Meta Learning Approach Based on Episodic Learning for Few-Shot Image Classification

Syeda Roohi Fatema * and Sumana Maradithaya

Department of Information Science and Engineering, Ramaiah Institute of Technology, Bengaluru, India
Email: syedaroohifatema@gmail.com (S.R.F.); sumana.m@msrit.edu (S.M.)
*Corresponding author

Abstract—**Deep learning encompasses the inherent properties of scattered data and acquires a more abstract representation of data than conventional machine learning techniques. Nevertheless, existing deep learning algorithms perform inadequately on novel problems such as image classification as they generally require an extensive number of samples intended to train the model. One of the efficient methods to resolve the indicated drawback is meta-learning, acclaimed as learning to learn. Using antecedent knowledge to aid in the learning of new tasks improves meta-capacity for generalization to unfamiliar tasks. Meta-learning determines previous assignments intending to discover a representation that is easily adaptive to unknown challenges. Meta learning methodologies help find these components through multitudinous learning episodes by learning to solve a set of tasks instead of solving a single task at a time. Episodic meta-learning seeks to imitate a realistic environment by producing small artificial tasks from a substantial set of training tasks for meta-training and then moving on to the related method for meta-testing. The research is evaluated with meta learning algorithms like Prototypical Network and proto-Model-Agnostic Meta-Learning (MAML) with episodic meta learning on SVHN and Omniglot dataset reporting compelling enhancements on public benchmarks. In this research, the obtained results demonstrate a notable improvement and enhancement compared to existing methodologies, indicating a successful and impactful improvisation in the proposed methodology.**

Keywords—**meta learning, deep learning, few shot learning, prototypical networks proto-Model-Agnostic Meta-Learning (MAML)**

I. INTRODUCTION

Deep learning has noticed tremendous success and has established itself as an effective method in extensive areas, including computer vision and natural language processing, even though it is extremely dependent on an extensive amount of labelled training data. Meta-learning aims to make it possible for models, specifically deep neural networks ascertain how to perform efficiently on new tasks from a barred amount of data samples. Meta learning solves new hidden tasks with a few sets of examples [1]. In classic machine learning domains, there is often a large dataset taken which is peculiar to a task and intended to train the model while using the dataset for regression/classification purposes. That differs significantly from the way people use their prior knowledge to efficiently acquire an additional skill from just a small number of examples. Technically, this entails using metadata of an algorithm to understand how the method of autonomous learning may become flexible in managing learning difficulties, which leads to improving the consummation of present learning algorithms. The inductive bias of the respective learning algorithm refers to the set of assumptions that it makes about the data. Humans are naturally capable of picking up new abilities fast. For instance, by observing a single knife, we can distinguish all knives from other cutlery pieces, such as spoons and forks. Our competencies are substantially more extensive than merely being able to recognize new items, learn a new language, or figure out how to use a new tool, being able to learn new skills and adapt to new situations quickly (revolve around only a few instances or demonstrations). Machines, particularly deep learning algorithms, however, often learn rather differently. They struggle with generalization and need huge quantities of information and computation. Humans excel at adapting and learning quickly because they use their existing experience and expertise to address novel problems. Similar to this, meta learning makes use of prior knowledge obtained from data to complete fresh tasks more quickly and effectively. Meta learning aims to enhance the outcomes and performance of the learning algorithm by adjusting certain properties of the algorithm in response to the results. Using meta-learning, researchers may determine algorithms that give accurate predictions from datasets [2, 3]. The primary goal is to develop a meta-learning approach based on episodic learning for few-shot image classification in order to enhance the model’s ability to rapidly adapt and generalize to novel classes with limited labeled examples.

II. LITERATURE REVIEW

Metadata from learning different algorithms is fed into meta- learning algorithms. They then formulate forecasts and offer data regarding the outcome of these learning algorithms. In a learning model, an image’s metadata could include things like its size, resolution, style, creation

Manuscript received October 7, 2023; revised December 27, 2023; accepted January 3, 2024; June 28, 2024.

doi: 10.18178/joig.12.2.205-214

date, and owner. The most significant difficulty in meta-learning is systematic experiment design. Meta-learning uses a small number of observations to generalise effectively to tasks that are unperceived. In conditions where there are few data points available, the goal is to simultaneously learn a representation and acclimate it to previously unrelated tasks [4]. Even while meta-models could be expounded over a wider range of learning problems the present literature refers to this process as "few-shot learning". It is thought to be the key environment for estimating meta-learning algorithms [5, 6]. Li *et al.* [7] introduced the idea of few-shot learning in 2003, highlighting the difficulty in applying previously acquired knowledge to the acquisition of new categories. Few-shot learning is planted entrenched on the concept of Bayesian framework in the initial phases of the research, and the sample's class probability reasoning is derived by integrating the model parameters with preceeding probability along with the posterior probability. Researchers advanced deep neural network models to address this issue of few-shot image classification as deep learning design has become more advanced. Most of the few-shot learning techniques currently in use have recently incorporated deep learning. The long-standing issue of the requirement for substantial and broad datasets is resolved by this method. Few-shot learning often stands in need to learn features of a training sample. Meta-learning algorithms that rely on a small number of data points might be brittle when the set of tasks is heterogeneous [8, 9].

