

Received 31 December 2022, accepted 23 January 2023, date of publication 25 January 2023, date of current version 10 February 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3239784

## TOPICAL REVIEW

# A Systematic Literature Review on Transfer Learning for Predictive Maintenance in Industry 4.0

MEHDI SAMAN AZARI<sup>1</sup>, FRANCESCO FLAMMINI<sup>1,2</sup>, (Senior Member, IEEE),  
STEFANIA SANTINI<sup>3</sup>, (Member, IEEE), AND MAURO CAPORUSCIO<sup>1</sup>

<sup>1</sup>Department of Computer Science and Media Technology, Linnaeus University, 351 95 Växjö, Sweden

<sup>2</sup>School of Innovation, Design, and Engineering, Mälardalens University, 632 20 Eskilstuna, Sweden

<sup>3</sup>Department of Electrical Engineering and Information Technology, University of Naples Federico II, 80125 Naples, Italy

Corresponding author: Francesco Flammini (francesco.flammini@mdu.se)

**ABSTRACT** The advent of Industry 4.0 has resulted in the widespread usage of novel paradigms and digital technologies within industrial production and manufacturing systems. The objective of making industrial operations monitoring easier also implied the usage of more effective data-driven predictive maintenance approaches, including those based on machine learning. Although those approaches are becoming increasingly popular, most of the traditional machine learning and deep learning algorithms experience the following three major challenges: 1) lack of training data (especially faulty data), 2) incompatible computation power, and 3) discrepancy in data distribution. A new data-driven technique, such as transfer learning, can be developed to overcome the issues related to traditional machine learning and deep learning for predictive maintenance. Motivated by the recent big interest towards transfer learning within computer science and artificial intelligence, in this paper we provide a systematic literature review addressing related research with a focus on predictive maintenance. The review aims to define transfer learning in the context of predictive maintenance by introducing a specific taxonomy based on relevant perspectives. We also discuss current advances, challenges, open-source datasets, and future directions of transfer learning applications in predictive maintenance from both theoretical and practical viewpoints.

**INDEX TERMS** Transfer learning, domain adaptation, fault diagnosis, fault detection, fault prognosis, predictive maintenance.

## I. INTRODUCTION

The Industry 4.0 era saw the introduction of new paradigms and technologies based on connectivity, data analytics, and novel devices, allowing for inventory reduction, customization, and controlled production [1]. In many manufacturing industries, profitability and effectiveness are reliant on developing high-quality products based on reliable systems. In fact, any unanticipated downtime of machinery or deterioration of equipment can lead to significant penalties and severe reputational losses for companies. Therefore, maintenance represents a critical activity with significant implications for the companies' capacity to compete on cost, quality, and

performance. The transition of maintenance strategies from reactive and preventive to predictive can be a concrete result within Industry 4.0. In this context, special attention must be paid to Artificial Intelligence (AI) approaches, due to their capability to manage high-dimensional and multivariate data as well as extract hidden correlations within data; as such, AI is especially suitable for enhancing the performance of Predictive Maintenance (PdM) [2].

Within the AI field, Machine Learning (ML) and, more recently, Deep Learning (DL), have emerged as effective techniques for developing PdM models due to their capability of performing failure prediction tasks such as estimating the remaining useful life of a machine [2]. However, despite their benefits, traditional ML and DL techniques suffer several limitations. To begin with, these methods are usually based

The associate editor coordinating the review of this manuscript and approving it for publication was Huiyan Zhang<sup>1</sup>.

on the assumption that both the training and testing datasets are drawn from the same distribution. When dealing with real-world applications, this assumption may not necessarily apply. Second, traditional ML and DL algorithms require a significant amount of historical fault data to learn the fault characteristics of machines, whereas in real-world applications it is not always feasible to run the equipment in failure mode for both safety and economic reasons. Finally, the cost of model training in terms of time is high. It takes a lot of time to adjust the weights and parameters of DL algorithms when training from scratch for a new operating condition.

To tackle the mentioned problems, Transfer Learning (TL) has recently emerged as a powerful AI technique [3]. More specifically, TL has proven to support target domain adaptation by leveraging knowledge acquired from the source domain, in order to cope with the lack of training data, especially faulty data, in the target domain. Furthermore, unlike traditional ML and DL methods, TL can be successfully applied even if the data distribution of the source and target domains differ. In addition, the amount of computing power and computation time required to train a model can be drastically decreased by leveraging pre-learned knowledge from various source domains [4]. Since TL methods can be effectively exploited for overcoming issues related to traditional ML and DL approaches, in this work we bridge the gap in the technical literature by providing the reference taxonomy, the state-of-the-art (SOTA), and the challenges arising from applying TL to PdM. In order to achieve those objectives, we leverage a method known as Systematic Literature Review (SLR). The main objectives and contributions of the paper can be summarized as follows:

- To clearly motivate the need for TL on the base of a structured PdM categorization.
- To provide a novel TL taxonomy in PdM from three perspectives: problem, solution, and application.
- To provide a SOTA with up-to-date references focused on the use of TL techniques for PdM tasks.
- To highlight the key challenges that must be tackled to successfully apply TL to PdM.
- To discuss recent advancements, the availability of open-source data sets, and the possible future directions of TL in PdM, from both scientific and industrial perspectives.

The remainder of the paper is organized as follows: Section II sets this study in the context of related works. The reference context, PdM categorization, and transfer learning definition are discussed in Section III. The motivation for the application of TL in the PdM context is discussed in Section IV. Section V described the systematic literature review approach, including the definition of research questions, the search process, and the selection and filtering of relevant publications. Section VI provides the responses to the research questions set in Section V, including the specific TL taxonomy and a discussion of methodological and industrial challenges. Finally, Section VII draws conclusions.

## II. RELATED WORK

Several survey papers exist addressing different aspects of fault detection, diagnosis, and prognosis within PdM. For example, Carlvalho et al. [2], conducted a systematic literature review of traditional ML methods used for PdM and discussed the efficiency of the current state-of-the-art traditional ML techniques. Furthermore, Jovani et al. [5] investigated the benefits that traditional ML algorithms may offer in PdM, and have also conducted an SLR to identify the implementation obstacles. Zhao et al. [6] conducted a systematic review of machine health monitoring systems based on deep learning algorithms, including Deep Belief Networks (DBN), Deep Auto-Encoders (DAE), Convolutional Neural Networks (CNN), and Recurrent Neural Networks (RNN). Olga et al. [7] presented a comprehensive assessment of current advancements, trends, challenges, and future direction of DL in the context of PHM. Lei et al. [8] introduced a study and a roadmap to comprehensively address the evolution of Intelligent Fault Diagnosis (IFD) as a result of developments in ML, as well as future perspectives. Zonta et al. [1] discussed current challenges and limitations in PdM and suggested a new taxonomy to categorize this study field in light of Industry 4.0 requirements. Juan et al. [9] conducted an SLR to summarize the current diagnostic and prognostic trends, as well as outline current challenges and research opportunities, with a specific focus on multi-model approaches. Keleko et al. conducted a bibliometric study to investigate and quantify the most important concepts, areas of application, methodologies, and significant trends of AI used in real-time predictive maintenance in Industry 4.0 [10].

Other survey papers are devoted to analyzing a specific component or equipment for fault diagnosis and prognosis, e.g., bearing [11], milling cutting tools [12] and rotating machinery [13]. Among them, Liu et al. [13] reviewed the main AI methods for fault diagnosis of rotating machines in industrial applications. In addition, a number of survey articles on fault prognosis are also available. For instance, Lei et al. [14] provided an evaluation of machinery prognostics based on four prognostic processes, including data acquisition, Health Indicator (HI) construction, Health Stage (HS) division, and remaining useful life (RUL) prediction. Although all those papers provide interesting reviews within fault diagnosis and prognosis, they focus on traditional ML and DL techniques, while none of them addresses TL as we do in this paper.

Only a few survey papers address the application of TL to maintenance (see, e.g., [15], [16]); however, they only focus on specific tasks for diagnosing machinery faults, while this work is more comprehensive. In addition, existing papers reviewed works published till 2019, while we found more than 120 related publications released after 2019. Indeed, only very recently TL has demonstrated its power within PdM and hence showed a rapid expansion. Also, in this study, we provided a novel TL taxonomy in PdM from three perspectives: problem, solution, and application.

### III. PRELIMINARY

#### A. BACKGROUND ON PREDICTIVE MAINTENANCE

The profit and competitiveness of the industrial sectors rely on designing and producing high-quality products and reliable systems. However, developing such sophisticated systems brings new challenges, including maintenance expenses. The main concern of the industry is reducing maintenance costs and minimizing business risks while improving asset reliability and safety. Three labels commonly used in the technical literature to indicate the main maintenance paradigms and related actions [17]:

- **Reactive Maintenance (RM).** This approach is also known as corrective maintenance or run-to-failure (R2F) since the repair is conducted following the occurrence of a fault. In this paradigm, a machinery component operates from installation until its failure occurs, at which point the whole machine is shut down for maintenance. However, because of the additional maintenance costs (due to the impact of the failure on others machinery parts) and unanticipated downtime, the cost and efficiency of this methodology are often unacceptable for industry sectors.
- **Preventive Maintenance (PM).** To prevent process or equipment breakdowns, this kind of maintenance is performed on a predetermined schedule in time or process iterations, so it is often called time-based or planned maintenance. Despite its effectiveness, in some cases, the PM may increase operational expenses due to the possibility of over-maintenance.
- **Predictive Maintenance (PdM).** PdM makes use of predictive technologies to determine when maintenance is required. It is built on the continuous monitoring of a machine's or process's integrity, thus allowing maintenance to be done only when required.

PdM stands out among the others in the industry 4.0 era owing to its capacity to optimize asset use and management [17]. PdM's main **purposes** are to avoid unexpected downtime, increase overall system reliability, and lower operational costs. The three main **principles** of predictive maintenance are fault detection, diagnosis, and prognosis. In other words, predictive maintenance involves detecting a developing defect (fault detection), in the case of faulty systems, isolating and identifying the specific type of fault (fault diagnostics), and forecasting the RUL or end life of the system (prognostics). PdM can be **implemented** online or offline. In an online implementation, information is collected, preprocessed, and analyzed in a real-time manner to raise alarm signals or adjust actions while the system is running. Offline implementation collects all operational data so that the maintenance team can analyze it afterward. Moreover, the **interaction** of predictive maintenance with manufacturing is based on the cyber system and service innovation. As a result, there is interaction with products and industrial processes of the system. It should be highlighted that Prognostics and Health Management (PHM) is an example of the possible predictive maintenance **extensions** often found in the literature.

Indeed, this term is frequently used even to replace predictive maintenance, and there is no consistency in the use of these keywords or their interaction in the maintenance area. In an attempt to clarify the keywords, Jimenez et al. [9], clearly define PHM as an extension of PdM aiming to increase asset predictability and life cycle management. According to [9], the predictive maintenance **approaches**, could be classified into two major groups: single-model approaches and multi-model approaches. Single-model methods can be further categorized as knowledge-based methods, data-driven methods, and physics-based methods, while the various multi-model approaches, sometimes referred to as hybrid models, arise from the different possible single-model method combinations. Data-driven machine health monitoring methods have grown increasingly appealing as sensors, and computer systems have advanced significantly. In this area, traditional ML, DL, and TL have emerged as effective techniques for developing intelligent prediction systems, especially in complex or large-scale and dynamic contexts. However, the success of these applications is contingent on the proper choice of the most suitable machine-learning technique according to the specific problem [5]. Finally, summarizing the main characteristics and features, the predictive maintenance overview, is depicted in Fig. 1 by using class diagrams.

#### B. TRANSFER LEARNING: MAIN CONCEPTS AND NOMENCLATURE

Transfer Learning (TL) aims to develop an effective model for a target domain with limited training data by leveraging and exploiting knowledge from different but related source domains. For the sake of completeness, the definitions commonly used in this field, and related notation, are introduced in the following.

*Definition 1 (Domain [18]):* A domain  $D = \{\mathcal{X}, P(X)\}$  includes two parts, namely a feature space, say  $\mathcal{X}$ , and a marginal distribution, say  $P(X)$ , where the symbol  $X$  specifies an instance set as  $X = \{x | x_i \in \mathcal{X}, i = 1, \dots, N\}$  that contains  $N$  instances.

*Definition 2 (Task [18]):* Given a domain  $D = \{\mathcal{X}, P(X)\}$ , a task  $\mathcal{T} = \{\mathcal{Y}, f\}$  consists of a label space  $\mathcal{Y}$  and a prediction function  $f(\cdot)$ , being the sample data composed of pairs  $\{x_i, y_i\}$  where  $x_i \in \mathcal{X}$  and  $y_i \in \mathcal{Y}$ . The prediction function  $f$  is designed to learn from sample data and predict the label of the future instance; it can be expressed as the following conditional distribution of instances:  $f(x_j) = P(Y|X) = \{P(y_k|x_j) | y_k \in \mathcal{Y}, k = 1, \dots, |\mathcal{Y}|\}$ .

The above definitions can be specified for source and target domains. Accordingly, let  $D_S = \{\mathcal{X}_S, P_S(X_S)\}$  and  $D_T = \{\mathcal{X}_T, P_T(X_T)\}$  be the source and the target domain, where  $\mathcal{X}_S$  and  $\mathcal{X}_T$  represent the source and target domains' feature spaces, respectively, while  $P_S(X_S)$  and  $P_T(X_T)$  represent the source and target domains' marginal probability distributions, respectively.

Correspondingly, a source task and a target task can be defined as  $\mathcal{T}_S = \{\mathcal{Y}_S, P_S(Y_S|X_S)\}$  and  $\mathcal{T}_T = \{\mathcal{Y}_T,$

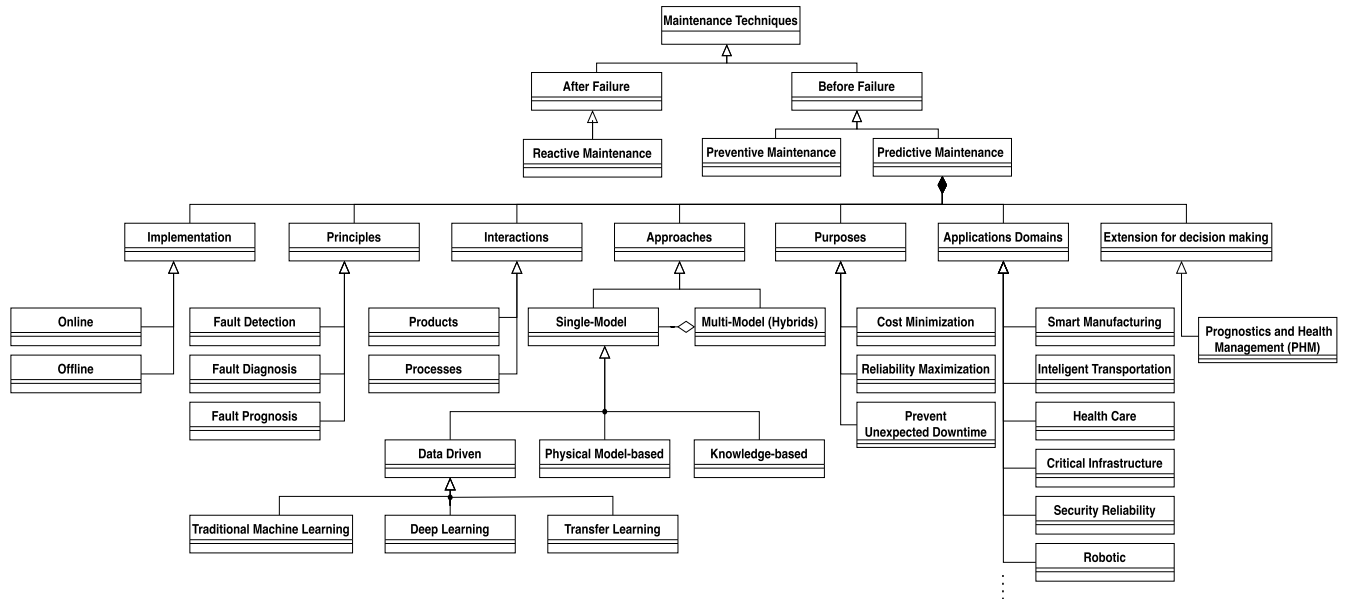


FIGURE 1. Overview of predictive maintenance using class diagrams.

$P_T(Y_T|X_T)$ , being  $\mathcal{Y}_S$  and  $\mathcal{Y}_T$  the source and target task label spaces, respectively,  $P_S(Y_S|X_S)$  and  $P_T(Y_T|X_T)$  the source and target domains' conditional probability distributions, respectively.

Given the above concepts, a unified definition of TL can be provided as follows.

**Definition 3 (Transfer Learning [18]):** Consider the source domain  $D_S = \{\mathcal{X}_S, P_S(X_S)\}$ , with relating source task  $T_S = \{\mathcal{Y}_S, P_S(Y_S|X_S)\}$ , and target domain  $D_T = \{\mathcal{X}_T, P_T(X_T)\}$ , with respective target task  $T_T = \{\mathcal{Y}_T, P_T(Y_T|X_T)\}$ . Transfer Learning tries to exploit the related knowledge embedded in  $D_S$  and  $T_S$  to improve the performance of the target prediction function  $f_T(\cdot)$  in the target domain  $D_T$  with target task  $T_T$ , being  $D_S \neq D_T$  or  $T_S \neq T_T$ . A crucial concept related to TL is the idea of domain adaptation, which has been developed to deal with cross-domain learning challenges under specific conditions. The aim is to adapt one or more source domains to exploit knowledge and improve the target learner's performance.

