frontiers | Frontiers in Mechanical Engineering

TYPE Original Research
PUBLISHED 29 August 2025
DOI 10.3389/fmech.2025.1647310

OPEN ACCESS

EDITED BY
Hamid Reza Karimi,
Polytechnic University of Milan, Italy

REVIEWED BY
Dip Saha,
Rajshahi University of Engineering and Technology, Bangladesh
Tuann Le,
RMIT University Vietnam, Vietnam
Chandrabhanu Malla,
Radhakrishna Institute of Technology and Engineering, India

*CORRESPONDENCE
Daiqiao Guo,
guodaiqiao1988@163.com

RECEIVED 15 June 2025
ACCEPTED 15 August 2025
PUBLISHED 29 August 2025

CITATION
Guo D (2025) Motor bearing fault diagnosis based on industrial internet of things and transfer learning.
Front. Mech. Eng. 11:1647310.
doi: 10.3389/fmech.2025.1647310

Motor bearing fault diagnosis based on industrial internet of things and transfer learning

Daiqiao Guo*

School of Intelligent Manufacturing, Hebi Polytechnic, Hebi, China

Introduction: Motor bearing faults seriously affect industrial safety and efficiency. Traditional diagnosis methods often lack adaptability across conditions and rely heavily on labeled data. This study proposes a fault diagnosis model integrating the Industrial Internet of Things (IIoT) and transfer learning to improve robustness and generalization.

Methods: The model combines CNN-BiGRU for spatio-temporal feature extraction, an adaptive multi-source feature fusion mechanism (Adaboost with dynamic cutting), and a dual-source domain transfer module based on joint maximum mean discrepancy (JMMD). Embedded in an IIoT platform, it supports real-time sensing, cross-domain adaptation, and closed-loop diagnosis-optimization.

Results: Using the Jiangnan University bearing dataset, the model achieved 94.7% diagnostic accuracy with a 6.8% false alarm rate, and a maximum state recognition match of 97.3%. In practical tests, accuracy remained above 93% under diverse conditions, with a response time of 19.6 s and GPU usage of only 8.9%. Compared with CNN, SqNet, and K-GCN, the proposed method showed superior robustness, efficiency, and cross-condition generalization.

Discussion: The IIoT-TL model effectively enhances predictive maintenance by reducing costs and improving equipment stability. While it depends on sensor configurations and has not been tested on more complex heterogeneous devices, it demonstrates strong potential for real-time deployment in industrial systems. Future work will explore lightweight frameworks and broader application scenarios.

KEYWORDS

IIoT, transfer learning, bearing fault diagnosis, CNN, sensor

1 Introduction

In modern intelligent industrial production, motors are the core power equipment, and the operating status of their bearings directly affects system safety. Faults in motor bearings cause not only direct economic losses but may also lead to production chain interruptions. Therefore, fault diagnosis of motor bearings is critical for safe and sustainable production (Choudhary et al., 2022). Traditional fault diagnosis methods mainly rely on vibration signal threshold monitoring and expert judgment. These methods have limitations such as insufficient real-time performance and poor generalization across equipment (Hu et al., 2023). The industrial Internet of Things (IIoT) collects motor bearing operational data through various sensors, enabling comprehensive and multidimensional extraction of fault information. This provides a data foundation for accurate bearing fault diagnosis (Chi et al., 2022; Hu et al., 2024). Transfer learning (TL) breaks data barriers under different working conditions by applying learned fault knowledge to new scenarios (Zhu et al., 2023).

Therefore, this study developed a motor fault diagnosis method by designing a multi-source feature fusion approach grounded in the industrial Internet of Things and driven by transfer learning as its underlying principle. A motor bearing fault diagnosis model was constructed by combining the principles of IIoT and transfer learning. This model combines IIoT's advantages of real-time multi-source data sensing and TL's cross-device generalization ability. It realizes multi-source heterogeneous data acquisition and high-precision fault identification across working conditions and devices. The model aims to improve motor bearing fault diagnosis accuracy, reduce economic losses for enterprises, and provide reliable technical support for intelligent industrial production.

The innovative aspects of this research are as follows:

1. A motor bearing fault diagnosis model based on multi-source feature fusion and transfer learning, Convolutional BiGRU with Adaptive Boosting and Feature Transfer Learning (CB-AD-FTL), is proposed. The model integrates the Convolutional-BiGRU (CNN-BiGRU) structure with an adaptive multi-source feature fusion mechanism Adaboost-Dynamic Cutting Mechanism (DCM), enhancing its ability to represent multi-modal industrial signals.
2. A dual-source domain feature transfer structure is constructed, combined with the Joint Maximum Mean Discrepancy (JMMD) criterion, to achieve high-robustness transfer diagnosis across different operating conditions.
3. The proposed CB-AD-FTL model is embedded in the IIoT platform, achieving an integrated closed-loop system for perception, diagnosis, and optimisation, thereby enhancing the model's practical deployment capabilities and industrial application value.

2 Related works

IIoT technology addressed complex problems in modern digital industries by coordinating multiple functional modules comprising the perception layer, the network layer, the platform layer, as well as the application layer. Experts and scholars worldwide conducted systematic research on this topic. For example, Khan et al. developed a blockchain-based protection scheme to tackle data security risks in medical Internet of Things. Experiments verified that the scheme effectively reduced data risks and ensured medical information security (Khan et al., 2023). Heidari et al. proposed a novel spanning tree construction method combining artificial bee colony algorithm, genetic operators, and density correlation to solve the problem of device mobility affecting spanning tree algorithms in IIoT. Simulation results demonstrated that this method significantly improved data collection reliability (Heidari et al., 2024). TL enabled knowledge transfer across tasks by transferring learned information from the source domain to the target domain. Researchers applied this technology in various fields. Theodoris et al. addressed the data scarcity issue in gene network research through the development of a deep learning framework that integrated TL and attention mechanisms. This model showed high prediction accuracy and robustness during pretraining (Theodoris et al., 2023). Facing limitations caused by data sparsity in software repair, Chen et al. put forward a TL-based security vulnerability repair method. This method conducted large-scale repair training followed by fine-tuning on small datasets. Experiments confirmed it improved software repair accuracy (Chen et al., 2022). Lotfollahi et al. addressed dataset limitations in single-cell atlas analysis by combining TL and parameter optimization techniques. They proposed a single-cell architecture surgery method that integrated few parameters across datasets and enabled model weight sharing, offering a reference for large-scale single-cell collaboration (Lotfollahi et al., 2022). Alijani et al. addressed the challenge of performance prediction difficulties caused by inconsistent data distributions between the training/validation phase and the deployment phase in deep learning models. They proposed applying visual Transformers to domain adaptation and domain generalisation scenarios, thereby enhancing the model's robustness and generalisation capabilities under distribution shifts (Alijani et al., 2024). Nancy et al. aimed to improve the efficiency of brain tumour segmentation and classification tasks by proposing an improved Visual Geometry Group (VGG)-19 model based on transfer learning, combined with an attribute-aware attention mechanism and model-agnostic feature extractors, thereby enhancing segmentation accuracy and classification accuracy (Nancy and Maheswari, 2025).

