Universal source-free domain adaptation method for cross-domain fault diagnosis of machines

Yongchao Zhang[a,b], Zhaohui Ren[a], Ke Feng[b], Kun Yu[c], Michael Beer[d,e,f], Zheng Liu[b]

[a] School of Mechanical Engineering and Automation, Northeastern University, Shenyang, Liaoning 110819, China
[b] School of Engineering, University of British Columbia, Kelowna, British Columbia V1V 1V7, Canada
[c] School of Information and Control Engineering, China University of Mining and Technology, Xuzhou, Jiangsu 221116, China
[d] Institute for Risk and Reliability, Leibniz Universität Hannover, Callinstr. 34, Hannover 30167, Germany
[e] Institute for Risk and Uncertainty, University of Liverpool, Peach Street, Liverpool L69 7ZF, United Kingdom
[f] International Joint Research Center for Resilient Infrastructure and International Joint Research Center for Engineering Reliability and Stochastic Mechanics, Tongji University, Shanghai 200092, China

Abstract

Cross-domain machinery fault diagnosis aims to transfer enriched diagnosis knowledge from a labeled source domain to a new unlabeled target domain. Most existing methods assume that the prior information on the fault modes of the target domain is known in advance. However, in engineering practice, prior knowledge of fault modes is rare in a new domain, in which there may be only partial source fault modes or some new fault modes. Furthermore, up to the present, almost all existing cross-domain fault diagnosis methods require the labeled source data during the model training process, which restricts their deployment on certain devices with limited computing resources. To this end, we propose a universal source-free domain adaptation method that can handle cross-domain fault diagnosis scenarios without access to the source data and is free of explicit assumptions about the target fault modes. More specifically, we develop a convolutional network with a Transformer as the attention module to extract discriminative feature information from the source data and then send the model and parameters to the target domain. In target domain training, we first propose a supervised contrastive learning strategy based on source class prototypes, which utilizes high-confident predictions to achieve source-free domain alignment and class alignment. Then, we also introduce a threshold-based entropy max-min loss to further align known class samples in the target domain or reject target outlier samples as an unknown class. Furthermore, we introduce self-supervised learning to further learn feature representations of the target domain to reduce the previous misclassification. A series of experiments on two rotating machine datasets demonstrate the effectiveness and practicability of the proposed method.

Keywords: Fault diagnosis, machinery, source-free, domain adaptation, supervised contrastive learning

1. Introduction

The fourth industrial revolution requires more automated, complex, and computerized industrial infrastructures to meet the growing manufacturing requirements, which brings higher requirements to the fault diagnosis and structural health monitoring of industrial systems and their sub-components [1, 2, 3]. Rotating machinery, as a key component of industrial systems, usually operates under a non-stationary and harsh working environment during its lifespan, which makes it easily subject to various faults [4, 5]. Unexpected failures may affect the productivity and product quality of the manufacturing process and even cause production accidents. Therefore, effective fault diagnosis solutions for these machines in modern industrial processes are essential, which can guarantee operating

safety, reduce downtime, and boost economic profits [6, 7, 8].

In the past years, deep learning technology has played a predominant role in the field of machinery fault diagnosis due to its strong ability to extract features from large-scale data [9, 10, 11, 12]. However, effective deep learning-based models need two essential prerequisites: 1) a large amount of labeled training data and 2) training data and test data follow the same or similar distribution [13]. These two prerequisites are generally difficult to satisfy in practical industrial applications. The reason for this situation is that the various operating conditions of the machine, such as rotating speed and operating load, may cause the distribution discrepancy between the test and training data [14, 15, 16]. Also, annotating data in all operating conditions is time-consuming, costly, or challenging. Consequently, there is a strong incentive to transfer the diagnosis knowledge from a well-labeled operating condition (i.e., source domain) to a different unlabeled operating condition (i.e., target domain), which is referred to as the domain shift issue.

Domain adaptation (DA) techniques have been proposed to deal with the domain shift issue. The DA aims to transfer knowledge from a label-rich source domain to the label-scarce target domain. Recently, numerous DA methods have been developed to boost model performances in various cross-domain recognition tasks such as fault diagnosis and structural health monitoring [17]. Their goals are to reduce the distribution discrepancy in the task-specific layer to mitigate the impact of the domain shift, such as different operating conditions or different structures [18]. However, existing cross-domain fault diagnosis methods are performed under strict assumptions, limiting their applicability to many practical industrial scenarios.

Most of the existing methods assume that the target domain's label set (i.e., the number of fault modes) is known in advance [19]. Suppose that, $\mathcal{C}_s$ and $\mathcal{C}_t$ define the label sets of the source domain and target domain, respectively. Recently, three common DA settings have been extended to various fault diagnosis scenarios, e.g., closed-set DA ($\mathcal{C}_s = \mathcal{C}_t$), partial DA ($\mathcal{C}_t \subset \mathcal{C}_s$), and open-set DA ($\mathcal{C}_s \subset \mathcal{C}_t$). Although some good results have been achieved in the setting with overlapped fault modes, it is hard to generalize them to practical industrial scenarios [20]. More specifically, the prior knowledge of the target label set is untrustworthy since we may not know which fault mode will occur during the real-world diagnosis process in advance. Unexpected fault modes shift generally impairs diagnostic model performance. This scenario is defined as universal DA, which is a quite challenging cross-domain fault diagnosis problem. The other strict assumption is that traditional DA methods require the source domain data during the training process. However, the scale of the source domain dataset is usually large, and training a diagnostic model jointly with source and target domain data usually requires large computing resources, limiting their deployment on certain platforms, especially portable devices in real-world applications.

To address the issues discussed above, we propose a universal source-free domain adaptation method for cross-domain fault diagnosis. The highlights and contributions of our work are summarized as follows.

- The proposed method can deal with the challenging cross-domain fault diagnosis scenarios where the source data are unavailable, and target fault modes are unknown. The model trained on the pre-collected dataset can be deployed at resource-limited devices to monitor their real-time operating condition by utilizing the proposed method.

- We develop a convolutional network with a Transformer as the attention module to conduct cross-domain

fault diagnosis. A high-confident supervised contrastive learning strategy based on source class prototypes is developed in the proposed network to achieve domain-invariant learning. A threshold-based entropy max-min loss is built to align samples of known classes or reject target outlier samples as an unknown class. Moreover, a self-supervised learning loss is introduced to further improve the fault diagnosis performance of the model.

- Massive experiments on two laboratory datasets are conducted to show that the proposed method stably crosses various domain adaptation settings and achieves higher performance than the advanced domain adaptation methods.

This work shows that the proposed method is a promising tool to address cross-domain fault diagnosis in real-world industrial applications.

The remainder of this paper starts with the related work in Section 2. The proposed framework is illustrated in Section 3. The experimental verification is in Section 4. The conclusions are drawn in Section 5.

2. Related work

In recent years, various deep learning approaches have been successfully applied to monitor mechanical systems [21]. Thanks to its robust feature extraction ability and impressive recognition performance, the convolutional neural network (CNN) has been studied extensively in various fault diagnosis tasks [22, 23]. For example, Wen et al. [24] proposed a deep learning model based on CNN for end-to-end fault diagnosis, in which the raw signal was converted into the two-dimensional image signal directly as the input of the deep learning model. Li et al. [25] proposed a convolutional residual network for bearing fault diagnosis, in which the augmented data in the time domain was adopted to improve the generalization of the model. Wang et al. [26] proposed a conditional variational neural network for fault diagnosis of planetary gearbox, and the Fourier domain data of the recorded vibration signal was also used as input to the network. Jiang et al. [27] proposed an intelligent diagnosis method for gearboxes based on a multi-scale CNN, in which a multi-layer convolution-pooling combination and a multi-scale learning mechanism were utilized to obtain complementary and rich diagnostic knowledge directly from the raw signal. Guo et al. [28] proposed a multi-task convolutional neural network based on information fusion, in which a multi-task CNN with a dynamic training rate was established. The two-dimensional time-frequency information of the vibration signal is used as the multi-dimensional input feature of the convolutional neural network, to realize the two tasks of fault diagnosis and localization at the same time. Guan et al. [29] proposed a multi-sensor and multi-scale CNN model for bearing fault diagnosis, and the correlation and complementarity among multi-level information were used to improve diagnostic performance. The experimental results in the existing literature show that CNN-based methods are well suited for fault diagnosis tasks in industrial big data scenarios. However, an essential hypothesis of these methods is that the feature distribution of training data and test data are the same or similar; once they have feature distribution discrepancies, the performance of the model will be seriously impaired [30].

For the industrial domain shift issue, DA methods have been developed to train a high generalization model that can transfer knowledge from a label-rich source domain to a label-unknown target domain. Domain adaptation has been applied to several engineering contents such as structural health monitoring (SHM) and fault diagnosis.

For example, Gardner et al. [18] applied transfer learning for diagnosing the unseen structure, in which the damage information learned from other structures in the population was utilized to improve the classification performance. Ritto et al. [31] proposed a new digital-twin framework for detecting localized torsional friction, in which a physics-based torsional model was used to obtain training data, and the domain adaptation method was introduced to reduce feature distribution discrepancy under structural changes. Gardner et al. [2] proposed a kernelized Bayesian transfer learning method for population-based SHM, in which two classification tasks were investigated to demonstrate that the proposed method can handle the inconsistent feature space scenario. The above-discussed research promoted the development of structural health monitoring, such as the aircraft wing and the drill-string. However, these studies were developed to solve the domain adaptation problem where the source domain data is available. While this paper focuses on source-free domain adaptation and the research object is rotating machinery. Therefore, the methodologies of cross-domain machinery fault diagnosis will be reviewed and discussed in the following.

More broadly in the field of fault diagnosis, existing DA methods mainly include three kinds of scenario settings: 1) Closed-set DA problem. For example, Lu et al. [32] proposed a deep domain adaptation network based on a deep neural network for cross-domain fault diagnosis of rolling bearing and gearbox, and introduced the maximum mean discrepancy (MMD) loss to achieve domain-invariant learning of source and target domains. Yang et al. [33] proposed a multi-kernel DA network based on CNN for cross-domain fault diagnosis, in which a multi-kernel MMD loss was established to minimize the marginal distribution discrepancy between the two domains. Wang et al. [34] proposed a subdomain adaptive method to reduce both marginal and conditional distribution discrepancy in cross-domain fault diagnosis. Han et al. [35] proposed a deep adversarial CNN network for mechanical fault diagnosis, in which a domain discriminator was introduced to identify whether the sample is from the source domain or the target domain. The goal of the training was to fool the domain discriminator to achieve domain alignment. Zhang et al. [36] proposed a joint domain alignment and class alignment approach for cross-machine fault diagnosis, in which the classifier discrepancy loss and contrastive loss were integrated to reduce the intra-class distance and increase the inter-class distance in the feature space. Jiao et al. [37] developed a domain adaptation method for cross-domain fault diagnosis with the one-dimensional residual network, which simultaneously reduces the joint distribution and marginal distribution discrepancy by jointing MMD and adversarial learning. 2) Partial DA problem. For example, Jiao et al. [38] constructed a classifier inconsistency-based partial DA network for machinery intelligent diagnosis, in which two different classifiers were designed to select the source samples belonging to the target class and emphasize the network to train these samples. Li et al. [39] proposed a class-weighted adversarial neural network based on partial DA for cross-domain fault diagnosis, in which the samples of shared classes were positively transferred, and the samples of source outliers were ignored. Jia et al. [40] proposed a weighted subdomain adaptation network for partial cross-domain fault diagnosis and used a weighted local MMD loss for the fine-grained transferable information. Li et al. [41] developed a weighting adversarial DA network to adaptively identify and filter source samples that were not related to the target classes by implementing a weighting learning strategy. Deng et al. [42] presented a double-layer attention-based adversarial network, in which two attention matrices were used to guide the network to know which samples should be trained or ignored. 3) Open-set DA problem. For example, Zhang et al. [43] developed an open-set DA method based on adversarial learning for the bearing diagnosis in the scenario where new fault modes occur in the target domain. An instance weighting

mechanism was introduced to calculate the similarity of the known classes of source and target and to identify unknown classes of the target domain. Zhu et al. [44] presented a cross-domain open-set diagnosis method, in which adversarial learning was used to achieve domain alignment, and multiple auxiliary classifiers were designed to identify known class samples and unknown class samples from target instances. Feng et al. [45] proposed a globally localized multisource DA method to address cross-domain fault diagnosis in case of class shift, in which the Wasserstein discrepancy and accumulative higher order were introduced into model training to reduce the shift on domain-level and class-level synchronously. Zhao et al. [46] showed a dual adversarial learning-based open-set DA network, in which the separated adversarial learning was used as a decision hyperplane to separate known and unknown classes.

