Rolling bearing fault diagnosis method based on adaptive signal diagnosis network and its application

Indexed by:
Web of Science Group

Jing Zhu[a,*], Ou Li[a], Minghui Chen[b], Bingbing Hu[b], EnHui Ma[b]

[a] School of Vehicle and Transportation Engineering, Henan University of Science and Technology, China
[b] College of Vehicle and Transportation Engineering, Henan University of Science and Technology, China

Highlights

- Rolling bearing fault diagnosis.

- Temporal Convolutional Network(TCN).

- Adaptive Signal Diagnostic Network.

Abstract

To address the issues of fixed convolution kernel sizes and weak targeting of extracted features in the application of deep learning for fault diagnosis, this paper develops a fault diagnosis framework called Adaptive Signal Diagnostic Network(ASDN), which integrates Adaptive Temporal Convolutional Neural Networks (AT-CNN) with Ensemble Empirical Mode Decomposition (EEMD) and Singular Spectrum Analysis (SSA). During the adaptive preprocessing stage, improvements are made to EEMD for multi-scale signal decomposition, capture the transient changes of the signal adaptively. Concurrently, enhancements are applied to SSA for further extraction of fault trends and periodic components, optimizing the signal representation. In the adaptive deep learning stage, an innovative Temporal Convolutional Network (TCN) with a dynamic adjustment mechanism is developed. This network constructs a novel dynamic convolution kernel adjustment mechanism, enabling the neural network to adapt its convolution kernel size based on different frequency components, thereby accurately processing signals with varying frequencies. Validation on datasets from Case Western Reserve University and Xi'an Jiaotong University demonstrates that the proposed method achieves superior performance, with diagnostic accuracies of 100% and 97.42% on the two public datasets, respectively.

Keywords
rolling bearing fault diagnosis, TCN, Adaptive Signal Diagnostic Network, Singular Spectrum Analysis

1. Introduction

The vibration signals generated by rolling bearings typically exhibit significant nonlinearity and non-stationary characteristics. The noise interference and the mixture of complex frequency components in the signals make it challenging to extract effective fault features from them[1].

Although deep learning methods have distinct advantages in handling large-scale data and automatic feature learning, the features automatically learned during fault feature extraction often lack specificity and interpretability. Consequently, models may perform poorly when confronted with unknown or varying

(*) Corresponding author.
E-mail addresses: J. Zhu (ORCID: 0000-0003-3146-875X) 1907175086@qq.com, O. Li (ORCID: 0000-0001-6268-6982) 3149372600@qq.com, M. Chen 867303403@qq.com, B. Hu 9906238@haust.edu.cn, E. Ma jingzhu244@gmail.com

fault characteristics.

Traditional signal processing methods can offer more intuitive and interpretable data analysis, aiding in the extraction of key fault features from nonlinear and non-stationary signals[2-4]. These methods can compensate for the limitations of deep learning approaches in fault diagnosis.

By extracting key features of fault signals through signal processing and subsequently inputting these features into deep learning models for further processing, this hybrid approach can fully leverage the advantages of both techniques. This improves the accuracy and robustness of fault diagnosis. Not only does this enhance model performance, but it also increases its adaptability to changing conditions in industrial environments.

Ensemble Empirical Mode Decomposition (EEMD), as an effective signal processing technique, can reveal subtle variations and dynamic characteristics within signals. Therefore, it exhibits outstanding performance in handling nonlinear and non-stationary signals [5-7]. Fang Dao [8] et al. proposed the Wavelet Thresholding and Ensemble Empirical Mode Decomposition (WT-EEMD) method for denoising acoustic vibration signals from hydro turbine runners under normal and sediment-laden flow conditions[8]. Feiyu Li et al. introduced a hybrid algorithm combining Ensemble Empirical Mode Decomposition (EEMD) and information entropy [9]. However, using EEMD alone has limitations in extracting the trend and periodic components of signals. The trend component typically reflects the long-term variations and evolution of signals, which may be related to bearing wear or gradually deteriorating fault conditions. Meanwhile, the periodic component can reveal features corresponding to the mechanical rotation period, such as specific damages or abnormalities in bearing components. Although EEMD is capable of capturing transient variations in signals, when dealing with complex time-series features with long-term trends and low-frequency components, it may overlook crucial time-series characteristics within signals, which are often key clues for predicting and diagnosing early faults.

To address this shortcoming, Singular Spectrum Analysis (SSA) is employed to further process the output of Ensemble Empirical Mode Decomposition (EEMD) [10]. By constructing a trajectory matrix and applying Singular Value Decomposition, SSA is not only capable of clearly identifying and extracting the principal trends and periodic components of the signal, but also effectively separates and reconstructs the noise and trend elements within the signal. This results in the provision of more precise fault diagnosis information.

Although EEMD and SSA each have theoretical advantages, their combined application is limited in practice due to the lack of an effective subcomponent selection optimization mechanism. In cases where signal characteristics are non-stationary, developing an effective optimization mechanism will greatly enhance the complementary strengths of these two methods, achieving maximum diagnostic efficacy. Moreover, an improved preprocessing procedure that effectively eliminates noise and interference frequency components will enhance the generalization capability and diagnostic accuracy of subsequent models.

Given the inherent characteristics of rolling bearing vibration signals—marked time dependency and periodic variations—this study selects the Temporal Convolutional Network (TCN) as the primary analytical tool[11]. TCN, through its unique dilated convolution architecture, can effectively capture long-term temporal dependencies, which is crucial for detecting periodic fault signals caused by subtle mechanical wear. Li Ding [12] proposed a novel method named "Self-Attention Mechanism with Temporal Convolutional Network and Soft Thresholding Algorithm (SAM-TCN-ST)" for intelligent fault recognition in rotating machinery. Chao Zhang [13] introduced a new fault diagnosis approach—the Bayesian Augmented Temporal Convolutional Network (BATCN)—to filter raw signals for defect detection in wind turbine pitch bearings. However, TCN's fixed kernel size presents a limitation in adapting to dynamic signal characteristics when dealing with multi-scale fault signals, and it exhibits insufficient sensitivity and responsiveness to changes in signal features under varying operational conditions.

To address this challenge, this paper proposes the Dynamic Adjustment Temporal Convolutional Network (DA-TCN) model, which introduces an innovative dynamic kernel adjustment mechanism. This mechanism enables the TCN to dynamically adjust its kernel size according to the actual frequency distribution of the signal.

This paper addresses the issues of strong noise interference and frequency aliasing in rolling bearing signals by proposing

a novel fault diagnosis framework based on the mixture of the signal processing and deep learning. The specific contributions and innovations are as follows:

(1) Dynamic convolutional Kernel Adjustment Mechanism : This study innovatively develops a dynamic adjustment mechanism based on the frequency characteristics of the input signal. Unlike traditional fixed kernel sizes, the convolutional kernel size of the DA-TCN can adaptively adjust according to signal frequency variations. This mechanism significantly enhances the model's adaptability and generalization to complex industrial signal changes.

(2) SSA Subcomponent Reconstruction Strategy: By analyzing the w-correlation graph, singular spectrum components of the same period are selected and merged. The variance contribution rate of the merged components is then calculated, and subcomponents containing substantial fault signal energy and information are chosen for input into the DA-TCN.

(3) EEMD Signal Selection and Reconstruction: Utilizing statistical indicators such as variance contribution rate, correlation coefficient, and permutation entropy, the signals decomposed by EEMD are finely selected and reconstructed. This optimization of signal component selection significantly

enhances the clarity of fault feature expression and the reliability of diagnosis.