The ongoing problem of the requirement for substantial and extensive datasets is resolved by this approach. Few-shot learning becomes competent in learning the features of a few labelled images for training samples to categorise new test images. Many image processing applications, namely image segmentation [10], image recognition [11], image classification, and image retrieval [12–14], currently make extensive use of few-shot learning. Additionally, a survey of few-shot picture classification has significant practical significance. Large-scale labelled data collection is challenging in industries like medicine [15] and public security [16], which hinders deep learning model's performance. Few-shot learning is an efficient methodology for the issue of some high-performance models being unable to generalise in new classes as a result of the lack of training data, allowing particular high-performance models to be designated in other disciplines. Few-shot learning has currently been reviewed by different scholars. The study advancement of few-shot classification models along with algorithms in line with techniques centred on data augmentation, model fine-tuning and transfer learning was explicitly introduced by Zhao *et al.* [17]. Few-shot learning has been extensively examined in the scholarly literature, and Wang *et al.*[18] have systematically classified and organized it into a cohesive taxonomy based on the perspectives of model, data, and algorithm.

Prototypical networks are also associated with the neural statistician considering the procreative modelling literature [19], which extends the variational autoencoder to learn generative models of the datasets instead of individual points [20–22]. The "statistic network" which converts a collection of data points caught up in a statistic vector, is one element of the neural statistician. So, to get an approximative posterior over the statistic vector, it does this by encoding each point inside a dataset, catching a sample mean, and implementing a subsequent processing network. By addressing each character as a distinct dataset and establishing predictions on the class whose approximate posterior over the statistic vector has the least KL divergence from the posterior constrained by the test point. The model was tested for one-shot classification on the Omniglot dataset. In a feedforward model that can yield effective results, researchers also generate a summary statistic about each class by adjusting the top layer weights. The tasks of the model with regards to the parameters: when sensitivity can learn an internal feature representation that is high, small local changes to parameters are suited for various tasks using a method of training the model's parameters [23, 24]. Contemporary data mining systems are only as powerful as their users. These technologies allow several algorithms incorporated into a single system, but their selection and combining ought to be completed before the system is launched, typically by a skilled user. According to several researchers, if machine learning systems are to be useful to non-specialists at all, the selection of learning and data transformation methods should be automated. Others argue that existing technology does not allow for complete automation in the aspect of the data mining process. An intermediate solution is the development of helper systems to assist in the selection of the appropriate learning algorithm(s). There appears to be an underlying consensus that meta knowledge should be smoothly included in the data mining system, regardless of the proposed approach. Meta-learning is scrutinised with the development and implementation of learning algorithms to acquire meta knowledge to assist machine learning users during the model selection process. Meta-learning for image classification holds great promise in overcoming the data scarcity challenge by enabling models to learn a generic understanding of tasks and quickly adapt to new ones. As research in this area progresses, meta-learning approaches are expected to play a vital role in enhancing the efficiency and applicability of image classification models across diverse domains.

This paper demonstrates that by establishing a better representation with episodic learning and better adaptability for the meta-learned model, employing a guided technique to designate the data greatly enhances the results on the workload. As investigations advance in this field, it is anticipated that meta-learning methodologies will play a crucial role in improving the effectiveness and adaptability of image classification models across a variety of domains. The resilience of meta-learning algorithms is further enhanced by this process when out-of-distribution tasks are available. At the time of meta-training and meta-testing, these techniques are utilized to generate the context and the query sets (in different combinations). The objective is to determine whether drawing the query and context sets in the form of episodes during meta-training can improve meta-training, and also assess the performance of actively drawing context sets on the rate of adaptation during meta-test. The combination of methods carried out during meta-train and meta-test time yields the best results, it also exhibits that the results still improve when the selection is just transferred for the adaptation at meta- test time in the form of episodes. We apply our methodology to proto–MAML [25] and prototypical networks [19], indicating that metric-based and optimization-based episodic meta-learning [26] can both benefit significantly from properly choosing the data. In-depth, experiment is conducted on Omniglot [27] and SVHN [28] datasets. The empirical findings support the theoretical study of the computational cost of the strategy, which has linear complexity in picking respective data points for episode configuration. The results of consequent studies signify that few-shot categorization methods are more accurate. Fine-grained image categorization and facial recognition have both benefited from the application of metric learning. The main goal is to discover an embedding function that indicates how similar samples within a category are to those within various classes. The query photos will be categorized when the embedding function has been learned. The Siamese network is made up of two identical sub-ConvNets that work together to decrease the distances between paired data that have the same label while maintaining a large distance between the data that have different labels. In contrast to absolute pair-wise distances, triplet loss aims to concentrate on relative distances. It's been used extensively in assignments with finer details.

