Open-Set Domain Adaptation in Machinery Fault Diagnostics Using Instance-Level Weighted Adversarial Learning

Wei Zhang, Xiang Li, Hui Ma, Zhong Luo, and Xu Li

***Abstract*—Data-driven machinery fault diagnosis methods have been successfully developed in the past decades. However, the cross-domain diagnostic problems have not been well addressed, where the training and testing data are collected under different operating conditions. Recently, domain adaptation approaches have been popularly used to bridge this gap, which extract domain-invariant features for diagnostics. Despite the effectiveness, most existing methods assume the label spaces of training and testing data are identical that indicates the fault mode sets are the same in different scenarios. In practice, new fault modes usually occur in testing, which makes the conventional methods focusing on marginal distribution alignment less effective. In order to address this problem, a deep learning-based open-set domain adaptation method is proposed in this study. Adversarial learning is introduced to extract generalized features, and an instance-level weighted mechanism is proposed to reflect the similarities of testing samples with known health states. The unknown fault mode can be effectively identified, and the known states can be also recognized. Entropy minimization scheme is further adopted to improve generalization. Experiments on two practical rotating machinery datasets validate the proposed method. The results suggest the proposed method is promising for open-set domain adaptation problems, which largely enhances the applicability of data-driven approaches in the real industries.**

***Index Terms*—Deep learning, fault diagnosis, open-set domain adaptation, rotating machines, transfer learning.**

Manuscript received October 16, 2020; revised January 4, 2021; accepted January 21, 2021. Date of publication January 26, 2021; date of current version July 26, 2021. This work was supported by grant (VCAME201906) from the Key Laboratory of Vibration, and Control of Aero-Propulsion System Ministry of Education, Northeastern University, under Grant 52005086 and Grant 11902202 from the National Natural Science Foundation of China, under Grant N2005010 and Grant N180708009 from the Fundamental Research Funds for the Central Universities, and under Grant 2020-BS-048 and Grant 2019-BS-184 from Liaoning Provincial Department of Science and Technology. Paper no. TII-20-4785. *(Corresponding author: Xiang Li.)*

Wei Zhang is with the School of Aerospace Engineering, Shenyang Aerospace University, Shenyang 110136, China and also with the Key Laboratory of Vibration, and Control of Aero-Propulsion System Ministry of Education, Northeastern University, Shenyang 110819, China (e-mail: 1012201003@tju.edu.cn).

Xiang Li is with the Key Laboratory of Vibration, and Control of Aero-Propulsion System Ministry of Education, and College of Sciences, Northeastern University, Shenyang 110819, China (e-mail: xiangli@mail.neu.edu.cn).

Hui Ma and Zhong Luo are with the School of Mechanical Engineering, and Automation, and Key Laboratory of Vibration, and Control of Aero-Propulsion System Ministry of Education, Northeastern University, Shenyang 110819, China (e-mail: mail006527@qq.com; 919953753@qq.com).

Xu Li is with the State Key Laboratory of Rolling, and Automation, Northeastern University, Shenyang 110819, China (e-mail: 812811071@qq.com).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TII.2021.3054651.

Digital Object Identifier 10.1109/TII.2021.3054651

I. Introduction

INTELLIGENT data-driven machinery fault diagnostic methods have been successfully developed and popularly applied in the practical scenarios in the past decades [1], [2]. The automatic implementation workflow, low requirement of special expertise and high diagnosis accuracy largely promote the development of such methods in both academic researches and industrial applications. As a representative intelligent computing approach, deep learning has been attracting growing attention in the recent years, and the fault diagnosis problems have been greatly benefited from this emerging technology, which holds great potential to overcome the tough challenges that cannot be well addressed by the conventional methods.

In general, supervised learning paradigm is mostly preferred in data-driven approaches. It is usually assumed that the training and testing data are sampled from the same distribution that indicates the data are supposed to be collected from the same machine under identical operating conditions. However, in the real industrial scenarios, similar machines are usually working in different regimes, due to the specific task demands and environments. Therefore, distribution discrepancy generally exists between the training and testing data, which are denoted as source and target domains, respectively. As a consequence, the fault diagnosis knowledge learned from the source domain is less effective in the target domain. This phenomenon is known as domain shift problem [3], and illustrated in Fig. 1.

In order to address this challenging issue, transfer learning methods have been proposed in the literature, which aim to transfer the data-driven knowledge across different domains, thus increasing the model generalization ability [4]. Specifically, domain adaptation techniques have been much preferred in fault diagnosis studies that can well bridge the domain gap through extraction of domain-invariant features for diagnostics. Especially, the deep neural network framework has shown great capability in learning shared features from different distributions [5] that is well suited for fault diagnosis problems.

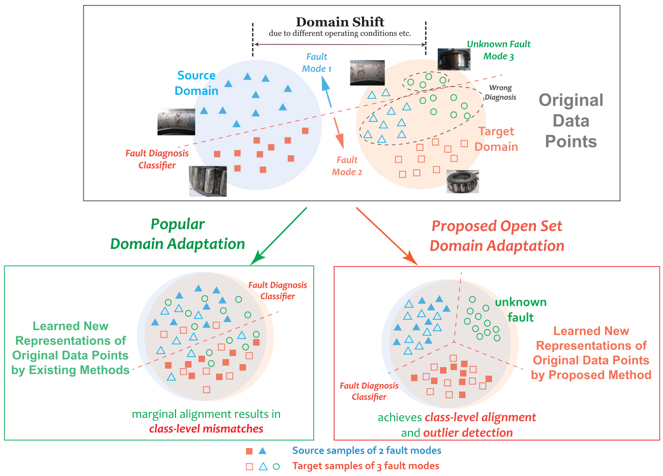


Fig. 1. Illustrations of domain shift phenomenon and open-set domain adaptation method.

One of the main assumptions in most domain adaptation methods is the identical label space across domains that suggests the training and testing data should cover the same machinery health states. This assumption facilitates fault diagnosis knowledge transfer through source and target alignment in marginal data distributions. However, it is difficult to satisfy this assumption in the real industries. Due to the significant economic costs and human labor in data collection, it is usually difficult to obtain labeled training data in a wide variety of machinery health states, resulting in limited source-domain classes. In the testing scenarios, new fault modes are possible to occur, which are not included in the source classes. That leads to a more challenging and practical domain adaptation problem, where the label space of the source domain is a subset of the target label space. This scenario is denoted as the open-set domain adaptation problem [6], [7]. For instance, with respect to rolling bearing fault diagnosis case, the training data may include bearing healthy and inner race fault conditions under 1500 r/min rotating speed, while the testing data contain bearing healthy, inner race fault and outer race fault conditions under 2000 r/min rotating speed. The outer race fault is not observed in the training data and, thus, should be treated as unknown health state.

As shown in Fig. 1, due to the target outlier classes, the existing popular domain adaptation methods which focus on the marginal distribution alignment cannot well achieve class-level diagnostic knowledge transfer in the open-set domain adaptation problem, thus resulting in deteriorations in model generalization ability [8]. This study contributes efforts to address this challenging and practical problem, and further develop transfer learning algorithms for better industrial applications. Specifically, the main novelties and contributions of this article are listed as follows:

1) The challenging open-set domain adaptation problem in machinery fault diagnostics is effectively addressed, which has been seldomly investigated in the literature.
2) A deep learning-based adversarial training scheme is introduced to extract domain-invariant features for diagnostic knowledge transfer. Instance-level weights for target-domain samples are proposed to describe the similarities with the source classes, which are derived from the domain discriminator scores.
3) Entropy minimization technique is applied to enhance the model performance in target outlier detection.
4) The proposed method can correctly identify the unknown fault modes in testing, and accurately diagnose the shared health states.
5) Extensive experiments on two real-world datasets validate the effectiveness and superiority of the proposed open-set domain adaptation method.

