Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Adaptive Wavelet Threshold Attention and Explainable Lightweight CNN for Strongly Time-Varying Fault Diagnosis and Feature Extraction Mechanism of Rotating Machinery

Tengji Xia
Nanjing University of Posts and Telecommunications
Pengfei Pang
664522431@qq.com
Army Engineering University

Article

Keywords: rotating machinery, fault diagnosis, strongly time-varying conditions, adaptive wavelet, lightweight

Posted Date: November 4th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11297250/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Adaptive Wavelet Threshold Attention and Explainable Lightweight CNN for Strongly Time-Varying Fault Diagnosis and Feature Extraction Mechanism of Rotating Machinery

Tengji Xia [1], Pengfei Pang[2,*]
1 Nanjing University of Posts and Telecommunications, Nanjing, 210007, China
2 Army Engineering University, Nanjing, 210001, China
*Correspondence: 664522431@qq.com (Pengfei Pang)
This work was supported by the National Natural Science Foundation of China under Grant 51705531.

ABSTRACT To address the problems of fault feature drift, weak information submerged by noise in rotating machinery under strongly time-varying conditions, as well as frequency domain rigidity, low-frequency distortion, and high deployment cost of traditional CNNs, an explainable lightweight CNN based on adaptive wavelet threshold attention is proposed. The model captures full-frequency features through an adaptive wavelet convolution layer, strengthens fault features and suppresses noise via an adaptive threshold temporal attention enhancement module, and achieves lightweight design with a single-layer architecture. On bearing variable-speed and gearbox variable-speed-variable-load datasets, the average diagnostic accuracy reaches 93.75%, and the F1-score exceeds 79% under low signal-to-noise ratio conditions. Its performance stems from the synergistic closed-loop of wavelet convolution, adaptive threshold temporal attention enhancement, and gradient feedback. Adaptable to industrial edge devices, the model provides a reliable technical solution for strongly time-varying fault diagnosis.

INDEX TERMS rotating machinery, fault diagnosis, strongly time-varying conditions, adaptive wavelet, lightweight

1. Introduction

Rotating machinery including gearboxes and motors acts as core power equipment in industrial production, whose stable operation directly determines production continuity and equipment operation & maintenance costs [1]. In practical industrial scenes, such devices are commonly affected by coupled strongly time-varying operation and heavy environmental noise: continuous fluctuations of rotating speed and load lead to drift of fault characteristic frequencies, while electromagnetic interference and vibration coupling from adjacent components submerge weak fault impulse features, resulting in highly non-stationary vibration signals [2]. Traditional diagnostic methods fail to capture stable fault information under such complex conditions, which creates urgent demand for targeted intelligent diagnostic solutions.

Fault diagnosis technology has been evolving around adapting to complex industrial operating conditions, having gone through three major stages.

The first stage corresponds to model-based fault diagnosis methods. This type of method Model-based diagnosis relies on equipment physical equations to build the mapping relationship between operating states and faults [3]. This method possesses clear physical interpretability, yet it requires massive structural parameters and manual calibration. Repeated parameter tuning is inevitable under variable working conditions, limiting its application in time-varying industrial sites [4].

The second stage refers to vibration signal-based fault diagnosis methods, which gradually became a research focus with the popularization of sensing technology [5]. Signal decomposition algorithms including Fourier Transform (FT) [6], wavelet transform (WT) [7], and variational mode decomposition (VMD) [8] are widely adopted to separate fault ingredients from original signals. However, fixed wavelet scales and artificial feature screening cannot adapt to dynamically drifted fault frequencies, leading to poor self-adaptability for field diagnosis [9].

The third stage corresponds to data-driven intelligent fault diagnosis [10]. Shallow models such as support vector machines (SVMs) [11] and random forests [12] lack sufficient feature representation ability for multi-fault coupling and strong time-varying interference [13]. Recent deep learning methods, including Autoencoders (AEs) [14], Restricted Boltzmann Machines (RBMs) [15], Convolutional Neural Networks (CNNs) [16], Graph Neural Networks (GNNs) [17], Recurrent Neural Networks (RNNs) [18], Long Short-Term Memory (LSTM) networks [19], Gated Recurrent Units (GRUs) [20], and Transformers [21], have been widely applied for fault diagnosis. However, most of them are either insensitive to local signal features or overly focused on global temporal dependence, making them less suitable for vibration feature extraction [22]. In contrast, 1-D CNNs excel in capturing local fault impulses from one-dimensional vibration signals, achieving superior feature extraction efficiency and diagnostic accuracy. Therefore, this study focuses on 1-D CNN-based diagnostic methods, which are uniformly referred to as CNN in the following text.

Despite the automatic feature extraction advantage of CNNs, two inherent defects restrict their application under strongly time-varying noisy conditions. Randomly initialized convolution kernels lack time-frequency prior knowledge of wavelet transform and cannot dynamically track drifting fault frequencies across full frequency bands [23]. Meanwhile, multi-layer stacked topology

causes cumulative distortion of low-frequency weak fault features, accompanied by excessive model parameters that raise high deployment barriers for industrial edge controllers and fail to meet real-time diagnosis demands [16].

To remedy the time-frequency rigidity of standard CNNs, researchers integrate wavelet multi-resolution analysis with convolutional networks and form two mainstream research routes, both of which have unreachable inherent drawbacks [24]. The first route adopts offline WT preprocessing before CNN input: wavelet decomposition is separated from network training, leading to decoupled end-to-end optimization and artificial fixed wavelet hyperparameters with poor adaptability [25]. The second route embeds wavelet functions as convolution kernels inside networks. W-CNN utilizes fixed-scale wavelet kernels that cannot adapt to frequency drift [26]; WKN introduces learnable wavelet layers but lacks gradient dynamic constraints, resulting in degraded time-frequency localization performance [27]; the Multi-scale Convolutional Neural Network with Channel Attention (CA-MCNN) combines multi-scale convolution and static channel attention, yet cannot eliminate low-frequency feature distortion brought by multi-layer stacked structures [28].

Overall, current wavelet-improved CNN frameworks have coupled defects under noisy time-varying conditions and face a unified core bottleneck: few integrated frameworks can synchronously realize learnable wavelet scales for dynamic frequency tracking, automatic denoising without artificial thresholds, low-distortion lightweight feature transmission for edge devices, as well as bidirectional gradient interaction between wavelet convolution and denoising modules. Existing studies only solve one or two of the above problems separately.

Based on the explainable lightweight CNN (ELCNN) proposed by Pang et al. [29], we propose an adaptive wavelet threshold attention ELCNN (AWTA-ELCNN). The model incorporates three core components: adaptive wavelet convolution for full-band dynamic feature extraction, the adaptive threshold temporal attention enhancement (ATTAE) module for directional weak fault enhancement and noise suppression, and a single-layer lightweight topology that eliminates low-frequency feature distortion to support embedded deployment.

Compared with mainstream wavelet CNN baselines (W-CNN, WKN, CA-MCNN), AWTA-ELCNN has three key innovations to solve the above bottlenecks. First, its equidistantly initialized adaptive wavelet convolution learns full-band scales to track speed-induced frequency drift, breaking the constraint of fixed wavelet scales in prior works. Second, the ATTAE module computes dynamic thresholds from local signal statistics and backpropagates weighted gradients to wavelet kernels, building a fault-enhancing positive feedback loop outperforming static channel attention. Third, the single-layer topology shortens feature transmission paths to reduce low-frequency feature loss. With only 13.54 K parameters and interpretable inter-module coordination, it differs from deep multi-layer wavelet CNNs.

Unlike most diagnostic studies that merely evaluate classification accuracy, this work balances performance validation and internal mechanism analysis. We thoroughly interpret the wavelet scale-frequency mapping rule, ATTAE gradient adjustment logic, and their closed-loop synergy, achieving dual advantages of high diagnosis accuracy and good model interpretability.

The rest of this paper is structured as follows. Section II presents the architecture and mathematical derivation of AWTA-ELCNN. Section III carries out comparative tests on two time-varying datasets to verify its overall accuracy, anti-noise performance and lightweight property. Section IV explores the feature extraction and collaborative optimization mechanism of core modules. Section V summarizes this work and prospects future research.

2. AWTA-ELCNN PROPOSED FOR STRONG TIME-VARYING VIBRATION SIGNALS

To address the challenge of feature tracking for lightweight models under strongly time-varying operating conditions, an AWTA-ELCNN based on ELCNN is proposed. As shown in Fig. 1, the model integrates wavelet multi-resolution analysis and temporal attention mechanism, employing an extremely simple single-layer topology comprising an input layer, an adaptive wavelet convolutional layer, an ATTAE module, and a classification layer. A data-driven hybrid diagnostic paradigm that fuses wavelet prior knowledge, adaptive threshold temporal enhancement, and single-layer convolution is thereby formed, simultaneously achieving multi-scale fault feature enhancement, temporal noise suppression, and adaptation to lightweight deployment. It provides technical support for the reliable diagnosis of rotating machinery under complex operating conditions.

2.1 Adaptive wavelet convolutional layer

The adaptive wavelet convolutional layer serves as a key component for the model to achieve multi-scale feature extraction. Its design primarily involves the incorporation of the time-frequency analysis capability of wavelet transform into convolutional operations, with fault features of different frequencies in the signal captured through the adaptive learning of scale factors [30]. The specific design is as follows.

