

Complex fault diagnosis in wind turbine bearings: a hybrid approach combining the improved feature mode decomposition and convolutional neural networks

Yongbin Du¹, Hengyu Wang², Yuanhai Zhao³, Kunwang Sun⁴, Yi Zhang⁵

College of Electrical and Control Engineering, Heilongjiang University of Science and Technology, Harbin, China

¹Corresponding author

E-mail: ¹duyongbin1@163.com, ²1461918620@qq.com, ³zhaoyuanhai0@163.com, ⁴1911657287@qq.com, ⁵2609700915@qq.com

Received 11 April 2025; accepted 4 July 2025; published online 12 August 2025
DOI <https://doi.org/10.21595/jve.2025.24974>



Copyright © 2025 Yongbin Du, et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract. The complex noise interference and diverse fault-induced signals in vibration data from wind turbine equipment pose significant challenges for bearing fault diagnosis, including cumbersome methodologies, prolonged processing times, and compromised accuracy. To address these limitations, this study proposes a novel composite fault diagnosis framework that integrates Feature Mode Decomposition (FMD), Fast Spectral Kurtosis (FSK), and Convolutional Neural Network (CNN). While conventional Empirical Mode Decomposition (EMD) exhibits limited noise robustness and struggles to extract subtle fault signatures in composite failure scenarios, our approach employs FMD to decompose fault-related intrinsic mode functions (IMFs) and further filters the IMF components using fast spectral cliffs with enhanced feature separability. Subsequently, the Short-Time Fourier Transform (STFT) is applied to derive time-frequency representations, followed by Fast Spectral Kurtosis analysis to identify optimal demodulation bands for non-stationary signals. The energy spectrum of denoised signals is converted into grayscale images, serving as input to a tailored CNN architecture for hierarchical feature learning. Experimental validation demonstrates that this hybrid methodology achieves a fault recognition accuracy of 98 % under compound fault conditions, outperforming conventional EMD-based approaches in terms of noise immunity and diagnostic precision. Comparative analysis reveals an 8 % improvement in detection reliability over standalone deep learning models, particularly in low signal-to-noise ratio (SNR) environments. The proposed framework offers a robust solution for multi-fault identification in industrial Bearing machinery, demonstrating superior generalization capability across varying operational conditions.

Keywords: rolling bearing, composite fault diagnosis, eigenmode decomposition, deep learning.

1. Introduction

Wind power is a renewable energy technique that transforms wind energy into electricity. It is both clean and renewable, while also playing a crucial role in diminishing greenhouse gas emissions and alleviating global climate change. The global population is projected to reach 9.8 billion by 2050, with annual power demand potentially surpassing 38,000 terawatt-hours (TWh). Wind energy, recognized as a primary source of clean energy, is anticipated to meet the majority of global electricity demand [1]. Nonetheless, as the number of wind farms has increased, the issue of wind turbine failures has become increasingly significant, given that areas abundant in wind resources are typically situated in remote locations characterized by a relatively pristine environment. The natural environment in these regions is comparatively severe. Wind Turbine Generators, as substantial machinery functioning over extended periods in the area, will be influenced by the natural environment for an extended duration. Consequently, blades, main bearings, gearboxes, and other critical components are prone to numerous failures; inadequate fault identification and remediation may lead to severe engineering catastrophes, resulting in

substantial economic losses. Real-time surveillance of bearings during operation can significantly prolong their lifespan [2]. By examining prevalent bearing failure modes and suggesting efficient problem diagnosis methods accordingly, mechanical equipment can be maintained in optimal operational condition [3].

Currently, numerous methodologies for bearing fault diagnosis, both domestically and internationally, predominantly rely on collected vibration and electrical signals [4]. Vibration signals are widely used in fault diagnosis due to their rich informational content and ease of acquisition, as the collection apparatus for vibration sensors and detection systems can be easily integrated into existing control systems.

The presence of noise, interference, and other issues necessitates the pre-processing of the acquired signal as a preliminary step for bearing fault diagnosis. Commonly employed signal processing techniques include Empirical Mode Decomposition (EMD) [5], Wavelet Packet Transform (WPT), and Hilbert-Huang Transform (HHT) [6]. Xiang Dan [7] et al., proposed a method based on empirical modal decomposition (EMD) entropy feature fusion to solve the rolling bearing fault diagnosis problem by taking the entropy of nonlinear dynamics parameter of the source signal as a feature; Wu Jichun [8] et al. used the Hilbert yellow transform method in combination with empirical modal decomposition to convert the decomposed screened and reconstructed sensitive signals into time-frequency diagrams, which were used as inputs to convolutional neural networks. data; Hoang [9] et al. proposed the direct conversion of vibration signals into ten two-dimensional grayscale maps, utilizing these images as inputs for a convolutional neural network, which attained commendable accuracy despite the lack of denoising. Li Zhinong [10] et al. integrated wavelet transform with empirical mode decomposition by constructing a wavelet filter to refine the delineated frequency spectrum and employed a Hilbert yellow transform to analyze the energy spectrum. Cheng X. [20] et al. developed a correlation indicator employing a Bayesian transformer to forecast the future lifespan of bearings based on health indicators. While EMD is suitable for linear data, there remains a likelihood of modal aliasing issues during data processing.

Variational Mode Decomposition (VMD) [11] is an adaptive signal processing technique that adeptly addresses the modal aliasing issue by identifying the optimal center frequency and bandwidth for each modal component. However, it necessitates the manual selection of a smoothing parameter, which can be excessively large or small, thereby influencing the modal components. Additionally, the VMD algorithm exhibits suboptimal performance when confronted with multi-modal signals, high noise environments, or non-smooth signals. The efficacy of the VMD algorithm diminishes when confronted with multi-modal signals, high noise environments, or non-stationary signals, leading to incomplete decomposition in composite fault scenarios. Consequently, it employs Feature Mode Decomposition (FMD) [13] for signal preprocessing; however, the FMD method remains susceptible to performance degradation in the presence of significant noise interference.

A Convolutional Neural Network (CNN) [12] is a deep learning model with significant relevance in image processing and speech recognition. The approach utilizes convolutional processes to extract local features from the input data, enabling automatic extraction of multi-level features and significantly diminishing the engineering demands for manual feature extraction. Nonetheless, the input data for the convolutional neural network must be denoised, making the handling of this data a crucial aspect in enhancing the network's accuracy.