The good results achieved by the aforementioned methods are based on the assumption that the target domain's label set is known in advance. However, it is usually unknown which fault modes have occurred in the unlabeled target domain in the real-world industry. This indicates that the testing target domain may have some known source fault modes and unknown fault modes. Unexpected fault mode shift causes the negative transfer, which reduces the model's performance. To our best knowledge, investigations on this practical universal DA cross-domain fault diagnosis problem are rare. A pioneer work is that a hybrid weighting mechanism method based on adversarial learning was proposed by Zhang et al. [20] to achieve cross-domain fault diagnosis in universal DA settings. However, this researcher has not considered DA fault diagnosis in the absence of source data due to computing resource problems.

In industry, with the rapid development of the industrial internet of things (IIoT), rapidly growing data volumes put pressure on servers with limited storage and computing resources. Edge computing is an emerging computing paradigm that relieves the pressure of centralized computing by distributing computing to each edge. Therefore, performing cross-domain fault diagnosis on the device side is the development trend of modern condition monitoring. However, the source domain datasets are usually large, and training a diagnostic model jointly with source and target domain data usually consumes large computing resources, which limits their deployment on certain platforms, especially portable devices.

In the computer vision field, some source-free domain adaptation methods were proposed to deal with the problem that the source data are not available during target adaptation. Source-free DA method means the adaptation step is source-free. The process of source-free DA is similar to that of the fine-tuning-based DA method, but they are different. In the fine-tuning-based DA method, a small amount of target data is used to update the pre-trained source model, thereby improving the performance of the model to the target domain [47]. In this training process, the target domain data with labels is required. Different from the fine-tuning-based DA method, the source-free DA method does not need to obtain the target domain label in advance during the target domain training process. For instance, Kim et al. [48] adopted the pre-trained model from the source domain to identify the target domain by further pseudo-labeling training in the target domain, which achieved higher performance than the traditional DA methods even if source domain data is unavailable. Liang et al. [49] proposed a Source HypOthesis Transfer approach, which updated the source feature extraction module by maximizing mutual information and self-supervised pseudo-labeling to achieve source data-absent DA. Ahmed et al. [50] proposed a novel source-free DA method, which can learn different weights for multiple source models to get the best target model. Recently, few initial attempts have been made on the universal source-free DA problem in computer vision scenarios. Kundu

et al. [51] proposed a universal source-free DA method based on a two-stage learning process, in which a novel instance-level weighting mechanism was defined to obtain an effective source-free adaptation objective. In the machine fault diagnosis field, source-free DA scenario, especially universal source-free DA scenario, widely exists in real-world industries. However, limited studies on machine fault diagnosis based on universal source-free DA can be found in the current literature. This study seeks to fill this gap for practical industrial applications.

Hence, we propose a universal source-free domain adaptation method, which can solve not only the universal DA problem but also the source-free DA problem. The proposed method can provide a promising perspective on a real-world industrial diagnosis.

3. Methodology

This section will introduce the proposed universal source-free domain adaptation method, including problem definition, proposed network structure, details of the training process, and the diagnosis procedure.

3.1. Problem definition

As a domain adaptation problem, let $\mathcal{D}^s = \{x_i^s, y_i^s\}_{i=1}^{n_s}$ denotes the labeled source domain consisting of n_s samples, where x^s and y^s represent the source domain samples and corresponding labels, respectively. Let $\mathcal{D}^t = \{x_i^t\}_{i=1}^{n_t}$ denotes the target domain consisting of n_t samples, where x^t represents target domain samples. The label sets of the source and target domains are defined as $\mathcal{C}_s$ and $\mathcal{C}_t$, respectively. This work mainly focuses on a more practical diagnostic scenario, i.e., $\overline{\mathcal{C}}_t = \mathcal{C}_t \backslash \mathcal{C}_s \neq \varnothing$, where $\overline{\mathcal{C}}_t$ represents target-private classes. The $\mathcal{C} = \mathcal{C}_s \cap \mathcal{C}_t$ represents the classes existing in both domains. In this work, the $\mathcal{D}^s$ and the $\mathcal{D}^t$ are used to train a model to classify the $\mathcal{D}^t$ into $\mathcal{C} + 1$ classes, where private samples are classified as "unknown" class. As a source-free domain adaptation method, the labeled $\mathcal{D}^s$ is only used to train a source model, and then the unlabeled $\mathcal{D}^t$ is used to train the well-trained source model to achieve domain alignment without any source instances.

3.2. Model architecture

In the proposed universal source-free DA method, we develop a convolution-based backbone network in which a Transformer is embedded as an attention module. The architecture and training steps of the proposed model are shown in Fig. 1. The model architecture consists of a feature extractor *G* for extracting high-level feature vectors and a fully connected classifier *C* for predicting labels. The training process of the proposed model consists of two steps, namely source domain training and target domain training. The model and parameters trained in the source domain are used as initialization for training in the target domain.

The network details of feature extractor *G* are shown in Fig. 2, which includes four one-dimensional convolutions (CNN) modules, a Transformer (T) module, and a fully connected (FC) module. Each convolution module includes a convolutional layer, a leaky rectified linear units activation function (LReLU) layer, a batch normalization (BN) layer, and a maximum pooling layer. The first convolution module receives one-dimensional raw vibration signal $x \in \mathbb{R}^L$ as model input, and the output after four convolution modules is expressed as $x \in \mathbb{R}^{C \times M}$, where *C* is the number of channels and *M* is the feature dimension of each channel. The local characteristics of the vibration signal are acquired using the CNN module. After the fourth convolution module, we embed a Transformer as an attention module to make the model focus on the global correlation information of the vibration

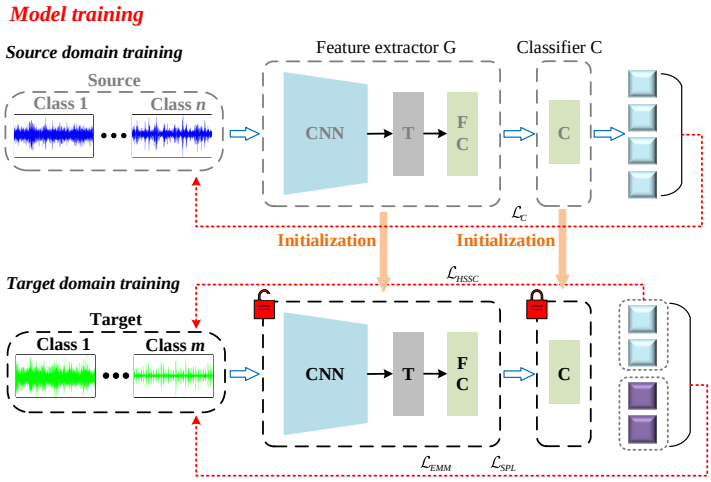


Figure 1: The architecture and training steps of the proposed model.

signal. Transformers with remarkable global representation capacities achieve competitive results for classification tasks [52]. By combining global and local feature representations, the generalization performance of the model for diagnosis under variable operation conditions is improved.

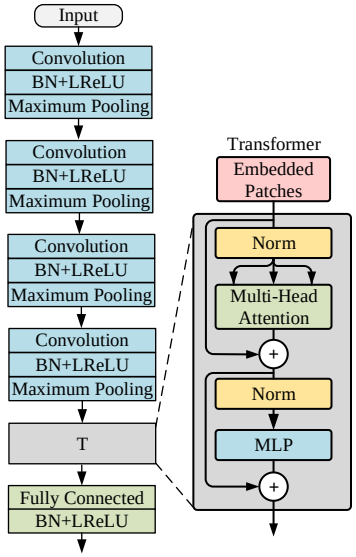


Figure 2: The network structure of the feature extractor.

The proposed Transformer module follows the original structure of Vision Transformer (ViT) [52]. ViT is the first example where the Transformer-based method outperforms CNN on image classification tasks. In our network, the feature $x \in \mathbb{R}^{C \times M}$ extracted by the convolution module is first converted to a sequence of patches $x_p \in \mathbb{R}^{C \times M}$, where C is the number of patches, M is the resolution of each patch. Similar to the class (CLS) token in BERT [52], a randomly trainable embedding CLS token is added to the sequence. In addition, since the position of self-attention in Transformer encoders is uncertain and classification tasks generally require position information, ViT embeds the position into each token, including CLS token. All tokens are passed through stacked Transformer encoders, and finally CLS tokens are used for feature representation. The output of ViT can be represented as:

$$
\begin{aligned}
z_0 &= \left[x_{\mathrm{CLS}}; x_p^1\mathrm{E}; x_p^2\mathrm{E}; \cdots; x_p^C\mathrm{E}\right] + \mathrm{E}_{pos} \\
z_l' &= \mathrm{MSA}\left(\mathrm{LayerNorm}\left(z_{l-1}\right)\right) + z_{l-1} \\
z_l &= \mathrm{FFN}\left(\mathrm{LayerNorm}\left(z_l'\right)\right) + z_l' \\
h &= z_L^0
\end{aligned}
\tag{1}
$$

where $x_{\mathrm{CLS}} \in \mathbb{R}^{1\times M}$, $\mathrm{E} \in \mathbb{R}^{C\times M}$, $\mathrm{E}_{pos} \in \mathbb{R}^{(C+1)\times M}$ is the position embedding, $\ell = 1, 2, \ldots, L$ denotes the *i*-th Transformer block, MSA denotes the multi-head self-attention mechanism, FFN denotes the feed-forward network. And then, one fully connected layer is introduced to obtain high-dimensional feature representations, which is expressed as: $x \in \mathbb{R}^M \rightarrow \mathbb{R}^D$, where *D* represents the high-level feature dimension. In the fully connected layer, the dropout operation is adopted to avoid over-fitting and improve the network's generalization ability.

Finally, classifier *C* consists of two fully connected layers to output prediction result $\hat{y} = C(G(x))$, which is expressed as: $x \in \mathbb{R}^D \rightarrow \mathbb{R}^K$, where *K* represents the number of health conditions of the source domain.

3.3. Model training in the source domain

The overall training process of the model is shown in Fig. 1. First, the cross-entropy loss is used on the source domain samples to learn a diagnostic model. The cross-entropy loss in the source domain can be expressed as:

$$
\mathcal{L}_C = -\mathbb{E}_{(x^s, y^s)\in\mathcal{D}^s}\left[\sum_{k=1}^{K} \mathbb{1}_{[k=y^s]} \log\left(\hat{y}^s\right)\right] \tag{2}
$$

where $\hat{y}^s$ denotes the class predictions, *K* denotes the number of health conditions of the source domain.

The $\mathcal{L}_C$ is minimized to seek optimal θ_G and θ_C in the training process, and this process can be described as:

$$
\left(\hat{\theta}_G, \hat{\theta}_C\right) = \arg\min_{\theta_G, \theta_C} \mathcal{L}_C \tag{3}
$$

where θ_G and θ_C represent the parameters of *G* and *C*, respectively.

3.4. Model training in the target domain

In the target domain, we further train the pre-trained source model by using the unlabeled target domain data. The model is parameterized by a neural network with two modules: feature extractor *G*: $x \in \mathbb{R}^L \rightarrow \mathbb{R}^D$ and classifier *C*: $x \in \mathbb{R}^D \rightarrow \mathbb{R}^K$, i.e., $\hat{y} = C(G(x))$. Most previous domain adaptation approaches use statistical moment matching or adversarial training to confuse the feature distribution in the shared feature space $x \in \mathbb{R}^D$ of two domains. However, those approaches are not applicable under source-free domain adaptation scenarios. Therefore, we design a unique strategy that not only reduces domain discrepancy but also identifies unknown class samples in the target domain.

3.4.1. High-confident source-free supervised contrastive loss

In the training stage of the target domain, we can get a pre-trained source model and the unlabeled target data. In the source classifier, the features before the classifier are multiplied by a weight matrix $\boldsymbol{W}^S \in \mathbb{R}^{D\times K}$, $\boldsymbol{W}^S = \left[\boldsymbol{w}_1^S, \boldsymbol{w}_2^S, \ldots, \boldsymbol{w}_K^S\right]$ to get the output of the model. For an input feature $G(x_i)$, the predicted probability can be expressed as a cosine similarity score $[s_{i,1}, s_{i,2}, \cdots, s_{i,K}]$ for the $G(x_i)$ and each weight vector, the

$s_{i,j} = \sigma(G(x_i)\boldsymbol{w}_j / \|G(x_i)\| \|\boldsymbol{w}_j\|)$ where σ is softmax function. For a well-trained model, the classifier can obtain predicted results by the cosine distance of the input features to the weight vector representing each class. Therefore, the well-trained weight vectors $\left[\boldsymbol{w}_1^S, \boldsymbol{w}_2^S, \ldots, \boldsymbol{w}_K^S\right]$ in source domain can be used as the class prototypes. To learn domain-invariant feature representations of source and target domains, we establish a supervised source-free contrastive loss based on class prototypes. For cross-domain fault diagnosis tasks, the sample features of the same class are generally close to each other while the sample features of the different classes are generally far away no matter which domain they come from. Based on this hypothesis, pulling together the same-label examples in the shared feature space and pushing apart the different-label examples can reduce domain differences.

