Few-Shot Bearing Fault Diagnosis Based on Model-Agnostic Meta-Learning

Zhang, Shen; Ye, Fei; Wang, Bingnan; Habetler, Thomas G

TR2021-081    July 12, 2021

Abstract
The rapid development of artificial intelligence and deep learning has provided many opportunities to further enhance the safety, stability, and accuracy of industrial CyberPhysical Systems (CPS). As indispensable components to many mission-critical CPS assets and equipment, mechanical bearings need to be monitored to identify any trace of abnormal conditions. Most of the data-driven approaches applied to bearing fault diagnosis up-to-date are trained using a large amount of fault data collected a priori. In many practical applications, however, it can be unsafe and time-consuming to collect sufficient data samples for each fault category, making it challenging to train a robust classifier. In this paper, we propose a few-shot learning framework for bearing fault diagnosis based on model-agnostic meta-learning (MAML), which targets for training an effective fault classifier using limited data. In addition, it can leverage the training data and learn to identify new fault scenarios more efficiently. Case studies on the generalization to new artificial faults show that the proposed framework achieves an overall accuracy up to 25% higher than a Siamese-network-based benchmark study. Finally, the robustness and the generalization capability of the proposed framework is further validated by applying it to identify real bearing damages using data from artificial damages, which compares favorably against 6 state-of-the-art few-shot learning algorithms using consistent test environments.

IEEE Transactions on Industry Applications
Abstract—The rapid development of artificial intelligence and deep learning has provided many opportunities to further enhance the safety, stability, and accuracy of industrial Cyber-Physical Systems (CPS). As indispensable components to many mission-critical CPS assets and equipment, mechanical bearings need to be monitored to identify any trace of abnormal conditions. Most of the data-driven approaches applied to bearing fault diagnosis up-to-date are trained using a large amount of fault data collected a priori. In many practical applications, however, it can be unsafe and time-consuming to collect sufficient data samples for each fault category, making it challenging to train a robust classifier. In this paper, we propose a few-shot learning framework for bearing fault diagnosis based on model-agnostic meta-learning (MAML), which targets for training an effective fault classifier using limited data. In addition, it can leverage the training data and learn to identify new fault scenarios more efficiently. Case studies on the generalization to new artificial faults show that the proposed framework achieves an overall accuracy up to 25% higher than a Siamese-network-based benchmark study. Finally, the robustness and the generalization capability of the proposed framework is further validated by applying it to identify real bearing damages using data from artificial damages, which compares favorably against 6 state-of-the-art few-shot learning algorithms using consistent test environments.

Index Terms—Bearings, fault diagnosis, few-shot, limited data, meta-learning.

I. INTRODUCTION

Cyber-Physical Systems (CPS) are a mixture of computation, networking, and physical processes, in which the embedded computational algorithms and networks have the power to monitor and control the physical components [1], [2]. The rapid development of artificial intelligence (AI) and the Internet of Things provides further opportunities to advance the technology frontier for sensing, monitoring, and interpreting CPS deployed in industry [3], [4]. As the essential component of many physical systems with rotating equipment, mechanical bearings are widely used in a variety of safety-critical applications such as airplanes, vehicles, production machinery, wind turbines, air-conditioning systems, and elevator hoists [5]. Therefore, the health conditions of bearings have major impacts on the reliability and performance of many industrial CPS [4].

While many data-driven and AI-based technologies have been applied to enhance the accuracy and reliability of bearing fault diagnosis [6]–[8], most of them require a large amount of training data such as vibration [9]–[11], acoustic [12], [13], and motor current [14], [15] signals. In real-world applications, however, it is often impossible to obtain sufficient data samples to train a robust classifier that identifies every fault type [10]. One of the reasons is that most bearing degradation would evolve slowly over time, a process that generally takes months or even years. Furthermore, certain safety-critical applications may not be allowed to run into faulty states [10]. Therefore, it can be expensive, unsafe, and often impractical to collect a sufficient amount of data at each bearing fault condition, which will inevitably result in data imbalance issues [16], [17]. All of these limitations on real-world bearing fault detection require the use of more effective algorithms that can leverage the limited data to train bearing fault classifiers with good generalization capabilities.

To achieve this goal, one approach is to leverage the limited data available at each class to perform data augmentation, using classical methods like signal translation and time stretching [18], or the more recent generative adversarial network (GAN) [17], [19]. The quality and temporal dependency of the generated time-series data is worth further validation [20]. It has been reported in [17] that “the quality of generated spectrum samples” from a GAN “isn’t good enough to provide auxiliary information”, which turns out to actually reduce the classifier accuracy after including these generated data into the training process. Therefore, the effectiveness of this method highly depends on the quality and amount of real data to generate high-quality “fake” data. Additionally, the well-known issues related to GAN training, such as instability, mode collapse, and weak gradient [21] can further impact their performance to monitor the health status of bearings in industrial CPS-based applications. Inspired by the GAN structure, [22] applies adversarial learning to extract generalized features and proposes an instance level weighted mechanism to reflect the similarities of testing samples with known health states. The unknown fault mode can be effectively identified through domain adaptation.

Besides data augmentation techniques, another promising approach to mitigate the limited data issue is to apply few-shot or one-shot learning methods [20], which has been successfully applied to a wide variety of tasks including few-shot image recognition, path planning of autonomous agents, and more recently in fault diagnosis [10], [11], [23]. Few-shot learning evaluates the model’s generalization capability
to classes not previously seen in the training process, given only a few samples of each new class [23]. Therefore, few-shot learning methods are well-suited to tackle the data imbalance issue as we can train a model that generalizes to the imbalanced classes.