III. Materials and Methods

The research aims to establish a comprehensive meta-learning approach based on episodic learning for few-shot image classification, contributing to advancements in the field. As much data the model can comprehend is typically delivered into machine learning applications [18]. This is to ensure that the model predicts more precisely in many machine learning applications. Few shot learning, on the contrary seeks to develop precise machine learning models with lesser training data. Few-shot learning, differs from traditional supervised learning due to the fact it makes predictions based on few data. Few-shot learning, which refers to acquiring new ideas from a small number of examples, is a skill that humans naturally have but that machines still lack. As suggested in Algorithm 1, improving on this feature might result in improved algorithms that flexibly grow their knowledge without requiring big labelled datasets [29]. The emphasis is on few-shot classification, which involves assigning previously unknown cases to one of N new 'test' classes using only a modest number of samples as references about each class.

The generic term for Few-Shot Image Classification (FSIC) task is an N-way K-shot problem [30]. There are many categories in the training set for few-shot learning, numerous examples are involved in each category. During the training phase, K samples (in N × K images) are arbitrarily chosen from each category of image samples of the training set to serve as a support set. The prediction object of the model, also known as a query set, is then chosen from the remaining samples from each of the N categories of data. The classification task is known as the few-shot image classification if K is extremely small (often K < 10); a task signifies a one-shot image classification when K = 1; and zero-shot image classification task is considered when K = 0. Episode training technique is typically applied for few-shot learning. The support set and the query set are both present in an episode. A few-shot learning activity is therefore equivalent to one episode. A few-shot image classification job aims to precisely categorise the images of the query set using support set that is already in existence. Hence, the model must learn how to differentiate between these N categories from N × K samples.

Algorithm 1: Extract n-way k-shot images

Require: n way, k shot, N classes, X training dataset
1. Select N number of samples randomly from the corresponding support set
2. random. shuffle (N) # each folder corresponding to a group
3. Train a model to predict the sample with an optimization algorithm
4. Update model parameters using the loss calculated
5. Train all K-shots
6. X.appen (images [0:k]) concerning K images in the corresponding group
7. Evaluate the model parameters on the query set.

A. Types of Meta Learning

The methodology of enhancing a learning algorithm over several learning episodes is termed meta-learning explained in Fig. 1. A model can understand how to swiftly adjust to new tasks using meta learning and minimum amount of training data. This is especially helpful for scenarios when the work distribution is flexible and there is a steady influx of new tasks. By allowing a model to gain knowledge from a small number of instances, meta learning can result in effective data utilization [31]. This is to ensure that meta learning algorithms can generalise to new tasks besides fewer training examples by learning to recognise the underlying structure. As described meta learning is categorized into different categories as shown in Table I.

TABLE I. Meta Learning Classification

Model Based	Metric Based	Optimization based
Memory Augmented Nueral Networks (MANN)	Convolutional Siamese Nueral network	LSTM Meta learner
Meta Networks	Matching Networks	MAML
	Prototypical Networks	Reptile
		Proto-MAML

1) Optimization based

Optimization-based meta-learning algorithms can be good at learning with just a few examples and adjusting optimization. As an instance, the gradient-based optimisation used by deep learning models to learn is not intended to handle a small number of training samples and does not converge within a limited number of optimization steps. However, this is the problem that meta-learning algorithms with an optimization-based approach aim to overcome, and these optimization techniques are based on hyperparameters. At the outset of the learning process, hyperparameters are established, and their values govern the entire learning process. Thus, hyperparameters have a direct effect on the effectiveness of the training process.

This approach assists well-designed models in adapting to unknown tasks during the testing stage. A Model-Agnostic Meta-Learning (MAML) method, which is considered as well-known example of meta-learning, was proposed by Finn *et al.* [32]. The core methodology of the MAML is learning a neural network initialization that follows a fast gradient direction to efficiently categorize novel classes. Furthermore, the Latent Embedding Optimization (LEO) method employed a learning algorithm similar to the MAML, incorporating an inner loop for task-specific parameter initialization and an outer loop for modernizing the parameter [33]. Instead of directly learning the explicit high-dimensional model parameters, LEO separated the gradient-based adaption process within a low-dimensional latent space and learned the productive distribution of model parameters. The prior meta-learning-based methodology simply proceeded with a pure meta-training criterion by training the model from scratch. In more recent picture recognition challenges, however, researchers have tried combining meta-learning and fine-tuning to create a hybrid solution. To maximize the use of the advantages of transfer-learning and meta-learning in the FSL scenario, Chikontwe *et al.* [34] developed a Meta-Transfer Learning (MTL) approach. A general meta-learning architecture called SNAIL is made of a casual attention layer and an interleaved time convolution [35]. To combine the data from previous experiences, the convolution network learns the feature vector from the training samples. To complete the few-shot learning tasks, the causal attention layer chooses knowledge from the accumulated experience to be popularised towards a new task. To address few-shot classification tasks, Chu *et al.* [36] suggested a reinforcement learning model contingent on the maximum entropy block sampling technique.

