machines

MDPI

Article

Research on Bearing Fault Diagnosis Method Based on MESO-TCN

Ruibin Gao [1], Jing Zhu [2,*], Yifan Wu [2], Kaiwen Xiao [3] and Yang Shen [3]

[1] Guoneng Jiangsu Electric Power Engineering Technology Co., Ltd., Nanjing 210019, China; 15139979182@163.com
[2] School of Vehicle and Transportation Engineering, Henan University of Science and Technology, Luoyang 471003, China; 9906238@haust.edu.cn
[3] School of Energy and Environment, Southeast University, Nanjing 210096, China
* Correspondence: jingzhu244@gmail.com; Tel.: +86-15139979182

Abstract

To address the issues of information redundancy, limited feature representation, and empirically set parameters in rolling bearing fault diagnosis, this paper proposes a Multi-Entropy Screening and Optimization Temporal Convolutional Network (MESO-TCN). The method integrates feature filtering, network modeling, and parameter optimization into a unified diagnostic framework. Specifically, ensemble empirical mode decomposition (EEMD) is combined with a hybrid entropy criterion to preprocess the raw vibration signals and suppress redundant noise. A kernel-extended temporal convolutional network (ETCN) is designed with multi-scale dilated convolution to extract diverse temporal fault patterns. Furthermore, an improved whale optimization algorithm incorporating a firefly-inspired mechanism is introduced to adaptively optimize key hyperparameters. Experimental results on datasets from Xi'an Jiaotong University and Southeast University demonstrate that MESO-TCN achieves average accuracies of 99.78% and 95.82%, respectively, outperforming mainstream baseline methods. These findings indicate the method's strong generalization ability, feature discriminability, and engineering applicability in intelligent fault diagnosis of rotating machinery.

Keywords: bearing fault diagnosis; ensemble empirical mode decomposition; mixed entropy; time convolution neural network; whale optimization algorithm

check for updates

Academic Editor: Davide Astolfi

Received: 30 May 2025
Revised: 12 June 2025
Accepted: 24 June 2025
Published: 27 June 2025

Citation: Gao, R.; Zhu, J.; Wu, Y.; Xiao, K.; Shen, Y. Research on Bearing Fault Diagnosis Method Based on MESO-TCN. *Machines* **2025**, *13*, 558. https://doi.org/10.3390/machines13070558

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Rolling bearings, as one of the most widely used core components in rotating machinery, are extensively deployed in various key equipment such as robot joints, spindles of CNC machine tools, axle boxes of high-speed trains, power motors, and wind turbine generators. Their health status directly affects the operational safety, production efficiency, and service life of the equipment. During long-term service, rolling bearings are highly prone to various types of faults such as spalling, wear and cracking due to factors like load impact, lubrication failure, material fatigue or manufacturing defects. If the equipment cannot be identified and dealt with in a timely and accurate manner, it will not only lead to the degradation of equipment performance and the increase of maintenance costs but also may cause serious system failures or even safety accidents. Therefore, the development of high-precision and automated fault diagnosis technology for rolling bearings has become an important technical support in the construction of health management systems for intelligent manufacturing, transportation, and energy equipment.

In practical operation, the vibration signals of rolling bearings are typically characterized by strong nonlinearity, non-stationarity, and high levels of noise interference. These signals often contain multiple frequency components, with relevant information sparsely distributed and feature representations highly entangled. Directly using raw, unscreened signals for model training tends to introduce redundant information, which interferes with the model's ability to learn meaningful patterns and ultimately degrades classification accuracy. Moreover, different types of bearing faults often exhibit similar time-domain characteristics, making it difficult for traditional feature extraction methods to distinguish them effectively and limiting the diagnostic capability of conventional models.

Traditional bearing fault diagnosis methods primarily rely on extracting informative features from vibration signals to identify fault types and severities. Classical approaches include variational mode decomposition (VMD) [1], wavelet transform [2], and Fourier transform [3], which enable feature extraction from the time, frequency, and time-frequency domains. These methods provide multiple perspectives for evaluating bearing health. However, they are highly dependent on expert knowledge and struggle to address the complexity and nonlinearity inherent in real-world signals, limiting their applicability in large-scale or automated diagnostic scenarios.

To address the limitations of traditional methods in handling complex signals, researchers have increasingly adopted deep learning techniques to replace the manual feature engineering process [4–6]. Among them, the Temporal Convolutional Network (TCN) [7] has demonstrated strong capabilities in modeling long-range dependencies by leveraging causal and dilated convolution structures. TCNs offer enhanced stability and accuracy in processing bearing vibration signals with temporal correlations [8–11].

In recent years, TCN-based models have been widely applied to bearing fault diagnosis and remaining life prediction tasks, often in combination with advanced mechanisms to further improve performance. Meanwhile, to better leverage distributed data collected across wind farms while preserving data privacy, some studies have started exploring the potential of federated learning (FL) in collaborative fault diagnosis settings [12,13]. For instance, Reference [14] proposed an end-to-end transfer prediction framework that combines TCN with a residual self-attention mechanism to address feature distribution shifts and generalization issues under varying working conditions. Reference [15] developed a TCN-based diagnostic model by incorporating soft thresholding and self-attention to enhance feature sparsity and discriminability, achieving intelligent fault identification in rotating machinery. In Reference [16], the input sequence was divided into multiple sub-blocks, each processed using a TCN to extract local features, which were then fused with global information through a multi-layer self-attention mechanism, achieving over 97% accuracy and recall on a chiller dataset. Reference [17] introduced dilated convolution and a multi-scale residual structure into a time-domain convolutional pooling network, enhancing local feature extraction with small convolution kernels and improving model accuracy by approximately 5%.

Nevertheless, the existing diagnostic methods based on TCN still have urgent problems to be solved in practical applications: (1) Most models are directly modeled based on the original signals, lacking an effective feature screening mechanism, and are vulnerable to redundancy and noise interference affecting diagnostic performance; (2) The model training process highly relies on human experience to set hyperparameters such as the size of convolution kernels, network depth, and learning rate, lacking a systematic adaptive optimization mechanism, which affects the training stability and the generalization performance of the model. (3) There is still room for improvement in the adaptability and robustness to multi-condition and multi-source data, which restricts the wide application

of the method in actual complex engineering scenarios. It can be seen from this that the fault diagnosis of rolling bearings still has important research value and innovation space.

In response to the above problems, this paper proposes a temporal convolutional neural network (MESO-TCN) model integrating multi-entropy feature screening and parameter optimization, aiming to systematically improve the feature extraction ability, model robustness and training adaptability of complex temporal signals in the fault diagnosis of rolling bearings. The main contributions of this article include:

(1) A unified diagnostic architecture (MESO-TCN) is proposed, which integrates multi-entropy-based signal preprocessing, kernel-extended temporal convolution, and adaptive parameter optimization. This end-to-end framework enhances the interpretability and adaptability of fault diagnosis under complex conditions.
(2) A hybrid entropy screening mechanism is developed by combining information entropy, relative entropy, and cross-entropy to evaluate intrinsic mode components derived from EEMD. This enables effective selection of high-value signal components while suppressing redundant or noisy features.
(3) A kernel-extended TCN structure with multi-scale dilated convolutions is constructed to capture diverse temporal fault patterns. In addition, an improved whale optimization algorithm incorporating a firefly-inspired attraction mechanism is introduced to jointly optimize key hyperparameters, thereby improving model convergence and robustness.

Extensive experiments on datasets from Xi'an Jiaotong University and Southeast University confirm that the proposed method outperforms conventional baselines in terms of accuracy, generalization, and stability.

2. Materials and Methods

The MESO-TCN model is a multi-scale optimized convolutional network designed for fault diagnosis of rolling bearings, aiming to enhance feature extraction and recognition capabilities under non-stationary and high-noise vibration conditions. The model consists of three components: the raw vibration signal is first processed using Ensemble Empirical Mode Decomposition (EEMD) combined with a hybrid entropy-based screening mechanism to suppress redundant noise and reconstruct informative feature sequences; then, an extended-kernel temporal convolutional network (ETCN) incorporating a Multi-Scale Dilated Convolution Block (MSDCB) within the residual module is employed to extract temporal features at multiple scales using causal convolutions with varying dilation rates, with learnable weight fusion to highlight critical patterns; finally, an improved whale optimization algorithm (IWOA) enhanced by a firefly-inspired brightness mechanism is adopted to globally optimize convolutional parameters such as learning rate and channel number. The complete structure of the MESO-TCN model is illustrated in Figure 1.

2.1. Signal Preprocessing and Feature Selection

In the diagnosis of rolling bearing faults, vibration signals usually exhibit strong non-stationary, multi-scale superposition, and high noise interference characteristics. Directly inputting the original signal into a deep model can easily introduce redundant components, mask key features, and affect recognition accuracy and convergence efficiency. Therefore, this article proposes a signal processing method that combines ensemble empirical mode decomposition [18] (EEMD) with a multi-entropy fusion evaluation mechanism to selectively screen and reconstruct input features.

