Review

Few-Shot Learning Approaches for Fault Diagnosis Using Vibration Data: A Comprehensive Review

Xiaoxia Liang [1,2], Ming Zhang [3,*], Guojin Feng [1,2,*], Duo Wang [4], Yuchun Xu [3] and Fengshou Gu [5]

1 College of Mechanical Engineering, Hebei University of Technology, Tianjin 300401, China
2 Advanced Equipment Research Institute Co., Ltd. of HEBUT, Tianjin 300401, China
3 College of Engineering and Physical Sciences, Aston University, Birmingham B4 7ET, UK
4 Beijing Institute of Control Engineering, Beijing 100190, China
5 Centre for Efficiency and Performance Engineering, University of Huddersfield, Huddersfield HD1 3DH, UK
* Correspondence: m.zhang21@aston.ac.uk (M.Z.); guojin.feng@outlook.com (G.F.)

Abstract: Fault detection and diagnosis play a crucial role in ensuring the reliability and safety of modern industrial systems. For safety and cost considerations, critical equipment and systems in industrial operations are typically not allowed to operate in severe fault states. Moreover, obtaining labeled samples for fault diagnosis often requires significant human effort. This results in limited labeled data for many application scenarios. Thus, the focus of attention has shifted towards learning from a small amount of data. Few-shot learning has emerged as a solution to this challenge, aiming to develop models that can effectively solve problems with only a few samples. This approach has gained significant traction in various fields, such as computer vision, natural language processing, audio and speech, reinforcement learning, robotics, and data analysis. Surprisingly, despite its wide applicability, there have been limited investigations or reviews on applying few-shot learning to the field of mechanical fault diagnosis. In this paper, we provide a comprehensive review of the relevant work on few-shot learning in mechanical fault diagnosis from 2018 to September 2023. By examining the existing research, we aimed to shed light on the potential of few-shot learning in this domain and offer valuable insights for future research directions.

Keywords: few-shot learning; meta-learning; metric-based meta-learning; vibration signal; fault diagnosis

1. Introduction

Rotating machinery, such as pumps, fans, generators, and compressors, is one of the most commonly used types of equipment in industrial production. The faults in these devices can lead to production downtime, losses, and safety accidents. Therefore, the timely and accurate diagnosis of faults in rotating machinery and its critical components is crucial for ensuring production safety and improving efficiency. Modern technology has advanced to the point where faults can be diagnosed using various signals, such as vibration, sound, temperature, and current. These techniques assist engineers in promptly detecting faults and taking measures to prevent production accidents and losses.

In the field of fault diagnosis for rotating machinery, machine learning and deep learning methods have been widely applied. For example, support vector machines (SVM) [1]; decision trees [2]; random forests [3]; and various shallow or deep neural networks, including convolutional neural networks (CNNs) [4], recurrent neural networks (RNNs) [5], long short-term memory (LSTM) networks [6], and autoencoders (AEs) [7], can automatically extract features from the signals of rotating machinery and perform fault classification and prediction. Among these methods, deep learning, with its powerful feature extraction and representation capabilities, is one of the important directions in current fault diagnosis research. Despite the significant successes of deep learning in many domains, it still has limitations. One of them is that deep learning models usually require a large amount of training data to achieve high performance. In the industrial

domain, the problem of a paucity of samples is common and challenging. In fields like fault diagnosis for rotating machinery, it is difficult to obtain a large number of fault samples due to the complexity of the equipment, the diversity of fault types, and the challenges in data collection. In such cases, few-shot learning methods can learn useful knowledge from limited data, thereby improving the model's generalization ability and prediction performance, reducing computational resource requirements, and better addressing real-world problems to enhance equipment operational efficiency and safety.

Few-shot learning has already achieved notable results in areas such as natural language processing and image classification. Several relevant review articles [8–12] can be found. Lu et al. [8] introduced the applications and progress of few-shot learning in fields such as image classification, object detection, semantic segmentation, and natural language processing. They categorized the main methods of few-shot learning into two categories: generative-model-based methods and discriminative-model-based methods. Wang et al. [9] classified few-shot learning from the perspectives of data, models, and algorithms. From the data perspective, it involves expanding the number of samples to increase the sample size. From the model perspective, it involves modeling with limited data by restricting model complexity and reducing the hypothesis space. From the algorithm perspective, it involves improving the search methods for the optimal solution in the hypothesis space. Regarding few-shot fault diagnosis, Pan et al. [10] reviewed 13 small-shot fault diagnosis methods based on generative adversarial networks (GANs) and categorized them into three types: deep GANs for data augmentation, adversarial training for transfer learning, and other methods. However, this review only focused on GANs and did not cover other generative models and few-shot fault diagnosis methods. Zhang et al. [11] classified fault diagnosis methods for small and imbalanced datasets according to the fault diagnosis process, dividing them into strategies based on data augmentation, feature extraction, and classifier design. Bhuiyan et al. [12] evaluated the current state of vision inspection systems by exploring various vision-based methods for defect detection and identification using vibration acoustic sensor data, with a focus on image processing, deep learning, machine learning, transfer learning, few-shot learning, and light-weight approaches. Continuing the work of previous researchers, this paper focuses on a comprehensive analysis of few-shot learning methods applied in the domain of mechanical fault diagnosis using vibration signals. The study particularly concentrates on advancements made within the past five years, spanning from 2018 to September 2023.

Our aim was to compile and synthesize the current state of knowledge on fault diagnosis using vibration data under conditions of extremely limited samples. By providing a comprehensive review of relevant work spanning from 2018 to September 2023, the article highlights the potential of few-shot learning in this domain. It not only fills a notable gap in the literature but also offers valuable insights for future research directions in mechanical fault diagnosis.

2. The Systematic Review Process

This section presents an outline of the literature review procedure concerning the implementation of few-shot learning in mechanical fault diagnosis. The search for relevant literature predominantly occurred in two widely recognized academic databases, namely Web of Science and Scopus. These databases encompass a substantial collection of interdisciplinary research papers subjected to rigorous peer review, rendering them well-suited for unearthing a significant volume of studies on the subject of few-shot learning.

To avoid overlapping papers that were double-counted, we employed a systematic approach during our database search and subsequent paper selection process. Firstly, we conducted a thorough search in both databases to identify relevant papers. In this stage, there existed overlapping papers between these two databases. Upon identifying potentially relevant papers, we used reference management software to carefully compare the papers from both databases. In cases where we encountered the same paper in both Web of Science and Scopus, we ensured that it was counted as a single reference rather than being

double-counted. In addition, we checked the abstract, introduction, and conclusion of each paper to make sure that all the papers we counted met our requirements. These processes allowed us to maintain the accuracy of our reference list and eliminate any duplications.

Specifically, during the literature review process, the following keywords were first used for the search: "Few-shot learning & fault diagnosis & vibration" and "Small sample learning & fault diagnosis & vibration", with a time span from 2018 to 30 September 2023. However, we identified more than 200 reference papers from the two databases. To be precise, using keywords "Small sample learning & fault diagnosis & vibration" and a time span from 2018 to 30 September 2023, we identified 102 papers from Scopus and 183 papers from Web of Science; using keywords "Few-shot learning & fault diagnosis & vibration" and the same time span, we identified 24 papers from Scopus and 46 papers from Web of Science.

Considering that few-shot learning refers to a machine learning paradigm in which a model is trained to recognize and generalize from a very limited number of examples or data points, for many papers, the number of examples used was one, five, or under twenty-five [13,14]. On the other hand, small sample learning does not include a clear definition regarding the sample size. Therefore, we ultimately applied the keywords "Few-shot learning & fault diagnosis & vibration" with a time span from 2018 to 30 September 2023.

In the next step, inclusion and exclusion criteria were established to systematically narrow down the scope and ensure high-quality evaluation. Work-in-progress papers, preprints, and other non-peer-reviewed publications were excluded. Only peer-reviewed journals and conference proceedings were considered, and works that did not align with the target domain or went beyond the defined scope were filtered based on abstracts and article browsing. The retrieved literature was manually filtered, and relevant target articles were selected. Specifically, the focus was solely on the application of few-shot learning in mechanical fault diagnosis, excluding its application in other fields, such as medicine. Regarding the type of signals considered, only vibration signals or multimodal signals containing vibrations were taken into account, while other signals such as images or speech were disregarded. After the screening process, a total of 41 papers were retained as references.