1) Structural design and mathematical expression

WT is primarily performed through time-frequency domain transformation of the mother wavelet $\psi(t)$ via scale factors a and translation factors b, enabling multi-scale decomposition of the signal [31], [32], [33]. The expression is given as follows:

$$WT_x(a,b) = \int_{-\infty}^{\infty} x(t)\psi\left(\frac{t-b}{a}\right)dt \tag{1}$$

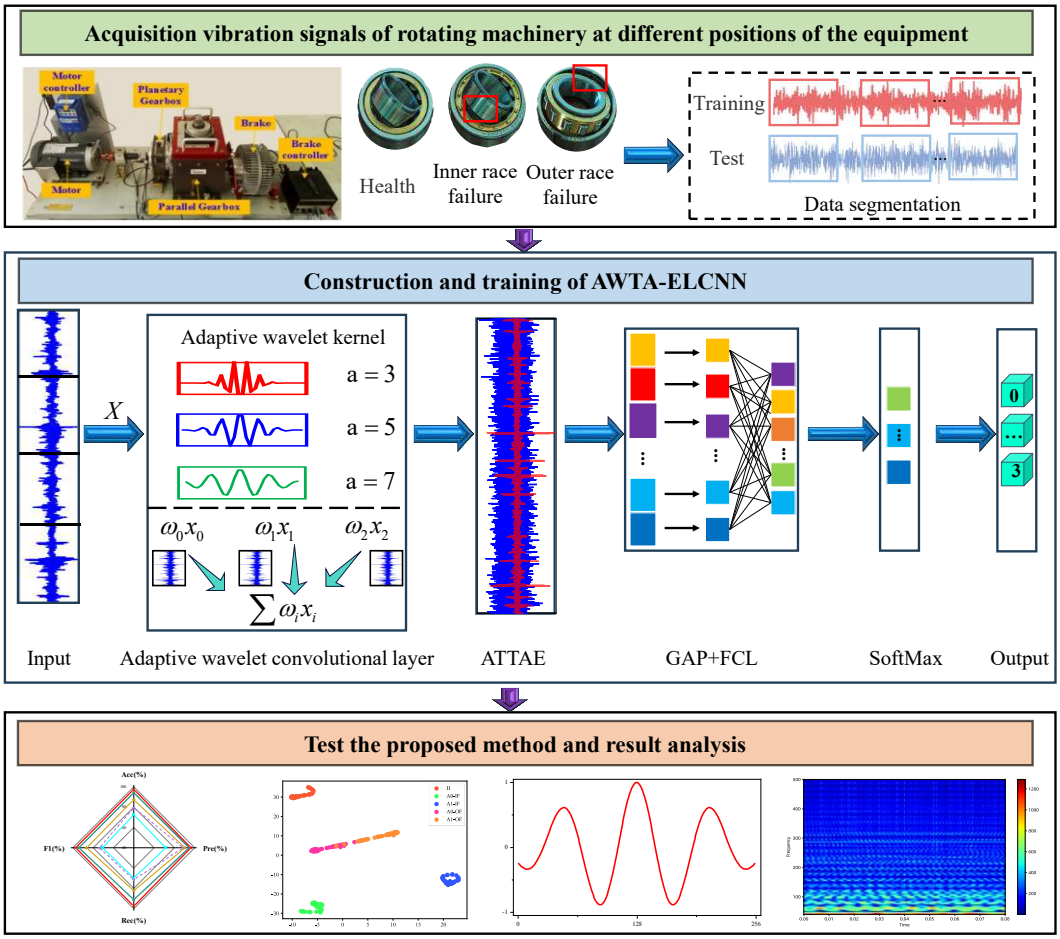


Fig. 1 AWTA-ELCNN-based Strongly Time-Varying Fault Diagnosis Flowchart for Rotating Machinery.

The adaptive wavelet convolutional layer is constructed by taking discrete wavelet functions as the convolutional kernels of AWTA-ELCNN.

$$y(n)=\sum_{k=0}^{L-1}x(n+k)\cdot\psi\left(\frac{k-b}{a}\right) \tag{2}$$

Where L denotes the convolutional kernel size, n denotes the convolutional sliding scan position, and k denotes the convolutional sliding step index. Since the sliding process of the convolutional kernel enables full-range position scanning of the signal, it overlaps with the function of the translation factor; in contrast, the scale factor is used to adjust the scale width of the wavelet function to capture features of different frequencies in the signal, complementing the convolutional function. Based on this, Equation 2 can be simplified to a form that retains only the scale factor.

$$y(n)=\sum_{k=0}^{L-1}x(n+k)\cdot\psi\left(\frac{k}{a}\right) \tag{3}$$

2) Initialization strategy of scale factor

To fully cover the main frequency range of the signal and ensure the integrity of multi-scale feature extraction, the scale factors are initialized using an equidistant initialization strategy, with C scale factors generated within the range $[a_{\min},a_{\max}]$. The expression is given as follows:

$$a_c=a_{\min}+\frac{(a_{\max}-a_{\min})\cdot(c-1)}{C-1},c=1,2,\cdots,C \tag{4}$$

Where $a_{\min}$ and $a_{\max}$ are the minimum and maximum scales preset according to the fault characteristic frequency range, and C denotes the number of scales. A foundation for subsequent adaptive learning is laid by this strategy through the reasonable distribution of initial scales, and the model's capability to capture full-frequency-band fault features is ensured.

Equidistant allocation is merely used for scale initialization. The upper and lower scale limits are determined by the frequency range of measured vibration signals, and initial scales are evenly spaced within this bound given the predefined scale number C. Scale coefficients are learnable and optimized via backpropagation. Equally spaced initialization ensures full frequency coverage, while trained intervals dynamically track frequency drift under time-varying conditions.

2.2 Adaptive threshold temporal attention enhancement module

ATTAE is regarded as a key unit in the model for feature enhancement and noise suppression. Through the synergistic effect of local statistical feature calculation, dynamic threshold generation, and temporal attention weighting, the output features of the wavelet convolutional layer are adaptively optimized. Its internal structure and operational logic are as follows.

1) Calculation of local statistical features

To capture the local trends and fluctuations of the feature map, a sliding window is employed to calculate the mean $\mu_{c,t}$ and standard deviation $\sigma_{c,t}$. The formulas are listed below.

$$\mu_{c,t} = \frac{1}{w}\sum_{i=0}^{w-1}\left|y_{c,t+i}\right| \tag{5}$$

$$\sigma_{c,t} = \sqrt{\frac{1}{w}\sum_{i=0}^{w-1}\left(y_{c,t+i}-\mu_{c,t}\right)^2+\varepsilon} \tag{6}$$

Where $y_{c,t}$ is the temporal mapping value of the c-th feature map at time t, ε is a small constant introduced to avoid a zero standard deviation, and w denotes the size of the sliding window. The window is empirically set to 100 sampling points to capture local transient fault impulses with affordable computational overhead.

2) Generation of dynamic threshold and enhancement coefficient

Based on local statistical features, adaptive threshold $\theta_{i,c}$, normalized value $z_{c,t}$, and dynamic enhancement factor $\gamma_{c,t}$ are designed in the module, thereby enabling the differential enhancement of features. The formulas are listed below.

$$\theta_{c,t} = \mu_{c,t} + k\sigma_{c,t} \tag{7}$$

$$z_{c,t} = \frac{\left|y_{c,t}\right|}{\beta\cdot\theta_{c,t}+\varepsilon} \tag{8}$$

$$\gamma_{c,t} = \alpha\cdot z_{c,t}^{\ 2} \tag{9}$$

Where k is a standard deviation weight coefficient, which is used to determine the adaptive range of the threshold; β is a threshold scaling coefficient, which is employed to indirectly adjust the stringency of the threshold; $z_{c,t}$ is a normalized value, which is designed to normalize the signal to the threshold scale; α is an enhancement intensity coefficient, which is responsible for reflecting the regulatory effect on the enhancement amplitude; $\gamma_{c,t}$ is a tensor that dynamically changes with the input signal (with different values at each position), and its main function is to assign a specific enhancement multiple to each signal point. The learnable parameters are initialized empirically to $k=4.0$, $\beta=2.0$ and $\alpha=3.0$. These coefficients are not fixed values but are optimized via gradient backpropagation during training to adjust the threshold constraint and feature enhancement degree adaptively for input time-varying vibration signals.

3) Generation and selection of temporal attention masks

To further highlight the features of key temporal positions, a temporal attention mask $m_{c,t}$ is introduced in the module, and a weighted output is generated through a screening mechanism. The relevant equations are given as follows.

$$m_{c,t} = \begin{cases} 1; \left|y_{c,t}\right| > \theta_{c,t} \\ 0; \left|y_{c,t}\right| \le \theta_{c,t} \end{cases} \tag{10}$$

$$h_{c,t} = \begin{cases} \mathrm{sign}(y_{c,t})\cdot\left|y_{c,t}\right|\cdot\gamma_{c,t}; m_{c,t}=1 \\ y_{c,t}; m_{c,t}=0 \end{cases} \tag{11}$$

Where $m_{c,t}$ is used to mark prominent features, $h_{c,t}$ is employed to enhance prominent features, and the original weak features are retained.

2.3 Single layer convolutional architecture

Considering that the iterative learning of multi-layer convolutional networks is prone to causing distortion in the transmission of underlying fault features, an extremely simple single-layer architecture consisting of adaptive wavelet convolution, the ATTAE module, and a classification layer is adopted in AWTA-ELCNN, thereby ensuring the accurate retention and adaptive optimization of fault features in the time-frequency domain. The expression of its overall architecture is given as follows.

$$output = \mathrm{Softmax}\left(\mathrm{Pooling}\left(h_{ATTEN}\left(\mathrm{Conv}\left(x;\ \omega_{\mathrm{wavelet}}\right)\right)\right)\right) \tag{12}$$

Where $\omega_{\mathrm{wavelet}}$ is the wavelet convolutional kernel, Conv is the convolutional operation, and h_{ATTEN} is the ATTAE module.