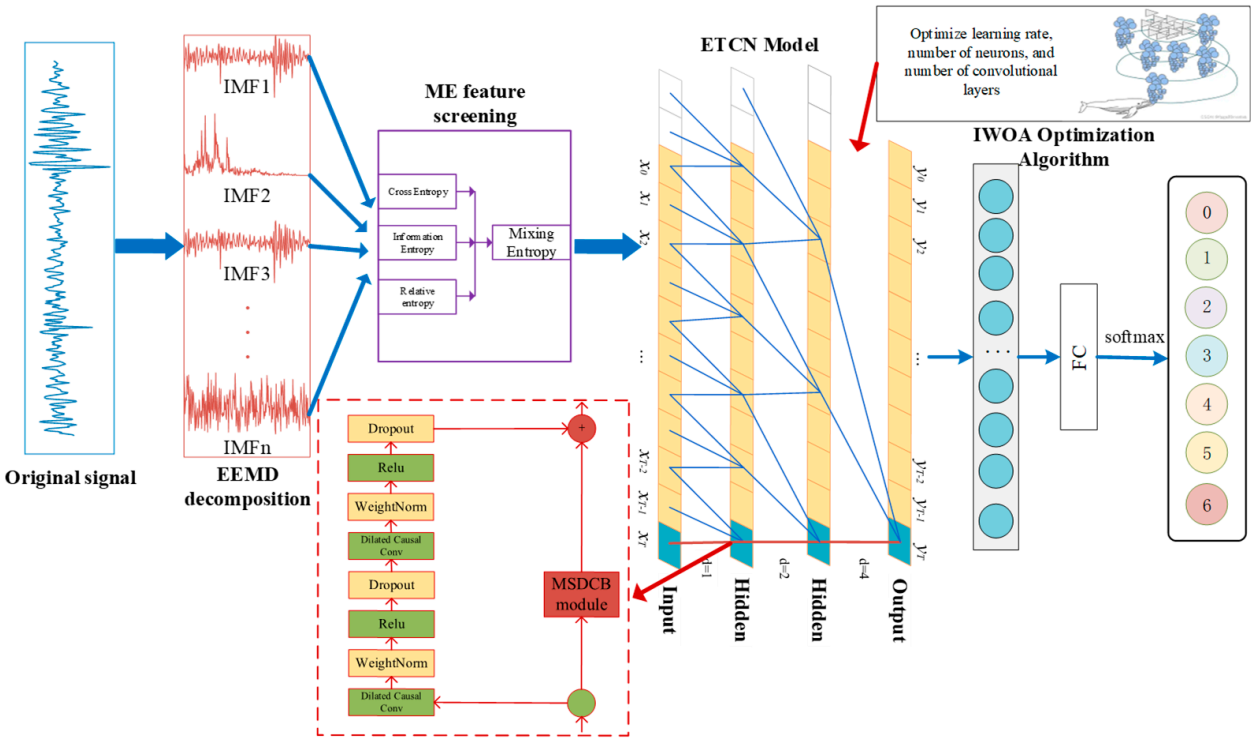


Figure 1. MESO-TCN model diagram.

Firstly, EEMD is adopted to decompose the original signal into several intrinsic mode functions (IMF), and a feature representation space with multiple frequency scales is constructed. EEMD achieves multiple EMDs by introducing Gaussian white noise and takes the average result, effectively alleviating the modal aliasing problem. Its decomposition process is as follows:

Superposition of the original signal and noise:

$$x_m(t) = x(t) + n_m(t) \tag{1}$$

The mth EMD result:

$$x_m(t) = \sum_{j=1}^{J} IMF_{m,j}(t) + r_m(t) \tag{2}$$

The average result of the sth IMF is:

$$C_s(t) = \frac{1}{M} \sum_{m=1}^{M} IMF_{m,j}(t) \tag{3}$$

After obtaining the complete IMF set, further screening of effective components is carried out through mixed entropy evaluation. To comprehensively measure signal complexity, information retention, and diagnostic potential, this paper introduces the definition of mixed entropy:

Assuming that $p(x)$ and $q(x)$ are two probability distributions of a discrete random variable X, the relative entropy of p to q is:

$$D_{kl}(p\|q) = \sum_i p(x_i) \log\left(\frac{p(x_i)}{q(x_i)}\right) \tag{4}$$

Further splitting the formula for calculating relative entropy:

$$\begin{aligned} D_{KL}(p\|q) &= \sum_i p(x_i) \log\left(\frac{p(x_i)}{q(x_i)}\right) \\ &= \sum_i p(x_i) \log p(x_i) - \sum_i p(x_i) \log q(x_i) \end{aligned} \tag{5}$$

Among them, cross entropy is defined as:

$$H(p,q) = -\sum_i p(x_i) \log q(x_i) \tag{6}$$

Shannon [19] defined the average information content of a source as information entropy.

$$H(X) = E[I(x_i)] = E - [-\log_a P(x_i)] = -\sum_{i=1}^{n} P_i \ln P_i \tag{7}$$

Among them, $P(x_i)$ is the probability of event x_i occurring. The calculation formula for mixed entropy is shown in Equation (8).

$$M = -D_{KL}(p\|q) - H(p\|q) + H(X) \tag{8}$$

The ME value is used to score and sort each *IMF* component, retaining only the components with high information content and strong consistency with the original signal structure, and performing signal reconstruction based on this:

$$X_{rec}(t) = \sum_{k \in K} IMF_k(t) \tag{9}$$

Among them, $K = \{k \mid ME_k > \theta\}$ represents the set of *IMF* component numbers for all mixed entropy ME_k greater than the threshold θ.

2.2. ETCN Network

To enhance the model's ability to extract multi-scale fault features in the vibration signals of rolling bearings, based on the Temporal Convolutional Network (TCN), in this paper, an Extended kernel Temporal Convolutional Network (ETCN) was designed. This model introduces a multi-scale convolutional path parallel mechanism through structural innovation, effectively expanding the receptive field range and enhancing the model's response ability to key features at different time scales.

The traditional TCN is mainly composed of causal convolution, dilated convolution and residual join, which can capture the long-term dependencies in the time series. However, in the standard TCN, the expansion rates of each layer are set-fixed, and the receptive field expansion relies on deep stacking, which easily leads to insufficient modeling ability of the model for local details and introduces a relatively high computational burden at the same time. In the fault diagnosis scenario of rolling bearings, the fault signals have obvious non-stationarity and multi-scale characteristics. It is difficult to take into account the pattern expression within different time windows by relying only on a single-scale convolution kernel.

To this end, in this paper, a Multi-Scale Dilated Convolution Block (MSDCB) is introduced into the residual module structure of TCN, as shown in Figure 2.

Specifically, each MSDCB module contains multiple causal convolution paths with different dilation rates (such as dilation rates of 1, 2, and 4), and models the input features

from different time scales, respectively. Let the input sequence be x and the convolution kernel corresponding to the i-th path be wi. The feature extraction process is as follows:

$$f_i(x) = Dropout(\sigma(Norm(Conv1Dd_i(x, wi)))) \tag{10}$$

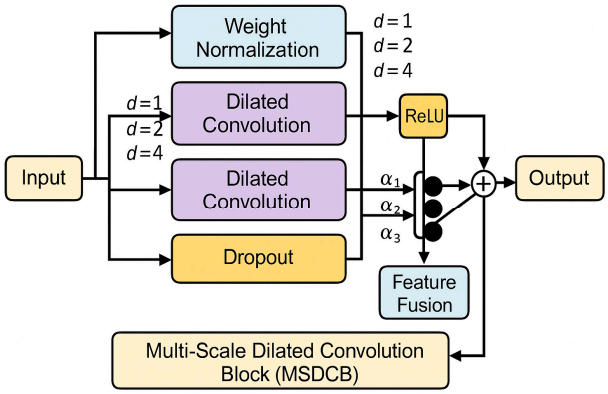


Figure 2. Structure of MSDCB.

Among them, $Conv1Dd_i$ represents the one-dimensional causal convolution operation with an expansion rate of d_i, Norm represents the weight normalization, ReLU is the activation function, and Dropout is the regularization processing.

After completing the feature extraction of multi-scale convolutional paths, ETCN fuses the outputs of all paths. To take into account the importance of different scales, learnable weights are adopted for the weighted sum of the outputs of each path. The final output is expressed as:

$$y = \sum_{i=1}^{N} \alpha_i f_i(x),\ where \sum_{i=1}^{N} \alpha_i = 1,\ \alpha_i \geq 0 \tag{11}$$

2.3. IWOA Algorithm

The Whale Optimization Algorithm [19] (Whale Optimization Algorithm, WOA) is a swarm intelligence optimization algorithm that simulates the behavior of humpback whales hunting their prey. Its core mechanism includes three operation strategies: "surrounding the prey", "spiral position update", and "global search" [20]. WOA has the advantages of simple structure and easy implementation, and shows good performance in most continuous optimization problems. However, the traditional WOA has problems such as getting trapped in local optimum and having a relatively slow convergence speed. To enhance its global search ability and optimization efficiency, this paper introduces the attraction degree and brightness mechanism in the Firefly Algorithm (FA) to improve WOA.

