# Integrating Expert Knowledge With Domain Adaptation for Unsupervised Fault Diagnosis

Qin Wang, Cees Taal, and Olga Fink, *Member, IEEE*

*Abstract*—Data-driven fault diagnosis methods often require abundant labeled examples for each fault type. On the contrary, real-world data is often unlabeled and consists of mostly healthy observations and only few samples of faulty conditions. The lack of labels and fault samples imposes a significant challenge for existing data-driven fault diagnosis methods. In this article, we aim to overcome this limitation by integrating expert knowledge with domain adaptation (DA) in a synthetic-to-real framework for unsupervised fault diagnosis. Motivated by the fact that domain experts often have a relatively good understanding on how different fault types affect healthy signals, in the first step of the proposed framework, a synthetic fault dataset is generated by augmenting real vibration samples of healthy bearings. This synthetic dataset integrates expert knowledge and encodes class information about the faults types. However, models trained solely based on the synthetic data often do not perform well because of the distinct distribution difference between the synthetically generated and real faults. To overcome this domain gap between the synthetic and real data, in the second step of the proposed framework, an imbalance-robust DA approach is proposed to adapt the model from synthetic faults (source) to the unlabeled real faults (target) which suffer from severe class imbalance. The framework is evaluated on two unsupervised fault diagnosis cases for bearings, the CWRU laboratory dataset and a real-world wind-turbine dataset. Experimental results demonstrate that the generated faults are effective for encoding fault type information and the DA is robust against the different levels of class imbalance between faults.

*Index Terms*—Adversarial training, deep learning, domain adaptation (DA), fault diagnosis, simulation to real, synthetic to real.

## I. INTRODUCTION

**D**ATA-DRIVEN fault diagnosis methods often require a large number of labeled data to generalize well. Faults are, however, rare in real-world complex and safety-critical systems. Therefore, a sufficient number of representative samples of faulty conditions is often impossible to be collected in real-world applications. Recordings from industry assets often consist of a majority of healthy states and only few faults. In addition, not all fault types may have been captured by the different assets. Moreover, these recordings are often unlabeled because precisely identifying when and which fault is emerging can be difficult even for experienced domain experts. These real-world restrictions make learning fault patterns from unlabeled real-world data a challenging task.

One potential solution to overcome these challenges in this unsupervised fault diagnosis setup is to use synthetic faults as the supervision for the data-driven diagnosis models. For example, for bearing fault diagnosis, given operating conditions and bearing characteristics as input, synthetic vibration signals can be generated by highly accurate physical models, e.g., [1]. By generating a large number of synthetic faults, a data-driven model can be trained solely based on the synthetic faults and then evaluated on the real target data [2]–[5]. However, this way of using synthetic faults has several drawbacks. First, detailed operating conditions can be unknown in order to achieve a realistic physical model. For example, in bearing vibration modeling, we typically see that only the bearing is modeled ignoring surrounding rotating equipment, which can have a strong impact on the diagnosis performance. Second, even advanced simulations are not perfect, and there will always be a distribution gap between the synthetic data and experimental measurements of mechanical systems [6]. This domain gap between synthetic source and real target often leads to a significant performance degradation [7] if the model is solely trained on the synthetic data. Third, this pure synthetic data approach fails to make use of the available unlabeled real data, which can be potentially useful for providing additional information on the real faults. Finally, in reality, developing accurate physical models for assets can be expensive and time-consuming. For complex systems and complex physics of failure dynamics, detailed physics-based models may not even be available.

To tackle the challenges of synthetic fault generation outlined above, we propose a novel framework for unsupervised fault diagnosis which relaxes the need of highly accurate physical models, while performing well on the real target data. Motivated by the fact that domain experts often have a good understanding on how different fault types affect healthy signals, we propose to integrate expert knowledge in synthetic data with imbalance-robust domain adaptation (DA) for unsupervised fault diagnosis. Unlike previous works [6] which use faults generated by highly accurate physical simulation models as the supervision, the proposed framework

Qin Wang and Olga Fink are with the Chair of Intelligent Maintenance Systems, ETH Zürich, 8092 Zürich, Switzerland (e-mail: qwang@ethz.ch; ofink@ethz.ch).

Cees Taal is with the SKF Group, 3992 AE Houten, The Netherlands (e-mail: cees.taal@skf.com).

uses a relatively simple process decided by the expert for fault generation by augmenting healthy samples. To compensate on the potentially large domain gap between synthetic and real faults, the idea of DA is adopted. We propose to align the conditional distributions between synthetic and real features. The proposed DA approach relies less on the quality of the synthetic fault simulator and is robust to the class imbalance in the target domain. Specifically, the proposed framework consists of two complementary parts. In the first stage, expert knowledge is used to generate synthetic data. In the second stage, DA is applied on the synthetic source and real target data to alleviate the domain gap under severe class imbalance.

In this work, we apply our framework to bearing fault diagnosis. More specifically, we rely on expert knowledge and use a relative simple approach where we synthesize vibration signals with fault-initiated pulse trains corresponding to certain surface defects [1], [8]. We include more realistic vibration disturbances from surrounding rotating equipment, by mixing synthetic signals with healthy conditions. As a consequence, a reasonable number of samples can be generated for each fault, forming a balanced synthetic dataset. This synthetic dataset thus makes use of expert knowledge and can then facilitate the training of a data-driven fault diagnosis model. This proposed generation process relaxes the need of highly accurate physical models.

Although mixing healthy with synthetic defect signals will result in more realistic samples, a domain gap is still inevitable. To overcome this, the idea of unsupervised DA [9]–[11] offers a potential solution. DA aims at learning a representative model from the labeled synthetic source domain, and at the same time leveraging unlabeled data from the target domain to improve the model's generalization ability on the target data. Popular adversarial DA approaches [7], [12] use a domain discriminator to directly align the unconditional source and target distributions and minimize the discrepancy between them. They were originally designed and validated for image classification which often consists of balanced classes.

However, these standard adversarial DA methods make the implicit assumption that source and target domains share similar class distributions and directly align the source and target unconditional data without using class information. While this assumption is often realistic for the original image datasets, this can be largely violated in the unsupervised fault diagnosis setups because of the nature of rare faults. Given an unlabeled dataset from real-world (target) domain, its class distribution is most likely largely skewed toward the healthy class. Moreover, no more information can be inferred about the target class distribution or its level of class imbalance because the dataset is unlabeled. On the synthetic (source) side, samples are usually generated for each fault class based on the healthy samples, leading to a class-balanced synthetic dataset. If we directly adopt the standard adversarial DA methods, the alignment is then performed between a class-balanced synthetic source domain and a highly imbalanced real-world target domain consisting of a majority of healthy data. This mismatch can further lead to a performance drop on the target data. On the one hand, the different faults in the source domain would mostly be aligned to the healthy state
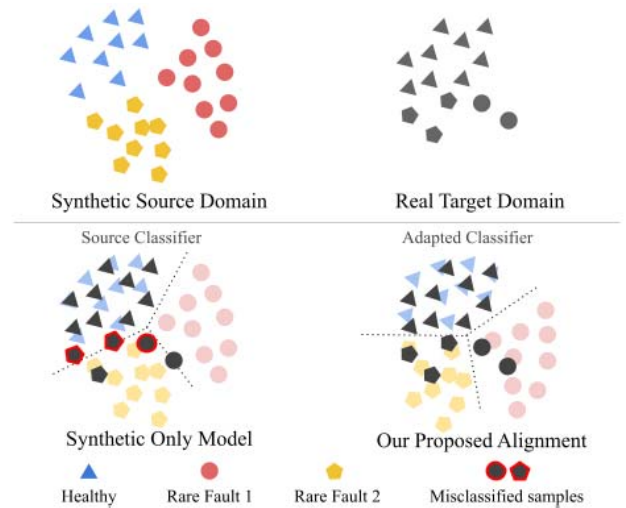


Fig. 1. Synthetic faults often suffer a distribution shift from the real faults, which leads to misclassified examples when models are trained only on the synthetic data. We propose to alleviate this problem by using DA to align the labeled synthetic domain with unlabeled target domain. The rare faults in the unlabeled target domain have very few samples, which makes the alignment harder. The proposed alignment method is robust to this class imbalance by aligning the conditional distributions. The model performance on the target domain can largely be improved by our approach.

which dominates the target data. On the other hand, rare faults in the target data would be poorly adapted.