In summary, this paper presents a fault diagnosis method that integrates FMD and CNN, simultaneously addressing signal processing and image recognition tasks. The primary contributions of this study are as follows: 1) To enhance the algorithm's applicability to non-periodic signals and its robustness in high-noise environments, a method for filtering FMD-decomposed signals based on FSK is developed. 2) A dataset for convolutional neural network input is constructed based on grayscale images, which effectively enhances image features while optimizing storage and processing efficiency, thereby improving the learning efficiency of the neural network. 3) A convolutional neural network model is trained to address

complex composite fault diagnosis issues. 4) The proposed method effectively decomposes significant IMF components with high accuracy, offering a novel approach to tackling composite fault problems. The feasibility of this method is validated through simulation experiments and the composite fault dataset from Xi'an Jiaotong University.

2. Signal pre-processing

2.1. Feature model decomposition (FMD)

Yonghao Miao [13] et al. published the FMD algorithm in IEEE Transactions on Industrial Electronics in 2023. The FMD algorithm's primary goal is to identify the defect information in the signal by utilizing the period estimation and designing an adaptive FIR filter bank. The FMD algorithm is not only more targeted but also more resilient to other disturbances and noise signals when decomposing the target signals. The process of FMD decomposition is as follows:

1) Input the acquired original signal, specifying the mode quantity and filter length.

2) Initialize the FIR filter bank using a Hanning window, constructing frequency-selective filters with predefined cutoff frequencies f_l, f_u :

$$f_l = \frac{k \cdot f_s}{2K}, \quad f_u = \frac{(k+1) \cdot f_s}{2K}, \quad k = 0, 1, \dots, K-1, \quad (1)$$

where: the sampling frequency (f_s) is defined by the system.

3) Calculate to obtain u_k^i (filtered signal):

$$u_k^i = x * f_k^i, \quad k = 1, 2, \dots, K, \quad (2)$$

where: $*$ denotes the convolution operation.

4) Maximize the filter parameter CK (Correlation Kurtosis) using the original signal, the filtered signal, and the estimated period to enhance modal separation effectiveness:

$$CK(u_k) = \frac{(\sum_{n=1}^N \prod_{m=0}^M u_k(n - mT_s))^2}{(\sum_{n=1}^N u_k(n)^2)^{M+1}}, \quad (3)$$

where: T_s is the estimation period, M is the shift order.

5) Calculate the correlation coefficient CC_{pq} between mode u_p and mode u_q . If the two modes exhibit high correlation, retain the mode with the larger CK (Correlation Kurtosis) value, thereby reducing redundancy:

$$CC_{pq} = \frac{\sum_{n=1}^N (u_p(n) - \bar{u}_p)(u_q(n) - \bar{u}_q)}{\sqrt{\sum_{n=1}^N (u_p(n) - \bar{u}_p)^2} \sqrt{\sum_{n=1}^N (u_q(n) - \bar{u}_q)^2}} \quad (4)$$

6) Iteratively repeat the filtering and mode selection process until the predefined number of modes is obtained.

2.2. Short time Fourier transform (STFT)

Short-Time Fourier Transform (STFT) [14] is a technique for analyzing time-frequency signals that combines time and frequency information to illustrate the frequency's evolution over time. At the same time, STFT will divide the signal into distinct time segments and subsequently employ a sliding window (such as the Hanning window) to intercept the signal that has been accurately segmented. The Fourier Transform of the signal for each time segment will be

performed. In comparison to the Hilbert yellow transform and other time-frequency analysis methods, the short-time Fourier transform is more straightforward and efficient, making it an ideal choice for practical engineering applications. The time-frequency diagram produced by the short-time Fourier transform is also capable of providing a highly intuitive representation of the signal's abrupt changes and characteristics, rendering it an ideal input for the convolutional neural network.

The feature extraction process for the Short-Time Fourier Transform (STFT) is as follows:

1) The input signal $x(t)$ is divided into localized time segments for analysis. During this process, a window function $w(t)$ (e.g., Hanning or Hamming window) is applied to each segment to mitigate spectral leakage caused by abrupt signal truncation. This window function acts as a weighting function, emphasizing the signal within the target time interval while attenuating the signal at the segment edges. The windowed signal segment at time position m is expressed as:

$$x(t) \cdot w(t - \tau), \quad (5)$$

where: τ denotes the center time index of the current segment, $w(t - \tau)$ is the window function (e.g., Gaussian, Hanning).

2) Perform the Fourier Transform on each windowed time segment to obtain its frequency-domain representation. The transformed signal at time index and frequency bin is given by:

$$X(\tau, \omega) = \int_{-\infty}^{+\infty} x(t) \cdot w(t - \tau) \cdot e^{-j\omega t} dt, \quad (6)$$

where: $X(\tau, \omega)$ denotes the signal spectrum at frequency bin ω and time index, $e^{-j\omega t}$ is the Fourier transform kernel function, which converts time-domain signals to frequency-domain representations.

3) Slide the window function to different positions, perform the Fourier Transform at each corresponding time position, and thereby obtain the time-frequency distribution of the entire signal.

4) Calculate the squared magnitude of the STFT result to generate the power spectrum, which serves as input to a convolutional neural network (CNN).

The power spectrum is defined as:

$$\text{Spectrogram}(\tau, \omega) = |X(\tau, \omega)|^2. \quad (7)$$

2.3. Fast spectral kurtosis

Fast Spectral Kurtosis (FSK) [15] analyzes the target signal by constructing a binary tree-structured filter bank and generates a kurtogram by computing the kurtosis of the signal within each frequency subband. Kurtosis [16], a dimensionless higher-order statistical measure, is defined as:

$$K = \frac{E(x - \mu)^4}{\sigma^4}, \quad (8)$$

where: x represents the target signal, μ is the mean value of x , and σ denotes the standard deviation of x .

This parameter indicates the degree of deviation of the target signal from Gaussian at a specific frequency. When a localized fault occurs in a rolling bearing, the probability density of the shock vibration signal induced by the fault increases, and the distribution of the signal amplitude will obviously deviate from a normal distribution, with the cragness value increasing [17].

2.4. Principles of improved FMD

The FMD method is more specifically designed to address the composite fault problem. However, in the presence of significant noise, the decomposed IMF components still contain fewer key features and a greater number of noise signals.

