Journal of Dynamics, Monitoring and Diagnostics, 2024, 3, 60-82
https://doi.org/10.37965/jdmd.2024.530

FUTURE DIRECTION PAPER

Transfer Learning for Prognostics and Health Management: Advances, Challenges, and Opportunities

Ruqiang Yan,[1] Weihua Li,[2] Siliang Lu,[3] Min Xia,[4] Zhuyun Chen,[2,5] Zheng Zhou,[1] Yasong Li,[1] and Jingfeng Lu[3]

[1]School of Mechanical Engineering, Xi'an Jiaotong University, Xi'an, China
[2]School of Mechanical & Automotive Engineering, South China University of Technology, Guangzhou, China
[3]College of Electrical Engineering and Automation, Anhui University, Hefei, China
[4]Department of Mechanical & Materials Engineering, Western University, Ontario, Canada
[5]State Key Laboratory of Precision Electronic Manufacturing Technology and Equipment, Guangdong University of Technology, Guangzhou 510006, China

(Received 05 April 2024; Revised 18 April 2024; Accepted 14 May 2024; Published online 14 May 2024)

Abstract: As failure data is usually scarce in practice upon preventive maintenance strategy in prognostics and health management (PHM) domain, transfer learning provides a fundamental solution to enhance generalization of data-driven methods. In this paper, we briefly discuss general idea and advances of various transfer learning techniques in PHM domain, including domain adaptation, domain generalization, federated learning, and knowledge-driven transfer learning. Based on the observations from state of the art, we provide extensive discussions on possible challenges and opportunities of transfer learning in PHM domain to direct future development.

Keywords: domain adaptation; domain generalization; federated learning; knowledge-driven; PHM; transfer learning

I. INTRODUCTION

This paper reflects the important aspects in the field of transfer learning for prognostics and health management (PHM) domain. Opportunities and future directions are discussed. Section II on domain adaptation for fault diagnosis and prognosis was completed by Professor Weihua Li and Assistant Professor Zhuyun Chen from South China University of Technology. Section III on domain generalization for PHM was written by Associate Professor Min Xia from Western University. Section IV on federated learning for PHM was presented by Professor Siliang Lu and PhD candidate Jingfeng Lu from Anhui University. Section V on knowledge-driven transfer learning for PHM was written by Professor Ruqiang Yan and PhD candidate Zheng Zhou and Yasong Li from Xi'an Jiaotong University.

II. DOMAIN ADAPTATION-BASED FOR FAULT DIAGNOSIS AND PROGNOSIS

A. OVERVIEW

In actual industrial scenarios, complex systems often operate in variable conditions, which can make it challenging to conduct intelligent fault diagnosis and prognosis. This difficulty is especially pronounced in instances where there is a significant distribution discrepancy in the training and testing data, resulting in a well-trained model on the training set failing to achieve good predictive results on the test set. Domain adaptation as a subsect of transfer learning is the method used to address the distribution discrepancy between the training set (source domain) and the test set (target domain). The principle of the domain adaptation method is illustrated in Fig. 1(a). The core of this method is to transfer domain-invariant knowledge from the source domain to the target domain using techniques such as model transfer, statistical criterion, and adversarial learning, enhancing the model's performance in the target domain. Therefore, domain adaptation is currently a focal point in the field of fault diagnosis and prognosis.

This section provides an overview of domain adaptation methods applied for intelligent maintenance, as illustrated in Fig. 1(b)–(d). It summarizes the latest advancements in domain adaptation from three perspectives: model transfer-based domain adaptation, statistical criterion-based domain adaptation, and adversarial transfer-based domain adaptation. Additionally, the challenges and future directions related to intelligent maintenance of complex systems are discussed.

B. ADVANCES OF DOMAIN ADAPTION

1. MODEL TRANSFER-BASED DOMAIN ADAPTATION. Model transfer-based domain adaptation refers to sharing the structure and parameters of a model between the source domain and the target domain. For the mechanical intelligent maintenance, model transfer-based domain adaptation involves transferring the knowledge learned from other machinery, components, or working conditions to enhance the reliability of the model for fault diagnosis and life prediction in a new environment [1]. When there is a large amount of data in the source domain and few samples in the target domain, the model transfer-based domain adaptation

Corresponding author: Ruqiang Yan; Weihua Li; Siliang Lu; Min Xia (e-mails: yanruqiang@xjtu.edu.cn; whlee@scut.edu.cn; silianglu@ahu.edu.cn; min.xia@uwo.ca).

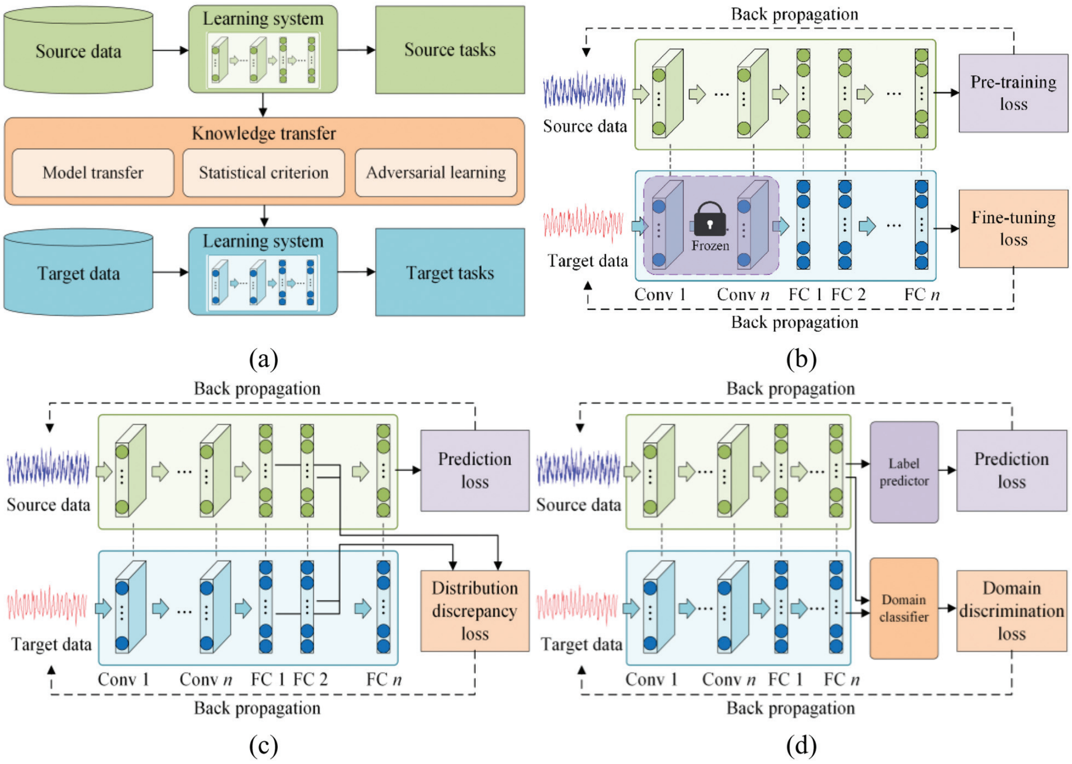

Fig. 1. Domain adaptation. (a) An illustration of domain adaptation. (b) Model-based domain adaptation. (c) Statistical criterion-based domain adaptation. (d) Adversarial learning-based domain adaptation.

can work effectively through pre-training and fine-tuning strategies. Specifically, as shown in Fig. 1(b), the model first undergoes supervised or unsupervised pre-training using data from the source domain to learn general parameters that are shared with the target domain, providing transferable prior knowledge. Subsequently, labeled data from the target domain is utilized for further supervised training of the model to appropriately adjust its parameters and enhance its adaptability to target tasks. The pre-training process can be defined as:

$$\theta^s = \arg\min_{\theta^s} \frac{1}{N^s} \sum_{i=1}^{N^s} \mathcal{L}^s(\mathcal{F}(x_i^s), y_i^s) \tag{1}$$

where x_i^s and y_i^s represent the samples and their corresponding labels in the source domain, respectively, N^s denotes the number of labeled source samples, $\mathcal{F}$ represents the model, $\mathcal{L}^s$ is the loss function during the pre-training stage, and θ^s signifies the model parameters obtained after pre-training. Additionally, the fine-tuning process can be defined as:

$$\begin{gathered}\theta^t = \arg\min_{\theta^t} \tfrac{1}{N^t} \sum_{i=1}^{N^t} \mathcal{L}^t(\mathcal{F}(x_i^t), y_i^t) \\ \text{s.t.} \quad \theta^t \text{ is initialized with } \theta^s\end{gathered} \tag{2}$$

where x_i^t and y_i^t represent the samples and their corresponding labels in the target domain, N^s denotes the number of labeled target samples, $\mathcal{L}^t$ stands for the loss function during the fine-tuning stage, and θ^t indicates the model parameters gained after fine-tuning.

The research progress of model transfer-based domain adaption in the field of fault diagnosis and life prediction task. Generally, the use of pre-training and fine-tuning strategies in model transfer-based domain adaptation can be divided into two implementation approaches: partial fine-tuning and full fine-tuning [2]. The characteristics of the summarized articles are shown in Table I.

Partial fine-tuning refers to the practice of freezing certain parameters of the pre-trained model during the fine-tuning process and updating only the remaining parameters, which is applicable to the situation where the data distribution of the source domain and the target domain is similar, or the labeled target data is relatively scarce. In the field of fault diagnosis and life prediction, domain adaptation methods employing partial fine-tuning strategies have yielded numerous research achievements in recent years. For instance, Chen *et al.* [3] proposed a transferable convolutional neural network (CNN) for diagnosing faults. Various datasets on rotating machinery were utilized as source domains to assist in pre-training the source model. The pre-trained model was segmented into multiple blocks, each of which was fine-tuned separately. Han *et al.* [4] introduced a domain adaptation framework based on CNN for gear and bearing fault diagnosis. Three different strategies for fine-tuning network layers were presented, and the performance of the model in different diagnosis scenarios was analyzed under different strategies. Zhong *et al.* [5] used a deep convolutional generative adversarial network (DCGAN) to generate abundant synthetic samples, which were considered as the source domain to pre-train VGG-16, while the original samples were treated as the target domain to fine-tune the deeper layers. In terms of mechanical life prediction, Berghout *et al.* [6] initially pre-trained a long short-term memory (LSTM) network with the source domain dataset and then fine-tuned the new added layers using a target domain dataset to improve remaining useful life (RUL) prediction performance of the target bearings.

In contrast, full fine-tuning means that updating all model parameters directly without freezing any during the

Table I. Solutions for model transfer-based domain adaptation

Types	Characteristic of methods	References	Application datasets	Tasks
Partial fine-tuning	Dividing the pre-trained model into multiple blocks and fine-tuning each block separately.	[3]	Gearbox dataset; bearing dataset	Fault diagnosis
	Providing three strategies for fine-tuning different layers to achieve domain adaptation under cross-condition and cross-class scenarios.	[4]	PHM09 gearbox dataset; gearbox dataset	Fault diagnosis
	Pre-training VGG-16 with generated samples by DCGAN, then fine-tuning the deeper layers with original samples.	[5]	CWRU bearing dataset; bearing dataset	Fault diagnosis
	Pre-training LSTM with source data and fine-tuning new added layers with target data.	[6]	PHM12 bearing dataset	RUL prediction
Full fine-tuning	Building NSTAE model for cross-condition fault diagnosis using full fine-tuning strategies.	[7]	Bearing dataset; gear dataset	Fault diagnosis
	Pre-training CNN with image dataset, then fine-tuning the CNN by wavelet scalograms.	[8]	CWRU bearing dataset; bearing dataset	Fault diagnosis
	Pre-training and full fine-tuning Bi-LSTM using data from different machinery to achieve effective RUL prediction in target domain.	[9]	XJTU bearing dataset; gearbox dataset	RUL prediction
	Introducing a feature alignment strategy based on KL divergence during pre-training and fine-tuning process to enhance the domain adaptation effects.	[10]	CNC machine tool dataset	RUL prediction

fine-tuning process. This approach is suitable for the situation where the data distribution of the source domain and the target domain is not very similar, or there is an ample amount of labeled target data. In the field of fault diagnosis and life prediction, domain adaptation methods using partial full fine-tuning strategies have shown excellent performance in certain scenarios. Shao *et al.* [7] developed a novel stacked transfer autoencoder (NSTAE) model for cross-domain fault diagnosis of bearings and gears. This model was first pre-trained using data from other rotating machinery, and then the pre-trained model was shared with the target, enhancing fault diagnosis performance. Zhong *et al.* [8] devised a domain adaptation approach that learns CNN parameters from an image dataset and then fine-tunes the pre-trained CNN using wavelet scalograms. To address the issue of predicting the RUL of turbine engines, Zhang *et al.* [9] proposed a domain adaptation model based on bidirectional LSTM (Bi-LSTM), which used different types of aero-engine sensing data to construct source domain and target domain, and implemented model transfer through pre-training and full fine-tuning strategies, effectively improving the RUL prediction performance of target tasks with scarce data. Sun *et al.* [10] designed a deep transfer learning network based on sparse autoencoder (SAE) for tool RUL prediction in cross-condition scenarios. During training, this network introduced a feature alignment strategy on top of model transfer to reinforce the adaptation between the source and target domains, thereby improving the accuracy of RUL prediction.

2. STATISTICAL CRITERION-BASED DOMAIN ADAPTATION. The statistical criterion-based method aims to enhance the model's performance on the target task by minimizing the discrepancy between the source and target domains. It achieves this by utilizing the distribution distance between the two domains as the loss function and leveraging deep neural networks to extract domain-invariant features. The principle of statistical criterion-based method is illustrated in Fig. 1(c). First, a statistical criterion is established for measuring distribution discrepancy between source and target domains. Then, raw data from the two domains is transformed into a shared latent feature space based on the criterion. Finally, the distribution discrepancy within the latent feature space is minimized to improve performance on the target task.

During the training process of deep neural networks, the model is able to update its parameters by minimizing the objective function. The objective function of statistical criterion-based domain adaptation consists of two components: one part is the classification loss function for labeled instance $R_C(\mathrm{X_L},\mathrm{Y_L})$, and the other part is the measurement of domain discrepancy based on statistical criteria. For example, the objective function for domain adaptation based on the statistical criterion maximum mean discrepancy (MMD) is as follows.

$$L = R_C(\mathrm{X_L},\mathrm{Y_L}) + \sum_{i=1}^{n} \lambda_i \mathrm{MMD}(\mathrm{h_i^S},\mathrm{h_i^T}) \tag{3}$$

where $\mathrm{MMD}(\mathrm{g})$ represents the statistical criterion based on MMD, $\mathrm{h^S}$ denotes the source domain, $\mathrm{h^T}$ denotes the target domain, n denotes the number of layers for adaptation, and λ_i denotes the penalty factor for the i-th adaptation layer.

