SCIENCE CHINA
Technological Sciences

•Article•

October 2023 Vol.66 No.10: 2862–2870
https://doi.org/10.1007/s11431-023-2447-4

# Adaptive inter-intradomain alignment network with class-aware sampling strategy for rolling bearing fault diagnosis

GAO QinHe[1*], HUANG Tong[1], ZHAO Ke[2], SHAO HaiDong[3], JIN Bo[4], LIU ZhiHao[1] & WANG Dong[1]

[1] *National Key Discipline Laboratory of Armament Launch Theory & Technology, Rocket Force University of Engineering, Xi'an 710025, China;*
[2] *Key Laboratory of Road Construction Technology & Equipment Ministry of Education, Chang'an University, Xi'an 710054, China;*
[3] *College of Mechanical and Vehicle Engineering, Hunan University, Changsha 410082, China;*
[4] *Institute of Systems and Robotics, Department of Electrical and Computer Engineering, University of Coimbra, Coimbra 3030-290, Portugal*

Received April 3, 2023; accepted June 12, 2023; published online September 21, 2023

Existing unsupervised domain adaptation approaches primarily focus on reducing the data distribution gap between the source and target domains, often neglecting the influence of class information, leading to inaccurate alignment outcomes. Guided by this observation, this paper proposes an adaptive inter-intra-domain discrepancy method to quantify the intra-class and inter-class discrepancies between the source and target domains. Furthermore, an adaptive factor is introduced to dynamically assess their relative importance. Building upon the proposed adaptive inter-intradomain discrepancy approach, we develop an inter-intra-domain alignment network with a class-aware sampling strategy (IDAN-CSS) to distill the feature representations. The class-aware sampling strategy, integrated within IDAN-CSS, facilitates more efficient training. Through multiple transfer diagnosis cases, we comprehensively demonstrate the feasibility and effectiveness of the proposed IDAN-CSS model.

**unsupervised domain adaptation, inter-class domain discrepancy, intra-class domain discrepancy, class-aware sampling strategy**

**Citation:** Gao Q H, Huang T, Zhao K, et al. Adaptive inter-intradomain alignment network with class-aware sampling strategy for rolling bearing fault diagnosis. Sci China Tech Sci, 2023, 66: 2862–2870, https://doi.org/10.1007/s11431-023-2447-4

## 1 Introduction

Rotating machinery is widely used in aerospace, marine surveying, material mining, and other fields, including large rotor systems (generators, engines, internal combustion engines, and steam turbines), as well as small equipment (electric motors and pumps) [1,2]. In practical engineering, the structure of rotating machinery is extremely complicated, so long-term operation in a harsh environment could easily trigger a series of problems. Therefore, developing efficient and precise fault diagnosis technology is necessary for improving the operational safety and maintenance economy of rotating machinery. Therefore, it is crucial to perform a rolling bearing fault diagnosis [3–6].

Data-driven intelligent fault diagnosis methods are important technical tools that aim to use deep learning techniques to learn the manifestations of faults and the rules for identifying failure modes from large amounts of equipment monitoring data [7,8]. The powerful learning capability of deep learning enables the tools to automatically learn fault characteristics from more complex and diverse mechanical signals and build end-to-end intelligent diagnostic models [9,10]. However, deep learning intelligent diagnosis methods face the following problems in practical engineering applications. (1) Massive labeled fault samples are a necessity [11,12]: The complex network structure of deep learning

*Corresponding author (email: qhgao201@163.com)

models places high demands on the training samples, and most methods rely on a labeled supervised learning paradigm, which imposes more demanding requirements on the training samples. (2) Strict distribution consistency assumption for training and testing samples [13,14]: For deep learning intelligent diagnostic models to truly perform the diagnostic function, training data and test data must have an equal probability distribution. However, it is challenging to ensure this result because the previous training samples in the case of variable test environments, fluctuating or changing mechanical equipment operating conditions that directly lead to fault detection. This gap will immediately cause the previously trained intelligent diagnostic model to perform worse or possibly fail completely, which will lead to a significant negative impact on the actual application effect.

A viable solution to the above problems is offered by the transfer learning hypothesis [15]. Transfer learning is a unique learning technique that uses familiarity with previous activities to aid learning of the current activity, which can effectively alleviate the domain shift phenomenon caused by differences in data distribution, thus broadening the application of data-driven intelligence-based models, and has received increased attention [16–18]. The three main types of transfer-intelligent diagnostic techniques are instance transfer-based techniques [19], model transfer-based techniques [20], and feature transfer-based techniques [21]. Unlike instance and model transfer techniques, feature transfer techniques consider comparable and shared feature space information [22]. In rolling bearing fault diagnosis, feature transfer-based techniques are often used because they can solve the problem of unpredictability. Lee's group [23] designed an MMD strategy for a convolutional neural network and used the transfer model to accurately identify the unlabeled data. Zhu et al. [24] conducted an MMD regulation term for a deep neural network loss and distilled the invariant representations of two domains. Lu et al. [25] developed a feature alignment approach based on MMD and achieved the cross-domain fault diagnosis. Li et al. [26] proposed a deep generative network and used sufficient cross-domain experiments to validate it. To diagnose bearing problems with variable condition transfer, Wu et al. [27] suggested a domain adaptation framework based on deep adaptation networks and lightweight models. A dynamic joint distribution alignment network was suggested by Shen et al. [28] to successfully identify bearing problems with varying operating circumstances. Zhao et al. [29] developed a transfer approach based on adversarial learning, and the proposed method was validated by many bearing transfer diagnosis cases with an average accuracy of 99.24%. Qin et al. [30] created a novel joint method that simultaneously aligns the marginal and conditional distributions.