For source-free domain adaptation scenarios, source sample instances are not available. We replace the source domain instances with the source class prototypes. We aim to pull the target predicted class and the corresponding source class prototype together and separate the target predicted class from other different class prototypes. Based on InfoNCE loss [53], we propose a supervised source-free contrastive loss by using source class prototypes and target instances. We freeze the *C*'s parameters and only train the *G* to maintain the source class prototypes while training in the target domain. The proposed source-free supervised contrastive (SSC) loss can be described as:

$$\mathcal{L}_{SSC}(g) = -\sum_{i \in I} \log \frac{\exp\left(g_i \cdot \boldsymbol{w}_j / \tau\right)}{\sum_{k \in K} \exp\left(g_i \cdot \boldsymbol{w}_k / \tau\right)} \tag{4}$$

where the $\cdot$ symbol indicates the inner dot product, index *i* is called the anchor and $i \in I \equiv \{1, 2, \ldots, \text{Batchsize}\}$,$\boldsymbol{w}_j$ is the positive class prototype for the anchor, $\boldsymbol{w}_k$ includes $\boldsymbol{w}_j$ and all negative class prototypes, and τ is a temperature hyper-parameter.

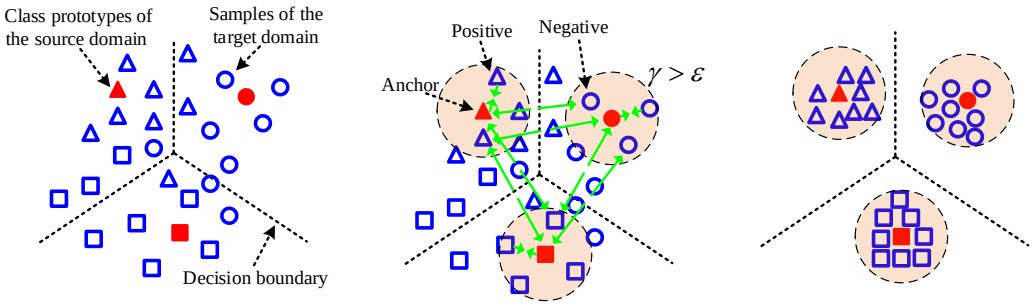
The SSC loss utilizes labeled information of the target domain. The incorrect target pseudo-labels may damage the performance of the SSC loss. The correctness of pseudo-labels in universal domain adaptation scenarios is impacted by two factors: domain distribution discrepancy and label shift. To mitigate the impact of false pseudo-labels, a predictive output probability threshold ε is defined into the source-free supervised contrastive loss, which utilizes the high-confident prediction samples to achieve domain alignment and known class alignment. We define the maximum probability value in the prediction output vector as γ, that is, $\gamma = \max\left[\hat{y}_{(1)}, \hat{y}_{(2)}, \ldots, \hat{y}_{(K)}\right]$. Thus, a larger γ indicates higher confidence in the prediction. Our idea is presented visually as Fig. 3. It can be seen from Fig. 3 (a) that using the source model directly to identify the target domain will generate many misclassified samples at the decision boundary. Once these misclassified samples are used for source-free supervised contrastive learning, the performance of the model will seriously deteriorate. Thus, we establish a high-confident source-free supervised contrastive learning process as shown in Fig. 3 (b), in which high-confident target prediction samples are selected for contrastive learning to prevent model performance from being negatively impacted by false pseudo-labels. In this way, source-free domain adaptation based on class alignment as shown in Fig. 3 (c) is achieved. The formulation of the proposed high-confident source-free supervised contrastive loss can be expressed as:

$$\begin{gathered} \mathcal{L}_{HSSC} = \sum_{i \in I} \mathcal{L}_{HSSC}\left(g_i\right), \\ \mathcal{L}_{HSSC}\left(g_i\right) = \begin{cases} -\log \frac{\exp(g_i \cdot \boldsymbol{w}_j / \tau)}{\sum_{k \in K} \exp(g_i \cdot \boldsymbol{w}_k / \tau)} & \gamma > \varepsilon, \\ 0 & \text{otherwise.} \end{cases} \end{gathered} \tag{5}$$

where ε indicates the threshold. In this way, the training process of the source-free supervised contrastive loss can

be expressed as:

$$\left(\hat{\theta}_G\right) = \arg\min_{\theta_G} \mathcal{L}_{HSSC} \tag{6}$$



(a) Initial state (b) Supervised contrastive training (c) Training completed

Figure 3: chematic diagram of the proposed high-confident source-free supervised contrastive learning.

3.4.2. Threshold-based entropy max-min loss

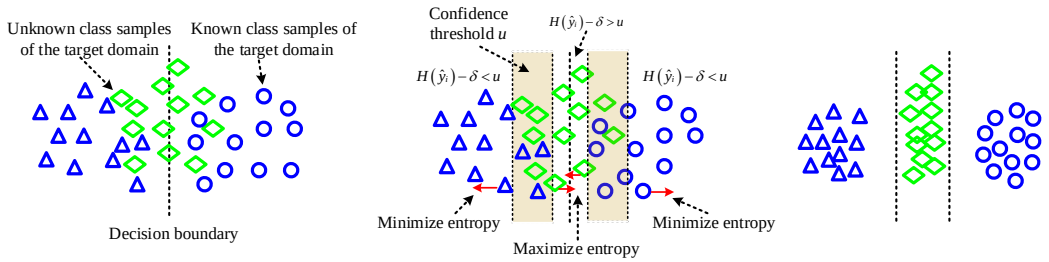
Source-free supervised contrastive loss encourages the high-confident prediction samples to be close to the same class prototypes, but there are low-confidence prediction samples or unknown class samples in the target domain, these samples are usually distributed in the decision-boundary and affect the diagnostic performance of the model. Thus, we need an explicit objective function to encourage alignment or reject these samples. Since the target samples of the unknown classes have no common features with the known source samples, the uncertainty of the predicted output of these unknown class samples is greater, that is, the predicted output has a larger entropy value than the known class samples.

Inspired by this, we introduce threshold-based entropy max-min loss. For a given threshold δ in advance, when the prediction entropy of the target sample is less than the δ, the sample is regarded as a known class sample. At this time, minimizing the prediction entropy can further reduce the feature distribution discrepancy between the source and target domains. Besides, when the prediction entropy of the target sample is greater than the δ, the sample is regarded as an unknown class sample. At this time, maximizing the prediction entropy can keep the unknown class samples away from the decision boundary and improve the recognition accuracy of unknown class samples. Since the maximum value of $H\left(\hat{y}_i\right)$ is $\log(K)$, δ is empirically defined as $\delta = \frac{\log(K)}{2}$, where K denotes the number of source health conditions [54]. The threshold may be ambiguous due to domain shift or sample label space shift. To improve the threshold confidence, a confidence threshold u is introduced. This idea is presented visually as Fig. 4. It can be seen from Fig. 4 (a) that a large number of unknown class samples of the target domain are distributed at the decision boundary, resulting in the low diagnostic performance of the model. The source-free supervised contrastive loss proposed above cannot effectively identify unknown class samples of the target domain. Thus, we establish a threshold-based entropy max-min loss as shown in Fig. 4 (b) to further train the proposed model. As shown in Fig. 4 (b) and Fig. 4 (c), optimizing this loss allows the classifier to reject samples of unknown classes in the target domain and reduce the feature distance between samples of the same class. In this way, the cross-domain diagnostic performance of the model is further improved under the source-free scenario. The threshold-based entropy max-min loss can be expressed as:

$$\mathcal{L}_{EMM} = \sum_{i \in I} \mathcal{L}_{EMM}\left(\hat{y}_i\right),$$
$$\mathcal{L}_{EMM}\left(\hat{y}_i\right) = \begin{cases} -\left|H\left(\hat{y}_i\right) - \delta\right| & \left|H\left(\hat{y}_i\right) - \delta\right| > u, \\ 0 & \text{otherwise.} \end{cases} \tag{7}$$

where $H\left(\hat{y}_i\right)$ indicates the prediction entropy. By maximizing $\left|H\left(\hat{y}_i\right) - \delta\right|$, we can simultaneously maximize and minimize entropy with the threshold as the boundary. Therefore, in the process of model training, minimizing $\mathcal{L}_{EMM}$ can increase the decision-making confidence of the network about known class samples and unknown class samples. During model testing, samples whose predicted entropy satisfies $\left|H\left(\hat{y}_i\right) - \delta\right| > u$ are regarded as unknown class samples. Thus, the training process of entropy max-min loss can be described as:

$$\left(\hat{\theta}_G\right) = \arg\min_{\theta_G} \mathcal{L}_{EMM} \tag{8}$$



(a) Initial state (b) Entropy max-min training (c) Training completed

Figure 4: Schematic diagram of the proposed entropy max-min learning.

3.4.3. Self-supervised pseudo-labeling learning

Although the source-free supervised contrastive loss and entropy max-min loss can reduce domain discrepancy and make target prediction results more confident and globally diverse, it is unavoidable that some target samples are matched to non-corresponding class prototypes due to wrong pseudo-labels. These mismatched samples have harmful effects on the training of the model. To alleviate this problem, one solution is to implement a self-labeling strategy of the target domain by learning the semantic representations. In this work, we introduce self-supervised pseudo-labeling learning [49] to further constrain the model. Specifically, we first compute the class prototypes for the target domain by weighted k-means clustering:

$$c_k^{(0)} = \frac{\sum_{x^t \in \mathcal{D}^t} \sigma_k\left(C\left(G\left(x^t\right)\right)\right) G\left(x^t\right)}{\sum_{x^t \in \mathcal{D}^t} \sigma_k\left(C\left(G\left(x^t\right)\right)\right)} \tag{9}$$

where $\sigma_k(\cdot)$ indicates the k-th element of the predicted softmax output.

Then, we generate initial pseudo-labels by the nearest centroid classifier:

$$\hat{y}^t = \arg\min_k 1 - \frac{G\left(x^t\right) c_k^{(0)}}{\left\|G\left(x^t\right)\right\|_2 \left\|c_k^{(0)}\right\|_2} \tag{10}$$

where $\| * \|_2$ indicates the $L2$-norm. Finally, we update the class centroids and pseudo labels:

$$
\begin{aligned}
c_k^{(1)} &= \frac{\sum_{x^t \in \mathcal{D}^t} \mathbb{1}\left(\hat{y}^t = k\right) G\left(x^t\right)}{\sum_{x^t \in \mathcal{D}^t} \mathbb{1}\left(\hat{y}^t = k\right)}, \\
\hat{y}^t &= \arg\min_k 1 - \frac{G\left(x^t\right) c_k^{(1)}}{\left\|G\left(x^t\right)\right\|_2 \left\|c_k^{(1)}\right\|_2}.
\end{aligned} \tag{11}
$$

Since the above process is trained in an unsupervised manner, the $\hat{y}^t$ is called self-supervised pseudo-labeling. Although the centroids and pseudo-labels can be trained multiple times, we find that good enough pseudo-labels can be obtained with just one round of training. The cross-entropy loss is used to calculate the pseudo-labeling loss. Thus, the self-supervised pseudo-labeling loss can be expressed as:

$$
\mathcal{L}_{SPL} = -\mathbb{E}_{(x^t) \in \mathcal{D}^t} \left[\sum_{k=1}^{K} \mathbb{1}_{[k=y^t]} \log \sigma_k \left(\hat{y}^t\right) \right] \tag{12}
$$

The training process of the self-supervised pseudo-labeling loss can be described as:

$$
\left(\hat{\theta}_G\right) = \arg\min_{\theta_G} \mathcal{L}_{SPL} \tag{13}
$$

3.4.4. Total loss in the target domain training

The total loss of model training in the target domain is expressed as:

$$
\mathcal{L}_{total} = \mathcal{L}_{HSSC} + \lambda \mathcal{L}_{EMM} + \beta \mathcal{L}_{SPL} \tag{14}
$$

where λ and β indicate tradeoff parameters.

3.5. Cross-domain diagnosis framework based on the proposed method.

In this work, a universal source-free cross-domain fault diagnosis framework, as shown in Fig 5, is established, and its detailed procedures are as follows.