Among the existing work that applies few-shot learning methods to bearing fault diagnosis [10], [11], [23], [24], [10] proposed a Siamese neural network-based model that demonstrated enhanced fault diagnosis performance when only using 9 training samples per class. Additionally, [23] applies a deep prototypical network-based method for few-shot bearing fault diagnosis. Despite demonstrating better performance when compared with supervised learning methods, all of the bearing fault types in the test set have shown up in the training set, as they are only differed by their defect diameters and operating conditions. In [11], an auto-encoder and a capsule network (CaAE) are implemented for the same purpose. However, the case studies in [11] are not formulated in the standard context of few-shot learning, as all of its identified bearing fault classes are already seen in the training process.

Another promising application is to leverage the generalization capability of few-shot learning to identify real bearing failures using data collected on artificially-damaged bearings, which can be a more convenient and cost-effective way when compared with collecting sufficient data on naturally-evolved failures. Despite its significance of the promising outcome, this Artificial-to-Real problem has only been investigated in a very recent few-shot learning work [24] using few-shot transfer learning algorithms such as feature transfer, frozen parameters, and relation nets. While many efforts are made to address this problem using transfer learning to perform domain-adaptation methods including parameter fine-tuning [25], deep inception net with atrous convolution [26], capsule networks [27], adversarial learning [22], metric-based meta learning [28], [29], and convolutional neural networks aided by conditional data alignment and unsupervised prediction consistency techniques [30], there is still room to further improve the model’s generalization capability when transferring to realistic failures. For example, it has been reported in [27] that when compared to Artificial-to-Artificial tasks that aim at generalizing from the trained artificial faults to new artificial ones, there is an obvious decrease in the average accuracy for Artificial-to-Real tasks, which may have been caused “by the differences between the artificial damage and the natural damage” [27].

Therefore, to further mitigate the limited data issue and improve the model’s generalization capability, this paper seeks to achieve effective fault diagnosis using the minimum amount of data with model-agnostic meta-learning (MAML) [1]. Beyond just generalizing to new tasks more effectively, MAML can also learn the process of learning itself, or learning to learn. Specifically, MAML is explicitly designed to train the model’s initial parameters such that “the model has maximal performance on a new task after the parameters have been updated through one or more gradient steps computed with a small amount of data from that new task” [31].

We summarize the main contributions of this paper as follows:

1) We propose a few-shot learning framework for bearing fault diagnosis with limited data, which is achieved by developing a diagnostic framework based on MAML.
2) We demonstrate that the MAML-based few-shot diagnostic framework consistently outperforms state-of-the-art results using the Siamese Network [10] when generalizing to new artificial faults. For example, with 3 unseen classes and using a small training set with only 9 samples per class, our method can achieve an average accuracy of 90.36% in comparison with 64.81% reported in [10].
3) We generalize the MAML-based diagnostic framework to identify real bearing damages using a model trained with artificial damages. Formulated as Artificial-to-Real tasks, the results indicate that the proposed framework has exceptional and robust generalization capabilities that are comparable to generalizing to new artificial faults.

The rest of the paper is organized as follows. In Section II, we introduce some background knowledge and underlying principles of MAML. Next, in Section III, we present the architecture of the proposed MAML-based few-shot bearing fault diagnostic framework, with detailed descriptions on establishing the test environment and model implementation. Section IV presents results for case studies performed on few-shot bearing fault diagnosis using both the Case Western Reserve University (CWRU) bearing dataset [32] and the Paderborn dataset [33], [34]. Section V concludes the paper by highlighting the effectiveness of the proposed model in bearing fault diagnosis with limited data and discussing future work directions.

II. PRINCIPLE OF MODEL-AGNOSTIC META-LEARNING

A. Meta-Learning

Meta learning is a general paradigm for few-shot learning. The objective of meta-learning is to learn a learning strategy to learn quickly on new tasks. In general, meta-learning algorithms involve two core processes: learning the transfer of knowledge across tasks and rapid adapting to new tasks [35]. Therefore, meta learning is intrinsically well-suited to perform anomaly detection or fault diagnosis tasks on previously unseen tasks. Specifically, meta-learning is able to find good initial parameters of a neural network model, such that with only a few gradient steps, the initial parameters can be updated to produce a model that provides a good performance for a new fault category.

In standard supervised learning, the goal is to learn a function that maps from some input $x$, which might be an image, to the label of that image $y$, as illustrated in Eqn. (1). In supervised meta-learning, the idea is very similar except that now it’s trying to map a training set $D_{\text{train}}$ (with corresponding data and labels) and a test input $x$ to the label of that test input, as described in Eqn. (2). Essentially, the goal of meta-learning is to train a model that when exposed to a training set, performs well on a corresponding test set of that task.

Supervised learning: $f(x) \rightarrow y$  \hspace{1cm} (1)
Supervised meta-learning: $f(D_{\text{train}}, x) \rightarrow y$  \hspace{1cm} (2)
The meta-training set $D_{\text{train}}$ is typically designed to contain a collection of little datasets of different categories [36]. At meta-test time, the goal is to identify new categories of labels not previously seen in the training process using a limited amount of test data $x$. The way to accomplish this goal is to find a model that can accurately match this function $f(D_{\text{train}}, x)$. Besides some successful deployment of sequence models in this effort such as recurrent neural networks [36] or temporal convolution networks [37], a very influential model was proposed by Finn et al. in [31], which was referred to as model-agnostic meta learning (MAML).