2. Proposed Approach

The proposed ASDN (Adaptive Signal Diagnosis Network), as illustrated in Figure 1, consists of two main components: the preprocessing stage and the deep learning stage.

In the preprocessing stage, EEMD and SSA are utilized for signal analysis. To enhance the accuracy and efficiency of signal processing, this study introduces an evaluation method for EEMD signal components based on variance contribution rate, correlation coefficient, and permutation entropy. Additionally, in the SSA process, a component selection and reconstruction strategy based on the w-correlation graph is incorporated. In the deep learning stage, dynamically Adjusted Temporal Convolutional Network (DA-TCN) is constructed innovatively. This model not only includes convolutional layers that can adaptively adjust their size based on the frequency characteristics of the input signal but also incorporates activation layers, normalization layers, and dropout layers to enhance the model's generalization capability and stability. The structure of the model is illustrated in the following diagram.
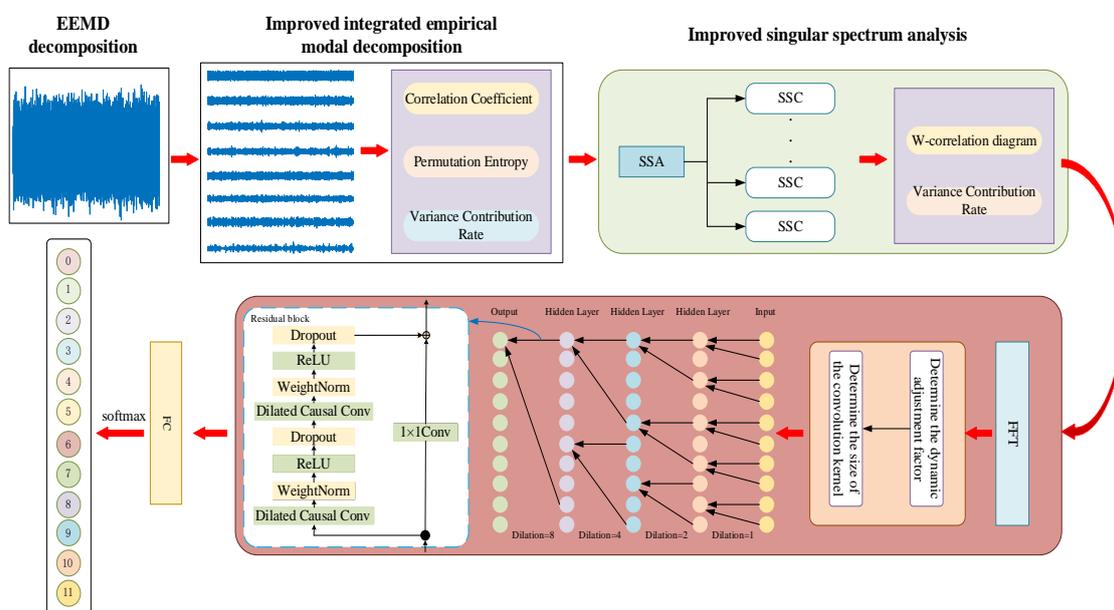

Figure1. The network architecture of ASDN proposed in this paper.

2.1. EEMD and Its Improvements

The limitation of EMD lies in the occurrence of different oscillatory components coexisting within a single IMF

(Intrinsic Mode Function), while very similar oscillations may exist in different IMFs, known as mode mixing. To address this issue, EEMD (Ensemble Empirical Mode Decomposition) was introduced. EEMD leverages the characteristics of Gaussian

white noise with a uniform frequency distribution and zero mean. It achieves this by introducing Gaussian white noise multiple times during the decomposition process to alter the extrema of the signal. Subsequently, the IMF components obtained from multiple EMD iterations are averaged to counteract the introduced Gaussian white noise, thereby suppressing the problem of mode mixing. The computational steps for a signal $X(t)$ are as follows:

(1) Adding Gaussian white noise to the original signal to obtain a series of new signals:

$$X_i(t) = X(t) + n_i(t) \tag{1}$$

Where $n_i(t)$ represents white noise signals of the same length as $X(t)$, and i = 1, 2, ..., M , where M is the number of ensemble experiments.

(2) Decomposing the obtained signals using EMD to obtain a series of IMF:

$$X_i(t) = \sum_{j=1}^{N} c_{ij}(t) + r_i(t) \tag{2}$$

Where j=1,2, …, N, with N denoting the number of IMFs, $c_{ij}(t)$ represents the IMFs (c_{i1}, c_{i2}，…，c_{iN}), and $r_i(t)$ denotes the residual of the i th trajectory.

(3) Performing ensemble averaging on the IMF components.

$$c_j(t) = \frac{1}{M} \sum_{i=1}^{M} c_{ij}(t) \tag{3}$$

During the decomposition process of EEMD, although multiple components are successfully extracted from the original signal, these components exhibit significant differences in their physical significance and contributions. Therefore, this study introduces three statistical indicators—variance contribution rate, correlation coefficient, and permutation entropy—as quantitative tools to assess the importance of each component. This comprehensive evaluation strategy is based on each component's contribution to the total signal variance, the linear correlation between components, and the time series complexity of the components, determining their utility in subsequent analysis.

Specifically, the variance contribution rate measures the importance of each component in signal reconstruction. A higher variance contribution rate indicates that the component plays a crucial role in the overall signal. The correlation

coefficient evaluates the linear correlation between the component and the target fault features. A higher correlation coefficient suggests a close association between the component and the fault state. Permutation entropy, as a nonlinear tool for measuring signal complexity, reveals the dynamic characteristics of the component's time series. High permutation entropy indicates that the component possesses high time series complexity and information content. By computing these three indicators and assessing their combined values, it is possible to effectively select the IMF components obtained from EEMD decomposition that are most informative for fault diagnosis.

In the calculation of permutation entropy, the choice of parameters m and t affects the computation results. If the value of m is too small, there will be fewer probability patterns in the reconstructed sequences. Conversely, if m is too large, it will increase computation time and decrease efficiency. Typically, the value of m falls between 3 and 7. In this study, m is chosen based on validation. The value of t is usually set to 1 under normal circumstances.

2.2. SSA and Its Improvements

SSA is a model-free spectral estimation method that decomposes a given signal into interpretable components, including slowly varying trends, oscillatory components, and unstructured noise. It is considered a signal processing method based on time series analysis and multivariate statistical principles. SSA mainly consists of four steps: embedding, decomposition of the trajectory matrix, grouping, and signal reconstruction.

(1) Embedding: For a time series $X=[X_1, X_2, ..., X_N]$ with length N, an appropriate window length L is chosen. Typically, the selection of L falls within the range (1，N/2) based on empirical evidence. Let $K=N-L+1$. Then, the time series is lagged and arranged to form a trajectory matrix :

$$X = \begin{bmatrix} X_1 & X_2 & \cdots & X_K \\ X_2 & X_3 & \cdots & X_{K+1} \\ \vdots & \vdots & & \vdots \\ X_L & X_{L+1} & \cdots & X_N \end{bmatrix} \tag{4}$$

(2) Decomposition: Perform Singular Value Decomposition on the trajectory matrix. The formula for singular value decomposition is as follows:

$$A_{m*n} = U_{m*m} \sum V_{n*n}^T \tag{5}$$

However, directly decomposing the trajectory matrix can be challenging. Therefore, constructing a covariance matrix is done first:

$$S = XX^T \tag{6}$$

The trajectory matrix can be represented as follows after decomposing S to obtain eigenvalues $\lambda_1 > \lambda_2 > ......\lambda_L \geq 0$ and their corresponding eigenvectors $U = [U_1, U_2, ..., U_L]$:

$$X = \sum_{i=1}^{L} \sqrt{\lambda_i} U_i V_i^T \tag{7}$$

(3) Grouping: Divide all components into m mutually exclusive groups $I_1, I_2, ..., I_m$. Then, the composite matrix is given by:

$$X = X_{I_1} + X_{I_2} + \cdots + X_{I_m} \tag{8}$$

Where X represents the reconstructed time series.