- Data sampling. Collect vibration signals of faults under various operating conditions on the experimental platform.
- Data preprocessing. Set the fault classes of the source and target domains to simulate different domain adaptation scenarios. And divide the samples according to the input data requirements of the model.
- Model initialization. Perform random parameter initialization for the proposed network structure.
- Model training in the source domain. Train the model using source domain labeled data to achieve model diagnostic performance. The objective function is expressed as Eq. 2.
- Model training in the target domain. The well-trained source model is further trained by using target domain unlabeled data to achieve domain-invariant of features and identify unknown class samples. The objective function is expressed as Eq. 14.

- Performance evaluation. Use test data to comprehensively evaluate the cross-domain fault diagnosis performance of the model. During model testing, samples whose predicted entropy satisfies $|H(\hat{y}_i) - \delta| > u$ are regarded as unknown class samples. The target diagnosis accuracy is used as an evaluation metric, which is defined as: $accuracy = \sum_{i=1}^{N} \psi(y_i^t, \hat{y}_i^t)/N$

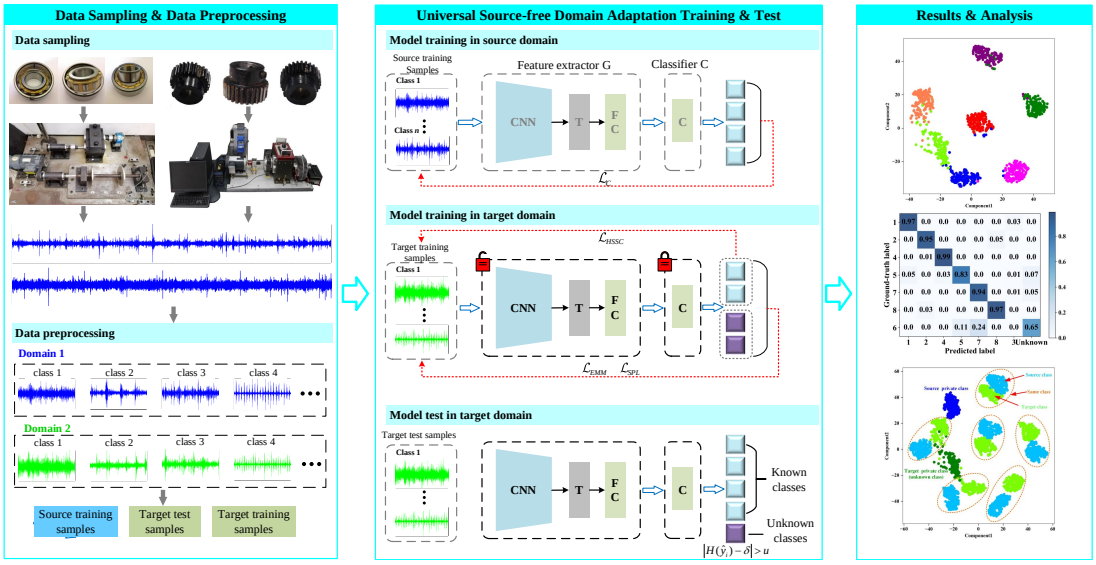


Figure 5: Universal source-free cross-domain fault diagnosis framework based on the proposed method.

4. Experiments

4.1. Experimental design and setting

4.1.1. Dataset description

As an important and failure-prone mechanical component in the industry, the gear and bearing systems are selected to evaluate the performance of the proposed method.

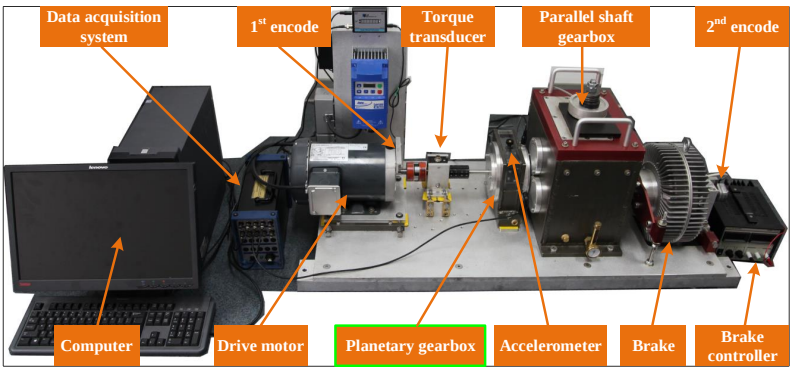


Figure 6: The DDS test rig.

1) The DDS test rig dataset. To demonstrate the effectiveness of the proposed method, the first experimental case is implemented on a gearbox platform, namely Dynamic Drivetrain Simulator (DDS) test rig. As shown in Fig. 6, the DDS test rig mainly includes a drive motor, a one-stage planetary gearbox, a parallel shaft gearbox, and a magnetic brake. In this work, various faults of the planetary gearbox are simulated. Specifically, the planetary

gearbox is composed of a 28-tooth sun gear, four 36-tooth planet gears, and a 100-tooth ring gear. An accelerometer was vertically mounted on the housing to collect the vibration signal with a sampling frequency of 30720 Hz. Eight different health conditions are studied, namely sun gear tooth missing (STM), sun gear tooth broken (STB), sun gear wearing (SW), sun gear crack (SC), planet gear tooth missing (PTM), planet gear tooth broken (PTB), planet gear crack (PC), and healthy (H). The actual health condition pictures are presented in Fig. 7. Three rotating speeds are conducted as 1800, 2400, and 3000 r/min to simulate three different domains. In summary, this DDS dataset consists of condition monitoring data collected in three domains for eight health conditions.

Figure 7: The different health conditions of the sun gear and planet gear.

2) The bearing test rig dataset. A practical rolling bearing test rig is used to further verify the performance of the proposed method. The test rig is supported by two rolling bearings, and the right bearing is used as the test bearing, which is shown in Fig. 8. The accelerometers attached to the bearing house collect the vibration signal with a sampling frequency of 20 kHz. The inner-race fault (IF), ball fault (BF), outer-race fault (OF), compound fault (OF-BF), and healthy (H) are manually customized. Similarly, three rotating speeds are conducted as 600, 1200, and 1800 r/min to simulate three different domains. Thus, this dataset consists of five health conditions data collected in three domains.

The detailed label information of the DDS test rig dataset and the bearing test rig dataset is shown in Table 1.

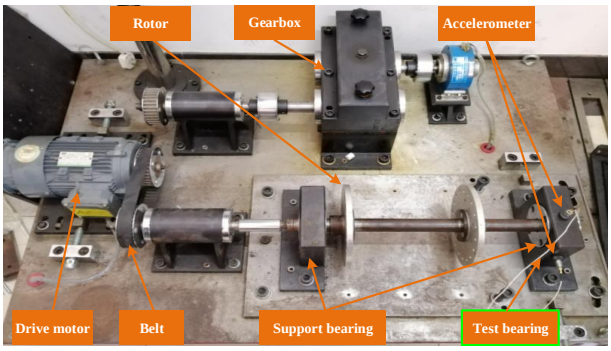


Figure 8: The bearing test rig.

Table 1: The health conditions and corresponding labels of the two datasets.

Dataset	Label	1	2	3	4	5	6	7	8
Gearbox	Health condition	STM	STB	SW	SC	PTM	PTB	PC	H
Bearing	Health condition	IF	BF	OF	OF-BF	H			

4.1.2. Experimental settings

1) Domain adaptation tasks. In this work, universal source-free DA tasks are investigated using the two datasets. At the same time, the closed-set DA task, the partial DA task, and the open-set DA task are also designed to evaluate the general performance of the proposed method. In the source domain and the target domain, we randomly selected different health conditions to form different DA tasks, in which the outliers of the target domain are regarded as an unknown class. In the DDS test rig dataset, the vibration signal of each health condition is divided into 500 samples. And the number of training samples, test samples, and verification samples for each health condition are 200, 150, and 150 respectively. In the bearing test rig dataset, the vibration signal of each health condition is divided into 700 samples. And the number of training samples, test samples, and verification samples for each health condition are 400, 150, and 150 respectively. The information of the DA tasks on the two datasets is shown in Table 2 and Table 3, respectively.

Table 2: The information of the DA tasks on the DDS test rig dataset.

Task	Source (r/min)	Target (r/min)	Source condition labels	Source training samples	Target condition labels	Target training samples	Target verification samples	Target test samples	DA problem	Source outlier conditions	Target outlier conditions
A_1	2400	3000	1,2,3,4,5,6,7,8	1600	1,2,3,4,5,6,7,8	1600	1200	1200	Closed set	0	0
A_2	1800	3000	1,2,3,4,5,6,7,8	1600	1,2,3,4,5,6,7,8	1600	1200	1200	Closed set	0	0
A_3	3000	1800	1,2,3,4,5,6,7,8	1600	1,2,3,4,5,6,8	1400	1050	1050	Partial	1	0
A_4	3000	2400	1,2,3,4,5,6,7,8	1600	1,2,3,4,6,8	1200	900	900	Partial	2	0
A_5	3000	2400	1,2,3,4,5,6,8	1400	1,2,3,4,5,6,7,8	1600	1200	1200	Open set	0	1
A_6	3000	1800	1,2,3,4,6,8	1200	1,2,3,4,5,6,7,8	1600	1200	1200	Open set	0	2
A_7	2400	3000	1,2,3,4,5,7,8	1400	1,2,4,5,6,7,8	1400	1050	1050	Universal	1	1
A_8	1800	3000	1,2,4,5,6,7,8	1400	1,2,3,4,6,7,8	1400	1050	1050	Universal	1	1
A_9	3000	2400	1,2,4,5,7,8	1200	1,2,3,5,6,7,8	1400	1050	1050	Universal	1	2
A_{10}	3000	1800	2,3,4,6,7,8	1200	1,2,3,4,5,7,8	1400	1050	1050	Universal	1	2
A_{11}	2400	3000	1,2,3,6,8	1000	1,3,4,5,6,7,8	1400	1050	1050	Universal	1	3
A_{12}	3000	1800	1,3,4,7,8	1000	1,2,3,4,5,6,8	1400	1050	1050	Universal	1	3

Table 3: The information of the DA tasks on the bearing test rig dataset.

Task	Source (r/min)	Target (r/min)	Source condition labels	Source training samples	Target condition labels	Target training samples	Target verification samples	Target test samples	DA problem	Source outlier conditions	Target outlier conditions
B_1	600	1200	1,2,3,4,5	2000	1,2,3,4,5	2000	750	750	Closed set	0	0
B_2	600	1800	1,2,3,4,5	2000	1,2,3,4,5	2000	750	750	Closed set	0	0
B_3	1800	600	1,2,3,4,5	2000	1,2,3,5	1600	600	600	Partial	1	0
B_4	1800	1200	1,2,3,4,5	2000	2,3,5	1200	450	450	Partial	2	0
B_5	1800	1200	1,2,3,5	1600	1,2,3,4,5	2000	750	750	Open set	0	1
B_6	1800	600	1,3,5	1200	1,2,3,4,5	2000	750	750	Open set	0	2
B_7	600	1200	1,2,3,5	1600	1,3,4,5	1600	600	600	Universal	1	1
B_8	1200	600	2,3,4,5	1600	1,2,3,5	1600	600	600	Universal	1	1
B_9	600	1800	1,2,3,5	1600	1,3,4,5	1600	600	600	Universal	1	1
B_{10}	600	1200	1,2,5	1200	2,3,4,5	1600	600	600	Universal	1	2
B_{11}	1800	1200	2,3,5	1200	1,3,4,5	1600	600	600	Universal	1	2
B_{12}	1800	600	1,3,5	1200	2,3,4,5	1600	600	600	Universal	1	2

2) Parameter settings. In this work, the network architecture and parameters are mostly determined by experiments on the task A_7 of the DDS test rig dataset. During the process of model training, one-dimensional vibration signal samples are directly fed into the model, and the size of each sample is 1×2048. The network architecture consists of four convolutional modules, a Transformer, a fully connected module, and a fully connected classifier. In the four convolutional layers, the kernel sizes are 9, the output channels are 4, 16, 64, and 128, respectively. The kernel sizes and strides for all maximum pooling layers are 2, and 2, respectively. For the Transformer mod-

ule, the number of Transformer blocks is 4, the number of attention heads is 5, and the hidden dimension is 122. The output channels for the fully connected module are 50. For the classifier, the output channels for two fully connected layers are 20 and *K*, where *K* is the number of health conditions of the source domain. Each method is performed 8 times, and its average value is taken as the final diagnosis result. After 200 epochs, the initial learning rate values of source training and target training are reduced to 10% of their initial values, respectively. The other related parameters are provided in Table 4.

Table 4: Main parameter settings.

