frontiers | Frontiers in Mechanical Engineering

TYPE Original Research
PUBLISHED 14 January 2026
DOI 10.3389/fmech.2025.1688598

Check for updates

OPEN ACCESS

EDITED BY
XJ Jing,
City University of Hong Kong, Hong Kong SAR, China

REVIEWED BY
Wu Deng,
Civil Aviation University of China, China
Junsheng Cheng,
Hunan University, China

*CORRESPONDENCE
Qiongyan Shi,
sqy20220122@163.com

RECEIVED 19 August 2025
REVISED 16 December 2025
ACCEPTED 23 December 2025
PUBLISHED 14 January 2026

CITATION
Shi Q and Yang F (2026) Composite fault diagnosis of rolling bearings based on EMD-AADPCI vibration images.
Front. Mech. Eng. 11:1688598.
doi: 10.3389/fmech.2025.1688598

COPYRIGHT
© 2026 Shi and Yang. 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.

Composite fault diagnosis of rolling bearings based on EMD-AADPCI vibration images

Qiongyan Shi[1]* and Fengbo Yang[2]

[1]Mechanical Engineering Institute, Changzhou Vocational Institute of Mechatronic Technology, Changzhou, China, [2]College of Mechanical and Electronic Engineering, Nanjing Forestry University, Nanjing, China

Introduction: Bearing fault detection and prevention are crucial. However, traditional diagnostic methods generally suffer from insufficient accuracy when dealing with complex bearing faults. Therefore, developing new methods that can effectively characterize complex fault features and achieve high-precision diagnosis has significant theoretical and engineering value.
Methods: This study proposes a vibration image generation method based on Empirical Mode Decomposition-Adaptive Angle Distribution Polar Image (EMD-AADPCI) and constructs a hybrid diagnostic model combining Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM). First, the vibration signal is processed by Empirical Mode Decomposition, and the designed adaptive angle distribution mechanism dynamically allocates polar coordinate angles according to the local features of the intrinsic mode functions, converting the signal into a two-dimensional vibration image containing rich fault information. Subsequently, a CNN-LSTM hybrid model is constructed. CNN extracts spatial and deep features from the image, and LSTM captures the temporal dependencies between features, ultimately achieving accurate classification of complex bearing faults.
Results: Experiments show that the proposed method significantly outperforms traditional methods. In terms of feature representation, the vibration images generated by EMD-AADPCI achieved a 20.25% improvement in fault classification accuracy compared to the comparative method MIC-SPCI (reaching 93.00%). The constructed CNN-LSTM model achieved a training accuracy of 94.88% with a loss rate as low as 1.43%. In the composite fault diagnosis task, the model achieved a classification accuracy of 98.00%. After 10 repeated experiments, the model achieved average accuracy, recall, and F1 score of 98.13%, 98.72%, and 98.33% for different composite fault diagnoses, respectively. Even in a low signal-to-noise ratio environment with strong noise interference (-4 dB), the model maintained a diagnostic accuracy of over 97%, demonstrating good robustness.
Discussion: The proposed EMD-AADPCI method can more effectively preserve and highlight fault-related information, while the CNN-LSTM hybrid model fully leverages the advantages of spatial feature extraction and time series modeling. Experimental results show that this method has extremely high accuracy and anti-interference ability in bearing composite fault diagnosis. This provides an effective and innovative solution for intelligent diagnosis and preventive maintenance of complex faults in bearings and other rotating machinery, and has good prospects for widespread application.

KEYWORDS

CNN-LSTM, composite fault, EMD-AADPCI, rolling bearings, vibration image

1 Introduction

Rolling bearings, as the core components of mechanical systems, are widely used in industrial machinery, wind turbines, rail transit and other fields (An et al., 2023). According to the 2023 Industrial Equipment Failure Statistics Report, bearing failures account for over 40% of rotating machinery failures, with composite failures accounting for over 60%. This type of fault is prone to misdiagnosis due to characteristic coupling and interference signals, which can lead to equipment shutdown. The daily loss of wind turbine shutdown due to composite faults can reach CNY 50,000 to 100,000, while car gearbox failures can easily trigger safety accidents (Agajie et al., 2023). The reliability of key rotating machinery is the foundation for the stable operation of energy systems, and the fault diagnosis technology is crucial for ensuring the efficiency and sustainability of renewable energy systems (Agajie et al., 2024). At present, the diagnostic methods for single bearing faults are relatively mature, but composite faults under complex working conditions still faces challenges. In traditional feature extraction methods, spectrograms convert time-domain vibration signals into frequency domains, visually displaying frequency components. The time-frequency diagram can reflect the variation of signal frequency over time. Gray level co-occurrence matrix images describe the statistical characteristics of vibration signals from a texture perspective (Liu et al., 2023; Bäßler et al., 2022). However, existing methods still have the following limitations: (1) Manual feature extraction is highly dependent, and traditional methods (such as SVM and ANN) require manual design of spectral peaks, entropy values, and other features, resulting in poor coupling feature extraction performance for composite faults (Toumi et al., 2022; Özmen and Karabacak, 2023); (2) High sensitivity to noise: Fixed transformation image generation methods such as spectrograms are prone to losing fault features under industrial noise, and have low diagnostic accuracy at low Signal-to-Noise Ratios (SNR<0 dB); (3) The generalization of composite faults is poor, and a single deep learning model cannot fully capture the "spatial coupling time evolution" characteristics of composite faults, resulting in low accuracy when generalizing to different types of composite faults (Preethi and Mamatha, 2023; Chen et al., 2021). Therefore, it is necessary to develop a more comprehensive and efficient system to address composite faults in bearing.

In terms of signal feature extraction, many feature extraction algorithms based on vibration signals have been built. To address the difficulties in obtaining labels for bearing data, Tao et al. built an unsupervised cross-domain fault diagnosis strategy. The method extracted features from signal decomposition and reconstruction. Compared with other algorithms, this method had higher robustness (Tao et al., 2023). Zhao et al. built a multi-scale feature fusion strategy to address the complex vibration signals and difficult impact characteristic extraction in composite faults. The method involved ensemble empirical mode decomposition, feature calculation, and fusion, followed by diagnosis using least squares SVM. This method could quantitatively characterize data, improve anti-interference ability, extract features, and identify fault types, with significantly higher accuracy than a single method (Zhao et al., 2022). To solve the challenge that the original permutation entropy algorithm ignores the amplitude relationship between adjacent signals, Zheng et al. proposed phase inverse permutation entropy, combined with whale optimization algorithm to SVM diagnosis. This method could effectively identify the location and degree of faults, with higher fault recognition rate (Zheng et al., 2023). In response to the difficulty of diagnosing multiple faults, Lv et al. built a multi-fault separation and identification method based on time-frequency spectra to improve the fast path optimization method and extract multiple transient component curves. The time-frequency masking was constructed, and signal reconstruction and fault detection were performed. This method effectively separated and identified multiple faults in rolling bearings (Lv et al., 2024). To address the certainty of feature extraction, Kaya et al. created an experimental bearing testing device. A feature extraction strategy based on the co-occurrence matrix was proposed. The new signal was obtained using a one-dimensional local binary pattern. After testing on three datasets, the success rate of this method reached 87.50%, 96.5%, and 99.30% for various speeds, fault sizes, and fault types, respectively (Kaya et al., 2021). Li X et al. proposed an intelligent diagnostic framework that combined acoustic vibration signals with graph neural networks to address the single signal diagnosis being affected by strong noise and traditional deep learning ignoring sample dependence. The framework integrated data using the correlation variance contribution method, optimized data representation using AcvGraph, and integrated features through enhanced DiffPool dimensionality reduction. Experimental results showed that it could effectively detect bearing faults and was superior to other intelligent technologies (Li et al., 2024). To address the difficulty in balancing accuracy and generalization ability in generalized bearing fault diagnosis, Li J et al. proposed the IBN-MixStyle network, which integrated discriminative and generalized features and combined them with a dynamic weighted invariant risk minimization strategy to optimize the model. The results showed that the average diagnostic accuracy was improved by 5.3%–15.47%, which was suitable for highly variable industrial environments (Li et al., 2025). To address the insufficient time-frequency resolution and energy diffusion of bearing fast time-varying signals, Deng W et al. proposed the parameterized iterative time-frequency multiple compression transform method, which improved time-frequency performance by optimizing kernel function parameters and iterative rearrangement strategy. The results showed that this method had excellent time-frequency energy concentration and could more accurately capture signal features to improve diagnostic performance (Deng et al., 2025a).

At present, many scholars have introduced deep learning algorithms into fault diagnosis. In response to the challenge that existing models cannot adaptively select features, Guo et al. built an end-to-end diagnosis strategy based on attention mechanism and CNN-Bidirectional Long Short-Term Memory Network (CNN-BiLSTM). This strategy had high accuracy and strong noise resistance on different datasets, and had universality and superiority (Guo et al., 2023a). To solve the single task learning being unable to mine complementary information, Wang et al. built a multi-task attention CNN, which achieved information sharing. The performance was superior to advanced deep learning methods (Wang et al., 2021). Xing et al. designed a new CNN-based to deal with the data imbalance in intelligent fault diagnosis. The experiment showed that this method could automatically extract

discriminative features and effectively deal with data imbalance problems (Xing et al., 2022). In response to the difficulties in fault identification caused by noise, Aljemely et al. proposed LSTM-large interval nearest neighbor algorithm, which used orthogonal weight initialization technology to remember key fault information and organize samples. Two bearing fault diagnosis experiments showed that this method outperformed existing methods (Aljemely et al., 2022). Amiri A F et al. proposed a two-step approach to address modeling difficulties and insufficient diagnostic accuracy in photovoltaic system fault detection: using MGWO algorithm to extract ODM parameter modeling, and then constructing a double random forest classifier. The modeling root mean squared error was 0.0122, and the accuracy of fault detection and diagnosis reached 99.4%, which was better than that of SVM (Amiri et al., 2024). To address the scarce train motor bearing fault samples and insufficient generalization ability under dynamic working conditions, Zhao H et al. proposed a few-sample cross-domain fault diagnosis method based on MAML-GA, which optimized the model through dynamic meta-task enhancement, parameter update operator, and lightweight feature extraction network. The results showed that the method effectively improved the cross-domain generalization ability. The diagnostic accuracy was significantly higher than that of the mainstream method (Zhao et al., 2025). To address the difficult identification of bearing fault pulse features under complex working conditions, Deng W et al. combined time redistribution multi-synchronous compression transformation, complex sparse learning dictionary, and mask decomposition algorithm to extract features. The results showed that the method had strong anti-noise interference and significantly improved the accuracy and robustness of fault frequency extraction (Deng et al., 2025b).