### B. Meta-Agnostic Meta-Learning

The MAML algorithm is model-agnostic. More specifically, it is agnostic both to the architecture of the neural network and also to the loss function. The backbone of MAML is to optimize for parameters that adapt quickly with gradient descent in two loops – an inner loop and an outer loop. All of these features provide many flexibilities for MAML, making it applicable to both reinforcement learning problems that maximize the expected cumulative reward function, and to supervised learning problems that minimize a certain loss function (cross-entropy, mean-squared error, etc.).

Performing a few-shot classification with MAML requires two stages – a meta-training stage and a meta-testing stage, as shown in Fig. 1. Formally, we consider MAML as a neural network $f_{θ}$ parameterized by $θ$, which will be updated to $φ_{i}$ using gradient descent when adapting to a new class $i$. During the meta-training stage, MAML operates in an inner loop and an outer loop. In the inner loop, MAML first computes the updated parameter vector $φ_{i}$ for each class $i$ using training data $D_{\text{train}}$, and then it evaluates the loss term on the validation data $D_{\text{valid}}$ sampled from the same class using the updated model parameters $φ_{i}$. The evaluated loss for each class $i$ can be written as

$$L (φ_{i}, D_{\text{valid}}) = L (θ - α \nabla_{θ} L (θ, D_{\text{train}}^{i}) , D_{\text{valid}}^{i})$$

(3)

where $φ_{i} ← θ - α \nabla_{θ} L (θ, D_{\text{train}}^{i})$ is the updated model parameter for class $i$. The loss term is typically the cross-entropy loss for classification tasks. In the context of bearing fault diagnosis, as illustrated in Fig. 1, different classes can represent different types of bearing defects, such as inner/outer race defects, ball defects, cage defects, among others.

In the outer loop, MAML aggregates the per-task post-update losses $L (φ_{i}, D_{\text{valid}}^{i})$ and performs a meta-gradient update on the original model parameter $θ$ as

$$θ ← θ - β \cdot \nabla_{θ} \sum_{i} L (φ_{i}, D_{\text{valid}}^{i})$$

(4)

where $β$ is the learning rate of the outer loop. At meta-test time, MAML is able to compute new model parameters based on only a few samples from unseen classes, and it uses the new model parameters to predict the label of a test sample from the same unseen class.

In summary, the essential idea of MAML is trying to find parameters of a neural network that does not necessarily have the optimal performance for different classes of data provided at the meta-training stage, but can quickly adapt to new (unseen) tasks.

### C. Meta-Agnostic Meta-Learning with Learnable Inner Loop Learning Rates

As illustrated in Fig. 1, there is a learning rate $lr$ in MAML for its inner loop gradient update, which is assigned as a fixed number $α$ in [31] and is shared among different classes for all update steps. However, this fixed and shared inner loop learning rate $lr$ can often affect MAML’s generalization capability and convergence speed [38], and the process of tuning this hyper-parameter $lr$ for a specific dataset can often be costly and computationally intensive.

Therefore, a variant of MAML is proposed in [38] to automatically learn the inner loop learning rate $lr$. Specifically, it tries to learn different learning rates for each layer of the neural network and for each step through back-propagation. By doing this, the learning rate $lr$ becomes a vector that accounts for different learning rates for each layer of the neural network. With this learnable $lr$ approach, elements in the learning rate vector $lr$ can learn to decrease their values as the training progresses, which may help promote faster convergence and alleviate overfitting. Therefore, the revised form of Eqn. (3) to compute the loss term of each class can be written as

$$L (φ_{i}, D_{\text{valid}}^{i}) = L (θ - lr_{i} \nabla_{θ} L (θ, D_{\text{train}}^{i}) , D_{\text{valid}}^{i})$$

(5)

### III. Few-Shot Bearing Fault Diagnostic Framework Based on MAML

#### A. Proposed Few-Shot Bearing Fault Diagnostic Framework

Few-shot classification is considered an instantiation of meta-learning in the field of supervised learning [39]. The standard few-shot learning is usually formulated as $N$-way $K$-shot problems, where $N$ is the number of new classes not seen in the meta-training process, while each class only has $K$ samples to train from.

The proposed few-shot bearing fault diagnostic framework based on MAML is illustrated in Fig. 2. At the meta-training stage, we’ll first optimize for a parameter set $θ$ of a neural network along the aggregated gradient descent direction of data from different bearing fault scenarios ($\nabla L_{1}, \nabla L_{2}, \nabla L_{3}$, etc.). As discussed in Section II, this parameter set $θ$ is optimized to achieve quick adaptation to new classes not previously seen at the meta-training stage, rather than achieving the optimal performance on classes it was directly trained on.

For example, as shown in Fig. 2, we can train the MAML-based diagnostic framework using data from bearing outer race defects and cage defects at different fault severity, and generalize it to detect new fault scenarios such as the ball defect and lubrication failure using a very small amount of data (e.g., 5 samples). This problem will be formulated as a 2-way 5-shot few-shot learning setting that is well-suited for MAML.