Accurate diagnosis of motor bearing faults is critical for industrial safety and has become a key research focus in intelligent industry. Brusamarello et al. introduced a smart diagnostic approach to identify faults in motor bearings by combining fiber Bragg grating sensors and support vector machine technology. Validation experiments showed the method had high accuracy and generality in diagnosing motor bearing faults (Brusamarello et al., 2022). To solve difficulties in diagnosing rotating machinery bearing faults, Ayas et al. introduced an innovative diagnostic approach grounded in time-domain signal to image conversion and a new deep residual network. Simulation results indicated the method overcame traditional limitations, achieving an average accuracy of 99.98% (Ayas and Ayas, 2022). Zhang et al. innovatively combined Gabor angle field image encoding with Convolutional Neural Networks (CNNs) and extreme learning machine to improve inefficient manual feature extraction of motor bearing faults. Experimental results demonstrated the method efficiently and accurately diagnosed motor bearing faults (Zhang et al., 2023). Xiong et al. introduced an intelligent diagnostic approach for petrochemical motor bearings through the application of complete ensemble empirical mode decomposition in conjunction with multidimensional indicator reconstruction, aiming to improve accuracy in nonlinear environments. Experiments verified the superior diagnostic performance of the method (Xiong et al., 2023). To tackle difficulties in extracting motor bearing fault features caused by noise interference, Zhen et al. introduced an innovative diagnostic approach that combines variational mode decomposition with a distributed control system. This method denoised signals via VMD and extracted fault frequencies through DCS demodulation, achieving precise fault diagnosis under noisy conditions (Zhen et al., 2022).

Table 1 summarizes the relevance of existing technologies to research. In summary, while existing research on motor bearing fault diagnosis has made significant progress, it still suffers from several shortcomings: First, most methods rely on a single modal signal, making it difficult to fully characterize fault characteristics under

TABLE 1 Summary and analysis of existing studies.

References	Method	Key technical features	Limitations	Relevance to this study
Khan et al. (2023)	Blockchain + IIoT security framework	Enhances data security, applied in healthcare IIoT	Does not involve fault diagnosis or signal feature extraction	Provides reference for secure IIoT deployment
Heidari et al. (2024)	ABC algorithm + Density correlation	Improves network reliability under device mobility	Cannot handle equipment status recognition or fault analysis	Inspires collaborative sensor data design
Theodoris et al. (2023)	Deep TL + attention Mechanism	Addresses data sparsity in gene networks	Difficult to adapt to industrial signal structures	Inspires few-shot transfer strategy design
Chen et al. (2022)	TL-based software vulnerability repair	Combines large-scale pretraining and fine-tuning	Not applicable to multi-source time-series modeling	Offers transferable structure for source–target domain learning
Lotfollahi et al. (2022)	Parameter sharing + single-cell data mapping	Enables cross-dataset model collaboration	Application is domain-specific; lacks IIoT applicability	Guides lightweight transfer module design
Alijani et al. (2024)	Vision Transformer + domain generalization	Improves robustness under distribution shifts	High computational complexity, unsuitable for edge deployment	Supports stability design in transfer modules
Nancy and Maheswari (2025)	VGG-19 + Attribute-aware attention	Enhances brain tumor segmentation and classification	Does not model temporal signals	Inspires improvements in feature fusion layers
Brusamarello et al. (2022)	FBG sensor + SVM	Achieves accurate bearing fault detection	Highly sensor-dependent, lacks generalization	Emphasizes the need for multi-source perception
Ayas and Ayas (2022)	Time-domain to Image + ResNet	Boosts performance in rotating machinery diagnostics	Loses signal details during transformation	Supports retention of original vibration characteristics
Zhang et al. (2023)	Gabor coding + CNN-ELM	Enhances image feature representation	Poor performance under small samples, weak transferability	Aligns with this study's focus on domain adaptation
Xiong et al. (2023)	EEMD + Multivariate indicator reconstruction	Effective in nonlinear fault environments	Complex structure, low real-time performance	Motivates optimization of deployment efficiency
Zhen et al. (2022)	VMD + DCS	Enhances denoising and fault feature extraction	Complex pipeline, not scalable for multi-source data	Reinforces the need for end-to-end fusion diagnosis models

complex operating conditions; second, many deep learning models lack the ability to migrate across different operating conditions or devices, resulting in limited generalization; and third, some models fail to fully consider the real-time nature and resource constraints of industrial deployments, making them difficult to adapt to application scenarios such as edge computing. Therefore, this paper proposes a multi-source dynamic domain-adaptive fault diagnosis model that integrates IIoT and TL. This model aims to achieve cross-condition knowledge transfer and stable operation in low-resource deployment environments, thereby addressing the shortcomings of existing methods in terms of generalization, practicality, and deployment performance.

3 Motor bearing fault diagnosis model development using IIoT and transfer learning

3.1 Design of multi-source feature extraction method

A crucial aspect of motor bearing fault diagnosis involved capturing significant features embedded in vibration signals. However, in the actual acquisition process, weak signals and noise interference often cause feature loss and reduce accuracy. IIoT achieves multi-source data fusion by deploying various types of sensors, which enhances the capability of feature representation. Therefore, this study implements comprehensive motor bearing fault feature extraction at the IIoT platform layer. CNN excels at handling local spatial features (Preethi and Mamatha, 2023), while BiGRU effectively captures temporal dependencies, compensating for CNN's limitation in processing time-series fault signals (Qiao et al., 2023). The combination of the two can achieve joint spatio-temporal feature extraction of multi-source industrial signals, improving fault identification performance in complex operating conditions. Thus, this study combines CNN and BiGRU to build a deep feature extraction method named CNN-BiGBR. The structure of this method is shown in Figure 1.

As shown in Figure 1, CNN-BiGBR consists of two branches. Branch 1 uses BiGBR to extract temporal correlations in motor fault signals. Branch 2 applies CNN to extract features from motor fault data. To facilitate feature extraction by CNN, short-time Fourier transform (STFT) converts one-dimensional mechanical vibration time series into two-dimensional time-frequency spectra. STFT uses a moving time window technique to segment continuous time-domain waveforms into local time segments, then performs spectral transformation on each segment to obtain the two-dimensional time-frequency representation of the original signal (Beeraka et al., 2022). The expression of STFT is given in Equation 1.

$$X_{(w,\tau)} = \int_{-\infty}^{\infty} x(t)w(t-\tau)e^{-jwt}dt \quad (1)$$

In Equation 1, w denotes the vibration frequency of the vibration signal, τ indicates the starting time of the current short-time window, t is the time variable, $x(t)$ is the original input signal, $X_{(w,\tau)}$ is the result of STFT, e^{-jwt} represents the complex exponential function, and dt is the integration variable. The processed two-

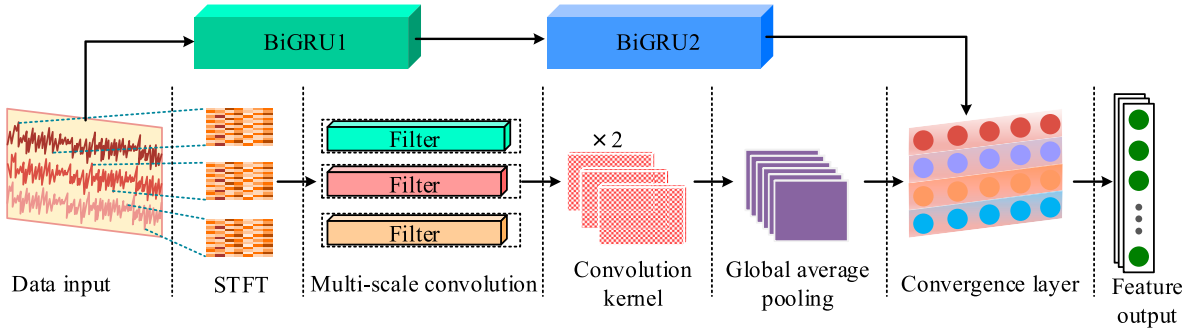

FIGURE 1
Flowchart of CNN-BiGBR deep feature extraction method (Source from: author self-drawn).