Conducting the literature search in the manner mentioned above, it was observed that few-shot learning can be applied to various types of mechanical equipment in the field of fault diagnosis. These mainly include bearings, gearboxes, engines, pumps, suspension systems on trains, and pipelines. The distribution of few-shot learning applications in different types of mechanical equipment is shown in Table 1.

Table 1. The distribution of few-shot learning applications in different types of mechanical equipment.

Application	Num. of Papers
Bearings	26
Bearings and gears	2
Bearings for electric machines	2
Bearings on machine tools	1
Freight train bearings	1
Magnetic flux leakage in a pipeline and bearings	1
Wind turbines (bearings and gearboxes)	1
Electrohydrostatic actuators	1
Electromechanical actuators	1
Industrial rotating machinery	1
Suspension systems on trains	1
Aviation hydraulic pumps	1
Autonomous underwater vehicles	1
Industrial robots	1

In addition to these classifications, a simple statistical analysis over time was performed, the results of which are shown in Figure 1. The statistics of the applied few-shot learning methods are show in Figure 2. From the figure, it can be observed that meta-learning [15–17] is a prominent area of interest in few-shot learning methods, with over 50% of the articles focused on meta-learning. Within the domain of meta-learning, MAML (model-agnostic meta-learning) and its variations, along with metric-based meta-learning methods, collectively accounted for approximately 90% of the research contributions.
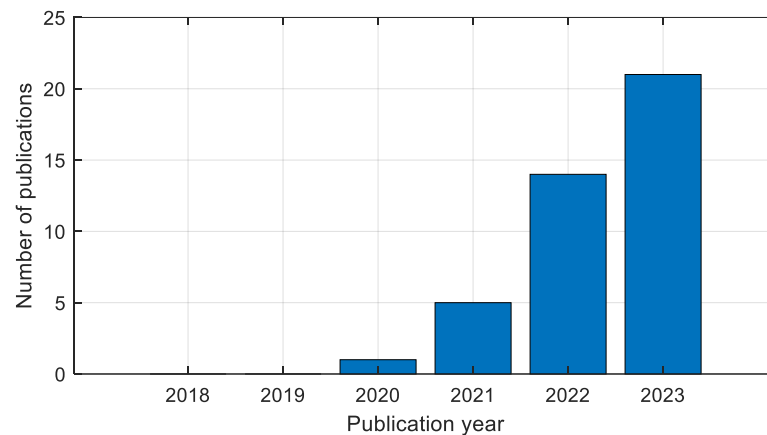

Figure 1. Number of publications by application of few-shot learning in machine fault diagnosis field using vibration data from 2018 to September 2023.
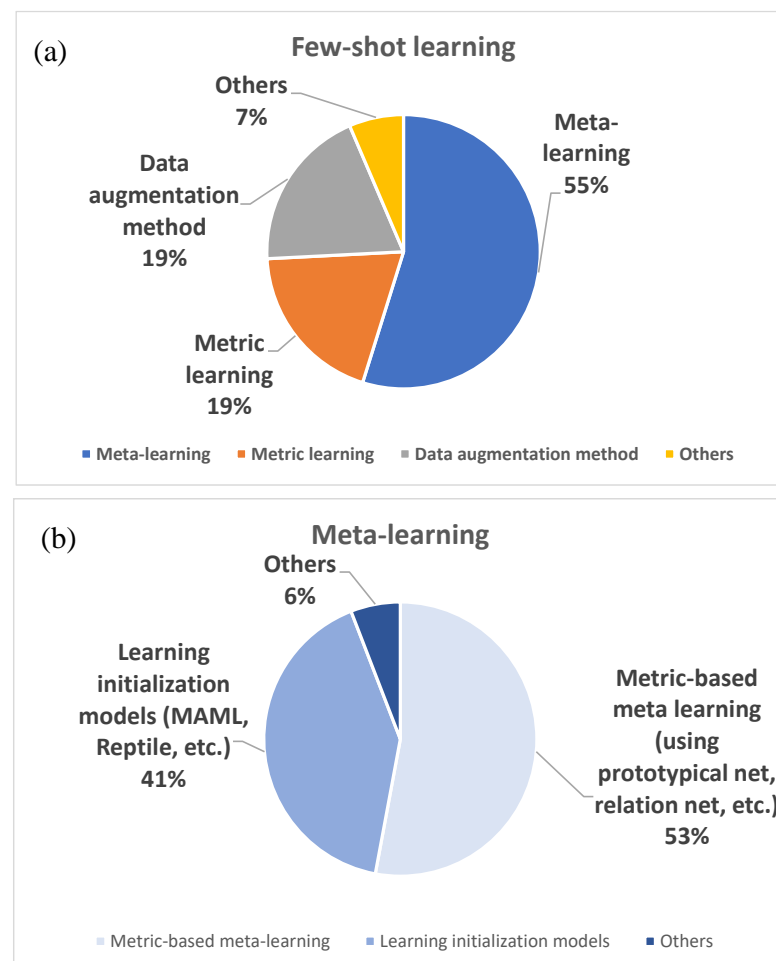

Figure 2. *Cont.*
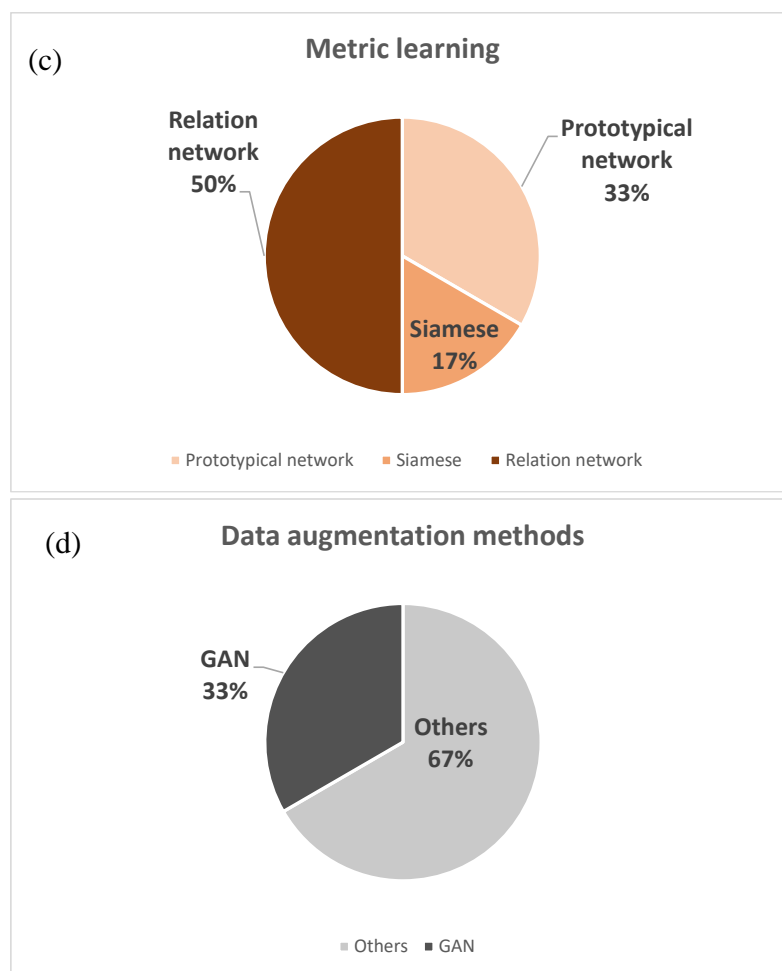
Figure 2. Statistics of few-shot learning methods in machine fault diagnosis field using vibration data from 2018 to September 2023: (**a**) few-shot learning methods; (**b**) meta-learning; (**c**) metric learning; (**d**) data augmentation methods.

The statistics regarding the different journals in which the reviewed papers were published are presented in Table 2. It can be seen from the figure that, from 2019 to 2023, the majority of research papers on few-shot fault diagnosis were published in Measurement Science and Technology, followed by ISA Transactions, Neural Computing and Applications, Measurement, IEEE Transactions on Industrial Informatics, IEEE Transactions on Instrumentation and Measurement, and Sensors. Other papers were also recognized by peers in other high-quality manufacturing journals and conferences.