Basic WOA model mechanism

① Encircling prey behavior: Whales approach their target prey by updating their current position. The mathematical model for position update is as follows:

$$X(t+1) = X_p(t) - A \cdot \left| C \cdot X_p(t) - X(t) \right| \tag{12}$$

In the formula, $X(t)$ is the current position vector of the whale; $X_p(t)$ is the current optimal individual position; A and C are coefficient vectors, generated respectively by random numbers and scaling factors:

$$A = 2a \cdot r_1 - a,\ C = 2 \cdot r_2 \tag{13}$$

In the formula, r_1 and r_2 are random numbers in the interval [0, 1]; a is the regulating parameter, which decreases linearly from 2 to 0 as the iteration proceeds.

② Update of the spiral position

When a whale approaches its prey, it simultaneously updates its position along a spiral trajectory. The update formula is:

$$X\left(t+1\right) = D \cdot e^{bl} \cdot \cos\left(2\pi l\right) + X_p\left(t\right) \tag{14}$$

$$D = \left|X_p(t) - X(t)\right| \tag{15}$$

In the formula, D is the distance between the whale and its prey; b is a constant used to define the spiral shape; l is a random number in the interval [0, 1]. p is a random number between [0, 1].

③ Global search

To avoid falling into local optimum, WOA introduces a global search mechanism, using the positions of randomly selected individuals in the current population as guidance. The update method is:

$$X(t+1) = X_{rand}(t) - A \cdot |C \cdot X_{rand}(t) - X(t)| \tag{16}$$

In the formula: $X_{rand}(t)$ represents the randomly selected individual positions in the current population.

(2) Improved strategies for introducing the FA mechanism

To enhance the population diversity and search ability of *WOA*, this paper improves it by combining the attraction mechanism of the firefly algorithm. As shown in Figure 3:

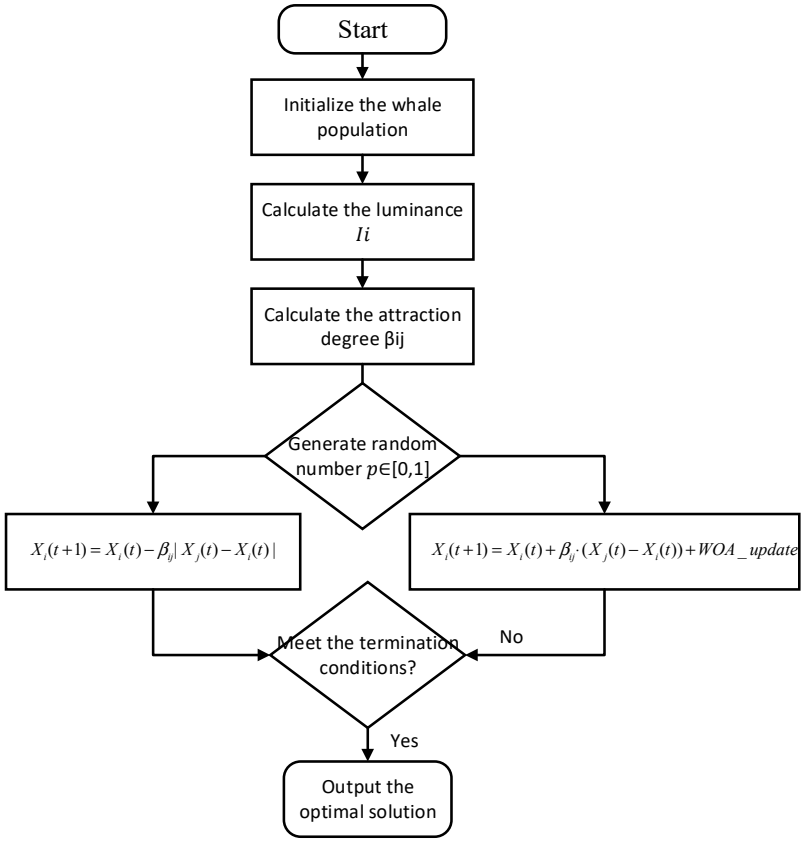


Figure 3. Flowchart of the IWOA Algorithm.

The specific steps are as follows:

(1) Initialization: Initialize the individual positions X_i of the whale group (i = 1, 2, . . ., n), and calculate the initial luminance Ii = f(X_i) of each individual and initialize the attraction parameters β_0 and γ of FA.

(2) Attractiveness calculation: In each round of iteration, calculate the luminance differences of all individuals in the group and define the attractiveness $\beta_{ij} = \beta_0 e^{-\gamma r_{ij}^2}$ accordingly to adjust the interaction influence among individuals.

(3) Position update integration: Before the position update of the standard *WOA*, the interaction weight of each individual is adjusted according to the attraction mechanism of FA to achieve the integrated optimization of local search and global exploration. That is:

$$X_i(t+1) = X_i(t) + \beta_{ij} \cdot (X_j(t) - X_i(t)) + WOA_update \quad (17)$$

(4) Termination condition: When the maximum number of iterations or the target accuracy is reached, the algorithm terminates and outputs the optimal solution.

After introducing the FA mechanism, the population guidance and individual homogeneity of the improved *WOA* algorithm (IWOA) during the optimization process have been significantly enhanced, further improving the convergence speed and global optimal ability of parameter optimization of the convolutional neural network.

2.4. Based on the MESO-TCN Intelligent Diagnosis Process

To achieve high-precision and intelligent fault identification for rolling bearings in wind power systems, this study proposes a diagnostic framework that integrates feature screening, deep modeling, and parameter optimization. As illustrated in Figure 4, the overall process consists of four key stages: signal preprocessing, feature screening, model construction, and parameter optimization.

(1) Signal decomposition and feature enhancement: The original vibration signal is decomposed using Ensemble Empirical Mode Decomposition (EEMD) to extract intrinsic mode functions (IMFs) at multiple frequency scales, which form the temporal basis for representing fault-related information.

(2) Feature screening and signal reconstruction: To suppress redundant and noisy components, a hybrid entropy index (ME), integrating information entropy, relative entropy, and cross-entropy, is employed to comprehensively evaluate each IMF. The most diagnostically relevant components are then selected and reconstructed to form a more targeted and discriminative input signal.

(3) Deep modeling and multi-scale feature extraction: The reconstructed signal is fed into the kernel-extended temporal convolutional network (ETCN), where various fault features—such as local impulses, periodic fluctuations, and long-term dependencies—are extracted in parallel using convolutional paths with diverse receptive fields. A residual fusion strategy is applied to achieve unified modeling across multiple temporal scales.

(4) Model tuning and parameter optimization: An improved whale optimization algorithm (IWOA), enhanced with a firefly-inspired brightness mechanism, is introduced to perform global search and automated tuning of key ETCN parameters, including learning rate, number of convolution kernels, and network depth. This enhances model adaptability and robustness under varying data distributions and operating conditions.

Finally, the optimized ETCN model performs fault classification and state prediction, completing the intelligent diagnosis of rolling bearing operational status. Experimental validation on multiple datasets demonstrates that the proposed method not only achieves

high classification accuracy but also exhibits superior feature extraction capacity, faster convergence, and better generalization performance compared to existing approaches.