(4) Reconstruction: Also known as diagonal averaging, each matrix X_{Ij} in Equation (9) is transformed into a new sequence of length N, which represents the decomposed sequence. The reconstruction formula is as follows:

$$y_i = \begin{cases} \dfrac{1}{i}\sum_{m=1}^{i} X_{m,i-m+1} & 1 \leq i < m \\ \dfrac{1}{L}\sum_{m=1}^{L} X_{m,i-m+1} & m \leq i < K \\ \dfrac{1}{N-i+1}\sum_{m=i-K+1}^{N-K+1} X_{m,i-m+1} & K \leq i < N \end{cases} \tag{9}$$

The sum of the time series y_i equals the original time series.

SSA decomposition may lead to the dispersion of signals with the same period into different subsequences, resulting in information loss and reducing the significance of the extracted fault features, thereby affecting the results of fault diagnosis. The application of w-correlation graphs allows for more accurate identification of the correlation between different frequency components present in rolling bearings, aiding in the selection of subsequences with the same period. This reduces information loss, making the final extracted components more complete and accurate, thereby enhancing the accuracy and reliability of fault diagnosis. As noise in signals appears random, it exhibits low correlation in w-correlation graphs, while signals with periodicity shows a strong correlation. Merging signals with the same period helps reduce the impact of noise on the final results, improving the separation of signals from noise. By merging signals with the same period, these fault features become more significant in the final results and are easier to detect.

(5) Threshold

The threshold is determined based on the calculation results of the mean square error. The formula for calculating the threshold is as follows:

$$S = \sqrt{\frac{\sum_{i=1}^{N}\left(X_i - \overline{X}\right)^2}{N}} \tag{10}$$

2.3. Dynamic Adjustment Temporal Convolutional Neural Network

(1) Dynamic Adjustment Mechanism

Traditional TCNs show certain limitations when dealing with time-varying signals, especially those with rich frequency characteristics. Firstly, in TCN models, the sizes of convolution kernels are predetermined and fixed, implying that the receptive field (i.e., the input data range covered by the convolution kernel) of each convolution layer remains constant. The size of the receptive field directly affects the network's ability to capture local features of the input data. When dealing with time-varying signals containing multiple frequency components, fixed-size convolution kernels may not be sufficient to simultaneously capture short-term and long-term dependencies. For instance, smaller convolution kernels are more suitable for capturing high-frequency details, while larger convolution kernels are better at capturing low-frequency long-term dependencies. However, fixed-size convolution kernels struggle to adapt between these two aspects.

Moreover, time-varying signals such as vibration data from rotating machinery often contain various frequency components ranging from low to high frequencies. These different frequency components reflect the motion states of different components of the machinery and possible fault patterns. When TCNs with fixed convolution kernel sizes handle such signals, some frequency components' features may be overlooked due to mismatches between the convolution kernel size and the signal's frequency characteristics. This oversight could impact the accuracy of fault diagnosis and the generalization capability of the model.

The dynamic convolution kernel adjustment mechanism is based on the spectral analysis of the signal, utilizing the energy

distribution in the frequency domain to guide the selection of the convolution kernel size. Let $X(t)$ denote the input time-domain signal, which is transformed into the frequency domain representation $X(f)$ through Fast Fourier Transform (FFT). The energy density function $E(f)$ of the signal can be expressed as:

Where $X(f)$ represents the amplitude of the signal at frequency f. Based on the energy density function, we define a dynamic adjustment factor , which is computed as:

$$\alpha = \frac{\int_{f_{low}}^{f_{high}} E(f)df}{\int_{0}^{f_{max}} E(f)df} \qquad (11)$$

Here, f_{low} and f_{high} are the lower and upper bounds of the considered frequency range, respectively, and f_{max} is the maximum frequency of the signal. The dynamic adjustment factor α reflects the proportion of energy of the signal within a specific frequency range and is used to adjust the size of the convolution kernel.

Based on the dynamic adjustment factor α the adjustment formula for the convolution kernel size k is as follows:

$$k = k_{\min} + \left(k_{\max} - k_{\min}\right)\Box\left(1-\alpha\right) \qquad (12)$$

Where k_{min} and k_{max} are the minimum and maximum values of the convolution kernel size, respectively. This formula ensures that when the main energy of the signal is concentrated in high frequencies, smaller convolution kernels are used to capture detailed features at high frequencies. Conversely, when the energy is mainly concentrated in low frequencies, larger convolution kernels are employed to cover a wider time range, enhancing the model's perception of low-frequency changes.

The dynamic convolution kernel adjustment mechanism proposed in this paper provides a new adaptive strategy for deep learning models in handling non-stationary time series data. By finely adjusting the convolution kernel size, this mechanism not only enhances the adaptability of the model to different frequency characteristics of the signal but also optimizes the accuracy of feature extraction and the responsiveness of the model. This theoretical innovation effectively addresses the shortcomings of traditional fixed convolution kernel models in handling complex industrial signals, demonstrating significant academic value and promising application prospects.

(2) Convolution Module

The convolution module of DA-TCN mainly includes causal convolution and dilated convolution. Causal convolution enables the network to effectively capture temporal features, while dilated convolution inserts zero values between convolution kernel elements, creating a distance between elements and expanding the receptive field of the convolution kernel. The residual module consists of two one-dimensional convolution units and one non-linear mapping unit. Each one-dimensional convolution unit includes one-dimensional causal dilated convolution, weight normalization, ReLU activation function, and Dropout layer, as shown in Figure2.
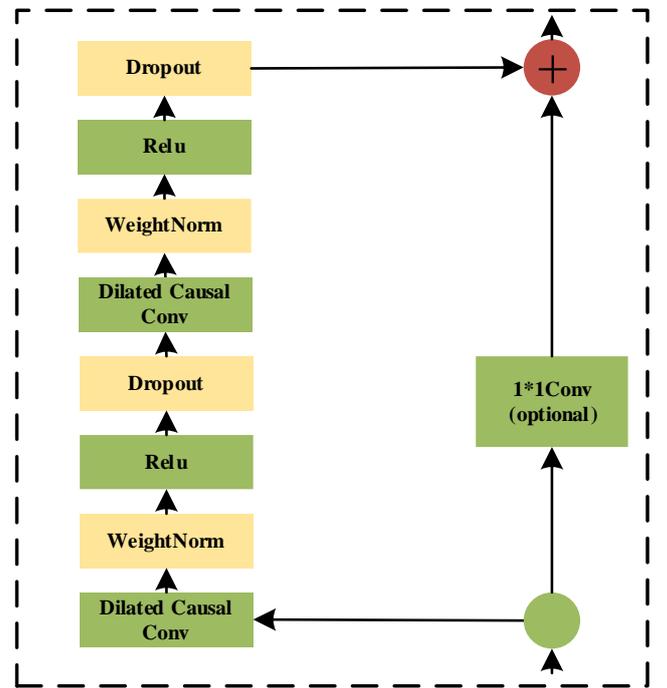

Figure 2. Structure of residual module.