The research progress of statistical criterion-based domain adaptation in the field of mechanical fault diagnosis and life prediction. Methods based on statistical criteria utilize means or higher-order moments to measure the discrepancy between domains. According to the different criterion types, these methods can be divided into five categories: (a) MMD and its variants: conditional maximum mean discrepancy (CMMD), multiple kernels MMD (MK-MMD), joint distribution adaptation (JDA), (b) correlation alignment (CORAL), (c) central moment discrepancy (CMD), (d) Wasserstein distances (WD), and (e) KL divergence. In recent years, various types of statistical criterion methods have achieved numerous research outcomes in the field of mechanical fault diagnosis and life prediction. Table II summarizes the main solutions of statistical criterion-based domain adaptation methods reported in the literature.

Table II. Solutions for statistical criterion-based domain adaptation

Algorithms used	Characteristic of methods	References	Application target/ datasets	Tasks
MMD	Employing five Gaussian kernels to construct a multi-kernel MMD.	[11]	C-MAPSS dataset, XJTU-SY dataset	RUL prediction
CMMD	Designing pseudo-labels of CMMD to construct WCMMD for evaluating conditional distribution distances.	[12]	Bearing dataset	Fault diagnosis
MK-MMD	Employing multiple kernels to construct a composite kernel in order to ascertain the optimal kernel.	[13]	IEEE PHM bearing dataset	RUL prediction
JDA	Integrating Marginal MMD and Conditional MMD with minimizing discrepancy in both marginal and conditional distributions.	[14]	Wind turbine dataset, CWRU dataset, gearbox dataset	Fault diagnosis
CORAL	Combining MMD and CORAL as a novel measure.	[15]	Custom-made RTS bearing dataset SWJTU bearing dataset	Fault diagnosis
CMD	Aligning the higher-order moments of distributions across domains.	[16]	Gearbox dataset	Fault diagnosis
WD	Aligning domain distributions by minimizing the Wasserstein distance between domains.	[17]	C-MAPSS dataset	RUL prediction
KL divergence	Measure both the first- and higher-order moments between distributions.	[18]	CWRU dataset, gearbox dataset	Fault diagnosis

Specifically, He *et al.* [11] proposed a joint MMD method to measure the distance between the marginal distributions and the conditional distributions of two domains. Shen *et al.* [12] utilized weighted CMMD to achieve intra-class adaptation, aligning equivalent samples between the source and target domains. Ding *et al.* [13] employed MK-MMD to quantify inter-domain discrepancy, facilitating the model to learn cross-domain feature representations. Han *et al.* [14] combined Marginal MMD and Conditional MMD to construct JDA, aligning both the marginal and conditional distributions between the source and target domains. Qian *et al.* [15] combined MMD and CORAL as a novel measure of distribution discrepancy to extract domain-invariant features. Xiong *et al.* [16] employed CMD as a domain adaptation metric regulator to learn domain-invariant features. Shi *et al.* [17] utilized the WD to measure the disparity between domains, aiming to reduce the distribution discrepancy between the source and target domains. Qian *et al.* [18] introduced the AHKL divergence to assess the discrepancy in both first- and higher-order moments of datasets.

3. ADVERSARIAL LEARNING-BASED DOMAIN ADAPTION. Adversarial learning-based domain adaptation is achieved by using an adversarial learning framework to learn domain-invariant representations [22]. In the research of predictive maintenance of electromechanical equipment, the domain adversarial migration model can learn the domain-invariant features under different parts and working conditions, overcoming the limitations of the traditional method based on labeled data, and improving the reliability of the model for predictive maintenance under unknown working conditions, equipment, and other complex environments. Specifically, as shown in Fig. 1(d), the domain adversarial network uses the labels of the source domain to guide the network learning, followed by the gradient reversal layer (GRL) to counter-train the generator and discriminator, and finally the source domain features are transformed to adapt to the target domain features, which enables the source domain to perform effective classification or regression on the target domain. The training objective of the domain adversarial network is to minimize the prediction error and maximize the domain classification error, and the specific process can be defined by equation (4):

$$\begin{aligned} E\left(\theta_f,\theta_y,\theta_d\right) &= \sum_{x_i\in D_s} L_y\left(G_y\left(G_f\left(x_i\right)\right),y_i\right) \\ &\quad -\lambda \sum_{x_i\in D_s\cup D_t} L_d\left(G_d\left(G_f\left(x_i\right)\right),d_i\right) \end{aligned} \tag{4}$$

where D_s and D_t are the source and target domains, respectively, x_i and y_i are the sample and the corresponding label, respectively, θ_f, θ_y, and θ_d represent the parameters of the feature extractor, label predictor, and domain discriminator, respectively, L_y and L_d are the loss of label predictor and domain discriminator, and d_i is the domain label.

The research progress of adversarial learning-based domain adaption in the field of fault diagnosis and life prediction task. Generally, domain adaptation methods based on adversarial transfer can be categorized based on different numbers of source and target domains as: a. Single-domain adversarial; b. Multi-domain adversarial; and C. Others. Table III summarizes the characteristics of different domain adversarial adaptation methods.

Single-domain adversarial aims to model and train the adversarial network in a single domain to try to learn inter-domain-invariant features. Therefore, some scholars' domain adaptation methods using single-domain adversarial have achieved numerous research results in recent years. Li *et al.* [23] proposed a novel weighted adversarial transfer network (WATN) for partial-domain fault diagnosis in rotating machinery, which used a weighted learning strategy and the domain discriminators to reduce the differences in the distribution of the shared classes among the domains. Li *et al.* [24] proposed a partial-domain adaptive RUL prediction method for incomplete target domain data of rolling bearings, which adopted a weighted degradation fusion scheme of source domain instances with an adversarial learning strategy as the main framework to achieve conditional domain adaptation under similar degradation levels. Zhu *et al.* [25] proposed an open-set adversarial

Table III. Solutions for adversarial learning-based domain adaption

Classification methodology	Algorithms used	Characteristic of methods	References	Application datasets	Task
Single-domain	WATN	Filtering out irrelevant source samples and minimizing cross-domain distribution differences in shared label space.	[23]	Bearing dataset, gearbox dataset	Fault diagnosis
	DANN	Modeling the data relationship across domains using instance level weighting mechanism	[24]	Bearing dataset	RUL prediction
	ANMAC	Using a weight module to evaluate domain knowledge in multiple auxiliary discriminators.	[25]	Bearing dataset	Fault diagnosis
	CR-DANN	Adding consistency-based regularization terms to eliminate negative effects of missing information in the target domain.	[26]	CMPASS dataset, XJTU-SY bearing dataset	RUL prediction
Multiple-domain	MDFN	Learning representations with space component analysis and entropy penalty strategy.	[27]	PU dataset, IMS dataset, CWRU dataset	Fault diagnosis
	MWDTN	Adaptive weighted learning from multiple complementary source domain datasets.	[28]	Bearing dataset, gearbox dataset	Fault diagnosis
	TS-MDAN	Combining MMD and CORAL metrics for cross-domain distribution adaptation.	[29]	IEEE PHM challenge bearing dataset	RUL prediction
	AOA	An adversarial generator is designed to maximize the variability and diversity of the generated pseudo-domains.	[30]	IEEE PHM challenge bearing dataset, XJTU-SY bearing dataset	RUL prediction
Others	DASMN	Integrating meta-learning and domain adaptive techniques for cross-domain fault identification.	[31]	CWRU dataset	Fault diagnosis
	KD-MDAN	Integrating knowledge distillation and adversarial networks to enable mutual learning between multiple-source models.	[32]	IEEE PHM challenge bearing dataset, XJTU-SY bearing dataset	RUL prediction

transfer method to solve the new fault identification in the target domain of rolling bearings problem by introducing a weighting module to distinguish the domain knowledge in multiple auxiliary classifiers. Siahpour *et al.* [26] proposed a new open-set RUL prediction method for rotating machinery based on domain adversarial transfer learning.

Multi-domain adversarial learning focuses on the rich underlying information in multiple domains and uses adversarial methods across multiple related domains in an attempt to learn inter-domain shared and private features. Therefore, several scholars have attempted to use multi-domain adversarial methods to solve cross-domain diagnostic and prediction tasks. Shi *et al.* [27] proposed a multi-source domain adversarial fault diagnostic network, which employed shared-space component analysis and entropy penalization strategy to extract the shared features of rolling bearings combining domain invariance and discriminability. Chen *et al.* [28] proposed a multi-source open-set domain adversarial fault diagnosis method, which employed multi-source domain data to extract fault information, and introduced a weighted learning strategy to weigh the importance of feature distribution alignment between known and unknown class samples. Ding *et al.* [29] proposed a new open-set RUL prediction framework for rolling bearings based on multi-source domain confrontation using a two-stage domain adaptation strategy, including domain-specific distributional adaptation and domain-specific regression adaptation. Ding *et al.* [30] proposed a multi-target domain adversarial method, which adopted an adversarial out-of-domain augmentation framework to improve the generalization ability of model in RUL prediction in rolling bearings.

In addition, some scholars have also tried to introduce some mainstream algorithms to optimize the domain adversarial training process to solve single- or multi-domain fault diagnosis and lifetime prediction tasks. Feng *et al.* [31] proposed a novel similarity-based meta-learning network with adversarial domain adaptation. Through meta-learning and adversarial learning, domain-invariant features are efficiently extracted from labeled source data and unlabeled target data, thus performing well on the target domain and solving the domain bias problem of rolling bearing fault identification in partial domains. Liu *et al.* [32] proposed a multi-source adversarial online knowledge distillation method for cross-device rolling bearing RUL prediction under multiple operating conditions within different machines.

C. CHALLENGES AND OPPORTUNITIES

Significant progress has been made in domain adaptation methods, yet many challenges remain for researchers to explore. In this section, the challenges in domain adaptation are discussed from three perspectives: model transfer-based statistical criterion-based and adversarial learning-based domain adaptation. It is believed that exploring these challenges will further advance the development and effective evaluation of domain adaptation methods and deepen theoretical research.

1) The main challenge in model-based domain adaptation is the difficulty of constructing a high-quality source domain to ensure effective model transfer. Model transfer-based domain adaptation can effectively utilize data from the source domain to assist model training, thereby avoiding overfitting issues that arise from directly training the model on target domain data. However, in real industrial scenarios, model transfer-based domain adaptation still faces a range of complex challenges. One noticeable challenge is that, to ensure the model performs well on target tasks, the source domain should bear a certain resemblance to the target domain and contain an adequate number of labeled samples for the model to undergo supervised training. Failure to meet this requirement will easily lead to the risk of negative transfer, resulting in a degradation of model performance. For tasks such as fault diagnosis and life prediction, it is hard to identify a suitable source domain, as the source data typically originate from other machinery, components, or working conditions. In real industrial environments, there are various forms of machinery, components, and working conditions, and it also requires substantial amounts of labor, time, and financial resources to collect and label samples, which makes it difficult to find a source domain that is similar to the target domain and has a large number of labeled samples. Therefore, to successfully implement model transfer-based domain adaptation in an industrial setting, overcoming the above-mentioned issues is undoubtedly a major challenge.

2) The challenges faced by statistical criterion-based domain adaptation primarily originate from two aspects: more complex tasks and theoretical research. a) Inadequate for more complex domain adaptation scenarios. Domain adaptation methods based on statistical criteria typically assume that the label spaces of the source and target domains are identical, implying both domains contain the same object categories. In such cases, the primary challenge revolves around addressing the distributional discrepancy between the source and target domains. However, in industrial scenarios, the label space of the source domain often differs from that of the target domain. In such cases, the problem to be addressed encompasses not only the distributional discrepancy between the source and target domains but also the discrepancy in their label spaces. Methods based on statistical criteria align the entire source and target domains. When the label spaces of the source and target domains are inconsistent, samples belonging to private label spaces lack corresponding samples for adaptation, thereby impacting the alignment effectiveness of samples belonging to shared label spaces and ultimately leading to negative transfer. Therefore, addressing domain adaptation problems in complex scenarios, such as when the label spaces of the source and target domains are inconsistent, is one of the challenges faced by statistical criterion-based domain adaptation. ***b) The absence of a universal statistical criterion framework.*** Various statistical criteria have shown good performance in fault diagnosis and RUL prediction tasks. However, these statistical criteria are isolated from each other, and there is limited research on the theoretical relationships between them, resulting in unclear understanding of the theoretical connections among various statistical criteria. Therefore, there is a lack of a theoretical framework to analyze and integrate the theoretical relationships among various statistical criteria.

3) Adversarial learning-based domain adaption still have some challenges interfering with the feature representation capabilities of the model and affecting the transferability of domain adversarial learning, originating from two aspects: difficulty of generalized knowledge transfer and limited defense against attacks, the more details are below. a) Difficulty of generalized knowledge transfer. At present, a large number of results have been achieved in the research of domain adversarial fault diagnosis and life prediction in partial domains and open domains, but most of them require the existence of labeling information in the target domain that is related to the source domain. In the actual industrial field, there exists a large amount of unknown label data and less common information with known fault labels or degradation states, the shared feature extraction between source and target domains is difficult. In addition, when there are multiple unknown target domains, the distribution differences between them can also cause the model to have difficulty in obtaining the common feature knowledge between the source and target domains, which affects the accuracy of fault diagnosis and RUL prediction. ***b) Limited defense against attacks.*** Domain adaptation models are susceptible to black-box or white-box attacks when performing adversarial training, resulting in difficulties for domain adversarial models to adapt the correct feature information in the target domain and achieve better transfer results on the target domain. For example, a black-box attacker cannot access the internal structure of the domain adversarial model but can try to trick the model by adding adversarial perturbations to the input samples in order to make it produce incorrect fault diagnosis and prediction results, whereas, a white-box attacker can use the model's internal training parameters, gradient information, and fault sample information, etc., to generate adversarial samples and interfere with the model domain adaptation process.

D. FUTURE RESEARCH DIRECTIONS

This section will discuss the future directions of three aspects of domain adaptation: model transfer-based, statistical criterion-based, and adversarial learning-based domain adaptation. These research directions will provide effective ways to address the current challenges.

1) The implementation of model transfer-based domain adaptation faces several challenges in fault diagnosis and life prediction. In actual industrial settings, differences between the source and target domains, along with the difficulty in acquiring sufficient labeled samples, put forward a test for the successful implementation of model transfer-based domain adaptation. ***In response, the rapidly growing trend of digital twin*** [19] ***technology in recent years may prove to be a valuable research direction***. By establishing a highly consistent mapping relationship between the physical and virtual worlds, digital twin technology enables to create a simulation environment in the digital realm that closely mirrors the target domain. Generating data in this simulation environment can yield a wealth of source data close to the actual target domain. Leveraging these data for pre-training not only promises better transfer performance but also eliminates the need to expend labor and financial resources to obtain labeled samples in the real environment, achieving significant results with half the effort. In addition, the recent high-profile large-scale model [20] technology is also worthy of attention. Large-scale model technology inherits the advantages of deep learning, and its massive scale allows the model to absorb and integrate vast amounts of data collected under various machinery, components, and working conditions, as well as simulation data generated through techniques like digital

twin. Consequently, the data scale of the source domain can be significantly increased, and richer information is offered for the model to learn, thus enhancing the effectiveness of the pre-training stage. Models that undergo pre-training on such large-scale data will possess more generic and superior parameters, enabling them to better adapt to the target domain during the subsequent fine-tuning stage and reducing the risk of overfitting.