To address the challenge of imbalanced datasets in unsupervised fault diagnostics setup, we propose a novel imbalance-robust DA approach which overcomes the imbalance problem by making better use of available class information for the alignment. Specifically, we design the discriminator to align the distributions conditioned on classes instead of the unconditional ones. This is achieved by feeding the class information encoded by pseudo-labels to the discriminator. In addition, to make the training of the discriminator more stable and provide more training examples from rare faults, we propose to use a mixup [13]-inspired augmentation to provide more support for the conditioned distributions on these rare classes. We illustrate the effect of the proposed DA approach in Fig. 1.

We demonstrate the effectiveness of our proposed framework on two different case studies. The first case study is based on the publicly available CWRU laboratory dataset. The second case study is based on the data collected from bearings of real wind turbine generators in the field. The proposed framework uses a relatively simple method for synthesizing the faults. However, it is still able to achieve a good performance with the help of the proposed DA approach.

While the evaluation of the proposed framework was performed on two bearing datasets, the proposed synthetic-to-real framework can be easily implemented for most industry assets. Moreover, the proposed imbalance-robust DA method can be generalized to other applications because it is in theory useful for any adaptation task which is facing the imbalanced data challenge between the source and target datasets.

Our contributions can be summarized as follows.

1) We propose an unsupervised fault diagnosis framework which builds on a simple fault generation process but

performs well with the help of DA. The generation-adaptation framework makes use of the power of both expert knowledge and DA to relieve the need of fault labels.

2) A novel imbalance-robust DA approach is proposed for unsupervised fault diagnosis which is robust against different imbalance levels between different health conditions.

3) To our knowledge, we provide the first publicly available synthetic dataset for bearing fault diagnosis. We also provide our code to facilitate further research.

## II. BACKGROUND

### A. Use of Synthetic Data for Fault Diagnosis

In the field of fault diagnosis, synthetic faults have been used as a reasonable substitute when data on real faults is not available [3]. For example, synthetic faults can be induced from healthy samples using an analytical model [2]. Faults can also be simulated by highly accurate physical simulation such as finite-element method (FEM) models [6]. In most works, the learning of the models is purely based on the synthetic faults, while the rare fault data collected from the real applications is preserved as a test set to evaluate the model performance [4], [5]. Some recent research studies, such as [14] make the assumption that they have access to a very small set of labeled real faults. A large number of synthetic faults is then generated to mimic the real faults by adversarial networks [15] or interpolating between samples [16]. The corresponding data-driven models can be trained based on these imitated faults. This approach is different from previous works which learn solely from the synthetic data, because the model now also have an implicit or explicit access to information of the small real set of labeled faults. However, in these works, the unavoidable domain shift between the synthetic and real faults is still overlooked. In addition, the requirement of access to the labeled real faults imposes a real limitation on the generalization ability. If the target operating condition is different from the observed one, the learned method may fail because of the existence of a further gap between operating conditions. To the best of our knowledge, these domain gaps between synthetic and real data in the field of fault diagnosis have been long overlooked. Very recently, Liu *et al.* [17] used both synthetic data and real data to learn a model for remaining useful life prediction in the related field of prognostics. Our work contributes to this by considering a mix of both the healthy and synthetic data which will result in a more realistic source domain.

### B. Domain Adaptation

Unsupervised DA [18] is a powerful tool to alleviate the domain shift between synthetic and real data. DA approaches often consider the case where there is a labeled source domain (synthetic in our case) and an unlabeled target (real) domain. The methods improve the performance on the target by making use of both labeled source data and additional unlabeled target data. DA has been widely studied in fields such as computer vision [19] and natural language

processing [20]. The alleviation of the distribution difference between the source and the target is often the motivation for applying DA approaches. By aligning the distributions, models can effectively benefit from both the source and target data. Following this motivation, a series of approaches have explored different ways of alignments. Discrepancies such as maximum mean discrepancy (MMD) were used [10] as a guide to align the distributions. Distributions can also be aligned in normalization layers [21]. Domain adversarial neural networks (DANNs) [7] use a domain discriminator to adversarially align the features from the source and target domains. The method usually has a classifier branch and a discriminator branch. Given a feature extractor $f$ parameterized by $\theta_f$, a classifier $g$ parameterized by $\theta_g$, and a discriminator $h$ parameterized by $\theta_h$, the DANN method is essentially solving the following equations:

$$\mathcal{L}(\theta_f, \theta_g, \theta_h) = \mathcal{L}_{\text{clf}}(g(f(x))) - \lambda_d \mathcal{L}_d(h(f(x))) \quad (1)$$

$$\hat{\theta}_f, \hat{\theta}_g = \arg\min_{\theta_f, \theta_g} \mathcal{L}(\theta_f, \theta_g, \hat{\theta}_h) \quad (2)$$

$$\hat{\theta}_h = \arg\max_{\theta_h} \mathcal{L}(\hat{\theta}_f, \hat{\theta}_g, \theta_h) \quad (3)$$

where $\mathcal{L}_{\text{clf}}$ is the cross-entropy loss function for the main classifier. $\mathcal{L}_d$ is the cross-entropy loss for the domain classification. The classifier branch is trained to minimize the classification loss $\mathcal{L}_{\text{clf}}$ on source data. The discriminator is trained to generate unbiased features for both the source and target data. One the one hand, the optimization is updating the discriminator's weight by minimizing discriminator loss $\mathcal{L}_d$ to make it distinguish the source and target features as well as possible. On the other hand, the equation is forcing the feature extractor $f$ to generate unbiased features such that the discriminator loss is large. This minimax game effectively aligns the distributions between the source and the target. CDAN [22] takes this concept one step further and conditions the adaptation models on discriminative information conveyed in the classifier predictions.

In recent years, input-space adaptation and self-training frameworks have also attracted many research interests. Aligning source and target data in the input space can be beneficial via Fourier transforms [23]. In self-training frameworks, model performance can be gradually improved by iteratively training the network using target pseudo-labels as ground truth. Following this motivation, PyCDA [24] improves adaptation performance by training with pseudo-labels from different scales. Zhang *et al.* [25] proposes to refine pseudo-labels by making use of prototypes. Wang *et al.* [26] find that explicitly learning the relationship between the main task and self-supervised auxiliary task can help improve DA performance. Mei *et al.* [27] find it beneficial for DA to integrate adversarial alignment with self-training.

### C. DA for Fault Diagnosis

In recent years, DA methods have raised a strong interest in the fault diagnosis community [28]–[32]. Most existing works focus on the adaptation between operating conditions and have found classic DA methods beneficial [33]. Adversarial alignment [34] is widely used by existing works [35], [36] and
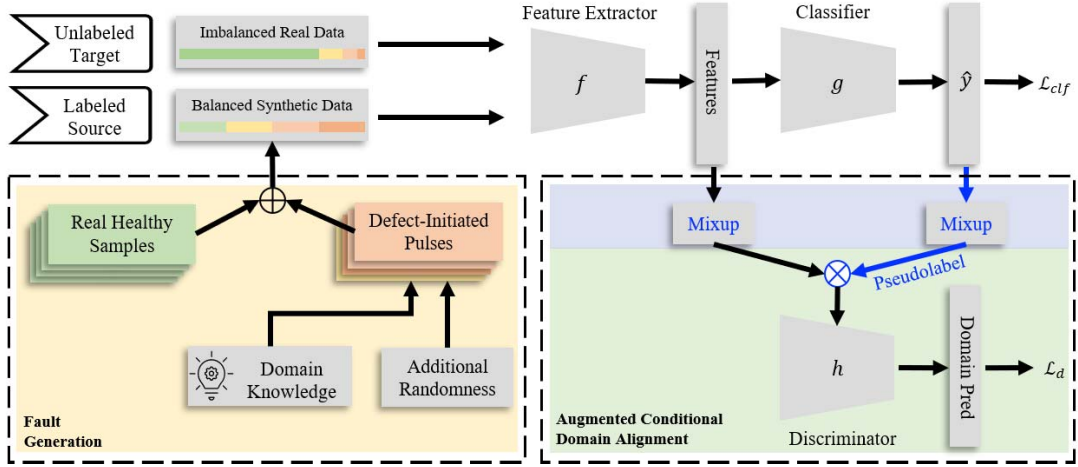
Fig. 2. Overview of the proposed method. (Left) The fault generation method we used in stage one. (Right) The proposed augmented conditional domain alignment method. As shown in the green block, both the feature and class information are fused by a multi-linear map before feeding into the discriminator to provide information on the class-conditioned distributions. In the standard adversarial approach DANN, the discriminator does not have access to class information $\hat{y}$. Our mixup-based distribution augmentation is shown in the purple block to provide additional distribution support for rare faults. The difference between our proposed augmented conditional domain alignment and DANN is shown in blue.

improved by conditional discriminators [37], [38]. Generative adversarial networks (GANs) are also explored to generate faults [39] to alleviate domain gap. Liu *et al.* [40] make use of optimal transport theory to align the distributions. Chen *et al.* [41] propose to also take contribution of individual data samples into account, in addition to the global data distribution. We refer the readers to [42] for an overview of existing methodologies and [43] for implementation collections. More recent works focus on making the DA setup more realistic. For example, Zhu *et al.* [44] focused on learning from multiple source domains and Zhang *et al.* [45] considered the more general case of universal DA [46].