Table 2. Number of publications by journal for few-shot learning in machine fault diagnosis field using vibration data from 2018 to 2023.

Journal Name	Number of Publications
Measurement Science and Technology	4
ISA Transactions	3
Neural Computing and Applications	2
Measurement	2
IEEE Transactions on Industrial Informatics	2
IEEE Transactions on Instrumentation and Measurement	2
Sensors	2

3. Methods for Few-Shot Fault Diagnosis in Mechanical Systems

This section provides a detailed overview of the methods used in few-shot learning for machine fault diagnosis, specifically utilizing vibration data. We organized the content of this section according to the results presented in Figure 2 of Section 2. The approaches to few-shot learning are categorized into meta-learning, metric-based learning, data augmentation, and other methods. Since researchers have predominantly employed meta-learning algorithms in their studies, this section further subdivides meta-learning and primarily focuses on the application of metric-based meta-learning and learning initialization methods.

3.1. Meta-Learning

The idea of meta-learning was proposed as early as the 1990s [8,18]. Serving as a primary approach in few-shot learning, meta-learning, or learning to learn, involves acquiring knowledge from multiple tasks and adapting to new ones. It aims to learn task-level knowledge instead of individual sample-level information, promoting task-agnostic learning systems over task-specific models. Meta-learning consists of two stages: meta-training and meta-testing. In the meta-training stage, the meta-learner learns from various tasks to capture underlying patterns or knowledge. In the meta-testing stage, the meta-learner adapts its parameters based on new tasks, leveraging the meta-knowledge to make predictions swiftly using available labeled data. In our investigations, meta-learning was classified into metric-based meta-learning, learning initialization methods, and others.

As shown in Figure 3a, machine learning can be summarized in three steps: The first step involves defining a function with unknown parameters θ, where these unknown parameters can represent weights and biases. The second step entails formulating a loss function concerning the unknown parameters θ. The third step involves an optimization process to find a θ that minimizes the loss. Gradient descent is one of the most commonly used optimization methods.
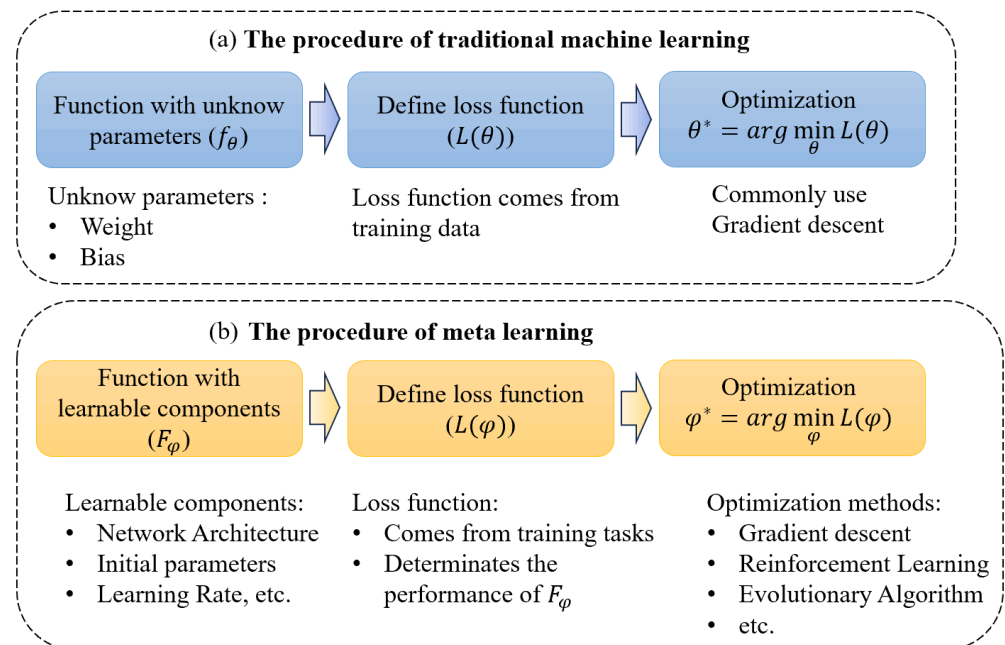
Figure 3. The differences between traditional learning and meta-learning: (**a**) the procedure of traditional machine learning, (**b**) the procedure of meta-learning.

The process of meta-learning is shown in Figure 3b. Unlike traditional machine learning, meta-learning involves learning how to learn. Learning itself can be viewed as a function, represented as a learning algorithm denoted by *F*. This function takes a set of data as the input and produces a trained learning model as the output. For example, in a fault

diagnosis task, the output would be a classifier. When test data are fed into the previously obtained classifier, it can classify the faults accordingly.

As depicted in the diagram, the first step in meta-learning is to define the learning algorithm F_φ, where the learnable components are denoted as φ. While traditional machine learning often focuses on learning weights and biases, in meta-learning, these learnable components can include network architectures, initial parameters, learning rates, and more. Different meta-learning methods aim to learn different learnable components.

The second step involves defining the loss function for the learning algorithm F_φ. In traditional machine learning, the loss function is derived from the training data, whereas in meta-learning, the loss function is derived from training tasks. This means that the loss function encapsulates the learning algorithm's ability to adapt and generalize across different tasks.

The third step is to use an optimization method to minimize the loss function. The choice of optimization algorithm depends on the nature of the learnable components φ. If the gradient $\partial L(\varphi)/\partial \varphi$ is computable, gradient descent can be used. However, if the gradient $\partial L(\varphi)/\partial \varphi$ is not computable, reinforcement learning algorithms or evolutionary algorithms can be employed to find solutions [19,20].

In summary, meta-learning aims to enhance the efficiency and adaptability of learning algorithms by enabling them to learn from multiple tasks and generalize better to new, unseen tasks. The optimization of the learnable components plays a crucial role in achieving this objective. According to Figure 2, meta-learning was further classified into metric-based meta-learning, learning initialization methods, and others.

3.1.1. Metric-Based Meta-Learning

Metric-based meta-learning is a specific type of meta-learning approach that utilizes metric learning to achieve rapid adaptation to new tasks. In metric-based meta-learning, it is customary to employ a metric learning method, such as a prototypical network, to learn an embedding space in which the similarity between tasks is associated with the distances between samples in the embedding space. The term "metric" in this context refers to the measurement of similarity between tasks, rather than the distance between data samples.

A typical metric-based meta-learning method for fault diagnosis is shown in Figure 4. In the initial stage, supervised learning is conducted on source domain data to train a classification model (upper branch). Subsequently, the feature extractor is kept fixed, and episodic training is performed on the metric embedding module using a set of few-shot tasks randomly sampled from the source domain. Finally, the trained feature extractor and metric embedding are applied to perform testing on target tasks (lower branch). Metric-based meta-learning enables models to adapt swiftly to new fault patterns, crucial for few-shot scenarios. It performs well when the number of labeled examples is limited, making it suitable for fault diagnosis with small datasets.

Well-known metric-based meta-learning networks include prototypical networks [21–25], relation networks [26,27], and matching networks [14].

Prototypical networks

Prototypical networks learn prototypes for each class based on feature representations of limited samples. Classification is performed by assigning a query sample to the class closest to its prototype in the feature space, making it effective for few-shot classification tasks.