With the proposed MAML-based few-shot bearing fault diagnostic framework, it is envisioned that we can mitigate both data scarcity and data imbalance issues discussed in [10] by adapting to these new classes at the meta-testing stage, which
can yield a satisfactory performance but only requires a limited amount of data. Additionally, another appealing application is to recognize naturally-evolved bearing defects using models that are only trained on data collected from artificial failures, since most bearing failures evolve slowly over time and it might take months if not years to collect a sufficient amount of data to train supervised learning-based fault classifiers.

### B. Objectives

In the proposed MAML-based few-shot bearing fault diagnostic framework, the objective is to validate the performance of MAML on few-shot bearing fault diagnosis from the following aspects:

1. **Training Data Size**: Investigate the influence of training data size on the performance of MAML-based few-shot bearing fault diagnosis.
2. **New Artificially-Induced Bearing Failures**: Validate the performance of MAML to predict previously unseen artificial bearing faults in the laboratory environment.
3. **New Realistic Bearing Failures with Accelerated Aging**: Explore the generalization capability of MAML to predict real bearing failures with accelerated lifetime tests using data collected from artificially-damaged bearings.

Both 1) and 2) have been investigated in [10] using the Siamese Network-based few-shot learning method on the CWRU dataset. In order to perform a fair comparison, we strive to keep the test environment consistent with the benchmark study by also leveraging the CWRU dataset and assigning the same fault labels. More details regarding the CWRU dataset can be found in their website [32].

A list of all 10 fault scenarios is presented in TABLE III-B. Specifically, different classes are identified based on the...
consists of a segmentation method as [10], in which each data segment covers speed and loading condition. We also adopt the same data segmentation strategies, the entire CWRU dataset is partitioned into 10 classes, with each class having 1,980 data segments.

Table I shows the different categories of bearing failures selected from the CWRU dataset [32].

<table>
<thead>
<tr>
<th>Class</th>
<th>Label</th>
<th>Fault Location</th>
<th>Fault Diameter (mils)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Healthy</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Ball</td>
<td>0.007</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Ball</td>
<td>0.014</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Ball</td>
<td>0.021</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Inner Race</td>
<td>0.007</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Inner Race</td>
<td>0.014</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Inner Race</td>
<td>0.021</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Outer Race</td>
<td>0.007</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Outer Race</td>
<td>0.014</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Outer Race</td>
<td>0.021</td>
<td></td>
</tr>
</tbody>
</table>

Each class contains fault data collected at 3 different operating conditions: 1) 1 hp load at 1772 rpm; 2) 2 hp load at 1750 rpm; and 3) 3 hp load at 1730 rpm.

To further investigate the generalization capability of MAML in predicting real bearing failures as described in [3], we’ll also apply the proposed MAML-based few-shot learning framework to the Paderborn dataset [33], since the CWRU dataset only contains artificially-induced defects. The Paderborn dataset includes data of 32 bearings under test, which were faulted using various methods, and among them, 6 are normal ones, 12 are with artificially-induced damages, and 14 are with real damages caused by accelerated aging tests. There are only inner and outer raceway defects present for both artificial and real bearing failures, while damages at the rolling elements were not observed. Each of the 32 fault categories contains data collected at 4 different operating conditions with different combinations of rotational speed, load torque, and radial force. The experimental setup used to collect the Paderborn dataset is shown in Fig. 3.

For the Paderborn bearing dataset, we also select 13 representative classes from the total 32 classes, with 1 at the healthy condition, 6 have manually-initiated bearing defects, and the rest have real bearing failures resulting from accelerated lifetime tests. The selected classes can be distinguished based on their fault locations, causes of failure, and fault severity.

The threshold values used to determine different levels of fault severity have been discussed in detail in [9], where level 1 corresponds to a defect length smaller than 2 mm, level 2 corresponds to defect lengths between 2 to 4.5 mm, and level 3 corresponds to 4.5 to 13.5 mm.

A complete list of the selected classes is presented in Table II, and the goal is to successfully identify real bearing failures (categories 8 to 13) using the healthy and artificial fault data (categories 1 to 7). The rest of the data segmentation process is consistent with that performed on the CWRU data.

<table>
<thead>
<tr>
<th>Label</th>
<th>Fault Location</th>
<th>Cause of Failure</th>
<th>Severity</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Healthy</td>
<td>N/A</td>
<td>0</td>
<td>K001</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Outer Race</td>
<td>EDM*</td>
<td>1</td>
<td>KA01</td>
</tr>
<tr>
<td>3</td>
<td>Outer Race</td>
<td>EE†</td>
<td>2</td>
<td>KA03</td>
</tr>
<tr>
<td>4</td>
<td>Outer Race</td>
<td>Drilling</td>
<td>1</td>
<td>KA07</td>
</tr>
<tr>
<td>5</td>
<td>Inner Race</td>
<td>EDM</td>
<td>2</td>
<td>KI01</td>
</tr>
<tr>
<td>6</td>
<td>Inner Race</td>
<td>EE†</td>
<td>1</td>
<td>KI03</td>
</tr>
<tr>
<td>7</td>
<td>Inner Race</td>
<td>EE†</td>
<td>2</td>
<td>KI07</td>
</tr>
<tr>
<td>8</td>
<td>Outer Race</td>
<td>Pitting</td>
<td>1</td>
<td>KA04</td>
</tr>
<tr>
<td>9</td>
<td>Outer Race</td>
<td>PD†</td>
<td>1</td>
<td>KA15</td>
</tr>
<tr>
<td>10</td>
<td>Outer Race</td>
<td>Pitting</td>
<td>2</td>
<td>KA16</td>
</tr>
<tr>
<td>11</td>
<td>Inner Race</td>
<td>Pitting</td>
<td>1</td>
<td>KI04</td>
</tr>
<tr>
<td>12</td>
<td>Inner Race</td>
<td>Pitting</td>
<td>3</td>
<td>KI16</td>
</tr>
<tr>
<td>13</td>
<td>Inner Race</td>
<td>Pitting</td>
<td>2</td>
<td>KI18</td>
</tr>
</tbody>
</table>

*EDM: Electrical discharge machining.
†EE: Electric engraver.
‡PD: Plastic deformation.

