



## OPEN ACCESS

## EDITED BY

Chengxi Zhang,  
Jiangnan University, China

## REVIEWED BY

Peilin Jia,  
Dalian University of Technology, China  
Lanhao Zhao,  
Beijing University of Technology, China

## \*CORRESPONDENCE

Xiaoqiao Qi,  
✉ [qixiaoqiao@126.com](mailto:qixiaoqiao@126.com)

RECEIVED 26 November 2025

REVISED 16 December 2025

ACCEPTED 22 December 2025

PUBLISHED 14 January 2026

## CITATION

Qi X, Yang Y, Han S, Bai G and Fang N (2026) Fault diagnosis of electromechanical systems considering noise suppression and multiscale signal features. *Front. Mech. Eng.* 11:1754564. doi: 10.3389/fmech.2025.1754564

## COPYRIGHT

© 2026 Qi, Yang, Han, Bai and Fang. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](#). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

# Fault diagnosis of electromechanical systems considering noise suppression and multiscale signal features

Xiaoqiao Qi\*, Yuance Yang, Shukui Han, Guangyu Bai and Nanxiang Fang

School of Mechanical and Electrical Engineering, North China Institute of Aerospace Engineering, Langfang, China

**Introduction:** In the electromechanical system, the performance of a direct current brushless motor is determined by its rolling bearings, which play a decisive role in ensuring the safe and smooth operation of the entire system. Thus, fault diagnosis of these bearings is of paramount importance. However, existing methods for diagnosing faults often suffer from low accuracy, particularly under complex noise conditions.

**Methods:** This study proposes an innovative approach to fault diagnosis that enhances the accuracy and robustness of detecting faults in brushless direct current motor rolling bearings. To achieve this goal, this study first employs wavelet threshold denoising to suppress noise in motor current signals and performs multiscale feature fusion. Additionally, a fault diagnosis method is developed by integrating a convolutional attention mechanism.

**Results:** The outcomes indicated that the proposed diagnostic method achieved a recall rate of 90.89% and a precision rate of 98.69%, both higher than those of the comparative methods. The suggested approach outperformed the comparison methods in all four fault categories, with diagnostic accuracy rates of 99.4%, 98.9%, 98.8%, and 99.3%.

**Discussion:** The findings of the experiments reveal that the proposed diagnostic method can effectively identify faults in rolling bearings of brushless direct current motors, providing a theoretical foundation for research in the field of electromechanical system fault diagnosis. The contributions of this research are in three aspects. First, the BLDCM rolling bearing current signal is reconstructed using a multiscale feature and wavelet threshold denoising. This significantly improves the signal quality and ability to extract fault features. Second, CBAM, residual network and Swin Transformer encoder are integrated into the fault diagnosis model. Compared with the existing methods, higher accuracy and precision are achieved. This study finally provides a solid theoretical foundation for further research in the field of electromechanical system fault diagnosis, particularly for BLDCM rolling bearing fault diagnosis under complex noise conditions.

## KEYWORDS

brushless direct current motors, convolutional attention, electromechanical systems, multiscale features, noise suppression, rolling bearings, wavelet threshold denoising

## 1 Background

Mechatronics is a technology system that integrates multiple disciplines such as mechanics and electronics (Shang et al., 2025; Zheng et al., 2025). The rolling bearings of brushless direct current motor (BLDCM) play a decisive role in the safe and smooth operation of electromechanical systems (EMSSs), and its fault diagnosis (FD) is extremely important. Many scholars have conducted relevant research. For example, Lu used a combination of literature analysis and case studies to conduct analysis in an effort to solve the problem of low efficiency in fault location of EMSSs (Lu, 2024). Zhao et al. developed a current assisted vibration fusion network to address the issues of low accuracy and low precision in current diagnosis methods for electromechanical drive systems (Zhao et al., 2024). Zhang et al. proposed a FD framework based on ensemble learning in their study on the low accuracy of actuator fault methods in aviation EMSSs (Zhang et al., 2024). Zhang et al. combined principal component analysis and belief rule library to establish a FD model for EMSSs, which was difficult to extract features caused by excessive noise and leads to poor accuracy in FD (Zhang et al., 2025). Zhao proposed a fault detection model that combined wavelet energy packet and improved support vector machine in his research on the difficulty of detecting faults in EMSSs (Zhao, 2023).

The accurate diagnosis of BLDCM rolling bearing faults determines whether the electromechanical drive system operates safely and smoothly. Wavelet threshold denoising (WTD) is a signal processing technique that has the advantages of multiscale analysis capability and noise whitening. It has been extensively utilized in the domains of EMS monitoring and audio processing (Das and Sahana, 2025). Convolutional block attention module (CBAM) is a deep learning technique that has advantages such as cross modal adaptation and strong flexibility, and has been widely used in fields such as feature extraction (Xu et al., 2023). Multiple experts have conducted relevant research. For example, Wang et al. constructed a compressed sensing reconstruction framework based on wavelet domain consistency constraints to address the issue of difficult noise removal (Wang et al., 2024a). To solve the problem of low accuracy brought on by noise in the existing stock price prediction systems, Singh et al. developed a technique based on discrete wavelet denoising (Singh et al., 2025). Sahoo et al. introduced a general wavelet selection method based on the sparsity of detail components in the wavelet domain (Sahoo et al., 2024). Bhuyan et al. combined residual networks with CBAM to construct a tea disease identification model. The model outperformed the comparison model, according to the comparative experimental data (Bhuyan et al., 2024). To solve the problem of low detection effectiveness in ground penetrating radar, Wang et al. built a radar detection system by combining CBAM with YOLOv8. The findings showed that the suggested system had higher detection efficiency compared to the original system (Wang et al., 2024b).

The above research results indicate that there are few methods for FD of BLDCM rolling bearings in EMSSs under complex noise conditions, and there is a problem of low accuracy. Therefore, initially, wavelet denoising techniques were used to reconstruct multiscale feature parameters (WP) from the biphasic current (BC) signals of BLDCM RB, in order to reduce noise and

improve fault feature extraction capability. Afterwards, the CBAM, residual network, and Swin Transformer (ST) encoder were integrated together to create an FD model for EMSS. To improve the accuracy of RB FD in EMSSs using BLDCMs, this model utilizes multiscale feature fusion with signal and noise suppression. This study is innovative as it uses WTD for multiscale feature algebraic reconstruction of BC signals from BLDCM RB, while combining CBAM, residual network, and ST encoder. The purpose of this method is to provide theoretical basis for EMSS FD research.

## 2 Methods and materials

### 2.1 Wavelet-based current signal noise suppression and multiscale feature fusion

Because of their high automation efficiency and precision control, EMSSs have been widely used in automotive and aerospace applications in recent years. FD for BLDCM RBs is especially important as a key component powering these systems. However, current diagnostic methods suffer from low accuracy due to noise interference. To suppress noise and improve fault feature extraction capabilities, this study uses a WTD method to recover multiscale feature parameters from BC waveforms. Before applying WTD to suppress noise in BLDCM RB current signals, it is essential to understand BC and its vector and algebraic reconstruction processes. The process is illustrated in Figure 1 (Sulistyo et al., 2025; Zangana and Mustafa, 2024).