Moreover, the advancements in intelligent sensing and diagnosis are not confined to mechanical systems. In other complex and high-risk fields, such as air traffic control, research has demonstrated the effectiveness of context aware speech recognition and multi-modal situational awareness systems in improving operational safety and efficiency through deep learning (Guo et al., 2025a; Guo et al., 2025b). These studies emphasize the potential of adaptive data-driven models to capture complex patterns in noisy, real-time environments, a challenge similar to that encountered in bearing composite fault diagnosis. Meanwhile, in the field of mechanical fault diagnosis, ongoing research continues to introduce innovative methods to overcome the persistent challenges in signal analysis. Recent developments have focused on enhancing the handling of non-stationary signals and improving feature extraction precision. For instance, techniques such as Empirical Fourier-Bessel Heuristic Denoising (EFBHD) provide enhanced frequency resolution for isolating fault-related narrowbands (Zhou et al., 2025). While iterative frameworks such as Shrinkage Sliding Fourier-Bessel Packet (SSFBP) can dynamically and centrally extract feature spectral components (Zhou et al., 2026). Beyond feature extraction, ensuring data integrity is equally vital. Efficient models such as Variable Scale Multi-layer Perceptron (VS-MLP) have effectively recovered abnormal vibration data with good computational efficiency (Fan et al., 2024). These efforts collectively reflect the ongoing push for more adaptive, precise, and computationally efficient signal processing.

In summary, scholars have proposed many solutions for bearing fault detection, but there are still problems such as difficulty in extracting complex faults and high dependence on samples. A vibration image generation method based on Empirical Mode Decomposition-Adaptive Angle Distribution Polar Coordinate Image (EMD-AADPCI) and a CNN-LSTM fault diagnosis model are built to deal with the composite fault diagnosis problem. The innovation of the research lies in: ① proposing the AADPCI polar coordinate transformation method, which uses an adaptive angle allocation mechanism (dynamically adjusting the angle factor based on the time-domain characteristics of IMF components) to solve the problem of fixed features in traditional polar coordinate imaging; ② Combining EMD with AADPCI, filtering IMF components and polar imaging to preserve richer fault features; ③ The CNN-LSTM hybrid model is constructed, and the spatial feature extraction and long time sequence dependency capture capabilities of the two are fused to solve the problem of incomplete feature extraction of a single model. The core goal of the research is to solve the problems of "feature loss and noise sensitivity" through the EMD-AADPCI method, and to solve the problems of "incomplete feature extraction and poor generalization" through the CNN-LSTM model, ultimately achieving high-precision diagnosis of composite faults in rolling bearings under complex working conditions and compensating for the limitations of existing methods.

2 Methods and materials

2.1 Vibration image generation based on EMD-AADPCI

Analyzing the vibration signals can obtain fault information of the bearings, but vibration signals are one-dimensional time series data, and the information display is relatively simple. Vibration images can express the amplitude, frequency, phase, and other information of vibration signals through features such as grayscale, color, and texture, making the feature information more intuitive and comprehensive (Li et al., 2022; Meng et al., 2021). Converting vibration signals into vibration images can enhance the stability of features and make it easier to use deep learning algorithms for fault classification. The study proposes the EMD-AADPCI vibration image generation method. Firstly, EMD decomposes the vibration signal into IMFs and a residual component, which contains local characteristic signals. The input vibration signal is $z(t)$. Firstly, all local maximum and minimum points of $z(t)$ are found. The cubic spline interpolation function is taken to match the maximum and minimum points, respectively, and obtain the upper envelope $e_{\max}(t)$ and lower envelope $e_{\min}(t)$ of the vibration signal $z(t)$. The average value is shown in Equation 1.

$$m_1(t) = \frac{e_{max}(t) + e_{min}(t)}{2} \tag{1}$$

In Equation 1, $m_1(t)$ signifies the mean of $e_{\max}(t)$ and $e_{\min}(t)$ at the first iteration. Subtracting the mean envelope yields the first IMF component, as shown in Equation 2.

$$h_1(t) = z(t) - m_1(t) \tag{2}$$

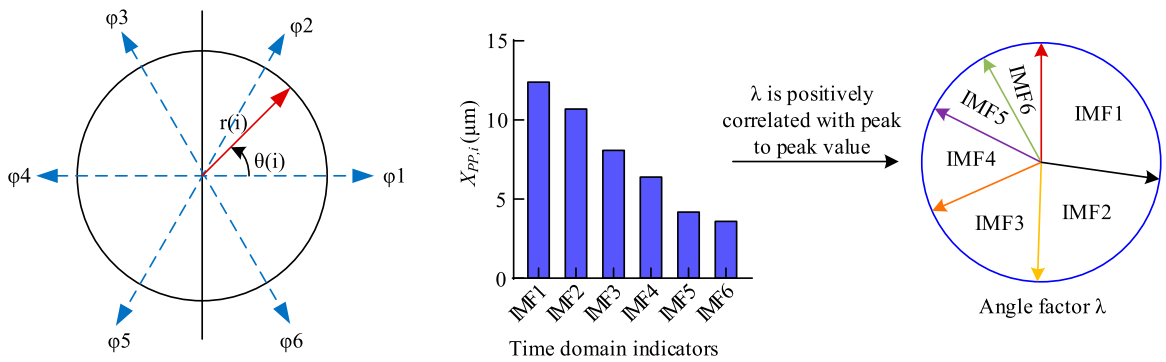

FIGURE 1
The process of AADPCI transformation.

If $h_1(t)$ satisfies the IMF is determined, that is, the number of extreme points and the quantity of zero crossings are same or differ by a maximum of 1, and are locally symmetrical. If it is not satisfied, $h_1(t)$ is taken as the new signal. The steps are repeated until $c_1(t)$ that meets the conditions is obtained. Subtracting the first IMF $c_1(t)$ yields the remaining signal, as presented in Equation 3.

$$r_1(t) = z(t) - c_1(t) \tag{3}$$

$r_1(t)$ signifies a new signal. The second IMF component $c_2(t)$ and the residual signal $r_2(t)$ are obtained. $r_2(t)$ signifies a new signal. The calculation is repeated until the residual signal $r_n(t)$ is a monotonic function or constant. Equation 4 presents the final obtained original vibration signal.

$$z(t) = \sum_{i=1}^{n} c_i(t) + r_n(t) \tag{4}$$

Among the multiple IMF components obtained from EMD, the component $c_i(t)$ containing the main fault information is selected for polar coordinate transformation. The transformation principle of AADPCI is shown in Figure 1.

In Figure 1, the length of $c_i(t)$ is N. The time series $t = 1, 2, 3, \cdots, N$ is mapped to the angle θ of the polar coordinate, and the signal amplitude $c_i(t)$ is mapped to the radius r of the polar coordinate. The angle is adaptively allocated. The definition of $r(i)$ in the transformed polar coordinate point $A(r(i), \theta(i))$ is shown in Equation 5.

$$r(i) = \frac{z_i - z_{\min}}{z_{\max} - z_{\min}} \tag{5}$$

In Equation 5, $z_{\min}$ signifies the minimum value of the vibration parameter, and $z_{\max}$ signifies the maximum value of the vibration parameter. The definition of $\theta(i)$ in polar coordinate point $A(r(i), \theta(i))$ is presented in Equation 6.

$$\theta(i) = \varphi + \frac{z_{i+q} - z_{\min}}{z_{\max} - z_{\min}} \times \lambda \times w \tag{6}$$

In Equation 6, φ represents the initial line angle, usually taken as 60°. q is the time interval, with values ranging from 3 to 10. w is the angle amplification factor, with values ranging from 20° to 60°. The values of q and w are determined based on controlled variable experiments. Experiments have shown that when q is within 3–10, it can balance computational complexity while ensuring image resolution and feature continuity. When q >10, the image details become redundant and the computation time significantly increase. When q <3, it leads to feature loss. Similarly, w can effectively distinguish the significance of different IMFs within 20°–60°. A too small w will result in insufficient sidelobe discrimination, while a large w will cause image distortion. The final value is determined through grid search within the specified range to achieve the best balance between feature retention rate and computational efficiency. λ represents the angle adaptation rate. The larger the λ, the larger the corresponding IMF component will be allocated to a larger angle factor. According to the point $A(r(i), \theta(i))$ transformed by polar coordinates, an image is drawn on the polar coordinate plane, and then the polar coordinate plane is discretized into several grids. The number of points or energy in each grid is counted and converted into grayscale values to obtain a two-dimensional vibration image. The mapping relationship between the characteristic matrix and the grayscale image is displayed in Figure 2.

In Figure 2, the original vibration signal data is changed into a vibration feature matrix. The minimum value min is mapped to level 0 in the grayscale image, and the maximum value max is mapped to level 255. The values of other elements in the matrix are mapped to corresponding grayscale values between 0–255 based on their relative positions between the maximum values, forming a visual grayscale image of the vibration signal characteristics for subsequent fault diagnosis and analysis. The EMD-AADPCI is shown in Figure 3.

From Figure 3, the vibration signal dataset is divided into N segments according to the fault state, with each segment being n in length. The -i segment data is subjected to EMD to obtain several IMF components. The Pearson Correlation Coefficient (PCC) between each IMF component and the original vibration signal is shown in Equation 7.

$$PCC = \frac{\sum_{t=1}^{N} (x(t) - \bar{x})(y(t) - \bar{y})}{\sqrt{\sum_{t=1}^{N} (x(t) - \bar{x})^2} \sqrt{\sum_{t=1}^{N} (y(t) - \bar{y})^2}} \tag{7}$$

In Equation 7, $x(t)$ is the original signal, $y(t)$ is the i th IMF component, and $\bar{x}$ and $\bar{y}$ are the average of the original signal and the IMF component. The IMFs are sorted from largest to smallest

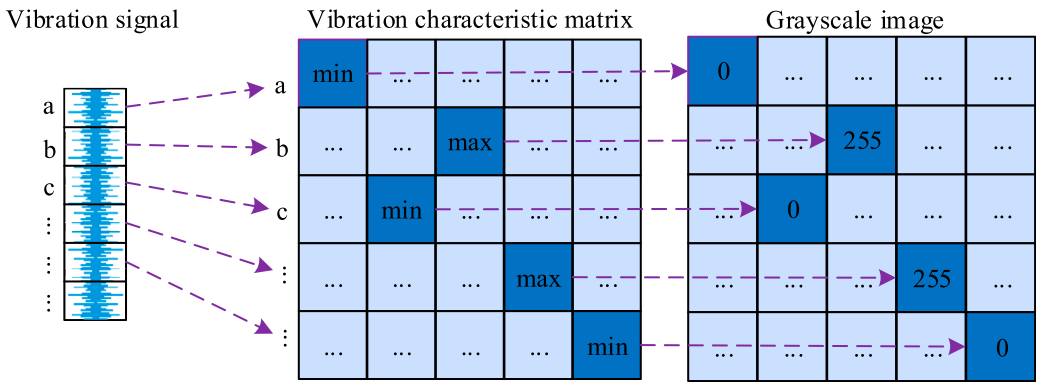

FIGURE 2
Mapping relationship.

