FAULT DETECTION AND DIAGNOSIS OF ROTATING EQUIPMENT BASED ON TEMPORAL-SPECTRAL FEATURE HYBRID STRATEGY

Lili Yi
Intelligent Manufacturing College, Anhui Xinhua University, Hefei, 230088, China

Abstract - Rotating equipment plays a core role in modern industry, and its fault detection and diagnosis are vital for production safety and efficiency. Traditional methods struggle to meet diagnostic demands under complex conditions due to insufficient fusion of time-frequency features and inadequate capture of temporal dependencies. Therefore, this paper proposes a fault detection and diagnosis model for rotating equipment based on a time-frequency feature fusion strategy to improve diagnostic accuracy. The model fully utilizes the advantages of the Temporal Convolutional Network and Bidirectional Gated Recurrent Unit to fuse and extract time-frequency signals. Experiments show that the model achieves a Pearson correlation coefficient of 0.9647 and a minimum loss value of 0.013, with faster convergence than comparison models. Meanwhile, its diagnostic accuracy reaches 94.5% and 95.6% in two factories. These results indicate that the proposed model effectively improves diagnostic performance through an adaptive hybrid of time-frequency features and bidirectional temporal dependency capture. This offers a new approach for rotating equipment fault detection and diagnosis under complex conditions and promotes intelligent and efficient development in this field.

Keywords: Temporal convolutional network; Bidirectional gated recurrent unit; Temporal-spectral feature fusion; Rotating equipment; Fault detection and diagnosis.

1. Introduction

Rotating equipment serves as a core component in industries such as power, chemical, and metallurgy, with operational stability directly affecting production safety and efficiency [1]. Traditional fault diagnosis methods rely on human expertise. However, with the development of deep learning technology, deep learning models can automatically learn features from large-scale data, bypassing the need for complex manual feature engineering, which effectively improves the efficiency and accuracy of fault diagnosis [2-3]. Among these, the Temporal Convolutional Network (TCN) captures temporal evolution and frequency feature distribution simultaneously through multi-scale dilated convolution and causal structure. It achieves effective correlation and integration of cross-domain features, enhancing the integrity and discriminability of fused time-frequency features [4]. The Bidirectional Gated Recurrent Unit (BiGRU) captures forward and backward contextual dependencies in temporal signals through a bidirectional recurrent structure and gating mechanisms. It dynamically selects key temporal features, strengthening the modeling of temporal dependencies in fault evolution [5]. Therefore, this paper proposes a time-frequency feature hybrid algorithm based on TCN, combined with BiGRU, to build a fault detection and diagnosis model for rotating equipment. The model aims to efficiently extract fault features from vibration signals and improve accuracy and generalization under complex conditions. The innovation lies in designing a TCN-based time-frequency feature hybrid algorithm that adaptively fuses dynamic temporal changes and spectral feature distributions. On this basis, an improved BiGRU network captures temporal dependencies of equipment faults, achieving efficient extraction and accurate classification of fault features, which enhances the performance of the diagnosis model.

2. Literature Review

Based on deep learning, feature extraction algorithms gradually became the preferred solution for complex real-world problems due to their ability to autonomously discover data patterns. Many scholars have conducted in-depth research on this. For example, Limouni T et al. leveraged the local

https://doi.org/10.17683/ijomam/issue22.v2.18

feature extraction capabilities of TCN to achieve high-precision one-step and multi-step ultra-short-term photovoltaic power forecasting. Their model performed better than single models under different seasons and weather conditions. However, their model parameters were complex to tune, and the computational cost was high, making real-time deployment efficiency uncertain [6]. Li W et al. proposed a TCN-based dissolved oxygen prediction model for water quality detection in aquaculture. Experiments showed their model had a mean absolute error of 0.236 and a Coefficient of Determination of 0.94. However, the model's generalization ability in complex aquaculture environments remained unverified [7]. Wei X's team combined TCN's temporal feature extraction with attention mechanisms for dynamic weight allocation, achieving high-precision human activity recognition. Their recognition accuracy improved by 1.13% to 1.83% compared to traditional models. But their model lacked sufficient validation of generalization in complex scenarios and might face computational resource bottlenecks during real-time deployment [8]. Rana M R R et al. proposed a BiGRU-based deep sentiment analysis model that captured text semantics and collaborative filtering logic through bidirectional sequence modeling. Their model reached 91.2% accuracy in customer review sentiment classification, improving 4.3% over traditional methods. However, it did not incorporate attention mechanisms to analyze sentiment intensity and implicit semantics, limiting model accuracy [9]. Dave E et al. combined pretrained language models, BiGRU, and conditional random fields to extract Indonesian personal finance entities with high accuracy. Their F1 score reached 89.7%, 5.2% higher than traditional methods. But they did not adopt data augmentation strategies, resulting in a recall rate of only 68.3% for low-frequency financial entities [10].

In the field of rotating equipment fault detection and diagnosis, as industrial demands for equipment reliability increased, data-driven intelligent diagnosis technology became a research hotspot. Jin Y's team proposed a time series Transformer model that used self-attention to capture temporal features for rotating machinery fault detection and diagnosis. Its accuracy outperformed traditional models, but performance dropped under small-sample and high-noise conditions, and its generalization in complex operating environments remained unverified [11]. Wan W et al. proposed a self-supervised simple siamese framework using unlabeled samples for rotating machinery fault detection and diagnosis. Their model achieved accuracy rates of 100%, 99.38%, and 98.87% on three datasets. However, the self-supervised task lacked targeted feature learning, causing performance decline in small-sample environments [12]. Jia N et al. developed a rotating machinery fault detection and diagnosis model based on wavelet time-frequency images and an optimized stacked denoising autoencoder. Their model achieved 98.7% recognition accuracy, improving 6.2% over traditional support vector machine models. Yet, it did not use attention mechanisms to analyze feature importance, and accuracy dropped to 89.3% when labeled data fell below 10% [13]. Kim T's team used autoencoders combined with adversarial training for cross-domain rotating machinery fault detection and diagnosis. Their accuracy on the Case Western Reserve University (CWRU) dataset reached 98.2%, 4.7% higher than traditional methods. But their model's performance in small-sample scenarios still needed further validation [14]. Jiang C et al. employed a dual-branch prototypical network to address few-shot fault detection and diagnosis in rotating machinery. Their model reached 97.6% accuracy on the CWRU dataset, improving 4.2% over single-branch models. However, accuracy dropped to 88.3% under strong noise, and real-time inference efficiency was not optimized [15].