The computational equations of TCN can be described as follows:

$$y = f\left(x;\theta\right) = d\left(g\left(x;\theta_g\right);\theta_d\right) \qquad (13)$$

Where x represents the input time series data, y represents the corresponding output. θ denotes the network parameters, f represents the overall mapping function of TCN, g represents the mapping function of the convolution part of TCN, and d represents the output mapping function of TCN
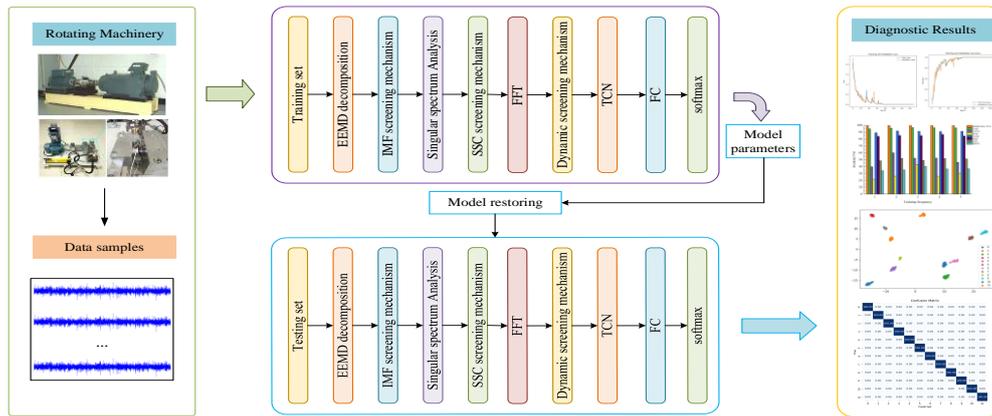
2.3. Diagnosis Flowchart Based on ASDN

.

Figure 3. The fault diagnosis process based on ASDN proposed in this paper.

(1) Signal Preprocessing Stage:

　　a) EEMD Decomposition and Subcomponent Selection: Perform ensemble empirical mode decomposition (EEMD) on the vibration signal to extract intrinsic mode functions (IMFs) effectively separating multi-frequency components. Utilize statistical indicators such as variance contribution rate, correlation coefficient, and permutation entropy to finely screen and reconstruct IMFs obtained from EEMD decomposition, selecting components with the most informative value for fault characteristics.

　　b) SSA Decomposition and Merging: Decompose the signal processed by EEMD using singular spectrum analysis (SSA), utilize the w-correlation graph to merge signals of the same period, and select the components with the highest information value based on variance contribution rate.

(2) Deep Learning Stage:

　　a) Divide the merged signal into training and testing datasets.

　　b) Dynamic Adjusted Temporal Convolutional Network (DA-TCN): Introduce a dynamic convolution kernel adjustment mechanism, allowing the convolution kernel size to adaptively adjust according to the frequency characteristics of the signal, thereby optimizing the accuracy of feature extraction and the responsiveness of the model. The model comprises adaptive convolutional layers, activation layers, normalization layers, and Dropout layers to enhance the model's generalization capability and stability.

(3) Model Training and Testing:

　　a) Model Training: Train the DA-TCN model using the training dataset, aiming to improve the model's fault diagnosis capability by optimizing the loss function.

　　b) Model Testing: After completing the model training, the model is evaluated using an independent test dataset. The test dataset contains new samples that were not involved in the training and is used to verify the generalization ability of the model.

3. Experimental Analysis of the Western Reserve University Rolling Bearing Data

(1) Data Introduction

　　This study utilizes data from the Case Western Reserve University (CWRU) Bearing Data Center for validation. The experimental setup includes a 2-horsepower motor, torque sensor, dynamometer, and accelerometer, with the latter installed at the fan end and drive end of the motor (sampled at 12 KHz or 48 KHz). The data simulates bearing conditions, including normal states, faults in the rolling elements, inner race, and outer race, with faults of three different diameters (0.007, 0.014, 0.021 inches). Focusing on the data sampled at 12 KHz, this study selects samples with faults in the inner race, outer race, and rolling elements, with a fault diameter of 0.007 inches, spanning a load range from 0Hp to 2Hp, covering 12 different bearing fault states labeled from 0 to 11. The types of faults and their corresponding category labels are listed in Table 1. The dataset is divided into 70% for training and 30% for testing to evaluate the performance of the proposed method in bearing fault diagnosis. This study aims to explore the effectiveness of the proposed approach in bearing fault detection, while examining the robustness of the model under strong noise conditions and its generalization performance on different datasets.

Table 1. Description of Each Fault Type in the CWRU Dataset.

Load	Annotation	Fault Type	Dataset Size
0（1797r/min）	0	Normal	97
	1	Inner Ring	105
	2	Rolling Element	118
	3	Outer Ring	130
1（1772r/min）	4	Normal	98
	5	Inner Ring	106
	6	Rolling Element	119
	7	Outer Ring	131
2（1750r/min）	8	Outer Ring	99
	9	Inner Ring	107
	10	Rolling Element	120
	11	Outer Ring	132

Traditional time-domain analysis methods struggle to accurately characterize the types of faults in rolling bearings, thus compromising the qualitative analysis. Therefore, this paper utilizes the advantage that EEMD can capture the transient changes of the signal, and also combines the advantage that SSA can extract the main trend and period components of the extracted signal to realize the comprehensive extraction of vibration signal information. Furthermore, SSA is utilized to further suppress noise over-decomposition by EEMD, reducing its impact and enhancing result accuracy. The time-domain plot post EEMD-SSA denoising is illustrated in Figure 4 (using label 1 as an example).
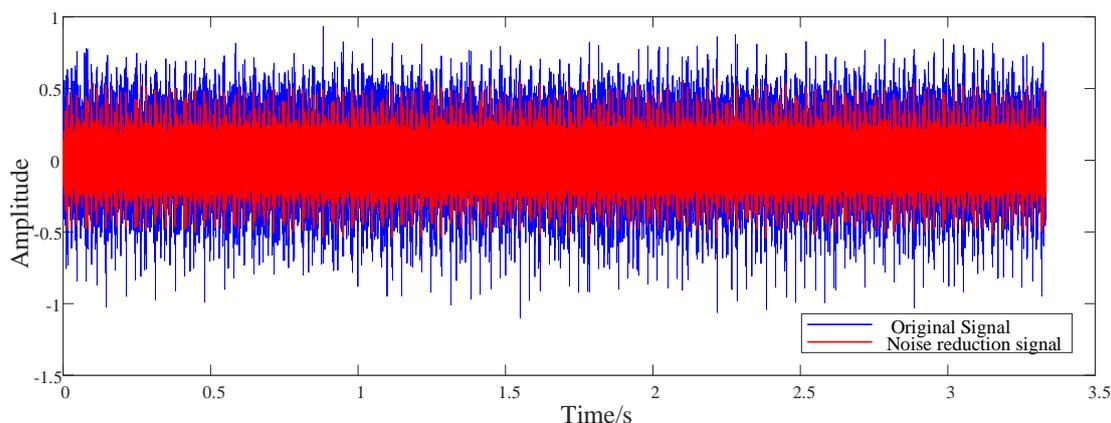

Figure 4. The time-domain waveforms.

From Figure 4, it can be observed that EEMD-SSA effectively eliminates noise while preserving the characteristics of the original signal.