The remainder of this article starts with the related works in Section II. Section III presents the preliminaries of this study. The proposed method is depicted in Section IV, and experimentally validated and investigated in Section V. Finally, Section VI concludes this article.

II. Related Works

In the past years, deep learning-based data-driven fault diagnosis methods have been popularly investigated, and the typical problems where the training and testing data are from similar distributions have been well addressed [9]–[11]. A deep residual neural network model was proposed in [12] for rolling element bearing fault diagnosis. The deep convolutional neural network-based structure achieves fairly high testing accuracies under the basic problem setting. A stacked denoising autoencoder model was proposed by Lu *et al.* [13], where the machine health condition classification task is investigated with ambient noises. A deep semisupervised learning method with the multiple association layers network was proposed by Zhang *et al.* [14] for planetary gearbox fault diagnosis. Both labeled and unlabeled data are exploited to increase recognition accuracy.

For the cross-domain fault diagnosis problems where the training and testing data are collected under different operating conditions, domain adaptation methods have been successfully integrated in the deep learning framework to enhance the model generalization ability [15], [16]. In general, shared features between source and target domains are expected to be learned in the high-level data subspace to achieve knowledge transfer [17]. Two popular learning schemes are mostly adopted in the literature, i.e., minimization of maximum mean discrepancy (MMD) and adversarial learning. The MMD metric between source and target domains measures the gap of the learned features, the minimization of which leads to cross-domain data alignment. Lu *et al.* [18] proposed an MMD-based domain adaptation method for fault diagnosis using deep learning. Besides optimization of the MMD loss, a weight term is also used for regularization. A sparse autoencoder deep neural network model was proposed by Wen *et al.* [19], where the MMD metric is also focused on for extracting invariant features.

An alternative approach for deep learning-based domain adaptation is domain adversarial learning, which is inspired by the generative adversarial networks. Through adversarial training between feature extractor and domain discriminator, domain-invariant features can be also effectively obtained [20]. Guo *et al.* [21] proposed a deep convolutional neural network model with domain adversarial learning for cross-domain machinery fault diagnosis. Minimization of MMD is also integrated to enhance the model performance. Li *et al.* [22] proposed an adversarial learning scheme to extract shared features from multiple domains, and the results show the fault diagnosis knowledge can be transferred across similar but different machines.

Despite the advances of transfer learning in fault diagnosis, the open-set domain adaptation problem has been seldomly investigated. As a similar problem, the partial domain adaptation issue has been recently focused on, where the target label space is contained in the source label space [23]. Jiao *et al.* [24] proposed an unsupervised intelligent diagnosis framework for partial transfer learning. The classifier inconsistency is utilized to guide the model to learn discriminative and domain-invariant features for accurate classification of unlabeled target data.

In computer vision problems, attempts have been made on the open-set domain adaptation scenarios. Saito *et al.* [25] proposed an adversarial learning method to separate unknown target instances from known target ones, and promising results are obtained in the image processing tasks. Liu *et al.* [6] proposed a separate to adapt scheme for open-set domain adaptation, where multibinary classifiers are introduced to separate target images into known and unknown classes. In the literature, limited studies can be found on the machinery fault diagnosis scenarios, and this study aims at bridging this gap for practical industrial applications.

III. Preliminaries

A. Problem Statement

In this study, the open-set domain adaptation problem in fault diagnosis is investigated. Let $\mathcal{D}_s = \{(\mathbf{x}_i^s, y_i^s)\}_{i=1}^{n_s}$ and $\mathcal{D}_t = \{(\mathbf{x}_i^t, y_i^t)\}_{i=1}^{n_t}$ denote the source domain with n_s labeled samples and target domain with n_t unlabeled samples, respectively. It is assumed in this study that the label space of the source domain Y_s is included in that of the target domain Y_t, i.e., $Y_s \subseteq Y_t$. Specifically, the source data are sampled from known machinery health states, while the target data can be possibly collected from additional unknown health states. The data in $\mathcal{D}_s$ and $\mathcal{D}_t$ are sampled from distributions P_s and P_t, respectively. Domain discrepancy exists between source and target data, which are collected under different machinery operating conditions. That indicates $P_s \neq P_t$ due to the domain shift phenomenon. To be concrete, it is assumed $P_s \neq P_{t,\text{shared}}$ where $P_{t,\text{source}}$ denotes the distribution of the target data belonging to the source label space Y_s. The openness for the open-set domain adaptation problem can be defined as $\mathcal{O}_{\text{open}} = 1 - \frac{|Y_s|}{|Y_t|}$, which describes the discrepancy in label spaces across domains [6].

This study aims to build a data-driven model, which can extract invariant features from the source and target domains for fault diagnosis. The target-domain data belonging to the shared health states with the source domain can be correctly classified, and the target outlier states are expected to be identified as unknown class. The main assumption in this study lies in the availabilities of both source and target data in the model training process. Through exploration of the data in both domains, the shared features in the common health states can be learned, while the target outliers can be neglected in domain adaptation.

B. Data Preprocessing

In the machinery fault diagnosis problems, vibration signals have been much preferred for analysis, which can well reflect different fault modes. In this study, fast Fourier transformation is first applied on the raw measured temporal data. Compared with the popular image data whose values are basically distributed in the fixed range, the frequency-domain vibration data have remarkable peaks in the spectrum that significant affect the data processing efficiency in the deep neural network. To reduce this effect, the min–max scaling preprocessing technique is used in this study. For each sample denoted as $\mathbf{x} = [x_1, x_2, \ldots, x_{N_{\text{input}}}]$ where N_{input} is the sample dimension, the preprocessing in the following is implemented:

$$\begin{aligned} x_i &= \ln(x_i - x_{\min} + 1) \\ x_{\min} &= \min x_i, i = 1, 2, \ldots, N_{\text{input}}. \end{aligned} \tag{1}$$

The logarithmic function is applied to alleviate the effects of the peaks, and project the data into the value range where they are more uniformly distributed.

C. Adversarial Learning in Domain Adaptation

In general, a feature extractor G can be first used to learn high-level features $G(\mathbf{x})$ from the input samples of different domains. Next, a domain discriminator D can be adopted, which takes the learned features as inputs, and outputs the predicted domain labels $D(G(\mathbf{x}))$. Let $\boldsymbol{\theta}_G$ and $\boldsymbol{\theta}_D$ denote the parameters of G and D, respectively. Adversarial learning is introduced for the feature extractor to learn invariant features across domains. Basically, the domain discriminator is trained to correctly identify the domain labels from the high-level representations, while the feature extractor is optimized to obtain generalized features that cannot be distinguished by the domain discriminator. Through adversarial training between G and D, the high-level features learned from G can be more and more domain-invariant, thus bridging the domain gap. In classification studies such as fault diagnosis, a classifier C can be further adopted to classify using $G(\mathbf{x})$, and its parameters are denoted as $\boldsymbol{\theta}_C$.

