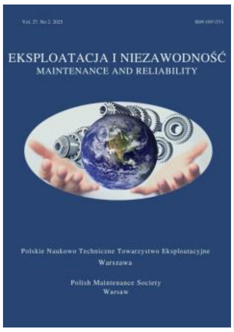


Eksploatacja i Niezawodnosc – Maintenance and Reliability
Volume 27 (2025), Issue 2
journal homepage: http://www.ein.org.pl

Article citation info:
Ge Y, Zhang F, Ren Y, A Federated Transfer Fault Diagnosis Method for Cross-Domain and Incomplete Data, Eksploatacja i Niezawodnosc – Maintenance and Reliability 2025: 27(2) http://doi.org/10.17531/ein/194182

Indexed by: Web of Science Group

A Federated Transfer Fault Diagnosis Method for Cross-Domain and Incomplete Data

Yang Ge[a,*], Fusheng Zhang[a], Yong Ren[a]

[a] Changshu Institute of Technology, China

Highlights

- Cross-domain federated fault diagnosis, exchanging only model parameters for privacy.
- We secure client data privacy by sharing only the parameters from local models.
- A relative distance-guided fine-tuning to improve diagnostics and avoid negative outcomes.

Abstract

Big data-driven intelligent fault diagnosis methods for device rely on a large amount of labeled data for centralized training. However, in practical engineering, it is difficult for a single client to collect enough labeled sample data, which is one of the reasons that limit the application of these methods. In fact, multiple clients often use similar devices and collect fault data separately, so joint multi-client collaborative fault diagnosis modeling can solve the problem of data scarcity, but this poses great challenges to data privacy protection. In this paper, we propose a federated transfer fault diagnosis method based on federated learning for cross-domain incomplete data. The proposed method only exchanges the parameters of the local training model, which achieves the privacy protection of the client's local data. We construct a multi-client collaborative learning framework to address the problem of weak generalization ability caused by the lack of terms in single client training samples. We also propose a targeted semi-supervised fine-tuning strategy based on relative distance to reduce the probability of negative fine-tuning of out-of-distribution samples and improve the accuracy of diagnostic models. The results of cross-condition and cross-equipment experiments demonstrate that the proposed method has obvious advantages over the existing fault diagnosis methods.

Keywords
fault diagnosis, data privacy, federated learning, transfer learning, prototypical network

1. Introduction

In recent years, data-driven intelligent fault diagnosis methods for mechanical equipment have achieved remarkable results [1], which can autonomously learn effective fault information from fault data and achieve a high fault recognition rate, and have gradually become a popular method in equipment health management. However, most intelligent diagnosis methods usually require sufficient high-quality operation monitoring data for model training [2, 3]. In reality, most of the existing fault sample data are collected in the laboratory, and it is often difficult to obtain the equipment operation data under the real fault state, and it often requires a lot of manpower and financial resources to accurately label the health state data. For a single client, it is challenging to obtain enough sample data, and it is even harder to obtain complete fault data, which is also the main factor limiting the large-scale

(*) Corresponding auth
E-mail addresses: Y. Ge (ORCID: 0000-0002-7722-5019) geyang@cslg.edu.cn, F. Zhang (ORCID: 0000-0001-8138-6862) mwzfs@163.com, Y. Ren reny_cumt@163.com,

engineering application of intelligent diagnosis methods.

In practical engineering, multiple clients often use the same device, and each customer may collect fault data. If the fault data is centralized, it can solve the problem of insufficient training samples faced by intelligent diagnosis methods. However, data privacy is becoming more and more important, and data sharing has the risk of information leakage. Although data can be protected by encryption, cleaning and other means [4, 5], there is a risk of data leakage once it leaves the local storage. Moreover, the transmission of data will also incur certain transmission costs. Therefore, the method of pooling data together to train intelligent diagnostic models has certain limitations.

To address the data privacy issue, the federated learning approach is applied to device fault diagnosis [6]. The usual practice is to train the fault diagnosis model on the client side using local data, and then upload the model parameters to the central server for averaging or weighting, and then distribute them back to all clients for further training, and so on until relevant training criteria are met. Finally, the trained model is used to diagnose the target sample [4, 7, 8]. To overcome the limited generalization ability of traditional federated learning in cross-domain learning, methods based on difference-weighted federal average (D-WFA) [9], federated diagnosis based on similarity collaboration (FedSC) [10], federated adversarial domain generalization [4], improved federated learning FA-FedAvg [11], and multi-method synthesis optimization aggregation strategy [12] have been proposed successively to enhance the training speed and generalization ability of the aggregated models.

Although federated learning can address the problem of multi-client coordinated fault diagnosis under data privacy constraints, most common federated learning methods assume that the client data distribution is roughly the same, that is, the working conditions are the same or similar [13]. However, in reality, the devices of different clients often work under different working conditions, the data distribution is significantly different, and even the same fault data features of different clients are hard to align, which poses a challenge for traditional federated learning. In recent years, transfer learning methods have been applied to fault diagnosis research under variable working conditions[14, 15], and have achieved good results [16-18]. Transfer learning methods based on domain adaptation usually require gathering source domain data and target domain data together to analyze their common features [19-21], which is not feasible for data privacy protection requirements. Although some methods such as identity recognition, privacy enhancement and fake data generation have been introduced into federated fault diagnosis [4, 22-24], they still require a large amount of data transmission, which incurs high data transmission costs and certain disclosure risks.

In fact, there may be very few labeled data in the target domain, which are called anchor point samples in this paper, and the use of these anchor point samples often plays a crucial role. One-shot learning is a common small-sample fault diagnosis method. The basic idea is to fine-tune the model trained in the source domain using a small number of samples [25-28]. Prototype network performs well in fault diagnosis methods with a small number of labeled samples, so it is used in many fault diagnosis methods with limited labeled samples [29-32]. Its basic principle is to measure the distance between the test sample and the known class centroid, and the prototype with a closer distance will be considered as the most likely class, so as to classify the sample. Since the prototype network only uses prototypes to represent classes in the training stage, rather than memorizing each specific sample feature, it has good generalization ability for small sample learning and zero sample learning tasks [33, 34]. However, the prototype network relies on the known complete class prototype, that is, the class center, which cannot accurately classify the client whose fault sample data is incomplete.

To address the problems in the previous research, we propose a multi-client collaborative federated transfer learning method for device intelligent fault diagnosis, which ensures data privacy. Our proposed method demonstrates robust performance even in scenarios where a single client's fault samples are incomplete, significant distribution differences exist between the source and target domains, and the calibration samples in the target domain are scarce. The key innovations of our method can be summarized as follows:

(1) The proposed method solely transmits the parameters of the local training model, achieving privacy protection for the client's local data. This approach reduces data transmission costs and enhances its applicability in engineering settings.

(2) We construct a multi-client collaborative learning framework to address the issue of weak model generalization caused by insufficient items in the training samples of individual clients.

(3) A targeted fine-tuning strategy is proposed based on the concept of a prototype network. We propose a training sample selection strategy based on relative distance to mitigate the negative impact of out-of-distribution (OOD) samples on model fine-tuning and enhance the final prediction accuracy of the model.

The rest of this paper is organized as follows: Section 2 introduces the related concepts and the formulation of the federated transfer diagnosis problem; Section 3 presents the basic principles; Section 4 describes the detailed procedure of the proposed method. Section 5 verifies the effectiveness of the method through experiments; Section 6 concludes the paper.

2. Problem Statement

2.1. Related concept

Cross-domain data refers to the state data collected under different working conditions or different device operating conditions. Due to the influence of working conditions, the data exhibits different distribution patterns [35].

Incomplete data means that in the same working condition (client), only some fault categories have labeled samples, while the data of other fault categories are missing, and the fault data of all fault categories are not available [36]. This is consistent with engineering practice, because the types of failures that occur in a piece of device are usually very limited.