The usage of gradient over lengthy inner optimizations causes a lot of computation and memory concerns. Many contemporary approaches are expressed as specific examples of a generalized inner loop meta-learning framework in a unified gradient-based meta-cognitive perspective. The prominent fundamental goal of the study of meta-learning is to comprehend an interplay between the mechanism of learning and the particular settings in which the mechanism is relevant. Due to the short amount of training samples in few-shot image classification tasks, the learner generally overfits, and before converging to yield a better result it is typically trained for millions of iterations. These issues not only impair the learner's performance but the model's ability to classify data accurately is also affected.

2) Metric-learning based methods

Metric-learning approaches use a straightforward methodology that compares the distances or similarities between the query image along the corresponding labelled image in the support set. To be more explicit, the complete support set is encoded into the latent representation space first. The query image is then projected into the above space, and then the similarity between each query image and support image is computed. The category of each inquiry image ought to be predicted using the correspondence measurement.

Prototypical network (ProtoNet) which is a classical metric-learning-based methodology [21]. As the prototype representation, the mean vector of feature embeddings of respective support classes was accurately calculated. The similarity between each query image and its corresponding prototype is then learned for categorization. In particular, the nearest-neighbour classifier is used for prediction during the testing step. Another exemplary metric-learning method was the relation network (Relation-net) suggested by Sung *et al.* [37]. The RelationNet proposed a covariance metric network (CovaM-Net) which arrogated new covariance metric concerning second-order local covariance representation for each class, rather than conventional first-order class representations (e.g. mean vector), alternately choosing a specific metric function. There are no data-independent parameters in the classifier considered in metric-learning approaches (for example, the nearest-neighbor classifier). As a result, there is no purpose to use a fine-tuning process during the testing stage. Metric-learning [38] or non-parametric methods have so far mostly been used in the well-known but particular few-shot meta-learning applications. By merely analyzing validation points with training points and predicting the label of matching training points, non-parametric learning is intended to be performed at the inner (task) level. This method has been achieved by using Siamese, matching, prototypical, relation, and graph neural networks along with a certain sequence [39–41]. In this scenario, metric learning (identification of a feature extractor) which depicts the data properly for comparison corresponds to outer-level learning. The extractor is utilized for target tasks as before after being learned on source tasks. A meta-learning model needs to be trained on numerous tasks before being further optimized for new tasks. A task is essentially a supervised learning issue (such as regression analysis or image classification). The concept is focused on how to extract prior information from a group of tasks, allowing for efficient learning on new tasks. For an image classification problem, an ideal setup would comprise many classes, each with at least a few samples.

3) Model learning based methods

A specialized type of meta-learning approach that concentrates on learning a model or representation of the underlying tasks or environments is known as model-based meta-learning. These meta-learning models are not based on any broad assumptions. Instead, they rely on models that are specifically developed for quick learning; these models quickly change their parameters with minimum training. This quick parameter change is accomplished using internal architectures managed by another meta-learner model.

Acquiring a model that successfully generalizes and adapts to new tasks with less data is the main objective. In model-based meta-learning, the meta-learner develops a model to represent the dynamics, structure, or patterns that are common to several connected tasks. This model can be used to produce simulated data, plan activities for new tasks, or make predictions. Model learning-based methods generally engage a meta-training approach on a series of few-shot tasks derived from base classes during the training stage. The goal of the model-based meta-learning is to develop the model parameters by leveraging the general knowledge acquired through the completion of multiple tasks, enabling the model to solve associated tasks adaptively and enhancing the performance of few-shot learning classification tasks. One crucial area of few-shot learning is model-based meta-learning approach. Andrychowicz *et al.* [42] develops meta-learner contingent on LSTM and demonstrates a method to convert the creation of an optimisation algorithm into a learning problem. Another LSTM-based meta learner was suggested by Ravi *et al.* [43] to learn important parameter updates and general learning model initialization. The Memory-Augmented Neural Network (MANN), which trains the Neural Turing Machine (NTM) [44] used as a metal earner, was proposed by Santoro *et al.* [45] as an alternative to LSTM. This neural network has better memory than usual. The sample feature information is retained in the shown external memory module, and the reading and writing processes are optimised by the meta learning algorithm.

IV. Proposed Approach

A. Meta Learning Process

Meta-learning, commonly referred to as learning to learn, focuses on teaching models how to learn and swiftly adapt to new tasks or domains. Meta-learning seeks to train models to learn how to learn rather than teaching them for a specific activity, allowing them to generalize across various tasks and increase their learning efficiency. Meta-training and meta-testing/adaptation are the two main stages in the meta-learning process [46]. The model first learns its parameters from a training dataset made up of photos from different classes and then utilizes those parameters as prior knowledge to fine-tune its parameters about a limited training set as mentioned in Fig. 1. Only a few instances from each respective class are used to train a model, and it is then tested against examples from those classes that were withheld from the original dataset, much like it will be tested when only a few training examples from novel classes are used. Each episode, or pair of train and test data points, makes up each training example in this scenario.