Specifically in model optimization, the feature extractor and classifier are optimized under source supervision, and the empirical classification error on the labeled source-domain data should be minimized. The domain discriminator is trained through minimization of the domain label prediction errors. Meanwhile, the feature extractor is also optimized to increase the domain label prediction errors in order to confuse the discriminator. Therefore, the general objective L_{adv} of adversarial learning in domain adaptation can be formulated as

$$\begin{aligned} L_{\text{adv}}(\boldsymbol{\theta}_G, \boldsymbol{\theta}_D, \boldsymbol{\theta}_C) = &\frac{1}{n_s} \sum_{\mathbf{x}_i \in \mathcal{D}_s} L_c(C(G(\mathbf{x}_i)), y_i) \\ &- \frac{\alpha_{\text{adv}}}{n_s + n_t} \sum_{\mathbf{x}_i \in (\mathcal{D}_s \cup \mathcal{D}_t)} L_d(D(G(\mathbf{x}_i)), d_i) \end{aligned} \tag{2}$$

where L_c represents the health state classification loss, and L_d denotes the domain label prediction loss. y_i and d_i are the health state and domain labels for the data sample $\mathbf{x}_i$, respectively. α_{adv} is the penalty coefficient for the two losses. The network parameters are adversarially trained to achieve

$$(\hat{\theta}_C, \hat{\theta}_G) = \arg\min_{\boldsymbol{\theta}_C, \boldsymbol{\theta}_G} L_{\text{adv}}(\boldsymbol{\theta}_C, \boldsymbol{\theta}_G, \hat{\theta}_D)$$

$$\hat{\theta}_D = \arg\max_{\boldsymbol{\theta}_D} L_{\text{adv}}(\hat{\theta}_C, \hat{\theta}_G, \theta_D) \tag{3}$$

where $\hat{\theta}_C$, $\hat{\theta}_G$, and $\hat{\theta}_D$ represent the optimal values of θ_C, $\boldsymbol{\theta}_G$, and $\boldsymbol{\theta}_D$, respectively.

IV. Proposed Method

A. Motivation

The open-set domain adaptation problem is investigated in this study, where the source label space is included in the target label space. That indicates along with the data distribution discrepancy, mismatches across domains in the label space also exist, which makes this problem quite challenging. In this literature, deep neural network-based domain adaptation methods have been successfully developed and achieved promising transfer learning effects. The domain adversarial learning strategy is one of the most effective approaches in such cases, where domain-invariant features can be learned through adversarial training. The marginal domain gap can be reduced, and the conventional domain adaptation problem can be well addressed. However, the open-set scenario has been paid much less attention, which is practical in the real industries. Direct application of most existing methods leads to class-level mismatches across domains, as illustrated in Fig. 1.

Specifically, based on the adversarial learning strategy, if all the source and target data instances are treated equally in (3), the discriminator cannot identify the original domains of the data that indicates all the source and target instances are projected into the same area in the high-level subspace. In the open-set problem, the target-domain outliers are supposed to be filtered out in adversarial learning. Promising class-level alignments can be expected, if the target data in the unknown health states are discarded in adversarial training. Therefore, it is motivated that additional weight items can be attached to the target-domain instances to indicate their class-level similarities with the source-domain health states. The target outliers can be effectively neglected in adversarial training as in (3), if larger weights are attached to the shared health states, and lower values are imposed to the outliers. In this way, domain adaptation can be selectively implemented between the shared classes across domains, rather than including the outliers. Correspondingly, an instance-level weighted mechanism in the adversarial learning framework is proposed in this article to address this issue.

It should be pointed out the class-level weights have also been employed in related studies in the literature. The partial domain adaptation scenario is more suited for class-level weights, where the target classes are included in the source classes. Labeled source data are available, and class-level weights can be used to indicate the similarities with the target domain. However,

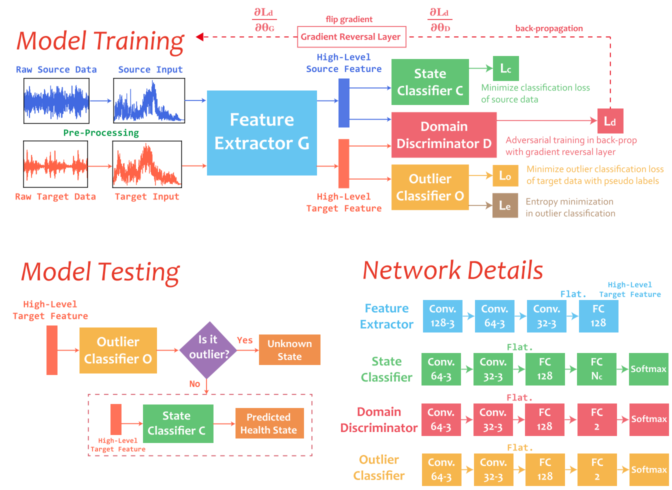


Fig. 2. Overview of the proposed method in training and testing, and the deep neural network architecture.

unlabeled target outliers exist in the open-set domain adaptation problem, and instance-level weights are, thus, more suitable to measure the similarity of each unlabeled target sample to the source domain.

B. Deep Neural Network Model

The overview of the proposed method and the deep neural network architecture are presented in Fig. 2. Generally, four modules are used, i.e., feature extractor G, domain discriminator D, state classifier C, and outlier classifier O. The feature extractor consists of 3 convolutional layers with filter number of 128, 64, and 32. After flattening, one fully connected layer with 128 neurons is adopted, which is the learned high-level representations of the raw data. The domain discriminator contains two convolutional layers with 64 and 32 filters. Two fully connected layers with 128 and two neurons are used next, and a softmax function is adopted to interpret the domain label prediction confidence values to probabilities.

In general, the state classifier and outlier classifier share similar network architectures with that of the domain discriminator. Two neurons are also used in the last layer of the outlier classifier to represent known and unknown classes. N_c neurons are adopted in the last layer of the state classifier to denote the confidence values of different health states, where N_c is the number of the source classes. Throughout the network, filter size of 3 is adopted for the convolutional layers, and the leaky rectified linear units functions are generally used for activations.

C. Instance-Level Weighted Domain Adaptation

In order to bridge the label space discrepancy in the open-set domain adaptation problem, an instance-level weighted mechanism is proposed to ignore the target outlier classes in domain alignment. Specifically, additional weights are attached to the target-domain instances, which represent their similarities with the source-domain classes. Recall that in adversarial learning, the domain discriminator aims to separate the input data into different domains, while the feature extractor is trained to obtain generalized features to make the discriminator less effective. Therefore, it is expected that the domain discriminator can better identify the outlier instances as the target-domain data, which are more different with the source data, and the target instances in the shared health states, which are more similar with the source data can be less recognized. This fact suggests the domain label prediction error can well serve as the indicator for the instance-level similarity weights.

In this study, the cross-entropy loss function between the predicted domain label and the ground truth is used to evaluate the prediction error. Concretely, the loss function $L^s_{d,i}$ for the ith source sample, and $L^t_{d,i}$ for the ith target sample can be defined as

$$L^s_{d,i} = -\sum_{k=1}^{2} 1\left\{d^s_i = k\right\} \log \frac{e^{x^s_{d,i,k}}}{\sum_{m=1}^{2} e^{x^s_{d,i,m}}}$$
$$L^t_{d,i} = -\sum_{k=1}^{2} 1\left\{d^t_i = k\right\} \log \frac{e^{x^t_{d,i,k}}}{\sum_{m=1}^{2} e^{x^t_{d,i,m}}} \quad (4)$$

where d^s_i and d^t_i denote the source and target domain labels, respectively. $x^s_{d,i,k}$ and $x^t_{d,i,k}$ are the kth elements of the output at the last layer in the domain discriminator for the source and target samples, respectively. In this way, the raw similarity weight $w_{r,i}$ for the ith target-domain sample with the source classes can be assigned as $L^t_{d,i}$. That corresponds with the fact that the target outlier samples can be generally better distinguished with small prediction errors, and the small similarity weights are, thus, attached.

Furthermore, it is noted that normalization is required for proper scaling the raw weight values, and the min–max normalization approach is used in this study

$$w_i = \gamma \frac{w_{r,i} - w_{\min}}{w_{\max} - w_{\min} + \varepsilon}$$
$$w_{\max} = \max(w_{r,i}), w_{\min} = \min(w_{r,i}) \quad (5)$$

where w_i denotes the normalized value of $w_{r,i}$, γ represents a scaling coefficient, and ε is a small positive value. In this study, γ is set as one for simplicity as indicated in Table IV, and it can be also changed based on the specific scenarios.