**Definition 4 (Domain Adaptation [16]):** Domain adaptation is a type of transfer learning problem in which the source and target tasks are assumed to be the same, i.e.  $T_S = T_T$ .

The definitions provided so far refer to single-source TL, which is the most common type of TL in current research. Since recent studies demonstrate that using several source domains and tasks improves the prediction function  $f_T$ , the following multi-source TL definition must be given.

**Definition 5 (Multiple-source Transfer Learning [19]):** Given some observations corresponding to  $m^S \in \mathbb{N}^+$  source domains and tasks (i.e.,  $\{(D_{S_i}, T_{S_i}) | i = 1, \dots, m^S\}$ ), and observation about target domain and task  $(D_T, T_T)$ , transfer

learning enhances the performance of the learned prediction function  $f_T$  in the target domain by leveraging the knowledge implied in the multiple source domains.

#### IV. MOTIVATIONS FOR APPLYING TL TO PdM

In recent years, many traditional ML and DL methods have made significant progress as a result of their superior representation learning and pattern recognition capabilities, and they have also been successfully used for the predictive maintenance of industrial systems [7], [20], [21]. However, intelligent PdM systems based on traditional ML and DL have to satisfy specific conditions for achieving outstanding performance. First, the feature extraction process of traditional ML methods can be inefficient and passive. As a result, manual feature extractor design not only necessitates the expertise of signal processing experts, but it is also a time-consuming process, as each handcrafted feature is only acceptable for specific working conditions, necessitating the operator's judgment in many situations [8]. Second, the network of DL and ML algorithms has to be trained with enough labeled data to discover representative features and be able to generalize adequately. Therefore, training such networks requires a sufficient amount of labeled data; however, in the real world, it is usually possible to collect only a significant amount of data under the normal operation of the system. Indeed, fault injection and run-to-failure experiments are constrained in real industrial applications for safety, security, and cost concerns. As a result, if PdM models are trained on limited or unbalanced datasets, they can easily overfit, making generalization to the test data difficult [4]. Finally, typically traditional ML and DL algorithms have acceptable performance under the condition that test data comes from the



same distribution as training data. However, real-world application complicates this issue because the working operation and environments of mechanical equipment (e.g., speed, load, and noise) might change over time. As a result, performance degrades due to the distribution discrepancy between the training dataset in a source domain and the test dataset in a target domain. In such cases, traditional ML and DL models are frequently reconstructed from scratch for a new working operation/environment, resulting in a waste of computational resources and training time. It follows that these issues make it challenging, in the case of traditional ML and DL, to generalize or adapt PdM for a new domain where there is a discrepancy in data distribution as well as a lack of data in the target domain.

In the real application world, maintenance technicians typically monitor the condition of one machine by using experience gained from different same-type machines instead of relying exclusively on insights gained from the target machine. So, in a real-world PdM application, training data can be collected from a variety of sources, such as different operation conditions, different same-type machines, failure simulation machines in the laboratory, or digital models. They will, however, have a different distribution than the data being examined in the target machine. However, owing to the same operating principle and failure mechanism, the data from the target machine and the data from the source machines should share the same fault characteristic information. [4]. These facts suggest that it may be useful to transfer knowledge across multiple related machines. In this context, TL has recently gained attention as a powerful technique to address all of these challenges.

The various advantages of adopting TL with respect to traditional ML and DL methods can be finally summarized as follows:

- Traditional ML and DL techniques rely on target domain data to train the model, while TL uses source domain data as a starting point, so requiring less target domain training data [19].
- Transfer learning-trained models can be easily adapted to other domains. Indeed, TL models are trained to recognize features, relations, instances, or parameters that may be applied to different domains [3].
- TL has the potential to make traditional machine learning and deep learning techniques more accessible. Indeed, TL algorithms can transfer to a new domain the maintenance knowledge coming from a model that has already been trained by traditional ML or DL algorithms [22].
- Unlike other learning methods, TL usually provides an optimal initial starting point, higher learning accuracy, and faster training for new domains [22].

## V. RESEARCH METHODOLOGY

The systematic literature review technique provided by Kitchenham et al. is considered throughout this paper [23].

The methodology has been applied according to the following main steps.

### A. RESEARCH QUESTIONS

The definition of proper research questions is a crucial part of a SLR [23]. Herein, the research questions aim to enable readers to better comprehend how TL can be effective in the PdM context. In so doing, we conducted preliminary research analyzing the publications useful for developing the research questions. At first, we have defined the Main Question (MQ) able to drive the search of open research challenges. Subsequently, based on the main question, we focused on Specific Questions (SQ) to highlight the existing approaches and to reveal gaps and opportunities for future studies. The main question and the specific questions are reported in Table. 1.

TABLE 1. Research questions.

Identifier	Issue
MQ	What are the current SOTA, challenges, and open issues regarding TL within PdM?
SQ1	What is the taxonomy of TL in the field of PdM?
SQ2	What is the current SOTA in methodologies and applications of TL in the field of PdM?
SQ3	Which datasets have been used and are available to apply TL in PdM?
SQ4	What are the main open issues, challenges, and opportunities of TL in the field of PdM?

### B. SEARCH PROCESS

The search process leverages two phases. The first is devoted to creating a search string, while the second aims to choose a source. The search string's construction necessitates a preliminary reading of handpicked publications related to the topic of interest. The considered search string is shown in Fig.2. Note that, on our first attempt to create the string we included keywords such as industry 4.0, cyber-physical systems, cloud computing, edge computing, internet of things, and so on. However, in this way, just a few papers have been obtained at the end of the process. So, we decided to remove these keywords from the search string in order to conduct a more comprehensive literature review in the field. Scopus has been chosen as the primary search database, due to its multi-publisher indexing (including Elsevier, Springer, IEEE, ACM, etc.), reliable metadata, powerful search engine, ease of use, extensive coverage, and quality filtering of sources [24].

### C. SELECTION OF RELEVANT STUDIES

The string presented in Fig. 2 has been applied to the Scopus database by the searching string in the title, keywords, and abs of papers. Following that, all research that is not related to the study's objectives has been removed. To this aim, the Exclusion Criteria (EC) in Table.2 have been applied.

TITLE-ABS-KEY (("predictive maintenance" OR "Prognostics and Health Management" OR "fault diagnosis" OR "fault detection" OR "fault prognosis" OR "remaining useful life") AND ("Transfer learning" OR "cross domain" OR "domain adaptation"))

FIGURE 2. Main search string.

TABLE 2. Exclusion criteria.

Section	Description
EC1	Papers are not in English language
EC2	Papers are not published in journals
EC3	Papers are not in computer science subject area

Subsequently, for further refinement, a manual filtering process has been performed, which is composed of the following steps.

- 1) Analysis of the title and abstract. It consists of reading the title and the abstract of the paper and determining whether or not it is adequate for our study. Papers, classified as surveys or reviews, are also excluded.
- 2) Entire text analysis. It consists of comprehensive reading. It is obviously essential when the title and abstract do not provide enough information about the suggested solution, while the proposed ideas appear to be viable for the objectives of the literature review.

In addition, we have further skimpily evaluated all papers at first removed on the base of the exclusion criteria, so to ensure that all relevant papers are considered in this study.

#### D. SEARCH RESULTS

This subsection illustrates the results of the search and selection process. An overview of the results is shown in Fig. 3.

In detail, 685 publications have been selected during the initial search step, and they have been then reduced to 224 exploiting the exclusion criteria listed in Table 2. In the last refinement phase, the selected publications underwent a text analysis based on the criteria listed in the section. V-C during the manual filtering phase. During this last phase, 72 studies were eliminated, leaving 152 papers resulting as the most relevant for this SLR. In addition, researchers added another 16 papers to the selected papers following a short assessment of the removed papers based on exclusion criteria.

The 168 papers obtained via the screening are categorized by year in Fig. 4, where the  $x$ -axis indicates the period of time included in this review, *i.e.*, from 2017 to 2021-22. Note that this area of research received little attention in 2017 and 2018, while in 2019 it grew to 20 papers. Furthermore, the publications show considerable growth in 2020, indicating that this will be the trend in the coming years.

#### VI. RESULTS AND DISCUSSION

The contributions of selected papers are discussed in this section. More specifically, each research question listed in

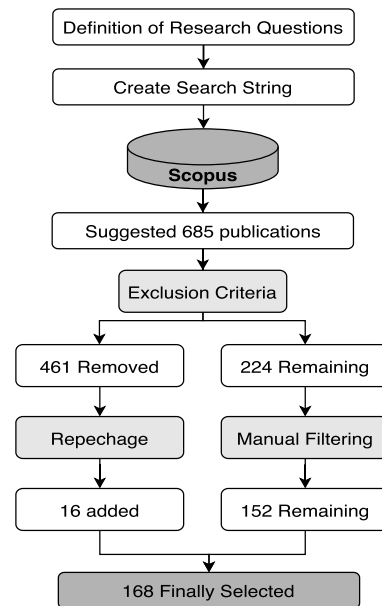


FIGURE 3. Screening of research.

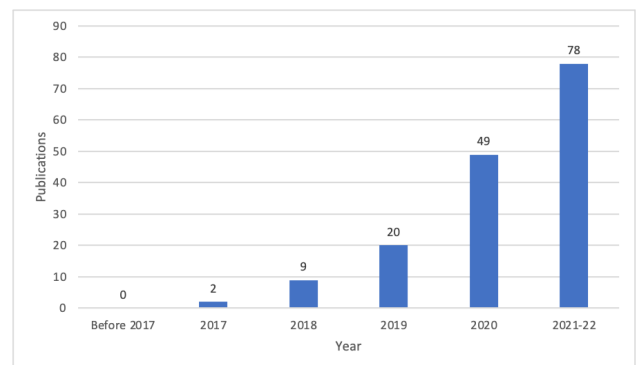


FIGURE 4. Distribution and tendency of publications by year.

Table 1 is addressed, taking into account the contributions of the publications found in the literature.

#### A. TAXONOMY OF TL IN THE FIELD OF PdM (SQ1)

Transfer learning can be classified according to a variety of factors. Different viewpoints and characteristics influence categorization. In this study, we classify the TL taxonomy by considering the problem, solution, and application criteria. According to the accessibility of labeled data in the source and target domains, the transfer learning problem could be split into three groups as follow [18]:

- **Transductive Learning.** The task is the same for the source and target domains, but domains may differ. The labeled data are only accessible in the source domain with this setting.
- **Inductive Learning.** Inductive learning differs from transductive learning in that regardless of whether the domains of the source and target are the same or not, the tasks are different. Despite whether labeled data is

available or not in the source domain, labeled data is generally accessible in the target domain in this setting.

- **Unsupervised Learning.** It's comparable to inductive TL except it's focused on the setting in which neither the source nor the target domains have a labeled dataset.

Table. 3 can be used as an index to assist researchers in identifying settings that are relevant to their interests.

**TABLE 3. Relationship between TL, and traditional ML and DL.**

Learning Settings	Domains	Tasks
Traditional ML & DL	$D_S = D_T$	$T_S = T_T$
Transductive TL	$D_S \neq D_T$	$T_S = T_T$
Inductive TL	$D_S = D_T$ OR $D_S \neq D_T$	$T_S \neq T_T$
Unsupervised TL	$D_S = D_T$ OR $D_S \neq D_T$	$T_S \neq T_T$

An alternative problem-based categorization can be based on feature and label space consistency among the source and target. On the basis of the discrepancy between the feature and label space of domains, the following types of TL can be defined: [19]:

- **Homogeneous Learning.** This concept has been developed and presented to deal with circumstances when the domains have identical labels and feature spaces. The closed-set transfer learning problem is another name for this circumstance. ( $X_S = X_T, Y_S = Y_T$ )
- **Heterogeneous Learning.** It indicates the process of transferring knowledge between domains that have different features or/and label spaces. It is more challenging than homogeneous learning since it requires feature or/and label space adaptation in addition to distribution adaptation. ( $X_S \neq X_T$  or/and  $Y_S \neq Y_T$ )

Furthermore, three forms of heterogeneous TL arise from the relation between the source and target label spaces:

- **Partial Transfer Learning.** It means that the label space of the target domain is a subset of the source domain label space. ( $Y_T \subset Y_S$ )
- **Open-set Transfer Learning.** The concept refers to the possibility of a new fault mode arising in the target domain which isn't included in the fault mode set of the source domain. As a result, the label space of the source domain is a subset of the target domain's label space. ( $Y_S \subset Y_T$ )
- **Universal Transfer Learning.** It deals with the case when there is no prior knowledge about the label space of the source and target domain. Herein, due to the lack of label space information, the universal TL problem provides a challenging and quite practical setting for general TL. ( $Y_S \neq Y_T$ )

Due to the high economic and labor expenses in real-world industries, it is generally difficult for a single source to collect enough high-quality data to build an efficient data-driven predictive maintenance model in the target domain. Therefore, depending on the number of source domains used for TL, we can have two sorts of transfers:

- **Single Source Domain.** This technique relies on knowledge from a single source.
- **Multiple Source Domain.** The multiple source domain transfer learning techniques transfer the knowledge from different multiple, but relevant sources.

Based on network structure, solutions to TL problems could be categorized into shallow and deep transfer learning approaches. The deep neural network-based TL methodology aims to learn more transferable representations by including TL methods in the deep learning pipeline. Based on the aspect of "what to transfer", transfer learning approaches are here split into the following five groups:

- **Instance-based.** This methodology attempts to reduce the conditional or/and marginal distribution difference across domains through re-weighting or applying importance sampling strategies.
- **Feature-based.** It intends to reduce the conditional or/and marginal distribution difference across domains based on two scenarios: asymmetric and symmetric. The source features are transformed to match the target characteristics in asymmetric methods. On the other hand, symmetric methods seek to find a shared latent feature space before transforming both source and target domain features into another distinct feature representation.
- **Parameter-based.** It aims to transfer knowledge at the level of the model or parameter.
- **Relational-based.** It is primarily concerned with problems in relational domains and transfers the source domain's functional relationships or rules to the target domain.
- **Hybrid-based.** It attempts to integrate two or more approaches to fulfill a single functional block of the TL model, and the integrated approaches work together to achieve desirable results.

Concerning applications, TL techniques could be exploited for accomplishing the main PdM tasks, i.e., fault detection, diagnosis, and prognosis. More specifically, TL can be leveraged for detecting a developing defect (fault detection) or in the case of faulty systems, isolating and identifying the specific type of fault (fault diagnostics), furthermore it can also be used for forecasting the RUL or the end life of a system (prognostics). Regarding application objects, predictive maintenance systems based on TL techniques could be used to provide safe and reliable operations of assets within critical systems, e.g., oil and gas, mining, aviation, industrial manufacturing components, and power plants.

The feasible transfer scenarios in this context can be divided into the following scenarios [4] (see also Table. 4):

- **Transfer in the Identical Machine (TIM).** In this transfer scenario, the source, and target domain data are collected on the same machine but under different operational conditions or working environments.
- **Transfer across Different related Machines (TDM).** The data for the source and target domains in this transfer scenario is collected from different but related machines. Compared to the TIM scenario, these data are

more complex because of differences in machine specs, structures, and so on. Hence, there is a significant data distribution discrepancy between the source and target domains. The intuitive motivation for these scenarios is that the probability of all failure modes happening during the previous operation of the target machine is low; hence gathering historical fault data from different related machines is a good option.

- **Transfer from Laboratory to Real Machine (TLRM).** Within this scenario, the source domain data is obtained from the laboratory machine to enable the identification of real-world machine fault modes in the target domain. This scenario is intuitively inspired by the fact that modeling failure modes in the lab are simpler, safer, and cheaper than gathering faulty data from a real-world machine.
- **Transfer from Virtual to Real Machine (TVRM).** Within this transfer scenario, the source domain data is collected from a machine's virtual model to provide transferable maintenance information for the target machine. The primary reason for this scenario is that, in real-world applications, the historical faulty data offered by physical machines can be limited. However, digital models, which could reflect the fundamental behaviors and principles of physical machines, can generate a significant amount of labeled data under different health and operating conditions.

**TABLE 4. Categorization of transfer scenarios in PdM.**

Transfer Scenarios	Source domain	Target domain	Reasons for discrepancy
TIM	Real-data Available	Real-data Limited	Different damage mode Different working conditions Different working environments
TDM	Real-data Available	Real-data Limited	Different damage mode Different machine specifications Different fault characteristics Different working conditions Different working environments
TLRM	Real-data Unavailable	Real-data Limited	Different damage mode Different fault characteristics Different working conditions Different working environments
TVRM	Real-data Unavailable	Real-data Limited	Different damage mode Different structure Different working environments

## B. STATE-OF-THE-ART OF TRANSFER LEARNING IN THE FIELD OF PREDICTIVE MAINTENANCE (SQ2)

In this subsection, the most relevant methodologies and applications of TL are briefly discussed on the basis of the characteristics and taxonomy depicted in Fig. 5.

### 1) APPROACH CATEGORIZATION

Table. 5 summarizes the surveyed papers based on the five primary TL approaches. The distribution of the reference

based on the approach categorization is summarized in Fig. 6. As expected, due to its capacity for projecting data into a shared feature space, where cross-domain inconsistencies may be minimized, feature-based TL is currently the most popular approach when referring to cross-domain predictive maintenance applications. In the following, we better detail and correlate the different TL approaches to the different PdM applications.