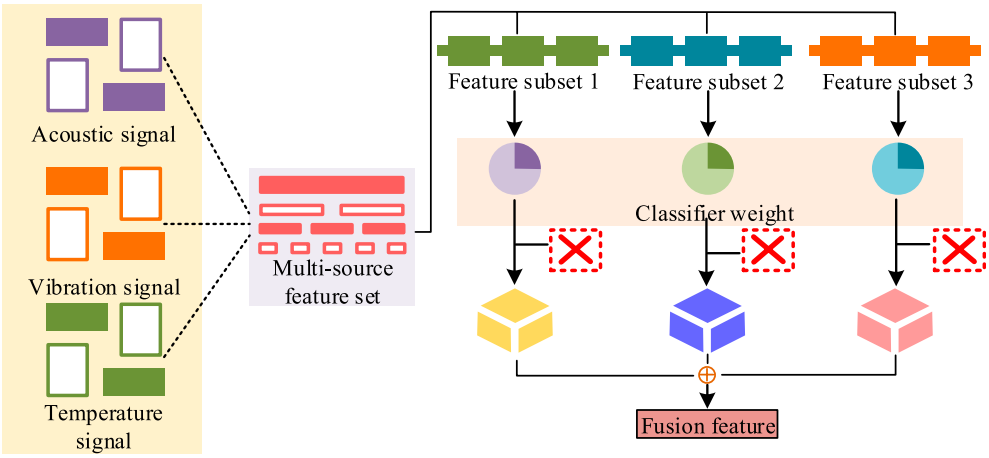

FIGURE 2
Structure of multi-source fault feature fusion method Adaboost-DCM (Source from: author self-drawn).

dimensional time-frequency images enter the multi-scale convolution module for feature extraction, then proceed to the fourth convolution kernel for further convolution operations. The convolution process is expressed in Equation 2.

$$\overset{l}{y}_n = f\sum_{i=1}^{m}\left(x_i^{l-1} \times w_n^l + b_n^l\right) \tag{2}$$

In Equation 2, $\overset{l}{y}_n$ represents the l-th feature map output by the n-th layer, x_i^{l-1} denotes the $l-1$-th feature map output by the i-th layer, w_n^l and b_n^l represent the weight matrix and bias matrix, and f is the activation function. After convolution, the pooling layer further reduces network parameters. This process is described by Equation 3.

$$p_{\max}^{l(i,j)} = \max_{(j-1)\sigma \le t \le j\sigma}\left\{u^{l(i,m)}\right\} \tag{3}$$

In Equation 3, $u^{l(i,t)}$ denotes the l-th neuron of the i-th feature map in the m-th layer, and σ is the depth of the convolution kernel. Multi-scale and temporal features are fused through the pooling layer, producing the final feature set. Because IIoT collects fault features from multiple sensors like vibration and temperature, which contain redundancy and complex nonlinear relations, feature purification and fusion are necessary. Adaboost dynamically adjusts sample and classifier weights to identify the importance of different feature sources, enabling efficient adaptive fusion of multi-source features. Based on Adaboost, this study adds a DCM and designs a multi-source fault feature fusion method called Adaboost-DCM. The structure of this method is shown in Figure 2.

As shown in Figure 2, Adaboost-DCM first integrates multi-source features to generate an initial feature set. Then it divides the feature set equally into subsets and filters effective features through weight calculation. Finally, the retained fault features are fused into the final fault feature set. During weight calculation, the fault subset samples undergo weight initialization, producing a weighted fault feature set. Then the feature set is trained to obtain the error rate of the weak learner, which determines the weight of the classifier. The weight coefficient of the fault feature subset is calculated by Equation 4.

$$\begin{cases} e_k = \sum_{i=1}^{m}\omega_{ki} I\left(G_k\left(x_i\right) \neq y_i\right), i = 1, 2, \cdots, m \\ a_k = \frac{1}{2}\ln\frac{1-e_k}{e_k} \end{cases} \tag{4}$$

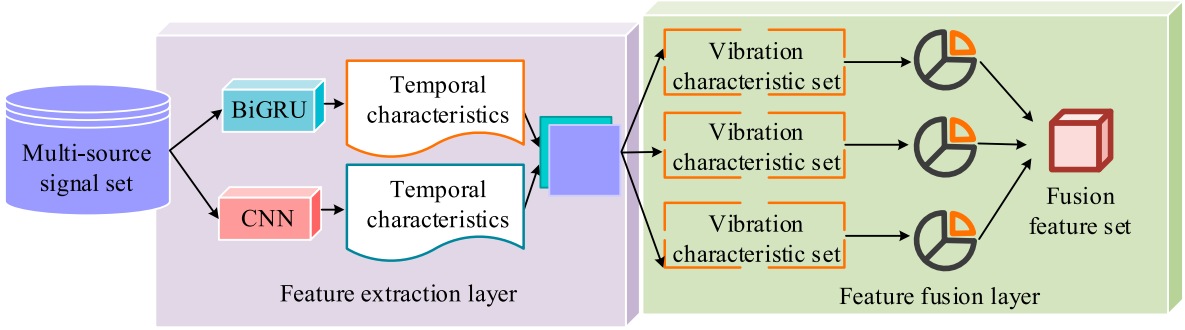

FIGURE 3
Framework of the CB-AD multi-source feature extraction process (Source from: author self-drawn).

In Equation 4, e_k and a_k represent the error rate and weight coefficient of the $G_k(x_i)$ weak learner, respectively. ω_{ki} is the initialized weight value of the k-th weak learner on the i-th sample. m is the number of samples. $I(G_k(x_i) \neq y_i$ equals 0 or 1, where 1 means incorrect and 0 means correct. After obtaining the weights of different weak learners, a cropping coefficient θ is set to decide whether the sample needs further learning. The cropping coefficient θ is expressed by Equation 5.

$$\theta = \begin{cases} \frac{1}{2}\tanh(2accuracy - 1) + \frac{1}{2}, accuracy \geq 50\% \\ 0, accuracy < 50\% \end{cases} \quad (5)$$

In Equation 5, $accuracy$ is the model accuracy. By comparing the cropping coefficient to ω_{ki}, when $\omega_{ki} \leq \theta$, it resets to 0, indicating the diagnostic accuracy of this fault feature reaches 100% and no further learning is needed. Otherwise, the sample weight is restored to $\theta = 0$. The weight distribution of fault features continuously updates through iterative training of weak learners, calculated by Equation 6.

$$\begin{cases} Z_k = \sum_{i=1}^{m} w_{ki} e^{-\alpha_k y_i G_k(x_i)} \\ \omega_{k+1,i} = \frac{\omega_{ki}}{Z_k} \exp(-\alpha_k y_i G_k(x_i)) \end{cases} \quad (6)$$

In Equation 6, Z_k is the normalization factor mapping ω into [0,1], and y_i is the true class label of sample x_i. Equation 6 improves accuracy for difficult-to-classify fault features. Finally, results from multiple weak classifiers fuse to form the final strong classifier. The fusion strategy is expressed by Equation 7.

$$G(x) = sign\left(\sum_{k=1}^{k} \alpha_k G_k(x)\right) \quad (7)$$

In Equation 7, $G(x)$ is the final strong classifier. By combining CNN-BiGBR deep feature extraction and Adaboost-DCM multi-source feature fusion methods, a multi-source feature fusion extraction method named CB-AD is obtained. The operation flow of this method is shown in Figure 3.

Figure 3 illustrates the operation flow of CB-AD, divided into multi-source feature extraction and multi-source feature fusion. The feature extraction part uses CNN and BiGRU to extract local spatial and temporal features from vibration, sound, and temperature signals, then fuses them into a global feature set. The fusion part uses weighted weak classifiers to assign dynamic weights and adaptively fuse the three types of features, finally outputting an optimized multi-source fused feature set. This set provides a feature data source for subsequent fault diagnosis.