Parameter	Value	Parameter	Value
Batch size	64	Optimizer	SGD
Source initial learning rate	0.01	λ	0.5
Target initial learning rate	0.001	β	0.5
Dropout	0.1	ε	0.99
Epochs	400	u	0.2

3) Comparative methods. In this work, we perform six comparative methods to demonstrate the superiority of the proposed method. The application details of the these methods are described as follows. To fairly highlight the advantages of the proposed method, the network structures and experimental configurations of all comparative methods are consistent with those of the proposed method. Also, the parameter optimization process of all comparative methods were performed on task A_7 to obtain high-performance model parameters.

- OSM. A baseline deep neural network based on source label data is performed for comparison, which is named only source model (OSM). Specifically, the source model trained by Eq. 2 is directly used for target domain diagnosis.

- MK-MMD. The state-of-the-art source-available DA methods are evaluated in this work. First, a multi-kernel MMD method (MK-MMD) [55] is introduced for comparison. In this method, MMD loss is added in the last three fully connected layers for domain-invariant learning.

- ALGR. According to the literature [35], [56], an adversarial learning method based on gradient reversal (ALGR) is introduced as the second source-available method for comparison, in which a domain discriminator is integrated into the shared fully connected layer after the Transformer module for domain-invariant learning. This domain discriminator consists of two fully connected layers whose output channels are 10 and 1, respectively.

- SAN. The selective adversarial network (SAN) [57] is introduced as the third source-available method for comparison. SAN can handle partial DA problems by introducing different domain discriminators. The domain discriminator structure and parameters of SAN are consistent with those of the ALGR.

- SHOT. Then, we compare the proposed method with a source-free DA method, namely source-free hypothesis transfer and labeling transfer method (SHOT) [49]. The SHOT is a popular source-free DA method for closed-set DA and open-set DA, in which information maximization and self-supervised learning are introduced into network training to achieve domain-invariant learning.

- USF. Finally, we compare the proposed method with a universal source-free domain adaptation method (USF) [51]. The USF is a state-of-the-art universal source-Free DA method, in which a novel instance-level weighting mechanism is defined into the target training process to achieve domain-invariant learning.

4.2. Experimental results and visualization analysis

4.2.1. Experimental results

Table 5: The statistic of the accuracy and standard deviation for the DDS test rig dataset (%).

Task	OSM	MK-MMD	ALGR	SAN	SHOT	USF	Proposed
A_1	66.38±3.021	96.12±0.369	97.05±0.459	95.70±0.687	80.24±2.231	81.44±1.860	95.32±0.353
A_2	61.85±3.440	92.47±0.513	91.78±0.733	91.74±0.956	70.05±3.247	74.96±2.429	90.77±0.780
A_3	62.05±3.376	87.96±0.993	88.73±1.055	94.14±0.783	72.95±3.014	80.53±1.661	91.56±0.742
A_4	68.83±2.745	92.45±0.732	93.01±1.121	97.23±0.784	82.59±2.206	85.10±2.068	95.92±0.579
A_5	63.56±3.587	74.33±1.951	73.57±2.325	78.26±2.058	76.35±2.893	77.73±2.215	92.35±0.898
A_6	59.37±3.444	70.19±2.346	69.78±2.597	73.34±2.219	72.07±3.221	73.86±2.344	88.76±0.993
A_7	60.32±2.598	71.52±2.136	68.51±2.692	70.36±2.367	73.56±3.156	80.21±1.874	90.06±0.845
A_8	57.48±2.884	65.02±2.340	63.33±2.779	67.44±2.657	67.03±3.446	74.96±2.377	85.51±1.240
A_9	52.61±3.985	53.32±2.256	51.26±2.894	50.26±3.024	58.36±3.256	70.51±2.661	84.23±1.059
A_{10}	46.34±4.133	50.24±2.447	51.67±3.330	50.37±3.115	54.01±3.241	63.32±3.657	78.55±1.557
A_{11}	48.25±3.856	47.59±2.563	45.53±3.013	40.43±3.125	50.27±3.450	62.14±3.848	79.04±1.456
A_{12}	40.67±4.787	41.34±3.211	40.29±3.883	42.12±3.578	44.33±4.446	53.23±4.152	70.24±1.956
Average	57.31±3.488	70.21±1.821	69.54±2.240	70.95±2.113	66.82±3.151	73.17±2.596	86.86±1.038

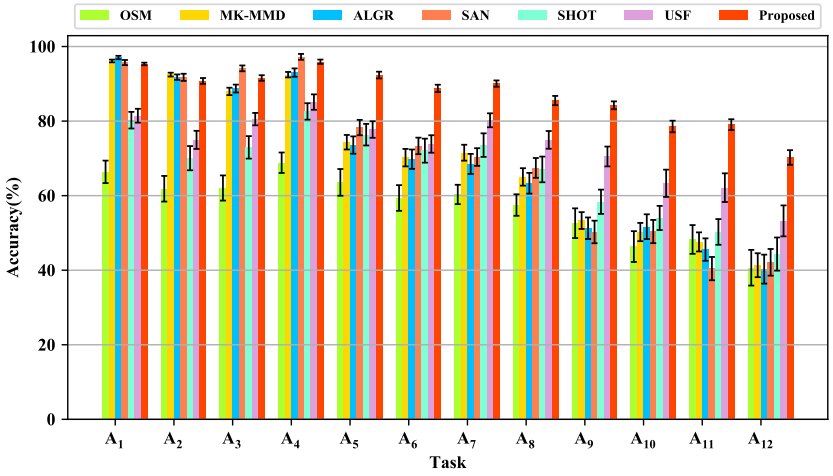


Figure 9: Diagnosis performance comparisons on the DDS test rig dataset.

The cross-domain fault diagnosis results in different tasks on the DDS test rig dataset are reported in Table 5 and Fig 9, whereas the results in different tasks on the bearing test rig dataset are reported in Table 6 and Fig 10. From the comparison results, it can be seen that the proposed method achieves the best average test accuracy in two cases, i.e., 86.86% on the DDS test rig dataset and 84.48% on the bearing test rig dataset.

For the non-domain adaptation method OSM, the diagnostic accuracy in all tasks is poor, which shows that there is a significant discrepancy in feature distribution between the source and target domains. In the closed-set

DA tasks A_1, A_2, B_1, and B_2, the MK-MMD method and ALGR method achieve slightly higher cross-domain diagnostic accuracy than the proposed method. In the partial DA tasks A_3, A_4, B_3, and B_4, the SAN method achieves slightly higher cross-domain diagnostic accuracy than the proposed method. The closed-set DA and partial DA tasks are designed to solve task-specific fault diagnosis problems, which are not practical in real-world industrial scenarios. However, the diagnostic performance obtained by the proposed source-free domain adaptation method is still competitive in the two extreme task scenarios. For the open-set DA task, the diagnostic accuracy obtained by the proposed method is higher than that of those source-available domain adaptation methods in tasks A_5, A_6, B_5, and B_6. There are two reasons for this phenomenon: 1) the samples of unknown classes are prone to negative transfer; 2) the samples of unknown classes are wrongly identified as known classes because those methods cannot have the ability to identify samples of unknown classes.

Table 6: The statistic of the accuracy and standard deviation for the bearing test rig dataset (%).

Task	OSM	MK-MMD	ALGR	SAN	SHOT	USF	Proposed
B_1	65.43±2.612	94.89±0.724	95.15±0.824	93.88±0.964	82.04±1.892	82.49±1.676	94.03±0.505
B_2	60.33±2.944	89.76±0.964	90.16±0.996	89.75±1.147	72.45±1.247	76.88±2.440	89.46±0.867
B_3	60.15±3.208	84.44±1.024	83.59±1.242	91.73±0.943	75.82±2.019	80.14±1.720	90.19±0.774
B_4	63.27±2.945	85.79±1.167	84.81±1.305	93.49±0.876	77.56±2.210	83.78±1.557	92.05±0.687
B_5	59.96±3.121	70.34±1.246	69.44±1.624	72.96±1.445	71.49±2.224	76.41±2.313	90.16±0.615
B_6	53.89±3.440	65.05±1.807	64.23±2.404	70.96±1.882	66.43±2.480	72.86±2.467	85.76±1.131
B_7	53.89±3.158	63.38±2.112	59.37±2.944	59.84±2.034	58.77±3.011	79.36±2.044	88.19±0.950
B_8	52.19±2.946	58.86±2.445	59.08±2.519	54.67±2.554	56.47±3.113	75.46±2.119	86.74±1.006
B_9	43.77±3.484	45.53±2.323	47.36±2.880	48.82±2.915	51.69±3.006	63.33±2.664	80.22±1.331
B_{10}	40.87±3.504	38.29±2.845	40.34±2.958	35.24±3.412	43.33±3.801	58.19±3.887	76.60±1.801
B_{11}	38.79±3.882	38.97±3.045	41.09±2.876	31.57±3.884	43.03±3.928	56.77±3.992	74.49±1.969
B_{12}	30.14±4.553	34.09±3.445	35.77±3.681	26.89±4.458	35.70±4.203	48.15±4.225	65.83±2.234
Average	51.89±3.316	64.12±1.929	64.20±2.188	64.15±2.210	61.23±2.761	71.15±2.592	84.48±1.156

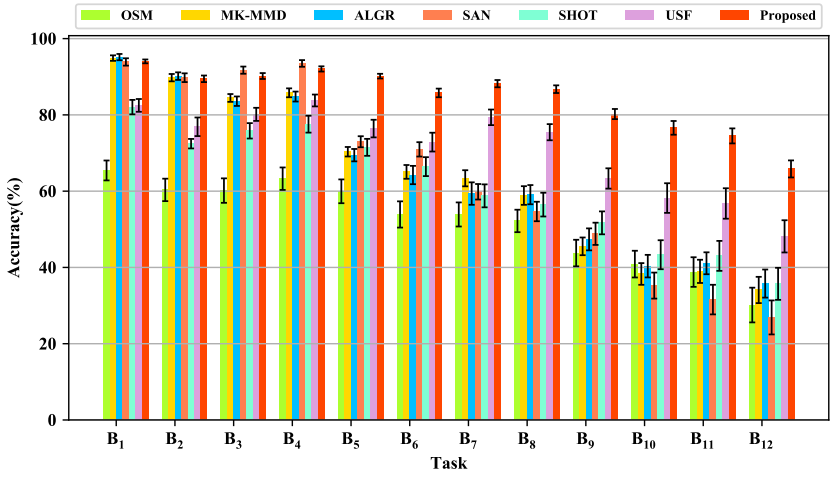


Figure 10: Diagnosis performance comparisons on the bearing test rig dataset.

In the universal DA tasks close to real-world industrial scenarios, the diagnostic accuracies obtained by the proposed method are significantly higher than that of the other 6 methods. It can be observed from Fig 9 and

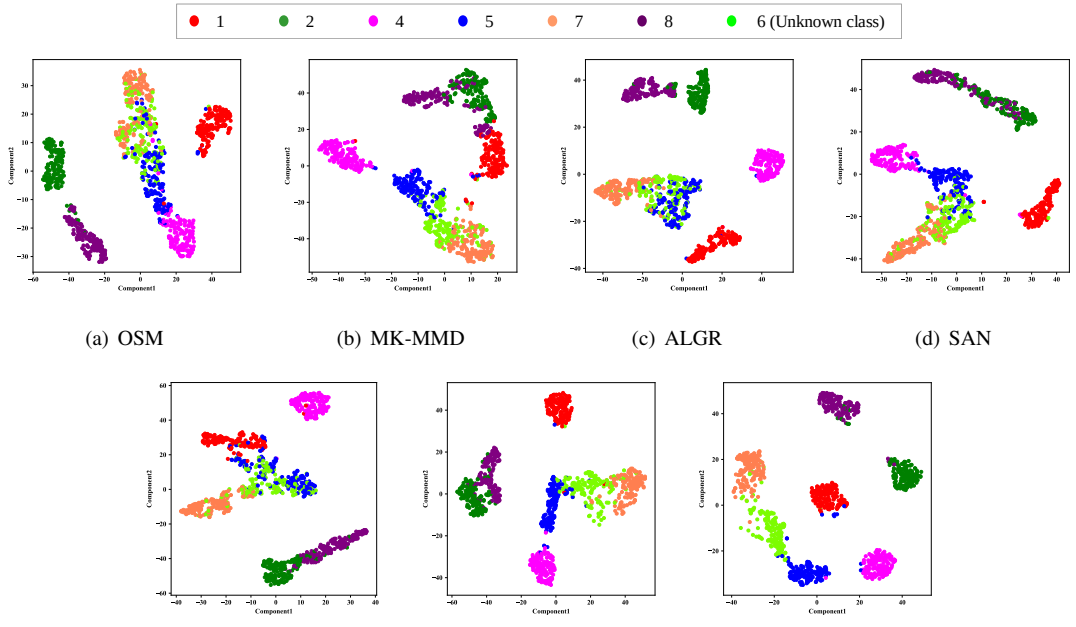
Fig 10 that the diagnostic performance of three source-available methods degrades severely, which can be attributed to the serious negative transfer caused by the simultaneous existence of source outliers and target outliers. As a popular source-free domain adaptation method, the diagnostic results of the SHOT have similar performance under the closed-set DA and partial DA scenarios and are significantly lower than those of the source-available methods. The diagnostic performance of the SHOT is somewhat better than the source-available methods in the universal DA scenario. Furthermore, as a state-of-the-art universal source-free DA method, USF achieves significant improvements in diagnostic accuracy in all universal DA tasks compared to other comparative methods. This is due to this method can classify outlier classes in the target domain into unknown classes, thereby improving the recognition performance of unknown classes. However, it can be seen from the diagnosis results that the current diagnostic accuracy obtained by USF is relatively low, especially with the increase of outlier classes, the diagnosis performance of the model degrades severely. Therefore, it can be seen that the existing methods can not well solve the source-free domain adaptation problem when the fault modes of the target are unknown. It is noteworthy that the proposed method achieves higher source-free cross-domain diagnostic accuracy on all universal DA tasks. Besides, it can be seen that with the increasing number of unknown classes, the diagnosis performance of the proposed method is stable. That is, it can be seen that the proposed method can effectively solve the source-free cross-domain fault diagnosis problem when there is a large label space biased between the source and target domains. In addition, the proposed method obtains the minimum standard deviation value in most tasks, which illustrates that the proposed method is more stable than the compared methods. Therefore, the effectiveness and superiority of the proposed method have been demonstrated in this section.

