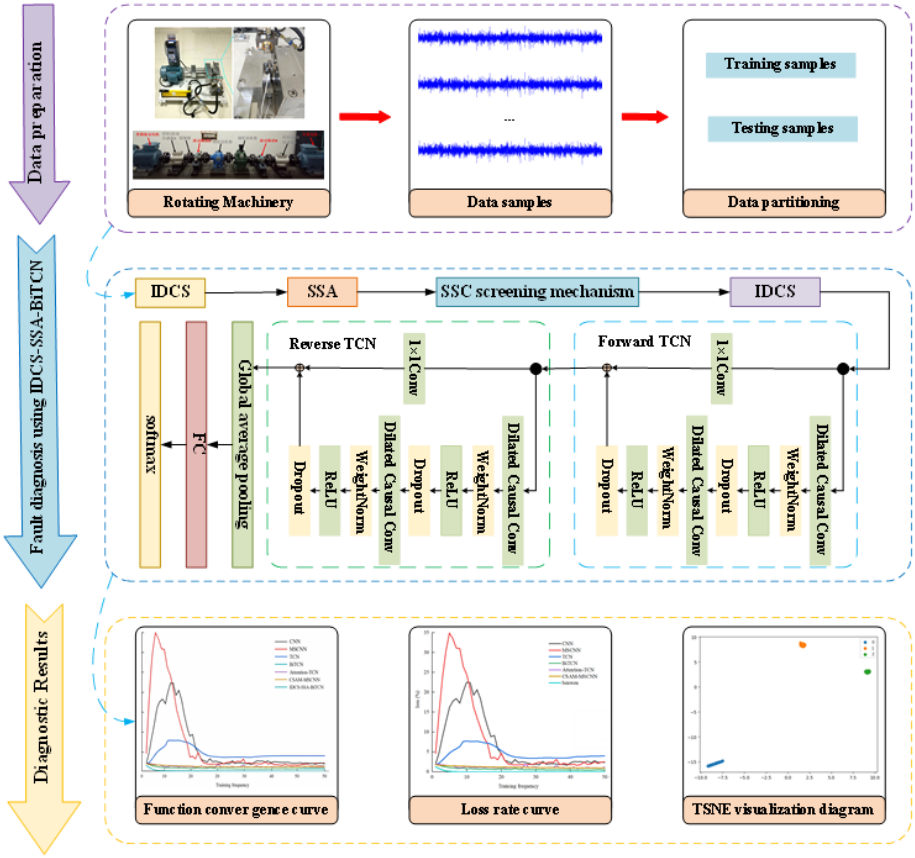


Fig. 2. Overall flow chart

3. EXPERIMENTAL VERIFICATION

3.1. IDCS algorithm performance test

In order to verify the performance of the proposed method, the particle swarm optimization algorithm (PSO), dynamic adaptive bee colony optimization algorithm (SCHO), goose optimization algorithm (GOOSE), differential creative search algorithm (DCS) and improved differential creative search algorithm (IDCS) were selected for comparative experiments. The maximum number of iterations was set to 500 times and the population size was set to 100. 12 test functions were selected for analysis.

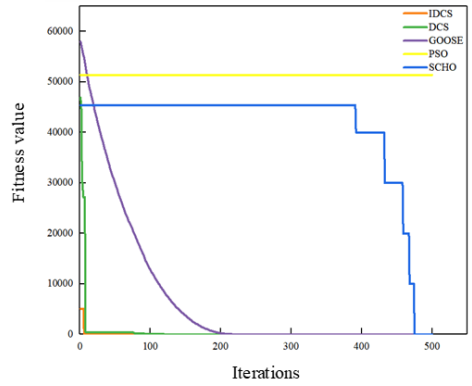
1) High-dimensional single objective function test

Table 1 shows 4 high-dimensional single objective test functions. Such functions have only one global optimal solution, which can test the convergence and local development performance of the algorithm. The curves of the 4 high-dimensional single objective functions are shown in fig. 3.

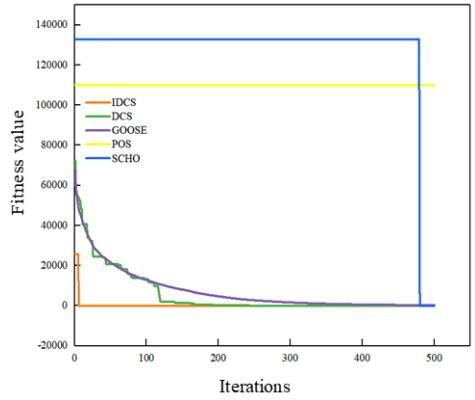
Table 1. Four high-dimensional single objective functions (Dim = 30)

function	scope	Minimum function
$f_1(x)=\sum_{i=1}^{n}x_i^2$	[-100,100]	0
$f_2(x)=\sum_{i=1}^{n}\left(\sum_{j=1}^{i}x_j\right)^2$	[-100,100]	0
$f_3(x)=max\{\lvert x_i\rvert, 1\le i\le n\}$	[-100,100]	0
$f_4(x)=\sum_{i=1}^{n}([x_i+0.5])^2$	[-100,100]	0

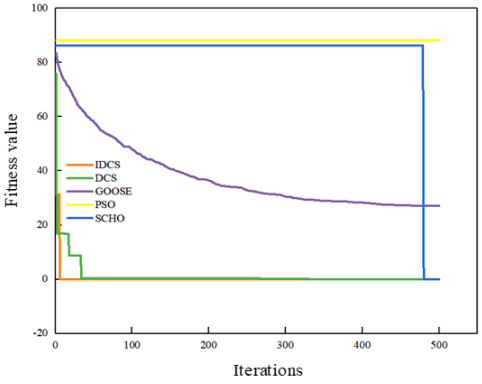
As can be seen from the figure, when the IDCS algorithm solves the four high-dimensional single-objective optimization problems of $f_1(x)$, $f_2(x)$, $f_3(x)$ and $f_4(x)$, it not only converges to the optimal solution quickly at a rate ahead of other existing algorithms, but also maintains a high degree of stability throughout the optimization process, reducing the risk of solution quality degradation caused by fluctuations during the iteration process.



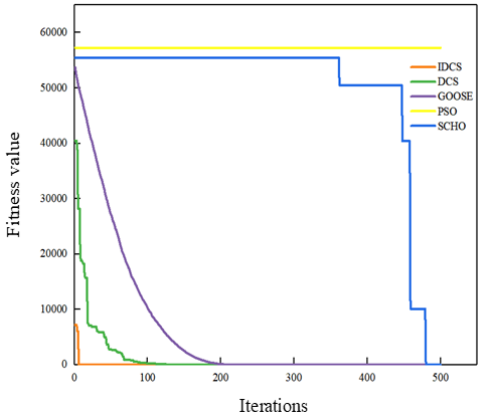
Convergence curve of function f1



Convergence curve of function f2



Convergence curve of function f3



Convergence curve of function f4

Fig. 3. Convergence curve of high-dimensional single objective function

2) High-dimensional multi-objective function test

Table 2 gives four high-dimensional multi-objective test functions, each with one global optimal solution and multiple local optimal solutions, which are used to test the local and global search capabilities of the algorithm. The curve of the high-dimensional single-objective function is shown in Fig. 4.