The main advantages of this single-layer architecture are reflected in two aspects.

First, information loss is reduced by shortening the feature transmission path, thereby ensuring the integrity of multi-scale fault features.

Second, the reduction of parameter scale and computational complexity is achieved through module simplification, thereby highlighting the lightweight advantage, which is suitable for the deployment requirements of vehicle-mounted embedded terminals for rotating machinery.

Single-layer topology, adaptive wavelet convolution and the ATTAE module cooperate to form a complete feature learning pipeline. Wavelet convolution extracts multi-band fault features, ATTAE enhances effective signals and suppresses noise, and the single-layer structure ensures efficient feature propagation. This collaborative framework adapts to dynamically changing fault features in strong time-varying environments while maintaining the model's lightweight characteristics and anti-noise capability for accurate fault diagnosis.

Fig. 2 shows the complete architecture of the proposed AWTA-ELCNN, including the sequential connection of the adaptive wavelet convolutional neural network (AWCNN), ATTAE module and lightweight single-layer classification module. The model topology and detailed parameter specifications are jointly demonstrated by Fig. 2 and TABLE 1.

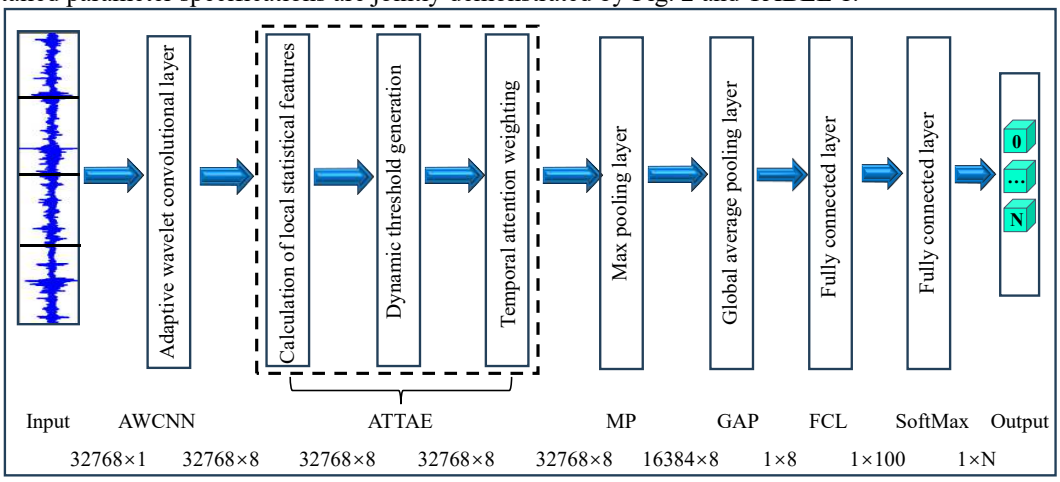


Fig. 2 AWTA-ELCNN architecture applied in this paper.

TABLE 1

Network structural parameters of AWTA-ELCNN.

Layer	Kernel size	Kernel number	Input size	Output size	Activation function	Bias	Stride	Padding	Local window size
Wavelet convolution	256	8	32768×1	32768×8	-	-	1	same	-
ATTAE	-	-	32768×8	32768×8	-	-	-	-	100
Max Pooling	2	-	32768×8	16384×8	-	-	-	-	-
Global Average Pooling	-	-	16384×8	1×8	-	-	-	-	-
Fully Connected	-	-	1×8	1×100	-	-	-	-	-
Fully Connected	-	-	1×100	1×N	SoftMax	-	-	-	-

3. Experiment

3.1 Data description

To verify the diagnostic performance under strongly time-varying working conditions, two types of data covering complex working conditions of variable speed and variable load are selected: the bearing vibration dataset from the University of Ottawa (Ottawa) [34], and the gearbox vibration dataset from MCC5 Group and Tsinghua University (MCC5-THU) [35].

1) Ottawa

Collected from an ER16K bearing platform under strongly time-varying speeds, with a sampling frequency of 200 kHz and 32,768 points per sample. Two speed profiles are included: A0 (increasing) and A1 (decreasing). The speed ranges are: A0-IF: 150–688 rpm, A0-OF: 888–1628 rpm, A1-IF: 1458–594 rpm, A1-OF: 1494–588 rpm. Three health states (Healthy (H), Inner Race Fault (IF), Outer Race Fault (OF)) form five classes, with 400 samples split 7:2:1 for training/validation/test (TABLE 2).

2) MCC5-THU

Collected from a gearbox platform under variable speed and load, with a sampling frequency of 12.8 kHz and 256,000 points per sample. The speed/load curves are shown in Fig. 3. Two working conditions are included: B0 (0→1000→500 rpm, 10 Nm), B1 (0→1000→500 rpm, 20 Nm). Three gear faults (Gear Pitting (GP), Missing Teeth (MT), Tooth Break (TB)) form six classes, with samples split 7:2:1 (TABLE 3).

TABLE 2

Description of the Ottawa dataset.

No.	Working conditions and rotational speed（rpm）	Label	No.	Working conditions and rotational speed（rpm）	Label
H	Health	0	A1-IF	The rotational speed is decreased from 1458 to 594 rpm	3
A0-IF	The rotational speed is increased from 150 to 688 rpm	1	A1-OF	The rotational speed is decreased from 1494 to 588 rpm	4
A0-OF	The rotational speed is increased from 888 to 1628 rpm	2			

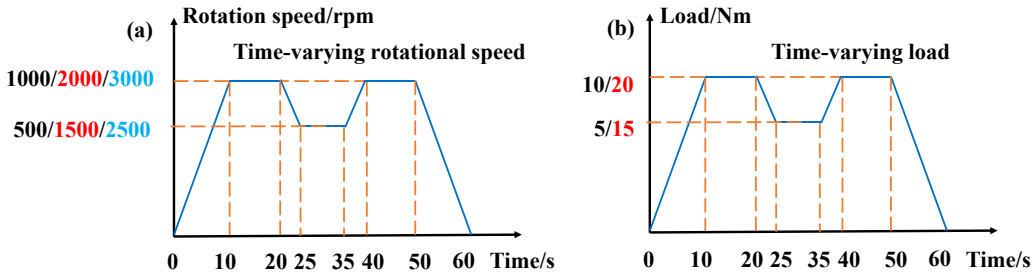


Fig. 3 Speed-time and load-time curve of the MCC5-THU dataset, (a) Speed-time curve, (b) Load-time curve

TABLE 3

Description of the MCC5-THU dataset.

No.	Working conditions and rotational speed（rpm）	Label	No.	Working conditions and rotational speed（rpm）	Label
B0-GP	Rotational speed: 0→1000→500 rpm, load: 10 Nm	0	B1-GP	Rotational speed: 0→1000→500 rpm, load: 20 Nm	3
B0-MT		1	B1-MT		4
B0-TB		2	B1-TB		5

3.2 Baseline and SOTA Methods

To systematically validate the superiority of AWTA-ELCNN for strongly time-varying fault diagnosis, representative improved CNNs and recent SOTA approaches are chosen for comprehensive comparison.

A vanilla CNN is set as the baseline. Wavelet Kernel-based CNN (W-CNN) [26] and WaveletKernelNet (WKN) [27] extract wavelet time–frequency features from vibration signals; Multi-scale CNN with Channel Attention (CA-MCNN) [28] employs attention to highlight distinguishable fault characteristics; Dual Path Convolution with Attention and Bidirectional Gated Recurrent Unit (DCA-BiGRU) [36] captures temporal variations of signals.

Several SOTA models published from 2023 to 2025 are also compared: Mechanism Character Generative Model-based CNN (MCGM-CNN) [37] produces GAN-generated virtual samples for few-shot diagnosis; improved Squeeze-and-Excitation CNN (SE-CNN) [38] improves feature extraction via SE channel attention; Wavelet-correlation Alignment Deep Canonical Correlation Analysis (WaveCORAL-DCCA) [39] conducts unsupervised cross-condition adaptation to mitigate distribution shifts; Attention-based Temporal Convolutional Network (ATCN) [40], originally proposed for RUL prediction, extracts temporal features under variable loads; Enhanced Transformer with Asymmetric Loss Function (ETALF) [41] alleviates label noise using Transformer self-attention and asymmetric loss.

All models are configured identically to guarantee fair evaluation. Four metrics, Accuracy (Acc), Precision (Pre), Recall (Rec) and F1 Score (F1) [24], are used to quantitatively evaluate classification performance from multiple dimensions.

3.3 Analysis of overall and anti-noise performance

TABLE 4 records the parameter size, computational overhead, inference time and classification accuracy of baseline and SOTA models at SNR=10dB. With only 13.54 K parameters and 0.56 M FLOPs, AWTA-ELCNN achieves an average diagnostic accuracy of 96.76% together with the shortest average inference latency of 7.72 s, achieving excellent equilibrium between prediction accuracy and computational consumption.

TABLE 4

Network scale, computational cost, diagnostic accuracy and inference latency of different models under SNR=10dB.

