Check for updates

Non-Uniformly Weighted Multisource Domain Adaptation Network For Fault Diagnosis Under Varying Working Conditions

Hongliang Zhang[1] · Yuteng Zhang[1] · Rui Wang[2] · Haiyang Pan[3] · Bin Chen[1]

Abstract

Most transfer learning-based fault diagnosis methods learn diagnostic information from the source domain to enhance performance in the target domain. However, in practical applications, usually there are multiple available source domains, and relying on diagnostic information from only a single source domain limits the transfer performance. To this end, a non-uniformly weighted multisource domain adaptation network is proposed to address the above challenge. In the proposed method, an intra-domain distribution alignment strategy is designed to eliminate multi-domain shifts and align each pair of source and target domains. Furthermore, a non-uniform weighting scheme is proposed for measuring the importance of different sources based on the similarity between the source and target domains. On this basis, a weighted multisource domain adversarial framework is designed to enhance multisource domain adaptation performance. Numerous experimental results on three datasets validate the effectiveness and superiority of the proposed method.

Keywords Rotating machinery · Fault diagnosis · Multisource domain adaptation · Non-uniform weight · Transfer learning

✉ Bin Chen
  binchenahut@163.com

  Hongliang Zhang
  hlzhang@ahut.edu.cn

  Yuteng Zhang
  ytzhangz@ahut.edu.cn

  Rui Wang
  wrfight96@163.com

  Haiyang Pan
  pansea@ahut.edu.cn

1 School of Management Science and Engineering, Anhui University of Technology, Maanshan 243032, China

2 School of Rail Transportation, Soochow University, Suzhou 215131, China

3 School of Mechanical Engineering, Anhui University of Technology, Maanshan 243032, China

🖄 Springer

1 Introduction

The rapid development of the manufacturing industry is closely related to various machinery equipment. Modern mechanical equipment is increasingly automated and intelligent, which has high requirements for the reliability and stability of equipment components [1]. Bearings are used to support the mechanical rotating body and play a crucial role in the operation of the equipment. Therefore, accurately monitoring the condition of bearings can ensure the regular operation of equipment and avoid economic losses and safety risks caused by equipment damage [2–4].

In recent years, with the rise of artificial intelligence such as machine learning, data-driven intelligent fault diagnosis (IFD) methods have attracted considerable attention [5]. Among them, IFD methods based on deep learning (DL) are widely used due to their good adaptability and learning ability [6]. DL-based IFD methods usually hold the assumption that a large amount of labeled and equally distributed data are available [7, 8]. However, the above assumptions do not always hold in complex practical scenarios, resulting in serious degradation of diagnostic performance [9, 10].

To address the above challenges, an important tool of the transfer learning technique, domain adaptation (DA), is introduced to IFD [11, 12]. DA facilitates the utilization of transferable diagnostic information learned from the source domain in the target domain through domain confusion, thereby effectively enhancing diagnostic accuracy in the target domain [13]. A common DA approach is realized by statistical matching [14, 15]. For instance, Schwendemann et al. [16] introduced a layered maximum mean discrepancy and integrated it as a loss function within the deep neural network, effectively mitigating the distribution discrepancy between the source and target domains. Similarly, Xiong et al. [17] designed a central moment discrepancy to minimize the distribution discrepancy between domains for cross-domain fault diagnosis. Another common DA approach is domain adversarial training [18, 19]. Ganin et al. [20] introduced a domain adversarial neural network (DANN) to extract domain invariant features in the target and source domains. Wang et al. [21] designed a deep adversarial DA model to effectively mitigate the distribution discrepancy between the source and target domains. Although the aforementioned methods effectively mitigate domain shift, they are limited to learning transferable information solely from a single source domain. A single source of diagnostic information may ultimately limit the cross-domain diagnostic performance of the model.

Generally, multiple source domains often exist because the equipment operates in complex and variable operating environments. More abundant diagnostic information can be exploited in a multisource domain dataset, which can compensate for the lack of diagnostic knowledge in a single source domain [22]. Traditional DA methods that learn only single-source information cannot fully utilize the shared information among multiple source domains [23]. Fortunately, multisource DA (MDA) methods have been successfully explored to solve this problem. For example, a novel adversarial DA network with classifier alignment was designed by Zhang et al. [24] for multisource domain problems that may arise in real industrial scenarios. Shi et al. [25] achieved knowledge transfer from multiple sources based on a MDA network with an entropy penalty strategy. Multi-adversarial learning was employed by Zhu et al. [26] to extract high-dimensional domain-invariant features. Wang et al. [27] constructed a sub-domain adaptation network that enables multisource information transfer by integrating source-private feature extractors and classifiers. Shi et al. [28] proposed an instance-adaptive multisource transfer method that mitigates multisource domain shifts and improves the ability of the model to learn multisource diagnostic information.

Although the above methods achieve effective MDA through multi-domain matching, they ignore the distribution discrepancy between each pair of source and target domains, which restricts the accurate matching of source and target domains. Moreover, since source domain data are collected from different working conditions, their respective contributions to the target domain also exhibit discrepancies. The source domain close to the target working condition should be given a larger weight, which indicates that it contains more common diagnostic information about the working condition of the target domain. Literature [29],[30] demonstrate that more competitive classification performance can be achieved by weighting multiple source domains. An improved LMMD was proposed by Tian et al. [31] for computing source-specific weight scores and combining the weights with a multi-classifier to consider the different contributions of the sources. According to the similarity of different working conditions, Wei et al. [32] constructed a specific discriminator for each source in order to implement non-uniform weighting. The successful implementation of the above methods proved the necessity of considering source weights. However, their proposed weighting scheme relies on constructing additional network models and parameters, increasing the required computational cost. Therefore, a more efficient weighted MDA approach needs to be developed for multi-source knowledge transfer.

To this end, a non-uniformly weighted MDA network (NWMDAN) is developed in this study. The proposed method eliminates multi-domain shifts from both the inter-domain perspective between source and target domains and the intra-domain perspective for each pair of source and target domains. To measure the contribution of multiple source domains effectively, a non-uniform weighting scheme is designed based on the similarity between each source domain and the target domain. Furthermore, a non-uniformly weighted adversarial training framework is proposed to enhance MDA performance. The main contributions of this paper are as follows.

(1) A more realistic fault diagnosis scenario, the multisource cross-domain fault diagnosis problem, is explored in this paper. To this end, a NWMDAN is proposed to enhance the MDA performance.
(2) To effectively eliminate multi-domain shifts, an intra-domain distribution alignment strategy is designed to eliminate the intra-domain distribution discrepancy for each pair of source and target domain.
(3) From the perspective of similarity between multiple source and target domains, a non-uniform weighting scheme is proposed for quantifying the contributions of different source domains.
(4) A non-uniformly weighted adversarial training framework is proposed to learn and combine the multisource information better.

The remainder of this paper is structured as follows. The problem definition and relevant theoretical background are provided in Sect. 2. Section 3 provides the framework and diagnostic flow of the proposed NWMDAN method. The experimental validation and performance analysis are carried out in Sect. 4. Finally, the conclusions of this study are given in Sect. 5.

2 Preliminaries

2.1 Problem Definition

In this paper, the problem of multisource cross-domain fault diagnosis is investigated. It conforms to the following basic assumptions: (a) source domain data are collected from

multiple working conditions; (b) target domain data are unlabeled; (c) the label space is shared by both the source and target domains.

Based on the above assumptions, the detailed symbolic definitions of the problem are given as follows. Assuming K source domains are available, and $D_k^s = \{(x_i^{s,k}, y_i^{s,k})\}_{i=1}^{n_{s,k}} (k = 1, 2, ..., K)$ represents the k-th source domain dataset with $n_{s,k}$ labeled samples, where $x_i^{s,k}$ and $y_i^{s,k}$ denote the i-th sample and corresponding condition label from the k-th source domain, respectively. Let $D^t = \{x_j^t\}_{j=1}^{n_t}$ be the target domain dataset with n_t unlabeled samples, where x_j^t denotes the j-th sample from the target domain. The purpose of this study is to build a reliable diagnosis model for multisource cross-domain tasks.

2.2 Maximum Mean Discrepancy

Maximum mean discrepancy (MMD) [32] is often used by statistics matching-based DA methods to calculate the distribution distance in different domains, which can measure the distribution discrepancies of different datasets [33, 34]. Given $X\{x_i | x_i \sim p, i = 1, 2, \ldots, n\}$ and $Y\{y_j | y_j \sim q, j = 1, 2, \ldots, m\}$, the MMD between X and Y is defined as follows:

$$D_{\mathcal{H}}(X, Y) \triangleq ||E_{X \sim p}[\Phi(x)] - E_{Y \sim q}[\Phi(y)]||_{\mathcal{H}} \tag{1}$$

where $\mathcal{H}$ represents the reproducing kernel Hilbert space (RKHS), and $\phi(\cdot)$ denotes the feature mapping function. By introducing the characteristic kernel, the squared value of MMD can be expressed as follows:

$$MMD_{(X,Y)}^2 = \frac{1}{n^2} \sum_{i,j=1}^{n} \mathcal{K}(x_i, x_j) - \frac{2}{mn} \sum_{i=1}^{n} \sum_{j=1}^{m} \mathcal{K}(x_i, y_j) + \frac{1}{m^2} \sum_{i,j=1}^{m} \mathcal{K}(y_i, y_j) \tag{2}$$

where $\mathcal{K}(\cdot, \cdot)$ represents the kernel function. If multiple kernels are available, the multi-kernel MMD (MK-MMD) can be used as a more effective metric to evaluate the differences in the different domains.