- **Feature-based.** This approach was introduced to tackle the mentioned problems by finding a common feature between domains in the latent space or decreasing the discrepancy between domain distributions by transferring features from the source to the target domain. Feature-based methods are frequently explored in the TDM scenario of TL because of their capacity to rectify significant across-domain discrepancies. For feature-based TL, there are two mainstreams. To reduce distribution shift, the first methodology, known as moment matching-based, measured statistic discrepancy by applying methods such as maximum mean discrepancy (MMD) and correlation alignment (CORAL). The second methodology is referred to as adversarial learning-based, where two parties compete to align distribution. Adversarial-based TL approaches, inspired by the adversarial learning process, have received much attention as a growing trend. Adversarial-based techniques can be split into two groups according to various strategies. The first is a generative-based technique [191], in which the primary idea is to use source data to build synthetic target data with ground-truth annotations and then use synthesized target data to enable cross-domain tasks. The second technique is the adversarial adaptation-based strategy [192], which uses a domain discriminator to adapt the representation distributions of the source and target domains.

In this framework, Tong et al. [29], for instance, developed a feature-based transfer learning method to deal with the fault diagnostic performance deterioration challenge under variable operating conditions of the bearings. A deep feature-based transfer network based on Linear Discriminant Analysis (LDA) and weighted MMD, is suggested by Wang et al. [64] for fault diagnosis of chemical processes. Different from existing research, for rolling bearing fault diagnosis, Li et al. [37] provided a new feature-based transfer network based on multi-Layer and multi-kernel MMD to reduce the domain shift problem. Wu et al. [91] proposed a deep transfer maximum classifier discrepancy approach to align the distributions of auxiliary samples produced by Batch-Normalized Long-Term Memory (BNLSTM) with unlabeled target domain data.

- **Parameter-based.**

This is also known as pre-train model-based transfer and involves fixing and transferring partial parameters of a trained network in the source domain while fine-tuning the remaining parameters in the target domain using



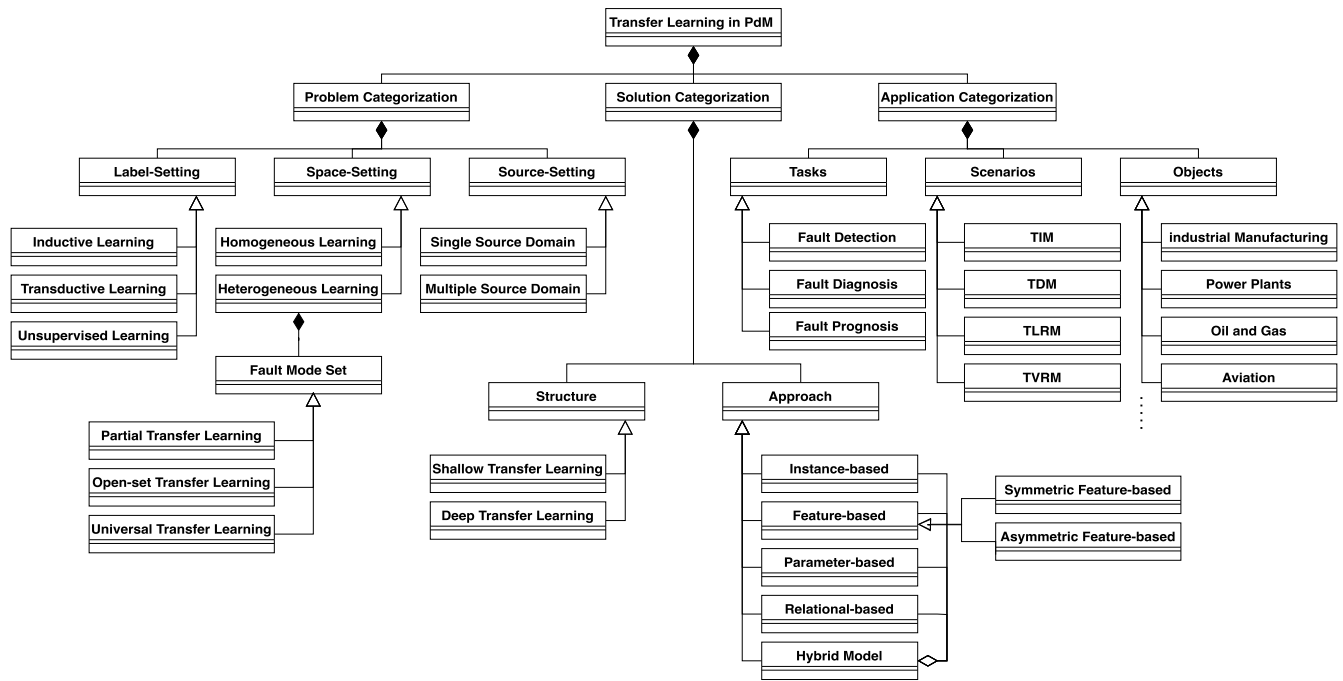


FIGURE 5. Structured taxonomy of transfer learning in predictive maintenance using class diagrams.

TABLE 5. Classification of references based on approach.

Approach	References
Instance-based	[25], [26], [27]
Feature-based	[28], [29], [30], [31], [32], [33], [34], [35], [36], [37], [38], [39], [40], [41], [42], [43], [44], [45], [46], [47], [48], [49], [50], [51], [52], [53], [54], [55], [56], [57], [58], [59], [60], [61], [62], [63], [64], [65], [66], [67], [68], [69], [70], [71], [72], [73], [74], [75], [76], [77], [78], [79], [80], [81], [82], [83], [84], [85], [86], [87], [88], [89], [90], [91], [92], [93], [94], [95], [96], [97], [98], [99], [100], [101], [102], [103], [104], [105], [106], [107], [108], [109], [110], [111], [112], [113], [114], [115], [116], [117], [118], [119], [120], [121], [122], [123], [124], [125], [126], [127], [128], [129], [130], [131], [132], [133], [134], [135], [136], [137], [138], [139], [140], [141], [142], [143], [144], [145]
Parameter-based	[146], [147], [148], [149], [150], [151], [152], [153], [154], [155], [156], [157], [158], [159], [160], [161], [162], [163], [164], [165], [166], [167], [168], [169], [170], [171], [172], [173], [174], [175], [176], [177], [178], [179], [180], [181]
Relational-based	[182],
Hybrid-based	[183], [184], [185], [186], [187], [188], [189], [190],

target data. In general, parameter-based approaches assume that common parameters exist across models from the source and target domains. In reality, it is not easy to ensure that the source and target domains share common parameters. As a result, these techniques are ineffective in cases where the domain shift is significant. In recent years, several researchers have mostly used this approach to accomplish transfer predictive maintenance of machines. Wen et al. [165], for instance, employed negative correlation learning to fine-tune the fully-connected and softmax classifier layers of a pre-trained ResNet-50 model by transforming raw time-domain data to RGB images. Liu et al. [155], based on LeNet-5, developed a CNN model to capture fault features from images, then used a fine-tuning technique to adopt the CNN model for the target domain, which is an

effective method of parameter-based TL. By fine-tuning the parameters of the transferred CNN model, which was trained with a significant amount of data, Han et al. [157] developed a parameter-based TL network for fault diagnosis of the target machine with new tasks. Even though this transfer technique is simple to comprehend and implement, the need for labeled data in the target domain remains a source of frustration. Hence, as a result of the unavailability of labeled data in real-world industrial applications, these TL approaches could face unexpected challenges. To deal with this issue, Li et al. [169], suggested a transfer network based on parameter TL and ensemble learning methods for identifying rolling bearing faults using target unlabeled data.

- **Instance-based.** This methodology aims to reduce the marginal or conditional distribution difference

across domains by leveraging importance sampling or reweighting methods. When applied to the target domain, the technique focuses on determining appropriate weights for labeled data from the source domain in order to learn the source task with the lowest possibility of negative transfer. Furthermore, instance-based techniques are founded on two fundamental assumptions:

a) Since only a certain number of training data from the source domain are relevant to the target domain, all source instances cannot be used to train the target model. As a result, it's critical to carefully choose samples that will support the target domain task

b) The conditional distributions of the source and target domains are identical. This assumption, however, may not always be valid. As a result, this strategy is both challenging and inefficient. It can be used to resolve discrepancies in TIM scenarios, but it cannot handle significant discrepancies in transfer scenarios, for instance, TDM scenarios, since these techniques may lack robust data fitting capacity.

Despite its merits, according to Table 5, few studies focus on the application of this approach to PdM. For example, Liu et al. [25] introduced a generic transfer framework to overcome the challenge of designing fault diagnosis with small samples. In the developed framework, the authors utilized an instance-based TL technique and the Adaptive Mixup (Admixup) method to transfer the vast number of fault samples in various working conditions. Lee et al. [26], designed an instance-based TL network that employs the Kullback-Leibler Divergence (KLD) and MMD indicators in combination.

- **Relational-based.** The goal is to find the correlation among data in the source domain and then apply that knowledge to the target domain task. Despite the other three methods of transferring, relation-based TL approaches do not require that the source data and the target data be distributed independently and identically. As a result, relation-based approaches are far more adaptable and robust than other approaches. Furthermore, most of these approaches are built on statistical learning techniques.

However, according to Table 5, relational-based TL is rarely exploited in the PdM field. Zhu et al. [182], for example, presented a flexible TL framework for transferring information from both a qualitative and quantitative perspective for monitoring the batch process. To begin, a statistical pattern clustering technique is designed for evaluating and distinguishing similar conditions. In addition, a multiphase bayesian network is built with nominal representations enabling qualitative knowledge transfer and statistical modeling.

- **Hybrid-based.** The hybrid techniques refer to the case when at least two approaches are integrated to perform one single functional block of a transfer and the combined models cooperate to achieve their outputs.

Nonetheless, integrating multiple TL methods into a hybrid approach provides an effective means of using historical information, but it also introduces new challenges.

Again, despite its benefits, the majority of surveyed studies (as summarized in Fig. 6) focus on one sort of TL approaches, such as feature-based and parameter-based TL, and only a few studies exploit hybrid-based transfer learning approaches. For instance, Hybrid Transfer Learning (HTL) has been recently proposed by Ma et al. [184] for predicting the behavior of Proton Exchange Membrane Fuel Cells (PEMFCs) based on intercell differences. Two types of TL, instance-based and parameter-based, are used to extract as much information as possible from past data and prior models to increase the similarity of the generated curve. Sun et al. [183] presented a deep hybrid-based TL network based on Sparse Autoencoder (SAE) and three transfer techniques, including weight transfer, feature TL, and weight updating, to transfer prognostic knowledge from a trained SAE to a new domain. These methods enable the prediction of a new domain without the need for labeled target data. For sparse target data, Han et al. [185], a proposed framework based on a hybrid approach. The proposed framework's overall goal is divided into two portions: multiple adversarial domain adaption and supervised fine-tuning. In order to overcome the open-set TL challenge in machinery fault diagnosis, Zhang et al. [186], employed a hybrid model based on adversarial learning and instance-level weighted techniques to capture generalized features and represent the similarities of testing samples with known health conditions. Li et al. [189] exploit both strategies of feature-based and parameter-based TL in their proposed Knowledge Mapping-based Adversarial Domain Adaptation (KMADA) structure, so as to achieve fast convergence and satisfying outcomes.

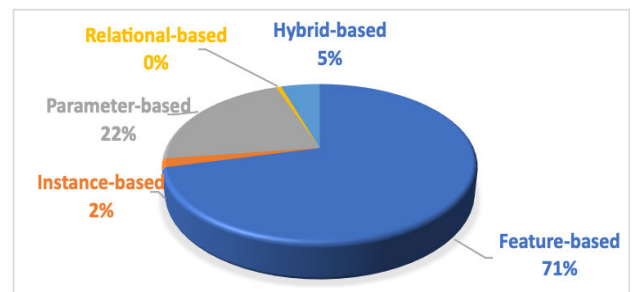


FIGURE 6. Distribution of references based on approach categorization.

## 2) PROBLEM CATEGORIZATION

In the following, the surveyed papers have been also classified and discussed based on different problem-setting in the source and target domains.

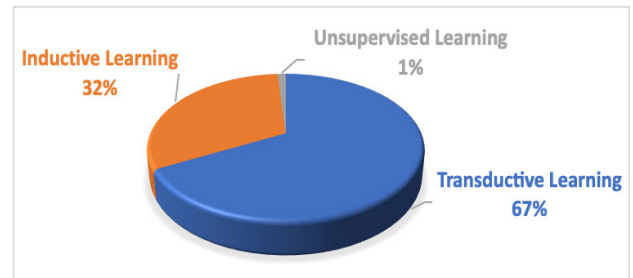
- **Label-setting Categorization.** The categorization of the selected papers on the basis of label-setting categorization is summarized in Table 6. The results in Fig. 7, clearly show that inductive and transductive TL has a significant body of contributions, and this can be clearly motivated by the fact that inductive and transductive TL are natural extensions of the very popular supervised learning approaches.

However, ML tasks can be often unsupervised in industrial applications, especially when monitoring complex industrial processes, since, the number of possible faults is uncountable due to the enormous number of parts and interacting components [193]. In addition, because industrial systems are designed to be reliable, obtaining a labeled dataset with sufficient samples of each potential fault, as the source or target domain, is often impossible in practice. It follows that often transductive and inductive algorithms cannot be exploited in practice, since they are not proper for handling conditions in which the source or/and target domains are unlabeled. Unsupervised TL, on the other hand, can deal with the absence of labeled data in both the source and target domains. So, it could be very useful for dealing with unique and special tasks for which there isn't enough labeled data from both the source and target domains. Despite its great benefits, unsupervised TL is still not popular in the PdM field, since learning good representations from a significant set of unlabeled data is a particularly challenging task that makes difficult its application to real-world tasks [194].

Among the few works focusing on Unsupervised Transfer Learning (UTL), it is worth noting the novel framework proposed by Mao et al. [102] for online early fault detection via UTL. The approach is aimed to improve the resilience of the detection model, resulting in a much lower false alarm rate, by combining robust state evaluation and coupled adversarial training of deep domain adaptable neural networks. Gabriel et al. [104], proposed a new framework for solving one-class classification issues. The approach is intended to detect anomalies in fleets of machines from the same manufacturer that are monitored by comparable sensors but are experiencing domain shifts due to differences in system setup, operation, or environment. However, the main drawback to this configuration is related to the difficulties arising from training the domain discriminator in an imbalanced configuration.

- **Source-setting Categorization.**

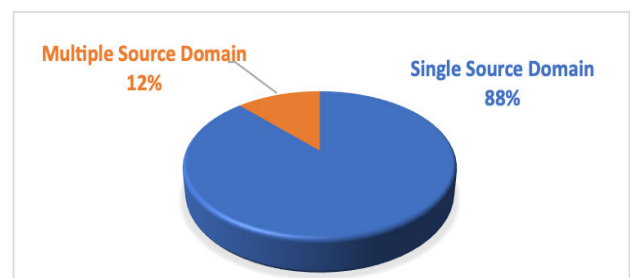
Table 7 summarizes the selected papers based on source-setting categorization, while their distribution in Fig. 8 discloses that most of the publications (namely, 88%) are focused on single source TL. Mainly for the sake of simplicity, current transfer learning-based fault diagnostic approaches mostly rely on transferring maintenance knowledge from one source domain to the target domain. Moreover, the accessibility of significant



**FIGURE 7.** Distribution of references based on label-setting categorization.

amounts of labeled training data from a single source allows for achieving excellent performance and testing accuracy.

Nevertheless, it is sometimes impossible for a single source to provide adequate labeled data for developing an efficient data-driven predictive maintenance model in real industries. On the contrary, it sometimes could be easier to collect labeled data from many different sources with comparable operating machines in order to expand the training dataset at a lower cost. Moreover, further advantages arise. First, by leveraging all the independent data collected from several source domains, the model will be able to transfer more comprehensive and general diagnostic and prognostic knowledge. Secondly, the possibility of over-fitting can be alleviated with more training samples, which could provide favorable performance in the target domain [105]. When shifting among multiple domains, even though greater data exploration could theoretically lead to higher model performance, in practice, it results to be hard aligning the distributions between all source domains and target domains. More specifically, as illustrated intuitively in Fig. 9, it is very challenging to completely remove the shift between a single source and target domain. So when attempting to align multiple sources and target domains, a considerable degree of mismatch commonly arises, which could adversely affect model performance. This further motivates the distribution in Fig. 8.