Model	Params (K)	FLOPs (M)	Ottawa Acc (%)	MCC5-THU Acc (%)	Avg Acc (%)	Avg Inference Time (s)
CNN	57.92	18.79	79.28±4.45	68.11±5.66	73.69±4.11	13.53±3.21
W-CNN	43.42	18.71	85.42±2.18	72.45±3.26	79.84±3.23	12.11±2.85
WKN	169.32	18.98	94.02±1.68	79.17±4.65	86.59±3.89	11.41±2.68
CA-MCNN	694.25	331.14	94.51±1.32	81.25±2.12	87.88±2.11	10.09±2.22
DCA-BiGRU	604.57	41.45	87.42±2.46	84.58±2.68	86.01±2.35	24.52±4.25
MCGM-CNN	751.69	363.24	88.26±3.68	80.78±3.68	84.52±3.17	40.45±4.66
SE-CNN	162.72	37.99	96.38±1.02	76.71±3.22	86.55±2.78	14.86±3.68
WaveCORAL-DCCA	123.88	71.15	96.27±1.44	77.45±2.89	86.86±2.22	13.79±3.33
ATCN	19.47	108.15	92.44±1.89	87.13±2.47	89.79±2.28	9.15±1.89
ETALF	261.51	19.01	87.44±2.68	78.33±3.58	82.88±2.98	28.09±4.89
AWTA-ELCNN	13.54	0.56	99.77±0.78	93.75±1.03	96.76±0.99	7.72±1.23

Fig. 4 and Fig. 5 compare the F1 scores of various methods under SNR=−5dB. The proposed model maintains F1 values higher than 79.2% on two datasets and outperforms all competitors considerably, confirming its prominent anti-noise ability and generalization against time-varying noisy samples.

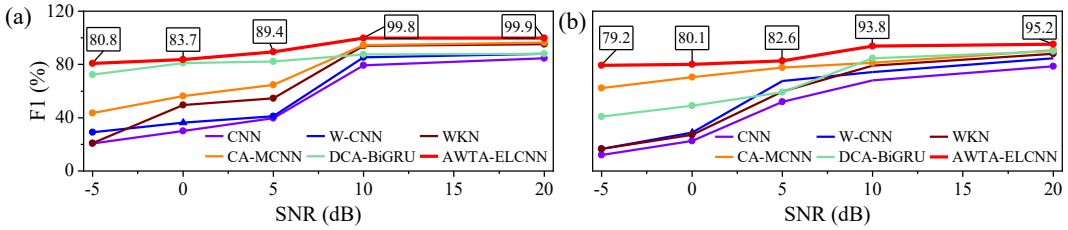


Fig. 4 Comparison of anti-noise performance of different representative improved CNNs，(a) Ottawa, (b) MCC5-THU

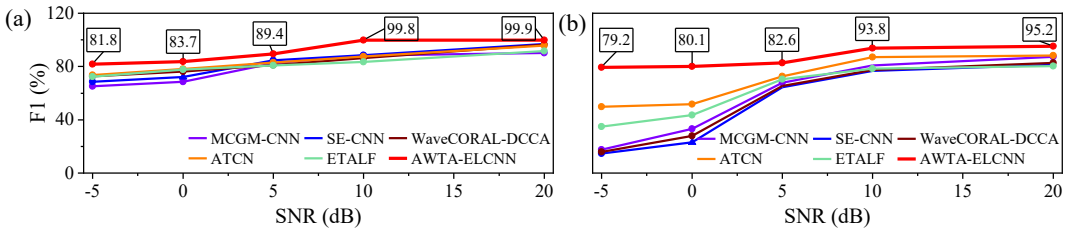


Fig. 5 Comparison of anti-noise performance of different SOTA methods，(a) Ottawa, (b) MCC5-THU

Two-sample t-test results in TABLE 5 list the p-values, accuracy gaps and Cohen's d values under clean (SNR=20dB) and noisy (SNR=−5dB) cases. Differences are statistically significant at $p < 0.05$ and effect magnitudes are graded as: tiny ($d < 0.2$), small ($0.2 \le d < 0.5$), medium ($0.5 \le d < 0.8$), large ($d > 0.8$).

Statistical results reveal insignificant marginal gains (small/medium effect, $p > 0.05$) in clean environments, whereas large, statistically significant improvements emerge under intense noise. This contrast verifies the strong anti-noise robustness of AWTA-ELCNN in practical industrial settings.

TABLE 5

Two-sample t-test results between AWTA-ELCNN and comparative models on Ottawa dataset.

Model	Ottawa SNR=20dB				Ottawa SNR=-5dB			
	p-value	Gap (%)	Cohen's d	Effect size level	p-value	Gap (%)	Cohen's d	Effect size level
CNN	0.000120	13.39	4.52	Large effect	0.000001	59.77	22.06	Large effect
W-CNN	0.000852	9.98	3.32	Large effect	0.000001	51.43	18.56	Large effect
WKN	0.272596	3.91	0.76	Medium effect	0.000001	59.53	22.16	Large effect
CA-MCNN	0.413018	2.10	0.54	Medium effect	0.000001	37.10	12.76	Large effect
DCA-BiGRU	0.000906	9.83	3.30	Large effect	0.001914	8.27	2.86	Large effect
MCGM-CNN	0.003921	7.64	2.57	Large effect	0.000028	15.54	5.38	Large effect
SE-CNN	0.473634	1.78	0.47	Small effect	0.000144	12.31	4.27	Large effect
WaveCORAL-DCCA	0.381245	2.27	0.58	Medium effect	0.002086	8.17	2.82	Large effect
ATCN	0.274212	3.74	0.76	Medium effect	0.004635	7.09	2.45	Large effect
ETALF	0.010248	6.31	2.14	Large effect	0.001945	8.26	2.86	Large effect

3.4 Effectiveness analysis of improvement strategies

Combining the quantitative metrics in TABLE 6 and the confusion matrices in Fig. 6(a)–(h), the validity of the AWTA-ELCNN improvement scheme is clearly reflected. The first five groups of the model parameters lead to low classification accuracy and large data deviations in TABLE 6, meanwhile the corresponding confusion matrices (a)–(e) exhibit severe inter-class misclassification with many non-diagonal elements. After parameter optimization, the last three experiments achieve gradually rising accuracy to 99.51% with stable convergence performance. Their matching confusion matrices (f)–(h) show far fewer misclassified samples, with most predictions concentrated on the matrix diagonal. The unified changing trend of numerical results and visual confusion distribution verifies that the combination of wavelet initialized convolution kernels, ATTAE module and single-layer convolution structure effectively reduce classification errors and improves the overall classification performance of AWTA-ELCNN.

TABLE 6

Influence of different improvement strategies of AWTA-ELCNN, among which ABC labeled red has the best performance. A, B and C represent different strategies used by the model, where A represents wavelet initialized convolution kernels, B represents ATTAE, C represents single-layer architecture. Correspondingly, A stands for randomly initialized convolution kernels, B stands for the model without the ATTAE, and C stands for a three-layer convolutional structure.

No.	Strategy Combination	Parameter Numbers	Convergence Epochs	Convergence Times(s)	Acc/%
1	ABC	36220	13±3	137.02±5.26	28.01±12.58
2	ABC	36221	40±4	191.94±6.87	30.01±19.94
3	ABC	36224	25±3	127.58±5.43	33.41±22.47
4	ABC	3453	33±2	338.71±8.98	43.51±24.68
5	ABC	36225	120±12	162.65±2.58	88.67±2.45
6	ABC	3454	81±4	114.75±4.25	93.51±2.95
7	ABC	3456	59±4	410.31±7.68	98.32±1.27
8	ABC	3457	50±2	341.83±3.68	99.51±0.74

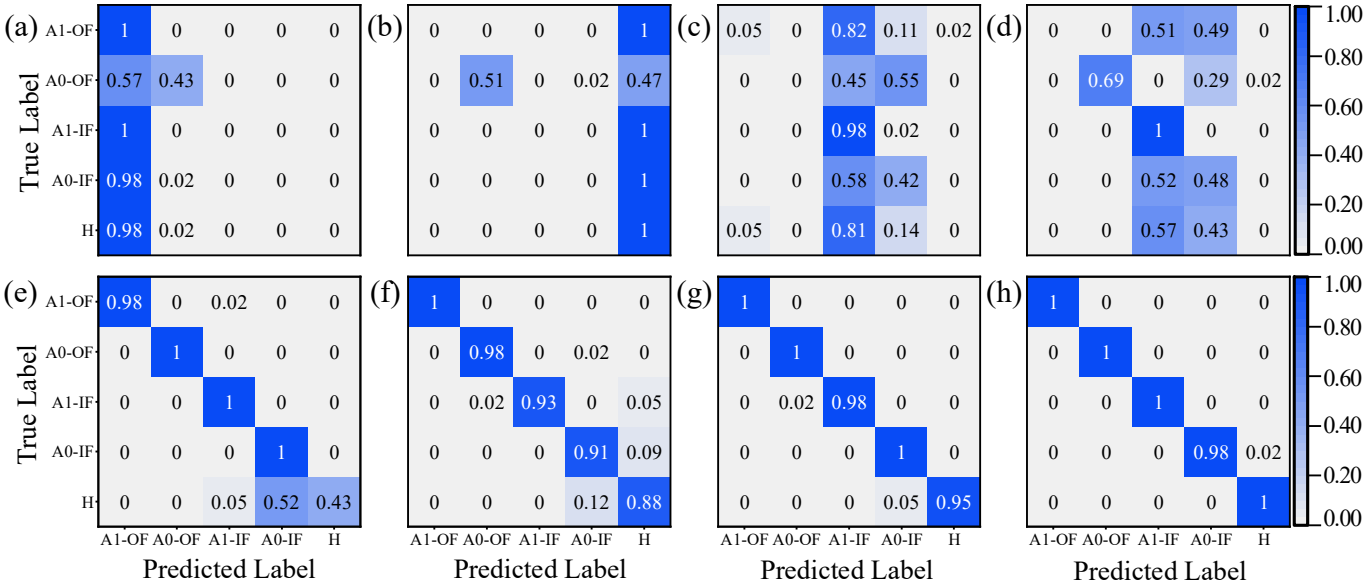


Fig. 6 Confusion matrix comparison of different improvement strategies, (a) ABC, (b) ABC, (c) ABC, (d) ABC, (e) ABC, (f) ABC, (g) ABC, (h) ABC

3.5 Analysis of feature visualization

To further validate the fault feature extraction capability of each model, Fig. 7 and Fig. 8 present the time-frequency spectra of A0-IF samples from the Ottawa dataset, which are obtained via the proposed improved CNN and other SOTA methods under strong noise (SNR=−5 dB). In the A0 operating condition, the IF rotating frequency ranges from 12.5 Hz to 27.8 Hz. This experiment selects samples corresponding to the initial rotational speed of 12.5 Hz, whose theoretical fault frequencies are definite: the fundamental fault frequency equals 67.9 Hz (calculated as 12.5 Hz×5.432, where 5.432 denotes the constant multiplier between fault frequency and rotational frequency). The sideband components at 30.4 Hz, 42.9 Hz and 67.9 Hz are the combination frequencies of the fundamental fault frequency and shaft rotating frequency.