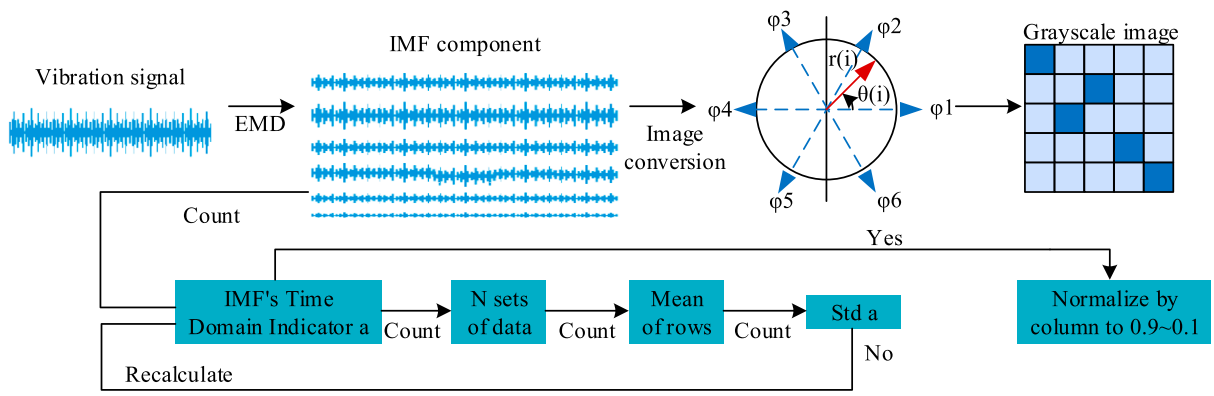

FIGURE 3
Schematic diagram of EMD-AADPCI.

TABLE 1 Influence of different numbers of IMFs on feature retention rate and calculation time (based on RBCFD).

Number of IMFs	Feature retention rate (%)	Calculation time (s/sample)
3	78.2	0.8
4	85.5	1.1
5	89.3	1.4
6	92.1	1.7
7	92.6	2.2
8	92.8	2.6

according to PCC, and the first six groups of IMFs are superimposed for noise reduction. The reasons are shown in Table 1.

Experimental verification with controlled variables shows that when the number of IMFs is ≥ 6, the fault feature retention rate remains stable at over 90%. When the number of IMFs is greater than 6, the feature retention rate does not improve significantly, but the computation time increases by over 30%. Therefore, six IMFs are selected to balance feature quality and computational efficiency. Next, common time-domain metrics (such as RMS value, peak value, kurtosis, and impulse factor) are calculated for the six selected IMF components to form a feature vector. The standard deviation of each time-domain metric across all six IMF components is then calculated. A larger standard deviation indicates a greater ability of the metric to distinguish between different IMFs. The time-domain metric with the largest standard deviation is selected and its values across IMF1-IMF6 are normalized. The normalized result is the adaptation rate angle λ corresponding to each IMF. Finally, IMF1-IMF6 are sequentially converted to the AADPCI.

2.2 Composite fault diagnosis based on CNN-LSTM

The vibration images produced by EMD-AADPCI contain a large amount of characteristic information of bearing composite faults, such as sharp pulses, amplitude changes, peak changes, etc. Next, a suitable model is selected to extract these features and complete the fault diagnosis task through continuous learning and training. The CNN-LSTM combines CNN and LSTM. CNN

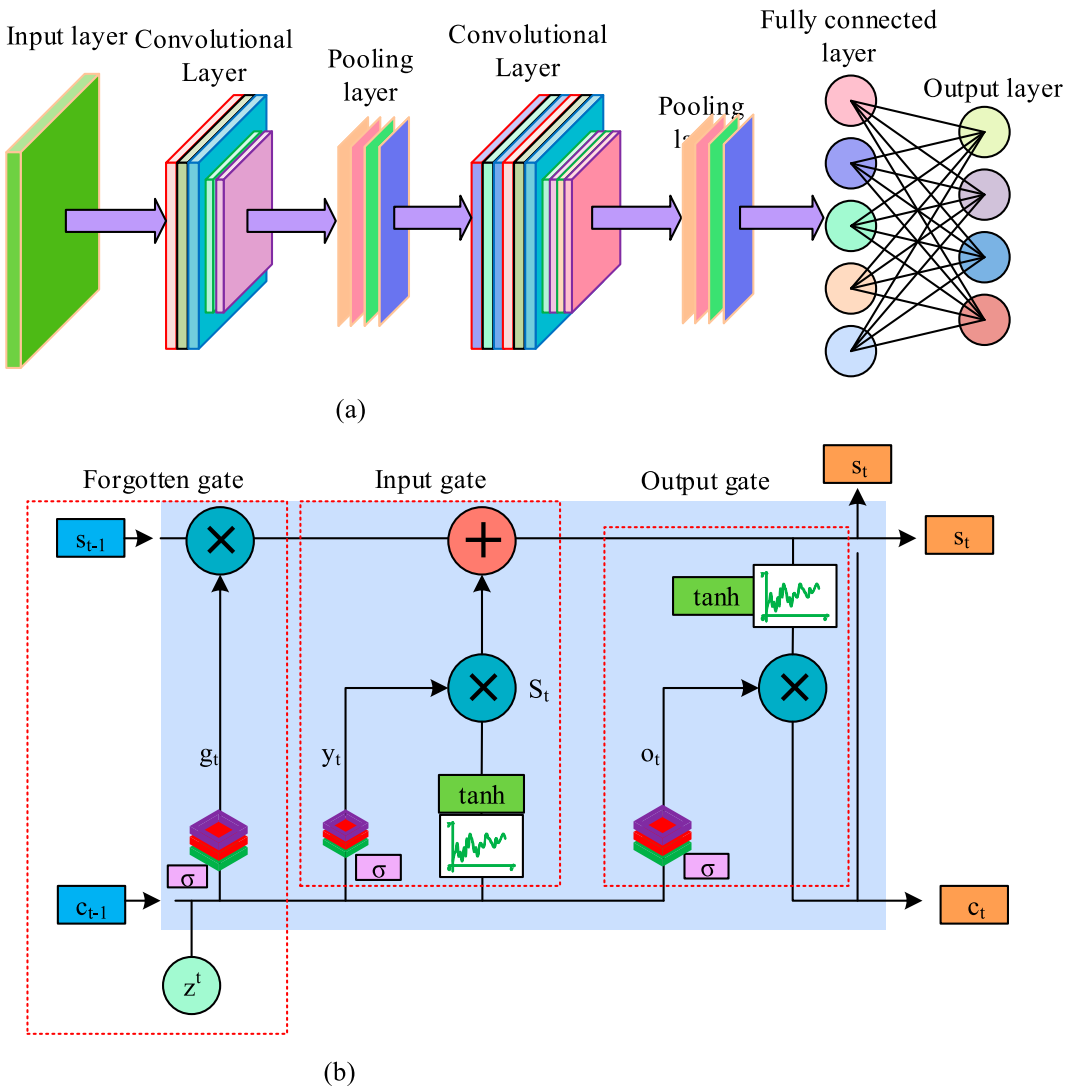

FIGURE 4
Network diagram of CNN and LSTM. **(a)** Network diagram of CNN **(b)** Network diagram of LSTM.

extracts local features, and then input into LSTM for sequence modeling. This model effectively integrates the feature extraction advantages of both to more comprehensively capture the fault features of bearings and improve the diagnosis accuracy. CNN includes convolutional computation and has deep structure, with representation learning ability, suitable for the field of fault classification (Guo et al., 2023b; Sun and Fan, 2023). The vibration signals and temperature signals generated during the operation of rolling bearings are typical time series data. LSTM can effectively process such sequence data and automatically learn various types of fault features. The diagrams of CNN and LSTM are presented in Figure 4.

In Figure 4a, the input layer can handle multidimensional data, such as vibration image data generated by EMD-AADPCI. The convolutional layer extracts features through convolutional kernel sliding scanning and generates a feature map. The convolution calculation of the output node is shown in Equation 8.

$$z_j^l = f\left(\sum_{i=1}^{m} v_{ij} * z_{mi+j}^{l-1} + \varepsilon_j^l\right) \tag{8}$$

In Equation 8, z_j^l signifies the j-th output node in the l-th layer. m signifies the quantity of convolution kernels. v_{ij} represents the i-th convolution kernel. ε signifies the bias term. * signifies the convolution operation. The feature map is subjected to dimensionality reduction through pooling layers, and then passed on to convolutional and pooling layers. As the hierarchy progresses, features are gradually extracted from simple to complex. The output of neurons after max pooling operation is shown in Equation 9.

$$y_j^{l(i)} = max_r\left[b_j^{t(i)}\right],\ (j-1)w + 1 \le r \le jw \tag{9}$$

In Equation 9, $y_j^{l(i)}$ signifies the j-th neuron in the i-th feature of the l-th. $b_j^{t(i)}$ signifies the activation value of the corresponding neuron. w signifies the width of the pooling area. Finally, all

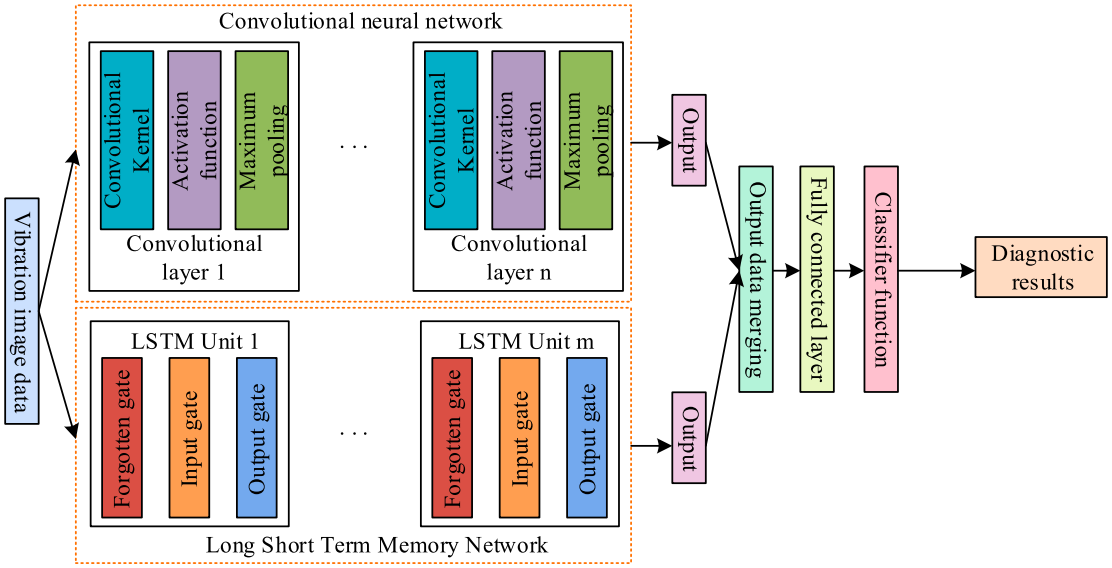

FIGURE 5
The framework for CNN-LSTM.