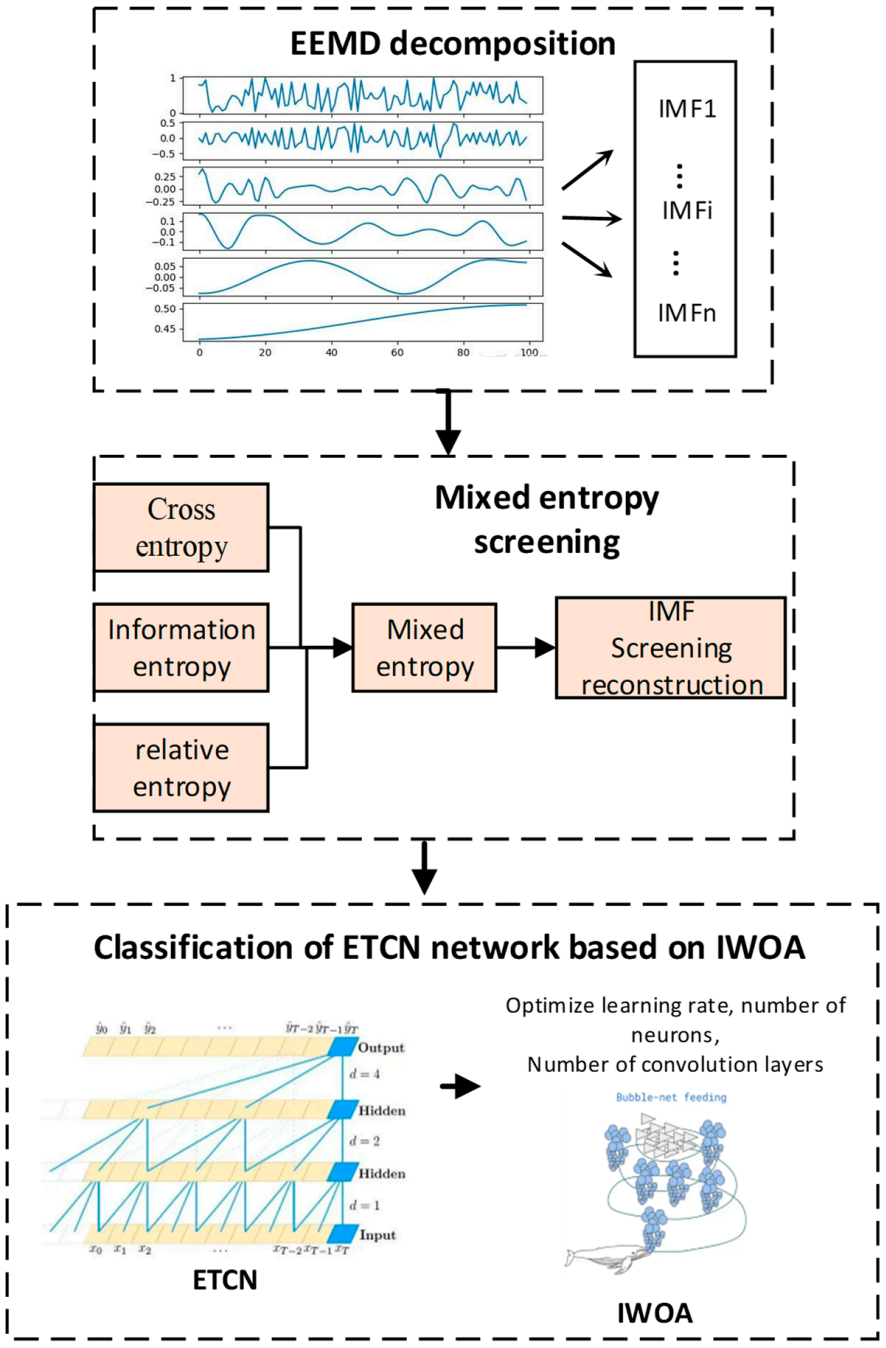


Figure 4. Fault diagnosis flow of rolling bearing based on MESO-TCN.

3. Results

To verify the universality and recognition performance of the proposed fault diagnosis method for rolling bearings, this paper selects two public datasets for experimental evaluation. Among them, the rolling bearing accelerated life test dataset of Xi'an Jiaotong University was used as the main test platform to evaluate the classification effect of the model in the typical bearing degradation process; meanwhile, the acoustic emission (AE) signal dataset of rolling bearings from Southeast University was introduced as auxiliary verification data to investigate the adaptability and robustness of the proposed method under changes in signal types and differences in sampling conditions.

3.1. Data Analysis of Rolling Bearings in Xi'an Jiaotong University

The data acquisition platform is composed of an AC motor, a motor speed controller, a rotating shaft, support bearings, a hydraulic loading system and test bearings, etc. The acquisition platform is shown in Figure 5.

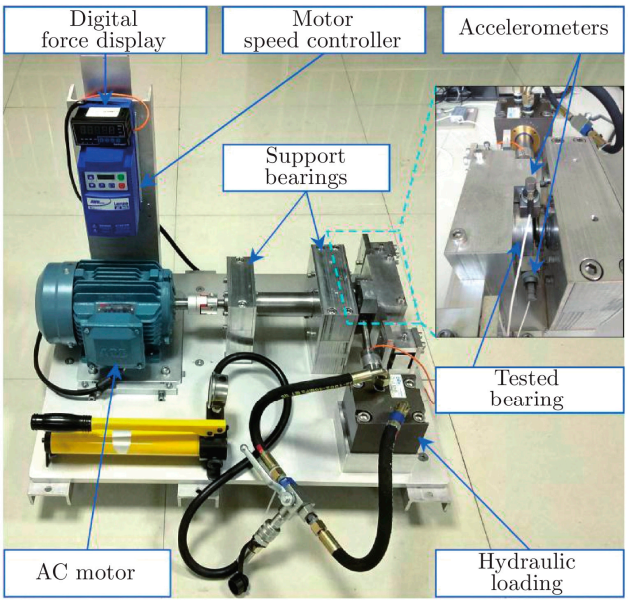


Figure 5. Xi'an Jiaotong University Rolling Bearing Simulation Failure Test Bed.

The bearings of this acquisition platform are LDK UER204 rolling bearings, and the platform uses a DT9837 portable dynamic signal collector to collect vibration signals. In the experiment, the sampling frequency was set at 25.6 kHz and the sampling interval was 1 min. A total of 3 types of working conditions were designed in the experiment, with 5 bearings in each type of working condition. This data set contains the full life cycle vibration signals of 15 rolling bearings in total. This paper selects seven of these datasets as the experimental datasets. The information of the selected dataset is shown in Table 1.

Table 1. Xi'an Jiaotong University dataset information.

Working Condition	Label	Data Set	Failure Position
1	0	Bearing 1_1	Outer circle
	1	Bearing 1_4	Cage
2	2	Bearing 2_1	Inner circle
	3	Bearing 2_2	Outer circle
	4	Bearing 2_3	Cage
3	5	Bearing 3_1	Outer circle
	6	Bearing 3_3	Inner circle

3.2. Data Preprocessing

Bearing 2_1 data subset under working condition 2 is taken as an example to carry out the preprocessing experiment. In this paper, the Integrated Empirical Mode Decomposition (EEMD) is adopted to decompose the original vibration signal, aiming to effectively distinguish the potential characteristic information from the high-frequency noise, thereby providing a more discriminative signal basis for the subsequent diagnostic modeling. Figure 6 shows several typical intrinsic mode function (IMF) components of this signal obtained under EEMD processing.

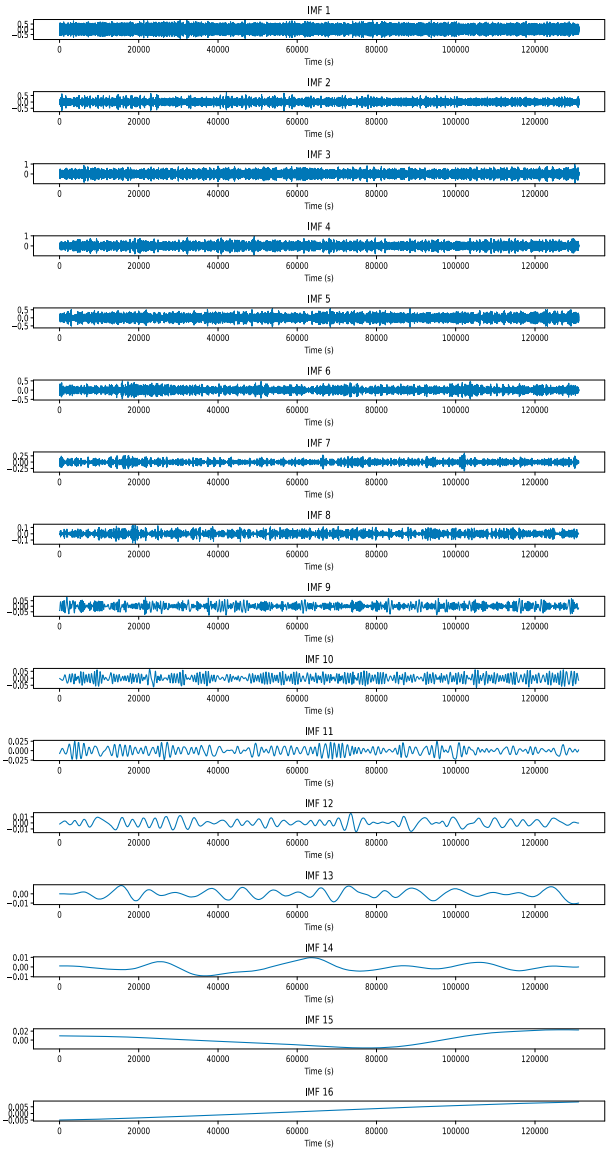


Figure 6. EEMD decomposition results.

From Figure 6, it can be observed that there are significant differences in frequency composition and amplitude variation among different IMF components, and the amount of information contained is also different. Some components still have strong noise interference. To evaluate the diagnostic value of each component, this paper quantitatively calculates all IMFs based on the mixed entropy (ME) index. The screening results indicate that components with ME values greater than 20,000 have significant advantages in information expression ability and are therefore retained for subsequent reconstruction and modeling.

In addition, to visually display the distribution of ME among various components, a bar chart is drawn as shown in Figure 7. The results indicate that the ME index can effectively distinguish between high information content and redundant components, providing a reliable basis for feature screening.

From Figure 7, it can be seen that the ME values of IMF_1, IMF_3, and IMF_4 are the highest. For the filtered *IMF* components, this article reconstructs them as follows:

$$signal = \sum_{i=1}^{n} IMF_i \tag{18}$$

In the formula, n represents the number of all IMF components with ME values greater than 20,000 compared to the original signal.

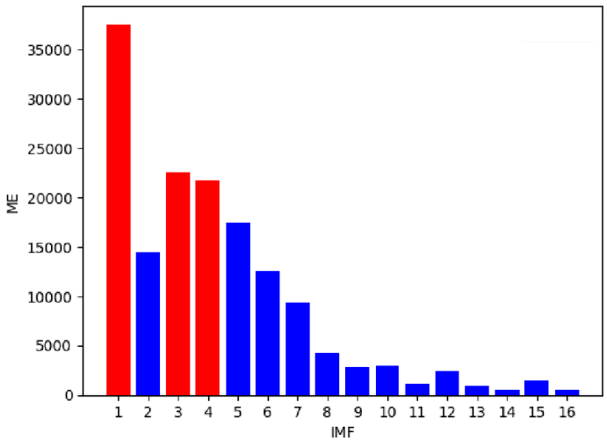


Figure 7. ME values for different IMF components. (The red bars represent IMF components with relatively large mixed entropy values, which are retained for signal reconstruction. The blue bars indicate components with lower entropy values, which are discarded as they contribute less meaningful information).