3.2 Fault diagnosis model design based on TL

Based on the extracted unified feature representation, this paper further constructs a TL-based fault diagnosis model. This model adopts a Feature Transfer Learning (FTL) architecture, pre-trains with source domain samples, and aligns the deep feature distribution between the source and target domains by introducing the JMMD loss function, achieving knowledge transfer and model generalization across working conditions and equipment scenarios. FTL solves the domain differences and data scarcity problems in motor bearing fault diagnosis by transferring knowledge learned in the source domain to the target domain to perform tasks (Abbasi et al., 2025; Chan et al., 2023). The specific architecture is shown in Figure 4.

As shown in Figure 4, in the source task branch, the FTL model uses a large number of source domain samples for pre-training and extracts features through convolution and fully connected layers. The target task uses training and testing samples from the target domain and also extracts fault features through convolution and fully connected layers. In both source and target tasks, JMMD calculates the joint distribution distance between the two-layer networks, which helps reduce the domain difference caused by the network structure and enables successful knowledge transfer. The JMMD loss function is defined in Equation 8.

$$L_{JMMD}(P,Q) = \left\| E_P\left(\otimes_{l=1}^{|L|} \phi^l(z^{sl})\right) - E_P\left(\otimes_{l=1}^{|L|} \phi^l(z^{tl})\right) \right\|^2_{\otimes_{l=1}^{|L|} H^l} \quad (8)$$

In Equation 8, H^l denotes the reproducing kernel Hilbert space of layer l, $|L|$ represents the number of corresponding layers, $\otimes_{l=1}^{|L|} \phi^l$ indicates the feature mapping of the tensor product in the reproducing kernel Hilbert space, z^{sl} and z^{tl} are the activations from the source and target domains in layer l, respectively. In real production scenarios, motor bearing fault data usually come from

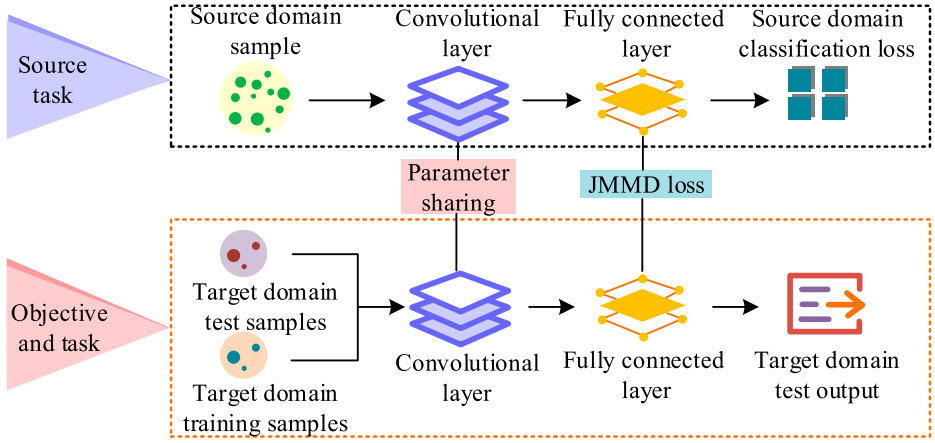

FIGURE 4
Architecture of the FTL) model (Source from: author self-drawn).

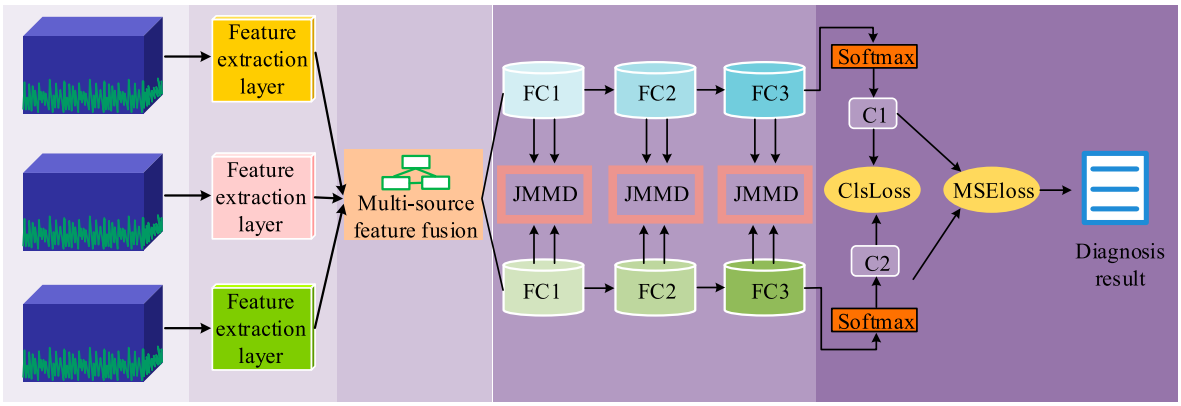

FIGURE 5
Structure of the CB-AD-FTL multi-source fault diagnosis model (Source from: author self-drawn).

multiple different source domains (Rajabioun et al., 2023). Multi-source data provide more comprehensive diagnostic information, which improves the accuracy of fault diagnosis (Wang et al., 2024). Based on the feature representation generated by the CB-AD module mentioned above, this study constructed a fault diagnosis model CB-AD-FTL that integrates multi-source domain adaptation and transfer learning. The model structure is shown in Figure 5.

As shown in Figure 5, the CB-AD-FTL model builds on the CB-AD feature extraction module and constructs a dual-source domain transfer module based on three sets of two-layer fully connected networks. By introducing the JMMD criterion, the source and target domain features are aligned, and the final diagnosis result is output through a shared Softmax classification layer. The Softmax function is shown in Equation 9.

$$p_i = \frac{e^{Z_i}}{\sum_{j=1}^{k} e^{Z_j}} \tag{9}$$

In Equation 9, p_i represents the classification probability of fault class i, Z_i is the feature value of fault class i, and k is the total number of fault classes. Because the input includes two source domains as transfer objects for the target domain, the diagnostic results in the target domain may differ. To reduce discrepancies among classifiers, this study applies a mean square error loss function, as shown in Equation 10.

$$L_{mse} = \frac{1}{n}\sum_{i=1}^{n}\left(y_i^{t1} - y_i^{t2}\right)^2 n \tag{10}$$

In Equation 10, n represents the scale of target domain samples, y_i^{t1} and y_i^{t1} are the predicted results from the corresponding classifiers in the target domain. L_c^1 and L_c^2 indicate the loss functions of classification results. Then, based on the loss function, the overall loss between the two source domains and the target domain is calculated using Equation 11.

$$\begin{cases} L_1 = L_c^1 + \lambda L_{JMMD}^1 \\ L_2 = L_c^2 + \lambda L_{JMMD}^2 \end{cases} \tag{11}$$

In Equation 11, L_c^1 and L_c^2 are the classification loss functions of the two source domains, λ is the balancing parameter of the loss

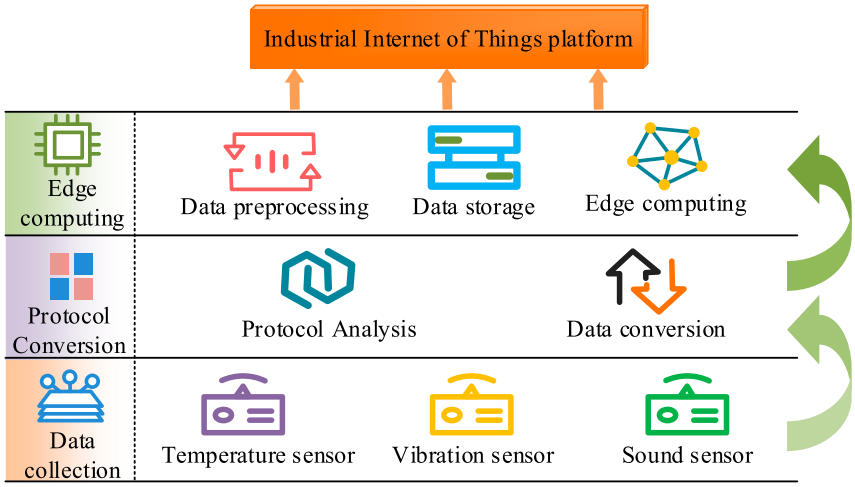

FIGURE 6
System architecture of the IIoT-TL fault diagnosis model (Source from: author self-drawn).

function, and L_1 and L_2 are the respective losses between the source and target domains. The global loss function of the CB-AD-FTL model is shown in Equation 12.

$$L = L_1 + L_2 + L_{mse} \tag{12}$$

In Equation 12, L represents the final loss function. The CB-AD-FTL model enables fault diagnosis of motor bearings by analyzing features from temperature, sound, and vibration signals. However, practical applications in industrial production still rely on IIoT functions such as data coordination, edge computing, and dynamic optimization. Therefore, this study embeds the CB-AD-FTL model into the IIoT and proposes an IIoT and TL-based fault diagnosis model for motor bearings, named IIoT-TL. The architecture of this model is given by Figure 6.