However, the above approaches can only overall align the source and target domain distributions without considering the corresponding classes, which inevitably affects the identification results. When the source domain samples are from various classes, and the target domain samples are not aligned, for instance, MMD and JMMD can still be reduced, and this may result in poor generalization of the learned decision boundaries to the target domain. Many suboptimal solutions exist near the decision boundary, which may have high recognition results for source domain data but poor recognition for the target domain. Therefore, to promote the domain-adaptive effect, the distribution alignment process should consider two measures. The first is transferability, which involves reducing the variation in distribution within a single category across many domains. The other is discriminability, or maximizing the distributional differences between different categories in different domains. On this basis, an adaptive inter-domain discrepancy method is designed to evaluate the intra-class and inter-class differences, and an adaptive factor is used to quantify the proportion between the two differences.

Therefore, an inter-intra-domain alignment network with a class-aware sampling strategy (IDAN-CSS) is constructed based on the proposed adaptive inter-intra-domain discrepancy method. The developed class-aware sampling strategy can help IDAN-CSS obtain a more efficient training way. The comparison is shown in Figure 1. Then, an IDAN-CSS is constructed to simultaneously minimize the inter-class domain differences and maximize the intra-class do-

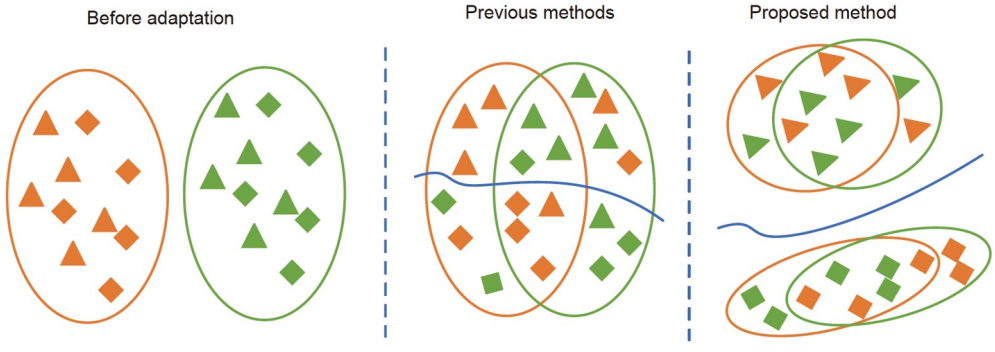


**Figure 1** (Color online) Comparison between the proposed and previous feature alignment methods.

main differences. Consider that during model training, for category A, the randomly selected small batch of samples may only come from a single domain, and the inter-class domain differences cannot be measured, affecting the final domain adaptation effect. Therefore, to increase training efficacy and ensure adaptive outputs, we collect data from both domains for each class within a randomly selected subset of classes in each iteration. The innovations and contributions of the suggested framework can be concluded as follows.

(1) The proposed inter-intra-domain discrepancy method can measure the inter-class and intra-class domain discrepancies between domains.

(2) The designed adaptive factor can quantify the proportion between the two discrepancies to achieve better domain adaption performance.

(3) The constructed inter-intra-domain alignment network with a class-aware sampling strategy can simultaneously increase the transferability and discriminability for better domain adaption performance.

(4) The class-aware sampling strategy can efficiently tackle the small batch training problem in the training process.

## 2 Primaries

### 2.1 Unsupervised domain adaptation

In an unsupervised domain adaptation, the labeled source data and the unlabeled target data are defined in eqs. (1) and (2), respectively.

$$D_s = \left\{(x_i^s, y_i^s)\right\}_{i=1}^{n_s}, \tag{1}$$

$$D_t = \left\{(x_j^t)\right\}_{j=1}^{n_t}, \tag{2}$$

where $D_s$ and $D_t$ denote the source and target domains, respectively. $x_i^s$ is the $i$th sample in the source domain, $x_j^t$ is the $i$th sample in the target domain, and $y_i^s$ means the corresponding label of the $i$th sample. $n_s$ and $n_t$ represent the total number of samples in the source and target domains, respectively. The source and target domains share the same label space $\{1, 2, \cdots, C\}$, and $C$ is the total number of failure modes, but with different distributions. Moreover, the current work trains a model with the labeled source data to accurately predict the unlabeled target data.

### 2.2 Maximum mean discrepancy

MMD is a crucial indicator of the distributional difference between common inter-class samples between the source and target domains and is frequently used in domain adaptation techniques. For marginal probability distributions $Q_s(X_s)$ and $Q_t(X_t)$, the MMD distance between them is expressed as

$$\mathrm{MMD}(Q_s(X_s), Q_t(X_t)) = \left\| \frac{1}{n_s}\sum_{i=1}^{n_s}\varphi(x_s^i) - \frac{1}{n_t}\sum_{j=1}^{n_t}\varphi(x_t^j) \right\|_{\mathbf{H}}, \tag{3}$$

where $\varphi$ stands for the feature map $H-\mathbf{H}$, and $\mathbf{H}$ means the Reproducing Kernel Hilbert Space.

For layer l, MMD adopting the selected kernel is calculated as

$$\begin{aligned}
&\mathrm{MMD}_l^2(Q_s(X_s), Q_t(X_t)) \\
&= \left\| \frac{1}{n_s}\sum_{i=1}^{n_s}\varphi(x_s^i) - \frac{1}{n_t}\sum_{j=1}^{n_t}\varphi(x_t^i) \right\|_{\mathbf{H}}^2 \\
&= \frac{1}{n_s n_s}\sum_{i=1}^{n_s}\sum_{j=1}^{n_s} k_l\left(x_s^i, x_s^j\right) + \frac{1}{n_t n_t}\sum_{i=1}^{n_t}\sum_{j=1}^{n_t} k_l\left(x_t^i, x_t^j\right) \\
&\quad - 2\frac{1}{n_s n_t}\sum_{i=1}^{n_s}\sum_{j=1}^{n_t} k_l\left(x_s^i, x_t^j\right),
\end{aligned} \tag{4}$$

where $x_s \in D_s^m \subset D_s$, $x_t \in D_t^m \subset D_t$, $n_s = \left|D_s^m\right|$, and $n_t = \left|D_t^m\right|$. $D_s^m$ and $D_t^m$ denote the mini-batch samples from $D_s$ and $D_t$ respectively, and $k_l$ is the kernel function corresponding to the network layer $l$.

