Review

Deep Transfer Learning Models for Industrial Fault Diagnosis Using Vibration and Acoustic Sensors Data: A Review

Md Roman Bhuiyan [1] and Jia Uddin [2,*]

[1] Faculty of Computing and Informatics, Multimedia University, Persiaran Multimedia, Cyberjaya 63100, Malaysia

[2] Artificial Intelligence and Big Data Department, Endicott College, Woosong University, 171 Dongdaejeon-ro (155-3 Jayang-dong), Daejeon 300718, Republic of Korea

* Correspondence: jia.uddin@wsu.ac.kr

Abstract: In order to evaluate final quality, nondestructive testing techniques for finding bearing flaws have grown in favor. The precision of image processing-based vision-based technology has greatly improved for defect identification, inspection, and classification. Deep Transfer Learning (DTL), a kind of machine learning, combines the superiority of Transfer Learning (TL) for knowledge transfer with the benefits of Deep Learning (DL) for feature representation. As a result, the discipline of Intelligent Fault Diagnosis has extensively developed and researched DTL approaches. They can improve the robustness, reliability, and usefulness of DL-based fault diagnosis techniques (IFD). IFD has been the subject of several thorough and excellent studies, although most of them have appraised important research from an algorithmic standpoint, neglecting real-world applications. DTL-based IFD strategies have also not yet undergone a full evaluation. It is necessary and imperative to go through the relevant DTL-based IFD publications in light of this. Readers will be able to grasp the most cutting-edge concepts and develop practical solutions to any IFD challenges they may encounter by doing this. The theory behind DTL is briefly discussed before describing how transfer learning algorithms may be included into deep learning models. This research study looks at a number of vision-based methods for defect detection and identification utilizing vibration acoustic sensor data. The goal of this review is to assess where vision inspection system research is right now. In this respect, image processing as well as deep learning, machine learning, transfer learning, few-shot learning, and light-weight approach and its selection were explored. This review addresses the creation of defect classifiers and vision-based fault detection systems.

Keywords: deep transfer learning; deep learning; intelligent fault diagnosis

1. Introduction

Safety-critical systems include aeroengines, vehicle dynamics, chemical processes, manufacturing systems, electric machines, wind energy conversion systems, and industrial electronic equipment, among others. Industrial systems prone to process irregularities and component breakdowns prioritize reliability and safety. Defect-tolerant operation may prevent performance drops and harmful circumstances by immediately recognizing abnormalities or errors.

Any variation from the expected/usual/standard behavior of a distinctive attribute or parameter of the system is considered a defect [1]. When a system component is disconnected or a sensor fails to work properly (for example, when a sensor gets stuck at one specific value or has a change in the sensor scalar factor), it may lead to problems. This means that, in order to classify the faults, they are often divided into actuator, sensor, and plant faults (also known as component or parameter faults). These faults either interrupt the controller's action on the plant, or produce significant measurement errors, or directly change the dynamic input/output properties of the plant, resulting in a decrease in the system's performance

and even system collapse [2]. Fault diagnostic is often used to monitor, find, and diagnose errors in a system utilizing the notion of redundancy, either hardware redundancy or software redundancy, with the goal of improving system dependability (or called analytical redundancy). Because the output signals from similar components with identical input signals can be compared, diagnostic choices may be made using a number of approaches such as limit checking, and majority voting, among others. In terms of reliability, hardware redundancy is costly, but it also adds weight and increases the amount of available space [3]. Hardware redundancy is required for important components, but not for the whole system because of the expense or complexity of installation when the space and/or weight are limited in terms of both size and weight. The analytical redundancy approach, whose schematic design may be illustrated in Figure 1, has been the major stream of fault diagnostic research since the 1980s, with the development of current control theory [4]. Assuming that the input and output data are consistent with the pre-existing knowledge of a healthy system for a controlled system, a fault diagnosis algorithm can be constructed using the input and output data in order to check for actuator, process/component, and sensor faults. A diagnostic decision can then be made using diagnostic logics. Due to external disturbances, inherent modeling inaccuracy, and the complexity of the system dynamics and control structure, analytical redundancy diagnostic techniques are more cost-effective than hardware redundancy approaches.

In today's industrial environment, it is critical to keep tabs on the health of machines. If these devices fail, they might result in huge financial losses and even put the lives of the individuals who work with them at risk. Machine health monitoring has always been in demand because it is necessary to maintain industrial equipment running at their peak performance and reliability [5,6]. With regard to rolling element bearings, their health problems may have a huge influence on the performance, stability and life duration of a rotating mechanism, for example by changing the diameters of the faults in various locations under different loads. Most often, a real-time vibration monitoring system is used while the rotating mechanism is operating [7,8].

Recent years have seen tremendous achievement in computer vision [9,10] and speech recognition [11,12] using deep learning approaches. The use of deep learning in machine health monitoring systems is not new. Auto encoders with three hidden layers were studied by Lu et al. in a comprehensive empirical investigation for defect detection of rotating machinery components [13]. Receptive input size and structure, deep structure, sparsity constraint, and denoising operation were all examined in their research. Instead of utilizing raw signals as inputs to acoustic emission (AE) models, some researchers employed frequency–domain characteristics as inputs. They entered time-series data from rotating equipment into SAE because frequency spectra may better distinguish the health issues of a rotating machine [14]. Ref. [15] Zhu and his colleagues suggested an SAE-based DNN for hydraulic pump problem diagnostics that employs Fourier transforms to create frequency characteristics. To train their DNN, Liu et al. employ a normalized STFT of a sound source as an input. Some studies [16,17] use SAE to combine statistical data from many domains, such as time, frequency, and time-frequency-domain.

Data with well-defined grid-like shapes such as 2D image data or 1D time series data may be processed using CNNs, which were initially suggested by LeCun [18]. CNNs have seen success in image recognition in the last several years. Several CNN architectures are available, including VGG-net, Res-net, and inception v4. Vibration signals in one dimension are not compatible with CNNs [19]. Using just $2 \times 3 \times 3$ weights, all of these models can achieve a receptive area the same size as a 5×5 convolutional kernel using two subsequent 3×3 convolutional layers. Two 3-by-1 convolutional layers may provide a 5-by-1 receptive area for a one-dimensional, signal, albeit at the expense of two 3-by-3 weights. Convolutional layers 1-7 and 7-1 are utilized for inception V4; however, for 1D vibration signals, 7-1 convolutional operations are not possible [20,21]. From the foregoing, it may be deduced that the standard 2D CNN model may not be as effective on 1D signals. For 1D vibration signals, a new CNN model is thus required. Using CNN for fault diagnosis has been employed since 2015. The DFT of the normalized vibration signal collected by two accelerators placed perpen-

dicular to one another and the outputs of four classification categories representing healthy bearing (HB), mildly inadequately lubricated bearing (MILB), and extremely inadequately lubricated bearing (EILB) were used by Janssens et al. to diagnose bearing health conditions [22]. Diagnosed bearing health conditions using the DFT of the normalized vibration signal collected by two perpendicularly placed accelerometers and the outputs of four classification categories representing healthy bearing (HB), mildly inadequately lubricated bearing (MILB), and extremely inadequately lubricated bearing (EILB). The input is the DFT of the normalized vibration signal, and the output is the DFT of the layer that is completely linked (200 neurons). Structural degradation can be diagnosed and localized using 1D CNNs [23]. For the experiment, thirty accelerometers were installed at thirty joints on the principal girders of a grandstand simulator. This paper proposes an adaptive 1D CNN architecture to integrate the feature extraction and learning (damage detection) phases of the raw accelerometer data. Using a 5-layer WDCNN model with large first-layer kernels and small subsequent kernels, as described in [24], it is feasible to diagnose bearing problems with a significant quantity of training data and high accuracy, even in noisy environments.

As a result, Deep Learning approaches are now being studied in the area of defect detection with the purpose of using a smaller number of training samples in order to learn more information and obtain better results than typical Machine Learning algorithms. Real-world circumstances are not consistent with this, since it is simple to gather big volumes of data in the industry as well as using data augmentation methods [25–27] to enhance the quantity of training samples in industry. Even if we use classical algorithms or deep learning algorithms to analyze the data, virtually all can achieve almost 100% accuracy. As a result, the algorithm's capacity to adapt in a real-world industrial setting has to be given additional consideration. The flexibility of an algorithm in the area of fault diagnostics may be measured in two ways: No matter how much the working load of a machine changes, the model trained under one loading may still be evaluated on data from another working load [28,29]. When testing on noisy data, a model trained on noise-free samples may attain a high accuracy rate [30,31] since noise is an inevitable issue. WDCNN's performance quickly declines without AdaBN, which needs statistical knowledge of all test data, in noisy situations. Even the state-of-the art DNN [14] fails to diagnose adequately in noisy environments. As a result, we may conclude that Deep Learning models have not yet come up with a solution to this issue.

1.1. Transfer Learning

Rotating equipment needs rolling element bearings to function. Bearing defects would impair the machine's performance and cause the whole rotary machine to break down, causing significant property damage, downtime, and injuries [32–34]. Four subcomponents make up a rolling element bearing: a rolling element, an inner race, an outer race, and a cage. Ninety percent of bearing failures are brought on by either the inner or outer race, while only ten percent are brought on by the cage or rolling element [35]. An emphasis on fault detection of rolling element bearings, particularly bearing race fault diagnostics, is necessary to increase the safety and dependability of mechanical equipment. The identification of bearing race faults at the fault site level is covered in this work. A single, isolated defect in a bearing subcomponent is sought to be identified.

For autonomous fault identification in complicated industrial processes, data-driven approaches starting with historical data are superior to well-developed model-based and signal-based techniques that need prior knowledge of models or signal patterns [36,37]. Three phases are commonly included in conventional data-driven techniques for issue detection: Data collection, human feature extraction and selection, and automated problem identification are the first three steps [38]. Despite the fact that conventional data-driven approaches use machine learning algorithms such as supporting vector machine (SVM), artificial neural network (ANN), and k-nearest neighbor (KNN) to accomplish automated diagnosis in defect detection procedures [39], feature extraction is still carried out by hand. Many feature extraction techniques have been proposed based on contemporary signal processing technologies, including those based on entropy, spectrum kurtosis (SK), and empirical mode decomposition

(EMD) [40]. Despite the effectiveness of these methods, manually extracting features might be tiresome and time-consuming [41]. Recent research focuses on end-to-end intelligent diagnosis based on deep learning techniques capable of hierarchical and nonlinear representation learning. These techniques include convolutional neural networks (CNN)-based diagnosis methods, Restricted Boltzmann Machines (RBM)-based diagnosis methods, and Auto-encoder (AE) and its variants [42]. Both conventional machine learning and deep learning consider the distribution of the training and test sets to be the same. As a result, data relevant to diverse health issues that were collected using the same tools and working environments and used in several research investigations to train and assess diagnosis models. However, it is frequently challenging to obtain extensive historical fault data from the equipment to be diagnosed in advance in real-world diagnostic situations, particularly online diagnosis; only a limited number of fault data of specific equipment operating under specific circumstances are accessible [43,44]. The diagnostic model would be vulnerable to overfitting and challenging to generalize if it were developed with such a small dataset [45].