In Figure 6, the IIoT-TL model consists of three main modules: multi-source data collection, protocol conversion, and edge computing. The data collection module gathers real-time vibration, temperature, and other signals using heterogeneous sensors. The protocol conversion layer enables cross-network communication. The edge computing module applies the CB-AD-FTL model for fault analysis. After uploading diagnostic results to the IIoT platform, parameter dynamic optimization completes a "perception–diagnosis–optimization" loop, enabling cross-condition knowledge transfer in motor bearing fault diagnosis.

4 Performance validation of IIoT-TL motor bearing fault diagnosis model

4.1 Experimental setup and model training performance validation

To assess the effectiveness of the IIoT-TL model raised for motor bearing fault diagnosis, two experiments were carried out in this study: model performance testing and practical application testing. The bearing fault dataset of Jiangsu Jiangnan University used in this study is a public dataset that is widely used in the field of industrial fault diagnosis and has good versatility and representativeness. This dataset collects operating data under different loads, speeds and fault types by building a standardized motor bearing test platform. The fault types include normal state, outer ring fault, inner ring fault and rolling element fault. Each type of fault is further divided into three levels: mild, moderate and severe, for a total of 10 working conditions. Each type of sample contains 1,200 three-axis vibration data, with a sampling frequency of 50 kHz and a single-segment signal acquisition time of 1 s, with obvious timing and spectrum characteristics. All signals are acquired synchronously from three orthogonal directions of the motor bearing seat through acceleration sensors, with strong adaptability to real working conditions and data completeness, providing a reliable data basis for the feature extraction and migration diagnosis of this research model. Details are shown in Table 2.

The experimental environment settings are shown in Table 3.

Based on the environment settings in Table 3, the IIoT-TL model was first trained through iterative experiments. The dataset was split with 70% used as the training set and 30% as the test set. Fault diagnosis verification for outer and inner ring faults was performed on both datasets. Among them, the key module parameter settings of the CB-AD-FTL model are shown in Table 4.

In Table 4, the STFT window size is set to 256 points, balancing the 50 kHz sampling frequency with the periodic characteristics of the vibration signal. This ensures sufficient spectral detail while maintaining temporal resolution. The BiGRU architecture adopts a one-layer, bidirectional structure with 64 units in each direction, determined by comprehensively considering model complexity and the ability to model temporal features. Furthermore, a combination of 3 × 3 and 5 × 5 convolution kernel sizes is used to balance local and mid-scale feature extraction, improving the model's adaptability to fault components of varying frequencies. The experimental results are given by Figure 7.

In Figure 7a, when the number of iterations reached 50, the fault identification accuracy for the training set (outer ring and inner ring) and the test set (outer ring and inner ring) was approximately

TABLE 2 Summary of Jiangnan University bearing fault dataset.

Condition type	Fault severity	Number of classes	Samples per class
Normal	—	1	1200
Outer ring fault	Mild/Moderate/Severe	3	1200
Inner ring fault	Mild/Moderate/Severe	3	1200
Ball fault	Mild/Moderate/Severe	3	1200
Total	—	10	12,000

TABLE 3 Experimental environment settings.

Parameter	Configuration
CPU	Intel i7
Programming language	Python 3.7
GPU	GeForce RTX 4070
Sensor	Three-axis acceleration sensor Current sensor Temperature sensor
Microcontroller	STM32H743
Local storage	8 GB TF

TABLE 4 Model parameter settings.

Module	Parameter	Value
CNN	Kernel size	3 × 3, 5 × 5
	Number of filters	32, 64
	Activation function	ReLU
	Pooling method	MaxPooling (2 × 2)
BiGRU	Units per direction	64
	Number of layers	1
	Dropout rate	0.3
Adaboost module	Number of weak learners	10
	Cropping threshold (q)	0.5
Fully connected	FC1/FC2 units	128/64
Transfer module	Loss function	JMMD + MSE + CE
	Balance coefficient λ	0.5
Optimizer	Optimizer type	Adam
	Initial learning rate	0.001
	Batch size	64
	Epochs	250

89.2%, 88.7%, 87.3%, and 86.8%. When the number of iterations increased to 250, the accuracy became stable, eventually reaching 94.7%, 94.5%, 93.4%, and 93.2%. According to Figure 7b, the false alarm rate decreased as the number of iterations increased. At 50 iterations, the false alarm rate for the training set (outer ring and inner ring) and the test set (outer ring and inner ring) was approximately 21.3%, 22.7%, 18.8%, and 19.3%, respectively. When the iterations increased to 250, these rates dropped to around 6.8%, 7.0%, 7.3%, and 7.5%. These results demonstrated that, after training, the proposed model achieved strong fault diagnosis performance. To further verify the effectiveness of the training process, the model's ability to recognize motor operating states was evaluated. The consistency between the model's predictions and the actual states for both training and test sets is shown in Figure 8.

As shown in Figures 8a,d, for the outer ring, the matching rate between the model's predictions and State 1 was 96.3% for the test set and 90.4% for the training set. For States 2 and 3, the matching rates were 96.2% and 94.1% (test set), and 95.9% and 89.9% (training set), respectively. According to Figures 8b,c, for the inner ring, the matching rate with State 1 was 95.3% (test set) and 91.4% (training set). For State 2, the rates were 94.9% and 92.4%, and for State 3, the rates reached 97.3% and 94.4%. In summary, after training, the IIoT-TL model accurately identified the motor bearing operating states with high precision. Finally, the study further analysed the interpretability of the CB-AD-FTL model and ranked the importance of input features, as shown in Figure 9.

As shown in Figure 9, the low-frequency energy feature (Vib_lowfreq) in the vibration signal has the highest average Shapley Additive Explanations (SHAP) value, indicating that it plays a decisive role in the model's fault identification. The Vib_lowfreq feature typically reflects frequency domain changes during equipment operation and is susceptible to factors such as structural damage. The temperature gradient change feature (Temp_Delta) ranks second, potentially reflecting the sensitive response of equipment thermal characteristics across different fault stages. Acoustic_Peak provide effective criteria during short-term impacts such as local structural collisions. In contrast, acoustic Acoustic_Mel-Frequency Cepstral Coefficients (MFCC) and temperature mean have relatively minor influences and serve as auxiliary discriminative features. Overall, the CB-AD-FTL model demonstrates superior effectiveness in multi-source signal fusion and feature learning, and its decision logic aligns to some extent with industrial fault mechanisms, thereby supporting the model's trustworthiness and deployment feasibility in practical applications.