In summary, many scholars conducted extensive exploration in rotating equipment fault detection and diagnosis and achieved significant progress. However, existing research lacked depth in fusing temporal and spectral vibration signal features. Simple feature concatenation could not achieve accurate modeling. Most methods failed to fully exploit the role of forward and backward information in capturing fault evolution dependencies and lacked a targeted design regarding feature importance. Therefore, this study proposed a TCN-based time-frequency feature hybrid algorithm and combined an improved BiGRU network to build a rotating equipment fault detection and diagnosis model. It aims to provide a new technical path for efficient fault detection and diagnosis by fusing time-frequency features and accurately capturing bidirectional temporal dependencies.

3. Rotating Equipment Fault Detection and Diagnosis Model Combining TCN-based Time-frequency Feature Hybrid Algorithm and Improved BiGRU

3.1 Design of TCN-based Time-frequency Feature Hybrid Algorithm

The vibration signals of rotating equipment often suffer from strong background noise and mechanical interference under complex conditions. Fault features are usually masked by noise, which directly affects subsequent feature extraction and diagnosis accuracy [16]. Empirical Mode Decomposition (EMD) decomposes signals into multiple Intrinsic Mode Functions (IMF), showing significant advantages in handling nonlinear sequence data. It obtains the difference function of the input original fault signal, calculated as shown in Equation (1).

$$H(t) = X(t) - M(t) \quad (1)$$

In Equation (1), $H(t)$ is the difference function of the fault signal, $X(t)$ represents the original fault signal, and $M(t)$ represents the average of the original signal's maxima and minima. The output difference function of the fault signal is then used as the new original signal and recalculated repeatedly until the fault signal in IMF form is obtained, as shown in Equation (2).

$$x(t) = \sum imf + r(t) \quad (2)$$

In Equation (2), $x(t)$ is the final IMF component output, $\sum imf$ represents the IMF fault signal that meets the IMF criteria, and $r(t)$ is the residual signal separated from the fault signal by the IMF component.

Although EMD can adaptively decompose non-stationary signals into intrinsic mode functions, it still suffers from mode mixing. Fast Independent Component Analysis (FastICA) is a blind source separation method based on the statistical independence of signals. It effectively separates mixed-mode components in EMD decomposition by exploiting component independence.

Therefore, this study combines EMD and FastICA to build the signal processing algorithm EMD-FastICA. Its flowchart is shown in Figure 1.

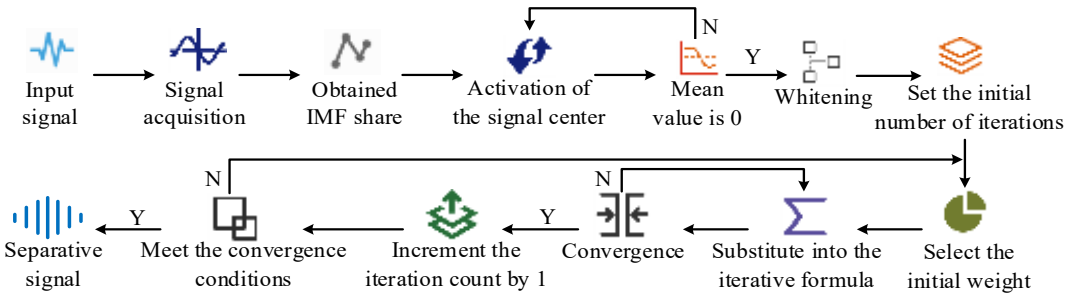


Figure 1: Signal processing algorithm EMD-FastICA operation process

As shown in Figure 1, the original fault vibration signal first enters the EMD module for decomposition. The difference function of the fault signal is calculated and used repeatedly as the new original signal until components meeting the IMF criteria are obtained. Then FastICA performs iterative optimization on the decomposition results. By solving the inverse matrix of the linear random mixing matrix, it processes the observed signals to finally separate independent source signals, achieving effective separation of mixed-mode components in vibration signals. FastICA can separate several independent source signals from multiple vibration signals, calculated as shown in Equation (3).

$$X(t) = A(t) * S(t) + V(t) \quad (3)$$

In Equation (3), $A(t)$ is the linear random mixing matrix, $S(t)$ is the unknown source signals, and $V(t)$ represents noise. Its purpose is to iteratively optimize and restore observed signals to source signals, calculated as shown in Equation (4).

$$Y(t) = W(t) * X(t) + V(t) \quad (4)$$

In Equation (4), $Y(t)$ is the observed signal, and $W(t)$ represents the inverse matrix of the linear random mixing matrix $A(t)$.

Although EMD-FastICA can separate independent source signals, it cannot effectively fuse the temporal and spectral features of vibration signals, limiting comprehensive fault signal representation. TCN captures temporal dynamics and frequency feature distribution simultaneously through multi-scale dilated convolution and causal structure, enhancing the integrity and discriminability of fused time-frequency features [17]. Dilated convolution is a key structure of TCN, expressed as shown in Equation (5).

$$F(s) = \sum_{i=0}^{k-1} f(i) X'_{s-d \cdot i} \quad (5)$$

In Equation (5), F represents the output value of dilated convolution, k represents the length of the convolution kernel, X' represents the input signal sequence, s represents the output position index, d represents the dilation rate, and i represents the kernel index. Although TCN achieves cross-domain correlation and integration of time-frequency features via multi-scale dilated convolution and causal structure, it cannot dynamically analyze the importance of different features for fault detection and diagnosis. Multi-head attention dynamically allocates weights to highlight key fault features, improving the algorithm's ability to capture effective features. The multi-head attention structure is shown in Figure 2.