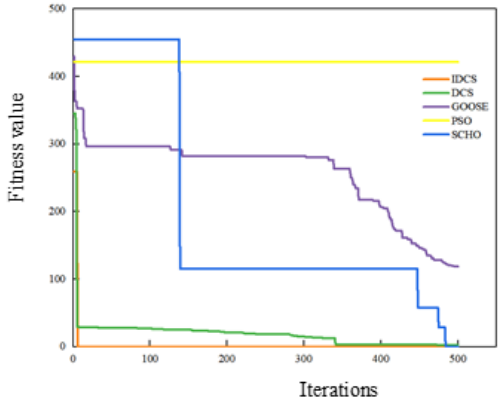
Table 2. Four high-dimensional multi-objective functions (Dim = 30)

function	scope	Minimum function
$f_5(x)=\sum_{i=1}^{n}\left[x_i^2-10\cos(2\pi x_i)+10\right]$	[-5.12, 5.12]	0
$f_6(x)$ $=-20\,exp\left(-0.2\sqrt{\frac{1}{n}\sum_{i=1}^{n}x_i^2}\right)$ $-exp\left(\frac{1}{n}\sum_{i=1}^{n}cos(2\pi x_i)\right)+20+e$	[-32,32]	0
$f_7(x)=\frac{1}{4000}\sum_{i=1}^{n}x_i^2-\prod_{i=1}^{n}cos\left(\frac{x_i}{\sqrt{i}}\right)+1$	[-600,600]	0
$f_8(x)=\frac{\pi}{n}\left\{10\,sin(\pi y_1)+\sum_{i=1}^{n-1}(y_i-1)^2[1+\right.$ $\left.+\sum_{i=1}^{n}u(x_i,10,100,4),y_i=\right.$ $u(x_i,a,k,m)=\begin{cases}k(x_i-a)^n\\0\\k(-x_i-a)\end{cases}$	[-50,50]	0

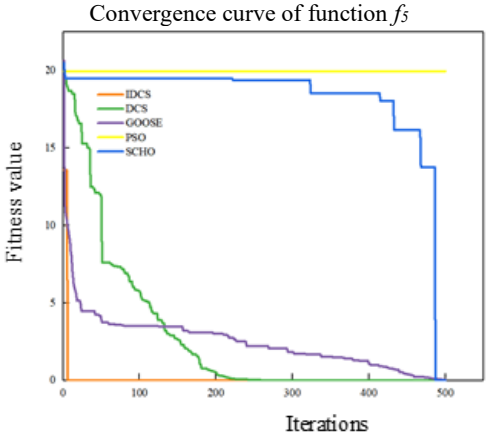
As can be seen from the figure, the convergence speed of the IDCS algorithm is significantly faster than other algorithms when dealing with the three high-dimensional multi-objective optimization problems of $f_5(x)$, $f_6(x)$ and $f_8(x)$. It is particularly noteworthy that even when facing the $f_7(x)$ function, the convergence efficiency of IDCS is still excellent, following closely behind the DCS and GOOSE algorithms.

3) Low-dimensional function test

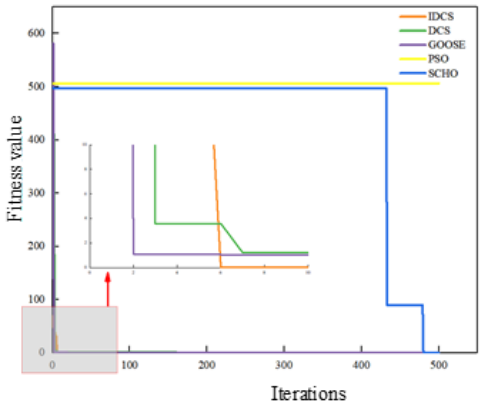
Table 3 gives four fixed-dimensional low-dimensional test functions to measure the global search and local development capabilities of the algorithm in the low-dimensional space, which are used to test and verify the convergence speed of the IDCA algorithm. The curve of the low-dimensional function is shown in fig. 5.



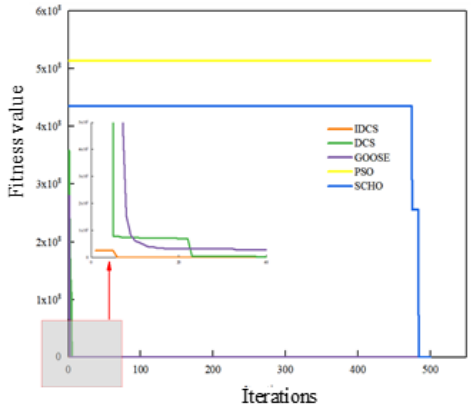
Convergence curve of function f_5



Convergence curve of function f_6



Convergence curve of function f_7

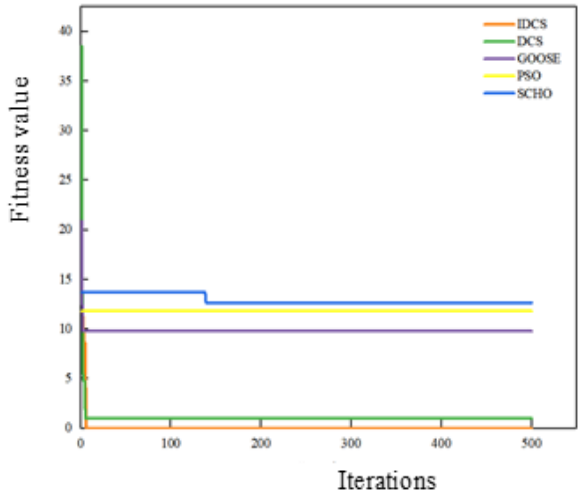


Convergence curve of function f_8

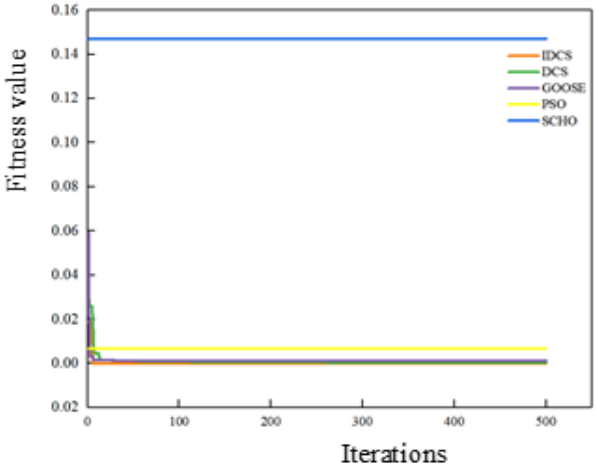
Fig. 4. Convergence curves of high-dimensional single-objective functions