Anchor samples refer to very few calibration fault data in the target domain (target client) [37]. This paper assumes that each fault category has only one labeled sample data, and the rest are unlabeled data, which is also consistent with engineering practice. After all, sample collection is very challenging.

2.2. Cross-domain Federation Transfer Diagnosis

Suppose there are N source domain clients $C_1, C_{2,} \cdots C_N$, each with a labeled dataset $\left(X_j^n, Y_j^n\right)$, where $j = 1,2,\cdots,N$, $n = 1,2,\cdots,m$, $m \leq R$, R is the total number of fault categories. Note that not every client has data for all R modes, meaning the client data is incomplete. There is only one target client, the target client only has one anchor point sample for each fault (X_{ok}^t, Y_{ok}^t), where $k = 1,2,\cdots,R$, and n_t unlabeled samples in the target domain $\{X_i^t | i = 1,2,\cdots,n_t\}$, $n_t \leq R$. Due to the different working conditions of the source and target domains, there are large differences in marginal distributions $P_{C_n}(x) \neq P_t(x)$ and conditional distributions $P_{C_n}(y|x) \neq P_t(y|x)$ between the datasets. Because of the scarcity of labeled samples in the target domain, it is hard to train a diagnostic model directly. To protect data privacy, we transfer the model trained by the source domain clients to the target client through federated learning across domains, and then fine-tune the model using one-shot learning and the proposed targeted fine-tuning strategy. Finally, we classify the samples to be diagnosed in the target domain.

3. Basic principles

3.1. Federated Learning

The federated learning framework proposed in this paper is illustrated in Figure 1. Several domain clients with incomplete data sources cooperate to train a global model, and then transfer the global model to the target client, who fine-tunes the model using one-shot learning and the proposed targeted fine-tuning strategy, and then applies it to diagnose test samples.

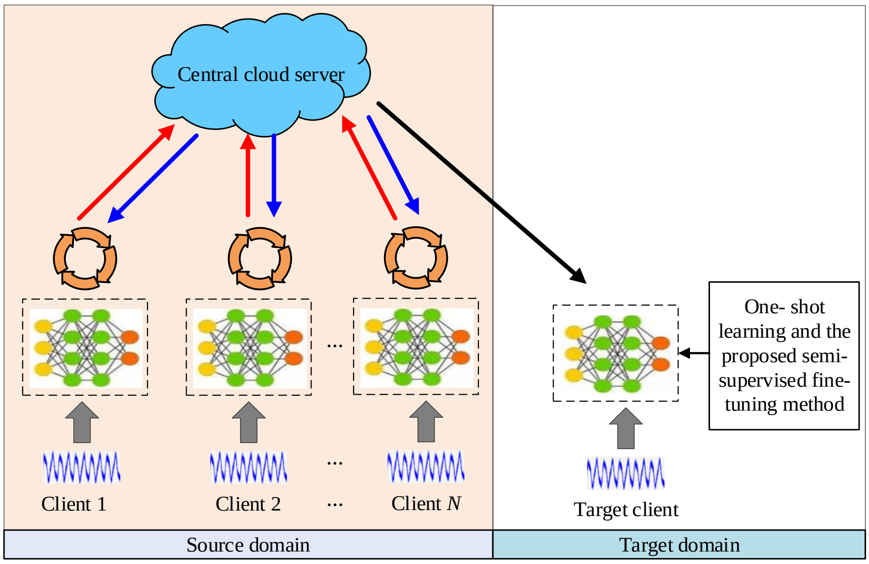


Fig. 1. The proposed federated learning framework.

Federated learning does not require sharing and accessing raw data with each other, but only trains models locally and then exchanges model parameters, thus protecting data privacy. In general, the central server of federated learning is used to coordinate training and compute the average of model parameters. The objective function of a typical federated averaging algorithm (FedAVG) is as follows:

$$\operatorname{argmin}\left\{F(\Theta) = \sum_{n\in N} p_n f_n(\Theta)\right\} \tag{1}$$

where $f_n(\Theta) = \mathbb{E}\left[\ell_n\left(\Theta; (X^n, Y^n)\right)\right]$ is the loss function of the nth client, Θ is the training model parameter, p_n is the aggregation weight of the client, usually $p_n = \frac{Q_n}{Q}$, Q_n is the sample size of the nth client participating in the training, and Q is the total sample size.

The federated learning training process proposed in this paper consists of the following steps:

1) The server selects the clients $C_1, C_{2,} \cdots C_N$ to participate in the training, and sends them the global model.

2) The clients participating in the training train locally and update their local model parameters.

3) The clients upload the updated local model parameter Θ_n^r to the server, where r represents the training round.

4) The server aggregates the uploaded model parameters and updates the global model parameters. The server checks whether the preset training round has been completed. If yes, it sends the updated global model to the target client. If no, it sends the updated global model to the source domain clients participating in the training, and repeats step 2).

3.2. Prototype classification network

Prototype network is a common meta-learning model that uses similarity to classify test data into different categories. After federated training, the global model acquires some meta-knowledge of classification, which can be applied by the prototype to classify test data. Suppose the anchor point samples of each fault category in the target domain are (X_{ok}^t, Y_{ok}^t), then their features as prototypes can be represented as:

$$S_k^t = f_\theta(X_{ok}^t) \tag{2}$$

where S_k^t is the feature prototype of class k fault samples, $f_\theta(\cdot)$ is the trained feature extractor, and θ is the model hyperparameters learned from the source domain. Since the target domain data may have a different distribution from the source domain data, θ also needs to be further trained and optimized to fit the classification needs of the target domain.

Then, the following formula can be used to calculate the probability of the test data belonging to class k fault.

$$P_\theta(y=k|x)=\frac{\exp\left(-d\left(f_\theta(X_i^t),S_k^t\right)\right)}{\sum_j \exp\left(-d\left(f_\theta(X_i^t),S_j^t\right)\right)} \tag{3}$$

where $d(u,v)$ is the distance between u and v, and the Euclidean distance is adopted in this paper. The smaller the distance between $f_\theta(X_i^t)$ and S_k^t, the higher the probability of class k fault.

4. The proposed method

4.1. Structure of the global model

Fig. 2 illustrates the structure of the global model that we propose in this paper. The model comprises a feature extractor and a classifier.

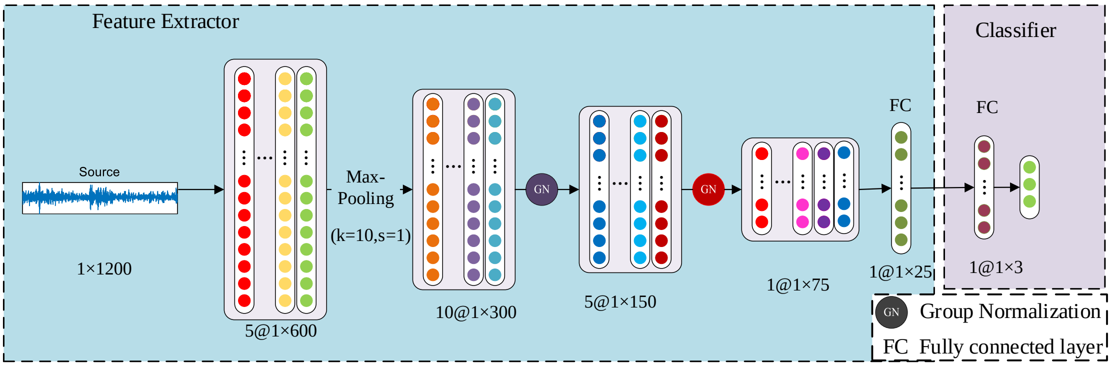


Fig. 2. Structure of the global model.

The feature extractor consists of four one-dimensional convolutional layers and one fully connected layer, with group normalization added to speed up the model convergence. The classifier has one fully connected layer. During the source domain training, the global model is trained on various clients, aiming to endow the feature extractor with adequate feature extraction ability.