2.3 Domain Adversarial Neural Network

As an adversarial training-based DA method, domain adversarial neural network (DANN) [35] has attracted considerable attention in the field of IFD [36, 37]. DANN is composed of a classifier C, a feature extractor F and a domain discriminator D. The key to achieving domain adversarial training is the game process of F and D. D is trained to distinguish samples from the source or target domain accurately, and conversely, the goal of F is to confuse D as could as possible. In conclusion, the optimization objective of the DANN is defined as following equation:

$$L_{DANN} = \frac{1}{n_s} \sum_{x_i \in D_s} L_c(C(F(x_i)), y_i) - \frac{\lambda}{n_s + n_t} \sum_{x_i \in (D_s \cup D_t)} L_d(D(F(x_i)), d_i) \tag{3}$$

where L_c and L_d represent the cross-entropy loss, n_s and n_t denote the number of samples in the D^s and D^t, y_i and d_i denote the true condition label and domain label of i-th sample, respectively, and λ represents the trade-off parameter. To avoid a phased training process, the gradient reversal layer (GRL) is implemented into DANN to optimize the training process, which does not add additional parameters and enables the gradient sign to change when back propagating.
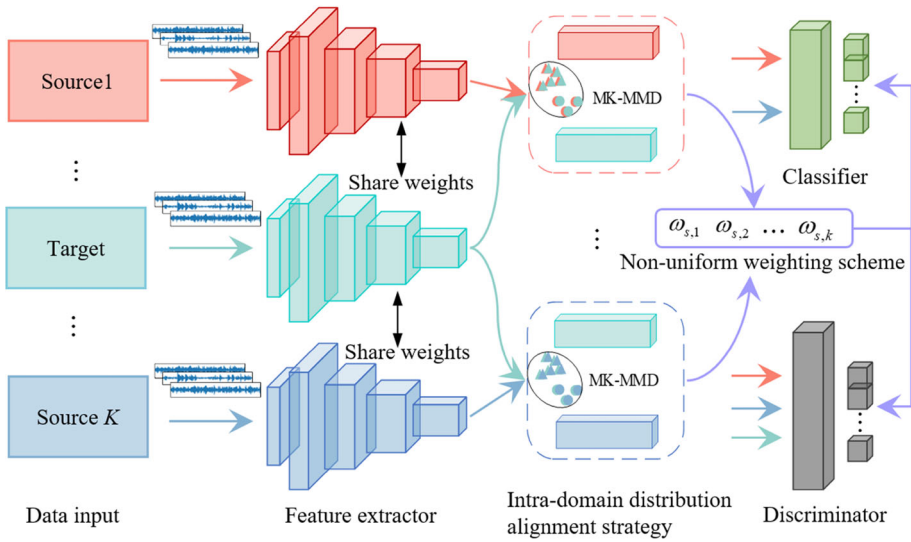
Fig. 1 The framework of the NWMDAN

3 Proposed method

The structure of the proposed NWMDAN is shown in Fig. 1. It consists of a feature extractor, an intra-domain distribution alignment strategy, a non-uniform weighting scheme, and a non-uniformly weighted adversarial training framework. Specifically, to clearly represent the data related to different sources, the corresponding colors are used in Fig. 1. For example, the red color indicates data regarding the first source domain. The structure and parameters of the proposed method are shown in Fig. 2. The Conv, BN, and FC shown in Fig. 2 denote the 1D convolutional layers, the batch normalization layers, and the fully connected layers. The three parameters of the Conv layer, for example, Conv (16,15,1), denote the number of input channels, the convolution kernel size and the stride, respectively. The parameters that follow the FC layer represent their output size. The detailed algorithm flowchart is shown in Algorithm 1.

3.1 Feature Extractor

In the field of IFD, convolutional neural networks (CNN) have been popularly adopted as feature extractors due to their simplicity of training and superior performance [38–40]. Based on this, a CNN with four convolutional layers and a fully connected layer is used as the feature extractor for the proposed method. In particular, batch normalization and dropout are used in the feature extractor to speed up model training and prevent overfitting.

3.2 Intra-Domain Distribution Alignment Strategy

The domain adversarial process described in Sect. 2.3 is able to map the source and target domain data to identical feature space and eliminate distribution differences between them. However, in multisource domain problems, it is challenging to directly eliminate multi-
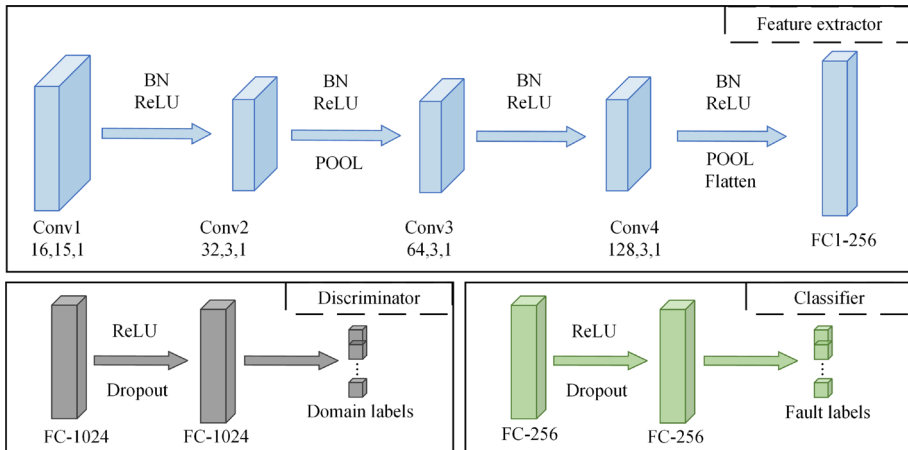
Fig. 2 The structural parameters of NWMDAN

Algorithm 1 NMWDAN

Require: source domain dataset $D_k^s = \{(x_i^{s,k}, y_i^{s,k})\}_{i=1}^{n_{s,k}}(k = 1, 2, \ldots, K)$ and target domain dataset $\{D^t = \{x_j^t\}_{j=1}^{n_t}\}$,initial NWMDAN$\{G_f, G_d, G_c\}$, maximum epoch N, tradeoff parameters λ and β, hardness coefficientη, batch size m.

Ensure: trained NWMDAN $\{G_f, G_c\}$.

1: **for** $k = 1 \to k = N$ **do**
2:    Random sample a batch $\{(x_i^{s,k}, y_i^{s,k})\}_{i=1}^{m/k}$ and $\{x_j^t\}_{j=1}^{m}$ from D_k^s and D^t.
3:    Samples from different source and target domains are fed into a shared feature extractor G_f to get feature representations $f_i^{s,k}$ and f_j^t.
4:    Feed $f_i^{s,k}$ and f_j^t into the intra-domain distribution alignment strategy and use Eq.5 to calculate the total alignment loss L_{intra} for each pair of source and target domains.
5:    Feed $f_i^{s,k}$ and f_j^t into a non-uniform weighting scheme to obtain source-specific weights $\omega_{s,k}$.
6:    Use G_d with source-specific weights to calculate the weighted domain discrimination loss L_d as Eq.7.
7:    Use G_c with source-specific weights to calculate the weighted classification loss L_c as Eq.8.
8:    Update the parameters $\theta_f, \theta_d, \theta_c$ of G_f, G_d, G_c by minimizing the total optimization objective in Eq.9.
9: **end for**

---

domain shifts by globally matching the multisource domain data with the target domain data. Because there are not only distribution discrepancies between the source and target domains but also distribution discrepancies observed among different source domains. Therefore, an intra-domain distribution alignment strategy is designed to align each source domain with the target domain. Specifically, in each iteration of the epoch, all samples from each source domain are selected from the data in each batch_size, and the distribution discrepancy between each source domain and the samples from the target domain in the batch_size is calculated. The proposed intra-domain distribution alignment strategy measures the distribution discrepancy in the two domains using MK-MMD with five Gaussian kernels. The square value of MK-MMD between the k-th source domain and the target domain can be calculated as follows:

$$MMD^2_{(f^{s,k}, f^t)} = \frac{1}{n_{s,k}^2}\sum_{i=1}^{n_{s,k}}\sum_{j=1}^{n_{s,k}}K(f_i^{s,k}, f_j^{s,k})$$

$$- \frac{2}{n_{s,k}n_t} \sum_{i=1}^{n_{s,k}} \sum_{j=1}^{n_t} K(f_i^{s,k}, f_j^t)$$

$$+ \frac{1}{n_t^2} \sum_{i=1}^{n_t} \sum_{j=1}^{n_t} K(f_i^t, f_j^t) \tag{4}$$

where $f_i^{s,k}$ represents the features of the i-th sample from D_k^s, and f_j^t represents the feature of the j-th sample from D^t. When the MK-MMD of all K sources and target domains are calculated, the intra-domain distribution alignment loss can be obtained as follows:

$$L_{\text{intra}} = \frac{1}{K} \sum_{k=1}^{K} MMD_{(f^{s,k}, f^t)}^2. \tag{5}$$

3.3 Non-Uniform Weighting Scheme

To measure the contribution of different source domains, a non-uniform weighting scheme is designed based on the statistical distance between source domains and target domains. Specifically, the MK-MMD distance is utilized in each epoch to calculate distinct source-specific weights. The similarity of domains calculated according to statistical distances has been demonstrated to be a reliable metric for the contribution of different source domains [29]. Consequently, the source domain that is closer to the target domain distribution should be assigned a larger weight. Based on the intra-domain distribution alignment loss accumulating the MMD values within the whole epoch, the weight of k-th source domain in the next epoch is formulated as follows:

$$\omega_{s,k} = \frac{\exp(-\eta MMD_{(f^{s,k}, f^t)}^2)}{\sum_{k=1}^{K} \exp(-\eta MMD_{(f^{s,k}, f^t)}^2)} \tag{6}$$

where η represents the hardness coefficient, it enables a more significant difference in the contributions of different sources; $f^{s,k}$ represents the sample features of D_k^s.