4.2 Practical performance validation of IIoT-TL motor bearing fault diagnosis model

To further validate the real-world effectiveness of the IIoT-TL model raised, the study conducted comparative experiments

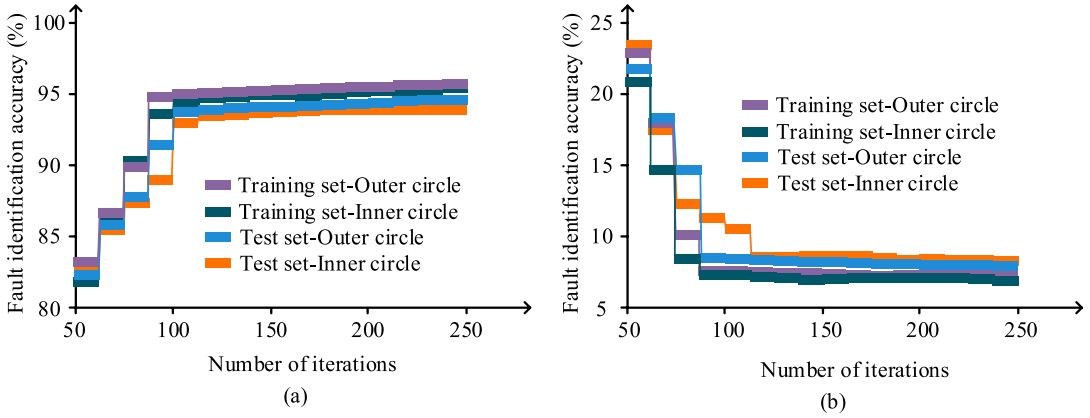

FIGURE 7
Changes in accuracy and false alarm rate under different numbers of iterations (Source from: author self-drawn). **(a)** Fault recognition rate **(b)** False alarm rate of faults.

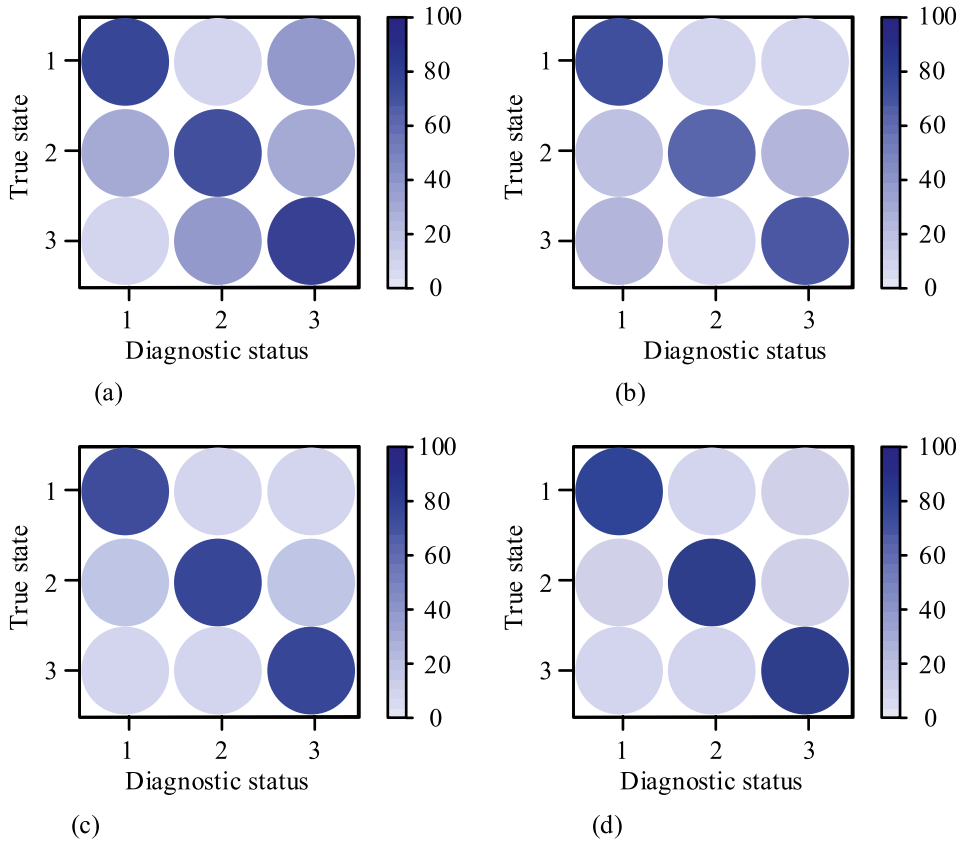

FIGURE 8
Model identification results of motor bearing operating status (Source from: author self-drawn). **(a)** Test set-outer ring **(b)** Test set-inner ring **(c)** Training set-inner ring **(d)** Training set-outer ring.

with other fault diagnosis models after the training process. The comparison included SqueezeNet: Lightweight Deep Neural Network (SqNet), CNN, and K-Nearest Neighbor Graph Convolutional Network (K-GCN). The first set of comparison experiments tested fault identification accuracy. The results are given by Figure 10.

As shown in Figure 10a, under Condition 1, the diagnosis accuracy of CNN ranged from 87.4% to 97.1%. K-GCN

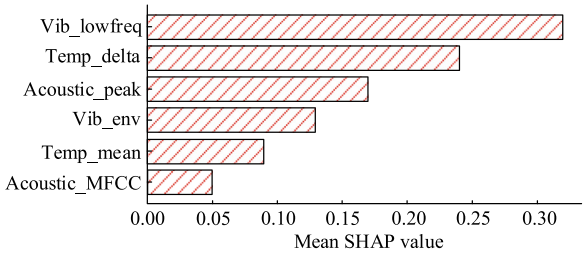

FIGURE 9
Input feature importance based on SHAP (Source from: author self-drawn).

performed slightly better, ranging from 91.5% to 96.2%. The accuracy of SqNet fluctuated between 86.8% and 97.2%. The IIoT-TL model achieved the most stable performance, with accuracy between 93.0% and 97.6%. According to Figure 10b, under Condition 2, CNN's accuracy ranged from 85.3% to 96.2%. K-GCN achieved an accuracy between 88.1% and 96.2%. SqNet showed more variation, ranging from 87.5% to 94.1%. In contrast, IIoT-TL maintained stable performance with accuracy between 93.5% and 98.1%. These results show that the IIoT-TL model can maintain a high and stable recognition accuracy under different equipment states or operating parameter changes, reflecting good cross-condition robustness and generalization ability. To further assess the practicality of IIoT-TL, the study compared the diagnosis time and GPU usage across the four models. The results are presented in Figure 11.

Figure 11a shows that as the data size increased from 5 GB to 25GB, the diagnosis time of CNN ranged from 22.3 s to 30.4 s. For K-GCN, the time varied from 9.8 s to 32.5 s. SqNet had diagnosis times between 18.4 s and 24.3 s. IIoT-TL performed the best, with diagnosis times between 13.5 s and 19.6 s. According to Figure 11b, as the data size increased, CNN's average GPU usage was 25.6%. K-GCN and SqNet had average usage rates of 14.8% and 11.2%, respectively. IIoT-TL required the least computational resources, with an average GPU usage of only 8.9%. These findings confirmed that IIoT-TL had the shortest diagnosis time and the lowest resource consumption, making it highly practical. Finally, the study conducted experiments to test diagnostic accuracy. The results are shown in Figure 12.

As shown in Figure 12a, when the noise level reached 50%, the accuracy of CNN and K-GCN gradually declined as the data volume increased, dropping to 92.1% and 90.2%, respectively, at 4 GB. SqNet and IIoT-TL also experienced a decline, with accuracy rates of 89.4% and 95.3%, respectively, at 4 GB. When the noise level dropped to 20%, CNN, K-GCN, and SqNet achieved diagnosis accuracy of 89.5%, 88.7%, and 90.6%, respectively. IIoT-TL still maintained a high accuracy of 94.6%. In summary, even under degraded signal quality, IIoT-TL maintained high diagnostic accuracy, demonstrating superior robustness and strong adaptability to complex fault diagnosis environments for motors.