The dataset denoised by EEMD-SSA was trained and tested on DA-TCN. The loss curve and accuracy curve after 150 iterations are shown in Figure 5.
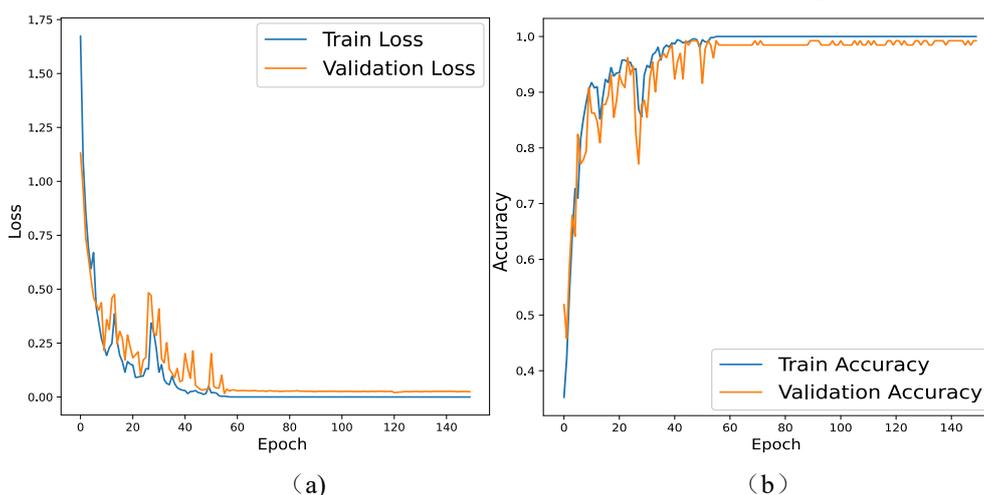

（a）　　　　　　　　　　（b）

Figure 5. Model Fault Recognition Curves: (a) Fault Recognition Loss Curve (b) Fault Recognition Accuracy Variation Curve.

The results depicted in the graphs show that after 150 iterations, the accuracy of the training set reached 100%, while that of the test set reached 99.66%. The loss value of the training set approached zero, and that of the test set decreased to 0.0033.

With the increase in the number of iterations, both the loss value curve and the accuracy variation curve tended to stabilize, indicating that the model reached a converged state, thus validating the effectiveness of the proposed method.

Visual comparison between the diagnostic results and the actual fault types is facilitated through a confusion matrix visualization (Figure 6), where the horizontal axis represents the diagnostic status and the vertical axis represents the actual fault status.
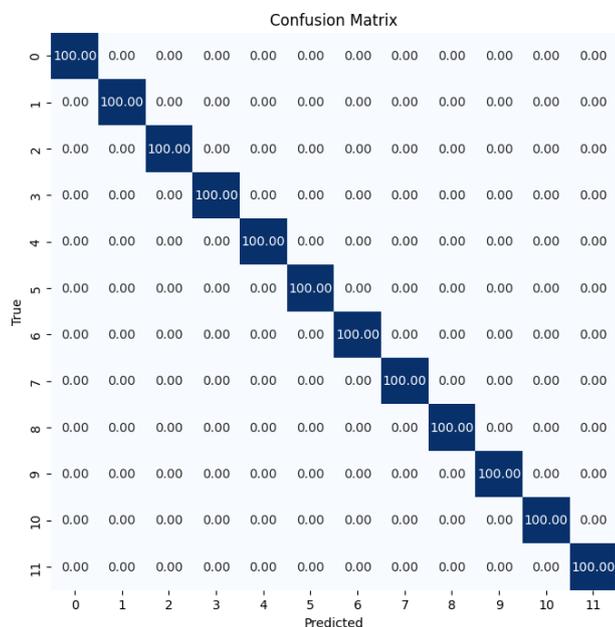

Figure 6. Confusion Matrix.

From Figure 6, it can be observed that the proposed method achieved 100% diagnostic accuracy.

(2) Comparative experiments

In order to evaluate the performance of the proposed method, this study primarily compares it with traditional MSCNN,TCN,

BiTCN, AlexNet, and two advanced methods Attention-TCN[14], CA-MCNN[15] in terms of diagnostic rate. To ensure fairness, the aforementioned methods are all trained using the same settings. That is, the size of the intercepted frame of data used is 500, the motion step of the intercepted frame is also 500, the learning rate at the beginning is 0.001, the batch size is 32, and the optimizer is Adam. The main evaluation criterion is diagnostic accuracy, which represents the ratio of correctly diagnosed samples to the total number of samples, reflecting the overall performance of the evaluation method. To reduce the impact of randomness, each experiment is repeated five times. The specific results are shown in Figure 7 and Table 2.
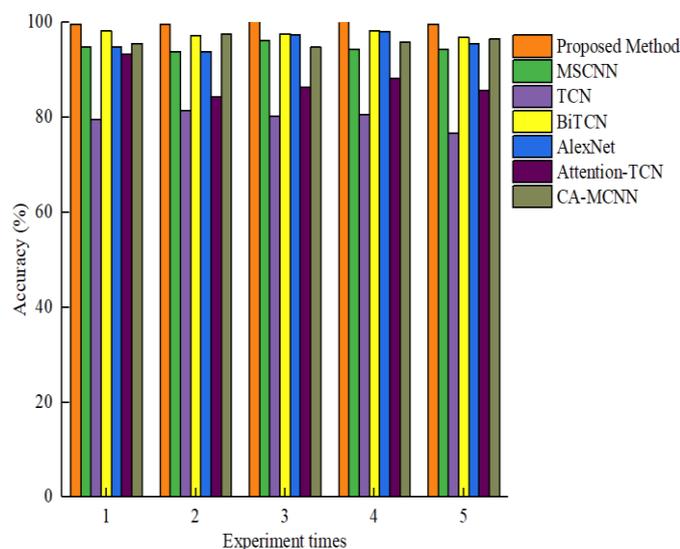

Figure 7. The accuracy of each method.

Table 2. The experimental results for the eight models.

Model	Highest Accuracy(%)	Lowest Accuracy(%)	Average Accuracy(%)
Proposed Method	**100**	**99.62**	**99.78$\pm$0.18**
MSCNN	96.18	93.92	94.76$\pm$0.77
TCN	81.53	76.74	79.72$\pm$1.62
BiTCN	98.26	96.88	97.64$\pm$0.55
AlexNet	98.10	93.80	95.94$\pm$1.59
Attention-TCN	93.40	84.38	87.64$\pm$3.13
CA-MCNN	97.57	94.79	96.04$\pm$0.95

The experimental results demonstrate that the method proposed in this paper exhibits outstanding performance in terms of diagnostic accuracy, reaching as high as 100%, with the lowest accuracy still achieving 99.62%. This performance even surpasses the highest diagnostic accuracy of other models. Compared to traditional MSCNN, TCN, BiTCN, AlexNet, and advanced Attention-TCN, CA-MCNN, the proposed method

achieves a significant improvement in diagnostic accuracy. Specifically, relative to these traditional methods, the diagnostic accuracy of the proposed method is increased by 5.02%, 20.06%, 2.14%, 3.84%, 12.14%, and 3.74%, respectively. Furthermore, this method demonstrates the best stability, showing the smallest mean square error. These results highlight the reliability and effectiveness of the proposed method in the

field of fault diagnosis.