obtained features are aggregated in the Fully Connected Layer (FCL) to classify. Equation 10 displays the output of the FCL.

$$p^l = f\left(\delta^l p^{l-1} + \eta^l\right) \tag{10}$$

In Equation 10, $f(\cdot)$ signifies the Softmax function. δ^l signifies the weight. η^l signifies the bias term. p^{l-1} represents the previous output. In Figure 4b, LSTM has Forget Gate (FG), Input Gate (IG), and Output Gate (OG). Given a time series $Z(z_1, z_2, z_3, \cdots, z_n)$, the hidden layer is $C(c_1, c_2, c_3, \cdots, c_n)$. FG checks the hidden state and current input of the previous moment. The Sigmoid outputs a value between 0-1, which is multiplied by the previous value in the memory unit to determine which information to discard from the memory unit, as presented in Equation 11.

$$g_t = Sigmoid\left(W_f \times c_{t-1} + U_f \times z_t + \xi_f\right) \tag{11}$$

In Equation 11, g_t represents the output value of the FG passing through the Sigmoid function at time t. W_f signifies the weight matrix connecting the FG to the output c_{t-1} of the previous hidden layer. U_f signifies the weight matrix connecting the FG to the input z_t at present. c_{t-1} signifies the output value at the previous time, carrying information from the previous time. ξ_f signifies the bias term of the FG, used to adjust the output of the activation function. IG: The Sigmoid layer decides information updates, the tanh layer produces new candidate values, and the two are combined to determine the new information that needs to be added to the storage unit. The updated memory unit is obtained by adding it to the memory unit value processed by the FG. The output y_t of the IG at time t is shown in Equation 12.

$$y_t = Sigmoid\left(W_i \times c_{t-1} + U_i \times z_t + \xi_i\right) \tag{12}$$

In Equation 12, W_i and U_i signify the weight matrices of the IG and c_{t-1} and z_t. ξ_i signifies the bias term of the IG. The candidate state value $\tilde{S}_t$ of the memory unit at time t is presented in Equation 13.

$$\tilde{S}_t = \tanh\left(W_C \times c_{t-1} + U_C \times z_t + \xi_C\right) \tag{13}$$

In Equation 13, the tanh function is used to compress the output value between −1 and 1. W_C and U_C are weight matrices related to candidate states for memory units. ξ_C signifies the bias term of the memory unit. By integrating the output results of the FG and IG, the current state value of the memory unit is obtained, as shown in Equation 14.

$$S_t = g_t{}^* S_{t-1} + y_t{}^* \tilde{S}_t \tag{14}$$

In Equation 14, S_{t-1} signifies the state value of the memory unit at time $t-1$. * signifies the product of vectors. OG: Based on the current hidden state and input, which information in the memory unit is determined to output through the Sigmoid function. Then, the memory unit value is transformed through the tanh function and multiplied with the Sigmoid output result to obtain the final output. The activation value O_t of the OG at time t is presented in Equation 15.

$$O_t = Sigmoid\left(W_O \times c_{t-1} + U_O \times z_t + \xi_O\right) \tag{15}$$

In Equation 15, ξ_O signifies the bias term of the OG. The output value c_t of the model at time t is presented in Equation 16.

$$c_t = O_t{}^* \tanh\left(S_t\right) \tag{16}$$

Finally, the CNN-LSTM is presented in Figure 5.

From Figure 5, the vibration image data are input into both the CNN module and the LSTM. The CNN has multiple convolutional layers. The kernel of each convolutional layer performs feature extraction, introduces nonlinearity into the activation function, and maximizes pooling to reduce data dimensionality and computational complexity. Multi-layer convolution operation gradually extracts spatial features from vibration images. The LSTM module consists of multiple LSTM units, each containing

TABLE 2 Structural parameter of CNN-LSTM.

Network layer	Sub - network	Number of layers	Kernel size	Number of output channels	Activation function	Pooling size	Number of hidden units	Dropout
Convolutional layer	CNN	3	3 × 3	32→64→128	ReLU	—	—	—
Max pooling layer		2	2 × 2	—	—	2 × 2	—	—
Fully connected layer		1	—	256	ReLU	—	—	—
LSTM layer	LSTM	2	—	—	Sigmoid	—	128	0.2
Fully connected layer		1	—	256	ReLU	—	256	—

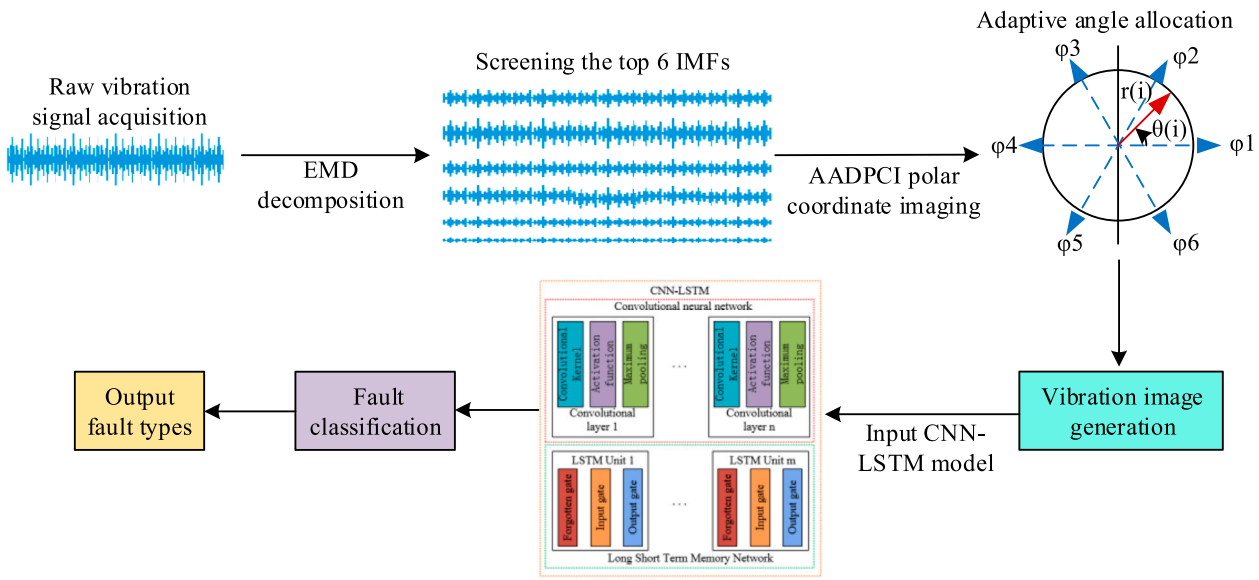

FIGURE 6
Complete flowchart for diagnosing composite faults in rolling bearings.

an FG, IG, and OG, used to process sequence data and capture long-term dependencies. The features extracted by CNN and LSTM are integrated, and then the features are integrated through an FCL. Finally, the diagnostic results are output through a classifier function to determine whether there is a fault in the bearing and its type. The structural parameters of the CNN-LSTM are shown in Table 2.

After constructing the CNN-LSTM model, the complete process of diagnosing composite faults in rolling bearings using EMD-AADPCI and CNN-LSTM is shown in Figure 6.

3 Results

3.1 Composite fault diagnosis analysis based on EMD-AADPCI vibration images

To verify that the vibration images generated by EMD-AADPCI contain richer and more accurate feature information, this study compares them with images generated by Multi-Information Capacity Simplified Polar Coordinate Image (MIC-SPCI) and validated them on the dataset using CNN. To balance model performance and operational efficiency, the fusion and training hyperparameters were determined as follows: Adam is chosen as the optimizer, which converges faster and has smaller loss fluctuations on small sample datasets compared to SGD and RMSprop. The learning rate is set to 0.001. If it is too high, it may overfit, while if it is too low, it may delay convergence. The batch size is set to 32 to balance hardware memory usage and training stability. The training epochs are set to 100, combined with an early stop strategy (terminate if the validation set loss does not decrease for five consecutive epochs). The data validation segmentation ratio is 20%, which not only avoids over-fitting, but also effectively evaluates generalization ability. The study used two self-built rolling bearing composite fault datasets for validation. Rolling Bearing Composite Fault Dataset (RBCFD): A single point fault was simulated on bearing 6,205 by performing electrical discharge

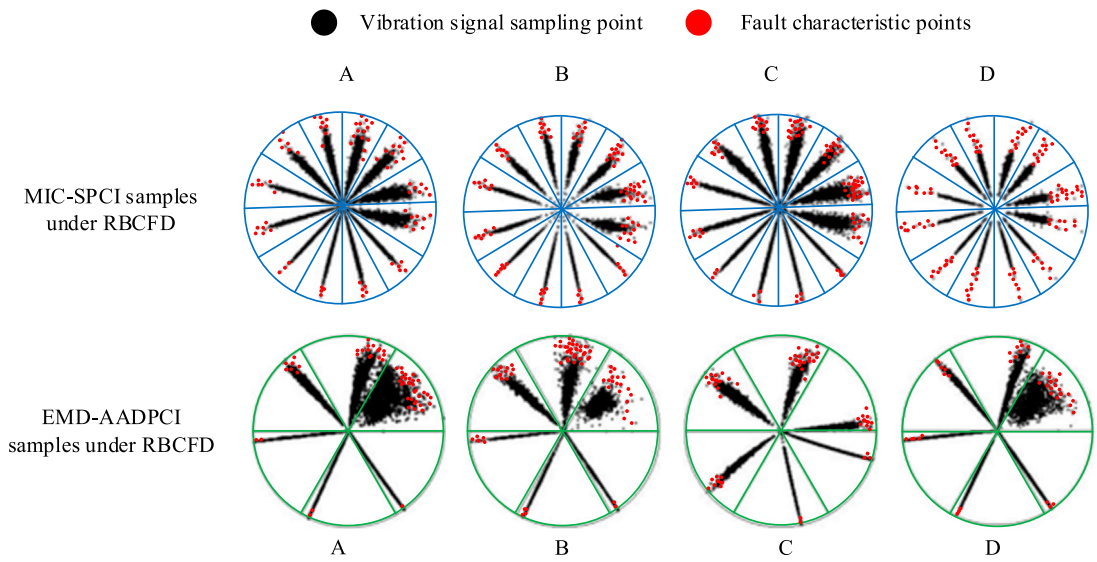

FIGURE 7
Comparison of MIC-SPCI and EMD-AADPCI vibration images for four types of faults (A-normal, B-inner ring + rolling element, C-outer ring + rolling element, and D-inner ring + outer ring) in the RBCFD.