As shown in Fig. 7 and Fig. 8, comparative methods are vulnerable to noise. Most cannot extract the 67.9 Hz fundamental frequency and its 42.9 Hz sideband, suffering from blurred features, low resolution and heavy high-frequency interference. In comparison, AWTA-ELCNN precisely locks the fault frequencies of 30.4 Hz, 42.9 Hz and 67.9 Hz and filters out irrelevant frequency disturbances.

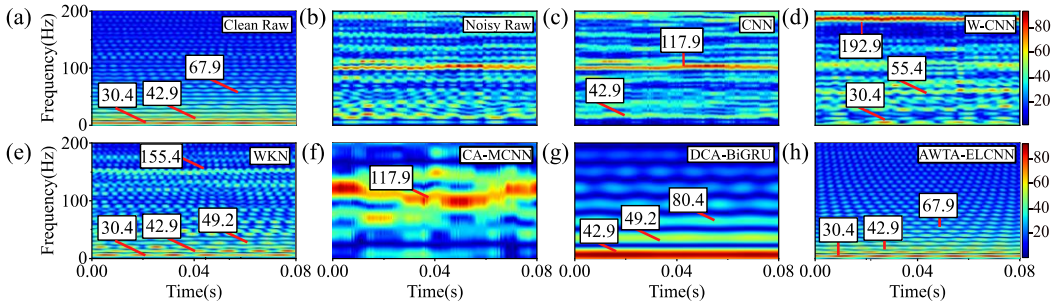


Fig. 7 Comparison of time-frequency domain features of A0-IF learned by different improved CNNs, (a)Clean Raw, (b)Noisy Raw, (c)CNN, (d)W-CNN, (e)WKN, (f)CA-MCNN, (g)DCA-BiGRU, (h)AWTA-ELCNN

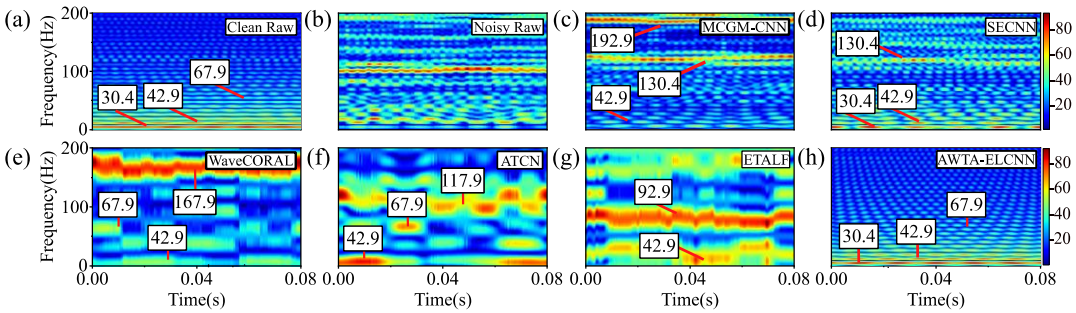


Fig. 8 Comparison of time-frequency domain features of A0-IF learned by different SOTA methods, (a)Clean Raw, (b)Noisy Raw, (c)MCGM-CNN, (d)SECNN, (e)WaveCORAL, (f)ATCN, (g)ETALF, (h)AWTA-ELCNN

These performance differences originate from structural limitations. Randomly initialized kernels in CNN, CA-MCNN, DCA-BiGRU, WaveCORAL, ETALF and ATC hardly capture dynamic fault features under varying operating conditions. W-CNN and WKN adopt fixed wavelet scales with narrow frequency coverage, incapable of coping with condition-induced frequency drift and vulnerable to weak feature loss. SECNN equipped with static attention fails to handle frequency offset and suppress noise effectively. MCGM-CNN uses manually designed multi-scale branches, exhibiting weak adaptability and extra noise during feature fusion. Both methods struggle to balance full-band feature extraction and noise reduction.

Instead, AWTA-ELCNN combines adaptive wavelet convolution kernels for full-band sensing, the ATTAE module for dynamic feature refinement, and a concise single-layer architecture for noise reduction. The synergy of the three designs enables robust and high-quality fault feature representation under complex fluctuating working conditions.

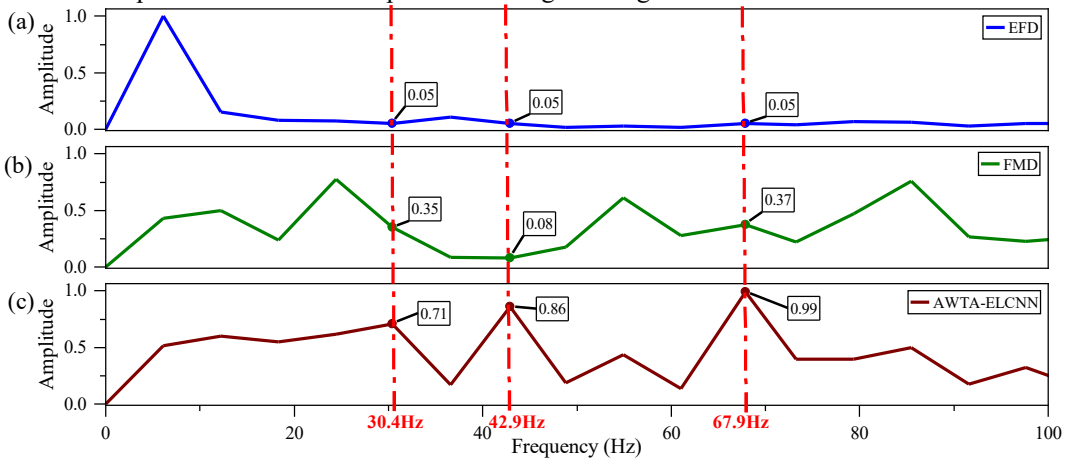


Fig. 9 Envelope spectrum comparison of A0-IF signals under SNR=−5 dB, (a)EFD, (b)FMD, (c)AWTA-ELCNN

Fig. 9 shows the envelope spectra of A0 inner race fault signals at SNR=−5dB. Two classic signal decomposition algorithms, Empirical Fourier Decomposition (EFD) [42] and Feature Mode Decomposition (FMD) [43], are adopted for comparative analysis. Obvious spectral dispersion can be seen in the processing results of EFD and FMD, where the characteristic peaks at 30.4 Hz, 42.9 Hz and 67.9 Hz are faint and covered by random noise. These defects stem from artificially preset decomposition parameters. Although internal filters support slight iterative optimization, the fixed filtering bandwidth cannot adapt to noisy signals dynamically, which disperses fault energy and generates a large number of false spectral components. In contrast, AWTA-ELCNN

produces sharp fault frequency peaks and suppresses noise effectively. Relying on data-driven adaptive wavelet filtering, the model concentrates fault energy and mitigates noise interference, delivering stronger anti-noise robustness for fault characteristic frequency extraction under noisy conditions.

4. Analysis of feature extraction mechanism

The performance advantages of AWTA-ELCNN stem from the closed-loop synergy between multi-scale feature extraction of the adaptive wavelet convolution layer and feature purification optimization of the ATTAE module. This section elaborates the internal working principle combined with theoretical derivation, visualized results and quantitative measured data.

4.1 Feature extraction mechanism of adaptive wavelet convolutional layer

Aiming at frequency drift and non-stationary characteristics of strongly time-varying vibration signals, the adaptive wavelet convolution integrates wavelet time-frequency analysis into convolution operation. The network adaptively optimizes scale parameters via backpropagation to realize automatic matching with real fault characteristic frequencies. The mechanism is analyzed from two aspects: wavelet scale-frequency mapping law and gradient regulation effect of the ATTAE module.

1) Wavelet scale-frequency mapping correlation

Scale-frequency mapping is the fundamental theoretical support for the convolutional layer to capture full-band fault features. The scaled expression of the mother wavelet is defined as follows.

$$\psi_a(t) = \frac{1}{\sqrt{a}}\psi(\frac{t}{a}) \tag{13}$$

where $\frac{1}{\sqrt{a}}$ is the normalization coefficient used to maintain energy conservation of wavelets under different scales. After Fourier transformation, the frequency-domain expression of the scaled wavelet is obtained.

$$\hat{\psi}_a(2\pi f) = \sqrt{a}\hat{\psi}(2\pi fa) \tag{14}$$

Wavelet scaling presents an inverse time–frequency correspondence: time-domain stretching compresses the response frequency band, while time-domain compression broadens the sensitive frequency range. Accordingly, the model can adjust its frequency response adaptively without manual kernel configuration to track drifted fault frequencies.