4.2.2. Visualization analysis

In this section, we consider task A_7 of the DDS test rig dataset and task B_7 of the bearing test rig dataset as two cases to perform feature visualization to better understand the advantages of the proposed method. Firstly, the t-SNE [58] algorithm is selected to visualize the target test features learned by the feature extractor. The t-SNE visualization results of all methods for the two tasks are shown in Fig. 11 and Fig. 12, respectively. As can be seen from Fig. 11 (a) and Fig. 12 (a), the features of the different classes obtained by OSM are mixed together, which indicates that there are a large number of misclassification samples when the source model is directly used for the target domain test. In addition, many features overlap in the different health conditions for MK-MMD, ALGR, SAN, SHOT, and USF. In contrast, the feature distribution map obtained by the proposed method is more discriminative. That is, the features of the same classes are gathered together, while the features of the different classes are well separated. At the same time, most unknown samples are clustered and separated from other classes.

Besides, the confusion matrix results of all methods for the two tasks are shown in Fig. 13 and Fig. 14, respectively. It can be seen from the confusion matrices of OSM, MK-MMD, ALGR, SAN, and SHOT that the target outlier classes are all misclassified, thus significantly affecting the overall recognition performance. USF can identify new failure modes in the target domain as an unknown class, thereby improving the performance of the model. It is worth noting that the proposed method can effectively identify both known and unknown classes in the target domain, achieving the highest diagnostic performance.

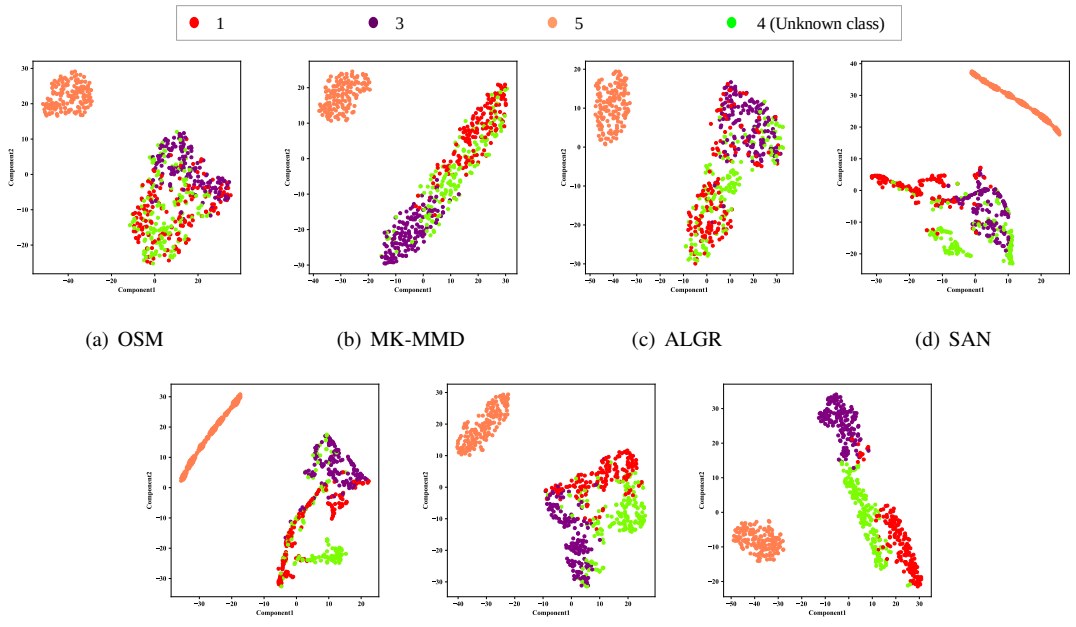
Further, we present the feature visualization result of the proposed method for the source domain training samples and target domain test samples under task A_7 and task B_7, and the results are shown in Fig. 15. It can



(a) OSM (b) MK-MMD (c) ALGR (d) SAN

(e) SHOT (f) USF (g) Proposed

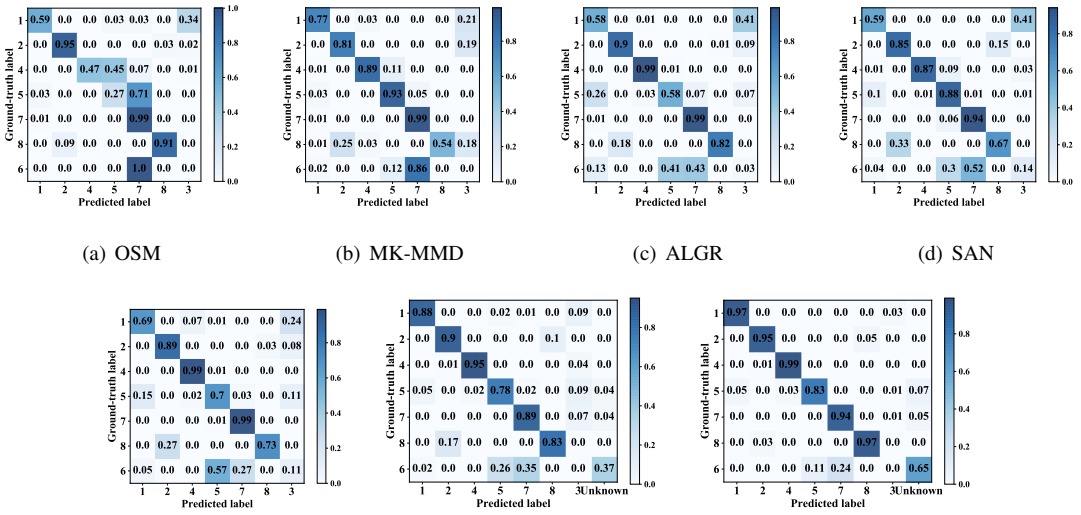
Figure 11: The t-SNE visualization of all methods on the task A_7.



(a) OSM (b) MK-MMD (c) ALGR (d) SAN

(e) SHOT (f) USF (g) Proposed

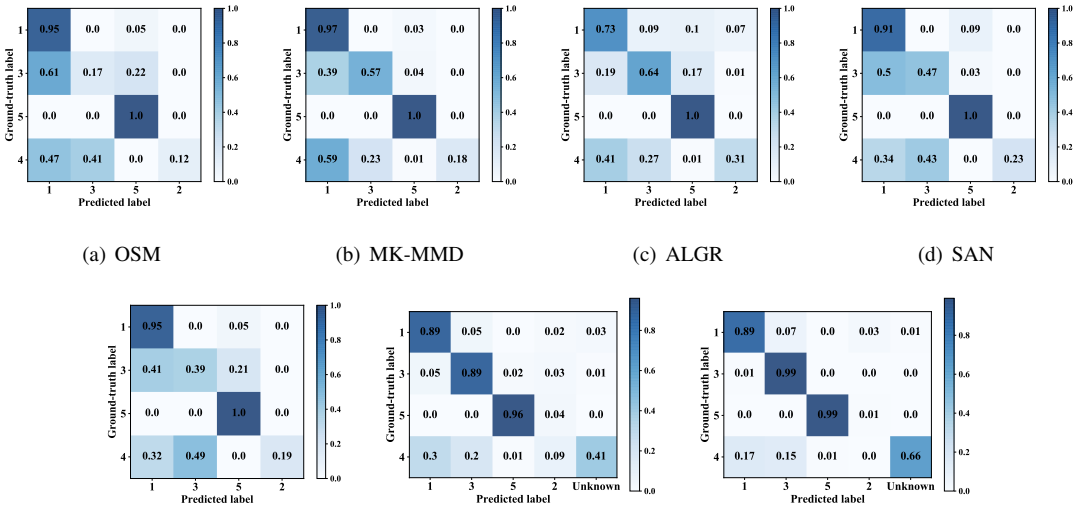
Figure 12: The t-SNE visualization of all methods on the task B_7.



(a) OSM (b) MK-MMD (c) ALGR (d) SAN

(e) SHOT (f) USF (g) Proposed

Figure 13: The confusion matrix of all methods on the task A_7.

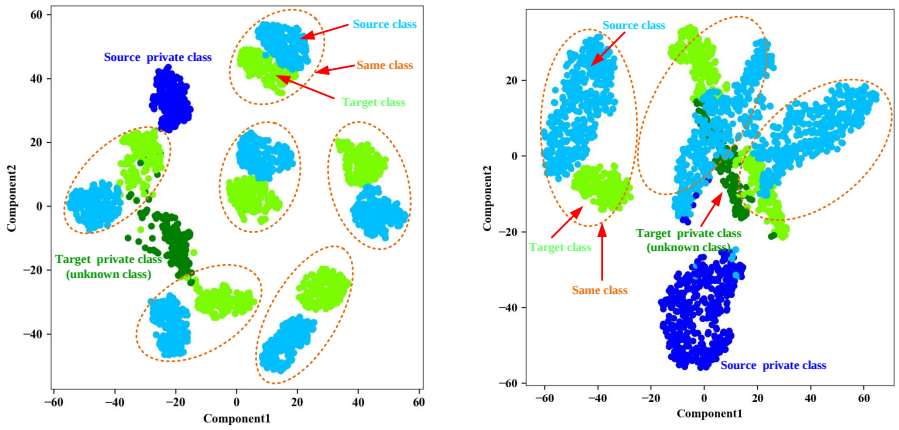


(a) OSM (b) MK-MMD (c) ALGR (d) SAN

(e) SHOT (f) USF (g) Proposed

Figure 14: The confusion matrix of all methods on the task B_7.

be seen that the proposed method achieves both promising domain alignment and class-level alignment, in which the same condition samples across domains are projected into similar regions in the new sub-space, and source and target outliers are well separated. Thus, this result demonstrates that the proposed method can achieve high-performance class-level alignment despite the inconsistency in the class spaces of the source and target domains. Therefore, these visualization results further illustrate the effectiveness of the proposed method.



(a) Task A_7 (b) Task B_7

Figure 15: The t-SNE visualization of source training samples and target test samples on the task A_7 and B_7.

4.3. Experimental analysis

4.3.1. Ablation study of the Transformer module

In the proposed network model, a Transformer module is embedded into the convolutional neural network. To analyze the advantages of the Transformer module, we design a common convolutional network model without the Transformer module for comparison. The diagnostic results of the proposed method and the method without the Transformer module (non-Transformer) on the DDS test rig dataset are presented in Table 7, in which the average diagnostic accuracies are 84.24% and 86.86%, respectively. And the diagnostic performance of the proposed method is better than that of the non-Transformer method on all tasks. Therefore, this result demonstrates that the Transformer module can improve the diagnostic performance of the model.

Table 7: The statistic of the accuracy for the non-Transformer method and proposed method (%).