2) Recognizing the existing challenges, the future research directions of statistical criterion-based domain adaptation are as follows. ***On the one hand***, significant progress has been made in addressing the issue of inconsistent labels space between the source and target domains. For example, Li *et al.* [21] systematically integrated transfer learning methods across different industrial application scenarios and provided recommendations for the selection of transfer methods in various scenarios. Combining methods based on statistical criteria with advanced methods in the field to fully leverage the strengths of different domain adaptation approaches, and addressing complex issues such as the inconsistent labels between source and target domain, will also be a research direction in the future. ***On the other hand,*** for domain adaptation tasks, experts and scholars have designed various statistical criteria-based methods from different perspectives. Theoretical analysis of these criteria, studying the theoretical relationships between different statistical criteria, and further unifying various statistical criteria on a theoretical basis can facilitate the theoretical development of statistical criteria. This approach can provide new perspectives for the design of statistical criteria.

3) Domain adversarial models need to leverage physical knowledge in order to improve model migration performance in complex scenarios, as well as to improve the model's defense capabilities against different types of attacks. The specific research directions are as follows. ***a) Supporting complex applications.*** Currently, domain adversarial models are less researched in complex domain scenarios where the target domain is unknown, so how to build a domain adversarial network that can match the unknown information space is a future research direction. For example, in the generalized domain scenario where the source and target domains do not have relevant label information, the knowledge and experience learned on multiple tasks can be adapted adversarial by combining multitask learning and meta-learning to improve the model's generalization ability on unknown domains. In addition, reinforcement learning can be used to solve the problem of unknown target domains by first using auxiliary tasks or self-supervised learning to generate reward signals and pseudo-domain labels, after which reinforcement learning optimizes the training process of the domain adversarial network using the reward signal as the learning objective to reduce the discrepancy between the pseudo-domain and the source domain. ***b) Building general defense system.*** No one defense method that can completely eliminate the risk of countering an attack. Attackers may continuously try new attack methods to interfere with the fault diagnosis and prediction process, so the combination of different defense methods is considered to build a more comprehensive defense system. For example, the input fault data is transformed using sine or wavelet transforms, which utilizes the structural similarity between different fault data to reduce the embedded noise in the samples. The model training process utilizes techniques such as gradient masking and neural network repair to enhance the robustness of the model against attacks. Finally, user access rights are set during model deployment to ensure that only authorized users or systems can interact with the model.

III. DOMAIN GENERALIZATION IN PHM

A. OVERVIEW

In practical PHM tasks, the training data usually comes from the same types of machines under different working conditions or different types of machines. Therefore, the domain shift affects the performance of data-driven PHM models; therefore, cross-domain models are needed. In recent years, domain generalization techniques have developed as an effective way to train deep learning-based models on multiple-source domains for PHM tasks. Domain generalization-based PHM models can reduce the dependence on target domain data and extract general features from source domains. These features are less sensitive to domain shifts for online condition monitoring. The domain generalization mechanism is divided into two main classes. One is to learn knowledge from different condition data of the same types of machine. Another is to extract shared PHM knowledge among multiple different machines with similar features. This section provides an overview of domain generalization for intelligent PHM with a focus on algorithms and applications. The paradigm of domain generalization applied in cross-domain PHM tasks is illustrated in Fig. 2.

B. ADVANCES OF DOMAIN GENERALIZATION

1. DATA MANIPULATION. The quantity, quality, and diversity of data can influence model performance including accuracy, generalization, etc. When training data is limited, domain generalization-based data processing methods can effectively enhance the diversity of data, which can increase the generalization of the model in fault diagnosis and RUL prediction tasks. Data processing methods include data augmentation and data generation methods, which can address insufficient faulty samples, poor cross-condition generalization, and vibration data distribution bias in PHM.

Data Augmentation: Data augmentation can effectively improve the training performance of the model. The basic method of data enhancement is to simulate domain shift. Specifically, this method transforms the original data into multiple forms but reserves the original labels [33]. Image data augmentation is the most widely used for many applications including PHM. The image augmentation consists of four methods, namely image transformation, adversarial gradients, model-based augmentation, and feature-based augmentation. In specific, image transformation increases the number of time-domain or frequency-domain images of the sensor signal by flipping, rotating, scaling, cropping, etc., to obtain more features in model training [34]. However, this can lead to label shifting, creating conflicts with other tasks. Adversarial gradients include the task adversarial method and the domain adversarial method, which can learn the domain-agnostic images to acquire more domain-invariant patterns. However, adversarial gradients can simulate real-world domain shifts,

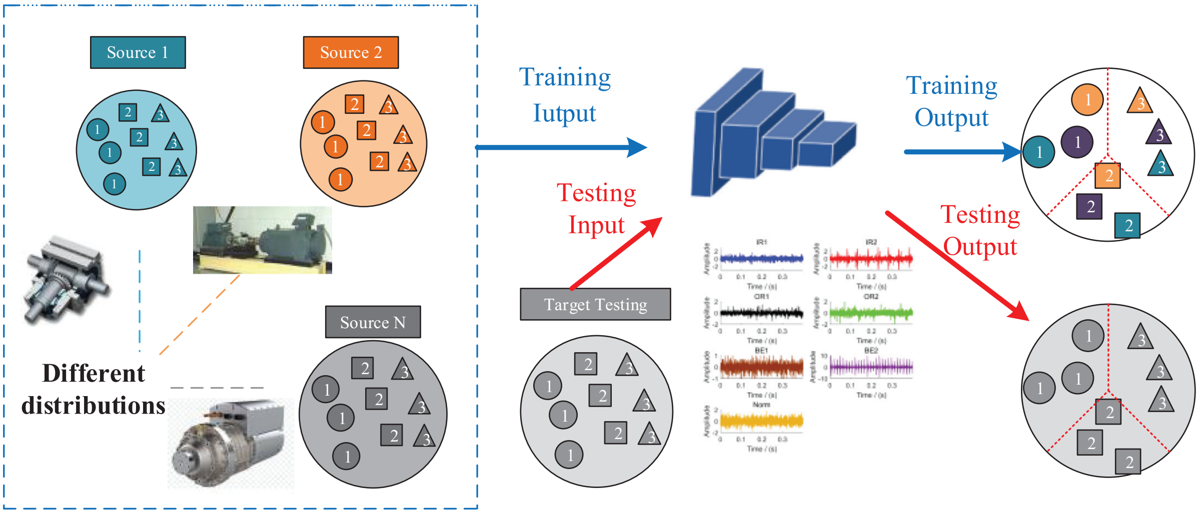

Fig. 2. Domain generalization applied in cross-domain PHM tasks.

which is more complex than salt-and-pepper noise in sensor signals. Model-based augmentation uses neural networks to generate new image data to mix with the original images for the source domain to improve the model performance of fault diagnosis. Feature-based augmentation utilizes the clustering method to extract the within-class and across-class prototypical representations for feature extractions [35].

Data generation: Data generation is one of the widely used techniques in domain generalization for PHM tasks. This method generates rich and diverse data samples, which can enhance the generalization of the model in multiple tasks. Currently, the typical methods are generating adversarial networks (GANs), variational autoencoder (VAE), and synthetic minority over-sampling techniques (SMOTE) [36,37]. A global optimization GAN was constructed for fault diagnosis, which applied AE as a generator to extract the fault features instead of the data sample [38]. Yan *et al.* established a VAE-based conditional Wasserstein GAN with a gradient penalty framework to generate synthetic faulty training samples for addressing the insufficient chiller faulty samples [39]. Chen *et al.* applied the adaptive synthesis as a novel SMOTE method to rich the fault samples for improving the accuracy of the wind turbine blade icing detection [40]. In addition, digital twins are a novel data generation method, which has obtained increasing attention. Xia *et al.* established a digital twin to build the simulation system for generating virtual data, which combined with the transfer learning method for cross-domain intelligence fault diagnosis [41].

2. REPRESENTATION LEARNING. Representation learning is a hot topic in machine learning and plays a significant role in domain generalization. The purpose of representation learning is to extract effective hidden features to construct accurate mapping relationships for fault diagnosis and RUL prediction. Representation learning can be divided into two categories: domain-invariant representation learning and invariant risk minimization respectively.

Domain-invariant representation learning. When the feature representation remains invariant across domains, this indicates that the features are general and can be applied to multiple domains. Therefore, the objective of domain generalization is to reduce the differences between feature representations of multiple-source domains to keep the domain invariant, which enhances the generalization of the model [42]. Currently, this method can be divided into three methods which are kernel-based approach, domain adversarial learning, and explicit feature alignment. Kernel-based methods are one of the primary methods for representation learning. The method uses kernel functions to map the original data to a high-dimensional space to achieve linear differentiability. The kernel methods widely used in domain generalization are polynomial kernel, Laplace kernel, and Gaussian radial basis kernel function. Gaussian kernel parameters are imported into one-class support vector machines to achieve fault detection, which selects the farthest and the nearest neighbors and "tightness" of the decision boundaries [43]. Cheng *et al.* proposed adaptive kernel spectral clustering to adaptively identify machine anomaly behaviors from multiple degradation features, which can outperform other feature extraction methods [44]. Domain adversarial learning has a generator and discriminator. Where the discriminator can distinguish the domains and the generator is applied to fool the discriminator to acquire the domain-invariant feature representations [45]. Chen *et al.* constructed a novel domain adversarial transfer network to deal with large distribution discrepancies across domains, which employed task-specific feature learning networks and domain adversarial training techniques [46]. A universal domain adaptation method was established for fault diagnosis, which applied source class-wise and target instance-wise weighting mechanisms to recognize the unknown fault modes [47]. Explicit feature alignment is applied to align the features across source domains by learning domain-invariant representations. Li *et al.* applied multitask instance normalization and batch normalization to enhance the informativeness of the extracted features to achieve the generalized bearing fault diagnostic framework [48]. A novel transfer learning-based method was constructed using local maximum mean difference and K-means to solve structural information in the unlabeled target samples [49].

Invariant risk minimization. Invariant risk minimization provides another way to learn the domain-invariance representation for domain generalization. The method

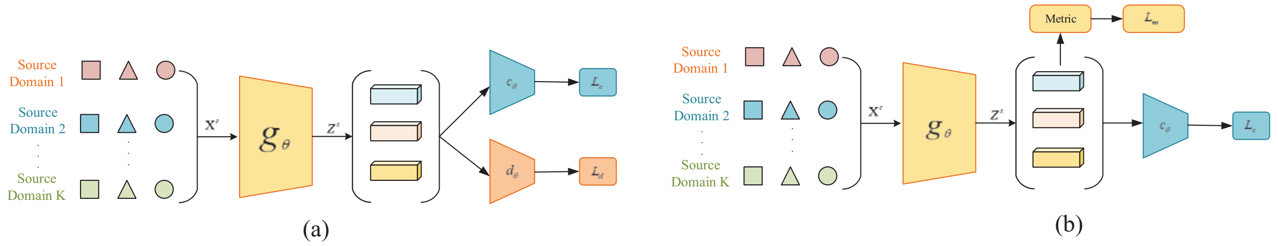

Fig. 3. Architectures of two typical methods of invariant representation learning. (a) Adversarial-based. (b) Metric-based.

considers the best classifier over the representation space to be the same across all domains, which does not require seeking to match the representation distribution across all domains [50]. The causal mechanism does not affect the mechanism, which can keep the domains invariant. Mo *et al.* established a sparsity-constrained invariant risk minimization framework, which applied models with better generalization for environmental disturbances in machinery fault diagnosis [51]. For graduating the invariance of optimal representation-level classifier, this learning method is acquired using minimizing the extrapolated risk among source domains, which can reduce the variance of source-domain risks [52]. Zhao *et al.* utilized the iterative min–max game of mutual information between the domain generation module and task diagnosis module to learn the generalized features for resisting the unknown domain shift, which can realize the domain-invariant representations from multiple-source domains [53].

The architectures of two typical methods of invariant representation learning are illustrated in Fig. 3.

3. LEARNING STRATEGY.

In parallel to data processing and representation learning, domain generalization can adopt different learning strategies in the general machine learning paradigms, including meta-learning, self-supervised learning, ensemble learning, etc.

Meta-learning. The key idea of meta-learning is to realize “learning from learning” and to construct general models using multiple methods, including optimization-based, metric learning-based, or model-based methods. This approach can learn extracted fragments for future learning in related tasks. Finn *et al.* first proposed a novel model that combined generalized domain with meta-learning, which can divide the training data into a meta-training set and a meta-testing set, and can learn features from the meta-training set to improve its performance [54]. The purpose of this method is to display the domain transfer of the model in training to better process the domain transfer in unknown domains. However, existing meta-learning-based domain generalization methods can only be employed for multi-source domains with labels and update the base model using second-order differentiation, which reduces the effectiveness and increases the computation cost in large-scale neural networks. Meta-learning has two important components, namely episodes and meta-representation. In particular, episodes can separate source domains into nonoverlapping meta-source and meta-target domains to simulate domain shift. And meta-representation is applied to denote the parameters for meta-learning.

In recent years, meta-learning methods have been rapidly developed in PHM, which are widely applied in complex working conditions, few-shot fault diagnosis, and RUL prediction. A novel hierarchical recurrent meta-learning-based method can be constructed to realize the fault diagnosis with small samples under different working conditions, which utilizes a recurrent meta-learning strategy with a one-shot learning way to train the proposed model [55]. Zhang *et al.* constructed a few-shot learning framework for bearing fault diagnosis using model-agnostic meta-learning (MAML), which can achieve the fault classifier with high performance and recognize the unknown faults using limited data [56]. Yang *et al.* established a MAML framework with a Gaussian process for RUL prediction, which can add kernel features as a regularization term to reduce the overfitting problem [57].

Self-supervised learning. Self-supervised learning can generate labels to learn from the data itself. This method can train models to predict transformations used in the image data, such as the rotation and patch-shuffling of the image. Self-supervised learning can learn generic features, which is beneficial to overcome overfitting domain-specific biases. And pretext task is an important part of self-supervised methods, which can automatically generate pseudo-labels to avoid manual labeling and realize unsupervised extraction. Overall, this method has good generalization, which is suitable for single-source scenarios and multi-source scenarios without requiring any domain labels. This method can effectively solve the insufficient labels of fault data samples, which can reduce the dependence on historical data. Currently, the self-supervised learning method has been widely used in several industrial scenarios, including incipient fault detection [58], fault diagnosis with unbalanced data, and feature extraction from monitoring data. Kong proposed a multitask self-supervised data mining approach, which can obtain massive diagnostic knowledge from unlabeled data to facilitate fault diagnosis [59]. Inter-instance and intratemporal self-supervised learning framework was constructed, which can process the unlabeled data integrated with a few labeled data to enrich the capacity of learnable data and ensure the stability of multitask optimization [60]. Mao *et al.* employed the generation of online RUL pseudo-values via fusing prior degradation information, which can establish a novel deep domain adversarial regression network with multilevel adaptation for estimating the online RUL values [61]. However, self-supervised learning-based methods were only applied to the object prediction and classifier tasks, but few literatures discussed the performance on out-of-distribution data generalization tasks. In addition, existing proxy tasks are specific rather than universal. For example, the rotation prediction task can mislead the model to capture rotation information instead of generic features, which can reduce generalization.