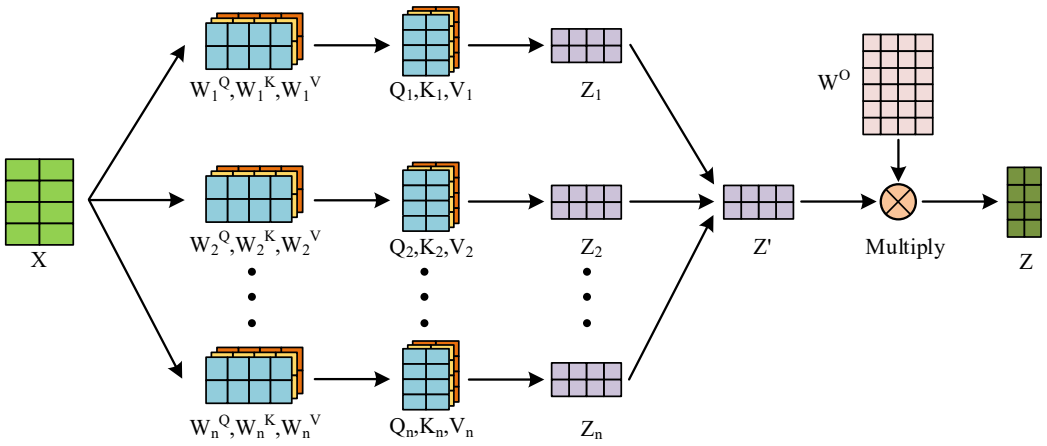


Figure 2: Multi-head attention structure diagram

In Figure 2, the input data first passes through three different linear transformation layers, generating query, key, and value matrices, respectively. These are divided into multiple heads, where the dot product between the query and key matrices is computed, followed by scaling and biasing to derive the attention weights. The outputs from all heads are concatenated and subsequently passed through a linear transformation, combining information from each head to generate the final multi-head attention output. During this process, the attention weight calculation equation is shown in Equation (6).

$$\begin{cases} V' = \hat{X}W^{V} \\ Q = \hat{X}W^{Q} \\ K = \hat{X}W^{K} \end{cases} \tag{6}$$

In Equation (6), $\hat{X}$ represents the input matrix, V', Q, and K represent the value vectors, query vectors, and key vectors, respectively. W is the weight matrix. Each head then calculates the attention weight function as shown in Equation (7).

$$A(Q,K,V') = Soft\max(\frac{QK^{T}}{\sqrt{d_{K}}})V' \tag{7}$$

In Equation (7), d_K represents the dimension of the key vector. Using the obtained attention function, each head's output is calculated, and the outputs of all heads are concatenated and transformed to generate the final output, calculated as shown in Equation (8).

$$b_i = W^{O}[b_1^{\hat{i}}; b_2^{\hat{i}}] \tag{8}$$

In Equation (8), b is the output of attention heads, W^O is the weight matrix of output transformation, $[b_1^{\hat{i}}; b_2^{\hat{i}}]$ represents the concatenation operation, and $\hat{i}$ represents the position in the input sequence. In summary, this study combines EMD-FastICA with Attention-TCN (ATCN) to build the EMD-FastICA-ATCN time-frequency feature hybrid algorithm. Its structural diagram is shown in Figure 3.

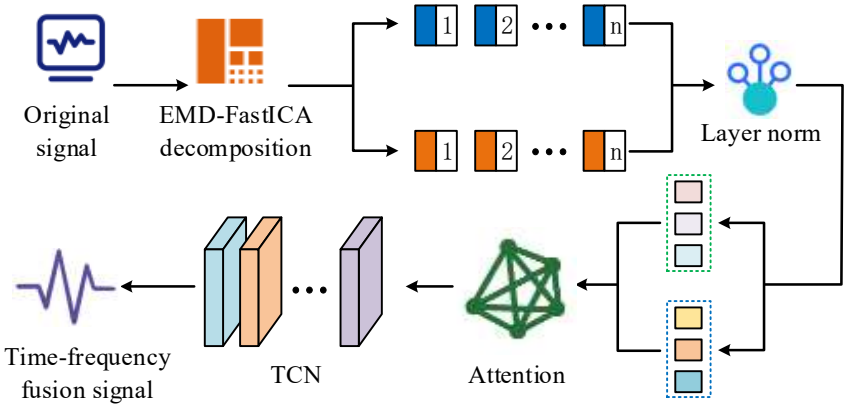


Figure 3: Flowchart of EMD-FastICA-ATCN time-frequency feature hybrid algorithm

As shown in Figure 3, during the operation of the EMD-FastICA-ATCN algorithm, the original fault vibration signal is first input into the EMD-FastICA module for effective preprocessing of rotating equipment vibration signals under complex conditions. Then, the features pass through three

different linear transformation layers to obtain query, key, and value matrices, which are split into multiple heads. The dot product between the query and key is computed, then scaled and normalized to generate attention weights, which emphasize the critical fault features. The outputs of all heads are concatenated and integrated through a linear transformation. After this, the TCN network captures temporal dynamics and frequency feature distribution simultaneously through multi-scale dilated convolution and causal structure, achieving the fusion of time-frequency features and providing high-quality feature input for subsequent fault detection and diagnosis.

3.2 Model Construction Combining EMD-FastICA-ATCN Algorithm with Improved BiGRU

The constructed EMD-FastICA-ATCN algorithm achieves cross-domain correlation and integration of temporal dynamic evolution and spectral feature distribution. However, it cannot perform fault detection and diagnosis. BiGRU, with its bidirectional recurrent structure and gating mechanisms, captures forward and backward dependencies in temporal signals and models temporal correlations in fault evolution [18].

It dynamically adjusts information retention and updates via update and reset gates. The reset gate is expressed as shown in Equation (9).

$$r_a = \sigma(W_r[h_{a-1}, x_a] + b_r) \tag{9}$$

In Equation (9), r represents the output of the reset gate, σ is the activation function, W_r represents the reset gate's weight matrix, h is the hidden state, and a is the time step. The reset gate computes the weights between the current input and the previous hidden state, as represented in Equation (10) [19].

$$z_a = \sigma(W_z[h_{a-1}, x_a] + b_z) \tag{10}$$

In Equation (10), z represents the output of the update gate, W_z represents the update gate's weight matrix, and b_z represents the update gate's bias. BiGRU extracts dynamic fault features by capturing forward and backward dependencies in temporal vibration signals. However, deep stacking can weaken feature learning due to gradient propagation issues. Residual Network (ResNet) alleviates gradient vanishing by directly passing gradients and original features via skip connections. The combined ResNet-BiGRU architecture is shown in Figure 4.