While many works make the assumption that both domains have a balanced number of samples for each class, Wang *et al.* [47] explore the scenario where the target data has missing fault classes. Another related line of work is to adapt models between different units of a fleet, particularly focusing on the complementary operating conditions [48]. Synthetic-to-real adaptation is relatively new and an unexplored research direction in the field of fault diagnostics. To the best of our knowledge, there has been no prior work which explores it for fault diagnosis where the imbalanced faults impose a great challenge to the DA task.

## III. PROPOSED METHODOLOGY

To exploit the possibility of effectively learning and adapting from synthetic data in real industrial scenarios, we consider a challenging and realistic setup where we only have access to segments of condition monitoring vibration recordings. Given unlabeled training data from the target real domain

$$\mathcal{D}_t = \left\{ x_t^1, \ldots, x_t^m \right\} \tag{4}$$

where $x_t^i$ is a vector of a real vibration recording. The target in this unsupervised fault diagnosis task is to train a model which performs well on this target domain.

### A. Overview of the Proposed Framework

The proposed framework integrates expert knowledge with imbalance-robust DA for unsupervised fault diagnosis tasks.

In the first part, the fault generation module makes use of expert knowledge on the fault types and generates synthetic faults based on the unlabeled real recordings of the bearings. Unlike previous works which utilize highly accurate physical models to synthesize the faults, our generation stage is deliberately designed to be simple such that the overall framework does not rely on highly accurate physical simulators.

In the second part, a novel imbalance-robust domain adaption approach is proposed to alleviate the distribution gap between the synthetic and real features, specifically for fault diagnosis. Unlike classic DA methods which typically deal with balanced data, the proposed method is able to align the features when the class distribution is different between the synthetic source and real target. This imbalance-robust framework is proposed for realistic scenarios as real datasets are often very imbalanced because of the rare faults. We would like to highlight that the proposed DA approach does not rely on specific generation process in the previous step. Thus, it can be easily generalized to other fault generation process or even to other classification problems that suffer from a class-imbalance between the source and target datasets.

Integrating both parts together leads to our proposed framework which makes effective use of both expert knowledge and unlabeled target data. The overview of our proposed framework is shown in Fig. 2.

### B. Bearing Fault Generation

The challenge of fault generation in our proposed framework is how to make the synthetic faults adaptable to the given

target (real) domain without using highly accurate physical models. It is important to take the limited information provided in the unlabeled target samples into account. As a prerequisite, we assume that we have access to some real healthy samples. This is usually achievable by using the early recordings of each asset where we can safely assume that the faults have not yet emerged. Thus, this is a realistic assumption also in real application conditions. The healthy real samples are then used as the base signal for the synthetic process and inject the fault patterns using expert knowledge. Since the base signal encodes information about the operating and environmental conditions, the generated signals can be more adaptable to the target domain.

To this end, a general procedure is followed to generate synthetic bearing defect signals as described in [49]. Let $s$ denote the oscillating waveform excited by a single impact due to over-rolling a surface defect with period $T$. The amplitude of the $i$th impact is denoted by $A_i$ with period $Q$, which mimics a modulation term due to a rotating inner ring or a rolling element defect. The terms $T$ and $Q$ can be directly calculated based on the kinematics of the bearing and a speed recording by a domain expert (see (1)–(4) in [8]). We can now describe our modeled vibration signal at time $t$ by

$$\epsilon(t) = \sum_{i=-\infty}^{\infty} A_i s(t - iT) + \beta n(t) \quad (5)$$

where the additive background noise term $n$ is taken equal to a vibration recording of a healthy bearing. To make our method robust for various signal-to-noise ratios (SNRs), $\beta$ is uniformly distributed between 0.25 and 2. A single impact $s$ is modeled by means of a Hann window with a duty period of 5% with respect to $T$. Since we do not know the transfer function, a wide-band bandpass filter is applied to $s$ instead. This filter mainly attenuates frequencies close to 0 Hz and the Nyquist frequency. As a result, our fault frequencies are visible in any frequency band, which facilitates the DA to bridge the gap with the real sensor data. The periodic amplitude modulation term $A_i$ is expressed as a sum of cosines

$$A_i = \gamma \sum_{k=0}^{K} \alpha_k \cos(iTk2\pi/Q) \quad (6)$$

such that we can control the number of side bands $k$ with corresponding amplitudes $\alpha$ in our envelope analysis directly. To introduce some natural randomness of the impacts, $\gamma$ is normally distributed with a unit mean and a standard deviation of 0.1. In our work, we use $\alpha = [1, 0.76, 0.38, 0.11, 0.05]$. The chosen $\alpha$ should describe a typical decaying sideband pattern for a radially loaded bearing. We would like to emphasize that $\alpha$ was not specifically optimized to improve the diagnosis performance on any of our datasets as we do not assume access to fault labels in the real data for training. Hence, a domain expert does not need exact knowledge of $\alpha$. However, in the case of an axially loaded bearing, one could consider reducing the number of side bands (see [50] for more details). The proposed generation method follows similar procedures as introduced in [49] and [51].
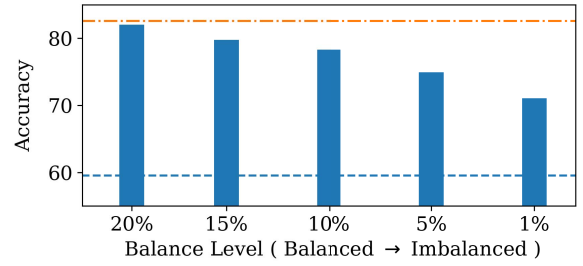


Fig. 3. Standard adversarial DA performance decreases more rapidly when the target domain becomes more imbalanced. A synthetic-to-real CWRU experiment with different levels of imbalance on the rolling element fault class. 1% on the $x$-axis means that only 1% samples are available for rare fault class, compared to the balanced case. The orange dashed-dotted line indicates the performance when the dataset is fully balanced (100%). Blue dashed line indicates the source-only baseline.

## C. Augmented Conditional Adversarial Alignment

The proposed generation approach provide us with the source dataset $\mathcal{D}_s$ which contains the synthetic faults. To make efficient use of the unlabeled data from the real target domain, we propose to re-formulate the unsupervised fault diagnosis task as a unsupervised DA problem.

1) Synthetic source domain data with balanced samples across all classes

$$\mathcal{D}_s = \left\{ (x_s^1, y_s^1), \ldots, (x_s^n, y_s^n) \right\}, \quad y_s^i \in Y. \quad (7)$$

2) Unlabeled real data from the target domain with an imbalance across classes

$$\mathcal{D}_t = \left\{ x_t^1, \ldots, x_t^m \right\} \quad (8)$$

where $Y$ is the set of discrete health states. Our aim is to improve the model performance in the target domain $\mathcal{D}_t$. The setup is now similar to transductive DA problems [52], [53].

*1) Imbalance Issue for Direct Alignment:* There is one essential difference between the standard DA commonly seen in image classification and our setup which could potentially harm the alignment quality. In most DA setups, the source and target domains are assumed to have the same class distributions, meaning that either both the source and target are class balanced, or both follow a similar class distribution. However, for our synthetic-to-real setup, the scenario is quite different. On the synthetic data side, since faults are all generated based on the healthy samples, each health condition has the exact same number of samples, leading to a balanced source dataset. On the target real data side, the class distribution contains a majority of healthy states and few faults, leading to an imbalanced dataset.