The motivations for solving the small sample problem can be roughly categorized into three groups: (1) generating new new data based on the existing data through data augmentation methods [46,47] or generation models [48,49]; (2) generating new data based on the existing data; and (3) generating new data based on the existing data. AE citeb49 and Generative Adversarial Networks (GAN) [49–52], (3) relying on the diagnosis knowledge learned from other datasets based on transfer learning [53–55] or meta learning [56–58], and (4) employing specific network structures or tricks to Inspiring by the second motivation, this study proposes A single point issue on the race of rolling element bearings often has a periodic effect, resulting in bearing characteristic frequencies (BCFs) in frequency domain and envelope characteristics in time domain in bearing vibration signals [59]. In actual vibration signals, however, in addition to these fault characteristics related to the bearing, there are numerous other complex signal components due to the influences of machine structure and numerous random factors, which are the specific characteristics of the equipment and working conditions. These components contribute to the complex distribution of real data by improving the vibration signal properties. When there is little available in-service data, the learning algorithm cannot learn the effective bearing failure characteristics, but rather other attributes that merely classify the unique training dataset, leading to a loss of generalization [60]. Theoretically disclosing the characteristics of a bearing under different operating conditions is made easier by physics-based modeling. In the bearing dynamic model, only bearing-related components are considered, and several random and nonlinear aspects are simplified. Thus, bearing fault features are substantially more obvious in simulation signals than in real signals. In addition, it is straightforward to acquire adequate vibration data for bearings of different kinds working under a range of conditions and failure patterns by selecting simulation parameters intelligently. Motivated by this, the proposed method proposes to employ simulation data-acquired diagnostic knowledge to help in the diagnosis of real-world events [61].

In order to solve the aforementioned deficiency, this study aims to provide a comprehensive assessment of DTL for defect detection in industrial settings. In contrast to previous research that either emphasized classical machine learning, deep learning, or transfer learning for IFD approaches, this review article will emphasize the new machine learning paradigm of DTL. Second, in contrast to prior review papers, this one classifies DTL-based intelligent failure diagnosis (IFD) systems in the context of real industrial conditions. This categorization might enable engineers to identify the most suitable algorithms for their specific industrial needs. The bulk of previously published review articles focused on works released before 2020. As the IFD attracts the interest of both academic and industry researchers, many new studies have recently been published. Nevertheless, the bulk of the most current state-of-the-art DTL approaches was already included in this review study before it was submitted for publication in the journal.

The major contributions of this research are below:

- In order to provide the reader with an algorithmic grasp of DTL, this article introduces the reader to the fundamental concepts and theories of DTL, including instance-based DTL, model-based DTL, and feature-based DTL.
- This article discusses a crucial case of defect identification where the diagnostic model may be used to previously unexplored regions without knowledge of their data distributions. The alternative is more appropriate for real-world diagnostic tasks than conventional data-driven methods.
- The well-designed approach can handle machinery defect diagnostic duties efficiently. An experimental study demonstrates the significance and superiority of each generalization aim.

1.2. Objectives and Organization

Based on these results, it is important to do a full evaluation of these updated and improved FDD approaches to make them more useful, effective, and efficient.

- The primary focus of this study is the modeling of systems using both basic principles and signal form. In this FDD subject, it is essential to comprehend the fundamentals of data-driven methodologies, to define issues, and to provide views.
- In order to provide academics and practitioners a thorough knowledge from both a theoretical and an application viewpoint, the second purpose is to undertake a systematic evaluation of the growing research effort, the so-called data-driven FDD approaches, for traction systems during the previous ten years.
- On our third attempt, we attempted to classify into three categories. Fault detection Datasets, Deep Learning Based Papers, and Transfer Learning Based Papers are among the categories.
- With an emphasis on real-world applications and contemporary data analysis tools, the ultimate objective is to provide research possibilities in data-driven FDD approaches for traction systems, as well as research challenges and future prospects.

The rest of the paper is organized based on Figure 1. Figure 1 is below. We have also added list of abbreviations in Table 1.

Table 1. List of abbreviations.

Short Form	Full Form
DTL	Deep Transfer
DL	Deep Learning
IFD	Intelligent Fault Diagnosis
DNN	Deep Neural Network
AE	Acoustic Emission
CNN	Convolutional Neural Network
HB	Health Bearing
SVM	Supporting Vector Machine
ANN	Artificial Neural Network
KNN	K-Nearest Neighbor
SK	Spectrum Kurtosis
EMD	Empirical Mode Decomposition
RMB	Restricted Boltzmann Machines
GAN	Generative Adversarial Networks
FDD	Feature-Driven Development
ASI	Acoustic Spectrum Imaging
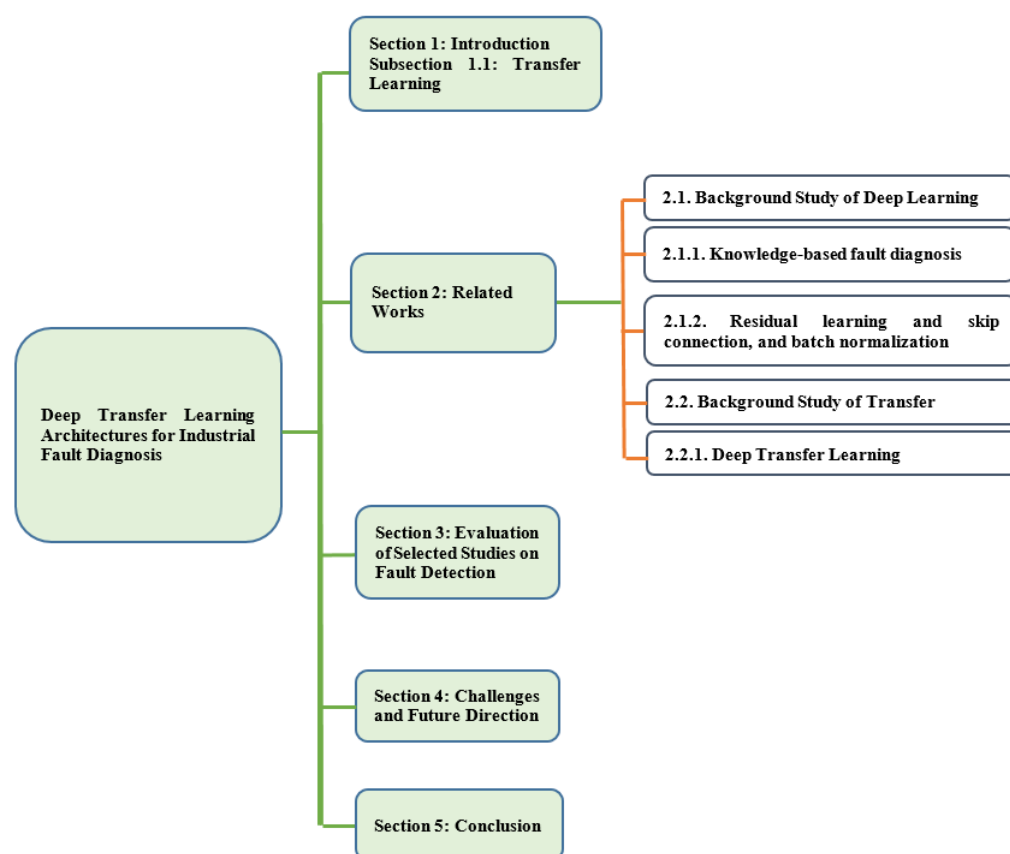
Figure 1. Road mapping of the full paper.

2. Related Works

2.1. Background Study of Deep Learning

Researchers have spent a lot of time and money trying to find a solution to the issue of bearing fault diagnostics in rotating equipment. Knowledge-based approaches are able to solve the bearing diagnostic problem with high accuracy and performance because they can analyze sensor and actuator data in an efficient way. Deep learning (DL) is one of the machine learning algorithms that have the benefit of not relying on feature extraction and instead giving precise diagnosis. In order to deal with this problem, a new DL-based model for the identification and classification of motor bearing faults is described in this study. First, a signal-to-image conversion technique for converting time-domain signals into pictures is presented. It then uses a revolutionary deep residual learning (DRL) network to develop an end-to-end mapping between pictures and the motor bearing health state [62].

In this section, we have divided the relevant works into two categories: Knowledge-based fault diagnosis is used in a new study on diagnosing problems with bearings. Two related subtopics related to the DRL network methodology are also discussed: residual learning with skip connection and batch normalization.

2.1.1. Knowledge-Based Fault Diagnosis

However, even though fault diagnosis approaches can be grouped into model-based, signal-based, and knowledge-based fault diagnosis, only knowledge-based studies are shown in this section because they can show the relationship between raw measured data and its result. The [63] complex industrial systems find it hard to have the models and signal patterns of simple systems [64].