Many researchers have applied this method for few-shot tasks. Tang et al. [21] presented an enhanced prototypical network with L2 prototype correction for few-shot cross-domain fault diagnosis. The method utilized L2 correction to refine prototype representations, enabling effective diagnosis across different domains with limited labeled data. Feng et al. [22] employed prototypical networks in the proposed semi-supervised meta-learning networks. The method learned class prototypes by averaging feature representations of limited labeled samples. Squeeze-and-excitation attention was integrated to enhance prototype quality, while the model was trained in a semi-supervised meta-learning framework for

improved performance. In [23], a wavelet-prototypical network was employed for few-shot fault diagnosis. It fused time and frequency domain information using wavelet transforms.
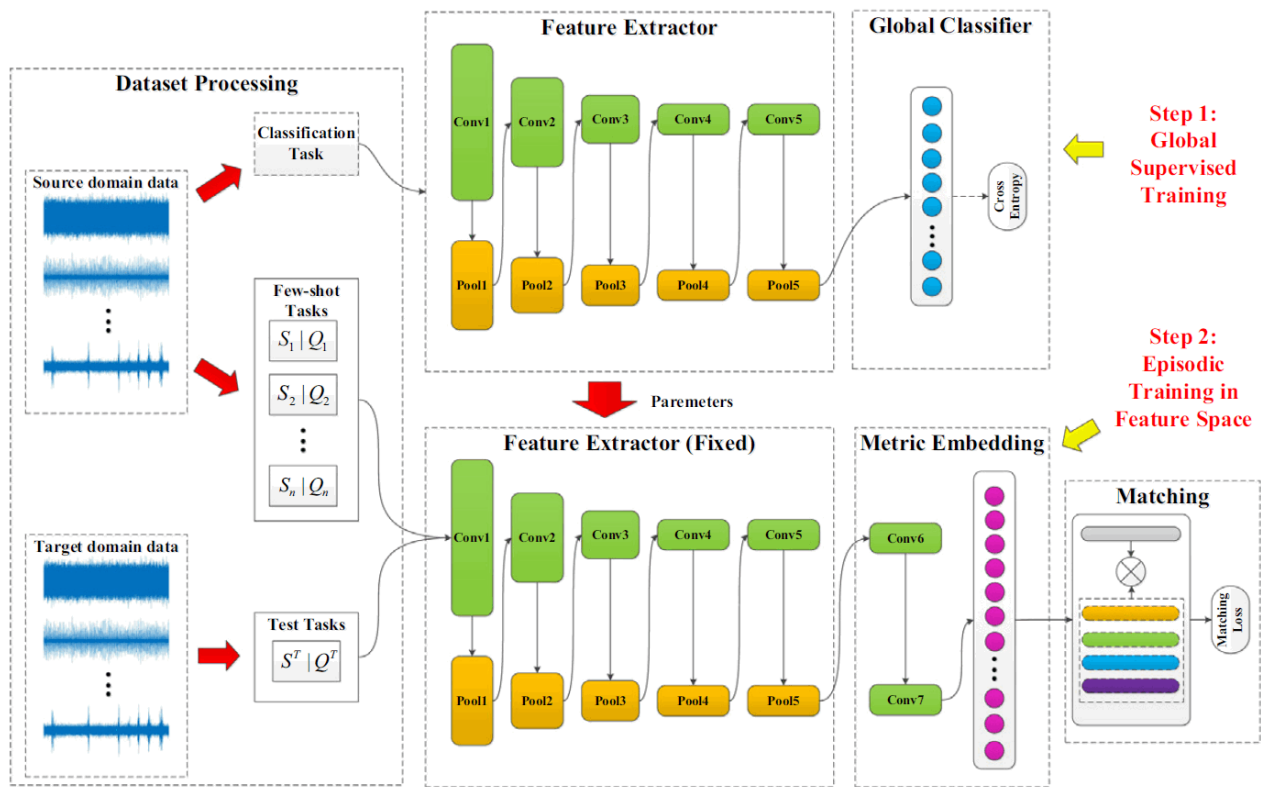

Figure 4. A typical metric-based meta-learning method for fault diagnosis [14].

Matching networks

Matching networks use attention mechanisms to establish correspondences between query and support samples. Attention weights are learned during training to classify query samples based on the labeled information aggregated from support samples, accommodating variable-length inputs in few-shot learning scenarios. Xu et al. [28] proposed a deep convolutional nearest-neighbor matching network (DC-NNMN) for cross-component few-shot fault diagnosis. The approach utilized matching networks with attention mechanisms to establish correspondences between query and support samples. DC-NNMN efficiently diagnosed faults across different components with limited labeled data, achieving accurate results. Wang et al. [14] proposed a feature space metric-based meta-learning model to distinguish failure attribution accurately under conditions of very limited data. A matching network and prototypical network were applied, respectively, in the proposed model to match the metric features to the support features using a public dataset. A similar approach was applied in [29] employing a matching network to match the metric features to the support features using experimental datasets.

Relation networks

Relation networks recognize relationships between samples using an embedding network to encode input samples into feature representations and a relation network to predict relationships between pairs of feature vectors, capturing complex sample relationships. Kang et al. [26] developed a few-shot rolling-bearing fault classification method using an improved relation network by introducing a residual shrinkage module and a scaled exponential linear unit activation function into the embedding module of the relation network. The approach enhanced the relation network's ability to capture complex relationships between samples in few-shot scenarios. Wang et al. [27] applied a relation network for the

few-shot multiscene fault diagnosis of rolling bearings under compound-variable working conditions. Further studies using relation networks can be found in [30,31].

Other researchers have applied metric-based meta-learning methods comprising a subspace network with shared representation learning [32] and a cross-level fusion neural network [33].

Meta-learning, in general, focuses on training models to learn from different tasks or datasets so that they can adapt quickly to new, unseen tasks or datasets. Metric-based meta-learning, in particular, emphasizes the use of a metric or distance function to enable this rapid adaptation. For classification purposes, metric-based meta-learning could be placed in either Section 3.1 (Meta-Learning) or Section 3.2 (Metric Learning). However, our examination of the relevant literature revealed that the papers addressing metric-based meta-learning methods mainly considered task adaptation. Metric-based meta-learning primarily focuses on the ability of a model to adapt quickly to new tasks or datasets. It achieves this by learning a metric or similarity measure during meta-training, which helps in task-specific adaptation. This aligns closely with the core objective of meta-learning, which is to enable models to generalize effectively to new tasks. Additionally, metric-based meta-learning is often applied in few-shot learning scenarios, where the model needs to make predictions with a very limited number of labeled examples in the support set. This is a typical use case within meta-learning, as one of its primary applications is to excel in few-shot or low-data scenarios. Therefore, we included metric-based meta-learning in the broader category of meta-learning.

3.1.2. Learning Initialization Methods

In meta-learning, learning initialization methods refer to a category of algorithms that leverage pretrained models or knowledge obtained from previous tasks to facilitate faster and more effective learning on new, unseen tasks. The core idea behind these methods is to utilize the knowledge acquired from related tasks as a starting point for learning new tasks, thus enabling the model to adapt and generalize quickly with limited data. As for learning initialization meta-learning, the prevailing networks are model-agnostic meta-learning (MAML) [34–39] and reptile [40] networks.

MAML

MAML [41], model-agnostic meta-learning, is a meta-learning algorithm that enables fast adaptation to new tasks with limited data. It optimizes a model's initial parameters such that it can efficiently adapt to a new task using just a few gradient updates. MAML learns a more general initialization that facilitates quick fine-tuning on unseen tasks. By using gradient-based optimization to adapt models across tasks, MAML promotes better generalization and transfer learning, making it a powerful and flexible approach for few-shot learning scenarios.

Several papers have applied this method. Liu et al. [35] used MAML to train a meta-baseline model capable of quickly diagnosing faults in wind turbines with limited labeled data. MAML enabled the model to efficiently adapt to new turbine faults with just a few gradient updates, improving few-shot fault diagnosis performance. Yang et al. [36] applied MAML for few-shot fault diagnosis tasks in high-speed train suspension systems. The approach improved fault diagnosis performance by efficiently leveraging knowledge across different suspension system faults. Yu et al. [37] applied MAML to a learnable and interpretable framework based on model-free DA methods, D3AFS, for industrial scenarios with limited data. The good performance of the proposed model was verified using a magnetic flux leakage dataset in a pipeline and bearing datasets. Further papers addressing the MAML concept can be found in [34,38,39].

Reptile

Reptile [42] is a meta-learning algorithm that facilitates quick adaptation to new tasks by fine-tuning a model's parameters. It operates by repeatedly sampling tasks, training the model on each task for a few gradient steps, and then updating the model's parameters

based on the changes accumulated across tasks. The algorithm encourages the model to learn a more general initialization that can be easily fine-tuned for different tasks. Unlike traditional meta-learning approaches, reptile does not maintain an explicit representation of meta-parameters, making it more straightforward and efficient. Reptile has shown promising results in few-shot learning scenarios, demonstrating improved generalization across tasks. Pei et al. [40] employed reptile to enhance the few-shot Wasserstein autoencoder (WAE). Reptile was utilized to further enhance the mapping ability of WAE from prior distribution to vibration signals when faced with a small dataset.