During the local training process of the source domain client, the objective is to maximize the accuracy of the classifier, adhering to the objective function (3).

$$L_{Class}=-\frac{1}{n_S}\sum_{i=1}^{n_S}\sum_{j=1}^{K}\left[y_{ij}(1-\epsilon)\log\left(p_{ij}+\epsilon\right)+y_{ij}\frac{\epsilon}{R}\sum_{l=1}^{K}\log(p_{il}+\epsilon)\right] \tag{4}$$

Eq. (3) is the smooth cross-entropy loss function, which is used to enhance the generalization ability of the model and reduce the overfitting of the model. In the equation, ϵ denotes the smoothness coefficient, set to 0.1 in this study. n_S is the number of samples, while R represents the number of fault classes. Moreover, y_{ij} indicates whether the jth class of the ith sample is a true label, and p_{ij} represents the prediction probability of the jth class for the ith sample. The first term in the equation corresponds to the prediction probability of the correct category, and the second term represents the average prediction probability across all categories.

4.2. Targeted fine-tuning strategy

To harness valuable insights from unlabeled samples in the target domain and enhance the model's classification capabilities, the trained global model is fine-tuned with sample data from the target domain after being sent by the server. We propose a training sample selection method based on similarity, outlined in Fig. 3. Let's consider classification 3 as an example: Initially, three calibrated anchor point samples (one for each fault type) are fed into the trained global model to obtain their features, which serve as prototypes for each fault type. Subsequently, the unlabeled samples from the target domain are passed through the trained feature extractor to extract their features. The features of these unlabeled samples are then compared to the features of the three anchor points, measuring their distance using the Euclidean distance metric in this study. The prototype with the closest distance indicates the highest probability of belonging to the corresponding fault category. As depicted in Fig. 3, for the three cases of ABC, the fault type of the unlabeled samples can be identified by their distance from the prototype; i.e., the one with the smallest distance corresponds to the same fault type. These samples are termed in-distribution samples (ID). However, there might be cases like DE, where the features of unlabeled samples are situated near or far from the three anchor points. These samples are referred to as out-of-distribution samples (OOD). Classifying OOD samples poses challenges and may lead to misjudgments. Therefore, it is essential to eliminate such samples during the training process. To differentiate OOD samples and ID samples, we construct a relative distance threshold screening model, represented by Eq. (4).

$$y_i^u = \begin{cases} \arg\min_k\{B_{ik}\} & if \ \min_k\{B_{ik}\} \le \beta \\ -1 & \text{otherwise} \end{cases} \quad (5)$$

where B_{ik} represents the relative distance, as shown in Eq. (5). β denotes the threshold, and a value of $\beta = 0.3$ is adopted in this study.

$$B_{ik} = \frac{d(f_\theta(X_i^t), S_k^t)}{\min_{j\neq k} d\left(f_\theta(X_i^t), S_j^t\right) + \delta} \quad (6)$$

$\min_{j\neq k} d\left(f_\theta(X_i^t), S_j^t\right)$ represents the minimum distance except $d(f_\theta(X_i^t), S_k^t)$. δ is a fixed value, introduced to avoid division by zero in Eq. (5). Specifically, δ is set to $\delta = 10^{-6}$.

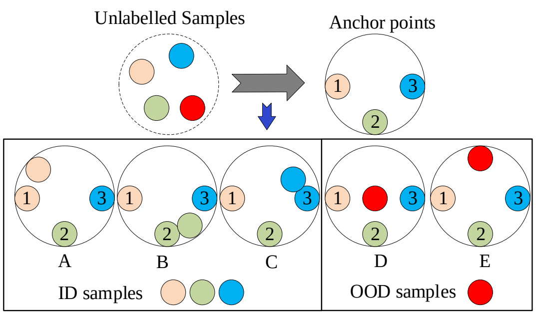


Fig. 3. Training strategies for unlabeled samples.

After the federated training of the source domain, the global model is sent to the target client. However, due to the variance in data distribution between the target client and the source domain client, the global model requires fine-tuning. As the target client possesses just one labeled anchor point sample for each fault type, extracting effective information from the unlabeled data becomes crucial. To achieve this, the features of the labeled data serve as anchor points for the fault prototype. Unlabeled samples are selected in batches using Eq. (4) for training. The objective of model fine-tuning is to align the extracted features closer to the real classification anchor while moving away from the unreal classification anchor. Hence, one of the loss functions employed during the training process is:

$$L_{dist}^t = \frac{1}{n_b}\sum_{i=1}^{n_b}\sum_{k=1}^{R}\left(\gamma d\left(f_\theta(X_i^t), S_k^t\right) + \log\sum_{j=1}^{R}\exp\left(-\gamma d\left(f_\theta(X_i^t), S_j^t\right)\right)\right) \quad (6)$$

In Eq. (6), where n_b represents the number of selected samples, and γ is the scaling factor. The first term aims to bring the extracted features closer to the real class anchor, while the second term ensures that the extracted features are pushed farther away from other class anchor points.

Additionally, the selected samples are assigned false labels using Eq. (7). Subsequently, the smoothing cross-entropy loss function from Eq. (3) is utilized as the constraint during the training of the second loss function. For prediction probability, the proposed attribution probability formula is employed, incorporating an enhanced stretching factor as depicted in Eq. (7).

$$P_{\theta,\gamma}(y = k|x) = \frac{\exp\left(-\gamma \cdot d\left(f_\theta(X_i^t), S_k^t\right)\right)}{\sum_j \exp\left(-\gamma \cdot d\left(f_\theta(X_i^t), S_j^t\right)\right)} \quad (7)$$

The smooth cross-entropy loss can be expressed in the following form, as shown in Eq. (8).

$$L_{Class}^t = -\frac{1}{n_b}\sum_{i=1}^{n_b}\sum_{j=1}^{R}\left[y_{ij}(1-\epsilon)\log\left(P_{\theta,\gamma}(y = j|x) + \epsilon\right) + y_{ij}\frac{\epsilon}{R}\sum_{l=1}^{R}\log\left(P_{\theta,\gamma}(y = l|x) + \epsilon\right)\right] \quad (8)$$

Hence, the comprehensive loss function for the local fine-tuning process of the diagnostic model can be formulated as follows:

$$L_{Total}^{S} = L_{dist}^{t} + L_{Class}^{t} \tag{9}$$

Eq. (8) ensures that the model can effectively adjust to specific nuances in the target domain. The use of these equations with specific parameters like ϵ and potentially γ, allows the model to maintain numerical stability and improve its generalization capabilities across different operational conditions and data distributions.

4.3. Federated transfer diagnostic steps

The overall flow of the proposed federated migration fault diagnosis, as depicted in Fig. 4, consists of three steps:

Step 1: Federated Global Model Training

Each source domain client employs its labeled training data for model training and then uploads the model parameters to the central server. The central server averages the model parameters and distributes them back to all clients for further training. This process continues until the specified number of cycles is reached. Once completed, the central server sends the trained global model to the target client. Both the feature extractor and classifier are trainable during this step.

Step 2: Global Model Localization Fine-Tuning

Upon receiving the global model, the target client utilizes labeled anchor point samples to fine-tune the model. Subsequently, unlabeled training data is used for further tuning, employing the targeting method outlined in Section 4.2. Throughout this process, the feature extractor remains trainable.

Step 3: Target Domain Test Sample Fault Prediction

Following the initial two training steps, the test sample from the target client is input into the trained model to predict its fault category. During this process, the feature extractor remains fixed, and Eq. (7) is utilized to determine the final classification of the test sample.