After the normalized weights are obtained, they are attached to the target instances in adversarial learning for domain adaptation. In this way, larger weights are employed on the target instances in the shared health states in domain alignment, and the outlier target instances can be ignored with smaller weights.

D. Classification Training

As shown in Fig. 2, three classification modules are considered in the proposed method. The state classifier aims to identify the machinery health states under source supervision, and the cross-entropy loss function $L^s_{c,i}$ for the ith source sample is defined as

$$L^s_{c,i} = -\sum_{k=1}^{N_c} 1\left\{y^s_i = k\right\} \log \frac{e^{x^s_{c,i,k}}}{\sum_{m=1}^{N_c} e^{x^s_{c,i,m}}} \quad (6)$$

where $x^s_{c,i,k}$ denotes the kth element of the output at the last layer in the state classifier.

The domain discriminator aims to classify the domains from which the data are sampled. Therefore, the two-class classification problem is a supervised task, and the cross-entropy loss functions for training the domain discriminator module are presented in (4).

Moreover, the outlier target states are supposed to be correctly identified by the outlier classifier O in the proposed method. In order to train this module in a supervised manner, labeled data are required. However, no target label space information is available. Therefore, in this study, pseudo outlier labels are proposed for the target instances. As described in Section IV-C, the domain classification error $L^t_{d,i}$ can be a promising indicator for outliers. The target instances from the shared classes with the source domain are generally difficult to be distinguished, resulting in larger errors. Meanwhile, the target outlier instances can be more different with the source samples, leading to small prediction errors.

Specifically at each training epoch, with respect to the ranking of the normalized target-domain similarity weights w_i, $i = 1, 2, \ldots, n_t$, the instances above the $1 - \rho$ percentile are considered from the shared classes, and those below the ρ percentile are the outliers. In this way, pseudo outlier labels can be attached to the unlabeled target instances for training the module O. It should be noted that only a portion of the target instances are used in this scenario, which can be generally identified with higher level of confidence. The loss function $L^t_{o,i}$ for the ith target sample can be, thus, formulated as

$$L^t_{o,i} = -\sum_{k=1}^{2} 1\left\{o_i = k\right\} \log \frac{e^{x_{o,i,k}}}{\sum_{m=1}^{2} e^{x_{o,i,m}}} \quad (7)$$

where o_i denotes the pseudo outlier label, and $x_{o,i,k}$ represents the kth element of the output at the last layer in the outlier classifier. $i \in S_{\text{pseudo}}$ where S_{pseudo} denotes the index set of the target instances, which are attached with pseudo outlier labels at each training epoch.

E. Entropy Minimization

While the pseudo outlier labels can mostly well indicate the class information, it is still difficult to guarantee the convergence for the module O in optimization, due to the uncertainty and instability of the pseudo labels. In this study, the entropy minimization principle [26], [27] is further adopted that encourages the low-density classification boundaries through minimization of the entropy of the class-conditional distribution. Specifically, the loss function L^i_e for the ith target instance is defined as

$$L^t_{e,i} = -\sum_{k=1}^{2} x_{o,i,k} \log x_{o,i,k}. \quad (8)$$

By minimizing the entropy loss function for all the target instances, the outlier classifier can be gradually trained to pass through the low-density area of the target data. That largely enhances the stability and class-level separability of the decision boundary in detecting target outlier classes.

F. Optimization Objectives

In summary, the general optimization objectives in this study can be formulated as

$$L_c = \frac{1}{n_s}\sum_{i=1}^{n_s} L_{c,i}^s$$
$$L_d = \frac{1}{n_s}\sum_{i=1}^{n_s} L_{d,i}^s + \frac{1}{n_t}\sum_{i=1}^{n_t} w_i L_{d,i}^t$$
$$L_e = \frac{1}{n_t}\sum_{i=1}^{n_t} L_{e,i}^t, L_o = \frac{1}{|S_{\text{pseudo}}|}\sum_{i\in S_{\text{pseudo}}} L_{o,i}^t. \tag{9}$$

The network optimization is carried out to achieve

$$\hat{\theta}_G = \arg\left\{\min_{\boldsymbol{\theta}_G} L_c + \alpha_e L_e + \alpha_o L_o, \max_{\boldsymbol{\theta}_G} \alpha_d L_d\right\}$$
$$\hat{\theta}_D = \arg\min_{\boldsymbol{\theta}_D} \alpha_d L_d, \hat{\theta}_C = \arg\min_{\boldsymbol{\theta}_C} L_c$$
$$\hat{\theta}_O = \arg\min_{\boldsymbol{\theta}_O} \alpha_o L_o + \alpha_e L_e \tag{10}$$

where α_d, α_e, and α_o denote the penalty coefficients for L_d, L_e, and L_o, respectively. θ_O is the parameters in the outlier classifier module, and $\hat{\theta}_O$ denotes the optimal value of $\boldsymbol{\theta}_O$. The optimization problem can be solved using the popular stochastic gradient descent-based methods, and the parameters can be updated at each training epoch as

$$\boldsymbol{\theta}_G \leftarrow \boldsymbol{\theta}_G - \delta\left(\frac{\partial L_c}{\partial \boldsymbol{\theta}_G} + \alpha_e\frac{\partial L_e}{\partial \boldsymbol{\theta}_G} + \alpha_o\frac{\partial L_o}{\partial \boldsymbol{\theta}_G} - \alpha_d\frac{\partial L_d}{\partial \theta_G}\right)$$
$$\boldsymbol{\theta}_D \leftarrow \boldsymbol{\theta}_D - \delta\alpha_d\frac{\partial L_d}{\partial \boldsymbol{\theta}_D}, \boldsymbol{\theta}_C \leftarrow \boldsymbol{\theta}_C - \delta\frac{\partial L_c}{\partial \boldsymbol{\theta}_C}$$
$$\boldsymbol{\theta}_O \leftarrow \boldsymbol{\theta}_O - \delta\left(\alpha_o\frac{\partial L_o}{\partial \boldsymbol{\theta}_O} + \alpha_e\frac{\partial L_e}{\partial \boldsymbol{\theta}_O}\right). \tag{11}$$

By using the gradient reversal layer, which flips the signs of gradients in back-propagation [22], the parameter updates can be simultaneously carried out in one step. In this way, the open-set domain adaptation problem can be well addressed, where the target outlier classes can be effectively filtered out in transfer learning, thus achieving better class-level alignments across domains.

V. Experimental Study

A. Dataset Descriptions

1) Case Western Reserve University (CWRU). First, the rolling element bearing dataset provided by the Bearing Data Center of CWRU [28] is used for validation. The time-domain vibration acceleration data are measured from the motor drive end with two rotating speeds, i.e., 1730 and 1797 r/min. Four bearing health states are considered, i.e., healthy (H), inner race fault (IF), ball fault (BF), and outer race fault (OF). Three levels of fault severities are also focused on, whose fault diameters are 7, 14, and 21 mils, respectively.

2) Train bogie. A more challenging experimental dataset is further used for validations, which is from a test rig of the high-speed multiunit train bogie rolling bearing system. Two operating conditions with rotating speeds of 1590 and 1950 r/min are implemented, which are associated with train speeds of 260 and 320 km/h, respectively. The vibration acceleration data are measured for analysis. Besides the healthy state (H), three fault modes are artificially created on different bearing components, i.e., IF, roller fault (RF), and OF. Three levels of fault severities are also considered, i.e., incipient, medium, and sever faults. Fig. 3 shows the test rig and different fault modes in this dataset. The detailed information of the two rotating machinery datasets in this study is presented in Table I.