In order to ensure the effectiveness of model training and the reliability of evaluation results, after completing signal preprocessing and reconstruction in this paper, the data set is divided into the training set and the test set in a ratio of 7:3. The training set accounts for 70% of the total sample and is used for model training and parameter optimization. The test set accounts for 30% and is used to evaluate the generalization ability of the model on unseen data. During the division process, the balanced distribution of samples of each fault category in the training set and the test set to avoid the bias problem caused by category imbalance is ensured. All experiments were conducted independently based on this division and the consistency of the division results was maintained to ensure fair comparability among the results of different models.

To strictly avoid the leakage of training-test information, this paper conducts data partitioning after the completion of data preprocessing to ensure that the samples of the training set and the test set do not overlap. During the model training stage, parameter updates are only based on the training set data, and the test set data does not participate in the training and tuning process at all. This ensures the independence and fairness of the evaluation results and effectively avoids the risk of performance overestimation caused by data leakage.

Before inputting the reconstructed signal into the neural network training, in this paper, the improved Whale Optimization Algorithm (IWOA) is adopted to adaptively optimize the key structural parameters of the ETCN network. The optimization results show that when the learning rate is set to 0.0001, the number of convolutional channels is 59, and the number of convolutional layers is 2, the model performance reaches the optimum. Therefore, this group of parameters is taken as the default configuration for the subsequent experiments. Although IWOA has successfully reduced the subjectivity of traditional empirical Settings, considering the computational complexity and experimental controllability, some key hyperparameters (such as the size of convolution kernels, batch size, the number of training rounds, and the selection of optimizers) have not yet been included in the automatic optimization, which may have a certain impact on the final performance.

Specifically, the remaining network parameters are set as follows: The convolution kernel sizes are 3, 5, and 7, respectively, which are used to construct multi-scale feature extraction paths; the batch size is 32; the training period is 100 rounds. The number of

output categories is 7; the verification set ratio is 0.2; The optimizer selects Adam to enhance the training stability and convergence speed. The end of the model contains the global average pooling layer and the fully connected classification layer, and the above parameters remain consistent in all experimental Settings.

To ensure the comparability and stability of experimental results, the above parameters were kept consistent in all comparative experiments. All experiments were conducted on Anaconda's TensorFlow 2.13 platform.

The filtered and reconstructed signals were input into the optimized ETCN model for training and testing. On the test set, the model achieved a classification accuracy of 99.78% and the loss value stably converged to about 0.0124, indicating its good classification performance and training stability in fault pattern recognition tasks. Meanwhile, the trend of accuracy and loss function changes during the training process show that the model has a fast convergence speed and low fluctuation amplitude, which verifies the optimization effect of the proposed structure under complex vibration signals. The relevant training process curve is shown in Figure 7.

As shown in Figure 8, the accuracy of the model rapidly improves with the number of iterations during the training process, and tends to saturate after about the 20th round, basically stabilizing at a level close to 100%. At the same time, the loss function shows a continuous downward trend and rapidly converges in the early stages, with the minimum value remaining at a low level, indicating that the constructed model performs well in both training efficiency and convergence performance.

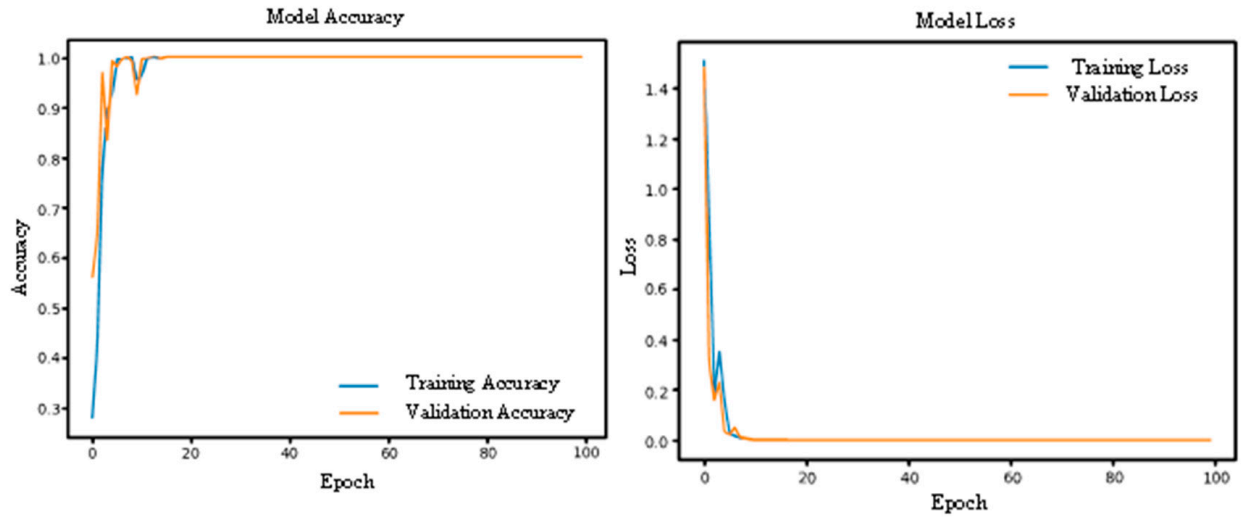


Figure 8. Iterative plot of model accuracy and loss rate based on MESO-TCN.

To further analyze the feature discrimination ability of the model, this paper uses the T-SNE method to reduce and visualize the high-dimensional feature vectors extracted by the network, as shown in Figure 9. The clear distribution of samples of different categories on the two-dimensional plane and the clear clustering boundaries indicate that the proposed method has good feature abstraction and category separability.

In addition, to evaluate the specific classification performance of the model on different types of faults, Figure 10 shows the confusion matrix results of the model on the test set. The horizontal and vertical axes in the figure correspond to the predicted and true labels of the model, respectively. It can be observed from the figure that the classification accuracy of each category sample is high, and the misjudgment rate is low, further verifying the reliability and robustness of the proposed method in multi-class fault diagnosis tasks.

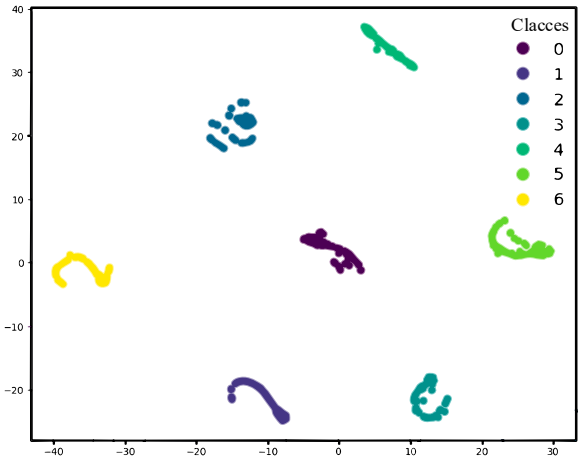


Figure 9. MESO-TCN based T-SNE classification.

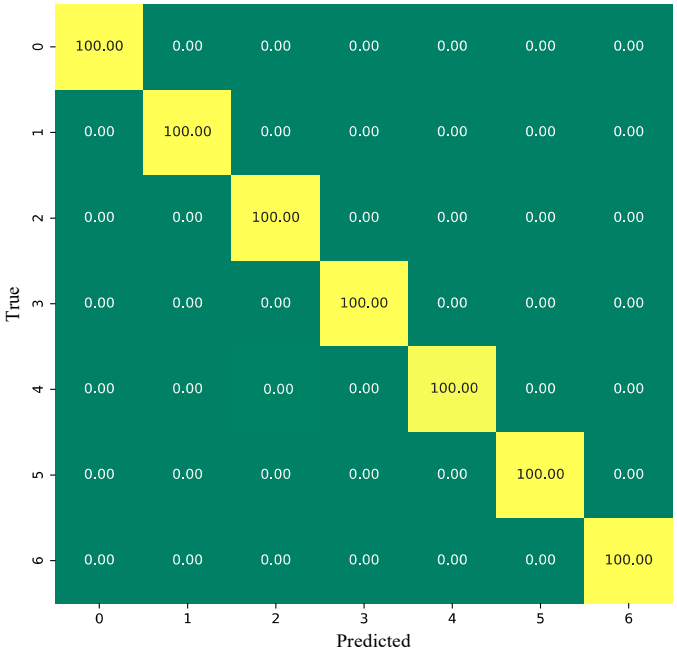


Figure 10. Confusion matrix based on MESO-TCN models.

Figure 9 shows the T-SNE visualization results of the proposed MESO-TCN method in the feature space. The samples of each category are densely distributed and have clear boundaries on the two-dimensional plane. Good intra class clustering and inter class separation have been achieved for all seven types of faults, verifying the discriminative ability of this method in high-dimensional feature extraction and dimensionality reduction mapping.