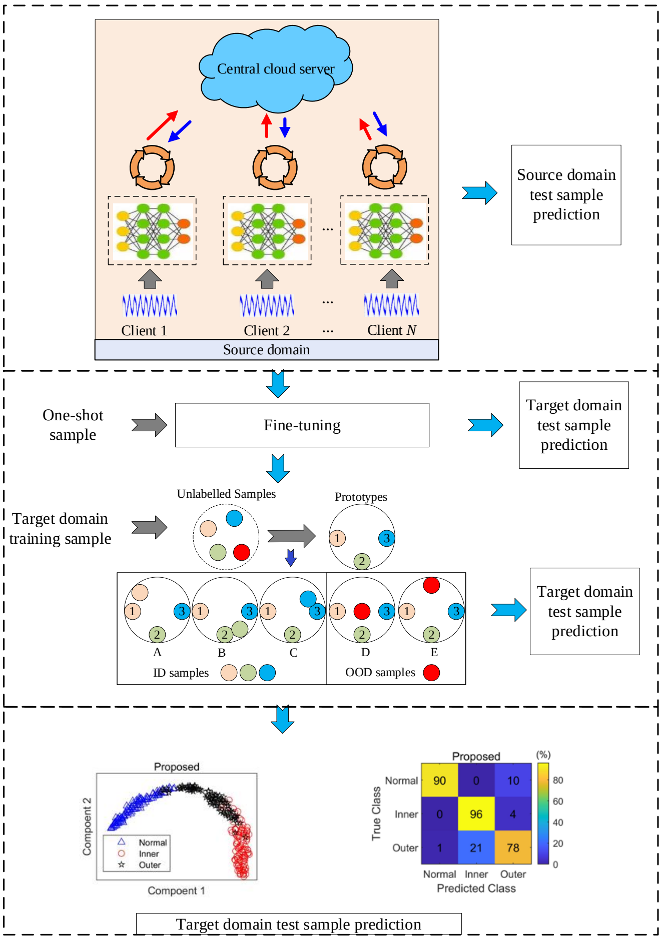


Fig. 4. Overall process.

To compare the training effects at each step, this paper includes corresponding test verifications after the completion of each training phase. The specific results can be found in Section 5 under the Experimental Analysis.

The steps for updating local model parameters are as follows:

Step 1: Initialize Prototype Network

Use the few labeled anchor samples from the target domain to initialize the prototypes by extracting their features with the globally trained model.

Step 2: Feature Extraction

Process the unlabeled samples from the target domain to obtain their feature representations using the feature extractor part of the global model.

Step 3: Distance Calculation and Sample Classification

Calculate the distances between the features of the unlabeled samples and the prototypes of each fault category. Temporarily classify the unlabeled samples into corresponding fault categories based on the principle of minimum distance.

Step 4: Construct Relative Distance Threshold Model

Define the calculation method for relative distance and set a threshold to distinguish between "ID" and "OOD" samples.

Step 5: Sample Selection

Select ID samples for further training based on the calculated relative distances and the set threshold.

Step 6: Model Fine-tuning

Fine-tune the global model using the selected ID samples to optimize the model to better adapt to the data from the target domain.

Step 7: Apply Loss Function

Guide the model learning in the fine-tuning process by applying specific loss functions as shown in Eq. (6) and (9).

Step 8: Iterative Optimization

Optimize the model parameters through multiple iterations of training until the model's performance on the target domain data is satisfactory.

Step 9: Model Evaluation

Evaluate the fine-tuned model on the test sample set from the target domain to ensure its accuracy and generalization capability.

5. Experiments

In this paper, two experimental scenarios are designed to validate the effectiveness of the proposed method. The first experiment focuses on transfer learning fault diagnosis under different working conditions of the same device. The second experiment, on the other hand, involves transfer learning fault diagnosis between different devices.

5.1. Fault Diagnosis of Cross Condition Federated Transfer

(1) Experimental Introduction

The experimental data for this study were obtained from the publicly available dataset of the SQ (Spectra Quest) experimental platform at Xi'an Jiaotong University [38].

Fig. 5. Spectra Quest test bed.

The test bench used in the experiments is depicted in Fig. 5. For the experiments, the bearing model employed is the NSK Company's 6203 bearing. The dataset replicates three failure modes of motor bearings: normal bearing, outer ring failure, and inner ring failure. Vibration signals from the motor bearings with varying degrees of fault (mild fault, moderate fault, and severe fault) were collected at three different rotation frequencies (19.05Hz, 29.05Hz, and 39.05Hz). To capture the motor bearing signals, a piezoelectric acceleration sensor was employed during the experiments. The data acquisition instrument used was the CoCo80 model, operating at a sampling frequency of 25.6KHz. The acceleration sensor utilized has a sensitivity of 50mv/g.

In this paper, the data collected at a rotation frequency of 29.05Hz is chosen as the source domain data, while the data at 39.05Hz rotation frequency serves as the target domain data. For this experiment, each source domain client possesses labeled data representing only one fault category. On the other hand, the target domain client possesses three anchor samples, each corresponding to a fault category, along with a substantial amount of unlabeled sample data, as depicted in Table 1. The samples in the table were obtained using the sliding window method, with sampling intervals of 600 points. Each sample consists of 1200 points.

Table 1. Relevant attributes of cross-working condition experimental data

Clients	Rotation frequency	Fault category	Fault Degree	Sample size	label
SC1		N	-	639	labeled
SC2	29.05Hz	I	mild	639	labeled
SC3		O	moderate	639	labeled
TC	39.05Hz	N/I/O	unknown	639*3	unlabeled

Remarks: "SC" refers to the source domain client, while "TC" refers to the target client. Additionally, "N" indicates normal, "I" represents inner ring fault, and "O" stands for outer ring fault. It is worth noting that the inner ring fault in the target client ("TC") is mild, whereas the outer ring fault is moderate.

As evident from Table 1, there exists a difference in the rotation frequency between the source domain client and the target domain client, indicating varying working conditions. The bearing exhibits three fault modes, and each source domain client contains only one sample corresponding to a particular fault type, resulting in incomplete data.

(2) Experimental setting

For the sake of ease in comparison, among the samples presented in Table 1, 100 samples of each fault category are reserved for both the source domain and the target domain as test samples. The remaining 539 samples are utilized as training samples. Consequently, both the source domain and the target domain possess 300 test samples each, which will be used to assess the model's effectiveness.

To validate the effectiveness of the proposed method in this paper, several conventional approaches are employed for simultaneous learning and prediction. In this paper, various parameter settings were tested multiple times, and for each method, the parameters yielding the best prediction results were carefully selected.

a) Baseline: Only target domain samples are used for training and prediction. There is no usage of source domain data or federation learning. The model utilizes the target domain anchor samples and unlabeled samples for training. After supervised training with labeled anchor sample data (60 times), the method proposed in Section 4.2 is applied for further targeted fine-tuning of the model (600 times).

b) FedAvg: This approach involves multiple federation training using labeled samples from the source domain client, according to the training steps in Section 3.1 and the model structure in Fig. 2. The global model training occurs on a central server and is then sent to the target client. Test samples are input into the trained global model for classification and prediction. The total number of training rounds is set to 300, with each client training 30 times per round.

c) FedAvg with Fine-tuning (FedAvgFT): Similar to FedAvg, but after transferring the global model to the target client, the classifier undergoes fine-tuning using the anchor samples of the target client. The test samples are then input into the fine-tuned model for classification and prediction. The number of training rounds in the source domain is the same as FedAvg (300 rounds), and the number of fine-tuning training on anchor samples is 30 times.

d) Centralized: All source domain and target domain client data are centralized on a central server for training.