For discrete sampled vibration signals, the actual physical center frequency of scaled wavelets is calculated combined with sampling frequency f_s.

$$f_a^{\text{real}} = \frac{f_0 f_s}{a} \tag{15}$$

The inverse formula for solving the scale parameter is written as follows.

$$f_a^{\text{real}} = \frac{f_0 f_s}{a} \Rightarrow a = \frac{f_0 f_s}{f_a^{\text{real}}} \tag{16}$$

Equations (15) and (16) quantify the physical connotation of scale learning and provide a basis for setting the initial range of scale values. As shown in Fig. 10, the learned scales match the measured high-frequency impact component (45570 Hz) and low-frequency modulation components (67.9 Hz, 80.4 Hz), which verifies the frequency tracking ability of adaptive scales under variable working conditions.

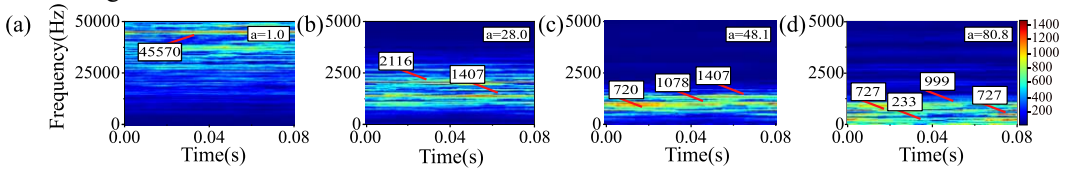


Fig. 10 Comparison of time-frequency domain characteristics of faults learned by wavelet convolutional kernels of different scales, (a)FM_1 (Feature map 1), (b)FM_2 (Feature map 2), (c)FM_3 (Feature map 3), (d)FM_4 (Feature map 4).

2) Gradient regulation of ATTAE on scale updating

The ATTAE module reconstructs the gradient update path of scale factors and forms a positive feedback loop of feature enhancement-gradient weighting-scale optimization, suppressing scale learning deviation caused by noise interference. According to the chain rule, the gradient formula of the scale factor is expressed as follows.

$$\frac{\partial \ell}{\partial a} = \frac{\partial \ell}{\partial h} \cdot \frac{\partial h}{\partial y} \cdot \frac{\partial y}{\partial a} \tag{17}$$

Where h is denoted as the weighted output of the ATTAE module, y as the output features of the adaptive wavelet convolutional layer, and $\delta = \frac{\partial h}{\partial y}$ as the gradient propagation gain term introduced by the ATTAE module.

For fault-related frequency components, ATTAE strengthens feature responses and amplifies corresponding scale gradients to drive scales to the optimal matching position. For noise components, invalid gradients are suppressed close to zero to avoid disturbing the updating of scale parameters. Two quantitative indicators quantify its actual regulatory capability.

First, convergence efficiency. The ATTAE module decreases wavelet scale convergence iterations from 99 to 81, improving training efficiency by 18% (Fig. 11).

Second, feature extraction performance under harsh noise. At SNR=−5 dB noisy condition, the intermediate feature SNR rises from 2.1 dB to 5.8 dB, recovering low-frequency fault constituents submerged by background noise (Fig. 12).

In summary, wavelet convolution breaks the fixed-frequency limitation of traditional convolution kernels and realizes full-band adaptive feature capture. The ATTAE module further optimizes the accuracy and convergence speed of scale learning. The coordinated work of the two modules builds an internal theoretical foundation for the excellent diagnostic performance reflected in TABLE 4 and statistical results in TABLE 5.

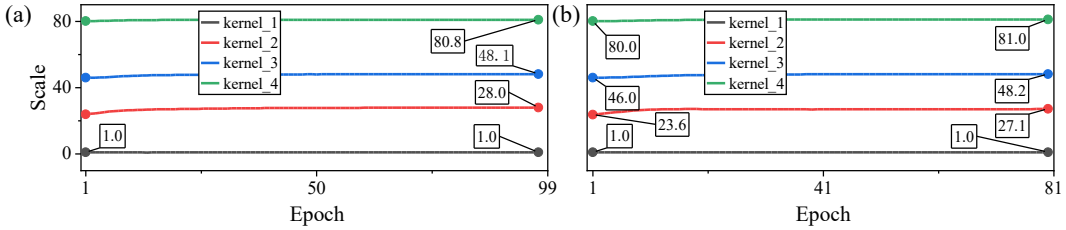


Fig. 11 Comparison of wavelet scale training process before and after using ATTAE, (a) No ATTAE, (b) ATTAE.

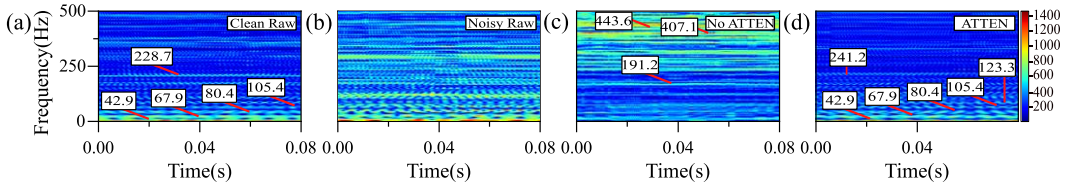


Fig. 12 Comparison of time-frequency domain characteristics of faults learned before and after using ATTAE, (a) Clean Raw, (b) Noisy Raw, (c) No ATTAE, (d) ATTAE.

4.2 Feature transformation mechanism of the ATTAE

Multi-scale features extracted by the wavelet convolution layer are inevitably mixed with background noise, resulting in serious aliasing between fault pulses and noise. Targeted at noise suppression and weak fault enhancement, the ATTAE module purifies features through dynamic threshold segmentation, signal normalization, adaptive enhancement and temporal mask focusing, and establishes a bidirectional gradient closed-loop relationship with the front-end wavelet convolution layer.

1) Theoretical basis of feature transformation

Vibration background noise follows a Gaussian distribution with amplitudes concentrated near the mean, while fault impulses exhibit sharp amplitude surges and concentrated energy. Accordingly, ATTAE employs a dynamic threshold based on local signal statistics to enhance fault-related transients and suppress noise, avoiding undesired noise amplification.

The complete multi-stage feature transformation procedure is illustrated in each subgraph of Fig. 13. The adaptive threshold segmentation in Fig. 13 (a) effectively distinguishes fault impulses from noise, achieving a fault recognition accuracy of 96.7% (misjudgment rate=3.3%), which substantially outperforms the 21.5% false alarm rate of fixed thresholds. The signal normalization in Fig. 13 (b) eliminates amplitude discrepancies caused by varying rotational speeds, yielding normalized amplitudes of 0.43–0.63 for fault regions and 0.03–0.26 for noise regions. The dynamic enhancement factor in Fig. 13 (c) adaptively assigns higher weights to fault peaks, with factors reaching 0.19–0.40 for fault areas and remaining below 0.06 for noise components. The optimized high-SNR fault features are finally output in Fig. 13 (d).

This multi-step workflow effectively enhances fault signatures and suppresses interference, providing high-quality purified features for subsequent classification.

2) Implementation of multi-stage feature optimization

First, dynamic threshold segmentation (Fig. 13 (a)). Sliding windows compute local statistical features to generate interval-adaptive thresholds. Low thresholds preserve faint fault signatures, whereas high thresholds filter random noise interference.

Dynamic thresholds achieve a fault recognition accuracy of 96.7% and a false alarm rate of only 3.3%, far outperforming fixed thresholds with a 21.5% misjudgment rate.

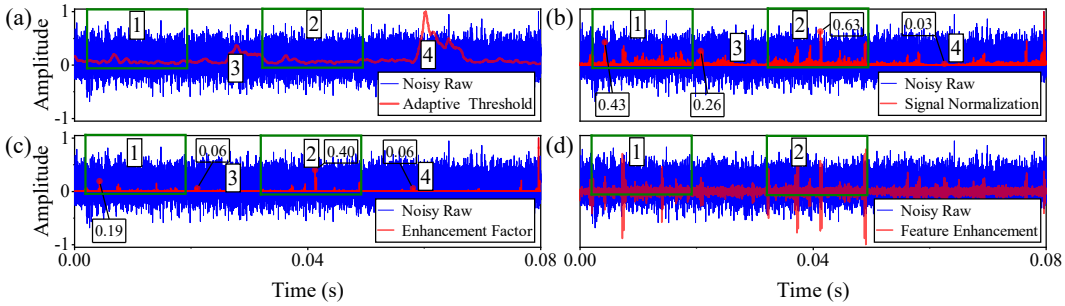


Fig. 13 Comparison of time-domain waveforms of faults in the ATTAE intermediate process, (a)Adaptive Threshold, (b)Signal Normalization, (c)Enhancement Factor, (d)Feature Enhancement

Second, signal normalization (Fig. 13 (b)). This procedure removes amplitude gaps between high-frequency shock pulses and low-frequency modulated components. Normalized amplitudes are confined to 0.43–0.63 for fault zones and suppressed to 0.03–0.26 for noisy regions, eliminating uneven enhancement caused by amplitude discrepancies across frequency bands.

Third, adaptive enhancement (Fig. 13 (c)). The customized enhancement factor simultaneously magnifies weak fault features and suppresses noise, lifting the overall network optimization efficiency by 18.2%. Enhancement factor peaks range from 0.19 to 0.40 over fault areas, compared with merely 0.06 within noise regions.