**FIGURE 8.** Distribution of references based on source-setting categorization.

As shown in Table 7, a few studies have recently exploited multiple source domain transfers to PdM. For instance, with respect to prediction accuracy,

**TABLE 6.** Classification of references based on label-setting.

Label-setting	References
Transductive Learning	[112], [28], [29], [30], [31], [32], [33], [84], [34], [183], [36], [37], [45], [39], [40], [35], [41], [42], [85], [43], [44], [60], [46], [48], [49], [65], [190], [86], [87], [89], [50], [51], [56], [92], [58], [93], [53], [59], [62], [57], [63], [64], [91], [66], [67], [95], [96], [97], [98], [81], [168], [68], [186], [47], [69], [70], [187], [99], [72], [73], [54], [100], [101], [74], [75], [169], [188], [94], [103], [76], [189], [77], [78], [106], [107], [88], [108], [109], [79], [27], [80], [110], [111], [82], [113], [114], [115], [116], [117], [118], [119], [120], [122], [123], [124], [125], [126], [127], [128], [129], [130], [131], [195], [133], [134], [135], [137], [138], [140], [141], [142], [143], [145]
Inductive Learning	[146], [148], [182], [149], [147], [150], [83], [153], [152], [38], [154], [196], [155], [157], [158], [25], [162], [159], [160], [90], [52], [161], [55], [163], [61], [164], [166], [167], [173], [185], [26], [71], [172], [105], [184], [151], [165], [170], [174], [175], [176], [156], [177], [121], [132], [136], [171], [178], [139], [179], [180], [181], [144]
Unsupervised Learning	[104], [102],

**TABLE 7.** Classification of references based on source-setting.

Source-setting	References
Single Source Domain	[146], [148], [112], [28], [182], [29], [30], [31], [147], [32], [150], [83], [153], [33], [84], [34], [183], [152], [37], [38], [196], [39], [155], [40], [35], [157], [41], [25], [42], [159], [85], [43], [44], [60], [46], [47], [48], [49], [65], [190], [86], [87], [89], [51], [52], [54], [161], [55], [91], [56], [81], [92], [58], [93], [162], [53], [59], [94], [163], [62], [57], [164], [63], [64], [66], [166], [67], [95], [96], [167], [97], [98], [168], [186], [69], [70], [71], [187], [73], [100], [101], [74], [75], [169], [188], [173], [26], [103], [76], [104], [189], [88], [77], [78], [172], [106], [107], [185], [108], [184], [109], [79], [27], [80], [111], [82], [113], [114], [118], [119], [120], [121], [122], [123], [125], [126], [127], [128], [129], [130], [131], [195], [133], [134], [135], [136], [137], [171], [138], [178], [139], [140], [179], [141], [143], [180], [181], [145]
Multiple Source Domain	[149], [102], [36], [160], [50], [90], [61], [68], [99], [72], [105], [110], [115], [116], [117], [124], [132], [142], [144]

Wen et al. [36] conducted analytical experiments to prove that the third dataset XTD might influence the DTL's final prediction accuracy. More specifically, they disclose that the prediction accuracy improves when the third dataset is more similar to the target dataset than the source dataset. Note that, transferring from multiple operating conditions to a single operating condition might also have a negative impact as shown in Zhang et al. [149]. Herein, the authors attempted to explain this negative effect by comparing sensor monitoring data values. More notably, they try to disclose what factors influence the impact of TL under various operating conditions and two factors have been identified so far. First, sensor data gathered under multiple operating conditions are more complex than data collected under single conditions. Second, the sensor value distribution differs from the sensor monitoring data with a single operating condition.

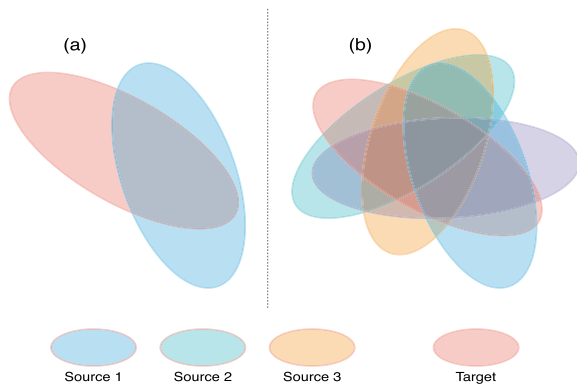
As discussed in [68], domain-invariant representations are difficult to understand when single-source domain adaptation is employed to explain the distribution of obtained data under various working conditions. To tackle this issue, Wen et al. [68] applied a technique based on multi-feature spatial domain adaptation. To achieve more accurate homogeneous transfer knowledge in the TIM scenario, Tian et al. [142] developed a multi-source subdomain adaptation transfer learning approach by employing a multi-branch network structure and local MMD method. The multiple source domain adaptation for machinery fault diagnosis has been very recently also addressed via Weighted Domain

Adaptation Neural Network (WDAN) in Wei et al. [110] by leveraging some criteria to find whether to conduct domain adaptation before starting the training or traditional supervised learning in order to prevent negative transfer. However, they are mostly focused on feature transferability and neglect the impact of sample transferability on domain adaptation. To this purpose, Shi et al. [116] introduced a unique unsupervised MDA-based TL scheme called Multisource Domain Factorization Network (MDFN), which learns generic diagnostic information from numerous sources and then applies it to diagnosing the target task. The proposed framework employs transferability-based entropy penalties and shared-space component analysis methods to deal with negative transfer from the perspectives of instance transferability and feature representation.

However, the theoretical frameworks proposed in the above studies are mainly designed under the main restrictive assumption that both source and target domains have identical features and label spaces. This is often unfeasible since the health condition sets in source and target domains are usually different in a real industrial environment. To deal with this issue, Li et al. [90] have instead developed a novel deep learning-based heterogeneous TL technique for diagnostics, in which diagnostic information, obtained from adequate labeled data of multiple rotating machines, is transferred to the target machine. However, the real-world application of the approach is limited since the required labeled target training samples are usually unavailable in practice. Finally, in order to address the issue of data



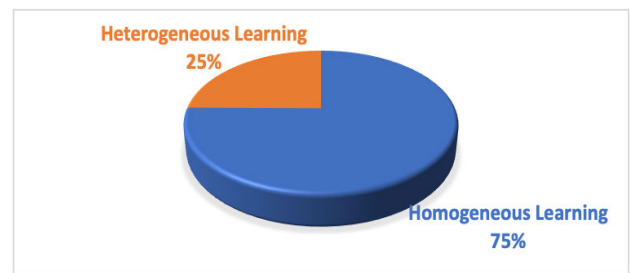
imbalances across health states in multi-source TL scenarios, Yang et al. [117] introduced a network that combines several partial distribution adaption sub-networks and a multi-source diagnostic knowledge fusion module to collect and leverage diagnostic knowledge from multiple source machines.



**FIGURE 9.** Demonstration of single-source and multi-source TL. (a) Single-source TL. (b) Multi-source TL.

- Space-setting Categorization.** In Table. 8, the surveyed publications are classified based on the space-setting categorization. Currently, due to ease of application, most transfer learning methods operate under the assumption the feature spaces of the data in the source and target domains are represented by the same attributes [197]. This also reflects in the PdM literature as highlighted by results in Fig. 10, where (75%) of the TL studies leverage homogeneous transfer learning. Nevertheless, regardless of the progress and achievement of homogeneous TL methods for PdM, there is still a challenge that cannot be neglected. Indeed, a challenging assumption is still required for homogenous approaches, i.e., they require a source dataset and target dataset sharing the same label space, that is, the health conditions sets in the two domains are identical. In real-world industrial applications, this assumption may not always be fulfilled, where collecting data under all the possible machine health conditions is very difficult. As shown in Fig. 11(a), since the homogeneous or closed-set TL assigns equal weights to the source and target instances, cross-domain data have an equal contribution to the adaptation process. It follows that most of the existing homogeneous TL techniques are exposed to negative transfer because they attempt to align all data, even outlier samples. In real-world industrial applications, since maintenance should be both effective and efficient, it is not reasonable that equipment cannot be monitored until data of the same category is obtained from the source domain and adapted to the target domain. This refers to a situation where the target label space is a subset of the source label space and is named a partial transfer learning problem.

As a result, compared to the scenario for which the closed-set TL is designed, this configuration is closer to engineering practice. This kind of TL is depicted in Fig. 11(b). With the aim of handling the partial TL problem in machinery fault diagnostics, Li et al. [95] developed a novel deep TL model based on class-weighted adversarial networks. Liu et al. [190] developed a deep partial adversarial domain adaptation network based on an SAE algorithm, to weigh and recognize common instance types from mixed source domain instances. For achieving a broad diagnostic model in the partial transfer applications, Deng et al. [109], introduced a Double-layer Attention-based Generative Adversarial Network (DA-GAN) to find target label space that should be considered for the transfer task, and to determine which samples for each sub-domain discriminator should be focused on. Finally, to overcome the challenges arising from the presence of unbalanced data across every health state of the target domain and heterogeneity of label-space in the partial mode, Yang et al. [78] presented an adversarial adaptation model called Deep Partial Transfer Learning Network (DPTL-Net) for machinery fault diagnosis.



**FIGURE 10.** Distribution of references based on space-setting categorization.

Often in practice, as shown in Fig. 11(c), new fault modes which not present in the source classes, may appear in the testing phases. A key factor in dealing with this situation might be finding source data with the same class as the target data so that these data may be used for domain adaptation and classifier training. Nevertheless, because the target data usually is fully unlabeled and the target label space is unknown, finding source data that is associated with the target label space might be challenging. This challenging scenario, in which the source domain's label space is a subset of the target domain's label space, is named open-set transfer learning in the technical literature. It is worth noting that, transfer learning approaches that rely on marginal distribution alignment cannot achieve class-level diagnostic knowledge transfer within the open-set transfer learning context, due to target outlier classes, thus often resulting in model adaptation capacity degradation. Despite its relevance, however, the open-set transfer learning problem has received only a little attention in the technical literature.

**TABLE 8.** Classification of references based on space-setting.

Space-setting	References
Homogeneous Learning	[148], [112], [28], [182], [29], [30], [31], [147], [32], [83], [153], [33], [84], [34], [183], [152], [36], [37], [38], [39], [155], [40], [35], [41], [25], [42], [85], [43], [44], [60], [47], [48], [49], [65], [160], [86], [87], [89], [54], [91], [56], [81], [92], [58], [93], [162], [53], [94], [163], [61], [62], [57], [64], [66], [166], [67], [96], [167], [97], [98], [168], [68], [69], [70], [71], [99], [100], [74], [75], [169], [188], [26], [123], [103], [76], [104], [55], [189], [77], [172], [105], [107], [108], [27], [110], [111], [82], [113], [114], [180], [115], [116], [119], [120], [121], [122], [126], [127], [128], [129], [130], [131], [195], [133], [136], [137], [171], [138], [178], [139], [140], [179], [141], [142], [181], [144]
Heterogeneous Learning	[146], [164], [149], [157], [150], [196], [159], [46], [106], [190], [50], [63], [51], [90], [95], [52], [161], [79], [59], [186], [187], [72], [173], [101], [73], [78], [185], [80], [109], [117], [118], [124], [125], [132], [134], [135], [143], [145]

For example, Wang et al. [52] introduce a CNN-based technique in conjunction with DA and prototype learning, to address the open-set case in fault diagnosis. To investigate open-set challenges within the TIM scenario, Zhu et al. [143] introduced a new adversarial domain adaptation network with diverse supplementary classifiers to diagnose the unknown and known fault categories in the target domain. Zhang et al. [186] also tackled this problem via a deep learning-based adversarial training strategy for extracting domain-invariant features from source and target domains. In detail, an instance-level weighted technique is suggested to identify the target outlier classes, representing the similarities of the target instances with the source classes.

Both partial and open set transfer learning techniques demand previous knowledge of label spaces across domains, hence implying that these approaches are suitable for dealing with off-line cross-domain problems. However, in a more general PdM scenario, obtaining the relationship between the source and target label spaces in advance is not feasible. Hence, a more difficult scenario, known as universal transfer learning, needs to be tackled. It follows that no previous knowledge about machine faults can be obtained in advance for the target domain and this suggests that testing machines may have both known and unknown source-domain faults (as shown in Fig. 11(d)). Note that this universal transfer learning scenario commonly arises when implementing predictive machine maintenance in practice. To address this challenging problem, Zhang et al. [187] exploited a deep adversarial learning technique so to bridge domain gaps across various operating conditions. More in detail, source class-wise, and target instance-wise weighting methods have been suggested for selective domain adaptation. In addition, novel outlier identifiers and reconstructor modules have been included to discover unknown fault modes while maintaining data information in processing.

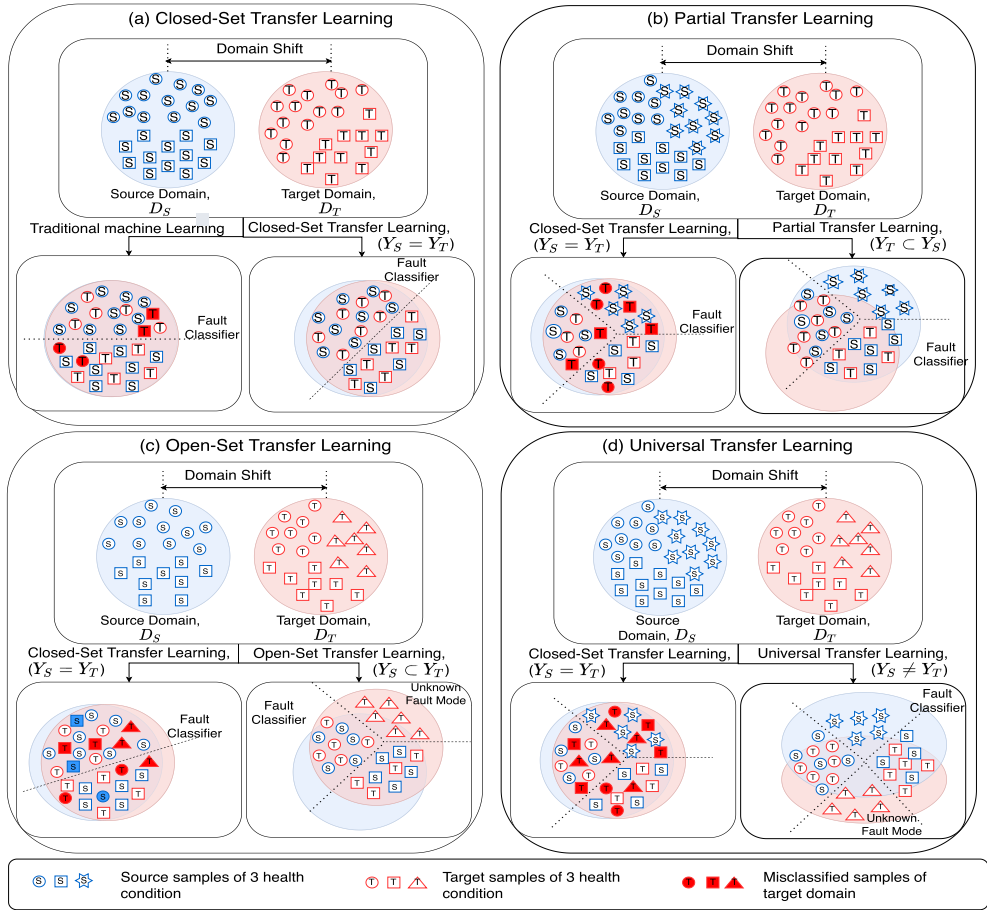
### 3) APPLICATION CATEGORIZATION

The selected papers have been classified and discussed in the following based on different application-setting in the source and target domains.

#### • Transfer Scenarios Categorization.

Depending on the application scenario, the selected papers have been classified in Table 9. It's important to note that in some works, the authors validated their proposed framework in more than one transfer scenario. Results in Fig. 12 confirm that, as expected, the majority of the studies validated their proposed approach in the simplest scenarios, i.e., TIM. The main weakness of these studies is that only TL within a machine is investigated and only transferred the maintenance knowledge from one operating condition to the next. Whereas, in the real world, the transfer of maintenance knowledge from the data collected from different related machines is the most realistic scenario for the PdM [4]. Moreover, within this challenging scenario, additional reasons, including fault machine specifications, characteristics, etc., can impact the discrepancy of the domains' data distribution in addition to the working condition. Along this line, Chen et al. [161] developed a transferable diagnostic model based on the CNN algorithm to enhance the performance of a target machine fault diagnosis model by using information obtained through different machines (historical data). However, in practice, when dealing with transferrable maintenance scenarios, only normal target domain samples are generally available, and this could significantly influence the proposed approach's performance. To tackle this issue, Zheng et al. [47], developed an innovative feature-based TL approach for fault diagnosis of gearbox machines, named Transfer Locality Preserving Projection-based Intelligent Fault Identification (TLPPIFI).

However, it is sometimes difficult to obtain enough labeled data, even from other real-world machines, because real-world machinery is typically maintained in good condition and failures are unusual. As a result, obtaining real faulty data is more difficult and time-consuming than collecting real normal data, and whereas real normal data is typically sufficient, real faulty data is sometimes inadequate. Furthermore, the fault types are often unknown during the running of real-world machines. It is impracticable to stop machines on a regular basis and then evaluate their health conditions based on the data obtained. Additionally, when data quantities grow at a fast rate, manual data labeling



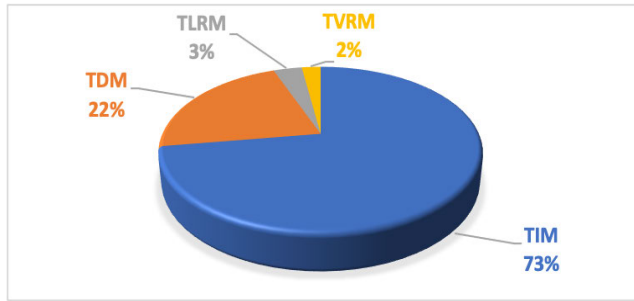
**FIGURE 11.** Demonstration of the (a) Closed-Set, (b) Partial, (c) Open-Set, and (d) Universal transfer learning problems.