The model structure is shown in Fig. 2. After the source domain training, the feature extractor and classifier directly input the test samples for classification prediction. This method lacks privacy protection and involves 30,000 training times for the model.

e) Centralized with Fine-tuning (CentrFT): Similar to the Centralized approach, but after training the source domain, the target client anchor samples are used to fine-tune the model. Test samples are then input into the fine-tuned model for classification and prediction. Like the Centralized method, this approach also lacks privacy protection, and the number of model fine-tuning training is set to 30 times.

f) Proposed method (Proposed): The steps in Fig. 4 involve training the global model with the central server and the source domain client. The trained global model is then sent to the target client. The target client fine-tunes the global model using the tagged anchor samples and further fine-tunes it using the untagged training samples to make it learn the distribution law of the target samples as much as possible. Finally, the target domain test samples are input into the trained model for fault prediction. The training times for the global model are the same as FedAvgFT, and the targeted fine-tuning is performed 90 times.

(3) Experimental results

The learning rate of each module of the above six methods is set to 0.0001, and the training batch is set to 32. Each method is repeated 5 times, and the average of the predicted results is taken, as presented in Table 2.

Table 2. Fault recognition rate of federated transfer learning across working conditions (%).

Methods	Source domain	Target domain
Baseline[39]	—	37.34±1.23
FedAvg[40]	75.87±1.17	34.20±0.18
FedAvgFT	Same as FedAvg	67.67±6.02
Centralized[41]	91.34±1.65	48.27±2.02
CentrFT	Same as Centralized	76.13±10.74
Proposed	Same as FedAvg	**86.20±1.91**

As shown in Table 2, it is evident that the proposed method outperforms the other methods significantly. From the perspective of the sample recognition rate in the source domain test, the central learning method demonstrates better performance compared to the federated learning method. This is because the average calculation method of model parameters in federated learning is less effective in learning the distribution law of data than the adaptive acquisition method used in central learning. Analyzing the recognition rate of test samples in the target domain, it becomes apparent that the model trained directly using the source domain yields very low recognition rates. This indicates that there are significant differences in the feature distribution between the source domain and the target domain, which is further supported by the feature visualization diagram (Fig. 6). Interestingly, even without source domain training, the targeted training method proposed in this paper achieves a recognition rate of 37.34%. This rate surpasses the recognition rate obtained by the global model through federated learning, showcasing the superiority of the proposed targeted training approach.

To further illustrate the feature extraction effects of the different methods, the features extracted by each method were reduced to 2 dimensions using the t-distributed neighborhood embedding algorithm (t-SNE) and then visualized in Fig. 6. Additionally, the specific results of the identification of three fault categories were plotted in a confusion matrix, as depicted in Fig. 7.

Upon observing Fig. 6 and Fig. 7, it is evident that, overall, the proposed method achieves the most balanced differentiation among the three types of faults. The prediction results of the Baseline method demonstrate the importance of transfer learning in fault diagnosis across various working conditions. Solely using target domain data does not lead to the desired prediction performance. Analyzing the prediction results of FedAvgFT and CentrFT, it becomes apparent that Fine-tuning plays a role in optimizing transfer learning, but its effectiveness is limited when dealing with a small sample size. Comparing the proposed method with the results from FedAvgFT, it is clear that the anchor point-based classification, using distance measurement, outperforms the classifier based on decision boundary classification when the number of fine-tuning samples is small.

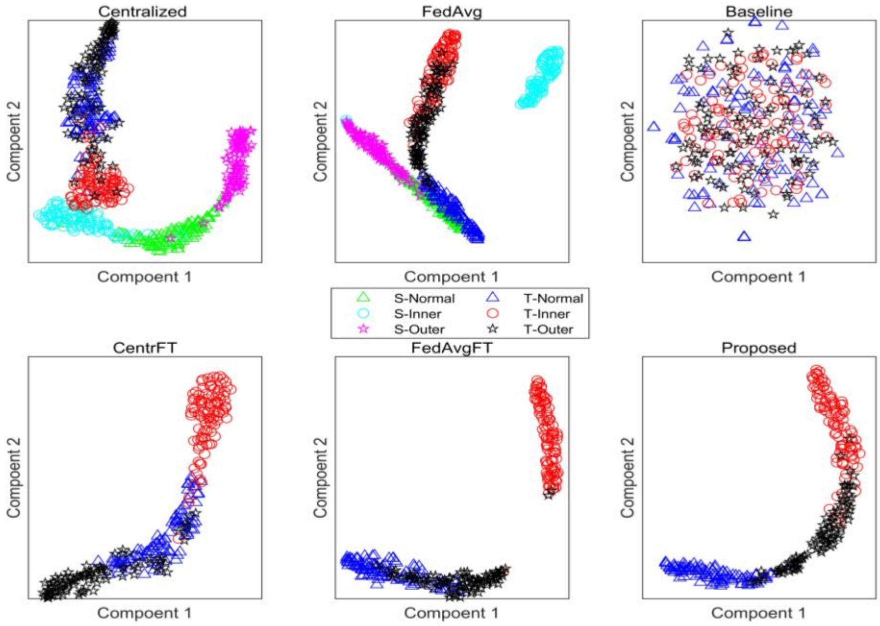


Fig. 6. Visual result of dimensionality reduction under cross-working conditions (in the legend, "S-" represents source domain feature, "T-" represents target domain feature).

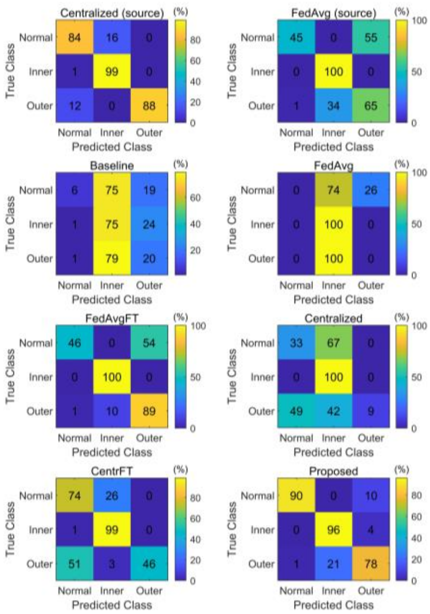


Fig. 7. Confusion matrix of identification results under cross-working conditions.

5.2. Cross-device Experiment

(1) Experimental Introduction

The source domain data for this experiment was obtained from the bearing dataset of Paderborn University in Germany [42], and the test bench used is depicted in Fig. 8. Similar to the previous experiment, the experimental bearing used is the 6203 bearing, and the vibration signals of the bearing seat are collected using a piezoelectric accelerometer with a sampling frequency of 64kHz.

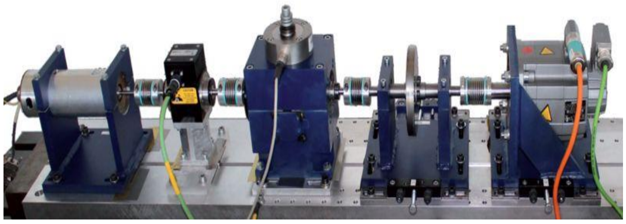


Fig. 8. Bearing data test bench of Paderborn University.

The experiment was conducted under various load torques, speeds, and radial forces. For this study, vibration data for normal, inner ring fault, and outer ring fault at 1500rpm (25Hz), with a load of 0.1Nm, and a radial force of 0.1N were selected as the source domain samples. The sliding window method was employed to collect 1200 data points, with an overlap of 600 points. The target domain samples are the same as those used in Experiment 5.1. All the attributes of the experimental data are presented in Table 3.

Table 3 related parameters of the cross-device experimental data

Clients	Rotation frequency	Fault category	Fault Degree	Sample size	label
SC1		N	-	852	labeled
SC2	25Hz	I	2	852	labeled
SC3		O	2	852	labeled
TC	39.05Hz	N/I/O	unknow	639*3	unlabeled

Even though the bearing model remains the same between the source and target domains, different devices operate under distinct load and speed conditions. Moreover, the data sampling frequency and acquisition sensors also vary across devices. Additionally, the fault severity level in the source domain is rated as level 2. However, in reality, the inner ring fault in the target domain is mild, and the outer ring fault is moderate. Under this cross-device scenario, fault diagnosis is transferred, and several methods used in Experiment 1 are employed for prediction, with the relevant parameters set accordingly, as in Experiment 1.