This paper employs the crag function as a discriminant benchmark to enhance the accuracy of the FMD method for composite bearing faults. It then employs the FSK method to calculate the crag values and fast spectral crag maps of the original function and each IMF component that undergoes FMD decomposition. By comparing the band intervals where the calculated crag values are located, the components that overlap with the intervals where the crags of the original signals are located are screened out and reconstructed.

This methodology capitalizes on the inherent capability of kurtosis functions to effectively characterize the probability density distribution of transient impulses within specific frequency bands. By establishing correspondence between original signals and IMF components through kurtosis domain alignment, the proposed method successfully preserves information-rich components containing predominant fault features. Consequently, it significantly improves the feature-to-noise ratio in reconstructed signals while enhancing their suitability as optimized inputs for convolutional neural networks in subsequent fault recognition tasks.

3. Convolutional neural network

The Convolutional Neural Network (CNN) [12] is a deep learning model that is highly applicable in applications such as speech recognition and image processing. Through convolutional operations, the algorithm extracts local features from the input data, thereby significantly reducing the engineering requirements for manual feature extraction. This process can automatically extract multi-level features from the data. Simultaneously, the multilayer architecture of deep learning can significantly enhance the capacity to extract small and concealed defect features. Simultaneously, the convolutional neural network is capable of analyzing the signal of the operating equipment in real time after training, rendering it highly appropriate for the actual working conditions.

Core Components of a Convolutional Neural Network:

1) Convolutional Layer: As the foundational component of a Convolutional Neural Network (CNN), the convolutional layer employs convolutional kernels to systematically traverse the input image (or feature map), extracting localized spatial features. The number of learnable kernels directly determines the layer's feature representation capacity-increasing the kernel count enhances both the diversity and granularity of captured features.

The mathematical operation of convolution is defined as:

$$(f * g)(i, j) = \sum_m \sum_n f(m, n) \cdot g(i - m, j - n), \quad (9)$$

where: f is the input feature map, g is the convolutional kernel.

2) Activation Function: The activation function is applied after each convolutional operation to introduce nonlinearity into the model, enabling the network to learn complex hierarchical patterns such as transient fault impulses in vibration signals or texture variations in images.

3) Fully Connected Layer: Positioned after convolutional and pooling layers, the fully connected (FC) layer integrates global features extracted from localized patterns. It maps flattened feature vectors into class probabilities through learnable weight matrices, facilitating decision-making in classification tasks. For instance, in bearing fault diagnosis, FC layers combine spectral energy distributions and temporal impulsivity features to distinguish between healthy and faulty states.

4) Loss Function: The loss function quantifies the discrepancy between predicted outputs and

ground-truth labels. Optimization algorithms iteratively minimize this loss to enhance model accuracy.

5) Backpropagation and Optimizer: Backpropagation computes gradients of the loss with respect to network parameters using the chain rule, while optimizers such as Adam (adaptive moment estimation) or SGD (stochastic gradient descent) update weights to minimize the loss. Adam, for example, adjusts learning rates dynamically based on gradient moments, improving convergence in non-stationary industrial data.

6) Pooling Layer: The pooling layer reduces feature map dimensions through downsampling operations (e.g., max-pooling), which retains critical spatial or temporal features while suppressing noise and redundant details. This dimensionality reduction mitigates computational complexity and overfitting, particularly in high-sample-rate vibration data.

4. Bearing composite fault diagnosis based on improved FMD and convolutional neural network

To address the challenges of improving noise robustness in the Frequency Modulation Demodulation (FMD) method and mitigating the risks of manual misjudgment and omission in bearing composite fault diagnosis, we propose a novel diagnostic framework integrating improved FMD and a Convolutional Neural Network (CNN). The workflow is illustrated in Fig. 1.

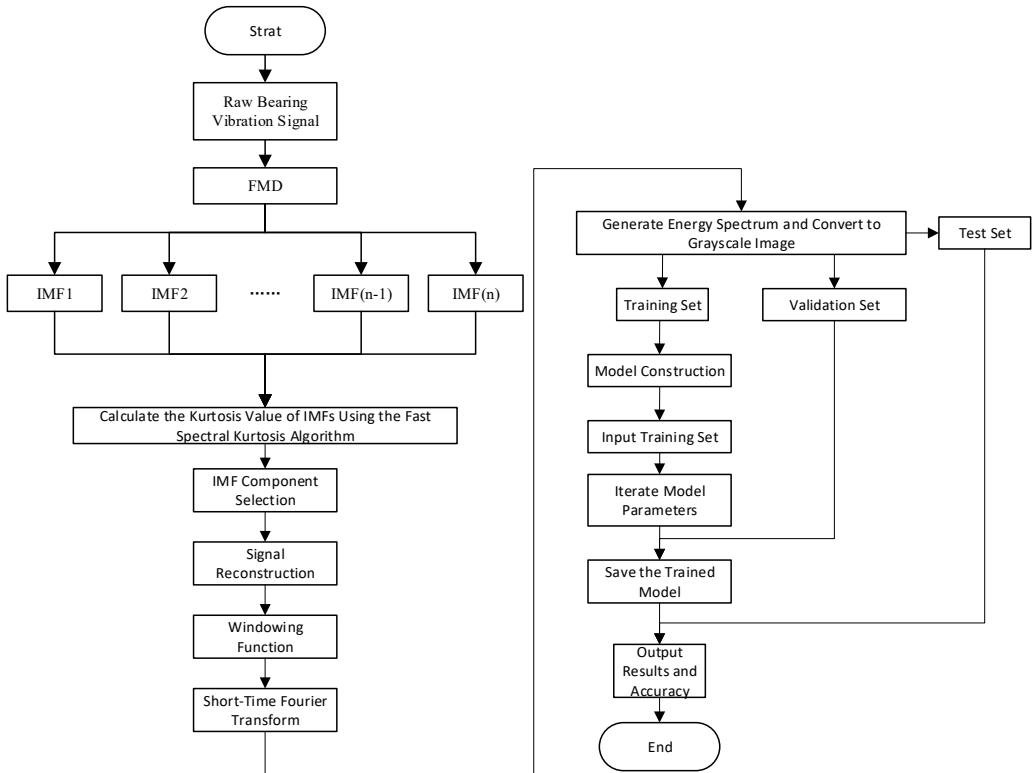


Fig. 1. Bearing compound fault diagnosis process based on FMD-FSK-CNN

Fig. 1 detailed steps outlined:

- 1) Signal Decomposition via FMD: The raw vibration signal is decomposed using FMD to extract multiple Intrinsic Mode Function (IMF) components.
- 2) Target IMF Component Selection: Calculate the spectral kurtosis of each IMF component to quantify its impulsiveness and non-Gaussianity. Identify the maximum kurtosis value and its

corresponding characteristic frequency for each IMF. Select IMF components whose characteristic frequencies fall within the original signal’s frequency band as target components.