Table 3. Four low-dimensional fixed functions

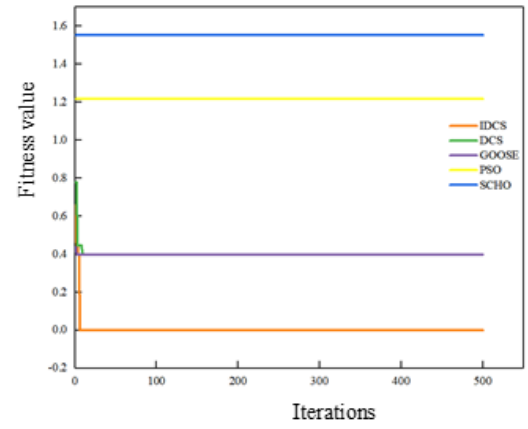
function	Dimensions	scope	Minimum function
$f_9(x) = \left(\frac{1}{500} + \sum_{j=1}^{25} \frac{1}{j + \sum_{i=1}^{2} (x_i -$	2	[-65,65]	1
$f_{10}(x) = \sum_{i=1}^{11} \left[ai - \frac{x_i(b_i^2 + b_i x_2)}{b_i^2 + b_i x_3 + x_4} \right]^2$	4	[-5,5]	0.0003
$f_{11}(x) = \left(x_2 - \frac{5.1}{4\pi^2 x_1^2} + \frac{5}{\pi} x_1 - 6 \right)^2 + 10 \left(1 - \frac{1}{8\pi} \right) \cos x_1 + 10$	2	[-5,5]	0.398
$f_{12}(x) = [1 + (x_1 + x_2 + 1)^2 (19 - 1 \times [30 + (2x_1 - 3x_2)^2 \times (18 - 32x_1$	2	[-5,5]	3



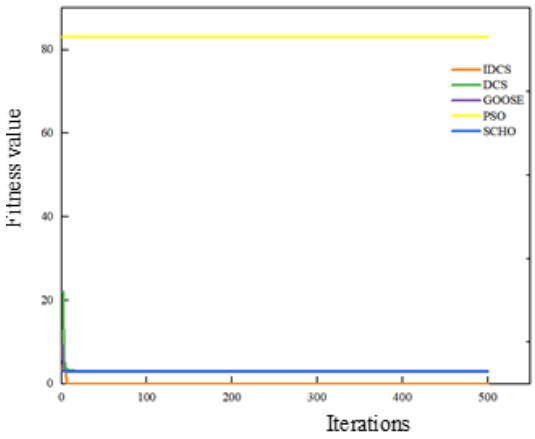
Convergence curve of function f_9



Convergence curve of function f_{10}



Convergence curve of function f_{11}



Convergence curve of function f_{12}

Fig. 5. Convergence curves of high-dimensional single-objective functions

As can be seen from the figure, when solving the four low-dimensional target optimization problems of $f_9(x)$, $f_{10}(x)$, $f_{11}(x)$ and $f_{12}(x)$, it still converges to the optimal solution quickly at a rate ahead of other existing algorithms.

3.2. Case 1: Rolling bearing data analysis at Xi'an Jiaotong University

3.2.1. Data introduction

This study uses the XJTU-SY rolling bearing dataset to verify the effectiveness of the proposed method. The test bench consists of an AC motor, a speed regulator, a support shaft, a bearing, etc. The sampling frequency is 25.6kHz, and the vibration signals are collected using LDK UER204 bearings and PCB 352C33 accelerometers. The dataset has a total of 917 samples, each containing 500, of which 60% are used for training, 10% for validation, and 30% for testing, covering five bearings under three operating conditions, as shown in Table 4. To strictly prevent data leakage, the data splitting was performed chronologically. The first 60% of the continuous vibration signal was exclusively used for model training, the subsequent 10% was used as a validation set for IDCS hyperparameter optimization, while the final completely unseen 30% was strictly isolated for testing, ensuring no temporal overlap between the sets.

Accordingly, all quantitative results reported in the following subsection were obtained on the held-out test set, whereas the validation set was used only for hyperparameter optimization during model development.

Table 4. XJTU-SY bearing data set information list

Working conditions	mark	Data	Fault type
1	0	Bearing 1_1	Outer ring
	1	Bearing 1_4	Cage
2	2	Bearing 2_1	Inner Circle
	3	Bearing 2_2	Outer ring
	4	Bearing 2_3	Cage
3	5	Bearing 3_1	Outer ring
	6	Bearing 3_3	Inner Circle

3.2.2. Experimental data analysis

1) IDCS-SSA

Since SSA has shown excellent decomposition performance and denoising ability in signal processing, SSA is first used to denoise the original signal and decompose and reconstruct it. However, the embedding dimension of traditional SSA needs to be manually selected based on experience, which is highly subjective and may lead to selection bias and unstable decomposition results. Therefore, IDCS is used to adaptively select the embedding dimension. After data decomposition, the reconstructed components are selected based on the w-correlation graph and correlation coefficient. Taking the cage failure data of the Bearing 2_3 data set as an example, the correlation matrix between each subsequence, i.e., the w-correlation graph, is shown in fig. 6.

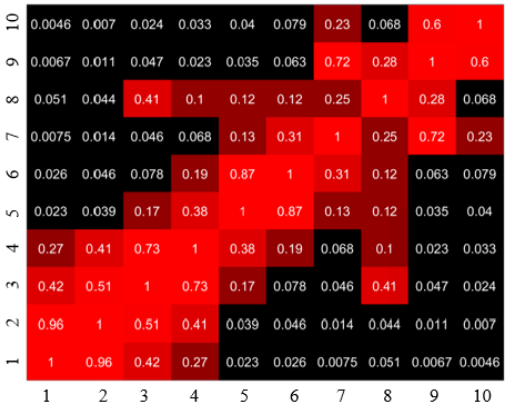


Fig. 6. w-correlation plot

Based on the analysis results of the w-collelation graph and the mean square error formula $S = \sqrt{\frac{\sum_{i=1}^{N}(x_i-\overline{X})^2}{N}}$, the threshold T=0.3650 is selected. Components with correlation coefficients greater than T are merged into the same periodic signal.

Table 5. Correlation coefficient of Bearing 2_3

SSD	SSD1	SSD2	SSD3	SSD4	SSD5	SSD6
Correlation coefficient	0.7424	0.7673	0.5468	0.6047	0.4274	0.5053

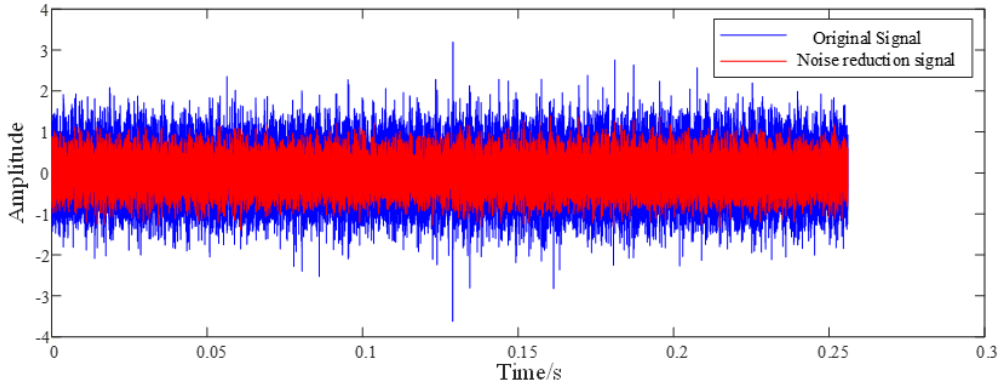


Fig. 7. Comparison of the original signal and the optimized IDCS-SSA denoising time domain waveform

Specifically, the following component combinations with correlations greater than T are observed: component 1 and component 2 (SSD1, correlation coefficient is 0.96); component 3 and component 4 (SSD2, correlation coefficient is 0.73); component 5 and component 6 (SSD3, correlation coefficient is 0.87); component 7 and component 9 (SSD4, correlation coefficient is 0.72); component 8 and component 10 have no subsequences greater than the threshold (then component 8 is SSD5 and component 10 is SSD6). Aggregate and merge the correlated components.