In Figure 1a, the BC is supplied by two independent alternating current (AC) power sources, namely, Phase 1 and Phase 2. Phase 1 and 2 share the same frequency but have a phase difference (PD) of 90°, causing their peaks to alternate in time and thus forming a BC. Equation 1 can be used to express the BC.

$$I = u + \lambda v \quad (1)$$

In Equation 1,  $I$  represents the magnitude of the BC in complex form.  $u$  denotes the magnitude of BC 1, with a phase angle of 0°.  $v$  denotes the magnitude of BC 2, with a phase angle of 90°.  $\lambda$  denotes the imaginary unit. To simplify the analysis and calculation of AC circuits, the magnitude and phase of BCs are typically represented by vectors in the complex plane, as shown in Equation 2.

$$\text{Vector} = \sqrt{u^2 + v^2} + (\arctan(u/v))\lambda \quad (2)$$

In Equation 2,  $\sqrt{u^2 + v^2}$  represents the vector magnitude.  $(\arctan(u/v))\lambda$  denotes the phase angle. Vector indicates the vector representation. Combining Figure 1a with Equation 2 reveals that under normal conditions, the PD between the two AC currents is 90°. If a circuit fault occurs, both the PD and current magnitude will change. Therefore, by reconstructing the current signal parameters, the fault characteristics of the data can be highlighted. Current signal parameter reconstruction not only involves changes in the magnitude of current 1 and current 2 through addition, subtraction, and multiplication operations, but also includes calculating the phase angle changes of current 1 and current 2 to diagnose faults in RBs. Figure 1b provides an illustration of the procedure. First, the BC is vectorized, and its

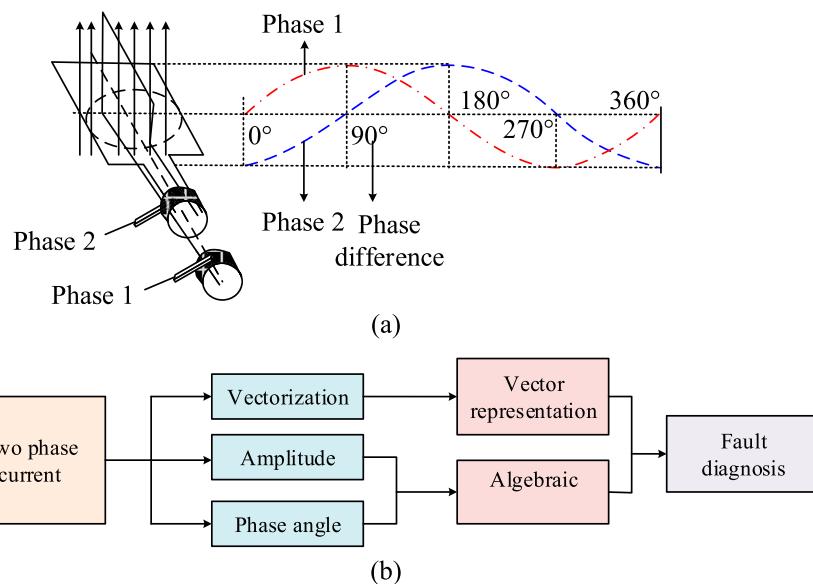


FIGURE 1

The BC and its vector and algebraic reconstruction process. (a) Two phase current. (b) Algebraic reconstruction and vector calculation process of current signal.

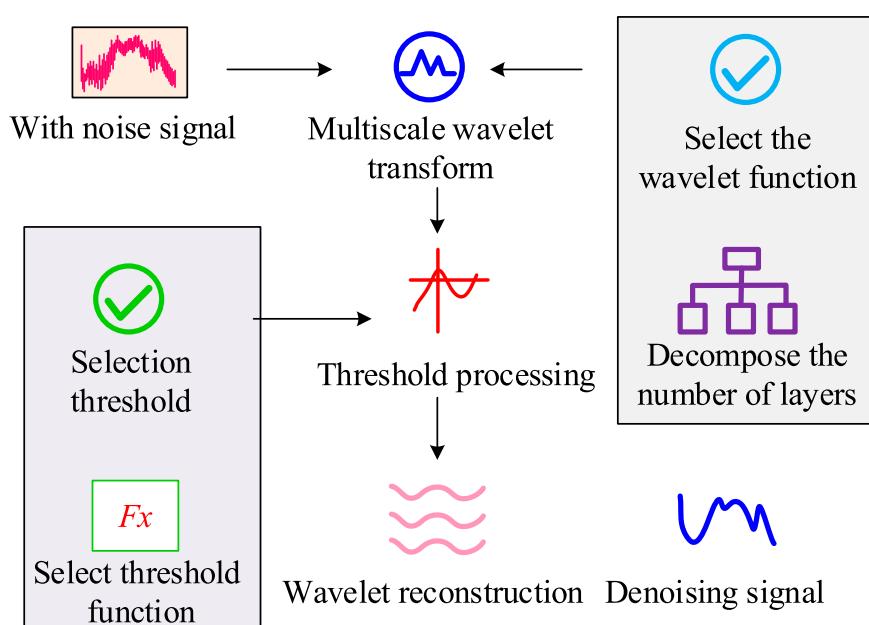
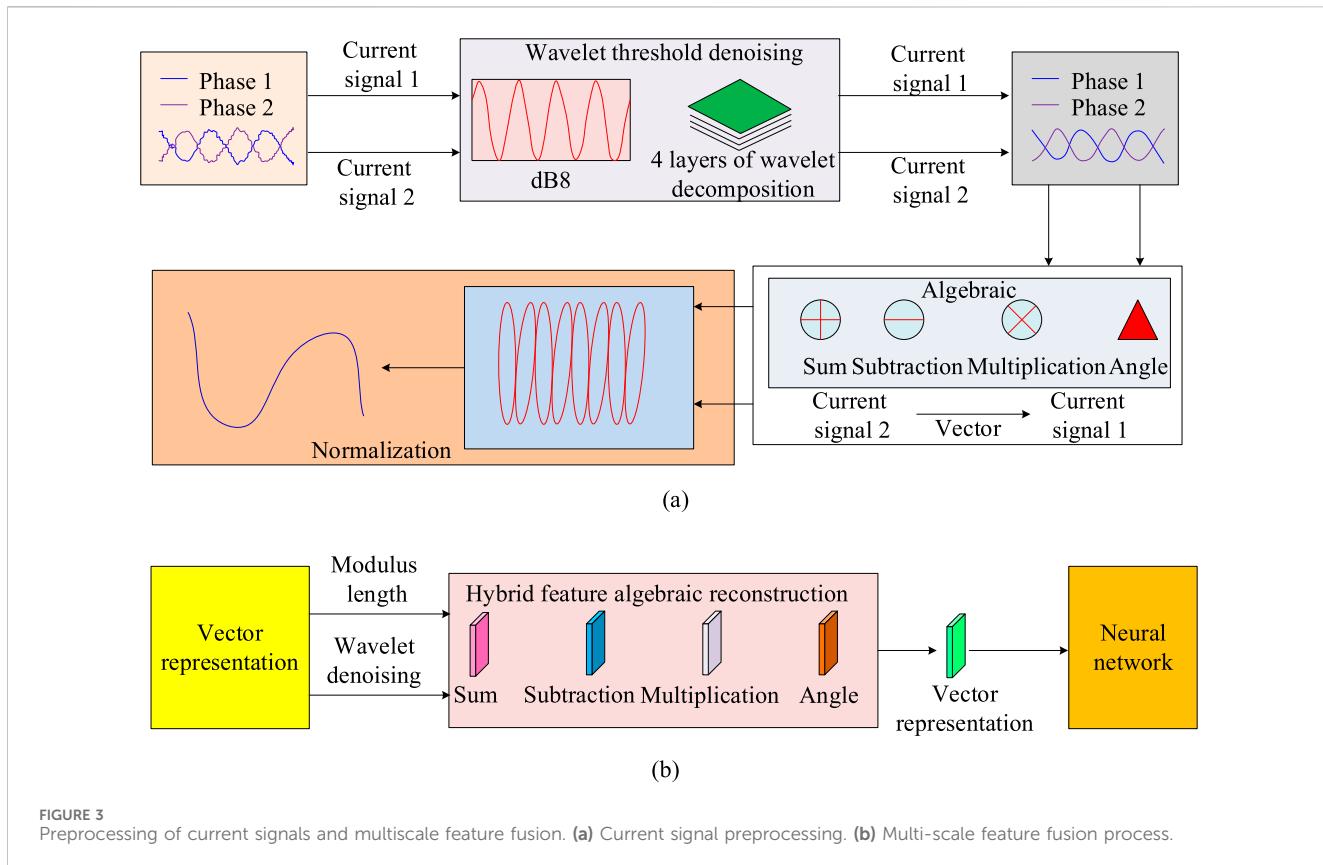


FIGURE 2  
Process of WTD method.

amplitude and phase angle are calculated. Subsequently, vector representation and algebraic reconstruction are performed. Finally, prominent features are obtained to diagnose faults through algorithms. To enhance the quality of algebraic reconstruction, this study combines relevant literature and employs a WTD method for algebraic reconstruction of the BC signal from BLDCM RBs. The WTD method is a signal processing technique based on wavelet transform (WT). It is extensively used in

domains like audio processing and EMS monitoring, where it uses threshold processing to isolate noise from the actual signal. Figure 1 depicts the WTD procedure.