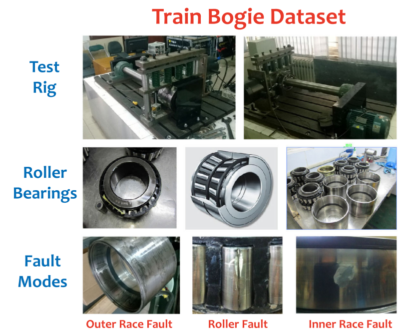


Fig. 3. Test rig and fault modes in the train bogie dataset [29].

TABLE I
INFORMATION OF THE TWO DATASETS

Dataset	Health State Label	1	2	3	4	5	6	7	8	9	10
CWRU	Fault Location	H	IF	IF	IF	BF	BF	BF	OF	OF	OF
	Fault Size (mil)	0	7	14	21	7	14	21	7	14	21
Train Bogie	Fault Location	H	IF	IF	IF	RF	RF	RF	OF	OF	OF
	Fault Severity	N/A	Inc.	Med.	Sev.	Inc.	Med.	Sev.	Inc.	Med.	Sev.

B. Compared Approaches

1) Basic

First, a basic deep learning method is implemented to serve as a baseline for comparisons, which follows the conventional supervised learning scheme. Only the cross-entropy loss function in (6) with source supervision is considered.

2) MMD-DA, correlation alignment (CORAL), and domain adversarial neural network (DANN)

The popular domain adaptation methods are evaluated in this study. First, the distance metric-based method aims at minimizing source and target domain distance in the feature space. The MMD-DA method is implemented, where the MMD metric at the high-level representations is optimized to achieve marginal distribution alignment across domains. The specific settings follow the popular related study in [18]. Next, the deep CORAL method [30] minimizes the domain shift through alignment of the second-order statistics of source and target features, which is well suited for unsupervised domain adaptation.

Meanwhile, adversarial training has been an effective alternative approach for transfer learning, and the DANN framework [31] is compared in this study. Similar with the proposed method, DANN focuses on adversarial learning

TABLE II
INFORMATION OF THE OPEN-SET DOMAIN ADAPTATION TASKS ON THE CWRU DATASET

Task	Source to Target	Source States	Openness
A_1	1730− > 1797	All	0
A_2	1730− > 1797	1,2,3,4,5,6,7,8,9	0.1
A_3	1730− > 1797	1,2,3,4,8,9,10	0.3
A_4	1730− > 1797	1,4,5,6,8,9	0.4
A_5	1730− > 1797	1,2,3,8,9,10	0.4
A_6	1730− > 1797	1,2,7,8	0.6
A_7	1730− > 1797	1,5,6	0.7
A_8	1797− > 1730	1,2,5,6,7,8,9,10	0.2
A_9	1797− > 1730	1,2,3,4,7,8	0.4
A_{10}	1797− > 1730	1,8,9,10	0.6

TABLE III
INFORMATION OF THE OPEN-SET DOMAIN ADAPTATION TASKS ON THE TRAIN BOGIE DATASET

Task	Source to Target	Source States	Openness
B_1	1590− > 1950	All	0
B_2	1590− > 1950	1,2,3,5,6,7,8,9,10	0.1
B_3	1590− > 1950	1,2,4,5,6,8,9,10	0.2
B_4	1590− > 1950	1,2,3,7,8,9,10	0.3
B_5	1590− > 1950	1,4,5,6,8,9	0.4
B_6	1590− > 1950	1,2,7,8,9	0.5
B_7	1590− > 1950	1,3,5,7,10	0.5
B_8	1950− > 1590	1,2,4,5,6,7,9,10	0.2
B_9	1950− > 1590	1,2,3,4,6,7,8	0.3
B_{10}	1950− > 1590	1,5,7,8,9,10	0.4

between source and target domains as (4) indicates, while the other techniques are not adopted.

3) Open-set support vector machine (SVM) (OSVM), OSVM-MMD, and OSVM-DANN
The OSVM [32] is an effective method for detecting outliers, which leverages the threshold probability to identify the samples as unknown when the predicted probability is smaller than the threshold. In this study, the proposed network structure is first used for training under the supervised learning paradigm. Afterward, the data representations learned by the feature extractor are focused on, and the OSVM algorithm is implemented.
In order to further integrate transfer learning techniques with OSVM, the OSVM-MMD, and OSVM-DANN are also evaluated. Specifically, in the first stage of feature learning, the MMD metric between source and target data is also minimized, and the OSVM algorithm is used afterward. This method is denoted as OSVM-MMD. Similarly, by combining the DANN approach and OSVM, the transfer learning effect can be also achieved with outlier detection ability that is denoted as OSVM-DANN.

4) NoEntropy
In order to evaluate the benefits of the proposed entropy minimization scheme, the NoEntropy method is carried out where the entropy minimization part is removed from the proposed method. Specifically, the objective in (8) is not considered in the model training.

C. Fault Diagnosis Tasks and Implementation Details

In this study, different fault diagnosis tasks for the open-set domain adaptation problem are investigated. Specifically, Tables II and III show the detailed settings of the concerned tasks.

TABLE IV
PARAMETERS USED IN THIS ARTICLE

Parameter	Value	Parameter	Value
Epochs	5000	α_d	1
batch size	256	α_e	1
δ	1e-5	α_o	1
N_{input}	256	ρ	0.1
ε	1e-4	γ	1

In general, it is assumed that 100 labeled samples are available for each source-domain class, and 100 unlabeled samples of each target-domain class are used for testing. In all the tasks, the target domain includes data of all the ten machine health states, while the source domain only contains partial classes. Different openness is also considered, ranging from 0, which is nonopen set problem, to 0.7, which is highly biased for domain adaptation.

In this article, the reported numerical results in this study are generally averaged by ten trials to reduce the effect of model randomness. The detailed parameters are presented in Table IV. They are basically determined from the validation task, where the CWRU dataset is used. The source domain is under the 1730 r/min rotating speed, and the target domain is in the 1750 r/min condition. The open-set domain adaptation task is randomly formulated, and the source-domain health states include the classes 1, 2, 4, 6, 7, and 10. The validation accuracy on the target domain is used as the metric for determination of the selected model parameters. It should be noted that the validation task is weakly related with the actual concerned tasks, since different regimes are considered.

D. Experimental Results and Performance Analysis

1) Cross-Domain Diagnosis Results: In this study, the source domain mostly covers partial health states in different tasks. To evaluate the testing performance, all the target outlier classes are considered as the unknown class. Fig. 4 shows the testing accuracies by different methods in different open-set domain adaptation tasks on the CWRU dataset. The general results are presented, which indicate the testing accuracies for all the target instances. The testing accuracies on the known and unknown classes are provided in Fig. 5, which show the performances for the target instances in the shared health states with the source domain, and those for the target outlier instances, respectively. The detailed fault diagnosis result on each machine fault mode in the task A_3, i.e., the confusion matrix, is shown in Fig. 6. The performances of the representative transfer learning method DANN and the proposed approach are presented for comparisons.

It can be observed that the proposed method generally outperforms the other approaches in all the concerned tasks. Higher than 80% testing accuracies can be mostly achieved by the proposed method. Due to the significant difficulty in open-set domain adaptation, the basic method implementing conventional supervised learning basically fails in different tasks. The MMD-DA, DANN, and CORAL methods have been proved in the recent literature to achieve promising transfer learning effects, and close to 100% testing accuracies are indeed obtained

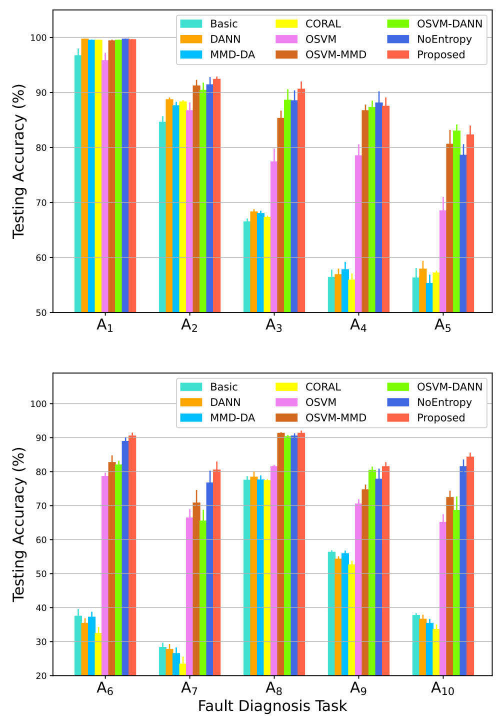


Fig. 4. Testing accuracies in the open-set domain adaptation tasks by different methods on the CWRU dataset.