3) Signal Reconstruction: Reconstruct the filtered IMF components to synthesize a noise-suppressed signal enhanced with fault-related features.

4) Time-Frequency Analysis: Apply Short-Time Fourier Transform (STFT) to the reconstructed signal and compute its energy spectrum to capture transient fault dynamics.

5) Dataset Preparation: Convert the energy spectrum into a grayscale image and partition it into training, validation, and test sets for model development.

6) CNN Model Training: Train the preconfigured CNN model using the training set to learn discriminative fault patterns. Save the optimized model weights after training completion.

7) Hyperparameter Tuning: Evaluate the model’s performance on the validation set to refine hyperparameters (e.g., learning rate, dropout rate). Store the fine-tuned model for subsequent testing.

8) Model Evaluation: Deploy the finalized model on the test set to assess its diagnostic accuracy and generalization capability.

5. Simulation analysis

A mixed simulation signal is constructed to evaluate the reasonableness of the proposed algorithm, which consists of a shock signal, a harmonic signal, and non-periodic signal. As demonstrated in Eq. (10):

$$x(t) = x_1(t) + x_2(t) + x_3(t),$$
$$x_1(t) = 5e^{-\theta_1 \cdot t} \cos(2\pi \cdot 1200 \cdot t), \quad x_2(t) = (1 + \cos(2\pi \cdot 50t)) \cdot \sin(2\pi 600t), \tag{10}$$

where $x_3(t)$ are noise signals with signal-to-noise ratio of 6. Their time domain waveforms are shown in Fig. 2

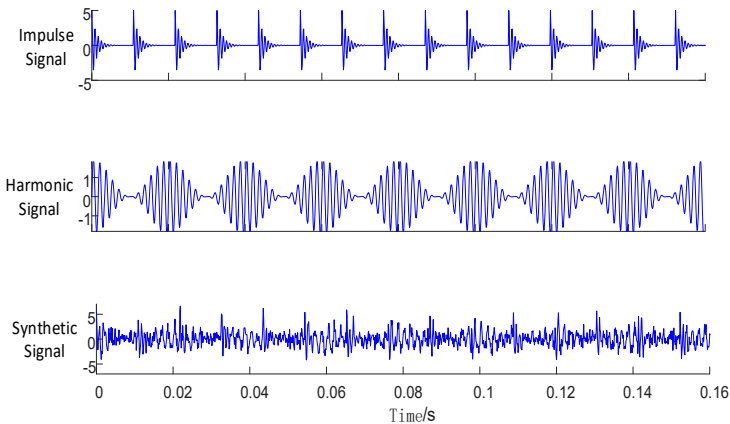


Fig. 2. Time-domain waveforms of the simulated signal and its components

The signal’s modal decomposition is executed utilizing EMD and FMD, respectively, while modal screening and reconstruction are conducted through the FSK method and the conventional correlation coefficient method, respectively. Envelope diagrams for the three methods will be provided to facilitate a comparative analysis of the decomposition and screening effects.

Fig. 3 illustrates that the signal envelope reconstructed through correlation coefficient screening following EMD algorithm decomposition exhibits a greater number of impactful spectral lines, accompanied by numerous trans-frequency sidebands adjacent to the eigenfrequency spectral lines, which lack sufficient clarity for convolutional neural network input.

Conversely, after FMD algorithm decomposition and correlation coefficient screening, the reconstructed signal envelope displays more distinct spectral lines; however, the eigenfrequency spectral lines remain inadequately pronounced. Following the decomposition of the FMD algorithm and the application of FSK screening for signal reconstruction, the envelope map exhibits reduced clutter. A notable increase in amplitude signifies a substantial enhancement of fault characteristics, thereby providing a distinct advantage in subsequent image recognition and fault diagnosis.

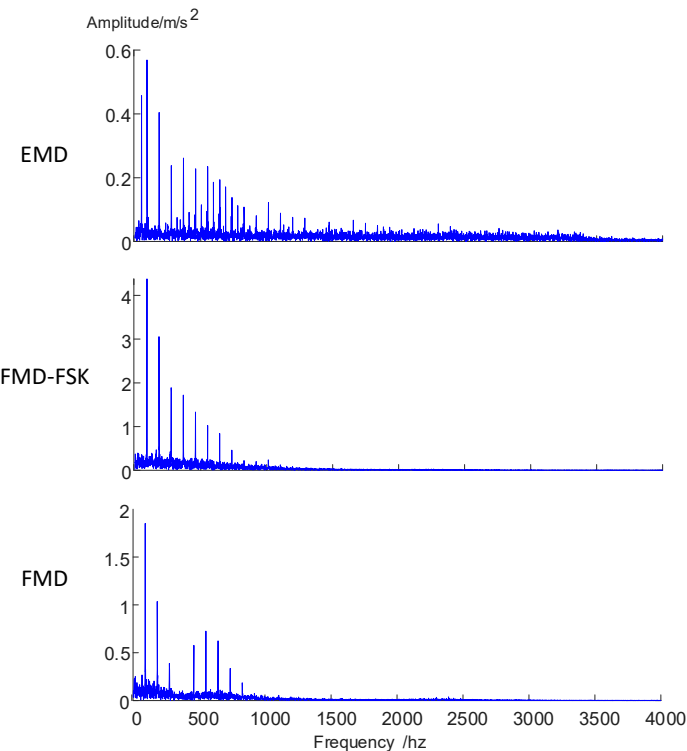


Fig. 3. Envelope diagrams of FMD, FMD-FSK, and EMD

6. Case study

6.1. Experimental data description

The experimental dataset utilized in this study is the open-access XJTU-SY Rolling Bearing Dataset [18], jointly developed by Xi'an Jiaotong University (XJTU) and ShengYang Technology Co., Ltd. under the supervision of Prof. Yaguo Lei. Data acquisition was performed using a customized bearing accelerated life test rig, which incorporates LDK UER204 bearings with the following specifications:

- Inner race diameter: 29.30 mm.
- Outer race diameter: 39.80 mm.
- Ball diameter: 7.92 mm.
- Operating speed: 2400 rpm.