3.4 Non-Uniformly Weighted Adversarial Training Framework

3.4.1 Non-Uniformly Weighted Discriminator

The domain discriminator is the key player in domain adversarial training. The discriminator of traditional DA methods receives the feature representations of the source and target domains and outputs a probability vector representing the domain labels of the samples. Besides the feature representations of the samples, in NWMDAN, the source-specific weights are also used as input elements to the domain discriminator, which makes the proposed method more effective in extracting diagnostic information. Therefore, the weighted domain discrimination loss L_d of the weighted discriminator is defined as follows:

$$L_d = \frac{1}{K} \sum_{k=1}^{K} \frac{\omega_{s,k}}{n_{s,k}} \sum_{(x_i^{s,k}, y_i^{s,k}) \in D_k^s} L_d(D(F(x_i^{s,k}), d_i^{s,k}) + \frac{1}{n_t} \sum_{x_i^t \in D^t} L_d(D(F(x_i^t), d_i^t) \tag{7}$$

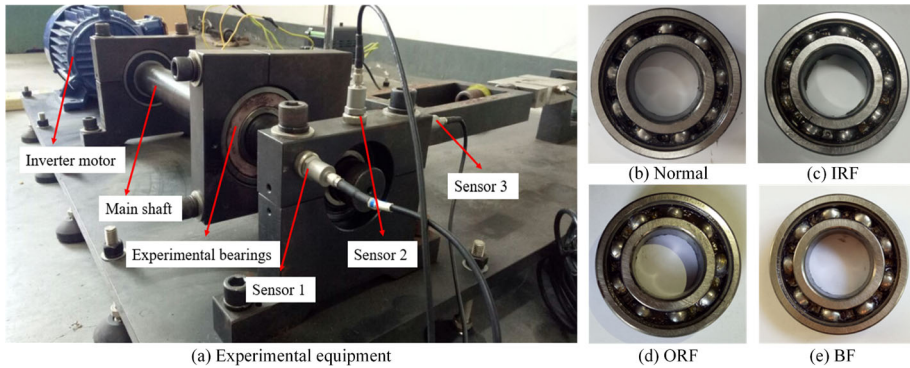where $d_i^{s,k}$ and d_i^t are the domain label of $x_i^{s,k}$ and x_i^t, respectively.

(a) Experimental equipment   (b) Normal   (c) IRF   (d) ORF   (e) BF

Fig. 3 Experimental equipment and test bearings of the Case1 dataset

3.4.2 Non-Uniformly Weighted Classifier

In adversarial training, the goal of the classifier is to identify the health status of samples as accurately as possible while ensuring that sample features remain domain-invariant. Considering that there are differences in the transferability of samples from different source domains, source-specific weights are integrated into the NWMDAN classifier. The weighted classification loss L_c of the weighted classifier can be formulated as follows:

$$L_c = \frac{1}{K} \sum_{k=1}^{K} \frac{\omega_{s,k}}{n_{s,k}} \sum_{(x_i^{s,k}, y_i^{s,k}) \in D_k^s} L_c(C(F(x_i^{s,k})), y_i^{s,k}). \tag{8}$$

3.5 Optimization Objective

The optimization objective contains intra-domain alignment loss L_{intra}, domain adversarial loss L_d and classification loss L_c, which is expressed as Eq. 9:

$$L = L_c + \lambda L_{\text{intra}} - \beta L_d \tag{9}$$

where λ and β are trade-off parameters to balance the loss terms. By minimizing L_c, samples of different fault categories can be predicted accurately, which is beneficial for extracting domain discriminative features. The distribution discrepancy in each pair of source and target domain is reduced by minimizing L_{intra}. Besides, maximizing L_d enables to align multiple source domains to the target domain. Let $\theta_f, \theta_c, \theta_d$ denote the parameters of feature generator, classifier and domain discriminator, respectively, and they can be updated as follows:

$$\theta_f \leftarrow \theta_f - \alpha \left(\frac{\partial L_c}{\partial \theta_f} + \frac{\lambda \partial L_{\text{intra}}}{\partial \theta_f} - \frac{\beta \partial L_d}{\partial \theta_f} \right)$$

$$\theta_c \leftarrow \theta_c - \alpha \frac{\partial L_c}{\partial \theta_c}$$

$$\theta_d \leftarrow \theta_d - \alpha \frac{\partial L_d}{\partial \theta_d} \tag{10}$$

where α is the learning rate. The Eq. 10 will be updated simultaneously in a training epoch by GRL, which avoids a staged training process.
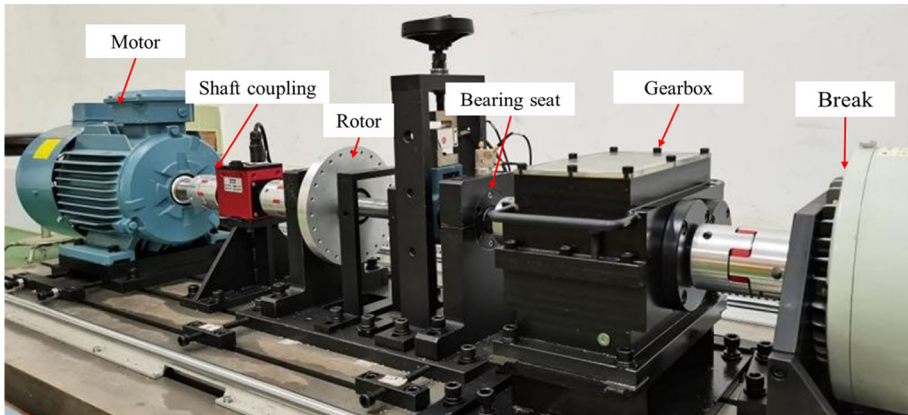
Fig. 4 Experimental equipment of the Case2 dataset

4 Case Study

4.1 Dataset Description

(1) Jiangnan University (JNU) bearing dataset: The JNU dataset is a famous open-source bearing dataset that is widely used as a benchmark to verify the feasibility of diagnostic methods. Four health statuses can be found in the JNU dataset, including normal state (Normal), inner race fault (IRF), outer race fault (ORF) and ball fault (BF). The vibration signals are collected at 50 kHz with three speeds (600, 800 and 1000 r/min).

(2) Self-made bearing test rig dataset-1(Case1): Case1 bearing vibration signals are collected from an experimental rig shown in Fig. 3(a). The experimental system is composed of an inverter motor, a main shaft, a supporting device, and the SKF 6206-2RS1/C3 test bearings. Three sensors distributed at different locations in space are used to collect vibration signals in multiple directions. Similar to the JNU dataset, the Case1 dataset also includes four state types (Normal, ORF, IRF and BF), and each fault has different fault sizes. Figure 3(b)–(d) show the different health conditions of the tested bearings. In addition, four speeds, including 150, 300, 900 and 1500 r/min, are applied to the bearings by adjusting the driving device.

(3) Self-made bearing test rig dataset-2(Case2): The Case2 bearing data are collected on a self-made experimental rig shown in Fig. 4, where, besides the ORF, IRF, and BF failure types, bearing retainer fracture failures (RF) are also simulated. The vibration signals of three loads, 0N, 500N and 1000N, at the speed of 300r/min are selected as experimental data for this experiment to further verify the effectiveness of the proposed method. The detailed descriptions of datasets are given in Table 1.

4.2 Comparison Methods

Several commonly used DA methods are chosen as comparison methods to validate the effectiveness and performance of the proposed network comprehensively. Specifically, the selected comparison methods have the same network structure and parameters as our proposed method. According to different states of the source domain data implemented in the

Table 1 Detailed description of the three datasets

Dataset	Condition	Fault size (mm)	Label	Working condition
JNU	Normal	–	0	600/800/1000 r/min
	IRF	–	1	
	ORF	–	2	
	BF	–	3	
Case1	Normal	–	0	150/300/900/1500 r/min
	BF	0.4	1	
	BF	0.2	2	
	IRF	0.3	3	
	IRF	0.4	4	
	ORF	0.2	5	
	ORF	0.3	6	
Case2	Normal	–	0	0/500/1000 N
	RF	–	1	
	BF	–	2	
	IRF	0.3	3	
	IRF	0.5	4	
	ORF	0.2	5	
	ORF	0.7	6	

comparison methods, the comparison experiments can be divided into the following three groups.