## 3 Methodology

### 3.1 Inter-intra domain discrepancy

Compared with other distance measurement methods, the proposed inter-intra-domain discrepancy method can measure the inter-class and intra-class domain discrepancies between domains. The transferability of domain adaptation can be improved by minimizing the distributional difference of the same class across domains, whereas the discriminability of domain adaptation can be improved by increasing the distributional difference of various classes across domains. The domain adaptation effect can be improved by optimizing both the intra-class and inter-class domain discrepancies. The proposed inter-intra-domain discrepancy approach uses MMD to compute the discrepancy between the conditional probability distributions of various domains, considering how convenient it is to compute the distribution discrepancy, $Q_s(X_s|Y_s)$ and $Q_t(X_t|Y_t)$.

Assuming $\mu_{cc'}(y, y') = \begin{cases} 1, & \text{if } y = c,\ y' = c' \\ 0, & \text{otherwise} \end{cases}$, for classes $c_1$ and $c_2$ (which can be identical or different), $\mathrm{MMD}^2(Q_s, Q_t)$ can be denoted as

$$\mathrm{MMD}_{c_1c_2}^2(\hat{y}_1^t, \hat{y}_2^t, \cdots, \hat{y}_{n_t}^t) = a_1 + a_2 - 2a_3, \tag{5}$$

$$a_1=\sum_{i=1}^{n_s}\sum_{j=1}^{n_s}\frac{\mu_{c_1c_1}\left(y_i^s,y_j^s\right)k\left(x_i^s,x_j^s\right)}{\sum_{i=1}^{n_s}\sum_{j=1}^{n_s}\mu_{c_1c_1}\left(y_i^s,y_j^s\right)}, \tag{6}$$

$$a_2=\sum_{i=1}^{n_t}\sum_{j=1}^{n_t}\frac{\mu_{c_2c_2}\left(\hat{y}_i^t,\hat{y}_j^t\right)k\left(x_i^t,x_j^t\right)}{\sum_{i=1}^{n_t}\sum_{j=1}^{n_t}\mu_{c_2c_2}\left(\hat{y}_i^t,\hat{y}_j^t\right)}, \tag{7}$$

$$a_3=\sum_{i=1}^{n_s}\sum_{j=1}^{n_t}\frac{\mu_{c_1c_2}\left(y_i^s,\hat{y}_j^t\right)k\left(x_i^s,x_j^t\right)}{\sum_{i=1}^{n_s}\sum_{j=1}^{n_t}\mu_{c_1c_2}\left(y_i^s,\hat{y}_j^t\right)}. \tag{8}$$

Note that when $c_1=c_2$, eq. (5) means the inter-class domain discrepancy, and when $c_1\neq c_2$, eq. (5) means the intra-class domain discrepancy. In addition, the calculation of $\mu_{c_2c_2}\left(\hat{y}_i^t,\hat{y}_j^t\right)$ and $\mu_{c_1c_2}\left(y_i^s,\hat{y}_j^t\right)$ requires the use of the target domain labels, so the pseudo-labels need to be obtained using the classifier.

### 3.2 Adaptive factor

Considering the dynamic change in the importance of transferability and discriminability during domain adaptation, the adaptive factor is designed to quantify the proportion between the two differences at each iteration. As the intra-class domain discrepancy is optimized in the direction opposite to the inter-class domain discrepancy, the inter-intra-domain discrepancy is expressed as

$$\begin{aligned}\mathcal{L}_{\mathrm{IDD}}&=\mu\mathcal{L}_{\mathrm{intra}}-(1-\mu)\mathcal{L}_{\mathrm{inter}}\\&=\frac{\mu}{C}\sum_{c=1}^{C}\mathrm{MMD}_{cc}^2\left(\hat{y}_1^t,\hat{y}_2^t,\cdots,\hat{y}_{n_t}^t\right)\\&\quad-\frac{(1-\mu)}{C(C-1)}\sum_{c=1}^{C}\sum_{\substack{c'=1\\c'\neq c}}^{C}\mathrm{MMD}_{cc'}^2\left(\hat{y}_1^t,\hat{y}_2^t,\cdots,\hat{y}_{n_t}^t\right).\end{aligned} \tag{9}$$

Note that the issue of unidentified target domain labels is one that this study addresses, so it is difficult to calculate $\mu$ directly. We adopt an alternative idea of approximating $\mu$ with the help of global and local domains. Specifically, A-distance is introduced as the basic measure, where A-distance stands for the mistake made when creating a classifier to differentiate between the two domains. Using $\varepsilon(f)$ denotes the error in distinguishing the domains $D_s$ and $D_t$ of linear classifier $f$. A-distance is then defined as

$$d_A(D_s,D_t)=2(1-2\varepsilon(f)). \tag{10}$$

The A-distance of the global domain, expressed as $d_G$, can be calculated directly using the above formula. For the A-distance of a local domain, using class $c$ as an example, it can be expressed as $d_c=d_A\left(D_s^{(c)},D_t^{(c)}\right)$, where $D_s^{(c)}$ and $D_t^{(c)}$ represent samples of class $c$ in $D_s$ and $D_t$, respectively. Eventually, we can obtain the value of $\mu$ using the following formula:

$$\mu\approx 1-\frac{d_G}{d_G+\sum_{c=1}^{C}d_c}. \tag{11}$$

This adaptive factor is carried out dynamically during the iterative process because domain adaptation affects the relative importance of transferability and discriminability. This dynamic way of measuring the importance of each distribution can undoubtedly achieve more accurate domain adaption effects and provide new ideas for future domain adaption research.