6.2. Digital signal preprocessing

The experiment utilises the segment of the dataset with a collection frequency of 10 kn, encompassing four fault types: inner-race fault, outer-race fault, composite fault of the inner and

outer race (IOF), and composite fault involving the inner and outer race along with the rolling elements (IORF). Fig. 5 illustrates the time-domain signal waveforms for each fault type, revealing that composite faults exhibit a higher amplitude compared to single fault types; however, distinguishing them based on waveform characteristics proves challenging.

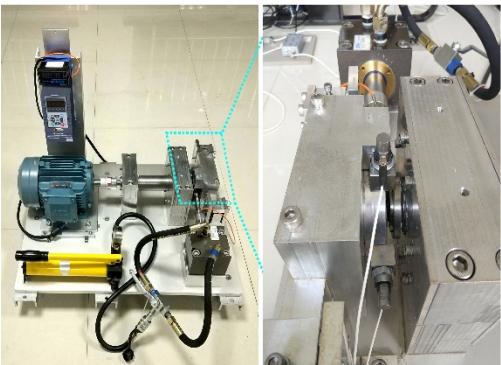


Fig. 4. Experimental platforms [18]

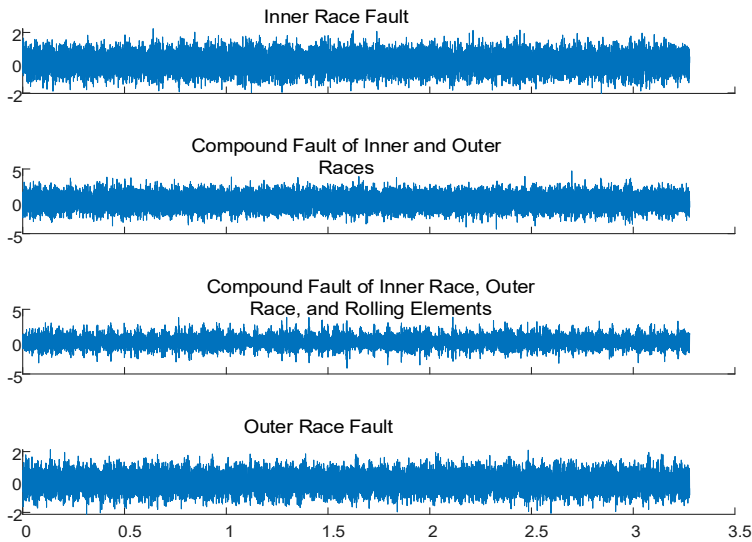


Fig. 5. Bearing fault signals

To address the issue, the signal in the dataset is initially dissected utilizing the FMD technique, as illustrated in Fig. 6, which displays the first four IMF components of a defective signal post-FMD decomposition. All four defective signals are subjected to this form of decomposition. Following signal decomposition, it is essential to compute the spectral gradient of both the original signal and each IMF component. The maximum spectral gradient must be identified, and a comparison should be made to ascertain whether the frequency band corresponding to the maximum gradient of each IMF component falls within the band of the maximum gradient of the original signal. This process is illustrated in Fig. 7, which presents the spectral gradient diagrams for the original signal and several IMF components. The maximum gradient of the original signal is 6.4, occurring within the frequency range of 4192 Hz to 5000 Hz. The IMF3 component exhibits a maximum gradient within the interval of 4220 Hz to 4335 Hz, which is situated within the original signal's maximum gradient range, thus classifying it as an effective component. Conversely, the IMF2 component's target interval does not overlap with that of the original signal and is therefore excluded. The approach is utilized for all vibration signals, and if several valid

IMF components exist, the chosen IMF components are then reconstructed.

To enhance the precision of subsequent experiments, the validated and reconstructed signals are converted into spectrograms via the short-time Fourier transform, and these spectrograms are normalized into visual grayscale maps to serve as inputs for the convolutional neural network.