Ensemble learning. Ensemble learning combines multiple models to enhance the capabilities of models in prediction and classification. This method can construct special network structures and develop training strategies to fuse relationships between multiple-source domains, which can effectively enhance domain generalization. The model regards the representation of a sample on multiple-source domains as an integration, which means that the results of prediction or classification are obtained by integrating multiple networks [62]. Ensemble learning includes multiple methods, such as Exemplar-SVM, domain-specific neural network, domain-specific batch normalization, and weight averaging, which has been in practical industrial applications. Yu *et al.* proposed the Bayesian network-based probabilistic ensemble learning to address a limited diagnostic effect for industrial processes, which can satisfy a particular application [63]. A novel model was constructed using an improved domain adaptation method, which can apply ensemble learning to integrate multiple classifiers for the final results of fault diagnosis [64]. Degradation-dependent weights are embedded into an ensemble learning-based prognostic approach for RUL prediction, which can achieve higher accuracy than other methods [65]. In summary, ensemble learning is an effective model in domain generalization, which can effectively enhance the performance of the model about fault diagnosis and RUL prediction. The reason is that the method integrates multiple models and reserves the diversity of features. However, ensemble learning needs to store multiple different models, which increases the training cost and increase the computation time.

The overall framework of domain generalization is shown in Fig. 4. Table IV summarizes the characteristics of different domain generalization methods.

C. CHALLENGES AND OPPORTUNITIES

1. DATA SAFETY OF DOMAIN GENERALIZATION. In real industrial scenarios, it is very difficult for a single user to acquire well-developed monitoring data. Limited fault data samples cause challenges in constructing an effective PHM framework. Typically, similar data with labels from equipment from multiple users can be employed for cross-domain learning, which can improve the performance of the model. However, sharing data among users creates security issues due to potential interest conflicts and data privacy regulations. Therefore, it is a promising research direction to construct novel models for jointly modeling data from multiple users and ensuring data security. We can try to improve the domain generalization model by adopting federated learning to enhance the effectiveness and security of model training.

2. MODEL CAPACITY FOR UNKNOWN CONDITIONS. The weights of the machine learning models as a feature extractor are fixed after training in the source domain. This may lead to the representational ability of a machine learning model restricted to the seen domains. It reduces the generalization when the unknown fault distribution is different from the existing fault distribution. Existing approaches develop dynamic architecture-based models for PHM tasks, such as dynamic filter networks and conditional CNNs, which have achieved good results in processing monitoring data and extracting hidden features. However, it is a concern whether dynamic architectures are suitable for domain generalization in domain shifts. In addition, the normalization layer has become one of the core modules of deep learning. In model training, the parameters of the layers are only able to represent the distribution of the training data, when computed within each instance or based on small batches. Therefore, the models need to ensure that the parameters adapt to these domains of unknown faults or operational environments.

3. OPEN-SET DOMAIN GENERALIZATION FOR MONITORING DATA. Collecting all-inclusive fault data for potential fault detection can consume a lot of time in the PHM task. In online testing, the system can develop new fault modes and work conditions, which shifts the labels between the training data and the testing data. Diagnostic or prognostic models are typically trained to deal with samples in the source label space. However, if these faults are outside the source domain, it is critical to identify the unknown conditions accurately for testing the performance of the model. Generally, the known class samples in the test set are closer to the source domain than the unknown class samples in the feature space. Therefore, it is also a key issue how to measure the uncertainty of the unknown samples, which can enhance the interpretability and reliability of the model.

4. IMBALANCE DOMAIN GENERALIZATION FOR FAULT DIAGNOSIS. The assumption of most fault diagnosis methods is a uniform distribution of data. However, imbalance is typically the case such as in the healthy vibration

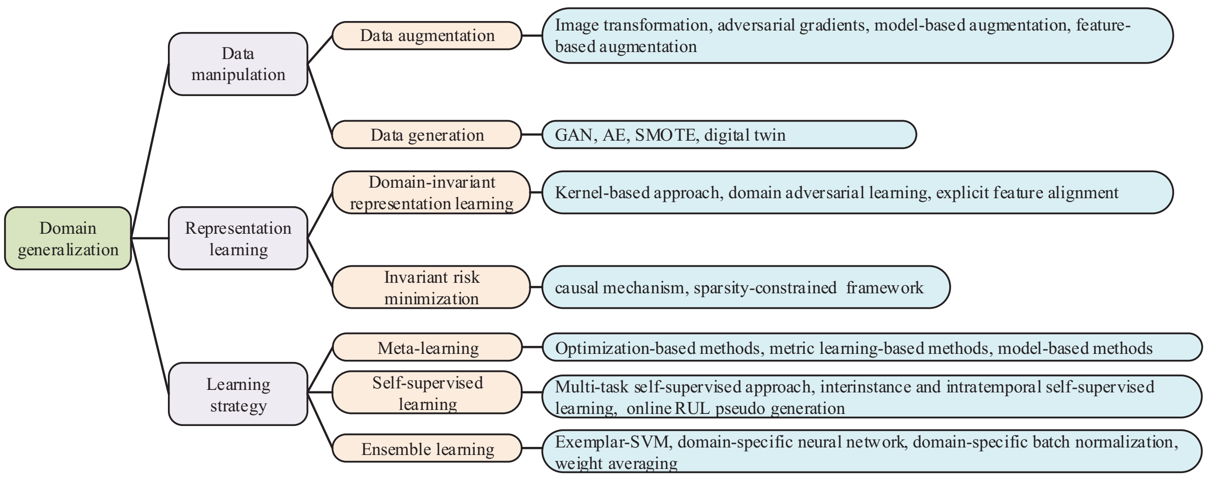

Fig. 4. The overall framework of domain generalization.

Table IV. Solutions for domain generalization in PHM

Algorithms used	Algorithms used	Characteristic of methods	References	Application datasets	Task
Data manipulation	CDAE	Extracting features to learn weights for further fine-tuning.	[37]	Laser-induced graphene production dataset	Condition monitoring
	GO-GAN	Using global optimization to generate more discriminant fault samples	[38]	Bearing dataset	Fault diagnosis
	ADASYN	Oversampling minority class to address the data imbalance.	[40]	Wind turbine fault data	Fault diagnosis
	Digital twin	Constructing simulation models to generate virtual data to train source domains.	[41]	Triplex pump fault dataset	Fault diagnosis
Representation learning	AKSC	Adaptive extraction of degradation features to detect abnormal states.	[44]	NSF I/UCR dataset	Condition monitoring
	DATN	Exploiting task-specific feature learning networks and domain adversarial training techniques for handling large distribution discrepancies across domains.	[46]	Bearing dataset, gearbox dataset	Fault diagnosis
	HWDAL	Recognizing unknown fault modes with class-level alignments without the target label set.	[47]	Bearing dataset	Fault diagnosis
	CDN	Minimizing the mutual information between subtask structures and capturing the causal invariant information for better generalization.	[48]	Bearing dataset	Fault diagnosis
Learning strategy	DRHRML	Using recurrent meta-learning strategy with one-shot learning way to train source domains.	[55]	CWRU dataset, bearing dataset	Fault diagnosis
	SDARA	Extracting temporal supervised information for the regression task through learning the degradation characteristics.	[61]	IEEE PHM challenge bearing dataset, XJTU-SY bearing dataset	RUL prediction
	DDW-EL	Assigning a degradation-dependent weight to each learner for better accuracy.	[65]	Rolling bearing data-set, C-MAPSS dataset	RUL prediction

data and fault data. Therefore, there is a need to improve the domain generalization capability of the model under unbalanced data. Existing data augmentation and data generation methods are mainly applied based on data distribution, which does not consider underlying physical information. Therefore, how to add physical constraints in domain generalization-based data processing is crucial.

D. FUTURE RESEARCH DIRECTIONS

1. INTERPRETABILITY AND CONTROLLABILITY. Data-driven based approaches have made significant progress in PHM tasks. However, the black-box nature of machine learning and deep learning increases the uncontrollability and non-interpretability of feature extraction. The solution for the above problems is to incorporate domain knowledge when the models and transferable representations are interpreted. For example, causal learning constructs are introduced for analyzing the generalized features of the model and eliminating unnecessary correlations. Evidence loss functions can also be introduced to analyze the uncertainty in the model for fault diagnosis to increase the reliability of the model in classification and prediction tasks.

2. GENERALIZATION WITHOUT DOMAIN LABELS. Most domain generalization-based methods need to utilize domain labels to implement PHM tasks. However, in real industrial scenarios, domain labels are difficult to acquire due to the complexity of machine operations. Especially, the data are from different users and are difficult to label due to differences between signal data features. Data lacking domain labels can decrease the performance of the domain generalization model. Currently, the literature about transfer learning-based models for PHM tasks rarely discusses about lack of domain labels, which lacks competitive inferiority to labeled domains in discussion. Because domain label learning is more effective and scalable, we encourage future work on the transfer learning-based PHM model to address this issue. For example, training models use data from labeled domains, which is used to evaluate the capabilities of the PHM model without domain labels, such as unknown fault types and operation conditions.

3. CONTINUOUS LEARNING AND SELF-EVOLUTION. In fault diagnosis and condition monitoring, the signal data is usually non-smooth. Therefore, we need to introduce continuous domain generalization to update the transfer learning-based model, which can reduce the negative impact of catastrophic forgetting to adapt to new data. Currently, the literature on continuous learning focuses on domain adaptive methods. However, few papers discuss continuous domain generalization. In addition, in condition monitoring, existing approaches usually assume smoothness in the health state, which is not suitable for practical industrial applications.

Therefore, novel models need to be designed, which have new domain generalization capabilities to adapt to unseen fault types. For example, pre-training and self-learning provide solutions for this topic. Existing pre-training methods cannot balance the prediction performance and training cost, so it is worth investigating how to design efficient domain generalization methods to help large-scale pre-training.

4. DATA REQUIREMENTS AND PERFORMANCE EVALUATION. The performance of existing domain generalization method models is similar to that of empirical risk minimization methods in PHM-related benchmark datasets. However, it cannot be proven that the performance of domain generalization does not have obvious advantages over traditional methods. The reason may be inappropriate evaluation methods, and datasets limit the performance. Therefore, we believe that researchers should address datasets with significant domain gaps. At the same time, the datasets should be general to accommodate the modeling demands of different scenarios, machine specifications, and PHM tasks. These databases can facilitate practical applications and promote the development of domain generalization-based methods.

IV. FEDERATED LEARNING IN PHM

A. OVERVIEW

Current mainstream PHM methods rely mainly on sensors to collect large amounts of real-time data from devices, and the real-time performance of such methods is one of the most important indicators of PHM tasks [66,67]. Owing to the problem of communication load, current PHM methods cannot be easily integrate directly into the Internet of things (IoT) technology [68]. In practical applications, data collection hardware typically uses open-source third-party communication protocols for compatibility reasons, which can pose a significant data privacy risk. The combination of FL technology with PHM models to solve the aforementioned problems has garnered considerable attention [69]. FL is a distributed computing technology based on the IoT and machine learning [70,71]. As shown in Fig. 5, the core idea behind the framework is to allow the clients to complete the partial model training directly using the locally collected data and the central server to no longer aggregate the original data of all the clients but receive only the trained parameters of each part of the model. After all the parameters are integrated and optimized, the global models are deployed to each client. The core technology of FL can be decomposed into two cores: (1) the architecture and learning strategy between the central end and the client and (2) the encryption technology that protects the transmission of the model parameters between the two ends [71].

Compared with traditional distributed machine learning techniques, FL has more unique advantages owing to its two core ideas. From an architectural perspective, the traditional distributed machine learning architecture is fully controlled by the central server for all the clients, with the data fully centralized on the central server [72]. The data storage and model update calculations are entirely undertaken by the central server, and the clients perform only the inference tasks. Meanwhile, the FL architecture emphasizes the data privacy of the client. In theory, in the FL architecture, the client will have complete local autonomy and can decide how and when to participate in collaborative learning [71]. In terms of computation, the FL architecture is more efficient than traditional architectures, because the client in an FL architecture can participate in the data storage and model update calculation tasks, and the communication content between the central end and the client will change from raw data to gradient information and model data parameters, thereby making it more advantageous in communication and energy consumption [73]. In terms of data privacy protection, in the FL framework, the model information transmitted between the client and the central end can be classified as indirect information of the working object [74]. When a hacker steals data from the IoT, they will only be able to infer the original information based on the model parameters or gradient information. Compared with the raw data transmitted through traditional distributed frameworks, those transmitted through an FL framework will encounter a natural barrier, which can considerably increase data theft difficulty. In addition, under an FL framework, clients will have high autonomy and be aligned with privacy protection agreements created in collaboration with various enterprises.

B. ADVANCES OF FL IN PHM

FL was first proposed by Google in 2016, which prompted many scholars to conduct research on the technique to propose effective methods [75–77]. The field of PHM

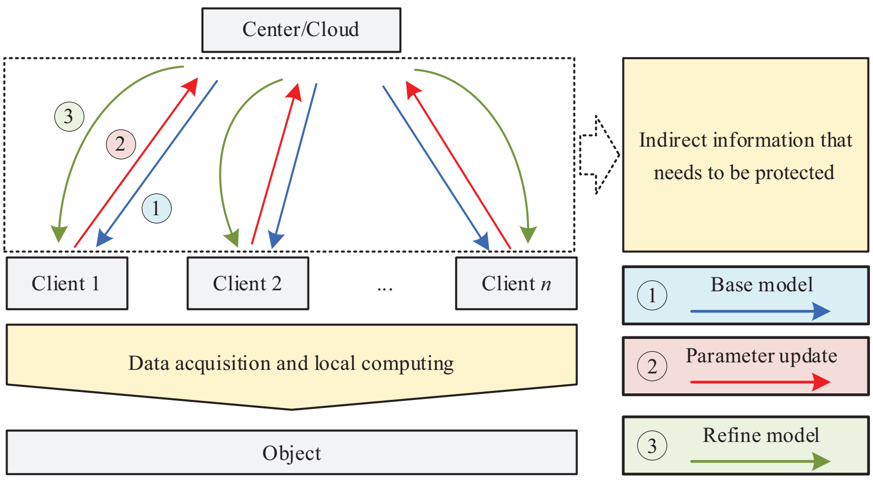

Fig. 5. Fundamental framework of FL.

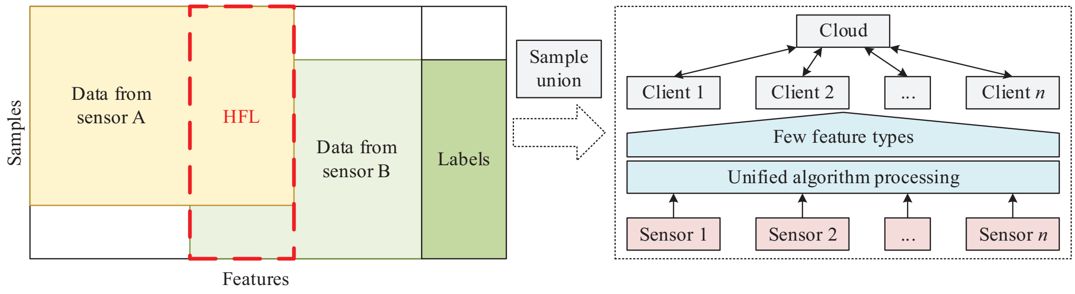

Fig. 6. Structure of HFL.

focuses mainly on the FL architecture, and some studies combined FL with PHM methods to address real-time and data privacy issues in the multidimensional fault diagnosis of complex electromechanical equipment in an IoT environment. Research on the FL framework can be grouped into three main categories [72,78].