A distribution of tasks is established during the meta-training phase. Each task comprises a particular learning challenge or a challenge that the model must overcome. The tasks can range in complexity and come from an extensive variety of fields. The model is exposed to a small labelled dataset or a few training samples for each task (few-shot learning setting). Based on these training samples, the model is tuned to function properly. In the adaptation process to improve the model's performance, the parameters are modified based on the training data. Finding a collection of initial parameters that enables the model to easily generalize to new tasks is the objective. On numerous tasks selected from the task distribution, the procedure is repeated. The model builds up knowledge and patterns that it can use on future tasks as a result of the variety of tasks it is assigned.

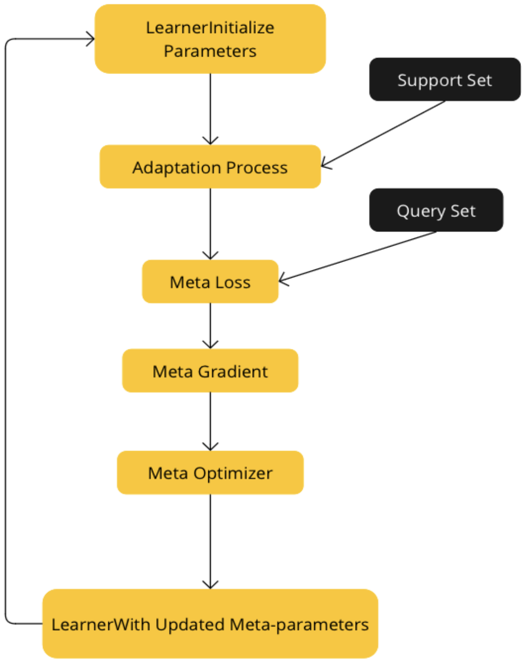

Fig. 1. Meta Learning process.

To quickly adjust to new tasks, the model requires to establishment of generalizable representations that incorporate similarities and transferable information across tasks. The model is evaluated using new, unsolved problems when the meta-training is finished. These tasks are distinct from those that were part of the meta-training. The model adjusts to the new task using the knowledge and parameters acquired during meta-training. Using the provided labelled data, it executes a few gradient updates or optimisation processes. The model adjusts its parameters for each new assignment to perform well on test examples and accumulate prior knowledge so that, at the time of inference, it can swiftly pick up knowledge unique to a certain task with just a few training instances. So forth the parameters are originally learned from a training dataset that is made up of images from different classes, and they are then used as prior knowledge to further fine-tune the characteristics according to the limited training set. Based on metrics like accuracy, error rate, or similar task-specific measurements, the performance of the updated model on the new task is determined. To enable models to learn is the basic methodology of meta-learning, which enables them to quickly adapt to new tasks or domains with minimum labelled data. The models develop generalised knowledge and optimisation techniques by training on a variety of tasks during meta-training. These techniques can then be applied to new tasks during meta-testing or adaptation [47–49]. This method is especially helpful in situations where getting labelled data for every new task might be difficult or expensive.

B. Episodic Meta Learning

Episodic meta-learning [26] is an approach in meta-learning that focuses on learning from episodes or tasks rather than individual data points It is particularly useful in instances of few-shot learning where the model must swiftly adapt to new tasks with an absence of labelled samples [50]. Episode comprises organizing training in a series of learning problems, each relying on small "support" and "query" sets to mimic the few-shot conditions encountered during evaluation as explained in Algorithm 2. A query data set is preowned to calculate and optimise a training loss given the specialised model. The context data set is utilised for model specialization and imitates tiny datasets used for adaptation at meta-test time as in Fig. 2. By generating a collection of small scale simulated tasks from a greater set of training tasks during meta-training and conducting meta-testing in an analogous way. Episodes are generated by a task distribution which has been sampled. Each episode represents a distinct assignment or instructional scenario as explained in Algorithm 2. The tasks can emanate from different fields and range in challenges. There are a few common parameters used to initialize the model. The generic knowledge that is anticipated to apply to a variety of jobs is captured by these characteristics [51]. There are a few common parameters used to initialize the model. The generic knowledge that is anticipated to apply to a variety of tasks is captured by these characteristics. The model's parameters are upgraded using gradient-based optimization techniques, for instance, Stochastic Gradient Descent (SGD), after adaptation on the support set. A conceptual change in machine learning is represented by episodic meta-learning, which employs episodes to train models to swiftly adapt to new tasks with less data. This novel method tackles the difficulties associated with few-shot learning scenarios, in which models have to efficiently generalize from a limited number of examples per class. Inspired by human learning, episodic meta-learning emphasizes the capacity to apply knowledge gained from a variety of experiences to new circumstances. In the realm of few-shot image classification, our research pioneers a transformative meta-learning approach grounded in the rich framework of episodic learning. By orchestrating learning episodes that mirror real-world few-shot scenarios, our model evolves dynamically, distilling a meta-knowledge.