CNN-based fault diagnosis was explored by Wen et al. [65]. The suggested CNN approach, which is based on LeNet-5, may be applied to a typical dataset of motor bearing raw signals. A collection of pictures was created by converting the 1D raw signals to

2D gray level images. Because of this, CNN-based methods for bearing failure detection seem to have a bright future ahead of them. To help with bearing failure identification, Xia et al. [65] created a CNN-based model that uses 1D raw vibration data. Using many sensors instead of a single one was proven by the authors. Unsupervised feature learning was used by Lei et al. [43] to develop a two-stage learning system for diagnosing bearing faults. Sparse filtering is used in the first step of the proposed method, which employs 1D raw vibration data. Bearings' health is then classified using softmax regression. Transfer learning-inspired deep convolutional transfer learning networks for bearing fault diagnostics have been suggested by [66]. For the first step, we propose to employ a 1D CNN to learn the discriminative features from a 1D raw signal and then classify the bearings' state. Researchers from Wen et al. [67] explored the use of 1D raw vibration data from bearings for bearing failure diagnostics. The discriminative features are extracted from the raw data using a three-layer sparse auto-encoder. On a well-known dataset of motor bearings, the suggested method outperformed existing machine learning algorithms in terms of classification accuracy. Feature learning and a deep stack autoencoder are used in conjunction with a deep adversarial domain adaption model to extract more useful fault features. The research shows that the suggested model's classification accuracy is superior to other current machine learning and deep learning approaches. A CNN model developed by Janssens et al. [23] was used to learn aspects of 1D raw vibration data from bearings and monitor their health. Rolling bearing defect identification was examined by Guo et al. using a hierarchical deep CNN model. Researchers came up with an adjustable learning rate for the model and found that it outperformed typical CNNs in terms of effectiveness. Raw vibration data were transformed to 2D pictures by Ding and He [7], who developed a bearing failure diagnostic approach based on deep convolutional networks. Eren [68] suggested a 1D CNN-based bearing defect detecting method. The network discovered a correlation between bearing health and raw vibration signals. It was proposed by Chen et al. [69] to enhance the classification performance of bearing defects using a DL-based method. Analysis of cyclic spectral patterns was used to identify the different kinds of bearing faults. The high-level feature learning capacity of a CNN model was then used to construct a classification model for mistakes. For bearing defect identification, Xu et al. [70] suggested a hybrid DL model that included both a CNN model and a deep forest model. Time-frequency data were processed using a CNN model. Vibration signal images and fault feature extractions loaded into a dense forest End-to-end DL methods for bearing defect identification have been examined by Khorram et al. [71]. The model was a mix of CNN and LSTM models, and it utilized raw vibration data directly without any preprocessing or modification methods. The results were impressive.

A new DRL-based network model for bearing defect identification is provided in this paper. An end-to-end mapping approach, rather than hand-crafted features with restricted representation capabilities, is presented as an alternative. For the experiment to be more reliable, the effects of the number of residual blocks and the BN layer are examined.

2.1.2. Residual Learning and Skip Connection, and Batch Normalization

The techniques of residual learning and skip connection, as well as batch normalization, are briefly discussed in the following sections.

Residual learning and skip connection:

As additional layers are stacked to address increasingly complicated problems, the model's training accuracy degrades. Furthermore, as the network becomes more complex, it becomes more difficult to train it. It is also difficult for the deep CNN model to achieve convergence because of the vanishing/exploding gradient issue during backpropagation. For the reasons stated above, He et al. [21] first developed a residual learning strategy for CNN structure. Instead of learning the intended original mapping, the residual mapping, which is the difference between the layer output and the ground truth, is discovered. Because of this, a DRL technique such as this makes deep CNN optimization easy while also improving accuracy. One or more layers might be skipped in order to obtain the residual

mapping. As a result, the outputs of the stacked layers and shortcut connections are added element-by-element [72]. It is simpler to optimize a network with these connections, and the network's categorization performance improves as a result. Relative learning schemes with identity skip connections are used in the suggested model's design. Relative learning's link to identity skip connection is also addressed in order to resolve the knowledge-based defect diagnostic issue that has not yet been solved utilizing residual learning.

Batch normalization: Mini-batch stochastic gradient descent (SGD) has been extensively employed as an optimization technique for minimizing the objective function in deep CNN models that need to optimize an objective function with regard to its parameters. SGD is simple and effective but model parameters and learning rates must be properly established at the start. A deep CNN must be able to adapt to its changing inputs during training since each layer's inputs are constantly changing. The issue of shifting internal covariates may be addressed by using a BN layer between convolutional layers and non-linearities. There has been no research towards using BN for knowledge-based defect diagnosis, as far as we can tell. The results of this investigation, on the other hand, show that batch normalization enhances classification accuracy, as previously reported.

A wide range of industries, including manufacturing, electric cars, and power production, have benefited from the deployment of PMSMs in recent years. A mechanical transmission is often used to connect the PMSM to the driving load in these cases. The transmission is vulnerable to failure because of its critical function in transferring energy from the PMSM to the driving load. Transmission problems may have serious ramifications, ranging from lost productivity to safety concerns. Because of this, there is a need for monitoring and early FDI of PMSM transmission to prevent failures and minimize downtime.

Because the transmission failure creates particular distinctive frequencies that modulate in the vibration spectrum, the vibration-based diagnostic approach is widely used. Vibrations are non-stationary and time-varying as a function of the gear's kinetics and arrangement. For this reason, signal processing techniques including Fourier series, short-time Fourier transform, and wavelet transform are often used on vibration signals. Using this vibration-based diagnostic procedure, on the other hand, would be costly. Continuous condition monitoring needs specialized accelerometer transducers and data gathering systems. Because the accelerometer transducers must be placed in a certain area, the vibration signal might be affected by outside disturbances. The accelerometer transducers cannot be installed in many places and drive-train layouts.

Mathematical models may be used to better understand PMSM behavior by exploiting the link between the input voltage signal and output current signal. For the purpose of calculating an estimated current signal, the voltage signal will be entered into a mathematical model. The predicted current signal will be removed from the actual current signal before being added back in. This allows for the transmission fault signature to be preserved in the residual current signal even after harmonics and noise from the voltage and any compensation from the controller inverter have been filtered out. In this research, the residual current signal is transferred from the time domain to the frequency domain, in contrast to the typical residual signal analysis. Instead of seeing the time-domain residual current signal, the fault signs are seen in the residual current spectrum. The PMSM's transmission defects may be seen as perturbations to the rotating PMSM. That particular frequency is disrupted by a malfunction in the transmission and might cause problems with the PMSM rotation harmonics. Detecting fault signs is easier when looking at data at a greater frequency rather than a slower rate. The model-based technique instead uses the residual current spectrum rather than the observed current spectrum, in contrast to MCSA.

2.2. Background Study of Transfer Learning

A novel approach to machine learning called Deep Transfer Learning (DTL) enhances performance by fusing the advantages of Deep Learning's (DL) feature representation and Transfer Learning's (TL) knowledge transfer. In the area of Intelligent Fault Detection, DTL techniques have been the focus of much study and development. This was carried out

to increase the effectiveness, dependability, and toughness of DL-based fault diagnosis techniques (IFD). IFD has been the focus of several thorough studies, although the majority of these justifications have focused on algorithms rather than how IFD is actually employed. IFD tactics based on DTL have not yet been subjected to a comprehensive study. A full review of relevant publications in the field of DTL-based IFD is necessary and even more significant for assisting readers in comprehending the current state of the art and creating a feasible solution to IFD difficulties in the real world. Before showing how transfer learning methods may be included into deep learning models, we briefly address the theoretical foundations of DTL. Finally, significant DTL applications and current advancements in IFD are studied in depth. As a result, proposals for implementing DTL algorithms in real-world applications as well as some potential challenges have been provided. The results of the investigation are then given. As a conclusion, we have reason to anticipate that the research reported in this article will inspire and serve as a resource for future researchers attempting to enhance IFD [73].

Deep Transfer Learning

Industry 4.0, the fourth industrial revolution, is now a part of production. The goal of Industry 4.0 is for machines to have exact self-awareness, make decisions on their own, and communicate with each other in an intelligent way throughout the manufacturing process [74–76]. Industrial equipment (IE) has been focusing on providing economic advantages including quality enhancement, efficiency improvement, energy conservation, and cost reduction in response to such developments and revolutions. The IE, on the other hand, is often requested to do Herculean duties while working in a difficult operational environment and delivering long-term services [77]. Health status monitoring and diagnosis of industrial environments (IEs) are critical to the safety and dependability of the industrial environment because it reduces downtime, formulates planned maintenance and increases economic advantages. When it comes to accurately diagnosing faults for IE in time, the production process complexity and dynamicity makes this a huge difficulty.

Throughout the last several decades, academics and business researchers have devoted increasing attention to the need of timely and accurate IFD, which has been cited as a major priority by many governments and organizations. However, academics and engineers have created a wide range of intelligent algorithms, including deep learning and transfer learning, to solve a variety of practical challenges in industrial situations, and have also made significant progress with intelligent failure detection (IFD) of IE.

Artificial intelligence (AI) researchers use the deep learning subfield of machine learning to extract higher-level representations from unstructured input data, such as pixel-based photos, audio files, text texts, etc. [78]. There are many methods for doing this, including deep neural networks, deep belief networks, recurrent neural networks, convolutional neural networks, and graph neural networks (GNNs). Due to deep learning technology's phenomenal advantages in massive data processing, discriminative feature learning, and efficient pattern recognition, building intelligent models by mapping relationships between IE health conditions and industrial data end-to-end has been shown to be a promising method in many manufacturing industry applications [79]. In contrast, the limits of deep learning technology make it impossible for it to progress and be applied in increasingly challenging real-world circumstances. The following are the ideal and hypothetical applications of deep learning:

- Deep learning requires many pre-labeled samples in order to train a model. As was already said, one disadvantage of deep learning is that it mostly on direct observation to teach abilities. On the other hand, deep learning systems rely heavily on enormous amounts of labeled training data; without them, the algorithms are prone to overfitting and are weak at generalization.
- For deep learning to work, certain requirements must be met by the distributions between the training and test sets of data. This implies that, if a deep learning model

is trained on data that does not match the planned data distribution, its performance will likely suffer greatly or perhaps fail.

Transfer learning, a branch of machine learning that focuses on the collection of common information from one or more related but diverse application environments, has emerged as a viable technique for assisting deep learning in overcoming the aforementioned restrictions. Transfer learning may improve an AI model's learning performance even when training material is limited and sparse, similar to how people might utilize a small number of examples or past knowledge to handle unanticipated situations [46]. In addition, it may provide an AI model with strong generalization performance from similar but distinct application contexts to a new scenario. Therefore, it may be difficult for traditional machine learning algorithms to acquire and recall the discriminative representations necessary for transfer learning.

A new machine learning paradigm that uses deep learning technologies to transmit information has evolved in recent years. This paradigm has the potential to discover complicated patterns and obtain transferable knowledge more quickly. DTL is more suited for real-world industrial applications since it is simple to integrate with IFD of IE's well developed deep learning models [80].