Table II shows the different categories of bearing failures selected from the Paderborn dataset.

C. Model Implementations

The model follows the same architecture as the embedding function used by [40], which has 4 modules with a 3 × 3 convolutions and 64 filters, followed by batch normalization, a ReLU nonlinearity, and 2 × 2 max-pooling. The bearing vibration signals are sampled with a dimension of 4096 and converted to 64 × 64, and the last layer is fed into a softmax layer. For N-way, K-shot classification tasks, each gradient is computed using a batch size of NK examples.

The N-way convolutional model is trained with 1 gradient step and a meta batch-size of 25 tasks. For MAML with a fixed learning rate, the learning coefficient is chosen as α = 0.01. For MAML with a learnable inner loop learning rate lr, the initial value is also kept at 0.01, and it will be optimized with the training step [38]. We used a meta batch-size of 1 task for both 1-shot and 5-shot testing. All models were trained for 1500 iterations.

IV. EXPERIMENTAL RESULTS

In this section, we seek to validate the performance of the MAML-based few-shot bearing fault diagnostic framework. As discussed in Section III, we’ll specifically investigate its performance with different training data size and different unseen fault categories using the CWRU dataset. Additionally, we’ll also leverage the Paderborn dataset to predict naturally-evolved bearing failures using data collected from artificially-damaged bearings. The results of the proposed MAML-based
few-shot classifier will be compared against those obtained using the Siamese Network in [10], and we strive to keep their test scenarios as consistent as possible by leveraging the open-source code provided in [10].

### A. Ablation Study

1) **Influence of the Inner Loop Learning Rate $l_r$ and the Optimizer Type:** To start with, we evaluate the influence of the inner loop learning rate $l_r$ and the optimizer type (e.g., Adam, RMSprop, and stochastic gradient descent (SGD)) on the performance of the proposed MAML-based few-shot bearing fault diagnostic framework. We conducted a series of comparison experiments by setting the last 3 classes of TABLE III-B as the test dataset. In this way, data related to the outer bearing defect is only contained in the meta-testing data rather than the meta-training data. Then this setting is formulated as a 3-way few-shot learning problem. We select 9 data samples per class at the meta-training stage, and in the end, we also perform a 5-shot case study at the meta-testing stage.

The meta-testing stage is repeated 10 times, and the mean accuracy and standard deviation are presented in TABLE III. It can be observed that in the proposed 3-way 5-shot learning setting with the unseen outer race defect during meta-training, the proposed MAML framework with learnable $l_r$ can achieve the optimal performance with the Adam optimizer and an inner loop learning rate of $l_r = 0.01$. Specifically, the margin of mean accuracy is at least 3% when compared with other cases.

2) **Influence of Training Data Size:** We then evaluate the influence of training data size with test setup similar to the previous section as a 3-way few-shot learning problem. The optimal combination of the Adam optimizer and $l_r = 0.01$ is also selected based on results presented in TABLE III. The training data size per class is selected to be 4, 6, 9, and 12, and these samples are randomly chosen from the 1980 segments in each class in TABLE III-B. The results obtained from a Siamese Network-based few-shot bearing diagnostic framework proposed in [10] are also provided as a benchmark. The original results in [10] only include the case of 9 training samples per class, and we obtained the results of the remaining cases using the open-source code provided in [10]. For each size of the training set, we repeated the algorithm training and testing stages 5 times to overcome the randomness of both algorithms. The average accuracy and the standard deviation are presented in TABLE IV with different numbers of training samples per class.

It can be observed that in the proposed 3-way few-shot learning setting, the proposed MAML-based framework with learnable $l_r$ consistently outperforms the benchmark Siamese network by 20% to 30%. In addition, the proposed method can achieve an average accuracy of 99.77% with 5-shot learning and 12 training samples per class. While this result obtained using 12 samples per class is indeed satisfactory, we decide to use 9 samples per class for later experiments. This is because we seek to further investigate the performance of the proposed framework predicting different unseen fault types, and a near-perfect accuracy may overshadow other intrinsic processes undergoing in MAML. Additionally, we can make the test environment more consistent with the benchmark study [10], which also presents its results using 9 training samples per class.

### B. Predicting New Artificially-Induced Bearing Defects

The experiments performed in this section seeks to investigate MAML’s performance to predict artificially-induced bearing faults that are not seen at the meta-training stage. We also extract data from the CWRU dataset to test 1-way to 5-way classifications with 1 and 5 shots. A total of 10 rounds of experiments are performed to compare with the benchmark in [10]. Since [10] only provides the 1-way to 3-way results, the 4-way and 5-way results are also obtained using the open-source code in [10]. We also followed the order of labels presented in TABLE III-B, thus a 5-way classification indicates we are deploying data with class labels 1 to 5 as the meta-training data, and the rest will serve as the meta-testing data.