**TABLE 9.** Classification of references based on application scenarios.

Scenario	References
TIM	[146], [148], [112], [28], [182], [29], [30], [31], [149], [147], [32], [150], [33], [84], [34], [183], [36], [37], [38], [196], [155], [40], [35], [157], [41], [25], [42], [159], [85], [43], [60], [46], [48], [49], [65], [190], [86], [87], [89], [50], [51], [52], [56], [81], [92], [58], [162], [53], [59], [163], [61], [62], [57], [164], [63], [64], [66], [166], [67], [95], [96], [167], [97], [98], [168], [68], [186], [71], [88], [187], [99], [72], [73], [100], [101], [74], [75], [188], [102], [26], [103], [76], [104], [189], [77], [78], [172], [106], [107], [185], [108], [184], [109], [79], [27], [110], [111], [113], [114], [115], [116], [117], [118], [119], [121], [124], [125], [126], [127], [129], [130], [131], [195], [132], [134], [136], [137], [171], [138], [140], [141], [142], [143], [180], [181], [144], [145]
TDM	[44], [60], [47], [160], [51], [90], [152], [52], [54], [161], [104], [55], [91], [59], [71], [94], [164], [67], [69], [169], [188], [105], [185], [108], [109], [80], [82], [116], [117], [119], [120], [123], [128], [132], [133], [135]
TLRM	[39], [51], [70], [79], [78], [122]
TVRM	[153], [93], [178], [173], [179], [139]

becomes an ineffective solution due to significant human labor costs and a high dependency on knowledge [4]. In order to address this problem, introducing various fault types into a lab-case machine and then acquiring a huge amount of labeled faulty data might be a feasible solution. The collected data from laboratory and real-world machines, on the other hand, have a dramatically different distribution, which is impacted by a number of reasons, such as the measurement environment, working conditions, fault characteristics, damage mode, and

others. As a result, maintenance models that were exclusively developed based on data collected from laboratory machines may not perform well when applied directly to real-world machinery. As shown in Fig. 12, only a few studies have been devoted to a deeper understanding of the above-mentioned TLRM scenario, and its related issues, via TL. For instance, with the aim of identifying the health conditions of wind turbine bearings, Lv et al. [70] introduced a deep transfer model based on a multi-kernel dynamic distribution adaptation method to



**FIGURE 12.** Distribution of references based on application scenarios categorization.

transfer diagnosis knowledge from a laboratory bearing. Furthermore, by using data gathered in labs from the gearbox and motor bearings, Yang et al. [39] introduced a feature-based transfer model based on a neural network, to diagnose the health conditions of locomotive bearings in real-world applications. Despite the remarkable success, researchers ignore the changes in the class weight of the target machine. Due to the controlled experimental circumstances, most lab machine datasets are balanced, but in real-world industrial applications, a considerable amount of training data belongs to normal conditions. To address this problem, Cao et al. [79] developed a pseudo-categorized MMD that takes into account the various class weight bias in the real-case machines dataset and exploits the category probability vector as a penalty term in the MMD.

However, in other cases, access to historical fault data generated by physical equipment during a machine's real-world application may be restricted, or the cost of obtaining faulty data through the lab-case machine may be too expensive. In all these cases, digital models, able to capture the underlying rules and behaviors of real-case machines, can produce the required large amounts of data under different operating and health conditions. As a matter of fact, even though both the virtual and real models are centered on the same machine health condition, there is always a distribution difference between them for a variety of factors. To begin with, it is practically unfeasible to integrate all features of a real-world machine into a digital model; as a consequence, certain simplifications must be made, and just the primary components are often included. Besides, digital modeling cannot fully explain the presence of uncertainty, noise, and random environmental influences. Finally, the distribution of vibration signals in real-world machines is affected by various factors (e.g., sensor type, calibration, installation and fixation technique, drift, structure transmission properties, and so on) that cannot be simulated in a digital model, e.g., via a simplified mathematical representation such as dynamical systems or differential equations. In a nutshell, addressing the distribution discrepancy challenge

between digital and real-world model data is critical for designing a predictive maintenance model that performs well in real-world machines. [4]. However, as shown in Fig. 12, this TVRM scenario is currently the topic of just a few research. For example, Xu et al. [153] introduced a deep TL model based on the digital twin technique to assist in fault diagnosis on the digital shop floor. Here, the source domain data is produced by the designed digital model, whereas the real shop floor data is considered as the target domain dataset. Li et al. [93] developed a deep TL approach to conduct cross-domain knowledge transfer from simulated data in order to identify anomalies in physical processes. Experiments on simulated Continuously Stirred Tank Reactors (CSTRs) and plant-wide pulp mill processes have shown that integrating computer simulation with TL for fault diagnosis could be effective. To enable simulation-to-real-world domain adaption, Liu et al. [139], developed the phenomenological model of bearing to generate the simulation data. Then, a domain adversarial neural network is developed to transfer the simulated data by the digital model to the target domain. The results of the experiments reveal that the proposed approach is capable of achieving high classification accuracy. With a specific focus on the limited real data problem, Don et al. [173] developed a novel intelligent diagnosis approach for the race fault of bearings. Herein, a dynamic bearing model is used to produce a huge amount of simulation data; results suggest that the diagnosis knowledge obtained from the virtual model can be effectively transferred to real-case rotating machines.

For the sake of completeness, Table. 10 provides an overview of application objects related to cross-domain predictive maintenance (Note that, some authors validated their proposed framework in more than one case study). Results related to the ongoing cross-domain PdM research and summarized in Fig.13 clearly show that bearing and gearbox are the two most commonly used case studies. The main reason is the availability of public open-source fault data of bearings and gearboxes gathered from different machines in different working operations, which makes cross-domain studies easier to be executed.

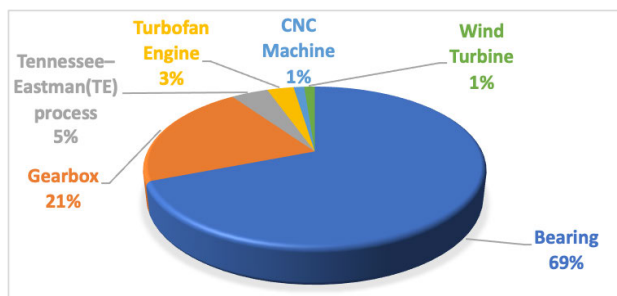
Only a few cross-domain PdM preliminary studies cope with different kinds of components or machines, such as Tennessee–Eastman (TE) process [104], Turbofan Engine [106], Chillers [80], Induction Motor [154], 3D Printer [32], Power Transmission Line Inspection [160], Fed-batch Penicillin Fermentation [182].

- **Tasks Categorization.** The surveyed publications have also been classified on the base of the application tasks in Table. 11. Herein, results in Fig. 14 confirm that only very few studies are currently focusing on detection and prognosis tasks, since gathering time-series data needed for prognosis purposes in predictive maintenance is much harder and more time-consuming

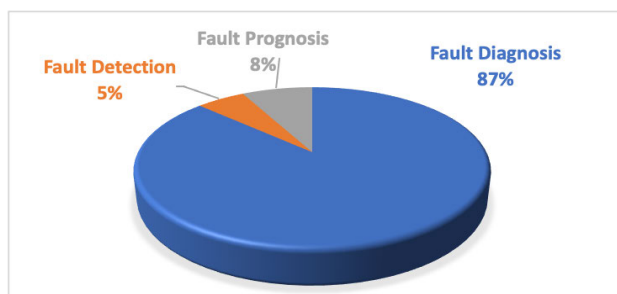


**TABLE 10.** Summary of the references based on application object categorization.

application object	References
Bearing	[146], [112], [29], [30], [31], [147], [150], [84], [152], [36], [154], [39], [155], [37], [40], [35], [41], [158], [42], [85], [43], [44], [46], [47], [48], [49], [65], [190], [86], [87], [89], [50], [51], [90], [52], [54], [161], [55], [91], [56], [92], [58], [162], [53], [59], [163], [61], [57], [164], [63], [66], [166], [95], [165], [96], [97], [98], [168], [68], [186], [70], [71], [187], [99], [73], [100], [101], [74], [75], [169], [188], [173], [102], [103], [104], [189], [77], [78], [172], [105], [107], [185], [108], [109], [79], [170], [27], [175], [177], [113], [114], [115], [116], [117], [118], [119], [120], [121], [122], [123], [124], [126], [127], [128], [129], [130], [131], [195], [132], [133], [134], [135], [137], [138], [139], [140], [141], [142], [143], [181]
Gearbox	[30], [31], [34], , [154], [39], [35], [151], [157], [47], [65], [50], [51], [161], [58], [53], [63], [166], [67], [97], [98], [71], [73], [100], [101], [103], [77], [78], [27], [110], [111], [113], [117], [125], [127], [131], [141], [142], [180], [145]
Tennessee–Eastman(TE) process	[28], [196], [25], [60], [46], [81], [64], [144]
Turbofan Engine	[149], [45], [159], [104], [106]
CNC Machine	[183], [88]
Wind Turbine	[50], [185]
Others	[148], [83], [153], [38], [93], [94], [62], [154], [160], [167], [69], [182], [72], [26], [33], [32], [76], [184], [80] [82], [156], [174], [176], [136], [171], [178], [179]

**FIGURE 13.** Distribution of references based on application object categorization.

with respect to gathering the data for classification issue.

**FIGURE 14.** Distribution of references based on application tasks categorization.

### C. DATASETS TO APPLY TL WITHIN PdM (SQ3)

High-quality data is the foundation for the implementation of TL approaches. Therefore, to design successful TL algorithms for PdM, a good collection of datasets is necessary. As already pointed out, collecting the dataset from real machines, is a very time-consuming process, since natural fault degradation is a slow process, that can take years. To tackle this issue, in some experiments, researchers gather

data by employing components with artificially induced faults or accelerated life cycle approaches. Nevertheless, data collection is still difficult and costly and, hence, several organizations have made their fault databases accessible to engineers and researchers, so that they can be exploited by the scientific community. These datasets can also be used as a baseline platform for evaluating and comparing different approaches since they are a common ground for scientists due to their widespread use. This section provides a brief summary of the open-source datasets in Table 12, where the sensor type of each dataset is shown in the second column and the number of sensors for each dataset is disclosed in the third column, while the remaining four columns refer to the sampling frequency, monitoring object, fault mode, and task of each dataset, respectively.

### D. OPEN ISSUES, CHALLENGES AND OPPORTUNITIES OF TL FOR PdM (SQ4)

The potential of transfer learning is becoming increasingly apparent, and many researchers believe that this technique can really improve the predictive maintenance industry [19]. However, considerable challenges must still be tackled and solved before reaching its full potential. The following are briefly mentioned important concerns, challenges, and opportunities:

#### 1) NEGATIVE TRANSFER

Transfer learning has been used to address the problem of a lack of training data in a target domain by leveraging knowledge from one or more source domains. Transferred knowledge, on the other hand, may not always have a positive impact on the target domain's task since its effectiveness is dependent on several assumptions [4]. For example, any violation of the following assumptions may result in negative transfer (NT). To begin with, the tasks in both domains should be related or similar. Second, the distribution of data across the domains should not be too diverse. Lastly, a proper TL

**TABLE 11. Summary of the references based on application tasks categorization.**

Tasks	References
Fault Detection	[148], [160], [61], [168], [102], [156], [76], [104], [136]
Fault Diagnosis	[146], [112], [28], [182], [29], [30], [31], [151], [147], [32], [150], [153], [33], [84], [34], [152], [36], [37], [38], [154], [196], [39], [155], [40], [35], [157], [41], [158], [25], [42], [83], [85], [43], [44], [60], [46], [47], [48], [49], [65], [190], [86], [89], [50], [51], [90], [52], [165], [87], [54], [161], [55], [91], [81], [92], [58], [93], [162], [53], [59], [94], [163], [62], [57], [164], [63], [64], [66], [166], [67], [95], [96], [167], [97], [98], [68], [186], [69], [70], [71], [187], [99], [72], [73], [100], [101], [74], [75], [169], [188], [173], [26], [103], [189], [77], [170], [78], [172], [105], [174], [107], [185], [108], [109], [79], [80], [110], [111], [82], [175], [176], [177], [113], [114], [115], [116], [117], [118], [119], [120], [122], [124], [125], [126], [127], [128], [129], [130], [131], [132], [133], [134], [135], [137], [178], [139], [140], [179], [141], [142], [143], [180], [181], [144], [145]
Fault Prognosis	[149], [183], [88], [45], [159], [56], [106], [184], [121], [123], [195], [171], [138]

**TABLE 12. Comparison of open-source fault datasets.**

Dataset (Download link)	Sensor type	Sensor number	Sample frequency	Machine type	Fault mode	Task
CWRU	Accelerometer	2	12/48 kHz	Bearing	Artificial	Diagnostics
PHM09	Accelerometer	2	66.67 kHz	gearbox	Real	Diagnostics
IMS	Accelerometer/ Thermocouple	2/4	20 kHz	Bearing	Real	Diagnostics/ Prognostics
Paderborn	Accelerometer/ Current/ Thermocouple	1/2/1	64 kHz	Bearing	Artificial- accelerated aging	Diagnostics
PRONOSTIA	Accelerometer/ Thermocouple	2/1	25.6 kHz	Bearing	Real	Prognostics
C-MAPSS	Temperature/ Pressure/ Speed, etc.	21	-	Turbofan Engine	Simulation	Prognostics
DRIG	Accelerometer	2	51.2 kHz	Bearing	Artificial	Diagnostics
XJTU-SY	Accelerometer	2	25.6 kHz	Bearing	Artificial	Prognostics
MFPT	Accelerometer/ Thermocouple	-	50 kHz	Bearing	Artificial	Diagnostics

technique should be used. However, despite the importance of the issue, just a few studies have described it informally, without proposing a comprehensive definition, extensive analysis, or systematic treatment [190]. As a result, the questions of what characterizes the formulation of NT, what attributes contribute to NT, and how to prevent or minimize NT remain unresolved.

## 2) UNSUPERVISED MULTIPLE SOURCE TRANSFER

As shown in Fig. 7, the current achievements of TL are largely focused on inductive and transductive TL. However, collecting a labeled dataset with sufficient samples of each possible fault type is quite challenging in practice. Furthermore, in real-world applications, the unlabeled data might come from multiple sources, such as various heterogeneous, related machines [117]. For these reasons, unsupervised multiple TL, aimed at accomplishing maintenance tasks by leveraging heterogeneous, unlabeled sources and target data, seems to be the most promising and significant research direction in the very near future for dealing with industrial applications in practice. Hence, it is crucial to understand and identify which factors influence the success of a TL model when exploiting knowledge from multiple unlabeled heterogeneous datasets.

## 3) UNSEEN TARGET DOMAINS

Despite the fact that TL has demonstrated good performance in PdM, most of the developed models are only applicable to off-line PdM tasks due to the following assumption: It was assumed that, during the training process, target domain data must be readily available and adaptation models must be specially trained with target data before being executed on target machines. As a result, this assumption can restrict the adaptation of existing approaches to the unseen target domains [4]. To overcome this limitation, it is essential to focus on the more practical scenario in which the target domain data is inaccessible. This is compatible with real-time cross-domain PdM and has more applicability in real-world machines. One possible solution to the challenge is to generalize the maintenance knowledge of the several source domains to an unseen target domain based on the meta-learning technique [198].

## 4) DIGITAL TWINS

A Digital Twin (DT) consists of a virtual model that is constantly updated to reflect the status of its physical counterpart. This technique allows engineers to collect a huge amount of useful component run-to-failure data [199].

In case of the unavailability of historical fault data of physical equipment, the combination of TL and DT is a very promising approach worth investigating, as the target domain can be dynamically augmented with virtual model data [173]. Despite its potential, as shown in Fig. 12, a few studies focused on the TVRM scenario, where there is no consensus or consolidation on how DT might be employed in TL applications. Another challenge that needs to be tackled is how engineers can adaptively update the operation conditions of DT to the new operating conditions of the target machine.

## 5) CROSS-MODALITY TRANSFER LEARNING

Indeed, most of the TL approaches typically need some sort of feature space connection across the source and target domains, and knowledge transfer is only feasible whenever the source and target data (such as image, audio, text, and so on) are both in the same modality. However, within Industry 4.0, different data sources (such as operation and maintenance logs, sensor measurements, design papers, etc.) can be used, and these might also provide valuable information for the implementation of PdM models [9]. It follows that Cross-Modality Transfer Learning (CMTL) is a hot topic in PdM aiming to cope with scenarios where the feature spaces of both domains are entirely distinct, e.g., when transferring knowledge from text to image, or from audio to text. So, in the future, it will be critical to investigate how maintenance knowledge could be transferred across cross-modality spaces.

## 6) TRANSFER LEARNING FOR PROGNOSIS TASK

As illustrated in Fig. 14, when limiting to TL application, the majority of the current research in the PdM field aims to fault diagnosis. Only a few studies are focused on prognosis, and they provide very preliminary results. As a consequence, the use of TL technology for prediction, decision-making, and proper scheduling of the maintenance work is still a crucial open issue.