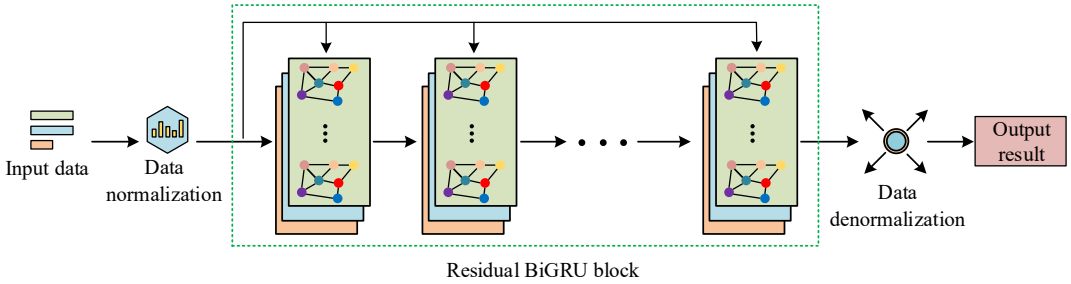


Figure 4: Network architecture diagram of ResNet-BiGRU

As shown in Figure 4, normalized input data first pass through the BiGRU layer's bidirectional gated recurrent structure to dynamically capture forward and backward dependencies and output hidden state sequences. Meanwhile, residual connections match and fuse BiGRU input features with lightweight convolutions, easing gradient vanishing during training. The update gate regulates the retention of historical hidden states and the incorporation of new information. The hidden state is expressed as shown in Equation (11).

$$h_a = (1 - z_a)h_{a-1} + z_a\tilde{h}_a \tag{11}$$

In Equation (11), h represents the hidden state, and $\tilde{h}$ represents the candidate hidden state. Then, skip connections directly transfer shallow features to deeper layers. The residual mapping equation is shown in Equation (12).

$$F'(p) = H'(p) + p \tag{12}$$

In Equation (12), F' represents the output, H' is the residual function, and p represents the input features. Despite this, ResNet-BiGRU does not sufficiently emphasize specific frequency components or amplitude jumps, affecting diagnostic accuracy. Therefore, the study introduces the Convolutional Block Attention Module (CBAM). CBAM captures channel semantic dependencies via global pooling and multi-layer perceptrons to assign weights to key features. It also focuses on spatial locations to highlight target regions and filter background interference, enhancing model generalization in complex scenarios [20]. The CBAM structure is shown in Figure 5.

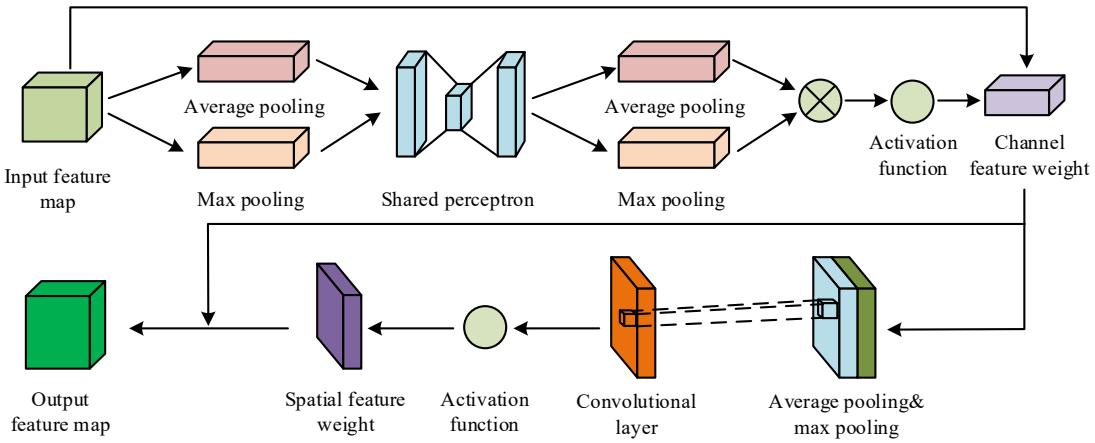


Figure 5: Schematic diagram of the CBAM structure

As illustrated in Figure 5, the input feature map undergoes both average pooling and max pooling to extract feature vectors, which are then fed into a shared multi-layer perceptron. The resulting outputs are summed and activated to produce channel attention weights. These weights are applied to the original feature map on a per-channel basis, yielding channel-enhanced features. The channel-enhanced feature map then undergoes average pooling and max pooling, followed by convolution to generate intermediate features. After activation, spatial attention weights are created and applied to the channel-enhanced features via element-wise multiplication to yield the final enhanced feature map. The channel attention calculation is shown in Equation (13).

$$M_c(\hat{F}) = \sigma(MLP(AvgPool(\hat{F})) + MLP(MaxPool(\hat{F}))) \quad (13)$$

In Equation (13), M_c is the output channel attention weights, $\hat{F}$ is the input feature map, MLP denotes the nonlinear transformation of two descriptors by the shared multi-layer perceptron. The spatial attention calculation is shown in Equation (14).

$$M_s(F') = \sigma(f^{7\times7}(AvgPool(F'); MaxPool(F'))) \quad (14)$$

In Equation (14), M_s represents the output spatial attention weights, f represents the convolution kernel, and F' represents the channel-attention-adjusted feature map.

In summary, this study combines the EMD-FastICA-ATCN time-frequency feature hybrid algorithm with the improved BiGRU to build the rotating equipment fault detection and diagnosis model (EF-ATCN-BiGRU). The EF-ATCN-BiGRU model flowchart is shown in Figure 6.