In terms of visual representation, the T-SNE algorithm is introduced to display the feature distribution of the eight models on the dataset. Different colors represent different health statuses of the devices. The results are shown in Figure 8
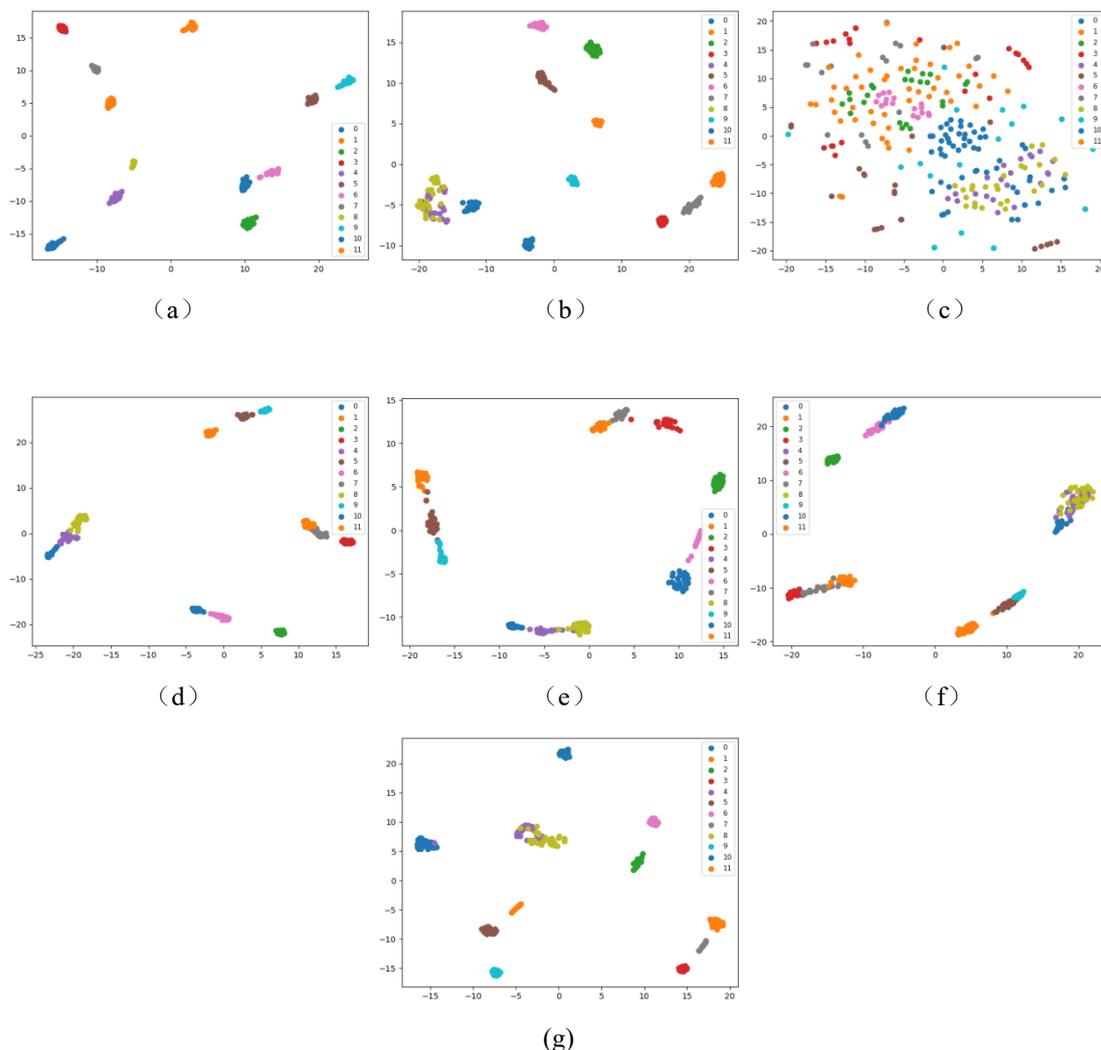

Figure 8. Feature visualization using t-SNE (a) This paper (b) MSCNN (c) TCN (d) BiTCN (e) AlexNet (f)Attention-TCN (g) CA-MCNN.

The results indicate that, apart from the method proposed in this paper, the feature distribution plots of the other seven methods are mixed, especially the TCN model, which is difficult to distinguish. In contrast, the feature distribution plot obtained by the proposed method exhibits greater discriminative capability, with features corresponding to the same health condition clustering together, while those from different health conditions are completely separated. This validates the advantage of the proposed method in feature extraction, effectively distinguishing various health states and thereby enhancing the accuracy and reliability of rolling bearing fault diagnosis.

(3) Adding Signal-to-Noise Ratio

In order to analyze the performance of the proposed method in strong noise environments, Gaussian white noise with different Signal-to-Noise Ratios (SNRs) was added to the original vibration signals to simulate real signals in industrial environments. Comparative evaluation with other learning models includes assessment criteria such as diagnostic accuracy and F-score, where the F-score reflects the performance of the evaluation method in each category. The final results are shown in Figure 9 and Table 3.
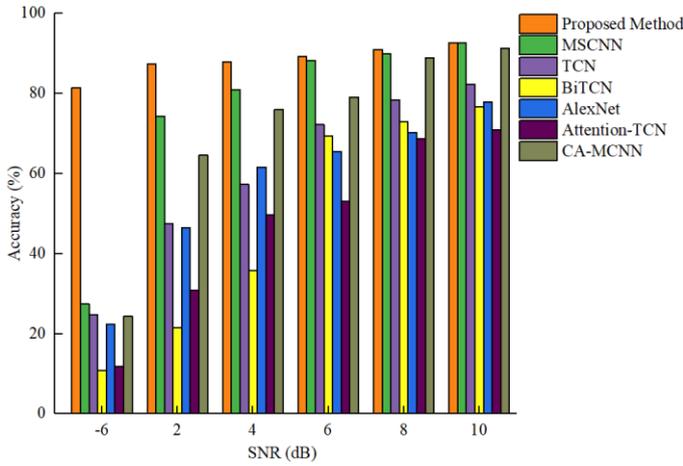
Figure 9. Diagnostic accuracy of eight models at various
SNRs.

Table 2. F-scores (%) of eight models at various SNRs.

Model	-6dB	2dB	4dB	6dB	8dB	10dB
Proposed Method	**79.33**	**86.99**	**87.99**	**89.31**	**89.94**	**92.72**
MSCNN	21.43	71.79	79.52	87.81	89.85	92.34
TCN	18.04	39.46	53.55	70.33	78.12	81.69
BiTCN	2.46	13.03	29.14	64.76	68.81	73.92
AlexNet	12.40	40.67	57.23	62.70	68.53	77.32
Attention-TCN	3.51	24.99	43.16	47.60	68.81	69.25
CA-MCNN	15.37	56.28	70.92	76.39	88.22	91.34

By examining Figure 9, it is evident that the diagnostic accuracy of the proposed method surpasses significantly that of the other seven methods. Additionally, from Table 3, it can be observed that the F-scores of the proposed method consistently outperform those of the other methods across all SNR conditions, demonstrating its superior performance in high-noise environments. This outcome underscores the high accuracy and robustness of the proposed method in strong noise environments.

(4) Ablation Study

In this study, ablation experiments were conducted to delve into the impact of each component of the proposed method on network performance. The study compared four network structures: a) without dynamic adjustment mechanism; b) without undergoing EEMD-SSA processing, directly inputting to DA-TCN; c) only processed by EEMD and then input into DA-TCN; d) replace DA-TCN with 1DCNN (its detailed model structure is shown in Table 4), and the original signal is directly input into 1DCNN; e) only processed by EEMD and then input into 1DCNN; f) The signal processed by EEMD-SSA is input

into 1DCNN; g) the complete network structure proposed in this paper. The fault categories to be diagnosed include the 12 types mentioned in the experiment (labeled 0-11). Table 5 shows the diagnostic accuracy, F-scores, and loss rates of the final results.

Table 4. 1DCNN model structure.