1. HORIZONTAL FL (HFL). As shown in Fig. 6, the core of horizontal FL (HFL) is data from different clients having a similar feature but different forms. In PHM, the clients are commonly computing systems that use sensors for the data collection. HFL is essentially sample union and widely used owing to its simple and effective structure and concept. In the field of PHM, the application of FL is based mainly on the fusion of the horizontal structure and other methods, that is, an FL client network with different sensors to perform multidimensional and high-precision PHM on an object, such as bearings, motors, and so on.

Currently, previous research focused on the method of combining an FL framework with a PHM model. Using the Case Western Reserve University rolling bearing dataset, [79] proposed a self-supervised model with FL to address the data island problem in industry IoT. Reference [80] used lightweight technology, including nonstructural pruning and fine-tuning, to enhance an FL framework and lay the foundation for FL deployment in edge computing (EC) devices. Besides, previous research focused on the FL framework structure and tested the compatibility of proposed frameworks with different algorithms to ensure their effectiveness and generalization ability. For instance, [81] represented an FL framework by combining a process description with a software architecture and verified the approach by using industrial datasets and different FL algorithms. Reference [82] presented on-demand FL as an enhanced HFL method, which is a client deployment approach for FL, to expand the available capacity of a client and improve the horizontality of the FL framework. Meanwhile, [83] designed a Paillier-based communication scheme to preserve the raw information of shipping agents, and [84] proposed a hierarchical FL framework and diagnosed the fault of a power transformer with privacy protection.

2. VERTICAL FL (VFL). In a vertical FL (VFL) network, the users are commonly similar and few, and different data features are combined, as shown in Fig. 7. In PHM, few sensors are used in the monitoring network, and a single type of data is collected. However, different extraction methods are employed to obtain the multiple features, which are combined to achieve the fault diagnosis and state monitoring. The statistical features of electromechanical equipment are derived from raw data conversion algorithms, rather than collected directly, such as users' age, income, expenditure, and movement trajectory, similar to mobile IoT [85]. Therefore, VFL in PHM will not typically form a distributed network with hardware but will focus mainly on the feature extraction algorithm. At the algorithmic level, VFL essentially involves using multiple feature extraction algorithms to extract as much information as possible from a single data source, then fusing it. Reference [86] introduced the method of combining a generic framework (i.e., FedMeta-FFD) with an easy-to-implement enhancement technique (i.e., AILR) to address the few-shot problems in fault diagnosis and regarded such problems as verticality enhancement in an FL framework. Meanwhile, [87] presented a new model alignment method based on data-free knowledge distillation to generate pseudo-features that can improve model performance, which can also address the verticality of an FL framework.

3. FEDERATED TRANSFER LEARNING (FTL). Federated transfer learning (FTL) is demanding and challenging, in which little overlap exists between the customer samples and the features, as shown in Fig. 8 [88]. In PHM, few sensors are involved, as well as single data types and limited feature extraction algorithms [89]. In other words, despite its difficulty, very few features can be extracted in a PHM

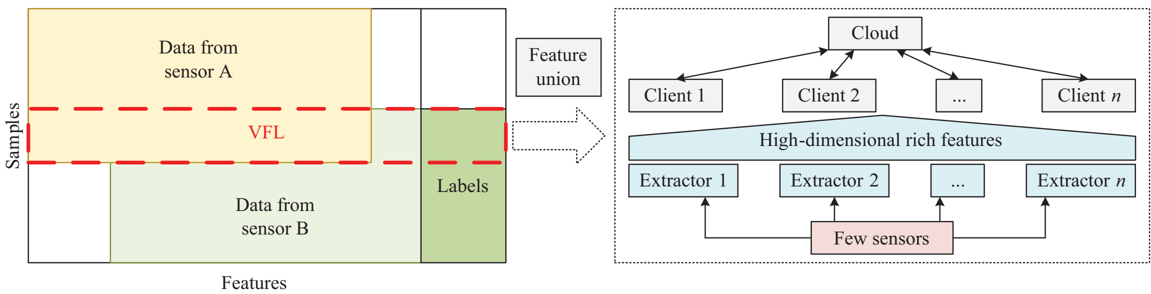

Fig. 7. Structure of VFL.

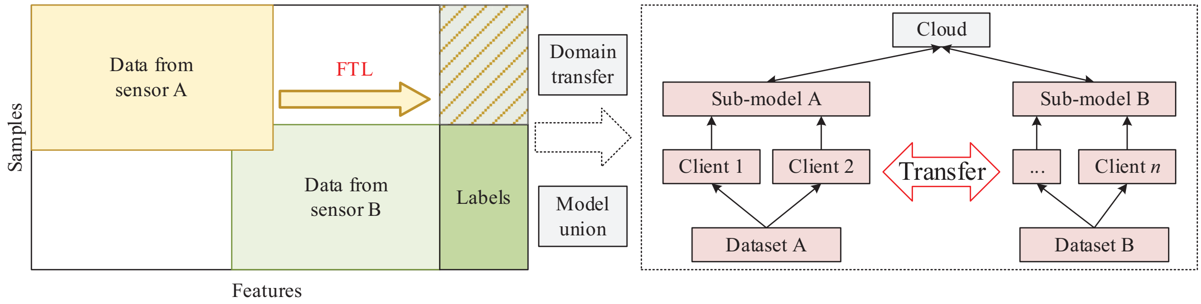

Fig. 8. Structure of FTL.

task. In common, special and important equipment with a complex data acquisition process and sensor placement will have strict requirements. In such a case, the entire distributed PHM model must be transferred based on the transfer learning and FL fusion using the common condition data of the devices as the source data domain [90]. FTL is one of the most effective methods for solving the problem of obtaining data for major equipment. Therefore, the current research focus is shifting gradually from HFL to FTL, which has become a key research direction for FL-based PHM. For example, [89] proposed an FTL framework with discrepancy-based weighted federated averaging to address the problem of potential domain shift in traditional federated averaging algorithms and validated the effectiveness and superiority of the method on self-made bearing datasets. Reference [91] developed the first unsupervised vertical FTL method for equipment fault diagnosis, which involved VFL and FTL. Reference [92] proposed a deep adversarial network-based FTL to address the domain shift phenomenon and data privacy problems.

4. FL-BASED CLOUD-EDGE COLLABORATION (CEC) COMPUTING. EC technology (ECT) and FL developed rapidly nearly during the same period [93,94]. The core idea behind ECT is the enhancement and simplification of deep learning models to complete inference tasks or model training tasks in edge devices with low computing resources to achieve high real-time performance [68,94,95]. The underlying PHM control hardware and data acquisition system commonly exploit low-process chips to resist complex electromagnetic interference and vibration interference, which can lead to limited computing resources. The introduction of ECT can effectively solve the application problems of high-performance models for PHM tasks, reduce the communication links of information transmission and results feedback, and improve real-time fault diagnosis performance and intelligent operation and maintenance. In addition, the characteristics of EC make it highly adaptable to FL. EC is the technical cornerstone of distributed training in FL, and FL is a distributed application scenario for EC [85,96]. Therefore, in recent years, an increasing number of scholars have begun to pay attention to FL-based cloud-edge collaboration (CEC) and exploit its advantages in computational performance, bandwidth reduction, and privacy protection to improve the performance of traditional PHM methods. For example, [97] effectively enhanced the EC performance of an FL framework and proposed a new FL framework, namely FedCAE, for bearing fault diagnosis. In an early research on CEC in RUL, [98] proposed an FL-based RUL prediction method. Moreover, to address the zero-shot problems in an ultra-supercritical thermal power group, Ref. [99] introduced a bidirectional alignment network and FL-based CEC, which was essentially a VFL framework.

5. SUMMARY OF FL APPLICATIONS FOR PHM. The development status of FL applications for PHM is summarized in Table V, and the current development status FL-based PHM can be summarized based on the characteristics of the objects, data types, and framework forms. FL-based PHM mainly involves the diagnosis of mechanical faults in objects such as bearings and gears, owing to the easy data acquisition process for such objects and the mature research on fault mechanisms. From a data perspective, vibration data are the main type of data driving the model, and establishing a highly intelligent PHM system would be monotonous for industrial IoT (IIoT). Thus, future studies should effectively combine temperature, current, voltage, and magnetic field strength information with an FL framework to present large PHM systems with continuous evolution capabilities. Moreover, few studies adopted a hybrid FL framework, and research on FL-based PHM framework integration demonstrates considerable potential. Existing studies on FL-based PHM focused mainly on single objects, such as bearings, which can limit the scale and breadth of FL frameworks. An ideal universal FL-based PHM framework should cover all the devices in an IIoT and perform intelligent maintenance on every working object, including gears, bearings, motors, and machine tools.

C. CHALLENGES AND OPPORTUNITIES

This study summarizes the future trends and challenges of combining PHM methods with FL by reviewing the development of different types of FL technologies. In addition, this study discusses the main directions for future research, including numerical calculation stability, system volume, and sustained intelligence.

1. FL-BASED FAULT DIAGNOSIS USING INFORMATION GENERATION MODELS. Sensor signal characteristics such as vibration, current, voltage, and magnetic field strength are extracted with algorithms or artificially defined. Thus, no differences exist between samples and features in FL-based PHM [100,101]. As sample generation techniques based on deep learning develop, signal sample and feature generation will exhibit convergence and become an essential information fusion form [102]. Information combination can substantially improve the horizontality and verticality of FL frameworks. However, as estimated values based on numerical statistics, generative information is naturally unstable and fatal for industrial applications. In addition, the clients in an FL framework are relatively

Table V. Summary of FL applications for PHM

Ref.	PHM object	Signal type	FL direction
[79]	Bearing	Vibration	HFL
[80]	Bearing, gearbox, and bogie	Vibration	HFL
[83]	Bearing	Vibration	HFL
[84]	Power transformer	Voltage	HFL
[86]	Bearing	Vibration	VFL
[87]	Bearing and gearbox	Vibration	VFL
[89]	Bearing	Vibration	FTL
[91]	Bearing and gearbox	Vibration	VFL, FTL
[92]	Bearing	Vibration	FTL
[97]	Bearing	Vibration	CEC
[98]	Milling cutter	Vibration	CEC
[99]	Ultra-supercritical thermal power group	Process variables	CEC, VFL

independent, and the lack of global information will lead to severe weak generalization problems in the model.

2. STABLE DISTRIBUTED FTL FRAMEWORKS FOR PHM. FTL is rarely combined with PHM models owing to the high demand for method stability in industrial scenarios, serious overfitting problems in data-driven models because of the periodicity of equipment signals, and the increasing instability of factors because of the parameter migration of the model and the distributed FL framework at the algorithmic and hardware levels [103]. Therefore, designing distributed TL models and FL systems has become a highly important challenge in the future owing to the absence of data on extreme operating conditions for major equipment [104].

3. EVOLUTIONARY PHM SYSTEMS BASED ON CONTINUOUS FL FRAMEWORK. As a continuous task, PHM requires the use of a CEC FL framework to further enhance the intelligence and autonomy of existing models, especially in the event of sudden unknown faults and unfamiliar working conditions during the PHM process [105]. Optional solutions include CEC-based self-updating models, unsupervised learning, and FL strategies that combine knowledge distillation, sample generation, and other optimization methods to form robust and generalized solutions that will consider the changes in the device environment [106].

V. KNOWLEDGE-DRIVEN TRANSFER LEARNING IN PHM

A. OVERVIEW

The purpose of transfer learning is to improve generalization ability in a new data domain with distribution shift. To solve this problem, the above domain adaption, domain generalization, and federated learning techniques generally aim to learn invariant representation from variable data domains to achieve generalization in joint distribution. Therefore, one of the key viewpoints of transfer learning is how to enable model to learn invariant representation with respect to domain shift. This is why we need to turn our attention to knowledge-driven model. In general, knowledge comes from our perception of physics phenomena and has strong adaptability to new data domain. It means knowledge itself is a kind of invariant representation with respect to domain shift, so it can generalize well to new domain. In addition, knowledge is usually independent of data acquisition. Therefore, compared to the above three techniques realizing transfer learning from data level, knowledge-driven transfer learning can provide extra perspective to learn invariant representation.

The sources of knowledge are diverse, including experiment, theory, computing science, and even data. From the viewpoint of experiment, knowledge can be concluded from the evolution law behind experimental results, like Newtonian mechanics. In this stage, knowledge is verified by observation and hypothesis testing. From perspective of theory, knowledge is derived by deductive inference and usually formulated as mathematical equations or model. In this stage, knowledge emphasizes universal laws and formalization. In the aspect of computing science, knowledge is based on complex mathematical model and solved by computing algorithms. In this stage, knowledge is reflected in information system. Currently, the most important source of knowledge is big data. The framework of data knowledge is built upon the above experimental observation, theory models, and computing science. In this stage, knowledge is usually implicit in complex models or algorithms and not easy to understand intuitively. Therefore, to make our direction paper more distinguishable, we limit the knowledge sources to the first three aspects, as shown in Fig. 9.

In this paper, we mainly illustrate application of three kinds knowledge in PHM, that is, signal processing, physics model, and experience induction. These three kinds knowledge have a long history of development in the PHM field and are supported by experimental observation, theory model, and computing science. In terms of signal processing, we can design various algorithms to extract interpretable feature from raw time series data. For example, we can use Fourier transform or wavelet transform to extract sparse impulse or period components from signal to investigate healthy condition of system. For physics model, it represents the basic understanding for the physical system of interest. In general, the first principle theory, like law of conservation of force or energy, is used to derive physics model. Depending on the nature of the concerned problem, different physics models can be constructed, such as lumped parameter model or finite element model. For experience induction, it has a broad definition, derived from the observation and summary of experiments, theories, and computing science. For example, for a prognostics

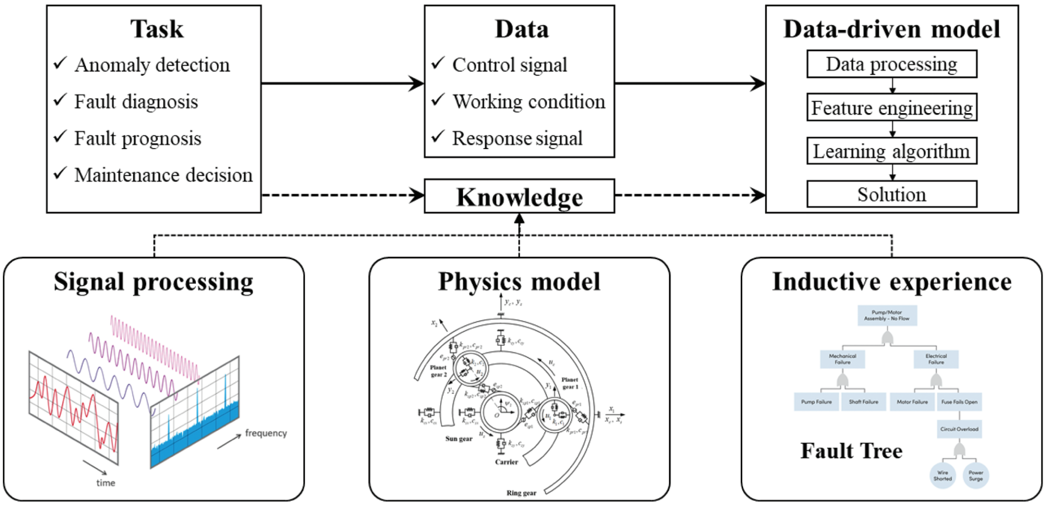

Fig. 9. Knowledge-driven transfer learning in PHM.

problem, we expect to build a monotonous, smooth, and trended healthy index to capture physical degradation. In the next part, we will summarize previous works on how to embed the knowledge from signal processing, physics model and experience induction into a data model, to realize knowledge-driven transfer learning.