### 3.3 Inter-intra-domain alignment network

Compared with other deep learning models, convolutional neural networks can extract more representative features from bearing vibration signals and are widely used in domain adaption models. Therefore, in this paper, the inter-intra-domain alignment network is constructed using a convolutional neural network (Table 1). The schematic diagram of our framework is presented in Figure 2. As the deep network extracts the key fault characteristics of the vibration signal layer by layer, this paper chooses to add an intra-class domain discrepancy to the last two fully connected layers. This addition can ensure that the network can be extracted effectively at a small computation cost. The domain-invariant features and domain-discriminative features are expressed as

$$\mathcal{L}_{\mathrm{IDD}}=\sum_{l=1}^{L}\mathcal{L}_{\mathrm{IDD}}^{l}. \tag{12}$$

In addition, the softmax cross-entropy loss function is used during network training to obtain accurate source domain prediction based on label information, denoted as

$$\mathcal{L}_C=-\frac{1}{M}\left[\sum_{m=1}^{M}\sum_{c=1}^{C}I\left\{y_m^s=c\right\}\log\left(P_c\left(x_m^s\right)\right)\right], \tag{13}$$

where $P_c(\cdot)$ represents the probability output for the $c$th class and $M$ means the number of training samples.

**Table 1** Structural parameters of IDAN-CSS

| Layer | Parameters | Output size |
|---|---|---|
| Input | – | (1, 784) |
| Conv1 | 20, 5×1 | (20, 780) |
| Pool1 | 20, 2×2 | (20, 390) |
| Conv2 | 20, 5×1 | (20, 386) |
| Pool2 | 20, 2×2 | (20, 193) |
| Conv3 | 40, 5×1 | (40, 189) |
| Pool3 | 40, 2×2 | (40, 94) |
| FC0 | ReLU | (1, 3760) |
| FC1 | ReLU | (1, 1000) |
| FC2 | ReLU | (1, 100) |
| Output | Softmax | (1, $C$) |

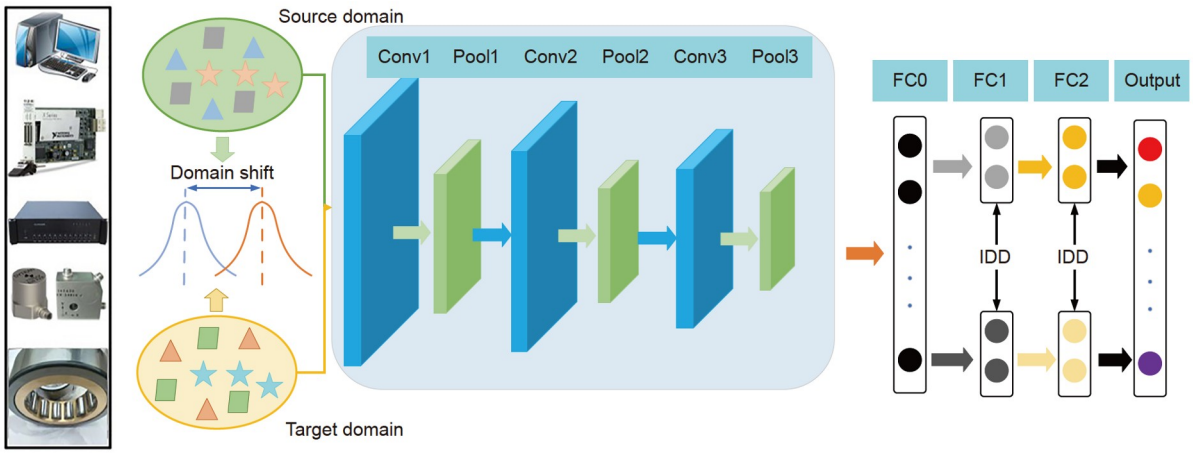


**Figure 2** (Color online) Proposed IDAN-CSS model for the transfer fault diagnosis.

Eventually, the final loss function of IDAN-CSS is denoted as

$$\mathcal{L} = \mathcal{L}_C + \lambda \mathcal{L}_{\text{IDD}} = \mathcal{L}_C + \lambda\mu\mathcal{L}_{\text{intra}} - \lambda(1-\mu)\mathcal{L}_{\text{inter}}, \tag{14}$$

where $\lambda$ is the balancing factor, which is calculated by $2/(1+2\times e^{(-10q)})-1$, where $q$ is from 0 to 1, as the training proceeds.

The conventional training approach for deep models samples a small batch of samples randomly at each iteration without following a specific class. If the conventional training method is chosen, the randomly selected samples may contain only a single domain. This option results in the inability to calculate the inter-class domain discrepancy and affects the subsequent calculation. If the conventional training method is applied directly, it inevitably affects the IDAN-CSS training efficiency and jeopardizes the final domain adaption effect. Therefore, a class-aware sampling strategy is developed to train IDAN-CSS to facilitate the computation of the inter-intra-domain discrepancy during domain adaptation. Specifically, we randomly select a subset of classes $T_e$ from the source domain target domain samples, respectively, and then sample source data and target data for each class in $T_e$. This option ensures that the randomly selected small batches of data for each training round can satisfy the requirement of inter-class-domain discrepancy calculation.

## 4 Experimental verification

### 4.1 Dataset illustration

SDUSTD [31]: The bearing failure locations are set as the inner ring, outer ring, and rolling element; the failure types are all pitting faults, and the failure levels include minor and serious, with a total of six failure conditions and one normal condition. The bearing fault test bench is shown in Figure 3. The sampling frequency is set to 25.6 kHz, and the load is set to 0 of 20, 40, and 60 N. Therefore, the vibration signals of seven conditions are measured, and three datasets are obtained. The raw vibration signals are transformed into frequency domain signals and then divided into 1400 samples with a sample dimension of 784. Finally, six cross-domain fault diagnosis cases are implemented based on A1, A2, and A3, and six sets of cross-domain diagnostic tasks are finally obtained.