DTL for fault diagnosis in manufacturing environments is the focus of this study, which aims to give readers interested in contributing to the growth and development of IFD with a thorough understanding of the technology. Fault diagnosis has been the subject of a number of systematic reviews in the past [20]. Traditional fault diagnostic and prognosis methods were heavily included in this literature study, which does not represent current methodologies [21]. Yan and Gao compiled a list of the most popular deep learning models used for machine health monitoring prior to 2019, including SAE, DBN, DBM, CNN, and RNN, and provided the data and code needed to replicate the results [81]. Using machine learning, Lei and Nandi [82] compiled an overview and road map for IFD approaches, dividing the evolution of IFD methods into three phases: classical machine learning, deep learning, and the future (transfer learning). Aside from those stated above, there are several related papers that concentrate on deep learning-based [83], transfer learning-based and convolutional neural network-based [55] approaches for machine defect diagnostics, which are not included here. Because of this, an overview of relevant DTL articles is more crucial than ever before in order to assist readers in swiftly grasping the most recent, cutting-edge concepts in IFD and devising practical solutions to common problems.

3. Evaluation of Selected Studies on Fault Detection

This section is divided into two categories: deep learning based paper and transfer learning based paper.

3.1. Deep Learning Based Paper

In [5], detecting and identifying any potential abnormalities or faults as early as possible and implementing real-time fault-tolerant operation in order to minimize performance degradation and avoid dangerous situations is becoming increasingly important because industrial systems are becoming increasingly complex and expensive. Throughout the past four decades, researchers have made significant progress in the fields of defect detection and fault-tolerant control. For the purpose of this three-part study, we are focusing on the last decade's findings in real-time defect diagnosis and fault tolerant control. In the first half of the paper, the methods and applications of model-based and signal-based defect diagnosis are looked at in depth.

In [84], a novel approach based on deep learning is offered to solve the issue in this research. The contributions consist of: Initially, we presented an end-to-end technique that accepts unprocessed temporal signals as inputs and so does not need any time-consuming denoising preprocessing. The model can attain a reasonable degree of precision in a noisy setting. Second, the model is independent of domain adaptation algorithms and does not need domain-specific knowledge. It can attain great precision regardless of the operating

load. To comprehend the suggested model, we will display the learnt characteristics and attempt to determine the causes for its great performance.

In [85], this study describes a model-based method for steady-state permanent magnet synchronous motor (PMSM) driving transmission defect detection and identification (FDI). The suggested architecture employs a PMSM state-space model and an approximation of a transmission model. Using Recursive Least-Square to create regression models for a parameter estimates (RLS) algorithm. The FDI is assessed via the residual current spectrum thresholding approach, the distinctive frequency and magnitude of the fault, as well as the parameter clustering. In the trials, transmissions with three distinct fault states are examined. As an introductory effort in condition monitoring of PMSM driving transmission, the findings of this work indicate a promising strategy that takes into account both residual current spectrum and parameter cluster, and accomplishes an acceptable level of decision-making in recognizing and diagnosing the defective situation.

In [86], deep learning has made amazing strides in the field of fault diagnosis. Nonetheless, the operating circumstances of an axial piston pump are variable, and the data distribution is not uniform, rendering the majority of deep learning models invalid. A deep multi-signal fusion adversarial model based transfer learning (MFAN) solution is given to address this issue. A multi-signal fusion module is created to assign weights to vibration signals and auditory signals, which increases the method's dynamic adjustment capability. In addition, the residual network is incorporated into the shared feature creation module in order to gather an abundance of feature data. According to the various operating loads of the axial piston pump, nine transfer scenarios are constructed, and the suggested approach is compared to five common diagnostic techniques.

In [87], this research is a continuation of our prior work on constructing induction motor inter-turn fault (ITF) models (Part I). In this research, an online ITF diagnostic technique for induction motors is presented by using the negative sequence current as a fault signal, based on the fault model developed in part I of a previous study. The ITF model analyzes the correlations between fault parameters, negative sequence current, and fault copper loss. The findings demonstrate that the fault severity index, which is a function of fault characteristics, has a direct relationship with the negative sequence and copper loss. Consequently, the model-based fault diagnostic approach suggested evaluates the fault severity index from the negative sequence current and identifies the ITF. With the predicted fault severity index, the thermal degradation-causing fault copper loss by the ITF may be computed. Finally, tests were conducted under several fault settings to validate the suggested approach for fault diagnosis.

Due [88] to the efficient processing of acquired sensor and actuator data, knowledge-based approaches are able to provide a potential solution to the bearing diagnostic issue with a high level of accuracy. Deep learning (DL) has the benefit among machine learning algorithms of disregarding feature extraction and offering accurate diagnosis. In order to solve this problem, a new DL-based model for failure detection and classification of motor bearing is described in this study. In this study, time domain signals are first transformed to images using a technique provided for signal-to-image conversion. The modified grayscale pictures are then sent into a unique deep residual learning (DRL) network designed to learn end-to-end mapping between images and motor bearing health state. Case Western Reserve University provides a frequently used actual vibration data set which is used to test the effectiveness of the planned DRL network (CWRU).

In [89], the purpose of this work is to provide interested readers with an overview of several model-based FD approaches created for linear time-invariant (LTI) systems in order for them to learn about current advancements in this area. Model-based FD approaches for LTI systems are categorized as parameter estimation techniques, parity space-based techniques, and observer-based techniques. In this study, the history, current advancements, and practical applications of each of these methodologies in relation to fault detection are explored.

In [90], existing failure diagnosis (FD) techniques cannot be used on different types of multiphase machines due to varying stator phase layouts. An enhanced machine-learning-based FD approach with adaptive secondary sample filtering is presented for multiphase drive systems in the age of big data and artificial intelligence. Experimental findings on both five-phase and six-phase motor driving platforms demonstrate the proposed method's generalizability as well as its excellent accuracy and robustness.

3.2. Transfer Learning Based Paper

In this section, we have tried to find out the research gap for transfer learning.

In [2], the proposed framework in uses convolutional neural networks (CNNs) and parameter transfer techniques to convey diagnostic data from enormous and varied simulation data generated by the dynamic model of bearings to the real environment. The usefulness of the technique is shown through a comprehensive analysis of three fault situations. The results show that the proposed method may discover more transferable features and reduce the disparity between the distribution of features based on simulation data and strategies for transferring parameters in CNN. This is a considerable increase in the performance of defect detection.

In [3] with this knowledge, it is even more crucial to conduct a thorough analysis of the pertinent DTL-based IFD publications. Doing so will help readers more rapidly understand the most cutting-edge techniques and come up with fast solutions for practical IFD problems. To demonstrate how transfer learning methods may be used with deep learning models, the theoretical foundations of deep learning transfer learning (DTL) are briefly presented. Finally, significant DTL applications and current IFD advancements are covered in-depth. It is also critical to emphasize that DTL algorithms are addressed in genuine applications and possible problems. Here are the survey's final findings. Finally, we have reason to assume that the study described in this article may assist and inspire scientists who desire to further IFD.

In [4], several partial distribution adaptation sub-networks (PDA-Subnets) are combined with a multi-source diagnostic knowledge fusion module to form a multi-source transfer learning network (MSTLN), which aggregates diagnostic information from multiple sources. By adjusting partial distributions of source and target pairings together, the former weights target samples using counter-balancing variables to release adverse effects owing to discrepancy across several source machines, while the latter fuses diagnostic conclusions from multiple PDA-Subnets even further. For unbalanced target samples, MSTLN may minimize misdiagnosis rates and provide higher transfer performance than standard approaches, according to two case studies.

In [91], this paper's primary objective was to use pre-trained deep transfer learning (DTL) structures and standard machine learning (ML) models as an automated method for diagnosing Parkinson's disease (PD) using sEMG data. To produce the discriminative feature vectors, we first stacked the extracted features from three deep pre-trained architectures, including AlexNet, VGG-f, and CaffeNet. Despite the enormous number of stacked features from all three deep structures, the correct features are successful in overcoming the difficulty of over-fitting and boosting resilience to varying amounts of additional noise. In order to minimize the size of the extracted features, a unique soft combination of subset feature selection approaches, including receiver operating characteristic (ROC), entropy, and signal-to-noise ratio (SNR) processes, was presented. Finally, we identified PD disorder using a support vector machine (SVM) with a radial basis function (RBF) kernel.

In [2], the proposed framework, convolutional neural network (CNN), and parameter transfer techniques are utilized to apply the diagnostic information derived from simulation data to real-world situations. A dynamic model of bearing is employed to generate enormous and varied simulation data. Three fault diagnosis cases are used to analyze and thoroughly describe the effectiveness of the recommended approach. The results show that the suggested strategy can learn more transferable features and reduce the feature

distribution discrepancy, significantly enhancing the fault detection performance, based on simulation data and parameter transfer approaches in CNN.

In [92], this study explains the design process of the built-in discrete time-series convolution neural network (DTCNN) and offers a hierarchical technique for TRU fault detection as well as a transfer learning-based fault diagnostic method as an alternative to training new models for various TRUs. Initially, the DTCNN construction was decided upon. Then, HDCNN's performance is confirmed. On this premise, the prerequisites for an appropriate source dataset for TRU defect diagnostics and the pre-trained HDCNN transfer layers are described. Comparing different algorithms with different noise settings showed that transfer learning is a good way to build a diagnostic network for similar equipment that can often lead to better performance.

In [93], CAE is utilized as the feature extractor in the proposed framework since it is capable of noise reduction. In addition, CORrelation ALignment (CORAL) loss and domain categorization loss are combined to amplify the impact of domain confusion. The suggested model is applied to the fault transfer diagnosis of planetary gearboxes operating under various working loads and noise levels, and it is compared to other conventional fault transfer diagnostic models. The experimental findings demonstrate that CAE-DTLN has greater diagnostic precision and more generalizability. The diagnostic accuracy of CAE-DTLN is typically greater than 99 percent. In addition, the transfer learning model provided has superior anti-noise performance.

In [94], this research provides a technique for defect diagnostics in building chillers based on transfer learning. On two water-cooled screw chillers, experiments were undertaken to capture both fault and fault-free data. Different transfer learning tasks, training instances, learning situations, and transfer learning implementation methodologies were considered for conducting experiments for fault diagnosis using transfer learning. The experimental findings demonstrate the use of transfer learning for FDD in building energy systems, particularly when there is limited experimental data available for model creation. The greatest accuracy increases for the two learning tasks were 12.63 and 8.18 percent. Based on transfer learning, the study's results can be used to help design FDD solutions for building energy systems.