To reduce loss or error on the query set, modifications must be made. A meta-objective, which frequently requires combining the losses or errors across numerous episodes is utilised to optimize the model parameters. The meta-objective directs the model's learning of parameters which help in efficient adaptation and task generalization. The model is tested on new episodes that were unseen during training after the meta-training is finished. The model receives a support set and a query set for every new episode. It must adjust to the support set and correctly predict or classify the query set.

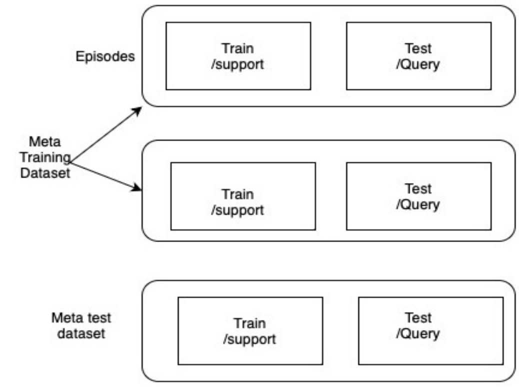

Fig. 2. Episodic Meta learning.

Algorithm 2: Episodic learning

Input: Episodic input samples Ex
Output: Episodic output samples Ey
1. Generate random samples and select k samples from Ex
2. sample subsets of Labels L
3. Sample images constituting support set and query set
4. Construct a support set and query set such that they contain classes of L
5. Train a series of mini-batches in which the image is affiliated to either support or query set
6. Learn the function f by minimizing the episodic loss
7. Generate a new sample Xn and learn the f1 by minimizing the update rule
8. Repeat the method until convergence
9. For each episode e: predict the label Ey for the occurrence of the input Ex Ey = f(Ex)

C. Prototypical Network

A metric-based meta-learning method which acts similarly concerning a k-nearest neighbor classification is called the Prototypical Network, or ProtoNet [19]. A new example is categorized using metric-based meta-learning techniques dependent on a distance function, d and all the components in the support set. How the prototype concept is learned in a feature space is how ProtoNets carries out this concept as explained in Fig. 3. Each input in the support set is initially encoded in an L-dimensional feature vector by ProtoNet using an embedding method.

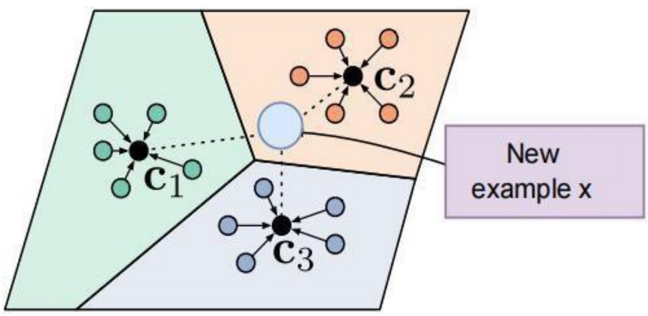

Fig. 3. Prototypical network.

The principle underlying the prototypical network is that there could be an entrenching in which a variety of points gather around a single prototype rendition for each respective class. It seeks towards learning sample-averaged per-class prototypes in the feature space. To be more precise, embedding functions with trainable parameters, prototypical networks calculate the M-dimensional representation or prototype for each class. Additionally, every prototype is its class's mean vector of embedded support points. Prototypical networks offer an interesting method for few-shot and zero-shot learning as they are more effective than previous meta-learning methods. By contriving distances to prototype representations for each class, they establish a metric space where classification is done. The black dots present next to the class label are the averages that do not seek to enable model recognition of the images in the prototypes, while the coloured dots show an encoded support training set followed by generalization to the test set. The prototype network learns to integrate these instances into a shared feature space given a set of labelled examples for each class. This is often accomplished by mapping the input examples to embedding vectors using a neural network design, such as a Convolutional Neural Network (CNN). The network determines the prototype for each present class by aggregating the embedded vectors of the labelled instances that correspond to that class. For each class in the embedding space, the prototypes act as centroids or as points of reference. The network computes additional examples' embeddings using the learned embedding function to classify them. The distances between each class's prototypes and embeddings are then calculated.

D. ProtoMAML

A combination of the prototypical networks and MAML is called protoMAML, according to Liu *et al.* [25]. It provides prototype networks with an adaptation mechanism akin to (fo) MAML and is demonstrated to perform significantly better on Meta-Dataset than foMAML. MAML has a challenge with how to construct the output classification layer. The output layer must be initialized with zeros or randomly in each iteration if each task has a distinct number of classes. We just begin with random predictions, even if there are always the same number of classes. To get a reasonable categorization outcome, this calls for numerous inner loop phases. To address this issue, Novais [49] proposed combining the advantages of Prototypical Networks and MAML. Specifically, prototypes are used to initialize the output layer, resulting in powerful initialization.