## VII. CONCLUSION

The goal of this paper has been to explore the main concepts, concerns, and potential of TL in the context of PdM. To that aim, an overview of PdM has been introduced and discussed together with the challenges arising from traditional ML and DL algorithms in real-world PdM applications, which motivated the need for TL. Through an SLR, we selected 168 recent journal publications that we further classified and analyzed. The taxonomy of the TL in the PdM context was first introduced with respect to three different problem settings: label, space, and source. Furthermore, the selected papers have been categorized and reviewed on the basis of the introduced taxonomy. Moreover, the PdM real-world applications that already have profited from TL have been clearly identified in this survey, as well as the open-source datasets introduced to assist the researchers in validating and comparing their approaches for cross-domain PdM challenges. Ultimately, open issues, challenges, and future research directions have been highlighted. We believe the SLR provided in

this paper represents a useful reference summarizing SOTA approaches as well as hints for developing and deploying TL for PdM in both academic research and industrial settings.

## REFERENCES

- [1] T. Zonta, C. A. da Costa, R. da Rosa Righi, M. J. de Lima, E. S. da Trindade, and G. P. Li, "Predictive maintenance in the industry 4.0: A systematic literature review," *Comput. Ind. Eng.*, vol. 150, Dec. 2020, Art. no. 106889.
- [2] T. P. Carvalho, F. A. A. M. N. Soares, R. Vita, R. D. P. Francisco, J. P. Basto, and S. G. S. Alcalá, "A systematic literature review of machine learning methods applied to predictive maintenance," *Comput. Ind. Eng.*, vol. 137, Nov. 2019, Art. no. 106024.
- [3] S. Niu, Y. Liu, J. Wang, and H. Song, "A decade survey of transfer learning (2010–2020)," *IEEE Trans. Artif. Intell.*, vol. 1, no. 2, pp. 151–166, Oct. 2020.
- [4] M. S. Azari, F. Flammini, and S. Santini, "Improving resilience in cyber-physical systems based on transfer learning," in *Proc. IEEE Int. Conf. Cyber Secur. Resilience (CSR)*, Jul. 2022, pp. 203–208.
- [5] J. Dalzochio, R. Kunst, E. Pignaton, A. Binotto, S. Sanyal, J. Favilla, and J. Barbosa, "Machine learning and reasoning for predictive maintenance in industry 4.0: Current status and challenges," *Comput. Ind.*, vol. 123, Dec. 2020, Art. no. 103298.
- [6] R. Zhao, R. Yan, Z. Chen, K. Mao, P. Wang, and R. X. Gao, "Deep learning and its applications to machine health monitoring," *Mech. Syst. Signal Process.*, vol. 115, pp. 213–237, Jan. 2019.
- [7] O. Fink, Q. Wang, M. Svensén, P. Dersin, W.-J. Lee, and M. Ducoffe, "Potential, challenges and future directions for deep learning in prognostics and health management applications," *Eng. Appl. Artif. Intell.*, vol. 92, Jun. 2020, Art. no. 103678.
- [8] Y. Lei, B. Yang, X. Jiang, F. Jia, N. Li, and A. K. Nandi, "Applications of machine learning to machine fault diagnosis: A review and roadmap," *Mech. Syst. Signal Process.*, vol. 138, Apr. 2020, Art. no. 106587.
- [9] J. J. M. Jimenez, S. Schwartz, R. Vingerhoeds, B. Grabot, and M. Salatiin, "Towards multi-model approaches to predictive maintenance: A systematic literature survey on diagnostics and prognostics," *J. Manuf. Syst.*, vol. 56, pp. 539–557, Jul. 2020.
- [10] A. T. Keleko, B. Kamsu-Foguem, R. H. Ngouna, and A. Tongne, "Artificial intelligence and real-time predictive maintenance in industry 4.0: A bibliometric analysis," *AI Ethics*, vol. 2, no. 4, pp. 1–25, 2022.
- [11] S. Zhang, S. Zhang, B. Wang, and T. G. Habetler, "Deep learning algorithms for bearing fault diagnostics—A comprehensive review," *IEEE Access*, vol. 8, pp. 29857–29881, 2020.
- [12] S. Sayyad, S. Kumar, A. Bongale, P. Kamat, S. Patil, and K. Kotecha, "Data-driven remaining useful life estimation for milling process: Sensors, algorithms, datasets, and future directions," *IEEE Access*, vol. 9, pp. 110255–110286, 2021.
- [13] R. Liu, B. Yang, E. Zio, and X. Chen, "Artificial intelligence for fault diagnosis of rotating machinery: A review," *Mech. Syst. Signal Process.*, vol. 108, pp. 33–47, Aug. 2018.
- [14] Y. Lei, N. Li, L. Guo, N. Li, T. Yan, and J. Lin, "Machinery health prognostics: A systematic review from data acquisition to RUL prediction," *Mech. Syst. Signal Process.*, vol. 104, pp. 799–834, May 2018.
- [15] C. Li, S. Zhang, Y. Qin, and E. Estupinan, "A systematic review of deep transfer learning for machinery fault diagnosis," *Neurocomputing*, vol. 407, pp. 121–135, Sep. 2020.
- [16] H. Zheng, R. Wang, Y. Yang, J. Yin, Y. Li, Y. Li, and M. Xu, "Cross-domain fault diagnosis using knowledge transfer strategy: A review," *IEEE Access*, vol. 7, pp. 129260–129290, 2019.
- [17] A. Kumar, R. B. Chinnam, and F. Tseng, "An HMM and polynomial regression based approach for remaining useful life and health state estimation of cutting tools," *Comput. Ind. Eng.*, vol. 128, pp. 1008–1014, Feb. 2019.
- [18] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, Oct. 2009.
- [19] F. Zhuang, Z. Qi, K. Duan, D. Xi, Y. Zhu, H. Zhu, H. Xiong, and Q. He, "A comprehensive survey on transfer learning," *Proc. IEEE*, vol. 109, no. 1, pp. 43–76, Jul. 2020.
- [20] S. Rajabi, M. S. Azari, S. Santini, and F. Flammini, "Fault diagnosis in industrial rotating equipment based on permutation entropy, signal processing and multi-output neuro-fuzzy classifier," *Expert Syst. Appl.*, vol. 206, Nov. 2022, Art. no. 117754.

- [21] M. S. Azari, F. Flammini, M. Caporuscio, and S. Santini, "Data-driven fault diagnosis of once-through Benson boilers," in *Proc. 4th Int. Conf. Syst. Rel. Saf. (ICSRS)*, Nov. 2019, pp. 345–354.
- [22] M. Ranaweera and Q. H. Mahmoud, "Virtual to real-world transfer learning: A systematic review," *Electronics*, vol. 10, no. 12, p. 1491, Jun. 2021.
- [23] B. Kitchenham, O. P. Brereton, D. Budgen, M. Turner, J. Bailey, and S. Linkman, "Systematic literature reviews in software engineering—A systematic literature review," *Inf. Softw. Technol.*, vol. 51, no. 1, pp. 7–15, 2009.
- [24] J. F. Burnham, "Scopus database: A review," *Biomed. Digit. Libraries*, vol. 3, no. 1, pp. 1–8, Dec. 2006.
- [25] J. Liu and Y. Ren, "A general transfer framework based on industrial process fault diagnosis under small samples," *IEEE Trans. Ind. Informat.*, vol. 17, no. 9, pp. 6073–6083, Sep. 2020.
- [26] K. Lee, S. Han, V. H. Pham, S. Cho, H.-J. Choi, J. Lee, I. Noh, and S. W. Lee, "Multi-objective instance weighting-based deep transfer learning network for intelligent fault diagnosis," *Appl. Sci.*, vol. 11, no. 5, p. 2370, Mar. 2021.
- [27] Z. Zhang, M. Shao, L. Wang, S. Shao, and C. Ma, "A novel domain adaptation-based intelligent fault diagnosis model to handle sample class imbalanced problem," *Sensors*, vol. 21, no. 10, p. 3382, May 2021.
- [28] X. Wang, J. Ren, and S. Liu, "Distribution adaptation and manifold alignment for complex processes fault diagnosis," *Knowl.-Based Syst.*, vol. 156, pp. 100–112, Sep. 2018.
- [29] Z. Tong, W. Li, B. Zhang, F. Jiang, and G. B. Zhou, "Bearing fault diagnosis under variable working conditions based on domain adaptation using feature transfer learning," *IEEE Access*, vol. 6, pp. 76187–76197, 2018.
- [30] W. Qian, S. Li, and J. Wang, "A new transfer learning method and its application on rotating machine fault diagnosis under variant working conditions," *IEEE Access*, vol. 6, pp. 69907–69917, 2018.
- [31] B. Zhang, W. Li, X.-L. Li, and S.-K. Ng, "Intelligent fault diagnosis under varying working conditions based on domain adaptive convolutional neural networks," *IEEE Access*, vol. 6, pp. 66367–66384, 2018.
- [32] J. Guo, J. Wu, Z. Sun, J. Long, and S. Zhang, "Fault diagnosis of delta 3D printers using transfer support vector machine with attitude signals," *IEEE Access*, vol. 7, pp. 40359–40368, 2019.
- [33] D. Xiao, Y. Huang, L. Zhao, C. Qin, H. Shi, and C. Liu, "Domain adaptive motor fault diagnosis using deep transfer learning," *IEEE Access*, vol. 7, pp. 80937–80949, 2019.
- [34] S. Pang and X. Yang, "A cross-domain stacked denoising autoencoders for rotating machinery fault diagnosis under different working conditions," *IEEE Access*, vol. 7, pp. 77277–77292, 2019.
- [35] W. Qian, S. Li, and X. Jiang, "Deep transfer network for rotating machine fault analysis," *Pattern Recognit.*, vol. 96, Dec. 2019, Art. no. 106993.
- [36] L. Wen, L. Gao, and X. Li, "A new deep transfer learning based on sparse auto-encoder for fault diagnosis," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 49, no. 1, pp. 136–144, Jan. 2019.
- [37] X. Li, W. Zhang, Q. Ding, and J.-Q. Sun, "Multi-layer domain adaptation method for rolling bearing fault diagnosis," *Signal Process.*, vol. 157, pp. 180–197, Apr. 2019.
- [38] A. Zhang and X. Gao, "Supervised dictionary-based transfer subspace learning and applications for fault diagnosis of sucker rod pumping systems," *Neurocomputing*, vol. 338, pp. 293–306, Apr. 2019.
- [39] B. Yang, Y. Lei, F. Jia, and S. Xing, "An intelligent fault diagnosis approach based on transfer learning from laboratory bearings to locomotive bearings," *Mech. Syst. Signal Process.*, vol. 122, pp. 692–706, May 2019.
- [40] Z. An, S. Li, J. Wang, Y. Xin, and K. Xu, "Generalization of deep neural network for bearing fault diagnosis under different working conditions using multiple kernel method," *Neurocomputing*, vol. 352, pp. 42–53, Aug. 2019.
- [41] Y. Yu, C. Zhang, Y. Li, and Y. Li, "A new transfer learning fault diagnosis method using TSC and JGSA under variable condition," *IEEE Access*, vol. 8, pp. 177287–177295, 2020.
- [42] Z. Xu, H. Darong, G. Sun, and W. Yongchao, "A fault diagnosis method based on improved adaptive filtering and joint distribution adaptation," *IEEE Access*, vol. 8, pp. 159683–159695, 2020.
- [43] W. Xu, Y. Wan, T.-Y. Zuo, and X.-M. Sha, "Transfer learning based data feature transfer for fault diagnosis," *IEEE Access*, vol. 8, pp. 76120–76129, 2020.
- [44] C. Zheng, X. Wang, Y. Hao, K. Wang, and X. Xiong, "Normalized recurrent dynamic adaption network: A new framework with dynamic alignment for intelligent fault diagnosis," *IEEE Access*, vol. 8, pp. 80243–80255, 2020.
- [45] Z. Wu, S. Yu, X. Zhu, Y. Ji, and M. Pecht, "A weighted deep domain adaptation method for industrial fault prognostics according to prior distribution of complex working conditions," *IEEE Access*, vol. 7, pp. 139802–139814, 2019.
- [46] Y. Tian, Y. Tang, and X. Peng, "Cross-task fault diagnosis based on deep domain adaptation with local feature learning," *IEEE Access*, vol. 8, pp. 127546–127559, 2020.
- [47] H. Zheng, R. Wang, J. Yin, Y. Li, H. Lu, and M. Xu, "A new intelligent fault identification method based on transfer locality preserving projection for actual diagnosis scenario of rotating machinery," *Mech. Syst. Signal Process.*, vol. 135, Jan. 2020, Art. no. 106344.
- [48] Z. An, S. Li, X. Jiang, Y. Xin, and J. Wang, "Adaptive cross-domain feature extraction method and its application on machinery intelligent fault diagnosis under different working conditions," *IEEE Access*, vol. 8, pp. 535–546, 2020.
- [49] S. Wu, X.-Y. Jing, Q. Zhang, F. Wu, H. Zhao, and Y. Dong, "Prediction consistency guided convolutional neural networks for cross-domain bearing fault diagnosis," *IEEE Access*, vol. 8, pp. 120089–120103, 2020.
- [50] 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 Trans.*, vol. 97, pp. 269–281, Feb. 2020.
- [51] J. Chen, J. Wang, J. Zhu, T. H. Lee, and C. W. de Silva, "Unsupervised cross-domain fault diagnosis using feature representation alignment networks for rotating machinery," *IEEE/ASME Trans. Mechatronics*, vol. 26, no. 5, pp. 2770–2781, Oct. 2020.
- [52] H. Wang, X. Bai, J. Tan, and J. Yang, "Deep prototypical networks based domain adaptation for fault diagnosis," *J. Intell. Manuf.*, vol. 33, pp. 973–983, 2020.
- [53] Z. Zhang, H. Chen, S. Li, Z. An, and J. Wang, "A novel geodesic flow kernel based domain adaptation approach for intelligent fault diagnosis under varying working condition," *Neurocomputing*, vol. 376, pp. 54–64, Feb. 2020.
- [54] K. Zhao, H. Jiang, Z. Wu, and T. Lu, "A novel transfer learning fault diagnosis method based on manifold embedded distribution alignment with a little labeled data," *J. Intell. Manuf.*, vol. 33, pp. 151–165, 2022.
- [55] X. Li, X.-D. Jia, W. Zhang, H. Ma, Z. Luo, and X. Li, "Intelligent cross-machine fault diagnosis approach with deep auto-encoder and domain adaptation," *Neurocomputing*, vol. 383, pp. 235–247, Mar. 2020.
- [56] J. Zhu, N. Chen, and C. Shen, "A new data-driven transferable remaining useful life prediction approach for bearing under different working conditions," *Mech. Syst. Signal Process.*, vol. 139, May 2020, Art. no. 106602.
- [57] C. Hu, Y. Wang, and J. Gu, "Cross-domain intelligent fault classification of bearings based on tensor-aligned invariant subspace learning and two-dimensional convolutional neural networks," *Knowl.-Based Syst.*, vol. 209, Dec. 2020, Art. no. 106214.
- [58] Z. Zhang, H. Chen, S. Li, and Z. An, "Sparse filtering based domain adaptation for mechanical fault diagnosis," *Neurocomputing*, vol. 393, pp. 101–111, Jun. 2020.
- [59] P. Ma, H. Zhang, W. Fan, and C. Wang, "A diagnosis framework based on domain adaptation for bearing fault diagnosis across diverse domains," *ISA Trans.*, vol. 99, pp. 465–478, Apr. 2020.
- [60] X. Wang and J. Zhao, "A complex process fault diagnosis method based on manifold distribution adaptation," *Eng. Appl. Artif. Intell.*, vol. 87, Jan. 2020, Art. no. 103267.
- [61] W. Mao, D. Zhang, S. Tian, and J. Tang, "Robust detection of bearing early fault based on deep transfer learning," *Electronics*, vol. 9, no. 2, p. 323, Feb. 2020.
- [62] M. Azamfar, X. Li, and J. Lee, "Intelligent ball screw fault diagnosis using a deep domain adaptation methodology," *Mech. Mach. Theory*, vol. 151, Sep. 2020, Art. no. 103932.
- [63] J. Jiao, M. Zhao, J. Lin, and C. Ding, "Classifier inconsistency-based domain adaptation network for partial transfer intelligent diagnosis," *IEEE Trans. Ind. Informat.*, vol. 16, no. 9, pp. 5965–5974, Sep. 2019.
- [64] Y. Wang, D. Wu, and X. Yuan, "LDA-based deep transfer learning for fault diagnosis in industrial chemical processes," *Comput. Chem. Eng.*, vol. 140, Sep. 2020, Art. no. 106964.
- [65] Y. Qin, X. Wang, Q. Qian, H. Pu, and J. Luo, "Multiscale transfer voting mechanism: A new strategy for domain adaption," *IEEE Trans. Ind. Informat.*, vol. 17, no. 10, pp. 7103–7113, Oct. 2020.