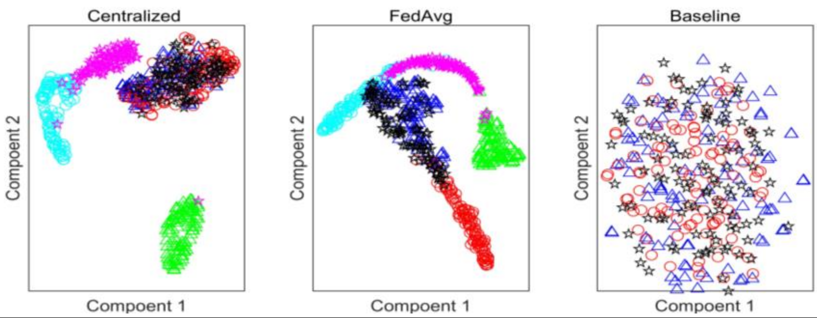
(2) Experimental Results

The six methods mentioned in Section 5.1 were applied to the experimental data, and each method was repeated five times. The average values of the predicted results are presented in Table 4.

Table 4 experiment results of cross-device

Methods	Source domain	Target damain
Baseline	—	37.34±1.23
FedAvg	92.53±1.98	39.67±4.67
FedAvgFT	Same as FedAvg	70.87±8.48
Centralized	94.40±2.61	33.33±0
CentrFT	Same as Centralized	44.07±5.35
Proposed	Same as FedAvg	**81.80±2.14**

It should be noted that since the Baseline method does not use the source domain samples for training, and the target domain data in both experiments are the same, the prediction results should also be identical. Hence, the results from Experiment 1 are directly used in Table 4. As shown in Table 4, the proposed method continues to exhibit the most effective recognition performance. The features extracted by each method were reduced to lower dimensions using t-SNE and visualized, as depicted in Fig. 9. Furthermore, the recognition results of each method were presented in the form of a confusion matrix, shown in Fig. 10.



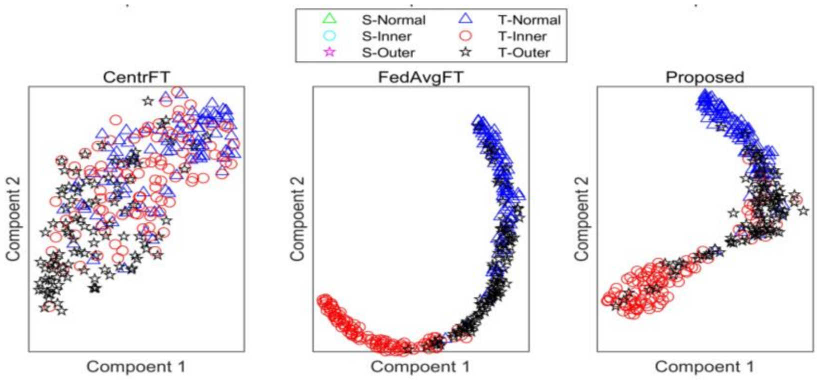


Fig. 9. Visual results of feature dimensionality reduction across devices.

In the context of cross-device transfer for fault diagnosis, it can be observed from Fig. 9 and 10 that whether the model obtained through centralized training or federated training is directly used for the prediction of target test samples, the feature aggregation appears to be relatively poor and dispersed. This suggests that the difference between the sample distributions in the source domain and the target domain is greater than that in Experiment 1. However, even under such challenging cross-device transfer learning conditions, the proposed method continues to perform admirably.

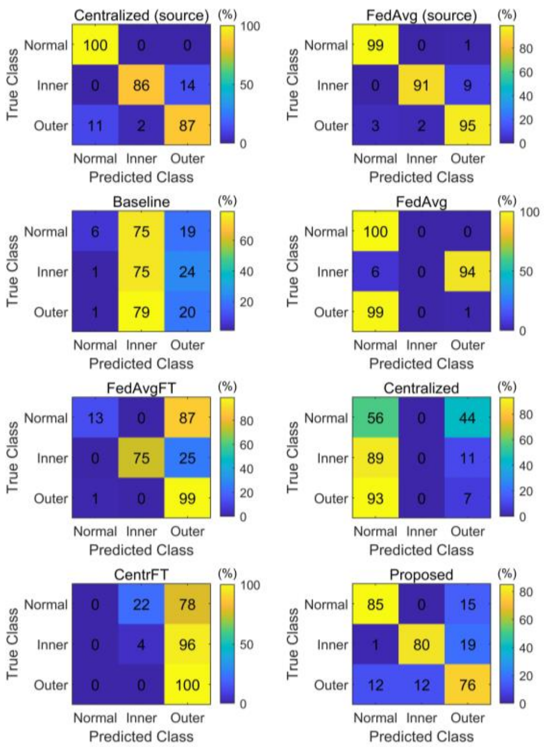


Fig. 10. Confusion matrix of identification results under cross-device conditions.

5.3. Analysis of Experimental Results

In the two experiments conducted in this paper, each source domain client only has one fault type calibration sample, making it an incomplete dataset. Due to the disparity in sample distribution between the source and target domains, traditional feature extractor and classifier methods yield suboptimal prediction results, even with a very small number of fine-tuned samples. Moreover, adopting the federated learning method of feature extraction combined with the prototype network results in poor model generalization due to the incomplete data samples from each client. To address this issue, this paper proposes a combination of federated learning and targeted tuning methods. The cross-device experiment and cross-device cross-condition experiment demonstrate that the proposed method performs well, even with few anchor samples. To verify the effect of the scaling factor γ in the proposed method, two experimental results under different scaling factors were plotted as a graph, as shown in Fig. 11.

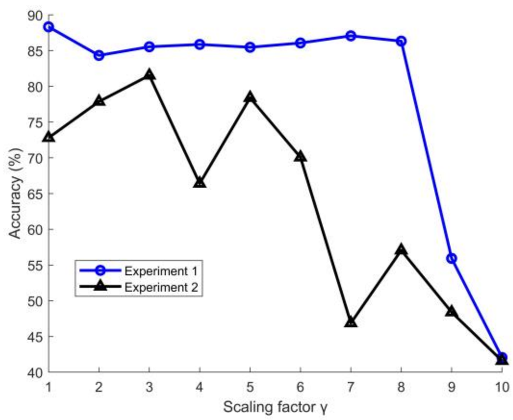


Fig. 11. Impact of scaling factors on prediction results.

The prediction result presented in Fig. 11 is the average value of the prediction results obtained from each run, conducted 5 times, under different scaling factors. It is evident that the prediction result is sensitive to the influence of scaling factors. The optimal scaling factor for Experiment 1 is 1, while for Experiment 2, it is 3. This suggests that the optimal scaling factor is not fixed and is dependent on the data distribution. It should be emphasized that the two experimental predictions mentioned in Table 2 and Table 4 in this paper were obtained when $\gamma = 3$.

The prediction results of the two experiments are drawn into a histogram, as shown in Fig. 12.

As depicted in Fig. 12, it is evident that the prediction accuracy of experiment 1 surpasses that of experiment 2 overall. This suggests that the transfer learning effect for fault diagnosis on the same device is superior to cross-device transfer learning. It is worth noting that the fault feature distribution varies to some extent among different devices, which explains the observed differences. Nevertheless, the proposed method introduces a valuable perspective for cross-device transfer learning.