The correlation coefficient results of each aggregated sequence and the original signal are shown in Table 5.

From the results of the correlation coefficient, it can be seen that the values of SSD1, SSD2, and SSD4 are much larger than the others, so the three components of SSD1, SSD2, and SSD4 contain the main components, and these three components are selected for reconstruction.

The time domain waveform of the original signal of the data set and the time domain waveform after IDCS-SSA decomposition and denoising are shown in fig. 7 (taking the Bearing 1_1 data set as an example).

It can be clearly observed from fig. 7 that the IDCS-SSA method has shown significant effects in noise removal and can retain the important features of the original signal, confirming the effectiveness of IDCS-SSA in signal denoising.

2) IDCS algorithm optimizes BiTCN

The performance of BiTCN depends strongly on several key hyperparameters, and manual tuning is often subjective, time-consuming, and difficult to reproduce. To address this issue, the IDCS algorithm is employed to optimize the main structural hyperparameters of BiTCN, including the convolution kernel size, number of convolution channels, number of hidden neurons, and network depth. This optimization strategy reduces human intervention and improves the data-driven adaptability of the diagnosis model.

3) Comparative experiment

In order to evaluate the performance of the proposed method, this paper mainly compares the algorithm diagnosis rate with traditional CNN, MSCNN, TCN, BiTCN and advanced Attention-TCN[17], CSAM-MSCNN[18] models. To ensure fairness, the above methods are trained using the same settings. The evaluation criteria are diagnostic accuracy, F1 score, precision, and recall. Diagnostic accuracy is the ratio of correctly diagnosed samples to the total number of samples, reflecting the overall performance of the evaluation method; F1 score reflects the performance of the evaluation method in each category; precision reflects the "accuracy" of the model in predicting positive class labels; recall reflects the ability of the model to identify all positive class samples, that is, the proportion of all actual positive samples found by the model. The final results are shown in Table 6.

Table 6. Test results of seven models

Model	Accuracy/%	F1/%	Precision/%	Recall/%	Training Time (s)	Testing Time (s)
CNN	58.33	58.59	61.16	58.33	45.26	0.85
MSCNN	55.80	56.61	61.97	55.80	58.42	1.12
TCN	36.23	36.38	37.04	36.23	52.18	0.94
BiTCN	80.80	79.42	79.27	80.80	86.55	1.48
Attention-TCN	47.83	39.69	50.90	47.83	112.30	1.86
CSAM-MSCNN	53.99	45.85	44.34	53.99	125.64	2.15
IDCS-SSA-BITCN	**99.20**	**99.05**	**99.10**	**99.10**	**1845.20**	**2.52**

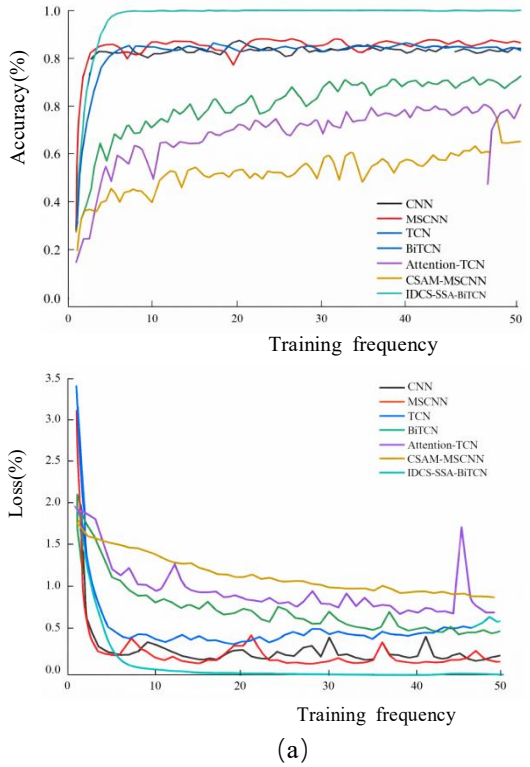
The results in Table 6 show that the proposed IDCS-SSA-BiTCN method consistently outperforms the six baseline models across all evaluation metrics. It achieves 99.20% accuracy, 99.05% F1-score, 99.10% precision, and 99.10% recall, indicating excellent fault classification capability on the XJTU-SY dataset. By contrast, the conventional CNN-, MSCNN-, and TCN-based models exhibit substantially lower diagnostic performance, while the unoptimized BiTCN still remains clearly inferior to the proposed optimized framework. These results confirm the effectiveness of combining adaptive signal enhancement with data-driven structural optimization.

In terms of computational efficiency, the proposed method requires a longer offline training time because of the global IDCS-based hyperparameter search and the SSA reconstruction process. Nevertheless, its inference time on the held-out test set is only 2.52 s, which remains comparable to that of the lightweight baseline models. This indicates that the computational burden is mainly concentrated in the offline optimization stage, whereas the trained model retains practical feasibility for real-time edge-side deployment.

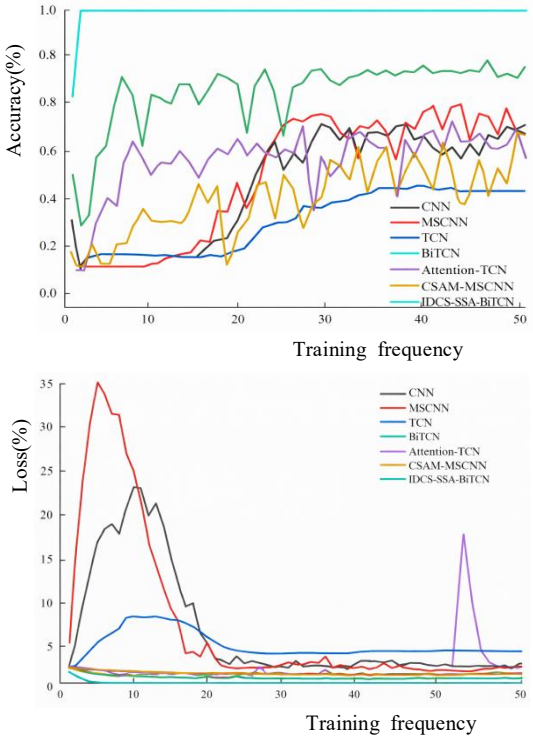
In addition to the diagnostic performance, the computational efficiency (time consumption) of the models was evaluated, as shown in the last two columns of Table 6. Due to the iterative global search mechanism of the IDCS algorithm and the signal reconstruction process of SSA, the proposed IDCS-SSA-BiTCN method requires a longer training time (1845.20s) compared to traditional baseline models. However, its inference time on the held-out test set is merely 2.52 s, which is highly competitive and on the same order of magnitude as other lightweight models. This indicates that the computationally expensive parameter optimization phase is strictly an offline process. Once the model is optimally trained, its online inference speed is entirely capable of meeting the real-time processing requirements of industrial on-site Edge devices.

Fig. 8 shows the accuracy and loss rate curves of each model after 50 iterations.

It can be seen from the figure that the convergence speed of the proposed method on the training set and the test set is significantly faster than that of other models. After iteration, it can achieve a high accuracy rate of nearly 99%, and the loss value approaches zero, indicating that the model is highly optimized and convergent stably. In contrast, the training set accuracy of CNN, MSCNN and TCN only reaches about 80% under the same number of iterations, and their performance on the test set drops significantly, with accuracy rates all not exceeding 60%, indicating a clear insufficiency in generalization ability. BiTCN remained at around 80% on both the training set and the test set, while Attention-TCN and CSAM-MSCNN were even worse, with accuracy rates not exceeding 50%.