- [66] J. Wu, T. Tang, M. Chen, Y. Wang, and K. Wang, "A study on adaptation lightweight architecture based deep learning models for bearing fault diagnosis under varying working conditions," *Expert Syst. Appl.*, vol. 160, Dec. 2020, Art. no. 113710.
- [67] X. Cao, B. Chen, and N. Zeng, "A deep domain adaption model with multi-task networks for planetary gearbox fault diagnosis," *Neurocomputing*, vol. 409, pp. 173–190, Oct. 2020.
- [68] T. Wen, R. Chen, and L. Tang, "Fault diagnosis of rolling bearings of different working conditions based on multi-feature spatial domain adaptation," *IEEE Access*, vol. 9, pp. 52404–52413, 2021.
- [69] G. Xiong, W. Ma, N. Zhao, J. Zhang, Z. Jiang, and Z. Mao, "Multi-type diesel engines operating condition recognition method based on stacked auto-encoder and feature transfer learning," *IEEE Access*, vol. 9, pp. 31043–31052, 2021.
- [70] M. Lv, S. Liu, X. Su, and C. Chen, "Deep transfer network with multi-kernel dynamic distribution adaptation for cross-machine fault diagnosis," *IEEE Access*, vol. 9, pp. 16392–16409, 2021.
- [71] Z. Dangdang, B. Han, G. Liu, Y. Li, and H. Yu, "Cross-domain intelligent fault diagnosis method of rotating machinery using multi-scale transfer fuzzy entropy," *IEEE Access*, vol. 9, pp. 95481–95492, 2021.
- [72] M. Liu, X. Li, K. Chakrabarty, and X. Gu, "Knowledge transfer in board-level functional fault diagnosis enabled by domain adaptation," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 41, no. 3, pp. 762–775, Mar. 2021.
- [73] S. Jia, J. Wang, X. Zhang, and B. Han, "A weighted subdomain adaptation network for partial transfer fault diagnosis of rotating machinery," *Entropy*, vol. 23, no. 4, p. 424, Apr. 2021.
- [74] 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. Ind. Informat.*, vol. 17, no. 4, pp. 2833–2841, Apr. 2020.
- [75] X. Li, Y. Hu, J. Zheng, M. Li, and W. Ma, "Central moment discrepancy based domain adaptation for intelligent bearing fault diagnosis," *Neurocomputing*, vol. 429, pp. 12–24, Mar. 2021.
- [76] Z. Li, R. Cai, H. W. Ng, M. Winslett, T. Z. J. Fu, B. Xu, X. Yang, and Z. Zhang, "Causal mechanism transfer network for time series domain adaptation in mechanical systems," *ACM Trans. Intell. Syst. Technol.*, vol. 12, no. 2, pp. 1–21, Apr. 2021.
- [77] W. Qian, S. Li, T. Yao, and K. Xu, "Discriminative feature-based adaptive distribution alignment (DFADA) for rotating machine fault diagnosis under variable working conditions," *Appl. Soft Comput.*, vol. 99, Feb. 2021, Art. no. 106886.
- [78] B. Yang, C.-G. Lee, Y. Lei, N. Li, and N. Lu, "Deep partial transfer learning network: A method to selectively transfer diagnostic knowledge across related machines," *Mech. Syst. Signal Process.*, vol. 156, Jul. 2021, Art. no. 107618.
- [79] X. Cao, Y. Wang, B. Chen, and N. Zeng, "Domain-adaptive intelligence for fault diagnosis based on deep transfer learning from scientific test rigs to industrial applications," *Neural Comput. Appl.*, vol. 33, no. 9, pp. 4483–4499, May 2021.
- [80] R. van de Sand, S. Corasaniti, and J. Reiff-Stephan, "Data-driven fault diagnosis for heterogeneous chillers using domain adaptation techniques," *Control Eng. Pract.*, vol. 112, Jul. 2021, Art. no. 104815.
- [81] H. Wu and J. Zhao, "Fault detection and diagnosis based on transfer learning for multimode chemical processes," *Comput. Chem. Eng.*, vol. 135, Apr. 2020, Art. no. 106731.
- [82] C. Shi, Z. Wu, X. Lv, and Y. Ji, "DGTL-Net: A deep generative transfer learning network for fault diagnostics on new hard disks," *Expert Syst. Appl.*, vol. 169, May 2021, Art. no. 114379.
- [83] S. Lu, T. Sirojan, B. T. Phung, D. Zhang, and E. Ambikairajah, "DA-DCGAN: An effective methodology for DC series arc fault diagnosis in photovoltaic systems," *IEEE Access*, vol. 7, pp. 45831–45840, 2019.
- [84] M. Zhang, D. Wang, W. Lu, J. Yang, Z. Li, and B. Liang, "A deep transfer model with Wasserstein distance guided multi-adversarial networks for bearing fault diagnosis under different working conditions," *IEEE Access*, vol. 7, pp. 65303–65318, 2019.
- [85] S. Jia, J. Wang, B. Han, G. Zhang, X. Wang, and J. He, "A novel transfer learning method for fault diagnosis using maximum classifier discrepancy with marginal probability distribution adaptation," *IEEE Access*, vol. 8, pp. 71475–71485, 2020.
- [86] J. Shao, Z. Huang, and J. Zhu, "Transfer learning method based on adversarial domain adaption for bearing fault diagnosis," *IEEE Access*, vol. 8, pp. 119421–119430, 2020.
- [87] J. Wang, S. Ji, B. Han, H. Bao, and X. Jiang, "Deep adaptive adversarial network-based method for mechanical fault diagnosis under different working conditions," *Complexity*, vol. 2020, pp. 1–11, Jul. 2020.
- [88] M. Marei, S. E. Zaatari, and W. Li, "Transfer learning enabled convolutional neural networks for estimating health state of cutting tools," *Robot. Computer-Integrated Manuf.*, vol. 71, Oct. 2021, Art. no. 102145.
- [89] D. She, M. Jia, and M. G. Pecht, "Weighted entropy minimization based deep conditional adversarial diagnosis approach under variable working conditions," *IEEE/ASME Trans. Mechatronics*, vol. 26, no. 5, pp. 2440–2450, Oct. 2020.
- [90] X. Li, W. Zhang, Q. Ding, and X. Li, "Diagnosing rotating machines with weakly supervised data using deep transfer learning," *IEEE Trans. Ind. Informat.*, vol. 16, no. 3, pp. 1688–1697, Mar. 2019.
- [91] Z. Wu, H. Jiang, T. Lu, and K. Zhao, "A deep transfer maximum classifier discrepancy method for rolling bearing fault diagnosis under few labeled data," *Knowl.-Based Syst.*, vol. 196, May 2020, Art. no. 105814.
- [92] X. Li, W. Zhang, H. Ma, Z. Luo, and X. Li, "Deep learning-based adversarial multi-classifier optimization for cross-domain machinery fault diagnostics," *J. Manuf. Syst.*, vol. 55, pp. 334–347, Apr. 2020.
- [93] 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, Aug. 2020, Art. no. 106904.
- [94] W. Wu, M. Peng, W. Chen, and S. Yan, "Unsupervised deep transfer learning for fault diagnosis in fog radio access networks," *IEEE Internet Things J.*, vol. 7, no. 9, pp. 8956–8966, Sep. 2020.
- [95] X. Li, W. Zhang, H. Ma, Z. Luo, and X. Li, "Partial transfer learning in machinery cross-domain fault diagnostics using class-weighted adversarial networks," *Neural Netw.*, vol. 129, pp. 313–322, Sep. 2020.
- [96] C. Cheng, B. Zhou, G. Ma, D. Wu, and Y. Yuan, "Wasserstein distance based deep adversarial transfer learning for intelligent fault diagnosis with unlabeled or insufficient labeled data," *Neurocomputing*, vol. 409, pp. 35–45, Oct. 2020.
- [97] J. Jiao, J. Lin, M. Zhao, and K. Liang, "Double-level adversarial domain adaptation network for intelligent fault diagnosis," *Knowl.-Based Syst.*, vol. 205, Oct. 2020, Art. no. 106236.
- [98] J. Jiao, M. Zhao, J. Lin, and K. Liang, "Residual joint adaptation adversarial network for intelligent transfer fault diagnosis," *Mech. Syst. Signal Process.*, vol. 145, Nov. 2020, Art. no. 106962.
- [99] W. Zhang and X. Li, "Federated transfer learning for intelligent fault diagnostics using deep adversarial networks with data privacy," *IEEE/ASME Trans. Mechatronics*, vol. 27, no. 1, pp. 430–439, Feb. 2022.
- [100] B. Zhao, X. Zhang, Z. Zhan, and Q. Wu, "Deep multi-scale adversarial network with attention: A novel domain adaptation method for intelligent fault diagnosis," *J. Manuf. Syst.*, vol. 59, pp. 565–576, Apr. 2021.
- [101] W. Li, Z. Chen, and G. He, "A novel weighted adversarial transfer network for partial domain fault diagnosis of machinery," *IEEE Trans. Ind. Informat.*, vol. 17, no. 3, pp. 1753–1762, Mar. 2020.
- [102] W. Mao, L. Ding, Y. Liu, S. S. Afshari, and X. Liang, "A new deep domain adaptation method with joint adversarial training for online detection of bearing early fault," *ISA Trans.*, vol. 122, pp. 444–458, Mar. 2021.
- [103] J. Jiao, M. Zhao, J. Lin, K. Liang, and C. Ding, "A mixed adversarial adaptation network for intelligent fault diagnosis," *J. Intell. Manuf.*, pp. 1–16, Dec. 2021.
- [104] G. Michau and O. Fink, "Unsupervised transfer learning for anomaly detection: Application to complementary operating condition transfer," *Knowl.-Based Syst.*, vol. 216, Mar. 2021, Art. no. 106816.
- [105] J. Zhu, N. Chen, and C. Shen, "A new multiple source domain adaptation fault diagnosis method between different rotating machines," *IEEE Trans. Ind. Informat.*, vol. 17, no. 7, pp. 4788–4797, Jul. 2021.
- [106] M. Ragab, Z. Chen, M. Wu, C. S. Foo, C. K. Kwok, R. Yan, and X. Li, "Contrastive adversarial domain adaptation for machine remaining useful life prediction," *IEEE Trans. Ind. Informat.*, vol. 17, no. 8, pp. 5239–5249, Oct. 2020.
- [107] K. Zhao, H. Jiang, K. Wang, and Z. Pei, "Joint distribution adaptation network with adversarial learning for rolling bearing fault diagnosis," *Knowl.-Based Syst.*, vol. 222, Jun. 2021, Art. no. 106974.
- [108] Y. Feng, J. Chen, Z. Yang, X. Song, Y. Chang, S. He, E. Xu, and Z. Zhou, "Similarity-based meta-learning network with adversarial domain adaptation for cross-domain fault identification," *Knowl.-Based Syst.*, vol. 217, Apr. 2021, Art. no. 106829.

- [109] Y. Deng, D. Huang, S. Du, G. Li, C. Zhao, and J. Lv, "A double-layer attention based adversarial network for partial transfer learning in machinery fault diagnosis," *Comput. Ind.*, vol. 127, May 2021, Art. no. 103399.
- [110] D. Wei, T. Han, F. Chu, and M. J. Zuo, "Weighted domain adaptation networks for machinery fault diagnosis," *Mech. Syst. Signal Process.*, vol. 158, Sep. 2021, Art. no. 107744.
- [111] Y. Qin, Q. Yao, Y. Wang, and Y. Mao, "Parameter sharing adversarial domain adaptation networks for fault transfer diagnosis of planetary gearboxes," *Mech. Syst. Signal Process.*, vol. 160, Nov. 2021, Art. no. 107936.
- [112] X. Li, W. Zhang, and Q. Ding, "A robust intelligent fault diagnosis method for rolling element bearings based on deep distance metric learning," *Neurocomputing*, vol. 310, pp. 77–95, Oct. 2018.
- [113] B. Zhao, X. Zhang, Z. Zhan, and Q. Wu, "Deep multi-scale separable convolutional network with triple attention mechanism: A novel multi-task domain adaptation method for intelligent fault diagnosis," *Expert Syst. Appl.*, vol. 182, Nov. 2021, Art. no. 115087.
- [114] Z. Chen, J. Wu, C. Deng, C. Wang, and Y. Wang, "Residual deep subdomain adaptation network: A new method for intelligent fault diagnosis of bearings across multiple domains," *Mech. Mach. Theory*, vol. 169, Mar. 2022, Art. no. 104635.
- [115] Y. Xia, C. Shen, D. Wang, Y. Shen, W. Huang, and Z. Zhu, "Moment matching-based intraclass multisource domain adaptation network for bearing fault diagnosis," *Mech. Syst. Signal Process.*, vol. 168, Apr. 2022, Art. no. 108697.
- [116] 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," *Mech. Syst. Signal Process.*, vol. 164, Feb. 2022, Art. no. 108219.
- [117] B. Yang, S. Xu, Y. Lei, C.-G. Lee, E. Stewart, and C. Roberts, "Multisource transfer learning network to complement knowledge for intelligent diagnosis of machines with unseen faults," *Mech. Syst. Signal Process.*, vol. 162, Jan. 2022, Art. no. 108095.
- [118] X. Yu, Z. Zhao, X. Zhang, Q. Zhang, Y. Liu, C. Sun, and X. Chen, "Deep-learning-based open set fault diagnosis by extreme value theory," *IEEE Trans. Ind. Informat.*, vol. 18, no. 1, pp. 185–196, Jan. 2021.
- [119] C. Yang, M. Jia, Z. Li, and M. Gabbouj, "Enhanced hierarchical symbolic dynamic entropy and maximum mean and covariance discrepancy-based transfer joint matching with welsch loss for intelligent cross-domain bearing health monitoring," *Mech. Syst. Signal Process.*, vol. 165, Feb. 2022, Art. no. 108343.
- [120] S. Schwendemann, Z. Amjad, and A. Sikora, "Bearing fault diagnosis with intermediate domain based layered maximum mean discrepancy: A new transfer learning approach," *Eng. Appl. Artif. Intell.*, vol. 105, Oct. 2021, Art. no. 104415.
- [121] F. Shen and R. Yan, "A new intermediate-domain SVM-based transfer model for rolling bearing RUL prediction," *IEEE/ASME Trans. Mechatronics*, vol. 27, no. 3, pp. 1357–1369, Jun. 2021.
- [122] N. Lu, H. Hu, T. Yin, Y. Lei, and S. Wang, "Transfer relation network for fault diagnosis of rotating machinery with small data," *IEEE Trans. Cybern.*, vol. 52, no. 11, pp. 11927–11941, Nov. 2021.
- [123] P. Xia, Y. Huang, P. Li, C. Liu, and L. Shi, "Fault knowledge transfer assisted ensemble method for remaining useful life prediction," *IEEE Trans. Ind. Informat.*, vol. 18, no. 3, pp. 1758–1769, Mar. 2021.
- [124] Z. Chai, C. Zhao, and B. Huang, "Multisource-refined transfer network for industrial fault diagnosis under domain and category inconsistencies," *IEEE Trans. Cybern.*, vol. 52, no. 9, pp. 9784–9796, Sep. 2021.
- [125] J. Li, R. Huang, G. He, Y. Liao, Z. Wang, and W. Li, "A two-stage transfer adversarial network for intelligent fault diagnosis of rotating machinery with multiple new faults," *IEEE/ASME Trans. Mechatronics*, vol. 26, no. 3, pp. 1591–1601, Jun. 2020.
- [126] Z.-H. Liu, B.-L. Lu, H.-L. Wei, L. Chen, X.-H. Li, and M. Ratsch, "Deep adversarial domain adaptation model for bearing fault diagnosis," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 51, no. 7, pp. 4217–4226, Jul. 2019.
- [127] Z. An, X. Jiang, J. Cao, R. Yang, and X. Li, "Self-learning transferable neural network for intelligent fault diagnosis of rotating machinery with unlabeled and imbalanced data," *Knowl.-Based Syst.*, vol. 230, Oct. 2021, Art. no. 107374.
- [128] T. Kim and J. Chai, "Pre-processing method to improve cross-domain fault diagnosis for bearing," *Sensors*, vol. 21, no. 15, p. 4970, Jul. 2021.
- [129] Z. Sun, X. Yuan, X. Fu, F. Zhou, and C. Zhang, "Multi-scale capsule attention network and joint distributed optimal transport for bearing fault diagnosis under different working loads," *Sensors*, vol. 21, no. 19, p. 6696, Oct. 2021.
- [130] K. Xu, S. Li, R. Li, J. Lu, X. Li, and M. Zeng, "Domain adaptation network with double adversarial mechanism for intelligent fault diagnosis," *Appl. Sci.*, vol. 11, no. 17, p. 7983, Aug. 2021.
- [131] G.-B. Jang and S.-B. Cho, "Feature space transformation for fault diagnosis of rotating machinery under different working conditions," *Sensors*, vol. 21, no. 4, p. 1417, Feb. 2021.
- [132] Y. Feng, J. Chen, S. He, T. Pan, and Z. Zhou, "Globally localized multi-source domain adaptation for cross-domain fault diagnosis with category shift," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Sep. 21, 2021, doi: 10.1109/TNNLS.2021.3111732.
- [133] Z. Yang, R. Yang, and M. Huang, "Rolling bearing incipient fault diagnosis method based on improved transfer learning with hybrid feature extraction," *Sensors*, vol. 21, no. 23, p. 7894, Nov. 2021.
- [134] Z. Yan, G. Liu, J. Wang, H. Bao, Z. Zhang, X. Zhang, and B. Han, "A new universal domain adaptive method for diagnosing unknown bearing faults," *Entropy*, vol. 23, no. 8, p. 1052, Aug. 2021.
- [135] Z. Zhu, L. Wang, G. Peng, and S. Li, "WDA: An improved Wasserstein distance-based transfer learning fault diagnosis method," *Sensors*, vol. 21, no. 13, p. 4394, Jun. 2021.
- [136] J. Shim, S. Cho, E. Kum, and S. Jeong, "Adaptive fault detection framework for recipe transition in semiconductor manufacturing," *Comput. Ind. Eng.*, vol. 161, Nov. 2021, Art. no. 107632.
- [137] T. Jin, C. Yan, C. Chen, Z. Yang, H. Tian, and J. Guo, "New domain adaptation method in shallow and deep layers of the CNN for bearing fault diagnosis under different working conditions," *Int. J. Adv. Manuf. Technol.*, vol. 124, pp. 3701–3712, 2021.
- [138] H. Cheng, X. Kong, Q. Wang, H. Ma, S. Yang, and G. Chen, "Deep transfer learning based on dynamic domain adaptation for remaining useful life prediction under different working conditions," *J. Intell. Manuf.*, vol. 34, pp. 587–613, 2021.
- [139] C. Liu and K. Gryllias, "Simulation-driven domain adaptation for rolling element bearing fault diagnosis," *IEEE Trans. Ind. Informat.*, vol. 18, no. 9, pp. 5760–5770, Sep. 2021.
- [140] P. Chen, R. Zhao, T. He, K. Wei, and Q. Yang, "Unsupervised domain adaptation of bearing fault diagnosis based on joint sliced Wasserstein distance," *ISA Trans.*, vol. 129, pp. 504–519, Oct. 2022.
- [141] M. Deng, A. Deng, Y. Shi, Y. Liu, and M. Xu, "Intelligent fault diagnosis based on sample weighted joint adversarial network," *Neurocomputing*, vol. 488, pp. 168–182, Jun. 2022.
- [142] J. Tian, D. Han, M. Li, and P. Shi, "A multi-source information transfer learning method with subdomain adaptation for cross-domain fault diagnosis," *Knowl.-Based Syst.*, vol. 243, May 2022, Art. no. 108466.
- [143] 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. Informat.*, vol. 18, no. 11, pp. 8077–8086, Nov. 2021.
- [144] Y. Ren, J. Liu, H. Zhang, and J. Wang, "TBDA-Net: A task-based bias domain adaptation network under industrial small samples," *IEEE Trans. Ind. Informat.*, vol. 18, no. 9, pp. 6109–6119, Sep. 2022.
- [145] Q. Qian, Y. Qin, J. Luo, and S. Wang, "Partial transfer fault diagnosis by multiscale weight-selection adversarial network," *IEEE/ASME Trans. Mechatronics*, vol. 27, no. 6, pp. 4798–4806, Dec. 2022.
- [146] R. Zhang, H. Tao, L. Wu, and Y. Guan, "Transfer learning with neural networks for bearing fault diagnosis in changing working conditions," *IEEE Access*, vol. 5, pp. 14347–14357, 2017.
- [147] M. J. Hasan and J.-M. Kim, "Bearing fault diagnosis under variable rotational speeds using stockwell transform-based vibration imaging and transfer learning," *Appl. Sci.*, vol. 8, no. 12, p. 2357, Nov. 2018.
- [148] Y. Xue and P. Beausery, "Transfer learning for one class SVM adaptation to limited data distribution change," *Pattern Recognit. Lett.*, vol. 100, pp. 117–123, Dec. 2017.
- [149] A. Zhang, H. Wang, S. Li, Y. Cui, Z. Liu, G. Yang, and J. Hu, "Transfer learning with deep recurrent neural networks for remaining useful life estimation," *Appl. Sci.*, vol. 8, no. 12, p. 2416, 2018.
- [150] H. Kim and B. D. Youn, "A new parameter repurposing method for parameter transfer with small dataset and its application in fault diagnosis of rolling element bearings," *IEEE Access*, vol. 7, pp. 46917–46930, 2019.