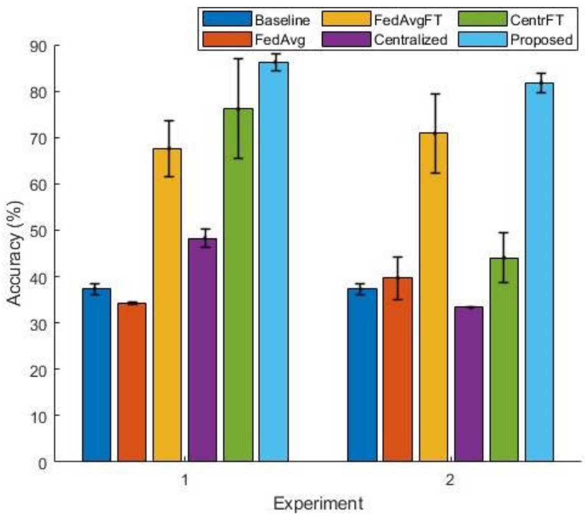


Fig. 12. Histogram of prediction results of two experiments.

6. Conclusion

In this paper, we address the challenge of multi-user collaborative intelligent diagnosis by presenting a federated transfer learning method that ensures data privacy. We establish a multi-client collaborative federated learning framework and introduce a targeted fine-tuning strategy for cross-working conditions and even cross-device transfer learning diagnosis models, thus enabling adaptive fault diagnosis across multiple clients under incomplete data conditions. Our proposed method offers several key advantages:

1) Data Privacy: By transmitting only the parameters of the local training model, we ensure the privacy protection of the client's local data. This approach reduces data transmission costs and promotes practical applicability.

2) Improved Generalization: We construct a multi-client collaborative learning framework to address the issue of weak model generalization caused by limited terms in single-client training samples.

3) Federated & Prototype Network Integration: By combining the strengths of federated learning and prototype network, we tackle the challenge of joint training and transfer with incomplete client data. Leveraging the advantages of prototype network in small-sample transfer learning significantly enhances the prediction accuracy of the transfer model.

4) Targeted Fine-Tuning Strategy: We propose a targeted fine-tuning strategy based on relative distance, which effectively reduces the probability of negative fine-tuning of out-of-distribution (OOD) samples and improves the recognition rate of the diagnostic model.

5) Broad Applicability: Our proposed method effectively addresses multi-user collaborative fault intelligent diagnosis through federated learning. It demonstrates exceptional performance in cross-working conditions and even cross-device migration diagnosis scenarios.

Looking ahead, future research will focus on studying the adaptive nature of federated learning models under self-supervision. We aim to overcome the limitation of insufficient labeled samples, which currently restricts the generalization ability of intelligent diagnosis models.

In conclusion, our federated transfer learning approach offers a promising solution for multi-user collaborative fault intelligent diagnosis, ensuring privacy and enhancing performance in various challenging scenarios. Subsequent efforts will continue to advance the field by addressing additional challenges in adaptive federated learning models and improving the generalization ability of intelligent diagnosis models in the absence of sufficient labeled data.

Funding

This work was supported in part by the Suzhou Science and Technology Foundation of China under Grant SYG202021 and SYG202351, Jiangsu Provincial Natural Science Research Foundation of China under Grant 21KJA510003.

Reference

1. L.Q. Xia, Y.S. Liang, P. Zheng, X. Huang, Residual-Hypergraph Convolution Network: A Model-Based and Data-Driven Integrated Approach for Fault Diagnosis in Complex Equipment, Ieee Transactions on Instrumentation and Measurement, 72 (2023). https://doi.org/10.1109/TIM.2022.3227609
2. Y. Wang, W.L. Sun, L.Q. Liu, B.K. Wang, S.H. Bao, R.B. Jiang, Fault Diagnosis of Wind Turbine Planetary Gear Based on a Digital Twin, Applied Sciences-Basel, 13 (2023).https://doi.org/10.3390/app13084776
3. Y.Y. Xiao, L.T. Chen, S.Y. Chen, Z.H. Hu, J.Y. Tang, Intelligent fault diagnosis of gear crack based on side frequency feature under different working conditions, Measurement Science and Technology, 34 (2023). https://doi.org/10.1088/1361-6501/acd9df
4. R. Wang, W.G. Huang, M.K. Shi, J. Wang, C.Q. Shen, Z.K. Zhu, Federated adversarial domain generalization network: A novel machinery fault diagnosis method with data privacy, Knowledge-Based Systems, 256 (2022). https://doi.org/10.1016/j.knosys.2022.109880
5. W. Zhang, Z.W. Wang, X. Li, Blockchain-based decentralized federated transfer learning methodology for collaborative machinery fault diagnosis, Reliability Engineering & System Safety, 229 (2023). https://doi.org/10.1016/j.ress.2022.108885
6. A. Kumar, A.M. Shaikh, Y. Li, H. Bilal, B.Q. Yin, Pruning filters with L1-norm and capped L1-norm for CNN compression, Applied Intelligence, 51 (2021) 1152-1160. https://doi.org/10.1007/s10489-020-01894-y
7. Q.Z. Wang, Q. Li, K. Wang, H. Wang, P. Zeng, Efficient federated learning for fault diagnosis in industrial cloud-edge computing, Computing, 103 (2021) 2319-2337. https://doi.org/10.1007/s00607-021-00970-6
8. K. Zhao, J.C. Hu, H.D. Shao, J.B. Hu, Federated multi-source domain adversarial adaptation framework for machinery fault diagnosis with data privacy, Reliability Engineering & System Safety, 236 (2023). https://doi.org/10.1016/j.ress.2023.109246
9. J.B. Chen, J.P. Li, R.Y. Huang, K. Yue, Z.Y. Chen, W.H. Li, Federated Transfer Learning for Bearing Fault Diagnosis With Discrepancy-Based Weighted Federated Averaging, Ieee Transactions on Instrumentation and Measurement, 71 (2022).https://doi.org/10.1109/TIM.2022.3180417
10. Y.H. Zhang, X.A. Xue, X.P. Zhao, L.H. Wang, Federated learning for intelligent fault diagnosis based on similarity collaboration, Measurement Science and Technology, 34 (2023). https://doi.org/10.1088/1361-6501/acab22
11. D.Q. Geng, H.W. He, X.C. Lan, C. Liu, Bearing fault diagnosis based on improved federated learning algorithm, Computing, 104 (2022) 1-19. https://doi.org/10.1007/s00607-021-01019-4
12. J.H. Du, N. Qin, D.Q. Huang, Y.M. Zhang, X.M. Jia, An Efficient Federated Learning Framework for Machinery Fault Diagnosis With Improved Model Aggregation and Local Model Training, Ieee Transactions on Neural Networks and Learning Systems, (2023).
13. H.F. Sun, S.Q. Li, F.R. Yu, Q. Qi, J.Y. Wang, J.X. Liao, Toward Communication-Efficient Federated Learning in the Internet of Things With Edge Computing, Ieee Internet of Things Journal, 7 (2020) 11053-11067. https://doi.org/10.1109/JIOT.2020.2994596
14. F.Y. Lu, Q.B. Tong, X.D. Jiang, Z.W. Feng, J.J. Xu, X. Wang, J.Y. Huo, A deep targeted transfer network with clustering pseudo-label learning for fault diagnosis across different Machines, Mechanical Systems and Signal Processing, 213 (2024).https://doi.org/10.1016/j.ymssp.2024.111344
15. F.Y. Lu, Q.B. Tong, J.J. Xu, Z.W. Feng, X. Wang, J.Y. Huo, Q.Z. Wan, Towards multi-scene learning: A novel cross-domain adaptation model based on sparse filter for traction motor bearing fault diagnosis in high-speed EMU, Advanced Engineering Informatics, 60 (2024).https://doi.org/10.1016/j.aei.2024.102536
16. X.H. Chen, R. Yang, Y.H. Xue, M.J. Huang, R. Ferrero, Z.D. Wang, Deep Transfer Learning for Bearing Fault Diagnosis: A Systematic Review Since 2016, Ieee Transactions on Instrumentation and Measurement, 72 (2023). https://doi.org/10.1109/TIM.2023.3244237
17. C.H. Qian, J.J. Zhu, Y.H. Shen, Q.S. Jiang, Q.K. Zhang, Deep Transfer Learning in Mechanical Intelligent Fault Diagnosis: Application and Challenge, Neural Processing Letters, 54 (2022) 2509-2531. https://doi.org/10.1007/s11063-021-10719-z
18. Y.S. Zou, K.M. Shi, Y.Z. Liu, G.F. Ding, K. Ding, Rolling bearing transfer fault diagnosis method based on adversarial variational