No.	Scale1
1	Conv1(3×1×32)
2	ReLU、BN
3	Max pooling(2×1,2)
4	Conv2(5×1×64)
5	ReLU、BN
6	Max pooling(2×1,2)
7	Conv3(7×1×128)
8	ReLU、BN
9	Max pooling(2×1,2)
10	FC
11	Softmax

Table 5. Different Models' Results.

Model	Average Diagnostic Accuracy/%	Average F-score /%	Average Loss Rate
a	96.11±2.83	96.01±2.98	0.1060
b	97.53±0.45	97.53±0.47	0.1156
c	99.08±0.17	99.08±0.17	0.0707
d	76.94±3.71	76.58±3.89	1.2024
e	81.94±1.42	82.06±1.27	0.6118
f	94.66±1.06	94.66±1.04	0.3575
g	**99.78±0.18**	**99.78±0.18**	**0.0099**

Table 5 shows that removing either structure in the ASDN model results in a decrease in the model's performance. When there is no EEMD-SSA structure, the accuracy and F-score both decrease by 2.25%, proving the importance of EEMD-SSA in in noise suppression and signal purity enhancement. When replacing DA-TCN with 1DCNN, the accuracy decreased by 3.29% and the F-score decreased by 2.87%; and when inputting the original signals into the two models, the results of 1DCNN decreased the accuracy by 22.84% and the F-score decreased by 23.20% with respect to DA-TCN, which proved the superiority of DA-TCN in extracting features. When the signal is input into DA-TCN and 1DCNN after only EEMD processing, the accuracy and F score drop by 0.70% and 12.72% respectively compared with DA-TCN and 1DCNN after EEMD-SSA processing. Compared with DA-TCN, the accuracy and F score of 1DCNN drop by 17.14%, proving the importance of SSA in signal reconstruction and feature enhancement. According to the error, it can be seen that the stability of the model also decreases after removing some structures, which further verifies

the robustness and stability of the method proposed in this paper.

4. Analysis of Rolling Bearing Data from Xi'an Jiaotong University

(1) Data Introduction

This study utilizes the XJTU-SY rolling bearing dataset to validate the proposed methodology. The experimental setup consists of an alternating current motor, a speed controller, a supporting shaft, bearings, among others. Vibration signals were collected using an LDK UER204 bearing and a PCB 352C33 accelerometer at a sampling frequency of 25.6 kHz. The dataset is divided into 70% for training and 30% for testing, covering five bearings under three operating conditions as shown in Table 6.

Table 6. Overview of the XJTU-SY Bearing Dataset Information.

Operating Condition	Annotation	Dataset	Fault Type
1	0	Bearing 1_1	Outer Race
	1	Bearing 1_4	Cage
2	2	Bearing 2_1	Inner Race
	3	Bearing 2_2	Outer Race
	4	Bearing 2_2	Cage
3	5	Bearing 3_1	Outer Race
	6	Bearing 3_3	Inner Race

The time-domain waveform plot of the dataset after denoising with EEMD-SSA is shown in Figure 10 (taking label 0 as an example).
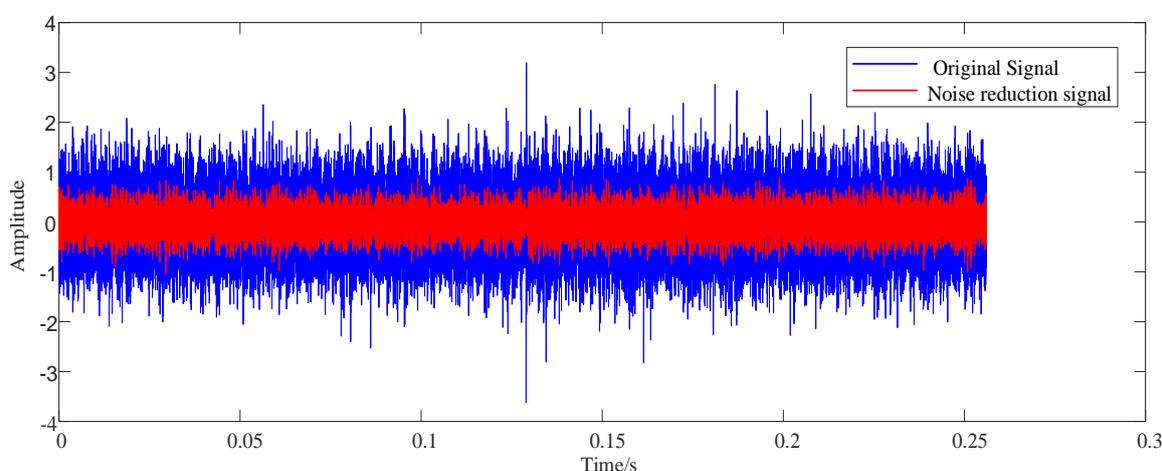

Figure 10. Time-domain waveform.

By Figure10, it can be observed that EEMD-SSA effectively eliminates noise while preserving the characteristics of the original signal.

(2) Comparative Experiment

Compare the proposed method with traditional MSCNN,TCN, BiTCN, AlexNet, and two advanced methods Attention-TCN, CA-MCNN in terms of diagnostic rate. Each method is run five times to ensure the reliability of the experiment. Table 7 and Figure 11 summarize the results of all models.

Table 7. Experimental Results of Eight Models.

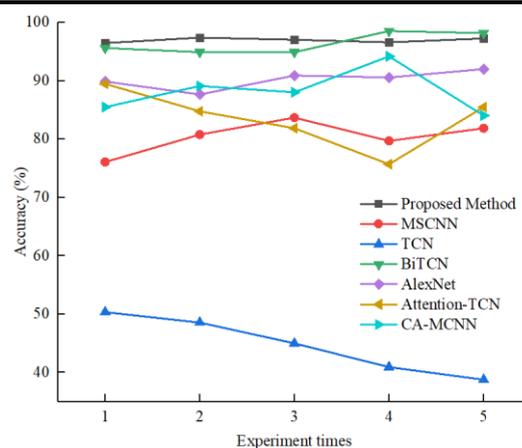| Model	Highest Accuracy / %	Lowest Accuracy / %	Average Accuracy /%	Average F-score / %
Proposed Method	**97.42**	**96.49**	**96.97±0.36**	**96.98±0.36**
MSCNN	83.70	76.09	80.44±2.54	80.33±2.65
TCN	50.35	38.77	44.73±4.39	43.62±5.37
BiTCN	98.55	94.93	96.45±1.59	96.44±1.60
AlexNet	92.03	87.68	90.23±1.45	89.92±1.63
Attention-TCN	89.49	75.72	83.48±4.58	83.25±4.98
CA-MCNN	94.20	84.06	88.19±3.50	87.81±3.80

Figure 11. Accuracy of Each Method.

The experimental results indicate that the proposed method outperforms other models in terms of average diagnostic accuracy Additionally, its average F-score also surpasses that of

other models. Theis finding underscore the efficiency and superior performance of the model in fault diagnosis. Figure11 illustrates the excellent performance stability demonstrated by our method.

(3) Ablation Study

Conducting ablation studies to delve deeper into the impact of each structure of the proposed method on network performance. The study includes comparisons of four network structures: a) without dynamic adjustment mechanism; b) without undergoing EEMD-SSA processing, directly inputting to DA-TCN; c) only processed by EEMD and then input into DA-TCN; d) replace DA-TCN with 1DCNN, and the original signal is directly input into 1DCNN; e) only processed by EEMD and then input into 1DCNN; f) The signal processed by EEMD-SSA is input into 1DCNN; g) the complete network structure proposed in this paper. This experiment examines seven different fault types (labeled as 0-6). Table 8 summarizes the performance of these structures in terms of diagnostic accuracy, F-score, and loss rate.