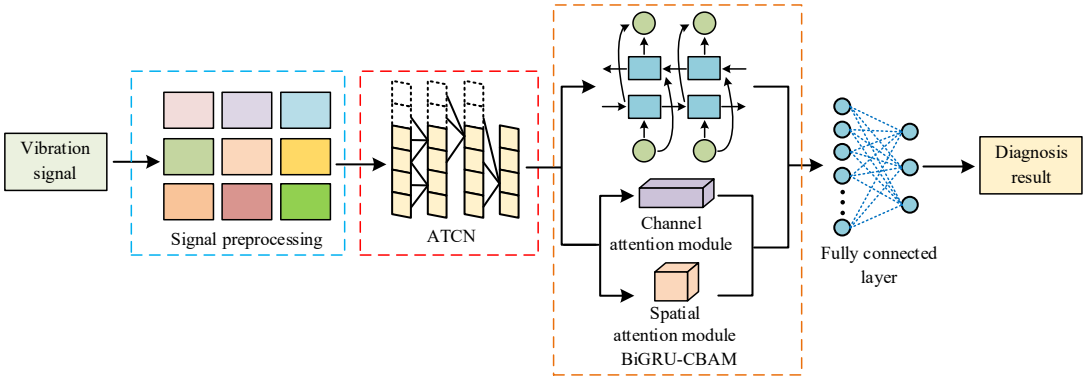


Figure 6: EF-ATCN-BiGRU model operation process diagram

As shown in Figure 6, in the EF-ATCN-BiGRU model flow, the original fault vibration signal first undergoes preprocessing to obtain a denoised signal. The feature tensor is then fed into the ATCN network, which captures local temporal features via multi-scale dilated convolution and models global spectral dependencies using self-attention, outputting a fused time-frequency feature vector. The feature vector enters the improved BiGRU network, first extracting forward and backward temporal dependencies through bidirectional gated recurrent units. Residual connections fuse input features with BiGRU outputs to ease gradient vanishing. The CBAM module weights and calibrates features in channel and spatial dimensions, enhancing fault-sensitive components. Finally, a fully

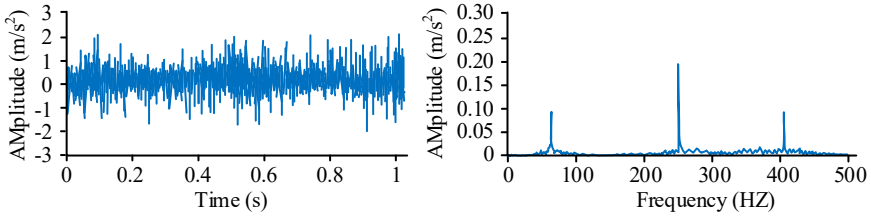
connected layer outputs fault class probabilities, achieving precise diagnosis of rotating equipment faults under multiple operating conditions.

4. Performance Validation of Rotating Equipment Fault Detection and Diagnosis Model based on Time-frequency Feature Hybrid Strategy

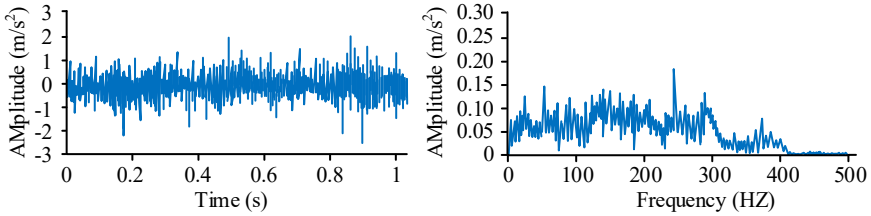
4.1 Performance Analysis of EMD-FastICA-ATCN Time-frequency Feature Hybrid Algorithm

To evaluate the performance of the EMD-FastICA-ATCN time-frequency feature hybrid algorithm, the study compared it with the Variational Mode Decomposition-Independent Component Analysis-Temporal Convolutional Network (VMD-ICA-TCN) algorithm, Ensemble Empirical Mode Decomposition-Principal Component Analysis-Bidirectional Long Short-Term Memory (EEMD-PCA-BiLSTM) algorithm, and Wavelet Packet Decomposition combined with Transformer (WPD-Trans) algorithm. The experiments ran on a Windows 10 system, using the PyTorch deep learning framework with the Adam optimizer.

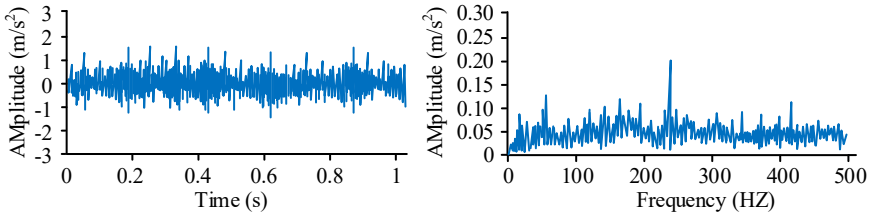
The CPU was AMD Ryzen 9 7950X3D@4.7GHz, the GPU was NVIDIA RTX 3090 (24GB), and the memory was 32GB. The Case Western Reserve University (CWRU) fault dataset was used, split into training, validation, and test sets by a ratio of 6:2:2. The maximum number of iterations was set to 150, and the learning rate was 0.0001. The denoising performance of the four algorithms was tested, with results shown in Figure 7.



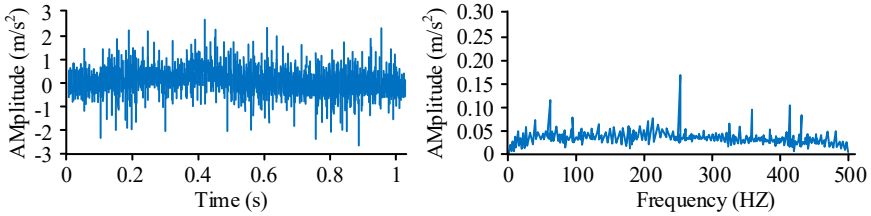
(a) Waveform diagram of the denoised signal by EMD-FastICA-ATCN



(b) Waveform diagram of the denoised signal by WPD-Trans



(c) Waveform diagram of the denoised signal by VMD-ICA-TCN



(d) Waveform diagram of the denoised signal by EEMD-PCA-BiLSTM

Figure 7: Comparison of noise reduction performance test results

Figure 7(a) showed that after processing the fault signals with the EMD-FastICA-ATCN algorithm, fault characteristic frequencies at 82Hz, 151Hz, and 408Hz were clearly identified, and noise interference was greatly reduced. Figures 7(b) and 7(c) indicated that WPD-Trans and VMD-ICA-TCN algorithms extracted fault frequencies but suffered from significant noise interference. Figure 7(d) showed that EEMD-PCA-BiLSTM identified fault frequencies but failed to effectively remove noise, which

interfered with feature extraction. Overall, the EMD-FastICA-ATCN algorithm, by adaptively decomposing nonlinear non-stationary vibration signals into IMF components through EMD and separating mixed modes with FastICA, better preserved the true signal and largely eliminated noise, enabling more accurate extraction of fault characteristic frequencies. To further validate the performance of EMD-FastICA-ATCN, the average loss curves of the four algorithms were compared, with results shown in Figure 8.