- [151] P. Cao, S. Zhang, and J. Tang, "Preprocessing-free gear fault diagnosis using small datasets with deep convolutional neural network-based transfer learning," *IEEE Access*, vol. 6, pp. 26241–26253, 2018.
- [152] X. Li, H. Jiang, K. Zhao, and R. Wang, "A deep transfer nonnegativity-constraint sparse autoencoder for rolling bearing fault diagnosis with few labeled data," *IEEE Access*, vol. 7, pp. 91216–91224, 2019.
- [153] Y. Xu, Y. Sun, X. Liu, and Y. Zheng, "A digital-twin-assisted fault diagnosis using deep transfer learning," *IEEE Access*, vol. 7, pp. 19990–19999, 2019.
- [154] S. Shao, S. McAleer, R. Yan, and P. Baldi, "Highly accurate machine fault diagnosis using deep transfer learning," *IEEE Trans. Ind. Informat.*, vol. 15, no. 4, pp. 2446–2455, Apr. 2019.
- [155] Q. Liu and C. Huang, "A fault diagnosis method based on transfer convolutional neural networks," *IEEE Access*, vol. 7, pp. 171423–171430, 2019.
- [156] M. Alfarhan, M. Deriche, and A. Maalej, "Robust concurrent detection of salt domes and faults in seismic surveys using an improved UNet architecture," *IEEE Access*, vol. 10, pp. 39424–39435, 2020.
- [157] T. Han, C. Liu, W. Yang, and D. Jiang, "Learning transferable features in deep convolutional neural networks for diagnosing unseen machine conditions," *ISA Trans.*, vol. 93, pp. 341–353, Oct. 2019.
- [158] T. Lu, F. Yu, B. Han, and J. Wang, "A generic intelligent bearing fault diagnosis system using convolutional neural networks with transfer learning," *IEEE Access*, vol. 8, pp. 164807–164814, 2020.
- [159] J. Li and D. He, "A Bayesian optimization AdaBN-DCNN method with self-optimized structure and hyperparameters for domain adaptation remaining useful life prediction," *IEEE Access*, vol. 8, pp. 41482–41501, 2020.
- [160] Y. Zhang, X. Wang, J. He, Y. Xu, F. Zhang, and Y. Luo, "A transfer learning-based high impedance fault detection method under a cloud-edge collaboration framework," *IEEE Access*, vol. 8, pp. 165099–165110, 2020.
- [161] Z. Chen, K. Gryllias, and W. Li, "Intelligent fault diagnosis for rotary machinery using transferable convolutional neural network," *IEEE Trans. Ind. Informat.*, vol. 16, no. 1, pp. 339–349, Jan. 2019.
- [162] C. Che, H. Wang, X. Ni, and Q. Fu, "Domain adaptive deep belief network for rolling bearing fault diagnosis," *Comput. Ind. Eng.*, vol. 143, May 2020, Art. no. 106427.
- [163] F. Yang, W. Zhang, L. Tao, and J. Ma, "Transfer learning strategies for deep learning-based PHM algorithms," *Appl. Sci.*, vol. 10, no. 7, p. 2361, Mar. 2020.
- [164] B. Zhao, X. Zhang, Z. Zhan, and S. Pang, "Deep multi-scale convolutional transfer learning network: A novel method for intelligent fault diagnosis of rolling bearings under variable working conditions and domains," *Neurocomputing*, vol. 407, pp. 24–38, Sep. 2020.
- [165] L. Wen, X. Li, and L. Gao, "A transfer convolutional neural network for fault diagnosis based on ResNet-50," *Neural Comput. Appl.*, vol. 32, no. 10, pp. 6111–6124, May 2020.
- [166] Z. He, H. Shao, X. Zhong, and X. Zhao, "Ensemble transfer CNNs driven by multi-channel signals for fault diagnosis of rotating machinery cross working conditions," *Knowl.-Based Syst.*, vol. 207, Nov. 2020, Art. no. 106396.
- [167] S. H. Cho, S. Kim, and J.-H. Choi, "Transfer learning-based fault diagnosis under data deficiency," *Appl. Sci.*, vol. 10, no. 21, p. 7768, Nov. 2020.
- [168] M. Liao, C. Liu, C. Wang, and J. Yang, "Research on a rolling bearing fault detection method with wavelet convolution deep transfer learning," *IEEE Access*, vol. 9, pp. 45175–45188, 2021.
- [169] X. Li, H. Jiang, R. Wang, and M. Niu, "Rolling bearing fault diagnosis using optimal ensemble deep transfer network," *Knowl.-Based Syst.*, vol. 213, Feb. 2021, Art. no. 106695.
- [170] Z. Chen, J. Cen, and J. Xiong, "Rolling bearing fault diagnosis using time-frequency analysis and deep transfer convolutional neural network," *IEEE Access*, vol. 8, pp. 150248–150261, 2020.
- [171] M. Marei and W. Li, "Cutting tool prognostics enabled by hybrid CNN-LSTM with transfer learning," *Int. J. Adv. Manuf. Technol.*, pp. 1–20, Jan. 2021.
- [172] H. Shao, M. Xia, G. Han, Y. Zhang, and J. Wan, "Intelligent fault diagnosis of rotor-bearing system under varying working conditions with modified transfer convolutional neural network and thermal images," *IEEE Trans. Ind. Informat.*, vol. 17, no. 5, pp. 3488–3496, May 2020.
- [173] Y. Dong, Y. Li, H. Zheng, R. Wang, and M. Xu, "A new dynamic model and transfer learning based intelligent fault diagnosis framework for rolling element bearings race faults: Solving the small sample problem," *ISA Trans.*, vol. 121, pp. 327–348, Feb. 2021.
- [174] F. Aziz, A. Ul Haq, S. Ahmad, Y. Mahmoud, M. Jalal, and U. Ali, "A novel convolutional neural network-based approach for fault classification in photovoltaic arrays," *IEEE Access*, vol. 8, pp. 41889–41904, 2020.
- [175] D. Zhang and T. Zhou, "Deep convolutional neural network using transfer learning for fault diagnosis," *IEEE Access*, vol. 9, pp. 43889–43897, 2021.
- [176] Y. Bai, J. Yang, J. Wang, Y. Zhao, and Q. Li, "Image representation of vibration signals and its application in intelligent compound fault diagnosis in railway vehicle wheelset-axlebox assemblies," *Mech. Syst. Signal Process.*, vol. 152, May 2021, Art. no. 107421.
- [177] C. Grover and N. Turk, "A novel fault diagnostic system for rolling element bearings using deep transfer learning on bispectrum contour maps," *Eng. Sci. Technol., Int. J.*, vol. 31, Jul. 2022, Art. no. 101049.
- [178] K. Guo, X. Wan, L. Liu, Z. Gao, and M. Yang, "Fault diagnosis of intelligent production line based on digital twin and improved random forest," *Appl. Sci.*, vol. 11, no. 16, p. 7733, Aug. 2021.
- [179] Z. Wang, W. Jia, K. Wang, Y. Wang, and Q. Hua, "Digital twins supported equipment maintenance model in intelligent water conservancy," *Comput. Electr. Eng.*, vol. 101, Jul. 2022, Art. no. 108033.
- [180] H. Bai, H. Yan, X. Zhan, L. Wen, and X. Jia, "Fault diagnosis method of planetary gearbox based on compressed sensing and transfer learning," *Electronics*, vol. 11, no. 11, p. 1708, May 2022.
- [181] H. Zhu, Z. Huang, B. Lu, F. Cheng, and C. Zhou, "Imbalance domain adaptation network with adversarial learning for fault diagnosis of rolling bearing," *Signal, Image Video Process.*, vol. 16, no. 8, pp. 2249–2257, 2022.
- [182] J. Zhu, Y. Yao, and F. Gao, "Transfer of qualitative and quantitative knowledge for similar batch process monitoring," *IEEE Access*, vol. 6, pp. 73856–73870, 2018.
- [183] 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. Informat.*, vol. 15, no. 4, pp. 2416–2425, Apr. 2018.
- [184] J. Ma, X. Liu, X. Zou, M. Yue, P. Shang, L. Kang, S. Jemei, C. Lu, Y. Ding, N. Zerhouni, and Y. Cheng, "Degradation prognosis for proton exchange membrane fuel cell based on hybrid transfer learning and intercell differences," *ISA Trans.*, vol. 113, pp. 149–165, Jul. 2021.
- [185] T. Han, C. Liu, R. Wu, and D. Jiang, "Deep transfer learning with limited data for machinery fault diagnosis," *Appl. Soft Comput.*, vol. 103, May 2021, Art. no. 107150.
- [186] W. Zhang, X. Li, H. Ma, Z. Luo, and X. Li, "Open-set domain adaptation in machinery fault diagnostics using instance-level weighted adversarial learning," *IEEE Trans. Ind. Informat.*, vol. 17, no. 11, pp. 7445–7455, Nov. 2021.
- [187] W. Zhang, X. Li, H. Ma, Z. Luo, and X. Li, "Universal domain adaptation in fault diagnostics with hybrid weighted deep adversarial learning," *IEEE Trans. Ind. Informat.*, vol. 17, no. 12, pp. 7957–7967, Dec. 2021.
- [188] N. Lu, H. Xiao, Y. Sun, M. Han, and Y. Wang, "A new method for intelligent fault diagnosis of machines based on unsupervised domain adaptation," *Neurocomputing*, vol. 427, pp. 96–109, Feb. 2021.
- [189] Q. Li, C. Shen, L. Chen, and Z. Zhu, "Knowledge mapping-based adversarial domain adaptation: A novel fault diagnosis method with high generalizability under variable working conditions," *Mech. Syst. Signal Process.*, vol. 147, Jan. 2021, Art. no. 107095.
- [190] Z.-H. Liu, B.-L. Lu, H.-L. Wei, L. Chen, X.-H. Li, and C.-T. Wang, "A stacked auto-encoder based partial adversarial domain adaptation model for intelligent fault diagnosis of rotating machines," *IEEE Trans. Ind. Informat.*, vol. 17, no. 10, pp. 6798–6809, Oct. 2020.
- [191] W. Hong, Z. Wang, M. Yang, and J. Yuan, "Conditional generative adversarial network for structured domain adaptation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 1335–1344.
- [192] E. Tzeng, J. Hoffman, K. Saenko, and T. Darrell, "Adversarial discriminative domain adaptation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jul. 2017, pp. 7167–7176.
- [193] G. Michau, Y. Hu, T. Palmé, and O. Fink, "Feature learning for fault detection in high-dimensional condition monitoring signals," *Proc. Inst. Mech. Eng., O, J. Risk Rel.*, vol. 234, no. 1, pp. 104–115, Feb. 2020.



- [194] G. M. Y. Dauphin, Y. Dauphin, X. Glorot, S. Rifai, Y. Bengio, I. Goodfellow, E. Lavoie, X. Muller, G. Desjardins, D. Warde-Farley, and P. Vincent, "Unsupervised and transfer learning challenge: A deep learning approach," in *Proc. ICML Workshop Unsupervised Transf. Learn.*, 2012, pp. 97–110.
- [195] W. Mao, J. He, B. Sun, and L. Wang, "Prediction of bearings remaining useful life across working conditions based on transfer learning and time series clustering," *IEEE Access*, vol. 9, pp. 135285–135303, 2021.
- [196] X. Wang, X. Liu, and Y. Li, "An incremental model transfer method for complex process fault diagnosis," *IEEE/CAA J. Automat. Sinica*, vol. 6, no. 5, pp. 1268–1280, Sep. 2019.
- [197] O. Day and T. M. Khoshgoftaar, "A survey on heterogeneous transfer learning," *J. Big Data*, vol. 4, no. 1, pp. 1–42, Dec. 2017.
- [198] 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. Ind. Informat.*, vol. 18, no. 3, pp. 1790–1800, Mar. 2021.
- [199] F. Flammini, "Digital twins as run-time predictive models for the resilience of cyber-physical systems: A conceptual framework," *Phil. Trans. Roy. Soc. A, Math., Phys. Eng. Sci.*, vol. 379, no. 2207, Oct. 2021, Art. no. 20200369.



**MEHDI SAMAN AZARI** received the B.Sc. degree in electrical engineering from the University of Tabriz, Tabriz, Iran, in 2013, and the M.Sc. degree in control engineering and industrial automation from Tarbiat Modares University, Tehran, Iran, in 2016. He is currently pursuing a Ph.D. degree with the Department of Computer Science, Linnaeus University, Sweden. His main research interests include artificial intelligence, specifically transfer learning and predictive maintenance, and their applicability to cyber-physical systems (CPS).



**FRANCESCO FLAMMINI** (Senior Member, IEEE) has worked in private and public companies on transportation and infrastructure safety and security projects, from 2003 to 2017. Since 2018, he has been working as an Associate Professor with Linnaeus University, Sweden, where he has chaired the cyber-physical systems (CPS) environment. He is currently a Full Professor of computer science with Mälardalen University, Eskilstuna, Sweden, where he is leading several research and educational projects with a focus on CPS security and Industry 4.0.



**STEFANIA SANTINI** (Member, IEEE) is an Associate Professor with the Department of Electrical Engineering and Information Technologies (DIETI), University of Naples Federico II, Napoli, Italy, where she leads the Distributed Automation Systems Laboratory (DAiSYLab). She is involved in many projects with industry, including small- and medium-sized enterprises. Her research interests include nonlinear control of nonlinear and cyber-physical systems, networked control with applications to energy, automotive engineering, and transportation technologies. She is the Vice-Chair of the IEEE Italy Intelligent Transportation Systems Society (ITSS)-Italian Chapter. She is currently an Associate Editor of IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS.



**MAURO CAPORUSCIO** received the Ph.D. degree in computer science from the University of L'Aquila, Italy, 2006. He was an Assistant Professor at the Politecnico di Milano and a Postdoctoral Researcher at INRIA Paris-Rocquencourt. He is a Full Professor of computer science with the Department of Computer Science and Media Technology, Linnaeus University, Växjö, Sweden. He has published several papers in the most important international journals and conferences and has served on the program committee of international conferences. His research interests include application of model-based software engineering methodologies and techniques to the field of self-adaptive systems, with a particular emphasis on decentralization and resiliency.

...