machining on the inner and outer rings and rolling elements of a normal bearing. Then, four states were constructed by combining faults from different parts: normal (A), inner ring + rolling element fault (B), outer ring + rolling element fault (C), and inner ring + outer ring fault (D). The vibration signal is collected by an acceleration sensor installed on the bearing seat, with a sampling frequency of 12 kHz, a motor speed of 1,800 rpm, and a load of 0.5 hp. 200 samples were collected for each state, with a length of 1,024 data points per sample, for a total of 800 samples. Rolling Bearing and Rotor Friction Composite Fault Dataset (RBRFCFD): On the same test bench, rotor friction faults were introduced and four additional states were constructed: normal (E), inner ring + rotor friction fault (F), outer ring + rotor friction fault (G), and rolling element + rotor friction fault (H). The collection settings are the same as that of RBCFD, and 200 samples are obtained for each state, totaling 800 samples. 80% of all datasets are used for training and 20% for testing. The vibration images generated by MIC-SPCI and EMD-AADPCI for various faults in the RBCFD are presented in Figure 7.

In Figure 7, the vibration image generated by MIC-SPCI had a total of 12 side lobes, and each side lobe had its own characteristic features, such as different thickness and deflection angles. However, the vibration images of faults A and C generated by MIC-SPCI were quite similar, which was not conducive to subsequent fault diagnosis tasks. The vibration images generated by EMD-AADPCI only retained the six most prominent side lobes, each with varying thickness, deflection angle, and distribution phenomenon. The distinguishing features of the four types of faults were relatively clear, which was beneficial for subsequent fault diagnosis tasks. The core reason why EMD-AADPCI is superior to MIC-SPCI is as follows: (1) MIC-SPCI adopts a fixed angle allocation of 12 side-lobes, which cannot be dynamically adjusted according to fault characteristics; While EMD-AADPCI allocates a larger angle factor to the IMF with more prominent fault features and a smaller angle to the IMF with fuzzy features according to the time-domain index allocation angle of the IMF to avoid irrelevant feature interference; (2) EMD can filter high-frequency noise in the original signal, while MIC-SPCI directly images the original signal. Noise can cause sidelobe features to be blurred, resulting in higher fault discrimination. 400 sample images are randomly selected from the RBCFD, with 100 images for each type of fault. The samples are subjected to MIC-SPCI and EMD-AADPCI operations. The confusion matrix combined with CNN diagnostic results is presented in Figure 8.

According to Figure 8a, the fault classification accuracy based on MIC-SPCI was only 72.75%. The classification accuracy of type A was the highest, at 75.00%. The classification accuracy of type C faults was relatively low, with 11 misclassified as type B faults, 10 misclassified as type D faults, and nine misclassified as type A. According to Figure 8b, the fault classification accuracy based on EMD-AADPCI was significantly higher than that of MIC-SPCI, reaching 93.00%. The model had a classification accuracy of over 90.00% for all four types of faults, with the highest classification accuracy for types A and D. The classification accuracy of type B was relatively low, with 4 misclassified as type C, 3 misclassified as type A, and 3 misclassified as type D. The experiment is repeated 10 times. Figure 9 displays the comparison results.

According to Figure 9a, after 10 experiments, the average accuracy of the MIC-SPCI algorithm was only 72.76%, the average recall was 72.97%, and the average F1 value was 71.64%. As shown in Figure 9b, the EMD-AADPCI had significantly higher indicators than the MIC-SPCI, with an average accuracy of 91.73%, which was 18.97% higher than that of the MIC-SPCI. The average recall rate and F1 value were 92.38% and 92.15%, respectively. The vibration images generated by the EMD-AADPCI algorithm have clearer features and help improve the diagnosis accuracy. To verify the independent contribution of each key module in the proposed method and quantitatively evaluate the effectiveness of the AADPCI

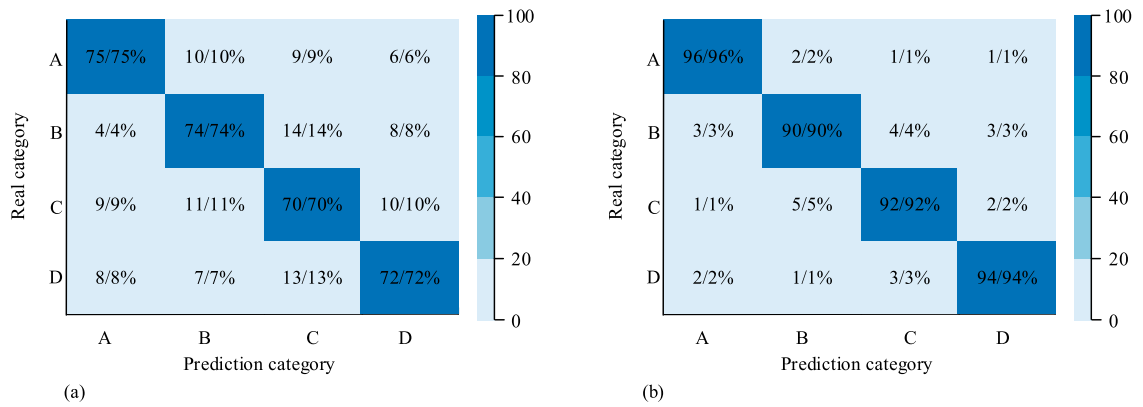

FIGURE 8
Confusion matrix of fault diagnosis results based on MIC-SPCI and EMD-AADPCI. **(a)** Sample testing confusion matrix for MIC-SPCI **(b)** Sample testing confusion matrix for EMD-AADPCI.

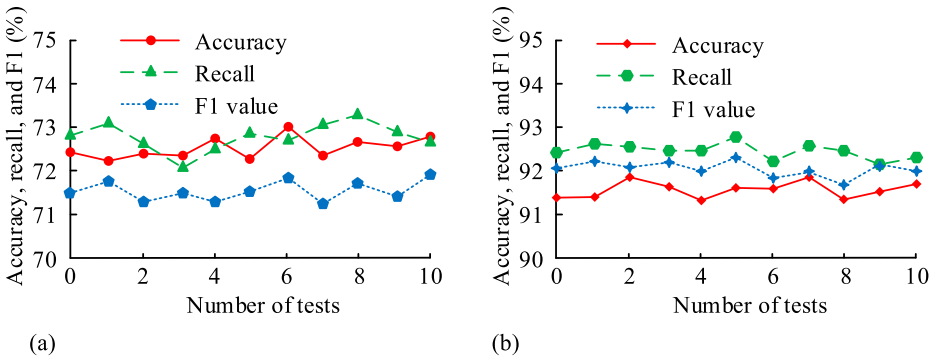

FIGURE 9
The average accuracy, recall, and F1 value of two models. **(a)** Various indicators based on MIC-SPCI model **(b)** Various indicators based on EMD-AADPCI model.

TABLE 3 Results of ablation experiments on the RBCFD.

Feature generation method	Classification model	Accuracy (%)	Recall (%)	F1-score (%)
MIC-SPCI	CNN-LSTM	72.75	72.97	71.64
EMD-FPCI	CNN-LSTM	86.50	86.80	86.25
EMD-AADPCI	CNN-LSTM	98.00	98.72	98.33
EMD-AADPCI	CNN	93.00	92.38	92.15
EMD-AADPCI	LSTM	85.25	85.32	85.17
EMD-AADPCI	CNN-LSTM	98.00	98.72	98.3

adaptive mechanism, ablation experiments were designed. On the RBCFD, the performance of the following feature generation methods was compared using the same CNN-LSTM classifier: (1) MIC-SPCI (baseline): Fixed angle polar coordinate imaging. (2) EMD-FPCI: Using EMD, but employing the same Fixed Polar Coordinate Imaging (FPCI) as MIC-SPCI for imaging, i.e. removing the adaptive mechanism of AADPCI. (3) EMD-AADPCI (Ours): Complete adaptive angle allocation polar coordinate imaging. In addition, the combination effect of different classifiers was evaluated: (4) EMD-AADPCI + CNN: Using only the CNN model. (5) EMD-AADPCI + LSTM: Using only the LSTM model. (6) EMD-AADPCI + CNN-LSTM: The proposed complete model. The experimental results are shown in Table 3.

According to Table 3, under different feature generation methods and the same CNN-LSTM classifier, the accuracy of

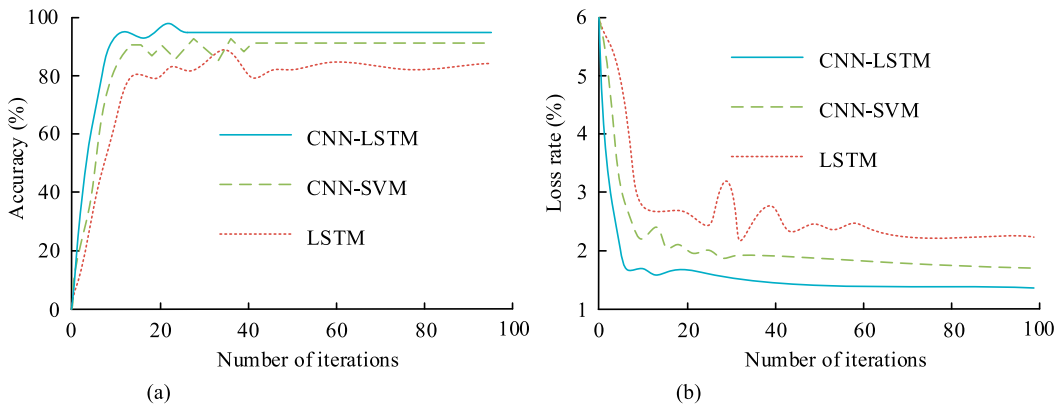

FIGURE 10
Training accuracy and loss rate of each model. **(a)** Accuracy of different algorithms **(b)** Loss rate of different algorithm.