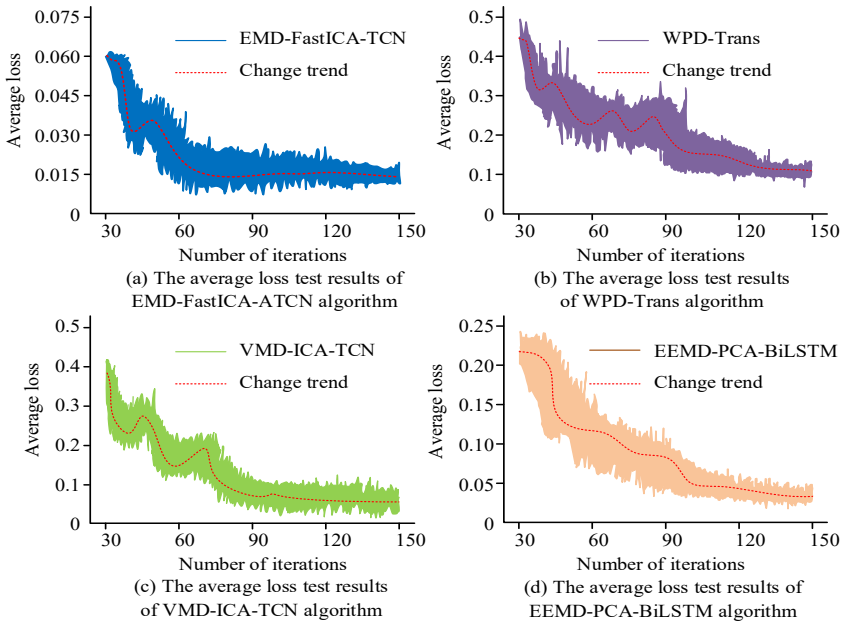


Figure 8: Comparison of average loss curve results

Figure 8(a) showed that the EMD-FastICA-ATCN algorithm reached an average loss of 0.015 at 120 iterations, and the average loss curve stabilized after 60 iterations. Figure 8(b) showed that WPD-Trans stabilized after 110 iterations. Figure 8(c) showed that at 120 iterations, the CNN-BiLSTM algorithm's average loss was 0.35 higher than that of EMD-FastICA-ATCN. Figure 8(d) showed that CNN-LSTM-Attention reached an average loss of 0.03 at 130 iterations, with the loss curve flattening after 100 iterations. In summary, the EMD-FastICA-ATCN algorithm, by combining attention mechanisms to emphasize key fault features and using TCN for cross-domain time-frequency feature integration, enhanced feature completeness and effective feature capture, resulting in a smaller average loss than the comparison algorithms. To further demonstrate performance, Mean Squared Error (MSE) and Peak Signal-to-Noise Ratio (PSNR) of the four algorithms were tested, with results shown in Figure 9.

Figure 9(a) showed that when the signal-to-noise ratio (SNR) exceeded 22 dB, the WPD-Trans model had a lower root mean square error than the LMMSE model. At 20 dB SNR, the EMD-FastICA-ATCN algorithm's root mean square error was 0.04. Figure 9(b) showed that at 25 dB SNR, EMD-FastICA-ATCN's PSNR reached 34 dB, and during the tests, its PSNR consistently surpassed other algorithms. These results indicated that EMD-FastICA-ATCN better preserved fault features in vibration signals, reduced time-frequency information distortion, and achieved an optimal balance between denoising fidelity and signal integrity, providing higher-quality time-frequency hybrid features for subsequent fault detection and diagnosis.

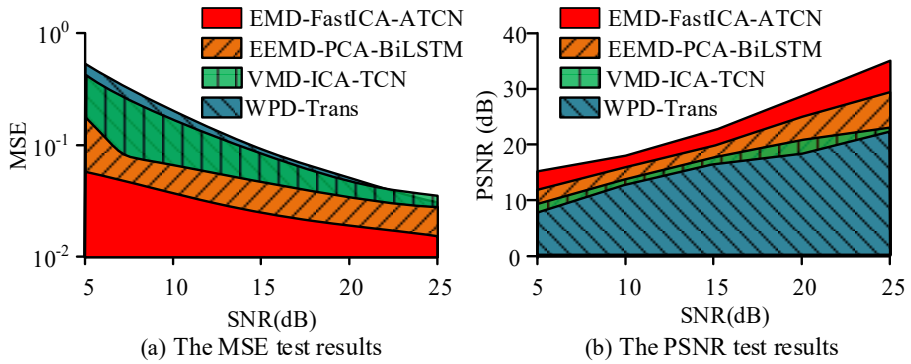


Figure 9: Comparison of experimental results of MSE and PSNR

4.2 Application Effect of EF-ATCN-BiGRU Rotating Equipment Fault Diagnosis Model

After verifying the performance of the EMD-FastICA-ATCN algorithm, the study further analyzed the EF-ATCN-BiGRU fault detection and diagnosis model constructed on its basis. The model was compared with Complete Ensemble Empirical Mode Decomposition with Adaptive Noise-Residual Network-Gated Recurrent Unit (CEEMDAN-ResNet-GRU), Multi-Branch Residual Shrinkage Convolution-Global Second Order Pooling-Network (MBRSC-GSoP-Net), and Efficient Channel Attention-Local Maximum Mean Discrepancy-Convolutional Neural Network (ECA-LMMD-CNN). The datasets used were CWRU and the Xian Jiao Tong University-Sum Young (XJTU-SY) mechanical fault detection and diagnosis dataset.