To further verify the performance advantages of the proposed IIoT-TL model on the Jiangnan University bearing fault dataset, two representative methods from existing literature using the same dataset were introduced for comparison. These are the fault recognition model based on SVM and FBG sensors proposed by Brusamarello et al. in reference (Brusamarello et al., 2022), and the Gabor feature encoding and CNN-ELM fusion model proposed by Zhang et al. in reference (Zhang et al., 2023). The comparison indicators include average diagnostic accuracy, 95% confidence interval, and statistical significance analysis results compared with the proposed model. The details are shown in Table 5.

Table 5 shows that the proposed IIoT-TL model achieved an average diagnostic accuracy of 94.7% on the Jiangnan University bearing dataset, significantly outperforming the SVM + FBG method proposed by Brusamarello et al. and the Gabor + CNN-ELM method proposed by Zhang et al. The IIoT-TL model improved its accuracy by 1.9% compared to the SVM + FBG method and by 1.6% compared to the CNN-ELM model. A two-sample t-test showed that the IIoT-TL model achieved statistical significance ($p < 0.05$) over both methods, demonstrating a significant advantage in diagnostic accuracy. Furthermore, the IIoT-TL model's accuracy within the 95% confidence interval

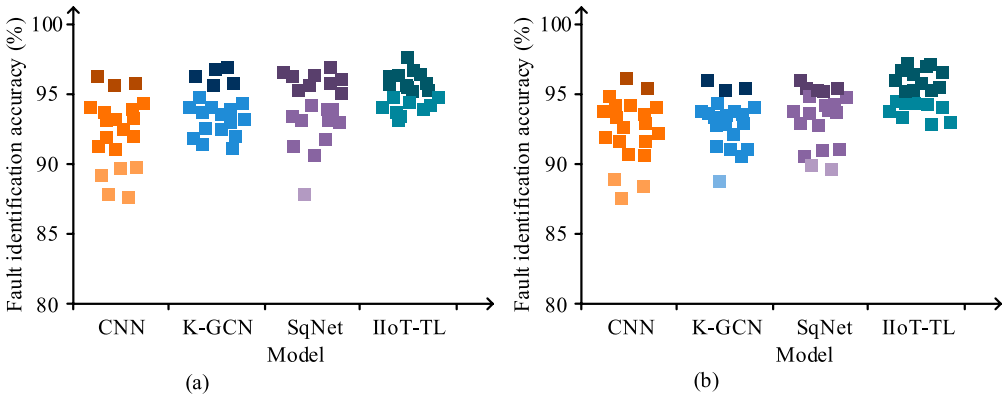

FIGURE 10
Comparison of multi-model fault diagnosis accuracy under different working conditions (Source from: author self-drawn). **(a)** Working Condition 1 **(b)** Working Condition 2.

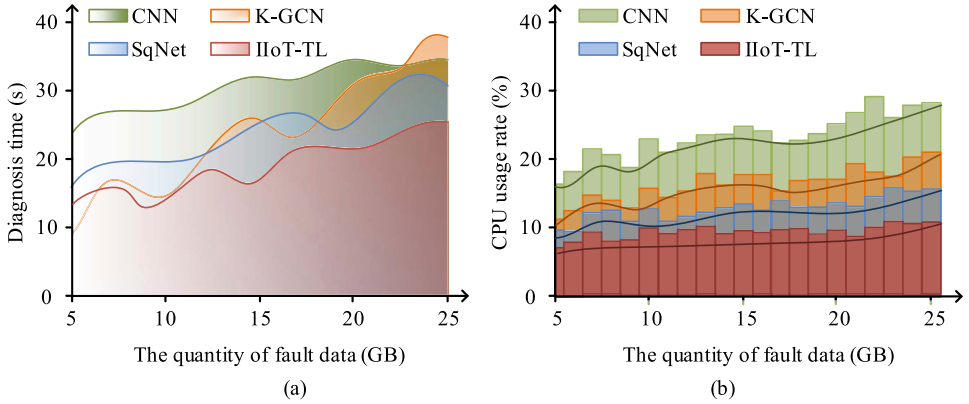

FIGURE 11
Diagnosis time and GPU usage of each model under different data volumes (Source from: author self-drawn). **(a)** Diagnosis time **(b)** Occupy computing resources.

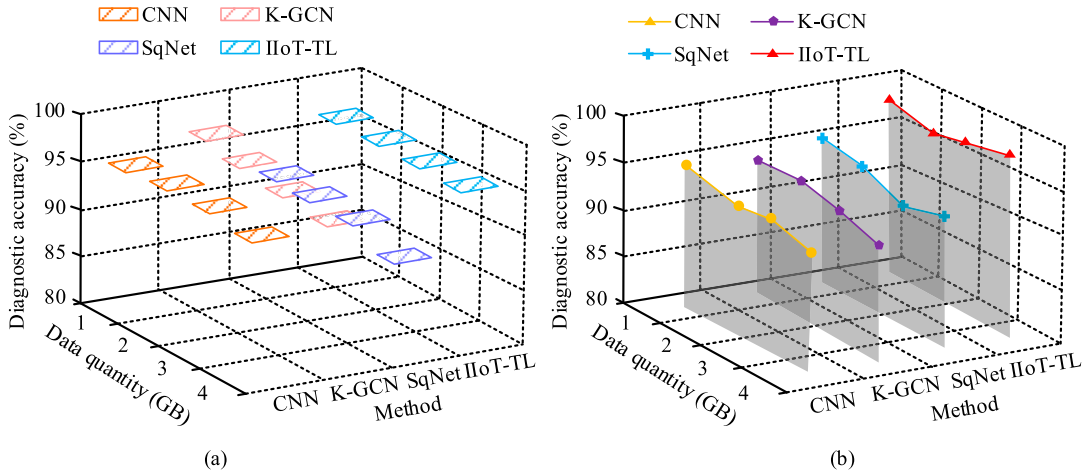

FIGURE 12
Diagnosis accuracy under different Signal-to-Noise Ratios (Source from: author self-drawn). **(a)** Noise content=50% **(b)** Noise content=20%.

TABLE 5 Performance comparison of different methods.

Method	Average accuracy (%)	95% confidence interval	*p*-value
SVM with FBG sensor	92.8	[91.20, 94.40]	0.021
Gabor + CNN-ELM	93.1	[91.75, 94.20]	0.035
IIoT-TL	94.7	[93.60, 95.50]	—

ranged from [93.60%–95.50%], a relatively narrow range that further validated its stability and generalization capabilities.

5 Conclusion

The increasing demand for intelligent operation and maintenance of modern industrial equipment has driven motor fault diagnosis technologies toward higher accuracy and lower latency. However, traditional fault diagnosis methods showed limitations such as poor transferability and heavy reliance on historical fault data. Therefore, this study put forward a multi-source dynamic domain adaptation model for motor fault diagnosis based on the IIoT and TL. Experimental results demonstrated that after training, the proposed IIoT-TL model achieved a fault diagnosis accuracy of 94.7% and a false alarm rate of 6.8%. The highest state matching rate reached 97.3%. In real-world application experiments, the diagnosis accuracy under various working

conditions remained above 93%. When processing 25 GB of data, the diagnosis time was only 19.6 s, and the average GPU usage was merely 8.9%. Even under 50% noise conditions, the model maintained an accuracy of 95.3%. Results show that IIoT-TL improves the model's adaptability under multiple operating conditions. The introduction of multi-source input and fusion mechanisms effectively enhances the model's robustness and fault tolerance. The IIoT-TL model provides strong diagnostic accuracy and portability, enabling reliable fault diagnosis in a variety of situations.