(1) Single source domain-based transfer methods (Single best). CNN, deep adaptation network (DAN) [41], deep coral (DCORAL) [42], joint adaptation network (JAN) [43], and DANN with a single-source domain. For this group of comparison methods, we select different source domains successively to construct transfer tasks and record the maximum diagnostic accuracy obtained using different source domains as the final result. Among them, the CNN model whose feature extractor and classifier have the same structure as the proposed method, is chosen as the benchmark for the comparison methods. The most widely used statistics matching-based DA methods, DAN and DCORAL, are chosen as comparison methods, which calculate the distribution discrepancy by MK-MMD and CORAL, respectively. Additionally, a JAN based on the joint maximum mean discrepancy is also used in the comparison experiments. Then, the adversarial learning-based DANN is chosen as one of the comparison methods, which is the original framework of our approach.

(2) Merged source domain-based transfer methods (Merged source domains). CNN, DAN, JAN, DCORAL, and DANN with merged source domains. Such group comparison methods provide improvements in source domain data compared with diagnosis methods based on the single source. Merging fault samples collected under different working conditions into the same source domain makes it possible to extract more diagnostic information.

(3) Multi-source domain transfer methods (Multisource domain). NWMDAN_a, NWM-DAN_b, MDA network (MDAN) [44] with multisource domain. As one of the most effective MDA approaches, MDAN is employed as a comparison method. NWMDAN_a and NWMDAN_b, variants of NWMDAN, serve as comparison methods to validate

Table 2 Detailed description of all diagnostic tasks

Dataset	Working condition	Domain name
JNU	Rotating speed:600 r/min	A
	Rotating speed:800 r/min	B
	Rotating speed:1000 r/min	C
Case1	Rotating speed:150 r/min	D
	Rotating speed:300 r/min	E
	Rotating speed:900 r/min	F
	Rotating speed:1500 r/min	G
Case2	Load:0 N	H
	Load:500 N	I
	Load:1000 N	J

the effectiveness of the proposed intra-domain distribution alignment strategy and non-uniformly weighted adversarial training framework. In NWMDAN_a, the intra-domain distribution alignment strategy is removed, while in NWMDAN_b, the non-uniform weights are omitted. However, both NWMDAN_a and NWMDAN_b retain the same structure and parameters as NWMDAN.

The utilized CNN model comprises a feature extractor and a classifier, whose structures are consistent with the corresponding components in the proposed method. Specifically, DAN, JAN, and DCORAL augment the CNN architecture by incorporating MK-MMD, joint MMD, and CORAL distance metrics respectively to address distribution discrepancies. In the DANN model, the feature extractor and classifier align with the proposed method, while the discriminator functions as a binary classifier. Extending the DANN, the MDAN model enhances the binary discriminator to enable the recognition of multiple domains. The network architectures of NWMDAN_a and NWMDAN_b align with the proposed method, with the distinction that the intra-domain distribution alignment strategy and non-uniform weighting scheme have been respectively omitted.

4.3 Experimental Details

The vibration data for different rotational speeds in the JNU and Case1 datasets are regarded as different domains, respectively, as shown in Table 2. The JNU dataset contains 2930 samples in each domain, which includes 1466 samples of healthy bearings and the remaining are samples of various faults. For the Case1 dataset, 1134 samples are available in each domain, which consists of seven categories with equal number of samples. The Case2 dataset contains 195 samples for each fault state. By default, the samples in the three datasets contain 1024 sampling points. For each task, 80% of samples from source and target domains are randomly selected as training samples, and the remaining 20% samples are used for testing. The details of all the transfer tasks on both datasets are shown in Table 3.

The trade-off parameters λ and β of NWMDAN, which correspond to domain alignment loss and discrimination loss in the optimization objective, respectively, may have a significant impact on performance. Therefore, the grid search strategy is employed to investigate the sensitivity of parameters λ and β on the network performance, and the experimental results in tasks B,C→A are illustrated in Fig. 5. According to Fig. 5, the highest diagnostic accuracy is

Table 3 Detailed description of all diagnostic tasks

Dataset	Transfer task	Source domain	Target domain
JNU	A+B→C	A+B	C
	A+C→B	A+C	B
	B+C→A	B+C	A
Case1	D+E+F→G	D+E+F	G
	D+E+G→F	D+E+G	F
	D+F+G→E	D+F+G	E
	E+F+G→D	E+F+G	D
Case2	H+I→J	H+I	J
	H+J→I	H+J	I
	I+J→H	I+J	H
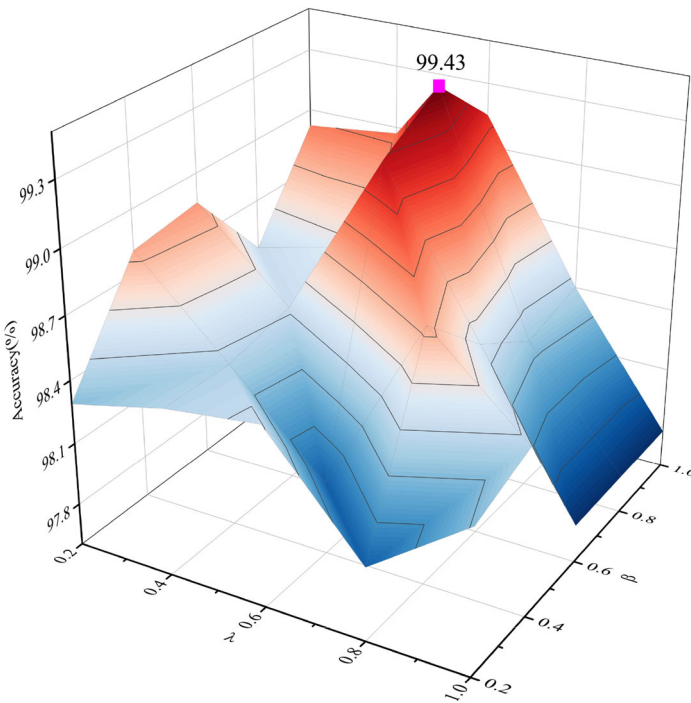
Fig. 5 Sensitivity analysis of trade-off parameters on task B,C→A

achieved at $\lambda = 0.8$ and $\beta = 0.6$, and the performance of NWMDAN degrades significantly when the parameter values close to 1 simultaneously.

Furthermore, the sensitivity of the diagnostic performance with respect to the hardness coefficient η is investigated on three tasks and the obtained results are shown in Fig. 6. The experimental results demonstrate that the proposed method achieves the highest diagnostic accuracy when $\eta = 20$. When the value of η is close to 100, the accuracy decreases on all diagnostic tasks, and the performance degradation is more obvious in some of the tasks with more challenging knowledge transfer, for example B, C→A. The detailed setting of the remaining parameters are given in Table 4. In particular, a fixed-step update strategy is

Table 4 Detailed parameter settings for the experiments

Parameters	Value	Parameters	Value
Maximum epoch	300	Weight decay	1e-5
Batch size	64	Trade-off parameters λ, β	0.8,0.6
Optimizer	Adam	Learning rate	1e-3
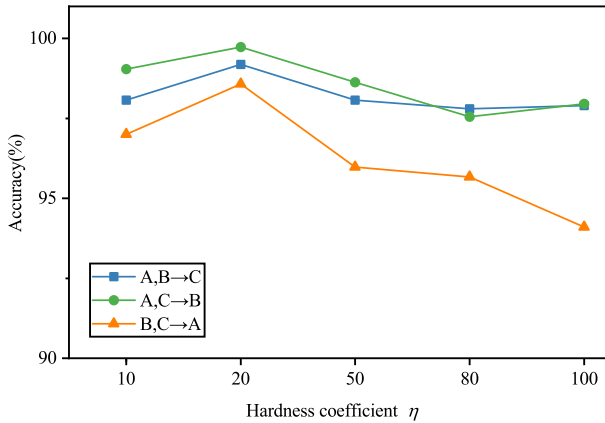
Fig. 6 Sensitivity analysis of hardness coefficient

used to adjust the learning rate, which decreases to 1e-4 and 1e-5 after 150 and 250 epochs, respectively.

Finally, the experimental accuracies of the test samples in the target domain are chosen as the model evaluation metrics. Each experiment is performed five times repeatedly to avoid the effect of randomness, and the averaged value of five experiments is regarded as the final diagnosis accuracy. All experiments are implemented through the Pytorch framework on a experimental equipment with Intel Core i5-11400K, NVIDIA GeForce RTX 3060.

4.4 Experimental Results and Analysis on JNU Case

4.4.1 Comparison of Classification Accuracy

The proposed method and the comparison methods discussed in the previous section are applied to the tasks in the JNU case, and the detailed diagnostic results are presented in Table 5. It is clear that more competitive diagnostic accuracy of NWMDAN is observed in all diagnostic tasks with an average diagnostic accuracy 99.16%, indicating its effectiveness in multisource cross-domain diagnostic tasks. By further analysis, the following conclusions can be found.