Using Algorithm 3, ProtoMAML is implemented. by sampling the batch of tasks, along with support and query set for each task, at the beginning of each training step. This simply implies that several support-query set pairings from our sampler in the case of few-shot classification are sampled. Adjust the current model on support set for each task. However, a copy of the model is improved because the initial parameters are remembered for other tasks such as outer loop gradient update and subsequent training steps. The model after it has been adjusted, the first-order gradients concerning the starting values are computed. Since they directly depend on, gradients in accordance to the output layer initialization is taken into account, or prototypes, in contrast to simple MAML.

The model proceeds in two ways to accomplish this. First, apply the original model to the support sets to enumerate the prototypes. Disconnect prototypes when initialising the output layer to stop the gradients. This is because gradients between prototypes and the original model are not taken into consideration in the inner loop. By including prototype networks, ProtoMAML expands the MAML framework to few-shot learning settings. Few-shot learning aims to acquire an understanding of recent ideas or classes from a small number of labelled examples. To resolve this issue, ProtoMAML makes use of the concept of prototype networks. Learning a set of task-specific prototypes that capture the distinctive qualities of each class or concept is the main goal of ProtoMAML. The benchmarks for categorizing new cases are these prototypes.

Algorithm 3: Proto MAML for Few-Shot Supervised Learning

Require: p(T): Tdistribution accross the tasks Require: α, β: step size hyperparameters
1: initialize θ
2: While not done do
3: Sample batch of tasks Ti ~p(T)
4: For each base task do
5: Evaluate the prototype by mean of embedded vector concerning K examples
6: Calculate adapted parameters concerning gradient descent: θ
7: Consideration of the gradients with regards to the output layer initialization, i.e. the prototypes, which directly depend on θ
8: end for
9: Calculate the distance between embedding and prototype for each class
10: Closest prototype is assigned as the predicted class for a new class
11. end while

V. Experimental Approach

The meta learning approach is scientifically validated in this section. To answer the questions, we activate the output layer with zeros on each iteration alternatively, as described above, we develop an experimental test-bed that matches the random. Even if the number of classes remains constant, we intend to implement episodic meta-learning. To start with arbitrary predictions, we assess a few few-shot scenarios. This necessitates a learning scenario. The process of training a model's number of inner loop steps to provide correct parameters such that only a few gradient steps or a single classification conclusion, are required. Researchers developed a gradient step, which delivers strong results on a novel job that can combine the advantages of MAML with a prototype seen from a feature learning approach as developing networks to handle this issue.

In particular, initialising internal delineation that is predominantly suitable for many tasks. Prototype networks to initialise our output layer strongly. The internal representation is adaptable to many scenarios, which can be demonstrated that the linear layer with softmax can fine-tune the parameters moderately.

A. Experimental Setup

In all attempts to learn the embedding, the same network of architecture is utilized. The neural network f is agitated of four convolutional blocks: a 22 max pooling layer, a batch normalization layer, a 64-filter 33 convolution, and a ReLU nonlinearity. This architecture was chosen to keep the experiment as close to earlier results while also demonstrating that the implemented technique does not necessitate significant modifications in the setup of the meta-learning algorithms. All the models were trained with the SGD optimizer and a learning rate of 0.05.

B. Algorithms

The episodic training method is preowned to run the meta-learning algorithms, Prototypical Networks and protoMAML.These algorithms were chosen because they are canonical representations of metric-based and optimization-based meta-learning. To assess the improvement obtained by constructing episodes with meta learning, we compare them to uniformly selected baselines.

C. Datasets

Omniglot and SVHN are two well-known and commonly utilized datasets. Omniglot is divided into individual sets of 30 training classes, 10 validation lessons, and 10 testing classes for alphabets. There are 50 alphabets in Omniglot. The Omniglot "background" images are used as training sets. We are going to train our model using a data source of 40,000 few-shot classification assignments. The training set alphabets and testing set's alphabets are employed in completely divergent scenarios. This ensures that the model will attain to categorize characters that were not observed during training when it becomes subjected to the test. The Street View Home Numbers (SVHN) dataset is a real-world image collection for home number detection. Although it has classes 0 to 9 like MNIST does, it is challenging due to its real-world environment and potential for distracting numbers to the left and right. Each image is assigned a class between 0 and 9 that corresponds to the image's prime digit.

VI. Result and Discussion

This section incorporates the experimental findings of few representative algorithms on universal datasets, as well as evaluation and conclusions, to compare the performance of various meta learning methods. The scenarios generated in this collection of studies are permitted to designate the context sets at meta-test, demonstrating that developing better context sets at test time is functional for meta-learning algorithms to acclimate and accomplish new tasks. In most circumstances, metric-based meta-learning algorithms outperform other algorithms. The fundamental cause for this effect is improved space coverage. Sample batch of tasks, along with support and query set for each task, at the beginning of each training step. This involves selecting several support-query set pairings from our sampler in case of few-shot classification. Thus, the present model is fine-tuned on the support set for respective activity. However, the original settings must be remembered for the remaining tasks. The outer loop gradient update along with the subsequent training steps must duplicate our model and only fine-tune the replica. Classification accuracy for 5-way 1-shot and 5-way 5-shot tests are considered as an evaluation criterion. Accuracy is a frequent evaluation metric used by researchers to evaluate the model's performance in few-shot image classification issues. The percentage of samples successfully categorised by the model among all samples is termed classification accuracy.