Table 8. Results of Different Models.

Model	Average Diagnostic Accuracy/%	Average F-score / %	Average Loss Rate
a	95.56±0.81	95.58±0.59	0.1406
b	92.85±2.14	92.82±2.15	0.6583
c	95.26±0.77	95.24±0.78	0.2367
d	85.15±1.39	85.03±1.36	1.4554
e	86.74±0.98	86.20±1.71	1.2062
f	91.45±1.98	91.32±2.05	0.7332
g	**96.97±0.36**	**96.98±0.36**	**0.2157**

According to the results in Table 8, it is proved again that removing any structure in the ASDN model will lead to a decrease in the performance of the model. When there is no EEMD-SSA structure, the accuracy is 4.12% and the F score decreases by 4.16%. When replacing DA-TCN with 1DCNN, the accuracy decreased by 5.43% and the F-score decreased by 5.66%; and when inputting the original signal into both models, the 1DCNN results decreased by 7.70% in accuracy and 7.79% in F-score relative to DA-TCN. When the signal is input into DA-TCN and 1DCNN after only EEMD processing, the accuracy decreases by 1.71% and 4.71% respectively, and the F score decreases by 1.74% and 5.12% respectively compared with DA-TCN and 1DCNN after EEMD-SSA processing. According to the error, it can be seen that the stability of the model also decreases after removing some structures, which

further verifies the robustness and stability of the method proposed in this paper.

5. Conclusion

To address the issues of fixed convolution kernel sizes and weak specificity in feature extraction in the application of deep learning for fault diagnosis, this paper develops a fault diagnosis framework that combines Ensemble Empirical Mode Decomposition (EEMD), Singular Spectrum Analysis (SSA), and a Dynamically Adjustable Temporal Convolutional Network (DA-TCN). In the adaptive preprocessing stage, adaptive improvements are made to the EEMD and SSA algorithms, and the signal is decomposed on multiple scales and fault components are extracted to optimize signal representation. In the adaptive deep learning stage, this paper proposes a dynamic convolution kernel adjustment mechanism, enabling the neural network to dynamically adjust the convolution kernel size according to the different frequency components in the signal, thus accurately processing signals with different frequencies. The proposed method is validated using publicly available datasets from Case Western Reserve University and Xi'an Jiaotong University, yielding the following conclusions:

(1) Adaptive Preprocessing Stage: Through adaptive improvements to EEMD and SSA, the limitations of EEMD in extracting signal trends and periodic components are successfully addressed. Time-domain waveform diagrams show that the signal processed with EEMD-SSA noise reduction effectively removes noise components and highlights fault characteristics. Ablation study results indicate that after adaptive EEMD-SSA processing, the fault diagnosis accuracy on the two datasets increased by 2.25% and 4.12%.

(2) Dynamic Convolution Kernel Adjustment Mechanism: This paper innovatively develops a dynamic convolution kernel adjustment mechanism, addressing the lack of sensitivity and responsiveness of TCN to signal characteristic changes under different operating conditions. Ablation study experimental results show that adding this mechanism increased fault diagnosis accuracy on the two datasets by 3.67% and 1.41%.

(3) Model Comparison: The proposed fault diagnosis framework is compared with existing methods including MSCNN, TCN, BiTCN, and advanced methods such as AlexNet, Attention-TCN and CA-MCNN. Results show that the

diagnostic accuracy on the two public datasets reached 100% and 97.42%, respectively, surpassing the accuracy of other models. Additionally, TSNE visualization of the classification results indicates that the proposed method exhibits the most distinct classification effect, clearly categorizing 12 fault types.

(4) Performance Validation in Noisy Environments: To verify the model's performance in noisy environments, Gaussian white noise with different Signal-to-Noise Ratios (SNR) (ranging from 2dB to 10dB) was added to the first dataset to simulate real industrial signals. Results show that, compared to the benchmark models, the proposed method achieved the highest diagnostic accuracy and F-score under various noise conditions.

The above validations demonstrate that the proposed fault diagnosis framework excels in handling complex vibration signals and in real industrial applications, significantly improving the accuracy and reliability of fault diagnosis.

References

1. Deng Y, Du S, Wang D, et al. A calibration-based hybrid transfer learning framework for RUL prediction of rolling bearing across different machines[J]. IEEE Transactions on Instrumentation and Measurement, 2023, 72: 1-15. https://doi.org/10.1109/TIM.2023.3260283

2. Huang J, Cui L, Zhang J. Novel morphological scale difference filter with application in localization diagnosis of outer raceway defect in rolling bearings[J]. Mechanism and Machine Theory, 2023, 184: 105288.

3. Li J, Cheng X, Zhang S, et al. Fault feature extraction method of rolling bearings based on coupled resonance system with vibrational resonance-assisted enhanced stochastic resonance[J]. Mechanical Systems and Signal Processing, 2024, 208: 111069.

4. Wei Y, Li Y, Wang X. Incipient bearing fault detection using adaptive fast iterative filtering decomposition and modified Laplace of Gaussian filter[J]. Structural Health Monitoring, 2024: 14759217241246985.

5. Li D, Jiang M R, Li M W, et al. A floating offshore platform motion forecasting approach based on EEMD hybrid ConvLSTM and chaotic quantum ALO[J]. Applied Soft Computing, 2023, 144: 110487.

6. Lu Y, Sheng B, Fu G, et al. Prophet-EEMD-LSTM based method for predicting energy consumption in the paint workshop[J]. Applied Soft Computing, 2023, 143: 110447.

7. Zhang H, Feng L, Wang J, et al. Development of technology predicting based on EEMD-GRU: An empirical study of aircraft assembly technology[J]. Expert Systems with Applications, 2024, 246: 123208.

8. Dao F, Zeng Y, Qian J. A novel denoising method of the hydro-turbine runner for fault signal based on WT-EEMD[J]. Measurement, 2023, 219: 113306.

9. Li F, Wan Z, Koch T, et al. Improving the accuracy of multi-step prediction of building energy consumption based on EEMD-PSO-Informer and long-time series[J]. Computers and Electrical Engineering, 2023, 110: 108845.

10. Shi P, Gao S, Tan H, et al. MGGSED-SSA: An improved sparse deconvolution method for rolling element bearing diagnosis[J]. Applied Acoustics, 2024, 220: 109960.

11. Guo J, Li D, Du B. A stacked ensemble method based on TCN and convolutional bi-directional GRU with multiple time windows for remaining useful life estimation[J]. Applied Soft Computing, 2024, 150: 111071.

12. Ding L, Li Q. Fault diagnosis of rotating machinery using novel self-attention mechanism TCN with soft thresholding method[J]. Measurement Science and Technology, 2024, 35(4): 047001. https://doi.org/10.1088/1361-6501/ad1eb3

13. Zhang C, Zhang L. Wind turbine pitch bearing fault detection with Bayesian augmented temporal convolutional networks[J]. Structural Health Monitoring, 2024, 23(2): 1089-1106. https://doi.org/10.1177/14759217231175886

14. Zhang H, Ge B, Han B. Real-time motor fault diagnosis based on tcn and attention[J]. Machines, 2022, 10(4): 249. https://doi.org/10.3390/machines10040249

15. Huang Y J, Liao A H, Hu D Y, et al. Multi-scale convolutional network with channel attention mechanism for rolling bearing fault diagnosis[J]. Measurement, 2022, 203: 111935.