PUD [32]: A specially designed bearing fault test rig is used to measure vibration signals across rotational speeds. Five bearing conditions are designed in this experiment. The bearing fault test bench and gear type adopted in this experiment are shown in Figure 4. The sampling frequency is set to 64 kHz. Four different operating conditions are set up for data collection, B1 (1500 r/min, 0.1 N m, and 1000 N), B2 (1500 r/min, 0.7 N m, and 400 N), and B3 (900 r/min, 0.7 N m, and 1000 N). The raw vibration signals are transformed into frequency domain signals and then divided into 1000 samples with a sample dimension of 784. Finally, six sets of cross-domain diagnostic tasks are finally obtained.

### 4.2 Result analysis of SDUSTD

To efficiently evaluate the diagnostic performance of IDAN-AUS, some approaches are used for comparison in the following part, including CNN, DANN [33], DAN [34], JAN [35], and the proposed method with no class-aware sampling strategy (IDAN). CNN is the conventional deep learning method without adding constraint terms, which can visualize the difficulty of cross-domain tasks. DANN, DCORAL, DAN, and JAN are the current mainstream domain adaption algorithms, which can fully illustrate the diagnostic performance. The diagnostic results are the average value of five runs, which guarantees the credibility of the experimental results.

The diagnostic accuracy of all methods on six cross-domain fault diagnosis cases is illustrated in Table 2, and the

concrete results are presented in Figure 5. Table 2 helps infer that the average accuracy of IDAN-CSS, CNN, DANN, DAN, JAN, and IDAN for the six cross-domain fault diagnosis cases is 98.89%, 83.12%, 92.28%, 94.70%, 95.60%, and 98.61%, respectively. The arguments can be obtained as (1) The diagnostic performance of IDAN-CSS significantly outperforms that of CNN, DANN, DAN, JAN, and IDAN, and the results show that the feature alignment capability of the proposed approach outperforms that of the mainstream domain adaption methods. (2) The poor diagnostic accuracy and robustness of CNN imply that the conventional deep approaches cannot deal with cross-domain fault diagnosis problems with various distributions. (3) The diagnostic robustness of IDAN-CSS is better than the other five methods, which further demonstrates the stability of IDAN-CSS. (4) The comparison between IDAN-CSS and IDAN further

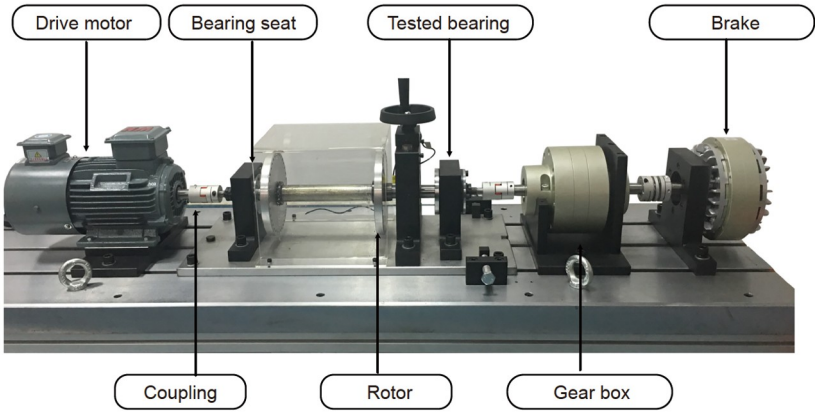


**Figure 3** (Color online) Experimental setup of SDUSTD [31].

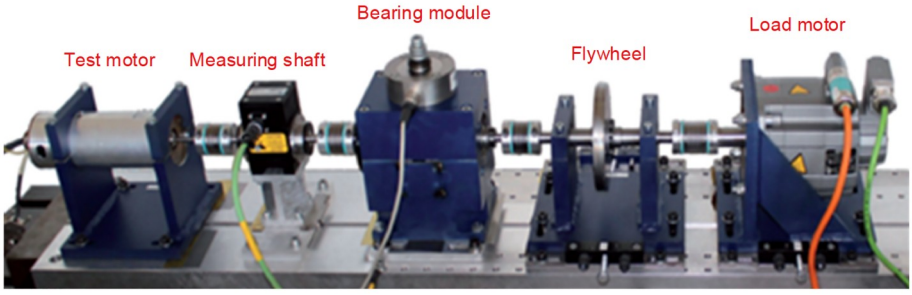


**Figure 4** (Color online) Test platform of PUD [32].

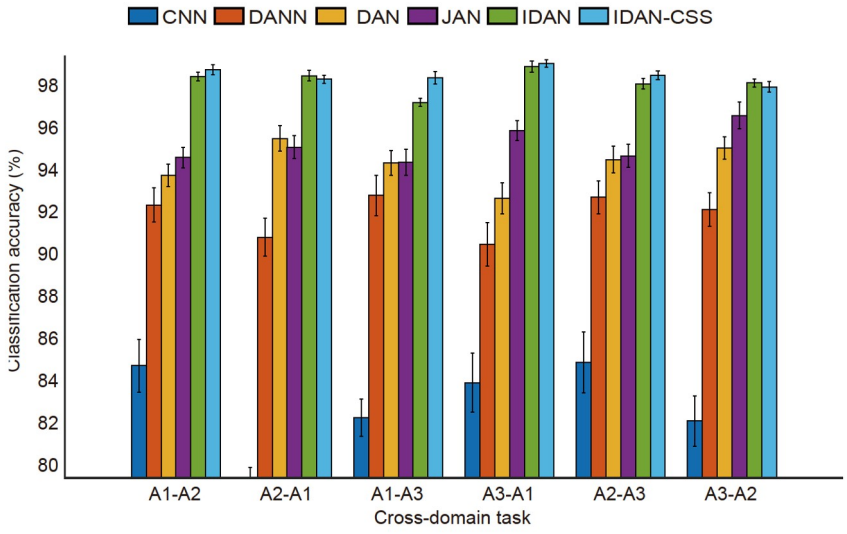


**Figure 5** (Color online) Fault diagnosis results of all approaches.

demonstrates the effectiveness of using a class-aware sampling strategy.