Finally, temporal mask focusing (Fig. 13 (d)). The temporal attention mask identifies effective fault information and blocks irrelevant noise components. Benefiting from mask screening, the final feature SNR increases from 2.1 dB to 7.2 dB, and low-frequency characteristic frequencies (42.9 Hz, 67.9 Hz) as well as their harmonic components can be clearly extracted.

The above four processing links gradually convert noisy original signals into high-SNR fault features, solving the problem of blurred fault characteristics under strong external interference.

3) Bidirectional collaborative logic between the two modules

The wavelet convolution layer and ATTAE are not merely cascaded modules but form a gradient closed-loop optimization system, and their synergistic effect is presented in Fig. 14.

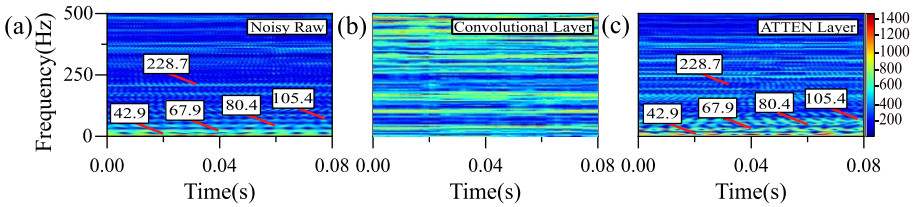


Fig. 14 Comparison of time-frequency domain characteristics of faults learned by AWTA-ELCNN at different levels, (a)Noisy Raw, (b)Convolutional Layer, (c)ATTAE Layer

First, the wavelet convolution captures full-band features ranging from 42.9 Hz to 45570 Hz through adaptive scale learning, whereas effective fault information is still submerged in noise.

Second, ATTAE screens and amplifies fault features, raising the low-frequency fault SNR to 7.2 dB and highlighting hidden harmonic components at 80.4 Hz and 105.4 Hz.

Last, optimized features feed valid gradients back to the convolution layer. The model adjusts the scale value from 80.8 to 81.0 to boost the response to the 67.9 Hz fault characteristic frequency, while gradients corresponding to noise bands are suppressed synchronously.

Essentially, wavelet convolution ensures full-range coverage of fault features, and ATTAE improves feature usability through denoising and feature enhancement. Benefiting from this gradient feedback loop, the model achieves stable performance in frequency drift tracking, noise suppression and weak fault identification.

The quantitative improvements induced by ATTAE are concluded below: the number of scale convergence iterations is reduced by 18%, the intermediate feature SNR increases from 2.1 dB to 5.8 dB, and the final optimized SNR reaches 7.2dB. Moreover, the learned wavelet scales can accurately fit the actual fault characteristic frequencies. Such internal feature improvements explain the high diagnostic accuracy listed in TABLE 4 and the prominent anti-noise capability validated via t-test and Cohen's d metrics in TABLE 5.

5. Conclusion

To address the main challenges of traditional models under strongly time-varying working conditions, including failure to dynamically track time-varying features and difficulty in synergizing lightweight performance, working condition adaptability, and anti-noise capability, the AWTA-ELCNN model is proposed based on the lightweight and basic anti-noise architecture of ELCNN. Through the addition of an adaptive wavelet convolutional layer, based on the linear correlation between fault characteristic frequency and rotational speed, equally spaced initialized wavelets are used to cover a wide frequency range. Dynamic learning of scale factors is employed to match frequency drift, enabling accurate capture of high-frequency impact and low-frequency periodic features. An ATTAE module is constructed to generate dynamic thresholds based on the mean and standard deviation of sliding windows. The threshold is reduced in fault-dense regions to retain details and increased in noise regions to suppress interference, thereby directionally enhancing fault features. Meanwhile, the single-layer topological architecture is strictly inherited, with only wavelet and attention-related parameters added, effectively controlling the growth of parameter quantity and achieving in-depth synergy among strong time-varying feature capture, noise suppression, and lightweight deployment.

Experiments demonstrate that proposed model has been validated on typical time-varying condition datasets such as variable-speed bearings and variable-load gearboxes, exhibiting excellent performance. Its diagnostic accuracy and anti-noise performance are significantly improved compared with comparative models, and the lightweight advantage continues to be prominent, fully meeting the requirements of embedded deployment in rotating machinery. The correlation logic between wavelet scales and fault frequencies constructs an interpretable frequency prior mechanism, providing a solid foundation for the optimization of related technologies.

The current model is based on single-sensor vibration signals and does not involve multimodal data fusion. In future work, fault scenarios and equipment types can be expanded, and feature extraction methods for multi-sensor fusion such as vibration, temperature, and sound pressure can be explored to further improve the model's generalization ability

Acknowledgements and Funding Information

This work was supported by the National Natural Science Foundation of China (Grant No.: 52375412).

Author contributions

Tengji Xia wrote the original draft of the manuscript; Pengfei Pang conceived the research idea and designed the study framework. All authors reviewed the manuscript.

Competing interest statements

The authors declare no competing interests.

Data availability

Ottawa Dataset: http://dx.doi.org/10.17632/v43hmbwxpm.1
MCC5-THU Dataset: https://data.mendeley.com/datasets/p92gj2732w/1

Code Availability

The custom code used in this study contains proprietary information related to ongoing research and potential patent applications. Therefore, the code is not publicly available at present. Data and analysis results supporting the findings are available from the corresponding author upon reasonable request.

References

[1] L. Lima Lopes, L. Travé-Massuyès, C. Jauberthie, and G. Alcalay, "A Review of Fault Diagnosis Techniques Applied to Aircraft Air Data Sensors," in 35th International Conference on Principles of Diagnosis and Resilient Systems (DX 2024), I. Pill, A. Natan, and F. Wotawa, Eds., in Open Access Series in Informatics (OASIcs), vol. 125. Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2024, p. 3:1-3:20. doi: 10.4230/OASIcs.DX.2024.3.

[2] D. Xiao, C. Qin, H. Yu, Y. Huang, C. Liu, and J. Zhang, "Unsupervised machine fault diagnosis for noisy domain adaptation using marginal denoising autoencoder based on acoustic signals," Measurement, vol. 176, p. 109186, May 2021, doi: 10.1016/j.measurement.2021.109186.

[3] M. Hamadache, J. H. Jung, J. Park, and B. D. Youn, "A comprehensive review of artificial intelligence-based approaches for rolling element bearing PHM: shallow and deep learning," JMST Adv., vol. 1, no. 1, pp. 125–151, Jun. 2019, doi: 10.1007/s42791-019-0016-y.

[4] J. Jiao, M. Zhao, J. Lin, and K. Liang, "A comprehensive review on convolutional neural network in machine fault diagnosis," Neurocomputing, vol. 417, pp. 36–63, Dec. 2020, doi: 10.1016/j.neucom.2020.07.088.

[5] M. Rhif, A. Ben Abbes, I. R. Farah, B. Martínez, and Y. Sang, "Wavelet Transform Application for/in Non-Stationary Time-Series Analysis: A Review," Applied Sciences, vol. 9, no. 7, Art. no. 7, Jan. 2019, doi: 10.3390/app9071345.

[6] Q. Zhang and L. Deng, "An Intelligent Fault Diagnosis Method of Rolling Bearings Based on Short-Time Fourier Transform and Convolutional Neural Network," J Fail. Anal. and Preven., vol. 23, no. 2, pp. 795–811, Apr. 2023, doi: 10.1007/s11668-023-01616-9.

[7] R. Yan et al., "Wavelet transform for rotary machine fault diagnosis:10 years revisited," Mechanical Systems and Signal Processing, vol. 200, p. 110545, Oct. 2023, doi: 10.1016/j.ymssp.2023.110545.

[8] Z. Jin, D. Chen, D. He, Y. Sun, and X. Yin, "Bearing Fault Diagnosis Based on VMD and Improved CNN," J Fail. Anal. and Preven., vol. 23, no. 1, pp. 1