The abovementioned methods offer a flexible way to initialize models, enabling them to adapt quickly to new tasks. However, training can be computationally intensive, requiring substantial resources.

3.1.3. Other Methods

Other researchers have applied ensemble techniques in meta-learning. For example, Li et al. [43] proposed a light gradient-boosting-machine-based multiscale weighted ensemble model to perform effective few-shot fault diagnosis without requiring cross-domain data. Chen et al. [44] designed a meta-self-attention multiscale convolution neural network for the actuator fault diagnosis of autonomous underwater vehicles. Che et al. [45] employed ensemble meta-learning, enabling the model to adapt quickly to new fault conditions with few samples. By combining multiple models and leveraging meta-learning techniques, the proposed method achieved enhanced diagnostic performance in variable working conditions.

The future of meta-learning holds the potential for emphasizing fine-grained meta-learning and crafting more robust meta-learning models with enhanced transferability and generalization capabilities. In the realm of fine-grained meta-learning, forthcoming research could concentrate on devising techniques that enable adaptability to even more nuanced variations in fault patterns. This entails the development of models capable of swiftly generalizing to hitherto unencountered fault types or levels of severity. In the pursuit of crafting more robust meta-learning models possessing heightened transferability and generalization capabilities, enhancing the transferability of meta-learned models across different machinery types and operating conditions is crucial. Researchers could delve into methodologies aimed at ensuring that insights garnered from one machinery category can be effectively applied to others.

3.2. Metric Learning

Metric learning [46] is a machine learning task that focuses on learning a distance or similarity metric between datapoints in a given feature space. A typical example of metric learning can be found in Figure 5. The objective of metric learning is to ensure that the learned metric reflects the actual similarity or dissimilarity between data samples, making it possible to measure how similar or dissimilar they are to each other. Metric learning can be used in traditional machine learning, deep learning, or meta-learning. In the current paper, the application of metric learning in few-shot learning is called metric-based meta-learning (Section 3.1.1).

The process of metric learning typically involves a set of training samples, each with a corresponding label or similarity information. The learning algorithm (i.e., Siamese, prototypical, or relation networks) then uses these labeled data to optimize the metric in a way that brings similar samples closer together and dissimilar samples farther apart in the feature space.

Lu et al. [47] applied multiple-kernel maximum mean discrepancy, which is an improved version of maximum mean discrepancy, for distribution discrepancy measurement and developed a multiview and multilevel network model for fault diagnosis. In another paper by the lead author [48], the fault diagnosis problem was treated as a similarity metric learning problem, with a transfer relation network (TRN) proposed to achieve this objective. The TRN model incorporated a relation learning module that captured knowledge and

patterns shared between the source and target domains. Further papers addressing metric learning can be found in [49–52].
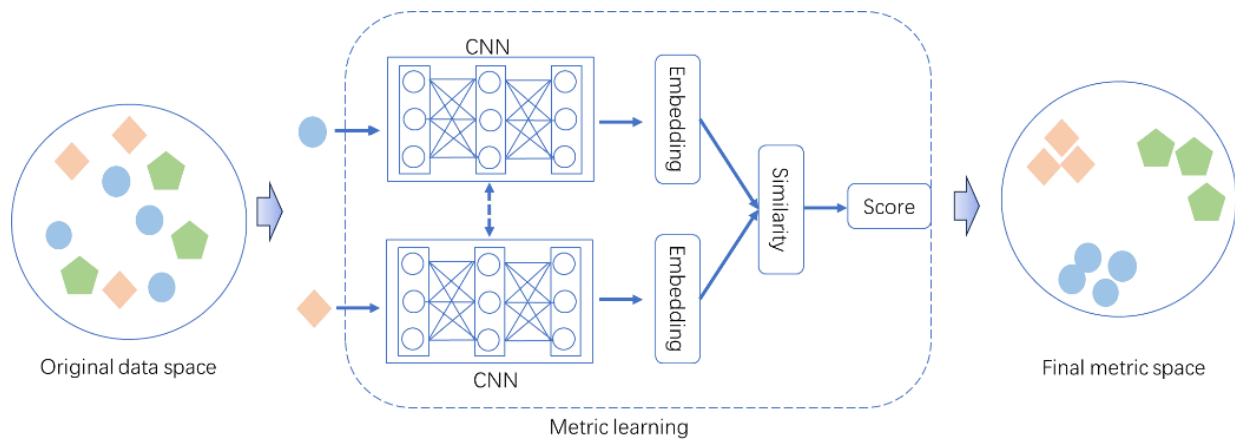

Figure 5. Typical example of metric learning.

Metric learning can lead to improved feature representations for fault diagnosis, emphasizing relevant patterns. Metric-based methods heavily depend on the quality and representativeness of the training data.

The future prospects for metric learning encompass two main directions: embedding for interpretability and the development of hybrid approaches. Embedding for interpretability involves the development of metric learning techniques that generate interpretable embeddings. This direction holds promise for gaining insights into the contributions of specific features to fault diagnosis, promoting transparency and actionable insights. The development of hybrid approaches entails the fusion of metric learning with other machine learning methods, such as deep learning and ensemble techniques, with the potential to enhance model performance. Hybrid models can effectively capture both global and local similarities in data, harnessing the strengths of diverse learning paradigms to improve fault diagnosis accuracy and robustness. These future developments in metric learning offer opportunities to enhance transparency and leverage complementary techniques for more effective fault detection and diagnosis in mechanical systems.

3.3. Data Augmentation Method

The core issue in few-shot fault diagnosis lies in the limited number of samples, which hinders the training of a reliable, highly generalizable diagnostic model. In the case of limited data, a generative model can be used to enhance sample diversity. According to the reviewed publications on the application of few-shot learning in the machine fault diagnosis field using vibration data from 2018 to 2023, the data augmentation methods are mainly GAN (generative adversarial network)-based methods [53,54].

3.3.1. GAN-Based Methods

A GAN is a type of unsupervised learning model that consists of two neural networks: a generator network and a discriminator network. The main idea behind GANs is to train the generator network to produce realistic data samples by learning from a training dataset. The generator takes random noise as the input and generates synthetic samples. The discriminator network, on the other hand, acts as a binary classifier that distinguishes between real data samples from the training set and the synthetic samples generated by the generator. During the training process, the generator and discriminator networks play a game against each other. The generator tries to produce realistic samples that can fool the discriminator, while the discriminator aims to correctly classify real and synthetic samples. The two networks are trained simultaneously, and their performance improves iteratively through an adversarial process.

As training progresses, the generator becomes better at generating realistic samples, and the discriminator becomes more skilled at distinguishing between real and synthetic data. Eventually, the generator network learns to generate samples that are difficult for the discriminator to differentiate from real data. It is worth mentioning that GANs face several challenges, including mode collapse (when the generator produces a limited set of similar samples) and instability during training [10]. However, they continue to be an active area of research, and many advancements and variations have been developed to address these issues and push the boundaries of generative modeling.

In the field of machine fault diagnosis, researchers usually use one-dimensional (1D) or 2D samples. One-dimensional samples include time-domain data, frequency-domain data, and some extracted data features. Li et al. [53] explored two few-shot learning techniques involving parameter fine-tuning and a conditional Wasserstein generative adversarial network (C-WGAN) for diagnosing faults in freight train rolling bearings. They applied data segmentation and transformed the signals into the frequency domain, resulting in 1D frequency signals. To automatically extract features from the bearing vibration signals and classify fault types, they developed a one-dimensional convolutional neural network (1D-CNN). Gao et al. [55] developed a data augmentation model based on an integrated convolutional transformer GAN to improve diagnostic performance under limited-data conditions by generating high-quality signals. Wan et al. [56] proposed an unsupervised fault diagnosis method based on a quick self-attention convolutional generative adversarial network. Chen et al. [54] utilized a Wasserstein deep convolutional generative adversarial network (WDCGAN) to improve the performance of few-shot fault diagnosis in electrohydrostatic actuators. They achieved this by transforming multidimensional experimental data into 2D grayscale data and extracting local features, effectively emphasizing the time-series characteristics and correlations among different signals.