TABLE 4 Performance indicators of different methods on RBCFD and RBRFCFD (average of 10 experiments).

Dataset	Method	Accuracy (%)	Recall (%)	F1-score (%)	Convergence epochs
RBCFD	MIC-SPCI + CNN	72.75	72.97	71.64	35
	EMD-AADPCI+LSTM	85.25	85.32	85.17	50
	EMD-AADPCI + CNN-SVM	91.25	91.67	91.25	40
	EMD-AADPCI + CNN-LSTM	98.00	98.72	98.33	22
RBRFCFD	MIC-SPCI + CNN	71.50	71.83	70.92	38
	EMD-AADPCI+LSTM	84.75	85.01	84.89	52
	EMD-AADPCI + CNN-SVM	90.80	91.22	91.01	42
	EMD-AADPCI + CNN-LSTM	97.80	98.55	98.18	23

EMD-FPCI was 13.75% higher that of MIC-SPCI, mainly due to the multi-scale decomposition and denoising effect of EMD on the signal. EMD-AADPCI has further improved by 11.50% compared to EMD-FPCI, which directly quantifies the performance gain brought by the adaptive angle allocation mechanism and proves its core innovative value. When using the same EMD-AADPCI features, the performance of the CNN-LSTM hybrid model was significantly better than that of a single CNN or LSTM model, verifying the effectiveness of model fusion.

3.2 Composite fault diagnosis analysis based on CNN-LSTM diagnostic model

After deciding to use EMD-AADPCI as the feature extraction algorithm, an algorithm for fault classification is further selected. Common fault classification algorithms include CNN, ANN, SVM, LSTM, etc. The study combines CNN with LSTM to design a CNN-LSTM composite diagnostic model for classification tasks. When verifying the effectiveness of vibration images generated by EMD-AADPCI, experiments have been conducted in conjunction with CNN. Therefore, the LSTM and CNN-SVM models are compared with the proposed CNN-LSTM, and verified on the RBRFCFD using EMD-AADPCI. Figure 10 presents the results.

In Figure 10a, the training accuracy of a single LSTM was low and fluctuated greatly. It gradually stabilized after 50 iterations, with a training accuracy of 81.07%. The training accuracy of the CNN-SVM was relatively high, stabilizing at 91.45% after 40 iterations. The CNN-LSTM model converged to 94.88% after only 22 iterations. In Figure 10b, the loss rate of the LSTM gradually converged to 2.53% after 60 iterations. The loss rate of the CNN-SVM converged to 1.92% after 30 iterations. The loss rate of the CNN-LSTM model stabilized at 1.43% after 15 iterations. The CNN-LSTM demonstrates better performance, higher training accuracy, and lower loss rate. The performance comparison of different methods on the RBCFD and RBRFCFD is shown in Table 4.

According to Table 4, the comprehensive performance of the proposed EMD-AADPCI + CNN-LSTM model is significantly better than other comparison methods on both datasets. On the RBCFD, the accuracy of the model reached 98.00%, the recall rate was 98.72%, and the F1 value was 98.33%. Compared with methods such as MIC-SPCI + CNN (accuracy 72.75%) and EMD-AADPCI + LSTM (accuracy 85.25%), all performance indicators were significantly improved. On the RBRFCFD, the model still maintained a leading performance, with accuracy, recall, and F1 score reaching 97.80%, 98.55%, and 98.18%, respectively, fully demonstrating good generalization ability. In addition, the training convergence efficiency of this model was also outstanding: It only

TABLE 5 Comparison of costs calculated by different methods.

Dataset	Method	Single-sample image generation time (ms)	Model training time (min/100 epochs)	Single-sample inference time (ms)	Number of parameters ($\times 10^4$)	FLOPs (G)
RBCFD	MIC-SPCI + CNN	8.2	12.5	3.1	128	2.8
	EMD-AADPCI+LSTM	17.3	18.6	4.5	156	3.5
	EMD-AADPCI + CNN-SVM	17.1	15.2	3.8	142	3.2
	EMD-AADPCI + CNN-LSTM	17.5	22.3	5.2	218	4.3
RBRFCFD	MIC-SPCI + CNN	8.5	13.1	3.3	128	2.8
	EMD-AADPCI+LSTM	17.8	19.2	4.7	156	3.5
	EMD-AADPCI + CNN-SVM	17.4	15.8	4.0	142	3.2
	EMD-AADPCI + CNN-LSTM	17.6	23.1	5.4	218	4.3

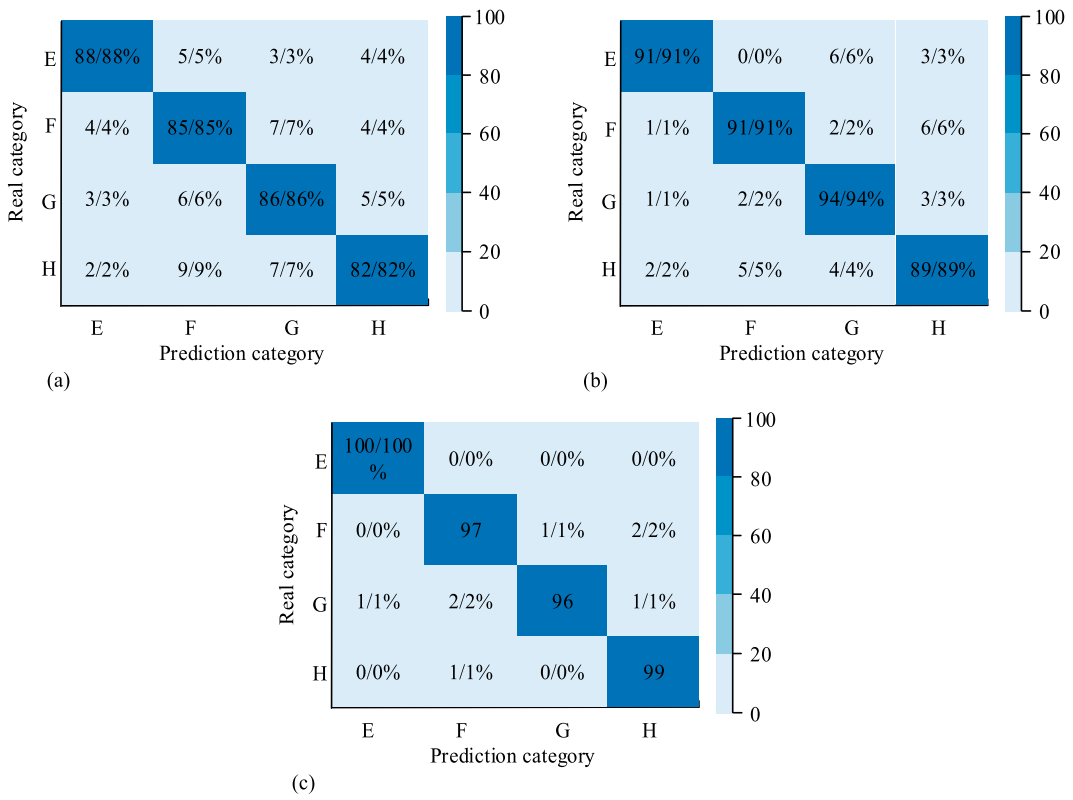

FIGURE 11
Confusion matrix of fault diagnosis results based on LSTM, CNN-SVM, and CNN-LSTM. **(a)** Sample testing confusion matrix for LSTM **(b)** Sample testing confusion matrix for CNN-SVM **(c)** Sample testing confusion matrix for CNN-LSTM.

required 22 rounds on the RBCFD and 23 rounds on the RBRFCFD, which was much lower than that of other comparison methods. To evaluate the practicality, the computational costs (including single-sample image generation time, model training time, and single-sample inference time) of different methods were tested based on a hardware environment with an Intel Core i7-12700H processor, 32 GB of memory, and an NVIDIA RTX 3060 graphics card. The results are shown in Table 5.

As shown in Table 5, the image generation time of EMD-AADPCI was slightly longer than that of MIC-SPCI (due to the

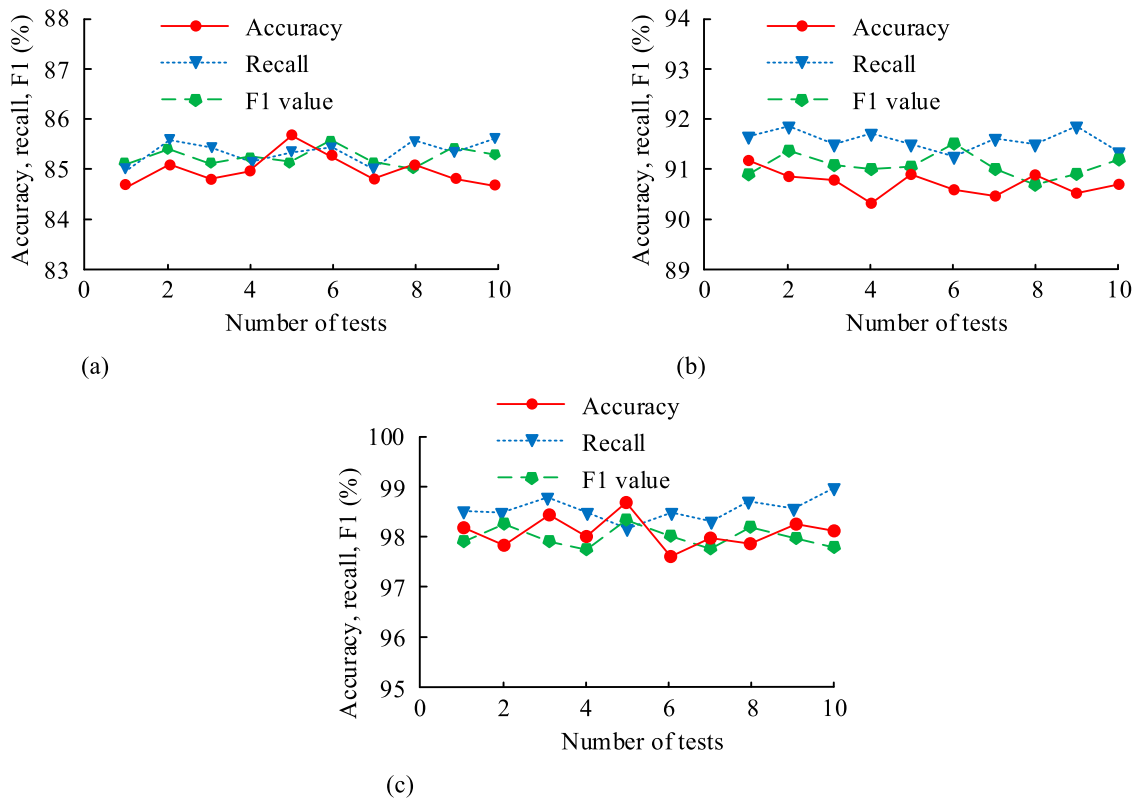

FIGURE 12
The average accuracy, recall, and F1 value of three models. **(a)** Various indicators based on LSTM model **(b)** Various indicators based on CNN-SVM model **(c)** Various indicators based on CNN-LSTM model.