In [95], this research proposes a defect detection approach based on parameter-based transfer learning and convolutional auto encoder (CAE) for wind turbines using small-scale data in order to maximize the use of this valuable data. The suggested approach is capable of transferring information from comparable wind turbines to the target wind turbine. The performance of the suggested technique is evaluated and compared to transfer and non-transfer methods. The comparative findings indicate that the suggested technique is superior for identifying flaws in small-scale wind turbine data.

In [96], the issue of transfer diagnosis with insufficient target data is addressed in a unique method in this study. The main concept is to pair the source and target data with the same machine condition and perform individual domain adaptation to mitigate the lack of target data, reduce the distribution discrepancy, and prevent negative transfer in consideration of the unclear data distribution described by the sparse data. Additionally, in our network, the problem of mismatched label spaces may be effectively resolved. The suggested method has been thoroughly tested on two case studies. All transfer variables, such as different working conditions and a range of machinery, are taken into account. Comprehensive testing shows that the suggested strategy works better than conventional transfer learning techniques.

Due to its high generalizability across a variety of industrial contexts [97], there has been a recent uptick in interest in transfer learning in equipment defect diagnostics. Existing techniques typically assume equal label spaces and suggest decreasing the marginal distribution difference between source and destination domains. This assumption, however, does not often hold true in real-world businesses, where testing data consist of a subspace of the source label space. Therefore, it is rational to transfer diagnostic information from a broad source area to a target domain with constrained machine conditions. In this paper,

a domain adaption strategy based on deep learning is used to tackle the difficult issue of partial transfer learning. It is suggested to use a class-weighted adversarial neural network to promote positive transfer of common classes and disregard source outliers. Experimental findings on two datasets of rotating equipment indicate that the suggested approach for partial transfer learning is promising.

Ref. [98] examines the latest developments in deep transfer learning for machine failure detection. There are several deep transfer structures and associated ideas, as well as an overview, categorization, and explanation of a number of relevant works. Based on this concept, this article discusses the key triumphs, problems, and prospective future research topics of deep transfer learning. This gives explicit guidelines for the architecture selection, design, and implementation of deep transfer learning for machine fault detection.

Ref. [99] propose more accurately aligning the marginal and conditional distributions of datasets, and this study introduces a novel transfer learning technique called improved joint distribution adaptation (IJDA). A working condition-resilient fault diagnostic approach based on vibration signals that fundamentally consists of three components is being developed in the meantime. Using noise to improve network performance, a novel data augmentation technique is created to give more relevant samples for unbalanced vibration signals. Second, the input dimension of IJDA is minimized via sparse filtering (SF).

Ref. [100] provides a dependable defect diagnostic technique based on acoustic spectrum imaging (ASI) of acoustic emission (AE) signals as an accurate health condition. Transfer learning is a method for machine learning that exchanges information with convolutional neural networks (CNN) for accurate diagnosis under different operating settings. ASI transforms the amplitudes of the spectral components of the time-domain windowed acoustic emission signal into spectrum imaging. ASI gives a visual representation of spectral properties of acoustic emission in photographs. This delivers improved spectrum pictures for transfer learning (TL) testing and training, as well as a robust classifier approach with excellent diagnostic precision.

Ref. [101] provides a brand-new taxonomy and investigates UDTL-based IFD in the context of various workloads. The transferability of features, the influence of backbones, negative transfer, physical priors, etc. seldom explored open and essential issues in UDTL-based IFD that are revealed through a comparative investigation of a number of popular methodologies and datasets. To facilitate future studies and establish the significance and repeatability of UDTL-based IFD, the whole test framework will be made available to the scientific community. In conclusion, the provided framework and comparison research may serve as an expanded interface and crucial findings for future work on UDTL-based IFD.

Ref. [102] describes a two-phase digital-twin-assisted fault detection approach based on deep transfer learning (DFDD) that enables fault identification in both the development and maintenance stages. Front running the high model in the virtual space identifies possible issues that were not addressed during design, while a deep neural network (DNN)-based diagnostic model is completely trained. In the second step, utilizing deep transfer learning, the previously trained diagnostic model may be transferred from the virtual environment to the physical space for real-time monitoring and predictive maintenance. This assures the correctness of the diagnosis and prevents time and knowledge waste.

Ref. [103] proposes a distance metric called polynomial kernel induced MMD (PKMMD) to solve the drawbacks. In conjunction with PK-MMD, a paradigm for reusing diagnostic information from one machine to another is developed. Two transfer learning instances are used to validate the proposed approaches, in which the health statuses of locomotive bearings are detected using data from laboratory motor bearings and gearbox bearings.

In this paper, we also discussed the latest datasets about fault detection. The datasets discussed below are:

Ref. [104] presents the "ToyADMOS" dataset, which is meant to identify anomalies in machine operation noises (ADMOS). According to their knowledge, there are no large-scale datasets available for ADMOS, despite the fact that large-scale datasets have contributed to recent breakthroughs in acoustic signal processing. This is because it is difficult to obtain

data on unusual sounds. To acquire abnormal operation noises of micro devices (toys) for the ADMOS database, we intentionally harmed them. The publicly available dataset includes three sub-datasets for machine-condition inspection: fault diagnosis of machines with geometrically fixed jobs and fault diagnosis of machines with changing tasks.

Ref. [105] provide a novel sound dataset for malfunctioning industrial machine inquiry and inspection (MIMII dataset). Different kinds of industrial machinery (e.g., valves, pumps, fans, and sliding rails) had normal noises recorded, and to simulate a real-world situation, numerous abnormal sounds were also recorded (e.g., contamination, leakage, rotational imbalance, and rail damage). The MIMII dataset is being made available to aid the machine-learning and signal-processing groups in the development of automated facility maintenance.

Ref. [106] offers the "ToyADMOS2" large-scale dataset for anomaly identification in machine operation noises (ADMOS). As we did for our earlier ToyADMOS dataset, we gathered a large number of operational sounds of small machines (toys) under normal and anomalous settings by intentionally destroying them, with the addition of controlled damage depth in anomalous samples. The ToyADMOS2 dataset is developed for testing systems under domain-shifting settings, since most ADMOS application scenarios demand robust performance under such conditions. The publicly available dataset includes two sub-datasets for machine-condition inspection: fault diagnosis of machines with geometrically fixed tasks and fault diagnosis of machines with moving tasks.

Ref. [107] provides MIMII Owing, a singular dataset for investigating and inspecting industrial machines that are not working properly with domain alterations brought on by operational and environmental changes. Standard techniques for recognizing anomalous sounds encounter practical difficulties since the distribution of characteristics changes between the training and operational phases (referred to as domain shift) as a result of several real-world factors. Unfortunately, there is currently not a dataset available that contains real domain shifts, making it impossible to assess the robustness against them. The new dataset contains the typical and anomalous operating noises of five different kinds of industrial machinery that operate in the source domain and target domain, respectively, under two different operational and environmental situations.

Ref. [108] renders denoising ineffective since it makes it hard to discover faults, which is a critical stage in defect identification and condition monitoring (CM). In the present study, background noise is efficiently removed, and the reliability of the defect identification method is increased by using the Singular Value Decomposition (SVD) and Hankel matrix based denoising approaches on the time domain vibration signals and spectra of ball bearings.

Ref. [109] presents and discusses empirical facts. The mode decomposition method (EMD) and the problem of mode mixing and its primary causes are discussed. Integrated EMD (EEMD) and, as treatments for EMD, methods for masking EMD (M-EMD) are described. The problem with blending models was that both upgraded procedures were used to examine and compare the results of an EEG signal. The M-EMD yielded the greatest results, successfully splitting the EEG signal into delta, theta, alpha, and beta bands [110]. In this paper, we use 1D to 2D conversion using Hilbert Transform, applied transfer learning with DCNN-LSTM architecture, and used a malfunctioning industrial machine investigation and inspection (MFPT) bearing vibration signals dataset. In this paper, we convert 1D to a 2D texture image, Gabor Filter with SVD for feature extraction, then SVM (Support Vector Machine) as a classifier, and used a Vibration signal dataset for validation [111]. We used 1D to 2D conversion for texture generation, Image Segmentation, and the Gabor filter for feature extraction used Acoustic Signal Dataset [112]. We applied GNS (Global Neighborhood Structure) for feature extraction, used multiclass SVM, and used vibration signals [113]. DNS map for feature extraction, multiclass SVM as a classifier, and used vibration signals [114], applied different 1D to 2D texture generation techniques with multiclass SVM, and used acoustic signals for validation [115,116].

To the best of our knowledge, deep transfer learning is best for fault detection. In this review, we have tried to compare deep transfer learning with other methods.

4. Challenges and Future Directions

In this section, we have tried to find the research gap, which is shown in Table 2.

Table 2. Find the research gap challenges and future directions.