3.3.2. Other Data Augmentation Models

Xia et al. [57] proposed an augmentation-based discriminative meta-learning method to address the issue of the few-shot cross-machine domain shift problem. During the meta-training process, signal transformation was introduced to enhance meta-task diversity for robust feature learning, and multi-scale learning was integrated for adaptive feature embedding; meanwhile, in the meta-testing phase, sparse labeled fault data boosted model generalization using quasi-meta-training via data augmentation, and a new hyperbolic prototypical loss was designed to ensure more distinct feature representation with a hyperbolic decision boundary for separable category prototypes.

Zhao et al. [58] introduced a data augmentation technique called randomized wavelet expansion to create a collection of synthetic samples that possessed comparable characteristics to the original samples. Subsequently, the synthesized samples were utilized as the training dataset for a deep CNN to achieve the few-shot fault diagnosis of aviation hydraulic pumps.

Wang et al. [59] proposed an extended convolutional adversarial autoencoder (ECAAE) as an end-to-end fault diagnosis approach for electromechanical actuators (EMAs) based merely on vibration signals. The ECAAE combined a set of CNNs for feature extraction with the data generation capabilities of adversarial autoencoders, allowing the model to utilize both unlabeled and unbalanced labeled samples. Through an adversarial training process and hyperparameter-free signal conversion method, the ECAAE achieved robust and precise fault diagnosis for EMAs even with varying working conditions, unbalanced samples, and few-shot situations.

Hu et al. [60] introduced a self-supervised learning framework that combined both inter-instance learning and intra-temporal learning. This innovative approach aimed to improve the model's ability to generalize from limited labeled data by leveraging the inherent structure and temporal information in the data.

In the field of few-shot fault diagnosis, the quantity and quality of data play a crucial role in the performance of models. However, in practical applications, acquiring a substan-

tial amount of high-quality annotated data can be prohibitively expensive and challenging. In such scenarios, generative models can serve as an effective solution. By leveraging generative models, it becomes possible to generate a large number of new samples based on the distribution of existing data. Although these samples are synthesized, they retain certain characteristics and the distribution of the original data to an extent. It is important to note that while generative models can enhance the diversity of samples, the generated samples may not fully represent the true distribution of real data. Therefore, caution is required when utilizing generated samples. Additionally, the quality and efficacy of generative models are influenced by the design and training process of the model. Consequently, when employing generative models for data augmentation, the careful adjustment of the model parameters is necessary to ensure that the generated samples are statistically similar to the real data, thereby guaranteeing the effectiveness of data augmentation. Data augmentation enhances the diversity of training samples, potentially reducing overfitting. It helps models generalize better to unseen fault patterns. It is worth noting that the effectiveness of data augmentation is limited by the quality and quantity of original data.

The future outlook for data augmentation methods encompasses two main aspects: cross-modal data augmentation and self-supervised learning. Cross-modal data augmentation involves the exploration of techniques facilitating the integration of data from diverse modalities, such as vibration signals, acoustic signals, and temperature data. This integration holds the potential to yield more comprehensive and informative representations for fault diagnosis, allowing for a richer understanding of system behavior. In the context of self-supervised learning, the amalgamation of self-supervised learning methods with data augmentation stands as a promising approach. This synergy could significantly reduce the dependence on labeled data, enabling models to acquire valuable representations from unlabeled or weakly labeled data, thereby advancing the efficiency and versatility of fault diagnosis processes. These future directions for data augmentation methods offer the prospect of enhancing the depth and breadth of insights available for fault detection and diagnosis in mechanical systems.

3.4. Other Few-Shot Learning Methods

Taking into account the impact of noise on vibration signals, Ma et al. [61] introduced a multi-order graph embedding model. They presented an enhanced sine cosine algorithm strategy to optimize the feature extraction capability of a stacked denoising autoencoder. This method was then employed to address the few-shot diagnosis issue. Chen et al. [62] introduce a multi-channel calibrated transformer (MCSwin-T) model with shifted windows to tackle few-shot fault diagnosis in scenarios with sharp speed variation. The model employed multi-channel calibration to align feature representations across speeds, enabling effective fault pattern recognition with limited labeled data. Ma et al. [31] developed a feature extractor named MiniNet that struck an optimal balance between channel count and network depth during fault feature extraction. Leveraging MiniNet, the authors introduced a fault diagnosis model for few-shot samples that effectively enhanced both transferability and discriminability. Wang et al. [63] proposed a method based on two-dimensional (2D) images and cross-domain few-shot learning for bearing fault diagnosis. In the paper, the authors did not mention meta-learning or metric-based methods, although a distance-based classifier was employed to improve the classification capacity of the few samples. Therefore, we included this paper in the category of other few-shot learning methods.

4. Discussion

The analysis of the literature revealed that meta-learning holds a central position in the realm of few-shot learning approaches, constituting more than 50% of the studied articles. Among the various meta-learning techniques, MAML (model-agnostic meta-learning) and its modifications, alongside metric-based meta-learning methods, collectively contributed to approximately 90% of the research endeavors in this field. In addition, through the

review, we found that transformers are increasingly popular for applications in few-shot learning; related applications can be found in [55,62,64].

Looking ahead, few-shot learning holds promising prospects in mechanical fault diagnosis.

Robustness and Generalization: Improving the robustness and generalization capabilities of few-shot learning models is an ongoing pursuit. Research will continue to address issues related to noisy data, domain shifts, and class imbalances to make few-shot learning more reliable in practical settings.

Unsupervised Few-shot Learning: Exploring unsupervised few-shot learning, where models can learn from unlabeled data in a few-shot scenario, is an exciting and challenging direction. This area has the potential to further reduce the reliance on labeled data.

Combining Few-shot Learning with Other Techniques: Integrating few-shot learning with other machine learning paradigms, such as transfer learning, semi-supervised learning, and reinforcement learning, could lead to more powerful and versatile models.

Few-shot Learning in Real-world Applications: Few-shot learning is valuable in real-world applications due to its data efficiency, rapid adaptation to new tasks, flexibility across domains, reduced annotation effort, and effective generalization to unseen data. These advantages make it a promising solution for addressing data scarcity and dynamic environments while minimizing manual labeling and achieving better performance.

5. Conclusions

Continuing the work of previous researchers, this paper focused on a comprehensive analysis of few-shot learning methods applied in the domain of mechanical fault diagnosis using vibration signals. The study particularly concentrated on advancements made within the past five years, spanning from 2018 to September 2023. We found that meta-learning is a prominent area of interest in few-shot learning methods, with over 50% of the articles focused on meta-learning. Within the domain of meta-learning, MAML (model-agnostic meta-learning) and its variations, along with metric-based meta-learning methods, collectively accounted for approximately 90% of the research contributions. In summary, the use of few-shot learning in real-world applications is motivated by its ability to handle data scarcity, adapt rapidly to new tasks, demonstrate flexibility across domains, reduce annotation effort, facilitate transfer learning, and effectively generalize to unseen data. These advantages make few-shot learning a promising approach for addressing various challenges in practical applications and pushing the boundaries of machine learning in real-world settings.