Figure 10 shows the confusion matrix results of MESO-TCN on the test set, indicating that the model achieved 100% recognition accuracy for all 7 types of faults without significant classification confusion. This result indicates that the proposed model not only has strong feature expression ability, but can also achieve high-precision recognition of multi-category fault states, with a good fault mode separation ability and practical application potential.

3.3. Ablation Test

To comprehensively evaluate the diagnostic performance of the proposed MESO-TCN method, this paper selects multiple comparative models for experimental comparison, including the original TCN, structurally optimized ETCN, EMD-based feature preprocessing EMD-TCN, EMD-ME-TCN with a mixed entropy screening mechanism, and EMD-ME-

CNN under a convolutional network structure. All methods are trained and tested under the same data conditions, and classification accuracy is uniformly used as the performance evaluation metric.

To reduce the impact of random errors on the results, each experiment was repeated 5 times, and the average accuracy and average loss value were calculated as the final evaluation results. The comparative experimental results are shown in Table 2.

Table 2. Accuracy and loss rate values for different comparison models.

Method	Accuracy Rate/%	Loss Rate	F1 Score/%	Precision/%	Recall/%
TCN	60.85	0.8370	68.16	65.90	60.72
ETCN	52.91	0.9980	60.05	69.84	63.50
EMD-TCN	66.95	0.6830	68.32	66.88	69.85
EMD-ETCN	61.36	0.8210	61.07	63.68	79.03
EMD-ME-TCN	99.40	0.0193	94.66	83.53	85.94
EMD-ME-CNN	99.47	0.0176	91.92	89.84	90.72
MES-TCN	98.45	0.0183	91.45	90.85	87.45
MESO-TCN	99.78	0.0124	98.47	97.45	94.78

According to Table 2, MESO-TCN achieved an average classification accuracy of 99.78% after 100 rounds of training, with a stable loss value of 0.0124, and overall performance better than all other comparison methods. Among them, compared to the traditional TCN model, the accuracy has been improved by 38.94%; compared to the ETCN, EMD-TCN, and EMD-ME-TCN models, the improvement rates are 46.88%, 32.83%, and 38.42%, respectively; even when compared to the EMD-ME-CNN method that is closer in structure, MESO-TCN still has a 0.38% improvement advantage, further verifying the effectiveness of the proposed multi-strategy fusion mechanism.

The above results indicate that the synergistic effect of feature selection, structural expansion, and parameter optimization significantly enhances the discriminative ability and training stability of the model, proving that the proposed MESO-TCN method has good accuracy and robustness in complex working conditions.

To further validate the performance of different methods in feature extraction ability, Figure 11 shows the feature distribution results based on T-SNE visualization.

It can be observed that there is a significant overlap between the traditional TCN and the structurally extended ETCN model in the feature space, and the clustering of samples of different categories is unclear, indicating that the extracted features have strong confounding. Although the EMD-TCN and EMD-ETCN methods introduce the signal decomposition mechanism, the clustering structure is still not ideal, and the category boundaries are ambiguous. Especially under the EMD-ME-TCN method, there is a significant overlap between the fault samples of Type 2 and Type 6, resulting in a decrease in diagnostic discrimination.

In contrast, the proposed MESO-TCN method shows good separability in the feature distribution under the T-SNE mapping: the samples of the same type of faults are compactly aggregated, and the boundaries between different fault categories are clear, demonstrating strong inter-class discriminability and intra-class consistency. This result indicates that MESO-TCN has a stronger separation ability in the feature expression space and can effectively enhance the discrimination effect of fault modes.

Based on the comprehensive comparison results of accuracy rate, loss value, and feature visualization, the comprehensive advantages of the proposed method in feature extraction, discriminant modeling, and classification performance were further verified, providing an effective solution for the multi-category fault identification of rolling bearings.

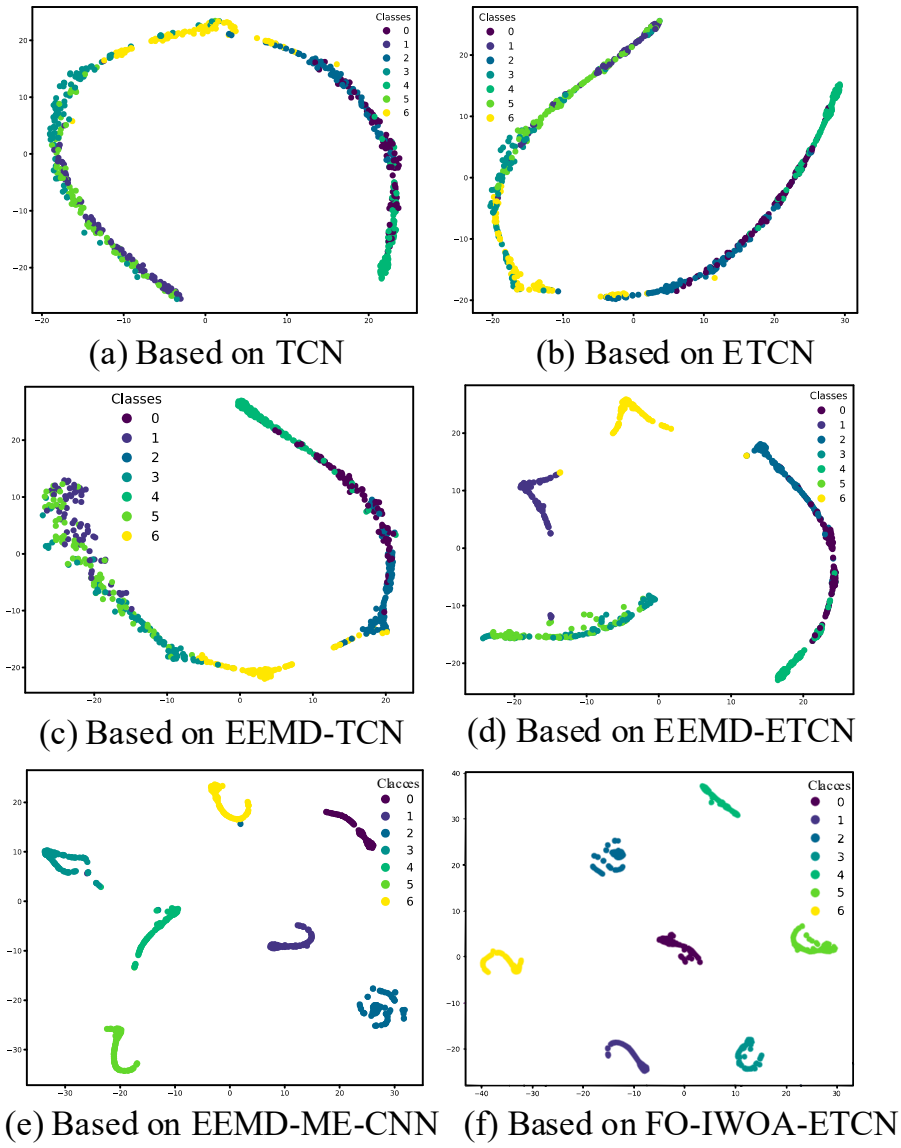


Figure 11. T-SNE diagrams of different models.

3.4. Analysis of AE Signal Data of Rolling Bearings in Southeast University

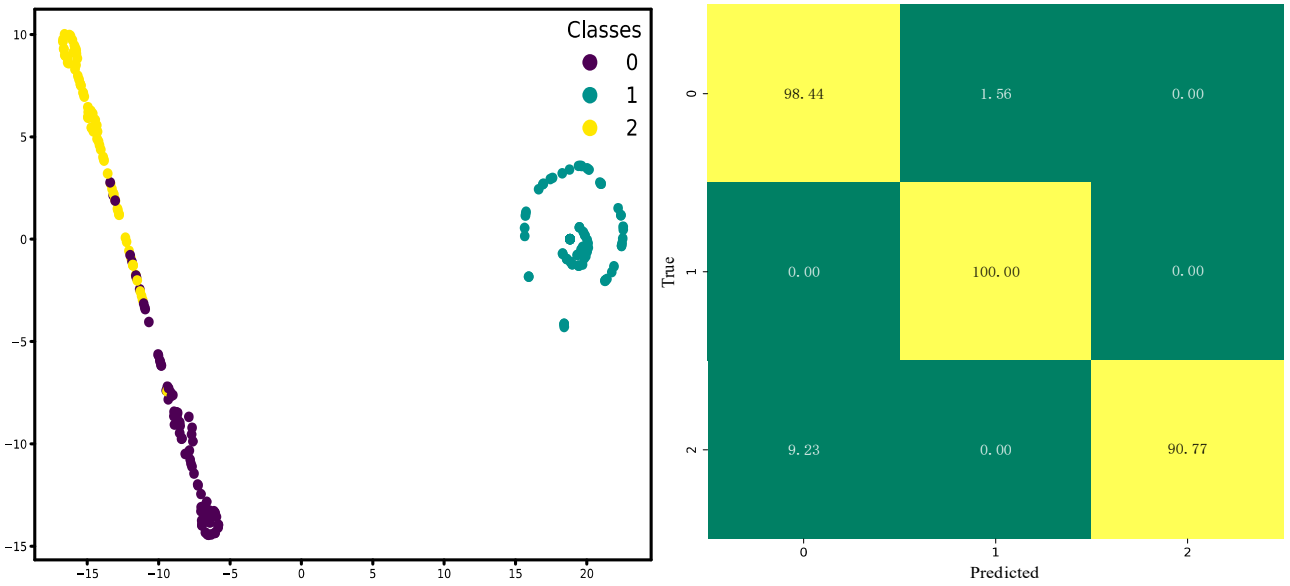
To further verify the adaptability and generalization ability of the proposed method under different signal types and test conditions, this paper selects the acoustic emission (AE) signal dataset of rolling bearings from Southeast University for external verification. This dataset adopts UC210 type rolling bearings, and the spindle speed is set at 30 r/min. During the experiment, three bearing states were installed in sequence on the test platform: normal bearings, faulty bearings with an inner ring crack width of 0.5 mm, and severely faulty bearings with a crack width of 0.8 mm. The AE signals of these bearings in the stable operation state were collected.