(a)



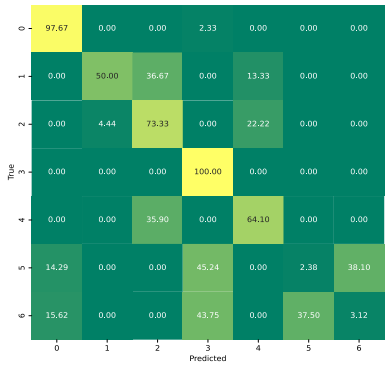
(b)

Fig. 8. Accuracy and loss rate curves of seven diagnostic models: (a) Training set accuracy and loss rate curve (b) Test set accuracy and loss rate curve

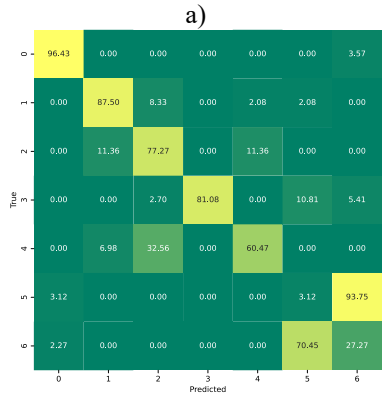
It can be seen from the curve trend that the proposed method not only has a fast optimization speed, but also a smooth curve, demonstrating good robustness and stability. The curves of the remaining models fluctuate greatly, further indicating that these models have deficiencies in optimization and generalization. This once again verifies that the

proposed method has high efficiency, reliability and strong generalization ability in the fault diagnosis of rolling bearings.

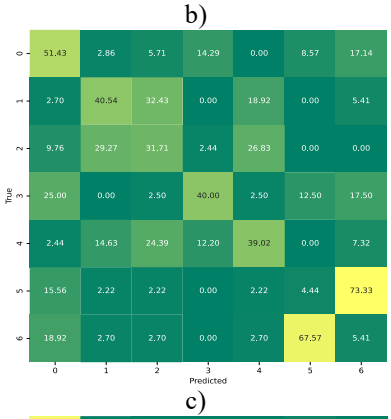
In order to further verify the rolling bearing fault classification ability of the proposed method, the confusion matrix of rolling bearing fault classification of each model diagnosis result is plotted, as shown in fig. 9.



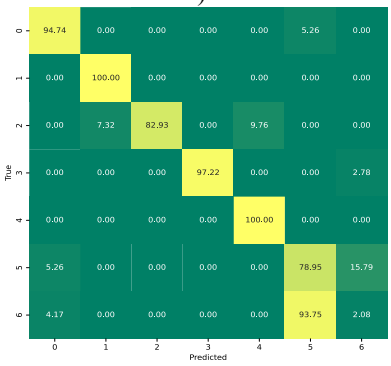
a)



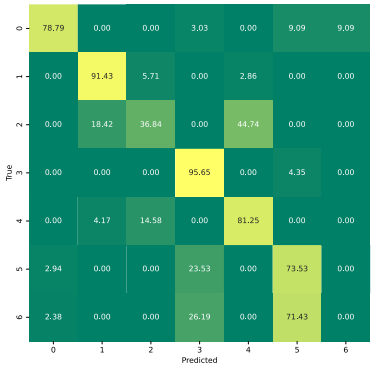
b)



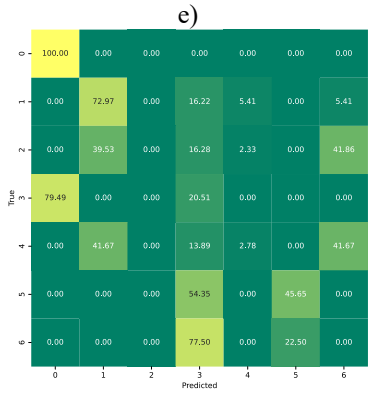
c)



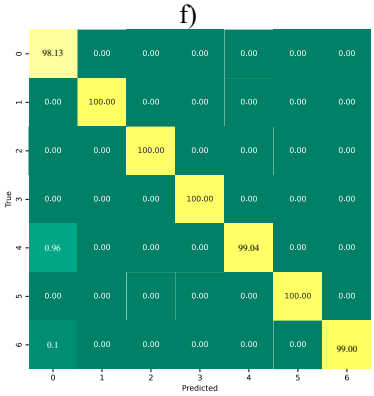
d)



e)



f)



(g)

Fig. 9. Confusion matrix of each model a) CNN, b) MSCNN, c) TCN, d) BiTCN, e) Attention-TCN, f) CSAM-MSCNN, g) IDCS-SSA-BiTCN

In the figure, coordinates 0 to 6 correspond to seven fault states respectively. The values on the diagonal represent the recognition accuracy of each state, and the remaining values represent the recognition error rate. Chart analysis shows that the method proposed in this paper has an overall accuracy rate as high as 99.20% in the seven types of state recognition tasks, with only a small amount of error among categories 1, 4, and 6, highlighting the excellent performance of this method in accurately distinguishing different fault states. Taking BiTCN as an example, we delved into the details of the confusion matrix of the BiTCN model: Although the main classification of category 0 is correct, 5.26% of the instances were still mistakenly classified as category 5. 7.32% of the second category was mistakenly regarded as the first category, and 9.76% of the fourth category was mistakenly regarded as the fourth category. 2.78% of

the three categories were mistakenly regarded as six categories. In the identification of category 5, 5.26% of the instances were mistakenly labeled as category 0, and as many as 15.79% of the instances were confused as category 6. The identification errors of Category 6 were mainly concentrated in being misidentified as Category 5 (93.75%), with a small portion being misidentified as Category 0 (4.17%), indicating that BiTCN encountered significant obstacles in distinguishing between Category 5 and Category 6. In conclusion, the method proposed in this study demonstrates extremely high classification accuracy and stability in the field of rolling bearing fault identification, effectively reducing confusion among categories and further verifying its high efficiency and reliability as a fault diagnosis technology.

3.3. Case 2: Southeast University Rolling Bearing Data Analysis

3.3.1. Dataset Introduction

Two bearings with inner-ring crack faults were manufactured using spark-machining technology, with crack widths of 0.5 mm and 0.8 mm, respectively. Vibration signals were collected using an accelerometer (MHS188U). Owing to its wider frequency-response range, the accelerometer is more sensitive to impact-induced vibration variations than velocity sensors and eddy-current sensors. The dataset was chronologically divided into 60% for training, 10% for validation, and 30% for testing. It includes three health conditions under the same operating condition: normal bearing state (label 0), inner-ring crack with a width of 0.5 mm (label 1), and inner-ring crack with a width of 0.8 mm (label 2). Similar to Case 1, a strict chronological split was adopted to define the training, validation, and testing subsets, ensuring that the test samples remained physically independent and were completely excluded from the IDCS-based hyperparameter selection process.

Accordingly, all quantitative results reported in the following subsection were obtained on the held-out test set, whereas the validation set was used only for hyperparameter optimization during model development.