In Figure 2, the WTD process begins by inputting the original noisy signal and performing WTs at multiple scales. During the multiscale WT, an appropriate wavelet basis function (WBF) pair is adopted to decompose the noisy original signal (OS), yielding the quantity of decomposition levels (DLs). Next, the threshold and



**FIGURE 3**  
Preprocessing of current signals and multiscale feature fusion. **(a)** Current signal preprocessing. **(b)** Multi-scale feature fusion process.

function for wavelet threshold processing are chosen and determined. Subsequently, reconstruction is performed via the inverse WT. Finally, the denoised current signal is obtained. The WT process can be expressed by [Equation 3](#).

$$y(t) = \sum_{j=1}^J \sum_{k=1}^{2^{j-1}} d_{j,k} \varphi_{j,k}(t) + \sum_{k=1}^{N-1} \omega_k \varphi_{J,k}(t) \quad (3)$$

In [Equation 3](#),  $j$  denotes the scale,  $k$  denotes the frequency.  $d_{j,k}$  represents the wavelet coefficient with scale  $j$  and frequency  $k$ .  $\varphi_{j,k}(t)$  denotes the WBF.  $y(t)$  represents the OS.  $\omega_k$  denotes the high-frequency noise coefficient.  $J$  is the quantity of scales in the wavelet decomposition.  $N$  is the quantity of samples in the high-frequency component obtained from the WT. Threshold processing  $d'_{j,k}$  is shown in [Equation 4](#).

$$d'_{j,k} = \begin{cases} d_{j,k}, & |d_{j,k}| > a_j \\ 0, & |d_{j,k}| \leq a_j \end{cases} \quad (4)$$

In [Equation 4](#),  $a_j$  represents the threshold. To ensure effective noise suppression, the study combines relevant literature and multiple experiments to ultimately determine the WBF as Db8 and the DL as 4. Therefore, the study employs the WTD method to suppress noise and fuse multiscale features in the BC signal of the BLDCM RB. The preprocessing and multiscale feature fusion of the current signal are illustrated in [Figure 3](#).

[Figure 3a](#) illustrates the preprocessing steps for the current signal. First, noise suppression is achieved using a four-level wavelet decomposition with a WTD method to extract the signal's approximate information curve. Subsequently, algebraic

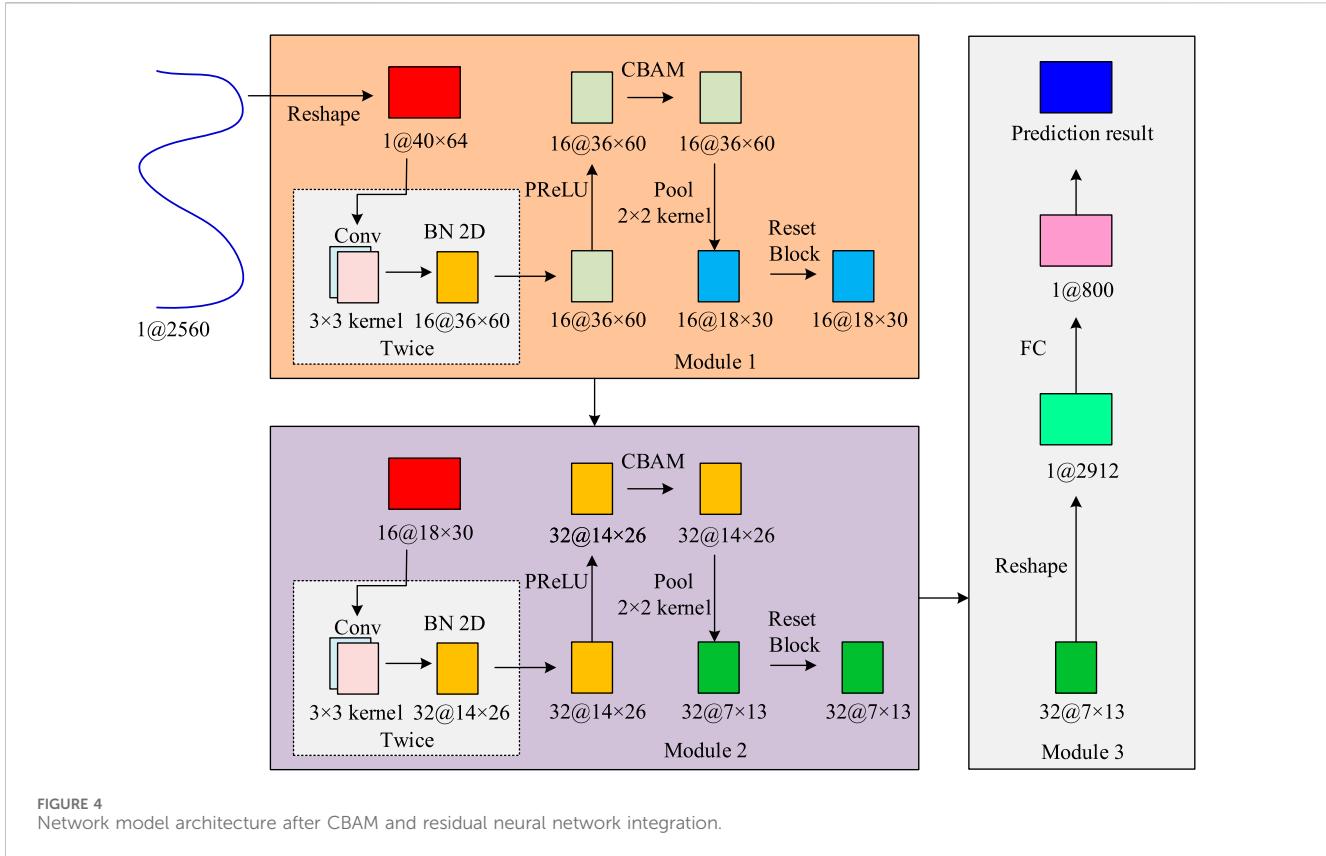
operations are performed on the signal to obtain four types of current signal features. Finally, the signal undergoes normalization using the maximum-minimum normalization method to yield the final current signal vector representation. This method is expressed by [Equation 5](#).

$$x_{nor} = \frac{2^{*}x - x_{min}}{(x_{max} - x_{min})} - 1 \quad (5)$$

In [Equation 5](#),  $x_{min}$  and  $x_{max}$  represent the lowest and highest values, respectively.  $x_{nor}$  denotes the normalized value.  $x$  is the original data value. Thus, after normalization via the WTD method, the current signal undergoes multiscale feature fusion. [Figure 3b](#) illustrates the multiscale feature fusion process for the current signal. First, the current signal is converted into a vector representation, and its magnitude is extracted. Next, wavelet denoising is applied to reduce noise interference. Specifically, algebraic reconstruction methods, including addition, subtraction, multiplication, and angle calculations, are employed to extract and fuse current signal features across four scales. Ultimately, a neural network receives these processed feature vectors for additional examination and diagnosis.