autoencoder network, Measurement Science and Technology, 32 (2021). https://doi.org/10.1088/1361-6501/ac1461
19. J.C. Kuang, G.H. Xu, T.F. Tao, Q.Q. Wu, C.C. Han, F. Wei, Domain Conditioned Joint Adaptation Network for Intelligent Bearing Fault Diagnosis Across Different Positions and Machines, Ieee Sensors Journal, 23 (2023) 4000-4010. https://doi.org/10.1109/JSEN.2023.3235370
20. B.Y. Yang, T.T. Wang, J.S. Xie, J.S. Yang, Deep Adversarial Hybrid Domain-Adaptation Network for Varying Working Conditions Fault Diagnosis of High-Speed Train Bogie, Ieee Transactions on Instrumentation and Measurement, 72 (2023). https://doi.org/10.1109/TIM.2023.3276007
21. Q. Hu, X.S. Si, A.S. Qin, Y.R. Lv, M. Liu, Balanced Adaptation Regularization Based Transfer Learning for Unsupervised Cross-Domain Fault Diagnosis, Ieee Sensors Journal, 22 (2022) 12139-12151. https://doi.org/10.1109/JSEN.2022.3174396
22. S.X. Lu, Z.W. Gao, Q.F. Xu, C.X. Jiang, A.H. Zhang, X.X. Wang, Class-Imbalance Privacy-Preserving Federated Learning for Decentralized Fault Diagnosis With Biometric Authentication, Ieee Transactions on Industrial Informatics, 18 (2022) 9101-9111. https://doi.org/10.1109/TII.2022.3190034
23. C.C. Che, H.W. Wang, M.L. Xiong, X.M. Ni, Few-shot fault diagnosis of rolling bearing under variable working conditions based on ensemble meta-learning, Digital Signal Processing, 131 (2022). https://doi.org/10.1016/j.dsp.2022.103777
24. P.Q. Wang, J.D. Li, S.B. Wang, F.S. Zhang, J.J. Shi, C.Q. Shen, A new meta-transfer learning method with freezing operation for few-shot bearing fault diagnosis, Measurement Science and Technology, 34 (2023). https://doi.org/10.1088/1361-6501/acc67b
25. X.B. Liu, H.T. Guo, Y.B. Liu, One-Shot Fault Diagnosis of Wind Turbines Based on Meta-Analogical Momentum Contrast Learning, Energies, 15 (2022). https://doi.org/10.3390/en15093133
26. Yang, Y.F. Zhou, X. Chen, C. Li, H. Song, Fault diagnosis for wind turbines with graph neural network model based on one-shot learning, Royal Society Open Science, 10 (2023). https://doi.org/10.1098/rsos.230706
27. Y.H. Li, K. Li, X. Liu, Y.X. Wang, L. Zhang, Lithium-ion battery capacity estimation-A pruned convolutional neural network approach assisted with transfer learning, Applied Energy, 285 (2021). https://doi.org/10.1016/j.apenergy.2020.116410
28. L. Cheng, X.W. Kong, J.Q. Zhang, M.Z. Yu, A Novel Adversarial One-Shot Cross-Domain Network for Machinery Fault Diagnosis With Limited Source Data, Ieee Transactions on Instrumentation and Measurement, 71 (2022). https://doi.org/10.1109/TIM.2022.3198486
29. X.L. Zhang, Z.Q. Su, X.L. Hu, Y. Han, S.X. Wang, Semisupervised Momentum Prototype Network for Gearbox Fault Diagnosis Under Limited Labeled Samples, Ieee Transactions on Industrial Informatics, 18 (2022) 6203-6213. https://doi.org/10.1109/TII.2022.3154486
30. C.X. Jiang, H. Chen, Q.F. Xu, X.X. Wang, Few-shot fault diagnosis of rotating machinery with two-branch prototypical networks, Journal of Intelligent Manufacturing, 34 (2023) 1667-1681. https://doi.org/10.1007/s10845-021-01904-x
31. Z.F. Xu, M. Bashir, Q.S. Liu, Z.F. Miao, X.Y. Wang, J. Wang, N. Ekere, A novel health indicator for intelligent prediction of rolling bearing remaining useful life based on unsupervised learning model, Computers & Industrial Engineering, 176 (2023). https://doi.org/10.1016/j.cie.2023.108999
32. Zhang, J.Y. Jiao, J. Lin, H. Li, J.D. Hua, D. He, Uncertainty-based contrastive prototype-matching network towards cross-domain fault diagnosis with small data, Knowledge-Based Systems, 254 (2022). https://doi.org/10.1016/j.knosys.2022.109651
33. X.P. Zhao, M.Y. Ma, F. Shao, Bearing fault diagnosis method based on improved Siamese neural network with small sample, Journal of Cloud Computing-Advances Systems and Applications, 11 (2022). https://doi.org/10.1186/s13677-022-00350-1
34. R.J. Hou, Z.Y. Chen, J.L. Chen, S.L. He, Z.T. Zhou, Imbalanced fault identification via embedding-augmented Gaussian prototype network with meta-learning perspective, Measurement Science and Technology, 33 (2022). https://doi.org/10.1088/1361-6501/ac3c1d
35. Y.C. Zhang, K. Feng, H. Ma, K. Yu, Z.H. Ren, Z. Liu, MMFNet: Multisensor Data and Multiscale Feature Fusion Model for Intelligent Cross-Domain Machinery Fault Diagnosis, Ieee Transactions on Instrumentation and Measurement, 71 (2022). https://doi.org/10.1109/TIM.2022.3213016
36. Q.B. Wang, Y.B. Xu, S.K. Yang, J.T. Chang, J.G. Zhang, X.G. Kong, A domain adaptation method for bearing fault diagnosis using multiple incomplete source data, Journal of Intelligent Manufacturing, (2023). https://doi.org/10.1007/s10845-023-02075-7
37. B. Yang, Y.G. Lei, X. Li, C. Roberts, Deep Targeted Transfer Learning Along Designable Adaptation Trajectory for Fault Diagnosis Across Different Machines, Ieee Transactions on Industrial Electronics, 70 (2023) 9463-9473. https://doi.org/10.1109/TIE.2022.3212415
38. L. haixin, SQ dataset, github, 2022, pp. SQ dataset.

39. J. Lin, H.D. Shao, Z.S. Min, J.J. Luo, Y.M. Xiao, S. Yan, J. Zhou, Cross-domain fault diagnosis of bearing using improved semi-supervised meta-learning towards interference of out-of-distribution samples, Knowledge-Based Systems, 252 (2022). https://doi.org/10.1016/j.knosys.2022.109493
40. R. Wang, F.C. Yan, L. Yu, C.Q. Shen, X. Hu, J. Chen, A federated transfer learning method with low-quality knowledge filtering and dynamic model aggregation for rolling bearing fault diagnosis, Mechanical Systems and Signal Processing, 198 (2023). https://doi.org/10.1016/j.ymssp.2023.110413
41. Z.X. Zhou, H. Wang, Z.X. Li, W. Chen, Fault diagnosis of rolling bearing based on deep convolutional neural network and gated recurrent unit, Journal of Advanced Mechanical Design Systems and Manufacturing, 17 (2023). https://doi.org/10.1299/jamdsm.2023jamdsm0017
42. C. Lessmeier, KAt-DataCenter, KAt-DataCenter, KAt-DataCenter, pp. KAt-DataCenter.