This mismatch between the source and target domain can in fact lead to a significant decline of the adaptation performance. We show empirically in Fig. 3 the performance of a naive synthetic-to-real adversarial adaptation based on the CWRU dataset using DANN [7]. In this setup, the source synthetic domain is balanced, while for the target domain, one of the faults (rolling element fault) has a smaller number of samples. The experiment shows that, compared to a fully balanced case, the stronger class imbalance in the target leads to a steeper performance decline.

To answer why such a performance decline occurs, we need to look back into the assumptions that the classical DA methods make. Classical DANN methods [7], [33], [54] are based on the following DA theory: the error function of the domain discriminator corresponds well to the discrepancy between unconditional feature distributions $P$ and $Q$. Thus, minimizing the error rate of the discriminator can lead to aligned feature distributions. However, when the joint distributions of feature and class, that is, $P(x_s, y_s)$ and $Q(x_t, y_t)$, are non-identical across domains, adapting only the feature representation may be insufficient [22]. This can be an especially large problem when the target class distributions in fault diagnosis are heavily skewed toward the healthy class. In this case, simply aligning the features, in fact, gives no theoretical guarantee that two different distributions are identical even if the discriminator is fully confused [55].

*2) Class Information Taken Into Account:* The performance degradation stems from the fact that the class distributions are mismatched and the discriminator fails to take any class information into account. We, thus, propose to use the class information of the representations as additional input to the discriminator in order to counteract the class mismatch from rare faults. By providing the generator hints on the class information, the discriminator can better align the distributions.

We use the multi-linear map to fuse the class information with features to improve the DA performance under severe class imbalance. As shown in Fig. 2, given a feature extractor $f$ parameterized by $\theta_f$, a classifier $g$ parameterized by $\theta_g$, and a discriminator $h$ parameterized by $\theta_h$, the loss function then becomes

$$\mathcal{L}(\theta_f, \theta_g, \theta_h) = \mathcal{L}_{clf}(g(f(x))) - \lambda_d \mathcal{L}_d(h(f(x) \otimes g(f(x)))) \tag{9}$$

where $\otimes$ is a multi-linear map, and $\hat{y} = g(f(x))$ represents the predicted pseudo-labels. Comparing to (1) of the standard adversarial training, the main difference is that instead of using features $f(x)$ alone as input to the discriminator, the information of features and classes (provided by pseudo-labels) are combined together via a multi-linear map. By applying the map, the discriminator can gain the information from both the feature distribution and class distribution and better align the class conditioned distributions.

The proposed method is inspired by conditional adversarial methods in other application fields. For example, for image generation tasks, conditional GANs [56] concatenate the class vector with the feature vector to generate images conditioned on the single classes. In DA for image classification, CDAN [22] uses multi-linear conditioning to align multi-modal distributions. These methods often deal with balanced datasets and do not focus on severe class-imbalanced scenarios, but the idea of aligning the conditional distribution is especially suitable for fault diagnosis.

*3) Augmented Distributions:* One potential issue with the above solution is that even though the important class information is provided for the discriminator, the fact that the rare fault classes have so few samples can potentially make the optimization of the discriminator unstable. This can also

lead to a decline of the alignment performance. To reinforce the alignment, it is essential to provide better distribution support for the conditional distributions on the rare classes. We propose to augment the features and pseudo-labels used as input to the discriminator. In particular, for a batch of $x$ from the target real dataset, we can get its corresponding feature embedding $e = f(x)$, and pseudo-label $\hat{y} = g(f(x))$ vectors. Whereby, $\hat{y}$ is generated by using the prediction from the model trained in the previous iteration. Then, the augmented interpolated sample can be represented as

$$\tilde{e} = \lambda e + (1 - \lambda)e[\text{idx}] \tag{10}$$
$$\tilde{y} = \lambda \hat{y} + (1 - \lambda)\hat{y}[\text{idx}] \tag{11}$$
$$\tilde{z} = \tilde{e} \otimes \tilde{y} \tag{12}$$

where idx is the shuffled index of a batch, $\lambda$ is generated from a prior $\beta$ distribution, that is, $\lambda \sim \beta(\alpha, \alpha)$ with $\alpha$ controls the shape of the $\beta$ distribution, $\tilde{e}$ is the interpolated features, $\tilde{y}$ is its class label, and $\tilde{z}$ is the multi-linear input to the discriminator. We use $\alpha = 1$ for all our experiments. This idea is inspired by MixUp [13] in image classification, where the authors mix input images and labels to provide more training samples in a fully supervised training setup. Instead of directly mixing the input images in MixUp, our proposed distribution augmentation is conducted on the feature space and used for our unsupervised DA setup.

The motivation behind this is to augment the conditional feature distributions of rare classes for fault diagnosis. By mixing up the features within a batch, the fault information becomes present in more samples within a batch. The interpolated samples enlarge the target training dataset for the rare fault classes, making the learning process for the discriminator more stable. We highlight the difference between DANN and our alignment method in blue on the right side of Fig. 2.

## IV. DATASETS

To facilitate our adaptation experiments, as described in Section III-B, we generate the synthetic faults for two datasets.

### A. CWRU Bearing Data

The Case Western Reserve University (CWRU) [57] bearing dataset is used. It is a benchmark dataset for DA in bearing fault diagnosis [32]–[34]. Drive-end accelerometer data of 12-kHz sampling rate is used. We consider four health states (classes) in this article: healthy, inner race fault (IF), rolling element fault (REF), and outer race fault (OF). We group the sub-fault-types of different spall sizes together. For each health state, we sample 1200 segments. Each signal segment contains 4096 points. This results in 4800 samples for the real dataset. Subsequently, our generation method is applied to the real healthy signals to synthesize the real defects. To avoid data leakage during evaluation, we split half of the real healthy samples to the synthetic source data and up-sample the number of samples in healthy class in both domains back to 1200.

Fig. 4 shows the full-wave rectified envelope spectrum of one synthetic and one real example of the IF. Defect
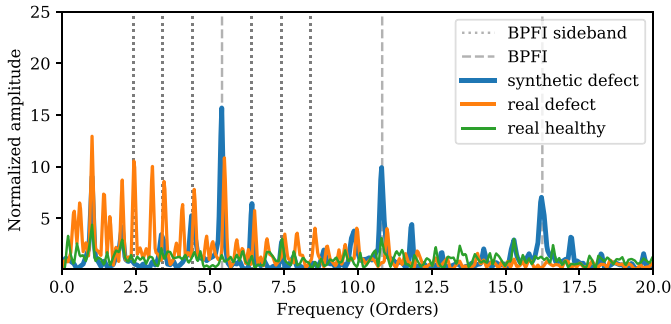
Fig. 4. CWRU normalized full-wave rectified envelope spectrum of an inner ring defect example with corresponding defect ball pass frequency (BPFI) denoted by the dashed vertical lines. The difference between the real and synthetic fault in BPFI frequencies and sidebands (showed by vertical dotted lines) motivates the use of DA.
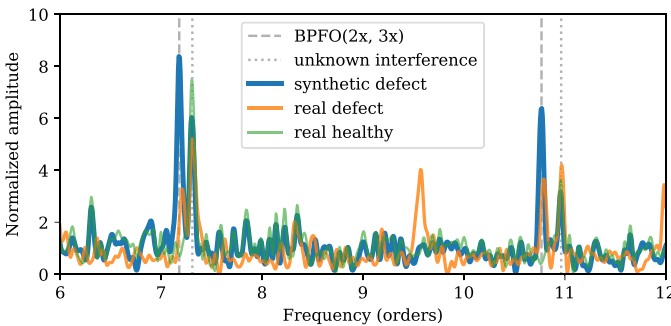


Fig. 5. Wind turbine generator dataset: normalized full-wave rectified envelope spectrum of an outer ring defect example. In this example, the unknown interference is realistically transferred from healthy to synthetic. However, the synthetic fault does not match perfectly with the real fault, and the difference motivates the use of DA. Notice the frequency range between 6 and 12 orders such that only two and three times the ball-pass frequency of the outer ring (BPFO) is visible.

frequencies are indicated by the dashed lines, where a dotted line shows the first three side bands of an inner ring defect around its ball-pass frequency (BPFI). The BPFI and its first harmonic in the synthetic and real examples are roughly aligned. However, differences can still be observed: 1) the real defect has more side bands around the BPFI; 2) the second harmonic is not present in the real defect; and 3) the defect frequencies have a slightly higher frequency compared to the analytic defect frequencies. These differences motivate the use of DA to bridge the gap between the real and synthetic signals.