## 2.2 FD method based on CBAM, noise suppression, and multiscale feature fusion

After performing multiscale feature fusion on the current signals of RBs in EMSs using BLDCM through WTD, a hybrid FD method combining CBAM, residual networks, and ST encoders is studied



**FIGURE 4**  
Network model architecture after CBAM and residual neural network integration.

and designed. The channel attention module (CAM) and the spatial attention module (SAM) make up the lightweight attention mechanism (AM) module known as CBAM. Global average pooling (GAP) and global max pooling (GMP) are used by the CAM to obtain global information for every channel. A multilayer perceptron (MLP) is then used to create CA weights. The CAM is expressed by [Equation 6](#).

$$Mc(F) = \delta\{MLP[\text{Avgpool}(F) + MLP[\text{Maxpool}(F)]]\} \quad (6)$$

In [Equation 6](#),  $F$  means the input feature map.  $\text{Avgpool}(F)$  and  $\text{Maxpool}(F)$  mean GAP and max pooling.  $MLP$  displays the MLP.  $\delta$  denotes the *Sigmoid* activation function (AF).  $Mc(F)$  denotes the output processed by the CAM. The SAM first obtains weights for each channel through the output of the CAM. [Equation 7](#) illustrates how these weights are subsequently transformed into SA weights by using GAP and GMP.

$$Ms(F) = \delta\{f^{7*7}[\text{Avgpool}(F); \text{Maxpool}(F)]\} \quad (7)$$

In [Equation 7](#),  $Ms(F)$  means the output processed by the SAM. The study incorporates the CBAM module into a residual neural network to take advantage of the AM's benefits and its strong feature learning power in an attempt to further enhance the model's feature representation capability and DA. [Figure 4](#) displays the network architecture following the integration of the residual neural network and CBAM.

In [Figure 4](#), BN denotes batch normalization, Conv represents the convolutional layer (CL), and FC signifies the fully connected layer (FCL). The notation  $1@40 \times 64$  indicates that the number 1 before the @ denotes the number of input channels, while the

number after the @ represents the input dimensions. In Module 1, the CL, the BN layer, and the first execution of AF PReLU all yield an output dimension of  $38 \times 62$ , with the CL employing 8 convolutional kernels (CKs). The output dimensions for the second execution are both  $36 \times 60$ , with 16 CKs in the CL. Both the CA and SA layers have an output size of  $36 \times 60$ , with kernel sizes of  $1 \times 1$  and  $7 \times 7$  respectively, and 16 and 1 kernels respectively. The max pooling layer and residual blocks both have an output size of  $18 \times 30$ , with kernel sizes of  $2 \times 2$  and  $3 \times 3$  respectively, and 16 kernels each. In Module 2, the initial execution sets the output dimensions of the CL, BN layer, and PReLU output layer to  $16 \times 28$ , with 24 CKs in the CL. The subsequent execution sets all output dimensions to  $14 \times 26$ , with 32 CKs in the CL. Both the CA and SA outputs maintain a dimension of  $14 \times 26$ , with unchanged kernel size, employing 32 and 1 CKs respectively. The maximum pooling layer and residual blocks both employ 32 CKs, with an output size of  $7 \times 13$ . The FCL in Module 3 has an output size of 800. This architecture extracts input features by utilizing Conv and BN, PReLU, pooling layers, and residual blocks. CBAM is employed to enhance feature channels and SA. After multi-layer processing, classification is performed via FC layers to output the final results. Due to the multiscale nature of fault features, the DA of models incorporating CBAM still falls short of requirements. In light of this, the study introduces the encoder module of the ST to improve DA. [Figure 5](#) illustrates an illustration of the ST model's encoder module.

[Figure 5](#) illustrates the processing flow of the ST model's encoder module. First, the input data is segmented into 160 blocks, each containing a specific number of features. Next, the processed blocks

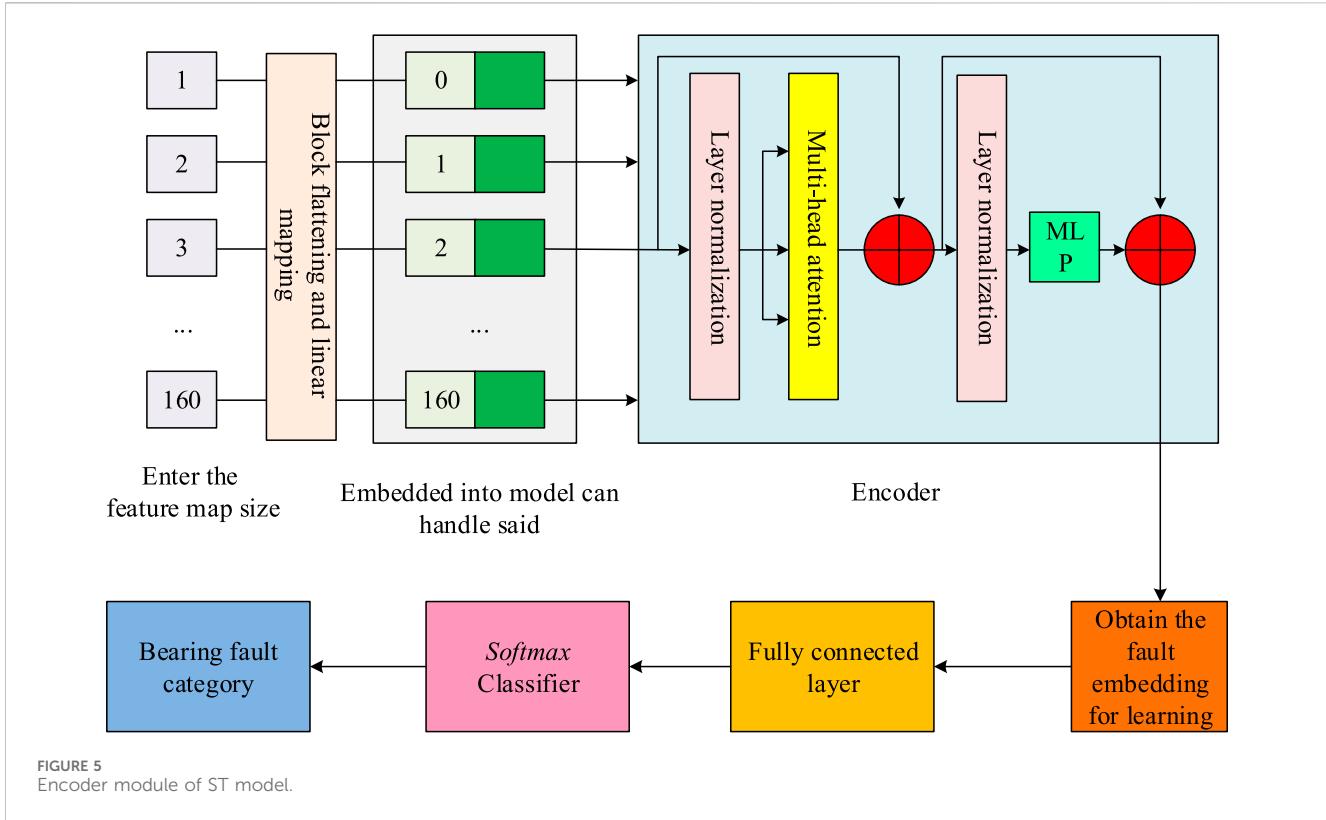


FIGURE 5  
Encoder module of ST model.

are embedded into the model. This includes positional embeddings and fault type embeddings. Positional embeddings represent the sequential relationships among each block. Fault type embeddings capture the structure and categories of different fault data. Subsequently, the data enters the encoder. The encoder normalizes the input data through layer normalization to stabilize and accelerate the training process. It utilizes multi-head attention (MHA) to capture different features within the input data. Simultaneously, it employs a multi-layer perceptron for further data processing, adding residual connections between each sub-layer to prevent gradient explosion. Subsequently, it outputs the fault features and categories learned from the data. Finally, the features and AF A are classified through a FCL to output the fault category. Among these, MHA serves as a crucial component of the ST model's encoder module. MHA excels at capturing complex dependencies between different positions, enabling the ST model to perform exceptionally well when processing long sequence data. The operation of MHA proceeds as follows: First, the input matrix  $X \in R^{n \times d}$  is transformed. Among them,  $d$  denotes the input feature dimension, and  $n$  represents the sequence length. Through linear transformation (LT), the query  $Q$ , key  $K$ , and  $V$  are obtained, as shown in Equation 8.