65–175, Feb. 2023, doi: 10.1007/s11668-022-01567-7.
[9] A. Mohammad-Alikhani, E. Jamshidpour, S. Dhale, M. Akrami, S. Pardhan, and B. Nahid-Mobarakeh, “Fault Diagnosis of Electric Motors by a Channel-Wise Regulated CNN and Differential of STFT,” IEEE Transactions on Industry Applications, vol. 61, no. 2, pp. 3066–3077, Mar. 2025, doi: 10.1109/TIA.2025.3532556.
[10] M. M. Quamar and A. Nasir, “Review on Fault Diagnosis and Fault-Tolerant Control Scheme for Robotic Manipulators: Recent Advances in AI, Machine Learning, and Digital Twin,” Feb. 05, 2024, arXiv: arXiv:2402.02980. doi: 10.48550/arXiv.2402.02980.
[11] R. Zhu et al., “Research on Vertical SEC Centrifugal Pump Multi-Fault Diagnosis Based on WPT–SVM,” Energies, vol. 16, no. 22, p. 7653, Nov. 2023, doi: 10.3390/en16227653.
[12] Y. Wu, Y. Bai, S. Yang, and C. Li, “Extracting random forest features with improved adaptive particle swarm optimization for industrial robot fault diagnosis,” Measurement, vol. 229, p. 114451, Apr. 2024, doi: 10.1016/j.measurement.2024.114451.
[13] Q. Zhou and J. Tang, “An Interpretable Parallel Spatial CNN-LSTM Architecture for Fault Diagnosis in Rotating Machinery,” IEEE Internet of Things Journal, vol. 11, no. 19, pp. 31730–31744, Oct. 2024, doi: 10.1109/JIOT.2024.3422969.
[14] Y. Xu, Y. Cai, and L. Song, “Latent Fault Detection and Diagnosis for Control Rods Drive Mechanisms in Nuclear Power Reactor Based on GRU-AE,” IEEE Sensors Journal, vol. 23, no. 6, pp. 6018–6026, Mar. 2023, doi: 10.1109/JSEN.2023.3241381.
[15] A. Nedaei, A. Eskandari, J. Milimonfared, B. D. Dimd, U. Cali, and M. Aghaei, “A Bi-Level Line-Line Fault Detection Model for Photovoltaic Arrays Using RBM-Based Automatic Feature Extraction,” in 2023 International Conference on Future Energy Solutions (FES), Jun. 2023, pp. 1–6. doi: 10.1109/FES57669.2023.10183027.
[16] C. Yao et al., “Online Open-Circuit Fault Diagnosis for ANPC Inverters Using Edge-Based Lightweight Two-Dimensional CNN,” IEEE Transactions on Power Electronics, vol. 39, no. 4, pp. 3979–3984, Apr. 2024, doi: 10.1109/TPEL.2024.3351911.
[17] T. Li, “The emerging graph neural networks for intelligent fault diagnostics and prognostics: A guideline and a benchmark study,” Mechanical Systems and Signal Processing, vol. 168, no. 1, p. 108653, 2022.
[18] T.-T. Vo, M.-K. Liu, and C.-L. Hsieh, “Hybrid 1D CNN-RNN Network for Fault Diagnosis in Induction Motors Using Electrical Signals,” in 2023 International Conference on System Science and Engineering (ICSSE), Jul. 2023, pp. 212–217. doi: 10.1109/ICSSE58758.2023.10227168.
[19] Q. Zhou and J. Tang, “An Interpretable Parallel Spatial CNN-LSTM Architecture for Fault Diagnosis in Rotating Machinery,” IEEE Internet Things J., vol. 11, no. 19, pp. 31730–31744, Oct. 2024, doi: 10.1109/JIOT.2024.3422969.
[20] A. Kumar, A. Parey, and P. K. Kankar, “A New Hybrid LSTM-GRU Model for Fault Diagnosis of Polymer Gears Using Vibration Signals,” J. Vib. Eng. Technol., vol. 12, no. 2, pp. 2729–2741, Feb. 2024, doi: 10.1007/s42417-023-01010-7.
[21] Y. Lian, “Residual attention guided vision transformer with acoustic-vibration signal feature fusion for cross-domain fault diagnosis,” Advanced Engineering Informatics, vol. 64, no. 1, pp. 497–535, 2025.
[22] K. Yin, C. Chen, Q. Shen, and J. Deng, “A Lightweight and Rapidly Converging Transformer Based on Separable Linear Self-Attention for Fault Diagnosis,” Meas. Sci. Technol., 2024, doi: 10.1088/1361-6501/ad9f89.
[23] T. Li et al., “WaveletKernelNet: An Interpretable Deep Neural Network for Industrial Intelligent Diagnosis,” IEEE Transactions on Systems, Man, and Cybernetics: Systems, vol. 52, no. 4, pp. 2302–2312, Apr. 2022, doi: 10.1109/TSMC.2020.3048950.
[24] P. Pang, J. Tang, J. Zhang, H. Fang, Y. Hou, and B. Fang, “Research on Fault Diagnosis and Feature Extraction Mechanism of Rotating Machinery Based on WGS-CNN,” IEEE Sensors Journal, vol. 25, no. 13, pp. 24485–24508, Jul. 2025, doi: 10.1109/JSEN.2025.3571866.
[25] B. Wang, “Remaining Useful Life Predictions for Bearings Using Spectrogram and Scalogram-Based Convolutional Neural Networks,” Université d’Ottawa / University of Ottawa, 2023. Accessed: Jul. 05, 2025.
[26] B. Ganguly et al., “Wavelet Kernel-Based Convolutional Neural Network for Localization of Partial Discharge Sources Within a Power Apparatus,” IEEE Transactions on Industrial Informatics, vol. 17, no. 3, pp. 1831–1841, Mar. 2021, doi: 10.1109/TII.2020.2991686.
[27] T. Li et al., “WaveletKernelNet: An Interpretable Deep Neural Network for Industrial Intelligent Diagnosis,” IEEE Transactions on Systems, Man, and Cybernetics: Systems, vol. 52, no. 4, pp. 2302–2312, Apr. 2022, doi: 10.1109/TSMC.2020.3048950.
[28] Y.-J. Huang, A.-H. Liao, D.-Y. Hu, W. Shi, and S.-B. Zheng, “Multi-scale convolutional network with channel attention mechanism for rolling bearing fault diagnosis,” Measurement, vol. 203, p. 111935, Nov. 2022, doi: 10.1016/j.measurement.2022.111935.
[29] P. Pang, J. Tang, J. Luo, M. Chen, H. Yuan, and L. Jiang, “An Explainable and Lightweight Improved 1-D CNN Model for Vibration Signals of Rotating Machinery,” IEEE Sensors Journal, vol. 24, no. 5, pp. 6976–6997, Mar. 2024, doi: 10.1109/JSEN.2023.3327783.
[30] D. D. N. De Silva, H. W. M. K. Vithanage, S. A. Xavier, I. T. S. Piyatilake, and S. Fernando, “Parameterized Wavelets for Convolutional Neural Networks,” in 2020 IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC), Apr. 2020, pp. 170–176. doi: 10.1109/ICARSC49921.2020.9096125.
[31] M.-Q. Tran, M.-K. Liu, Q.-V. Tran, and T.-K. Nguyen, “Effective Fault Diagnosis Based on Wavelet and Convolutional Attention Neural Network for Induction Motors,” IEEE Transactions on Instrumentation and Measurement, vol. 71, pp. 1–13, 2022, doi: 10.1109/TIM.2021.3139706.
[32] Z. Shang, Z. Zhao, and R. Yan, “Denoising Fault-Aware Wavelet Network: A Signal Processing Informed Neural Network for Fault Diagnosis,” Chin. J. Mech. Eng., vol. 36, no. 1, p. 9, Jan. 2023, doi: 10.1186/s10033-023-00838-0.
[33] T. Li, C. Sun, S. Li, Z. Wang, X. Chen, and R. Yan, “Explainable Graph Wavelet Denoising Network for Intelligent Fault Diagnosis,” IEEE Trans. Neural Netw. Learning Syst., vol. 35, no. 6, pp. 8535–8548, Jun. 2024, doi: 10.1109/TNNLS.2022.3230458.
[34] H. Huang and N. Baddour, “Bearing vibration data collected under time-varying rotational speed conditions,” Data in Brief, vol. 21, pp. 1745–1749, Dec. 2018, doi: 10.1016/j.dib.2018.11.019.
[35] S. Chen, Z. Liu, X. He, D. Zou, and D. Zhou, “Multi-mode fault diagnosis datasets of gearbox under variable working conditions,” Data in Brief, vol. 54, p. 110453, Jun. 2024, doi: 10.1016/j.dib.2024.110453.
[36] X. Zhang, C. He, Y. Lu, B. Chen, L. Zhu, and L. Zhang, “Fault diagnosis for small samples based on attention mechanism,” Measurement, vol. 187, p. 110242, Jan. 2022, doi: 10.1016/j.measurement.2021.110242.
[37] G. Yu, P. Wu, Z. Lv, J. Hou, B. Ma, and Y. Han, “Few-Shot Fault Diagnosis Method of Rotating Machinery Using Novel MCGM Based CNN,” IEEE Trans. Ind. Inf., vol. 19, no. 11, pp. 10944–10955, Nov. 2023, doi: 10.1109/TII.2023.3242813.
[38] Wei Ai, “An Improved SE-CNN-Based Method for Rolling Bearing Fault Diagnosis,” Scientific Innovation in Asia, vol. 3, no. 1, p. 1, 2025, doi: 10.12410/sia0303005.
[39] N. Rezazadeh, M. De Oliveira, G. Lamanna, D. Perfetto, and A. De Luca, “WaveCORAL-DCCA: A Scalable Solution for Rotor Fault Diagnosis Across Operational Variabilities,” Electronics, vol. 14, no. 15, p. 3146, Jan. 2025, doi: 10.3390/electronics14153146.
[40] Q. Zhang, Q. Liu, and Q. Ye, “An attention-based temporal convolutional network method for predicting remaining useful life of aero-engine,” Engineering Applications of Artificial Intelligence, vol. 127, p. 107241, Jan. 2024, doi: 10.1016/j.engappai.2023.107241.
[41] H. Wang et al., “A novel transformer-based few-shot learning method for intelligent fault diagnosis with noisy labels under varying working conditions,” Reliability Engineering & System Safety, vol. 251, p. 110400, Nov. 2024, doi: 10.1016/j.ress.2024.110400.
[42] W. Zhou, Z. Feng, Y. F. Xu, X. Wang, and H. Lv, “Empirical Fourier decomposition: An accurate signal decomposition method for nonlinear and non-stationary time series analysis,” Mechanical Systems and Signal Processing, vol. 163, p. 108155, Jan. 2022, doi: 10.1016/j.ymssp.2021.108155.
[43] Y. Miao, B. Zhang, C. Li, J. Lin, and D. Zhang, “Feature Mode Decomposition: New Decomposition Theory for Rotating Machinery Fault Diagnosis,” IEEE Trans. Ind. Electron., vol. 70, no. 2, pp. 1949–1960, Feb. 2023, doi: 10.1109/TIE.2022.3156156