Author Contributions: Conceptualization, D.W., M.Z. and X.L.; methodology, D.W., M.Z. and X.L.; writing—original draft preparation, X.L. and G.F; writing—review and editing, X.L., M.Z., G.F., D.W., Y.X. and F.G.; supervision, Y.X. and F.G.; project administration, X.L., M.Z. and G.F.; funding acquisition, X.L., M.Z., G.F., Y.X. and F.G. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Yin, Z.; Hou, J. Recent advances on SVM based fault diagnosis and process monitoring in complicated industrial processes. *Neurocomputing* **2016**, *174*, 643–650. [CrossRef]
2. Li, Z.; Zhang, Y.; Abu-Siada, A.; Chen, X.; Li, Z.; Xu, Y.; Zhang, L.; Tong, Y. Fault diagnosis of transformer windings based on decision tree and fully connected neural network. *Energies* **2021**, *14*, 1531. [CrossRef]
3. Hu, Q.; Si, X.-S.; Zhang, Q.-H.; Qin, A.-S. A rotating machinery fault diagnosis method based on multi-scale dimensionless indicators and random forests. *Mech. Syst. Signal Process.* **2020**, *139*, 106609. [CrossRef]
4. Jiao, J.; Zhao, M.; Lin, J.; Liang, K. A comprehensive review on convolutional neural network in machine fault diagnosis. *Neurocomputing* **2020**, *417*, 36–63. [CrossRef]
5. Zhang, Y.; Zhou, T.; Huang, X.; Cao, L.; Zhou, Q. Fault diagnosis of rotating machinery based on recurrent neural networks. *Measurement* **2021**, *171*, 108774. [CrossRef]
6. Zhao, H.; Sun, S.; Jin, B. Sequential fault diagnosis based on LSTM neural network. *IEEE Access* **2018**, *6*, 12929–12939. [CrossRef]
7. Jiang, G.; He, H.; Xie, P.; Tang, Y. Stacked multilevel-denoising autoencoders: A new representation learning approach for wind turbine gearbox fault diagnosis. *IEEE Trans. Instrum. Meas.* **2017**, *66*, 2391–2402. [CrossRef]
8. Lu, J.; Gong, P.; Ye, J.; Zhang, C. Learning from very few samples: A survey. *arXiv* **2020**, arXiv:2009.02653.
9. Wang, Y.; Yao, Q.; Kwok, J.; Ni, L.M. Generalizing from a Few Examples: A Survey on Few-Shot Learning. *arXiv* **2020**, arXiv:1904.05046. Available online: http://arxiv.org/abs/1904.05046 (accessed on 6 February 2023). [CrossRef]
10. Pan, T.; Chen, J.; Zhang, T.; Liu, S.; He, S.; Lv, H. Generative adversarial network in mechanical fault diagnosis under small sample: A systematic review on applications and future perspectives. *ISA Trans.* **2022**, *128*, 1–10. [CrossRef]
11. Zhang, T.; Chen, J.; Li, F.; Zhang, K.; Lv, H.; He, S.; Xu, E. Intelligent fault diagnosis of machines with small & imbalanced data: A state-of-the-art review and possible extensions. *ISA Trans.* **2022**, *119*, 152–171. [PubMed]
12. Bhuiyan, M.R.; Uddin, J. Deep transfer learning models for industrial fault diagnosis using vibration and acoustic sensors data: A review. *Vibration* **2023**, *6*, 218–238. [CrossRef]
13. Ravi, S.; Larochelle, H. Optimization as a model for few-shot learning. In Proceedings of the 5th International Conference on learning Representations, Toulon, France, 24–26 April 2016.
14. Wang, D.; Zhang, M.; Xu, Y.; Lu, W.; Yang, J.; Zhang, T. Metric-based meta-learning model for few-shot fault diagnosis under multiple limited data conditions. *Mech. Syst. Signal Process.* **2021**, *155*, 107510. [CrossRef]
15. Wang, P.; Li, J.; Wang, S.; Zhang, F.; Shi, J.; Shen, C. A new meta-transfer learning method with freezing operation for few-shot bearing fault diagnosis. *Meas. Sci. Technol.* **2023**, *34*, 074005. [CrossRef]
16. Ke, W.; Yukang, N.; Wu, J.; Yuanhang, W. Prior Knowledge-based Self-supervised Learning for Intelligent Bearing Fault Diagnosis with Few Fault Samples. *Meas. Sci. Technol.* **2023**, *34*, 105104.
17. Zhang, Y.; Han, D.; Tian, J.; Shi, P. Domain adaptation meta-learning network with discard-supplement module for few-shot cross-domain rotating machinery fault diagnosis. *Knowl.-Based Syst.* **2023**, *268*, 110484. [CrossRef]
18. Naik, D.K.; Mammone, R.J. Meta-neural networks that learn by learning. In Proceedings of the IJCNN International Joint Conference on Neural Networks, Baltimore, MD, USA, 7–11 June 1992; Volume 1, pp. 437–442.
19. Baker, B.; Gupta, O.; Naik, N.; Raskar, R. Designing neural network architectures using reinforcement learning. *arXiv* **2016**, arXiv:1611.02167.
20. Lu, Z.; Whalen, I.; Boddeti, V.; Dhebar, Y.; Deb, K.; Goodman, E.; Banzhaf, W. Nsga-net: Neural architecture search using multi-objective genetic algorithm. In Proceedings of the Genetic and Evolutionary Computation Conference, Prague, Czech Republic, 13–17 July 2019; pp. 419–427.
21. Tang, T.; Wang, J.; Yang, T.; Qiu, C.; Zhao, J.; Chen, M.; Wang, L. An improved prototypical network with L2 prototype correction for few-shot cross-domain fault diagnosis. *Measurement* **2023**, *217*, 113065. [CrossRef]
22. Feng, Y.; Chen, J.; Zhang, T.; He, S.; Xu, E.; Zhou, Z. Semi-supervised meta-learning networks with squeeze-and-excitation attention for few-shot fault diagnosis. *ISA Trans.* **2022**, *120*, 383–401. [CrossRef]
23. Wang, Y.; Chen, L.; Liu, Y.; Gao, L. Wavelet-prototypical network based on fusion of time and frequency domain for fault diagnosis. *Sensors* **2021**, *21*, 1483. [CrossRef]
24. Tnani, M.-A.; Subarnaduti, P.; Diepold, K. Efficient feature learning approach for raw industrial vibration data using two-stage learning framework. *Sensors* **2022**, *22*, 4813. [CrossRef] [PubMed]
25. Han, Y.; Li, C.; Huang, Q.; Wen, R.; Zhang, Y. Boundary-enhanced prototype network with time-series attention for gearbox fault diagnosis under limited samples. *J. Electron. Meas. Instrum.* **2023**, *37*, 90–98.
26. Kang, S.; Liang, X.; Wang, Y.; Wang, Q.; Qiao, C.; Mikulovich, V.I. Few-shot rolling bearing fault classification method based on improved relation network. *Meas. Sci. Technol.* **2022**, *33*, 125020. [CrossRef]
27. Wang, S.; Wang, D.; Kong, D.; Li, W.; Wang, J.; Wang, H. Few-shot multiscene fault diagnosis of rolling bearing under compound variable working conditions. *IET Control. Theory Appl.* **2022**, *16*, 1405–1416. [CrossRef]
28. Xu, J.; Xu, P.; Wei, Z.; Ding, X.; Shi, L. DC-NNMN: Across components fault diagnosis based on deep few-shot learning. *Shock. Vib.* **2020**, *2020*, 1–11. [CrossRef]
29. Liang, X.; Zhang, M.; Feng, G.; Xu, Y.; Zhen, D.; Gu, F. A Novel Deep Model with Meta-learning for Rolling Bearing Few-shot Fault Diagnosis. *J. Dyn. Monit. Diagn.* **2023**, *2*, 102–114. [CrossRef]

30. Wei, P.; Liu, M.; Wang, X. Few-shot bearing fault diagnosis using GAVMD–PWVD time–frequency image based on meta-transfer learning. *J Braz. Soc. Mech. Sci. Eng.* **2023**, *45*, 277. [CrossRef]

31. Ma, W.; Zhang, Y.; Ma, L.; Liu, R.; Yan, S. An unsupervised domain adaptation approach with enhanced transferability and discriminability for bearing fault diagnosis under few-shot samples. *Expert Syst. Appl.* **2023**, *225*, 120084. [CrossRef]

32. Liu, S.; Chen, J.; He, S.; Shi, Z.; Zhou, Z. Subspace Network with Shared Representation learning for intelligent fault diagnosis of machine under speed transient conditions with few samples. *ISA Trans.* **2023**, *128*, 531–544. [CrossRef]

33. Wang, S.; Wang, D.; Kong, D.; Li, W.; Wang, H.; Pecht, M. Cross-Level fusion for rotating machinery fault diagnosis under compound variable working conditions. *Measurement* **2022**, *199*, 111455. [CrossRef]

34. Li, C.; Li, S.; Wang, H.; Gu, F.; Ball, A.D. Attention-based deep meta-transfer learning for few-shot fine-grained fault diagnosis. *Knowl.-Based Syst.* **2023**, *264*, 110345. [CrossRef]

35. Liu, X.; Teng, W.; Liu, Y. A model-agnostic meta-baseline method for few-shot fault diagnosis of wind turbines. *Sensors* **2022**, *22*, 3288. [CrossRef] [PubMed]

36. Yang, F.; Lv, L.; Hua, C.; Xiong, L.; Dong, D. Fault diagnosis of suspension system of high-speed train based on model-agnostic meta-learning. In Proceedings of the 2022 Global Reliability and Prognostics and Health Management (PHM-Yantai), Yantai, China, 13–16 October 2022; pp. 1–6.