Method	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8	A_9	A_{10}	A_{11}	A_{12}	Average
Non-Transformer	94.56	88.96	90.04	94.89	90.66	86.03	86.79	83.17	80.28	75.31	75.01	65.22	84.24
Proposed	95.32	90.77	91.56	95.92	92.35	88.76	90.06	85.51	84.23	78.55	79.04	70.24	86.86

4.3.2. Ablation study of target loss terms

In the proposed method, the objective of model training in the target domain contains three terms, namely high-confident source-free supervised contrastive loss ($\mathcal{L}_{HSSC}$), threshold-based entropy max-min loss ($\mathcal{L}_{EMM}$), and self-supervised pseudo-labeling loss ($\mathcal{L}_{SPL}$). $\mathcal{L}_{HSSC}$ is used to achieve domain-invariant learning with known classes in the target domain. $\mathcal{L}_{EMM}$ is used to align samples of known classes or reject target outlier samples as an unknown

class. Since both $\mathcal{L}_{HSSC}$ and $\mathcal{L}_{EMM}$ need to use the target domain pseudo labels, we use $\mathcal{L}_{SPL}$ to further reduce the number of false pseudo-labels. Thus, $\mathcal{L}_{HSSC}$ and $\mathcal{L}_{EMM}$ play a major role in our proposed method. To illustrate the advantages of the proposed loss functions, five variants are considered for ablation study: 1) HSSC, only $\mathcal{L}_{HSSC}$ is used during target training, 2) EMM, only $\mathcal{L}_{EMM}$ is used during target training, 3) HSSC_SPL, $\mathcal{L}_{HSSC}$ and $\mathcal{L}_{SPL}$ are used during target training, 4) EMM_SPL, $\mathcal{L}_{EMM}$ and $\mathcal{L}_{SPL}$ are used during target training, and 5) HSSC_EEM, $\mathcal{L}_{HSSC}$ and $\mathcal{L}_{EMM}$ are used during target training. The comparison results of the proposed method and the five variants on task A_7 are given in Fig. 16, their average diagnostic accuracy is 85.15%, 80.02%, 86.67%, 83.06%, 88.11%, and 90.06%, respectively. And it can be seen that the diagnostic results of the proposed method are better than the five variants methods in almost all trials. Therefore, this experiment demonstrates the effectiveness of the proposed loss term.

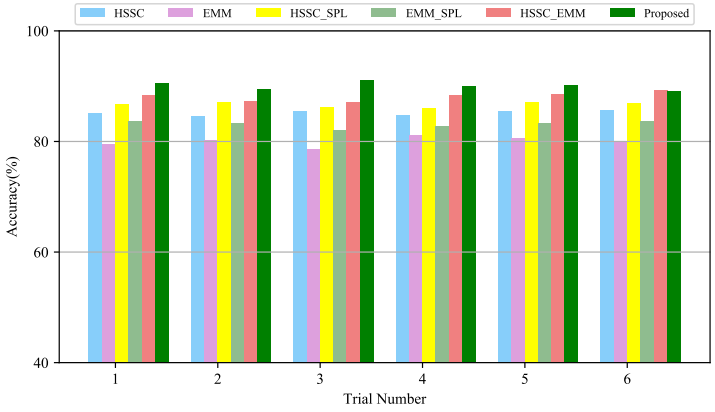


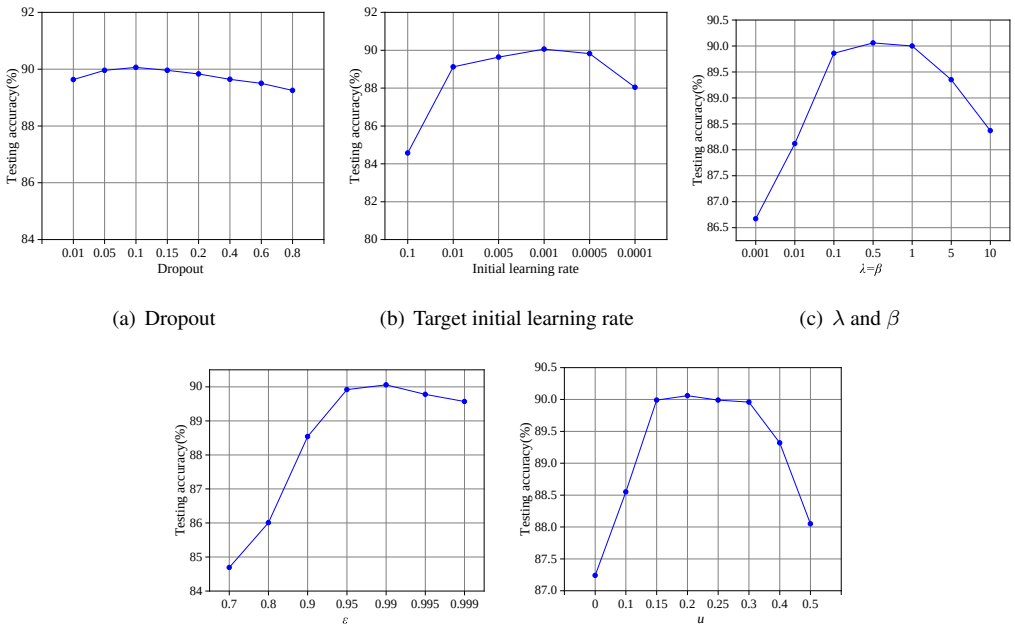
Figure 16: A comparison of the diagnostic results using different loss terms in task A_7.

4.3.3. Hyper-parameter selection

In this study, the grid search technique is used to experimentally validate the hyper-parameters. The proposed method's major hyper-parameters, including dropout, target initial learning rate, tradeoff parameter λ of $\mathcal{L}_{EMM}$, tradeoff parameter β of $\mathcal{L}_{SPL}$, threshold parameter ε of $\mathcal{L}_{HSSL}$, and confidence threshold u of $\mathcal{L}_{EMM}$ were chosen based on task A_7. To reduce the number of hyper-parameters, λ and β are equal in our experiment. In the parameter selection process, we first set the corresponding parameter range based on experience, and then perform a grid search on the pre-set parameter range by automatically training the proposed model. Fig. 17 presents the accuracy curve of the proposed model during parameter selection processes. It can be seen that the optimal diagnostic accuracy of the model is obtained under the parameters set in this paper. In addition, it can be observed that the diagnostic accuracy obtained around the parameters given in this work does not change significantly. Therefore, it illustrates that the hyper-parameters of the proposed method are not sensitive within a reasonable range.

4.3.4. Parameter freezing position analysis

During training in the target domain, we freeze the classifier's parameters and only train the feature extractor to maintain the source class prototypes. In this section, the effect of the parameter freezing position on the diagnostic performance of the model is investigated. Our model consists of a feature extractor and a classifier, and the feature extractor consists of a CNN module, a Transformer module, and a fully connected module. Therefore, in

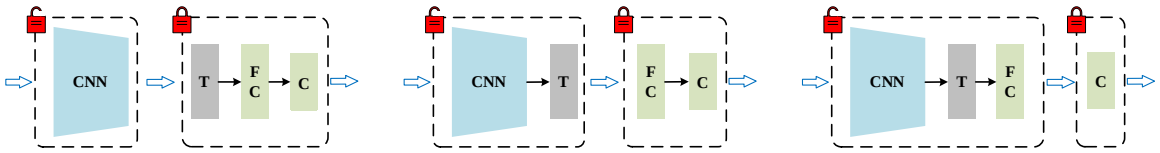


(a) Dropout (b) Target initial learning rate (c) λ and β

(d) ε (e) u

Figure 17: Test accuracy curves for hyper-parameters.

addition to the parameter freezing position proposed in this paper, two other freezing methods can be considered. A schematic diagram of the three parameter freezing positions is shown in Fig. 18, where (a) freezing the Transformer module, the fully connected module, and the classifier; (b) freezing the fully connected module and classifier; (b) freezing the classifier (proposed method). The diagnostic results of the three different parameter freezing methods on the DDS test rig dataset are shown in Table 8, in which the average diagnostic accuracies are 84.77%, 86.33%, and 86.86%, respectively. The diagnostic performance of the proposed method is better than that of the other two freezing methods on all tasks. Therefore, this result shows that the freeze position proposed in this paper is optimal.



(a) Freeze position A (b) Freeze position B (c) Freeze position C (Proposed)

Figure 18: A schematic diagram of the three parameter freezing positions.

Table 8: The statistic of the accuracy for the three different parameter freezing methods (%).

Method	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8	A_9	A_{10}	A_{11}	A_{12}	Average
Freeze position A	93.24	88.80	89.47	93.64	90.20	86.91	87.13	83.65	82.37	76.63	77.09	68.07	84.77
Freeze position B	94.58	90.14	91.18	95.33	91.96	88.29	89.57	85.42	83.79	78.01	78.43	69.20	86.33
Proposed	95.32	90.77	91.56	95.92	92.35	88.76	90.06	85.51	84.23	78.55	79.04	70.24	86.86

5. Conclusion

This work developed a novel solution for the universal source-free domain adaptation problem in the field of machinery fault diagnosis. To the best of our knowledge, no existing research has been reported on the universal source-free cross-domain machinery fault diagnosis. Thus, this paper provided a novel perspective for fault diagnosis in real-world industrial scenarios. The proposed method merely requires a well-trained source model and offers the feasibility of cross-domain fault diagnosis without access to source data. Specifically, a convolutional network with a Transformer module was developed to pay more attention to the discriminative feature. To achieve source-free domain adaptation, a supervised contrastive learning strategy based on source class prototypes was designed by using high-confident prediction samples. In absence of prior information on the target label set, a threshold-based entropy max-min loss was built to further align samples of known classes or reject target outlier samples as an unknown class. Self-supervised learning technique was further used to provide more accurate pseudo labels. Experiments under diverse domain adaptation settings were performed on two rotating machinery datasets. The experimental results demonstrated the effectiveness and superiority of the proposed method. Thus, the proposed method can bring a powerful tool to overcome diagnostic difficulties when the source domain data are unavailable and target domain fault modes are unknown. This would ensure the safe and efficient operation of complicated industrial systems, which would bring significant economic benefits to industry practices by reducing maintenance costs and unnecessary shutdowns.

6. Acknowledgments

This work was partially supported by the China Scholarship Council (CSC) under Grant 202106080066. The authors are grateful for the support.

References

[1] Y. Lei, B. Yang, X. Jiang, F. Jia, N. Li, A. K. Nandi, Applications of machine learning to machine fault diagnosis: A review and roadmap, Mechanical Systems and Signal Processing 138 (2020) 106587, https://doi.org/10.1016/j.ymssp.2019.106587.

[2] P. Gardner, L. Bull, N. Dervilis, K. Worden, On the application of kernelised bayesian transfer learning to population-based structural health monitoring, Mechanical Systems and Signal Processing 167 (2022) 108519, https://doi.org/10.1016/j.ymssp.2021.108519.

[3] A. Kamariotis, E. Chatzi, D. Straub, Value of information from vibration-based structural health monitoring extracted via bayesian model updating, Mechanical Systems and Signal Processing 166 (2022) 108465, https://doi.org/10.1016/j.ymssp.2021.108465.

[4] M. M. Islam, J.-M. Kim, Reliable multiple combined fault diagnosis of bearings using heterogeneous feature models and multiclass support vector machines, Reliability Engineering & System Safety 184 (2019) 55–66, https://doi.org/10.1016/j.ress.2018.02.012.

[5] R. Yan, F. Shen, C. Sun, X. Chen, Knowledge transfer for rotary machine fault diagnosis, IEEE Sensors Journal 20 (15) (2019) 8374–8393, https://doi.org/10.1109/JSEN.2019.2949057.

[6] N. Dervilis, M. Choi, S. Taylor, R. Barthorpe, G. Park, C. Farrar, K. Worden, On damage diagnosis for a wind turbine blade using pattern recognition, Journal of sound and vibration 333 (6) (2014) 1833–1850, https://doi.org/10.1016/j.jsv.2013.11.015.

[7] K. Feng, J. Ji, Q. Ni, M. Beer, A review of vibration-based gear wear monitoring and prediction techniques, Mechanical Systems and Signal Processing 182 (2023) 109605, https://doi.org/10.1016/j.ymssp.2022.109605.

[8] M. Xia, H. Shao, D. Williams, S. Lu, L. Shu, C. W. de Silva, Intelligent fault diagnosis of machinery using digital twin-assisted deep transfer learning, Reliability Engineering & System Safety 215 (2021) 107938, https://doi.org/10.1016/j.ress.2021.107938.

[9] N. Dervilis, M. Choi, I. Antoniadou, K. Farinholt, S. Taylor, R. J. Barthorpe, G. Park, C. R. Farrar, K. Worden, Machine learning applications for a wind turbine blade under continuous fatigue loading, in: Key Engineering Materials, Vol. 588, Trans Tech Publ, 2014, pp. 166–174, https://doi.org/10.4028/www.scientific.net/KEM.588.166.

[10] T. Han, C. Liu, L. Wu, S. Sarkar, D. Jiang, An adaptive spatiotemporal feature learning approach for fault diagnosis in complex systems, Mechanical Systems and Signal Processing 117 (2019) 170–187, https://doi.org/10.1016/j.ymssp.2018.07.048.