(1) Compared with single source-based transfer methods, merged source domains and multisource-based methods achieve higher diagnostic accuracy in most tasks, but negative transfers are observed in some tasks, for example, task A+B→C. These results demonstrate that more diagnostic information can be found in multisource domains than in single-source domains, but extracting them is a complicated task.
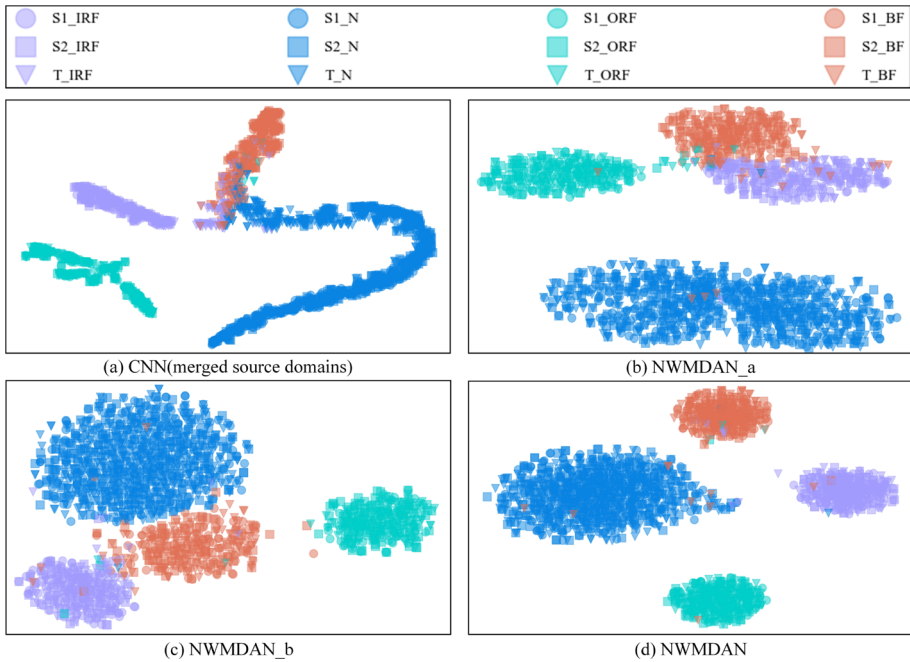
Table 5 Classification accuracy (%) on the JNU case with different methods

Groups	Methods	Transfer tasks			Average
		A+B→C	A+C→B	B+C→A	
Single best	CNN	94.92	94.67	79.71	89.77
	DAN	96.55	96.79	95.29	96.21
	JAN	97.71	97.39	96.76	97.29
	DCORAL	97.30	95.37	89.74	94.14
	DANN	97.18	96.42	91.06	94.89
Merged source domains	CNN	95.33	96.43	83.96	91.91
	DAN	96.79	97.65	94.95	96.46
	JAN	97.54	98.09	96.45	97.36
	DCORAL	96.17	96.79	89.86	94.27
	DANN	97.88	97.27	93.33	96.16
Multisource domain	MDAN	96.31	97.58	94.83	96.24
	NWMDAN_a	97.47	98.26	96.11	97.28
	NWMDAN_b	97.75	98.15	97.00	97.63
Proposed method	NWMDAN	99.18	99.73	98.57	99.16

(2) Although the merged source domain approaches achieve an improvement in diagnostic accuracy compared with the single source domain-based approaches, their performance is lower than the proposed approach. It indicates that the multisource invariant features are not sufficiently extracted by the methods of the merged source domains. For the proposed NWMDAN, the average diagnostic accuracy remains above 99% in all tasks, further demonstrating that the proposed method achieves efficient and accurate knowledge transfer.

(3) The proposed NWMDAN outperforms MDAN in terms of diagnostic accuracy on the multisource tasks. Furthermore, the results validate the advantages and necessity of the proposed alignment strategy and non-uniform weighting scheme. Despite NWMDAN_a achieves more promising diagnostic results than MDAN, it is less competitive compared with the proposed method. Similarly, when the weighting scheme is removed, the diagnostic accuracy of the NWMDAN_b decreases on different tasks compared with the results of the proposed method.

4.4.2 Feature Visualization Results

To analyze the model performance and effectiveness more intuitively, the high-dimensional feature representations are visualized using the distributed stochastic neighbor embedding (T-SNE) technology. The feature visualization results of task B+C→A are given in Fig. 7. As shown in Fig. 7(a), the CNN without using the transfer learning strategy has serious class confusion and unclear classification boundaries, which confirms the poor diagnostic results of the CNN in Table 5. Poor clustering performance and significant overlaps between classes are found in Fig. 7(b). This indicates that it is difficult to eliminate multi-domain shifts only by adversarial training methods. In Fig. 7(c), although a promising clustering performance is achieved, the samples of different categories are not well separated, which leads to a decline in classification accuracy. By contrast, separable boundaries and fused features among samples of different health statuses are observed clearly in Fig. 7 (d), which indicates the excellent properties of the proposed method in eliminating multi-domain shifts.

○ S1_IRF	● S1_N	● S1_ORF	● S1_BF
■ S2_IRF	■ S2_N	■ S2_ORF	■ S2_BF
▼ T_IRF	▼ T_N	▼ T_ORF	▼ T_BF

(a) CNN(merged source domains)　　　　(b) NWMDAN_a

(c) NWMDAN_b　　　　(d) NWMDAN

Fig. 7 T-SNE visualization results of different methods on task B+C→A

4.4.3 Weight Analysis

To analyze the effectiveness of the weighting scheme, the weight values assigned to multiple source domains in different diagnostic tasks are recorded in Fig. 8, where w_1 and w_2 denote the weight values of the first and second source domain, respectively.

As can be seen from Fig. 8, the weights estimated based on the MMD values are significantly correlated with the physical significance of the source domain working conditions. The source domains that have a smaller gap with the working conditions of the target domain are assigned a larger weight. The results in Figs. 8 and 9 show a similar trend between the weights and the single source transfer results. When a single source domain is used for transfer and achieves higher diagnostic accuracy, it is often given a higher weight in MDA scenario. In the tasks A+B→C and B+C→A, where the speed of A, B, C is 600, 800 and 1000 r/min, respectively, the B source domain whose speed is closest to the target domain is given the largest weight. In the diagnostic task A+C→B where the gaps in working conditions between the source and target domains are equal, the A source domain with lower speed is selected as the best source. This suggests that the A source domain within the JNU dataset holds diagnostic information closely related to the target working condition.

4.5 Experimental Results and Analysis on Case1

4.5.1 Comparison of Classification Accuracy

Experiments on the Case1, which keeps the same experimental setup as the JNU case, are constructed for further validation of the effectiveness of NWMDAN. The detailed experimen-
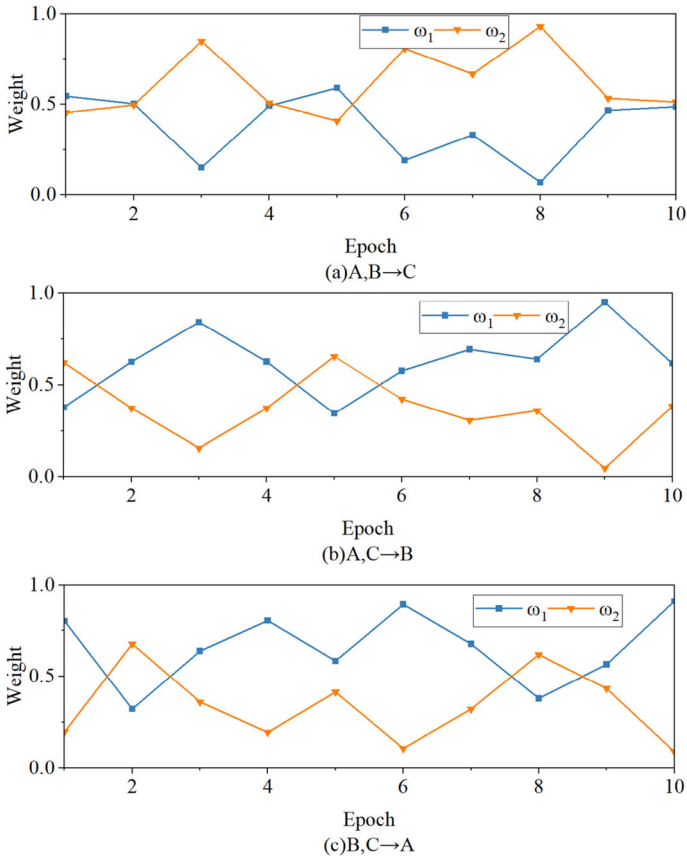
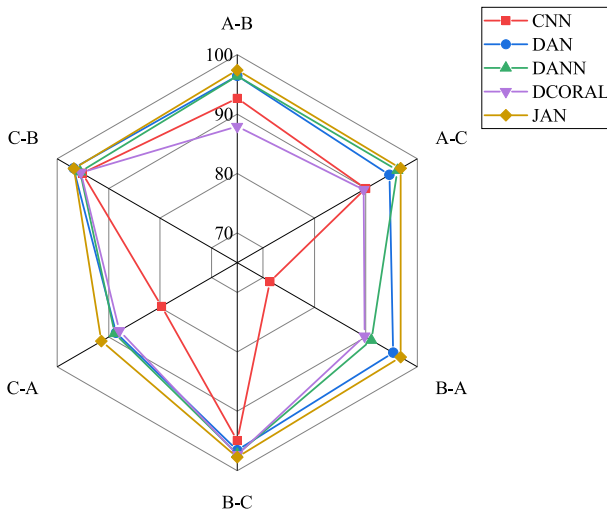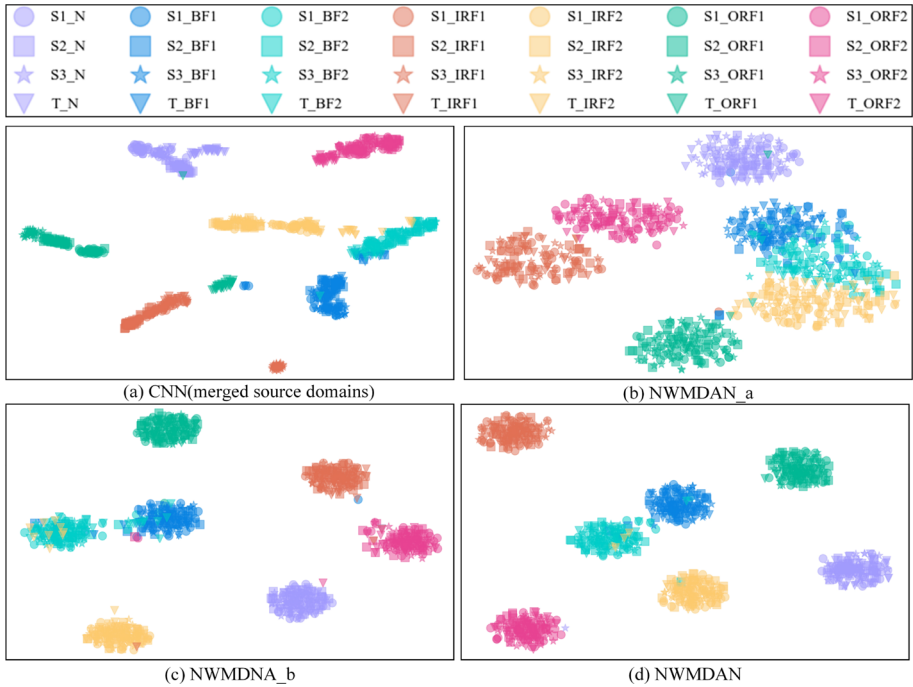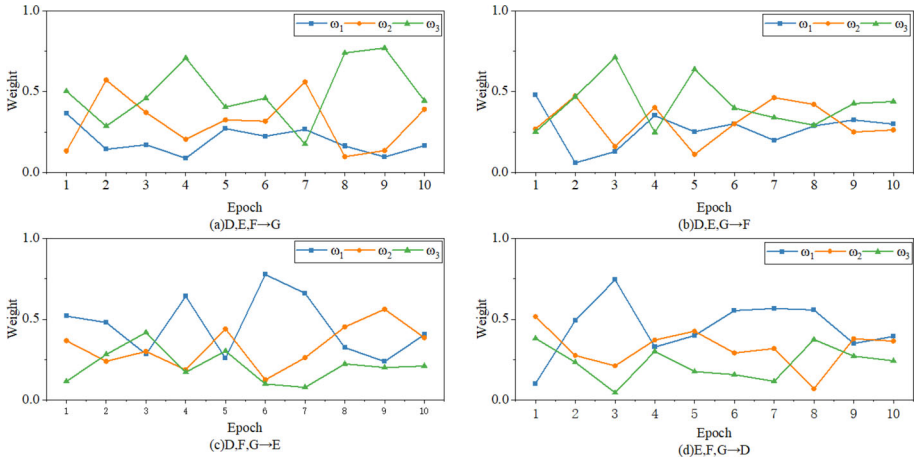**Fig. 8** Weight values of source domains in different tasks for JNU case