3.3.2. Experimental data analysis

The IDCS-SSA method is used to deeply decompose and reconstruct the Southeast University bearing dataset, which effectively enhances the purity and feature information of the signal. Subsequently, the reconstructed high-quality data is input into the BiTCN model carefully tuned by the IDCS algorithm for fine classification and diagnosis of rolling bearing faults. The performance comparison with a series of classic models - traditional CNN, MSCNN, TCN, BiTCN and advanced Attention-TCN, CSAM-MSCNN models further verifies its effectiveness. The final comparison results are summarized in Table 7. The accuracy and loss rate curves of the seven diagnostic models are shown in fig. 11.

The results in Table 7 indicate that the proposed method achieves the best overall classification performance on the SEU dataset. All four evaluation

metrics are close to 98.7%, which is higher than those of all comparison models. Among the baseline methods, only BiTCN and Attention-TCN achieve diagnostic accuracies above 90%, whereas the remaining models perform substantially worse. These results further demonstrate the effectiveness and robustness of the proposed framework under a different dataset and fault configuration.

Furthermore, the time consumption data in Table 7 confirms the aforementioned conclusion. Although the offline training phase of IDCS-SSA-BiTCN takes 1560.85 s, the testing time is only 2.05 s. This highly efficient online inference capability further validates its feasibility for real-time edge deployment under varying industrial conditions.

Further observation of fig. 11(a) shows that during the training stage, except for BiTCN and Attention-TCN, the accuracy rates of the training sets of other models are also close to 100%, and the loss rates approach zero, demonstrating a good fitting effect. However, in the test set evaluation, except for the proposed method which still maintained a high accuracy rate of approximately 98.7%, the performance of the other models all declined significantly, exposing the problem of insufficient generalization ability. It is particularly worth noting that through the shape of the loss rate curve, it can be seen that CNN, MSCNN and TCN models show obvious signs of overfitting during the training process. They overlearn the random features and noise in the training data, resulting in a significant decline in performance when facing new data. It once again verified the significant advantages of the proposed method in avoiding overfitting and maintaining strong generalization ability.

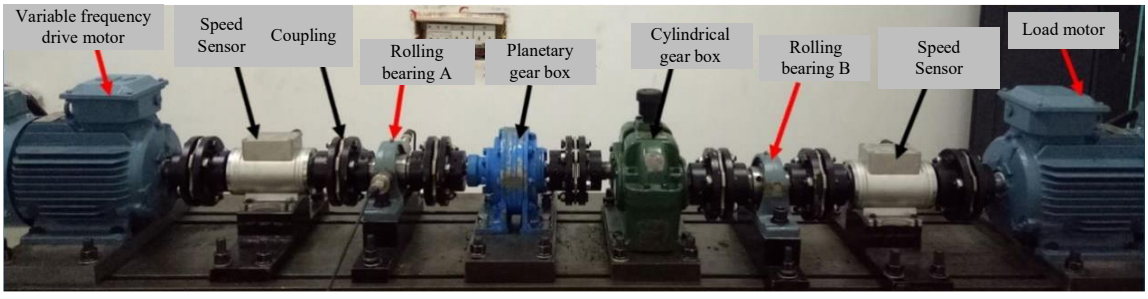


Fig. 10. Bearing Experimental Platform of Southeast University

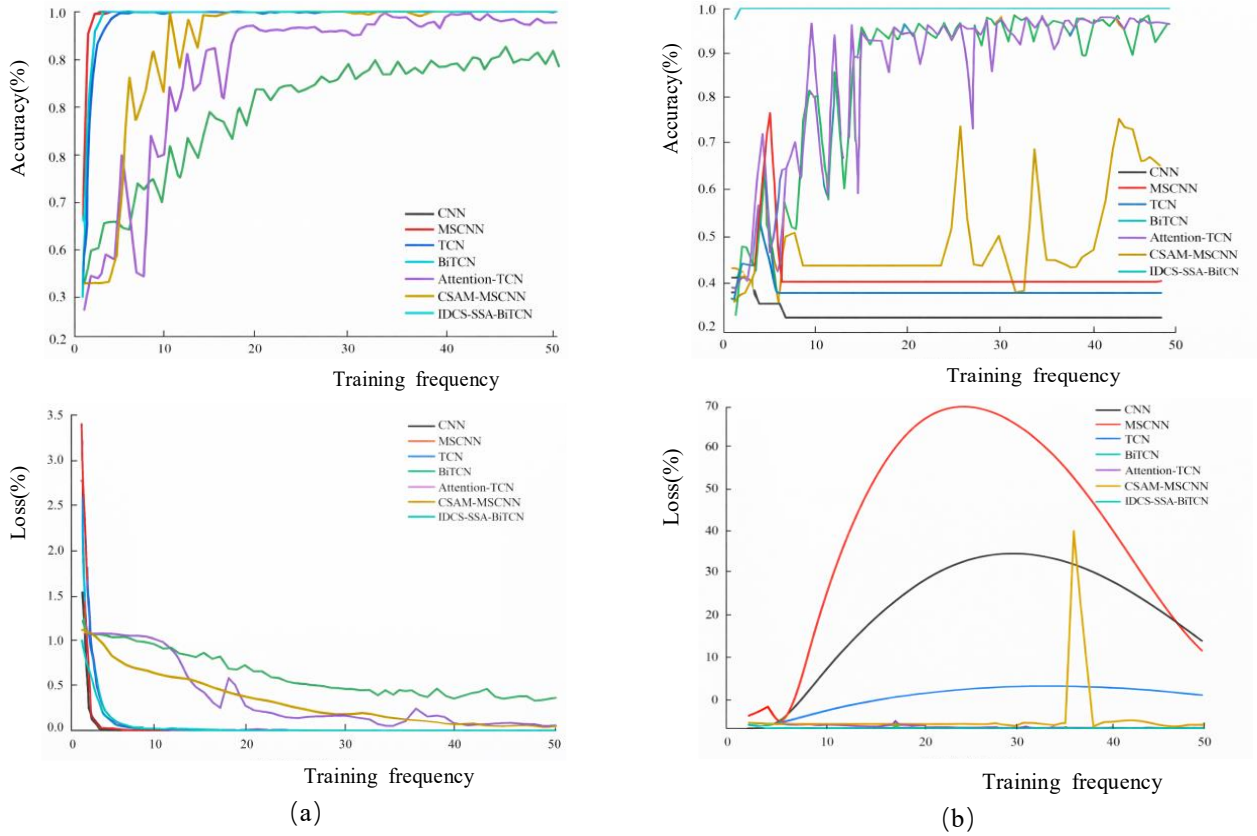


Fig. 11. Accuracy and loss rate curves of seven diagnostic models a)Training set accuracy and loss rate curve (b) Test set accuracy and loss rate curve

Table 7. Test results of seven models

Model	Accuracy/%	F1/%	Precision/%	Recall/%	Training Time (s)	Testing Time (s)
CNN	26.09	10.79	6.81	26.09	38.65	0.62
MSCNN	35.87	20.21	49.19	35.87	49.31	0.88
TCN	32.61	16.04	10.63	32.61	44.75	0.76
BiTCN	93.48	93.59	94.54	93.48	75.24	1.15
Attention-TCN	94.57	94.52	94.90	94.57	98.42	1.54
CSAM-MSCNN	61.96	54.98	72.82	61.96	110.58	1.82
IDCS-SSA-BITCN	**98.75**	**98.60**	**98.70**	**98.70**	**1560.85**	**2.05**

In order to further explore and visually verify the excellent performance of this method in rolling bearing fault classification, the T-SNE algorithm is used to map and display the spatial distribution of features extracted by the seven models when processing the data set. As shown in fig. 12, different colors represent different health states of the equipment, which intuitively shows the expressiveness of each model in feature differentiation and category clustering.