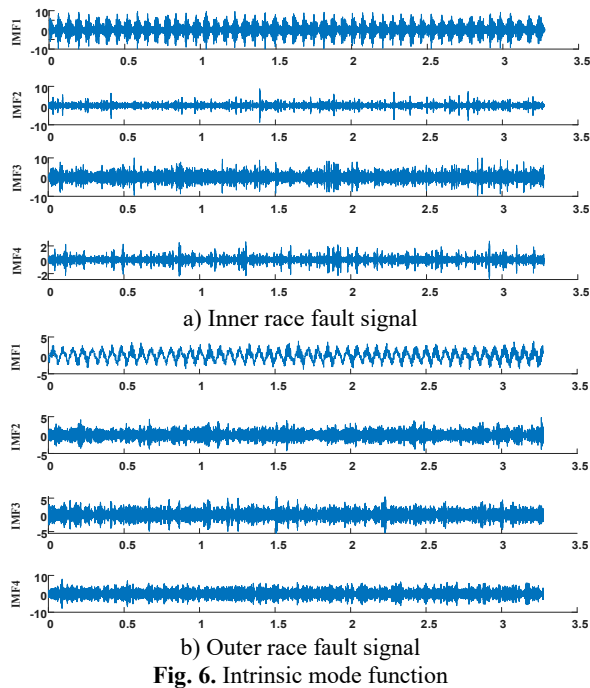


Fig. 6. Intrinsic mode function

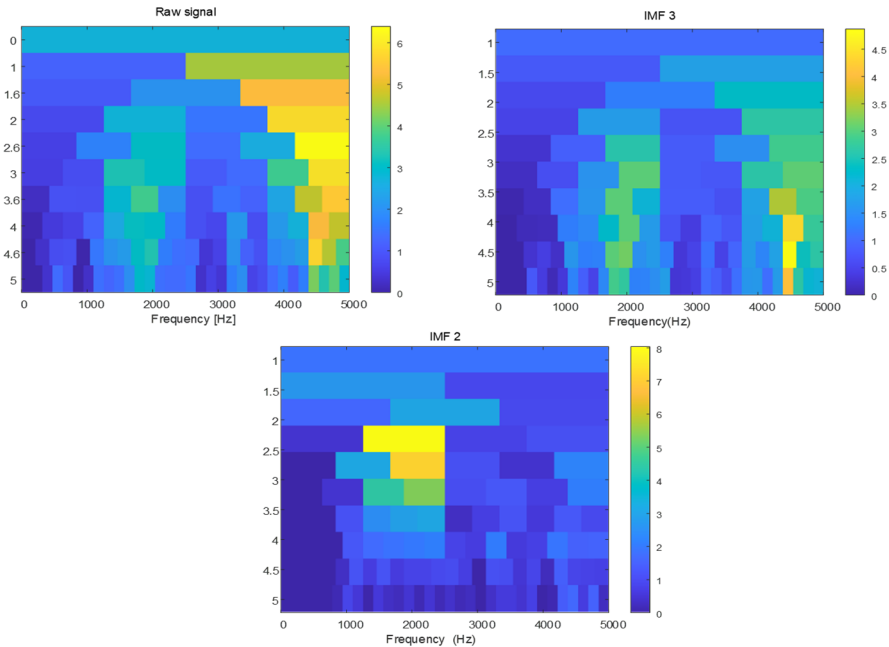


Fig. 7. Spectral kurtosis diagrams of the raw signal and IMF components

To enhance the accuracy of subsequent experiments, the screened valid signals and

reconstructed signals are converted into time-frequency spectrograms using the STFT. The generated spectrograms are then normalized into grayscale images, which serve as standardized inputs to the CNN.

6.3. Convolutional neural network parameters

To augment the experimental efficacy and boost the precision of the convolutional neural network, the AlexNet model [19] serves as a foundation for modifying certain model parameters, optimizing performance, and identifying parameters conducive to bearing defect diagnostics. The precise characteristics are described in Table 1.

Table 1. Convolutional network parameters

Network structure type	Structural parameters
Input layer	256×256 image
Convolutional layer C1	32 3×3 kernels
Pooling layer P1	Max pooling with 2×2 kernel
Convolutional layer C2	64 3×3 kernels
Pooling layer P2	Max pooling with 2×2 kernel
Convolutional layer C3	128 3×3 kernels
Pooling layer P3	Max pooling with 2×2 kernel
AlphaDropout layer	0.5

6.4. Results

This experiment classifies bearing failure into four categories: inner ring failure, outer ring failure, composite failure of the inner and outer rings, and composite failure of the inner ring and rolling bodies. This categorization occurs during the generation of the grayscale map through signal preprocessing. Subsequently, 500 randomly selected grayscale maps are allocated into a training set, validation set, and test set, with the training set comprising 70 %, the validation set 10 %, and the test set 20 %.

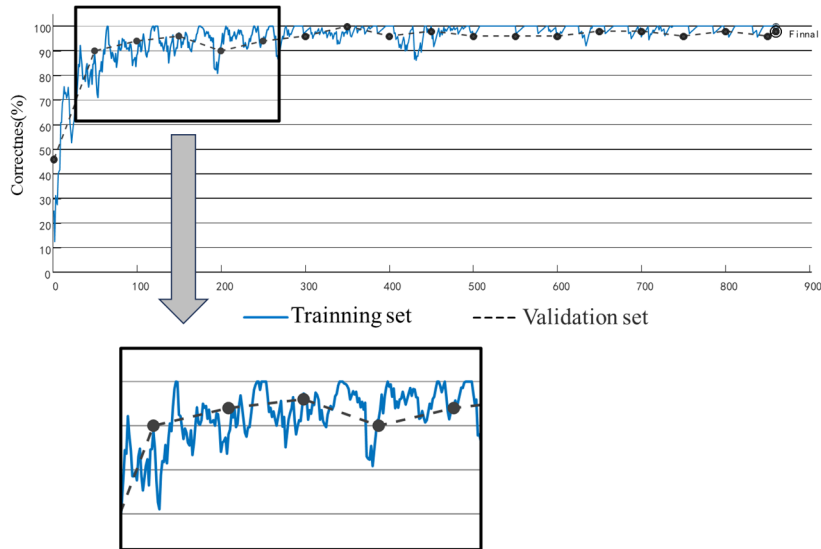


Fig. 8. Spectral kurtosis diagrams of the raw signal and IMF components

The apparatus utilized for the experiment comprises an 11th Gen Intel(R) Core(TM) i7-11800H processor and an NVIDIA GeForce RTX 3060 Laptop GPU, with a total of 800 iterations scheduled. Fig. 8 illustrates the training outcomes, indicating that the accuracy of the

experimental samples ascends to 100 % for the first time at the 90th iteration, while the accuracy of the validation set concurrently achieves 100 % without exhibiting overfitting conditions. The remaining test sets are incorporated into the finalized training model for prediction, with results illustrated in Fig. 9. The horizontal axis represents the predicted fault type, while the vertical axis denotes the actual fault type. The model achieves an average accuracy of 98 %, with a diagnostic accuracy of 100 % for mixed fault types. This demonstrates that the methodology presented in the paper effectively diagnoses composite faults of varying types, accurately identifies the fault signals associated with each type, and subsequently diagnoses the fault signals of different fault types. defects, identifying the attributes of each defect signal.

Inner Race Fault	35			1
IOF	1	9		
IORF	1		32	
Outer Race Fault	1			21
	Inner Race Fault	IOF	IORF	Outer Race Fault

Fig. 9. Model prediction results diagram

6.5. Comparison experiment

This research will employ various methodologies for experimental comparison to validate the superiority of the proposed approach.

Fig. 10 illustrates that the FMD-FSK-CNN and FMD-CNN exhibit minimal disparity in the convergence speed of the validation set's accuracy. Although the initial accuracy of the validation set is slightly different, the FMD-FSK-CNN method sustains an accuracy of 98 % and achieves 100 % at certain intervals during subsequent iterations. The accuracy of FMD-CNN varies between 90 % and 95 %. As illustrated in Fig. 11(a) and Fig. 11(b), both techniques achieve an accuracy of one hundred percent for composite defects, with minimal error observed in instances of inner and outer ring faults. This comparative experiment demonstrates that fault diagnosis with the FMD approach exhibits superior diagnostic capabilities for composite fault issues, while the accuracy of the enhanced FMD method has been further elevated, along with a notable improvement in stability.