$$\begin{cases} Q_i = XW_i^Q \\ K_i = XW_i^K \\ V_i = XW_i^V \end{cases} \quad (8)$$

In Equation 8,  $W_i^Q$ ,  $W_i^K$ , and  $W_i^V$  represent the weight matrices for the  $i$  th head. Next, the attention score  $Attention_i(Q_i, K_i, V_i)$  for the  $i$  th head is computed as shown in Equation 9.

$$Attention_i(Q_i, K_i, V_i) = \text{Softmax}\left(\frac{Q_i K_i^T}{\sqrt{d_k}}\right) V_i \quad (9)$$

In Equation 9,  $d_k$  is the scaling factor. Then, all heads are concatenated as shown in Equation 10.

$$\begin{aligned} \text{Concat}(&Attention_1, Attention_2, \dots, Attention_h) \\ &= [Attention_1, Attention_2, \dots, Attention_h] \end{aligned} \quad (10)$$

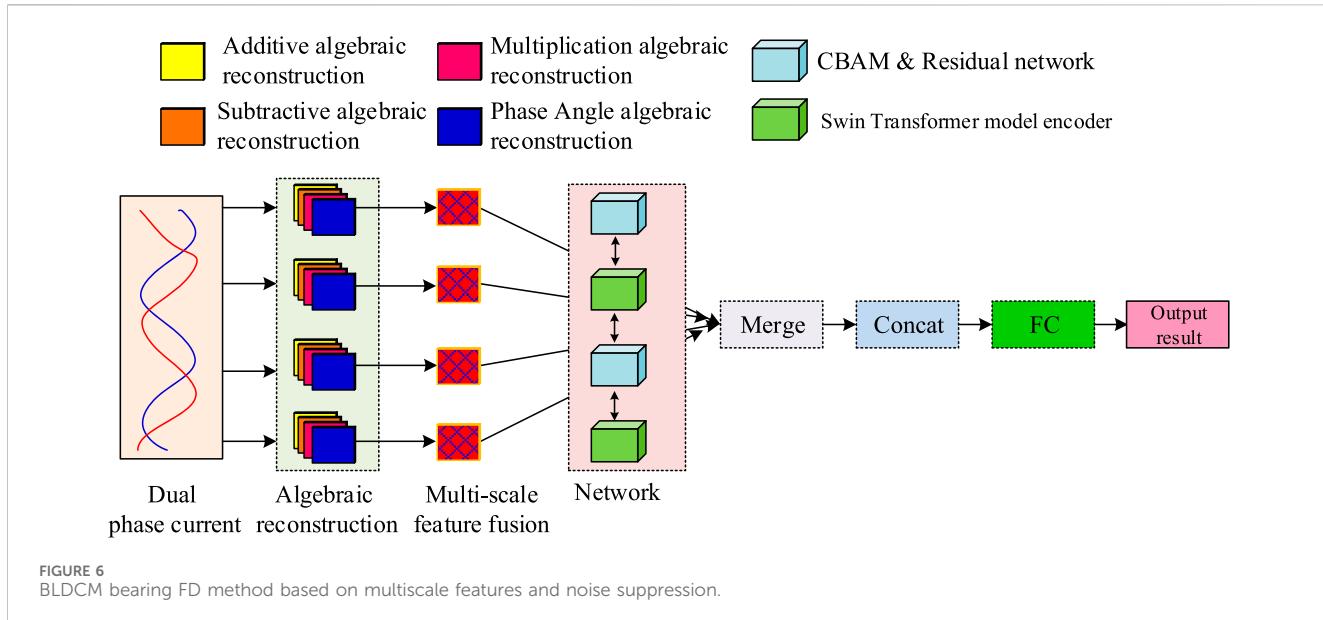
In Equation 10,  $h$  means the number of heads. Finally, a LT yields the final MHA output  $MultiHead(Q, K, V)$ , as expressed in Equation 11.

$$MultiHead(Q, K, V) = \text{Concat}(Attention_1, Attention_2, \dots, Attention_h)W^o \quad (11)$$

In Equation 11,  $W^o$  represents the weight matrix. Furthermore, the study uses a cross-loss entropy function, as indicated by Equation 12, to improve the accuracy of the model.

$$\begin{cases} CEL_x = \frac{1}{N} \sum_j -[y_j * \log(p_j) + (1 - y_j) * \log(1 - p_j)] \\ CEL_y = -\frac{1}{N} \sum_j \sum_{c=1}^M y_{ic} \log(p_{ic}) \end{cases} \quad (12)$$

In Equation 12,  $CEL_x$  and  $CEL_y$  represent the cross-entropy losses for input feature  $x$  and output feature  $y$ .  $N$  means the total quantity of samples.  $M$  means the total quantity of classes.  $p_j$  means the probability that a sample is positive.  $y_{ic}$  means the true label (TL) of the  $c$  th class in the  $j$  th sample.  $p_{ic}$  means the probability that the



$j$  th sample belongs to the  $c$  th category.  $y_j$  represents the TL of the  $j$  th sample. Ultimately, based on the above, this study constructs a BLDCM bearing FD method incorporating multiscale features and noise suppression. This method is illustrated in Figure 6.

Figure 6 illustrates the process of this method. First, the BC signal is reconstructed to obtain comprehensive signal features. Next, preprocessing is performed using WTD, followed by further extraction of multiscale features through algebraic reconstruction methods such as addition, multiplication, subtraction, and phase angle. These multiscale features undergo feature fusion via the CBAM and residual network modules, as well as the encoder of the ST model. Finally, the fused features undergo merging and concatenation before being processed through a FCL to output the BLDCM bearing FD results for the EMS.

### 3 Results and analysis

#### 3.1 Performance analysis of RB FD methods for BLDCM EMSs

After establishing the BLDCM RB FD method for EMSs based on noise suppression and multiscale analysis, a comparative analysis of its performance is conducted. The comparison algorithms are convolutional neural network-long short-term memory (CNN-LSTM), variational mode decomposition-genetic algorithm-support vector machine (VMD-GA-SVM), and variational mode decomposition-continuous WT-convolutional neural network (VMD-CWT-CNN). Model parameter settings employ stochastic gradient descent as the optimizer with a learning rate of 0.001 and momentum of 0.9. The ST features 4 attention heads, a window size of 5, 5 classification heads, 5 input channels, and an input feature dimension of 160. The cross-entropy loss function is utilized. The parameter settings for CBAM and the residual module follow consistent methodologies. Data is sourced from the PU dataset at the University of Paderborn, Germany. The current signal in the PU

TABLE 1 Experimental environment configuration.