### B. Wind Turbine Bearing Data

A real-world dataset from generator bearings of multiple wind turbines is collected and used to evaluate the method. This data origins from a condition monitoring service, where bearings are monitored by human experts aided by analysis software. For each turbine, historic data is available varying between three and five years, where on a daily basis two recordings are made. Each vibration recording has a length of 1.28 s and a sample rate of 12.8 kHz. Rotational speeds vary between 900 and 1700 rpm. Speed recordings are available via a tachometer. There are three health states (classes): healthy, outer-ring defects, and inner-ring defects. Example signals from the wind turbine bearing dataset are presented in Fig. 5.

TABLE I
NETWORK COMPONENTS USED IN ALL OUR EXPERIMENTS

| Input $x$ | | |
|---|---|---|
| **Feature Extractor $f$** | | |
| Conv0 | filter size 3; channel num 10 relu; dropout | |
| Conv1 | filter size 3; channel num 10 relu; dropout | |
| Conv2 | filter size 3; channel num 10 relu; dropout; flatten | |
| Dense | channel num 256 relu; dropout | |
| **256-D Features $f(x)$** | | |
| **Classifier $g$** | | **Discriminator $h$** |
| Dense | channel num 256 relu | Dense | channel num 512 relu |
| Dense | num of classes softmax | Dense | channel num 2 softmax |
| **Class Prediction** | | **Domain Prediction** |

Based on a combination of manual analysis of vibration and process data, knowledge of bearing replacements and anonymized customer feedback on bearing defects, the data was carefully labeled. In total, we collected 1643 samples for the healthy state, 2990 samples for the outer-ring defects, and 192 samples for the inner-ring defects from the generator drive-end bearings each on a different wind turbine. Note that during the labeling of the data, we had access to additional process data such as electrical power and wind-speed but were also able to inspect trends and high SNR averaged spectra of many vibration recordings. As a consequence, a single vibration recording, as used by our proposed DA method, is only a fraction of the information used in the labeling procedure. In many situations, all these additional information sources are unavailable, which motivates the use of single vibration recording predictions as in this research.

## V. EVALUATION

### A. Implementation

*1) Model Architecture:* The network consists of a feature extractor, classifier, and a discriminator. The architecture of the feature extractor and classifier are taken from [39]. The details are summarized in Table I.

*2) Data Preprocessing:* A preprocessing step is applied to all data. First, the time-domain waveforms are normalized to unit standard deviation and subsequently converted to a full-wave rectified envelope spectrum. It is obtained by taking the Fourier transform magnitudes of the full-wave rectified and bandpass filtered signals. A bandpass filter with a passband between 500 and 4000 Hz is used following a standard frequency band for bearing diagnostics as described in [58]. We filter at 4000 Hz to attenuate the signal below the Nyquist frequency. Although more optimal bearing diagnostic signal representations are available, for example, as used in [59], this simple representation is sufficient to make a valid comparison between the different DA methods. As no labeled real-world faulty signal is available, Kurtogram-based methods [60], [61] cannot be used to optimize the pass bands. Subsequently, the full-wave rectified envelope spectrum is then interpolated

to a speed-normalized axis of 1000 values between 0 and 30 repetition orders. The same data preprocessing is applied on all experiments across all methods.

*3) Evaluation Metric:* Since the classes are imbalanced, we report a balanced version of the accuracy. This metric was used by imbalanced classification tasks such as [62]. For a dataset with $K$ classes, the reported accuracy is defined as

$$\frac{1}{K} \sum_{k=1}^{K} \frac{P_k}{M_k}$$

where $M_k$ is the number of test samples with class $k$ as ground truth, and $P_k$ is the number of correct predictions for the given $M_k$ samples in class $k$. This metric provides a fair comparison also in highly imbalanced datasets. All reported results are based on the average of ten runs using this metric. We additionally report the F1 score and Cohen's kappa [63]. We use the implementation from Scikit-learn [64] for all evaluation metrics.

*4) Training and Test Details:* We train our neural architecture specified in Table I end-to-end for 100 epochs with a batch size of 128 and a learning rate of 0.001. All models are trained using the Adam optimizer. The same architecture and training parameters are used for all methods in our experiments to ensure a fair comparison. During training, only the source labeled data $\mathcal{D}_s$ is used to update the classifier $g$ via $\mathcal{L}_{\text{clf}}$ because we do not have labels for the target real data $\mathcal{D}_t$. Both the synthetic and real data contribute to the learning of the feature extractor $f$ and discriminator $h$ via the combined loss defined in (9). For inference, we feed the target data to the feature extractor $f$ and classifier $g$ to make the prediction. The discriminator is not needed in test time. All models are trained and tested under different random seeds for ten times and we report the average performance. A flowchart of the proposed framework is provided in Fig. 6.

### B. CWRU Synthetic-to-Real Experiments With Different Levels of Imbalance on One Fault

We evaluate the proposed method against a source-only baseline and the standard DANN alignment. In this ablation study, we reduce the complexity by assuming that all classes have the same number of samples except the rolling element fault. More realistic setups are evaluated in later sections. To show the effectiveness of the proposed methodology against different levels of class imbalance, we change the number of samples with the rolling element fault. For example, when the balance level is 1%, the unlabeled target training data contains $1\% * 1200 = 12$ samples for the rolling element fault class, while 1200 samples are available for every other class. In the 10% setup, the target set contains $10\% * 1200 = 120$ samples for the rolling element fault class. A smaller value of level of balance (1%) indicates a larger level of class imbalance as the number of faults is considerably smaller compared to the number of healthy class samples.

As shown in Table II, without any alignment, the source-only accuracy is 59.55% in this four-class-classification problem. This is already higher than the random guess performance of 25% and proves that our synthetic faults encode the fault
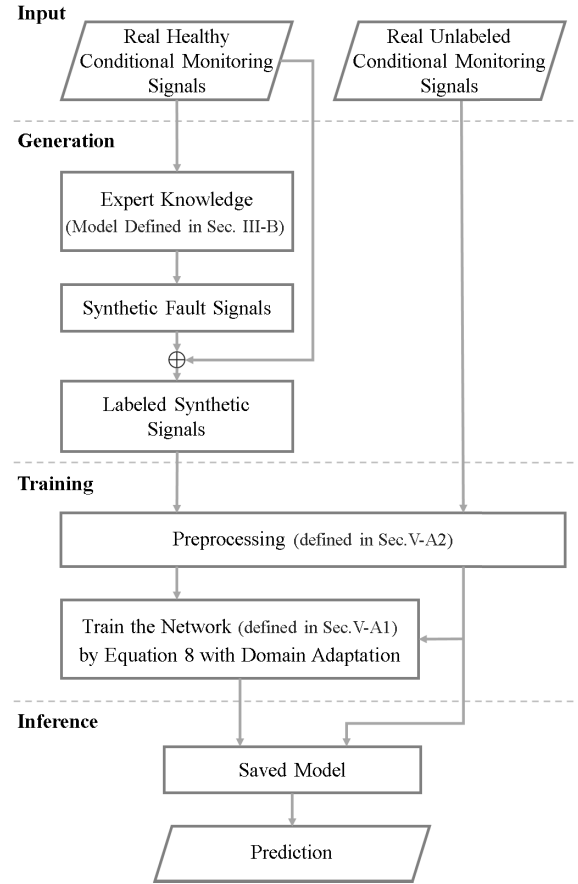


Fig. 6. Flowchart of the proposed framework.