In comparing the FMD-FSK method with the EMD-FSK method, as illustrated in Fig. 10, the EMD algorithm exhibits a smoother curve post-convergence of the validation set; however, its accuracy stabilizes at approximately 90 percent. Furthermore, the confusion matrix in Fig. 11(c) indicates that the EMD method struggles to effectively decompose features when confronted with multiple composite fault scenarios, resulting in a higher incidence of misdiagnosis across various fault types.

The FMD-FSK-CNN method is juxtaposed with the FMD-FSK-BP method. As illustrated in Fig. 11(d), the convergence rate of the BP neural network is markedly inferior to that of the convolutional neural network. Furthermore, the accuracy achieved post-convergence fails to meet expectations, while the validation set accuracy exhibits significant fluctuations. The confusion matrix indicates a substantial number of errors in the model's performance following training, thereby demonstrating the inadequacy of the BP neural network when addressing two-dimensional images.

In comparing the FMD-FSK-CNN method with the VMD-CNN method, as illustrated in Fig. 10, the conventional VMD method exhibits a lower initial accuracy on the validation set compared to the FMD method. Additionally, its convergence rate is slower than that of the

FMD-FSK-CNN method. Although it demonstrates strong stability post-convergence, its accuracy is only sustained at 90 %. Fig. 11(e) illustrates that the conventional VMD approach remains inadequate for the composite fault issue, exhibiting a higher rate of misdiagnosis and demonstrating marginally inferior overall accuracy compared to the FMD-FSK-CNN method.

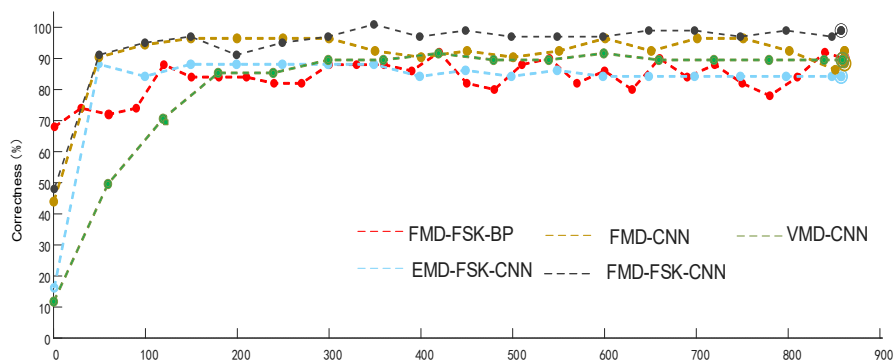


Fig. 10. Correctness of the validation set for each method

Inner Race Fault	35			1
IOF	1	9		
IORF	1		32	
Outer Race Fault	1			21
	Inner Race Fault	IOF	IORF	Outer Race Fault

a) Confusion matrix for FMD-CNN accuracy

Inner Race Fault	29	3		4
IOF	2	8		1
IORF			21	
Outer Race Fault	3		1	29
	Inner Race Fault	IOF	IORF	Outer Race Fault

c) Confusion matrix for EMD-FSK-CNN accuracy

Inner Race Fault	21			2
IOF	3	11		
IORF			33	
Outer Race Fault	3			27
	Inner Race Fault	IOF	IORF	Outer Race Fault

b) Confusion matrix for FMD-FSK-CNN accuracy

Inner Race Fault	21			2
IOF	3	11		
IORF			33	1
Outer Race Fault	3		1	27
	Inner Race Fault	IOF	IORF	Outer Race Fault

d) Confusion matrix for FMD-FSK-BP accuracy

Inner Race Fault	26			
IOF		8		
IORF		2	30	1
Outer Race Fault		1	4	29
	Inner Race Fault	IOF	IORF	Outer Race Fault

e) Confusion matrix for VMD-CNN accuracy

Fig. 11. Confusion matrix for each method

Table 2. Test set accuracy for each method

Usage	Correctness / %
FMD-FSK-CNN	96.2 %
FMD-CNN	92 %
EMD-FSK-CNN	86.1 %
FMD-FSK-BP	90.0 %
VMD-CNN	92.1 %

7. Conclusions

The article introduces the FMD-FSK-CNN approach for bearing failure diagnosis, validated through the integration of the FMD algorithm with FSK and the application of convolutional neural networks.

The FMD algorithm and FSK, when combined, effectively screen out components with strong features, further enhancing the processing capability for composite fault problems and strengthening the noise immunity capability. This method is more applicable and can better distinguish the fault features in a strong noise environment, while also amplifying the feature signals to make them serve as high-quality CNN inputs. In comparison to traditional EMD and other methods. Information.

The integration of CNN and FMD-FSK methods and their comparison experiments have resulted in an 8 % increase in the accuracy of the verification set, which can more accurately identify the type of each composite fault. This demonstrates the method's intelligence, superiority, and effectiveness.

Future Outlook: Nevertheless, the method described in this paper still necessitates the manual adjustment of certain parameters in the context of FMD. The modification of this parameter has a significant impact on the effect's functionality and is contingent upon the user's experience. There is no more consistent evaluation standard, necessitating further optimization of the FMD method to meet the demand. However, the screening of IMF components decomposed by FMD using the FSK technique requires a longer time than the EMD method, necessitating further optimization.

Acknowledgements

The authors have not disclosed any funding.

Data availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Author contributions

Yongbin Du: conceptualization, methodology, writing-review and editing, supervision, project administration and funding acquisition. Hengyu Wang: conceptualization, methodology, software, writing-original draft preparation, writing-review and editing. Yuanhai Zhao: software, writing-review and editing. Kunwang Sun: software, validation, formal analysis, data curation. Yi Zhang: validation, formal analysis, investigation.

Conflict of interest

The authors declare that they have no conflict of interest.