EMD and adaptive angle calculation), but the single-sample generation time was still controlled within 20 ms, meeting the real-time requirements. Although the number of parameters and FLOPs of the CNN-LSTM model was higher than that of a single model, the actual training time (convergence in 22–23 rounds) was only 8–10 min, and the single-sample inference time was about 5 ms, which was far below the threshold (50 ms) for real-time diagnosis in industrial scenarios. Compared with the Transformer model (approximately five million parameters and approximately 12G of FLOPs), the computational cost of the EMD-AADPCI + CNN-LSTM was reduced by more than 60%, making it more suitable for deployment on edge computing devices. 400 sample images are randomly selected from the RBRFCFD, with 100 images for each type of fault. The confusion matrix diagnosed by the three types of models is shown in Figure 11.

In Figure 11a, the fault classification accuracy of a single LSTM was relatively low, only 85.25%. The classification accuracy for types E and G was relatively high, at 88.00% and 86.00%, respectively. According to Figure 11b, the classification accuracy of the CNN-SVM was 91.25%, with the highest classification accuracy of 94.00% for type G. In Figure 11c, the CNN-LSTM was 98.00%. The classification accuracy of this model for four types of faults was 100%, 97.00%, 96.00%, and 99.00%, respectively. The experiment conducted was 10 times. The average accuracy, recall, and F1 value are obtained, as presented in Figure 12.

According to Figure 12a, the average classification accuracy of the LSTM was only 84.79%, with an average recall and F1 value of 85.32% and 85.17%, respectively. In Figure 12b, compared with the LSTM model, the CNN-SVM had significantly improved performance, with an average classification accuracy, recall, and F1 value of 90.89%, 91.67%, and 91.25%. In Figure 12c, the CNN-LSTM had significantly higher performance indicators than the other two models, with an average classification accuracy, recall, and F1 value of 98.13%, 98.72%, and 98.33%. The CNN-LSTM demonstrates better model performance and higher classification accuracy. Due to the diverse types of composite faults in bearings under actual working conditions, the noise mixed in their signals is also diverse. To verify the noise resistance of these three models, Gaussian white noise of −4~4 dB was added to the RBCFD and RBRFCFD, and comparative experiments were conducted. The classification accuracy of three models on the dataset with added noise is presented in Figure 13.

In Figure 13a, on the RBCFD, LSTM and CNN-SVM had lower classification accuracy when SNR<0 dB. However, when SNR was between 0 and 4dB, the accuracy of these two models slightly improved, stabilizing at around 85.00% and 90.00%, respectively. The classification accuracy of CNN-LSTM was relatively stable at different SNR values, reaching 98.89% at SNR = 4 dB. In Figure 13b, on the RBRFCFD, the classification accuracy of LSTM and CNN-SVM was significantly lower than that of CNN-LSTM. When the SNR was -2 dB and 4dB, the accuracy of CNN-LSTM was relatively high, reaching 97.63% and 98.12%, respectively. The reasons why CNN-LSTM still maintains high accuracy at low SNR are as follows: ① EMD can suppress more than 80% of high-frequency noise, while

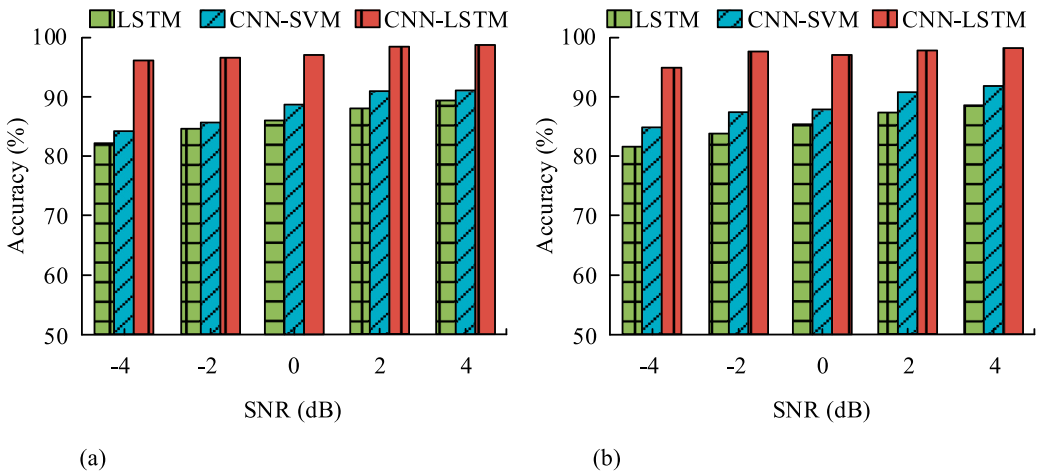

FIGURE 13
Comparison of SNR and fault classification accuracy. **(a)** Fault classification accuracy on RBCFD **(b)** Fault classification accuracy on RBRFCFD.

AADPCI disperses the remaining noise to low-weight regions in polar coordinates to avoid noise masking fault features; ② The convolutional layer of CNN can filter isolated noise points in the image by focusing on the spatial features of faults through local receptive fields; The forget gate of LSTM can ignore temporal noise interference and only retain the temporal evolution law of fault features. This result indicates that the proposed model has stronger anti-noise performance, which is suitable for fault diagnosis tasks under complex working conditions.

4 Discussion and future work

4.1 Discussion and conclusion

For composite fault diagnosis, the EMD-AADPCI vibration image generation method was first proposed, followed by the CNN-LSTM fault diagnosis model. The results indicated the vibration image generated by EMD-AADPCI contained richer features, and the image discrimination of various faults was more obvious. After combining MIC-SPCI and EMD-AADPCI with CNN for validation on the RBCFD, the classification accuracy of MIC-SPCI was only 72.75%. The average classification accuracy, average recall, and F1 value after 10 experiments were 72.76%, 72.97%, and 71.64%, respectively. The classification accuracy of EMD-AADPCI was 93.00%. After 10 experiments, the average classification accuracy, average recall, and F1 value were 91.73%, 92.38%, and 92.15%, respectively. The vibration images generated by EMD-AADPCI had better performance and more accurate feature extraction. The reason is that the angle allocation of AADPCI is adaptive and dynamically adjusted based on the local characteristics of IMF components. This adaptive approach enables the generated polar coordinate image to contain richer and more accurate fault feature information. The proposed EMD-AADPCI is compared with existing vibration image generation methods. Spectrum diagram calculation is simple, but it can only present frequency domain information, making it difficult to distinguish composite faults. The Gray level Co-occurrence Matrix (GCM) requires a large amount of computation, and the feature representation effect depends on parameter settings. MIC-SPCI improves feature visibility through polar coordinates, but due to fixed angle allocation, it is difficult to adapt to subtle differences in different faults. EMD-AADPCI first performs adaptive decomposition on the signal through EMD, highlighting multi-scale fault characteristics, and then uses adaptive angle allocation driven by time-domain indicators to amplify the differences between faults. Although its computational complexity is slightly higher than that of spectrograms, it is significantly lower than that of GCM and has an absolute advantage in diagnostic accuracy.

In terms of fault classification, the study compared CNN-LSTM with LSTM and CNN-SVM. The results showed that CNN-LSTM had faster convergence speed, higher training accuracy, and lower loss rate, which were 94.88% and 1.43%, respectively. The fault classification accuracy of LSTM was 85.25%, the classification accuracy of CNN-SVM was 91.25%, and the overall classification accuracy of CNN-LSTM was as high as 98.00%. After 10 experiments, the average classification accuracy, average recall, and F1 value of CNN-LSTM were 98.13%, 98.72%, and 98.33%, respectively. Meanwhile, the anti-noise performance of CNN-LSTM was also stronger. The CNN-LSTM model demonstrates better performance and higher classification accuracy for various types of faults. The reason is that CNN-LSTM combines the characteristics of CNN in extracting local spatial features and LSTM in extracting long-term sequence dependencies, thereby more effectively identifying the features of various faults and improving classification accuracy. In addition, compared with the proposed CNN-LSTM with other advanced intelligent diagnostic methods, although the EEMD-based method can alleviate modal aliasing, it has higher computational cost and manual selection of sensitive IMF. Models based on attention mechanism or Transformer are good at capturing long-range dependencies, but the model structure is complex, requiring not only a large amount of data to support training, but also higher requirements for computing resources. The Graph Convolutional Network (GCN) is a type of neural network method that can mine the correlation features between samples, but it relies on the sample topology

structure and has strict adaptability to data formats. It also additional preprocesses topological relationships, which raises the threshold for practical deployment. The CNN-LSTM architecture is clear and efficient, with CNN responsible for extracting spatial features and LSTM capturing temporal dynamic patterns. It can effectively fuse spatiotemporal information without introducing complex attention modules, ensuring high diagnostic accuracy while controlling model complexity and training costs within a reasonable range.

In summary, the method can solve the composite fault diagnosis, providing new technical support for the normal operation. The performance of this method demonstrates its potential for application in practical industrial environments, which can assist in predictive maintenance of equipment. Although the research focuses on rolling bearings, the proposed EMD-AADPCI generation concept and CNN-LSTM diagnostic framework have excellent scalability and are expected to be applied in the field of fault diagnosis of other rotating machinery such as gearboxes and gears in the future.

4.2 Limitations and future work

Although the proposed EMD-AADPCI performs well in the diagnosis of composite faults in rolling bearings, there are still limitations: (1) Limited coverage of operating conditions: Experimental data are mainly obtained under fixed load and speed conditions, and have not fully covered more complex industrial operating conditions such as variable loads and speeds. This may affect the model's generalization ability in dynamic environments. (2) Real-time challenge: The serial computation process of EMD and CNN-LSTM model results in relatively high computational costs. In online monitoring scenarios that require extremely high real-time performance, the inference delay of current methods may become a deployment bottleneck. (3) Diversity of fault types: The experiment mainly focuses on several preset composite fault modes, and the diagnostic ability for more diverse or unknown composite fault combinations needs further verification.