TABLE II

SYNTHETIC-TO-REAL CWRU ADAPTATION WITH DIFFERENT BALANCE LEVELS ON ROLLING ELEMENT FAULTS

| Balance Level* | DANN [54] | Conditional(ours) | Proposed(ours) | Δ |
|---|---|---|---|---|
| 20% | 82.04% | 82.11% | 83.49% | +1.44% |
| 15% | 79.77% | 81.90% | 83.04% | +3.27% |
| 10% | 78.30% | 82.88% | 83.32% | +5.02% |
| 5% | 74.96% | 76.90% | 83.30% | +8.34% |
| 1% | 71.05% | 73.93% | 82.34% | +11.29% |
| Source-only | 59.55% | Balanced DANN | | 82.59% |
| Balanced CDAN | 80.32% | Balanced Ours | | 84.18% |

* Smaller balance level indicates fewer samples and larger imbalance.

information in a meaningful way. On the other hand, DANN can perform quite well when all the classes are balanced, but degrades dramatically when the level of balance decreases. By taking the class information into account, conditional alignment alone shows resistance to the degradation from the class imbalance. However, the performance gap still exists, and the performance becomes less stable when the balance level decreases. By further augmenting the conditional distributions, the proposed method can significantly strengthen the quality of the alignment and lead to a performance that is at a similar level as for the balanced dataset, even when there are only 1% rolling element fault samples in the target dataset. This demonstrates the effectiveness of our proposed method.

TABLE III

EXPERIMENT RESULTS ON THE SYNTHETIC-TO-REAL TASK FOR CWRU DATASET WHEN LEVEL OF BALANCE IS DIFFERENT AMONG FAULTS. THE TARGET DATA CONTAINS 1200 (100%) HEALTHY SAMPLES, 120 (10%) OF SAMPLES, 60 (5%) IF SAMPLES, AND 12 (1%) REF SAMPLES

| Metric | Source | DANN [54] | IAST [27] | Condi. | Proposed |
|---|---|---|---|---|---|
| Balanced acc. | 0.596 | 0.726 | 0.760 | 0.754 | **0.823** |
| F1 (Macro) | 0.402 | 0.599 | 0.681 | 0.678 | **0.814** |
| F1 (Micro) | 0.488 | 0.630 | 0.697 | 0.688 | **0.787** |
| Cohen's kappa | 0.352 | 0.476 | 0.574 | 0.560 | **0.696** |

TABLE IV

EXPERIMENTAL RESULTS ON THE WIND TURBINE DATASET

| Metric | Source | DANN [54] | IAST [27] | Condi. | Proposed |
|---|---|---|---|---|---|
| Balanced acc. | 0.609 | 0.645 | 0.679 | 0.683 | **0.708** |
| F1 (Macro) | 0.553 | 0.570 | 0.666 | 0.671 | **0.683** |
| F1 (Micro) | 0.681 | 0.768 | 0.842 | 0.840 | **0.845** |
| Cohen's kappa | 0.411 | 0.524 | 0.643 | 0.641 | **0.650** |

We also report the improvement of the proposed method over the standard DANN. An interesting observation is that with an increasing level of imbalance, our method provides an increasing amount of improvement over DANN. This suggests that the method can be especially beneficial when imbalance in the target dataset is severe. Furthermore, the proposed method appears to be robust to the degree of imbalance. The results for all the degrees of imbalance demonstrate a similar accuracy. Even with a relatively strong imbalance (1% balance level), the performance is similar to the fully balanced setup of the standard DANN. Therefore, the methodology is suitable for cases where the degree of class imbalance in the target dataset is not known which is a typical setup in real applications. Thus, the method is quite suitable for real fault diagnosis tasks.

*1) Performance Under a Balanced Setting:* To further analyze the effectiveness of the proposed method, we report the performance under the balanced CWRU setup in Table II. Under this setup, all classes in the target domain have equal number of samples. We observe that all alignment methods work well comparing to the source-only baseline. The DANN methods achieves 82.59%, while our method provides further improvement and achieves 84.18%. This shows that the proposed augmented conditional alignment method is complementary to the adversarial alignment, even when all classes are balanced. This is most likely due to our feature distribution augmentation strategy which better facilitates the adversarial alignment.

### C. CWRU Synthetic-to-Real Experiment With a More Realistic Class Imbalance Among Faults

In reality, the imbalance does not only exist for one single fault, but also exists between different faults. To mimic also the imbalance between several faults in the target domain, we consider the following setup. The unlabeled target dataset consists of $N = 1200$ healthy samples, $10\% \times N$ samples for the OF, $5\% \times N$ samples for the IF, and $1\% \times N$ sample for the REF. This constructs a highly imbalanced target dataset. All labels are again removed during the training. This setup makes it a more challenging task as the imbalance levels of the three faults are different.

Results in Table III show that when class imbalance is severe among different faults, the proposed method can still provide an improved alignment for the different classes. Compared to the simple DANN [54] alignment method which

does not consider the imbalance between different classes, both the conditional alignment method CDAN [22] and our proposed mixup augmentation provide a significant improvement. Conditional alignment alone provides a 2.8% absolute improvement in balanced accuracy over DANN by simply providing additional class information. This improvement is also validated by the F1-score and Cohen's kappa. By combining the proposed augmentation with the class conditional alignment, we achieve a much stronger balanced accuracy of 82.3% on the target data. This is a 9.7% absolute improvement over the DANN baseline, and the improvement is validated also by the other metrics. The overall performance of the proposed methodology is similar to that of DANN in the ideal fully balanced case. This makes the proposed framework applicable for unknown imbalance levels achieving same level of performance as in a balanced dataset. This finding is particularly encouraging for real applications where the imbalance level is unknown *a priori*.

We additionally compare our method with the state-of-the-art DA method IAST [27]. IAST constitutes a stronger competing method as it combines the advantages of adversarial learning and self-training. Under this setup with severe imbalance, our proposed method outperforms IAST by a large margin in terms of all metrics including balanced accuracy, F1-score, and Cohen's kappa.

### D. Wind Turbine Generator Synthetic-to-Real Experiment

To evaluate the effectiveness of the proposed methodology in a more realistic scenario, we additionally conduct an experiment on the real-world wind turbine bearings. As described in the dataset section, the OF has a significantly larger number of samples than the IF. Results is listed in Table IV. In this setup, the source-only baseline achieves an accuracy of 60.85%, showing that the synthetic data we generated for the wind turbine bearing is meaningful. The DANN method can improve the performance and achieve 64.5%. By providing class information to the discriminator, we can achieve 68.3%. If we additionally use our augmentation to enhance the distribution support, the final proposed method can achieve an accuracy of 70.8%, yielding an almost 10% absolute improvement on this real dataset, compared to the source-only baseline. In terms of F1-score and Cohen-s Kappa, the proposed method also shows improvement over the source only baseline and DANN approach. The performance improvement is large over the baselines, but less significant than the setup in Section V-C. This is mostly likely because the class imbalance is less severe in the wind turbine dataset than in the setup we used in Section V-C. We additionally compare our method with the

state-of-the-art DA method IAST [27]. The proposed method is able to outperform IAST in terms of all four metrics. This validates the effectiveness and competitiveness of our proposed augmented conditional distribution alignment.

We would like to emphasize that in all experiments, we only use knowledge of the healthy labels from the target domain. All class information on the faults is learned from the synthetic data where the expert knowledge is encoded. This result shows that given an unlabeled target bearing dataset, it is possible to make use of expert knowledge and train a data-driven fault diagnosis model for it. This is achieved by our proposed framework that combines a fault generation process and our proposed synthetic-to-real adaptation approach which is specifically designed for the imbalanced target data.

### E. Experiment Summary

One of the main findings in the experiments is that, even though the popular DA methods such as DANN and CDAN work well in the fully balanced setup, their performance can suffer significantly when class imbalance becomes larger. By providing class information to the discriminator and augment the features, we are able to alleviate this negative effect. Interestingly, we notice that the feature distribution augmentation is an important design to make the conditional alignment work well. This is likely because without the distribution augmentation, the training of the discriminator is also severely biased because of the imbalance between the classes. Another noteworthy finding is that even though a relatively simple model for generating the synthetic faults was chosen, it still successfully encodes the key fault information, and the proposed DA method was able to transfer the expert knowledge from the synthetic faults to a diagnosis model.

## VI. Conclusion

We proposed a novel bearing fault diagnosis framework which can learn effective models from unlabeled real bearing data. In particular, we showed that by generating synthetic faults using expert knowledge and conducting imbalance-robust DA, a fault diagnosis model can be learned without any supervision from the real faults. We showed that a good approach to compensate the imbalance from rare target faults is key to the synthetic-to-real adaptation performance. A class conditioned adversarial adaptation method is thus proposed to address this issue. An additional augmentation based on the mixup approach was further proposed to deal with the limited number of fault samples and bridge the class distribution gap. The two components for DA can be easily applied in combination with other fault generation frameworks. The proposed methodology does not require any assumption on the degree of the underlying class imbalance and achieves a similar performance in the imbalanced setup as the standard DANN on the fully balanced setup, demonstrating its robustness to different imbalance levels. Experiments on the benchmark CWRU bearing dataset and a wind turbine generator bearing dataset have validated the effectiveness of

our approach also in real applications and under realistic assumptions. The framework can be implemented easily and have the potential to be applied on other industry assets.