[11] M. Zhao, S. Zhong, X. Fu, B. Tang, M. Pecht, Deep residual shrinkage networks for fault diagnosis, IEEE Transactions on Industrial Informatics 16 (7) (2019) 4681–4690, https://doi.org/10.1109/TII.2019.2943898.

[12] H. Shao, H. Jiang, H. Zhao, F. Wang, A novel deep autoencoder feature learning method for rotating machinery fault diagnosis, Mechanical Systems and Signal Processing 95 (2017) 187–204, https://doi.org/10.1016/j.ymssp.2017.03.034.

[13] Y. Lei, X. L. Liang, F. Chaari, Special feature on rotating machinery condition monitoring by connecting physics-based and data-driven methods, Measurement Science and Technology 33 (1) (2021) 010103.

[14] L. D. Avendano-Valencia, E. N. Chatzi, D. Tcherniak, Gaussian process models for mitigation of operational variability in the structural health monitoring of wind turbines, Mechanical Systems and Signal Processing 142 (2020) 106686, https://doi.org/10.1016/j.ymssp.2020.106686.

[15] X. Zhao, J. Yao, W. Deng, P. Ding, Y. Ding, M. Jia, Z. Liu, Intelligent fault diagnosis of gearbox under variable working conditions with adaptive intraclass and interclass convolutional neural network, IEEE Transactions on Neural Networks and Learning Systemshttps://doi.org/10.1109/TNNLS.2021.3135877.

[16] Y. Zhang, K. Feng, H. Ma, K. Yu, Z. Ren, Z. Liu, Mmfnet: Multi-sensor data and multi-scale feature fusion model for intelligent cross-domain machinery fault diagnosis, IEEE Transactions on Instrumentation and Measurementhttps://doi.org/10.1109/TIM.2022.3213016.

[17] R. Huang, J. Li, Y. Liao, J. Chen, Z. Wang, W. Li, Deep adversarial capsule network for compound fault diagnosis of machinery toward multidomain generalization task, IEEE Transactions on Instrumentation and Measurement 70 (2020) 1–11, https://doi.org/10.1109/TIM.2020.3042300.

[18] P. Gardner, L. Bull, J. Gosliga, J. Poole, N. Dervilis, K. Worden, A population-based shm methodology for heterogeneous structures: Transferring damage localisation knowledge between different aircraft wings, Mechanical Systems and Signal Processing 172 (2022) 108918, https://doi.org/10.1016/j.ymssp.2022.108918.

[19] Y. Zhang, Z. Ren, S. Zhou, K. Feng, K. Yu, Z. Liu, Supervised contrastive learning-based domain adaptation network for intelligent unsupervised fault diagnosis of rolling bearing, IEEE/ASME Transactions on Mechatronicshttps://doi.org/10.1109/TMECH.2022.3179289.

[20] W. Zhang, X. Li, H. Ma, Z. Luo, X. Li, Universal domain adaptation in fault diagnostics with hybrid weighted deep adversarial learning, IEEE Transactions on Industrial Informatics 17 (12) (2021) 7957–7967, https://doi.org/10.1109/TII.2021.3064377.

[21] X. Zhao, J. Yao, W. Deng, M. Jia, Z. Liu, Normalized conditional variational auto-encoder with adaptive focal loss for imbalanced fault diagnosis of bearing-rotor system, Mechanical Systems and Signal Processing 170 (2022) 108826, https://doi.org/10.1016/j.ymssp.2022.108826.

[22] M. Zhao, S. Zhong, X. Fu, B. Tang, S. Dong, M. Pecht, Deep residual networks with adaptively parametric rectifier linear units for fault diagnosis, IEEE Transactions on Industrial Electronics 68 (3) (2020) 2587–2597, https://doi.org/10.1109/TIE.2020.2972458.

[23] Y. Ding, J. Zhuang, P. Ding, M. Jia, Self-supervised pretraining via contrast learning for intelligent incipient fault detection of bearings, Reliability Engineering & System Safety 218 (2022) 108126, https://doi.org/10.1016/j.ress.2021.108126.

[24] L. Wen, X. Li, L. Gao, Y. Zhang, A new convolutional neural network-based data-driven fault diagnosis method, IEEE Transactions on Industrial Electronics 65 (7) (2017) 5990–5998, https://doi.org/10.1109/TIE.2017.2774777.

[25] X. Li, W. Zhang, Q. Ding, J.-Q. Sun, Intelligent rotating machinery fault diagnosis based on deep learning using data augmentation, Journal of Intelligent Manufacturing 31 (2) (2020) 433–452, https://doi.org/10.1007/s10845-018-1456-1.

[26] Y.-r. Wang, Q. Jin, G.-d. Sun, C.-f. Sun, Planetary gearbox fault feature learning using conditional variational neural networks under noise environment, Knowledge-Based Systems 163 (2019) 438–449, https://doi.org/10.1016/j.knosys.2018.09.005.

[27] G. Jiang, H. He, J. Yan, P. Xie, Multiscale convolutional neural networks for fault diagnosis of wind turbine gearbox, IEEE Transactions on Industrial Electronics 66 (4) (2018) 3196–3207, https://doi.org/10.1109/TIE.2018.2844805.

[28] S. Guo, B. Zhang, T. Yang, D. Lyu, W. Gao, Multitask convolutional neural network with information fusion for bearing fault diagnosis and localization, IEEE Transactions on Industrial Electronics 67 (9) (2019) 8005–8015, https://doi.org/10.1109/TIE.2019.2942548.

[29] Y. Guan, Z. Meng, D. Sun, J. Liu, F. Fan, 2mnet: Multi-sensor and multi-scale model toward accurate fault diagnosis of rolling bearing, Reliability Engineering & System Safety 216 (2021) 108017, https://doi.org/10.1016/j.ress.2021.108017.

[30] W. Li, R. Huang, J. Li, Y. Liao, Z. Chen, G. He, R. Yan, K. Gryllias, A perspective survey on deep transfer learning for fault diagnosis in industrial scenarios: Theories, applications and challenges, Mechanical Systems and Signal Processing 167 (2022) 108487, https://doi.org/10.1016/j.ymssp.2021.108487.

[31] T. Ritto, K. Worden, D. Wagg, F. Rochinha, P. Gardner, A transfer learning-based digital twin for detecting localised torsional friction in deviated wells, Mechanical Systems and Signal Processing 173 (2022) 109000, https://doi.org/10.1016/j.ymssp.2022.109000.

[32] W. Lu, B. Liang, Y. Cheng, D. Meng, J. Yang, T. Zhang, Deep model based domain adaptation for fault diagnosis, IEEE Transactions on Industrial Electronics 64 (3) (2016) 2296–2305, https://doi.org/10.1109/TIE.2016.2627020.

[33] B. Yang, Y. Lei, F. Jia, S. Xing, An intelligent fault diagnosis approach based on transfer learning from laboratory bearings to locomotive bearings, Mechanical Systems and Signal Processing 122 (2019) 692–706, https://doi.org/10.1016/j.ymssp.2018.12.051.

[34] Z. Wang, X. He, B. Yang, N. Li, Subdomain adaptation transfer learning network for fault diagnosis of roller bearings, IEEE Transactions on Industrial Electronicshttps://doi.org/10.1109/TIE.2021.3108726.

[35] T. Han, C. Liu, W. Yang, D. Jiang, A novel adversarial learning framework in deep convolutional neural network for intelligent diagnosis of mechanical faults, Knowledge-based systems 165 (2019) 474–487, https://doi.org/10.1016/j.knosys.2018.12.019.

[36] Y. Zhang, K. Yu, Z. Ren, S. Zhou, Joint domain alignment and class alignment method for cross-domain fault diagnosis of rotating machinery, IEEE Transactions on Instrumentation and Measurement 70 (2021) 1–12, https://doi.org/10.1109/TIM.2021.3120790.

[37] J. Jiao, M. Zhao, J. Lin, K. Liang, Residual joint adaptation adversarial network for intelligent transfer fault diagnosis, Mechanical Systems and Signal Processing 145 (2020) 106962, https://doi.org/10.1016/j.ymssp.2020.106962.

[38] J. Jiao, M. Zhao, J. Lin, C. Ding, Classifier inconsistency-based domain adaptation network for partial transfer intelligent diagnosis, IEEE Transactions on Industrial Informatics 16 (9) (2019) 5965–5974, https://doi.org/10.1109/TII.2019.2956294.

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

[40] S. Jia, J. Wang, X. Zhang, B. Han, A weighted subdomain adaptation network for partial transfer fault diagnosis of rotating machinery, Entropy 23 (4) (2021) 424, https://doi.org/10.3390/e23040424.

[41] W. Li, Z. Chen, G. He, A novel weighted adversarial transfer network for partial domain fault diagnosis of machinery, IEEE Transactions on Industrial Informatics 17 (3) (2020) 1753–1762, https://doi.org/10.1109/TII.2020.2994621.

[42] Y. Deng, D. Huang, S. Du, G. Li, C. Zhao, J. Lv, A double-layer attention based adversarial network for partial transfer learning in machinery fault diagnosis, Computers in Industry 127 (2021) 103399, https://doi.org/10.1016/j.compind.2021.103399.

[43] W. Zhang, X. Li, H. Ma, Z. Luo, X. Li, Open-set domain adaptation in machinery fault diagnostics using instance-level weighted adversarial learning, IEEE Transactions on Industrial Informatics 17 (11) (2021) 7445–7455, https://doi.org/10.1109/TII.2021.3054651.

[44] J. Zhu, C. Huang, C. Shen, Y. Shen, Cross-domain open set machinery fault diagnosis based on adversarial network with multiple auxiliary classifiers, IEEE Transactions on Industrial Informaticshttps://doi.org/10.1109/TII.2021.3138558.

[45] Y. Feng, J. Chen, S. He, T. Pan, Z. Zhou, Globally localized multisource domain adaptation for cross-domain fault diagnosis with category shift, IEEE Transactions on Neural Networks and Learning Systemshttps://doi.org/10.1109/TNNLS.2021.3111732.

[46] C. Zhao, W. Shen, Dual adversarial network for cross-domain open set fault diagnosis, Reliability Engineering & System Safety 221 (2022) 108358, https://doi.org/10.1016/j.ress.2022.108358.

[47] F. Zhuang, Z. Qi, K. Duan, D. Xi, Y. Zhu, H. Zhu, H. Xiong, Q. He, A comprehensive survey on transfer learning, Proceedings of the IEEE 109 (1) (2020) 43–76, https://doi.org/10.1109/JPROC.2020.3004555.

[48] Y. Kim, D. Cho, K. Han, P. Panda, S. Hong, Domain adaptation without source data, arXiv preprint arXiv:2007.01524https://arxiv.org/abs/2007.01524.

[49] J. Liang, D. Hu, Y. Wang, R. He, J. Feng, Source data-absent unsupervised domain adaptation through hypothesis transfer and labeling transfer, IEEE Transactions on Pattern Analysis and Machine Intelligencehttps://doi.org/10.1109/TPAMI.2021.3103390.

[50] S. M. Ahmed, D. S. Raychaudhuri, S. Paul, S. Oymak, A. K. Roy-Chowdhury, Unsupervised multi-source domain adaptation without access to source data, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, pp. 10103–10112.

[51] J. N. Kundu, N. Venkat, R. V. Babu, et al., Universal source-free domain adaptation, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 4544–4553.

[52] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, et al., An image is worth 16x16 words: Transformers for image recognition at scale, arXiv preprint arXiv:2010.11929https://arxiv.org/abs/2010.11929.

[53] A. Van den Oord, Y. Li, O. Vinyals, et al., Representation learning with contrastive predictive coding, arXiv preprint arXiv:1807.03748 2 (3) (2018) 4.

[54] K. Saito, D. Kim, S. Sclaroff, K. Saenko, Universal domain adaptation through self supervision, Advances in neural information processing systems 33 (2020) 16282–16292.

[55] X. Li, W. Zhang, Q. Ding, J.-Q. Sun, Multi-layer domain adaptation method for rolling bearing fault diagnosis, Signal processing 157 (2019) 180–197, https://doi.org/10.1016/j.sigpro.2018.12.005.

[56] X. Li, W. Zhang, N.-X. Xu, Q. Ding, Deep learning-based machinery fault diagnostics with domain adaptation across sensors at different places, IEEE Transactions on Industrial Electronics 67 (8) (2019) 6785–6794, https://doi.org/10.1109/TIE.2019.2935987.

[57] Z. Cao, M. Long, J. Wang, M. I. Jordan, Partial transfer learning with selective adversarial networks, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2018, pp. 2724–2732.

[58] L. Van der Maaten, G. Hinton, Visualizing data using t-sne., Journal of machine learning research 9 (11).