To further illustrate the feature adaptation performance of our approach, t-SNE is adopted to directly visualize the feature alignment results. For the transfer task A1–A2, the 2D and 3D visualization results are displayed in Figure 6. From the visualization results, the source and target domain features learned by the proposed method can be easily distinguished and overlap well. This is because the proposed method not only considers the marginal and conditional distribution adaptation but also simultaneously increases the transferability and discriminability for excellent domain adaption performance. However, it is also clear from Figure 6 that the features of a small number of samples partially overlap. The reason for this observation is the influence of various uncertainties in the actual operation process, which results in too significant differences in the distribution characteristics of the source and target domains.

### 4.3 Result analysis of PUD

To further verify the effectiveness of the proposed method, six sets of cross-domain fault diagnosis tasks were implemented using the PUD dataset. The diagnosis results of all methods are shown in Table 3 and Figure 7.

The average diagnostic accuracy of IDAN-CSS is 98.37% under six fault diagnosis tasks, which is higher than that of the other five methods. IDAN-CSS achieves optimal diagnostic accuracy under transfer tasks B1–B2, B1–B3, and B3–B2. The overall diagnostic accuracy is much better than that of DANN, DAN, and JAN and is slightly better than that of IDAN. From the two bearing diagnosis experiments, the conclusions that can be drawn include: (1) The diagnostic accuracy and diagnostic robustness of IDAN-CSS are significantly better than those of other domain adaption approaches, which means the proposed approach has a superior feature alignment capability. (2) For cross-domain fault diagnosis problems, transfer learning has obvious advantages over deep learning.

For the transfer task B1–B2, the 2D and 3D visualization results are displayed in Figure 8. From the visualization results, the source and target domain features learned by the proposed method can be easily distinguished and overlap well. This is because the proposed method not only considers the marginal and conditional distribution adaptations but also simultaneously increases the transferability and the discriminability for excellent domain adaption performance. The visualization results fully verify the efficacy of our feature alignment method.

## 5 Conclusions

In this study, an inter-intra-domain alignment network with a

**Table 2** Diagnostic results of all methods

| Transfer task | CNN | DANN | DAN | JAN | IDAN | IDAN-CSS |
|---|---|---|---|---|---|---|
| A1–A2 | 85.13±1.26 | 92.75±0.82 | 94.16±0.53 | 95.00±0.48 | 98.84±0.21 | 99.16±0.24 |
| A2–A1 | 78.80±1.51 | 91.22±0.89 | 95.90±0.60 | 95.49±0.55 | 98.88±0.24 | 98.71±0.20 |
| A1–A3 | 82.67±0.88 | 93.20±0.95 | 94.73±0.59 | 94.76±0.62 | 97.61±0.20 | 98.77±0.29 |
| A3–A1 | 84.32±1.40 | 90.88±1.04 | 93.07±0.74 | 96.27±0.47 | 99.30±0.26 | 99.46±0.18 |
| A2–A3 | 85.28±1.44 | 93.11±0.77 | 94.90±0.63 | 95.08±0.55 | 98.49±0.25 | 98.90±0.20 |
| A3–A2 | 82.51±1.19 | 92.54±0.80 | 95.44±0.53 | 96.99±0.63 | 98.53±0.19 | 98.34±0.25 |
| Average | 83.12 | 92.28 | 94.70 | 95.60 | 98.61 | 98.89 |

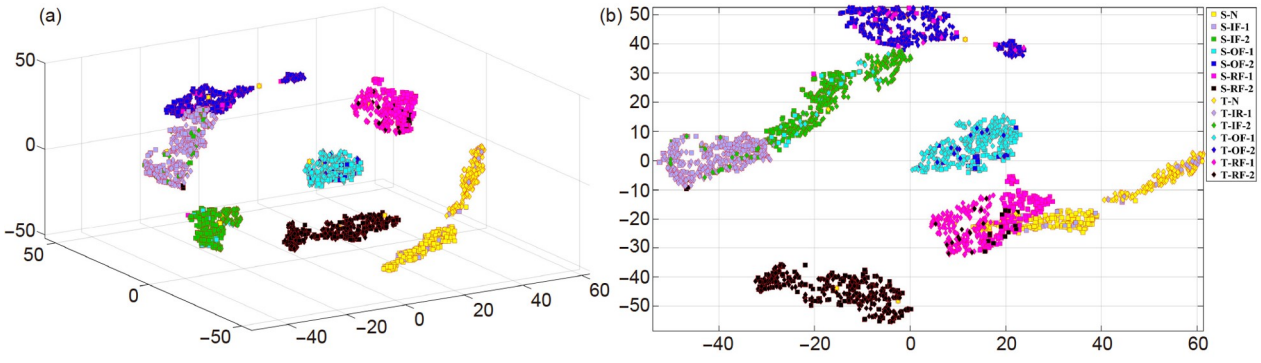


**Figure 6** Visualization results of the source and target domains on A1–A2. (a) 3D visualization results; (b) 2D visualization results.