B. ADVANCES OF KNOWLEDGE-DRIVEN TRANSFER LEARNING IN PHM

Many scholars have begun to study knowledge-driven transfer learning methods in the realm of PHM and have made certain progress. According to the existing research [107], knowledge is defined as information about a domain that can be utilized to solve issues in that domain. For the PHM domain, this paper subdivides knowledge into physical knowledge, signal processing knowledge, and inductive knowledge. By combining these knowledges, the model can obtain stronger transferability and generalization ability.

1. SIGNAL PROCESSING KNOWLEDGE. Signal processing knowledge is usually extracted from signal analysis and utilized to improve preprocessing methods or guide model design to eliminate domain shift. Kim *et al.* [108] employed low-pass filter and resample strategy to convert vibration signals from two mechanical systems into common pattern space. This can reduce feature distribution discrepancy between different datasets, thereby promoting invariant feature learning. Yu *et al.* [109] employed wavelet packet transform to construct time-frequency feature map, in which zigzag stitching reorders the coefficient matrix of leaf node. The constructed two-dimensional input is more sensitive to fault frequency and can be better combined with transfer algorithms. Kim *et al.* [110] considers that vibration signals from bearings can be described by impulse excitation with fault characteristic order, which is affected by rotational frequency and centrifugal force. Therefore, in the face of fault diagnosis problems under continuous nonstationary working conditions, multiple preprocessing methods are combined to reduce the influence of these two factors. These methods include speed normalization, angular resampling, envelope extraction and spectral analysis. The preprocessed signals from different domains can exhibit domain invariance. Li *et al.* [111] argued that the consistency of multi-scale entropy for the same class is capable of assisting transfer learning. For each sample, multi-scale symbolic dynamic entropy is calculated and thereby fed into the network. By combining two commonly used regularization loss terms in domain adaptation, the model can accurately diagnose fault types across domains. Obviously, such methods typically mine knowledge from the mechanisms of signal generation and attempt to construct invariant feature representations for different domains.

There are also some methods that promote the model to learn generalizable feature representations from the perspective of model construction. For instance, Liu *et al.* [112] developed a time-scattering convolutional network (TScatNet), in which Morlet wavelet kernel replaced the traditional convolution kernel to learn fault features. The predefined Morlet wavelet module possesses translation invariance and the deformation stability, allowing the learned scattering features from different working conditions to obtain similar distributions. Experiment results have indicated that the network can learn domain-invariant feature representations without applying any transfer learning techniques. Liu *et al.* [113] further proposed the normalized TScatNet (NTScatNet) for domain generalization tasks across different transmission paths. For a linear time-invariant system, it has been theoretically proven that the feature output from a scattering normalization layer is domain-invariant. He *et al.* [114] simplified the design of the network and only applied wavelet kernel convolution in the first layer of CNN. Specifically, the model enhances the transferability by initializing weights of the first layer with optimized wavelet weights. Experiments have revealed that this embedded knowledge can assist the model generalization. Yin *et al.* [115] embedded band-pass filters into convolutional layer to encourage the network to learn from regions with concentrated fault frequency. The prior knowledge about fault frequency improves the transferability of the network.

Overall, signal processing knowledge is closely related to the mechanism of fault occurrence, and reasonable application can enhance the cross-domain generalization ability of the model.

2. PHYSIC KNOWLEDGE. Physical knowledge is derived from objective physical laws and can be expressed through mathematical tools. This work divides the physical knowledge that drives transfer learning into two types: one is to establish a dynamic model to generate virtual data, and the other is to build physics-informed neural network (PINN). For the latter, Liu *et al.* [116] constructed a phenomenological model for the vibration response of bearings. On this basic, the full life cycle data of bearings can be obtained using exponential degradation. Finally, using simulation data as the source domain and real data as the target domain, domain adversarial neural network was combined to realize the model transfer. Zhang *et al.* [117] acquired life cycle simulation data of bearings by using a dynamic model with 5-freedom, in which degradation process is divided into four stages. Through the evolution of the bearing fault size, the vibration signals can show degradation trend. Domain adaptation techniques were utilized to transfer models from the virtual domain to the real domain. For the data scarcity problem in RUL prediction, by establishing simulation models, data sources can be enriched and the adaptability of the model in real environments can be significantly improved. Some studies adopt similar frameworks to model the objects to be monitored and combine them with transfer learning for training, including gas turbine [118], continuously stirred tank reactor [119], and triplex pump [120], etc. For the former, Gong *et al.* [121] proposed a PINN according to the physics mechanism of aerospace control actuator gyro. The association relationships between telemetry signals is modeled with neural networks, and changes in this relationship are used to detect anomalies. The indicators for anomaly monitoring are constructed through fine-tuning transfer learning strategy. Borate *et al.* [122] proposed a PINN to predict lab earthquakes. Fault physics is encoded into the loss function to improve predicted performance under the unknown condition. Lin *et al.* [123] constructed PINN using the physical topology structure of the power grid system, and spectral graph convolution network is applied to modeling node relationships. Through transfer learning technology, the network can achieve online evaluation for power system transient stability. Zhou *et al.* [124] designed PINN by introducing intrinsic correlation of physics between labeled gear fault data and unlabeled gear fault data. On this basis, this study further combines generative adversarial network, which significantly improves the generalization ability of the method.

The integration of physics knowledge usually improves the transparency of model decision-making and makes them more interpretable. This knowledge can drive the network to transfer to some unknown environments to solve problems. However, it should also be noted that the modeling of complex systems cannot be accurate enough and requires careful design to achieve excellent generalization effects.

3. INDUCTIVE KNOWLEDGE. Inductive knowledge is usually derived from experimental data or phenomena. By embedding inductive prior knowledge in the transfer framework, it can help the model generalize to different domains. For example, Mao *et al.* [125] considered that the data from bearing entities with similar degradation tendencies as the target domain is more conducive to transfer. Moreover, a transfer domain validity index is designed to quantify the contributions from different degradation features. Further research from Mao *et al.* [126] showed that feature transferability is affected by fault mode and degradation characteristic and a new indicator is constructed to assist the model selectively transfer features. Experimental results and theory analysis indicate that the proposed method is effective for cross-condition RUL prediction. Similarly, Zhu *et al.* [127] believed that uncertainty can also reflect whether features are suitable for transfer. In the feature space, samples with low uncertainty is generally located in a region of rich information. Therefore, a Bayesian neural network is designed to estimate sample uncertainty from the source domain and is utilized to guide samples to selectively transfer. Some researchers have proposed that sparse model parameters can better extract the feature representations of vibration signals. Xing *et al.* [128] designed a periodic cyclic sparse pattern in the convolutional layer and fully connected layer, in which a large number of model parameters are set to 0. Experiments illustrated that a reasonable sparse induction prior is beneficial to model cross-domain diagnosis. Li *et al.* [129] constructed a variational sparse attention layer in transformer network, in which Dirichlet distribution is set as the prior distribution to ensure sparsity. The visualization attention map reveals that high attention weights can correspond to fault-related features. Domain generalization experiment conducted in a bevel gear dataset verified the superiority of this inductive prior knowledge. Causal models are also often applied in conjunction with transfer learning. Guo *et al.* [130] attempted to guide domain generalization through causal metric interaction between two CNN models. To be specific, a CNN model selects the most causal regions as new data to assist the other CNN model train, in which conditional mutual information is employed to evaluate causal relationships. By learning the causal relationships among features, the network acquires stronger domain generalization capabilities.

Inductive knowledge is usually a phenomenon artificially summarized, rather than having strict definitions and rules like physical knowledge. Therefore, the fusion with transfer learning often does not require excessive redundant design, which makes such methods easy to train. But it should also be emphasized that inductive knowledge is not necessarily entirely correct, and unreasonable application may lead to negative transfer problem. Table VI summarizes the above literature on PHM objects, knowledge sources, and knowledge embedding methods.

C. CHALLENGES AND OPPORTUNITIES

According to the literature review above, knowledge-driven transfer learning has achieved some advancements. Despite this, such research still faces some challenges that need to be addressed by researchers.

1. COMPLEX KNOWLEDGE REPRESENTATION. For complex monitoring objects, domain knowledge is often unstructured. It is challenging to represent this knowledge and embed it into models in a reasonable way. For example, for a mechanical system, the established dynamic model is difficult to approximate the real physical model. Virtual data generated using dynamic models is likely to lead to negative transfer phenomena. Or, the transmission paths of different sensor signals are inconsistent, making it difficult to express knowledge under unknown conditions. This work suggests that future research should attempt to explore cross-domain-invariant knowledge and represent it using

Table VI. Summary of knowledge-driven transfer learning for PHM

Ref.	PHM object	Knowledge source	Embedding method
[108]	Bearing	Signal processing	Data preprocessing
[109]	Bearing	Signal processing	Data preprocessing
[110]	Bearing	Signal processing	Data preprocessing
[111]	Bearing	Signal processing	Data preprocessing
[112]	Bearing	Signal processing	Model design
[113]	Bearing	Signal processing	Model design
[114]	Bearing	Signal processing	Model design
[115]	Bearing	Signal processing	Model design
[116]	Bearing	Physics	Physical simulation model
[117]	Bearing	Physics	Physical simulation model
[118]	Gas turbines	Physics	Physical simulation model
[119]	Stirred tank reactor	Physics	Physical simulation model
[120]	Triplex pump	Physics	Physical simulation model
[121]	Aerospace control moment gyro	Physics	PINN design
[122]	Earthquakes	Physics	PINN design
[123]	Power system	Physics	PINN design
[124]	Gear	Physics	PINN design
[125]	Bearings	Inductive	Indicator construction
[126]	Bearings	Inductive	Indicator construction
[127]	Bearings, lithium-ion battery	Inductive	Indicator construction
[128]	Bearing	Inductive	Model design
[129]	Bearing, gear	Inductive	Regular term design
[130]	Gear	Inductive	Model design

reasonable embedding methods. Digital twin-driven transfer learning should be given attention, as it realizes information exchange between physical and virtual domains. Digital twins can mine the physical knowledge of monitoring objects and update the evolution process of objects in real time, achieving generalization in the time domain.

2. GENERALIZABLE INTERPRETABILITY. Embedding knowledge as domain information into a transfer learning framework is often accompanied by a focus on interpretability. Interpretability means that humans can understand the decision-making basis of the method, which requires a transparent and trustworthy model. Although some of the research mentioned above, especially in the field of physical knowledge, involves interpretability, it is currently not closely connected to transfer learning. In particular, few studies focus on the interpretability of models generalizing to unknown conditions. Sometimes, the accuracy of the model does not decrease significantly, but the decision logic may change due to domain drift. This review suggests that future research should try to focus on generalizable interpretability. The method can not only learn domain-invariant feature representations but also learn domain-invariant decision-making logic, which is beneficial to the application of the method in the industrial field. This puts forward higher requirements for the representation and integration of domain knowledge. Moreover, how to achieve doubly fed optimization between interpretability and generalization also needs attention. In other words, there is a need to study the work of interpretability guiding model generalization and model generalization embodying interpretability.

D. FUTURE RESEARCH DIRECTIONS

To bring development of knowledge-driven transfer learning to a better level and enable it with stronger generalization ability, better flexibility, and more interpretability, there are mainly two directions from both theory and application to further research in this topic.

In theory part, we should concern the acquisition approach of knowledge, the way that knowledge is represented, and how to embed knowledge into data models. In terms of knowledge acquisition, knowledge usually requires expensive labor and time costs to condense and requires trial and error to verify in experiments. So how to optimize the acquisition pipeline of knowledge is the first problem we should solve. The next problem is how to represent knowledge. As knowledge is often expressed as complex and unstructured relation and sometimes there are no explicit formulas, we have to utilize several techniques to capture the property of knowledge, such as knowledge graph or fuzzy rule theory. The technical part is to build a universal framework of knowledge embedding approach to data models, which is also the mainstream trend in recent researches. Generally, knowledge is viewed as some kinds of constraint for data models, and there are two strategies to realize such constraint. The first strategy is to modify the model structure to confirm knowledge, and then the prediction logic of model will be constrained by knowledge. For example, we can utilize Fourier transform or wavelet transform to design the structure of neural network. and then the feature extraction module will follow the basic logic in signal processing. The second strategy is to construct an optimization regularization from knowledge, and then the solution of data model will be limited in a regularized space. PINN is a typical example of such strategy.

In application part, an important thing is to define an evaluation approach to verify the effect of knowledge-driven transfer learning in generalization ability. As claimed that knowledge is an invariant representation with respect to domain shift, knowledge-driven data models are expected to generalize well in new domain, such as varying working condition or multi-modal noise. However, we can only review little literatures considering generalization ability of knowledge-driven model. Therefore, a uniform evaluation framework will greatly improve the development of knowledge-driven transfer learning and extend its downstream tasks in PHM.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

REFERENCES

[1] X. Chen, R. Yang, Y. Xue, M. Huang, R. Ferrero, and Z. Wang, "Deep transfer learning for bearing fault diagnosis: a systematic review since 2016," *IEEE Trans. Instrum. Meas.*, vol. 72, pp. 1–23, 2023.

[2] W. Li et al., "A perspective survey on deep transfer learning for fault diagnosis in industrial scenarios: theories, applications and challenges," *Mechanical Systems and Signal Processing*, vol. 167, p. 108487, 2022.

[3] Z. Chen, K. Gryllias, and W. Li, "Intelligent fault diagnosis for rotary machinery using transferable convolutional neural network," *IEEE Trans. Ind. Inf.*, vol. 16, no. 1, pp. 339–349, 2020.

[4] T. Han, C. Liu, W. Yang, and D. Jiang, "Learning transferable features in deep convolutional neural networks for diagnosing unseen machine conditions," *ISA Transactions*, vol. 93, pp. 341–353, 2019.

[5] H. Zhong, S. Yu, H. Trinh, Y. Lv, R. Yuan, and Y. Wang, "Fine-tuning transfer learning based on DCGAN integrated with self-attention and spectral normalization for bearing fault diagnosis," *Measurement*, vol. 210, p. 112421, 2023.

[6] T. Berghout, L.-H. Mouss, T. Bentrcia, and M. Benbouzid, "A semi-supervised deep transfer learning approach for rolling-element bearing remaining useful life prediction," *IEEE Trans. Energy Convers.*, vol. 37, no. 2, pp. 1200–1210, 2022.