While this research focuses on motor bearing fault diagnosis, the model still has some dependencies on sensor configuration and acquisition environment, and does not yet cover complex scenarios such as complex faults and multi-source heterogeneous devices. Experiments were conducted using the Jiangnan University bearing dataset, which features multiple operating conditions and multi-level labels, making it suitable for validating model transfer capabilities. However, different datasets vary significantly in signal modality and sampling methods. Direct transfer can lead to inconsistent feature domains, compromising the comprehensiveness of model robustness assessments. Furthermore, to meet the demands of edge computing and low-power deployment in IIoT scenarios, this model was designed with controlled parameter size and good embedded adaptability. Future work will further test its deployment performance in lightweight frameworks such as TensorRT, ONNX, and PyTorch Mobile, and evaluate its real-time diagnostic capabilities on platforms such as the Raspberry Pi and Jetson Nano. Future research will also expand the model's adaptability to diverse device types, complex faults, and multi-machine systems, improving its practicality and generalization performance.

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

DG: Writing – original draft, Writing – review and editing.

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

Conflict of interest

The author declares that the research 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) declare that no Generative AI was 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

References

Abbasi, M. A., Huang, S., and Khan, A. S. (2025). Hybrid fault diagnostics and prognostics system development for motor bearings health monitoring. *Qual. Eng.* 37 (2), 292–307. doi:10.1080/08982112.2024.2385920

Alijani, S., Fayyad, J., and Najjaran, H. (2024). Vision transformers in domain adaptation and domain generalization: a study of robustness. *Neural Comput. Appl.* 36 (29), 17979–18007. doi:10.1007/s00521-024-10353-5

Ayas, S., and Ayas, M. S. (2022). A novel bearing fault diagnosis method using deep residual learning network. *Multimedia Tools Appl.* 81 (16), 22407–22423. doi:10.1007/s11042-021-11617-1

Beeraka, S. M., Kumar, A., Sameer, M., Ghosh, S., and Gupta, B. (2022). Accuracy enhancement of epileptic seizure detection: a deep learning approach with hardware realization of STFT. *Circuits, Syst. Signal Process.* 41 (1), 461–484. doi:10.1007/s00034-021-01789-4

Brusamarello, B., Da Silva, J. C. C., de Morais Sousa, K., and Guarneri, G. A. (2022). Bearing fault detection in three-phase induction motors using support vector machine and fiber bragg grating. *IEEE Sensors J.* 23 (5), 4413–4421. doi:10.1109/JSEN.2022.3167632

Chan, J. Y. L., Bea, K. T., Leow, S. M. H., Phoong, S. W., and Cheng, W. K. (2023). State of the art: a review of sentiment analysis based on sequential transfer learning. *Artif. Intell. Rev.* 56 (1), 749–780. doi:10.1007/s10462-022-10183-8

Chen, Z., Kommrusch, S., and Monperrus, M. (2022). Neural transfer learning for repairing security vulnerabilities in c code. *IEEE Trans. Softw. Eng.* 49 (1), 147–165. doi:10.1109/tse.2022.3147265

Chi, H. R., Wu, C. K., Huang, N. F., Tsang, K. F., and Radwan, A. (2022). A survey of network automation for industrial internet-of-things toward industry 5.0. *IEEE Trans. Industrial Inf.* 19 (2), 2065–2077. doi:10.1109/tii.2022.3215231

Choudhary, A., Mian, T., Fatima, S., and Panigrahi, B. K. (2022). Passive thermography based bearing fault diagnosis using transfer learning with varying working conditions. *IEEE Sensors J.* 23 (5), 4628–4637. doi:10.1109/jsen.2022.3164430

Heidari, A., Shishehlou, H., Darbandi, M., Navimipour, N. J., and Yalcin, S. (2024). A reliable method for data aggregation on the industrial internet of things using a hybrid optimization algorithm and density correlation degree. *Clust. Comput.* 27 (6), 7521–7539. doi:10.1007/s10586-024-04351-4

Hu, W., Xin, G., Wu, J., An, G., Li, Y., Feng, K., et al. (2023). Vibration-based bearing fault diagnosis of high-speed trains: a literature review. *High-speed Railw.* 1 (4), 219–223. doi:10.1016/j.hspr.2023.11.001

Hu, Y., Jia, Q., Yao, Y., Lee, Y., Lee, M., Wang, C., et al. (2024). Industrial internet of things intelligence empowering smart manufacturing: a literature review. *IEEE Internet Things J.* 11 (11), 19143–19167. doi:10.1109/jiot.2024.3367692

Khan, A. A., Bourouis, S., Kamruzzaman, M. M., Hadjouni, M., Shaikh, Z. A., Laghari, A. A., et al. (2023). Data security in healthcare industrial internet of things with blockchain. *IEEE Sensors J.* 23 (20), 25144–25151. doi:10.1109/jsen.2023.3273851

Lotfollahi, M., Naghipourfar, M., Luecken, M. D., Khajavi, M., Büttner, M., Wagenstetter, M., et al. (2022). Mapping single-cell data to reference atlases by transfer learning. *Nat. Biotechnol.* 40 (1), 121–130. doi:10.1038/s41587-021-01001-7

Nancy, A. M., and Maheswari, R. (2025). Brain tumor segmentation and classification using transfer learning based CNN model with model agnostic concept interpretation. *Multimedia Tools Appl.* 84 (5), 2509–2538. doi:10.1007/s11042-024-20353-1

Preethi, P., and Mamatha, H. R. (2023). Region-based convolutional neural network for segmenting text in epigraphical images. *Artif. Intell. Appl.* 1 (2), 103–111. doi:10.47852/bonviewaia2202293

Qiao, Y., Xu, H. M., Zhou, W. J., Peng, B., Hu, B., and Guo, X. (2023). A BiGRU joint optimized attention network for recognition of drilling conditions. *Petroleum Sci.* 20 (6), 3624–3637. doi:10.1016/j.petsci.2023.05.021

Rajabioun, R., Afshar, M., Mete, M., Atan, Ö., and Akin, B. (2023). Distributed bearing fault classification of induction motors using 2-D deep learning model. *IEEE J. Emerg. Sel. Top. Industrial Electron.* 5 (1), 115–125. doi:10.1109/jestie.2023.3323253

Theodoris, C. V., Xiao, L., Chopra, A., Chaffin, M. D., Al, S. Z. R., Hill, M. C., et al. (2023). Transfer learning enables predictions in network biology. *Nature* 618 (7965), 616–624. doi:10.1038/s41586-023-06139-9

Wang, M. H., Xing, L., Pan, Y., Gu, F., Fang, J., Yu, X., et al. (2024). AI-based advanced approaches and dry eye disease detection based on multi-source evidence: cases, applications, issues, and future directions. *Big Data Min. Anal.* 7 (2), 445–484. doi:10.26599/bdma.2023.9020024

Xiong, J., Liu, M., Li, C., Cen, J., Zhang, Q., and Liu, Q. (2023). A bearing fault diagnosis method based on improved mutual dimensionless and deep learning. *IEEE Sensors J.* 23 (16), 18338–18348. doi:10.1109/jsen.2023.3264870

Zhang, Y., Shang, L., Gao, H., He, Y., Xu, X., and Chen, Y. (2023). A new method for diagnosing motor bearing faults based on Gramian angular field image coding and improved CNN-ELM. *IEEE Access* 11 (3), 11337–11349. doi:10.1109/access.2023.3241367

Zhen, D., Li, D., Feng, G., Zhang, H., and Gu, F. (2022). Rolling bearing fault diagnosis based on VMD reconstruction and DCS demodulation. *Int. J. Hydromechatronics* 5 (3), 205–225. doi:10.1504/ijhm.2022.125092

Zhu, Z., Lin, K., Jain, A. K., and Zhou, J. (2023). Transfer learning in deep reinforcement learning: a survey. *IEEE Trans. Pattern Analysis Mach. Intell.* 45 (11), 13344–13362. doi:10.1109/tpami.2023.3292075