References

- [1] P. Veers et al., "Grand challenges in the science of wind energy," *Science*, Vol. 366, No. 6464, Oct. 2019, <https://doi.org/10.1126/science.aau2027>
- [2] C. H. Xie, X. Wang, and And F. H. Xiao, "A review of bearing fault modes and fault diagnosis methods," (in Chinese), *Computer Measurement and Control*, 2025, <https://doi.org/10.16526/j.cnki.11-4762/tp.2025.04.001>
- [3] H. X. Wei, "Research methods and advances in fault diagnosis of rolling bearings," (in Chinese), *Journal of Tianjin University of Technology*, Vol. 37, No. 2, pp. 36–40, 2021, <https://doi.org/10.3969/j.issn.1673-095x.2021.02.008>
- [4] J. Q. Li, Z. F. Wang, and And L. Wang, "A motor bearing fault diagnosis method based on current signals and deep reinforcement learning," (in Chinese), *Electric Power Science and Engineering*, Vol. 39, No. 3, pp. 61–70, 2023, <https://doi.org/10.3969/j.issn.1672-0792.2023.03.008>
- [5] N. E. Huang et al., "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, Vol. 454, No. 1971, pp. 903–995, Mar. 1998, <https://doi.org/10.1098/rspa.1998.0193>
- [6] N. E. Huang, Z. Shen, and S. R. Long, "A new view of nonlinear water waves: the Hilbert spectrum," *Annual Review of Fluid Mechanics*, Vol. 31, No. 1, pp. 417–457, Jan. 1999, <https://doi.org/10.1146/annurev.fluid.31.1.417>
- [7] Y. F. Gao and And J. P. Zhang, "Bearing fault diagnosis based on EMD and FastICA," (in Chinese), *Machine Design and Manufacturing*, Vol. 6, pp. 48–52, 2024, <https://doi.org/10.19356/j.cnki.1001-3997.2024.06.004>
- [8] J. C. Wu, G. X. Yang, and K. Xu, "Tool wear state recognition based on EEMD-FK and attention CNN network," (in Chinese), *Journal of Computer Integrated Manufacturing Systems*, Vol. 29, No. 10, 2023, <https://doi.org/10.13196/j.cims.2023.10.000>
- [9] D.-T. Hoang and H.-J. Kang, "Rolling element bearing fault diagnosis using convolutional neural network and vibration image," *Cognitive Systems Research*, Vol. 53, pp. 42–50, Jan. 2019, <https://doi.org/10.1016/j.cogsys.2018.03.002>
- [10] Z. N. Li, M. Zhu, and F. L. Chu, "Research on mechanical fault diagnosis method based on empirical wavelet trans-form," (in Chinese), *Journal of Instrumentation and Measurement*, Vol. 35, No. 11, 2014, <https://doi.org/10.19650/j.cnki.cjsi.2014.11.003>
- [11] K. Dragomiretskiy and D. Zosso, "Variational mode decomposition," *IEEE Transactions on Signal Processing*, Vol. 62, No. 3, pp. 531–544, Feb. 2014, <https://doi.org/10.1109/tsp.2013.2288675>
- [12] K. Fukushima, "Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position," *Biological Cybernetics*, Vol. 36, No. 4, pp. 193–202, Apr. 1980, <https://doi.org/10.1007/bf00344251>
- [13] Y. Miao, B. Zhang, C. Li, J. Lin, and D. Zhang, "Feature mode decomposition: new decomposition theory for rotating machinery fault diagnosis," *IEEE Transactions on Industrial Electronics*, Vol. 70, No. 2, pp. 1949–1960, Feb. 2023, <https://doi.org/10.1109/tie.2022.3156156>
- [14] D. Griffin and Jae Lim, "Signal estimation from modified short-time Fourier transform," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol. 32, No. 2, pp. 236–243, Apr. 1984, <https://doi.org/10.1109/tassp.1984.1164317>
- [15] J. Antoni, "Fast computation of the kurtogram for the detection of transient faults," *Mechanical Systems and Signal Processing*, Vol. 21, No. 1, pp. 108–124, Jan. 2007, <https://doi.org/10.1016/j.ymssp.2005.12.002>
- [16] J. Antoni, "The spectral kurtosis: a useful tool for characterising non-stationary signals," *Mechanical Systems and Signal Processing*, Vol. 20, No. 2, pp. 282–307, Feb. 2006, <https://doi.org/10.1016/j.ymssp.2004.09.001>
- [17] X. T. Wu, M. Yang, and And X. H. Yuan, "Bearing fault diagnosis based on kurtosis criterion EEMD and improved morpho-logical filtering method," (in Chinese), *Journal of Vibration and Shock*, Vol. 34, No. 2, pp. 38–44, 2015, <https://doi.org/10.13465/j.cnki.jvs.2015.02.007>
- [18] Y. G. Lei, T. Han, B. Wang, N. Li, T. Yan, and J. Yang, "XJTU-SY rolling element bearing accelerated life test datasets: a tutorial," *Journal of Mechanical Engineering*, Vol. 55, No. 16, pp. 1–6, Jan. 2019, <https://doi.org/10.3901/jme.2019.16.001>

- [19] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Communications of the ACM*, Vol. 60, No. 6, pp. 84–90, May 2017, <https://doi.org/10.1145/3065386>
- [20] X. Chen, K. Li, S. Wang, and H. Liu, "A hybrid prognostic approach combined with deep Bayesian transformer and enhanced particle filter for remaining useful life prediction of bearings," *Measurement*, Vol. 252, p. 117184, Aug. 2025, <https://doi.org/10.1016/j.measurement.2025.117184>



Yongbin Du, Associate Professor at Heilongjiang University of Science and Technology. He obtained his bachelor's degree from Zhengzhou University in 2000, his master's degree from Harbin University of Science and Technology in 2015, and his Ph.D. from Harbin Institute of Technology in 2022. His current main research direction is fault diagnosis and sensor calibration.



Hengyu Wang is a master's degree holder from Heilongjiang University of Science and Technology. He received his B.S. degree in engineering from Northeast Electric Power University in 2022. His current main interest in research is bearing fault diagnosis.



Yuanhai Zhao is a graduate student at the School of Electrical and Control Engineering, majoring in Electrical Engineering from Heilongjiang University of Science and Technology. He received his bachelor's degree from Henan College of Urban Construction in 2023. His current research focuses on the design of flexible ultrasonic transducer arrays.



Kunwang Sun is a master's student in the School of Electrical and Control Engineering at Heilongjiang University of Science and Technology. He received his B.E. degree from School of Electrical and Control Engineering at Heilongjiang University of Science and Technology University in 2023. His current major research direction is flexible ultrasonic array.



Yi Zhang is a master's student in the School of Electrical and Control Engineering at Heilongjiang University of Science and Technology. He received his B.E. degree from Shandong Agricultural University in 2020. His current major research direction is short-term power load forecasting.