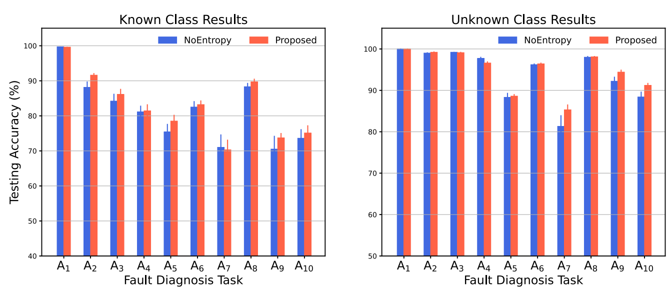


Fig. 5. Testing accuracies for the known and unknown health states in the open-set domain adaptation tasks on the CWRU dataset.

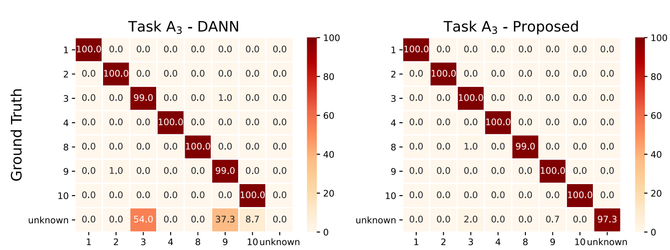


Fig. 6. Confusion matrix in the task A_3 by the proposed approach and the DANN method.

in the nonopen-set domain adaptation task A_1. However, with the interference of the target outliers, all the methods focusing on the marginal distribution alignment between source and target domains, cannot achieve the good class-level alignment. As a consequence, low testing accuracies are obtained in the open-set tasks.

The OSVM method is an effective approach for outlier detection, and it obtains noticeable improvements in the testing

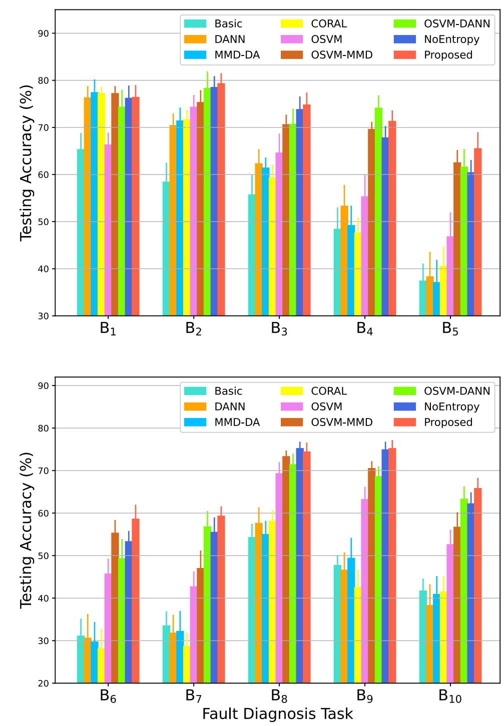


Fig. 7. Testing accuracies in the open-set domain adaptation tasks by different methods on the train bogie dataset.

accuracies compared with the baselines in most cases. However, the domain shift issue cannot be well addressed. By integrating transfer learning techniques with OSVM, the OSVM-MMD, and OSVM-DANN methods have achieved significant improvements. They can identify the known health states by extracting shared features across domain for fault diagnosis, and also recognize the unknown class. Promising results have been obtained in most cases. However, since the knowledge transfer part in both the methods is focused on the marginal distributions, they are still less competitive compared with the proposed frameworks. The confusion matrix also validates the effectiveness of the proposed method for the specific predictions in each case.

On the other hand, the proposed method can well address this challenging problem. High classification accuracies on the target instances can be achieved for the known health states, and the outlier states also can be effectively detected. The benefits of the proposed entropy minimization scheme are also shown in the results. The noticeable performance increases can be observed between the proposed method and the NoEntropy approach with respect to both known and unknown classes, which further validate the proposed open-set domain adaptation algorithm.

The experimental results of the train bogie dataset are shown in Figs. 7–9. In general, the similar performance patterns are still observed with those on the CWRU dataset. The proposed method significantly outperforms the compared approaches. The conventional domain adaptation methods MMD-DA and DANN basically fail in the open-set cross-domain fault diagnosis problems, and limited improvements are obtained compared with the baseline method Basic. The proposed framework is well

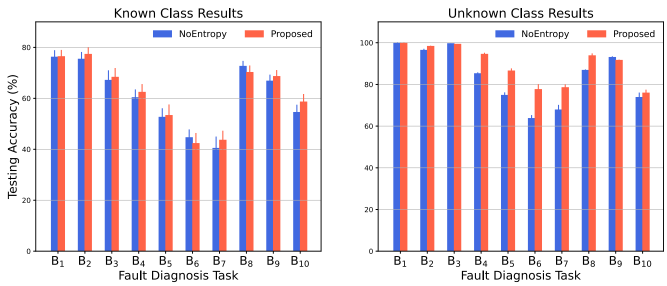


Fig. 8. Testing accuracies for the known and unknown health states in the open-set domain adaptation tasks on the train bogie dataset.

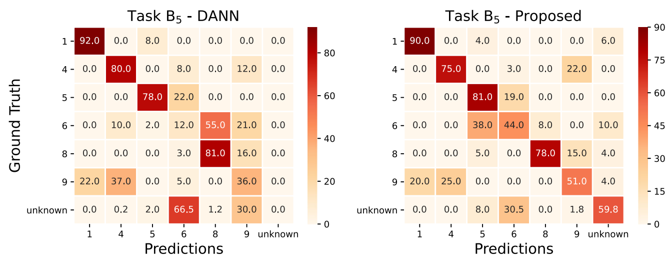


Fig. 9. Confusion matrix in the task B_5 by the proposed approach and the DANN method.

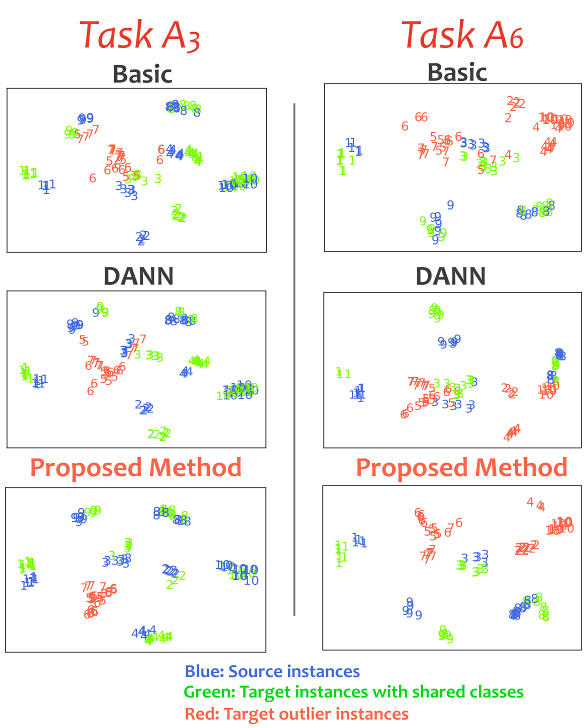


Fig. 10. Visualizations of the learned features by different methods in the tasks A_3 and A_6. The numbers denote the class labels of the source and target-domain instances.

suited for the challenging problem, and the positive effects of the entropy minimization technique are further shown.