Fig. 9 Single-source domain diagnosis accuracy (%) on JNU case

● S1_N	● S1_BF1	● S1_BF2	● S1_IRF1	● S1_IRF2	● S1_ORF1	● S1_ORF2
■ S2_N	■ S2_BF1	■ S2_BF2	■ S2_IRF1	■ S2_IRF2	■ S2_ORF1	■ S2_ORF2
★ S3_N	★ S3_BF1	★ S3_BF2	★ S3_IRF1	★ S3_IRF2	★ S3_ORF1	★ S3_ORF2
▼ T_N	▼ T_BF1	▼ T_BF2	▼ T_IRF1	▼ T_IRF2	▼ T_ORF1	▼ T_ORF2

(a) CNN(merged source domains)   (b) NWMDAN_a

(c) NWMDNA_b   (d) NWMDAN

Fig. 10 T-SNE visualization results of different methods on task E+F+G→D

tal results are recorded in Table 6. It can be seen that NWMDAN achieves a more competitive diagnostic performance than the comparison methods. More concretely, an average diagnostic accuracy 99.41% is observed for the proposed method in all tasks. Meanwhile, the negative transfer is not observed in the diagnosis results of the NWMDAN. It is also observed that the proposed network retains a superior diagnostic performance compared to NWMDAN_a and NWMDAN_b, further validating the effectiveness of the intra-domain distribution alignment strategy and weighting scheme.

4.5.2 Feature Visualization Results

Similar to the JNU case, the visualization of experimental results for the transfer task E+F+G→D are illustrated in Fig. 10. The promising clustering performance can be observed in the visualization result of NWMDAN shown in Fig. 10(d), where different states of fault samples are obviously separated. By contrast, from Fig. 10(a), it is clear that there is a significant overlap among the classes, and poor alignment performance of the source and target domains is achieved. Figure 10(b) and (c) show clear changes in the distribution of features across categories, but there is a slight overlap between categories, which leads to a decrease in diagnosis performance.

4.5.3 Weight Analysis

The diagnosis tasks in the Case1 contain three source domains, which are different from the JNU case. In Fig. 11, the weight curves of the source domains on the four transfer

Table 6 Classification accuracy (%) on the Case1 with different methods

Groups	Methods	Transfer tasks				Average
		D+E+F→G	D+E+G→F	D+F+G→E	E+F+G→D	
Single best	CNN	96.67	81.37	90.63	87.50	89.04
	DAN	98.94	97.09	95.95	92.16	96.04
	JAN	99.21	97.95	97.18	91.19	96.38
	DCORAL	98.68	94.36	96.04	90.40	94.87
	DANN	98.86	93.83	96.48	92.51	95.42
Merged source domains	CNN	96.48	94.80	91.37	86.61	92.32
	DAN	98.87	96.92	96.21	91.81	95.95
	JAN	98.86	97.36	96.04	94.45	96.68
	DCORAL	98.42	96.74	94.10	91.72	95.25
	DANN	98.77	97.18	96.39	94.36	96.68
Multisource domain	MDAN	98.15	96.74	95.59	91.63	95.53
	NWMDAN_a	99.30	98.06	98.15	94.27	97.45
	NWMDAN_b	99.47	98.42	97.80	96.48	98.04
Proposed method	NWMDAN	99.91	99.47	99.74	98.98	99.41

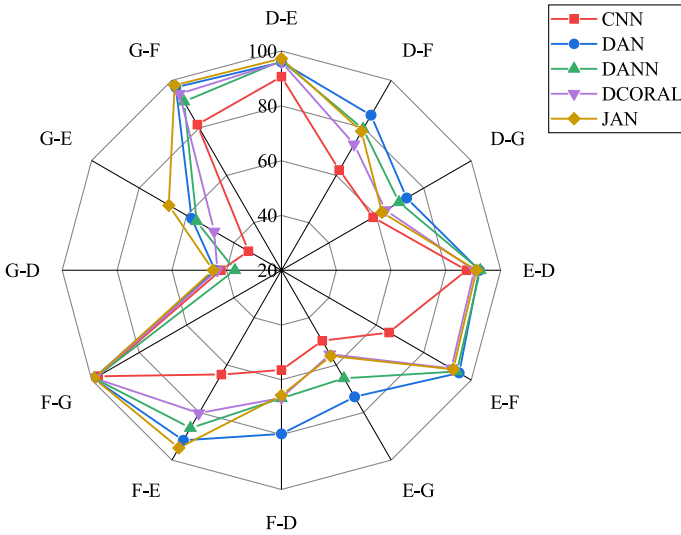**Fig. 11** Weight values of source domains in different tasks for Case1

Fig. 12 Single-source domain diagnosis accuracy (%) on Case1

tasks are recorded, where w_1, w_2, w_3 represent the weights of different source domains, respectively. Combined with the single-source domain transfer diagnosis results shown in Fig. 12, it can be observed that the assignment of weights is closely related to the working condition differences between the source and target domains. The F, G, D, and E source domains are given the largest weights of the four diagnostic tasks, respectively, which is consistent with the ranking of diagnostic results based on the single domain.

Table 7 Classification accuracy (%) on the Case2 with different methods

Groups	Methods	Transfer tasks			Average
		H+I→J	H+J→I	I+J→H	
Single best	CNN	96.83	96.46	96.41	96.56
	DAN	98.05	97.92	97.80	97.92
	JAN	98.08	97.74	96.94	97.59
	DCORAL	98.17	97.80	97.75	97.87
	DANN	96.68	97.67	96.58	96.98
Merged source domains	CNN	96.82	96.79	95.97	96.53
	DAN	97.44	99.51	96.70	97.88
	JAN	98.05	98.41	98.53	98.33
	DCORAL	97.93	98.29	96.94	97.72
	DANN	97.56	98.66	97.07	97.76
Multisource domain	MDAN	98.35	98.77	97.90	98.34
	NWMDAN_a	97.66	99.02	98.40	98.36
	NWMDAN_b	98.97	99.27	99.02	99.09
Proposed method	NWMDAN	99.88	100.00	99.39	99.76

4.6 Experimental Results and Analysis on Case2

4.6.1 Comparison of Classification Accuracy

The diagnostic accuracy of the proposed NWMDAN and comparison methods on the Case2 dataset are shown in Table 7. It can be observed from Table 7 that NWMDAN has an average diagnostic accuracy of 99.76% on the three tasks, outperforming the comparison methods. On the task H+J→I, NWMDAN achieves 100% classification accuracy, while the single best and merged source domain approaches cause degradation in classification accuracy. The classification performance of NWMDAN_a and NWMDAN_b outperforms MDAN, further validating the effectiveness of the proposed intra-domain distribution alignment strategy and non-uniform weighted domain adversarial framework. Compared to the JNU and Case1 datasets, the load variations on the Case2 dataset do not significantly increase the knowledge transfer challenge, and the NWMDAN and comparison methods show high accuracy of classification on all three tasks.