By randomly selecting 9 samples from the 1,980 available ones for each class, the complete results on $N$-way $K$-shot classification are presented in TABLE IV. It can be observed that MAML with a fixed $l_r$ is able to achieve 10% to 20% enhancement in average accuracy when compared with the benchmark Siamese Network, while MAML with learnable $l_r$ is able to achieve even larger improvements ranging from 20% to 30%. A comparison study for MAML with fixed and learnable $l_r$ is illustrated in Fig. 4, where the validation accuracy of the learnable inner update $l_r$ consistently outperforms, and is also more stable than the fixed $l_r$ case after 600 training epochs. This observation can be interpreted in such a way that the learnable $l_r$ can learn to decrease the learning rates
with larger training epochs and getting closer to the local optimum, which may help alleviate overfitting and promote faster convergence [38].

Besides examining the performance of MAML in N-way classification, it is also of practical value to investigate its performance at predicting different types of bearing faults not previously seen at the meta-training stage. An example of an unseen outer race defect is in fact the 3-way example demonstrated in TABLE IV, where the term “3-way” corresponds to classes 8, 9, 10 in TABLE III-B, and all of them contain data with bearing outer race defects that are not included in the meta-training data.

The results of MAML predicting other types of artificial bearing failures previously unseen at the meta-training stage are presented in TABLE VI, where the meta-testing data for the inner race fault are sampled from classes 5, 6, 7, and data for the ball failure are sampled from classes 2, 3, 4, as listed in TABLE III-B. It is shown in these 3 cases that MAML consistently outperforms the benchmark Siamese Network by around 20% with a fixed lr and by 25% with a learnable lr, when using the same training and test set over 1,500 training epochs. The 5-shot accuracy with learnable lr is around 90% for all 3 cases, validating the effectiveness of MAML predicting new types of bearing defects using a limited amount of data.

Another interesting case study is to predict bearing faults with higher levels of fault severity, or class (4, 7, 10) shown in TABLE III-B, while their fault types are not new to MAML, since the training data (classes 1, 2, 3, 5, 6, 8, 9) already covers all of the fault categories. This setting can still be formulated as a 3-way few-shot learning problem, and we are able to obtain around 10% improvement in average accuracy when compared with the earlier 3 cases with completely new fault categories. This finding can be explained as the task distribution of the same type of fault with different levels of fault severity is closer than the distribution of different fault types. This

<table>
<thead>
<tr>
<th>Number of Training Samples Per Class</th>
<th>4</th>
<th>6</th>
<th>9</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Siamese Network 1-shot [10]</td>
<td>53.22 ± 4.38%</td>
<td>59.69 ± 6.25%</td>
<td>63.79 ± 4.45%</td>
<td>66.08 ± 5.56%</td>
</tr>
<tr>
<td>Siamese Network 5-shot [10]</td>
<td>53.44 ± 4.93%</td>
<td>59.88 ± 6.04%</td>
<td>64.81 ± 6.32%</td>
<td>66.57 ± 5.74%</td>
</tr>
<tr>
<td>MAML (learnable lr) 1-shot</td>
<td>78.45 ± 2.94%</td>
<td>84.71 ± 2.53%</td>
<td>88.39 ± 2.95%</td>
<td>96.29 ± 1.66%</td>
</tr>
<tr>
<td>MAML (learnable lr) 5-shot</td>
<td>78.54 ± 2.41%</td>
<td>86.79 ± 1.21%</td>
<td>90.36 ± 1.38%</td>
<td>99.77 ± 0.20%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TABLE IV</th>
<th>3-Way K-Shot Classification Results Predicting the Unseen Outer Race Defect Using Different Numbers of Training Samples.</th>
</tr>
</thead>
<tbody>
<tr>
<td>N-way Accuracy</td>
<td>5-way Accuracy</td>
</tr>
<tr>
<td>Siamese-nets [10]</td>
<td>46.95 ± 4.23%</td>
</tr>
<tr>
<td>MAML (fixed lr)</td>
<td>56.31 ± 2.07%</td>
</tr>
<tr>
<td>MAML (learnable lr)</td>
<td>76.55 ± 1.57%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TABLE V</th>
<th>N-Way K-Shot Classification Results Predicting New Types of Artificial Bearing Faults (9 Training Samples Per Class).</th>
</tr>
</thead>
<tbody>
<tr>
<td>N-way Accuracy</td>
<td>3-way Accuracy</td>
</tr>
<tr>
<td>Siamese-nets [10]</td>
<td>63.79 ± 4.45%</td>
</tr>
<tr>
<td>MAML (fixed lr)</td>
<td>82.73 ± 3.12%</td>
</tr>
<tr>
<td>MAML (learnable lr)</td>
<td>88.39 ± 2.95%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TABLE VI</th>
<th>3-Way K-Shot Classification Results on the CWRU Dataset Predicting New Types of Artificial Bearing Faults.</th>
</tr>
</thead>
<tbody>
<tr>
<td>New Fault Category</td>
<td>Outer Race Defect</td>
</tr>
<tr>
<td>Siamese-nets [10]</td>
<td>63.79 ± 4.45%</td>
</tr>
<tr>
<td>MAML (fixed lr)</td>
<td>82.73 ± 3.12%</td>
</tr>
<tr>
<td>MAML (learnable lr)</td>
<td>88.39 ± 2.95%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>New Fault Category</th>
<th>Ball Defect</th>
<th>No (New Classes (4, 7, 10) only with Different Defect Sizes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Siamese-nets [10]</td>
<td>63.79 ± 4.45%</td>
<td>64.81 ± 6.32%</td>
</tr>
<tr>
<td>MAML (fixed lr)</td>
<td>78.53 ± 5.12%</td>
<td>81.57 ± 3.72%</td>
</tr>
<tr>
<td>MAML (learnable lr)</td>
<td>88.02 ± 4.03%</td>
<td>88.86 ± 1.85%</td>
</tr>
</tbody>
</table>
observation leads to a potential guideline of selecting the meta-training data from as many fault categories as possible when employing the MAML-based few-shot bearing fault diagnostic framework.