In this work, we mainly considered an unsupervised setup, where we do not have any access to the real fault labels/supervision. This is a relatively challenging task and could be unnecessarily difficult for some practical use cases. One possible direction for further improvement is to consider relaxed setups such as semi-supervised DA [65]. In these setups, a few real samples for each fault type are accessible to provide additional supervision. This is often a realistic relaxation when the assets in consideration have been running for a while. We leave this for future research.

## References

[1] S. Singh, C. Q. Howard, and C. H. Hansen, "An extensive review of vibration modelling of rolling element bearings with localised and extended defects," *J. Sound Vib.*, vol. 357, pp. 300–330, Nov. 2015.

[2] J. Q. Zhang and Y. Yan, "A wavelet-based approach to abrupt fault detection and diagnosis of sensors," *IEEE Trans. Instrum. Meas.*, vol. 50, no. 5, pp. 1389–1396, Oct. 2001.

[3] N. H. W. Eklund, "Using synthetic data to train an accurate real-world fault detection system," in *Proc. Multiconf. Comput. Eng. Syst. Appl.*, Oct. 2006, pp. 483–488.

[4] K. C. Gryllias and I. A. Antoniadis, "A support vector machine approach based on physical model training for rolling element bearing fault detection in industrial environments," *Eng. Appl. Artif. Intell.*, vol. 25, no. 2, pp. 326–344, Mar. 2012.

[5] C. Sobie, C. Freitas, and M. Nicolai, "Simulation-driven machine learning: Bearing fault classification," *Mech. Syst. Signal Process.*, vol. 99, pp. 403–419, Jan. 2018.

[6] Y. Gao, X. Liu, and J. Xiang, "FEM simulation-based generative adversarial networks to detect bearing faults," *IEEE Trans. Ind. Informat.*, vol. 16, no. 7, pp. 4961–4971, Jul. 2020.

[7] Y. Ganin and V. Lempitsky, "Unsupervised domain adaptation by backpropagation," in *Proc. Int. Conf. Mach. Learn.*, 2015, pp. 1180–1189.

[8] R. B. Randall and J. Antoni, "Rolling element bearing diagnostics—A tutorial," *Mech. Syst. Signal Process.*, vol. 25, no. 2, pp. 485–520, 2011.

[9] B. Fernando, A. Habrard, M. Sebban, and T. Tuytelaars, "Unsupervised visual domain adaptation using subspace alignment," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2013, pp. 2960–2967.

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

[11] B. Sun and K. Saenko, "Subspace distribution alignment for unsupervised domain adaptation," in *Proc. Brit. Mach. Vis. Conf.*, 2015, pp. 1–24.

[12] Y.-H. Tsai, W.-C. Hung, S. Schulter, K. Sohn, M.-H. Yang, and M. Chandraker, "Learning to adapt structured output space for semantic segmentation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 7472–7481.

[13] H. Zhang, M. Cisse, Y. N. Dauphin, and D. Lopez-Paz, "Mixup: Beyond empirical risk minimization," 2017, *arXiv:1710.09412*.

[14] J. Liu, F. Qu, X. Hong, and H. Zhang, "A small-sample wind turbine fault detection method with synthetic fault data using generative adversarial nets," *IEEE Trans. Ind. Informat.*, vol. 15, no. 7, pp. 3877–3888, Jul. 2019.

[15] W. Mao, Y. Liu, L. Ding, and Y. Li, "Imbalanced fault diagnosis of rolling bearing based on generative adversarial network: A comparative study," *IEEE Access*, vol. 7, pp. 9515–9530, 2019.

[16] Y. Zhang, X. Li, L. Gao, L. Wang, and L. Wen, "Imbalanced data fault diagnosis of rotating machinery using synthetic oversampling and feature learning," *J. Manuf. Syst.*, vol. 48, pp. 34–50, Jul. 2018.

[17] C. Liu, A. Mauricio, J. Qi, D. Peng, and K. Gryllias, "Domain adaptation digital twin for rolling element bearing prognostics," in *Proc. Annu. Conf. PHM Soc.*, vol. 12, 2020, p. 10.

[18] S. J. Pan, I. W. Tsang, J. T. Kwok, and Q. Yang, "Domain adaptation via transfer component analysis," *IEEE Trans. Neural Netw.*, vol. 22, no. 2, pp. 199–210, Feb. 2011.

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

[20] J. Jiang and C. Zhai, "Instance weighting for domain adaptation in NLP," in *Proc. 45th Annu. Meeting Assoc. Comput. linguistics*, 2007, pp. 264–271.

[21] Y. Li, N. Wang, J. Shi, J. Liu, and X. Hou, "Revisiting batch normalization for practical domain adaptation," 2016, *arXiv:1603.04779*.

[22] M. Long, Z. Cao, J. Wang, and M. I. Jordan, "Conditional adversarial domain adaptation," in *Proc. Adv. Neural Inf. Process. Syst.*, 2018, pp. 1640–1650.

[23] Y. Yang and S. Soatto, "FDA: Fourier domain adaptation for semantic segmentation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 4085–4095.

[24] Q. Lian, L. Duan, F. Lv, and B. Gong, "Constructing self-motivated pyramid curriculums for cross-domain semantic segmentation: A non-adversarial approach," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 6758–6767.

[25] P. Zhang, B. Zhang, T. Zhang, D. Chen, Y. Wang, and F. Wen, "Prototypical pseudo label denoising and target structure learning for domain adaptive semantic segmentation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 12414–12424.

[26] Q. Wang, D. Dai, L. Hoyer, O. Fink, and L. Van Gool, "Domain adaptive semantic segmentation with self-supervised depth estimation," in *Proc. Int. Conf. Comput. Vis.*, Oct. 2021, pp. 8515–8525.

[27] K. Mei, C. Zhu, J. Zou, and S. Zhang, "Instance adaptive self-training for unsupervised domain adaptation," in *Proc. 16th Eur. Conf. Comput. Vis.*, Glasgow, U.K.: Springer, Aug. 2020, pp. 415–430.

[28] W. Zhang, G. Peng, C. Li, Y. Chen, and Z. Zhang, "A new deep learning model for fault diagnosis with good anti-noise and domain adaptation ability on raw vibration signals," *Sensors*, vol. 17, no. 2, p. 425, 2017.

[29] X. Li, W. Zhang, Q. Ding, and J.-Q. Sun, "Multi-layer domain adaptation method for rolling bearing fault diagnosis," *Signal Process.*, vol. 157, pp. 180–197, Apr. 2019.

[30] B. Yang, Y. Lei, F. Jia, and S. Xing, "An intelligent fault diagnosis approach based on transfer learning from laboratory bearings to locomotive bearings," *Mech. Syst. Signal Process.*, vol. 122, pp. 692–706, May 2019.

[31] H. Zheng, R. Wang, Y. Yang, Y. Li, and M. Xu, "Intelligent fault identification based on multisource domain generalization towards actual diagnosis scenario," *IEEE Trans. Ind. Electron.*, vol. 67, no. 2, pp. 1293–1304, Feb. 2020.

[32] 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.

[33] Q. Wang, G. Michau, and O. Fink, "Domain adaptive transfer learning for fault diagnosis," in *Proc. Prognostics Syst. Health Manage. Conf. (PHM-Paris)*, May 2019, pp. 279–285.

[34] B. Zhang, W. Li, J. Hao, X.-L. Li, and M. Zhang, "Adversarial adaptive 1-D convolutional neural networks for bearing fault diagnosis under varying working condition," 2018, *arXiv:1805.00778*.

[35] Y. Wang, X. Sun, J. Li, and Y. Yang, "Intelligent fault diagnosis with deep adversarial domain adaptation," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–9, 2021.

[36] T. Li, Z. Zhao, C. Sun, R. Yan, and X. Chen, "Domain adversarial graph convolutional network for fault diagnosis under variable working conditions," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–10, 2021.