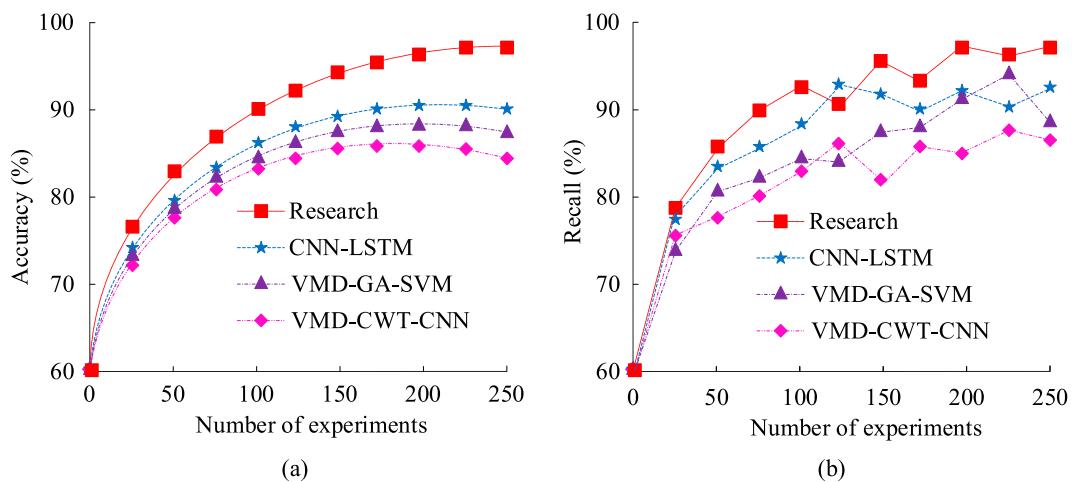
Parameter names	Parameter
Data analysis software	Spss24.0
Operating system	Windows 10 64
Matlab version	Matlab 2022a
Hard disk capacity	500 GB
Internal memory	32 GB
Main frequency	5.8 GHz
Processor	Intel core i9-13900K

dataset is a BC, a type of AC consisting of two sinusoidal currents that are  $90^\circ$  out of phase with each other. This is consistent with the BC signal of the rolling bearing proposed in the research. A total of 3,478 data points are selected, with current signals acquired over 4 s at a sampling frequency of 64 kHz. With 300 training iterations, the dataset consists of an 80% training set and a 20% test set. Accuracy, recall, and F1 score are evaluation metrics. Table 1 provides specifics on the experimental setup.

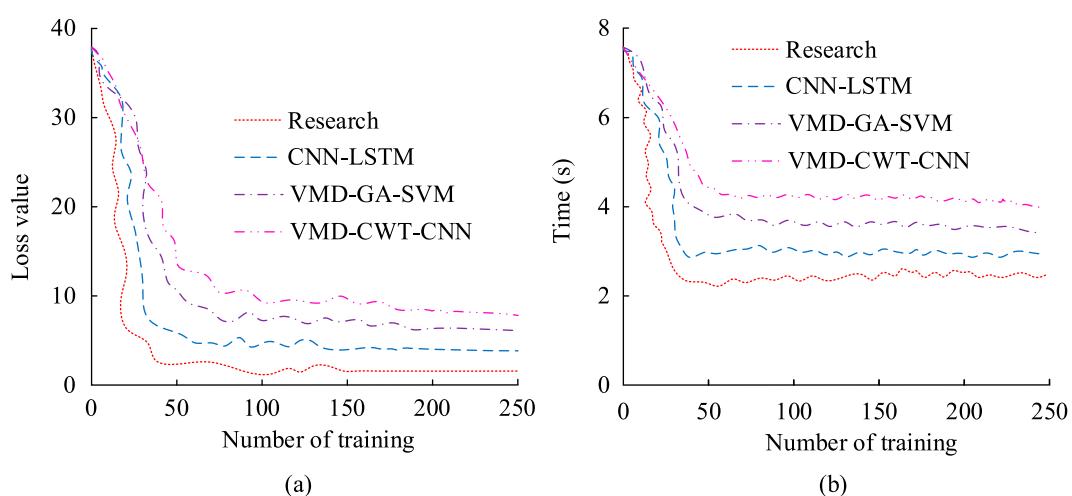
The study initially compares the accuracy and recall rates (RRs) of each approach in the previously specified context. The experimental results are displayed in Figure 7.

In Figure 7a, the suggested FD method achieves the highest accuracy of 98.97%. The accuracy of CNN-LSTM is 90.89%, VMD-GA-SVM is 88.26%, and VMD-CWT-CNN is 85.23%. In Figure 7b, the RRs of the proposed FD method, CNN-LSTM, VMD-GA-SVM, and VMD-CWT-CNN are 98.69%, 93.12%, 89.34%, and 89.15%. The suggested FD approach outperforms the others in terms of RR. The results presented above indicate that the recommended FD technique has the best accuracy and recall. The comparison of loss values and running times (RTs) among the methods is shown in Figure 8.

In Figure 8a, the suggested FD method achieves the earliest convergence in the loss curve, with a loss value of 0.74,



**FIGURE 7**  
Comparison of recall rate and accuracy rate. **(a)** Accuracy rate comparison results. **(b)** Recall rate comparison results.



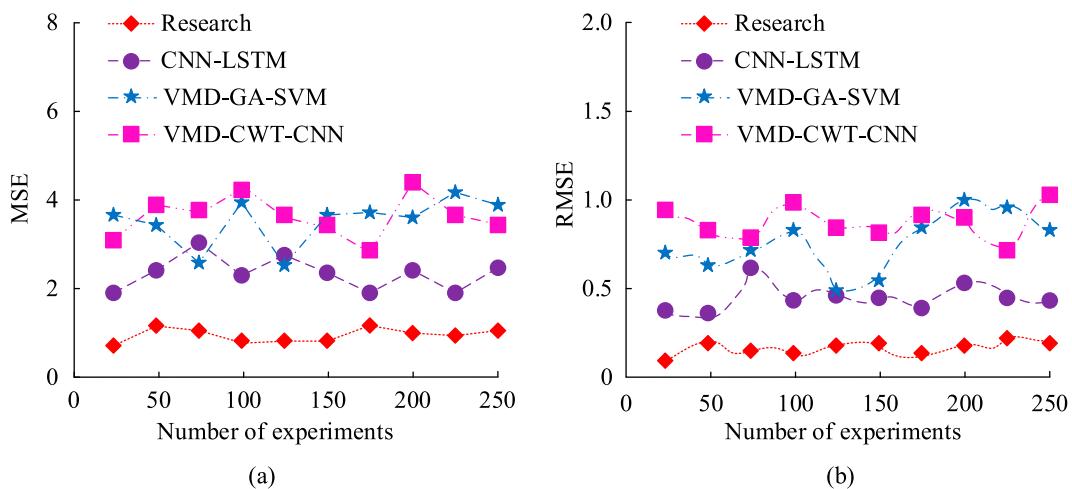
**FIGURE 8**  
Comparison results of loss values and running time. **(a)** Loss value comparison results. **(b)** Recall rate comparison results.

significantly lower than the 4.76 of CNN-LSTM, 8.27 of VMD-GA-SVM, and 10.62 of VMD-CWT-CNN. In Figure 8b, the average RTs for the proposed FD method, CNN-LSTM, VMD-GA-SVM, and VMD-CWT-CNN are 2.38s, 3.16s, 3.98s, and 4.48s, respectively. Among these, the proposed method exhibits the shortest average RT. In summary, from the perspectives of loss values and average RT, the proposed FD method outperforms the comparison methods. Figure 9 displays each method's mean square error (MSE) and root mean square error (RMSE) outcomes.

In Figure 9a, the average MSE values for the proposed FD method, CNN-LSTM, VMD-GA-SVM, and VMD-CWT-CNN are 1.18, 2.46, 3.74, and 4.02, respectively. Among these, the proposed FD method exhibits the lowest average MSE. In Figure 9b, the proposed FD method reaches an average RMSE of 0.31, significantly

lower than the 0.46 for CNN-LSTM, 0.63 for VMD-GA-SVM, and 0.87 for VMD-CWT-CNN. Among these, the proposed method exhibits the lowest average RMSE. In summary, in terms of both RMSE and MSE values, the recommended FD technique outperforms the reference approaches. To verify which module made the highest contribution to the model, the study conducts an ablation experiment on it. The experimental results are shown in Table 2.

In Table 2, with the removal of the model module, its F1 score and AUC value decrease rapidly. When the CBAM and ST modules are removed, the F1 score drops from 98.62% to 82.14%, and the AUC value drops from 0.986 to 0.836. Among them, when the ST module is removed, the decrease intervals of F1 score and AUC value are the largest. The above results indicate that the ST module plays a key role in the model.