C. Predicting New Realistic Bearing Defects

This study also goes beyond identifying artificially-induced bearing defects by further exploring the generalization capability of MAML in predicting real bearing failures, which are formulated as Artificial-to-Real tasks. The objective is to use a combination of artificially-damaged bearings and healthy bearings to identify those with real damages. Due to intrinsic differences between data collected from these two scenarios, standard supervised learning methods can only achieve accuracies lower than 75% [9]. Additionally, these differences will also cause the transfer learning-based method to experience an obvious decrease in the average accuracy for Artificial-to-Real tasks when compared with generalizing to other artificial tasks [27].

The selected representative classes from the Paderborn dataset [33] are listed in TABLE II, and the order of which will be strictly followed while performing different N-way case studies. We can conduct a maximum of 6-way meta-testing by adapting to all of the 6 real bearing faults (classes 8 to 13). The source of meta-training data will only be the healthy (class 1) and the artificially-damaged bearings (classes 2 to 7). Both 1-shot and 5-shot tests are conducted using MAML with a fixed scalar lr and with a learnable lr vector.

The results of identifying real bearing defects are presented in TABLE VII, which demonstrates satisfactory 5-shot results with over 97% accuracy when dealing with 3 or fewer new realistic bearing failures using MAML with learnable lr. Fig.

### TABLE VII

<table>
<thead>
<tr>
<th>N-way Accuracy</th>
<th>6-way Accuracy</th>
<th>5-way Accuracy</th>
<th>4-way Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1-shot</td>
<td>5-shot</td>
<td>1-shot</td>
</tr>
<tr>
<td>MAML (fixed lr)</td>
<td>44.36 ± 4.02%</td>
<td>47.30 ± 2.29%</td>
<td>55.51 ± 3.02%</td>
</tr>
<tr>
<td>MAML (learnable lr)</td>
<td>55.21 ± 3.01%</td>
<td>62.58 ± 2.78%</td>
<td>73.63 ± 3.75%</td>
</tr>
</tbody>
</table>

### TABLE VIII

<table>
<thead>
<tr>
<th>Artificial damages (source domain)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bearing Code</td>
</tr>
<tr>
<td>KA01</td>
</tr>
<tr>
<td>KA03</td>
</tr>
<tr>
<td>KA05</td>
</tr>
<tr>
<td>KA07</td>
</tr>
<tr>
<td>KA08</td>
</tr>
<tr>
<td>KI01</td>
</tr>
<tr>
<td>KI03</td>
</tr>
<tr>
<td>KI05</td>
</tr>
<tr>
<td>K001</td>
</tr>
<tr>
<td>KA04</td>
</tr>
<tr>
<td>KB23</td>
</tr>
<tr>
<td>KB27</td>
</tr>
<tr>
<td>KI04</td>
</tr>
</tbody>
</table>

* EDM: Electrical discharge machining.
† EE: Electric engraver.
†† PD: Plastic deform.
5 illustrates a comparison study between TABLE VII and
the similar $N$-way 5-shot results obtained on the artificially-
induced bearing faults in TABLE V, and it can be observed that
MAML’s adaptation capability to real bearing failures delivers
similar, if not better results than adapting to artificially-induced
faults. This robust generalization capability, though much
desired, can be difficult to accomplish using other methods
such as [27] that do not involve leveraging past experience
and data to learn new tasks more quickly.

To further validate the generalization capability of MAML
in Artificial-to-Real tasks, we’ll compare it with some state-
of-the-art few-shot or transfer learning algorithms, such as the
Unfrozen Fine-tuning Net [24], the Feature Transfer Net [41],
and the Meta Relation Net [42]. To make a fair comparison,
we used the same Paderborn University dataset, the same data
pre-processing methods (fast Fourier transform, etc), and the
same source domain and target domain fault categories as [24],
which is presented in TABLE VIII. Four case studies with 1-
shot, 3-shot, 5-shot, and 10-shot settings are conducted.

The fault classification accuracy and standard deviation
obtained using the Paderborn dataset is shown in TABLE IX.
The results for the 6 few-shot transfer learning benchmark
studies are obtained by running the open source code provided
in [24]. To reduce the effect of randomness in data sampling,
the reported numerical results in this table are obtained by
taking the average of 10 trials.

It can be observed from TABLE IX that the best-performing
algorithm is the proposed MAML-based few-shot bearing fault
diagnostic framework, though not by a large margin. This is
because for the Feature Transfer Net and the 4 Unfrozen Fine-
Tuning are Nets, all of the $N$ way $K$ shot data are trained with
50 epochs, or their convolutional layer weights are updated
50 times to reach a good domain adaptation performance.
The Relation Net has been also updated using 100 epochs as
described in [24]. However, the initial parameters of MAML
are only updated once. This further validates MAML’s intrinsic
capability to quickly adapt to new tasks. In addition, the
standard deviation of the proposed MAML-based bearing fault
diagnostic framework is also 2 to 20 times lower than the
benchmark studies, demonstrating the robustness and consist-
ency of the proposed framework.