Objectives and Challenges	Method	Dataset
Dynamic model of bearings was used to apply the diagnostic information from the simulation data to a real situation [2].	Intelligent fault diagnosis, Transfer learning, Dynamic model and Convolutional neural network	CWRU data and MFPT data
Transforming deep learning models into transfer learning approaches begins with a quick overview of the theoretical foundations of DTL. Following that, we go through some of the most important DTL applications and the most current DTL improvements in IFD [3].	Fault diagnosis, Deep learning, Transfer learning, Domain adaptation, Deep transfer learning	MIMII
A multi-source transfer learning network (MSTLN) structure is proposed in this study to aggregate and transmit diagnostic information from many sources. A multi-source diagnostic knowledge fusion module is used in conjunction with many partial distribution adaption sub-networks (PDA-Subnets) [4].	Intelligent fault diagnosis, Rotating machines, Multi-source transfer learning and Deep transfer learning	Bearing datasets, Planetary gearbox datasets
Using a reinforcement ensemble deep transfer learning network for defect detection with numerous sources is recommended (REDTLN) [74].	Multi-source domains and Reinforcement ensemble deep transfer network	Bearing dataset
This paper's primary objective was to use pre-trained deep transfer learning (DTL) structures and standard machine learning (ML) models as an automated method for diagnosing Parkinson's disease (PD) using sEMG data [2].	Deep transfer learning and Ensembling feature selection	Prosthetic fingers and Gait rhythmicity datasets.
This study explains the design process of the built-in discrete time-series convolution neural network (DTCNN) and provides a hierarchical technique for TRUs fault detection as well as a transfer learning-based fault diagnostic method rather than training new models for distinct TRUs [76].	Transformer rectifier units, Intelligent fault diagnosis, Convolutional neural network and Transfer learning.	Asy-24 and Sy-24.
Researchers are employing transfer learning to identify faults due to a shortage of fault data. This paper investigates the advantages of transfer learning for AI-based fault-detection problems [77].	Fault diagnosis, feature extraction, feature transfer and sensors.	Two data sets, denoted Group A and Group B, are used in the comparative experiments. Group A depicts the transition of the Hp1 fault prediction model to Hp2. Group B reflects the transition of the Hp3 fault prediction model to Hp4.
This approach can be immediately applied to transient data while preserving accuracy without the need for a steady-state detector, allowing for early defect diagnosis. The transformer design employs a new multi-head attention mechanism devoid of convolutional and recurrent layers, as is the case with standard deep learning techniques [78].	Transformer architecture and Deep learning method	N/A.
By assessing the attributes of great observations from sensors put in traction systems, the typical processes, difficulties that may limit future FDD installations are analyzed in detail, as are realistic high-speed trains. Using the theoretical advancements of data-driven FDD techniques, further enlightening insights on this topic are provided. Orally generated by embracing FDD-based model-based issues. High-speed train traction system approaches for system identification and new machine learning technologies that provide a range of interesting solutions to FDD strategies [79].	Data-driven, fault diagnosis and detection (FDD), traction systems, and high-speed trains.	N/A.
This study presents a domain generalization-based hybrid diagnostic network for deployment under unanticipated working settings in order to address this difficulty. Using both intrinsic and extrinsic generalization objectives, the deep network's discriminant structure is intended to be made more regular. This allows the diagnostic model to acquire robust traits and then apply them to previously unexplored areas [80].	Keywords: deep learning, domain generalization, intelligent failure detection, rotating equipment, and vibration signals. and a record of problems with the gearbox.	N/A.

5. Conclusions

This review provides an overview of the theory and practices of DTL techniques from the perspective of the algorithm. This review defines the terminologies for DTL and describes how TL technologies may be used to enhance the performance of DL models. Real-world industrial applications have also been taken into consideration while examining the state-of-the-art uses of DTL-based IFD techniques. The four key application scenarios—improving generalization performance, diagnosing flaws in partial domains, identifying faults that are just beginning to appear, and decoupling problems that are driven by other faults—have been defined and discussed in depth. Then, tactics for selecting DTL algorithms for a new IFD project are thoroughly outlined, and potential challenges and trends for the future are also examined. This article offers suggestions for what academics who wish to assist IFD in developing and improving may do.

References

1. Schrick, D. Remarks on terminology in the field of supervision, fault detection and diagnosis. *IFAC Proc.* **1997**, *30*, 959–964. [CrossRef]
2. Dong, Y.; Li, Y.; Zheng, H.; Wang, R.; Xu, M. 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.* **2022**, *121*, 327–348. [CrossRef]
3. Li, W.; Huang, R.; Li, J.; Liao, Y.; Chen, Z.; He, G.; Yan, R.; Gryllias, K. A perspective survey on deep transfer learning for fault diagnosis in industrial scenarios: Theories, applications and challenges. *Mech. Syst. Signal Process.* **2022**, *167*, 108487. [CrossRef]
4. Yang, B.; Xu, S.; Lei, Y.; Lee, C.G.; Stewart, E.; Roberts, C. Multi-source transfer learning network to complement knowledge for intelligent diagnosis of machines with unseen faults. *Mech. Syst. Signal Process.* **2022**, *162*, 108095. [CrossRef]
5. Gao, Z.; Cecati, C.; Ding, S.X. A survey of fault diagnosis and fault-tolerant techniques—Part I: Fault diagnosis with model-based and signal-based approaches. *IEEE Trans. Ind. Electron.* **2015**, *62*, 3757–3767. [CrossRef]
6. Lei, Y.; Lin, J.; Zuo, M.; He, Z. Condition monitoring and fault diagnosis of planetary gearboxes: A review. *Measurement* **2014**, *48*, 292–305. [CrossRef]
7. Fong, A.C.M.; Hui, S. An intelligent online machine fault diagnosis system. *Comput. Control Eng. J.* **2001**, *12*, 217–223. [CrossRef]
8. Li, B.; Chow, M.; Tipsuwan, Y.; Hung, J. Neural-network-based motor rolling bearing fault diagnosis. *IEEE Trans. Ind. Electron.* **2000**, *47*, 1060–1069. [CrossRef]
9. Muruganatham, B.; Sanjith, M.; Krishnakumar, B.; Murty, S. Roller element bearing fault diagnosis using singular spectrum analysis. *Mech. Syst. Signal Process.* **2013**, *35*, 150–166. [CrossRef]
10. Socher, R.; Huval, B.; Bath, B.; Manning, C.; Ng, A. Convolutional-Recursive Deep Learning for 3D Object Classification. *NIPS* **2012**, *3*, 8.
11. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [CrossRef]
12. Deng, L.; Li, J.; Huang, J.; Yao, K.; Yu, D.; Seide, F.; Seltzer, M.; Zweig, G.; He, X.; Williams, J.; et al. Recent advances in deep learning for speech research at Microsoft. In Proceedings of the 2013 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Vancouver, BC, Canada, 26–30 May 2013; pp. 8604–8608.
13. Hinton, G.; Deng, L.; Yu, D.; Dahl, G.; Mohamed, A.; Jaitly, N.; Senior, A.; Vanhoucke, V.; Nguyen, P.; Sainath, T.; et al. Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *IEEE Signal Process. Magaz.* **2012**, *29*, 82–97. [CrossRef]
14. Lu, C.; Wang, Z.; Qin, W.; Ma, J. Fault diagnosis of rotary machinery components using a stacked denoising autoencoder-based health state identification. *Signal Process.* **2017**, *130*, 377–388. [CrossRef]
15. Jia, F.; Lei, Y.; Lin, J.; Zhou, X.; Lu, N. Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data. *Mech. Syst. Signal Process.* **2016**, *72*, 303–315. [CrossRef]

16. Huijie, Z.; Ting, R.; Xinqing, W.; You, Z.; Husheng, F. Fault diagnosis of hydraulic pump based on stacked autoencoders. In Proceedings of the 2015 12th IEEE International Conference on Electronic Measurement & Instruments (ICEMI), Qingdao, China, 6–18 July 2015; Volume 1, pp. 58–62.

17. Guo, L.; Gao, H.; Huang, H.; He, X.; Li, S. Multifeatures fusion and nonlinear dimension reduction for intelligent bearing condition monitoring. *Shock Vib.* **2016**, *2016*, 4632562. [CrossRef]

18. Verma, N.; Gupta, V.; Sharma, M.; Sevakula, R. Intelligent condition based monitoring of rotating machines using sparse auto-encoders. In Proceedings of the 2013 IEEE Conference on Prognostics and Health Management (PHM), Gaithersburg, MD, USA, 24–27 June 2013; pp. 1–7.

19. Le Cun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. Gradient-based learning applied to document recognition. *Proc. IEEE* **1998**, *86*, 2278–2324. [CrossRef]

20. Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv* **2014**, arXiv:1409.1556.

21. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.

22. Szegedy, C.; Ioffe, S.; Vanhoucke, V.; Alemi, A. Inception-v4, Inception-Resnet In addition, The Impact of Residual Connections On Learning. *arXiv* **2016**, arXiv:1602.07261.

23. Janssens, O.; Slavkovikj, V.; Vervisch, B.; Stockman, K.; Loccufier, M.; Verstockt, S.; Van de Walle, R.; Van Hoecke, S. Convolutional neural network based fault detection for rotating machinery. *J. Sound Vib.* **2016**, *377*, 331–345. [CrossRef]

24. Abdeljaber, O.; Avci, O.; Kiranyaz, S.; Gabbouj, M.; Inman, D. Real-time vibration-based structural damage detection using one-dimensional convolutional neural networks. *J. Sound Vib.* **2017**, *388*, 154–170. [CrossRef]

25. Zhang, W.; Peng, G.; Li, C.; Chen, Y.; Zhang, Z. A new deep learning model for fault diagnosis with good anti-noise and domain adaptation ability on raw vibration signals. *Sensors* **2017**, *17*, 425. [CrossRef]

26. Gan, M.; Wang, C. Construction of hierarchical diagnosis network based on deep learning and its application in the fault pattern recognition of rolling element bearings. *Mech. Syst. Signal Process.* **2016**, *72*, 92–104. [CrossRef]

27. Li, C.; Sánchez, R.; Zurita, G.; Cerrada, M.; Cabrera, D. Fault diagnosis for rotating machinery using vibration measurement deep statistical feature learning. *Sensors* **2016**, *16*, 895. [CrossRef]

28. Tang, Y. Deep Learning Using Linear Support Vector Machines. *arXiv* **2013**, arXiv:1306.0239.

29. Shen, F.; Chen, C.; Yan, R.; Gao, R. Bearing fault diagnosis based on svd feature extraction and transfer learning classification. In Proceedings of the Prognostics and System Health Management Conference (PHM), Beijing, China, 21–23 October 2015; pp. 1–6.

30. Lu, W.; Liang, B.; Cheng, Y.; Meng, D.; Yang, J.; Zhang, T. Deep model based domain adaptation for fault diagnosis. *IEEE Trans. Ind. Electron.* **2016**, *64*, 2296–2305. [CrossRef]

31. Amar, M.; Gondal, I.; Wilson, C. Vibration spectrum imaging: A novel bearing fault classification approach. *IEEE Trans. Ind. Electron.* **2015**, *62*, 494–502. [CrossRef]

32. Combastel, C. Merging kalman filtering and zonotopic state bounding for robust fault detection under noisy environment. *IFAC-PapersOnLine* **2015**, *48*, 289–295. [CrossRef]

33. Yang, B.; Lei, Y.; Jia, F.; Xing, S. An intelligent fault diagnosis approach based on transfer learning from laboratory bearings to locomotive bearings. *Mech. Syst. Signal Process* **2019**, *122*, 692–706. 2018.12.051. [CrossRef]

34. Hoang, D.T.; Kang, H.J. A survey on deep learning based bearing fault diagnosis. *Neurocomputing* **2019**, *335*, 327–335. neucom.2018.06.078. [CrossRef]

35. Bently, D. Predictive maintenance through the monitoring and diagnostics of rolling element bearings. *Bently Nev. Co. Appl. Note* **1989**, *44*, 2–8.