The AE signal acquisition system is composed of the UT-1000 wideband acoustic emission sensor, a 40 dB preamplifier, a PCI-2 acoustic emission acquisition card and an industrial control computer. The signal samples collected for each type of state are consistent in both quantity and duration to ensure data balance.

In this test, the model structure and parameter Settings adopted were consistent with those of the aforementioned main experiment, with only the number of output categories adjusted from 7 to 3. The diagnostic results and the visualization of their characteristics on this dataset are shown in Table 3 and Figure 12.

Table 3. Comparative experiments of multiple methods based on AE data from Southeast University.

Method	Accuracy Rate/%	Loss Rate	F1 Score/%	Precision/%	Recall/%
TCN	60.89	0.7370	57.45	65.90	64.77
ETCN	59.35	0.8280	64.12	45.75	45.12
EMD-TCN	62.77	0.6385	65.45	54.41	75.45
EMD-ETCN	59.63	0.6690	62.33	66.11	67.41
EMD-ME-TCN	94.87	0.1270	86.45	85.44	89.44
EMD-ME-CNN	94.87	0.1370	90.85	86.45	87.41
MES-TCN	93.45	0.0173	90.31	91.85	86.45
MESO-TCN	95.82	0.1770	97.45	98.91	96.11



(**a**) T-SNE (**b**) Confusion matrix

Figure 12. Classification T-SNE and Confusion Matrix of AE data of Southeast University based on the method of this paper.

It can be seen from Table 3 that the average classification accuracy of the proposed MESO-TCN model on the AE signal bearing dataset of Southeast University reaches 95.82%, showing better recognition performance compared with traditional methods, verifying the robustness and adaptability of this method under different signal types and sampling conditions.

Figure 12 further shows the T-SNE feature visualization and confusion matrix results of the model on this dataset. It can be seen from Figure 12a that the healthy state and the fault state have good separability in the low-dimensional feature space, and the boundaries between samples are clear. The confusion matrix shown in Figure 12b indicates that the model can stably distinguish three types of fault categories, and the classification results basically have no misjudgment or missed judgment.

The above results show that the proposed model not only has excellent performance on the accelerated lifetime dataset of Xi'an Jiaotong University, but can also maintain a high diagnostic accuracy rate under cross-platform and heterogeneous signal conditions, and has a strong generalization ability and engineering application potential.

3.5. PRONOSTIA Bearing Failure Dataset

The PRONOSTIA bearing failure dataset was released by the FEMTO-ST Institute in France. This platform builds a highly controllable and realistic bearing testing environment by simulating the real working conditions of industrial sites, including constant and adjustable loads, high-speed rotation, temperature control and other environments.

PRONOSTIA data features real degradation trajectories, clear fault labels, and a high sampling frequency (1 s of vibration data is collected every 10 s, with a sampling rate of 25.6 kHz), including inner ring spalling faults, outer ring spalling faults, rolling element fatigue faults, natural compound faults, and compound faults.

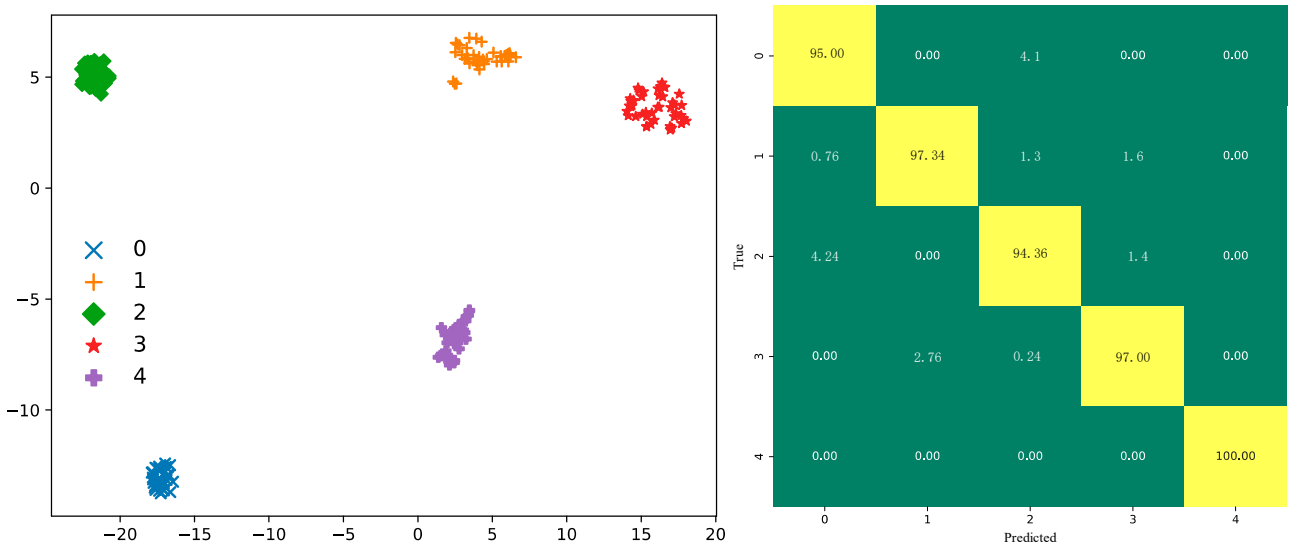
In this test, the model structure and parameter Settings adopted were consistent with those of the aforementioned main experiment, with only the number of output categories adjusted from 7 to 5. The diagnostic results and the visualization of their characteristics on this dataset are shown in Table 4 and Figure 12.

Table 4. Comparison of Multiple Methods for the PRONOSTIA bearing failure Dataset.

Method	Accuracy Rate/%	Loss Rate	Time/%	Precision/%	Recall/%	Method
TCN	57.46	0.7841	45	45.12	56.41	75.45
ETCN	65.45	0.9516	67	75.45	45.12	72.42
EMD-TCN	67.12	0.8644	135	71.64	85.75	76.34
EMD-ETCN	57.66	0.7341	168	63.44	56.34	73.45
EMD-ME-TCN	93.85	0.2270	149	61.23	56.41	75.99
EMD-ME-CNN	94.75	0.3451	167	67.89	75.94	78.41
MES-TCN	93.45	0.2574	150	78.45	71.94	75.41
MESO-TCN	96.73	0.1341	231	90.54	98.15	95.44

It can be seen from Table 4 that the average classification accuracy of the proposed MESO-TCN model on the PRONOSTIA bearing fault dataset reaches 96.73%, showing better recognition performance compared with traditional methods, verifying the robustness and adaptability of this method under different signal types and sampling conditions.

Figure 13 further presents the T-SNE feature visualization and confusion matrix results of the model on this dataset. It can be seen from Figure 13a that the healthy state and the fault state have good separability in the low-dimensional feature space, and the boundaries between samples are clear. The confusion matrix shown in Figure 13b indicates that the model can stably distinguish five types of fault categories, and the classification results basically have no misjudgment or missed judgment.



(**a**) T-SNE (**b**) Confusion matrix

Figure 13. The T-SNE and confusion matrix of the PRONOSTIA bearing fault dataset based on the method proposed in this paper.

The above results show that the proposed model not only has excellent performance on the experimental dataset but can also maintain a high diagnostic accuracy rate under

the conditions of the actual industrial dataset, and has strong generalization ability and engineering application potential.

3.6. Result Summary

To further verify the diagnostic performance of the proposed MESO-TCN model, this paper compares it with some representative methods that have existed on the rolling bearing dataset of Xi'an Jiaotong University in recent years. The MTF-CBAM-LCNN method proposed in reference [21] achieved a fault diagnosis accuracy rate of 99.47% on this dataset. Reference [22] combines the convolutional Denoising Autoencoder (CDAE) with the Bidirectional Long Short-Term Memory Network (Bi-LSTM), and the model accuracy rate reaches 88.65%. The Adaptive Signal Diagnosis Network (ASDN) framework developed in reference [23] achieved an accuracy rate of 97.42% on the same dataset. In contrast, the MESO-TCN model proposed in this paper has an average diagnostic accuracy rate of 99.78% on the same dataset and is superior to the existing methods in terms of recognition accuracy.