In response to the above limitations, future research work will be carried out from the following aspects: (1) Extended operating condition verification: To verify the robustness and generalization ability of the method under different loads, speeds, and more diverse composite fault types. (2) Model lightweight and deployment optimization: Exploring model lightweight technologies, such as using lightweight backbone networks like MobileNet, or using model pruning, quantification and other means to reduce computing complexity and memory occupation, and meet the deployment needs of edge computing devices. (3) Enhanced feature extraction capability: Attempting to integrate attention mechanisms (such as SE module and CBAM) into CNN or LSTM modules to enhance the model's ability to focus on key fault features, further improving diagnostic accuracy and anti-interference. (4) Exploring end-to-end architecture: Researching the joint optimization of signal preprocessing (EMD) and diagnostic networks or designing lightweight end-to-end networks to reduce process steps and improve overall efficiency.

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

QS: Conceptualization, Methodology, Validation, Visualization, Writing – original draft. FY: Funding acquisition, Investigation, Supervision, Writing – review and editing.

Funding

The author(s) declared that financial support was received for this work and/or its publication. The research is supported by: Cultivation Object of the Seventh Phase of Jiangsu Province's "333 Project". No. (2024) 3-0604.

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.

References

Agajie, T. F., Ali, A., Fopah-Lele, A., Amoussou, I., Khan, B., Velasco, C. L. R., et al. (2023). A comprehensive review on techno-economic analysis and optimal sizing of hybrid renewable energy sources with energy storage systems. *Energies* 16 (2), 642. doi:10.3390/en16020642

Agajie, T. F., Fopah-Lele, A., Amoussou, I., Khan, B., Bajaj, M., Zaitsev, I., et al. (2024). Enhancing Ethiopian power distribution with novel hybrid renewable energy systems for sustainable reliability and cost efficiency. *Sci. Rep.* 14 (1), 10711. doi:10.1038/s41598-024-61413-8

Aljemely, A. H., Xuan, J., Al-Azzawi, O., and Jawad, F. K. (2022). Intelligent fault diagnosis of rolling bearings based on LSTM with large margin nearest neighbor algorithm. *Neural Comput. Appl.* 34 (22), 19401–19421. doi:10.1007/s00521-022-07353-8

Amiri, A. F., Oudira, H., Chouder, A., and Kichou, S. (2024). Faults detection and diagnosis of PV systems based on machine learning approach using random forest classifier. *Energy Convers. Manag.* 301, 118076. doi:10.1016/j.enconman.2024.118076

An, Z., Wu, F., Zhang, C., Ma, J., Sun, B., Tang, B., et al. (2023). Deep learning-based composite fault diagnosis. *IEEE J. Emerg. Sel. Top. Circuits Syst.* 13 (2), 572–581. doi:10.1109/jetcas.2023.3262241

Bäßler, R., Bäßler, T., and Kley, M. (2022). Classification of load and rotational speed at wire-race bearings using convolutional neural networks with vibration spectrograms. *Tm-Technisches Mess.* 89 (5), 352–362. doi:10.1515/teme-2021-0143

Chen, X., Zhang, B., and Gao, D. (2021). Bearing fault diagnosis base on multi-scale CNN and LSTM model. *J. Intelligent Manufacturing* 32 (4), 971–987. doi:10.1007/s10845-020-01600-2

Deng, W., Guan, H., and Zhao, H. (2025a). Parameterized iterative time-frequency-multisqueezing transform for bearing fault diagnosis. *IEEE Trans. Instrum. Meas.* 74, 1–11. doi:10.1109/tim.2025.3561399

Deng, W., Li, H., and Zhao, H. (2025b). Anti-noise bearing fault diagnosis using time-reassigned multisynchrosqueezing transform and complex sparse learning dictionary. *IEEE Trans. Instrum. Meas.* 74, 1–10. doi:10.1109/tim.2025.3604987

Fan, C., Peng, Y., Shen, Y., Guo, Y., Zhao, S., Zhou, J., et al. (2024). Variable scale multilayer perceptron for helicopter transmission system vibration data abnormity beyond efficient recovery. *Eng. Appl. Artif. Intell.* 133, 108184. doi:10.1016/j.engappai.2024.108184

Guo, Y., Mao, J., and Zhao, M. (2023a). Rolling bearing fault diagnosis method based on attention CNN and BiLSTM network. *Neural Processing Letters* 55 (3), 3377–3410. doi:10.1007/s11063-022-11013-2

Guo, J., Wang, J., Wang, Z., Gong, Y., Qi, J., Wang, G., et al. (2023b). A CNN-BiLSTM-Bootstrap integrated method for remaining useful life prediction of rolling bearings. *Qual. Reliab. Eng. Int.* 39 (5), 1796–1813. doi:10.1002/qre.3314

Guo, D., Zhang, S., Zhang, J., Yang, B., and Lin, Y. (2025a). Exploring contextual knowledge-enhanced speech recognition in air traffic control communication: a comparative study. *IEEE Trans. Neural Netw. Learn. Syst.* 36, 16085–16099. doi:10.1109/TNNLS.2025.3569776

Guo, D., Zhang, J., Yang, B., and Lin, Y. (2025b). Multi-modal intelligent situation awareness in real-time air traffic control: control intent understanding and flight trajectory prediction. *Chin. J. Aeronautics* 38 (6), 10337. doi:10.1016/j.cja.2024.103376

Kaya, Y., Kuncan, M., Kaplan, K., Minaz, M. R., and Ertunç, H. M. (2021). A new feature extraction approach based on one dimensional gray level co-occurrence matrices for bearing fault classification. *J. Exp. Theor. Artif. Intell.* 33 (1), 161–178. doi:10.1080/0952813x.2020.1735530

Li, Y., Zhou, J., Li, H., Meng, G., and Bian, J. (2022). A fast and adaptive empirical mode decomposition method and its application in rolling bearing fault diagnosis. *IEEE Sensors J.* 23 (1), 567–576. doi:10.1109/jsen.2022.3223980

Li, X., Wang, Y., Yao, J., Li, M., and Gao, Z. (2024). Multi-sensor fusion fault diagnosis method of wind turbine bearing based on adaptive convergent viewable neural networks. *Reliab. Eng. Syst. Saf.* 245, 109980. doi:10.1016/j.ress.2024.109980

Li, J., Deng, W., Ding, J., and Zhao, H. (2025). IBN-MixStyle network with dynamic weighted invariant risk minimization for domain-generalized bearing fault diagnosis. *IEEE Trans. Consumer Electron.* 71, 9929–9939. doi:10.1109/tce.2025.3607134

Liu, Y., Kang, J., Bai, Y., and Guo, C. (2023). Research on the health status evaluation method of rolling bearing based on EMD-GA-BP. *Qual. Reliab. Eng. Int.* 39 (5), 2069–2080. doi:10.1002/qre.3350

Lv, M., Yan, C., Kang, J., Meng, J., Wang, Z., Li, S., et al. (2024). Multiple faults separation and identification of rolling bearings based on time-frequency spectrogram. *Struct. Health Monit.* 23 (4), 2040–2067. doi:10.1177/14759217231197110

Meng, D., Wang, H., Yang, S., Lv, Z., Hu, Z., and Wang, Z. (2021). Fault analysis of wind power rolling bearing based on EMD feature extraction. *CMES-Computer Model. Eng. Sci.* 130 (1), 543–558. doi:10.32604/cmes.2022.018123

Özmen, N. G., and Karabacak, Y. E. (2023). A new bearing fault diagnosis approach based on common spatial pattern features. *Niğde Ömer Halisdemir Üniversitesi Mühendislik Bilim. Derg.* 12 (4), 1545–1557. doi:10.28948/ngumuh.1330864

Preethi, P., and Mamatha, H. R. (2023). Region - based convolutional neural network for segmenting text in epigraphical images. *Artif. Intell. Appl.* 1 (2), 119–127. doi:10.47852/bonviewaia2202293

Sun, H. B., and Fan, Y. G. (2023). Fault diagnosis of rolling bearings based on CNN and LSTM networks under mixed load and noise. *Multimedia Tools Appl.* 82 (28), 43543–43567. doi:10.1007/s11042-023-15325-w

Tao, H., Qiu, J., Chen, Y., Stojanovic, V., and Cheng, L. (2023). Unsupervised cross-domain rolling bearing fault diagnosis based on time-frequency information fusion. *J. Frankl. Inst.* 360 (2), 1454–1477. doi:10.1016/j.jfranklin.2022.11.004

Toumi, Y., Bengherbia, B., Lachenani, S., and Ould Zmirli, M. (2022). FPGA implementation of a bearing fault classification system based on an envelope analysis and artificial neural network. *Arabian J. Sci. Eng.* 47 (11), 13955–13977. doi:10.1007/s13369-022-06599-7

Wang, H., Liu, Z., Peng, D., Yang, M., and Qin, Y. (2021). Feature-level attention-guided multitask CNN for fault diagnosis and working conditions identification of rolling bearing. *IEEE Transactions Neural Networks Learning Systems* 33 (9), 4757–4769. doi:10.1109/TNNLS.2021.3060494

Xing, Z., Zhao, R., Wu, Y., and He, T. (2022). Intelligent fault diagnosis of rolling bearing based on novel CNN model considering data imbalance. *Appl. Intell.* 52 (14), 16281–16293. doi:10.1007/s10489-022-03196-x

Zhao, Y., Fan, Y., Li, H., and Gao, X. (2022). Rolling bearing composite fault diagnosis method based on EEMD fusion feature. *J. Mech. Sci. Technol.* 36 (9), 4563–4570. doi:10.1007/s12206-022-0819-x

Zhao, H., Liu, C., Dang, X., Xu, J., and Deng, W. (2025). Few-shot cross-domain fault diagnosis of transportation motor bearings using MAML-GA. *IEEE Trans. Transp. Electrification*, 1. doi:10.1109/tte.2025.3625779

Zheng, J., Chen, Y., Pan, H., and Tong, J. (2023). Composite multi-scale phase reverse permutation entropy and its application to fault diagnosis of rolling bearing. *Nonlinear Dyn.* 111 (1), 459–479. doi:10.1007/s11071-022-07847-z

Zhou, J., Peng, Y., Shao, H., Shen, Y., Bin, G., Zheng, J., et al. (2025). Empirical fourier-bessel heuristic denoising and its application to gear fault diagnosis. *ISA Transactions* 167, 1906–1924. doi:10.1016/j.isatra.2025.10.003

Zhou, J., Bin, G., Peng, Y., Shao, H., Shen, Y., Zheng, J., et al. (2026). Shrinkage sliding fourier-bessel packet: a coarse-to-fine local narrowband filtering method. *Adv. Eng. Inf.* 69, 104074. doi:10.1016/j.aei.2025.104074