[7] S. Haidong, D. Ziyang, C. Junsheng, and J. Hongkai, "Intelligent fault diagnosis among different rotating machines using novel stacked transfer auto-encoder optimized by PSO," *ISA Transactions*, vol. 105, pp. 308–319 2020.

[8] H. Zhong, Y. Lv, R. Yuan, and D. Yang, "Bearing fault diagnosis using transfer learning and self-attention ensemble lightweight convolutional neural network," *Neurocomputing*, vol. 501, pp. 765–777, 2022.

[9] H. Zhang, Q. Zhang, S. Shao, T. Niu, X. Yang, and H. Ding, "Sequential network with residual neural network for rotatory machine remaining useful life prediction using deep transfer learning," *Shock and Vibration, vol.* 2020, pp. 1–16, 2020.

[10] C. Sun, M. Ma, Z. Zhao, S. Tian, R. Yan, and X. Chen, "Deep transfer learning based on sparse autoencoder for remaining useful life prediction of tool in manufacturing," *IEEE Trans. Ind. Inf.*, vol. 15, no. 4, pp. 2416–2425, 2019.

[11] R. He, Z. Tian, and M. Zuo, "A transferable neural network method for remaining useful life prediction," *Mechanical Systems and Signal Processing*, vol. 183, p. 109608, 2023.

[12] C. Shen, X. Wang, D. Wang, Y. Li, J. Zhu, and M. Gong, "Dynamic joint distribution alignment network for bearing fault diagnosis under variable working conditions," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–13, 2021.

[13] Y. Ding, P. Ding, and M. Jia, "A novel remaining useful life prediction method of rolling bearings based on deep transfer auto-encoder," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–12, 2021.

[14] T. Han, C. Liu, W. Yang, and D. Jiang, "Deep transfer network with joint distribution adaptation: a new intelligent fault diagnosis framework for industry application," *ISA Transactions*, vol. 97, pp. 269–281, 2020.

[15] Q. Qian, Y. Qin, J. Luo, Y. Wang, and F. Wu, "Deep discriminative transfer learning network for cross-machine fault diagnosis," *Mechanical Systems and Signal Processing*, vol. 186, p. 109884, 2023.

[16] P. Xiong, B. Tang, L. Deng, M. Zhao, and X. Yu, "Multi-block domain adaptation with central moment discrepancy for fault diagnosis," *Measurement*, vol. 169, p. 108516, 2021.

[17] H. Shi, C. Huang, X. Zhang, J. Zhao, and S. Li, "Wasserstein distance based multi-scale adversarial domain adaptation method for remaining useful life prediction," *Appl Intell*, vol. 53, no. 3, pp. 3622–3637, 2023.

[18] W. Qian, S. Li, and X. Jiang, "Deep transfer network for rotating machine fault analysis," *Pattern Recognition*, vol. 96, p. 106993, 2019.

[19] L. Ma, B. Jiang, L. Xiao, N. Lu, "Digital twin-assisted enhanced meta-transfer learning for rolling bearing fault diagnosis," *Mechanical Systems and Signal Processing*, vol. 200, p. 110490, 2023.

[20] N. Ding et al., "Parameter-efficient fine-tuning of large-scale pre-trained language models," *Nat Mach Intell*, vol. 5, no. 3, pp. 220–235, 2023.

[21] W. Li et al., "A perspective survey on deep transfer learning for fault diagnosis in industrial scenarios: theories, applications and challenges," *Mechanical Systems and Signal Processing*, vol. 167, p. 108487, 2022.

[22] Y. Ganin et al., "Domain-adversarial training of neural networks," *Journal of Machine Learning Research*, vol. 17, no. 59, pp. 1–35, 2016.

[23] W. Li, Z. Chen, and G. He, "A novel weighted adversarial transfer network for partial domain fault diagnosis of machinery," *IEEE Trans. Ind. Inf.*, vol. 17, no. 3, pp. 1753–1762, 2021.

[24] X. Li, W. Zhang, X. Li, and H. Hao, "Partial domain adaptation in remaining useful life prediction with incomplete target data," *IEEE/ASME Trans. Mechatron.*, pp. 1–11, 2024.

[25] J. Zhu, C.-G. Huang, C. Shen, and Y. Shen, "Cross-domain open-set machinery fault diagnosis based on adversarial network with multiple auxiliary classifiers," *IEEE Trans. Ind. Inf.*, vol. 18, no. 11, pp. 8077–8086, 2022.

[26] S. Siahpour, X. Li, and J. Lee, "A novel transfer learning approach in remaining useful life prediction for incomplete dataset," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–11, 2022.

[27] Y. Shi, A. Deng, X. Ding, S. Zhang, S. Xu, and J. Li, "Multisource domain factorization network for cross-domain fault diagnosis of rotating machinery: an unsupervised multisource domain adaptation method," *Mechanical Systems and Signal Processing*, vol. 164, p. 108219, 2022.

[28] Z. Chen et al., "A multi-source weighted deep transfer network for open-set fault diagnosis of rotary machinery," *IEEE Trans. Cybern.*, vol. 53, no. 3, pp. 1982–1993, 2023.

[29] Y. Ding, P. Ding, X. Zhao, Y. Cao, and M. Jia, "Transfer learning for remaining useful life prediction across operating conditions based on multisource domain adaptation," *IEEE/ASME Trans. Mechatron.*, vol. 27, no. 5, pp. 4143–4152, 2022.

[30] Y. Ding, M. Jia, Y. Cao, P. Ding, X. Zhao, and C.-G. Lee, "Domain generalization via adversarial out-domain augmentation for remaining useful life prediction of bearings under unseen conditions," *Knowledge-Based Systems*, vol. 261, p. 110199, 2023.

[31] Y. Feng et al., "Similarity-based meta-learning network with adversarial domain adaptation for cross-domain fault identification," *Knowledge-Based Systems*, vol. 217, p. 106829, 2021.

[32] K. Liu and Y. Li, "Remaining useful life prediction across machines using multi-source adversarial online knowledge distillation," *Engineering Applications of Artificial Intelligence*, vol. 130, p. 107726, 2024.

[33] N. Honarvar Nazari and A. Kovashka, "Domain generalization using shape representation," In *European Conference on Computer Vision*. Cham: Springer, 2020.

[34] R. Volpi and V. Murino, "Addressing model vulnerability to distributional shifts over image transformation sets," *Proceedings of the IEEE/CVF International Conference on Computer Vision*. vol. 0, pp. 7979–7988, 2019.

[35] C.-W. Kuo, C.-Y. Ma, J.-B. Huang, and Z. Kira, "Featmatch: feature-based augmentation for semi-supervised learning," In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XVIII*. Heidelberg: Springer, 2020.

[36] J. Li, C. Shen, L. Kong, D. Wang, M. Xia, and Z. Zhu, "A new adversarial domain generalization network based on class boundary feature detection for bearing fault diagnosis," *IEEE Trans Instrum Meas*, vol. 71, pp. 1–9, 2022.

[37] M. Xia, H. Shao, Z. Huang, Z. Zhao, F. Jiang, and Y. Hu, "Intelligent process monitoring of laser-induced graphene production with deep transfer learning," *IEEE Trans Instrum Meas*, vol. 71, pp. 1–9, 2022.

[38] F. Zhou, S. Yang, H. Fujita, D. Chen, and C. Wen, "Deep learning fault diagnosis method based on global optimization GAN for unbalanced data," *Knowl Based Syst*, vol. 187, p. 104837, 2020.

[39] K. Yan, J. Su, J. Huang, and Y. Mo, "Chiller fault diagnosis based on VAE-enabled generative adversarial networks," *IEEE Transactions on Automation Science and Engineering*, vol. 19, no. 1, pp. 387–395, 2022.

[40] W. Chen, L. Cheng, Z. Chang, B. Wen, and P. Li, "Wind turbine blade icing detection using a novel bidirectional gated recurrent unit with temporal pattern attention and improved coot optimization algorithm," *Meas Sci Technol*, vol. 34, no. 1, p. 014004, 2022.

[41] M. Xia, H. Shao, D. Williams, S. Lu, L. Shu, and C. W. de Silva, "Intelligent fault diagnosis of machinery using digital twin-assisted deep transfer learning," *Reliab Eng Syst Saf*, vol. 215, p. 107938, 2021.

[42] L. Chen, Q. Li, C. Shen, J. Zhu, D. Wang, and M. Xia, "Adversarial domain-invariant generalization: a generic domain-regressive framework for bearing fault diagnosis under unseen conditions," *IEEE Trans Industr Inform*, vol. 18, no. 3, pp. 1790–1800, 2022.

[43] Y. Xiao, H. Wang, L. Zhang, and W. Xu, "Two methods of selecting Gaussian kernel parameters for one-class SVM and their application to fault detection," *Knowl Based Syst*, vol. 59, pp. 75–84, 2014.

[44] Y. Cheng, H. Zhu, J. Wu, and X. Shao, "Machine health monitoring using adaptive kernel spectral clustering and deep long short-term memory recurrent neural networks," *IEEE Trans Industr Inform*, vol. 15, no. 2, pp. 987–997, 2019.

[45] Y. Ganin and V. Lempitsky, "Unsupervised domain adaptation by backpropagation," *International Conference on Machine Learning, PMLR*, vol. 37, pp. 1180–1189, 2015.

[46] Z. Chen, G. He, J. Li, Y. Liao, K. Gryllias, and W. Li, "Domain adversarial transfer network for cross-domain fault diagnosis of rotary machinery," *IEEE Trans Instrum Meas*, vol. 69, no. 11, pp. 8702–8712, 2020.

[47] 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 Industr Inform*, vol. 17, no. 12, pp. 7957–7967, 2021.

[48] J. Li, Y. Wang, Y. Zi, H. Zhang, and Z. Wan, "Causal disentanglement: a generalized bearing fault diagnostic framework in continuous degradation mode," *IEEE Trans Neural Netw Learn Syst*, vol. 34, no. 9, pp. 6250–6262, 2023.

[49] X. Zhang, L. He, X. Wang, J. Wang, and P. Cheng, "Transfer fault diagnosis based on local maximum mean difference and K-means," *Comput Ind Eng*, vol. 172, p. 108568, 2022.

[50] M. Arjovsky, L. Bottou, I. Gulrajani, and D. Lopez-Paz, "Invariant risk minimization," arXiv preprint arXiv:1907.02893, 2019.

[51] Z. Mo, Z. Zhang, Q. Miao, and K.-L. Tsui, "Sparsity-constrained invariant risk minimization for domain generalization with application to machinery fault diagnosis modeling," *IEEE Trans Cybern*, vol. 54, no. 3, pp. 1547–1559, 2024.

[52] D. Krueger et al., "Out-of-distribution generalization via risk extrapolation (rex)," *International Conference on Machine Learning, PMLR*, vol. 139, pp. 5815–5826, 2021.

[53] C. Zhao and W. Shen, "Adversarial mutual information-guided single domain generalization network for intelligent fault diagnosis," *IEEE Trans Industr Inform*, vol. 19, no. 3, pp. 2909–2918, 2023.

[54] C. Finn, P. Abbeel, and S. Levine, "Model-agnostic meta-learning for fast adaptation of deep networks," *International conference on machine learning, PMLR*, vol. 70, pp. 1126–1135, 2017.

[55] H. Su, L. Xiang, A. Hu, Y. Xu, and X. Yang, "A novel method based on meta-learning for bearing fault diagnosis with small sample learning under different working conditions," *Mech Syst Signal Process*, vol. 169, p. 108765, 2022.

[56] S. Zhang, F. Ye, B. Wang, and T. G. Habetler, "Few-shot bearing fault diagnosis based on model-agnostic meta-learning," *IEEE Trans Ind Appl*, vol. 57, no. 5, pp. 4754–4764, 2021.

[57] J. Yang, X. Wang, and Z. Luo, "Few-shot remaining useful life prediction based on meta-learning with deep sparse kernel network," *Inf Sci (N Y)*, vol. 653, p. 119795, 2024.

[58] Y. Ding, J. Zhuang, P. Ding, and M. Jia, "Self-supervised pretraining via contrast learning for intelligent incipient fault detection of bearings," *Reliab Eng Syst Saf*, vol. 218, p. 108126, 2022.

[59] D. Kong, W. Huang, L. Zhao, J. Ding, H. Wu, and G. Yang, "Mining knowledge from unlabeled data for fault diagnosis:

a multi-task self-supervised approach," *Mech Syst Signal Process*, vol. 211, p. 111189, 2024.
[60] C. Hu, J. Wu, C. Sun, R. Yan, and X. Chen, "Inter-instance and intra-temporal self-supervised learning with few labeled data for fault diagnosis," *IEEE Trans Industr Inform*, vol. 19, no. 5, pp. 6502–6512, 2022.
[61] W. Mao, J. Chen, J. Liu, and X. Liang, "Self-supervised deep domain-adversarial regression adaptation for online remaining useful life prediction of rolling bearing under unknown working condition," *IEEE Trans Industr Inform*, vol. 19, no. 2, pp. 1227–1237, 2023.
[62] J. Wen, H. Yao, Z. Ji, B. Wu, and M. Xia, "On fault diagnosis for high-g accelerometers via data-driven models," *IEEE Sens J*, vol. 21, no. 2, pp. 1359–1368, 2021.
[63] W. Yu and C. Zhao, "Online fault diagnosis for industrial processes with bayesian network-based probabilistic ensemble learning strategy," *IEEE Transactions on Automation Science and Engineering*, vol. 16, no. 4, pp. 1922–1932, 2019.
[64] Y. Li, Y. Song, L. Jia, S. Gao, Q. Li, and M. Qiu, "Intelligent fault diagnosis by fusing domain adversarial training and maximum mean discrepancy via ensemble learning," *IEEE Trans Industr Inform*, vol. 17, no. 4, pp. 2833–2841, 2021.
[65] Z. Li, D. Wu, C. Hu, and J. Terpenny, "An ensemble learning-based prognostic approach with degradation-dependent weights for remaining useful life prediction," *Reliab Eng Syst Saf*, vol. 184, pp. 110–122, 2019.
[66] Y. H. Che, X. S. Hu, X. K. Lin, J. Guo, and R. Teodorescu, "Health prognostics for lithium-ion batteries: mechanisms, methods, and prospects," *Energy & Environmental Science*, vol. 16, no. 2, pp. 338–371, 2023.
[67] T. B. Xia, Y. F. Dong, L. Xiao, S. C. Du, E. S. Pan, and L. F. Xi, "Recent advances in prognostics and health management for advanced manufacturing paradigms," *Reliability Engineering & System Safety*, vol. 178, pp. 255–268, 2018.
[68] S. L. Lu, J. F. Lu, K. An, X. X. Wang, and Q. B. He, "Edge computing on IoT for machine signal processing and fault diagnosis: a review," *IEEE Internet of Things Journal*, vol. 10, no. 13, pp. 11093–11116, 2023.
[69] X. F. Yin, Y. M. Zhu, and J. K. Hu, "A comprehensive survey of privacy-preserving federated learning: a taxonomy, review, and future directions," *ACM Computing Surveys*, vol. 54, no. 6, p. 131, 2021.
[70] L. Li, Y. X. Fan, M. Tse, and K. Y. Lin, "A review of applications in federated learning," *Computers & Industrial Engineering*, vol. 149, p. 106854, 2020.
[71] S. Banabilah, M. Aloqaily, E. Alsayed, N. Malik, and Y. Jararweh, "Federated learning review: fundamentals, enabling technologies, and future applications," *Information Processing & Management*, vol. 59, no. 6, p. 103061, 2022.
[72] Q. Yang, Y. Liu, T. J. Chen, and Y. X. Tong, "Federated machine learning: concept and applications," *ACM Transactions on Intelligent Systems and Technology*, vol. 10, no. 2, p. 12, 2019.
[73] S. W. Yu and J. W. Ma, "Deep learning for geophysics: current and future trends," *Reviews of Geophysics*, vol. 59, no. 3, p. e2021RG000742, 2021.
[74] G. A. Kaissis, M. R. Makowski, D. Rückert, and R. F. Braren, "Secure, privacy-preserving and federated machine learning in medical imaging," *Nature Machine Intelligence*, vol. 2, no. 6, pp. 305–311, 2020.
[75] H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. Agüera y Arcas, "Communication-Efficient Learning of Deep Networks from Decentralized Data," p. arXiv:1602.05629, 2016. doi: 10.48550/arXiv.1602.05629.
[76] J. Konečný, H. Brendan McMahan, F. X. Yu, P. Richtárik, A. Theertha Suresh, and D. Bacon, "Federated Learning: Strategies for Improving Communication Efficiency," p. arXiv:1610.05492, 2016. doi: 10.48550/arXiv.1610.05492.
[77] J. Konečný, H. Brendan McMahan, D. Ramage, and P. Richtárik, "Federated Optimization: Distributed Machine Learning for On-Device Intelligence," p. arXiv: 1610.02527, 2016. doi: 10.48550/arXiv.1610.02527.
[78] T. Berghout, M. Benbouzid, T. Bentrcia, W. H. Lim, and Y. Amirat, "Federated learning for condition monitoring of industrial processes: a review on fault diagnosis methods, challenges, and prospects," *Electronics*, vol. 12, no. 1, p. 158, 2023.
[79] W. Zhang, X. Li, H. Ma, Z. Luo, and X. Li, "Federated learning for machinery fault diagnosis with dynamic validation and self-supervision," *Knowledge-Based Systems*, vol. 213, p. 106679, 2021.
[80] J. Du, N. Qin, D. Huang, X. Jia, and Y. Zhang, "Lightweight FL: a low-cost federated learning framework for mechanical fault diagnosis with training optimization and model pruning," *IEEE Transactions on Instrumentation and Measurement*, vol. 73, p. 3504014, 2024.
[81] T. Hiessl, S. Rezapour Lakani, J. Kemnitz, D. Schall, and S. Schulte, "Cohort-based federated learning services for industrial collaboration on the edge," *Journal of Parallel and Distributed Computing*, vol. 167, pp. 64–76, 2022.
[82] M. Chahoud, H. Sami, A. Mourad, S. Otoum, H. Otrok, J. Bentahar, and M. Guizani, "On-demand-FL: a dynamic and efficient multicriteria federated learning client deployment scheme," *IEEE Internet of Things Journal*, vol. 10, no. 18, pp. 15822–15834, 2023.
[83] Z. Zhang, C. Guan, H. Chen, X. Yang, W. Gong, and A. Yang, "Adaptive privacy-preserving federated learning for fault diagnosis in internet of ships," *IEEE Internet of Things Journal*, vol. 9, no. 9, pp. 6844–6854, 2022.
[84] J. Lin, J. Ma, and J. Zhu, "Hierarchical federated learning for power transformer fault diagnosis," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, p. 3520611, 2022.
[85] G. Qu, N. Cui, H. Wu, R. Li, and Y. Ding, "ChainFL: a simulation platform for joint federated learning and blockchain in edge/cloud computing environments," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 5, pp. 3572–3581, 2022.
[86] J. Chen, J. Tang, and W. Li, "Industrial edge intelligence: federated-meta learning framework for few-shot fault diagnosis," *IEEE Transactions on Network Science and Engineering*, vol. 10, no. 6, pp. 3561–3573, 2023.
[87] W. Sun, R. Yan, R. Jin, R. Zhao, and Z. Chen, "FedAlign: federated model alignment via data-free knowledge distillation for machine fault diagnosis," *IEEE Transactions on Instrumentation and Measurement*, vol. 73, p. 3506112, 2024.
[88] W. Zhang and X. Li, "Data privacy preserving federated transfer learning in machinery fault diagnostics using prior distributions," *Structural Health Monitoring-an International Journal*, vol. 21, no. 4, pp. 1329–1344, 2022.
[89] J. B. Chen, J. P. Li, R. Y. Huang, K. Yue, Z. Y. Chen, and W. H. Li, "Federated transfer learning for bearing fault diagnosis with discrepancy-based weighted federated averaging," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, p. 3514911, 2022.

[90] X. T. Jin, J. J. Bu, Z. Yu, H. Zhang, and Y. A. Wang, "FedCrack: federated transfer learning with unsupervised representation for crack detection," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 10, pp. 11171–11184, 2023.
[91] J. Shen, S. Yang, C. Zhao, X. Ren, P. Zhao, Y. Yang, Q. Han, and S. Wu, "FedLED: Label-free equipment fault diagnosis with vertical federated transfer learning," *IEEE Transactions on Instrumentation and Measurement*, vol. 73, p. 3509910, 2024.
[92] W. Zhang and X. Li, "Federated transfer learning for intelligent fault diagnostics using deep adversarial networks with data privacy," *IEEE/ASME Transactions on Mechatronics*, vol. 27, no. 1, pp. 430–439, 2022.
[93] H. G. Abreha, M. Hayajneh, and M. A. Serhani, "Federated learning in edge computing: a systematic survey," *Sensors*, vol. 22, no. 2, p. 450, 2022.
[94] W. S. Shi, J. Cao, Q. Zhang, Y. H. Z. Li, and L. Y. Xu, "Edge computing: vision and challenges," *IEEE Internet of Things Journal*, vol. 3, no. 5, pp. 637–646, 2016.
[95] G. M. Bao and P. Guo, "Federated learning in cloud-edge collaborative architecture: key technologies, applications and challenges," *Journal of Cloud Computing-Advances Systems and Applications*, vol. 11, no. 1, p. 94, 2022.
[96] G. Li, J. Cai, C. He, X. Zhang, and H. Chen, "Online incentive mechanism designs for asynchronous federated learning in edge computing," *IEEE Internet of Things Journal*, vol. 11, no. 5, pp. 7787–7804, 2024.
[97] Y. Yu, L. Guo, H. Gao, Y. He, Z. You, and A. Duan, "FedCAE: a new federated learning framework for edge-cloud collaboration based machine fault diagnosis," *IEEE Transactions on Industrial Electronics*, vol. 71, no. 4, pp. 4108–4119, 2024.
[98] L. Guo, Y. Yu, M. Qian, R. Zhang, H. Gao, and Z. Cheng, "FedRUL: a new federated learning method for edge-cloud collaboration based remaining useful life prediction of machines," *IEEE/ASME Transactions on Mechatronics*, vol. 28, no. 1, pp. 350–359, 2023.
[99] B. Li and C. Zhao, "Federated zero-shot industrial fault diagnosis with cloud-shared semantic knowledge base," *IEEE Internet of Things Journal*, vol. 10, no. 13, pp. 11619–11630, 2023.
[100] Y. Li, J. Z. Li, and Y. Wang, "Privacy-preserving spatiotemporal scenario generation of renewable energies: a federated deep generative learning approach," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 4, pp. 2310–2320, 2022.
[101] C. Y. She, C. J. Sun, Z. Y. Gu, Y. H. Li, C. Y. Yang, H. V. Poor, and B. Vucetic, "A tutorial on ultrareliable and low-latency communications in 6G: integrating domain knowledge into deep learning," *Proceedings of the IEEE*, vol. 109, no. 3, pp. 204–246, 2021.
[102] Z. Q. Geng, C. J. Shi, and Y. M. Han, "Intelligent small sample defect detection of water walls in power plants using novel deep learning integrating deep convolutional GAN," *IEEE Transactions on Industrial Informatics*, vol. 19, no. 6, pp. 7489–7497, 2023.
[103] Y. Liu, Y. Kang, C. P. Xing, T. J. Chen, and Q. Yang, "A secure federated transfer learning framework," *IEEE Intelligent Systems*, vol. 35, no. 4, pp. 70–82, 2020.
[104] W. Zhang and X. Li, "Federated transfer learning for intelligent fault diagnostics using deep adversarial networks with data privacy," *IEEE-ASME Transactions on Mechatronics*, vol. 27, no. 1, pp. 430–439, 2022.
[105] Y. Fu, H. R. Cao, X. F. Chen, and J. M. Ding, "Broad autoencoder for machinery intelligent fault diagnosis with incremental fault samples and fault modes," *Mechanical Systems and Signal Processing*, vol. 178, p. 109353, 2022.
[106] H. Zhou, H. P. Yin, D. D. Zhao, and L. Cai, "Incremental learning and conditional drift adaptation for nonstationary industrial process fault diagnosis," *IEEE Transactions on Industrial Informatics*, vol. 19, no. 4, pp. 5935–5944, 2023.
[107] F. Farbiz, M. S. Habibullah, B. Hamadicharef, T. Maszczyk, and S. Aggarwal, "Knowledge-embedded machine learning and its applications in smart manufacturing," *J. Intell. Manuf.*, vol. 34, no. 7, pp. 2889–2906, 2023.
[108] T. Kim and J. Chai, "Pre-processing method to improve cross-domain fault diagnosis for bearing," *Sensors*, vol. 21, no. 15, p. 4970, 2021.
[109] X. Yu et al., "A wavelet packet transform-based deep feature transfer learning method for bearing fault diagnosis under different working conditions," *Measurement*, vol. 201, p. 111597, 2022.
[110] I. Kim et al., "Single domain generalizable and physically interpretable bearing fault diagnosis for unseen working conditions," *Expert Syst. Appl.*, vol. 241, p. 122455, 2024.
[111] Y. Li, Y. Ren, H. Zheng, Z. Deng, and S. Wang, "A novel cross-domain intelligent fault diagnosis method based on entropy features and transfer learning," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–14, 2021.
[112] C. Liu, C. Qin, X. Shi, Z. Wang, G. Zhang, and Y. Han, "TScatNet: an interpretable cross-domain intelligent diagnosis model with antinoise and few-shot learning capability," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–10, 2021.
[113] C. Liu, X. Ma, T. Han, X. Shi, C. Qin, and S. Hu, "NTScatNet: an interpretable convolutional neural network for domain generalization diagnosis across different transmission paths," *Measurement*, vol. 204, p. 112041. 2022.
[114] C. He, H. Shi, X. Liu, and J. Li, "Interpretable physics-informed domain adaptation paradigm for cross-machine transfer diagnosis," *Knowledge-Based Syst.*, vol. 288, p. 111499. 2024.
[115] T. Yin, N. Lu, G. Guo, Y. Lei, S. Wang, and X. Guan, "Knowledge and data dual-driven transfer network for industrial robot fault diagnosis," *Mech. Syst. Signal Proc.*, vol. 182, p. 109597, 2023.
[116] C. Liu, A. Ricardo Mauricio, J. Qi, D. Peng, and K. Gryllias, "Domain adaptation digital twin for rolling element bearing prognostics," *Annual Conference of the PHM Society*, vol. 12, pp. 1–10, 2020.
[117] Y. Zhang, K. Feng, J. C. Ji, K. Yu, Z. Ren, and Z. Liu, "Dynamic model-assisted bearing remaining useful life prediction using the cross-domain transformer network," *IEEE/ASME Trans. on Mech.*, vol. 28, no. 2, pp. 1070–1080, 2023.
[118] D. Zhou, J. Hao, D. Huang, X. Jia, and H. Zhang, "Dynamic simulation of gas turbines via feature similarity-based transfer learning," *Front. Energy*, vol. 14, no. 4, pp. 817–835, 2020.
[119] W. Li, S. Gu, X. Zhang, and T. Chen, "Transfer learning for process fault diagnosis: knowledge transfer from simulation to physical processes," *Comput. Chem. Eng.*, vol. 139, p. 106904, 2020.
[120] J. Wang, Z. Zhang, Z. Liu, B. Han, H. Bao, and S. Ji, "Digital twin aided adversarial transfer learning method for domain adaptation fault diagnosis," *Reliab. Eng. Syst. Saf.*, vol. 234, p. 109152, 2023.
[121] M. Gong, D. Tang, J. Yu, and L. Tian, "A physics-informed transfer learning approach for anomaly detection of

aerospace CMG with limited telemetry data," *Proc. Global Rel. Prognostics Health Manage. (PHM-Nanjing)*, pp. 1–5, 2021.

[122] Borate P, Rivière J, Marone C, et al, "Using a physics-informed neural network and fault zone acoustic monitoring to predict lab earthquakes," *Nat. Commun.*, vol. 14, no.1 pp. 3693, 2023.

[123] Z. Liu, Z. Ding, X. Huang, and P. Zhang, "An online power system transient stability assessment method based on graph neural network and central moment discrepancy," *Front. Energy Res.*, vol. 11, p. 1082534, 2023.

[124] K. Zhou, E. Diehl, and J. Tang, "Deep convolutional generative adversarial network with semi-supervised learning enabled physics elucidation for extended gear fault diagnosis under data limitations," *Mech. Syst. Signal Proc.*, vol. 185, p. 109772, 2023.

[125] W. Mao, J. Liu, J. Chen, and X. Liang, "An interpretable deep transfer learning-based remaining useful life prediction approach for bearings with selective degradation knowledge fusion," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–16, 2022.

[126] W. Mao, W. Zhang, K. Feng, M. Beer, and C. Yang, "Tensor representation-based transferability analytics and selective transfer learning of prognostic knowledge for remaining useful life prediction across machines," *Reliab. Eng. Syst. Saf.*, vol. 242, p. 109695, 2024.

[127] R. Zhu, W. Peng, D. Wang, and C. Huang, "Bayesian transfer learning with active querying for intelligent cross-machine fault prognosis under limited data," *Mech. Syst. Signal Proc.*, vol. 183, p. 109628, 2023.

[128] Z. Xing, C. Yi, J. Lin, and Q. Zhou, "A novel periodic cyclic sparse network with entire fomain adaptation for deep transfer fault diagnosis of rolling bearing," *IEEE Sens. J.*, vol. 23, no. 12, pp. 13452–13468, 2023.

[129] Y. Li, Z. Zhou, C. Sun, X. Chen, and R. Yan, "Variational attention-based interpretable transformer network for rotary machine fault diagnosis," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 5, pp. 6180–6193, 2022.

[130] C. Guo, Z. Zhao, J. Ren, S. Wang, Y. Liu, and X. Chen, "Causal explaining guided domain generalization for rotating machinery intelligent fault diagnosis," *Expert Syst. Appl.*, vol. 243, p. 122806, 2024.