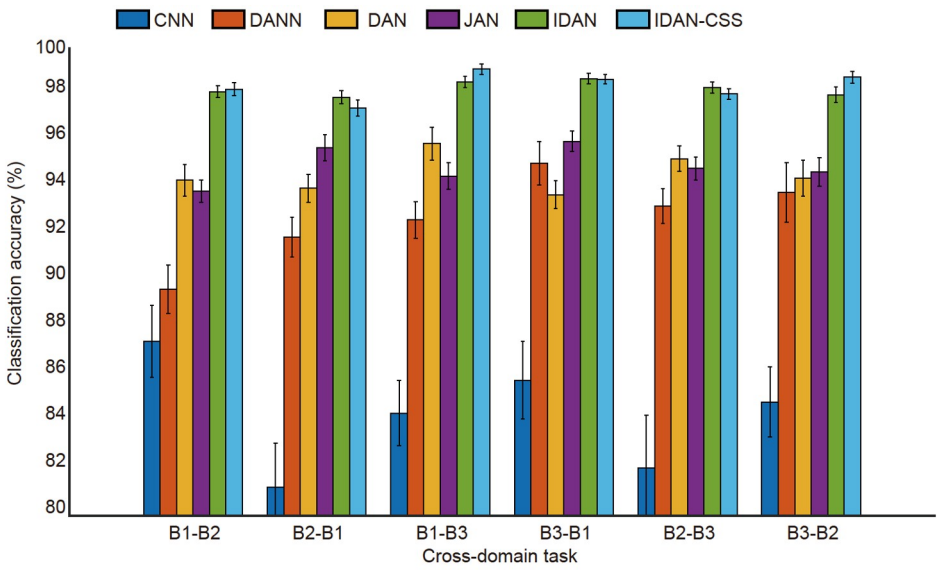


**Figure 7** Fault diagnosis results of all approaches.

**Table 3** Diagnostic results of all methods

| Transfer task | CNN | DANN | DAN | JAN | IDAN | IDAN-CSS |
|---|---|---|---|---|---|---|
| B1–B2 | 87.46±1.54 | 89.69±1.03 | 94.35±0.67 | 93.88±0.49 | 98.13±0.25 | 98.24±0.28 |
| B2–B1 | 81.22±1.89 | 91.92±0.84 | 94.00±0.59 | 95.74±0.55 | 97.89±0.28 | 97.43±0.35 |
| B1–B3 | 84.39±1.40 | 92.65±0.79 | 95.92±0.70 | 94.53±0.58 | 98.55±0.25 | 99.10±0.22 |
| B3–B1 | 85.80±1.66 | 95.07±0.92 | 93.73±0.60 | 96.01±0.44 | 98.69±0.22 | 98.66±0.20 |
| B2–B3 | 82.05±2.24 | 93.24±0.75 | 95.27±0.55 | 94.85±0.50 | 98.31±0.24 | 98.04±0.22 |
| B3–B2 | 84.87±1.51 | 93.83±1.28 | 94.44±0.77 | 94.69±0.61 | 98.00±0.33 | 98.75±0.25 |
| Average | 84.30 | 92.73 | 94.62 | 94.95 | 98.26 | 98.37 |

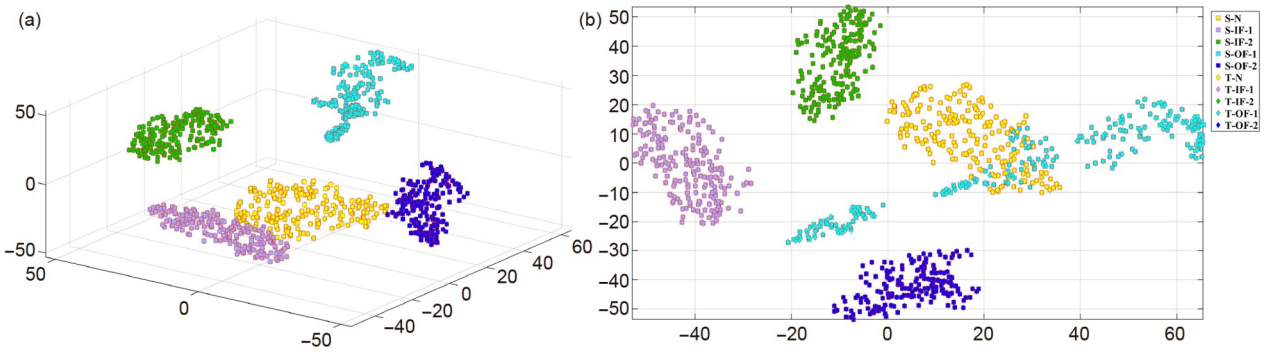


**Figure 8** Visualization results of the source and target domains on B1–B2. (a) 3D visualization results; (b) 2D visualization results.

class-aware sampling strategy is developed to solve the cross-domain fault diagnosis problem of rotating machinery. The proposed adaptive inter-intra-domain discrepancy method can dynamically evaluate the importance between inter-class domain discrepancy and intra-class domain discrepancy. Extensive experiments prove that IDAN-CSS significantly outperforms other fault diagnosis methods. Therefore, IDAN-CSS is a promising solution to the unsupervised cross-domain fault diagnosis problem.

However, IDAN-CSS mainly contributes to cross-domain fault diagnosis problems with the same class space, which limits its wide application in practical engineering. In the future, we will conduct in-depth research to expand the application scenarios of IDAN-CSS.

*This work was supported by the National Natural Science Foundation of China (Grant Nos. 52275104, 51905160) and the Natural Science Fund for Excellent Young Scholars of Hunan Province (Grant No. 2021JJ20017).*