4.6.2 Feature Visualization Results

The results of the T-SNE visualizations on the Case2 dataset are presented in Fig. 13. The visualization results of NWMDAN shown in Fig. 13(d) show promising classification results with clear classification boundaries for different classes, and sample features in the target domain can be accurately matched to the corresponding source domain. In contrast, the visualizations shown in Fig. 13(a) and (b) exhibit poor distinguishability between the classes, and Fig. 13(c) presents a partial confusion of the classes.

4.6.3 Weight Analysis

The non-uniform weights during the training of the transfer task for the Case2 dataset are recorded in Fig. 14, and the diagnostic accuracy of single source domain-based methods is

S1_N	S1_RF	S1_BF	S1_IRF1	S1_IRF2	S1_ORF1	S1_ORF2
S2_N	S2_RF	S2_BF	S2_IRF1	S2_IRF2	S2_ORF1	S2_ORF2
T_N	T_RF	T_BF	T_IRF1	T_IRF2	T_ORF1	T_ORF2

(a) CNN(merged source domains)   (b) NWMDAN_a

(c) NWMDNA_b   (d) NWMDAN
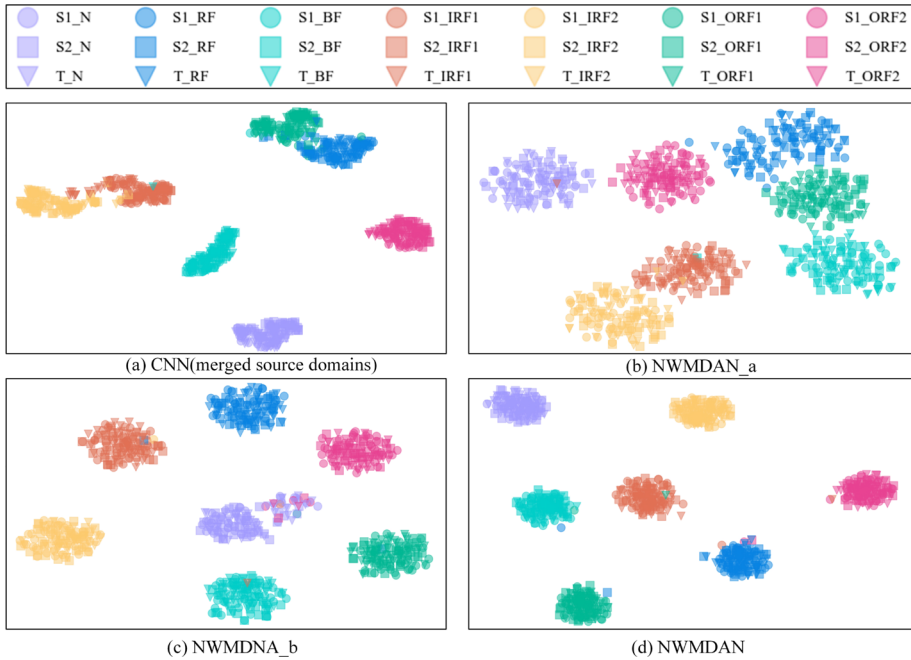
Fig. 13 T-SNE visualization results of different methods on task I+J→H

shown in Fig. 15. Different from the JNU dataset and Case1 dataset, the variations of load on the Case2 dataset are considered. The results in Fig. 14 are similar to JNU and Case1, showing that the weights of different sources are are physically correlated. Sources with high similarity to the target domain working conditions tend to be assigned higher weights. Conversely, the contributions of source domains that differ significantly from the target working conditions are reduced.

5 Conclusion

In this study, a multisource domain IFD method, NWMDAN, considering non-uniform weights of source domains and domain shift between each pair of source and target domain is proposed. The proposed NWMDAN can precisely align multisource domains with the target domain and extract multi-domain invariant features effectively. In particular, the proposed NMWDAN eliminates the domain shifts between each pair of source and target domains by an intra-domain distribution alignment strategy proposed in this study. Considering the different contribution of different source domains to the target domains, a non-uniform weighting scheme is designed to measure the relative importance of different source domains. Moreover, to further learn and combine multisource diagnostic information, a non-uniformly weighted domain adversarial framework is designed. The experimental results on three datasets demonstrate the effectiveness of the proposed NWMDAN. Comprehensive results show that the proposed network obtains promising DA effect and more competitive diagnostic performance compared with the comparison methods.
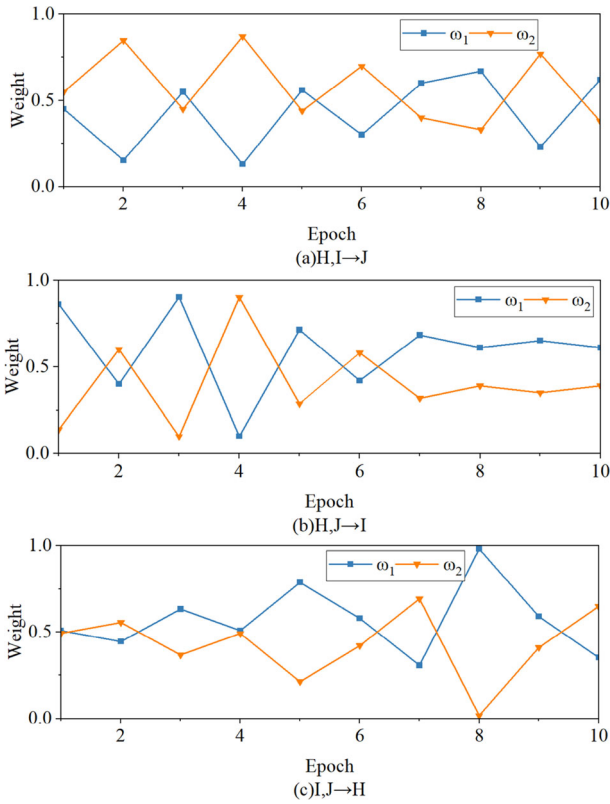
Fig. 14 Weight values of source domains in different tasks for Case2
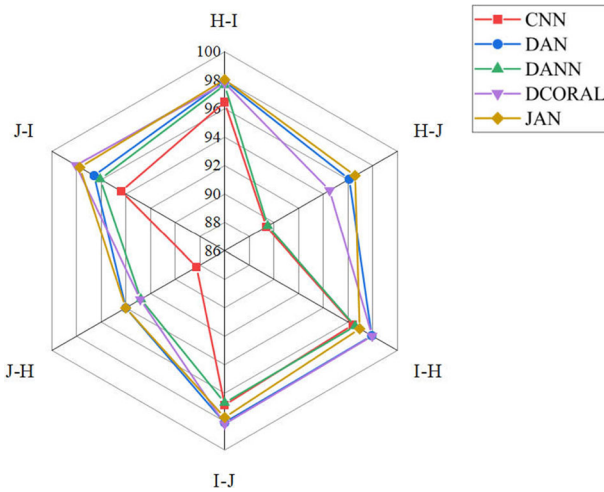

Fig. 15 Single-source domain diagnosis accuracy (%) on Case2

Although the proposed method is expected to achieve good MDA performance, it is limited by the assumption that the label spaces of the source and target domains are equivalent. Future research will focus on a more complex multisource cross-domain diagnostic scenario with asymmetric labeling spaces of the source and target domains.

Declarations

confcts of interest The author(s) declared no potential confcts of interest concerning the research, authorship, and/or publication of this article.

References

1. Yu G, Lin T, Wang Z, Li Y (2021) Time-reassigned multisynchrosqueezing transform for bearing fault diagnosis of rotating machinery. IEEE Trans Industr Electron 68(2):1486–1496. https://doi.org/10.1109/TIE.2020.2970571

2. Zhu D, Cheng X, Yang L, Chen Y, Yang SX (2022) Information fusion fault diagnosis method for deep-sea human occupied vehicle thruster based on deep belief network. IEEE Trans Cybern 52(9):9414–9427. https://doi.org/10.1109/TCYB.2021.3055770

3. Xia Y, Shen C, Wang D, Shen Y, Huang W, Zhu Z (2022) Moment matching-based intraclass multisource domain adaptation network for bearing fault diagnosis. Mech Syst Signal Process168, 108697. https://doi.org/10.1016/j.ymssp.2021.108697

4. Hasan MJ, Sohaib M, Kim J-M (2021) An explainable ai-based fault diagnosis model for bearings. Sensors 21(12):4070. https://doi.org/10.3390/s21124070

5. Pan H, Xu H, Liu Q, Zheng J, Tong J (2022) An intelligent fault diagnosis method based on adaptive maximal margin tensor machine. Measurement 198, 111337. https://doi.org/10.1016/j.measurement.2022.111337

6. Cao H, Shao H, Liu B, Cai B, Cheng J (2022) Clustering-guided novel unsupervised domain adversarial network for partial transfer fault diagnosis of rotating machinery. IEEE Sens J 22(14):14387–14396. https://doi.org/10.1109/JSEN.2022.3182727