The feature distribution comparison in fig. 12 clearly demonstrates the excellent performance of the proposed method: its feature point distribution shows obvious clarity, the boundaries of each category are clear, and there is no overlap. After the high-dimensional features are reduced to a two-dimensional view, the fault state has a very high distinguishability. This proves the ability of the proposed method to effectively extract and maintain classification differences in complex feature spaces. In contrast, although the feature distribution of the BiTCN and Attention-TCN models shows a certain clarity, a small number of classification errors can still be observed, indicating that they are slightly insufficient in distinguishing certain health states. The feature distribution of the three traditional models, CNN, MSCNN and TCN, is particularly fuzzy, and the boundaries between different health states are unclear, which fully reflects the limitations and challenges of these models in distinguishing fault states, especially in feature extraction and maintaining the distinction between categories. In summary, the proposed method shows significant advantages in the accuracy of feature extraction and fault classification. It not only effectively distinguishes various health states, but also greatly improves the accuracy and reliability of rolling bearing fault diagnosis.

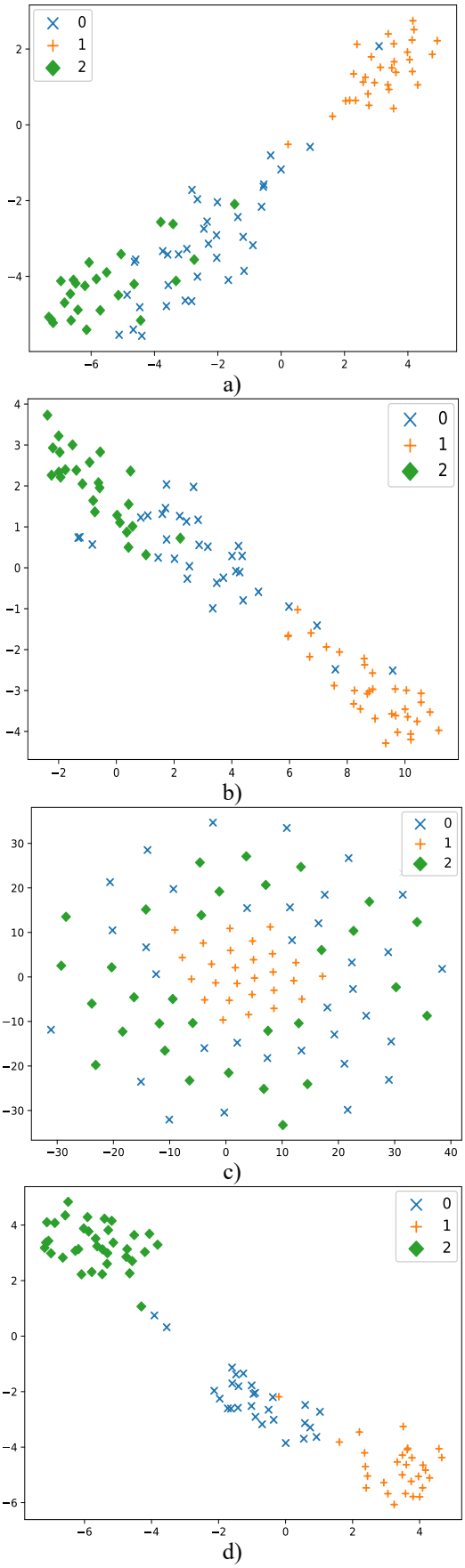


Fig. 12. Feature visualization using TSNE (a) CNN (b) MSCNN (c) TCN (d) BiTCN (e) Attention-TCN (f) CSAM-MSCNN (g) IDCS-SSA-BiTCN

4. CONCLUSIONS

This study addresses the challenges in rolling bearing fault diagnosis, including the reliance on manual signal preprocessing, the difficulty of tuning deep learning models, and the lack of adaptability in integrated diagnostic systems. A novel intelligent diagnosis framework is proposed, which combines adaptive signal enhancement and structural optimization. Specifically, an improved Singular Spectrum Analysis (SSA) method is introduced, where the embedding dimension is optimized by a modified Differential Creative Search (IDCS) algorithm. Additionally, a w-correlation-based component selection strategy is employed to retain cyclic signal components with high consistency. In the modeling stage, a Bidirectional Temporal

Convolutional Network (BiTCN) is adopted to extract multi-scale temporal features, and IDCS is applied again to perform global optimization of key structural parameters. Extensive experiments on multiple public bearing datasets validate the effectiveness and robustness of the proposed method.

The main conclusions are summarized as follows:

(1) In the signal processing stage, IDCS enables optimal embedding dimension selection in SSA, improving both feature retention and noise suppression. The w-correlation-based component reconstruction strategy effectively identifies periodic components with strong inter-correlation, enhancing the discriminability of the reconstructed signals. Visualization using t-SNE shows that the optimized signals produce more clearly separated clusters, improving the quality of the model input.

(2) In the structural optimization stage, IDCS establishes a global search mechanism in the BiTCN hyperparameter space, enabling collaborative optimization of kernel size, channel number, and network depth. As a result, the optimized model achieves 99.20% accuracy on the XJTU-SY dataset and 98.75% accuracy on the SEU dataset, demonstrating strong diagnostic capability and satisfactory generalization performance across different experimental conditions.

(3) Regarding the optimizer itself, the improved IDCS incorporates nonlinear parameter control and Lévy flight strategies, which enhance both convergence efficiency and global search capability. The benchmark-function experiments show that IDCS converges faster and achieves more stable optimization performance than the compared algorithms, confirming its effectiveness for complex high-dimensional optimization tasks.

In summary, the proposed approach achieves methodological innovation and system-level integration in both feature extraction and model construction.

Future work will focus on extending this framework to real-world industrial machines, specifically targeting fault diagnosis for wind turbine drivetrains under varying operating conditions. To validate the model's robustness against highly dynamic loads and complex environmental noise, and to overcome the computational burden on resource-constrained edge devices, a cloud-edge collaboration paradigm will be explored. Specifically, the heavy IDCS parameter search and model training will be executed offline on high-performance servers, while the optimized, lightweight BiTCN model will be deployed to the edge for real-time inference (which requires less than 3 seconds). Furthermore, to enable real-time on-site adaptation under these varying loads, multi-sensor fusion, lightweight transfer learning, and domain adaptation strategies will be integrated. This allows the edge device to fine-tune only the classifier layer without triggering a full computationally expensive optimization, thereby significantly advancing the practical application and edge deployment of intelligent diagnosis technologies in complex industrial scenarios.