2) Visualizations: In this section, the performance of the proposed method is intuitively investigated with visualizations. Specifically, the learned high-level data representations by the feature extractor module are focused on with the t-SNE method.

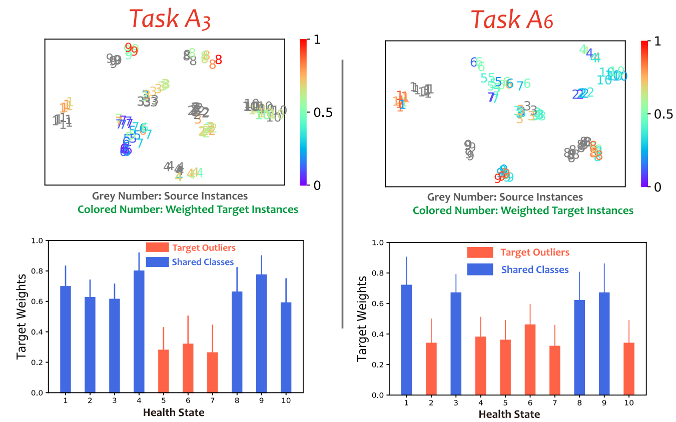


Fig. 11. Upper panel shows the visualizations of the learned features and weights by the proposed method in the tasks A_3 and A_6. The numbers denote the class labels of the instances. The lower panel shows the corresponding mean values and standard deviations of the weights of the target instances in each class.

Fig. 10 shows the visualization results by different methods on the open-set domain adaptation tasks A_3 and A_6.

It can be observed from the results in the task A_3 that the proposed method is able to achieve promising features for class-level alignment across domains. The source and target-domain instances in the shared health states can be effectively projected into the same regions in the high-level representation space, and the target outlier classes can be properly isolated in separate areas. In this way, the diagnostic knowledge learned from the labeled source domain can be well transferred to the target domain, and the unknown target classes can be also identified. However, for the basic and DANN methods, noticeable overlappings between different classes across domains are observed, which result in the less effective cross-domain diagnosis performance under source supervision. Similar patterns are still observed in the task A_6 with larger openness. The visualizations of the learned features in this section intuitively validate the proposed method in the open-set domain adaptation problems.

3) Performance Investigations: Next, the learned instance-level weights of the target-domain data in different tasks are investigated, and the results are presented in Fig. 11. It is observed that in general, large weights are learned for the target instances, which share the same health state labels with the source-domain data. For the target outlier classes, small weights are usually obtained. For instance, with respect to the task A_3, the target outlier instances with labels 5–7 generally achieve low level of weight values, and they are, thus, less considered in the source and target distribution alignments. The other target instances of the shared labels with the source domain basically have higher level of weight values, and they are paid more attention in adversarial learning for domain adaptation. The statistics of the instance-level weights in different classes further validate the proposed method. The results of the task A_6 also follow similar patterns.

Moreover, the model sensitivities to the parameters are investigated, and the results in the validation task are presented in Fig. 12. The influence of multiple network parameters is

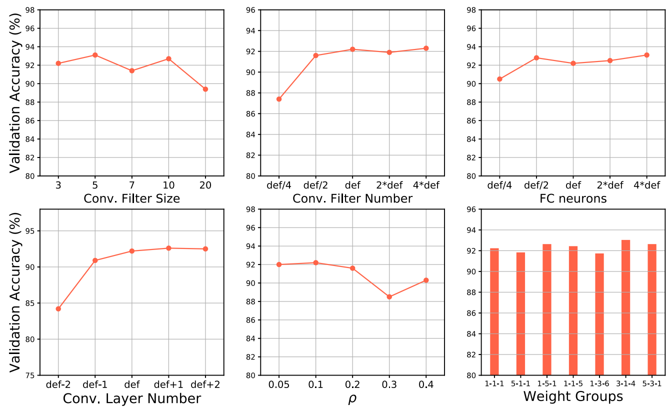


Fig. 12. Effects of different parameters on the model performance. "def" in the x-axis labels denote the default values of different parameters, and different times of the default values are investigated. The variations of the convolutional filter size, filter numbers, fully connected (FC) neurons, and convolutional layer numbers represent those for the entire network. For instance, the weight group 1–3–6 represents $\alpha_d = 1$, $\alpha_e = 3$, and $\alpha_o = 6$.

examined. It can be observed that the detailed neural network architecture has limited influence on the model performance. When sufficiently large network is used, the promising and stable performance can be generally achieved. However, the network with insufficient learning capability results in the low model performance. The effects of the threshold parameter ρ in pseudo outlier label setting and different weight coefficients α_d, α_e, and α_o in the optimization objectives are also investigated. The results indicate that relatively small ρ leads to a better model performance in general. No significant influence of the weight coefficients is observed under the condition that they are in a reasonable range with similar scale.

VI. Conclusion

In this article, a deep learning-based open-set domain adaptation method was proposed for machinery fault diagnosis problems. The adversarial learning scheme was introduced to obtain domain-invariant features from source and target domains. To identify the target outlier classes, an instance-level weighted mechanism was proposed, which reflected the similarities of the target instances with the source classes. The weights were obtained using the score of the discriminator in deep learning framework. Entropy minimization technique was further adopted to enhance the model generalization ability. Experiments on two practical rotating machine datasets were carried out for validations. It is observed that high cross-domain fault diagnosis testing accuracies can be achieved by the proposed method, and the target outlier classes can be effectively recognized.

Despite the promising results, it should be pointed out that this study focused on the open-set domain adaptation problem where the source classes were included in the target label space. With respect to generalized transfer learning, the partial domain adaptation problem was also practical in the real industrial cases, where the target classes were contained in the source label space. More generalized framework of domain adaptation considering both scenarios is supposed to be developed in the following works. Meanwhile, the unlabeled target-domain data are assumed to be available in the proposed framework, which indicates the methodology is more suited for offline cross-domain fault diagnosis. The following research works will be also focused on the online fault diagnosis paradigm, which is more practical in the real industries.

References

[1] Y. Lei, B. Yang, X. Jiang, F. Jia, N. Li, and A. K. Nandi, "Applications of machine learning to machine fault diagnosis: A review and roadmap," *Mech. Syst. Signal Process.*, vol. 138, 2020, Art. no. 106587.

[2] W. Zhang, X. Li, H. Ma, Z. Luo, and X. Li, "Federated learning for machinery fault diagnosis with dynamic validation and self-supervision," *Knowl.-Based Syst.*, vol. 213, 2021, Art. no. 106679.

[3] G. Csurka, "Domain adaptation for visual applications: A comprehensive survey," 2017, *arXiv:1702.05374*.

[4] M. Long, Y. Cao, J. Wang, and M. Jordan, "Learning transferable features with deep adaptation networks," in *Proc. 32nd Int. Conf. Mach. Learn.*, 2015, vol. 37, pp. 97–105.

[5] J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, "How transferable are features in deep neural networks?" in *Proc. 27th Int. Conf. Neural Infor. Process. Syst.*, 2014, pp. 3320–3328.

[6] H. Liu, Z. Cao, M. Long, J. Wang, and Q. Yang, "Separate to adapt: Open set domain adaptation via progressive separation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 2927–2936.

[7] Z. Fang, J. Lu, F. Liu, J. Xuan, and G. Zhang, "Open set domain adaptation: Theoretical bound and algorithm," 2019, *arXiv:1907.08375*.