[37] Q. Zhang *et al.*, "Conditional adversarial domain generalization with a single discriminator for bearing fault diagnosis," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–15, 2021.

[38] X. Yu *et al.*, "Conditional adversarial domain adaptation with discrimination embedding for locomotive fault diagnosis," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–12, 2021.

[39] X. Li, W. Zhang, and Q. Ding, "Cross-domain fault diagnosis of rolling element bearings using deep generative neural networks," *IEEE Trans. Ind. Electron.*, vol. 66, no. 7, pp. 5525–5534, Jul. 2019.

[40] Z.-H. Liu, L.-B. Jiang, H.-L. Wei, L. Chen, and X.-H. Li, "Optimal transport-based deep domain adaptation approach for fault diagnosis of rotating machine," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–12, 2021.

[41] C. Chen, F. Shen, J. Xu, and R. Yan, "Domain adaptation-based transfer learning for gear fault diagnosis under varying working conditions," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–10, 2021.

[42] R. Yan, F. Shen, C. Sun, and X. Chen, "Knowledge transfer for rotary machine fault diagnosis," *IEEE Sensors J.*, vol. 20, no. 15, pp. 8374–8393, Aug. 2020.

[43] Z. Zhao *et al.*, "Unsupervised deep transfer learning for intelligent fault diagnosis: An open source and comparative study," 2019, *arXiv:1912.12528*.

[44] J. Zhu, N. Chen, and C. Shen, "A new multiple source domain adaptation fault diagnosis method between different rotating machines," *IEEE Trans. Ind. Informat.*, vol. 17, no. 7, pp. 4788–4797, Jul. 2021.

[45] W. Zhang, X. Li, H. Ma, Z. Luo, and X. Li, "Universal domain adaptation in fault diagnostics with hybrid weighted deep adversarial learning," *IEEE Trans. Ind. Informat.*, vol. 17, no. 12, pp. 7957–7967, Dec. 2021.

[46] K. You, M. Long, Z. Cao, J. Wang, and M. I. Jordan, "Universal domain adaptation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2019, pp. 2720–2729.

[47] Q. Wang, G. Michau, and O. Fink, "Missing-class-robust domain adaptation by unilateral alignment," *IEEE Trans. Ind. Electron.*, vol. 68, no. 1, pp. 663–671, Jan. 2021.

[48] G. Michau and O. Fink, "Unsupervised transfer learning for anomaly detection: Application to complementary operating condition transfer," *Knowl.-Based Syst.*, vol. 216, Mar. 2021, Art. no. 106816.

[49] R. B. Randall, J. Antoni, and S. Chobsaard, "The relationship between spectral correlation and envelope analysis in the diagnostics of bearing faults and other cyclostationary machine signals," *Mech. Syst. Signal Process.*, vol. 15, no. 5, pp. 945–962, 2001.

[50] P. D. McFadden and J. D. Smith, "Model for the vibration produced by a single point defect in a rolling element bearing," *J. Sound Vib.*, vol. 96, no. 1, pp. 69–82, Sep. 1984.

[51] P. Borghesani, R. Ricci, S. Chatterton, and P. Pennacchi, "A new procedure for using envelope analysis for rolling element bearing diagnostics in variable operating conditions," *Mech. Syst. Signal Process.*, vol. 38, no. 1, pp. 23–35, Jul. 2013.

[52] A. Gammerman, V. Vovk, and V. Vapnik, "Learning by transduction," in *Proc. 14th Conf. Uncertainty Artif. Intell.*, San Mateo, CA, USA: Morgan Kaufmann, 1998, pp. 148–155.

[53] A. Arnold, R. Nallapati, and W. W. Cohen, "A comparative study of methods for transductive transfer learning," in *Proc. 7th IEEE Int. Conf. Data Mining Workshops (ICDMW)*, Oct. 2007, pp. 77–82.

[54] Y. Ganin *et al.*, "Domain-adversarial training of neural networks," *J. Mach. Learn. Res.*, vol. 17, no. 1, pp. 2030–2096, May 2015.

[55] S. Arora, R. Ge, Y. Liang, T. Ma, and Y. Zhang, "Generalization and equilibrium in generative adversarial nets (GANs)," in *Proc. Int. Conf. Mach. Learn.*, 2017, pp. 224–232.

[56] M. Mirza and S. Osindero, "Conditional generative adversarial nets," 2014, *arXiv:1411.1784*.

[57] 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, pp. 100–131, Dec. 2015.

[58] S. Morris, "SKF pulp & paper practices," *SKF Global Pulp & Paper Segment*, to be published. [Online]. Available: https://www.skf.com/binaries/pub12/Images/0901d19680434822-Compilation-of-the-SKF-Pulp-and-Paper-Practices-issues-1-15-11147-EN_tcm_12-262199.pdf

[59] Z. Chen, A. Mauricio, W. Li, and K. Gryllias, "A deep learning method for bearing fault diagnosis based on cyclic spectral coherence and convolutional neural networks," *Mech. Syst. Signal Process.*, vol. 140, Jun. 2020, Art. no. 106683.

[60] J. Antoni, "The infogram: Entropic evidence of the signature of repetitive transients," *Mech. Syst. Signal Process.*, vol. 74, pp. 73–94, Jun. 2016.

[61] Y. Wang, J. Xiang, R. Markert, and M. Liang, "Spectral kurtosis for fault detection, diagnosis and prognostics of rotating machines: A review with applications," *Mech. Syst. Signal Process.*, vol. 66, pp. 679–698, Jan. 2016.

[62] W. Li, L. Wang, W. Li, E. Agustsson, and L. Van Gool, "WebVision database: Visual learning and understanding from web data," 2017, *arXiv:1708.02862*.

[63] J. Cohen, "A coefficient of agreement for nominal scales," *Educ. Psychol. Meas.*, vol. 20, no. 1, pp. 37–46, 1960.

[64] F. Pedregosa *et al.*, "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, Oct. 2011.

[65] X. Yue *et al.*, "Prototypical cross-domain self-supervised learning for few-shot unsupervised domain adaptation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 13834–13844.

**Qin Wang** received the bachelor's degree from the Department of Electrical Engineering, Tsinghua University, Beijing, China, in 2015, and the master's degree from the Department of Information Technology and Electrical Engineering, ETH Zürich, Zürich, Switzerland, in 2018. He is currently pursuing the Ph.D. degree with ETH Zürich.

His research interests include weakly supervised learning and domain adaptation.

**Olga Fink** (Member, IEEE) received the Diploma degree in industrial engineering from the Hamburg University of Technology, Hamburg, Germany, in 2008, and the Ph.D. degree from ETH Zürich, Zürich, Switzerland, in 2014, with the thesis on "Failure and Degradation Prediction by Artificial Neural Networks: Applications to Railway Systems."

She has been an Assistant Professor of intelligent maintenance systems with ETH Zürich, since October 2018, being awarded the Prestigious Professorship Grant of the Swiss National Science Foundation (SNSF). She is also a Research Affiliate with the Massachusetts Institute of Technology, Cambridge, MA, USA, and an Expert of the Innosuisse in the field of ICT. Her research focuses on intelligent maintenance systems, data-driven condition-based and predictive maintenance, hybrid approaches fusing physical performance models and deep learning algorithms, deep learning, and decision support algorithms for fault detection and diagnostics of complex infrastructure and industrial assets. Before joining ETH faculty, she was the Head of the Research Group "Smart Maintenance," Zurich University of Applied Sciences (ZHAW), Zürich, from 2014 to 2018. She has gained valuable industrial experience as a Reliability Engineer with Stadler Bussnang, Bussnang, Switzerland, AG and a Reliability and Maintenance Expert with Pöyry Switzerland Ltd., Zurich.

**Cees Taal** has a research background with over ten years of experience in the field of sensor signal processing and machine learning. He has worked in academia (Delft University of Technology, Delft, The Netherlands, and the KTH Royal Institute of Technology, Stockholm, Sweden) on audio and speech processing, after which he held various industrial positions in biomedical engineering (Philips Research, Eindhoven, Netherlands) and the energy domain (Eneco, Rotterdam, Netherlands). He is currently appointed as the Technologist of SKF, The Netherlands, where he is responsible for defining a research strategy in the field of bearing diagnostics and prognostics.

Dr. Taal received the IEEE Signal Processing Society Best Paper Award in 2016.