Source of funding: *This research was funded by the Jiangsu Province Carbon Peak and Carbon Neutrality Science and Technology Innovation Special Fund (BT2024004, BE2023854), and the Special Fund for Basic Scientific Research Business Expenses of Central Universities (2242025K30015). The article processing charge (APC) was funded by the same sources.*

Authors contributions: *Research concept and design, X.J., B.W.; Collection and/or assembly of data, X.J., B.W., G.W., J.H., G.Y; Data analysis and interpretation, X.J., B.W., G.W., J.H., G.Y., H.Y., X.H.; Writing the article, G.W., J.H.; G.Y., H.Y., X.H. Critical revision of the article, H.Y., X.H.; Final approval of the article, J.Z.*

Declaration of competing interest: *The author declares no conflict of interest.*

REFERENCES

1. Chen Z, Yang Y, He C, et al. Feature extraction based on hierarchical improved envelope spectrum entropy for rolling bearing fault diagnosis. IEEE Transactions on Instrumentation and Measurement. 2023;72:1-12. https://doi.org/10.1109/TIM.2023.3277938.
2. Zhenxing L. Research on fault diagnosis of rolling bearing based on Variational Mode Decomposition and Hilbert envelope spectrum analysis. 36th Chinese Control and Decision Conference (CCDC). IEEE. 2024:4816-4820. https://doi.org/10.1109/CCDC62350.2024.10587534.
3. Touti W, Salah M, Sheng S, et al. An envelope time synchronous averaging for wind turbine gearbox fault diagnosis. Journal of Vibration Engineering & Technologies. 2024;12(4):6513-6525. https://doi.org/10.1007/s42417-023-01267-y.
4. Zhang Q, Huo R, Zheng H, et al. A fault diagnosis method with bitask-based time-and frequency-domain feature learning. IEEE Transactions on Instrumentation and Measurement. 2023;72:1-11. https://doi.org/10.1109/TIM.2023.3305652.
5. Lin L, Zhou K, Li D, et al. Experimental study on train axle fatigue crack acoustic emission signals recognition based on a one-dimensional convolutional neural network. Nondestructive Testing and Evaluation. 2024:1-32. https://doi.org/10.1080/10589759.2024.2389947.
6. Jiang L, Shi C, Sheng H, et al. Lightweight CNN architecture design for rolling bearing fault diagnosis. Measurement Science and Technology. 2024;35(12): 126142. https://doi.org/10.1088/1361-6501/ad7a1a.
7. Saeed S, Reddy B. Diagnostics of vibrations due to looseness fault and unbalance in rotating machinery with neural network. Diagnostyka. 2025;26(4): 2025403. https://doi.org/10.29354/diag/211587.
8. Ibrahim MA, Hadid IN. Intelligent fault detection and classification in overcurrent protection systems based

on artificial neural networks. Diagnostyka. 2025; 26(4):2025415. https://doi.org/10.29354/diag/215701.
9. Yu Z, Zhang L, Kim J. The performance analysis of PSO-ResNet for the fault diagnosis of vibration signals based on the pipeline robot. Sensors. 2023; 23(9):4289. https://doi.org/10.3390/s23094289.
10. Li X, Ma Z, Yuan Z, et al. A review on convolutional neural network in rolling bearing fault diagnosis. Measurement Science and Technology. 2024. https://doi.org/10.1088/1361-6501/ad356e.
11. Yu X, Wang S, Xu H, et al. Intelligent fault diagnosis of rotating machinery under variable working conditions based on deep transfer learning with fusion of local and global time–frequency features. Structural Health Monitoring. 2024;23(4):2238-2254. https://doi.org/10.1177/14759217231199427.
12. Yu X, Wang Y, Liang Z, et al. An adaptive domain adaptation method for rolling bearings' fault diagnosis fusing deep convolution and self-attention networks. IEEE Transactions on Instrumentation and Measurement. 2023;72:1-14. https://doi.org/10.1109/TIM.2023.3246494.
13. Hamdaoui H, Ngiejungbwen L A, Gu J, et al. Improved signal processing for bearing fault diagnosis in noisy environments using signal denoising, time–frequency transform, and deep learning. Journal of the Brazilian Society of Mechanical Sciences and Engineering. 2023;45(11): 576. https://doi.org/10.1007/s40430-023-04471-9.
14. Zhang Y, Cheng Z, Wu Z, et al. Research on electronic circuit fault diagnosis method based on SWT and DCNN-ELM. IEEE Access. 2023;11: 71301-71313. https://doi.org/10.1109/ACCESS.2023.3292247.
15. Jiang X, Zheng J, Chen Z, et al. Leveraging transfer learning for data augmentation in fault diagnosis of imbalanced time-frequency images. IEEE Transactions on Automation Science and Engineering. 2024. https://doi.org/10.1109/TASE.2024.3454418.
16. Ren C, Li X, Wang W, et al. A wavelet packet transform-based domain adaptation method for rolling bearings fault diagnosis with fusion of local and global time–frequency features. Measurement Science and Technology. 2025;36(2):025022. https://doi.org/10.1088/1361-6501/ada6eb.
17. Truong G T, Choi K K, Nguyen T H, et al. Prediction of shear strength of RC deep beams using XGBoost regression with Bayesian optimization. European Journal of Environmental and Civil Engineering. 2023;27(14):4046-4066. https://doi.org/10.1080/19648189.2023.2169357.
18. Zhang H, Ge B, Han B. Real-time motor fault diagnosis based on tcn and attention. Machines. 2022; 10(4):249. https://doi.org/10.3390/machines10040249.
19. Xu Q, Jiang H, Zhang X, et al. Multiscale convolutional neural network based on channel space attention for gearbox compound fault diagnosis. Sensors. 2023;23(8):3827. https://doi.org/10.3390/s23083827.

Xinjun JIANG is currently with Guoneng Changzhou Second Power Generation Co., Ltd., Changzhou, Jiangsu, China. His research interests include rotating machinery fault diagnosis, signal processing, intelligent optimization algorithms, and condition monitoring of power generation equipment.
e-mail: 15139979182@163.com

Baohua WANG is currently with Guoneng Changzhou Second Power Generation Co., Ltd., Changzhou, Jiangsu, China. His research interests include power generation equipment maintenance, fault diagnosis of rotating machinery, vibration signal analysis, and intelligent operation and maintenance.
e-mail: jingzhu244@gmail.com

Guoxing WU is currently with Guoneng Changzhou Second Power Generation Co., Ltd., Changzhou, Jiangsu, China. His research interests include equipment condition monitoring, bearing fault diagnosis, vibration analysis, and intelligent maintenance technology for thermal power units.
e-mail: 18205179811@163.com

Jun HU is currently with Guoneng Changzhou Second Power Generation Co., Ltd., Changzhou, Jiangsu, China. His research interests include mechanical equipment fault diagnosis, vibration signal processing, reliability analysis, and operation safety of power generation systems.
e-mail: 18205179811@163.com

Ge YIN is currently with Guoneng Changzhou Second Power Generation Co., Ltd., Changzhou, Jiangsu, China. His research interests include intelligent diagnosis of rotating machinery, equipment reliability assessment, signal denoising, and power plant operation and maintenance.
e-mail: 9906238@haust.edu.cn

Han YAN is currently with Guoneng Changzhou Second Power Generation Co., Ltd., Changzhou, Jiangsu, China. His research interests include fault detection, vibration signal analysis, intelligent optimization, and condition monitoring of industrial equipment.
e-mail: 207960008@qq.com

Xinrong HE is currently with Guoneng Changzhou Second Power Generation Co., Ltd., Changzhou, Jiangsu, China. His research interests include rolling bearing fault diagnosis, intelligent maintenance, signal processing, and reliability improvement of power generation equipment.
e-mail: 2067638513@qq.com

Jing ZHU is currently with Henan University of Science and Technology, Luoyang, Henan, China. Her research interests include intelligent fault diagnosis, deep learning, adaptive signal processing, and condition monitoring of rotating machinery. She is the corresponding author of this paper.
e-mail: 1907175086@qq.com