V. CONCLUSION

This paper proposed a MAML-based few-shot bearing fault
diagnostic framework, which can be generalized to identify
new fault scenarios not present in the meta-training data. The
effectiveness of the proposed framework in identifying new
artificial bearing faults was validated using the CWRU dataset,
and the results demonstrated a distinct advantage over the
benchmark study based on the Siamese Network. Specifically,
this advantage can be up to 25% when using MAML with
learnable inner loop learning rates lr and a small meta-training
set.

Since most of the real-world bearing failures are evolved
naturally over time, we also applied the proposed method to
the Paderborn dataset to further explore the generalization ca-
pability of MAML when adapting to real bearing failures. The
results demonstrate that MAML is able to deliver comparable
performance as adapting to artificial bearing faults. Addition-
ally, MAML is also able to deliver competitive performance
for Artificial-to-Real tasks when compared with several few-
shot learning methods. These results offer promising prospects
for identifying naturally-evolved bearing failures using data
collected from laboratory tests with artificially-induced faults.

It is also worthwhile to mention that the vanilla MAML
algorithm in [31] may still suffer from a variety of issues,
such as being sensitive to the inner loop hyper-parameters,
having unstable training processes and limited generalization
capabilities, among others. These issues have been extensively
discussed in [38], which also provides 6 different approaches
to mitigate them. This work primarily focused on applying one
of the approaches, namely learning per-step learning rates, to
the vanilla MAML algorithm, and has achieved noticeable im-
provement in performance. Therefore, a desirable future work
direction is to implement other optimization strategies in [38],
most notably on applying batch normalization parameters and
optimizing on per-step target losses, to develop a more robust
and strongly generalizable few-shot bearing fault diagnostic
framework.

REFERENCES

anomaly detection via model-agnostic meta-learning,” in Proc. Int. Conf.
of intelligence,” [Online]. Available: https://www.technative.io/cyber-
physical-systems-the-new-and-emerging-systems-of-intelligence/, 2020,
edge computing framework for cyber-physical-social services,” IEEE
remaining useful life with deep convolution neural network,” IEEE


Shen Zhang (S’13–M’19) received the B.S. degree with the highest distinction in electrical engineering from Harbin Institute of Technology, Harbin, China, in 2014, and the M.S. and Ph.D. degrees in electrical and computer engineering from Georgia Institute of Technology, Atlanta, GA, USA, in 2017 and 2019, respectively. Since August 2019, he has been with Joby Aviation, San Carlos, CA. His research interests include design, control, condition monitoring, and fault diagnostics of electric machines, control of power electronics, powertrain engineering for electric propulsion, machine learning and reinforcement learning applied to energy systems.
Fei Ye (S’16-M’19) received the M.S. degree from Northeastern University, Boston, MA, USA, in 2014, and the Ph.D. degree from University of California, Riverside, CA, USA, in 2019, both in electrical and computer engineering. She is currently a post-doctoral researcher at University of California, Berkeley, CA, USA.

Her research interests include intelligent vehicles and transportation, deep learning, deep reinforcement learning with domain adaptation and their applications in behavior prediction, decision making and trajectory planning.

Bingnan Wang (M’12-SM’15) received his B.S. degree from Fudan University, Shanghai, China, in 2003, and Ph.D. degree from Iowa State University, Ames, IA, USA, in 2009, both in Physics. He has been with Mitsubishi Electric Research Laboratories (MERL), located in Cambridge, Massachusetts since then, and is now a Senior Principal Research Scientist.

His research interests include electromagnetics and photonics, and their applications to wireless communications, wireless power transfer, sensing, electric machines, and energy systems. He has published over 60 technical papers in related fields and filed over 30 patents and applications.

Thomas G. Habetler (S’82-M’83-SM’92-F’02) received the B.S.E.E. degree in 1981 and the M.S. degree in 1984, both in electrical engineering, from Marquette University, Milwaukee, WI, and the Ph.D. degree from the University of Wisconsin-Madison, in 1989.

From 1983 to 1985 he was employed by the Electro-Motive Division of General Motors as a Project Engineer. Since 1989, he has been with the Georgia Institute of Technology, Atlanta, where he is currently a Professor of Electrical and Computer Engineering. His research interests are in electric machine protection and condition monitoring, switching converter technology, and drives. He has published over 300 technical papers in the field. He is a regular consultant to industry in the field of condition-based diagnostics for electrical systems.

Dr. Habetler was the inaugural recipient of the IEEE-PELS “Diagnostics Achievement Award,” and a recipient of the EPE-PESC “Outstanding Achievement Award.” the 2012 IEEE Power Electronics Society Harry A. Owen Distinguished Service Award, the 2012 IEEE Industry Application Society Gerald Kliman Innovator Award. He has also received one Transactions and four conference prize paper awards from the Industry Applications Society. He has served on the IEEE Board of Directors as the Division II Director, and on the Technical Activities Board, the Member and Geographic Activities Board, as a Director of IEEE-USA, and is a past president of the Power Electronics Society. He has also served as an Associate Editor for the IEEE TRANSACTIONS ON POWER ELECTRONICS.