1 Shao H, Li W, Cai B, et al. Dual-threshold attention-guided GAN and limited infrared thermal images for rotating machinery fault diagnosis under speed fluctuation. IEEE Trans Ind Inf, 2023, 19: 9933–9942
2 Zhao B, Zhang X, Wu Q, et al. A novel unsupervised directed hierarchical graph network with clustering representation for intelligent fault diagnosis of machines. Mech Syst Signal Processing, 2023, 183: 109615
3 Zhang Z, Wang J, Li S, et al. Fast nonlinear blind deconvolution for rotating machinery fault diagnosis. Mech Syst Signal Processing, 2023, 187: 109918
4 Di Z Y, Shao H D, Xiang J W. Ensemble deep transfer learning driven by multisensor signals for the fault diagnosis of bevel-gear cross-operation conditions. Sci China Tech Sci, 2021, 64: 481–492
5 Wang J, Zhang Z, Liu Z, et al. Digital twin aided adversarial transfer learning method for domain adaptation fault diagnosis. Reliability Eng Syst Saf, 2023, 234: 109152
6 Zhao K, Hu J, Shao H, et al. Federated multi-source domain adversarial adaptation framework for machinery fault diagnosis with data privacy. Reliability Eng Syst Saf, 2023, 236: 109246
7 Hou W, Zhang C, Jiang Y, et al. A new bearing fault diagnosis method via simulation data driving transfer learning without target fault data. Measurement, 2023, 215: 112879
8 Zhang W, Li X, Ma H, et al. Open-set domain adaptation in machinery fault diagnostics using instance-level weighted adversarial learning. IEEE Trans Ind Inf, 2021, 17: 7445–7455
9 Zhang W, Li X, Li X. Deep learning-based prognostic approach for lithium-ion batteries with adaptive time-series prediction and on-line validation. Measurement, 2020, 164: 108052
10 Yu S, Wang M, Pang S, et al. TDMSAE: A transferable decoupling multi-scale autoencoder for mechanical fault diagnosis. Mech Syst Signal Processing, 2023, 185: 109789
11 Wang X, Shen C, Xia M, et al. Multi-scale deep intra-class transfer learning for bearing fault diagnosis. Reliability Eng Syst Saf, 2020, 202: 107050
12 Sun M, Wang H, Liu P, et al. Stack autoencoder transfer learning algorithm for bearing fault diagnosis based on class separation and domain fusion. IEEE Trans Ind Elec, 2022, 69: 3047–3058
13 Zhao K, Jia F, Shao H. A novel conditional weighting transfer Wasserstein auto-encoder for rolling bearing fault diagnosis with multi-source domains. Knowledge-Based Syst, 2023, 262: 110203
14 Zhu J, Huang C, Shen C, et al. Cross-domain open-set machinery fault diagnosis based on adversarial network with multiple auxiliary classifiers. IEEE Trans Ind Inf, 2022, 18: 8077–8086
15 Zhu Z, Lei Y, Qi G, et al. A review of the application of deep learning in intelligent fault diagnosis of rotating machinery. Measurement, 2023, 206: 112346
16 Qian Q, Qin Y, Luo J, et al. Deep discriminative transfer learning network for cross-machine fault diagnosis. Mech Syst Signal Processing, 2023, 186: 109884
17 Zhang W, Wang Z, Li X. Blockchain-based decentralized federated transfer learning methodology for collaborative machinery fault diagnosis. Reliability Eng Syst Saf, 2023, 229: 108885
18 Shi Y, Deng A, Deng M, et al. Transferable adaptive channel attention module for unsupervised cross-domain fault diagnosis. Reliability Eng Syst Saf, 2022, 226: 108684
19 Li W, Huang R, Li J, et al. A perspective survey on deep transfer learning for fault diagnosis in industrial scenarios: Theories, applications and challenges. Mech Syst Signal Processing, 2022, 167: 108487
20 Wu Z, Jiang H, Zhao K, et al. An adaptive deep transfer learning method for bearing fault diagnosis. Measurement, 2020, 151: 107227
21 Wang P, Gao R X. Transfer learning for enhanced machine fault diagnosis in manufacturing. CIRP Ann, 2020, 69: 413–416
22 Wu Z, Zhang H, Guo J, et al. Imbalanced bearing fault diagnosis under variant working conditions using cost-sensitive deep domain adaptation network. Expert Syst Appl, 2022, 193: 116459
23 Azamfar M, Li X, Lee J. Intelligent ball screw fault diagnosis using a deep domain adaptation methodology. Mechanism Machine Theor, 2020, 151: 103932
24 Zhu J, Chen N, Shen C. A new deep transfer learning method for bearing fault diagnosis under different working conditions. IEEE Sens J, 2020, 20: 8394–8402
25 Lu W, Liang B, Cheng Y, et al. Deep model based domain adaptation for fault diagnosis. IEEE Trans Ind Electron, 2017, 64: 2296–2305
26 Li X, Zhang W, Ding Q. Cross-domain fault diagnosis of rolling element bearings using deep generative neural networks. IEEE Trans Ind Electron, 2019, 66: 5525–5534
27 Wu J, Tang T, Chen M, et al. A study on adaptation lightweight architecture based deep learning models for bearing fault diagnosis under varying working conditions. Expert Syst Appl, 2020, 160: 113710
28 Shen C, Wang X, Wang D, et al. Dynamic joint distribution alignment network for bearing fault diagnosis under variable working conditions. IEEE Trans Instrum Measure, 2021, 70: 3510813
29 Zhao K, Jiang H, Wang K, et al. Joint distribution adaptation network with adversarial learning for rolling bearing fault diagnosis. Knowledge-Based Syst, 2021, 222: 106974
30 Qin Y, Qian Q, Luo J, et al. Deep joint distribution alignment: A novel enhanced-domain adaptation mechanism for fault transfer diagnosis. IEEE Trans Cybern, 2023, 53: 3128–3138
31 Jia M, Wang J, Zhang Z, et al. A novel method for diagnosing bearing transfer faults based on a maximum mean discrepancies guided domain-adversarial mechanism. Meas Sci Technol, 2022, 33: 015109
32 Lessmeier C, Kimotho J, Zimmer D, et al. Condition monitoring of bearing damage in electromechanical drive systems by using motor current signals of electric motors: A benchmark data set for data-driven classification. In: European Conference of the Prognostics and Health Management Society. Bilbao, 2016
33 Jiao J, Zhao M, Lin J. Unsupervised adversarial adaptation network for intelligent fault diagnosis. IEEE Trans Ind Electron, 2020, 67: 9904–9913
34 Long M, Cao Y, Cao Z, et al. Transferable representation learning with deep adaptation networks. IEEE Trans Pattern Anal Mach Intell, 2019, 41: 3071–3085
35 Wang P, Lu L, Li J, et al. Transfer learning with joint distribution adaptation and maximum margin criterion. J Phys: Conf Series, 2019, 1169: 230–237