4. Conclusions

This paper proposes a novel fault diagnosis method for rolling bearings—MESO-TCN—which integrates multi-scale feature screening, network structure optimization, and adaptive parameter tuning. The method combines Ensemble Empirical Mode Decomposition (EEMD) with a hybrid entropy mechanism for signal preprocessing, employs an extended-kernel temporal convolutional network (ETCN) for multi-scale temporal feature extraction, and introduces an improved whale optimization algorithm (IWOA) for the joint optimization of structural parameters. Extensive experiments were conducted on rolling bearing datasets from Xi'an Jiaotong University and Southeast University, yielding the following conclusions:

(1) On the accelerated life test dataset from Xi'an Jiaotong University, the MESO-TCN model achieved an average diagnostic accuracy of 99.78%, with the loss value stably converging to 0.0124. Compared with baseline models such as TCN, ETCN, EMD-TCN, EMD-ME-TCN, and EMD-ME-CNN, the proposed model improved accuracy by 38.94%, 46.88%, 32.83%, 38.42%, and 0.38%, respectively, demonstrating superior diagnostic performance. On the acoustic emission (AE) signal dataset from Southeast University, the model also achieved an average accuracy of 95.82%, validating its robustness and generalization capability under heterogeneous signal conditions.

(2) Feature visualization results based on t-SNE indicate that the features extracted by MESO-TCN exhibit clear inter-class separation and tight intra-class clustering, reflecting strong discriminative power and interpretability. The confusion matrix further confirms that the model can accurately and consistently identify all fault categories with extremely low misclassification rates, highlighting its expressive strength and reliability in both feature extraction and fault classification.

(3) The results of comparative experiments confirm the effectiveness of the multi-strategy fusion framework proposed in this study. The hybrid entropy-based feature screening significantly enhances the discriminability of input signals; the extended-kernel TCN improves the model's capacity to capture multi-scale dynamic features; and the IWOA optimization mechanism ensures efficient and stable parameter tuning. Overall, MESO-TCN achieves high accuracy, rapid convergence, and strong generalization in multi-class fault diagnosis tasks, showing promising potential for engineering applications.

Author Contributions: Methodology, R.G.; Software, J.Z.; Formal analysis, Y.W.; Data curation, K.X.; Writing—review & editing, Y.S. All authors have read and agreed to the published version of the manuscript.

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 APC was funded by the same sources.

Data Availability Statement: The data used in this study are publicly available benchmark datasets. Specifically, the Xi'an Jiaotong University dataset, Southeast University acoustic emission dataset, and the PRONOSTIA bearing dataset were employed. No new data were generated in this study. These datasets can be accessed from their respective official sources or upon reasonable request.

Conflicts of Interest: Author Ruibin Gao was employed by the company Guoneng Jiangsu Electric Power Engineering Technology Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Lu, Z.; Li, L.; Zhang, C.; Zhao, S.; Gong, L. Fault Feature Extraction Based on Variational Modal Decomposition and Lifting Wavelet Transform: Application in Gear of Mine Scraper Conveyor Gearbox. *Machines* **2024**, *12*, 871. [CrossRef]
2. Zhang, X.; Wang, H.X. A Bearing Fault Diagnosis Model Based on Synchronous Squeeze Wavelet Transform and Transformer. *Electromech. Eng.* **2024**, *3*, 1–10.
3. Dong, S.; Zhang, Y.; Wang, S. Multisensor Fault Diagnosis of Rolling Bearing with Noisy Unbalanced Data via Intuitionistic Fuzzy Weighted Least Squares Twin Support Higher-Order Tensor Machine. *Machines* **2025**, *13*, 445. [CrossRef]
4. Zeng, M.; Li, S.; Li, R.; Zhang, Y.; Wang, Q.; Liu, H.; Chen, X.; Zhao, L.; Huang, J.; Xu, F. Intelligent Fault Diagnosis Method for Rolling Bearings Based on Stacked Sparse Discriminant Autocoding. *Bearing* **2024**, *3*, 77–83, 98.
5. Fu, W.; Jiang, X.; Li, B.; Tan, C.; Chen, B.; Chen, X. Rolling Bearing Fault Diagnosis Based on 2D Time-Frequency Images and Data Augmentation Technique. *Meas. Sci. Technol.* **2023**, *34*, 045005. [CrossRef]
6. Gu, Y.K.; Shi, C.W.; Chen, J.F. Fault Diagnosis of Planetary Gearbox Based on Gram Angle Field and Deep Convolution Generative Adversarial Network. *Noise Vib. Control.* **2024**, *44*, 111–118.
7. Wang, L.; Wang, Y.; Guo, F.; Yan, H.; Zhao, F. Lower Limb Joint Angle Prediction Based on Multistream Signaling and Quantile Regression, Temporal Convolution Network–Bidirectional Long Short-Term Memory Network Neural Network. *Machines* **2024**, *12*, 901. [CrossRef]
8. Xiaojing, U.; Xue, R.M. Short Term Electricity Price Prediction of Bi-LSTM-TCN Based on Wavelet Transform. *Adv. Technol. Electr. Eng. Energy* **2020**, *24*, 16453–16482.
9. Wang, X.-B.; Zhang, X.; Li, Z.; Wu, J. Ensemble Extreme Learning Machines for Compound-Fault Diagnosis of Rotating Machinery. *Knowl.-Based Syst.* **2020**, *188*, 105012. [CrossRef]
10. Fan, J.; Zhang, K.; Huang, Y.; Zhu, Y.; Chen, B. Parallel Spatio-Temporal Attention-Based TCN for Multivariate Time Series Prediction. *Neural Comput. Appl.* **2023**, *35*, 13109–13118. [CrossRef]
11. Zhang, B.; Wang, S.; Deng, L.; Jia, M.; Xu, J. Ship Motion Attitude Prediction Model Based on IWOA-TCN-Attention. *Ocean. Eng.* **2023**, *272*, 113911. [CrossRef]
12. Gao, Z.W. Units Collaborative Diagnosis with Label Heterogeneity and Communication Redundancy. *Eng. Appl. Artif. Intell.* **2025**, *152*, 110724. [CrossRef]
13. Swain, S.; Khilar, P.M.; Swain, R.R.; Senapati, B.R. A Federated Learning Based Fault Diagnosis in UAV-Reliable Sensor Network. In Proceedings of the 2023 IEEE 20th India Council International Conference (INDICON), Hyderabad, India, 14–17 December 2023; IEEE: Piscataway, NJ, USA, 2024; pp. 1386–1391.
14. Pan, X.J.; Dong, S.J.; Zhu, P.; Zhou, C.; Song, K. Prediction of Remaining Life Migration of Rolling Bearings under Variable Operating Conditions Based on TCN and Residual Self Attention. *Vib. Shock.* **2024**, *43*, 238–259.
15. Ding, L.; Li, Q. Fault Diagnosis of Rotating Machinery Using Novel Self-Attention Mechanism TCN with Soft Thresholding Method. *Meas. Sci. Technol.* **2024**, *35*, 047001. [CrossRef]
16. Sun, Y.; Ding, Q.; Xia, Y.; Li, C. Fault Diagnosis of Chiller Based on the Combination of Multiple Blocks and Self-Attention TCN. *J. Process Eng.* **2024**, *24*, 162–171.
17. Hu, C.; Li, G.; Ma, L.; Yan, X.; Wei, H. Improve the Fault Diagnosis Method of TCPN Variable Working Condition Bearings. *Noise Vib. Control.* **2022**, *42*, 134–141.

18. Li, S.; Yang, Y.; Li, C.; He, H.; Zhang, Q.; Zhao, S. Research on Signal Processing Technology of Ultrasonic Non-Destructive Testing Based on EEMD Combined with Wavelet Packet. *IEEJ Trans. Electr. Electron. Eng.* **2023**, *18*, 686–700. [CrossRef]
19. Shannon, C.E. A Mathematical Theory of Communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [CrossRef]
20. Nadimi-Shahraki, M.H.; Zamani, H.; Varzaneh, Z.A.; Mirjalili, S. A Systematic Review of the Whale Optimization Algorithm: Theoretical Foundation, Improvements, and Hybridizations. *Arch. Comput. Methods Eng.* **2023**, *30*, 4113–4159. [CrossRef]
21. Liu, W.; Liu, S.; He, Y.; Wang, J.; Gu, Y. Rolling Bearing Fault Diagnosis Based on MTF Encoding and CBAM-LCNN Mechanism. *Comput. Mater. Contin.* **2025**, *82*, 4863. [CrossRef]
22. Yao, X.; Zhu, J.; Jiang, Q.; Yao, Q.; Shen, Y.; Zhu, Q. RUL Prediction Method for Rolling Bearing Using Convolutional Denoising Autoencoder and Bidirectional LSTM. *Meas. Sci. Technol.* **2023**, *35*, 035111. [CrossRef]
23. Li, A.; Yao, D.; Yang, J.; Chang, M.; Zhou, T. Bearing Diagnosis Using an Anti-Noise Neural Network Based on Selectable Branch Multiscale Modules and Attention Mechanisms. *IEEE Sens. J.* **2023**, *24*, 5830–5840. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.