The batch size was 64, the learning rate was 0.0001, and the training iterations were 250. Loss curves of the four models were tested, with results shown in Figure 10.

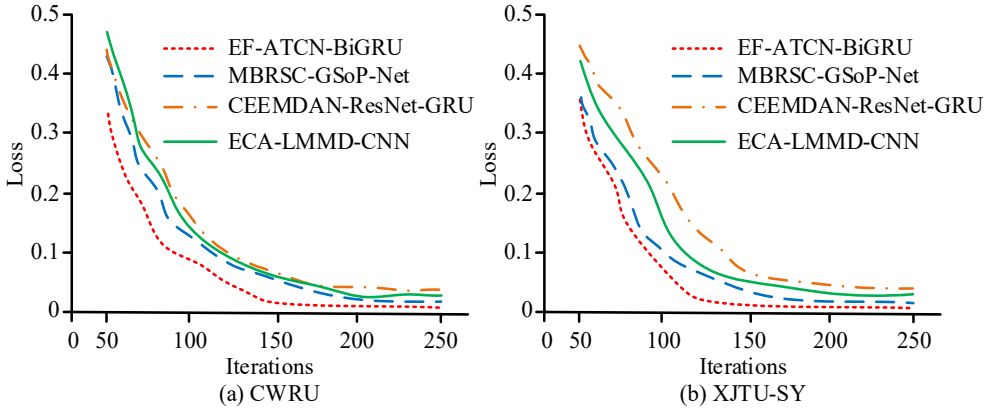


Figure 10: Comparison of experimental loss curve results

Figure 10(a) showed that at 90 iterations, EF-ATCN-BiGRU had a loss value of 0.014, while MBRSC-GSoP-Net and ECA-LMMD-CNN had losses of 0.022 and 0.026, respectively. Figure 10(b) shows that at 120 iterations, EF-ATCN-BiGRU's loss was 0.013, lower than that of CEEMDAN-ResNet-GRU and MBRSC-GSoP-Net by 0.020 and 0.015, respectively. Overall, EF-ATCN-BiGRU had the fastest convergence and lowest loss, indicating better diagnostic accuracy. To further analyze its performance, Pearson correlation coefficients of the four models were tested, with results shown in Figure 11.

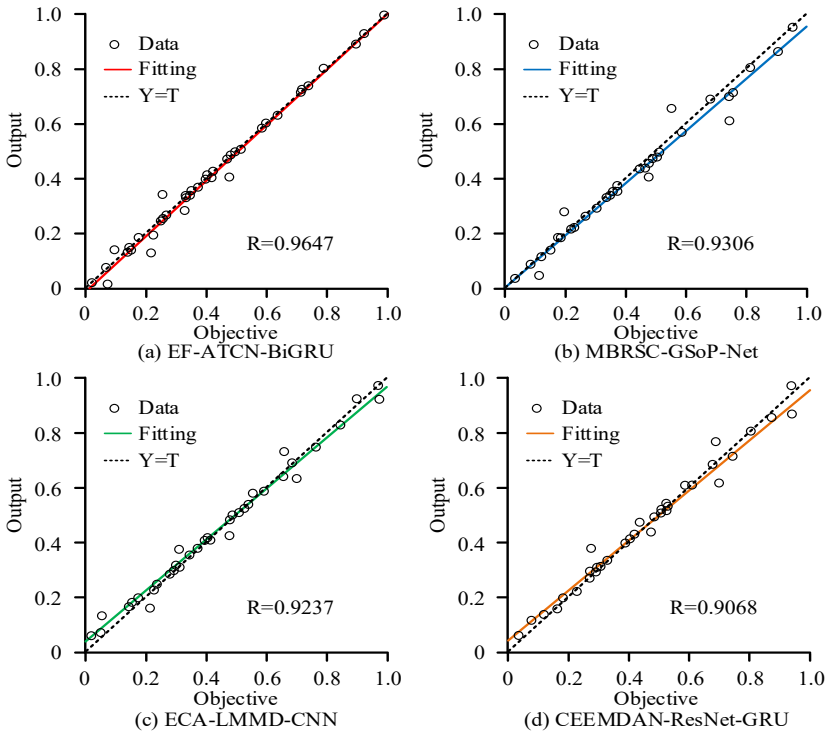


Figure 11: Comparison of Pearson correlation coefficient experimental results

Figure 11(a) showed that EF-ATCN-BiGRU had a strong fit, with data points closely aligned to the fitted curve and a correlation coefficient of 0.9647. Figures 11(b) and 11(c) showed that MBRSC-GSoP-Net and ECA-LMMD-CNN had lower correlation coefficients than the proposed model. Figure 11(d) showed that CEEMDAN-ResNet-GRU's correlation coefficient dropped by 0.0579. These results indicated that EF-ATCN-BiGRU diagnosed rotating equipment faults accurately and had significantly better linear consistency with manual diagnosis compared to other models. To further evaluate EF-ATCN-BiGRU's application performance, diagnostic precision, and recall of the four models were tested in two factories (A and B) in southern China, with results shown in Table 1.

Table 1. Comparison of precision and recall results

Data set	Model	Precision (%)	Recall (%)
Factory A	EF-ATCN-BiGRU	94.5	93.7
	MBRSC-GSoP-Net	92.7	92.2
	ECA-LMMD-CNN	91.8	90.9
	CEEMDAN-ResNet-GRU	90.3	89.5
Factory B	EF-ATCN-BiGRU	95.6	94.6
	MBRSC-GSoP-Net	93.1	92.5
	ECA-LMMD-CNN	92.4	91.8
	CEEMDAN-ResNet-GRU	90.7	90.1

In Table 1, EF-ATCN-BiGRU achieved a diagnostic precision of 94.5% and a recall of 93.7% in factory A. In factory B, EF-ATCN-BiGRU's precision was 95.6%, while MBRSC-GSoP-Net and ECA-LMMD-CNN had precisions of 93.1% and 92.4%. EF-ATCN-BiGRU's recall was 94.6%, which improved by 2.1% compared to MBRSC-GSoP-Net. Overall, EF-ATCN-BiGRU maintained good diagnostic precision and recall across different factories.