**FIGURE 9**  
The MSE and RMSE comparison. **(a)** MSE comparison results. **(b)** RMSE comparison results.

**TABLE 2** Results of ablation experiment.

Model variant	F1 score	AUC value
Complete model	98.62%	0.986
No CBAM	97.03%	0.976
No residual network	96.84%	0.967
No swin transformer	90.13%	0.912
No CBAM and residual network	89.26%	0.884
No CBAM and swin transformer	82.14%	0.836

### 3.2 Diagnostic effect analysis

After verifying the performance of the suggested FD method, a comparative analysis of its diagnostic effectiveness is carried out. A BLDCM RB from an automotive EMS is selected for FD. The data sources are divided into two parts. Some of the signal information comes from the actual collected signals stored when the BLDCM rolling bearing malfunctions and is repaired. These signals provide the original data under the fault state for the research. To further supplement the signal characteristics in the fault state. Another part of the data comes from the BLDCM that is under repair. During the signal acquisition process, current sensors are used to take measurements, and a low-pass filter is employed to remove high-frequency noise from the current signal while retaining useful information related to FD. Finally, the filtered signal is converted into a digital signal with a frequency of 64 kHz. Diagnosis is performed for four fault categories: normal state (A), inner ring fault (IRF) (B), outer ring fault (ORF) (C), and combined inner and outer ring faults (CIORF) (D). The diagnostic results for each method are shown in Figure 10.

In Figure 10a, the proposed FD method achieves DA rates of 99.4%, 98.9%, 98.8%, and 99.3% for normal state, IRF, ORF, and CIORF, respectively. These values surpass the 86.4%, 88.3%, 85.2%, and 90.4% achieved by the CNN-LSTM approach in Figure 10b.

Moreover, it outperforms the VMD-GA-SVM method in Figure 10c (84.8%, 86.2%, 83.8%, 82.7%) and the VMD-CWT-CNN method in Figure 10d (85.7%, 83.7%, 83.6%, 89.7%). In conclusion, the suggested FD approach performs better on each of the four FD accuracy metrics. The comparison results for precision, area under the curve (AUC) values, and central processing unit (CPU) utilization among the various methods are presented in Table 3.

As shown in Table 3, the proposed diagnostic method achieves a precision rate of 99.68% and an AUC value of 0.986. These results are significantly higher than those of CNN-LSTM (90.08% and 0.874), VMD-GA-SVM (86.78% and 0.882), and VMD-CWT-CNN (89.23% and 0.903). The proposed diagnostic method exhibits the lowest CPU utilization at 42.14% among all approaches. These findings reveal that the suggested FD approach performs best in terms of CPU usage, accuracy, and AUC value. The Transformer + ResNet hybrid model proposed in the research shows better convergence speed and CPU occupancy rate compared with the more lightweight CNN-LSTM model. The causes of this phenomenon can be attributed to several key factors. First, this model adopts an efficient self-attention mechanism and residual connection. The self-attention mechanism allows the model to swiftly identify long-distance dependencies in the data, which is essential for comprehending intricate data patterns. Residual connections help alleviate the vanishing gradient problem, which is a common challenge in deep learning models, especially when dealing with deep networks. These architectural advantages jointly promote the improvement of the model's efficiency during the training process. Second, an adaptive learning rate and a mixed-precision training strategy were adopted in the model training. The adaptive learning rate can be dynamically adjusted according to the training progress of the model, thus converging rapidly in the early stage of training and fine-tuning in the later stage of training to improve accuracy. Mixed-precision training combines single- and half-precision computations to reduce memory usage and computational requirements while maintaining model accuracy. This further accelerates convergence and optimizes CPU usage. In addition, the dataset has undergone excellent preprocessing to ensure the quality and diversity of the data.

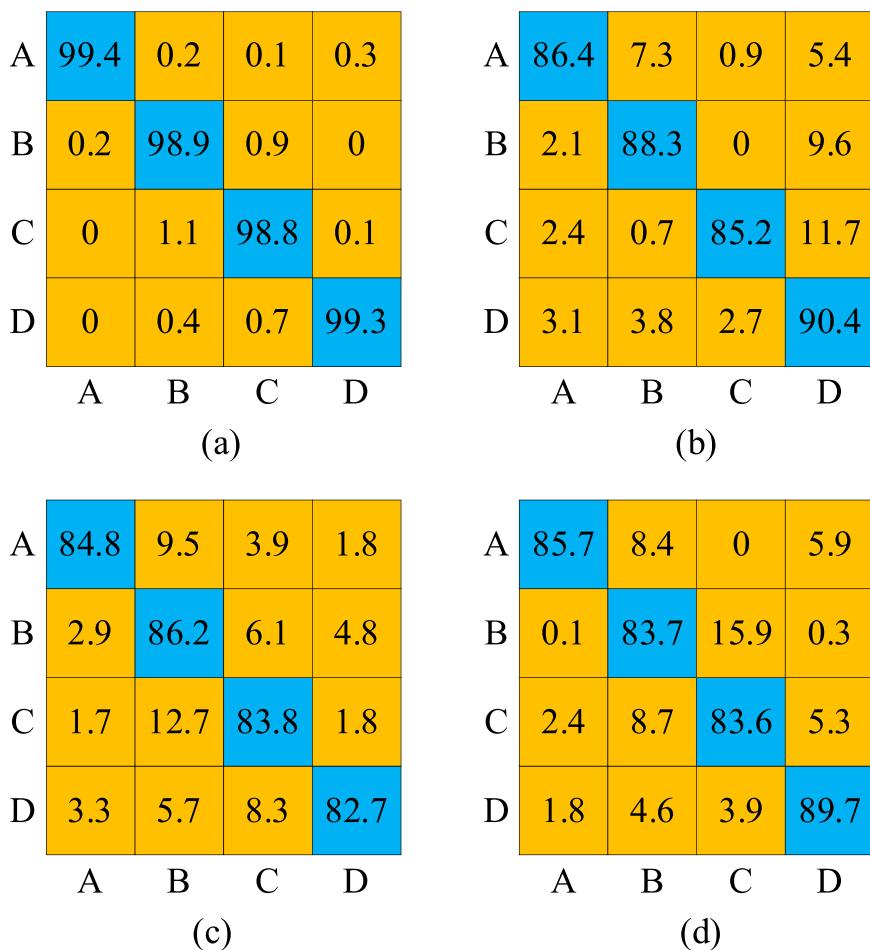


FIGURE 10  
Results of diagnostic effects. (a) Research. (b) CNN-LSTM. (c) VMD-GA-SVM. (d) VMD-CWT-CNN.

TABLE 3 Comparison results.

Method	Precision	AUC	CPU occupancy rate
Research	99.68%	0.986	42.14%
CNN-LSTM	90.08%	0.874	58.78%
VMD-GA-SVM	86.78%	0.882	60.14%
VMD-CWT-CNN	89.23%	0.903	76.34%

Good data preprocessing includes not only cleaning and normalization, but also feature selection and enhancement. This helps the model learn effective feature representations more quickly. High-quality data input is another key factor for the model to converge rapidly. These factors work together to make the model perform well in handling complex tasks.