7. Zhao MH, Zhong SS, Fu XY, Tang BP, Pecht M (2020) Deep residual shrinkage networks for fault diagnosis. IEEE Trans Industr Inf 16(7):4681–4690. https://doi.org/10.1109/tii.2019.2943898

8. Wang Y, Ning D, Lu J (2022) A novel transfer capsule network based on domain-adversarial training for fault diagnosis. Neural Process Lett 54(5):4171–4188. https://doi.org/10.1007/s11063-022-10803-y

9. Xu H, Wang J, Liu J, Peng X, He C (2022) A novel joint distinct subspace learning and dynamic distribution adaptation method for fault transfer diagnosis. Measurement 203, 111986. https://doi.org/10.1016/j.measurement.2022.111986

10. Qian C, Zhu J, Shen Y, Jiang Q, Zhang Q (2022) Deep transfer learning in mechanical intelligent fault diagnosis: application and challenge. Neural Process Lett 54(3):2509–2531. https://doi.org/10.1007/s11063-021-10719-z

11. Qin Y, Yao Q, Wang Y, Mao Y (2021) Parameter sharing adversarial domain adaptation networks for fault transfer diagnosis of planetary gearboxes. Mech Syst Signal Process 160, 107936. https://doi.org/10.1016/j.ymssp.2021.107936

12. Vashishtha G, Kumar R. Unsupervised learning model of sparse filtering enhanced using wasserstein distance for intelligent fault diagnosis. J Vib Eng Tech. https://doi.org/10.1007/s42417-022-00725-3

13. Li X, Hu Y, Zheng J, Li M, Ma W (2021) Central moment discrepancy based domain adaptation for intelligent bearing fault diagnosis. Neurocomputing 429, 12–24. https://doi.org/10.1016/j.neucom.2020.11.063

14. Li C, Zhang S, Qin Y, Estupinan E (2020) A systematic review of deep transfer learning for machinery fault diagnosis. Neurocomputing 407, 121–135. https://doi.org/10.1016/j.neucom.2020.04.045

15. Wang Y, Ning D, Lu J (2022) A novel transfer capsule network based on domain-adversarial training for fault diagnosis. Neural Process Lett 54(5):4171–4188. https://doi.org/10.1007/s11063-022-10803-y

16. Schwendemann S, Amjad Z, Sikora A (2021) Bearing fault diagnosis with intermediate domain based layered maximum mean discrepancy: a new transfer learning approach. Eng Appl Artif Intell 105, 104415. https://doi.org/10.1016/j.engappai.2021.104415

17. Xiong P, Tang B, Deng L, Zhao M, Yu X (2021) Multi-block domain adaptation with central moment discrepancy for fault diagnosis. Measurement 169, 108516. https://doi.org/10.1016/j.measurement.2020.108516

18. Wang Z, He X, Yang B, Li N (2022) Subdomain adaptation transfer learning network for fault diagnosis of roller bearings. IEEE Trans Industr Electron 69(8):8430–8439. https://doi.org/10.1109/TIE.2021.3108726

19. Wang Y, Ning D, Lu J (2022) A novel transfer capsule network based on domain-adversarial training for fault diagnosis. Neural Process Lett 54(5):4171–4188. https://doi.org/10.1007/s11063-022-10803-y

20. Ganin Y, Ustinova E, Ajakan H, Germain P, Larochelle H, Laviolette F, Marchand M, Lempitsky V (2016) Domain-adversarial training of neural networks. The journal of machine learning research 17(1):2096–2030

21. Wang Y, Sun X, Li J, Yang Y (2021) Intelligent fault diagnosis with deep adversarial domain adaptation. IEEE Trans Instrum Meas 70, 1–9. https://doi.org/10.1109/TIM.2020.3035385

22. Sun S, Shi H, Wu Y (2015) A survey of multi-source domain adaptation. Inf Fusion 24, 84–92. https://doi.org/10.1016/j.inffus.2014.12.003

23. Wang R, Huang W, Wang J, Shen C, Zhu Z (2022) Multisource domain feature adaptation network for bearing fault diagnosis under time-varying working conditions. IEEE Trans Instrum Meas 71, 1–109761066. https://doi.org/10.1109/TIM.2022.3168903

24. Zhang Y, Ren Z, Zhou S, Yu T (2020) Adversarial domain adaptation with classifier alignment for cross-domain intelligent fault diagnosis of multiple source domains. Meas Sci Technol 32(3):035102. https://doi.org/10.1088/1361-6501/abcad4

25. Shi Y, Deng A, Ding X, Zhang S, Xu S, Li J (2022) Multisource domain factorization network for cross-domain fault diagnosis of rotating machinery: an unsupervised multisource domain adaptation method. Mech Syst and Signal Process 164, 108219. https://doi.org/10.1016/j.ymssp.2021.108219

26. Zhu J, Chen N, Shen CQ (2021) A new multiple source domain adaptation fault diagnosis method between different rotating machines. IEEE Trans Industr Inf 17(7):4788–4797. https://doi.org/10.1109/tii.2020.3021406

27. Wang ZC, Huang WT, Chen Y, Jiang YC, Peng GL (2022) Multisource cross-domain fault diagnosis of rolling bearing based on subdomain adaptation network. Meas Sci Technol 33(10). https://doi.org/10.1088/1361-6501/ac7941

28. Shi YW, Deng AD, Deng MQ, Xu M, Liu Y, Ding X, Bian WB(2022) Instance adaptive multisource transfer for fault diagnosis of rotating machinery under variable working conditions. Measurement 202. https://doi.org/10.1016/j.measurement.2022.111797

29. Blitzer J, Crammer K, Kulesza A, Pereira F, Wortman J (2007) Learning bounds for domain adaptation. Adv Neural Inf Process Syst 20:129–136

30. Ge L, Gao J, Ngo H, Li K, Zhang A (2014) On handling negative transfer and imbalanced distributions in multiple source transfer learning. Stat Anal Data Mining ASA Data Sci J 7(4):254–271. https://doi.org/10.1002/sam.11217

31. Tian J, Han D, Li M, Shi P (2022) A multi-source information transfer learning method with subdomain adaptation for cross-domain fault diagnosis. Knowl Based Syst 243, 108466. https://doi.org/10.1016/j.knosys.2022.108466

32. Wei D, Han T, Chu F, Zuo MJ (2021) Weighted domain adaptation networks for machinery fault diagnosis. Mech Syst Signal Process 158, 107744. https://doi.org/10.1016/j.ymssp.2021.107744

33. Borgwardt KM, Gretton A, Rasch MJ, Kriegel H-P, Schölkopf B, Smola AJ (2006) Integrating structured biological data by kernel maximum mean discrepancy. Bioinformatics 22(14):49–57. https://doi.org/10.1093/bioinformatics/btl242

34. Arora JK, Rajagopalan S, Singh J, Purohit A. Low-frequency adaptation-deep neural network-based domain adaptation approach for shaft imbalance fault diagnosis. J Vib Eng Technol. https://doi.org/10.1007/s42417-022-00848-7

35. Ghifary M, Kleijn WB, Zhang M (2014) Domain adaptive neural networks for object recognition. In: 13th Pacific rim international conference on artificial intelligence (PRICAI). Lect Notes Artif Intell 8862, 898–904. https://doi.org/10.1007/978-3-319-13560-1_76

36. Li X, Zhang W, Ma H, Luo Z, Li X (2020) Partial transfer learning in machinery cross-domain fault diagnostics using class-weighted adversarial networks. Neural Netw 129, 313–322. https://doi.org/10.1016/j.neunet.2020.06.014

37. Han T, Liu C, Yang W, Jiang D (2019) A novel adversarial learning framework in deep convolutional neural network for intelligent diagnosis of mechanical faults. Knowl Based Syst 165, 474–487. https://doi.org/10.1016/j.knosys.2018.12.019

38. Zhang L, Lv Y, Huang WY, Yi CC (2022) Bearing fault diagnosis under various operation conditions using synchrosqueezing transform and improved two-dimensional convolutional neural network. Meas Sci Technol 33(8):085002. https://doi.org/10.1088/1361-6501/ac69b1

39. Lee CK, Shin YJ (2021) Detection and assessment of i&c cable faults using time-frequency r-cnn-based reflectometry. IEEE Trans Industr Electron 68(2):1581–1590. https://doi.org/10.1109/tie.2020.2970677

40. Zhu ZY, Peng GL, Chen YH, Gao HJ (2019) A convolutional neural network based on a capsule network with strong generalization for bearing fault diagnosis. Neurocomputing 323, 62–75. https://doi.org/10.1016/j.neucom.2018.09.050

41. Long M, Cao Y, Wang J, Jordan M. Learning transferable features with deep adaptation networks. In: International conference on machine learning, pp 97–105. PMLR

42. Sun B, Saenko K (2016) Deep coral: Correlation alignment for deep domain adaptation. In: 14th European conference on computer vision (ECCV). Lect Notes Comput Sci 9915, 443–450. https://doi.org/10.1007/978-3-319-49409-8_35

43. Long M, Zhu H, Wang J, Jordan MI (2017) Deep transfer learning with joint adaptation networks. In: 34th International conference on machine learning. Proc Mach Learn Res 70:2208–2217

44. Dai Y, Liu J, Ren X, Xu Z (2020) Adversarial training based multi-source unsupervised domain adaptation for sentiment analysis. Proc AAAI Conf Artif Intell 34, 7618–7625 . https://doi.org/10.1609/aaai.v34i05.6262