5. Summary and Future Work

In modern industry, rotating equipment fault detection and diagnosis play a crucial role in production safety and efficiency. Traditional methods had problems such as insufficient fusion of time-frequency features and inadequate capture of temporal dependencies, making it difficult to meet the diagnostic needs under complex operating conditions. Therefore, this study proposed a rotating equipment fault detection and diagnosis model based on time-frequency feature fusion. The study first applied EMD and FastICA to denoise the raw signals. Then, by combining attention mechanisms with TCN's multi-scale dilated convolution, the EMD-FastICA-ATCN algorithm was constructed to achieve cross-domain time-frequency feature fusion. On this basis, the study introduced a CBAM-enhanced BiGRU to capture forward and backward temporal dependencies of fault evolution and built the EF-ATCN-BiGRU diagnostic model. Experimental results showed that the EMD-FastICA-ATCN algorithm reached a low loss of 0.015 after 120 iterations, and its PSNR consistently outperformed comparison algorithms, accurately preserving fault characteristic frequencies. Additionally, the EF-ATCN-BiGRU model achieved diagnostic recall rates of 93.7% and 92.5% in different factories, with a Pearson correlation coefficient of 0.9647, significantly surpassing comparison models. In summary, the proposed model effectively improved rotating equipment fault detection and diagnosis performance through an adaptive hybrid of time-frequency features and precise capture of bidirectional temporal dependencies. However, the study tested limited operating conditions and fault types, and the model's generalization capability under extreme noise and heavy load environments still requires verification. Future work will expand testing scenarios and further optimize the model's robustness and adaptability.

Acknowledgement

The research is supported by: Key University-Level Scientific Research Project “Research on Real-Time Monitoring and Control Strategies for Tightening Quality in Intelligent Manufacturing Environments”, Project ID: 2024zr005

References

[1] Althubaiti A, Elasha F, Teixeira J A. Fault diagnosis and health management of bearings in rotating equipment based on vibration analysis–a review. Journal of Vibroengineering, 2022, 24(1): 46-74.

[2] Tama B A, Vania M, Lee S, Lim S. Recent advances in the application of deep learning for fault diagnosis of rotating machinery using vibration signals. Artificial Intelligence Review, 2023, 56(5): 4667-4709.

[3] Liu D, Cui L, Wang H. Rotating machinery fault diagnosis under time-varying speeds: A review. IEEE Sensors Journal, 2023, 23(24): 29969-29990.

[4] Fan J, Zhang K, Huang Y, Zhu Y, Chen B. Parallel spatio-temporal attention-based TCN for multivariate time series prediction. Neural Computing and Applications, 2023, 35(18): 13109-13118.

[5] Qiao Y, Xu H M, Zhou W J, Peng B, Hu B, Guo X. A BiGRU joint optimized attention network for recognition of drilling conditions. Petroleum Science, 2023, 20(6): 3624-3637.

[6] Limouni T, Yaagoubi R, Bouziane K, Guissi K, Baali E H. Accurate one step and multistep forecasting of very short-term PV power using LSTM-TCN model. Renewable Energy, 2023, 205: 1010-1024.

[7] Li W, Wei Y, an D, Jiao Y, Wei Q. LSTM-TCN: Dissolved oxygen prediction in aquaculture, based on combined model of long short-term memory network and temporal convolutional network. Environmental Science and Pollution Research, 2022, 29(26): 39545-39556.

[8] Wei X, Wang Z. TCN-attention-HAR: Human activity recognition based on attention mechanism time convolutional network. Scientific Reports, 2024, 14(1): 7414-7427.

[9] Rana M R R, Nawaz A, Ali T, El-Sherbeeny A M, Ali W. A BiLSTM-CF and BiGRU-based deep sentiment analysis model to explore customer reviews for effective recommendations. Engineering, Technology & Applied Science Research, 2023, 13(5): 11739-11746.

[10] Dave E, Chowanda A. IPerFEX-2023: Indonesian personal financial entity extraction using indoBERT-BiGRU-CRF model. Journal of Big Data, 2024, 11(1): 139-162.

[11] Jin Y, Hou L, Chen Y. A time series transformer based method for the rotating machinery fault diagnosis. Neurocomputing, 2022, 494: 379-395.

[12] Wan W, Chen J, Zhou Z, Shi Z. Self-supervised simple siamese framework for fault diagnosis of rotating machinery with unlabeled samples. IEEE Transactions on Neural Networks and Learning Systems, 2022, 35(5): 6380-6392.

[13] Jia N, Cheng Y, Liu Y, Tian Y. Intelligent fault diagnosis of rotating machines based on wavelet time-frequency diagram and optimized stacked denoising auto-encoder. IEEE Sensors Journal, 2022, 22(17): 17139-17150.

[14] Kim T, Lee S. A novel unsupervised clustering and domain adaptation framework for rotating machinery fault diagnosis. IEEE Transactions on Industrial Informatics, 2022, 19(9): 9404-9412.

[15] Jiang C, Chen H, Xu Q, Wang X. Few-shot fault diagnosis of rotating machinery with two-branch prototypical networks. Journal of Intelligent Manufacturing, 2023, 34(4): 1667-1681.

[16] Wang Y, Zhou J, Zheng L, Gogu C. An end-to-end fault diagnostics method based on convolutional neural network for rotating machinery with multiple case studies. Journal of Intelligent Manufacturing, 2022, 33(3): 809-830.

[17] Li L, Li Y, Mao R, Li L, Hua W, Zhang J. Remaining useful life prediction for lithium-ion batteries with a hybrid model based on TCN-GRU-DNN and dual attention mechanism. IEEE Transactions on Transportation Electrification, 2023, 9(3): 4726-4740.

[18] Gheisari M, Hamidpour H, Liu Y, Saedi P, Raza A, Jalili A, Rokhsati H, Amin R. Data Mining Techniques for Web Mining: A Survey. Artificial Intelligence and Applications, 2023, 1(1): 3-10.

[19] Amriza R N S, Ngafidin K N M. BiGRU-CNN-AT: classifiying emotion on social media. Data Technologies and Applications, 2025, 59(2): 250-275.

[20] Wang W, Tan X, Zhang P, Wang X. A CBAM based multiscale transformer fusion approach for remote sensing image change detection. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2022, 15: 6817-6825.