## 4 Discussion and interpretation

This study compared the performance and diagnostic effectiveness of the suggested FD method. In the accuracy

comparison analysis, the proposed FD method, CNN-LSTM, VMD-GA-SVM, and VMD-CWT-CNN had accuracy of 98.97%, 90.89%, 88.26%, and 85.23%, respectively. The suggested FD method outperforms other methods in terms of accuracy. These results were comparable to those published in relevant studies by [Sahu and Rai. \(2023\)](#). In RT comparative analysis, the recommended diagnostic methods CNN-LSTM, VMD-GA-SVM, and VMD-CWT-CNN had average RT values of 2.38 s, 3.16 s, 3.98 s, and 4.48 s, respectively. This result was consistent with the research findings of [Zhu and Liu. \(2023\)](#). In addition, the proposed FD method achieved 98.69% RR, 0.74 loss value, 1.18 average MSE, and 0.31 average RMSE, which all outperformed the comparing techniques. This result was align with the research findings of [Guo et al. \(2024\)](#). In the analysis of application effectiveness, the recommended FD method outperformed the comparative method in every situation. This result was similar to the findings of Zhang and Wang in their study in 2024 ([Zhang and Wang, 2024](#)). This study is constrained by the fact that EMS operates in more complex real-world environments, where other factors may influence BLDCM RB failures. Future research may move forward in a positive path if FD incorporates a more complete range of affecting factors.

## 5 Summary

In response to the problem of low accuracy caused by noise in the FD method of BLDCM rolling bearings in EMSs, this study introduced WTD method to suppress noise and fuse multiscale features of BLDCM rolling bearing current signals. At the same time, CBAM and ST encoders were used to create a problem detection technique for BLDCM rolling bearings based on noise reduction and multi-scale feature fusion. The effectiveness of the suggested diagnostic techniques was examined and contrasted. According to the results, the recommended strategy outperformed the comparative strategies in terms of RT, accuracy, recall, and loss value. It was indicated that in addition to correctly identifying four distinct defect categories, the proposed diagnostic approach outperformed the comparative alternatives in terms of accuracy, CPU usage, and AUC value. The aforementioned findings demonstrate the efficacy of the FD approach suggested in the study in locating defects in the EMS's BLDCM rolling bearings.

## Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

## Author contributions

XQ: Conceptualization, Writing – original draft. YY: Methodology, Writing – original draft. SH: Data curation, Project administration, Writing – original draft. GB: Investigation, Software, Writing – review and editing. NF: Formal Analysis, Validation, Writing – review and editing.

## References

Bhuyan, P., Singh, P. K., and Das, S. K. (2024). Res4net-CBAM: a deep cnn with convolution block attention module for tea leaf disease diagnosis. *Multimedia Tools Appl.* 83 (16), 48925–48947. doi:10.1007/s11042-023-17472-6

Das, M., and Sahana, B. C. (2025). Optimized orthogonal wavelet-based filtering method for electrocardiogram signal denoising. *J. Institution Eng. (India) Ser. B* 106 (3), 965–978. doi:10.1007/s40031-022-00796-6

Guo, J., Wang, Z., Yang, Y., Song, Y., Wan, J. L., and Huang, C. G. (2024). A dual-channel transferable RUL prediction method integrated with Bayesian deep learning and domain adaptation for rolling bearings. *Qual. Reliab. Eng. Int.* 40 (5), 2348–2366. doi:10.1002/qre.3539

Lu, Q. (2024). Application of data visualization technology in fault diagnosis and maintenance of intelligent electromechanical systems. *Procedia Comput. Sci.* 243 (1), 716–723. doi:10.1016/j.procs.2024.09.010

Sahoo, G. R., Freed, J. H., and Srivastava, M. (2024). Optimal wavelet selection for signal denoising. *IEEE Access* 12 (1), 45369–45380. doi:10.1109/ACCESS.2024.3377664

Sahu, P. K., and Rai, R. N. (2023). Fault diagnosis of rolling bearing based on an improved denoising technique using complete ensemble empirical mode decomposition and adaptive thresholding method. *J. Vib. Eng. & Technol.* 11 (2), 513–535. doi:10.1007/s42417-022-00591-z

Shang, L., Zhang, Z., Tang, F., Cao, Q., Pan, H., and Lin, Z. (2025). Signal process of ultrasonic guided wave for damage detection of localized defects in plates: from shallow learning to deep learning. *J. Data Sci. Intelligent Syst.* 3 (2), 149–164. doi:10.47852/bonviewJDSIS32021771

Singh, S., Parmar, K. S., and Kumar, J. (2025). Development of multi-forecasting model using Monte Carlo simulation coupled with wavelet denoising-ARIMA model. *Math. Comput. Simul.* 230 (1), 517–540. doi:10.1016/j.matcom.2024.10.040

Sulistyo, M. E., Susilo, D. D., Nizam, M., and Ubaidillah, U. (2025). A literature review: bearing fault in BLDC motor based on vibration and thermal signals. *J. Electr. Electron. Inf. Commun. Technol.* 7 (1), 10–15. doi:10.20961/jeeict.7.1.100165

Wang, X., Zhao, L., Zhang, J., Wang, A., and Bai, H. (2024a). A wavelet-domain consistency-constrained compressive sensing framework based on memory-boosted guidance filtering. *IEEE Trans. Instrum. Meas.* 73 (1), 1–6. doi:10.1109/TIM.2024.3398096

Wang, N., Zhang, Z., Hu, H., Li, B., and Lei, J. (2024b). Underground defects detection based on GPR by fusing simple linear iterative clustering phash (SLIC-Phash) and convolutional block attention module (CBAM)-YOLOv8. *IEEE Access* 12 (1), 25888–25905. doi:10.1109/ACCESS.2024.3365959

Xu, Q., Gao, Y., Shen, J., Li, Y., Ran, X., Tang, H., et al. (2023). Enhancing adaptive history reserving by spiking convolutional block attention module in recurrent neural networks. *Adv. Neural Inf. Process. Syst.* 36 (1), 58890–58901. doi:10.48550/arXiv.2401.03719

Zangana, H. M., and Mustafa, F. M. (2024). From classical to deep learning: a systematic review of image denoising techniques. *Jurnal Ilmiah Comput. Sci.* 3 (1), 50–65. doi:10.58602/jics.v3i1.36

Zhang, T., and Wang, Y. (2024). “Fault diagnosis of wind turbine gearbox based on wavelet packet denoising and CNN-Swin Transformer-LSTM,” *Third Int. Conf. Image Process. Object Detect. Track.* 13396. SPIE, 161–172. doi:10.1117/12.3050446

## Funding

The author(s) declared that financial support was received for this work and/or its publication. The research is supported by 2022 North China Institute of Aerospace Engineering Doctoral Research Startup Fund Project: “Research on Vibration and Noise Reduction Characteristics of Plate-Rod Phononic Crystals (Funding Number: BKY-2022-11)”.

## Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Generative AI statement

The author(s) declared that generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Zhang, J., Liu, M., Deng, W., Zhang, Z., Jiang, X., and Liu, G. (2024). Research on electro-mechanical actuator fault diagnosis based on ensemble learning method. *Int. J. Hydromechatronics* 7 (2), 113–131. doi:10.1504/IJHM.2024.138231

Zhang, Z., Liu, W., Xiao, G., Xu, X., Li, M., Cheng, Z., et al. (2025). A two-phase features extraction approach for BRB based fault diagnosis of electromechanical system. *Int. J. Adapt. Control Signal Process.* 39 (7), 1451–1468. doi:10.1002/acs.3862

Zhao, Y. (2023). Precision local anomaly positioning technology for large complex electromechanical systems. *J. Meas. Eng.* 11 (4), 373–387. doi:10.21595/jme.2023.23319

Zhao, R., Jiang, G., He, Q., Hang, X., and Xie, P. (2024). Current-aided vibration fusion network for fault diagnosis in electromechanical drive system. *IEEE Trans. Instrum. Meas.* 73 (1), 1–10. doi:10.1109/TIM.2024.3363791

Zheng, S., Chen, L., and Lu, J. (2025). Numerical analysis of a fractional micro/nanobeam-based micro-electromechanical system. *FRACTALS (fractals)* 33 (1), 1–10. doi:10.1142/S0218348X25500288

Zhu, J., and Liu, T. (2023). Bidirectional current WP and CBAR neural network model-based bearing fault diagnosis. *IEEE Access* 11 (1), 143635–143648. doi:10.1109/ACCESS.2023.3343157