[8] P. P. Busto and J. Gall, "Open set domain adaptation," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2017, pp. 754–763.

[9] W. Zhang, X. Li, X.-D. Jia, H. Ma, Z. Luo, and X. Li, "Machinery fault diagnosis with imbalanced data using deep generative adversarial networks," *Measurement*, vol. 152, 2020, Art. no. 107377.

[10] R. Razavi-Far, E. Hallaji, M. Farajzadeh-Zanjani, and M. Saif, "A semi-supervised diagnostic framework based on the surface estimation of faulty distributions," *IEEE Trans. Ind. Informat.*, vol. 15, no. 3, pp. 1277–1286, Mar. 2019.

[11] W. Zhang, X. Li, and X. Li, "Deep learning-based prognostic approach for lithium-ion batteries with adaptive time-series prediction and on-line validation," *Measurement*, vol. 164, 2020, Art. no. 108052.

[12] W. Zhang, X. Li, and Q. Ding, "Deep residual learning-based fault diagnosis method for rotating machinery," *ISA Trans.*, vol. 95, pp. 295–305, 2019.

[13] C. Lu, Z. Y. Wang, W. L. Qin, and J. Ma, "Fault diagnosis of rotary machinery components using a stacked denoising autoencoder-based health state identification," *Signal Process.*, vol. 130, pp. 377–388, 2017.

[14] K. Zhang, B. Tang, Y. Qin, and L. Deng, "Fault diagnosis of planetary gearbox using a novel semi-supervised method of multiple association layers networks," *Mech. Syst. Signal Process.*, vol. 131, pp. 243–260, 2019.

[15] X. Wang and J. Schneider, "Flexible transfer learning under support and model shift," *in Proc. 27th Int. Conf. Neural Inf. Process. Syst.*, 2014, pp. 1898–1906.

[16] X. Li and W. Zhang, "Deep learning-based partial domain adaptation method on intelligent machinery fault diagnostics," *IEEE Trans. Ind. Electron.*, to be published, doi: 10.1109/TIE.2020.2984968.

[17] Z. Chen, K. Gryllias, and W. Li, "Intelligent fault diagnosis for rotary machinery using transferable convolutional neural network," *IEEE Trans. Ind. Informat.*, vol. 16, no. 1, pp. 339–349, Jan. 2019.

[18] W. Lu, B. Liang, Y. Cheng, D. Meng, J. Yang, and T. Zhang, "Deep model based domain adaptation for fault diagnosis," *IEEE Trans. Ind. Electron.*, vol. 64, no. 3, pp. 2296–2305, Mar. 2017.

[19] L. Wen, L. Gao, and X. Li, "A new deep transfer learning based on sparse auto-encoder for fault diagnosis," *IEEE Trans. Syst., Man, Cybern. Syst.*vol. 49, no. 1, pp. 136–144, Jan. 2019.

[20] E. Tzeng, J. Hoffman, K. Saenko, and T. Darrell, "Adversarial discriminative domain adaptation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 2962–2971.

[21] L. Guo, Y. Lei, S. Xing, T. Yan, and N. Li, "Deep convolutional transfer learning network: A new method for intelligent fault diagnosis of machines with unlabeled data," *IEEE Trans. Ind. Electron.*, vol. 66, no. 9, pp. 7316–7325, Sep. 2019.

[22] X. Li, W. Zhang, Q. Ding, and X. Li, "Diagnosing rotating machines with weakly supervised data using deep transfer learning," *IEEE Trans. Ind. Informat.*, vol. 16, no. 3, pp. 1688–1697, Mar. 2020.

[23] Z. Cao, M. Long, J. Wang, and M. Jordan, "Partial transfer learning with selective adversarial networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 2724–2732.

[24] J. Jiao, M. Zhao, J. Lin, and C. Ding, "Classifier inconsistency based domain adaptation network for partial transfer intelligent diagnosis," *IEEE Trans. Ind. Informat.*, vol. 16, no. 9, pp. 5965–5974, Sep. 2020.

[25] K. Saito, S. Yamamoto, Y. Ushiku, and T. Harada, "Open set domain adaptation by backpropagation," in *Proc. Eur. Conf. Comput. Vis.*, 2018, pp. 153–168.

[26] M. Long, H. Zhu, J. Wang, and M. I. Jordan, "Unsupervised domain adaptation with residual transfer networks," in *Proc. 30th Conf. Neural Inf. Process. Syst.*, 2016, pp. 136–144.

[27] Y. Grandvalet and Y. Bengio, "Semi-supervised learning by entropy minimization," in *Proc. 17th Int. Conf. Neural Inf. Process. Syst.*, 2004, pp. 529–536.

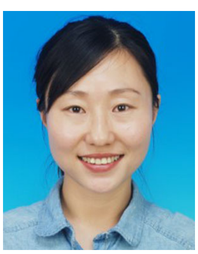
[28] W. A. Smith and R. B. Randall, "Rolling element bearing diagnostics using the case western reserve university data: A benchmark study," *Mech. Syst. Signal Process.*, vol. 64–65, pp. 100–131, 2015.

[29] X. Li, W. Zhang, N.-X. Xu, and Q. Ding, "Deep learning-based machinery fault diagnostics with domain adaptation across sensors at different places," *IEEE Trans. Ind. Electron.*, vol. 67, no. 8, pp. 6785–6794, Aug. 2020.

[30] B. Sun and K. Saenko, "Deep CORAL: Correlation alignment for deep domain adaptation," in *Proc. Eur. Conf. Comput. Vis.*, 2016, pp. 443–450.

[31] Y. Ganin and V. Lempitsky, "Unsupervised domain adaptation by backpropagation,"2015, *arXiv:1409.7495*.

[32] L. P. Jain, W. J. Scheirer, and T. E. Boult, "Multi-class open set recognition using probability of inclusion," in *Proc. Eur. Conf. Comput. Vis.*, 2014, pp. 393–409.



Wei Zhang received the Ph.D. degree in mechanics from Tianjin University, Tianjin, China, in 2017.

She was a Visiting Scholar with Texas A & M University, College Station, TX, USA, in 2015. She is currently an Assistant Professor with the School of Aerospace Engineering, Shenyang Aerospace University, Shenyang, China. Her research interests include rotor dynamics, fault diagnosis, multiobjective optimization algorithm, and squeeze film damper.

Xiang Li received the B.S. degree in engineering mechanics and the Ph.D. degree in mechanics from Tianjin University, Tianjin, China, in 2012 and 2017, respectively.

He is currently an Associate Professor with College of Sciences, and Key Laboratory of Vibration and Control of Aero-Propulsion System Ministry of Education, Northeastern University, Shenyang, China. His research interests include industrial AI, deep learning, fault diagnosis, and prognostics and health management.

Hui Ma received the M.Sc. and Ph.D. degrees in mechanical engineering from Northeastern University, Shenyang, China, in 2004 and 2007, respectively.

He is currently a Professor with the School of Mechanical Engineering and Automation, Northeastern University. His research interests include dynamic modeling and fault diagnosis of rotating machinery, vibration characteristic analysis of rotating machinery, and machinery fault diagnosis.

Zhong Luo received the M.Sc. and Ph.D. degrees in mechanical engineering from Northeastern University, Shenyang, China, in 2004 and 2007, respectively.

He is currently a Professor with the School of Mechanical Engineering and Automation, Northeastern University. His research interests include dynamics of machinery and control, theory of model experiment, and machinery fault diagnosis.

Xu Li received the Ph.D. degree in material processing engineering from Northeastern University, Shenyang, China, in 2009.

He is currently a Full Professor with the State Key Laboratory of Rolling and Automation, Northeastern University. His research interests include rolling process, fault diagnosis, automation systems, and intelligent control.