36. Yin, S.; Li, X.; Gao, H.; Kaynak, O. Data-based techniques focused on modern industry: An overview. *IEEE Trans. Ind. Electron.* **2015**, *62*, 657–667. [CrossRef]

37. Dai, X.; Gao, Z. From model, signal to knowledge: A data-driven perspective of fault detection and diagnosis. *IEEE Trans. Ind. Inform.* **2013**, *9*, 2226–2638. [CrossRef]

38. Gao, Z.; Cecati, C.; Ding, S.X. A survey of fault diagnosis and fault-tolerant techniques-part II: Fault diagnosis with knowledge-based and hybrid/active approaches. *IEEE Trans. Ind. Electron.* **2015**, *62*, 3768–3774. [CrossRef]

39. Liu, R.; Yang, B.; Zio, E.; Chen, X. Artificial intelligence for fault diagnosis of rotating machinery: A review. *Mech. Syst. Signal Process.* **2018**, *108*, 33–47. [CrossRef]

40. Li, Y.; Wang, X.; Liu, Z.; Liang, X.; Si, S. The entropy algorithm and its variants in the fault diagnosis of rotating machinery: A review. *IEEE Access* **2018**, *6*, 66723–66741. [CrossRef]

41. Wang, Y.; Xiang, J.; Markert, R.; Liang, M. Spectral kurtosis for fault detection, diagnosis and prognostics of rotating machines: A review with applications. *Mech. Syst. Signal Process.* **2016**, *66–67*, 679–698. [CrossRef]

42. Li, Y.; Xu, M.; Liang, X.; Huang, W. Application of bandwidth EMD and adaptive multiscale morphology analysis for incipient fault diagnosis of rolling bearings. *IEEE Trans. Ind. Electron.* **2017**, *64*, 6506–6517. [CrossRef]

43. Lei, Y.; Jia, F.; Lin, J.; Xing, S.; Ding, S.X. An intelligent fault diagnosis method using unsupervised feature learning towards mechanical big data. *IEEE Trans. Ind. Electron.* **2016**, *63*, 3137–3147. [CrossRef]

44. Khan, S.; Yairi, T. A review on the application of deep learning in system health management. *Mech. Syst. Signal Process.* **2018**, *107*, 241–265. [CrossRef]

45. Hu, X.; Zhang, K.; Liu, K.; Lin, X.; Dey, S.; Onori, S. Advanced fault diagnosis for lithium-ion battery systems. *IEEE Ind. Electron. Mag.* **2020**, *14*, 65–91. [CrossRef]

46. Lei, Y.; Yang, B.; Jiang, X.; Jia, F.; Li, N.; Nandi, A.K. Applications of machine learning to machine fault diagnosis: A review and roadmap. *Mech. Syst. Signal Process.* **2020**, *138*, 106587. [CrossRef]

47. Li, X.; Li, X.; Ma, H. Deep representation clustering-based fault diagnosis method with unsupervised data applied to rotating machinery. *Mech. Syst. Signal Process.* **2020**, *143*, 106825. [CrossRef]

48. Meng, Z.; Guo, X.; Pan, Z.; Sun, D.; Liu, S. Data segmentation and augmentation methods based on raw data using deep neural networks approach for rotating machinery fault diagnosis. *IEEE Access* **2019**, *7*, 79510–79522. [CrossRef]

49. Hu, T.; Tang, T.; Lin, R.; Chen, M.; Han, S.; Wu, J. A simple data aug- mentation algorithm and a self-adaptive convolutional architecture for few-shot fault diagnosis under different working conditions. *Measurement* **2020**, *156*, 107539. [CrossRef]

50. Zhao, D.; Liu, S.; Gu, D.; Sun, X.; Wang, L.; Wei, Y.; Zhang, H. Enhanced data-driven fault diagnosis for machines with small and unbalanced data based on variational auto-encoder. *Meas. Sci. Technol.* **2019**, *31*, 035004. [CrossRef]

51. Ding, Y.; Ma, L.; Ma, J.; Wang, C.; Lu, C. A generative adversarial network based intelligent fault diagnosis method for rotating machinery under small sample size conditions. *IEEE Access* **2019**, *7*, 149736–149749. [CrossRef]

52. Wang, Z.; Wang, J.; Wang, Y. An intelligent diagnosis scheme based on generative adversarial learning deep neural networks and its application to planetary gearbox fault pattern recognition. *Neurocomputing* **2018**, *310*, 213–222. [CrossRef]

53. Liu, J.; Qu, F.; Hong, X.; Zhang, H. A small-sample wind turbine fault de- tection method with synthetic fault data using generative adversarial nets. *IEEE Trans. Ind. Inform.* **2019**, *15*, 3877–3888. [CrossRef]

54. Xiao, D.; Huang, Y.; Qin, C.; Liu, Z.; Li, Y.; Liu, C. Transfer learning with convolutional neural networks for small sample size problem in machinery fault diagnosis. *Proc. Inst. Mech. Eng. C* **2019**, *233*, 5131–5143. [CrossRef]

55. Kim, H.; Youn, B.D. A New parameter repurposing method for parameter transfer with small dataset and its application in fault diagnosis of rolling element bearings. *IEEE Access* **2019**, *7*, 46917–46930. [CrossRef]

56. Cao, P.; Zhang, S.; Tang, J. Preprocessing-free gear fault diagnosis using small datasets with deep convolutional neural network-based transfer learning. *IEEE Access* **2018**, *6*, 26241–26253. [CrossRef]

57. Yu, C.; Ning, Y.; Qin, Y.; Su, W.; Zhao, X. Multi-label fault diagnosis of rolling bearing based on meta-learning. *Neural. Comput. Appl.* **2020**, *33*, 5393–5407. [CrossRef]

58. Wang, S.; Wang, D.; Kong, D.; Wang, J.; Li, W.; Zhou, S. Few-shot rolling bearing fault diagnosis with metric-based meta learning. *Sensors* **2020**, *20*, 6437. [CrossRef] [PubMed]

59. Wu, J.; Zhao, Z.; Sun, C.; Yan, R.; Chen, X. Few-shot transfer learning for intelligent fault diagnosis of machine. *Measurement* **2020**, *166*, 108202. [CrossRef]

60. Plakias, S.; Boutalis, Y.S. Fault detection and identification of rolling element bearings with Attentive Dense CNN. *Neurocomputing* **2020**, *405*, 208–217. [CrossRef]

61. Mishra, C.; Samantaray, A.K.; Chakraborty, G. Ball bearing defect models: A study of simulated and experimental fault signatures. *J. Sound. Vib.* **2017**, *400*, 86–112. [CrossRef]

62. Zhang, W.; Li, C.; Peng, G.; Chen, Y.; Zhang, Z. A deep convolutional neural network with new training methods for bearing fault diagnosis under noisy environments and different working loads. *Mech. Syst. Signal Process.* **2018**, *100*, 439–453. [CrossRef]

63. Purbowaskito, W.; Wu, P.-Y.; Lan, C.-Y. Permanent Magnet Synchronous Motor Driving Mechanical Transmission Fault Detection and Identification: A Model-Based Diagnosis Approach. *Electronics* **2022**, *11*, 1356. [CrossRef]

64. He, Y.; Tang, H.; Ren, Y.; Kumar, A. A deep multi-signal fusion adversarial model based transfer learning and residual network for axial piston pump fault diagnosis. *Measurement* **2022**, *2022*, 110889. [CrossRef]

65. Im, S.-H.; Gu, B.-G. Study of Induction Motor Inter-Turn Fault Part II: Online Model-Based Fault Diagnosis Method. *Energies* **2022**, *15*, 977. [CrossRef]

66. Ayas, S.; Ayas, M.S. A novel bearing fault diagnosis method using deep residual learning network. *Multimed. Tools Appl.* **2022**, *81*, 22407–22423. [CrossRef]

67. Ahmad, M.; Mohd-Mokhtar, R. A Survey on Model-based Fault Detection Techniques for Linear Time-Invariant Systems with Numerical Analysis. *Pertanika J. Sci. Technol.* **2022**, *30*. [CrossRef]

68. Liu, Z.; Fang, L.; Jiang, D.; Qu, R. A machine-learning based fault diagnosis method with adaptive secondary sampling for multiphase drive systems. *IEEE Trans. Power Electron.* **2022**, *37*, 8767–8772. [CrossRef]

69. Koizumi, Y.; Saito, S.; Uematsu, H.; Harada, N.; Imoto, K. ToyADMOS: A dataset of miniature-machine operating sounds for anomalous sound detection. In Proceedings of the 2019 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA), New Paltz, NY, USA, 20–23 October 2019; pp. 313–317.

70. Purohit, H.; Tanabe, R.; Ichige, K.; Endo, T.; Nikaido, Y.; Suefusa, K.; Kawaguchi, Y. MIMII Dataset: Sound dataset for malfunctioning industrial machine investigation and inspection. *arXiv* **2019**, arXiv:1909.09347.

71. Harada, N.; Niizumi, D.; Takeuchi, D.; Ohishi, Y.; Yasuda, M.; Saito, S. ToyADMOS2: Another dataset of miniature-machine operating sounds for anomalous sound detection under domain shift conditions. *arXiv* **2021**, arXiv:2106.02369.

72. Tanabe, R.; Purohit, H.; Dohi, K.; Endo, T.; Nikaido, Y.; Nakamura, T.; Kawaguchi, Y. MIMII DUE: Sound dataset for malfunctioning industrial machine investigation and inspection with domain shifts due to changes in operational and environmental conditions. In Proceedings of the 2021 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA), New Paltz, NY, USA, 17–20 October 2021; pp. 21–25.