37. Yu, G.; Wang, D.; Liu, J.; Zhang, X. Distribution-Agnostic Few-Shot Industrial Fault Diagnosis via Adaptation-Aware Optimal Feature Transport. *IEEE Trans. Ind. Inform.* **2022**, *19*, 5623–5632. [CrossRef]

38. Chen, J.; Hu, W.; Cao, D.; Zhang, Z.; Chen, Z.; Blaabjerg, F. A meta-learning method for electric machine bearing fault diagnosis under varying working conditions with limited data. *IEEE Trans. Ind. Inform.* **2022**, *19*, 2552–2564. [CrossRef]

39. Yu, C.; Ning, Y.; Qin, Y.; Su, W.; Zhao, X. Multi-label fault diagnosis of rolling bearing based on meta-learning. *Neural Comput. Applic.* **2021**, *33*, 5393–5407. [CrossRef]

40. Pei, Z.; Jiang, H.; Li, X.; Zhang, J.; Liu, S. Data augmentation for rolling bearing fault diagnosis using an enhanced few-shot Wasserstein auto-encoder with meta-learning. *Meas. Sci. Technol.* **2021**, *32*, 084007. [CrossRef]

41. Finn, C.; Abbeel, P.; Levine, S. Model-agnostic meta-learning for fast adaptation of deep networks. *Int. Conf. Mach. Learn. PMLR* **2017**, *70*, 1126–1135.

42. Nichol, A.; Achiam, J.; Schulman, J. On first-order meta-learning algorithms. *arXiv* **2018**, arXiv:1803.02999.

43. Li, W.; He, J.; Lin, H.; Huang, R.; He, G.; Chen, Z. A LightGBM-based Multi-scale Weighted Ensemble Model for Few-shot Fault Diagnosis. *IEEE Trans. Instrum. Meas.* **2023**, *72*, 3523014. [CrossRef]

44. Chen, Y.; Wang, Y.; Yu, Y.; Wang, J.; Gao, J. A Fault Diagnosis Method for the Autonomous Underwater Vehicle via Meta-Self-Attention Multi-Scale CNN. *J. Mar. Sci. Eng.* **2023**, *11*, 1121. [CrossRef]

45. Che, C.; Wang, H.; Xiong, M.; Ni, X. Few-shot fault diagnosis of rolling bearing under variable working conditions based on ensemble meta-learning. *Digit. Signal Process.* **2022**, *131*, 103777. [CrossRef]

46. Kaya, M.; Bilge, H.Ş. Deep metric learning: A survey. *Symmetry* **2019**, *11*, 1066. [CrossRef]

47. Lu, N.; Cui, Z.; Hu, H.; Yin, T. Multi-view and Multi-level network for fault diagnosis accommodating feature transferability. *Expert Syst. Appl.* **2023**, *213*, 119057. [CrossRef]

48. Lu, N.; Hu, H.; Yin, T.; Lei, Y.; Wang, S. Transfer relation network for fault diagnosis of rotating machinery with small data. *IEEE Trans. Cybern.* **2021**, *52*, 11927–11941. [CrossRef] [PubMed]

49. Jiang, C.; Chen, H.; Xu, Q.; Wang, X. Few-shot fault diagnosis of rotating machinery with two-branch prototypical networks. *J. Intell. Manuf.* **2023**, *34*, 1667–1681. [CrossRef]

50. Shen, H.; Zhao, D.; Wang, L.; Liu, Q. Bearing fault diagnosis based on prototypical network. In Proceedings of the International Conference on Mechatronics Engineering and Artificial Intelligence (MEAI 2022), SPIE, Changsha, China, 28 February 2023; pp. 79–84.

51. Fang, Q.; Wu, D. ANS-net: Anti-noise Siamese network for bearing fault diagnosis with a few data. *Nonlinear Dyn.* **2021**, *104*, 2497–2514. [CrossRef]

52. Hu, Y.; Xiong, Q.; Zhu, Q.; Yang, Z.; Zhang, Z.; Wu, D.; Wu, Z. Few-shot transfer learning with attention for intelligent fault diagnosis of bearing. *J. Mech. Sci. Technol.* **2022**, *36*, 6181–6192. [CrossRef]

53. Li, C.; Yang, K.; Tang, H.; Wang, P.; Li, J.; He, Q. Fault diagnosis for rolling bearings of a freight train under limited fault data: Few-shot learning method. *J. Transp. Eng. Part A Syst.* **2021**, *147*, 04021041. [CrossRef]

54. Chen, H.; Miao, X.; Mao, W.; Zhao, S.; Yang, G.; Bo, Y. Fault diagnosis of EHA with few-shot data augmentation technique. *Smart Mater. Struct.* **2023**, *32*, 044005. [CrossRef]

55. Gao, H.; Zhang, X.; Gao, X.; Li, F.; Han, H. ICoT-GAN: Integrated Convolutional Transformer GAN for Rolling Bearings Fault Diagnosis under Limited Data Condition. *IEEE Trans. Instrum. Meas.* **2023**, *72*, 3515114. [CrossRef]

56. Wan, W.; He, S.; Chen, J.; Li, A.; Feng, Y. QSCGAN: An un-supervised quick self-attention convolutional GAN for LRE bearing fault diagnosis under limited label-lacked data. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 1–16. [CrossRef]

57. Xia, P.; Huang, Y.; Wang, Y.; Liu, C.; Liu, J. Augmentation-based discriminative meta-learning for cross-machine few-shot fault diagnosis. *Sci. China Technol. Sci.* **2023**, *66*, 1698–1716. [CrossRef]

58. Zhao, M.; Fu, X.; Zhang, Y.; Meng, L.; Zhong, S. Data augmentation via randomized wavelet expansion and its application in few-shot fault diagnosis of aviation hydraulic pumps. *IEEE Trans. Instrum. Meas.* **2021**, *71*, 1–13. [CrossRef]

59. Wang, C.; Tao, L.; Ding, Y.; Lu, C.; Ma, J. An adversarial model for electromechanical actuator fault diagnosis under nonideal data conditions. *Neural Comput. Appl.* **2022**, *34*, 5883–5904. [CrossRef]
60. Hu, C.; Wu, J.; Sun, C.; Yan, R.; Chen, X. Inter-Instance and Intra-Temporal Self-Supervised Learning with Few Labeled Data for Fault Diagnosis. *IEEE Trans. Ind. Inform.* **2022**, *19*, 6502–6512. [CrossRef]
61. Ma, W.; Liu, R.; Guo, J.; Wang, Z.; Ma, L. A collaborative central domain adaptation approach with multi-order graph embedding for bearing fault diagnosis under few-shot samples. *Appl. Soft Comput.* **2023**, *140*, 110243. [CrossRef]
62. Chen, Z.; Chen, J.; Liu, S.; Feng, Y.; He, S.; Xu, E. Multi-channel Calibrated Transformer with Shifted Windows for few-shot fault diagnosis under sharp speed variation. *ISA Trans.* **2022**, *131*, 501–515. [CrossRef]
63. Wang, T.; Chen, C.; Dong, X.; Liu, H. A Novel Method of Production Line Bearing Fault Diagnosis Based on 2D Image and Cross-Domain Few-Shot Learning. *Appl. Sci.* **2023**, *13*, 1809. [CrossRef]
64. Chen, C.; Wang, T.; Liu, C.; Liu, Y.; Cheng, L. Lightweight Convolutional Transformers Enhanced Meta Learning for Compound Fault Diagnosis of Industrial Robot. *IEEE Trans. Instrum. Meas.* **2023**, *72*, 3520612. [CrossRef]