73. Qiu, H.; Lee, J.; Lin, J. Wavelet Filter-based Weak Signature Detection Method and its Application on Roller Bearing Prognostics. *J. Sound Vib.* **2006**, *289*, 1066–1090 [CrossRef]

74. Rezaee, K.; Savarkar, S.; Yu, X.; Zhang, J. A hybrid deep transfer learning-based approach for Parkinson's disease classification in surface electromyography signals. *Biomed. Signal Process. Control* **2022**, *71*, 103161. [CrossRef]

75. Chen, S.; Ge, H.; Li, H.; Sun, Y.; Qian, X. Hierarchical deep convolution neural networks based on transfer learning for transformer rectifier unit fault diagnosis. *Measurement* **2021**, *167*, 108257. [CrossRef]

76. Qian, Q.; Qin, Y.; Wang, Y.; Liu, F. A new deep transfer learning network based on convolutional auto-encoder for mechanical fault diagnosis. *Measurement* **2021**, *178*, 109352. [CrossRef]

77. Liu, J.; Zhang, Q.; Li, X.; Li, G.; Liu, Z.; Xie, Y.; Li, K.; Liu, B. Transfer learning-based strategies for fault diagnosis in building energy systems. *Energy Build.* **2021**, *250*, 111256. [CrossRef]

78. Li, Y.; Jiang, W.; Zhang, G.; Shu, L. Wind turbine fault diagnosis based on transfer learning and convolutional autoencoder with small-scale data. *Renew. Energy* **2021**, *171*, 103–115. [CrossRef]

79. Han, T.; Liu, C.; Wu, R.; Jiang, D. Deep transfer learning with limited data for machinery fault diagnosis. *Appl. Soft Comput.* **2021**, *103*, 107150. [CrossRef]

80. Li, X.; Zhang, W.; Ma, H.; Luo, Z.; Li, X. Partial transfer learning in machinery cross-domain fault diagnostics using class-weighted adversarial networks. *Neural Networks* **2020**, *129*, 313–322. [CrossRef] [PubMed]

81. Li, C.; Zhang, S.; Qin, Y.; Estupinan, E. A systematic review of deep transfer learning for machinery fault diagnosis. *Neurocomputing* **2020**, *407*, 121–135. [CrossRef]

82. Qian, W.; Li, S.; Yi, P.; Zhang, K. A novel transfer learning method for robust fault diagnosis of rotating machines under variable working conditions. *Measurement* **2019**, *138*, 514–525. [CrossRef]

83. Hasan, M.J.; Islam, M.; Kim, J.-M. Acoustic spectral imaging and transfer learning for reliable bearing fault diagnosis under variable speed conditions. *Measurement* **2019**, *138*, 620–631. [CrossRef]

84. Zhao, Z.; Zhang, Q.; Yu, X.; Sun, C.; Wang, S.; Yan, R.; Chen, X. Applications of unsupervised deep transfer learning to intelligent fault diagnosis: A survey and comparative study. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 1–28. [CrossRef]

85. Xu, Y.; Sun, Y.; Liu, X.; Zheng, Y. A digital-twin-assisted fault diagnosis using deep transfer learning. *IEEE Access* **2019**, *7*, 19990–19999. [CrossRef]

86. Yang, B.; Lei, Y.; Jia, F.; Li, N.; Du, Z. A polynomial kernel induced distance metric to improve deep transfer learning for fault diagnosis of machines. *IEEE Trans. Ind. Electron.* **2019**, *67*, 9747–9757. [CrossRef]

87. Baraldi, P.; Cannarile, F.; Di Maio, F.; Zio, E. Hierarchical k-nearest neighbours classification and binary differential evolution for fault diagnostics of automotive bearings operating under variable conditions. *Eng. Appl. Artif. Intell.* **2016**, *56*, 1–13. [CrossRef]

88. Jia, F.; Lei, Y.; Guo, L.; Lin, J.; Xing, S. A neural network constructed by deep learning technique and its application to intelligent fault diagnosis of machines. *Neurocomputing* **2018**, *272*, 619–628. [CrossRef]

89. Pons-Llinares, J.; Antonino-Daviu, J.A.; Riera-Guasp, M.; Lee, S.B.; Kang, T.-J.; Yang, C. Advanced induction motor rotor fault diagnosis via continuous and discrete time–frequency tools.*IEEE Trans. Ind. Electron.* **2014**, *62*, 1791–1802. [CrossRef]

90. Wen, L.; Li, X.; Gao, L.; Zhang, Y. A new convolutional neural network-based data-driven fault diagnosis method. *IEEE Trans. Ind. Electron.* **2017**, *65*, 5990–5998. [CrossRef]

91. Xia, M.; Li, T.; Xu, L.; Liu, L.; De Silva, C.W. Fault diagnosis for rotating machinery using multiple sensors and convolutional neural networks. *IEEE/ASME Trans. Mechatronics* **2017**, *23*, 101–110. [CrossRef]

92. Guo, L.; Lei, Y.; Xing, S.; Yan, T.; Li, N. Deep convolutional transfer learning network: A new method for intelligent fault diagnosis of machines with unlabeled data. *IEEE Transactions on Industrial Electronics* **2018**, *66*, 7316–7325. [CrossRef]

93. Wen, L.; Gao, L.; Li, X. A new deep transfer learning based on sparse auto-encoder for fault diagnosis. *IEEE Transactions on systems, man, and cybernetics: Systems* **2017**, *49*, 136–144. [CrossRef]

94. Eren, L. Bearing fault detection by one-dimensional convolutional neural networks. *Math. Probl. Eng.* **2017**, *2017*, 1–9. [CrossRef]

95. Chen, Z.; Mauricio, A.; Li, W.; Gryllias, K. A deep learning method for bearing fault diagnosis based on cyclic spectral coherence and convolutional neural networks. *Mechanical Systems and Signal Processing* **2020**, *140*, 106683. [CrossRef]

96. Xu, Y.; Li, Z.; Wang, S.; Li, W. Thompson Sarkodie-Gyan, and Shizhe Feng. A hybrid deep-learning model for fault diagnosis of rolling bearings. *Measurement* **2021**, *169*, 108502. [CrossRef]

97. Khorram, A.; Khalooei, M.; Rezghi, M. End-to-end CNN+ LSTM deep learning approach for bearing fault diagnosis. *Appl. Intell.* **2021**, *51*, 736–751. [CrossRef]

98. Venables, W.N.; Ripley, B.D. *Modern Applied Statistics with S-PLUS*; Springer Science& Business Media: Berlin/Heidelberg, Germany, 2013.

99. Kankar, P.K.; Sharma, S.; Harsha, S.P. Fault diagnosis of ball bearings using machine learning methods. *Expert Syst. Appl.* **2011**, *38*, 1876–1886. [CrossRef]

100. Cerrada, M.; Zurita, G.; Cabrera, D.; Sánchez, R.V.; Artés, M.; Li, C. Fault diagnosis in spur gears based on genetic algorithm and random forest. *Mech. Syst. Signal Process.* **2016**, *70*, 87–103. [CrossRef]

101. Widodo, A.; Yang, B.-S. Support vector machine in machine condition monitoring and fault diagnosis. *Mech. Syst. Signal Process.* **2007**, *21*, 2560–2574. [CrossRef]

102. Tong, C.; Wang, S.; Selesnick, I.; Yan, R.; Chen, X. Ridge-aware weighted sparse time-frequency representation. *IEEE Trans. Signal Process.* **2020**, *69*, 136–149. [CrossRef]

103. Min, S.; Lee, B.; Yoon, S. Deep learning in bioinformatics. *Briefings Bioinform.* **2017**, *18*, 851–869. [CrossRef] [PubMed]

104. Krizhevsky, A.; Sutskever, I.; Hinton, G. Imagenet classification with deep convolutional neural networks. *Adv. Neural Inf. Process. Syst.* **2012**, *60*, 1097–1105. [CrossRef]

105. Li, X.; Zhang, W.; Ding, Q.; Sun, J.-Q. Multi-layer domain adaptation method for rolling bearing fault diagnosis. *Signal Process.* **2019**, *157*, 180–197. [CrossRef]

106. Zhang, R.; Tao, H.; Wu, L.; Guan, Y. Transfer learning with neural networks for bearing fault diagnosis in changing working conditions. *IEEE Access* **2017**, *5*, 14347–14357. [CrossRef]

107. Zhang, C.; Xu, L.; Li, X.; Wang, H. A method of fault diagnosis for rotary equipment based on deep learning. In Proceedings of the 2018 Prognostics and System Health Management Conference (PHMChongqing), Chongqing, China, 26–28 October 2018; pp. 958–962.

108. Chen, D.; Yang, S.; Zhou, F. Incipient fault diagnosis based on dnn with transfer learning. In Proceedings of the 2018 International Conference on Control, Automation and Information Sciences (ICCAIS), Hangzhou, China, 24–27 October 2018; pp. 303–308.

109. Golafshan, R.; Sanliturk, K.Y. SVD and Hankel matrix based de-noising approach for ball bearing fault detection and its assessment using artificial faults. *Mech. Syst. Signal Process.* **2016**, *70*, 36–50. [CrossRef]

110. Ho, R.; Hung, K. A comparative investigation of mode mixing in EEG decomposition using EMD, EEMD and M-EMD. In Proceedings of the 2020 IEEE 10th Symposium on Computer Applications & Industrial Electronics (ISCAIE), Penang, Malaysia, 18–19 April 2020; pp. 203–210.

111. Zabin, M.; Choi, H.J.; Uddin, J. Hybrid deep transfer learning architecture for industrial fault diagnosis using Hilbert transform and DCNN–LSTM. *J. Supercomput.* **2022**, *2022*, 1–20. [CrossRef]

112. Islam, R.; Uddin, J.; Kim, J.M. Texture analysis based feature extraction using Gabor filter and SVD for reliable fault diagnosis of an induction motor. *Int. J. Inf. Technol. Manag.* **2018**, *17*, 20–32. [CrossRef]

113. Uddin, J.; Islam, M.R.; Kim, J.M.; Kim, C.H. A two-dimensional fault diagnosis model of induction motors using a gabor filter on segmented images. *Int. J. Control. Autom.* **2016**, *9*, 11–22. [CrossRef]

114. Uddin, J.; Van, D.N.; Kim, J.M. Accelerating 2D Fault Diagnosis of an Induction Motor using a Graphics Processing Unit. *Int. J. Multimed. Ubiquitous Eng.* **2015**, *10*, 341–352. [CrossRef]

115. Uddin, J.; Nguyen, D.; Kim, J. A reliable fault detection and classification model of induction motors using texture features and multi-class support vector machines.*J. Math. Probl. Eng.* **2014**, *2014*, 1–9.

116. Uddin, J.; Islam, R.; Kim, J.M. Texture feature extraction techniques for fault diagnosis of induction motors. *JoC* **2014**, *5*, 15–20.