$$\text{Accuracy} = \frac{\text{Number of correctly classified samples}}{\text{Number of total samples}} \times 100\%$$

The model is trained to ingeminate through a large number of arbitrarily generated few-shot classification tasks, with the fit method updating the model via episodic meta-learning after each task. As previously mentioned, this is known as episodic training. ProtoMAML can outperform ProtoNet, as seen in Figs. 4 and 5. This is because as sample sizes increase, it becomes more important to modify the parameters of the underlying model. ProtoMAML, however, performs worse than ProtoNet for $K = 2$. Since there is a risk of overfitting with more updates, it is likely related to the decision to use 200 inner loop updates. Nevertheless, it is challenging to arrive at any statistically significant conclusions due to the huge standard deviation for $K = 2$.

In this research, the obtained results demonstrate a notable improvement and enhancement compared to existing methodologies, indicating a successful and impactful improvisation in the proposed methodology as mentioned in Table II.

TABLE II. Accuracy of the Proposed Algorithm (5-Way Accuracy)

Algorithm	Dataset	1-shot (%)	5-shot (%)
ProtoNet	Omniglot	99.07±0.16	99.07±0.56
ProtoMAML	SVHN	61.20±1.80	75.50±0.80

Fig. 4. Few shot performance protonet and proto maml on omniglot dataset.

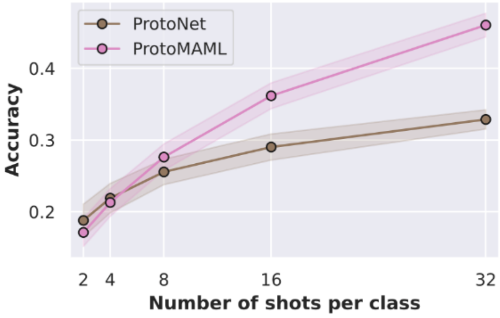

Fig. 5. Few shot performance protonet and proto maml on SVHN dataset.

VII. Conclusion

In this research, we explore the function of episodes in popular few-shot learning methods to understand the reasons behind the impoverished competitiveness of meta-learning approaches about traditional baselines. Under the research, the hyperparameters utilized to sample these events have a significant impact on the performance. By combining the Prototypical Networks and MAML with the closely related episodic analysis, we were able to ignore these hyperparameters, while improvising the few-shot classification accuracy. ProtoNet offers alternative advantages compared to ProtoMAML, specifically a very cheap training and test cost as well as an uncomplicated implementation. The few-shot image classification is a research topic combined with practical applications. The performance is closely related to the scale and amenity of the dataset. Through the utilization of meta-learning principles, the model demonstrates a capacity to rapidly adapt to novel classes with limited examples, showcasing its potential for real-world applications where data scarcity is a common challenge. The episodic learning framework enables the model to generalize effectively from a small number of support examples, exhibiting a form of memory and abstraction akin to human-like learning. This not only enhances the model's ability to handle novel tasks but also contributes to the efficiency and scalability of the few-shot learning process.

Furthermore, current few-shot image classification methods are designed particularly for large datasets. Due to data security restrictions and the complexity of data collection, there are extremely few segments of data research in these special disciplines. As a result, constructing a more appropriate multimodal fusion strategy to further boost the classification is a few-shot image classification research trend.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Syeda Roohi Fatema and Sumana Maradithaya conducted the research and analyzed the data. Syeda Roohi Fatema wrote the paper. All authors have approved the final version.

References

[1] C. Cao, Y. Li, Q. Lv, P. Wang, and Y. Zhang, "Few-shot action recognition with implicit temporal alignment and pair similarity optimization," *Computer Vision and Image Understanding*, vol. 210, 103250, 2021.

[2] G. S. Dhillon, P. Chaudhari, A. Ravichandran, and S. Soatto, "A baseline for few-shot image classification," arXiv preprint, arXiv:1909.02729, 2020.

[3] T. Hospedales, A. Antoniou, P. Micaelli, and A. Storkey, "Meta-learning in neural networks: A survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 9, pp. 5149–5169, 2021.

[4] I. H. Sarker, "Deep learning: A comprehensive overview on techniques, taxonomy, applications and research directions," *SN Computer Science*, vol. 2, no. 6, 2021. https://doi.org/10.1007/s42979-021-00815-1

[5] H.-J. Ye and W.-L. Chao, "How to train your MAML to excel in few-shot classification," arXiv preprint, arXiv:2106.16245, 2021.

[6] D. Sahoo, H. Le, C. Liu, and S. C. Hoi, "Meta-learning with domain adaptation for few-shot learning under domain shift," in *Proc. ICLR*, 2019.

[7] L. F.-F., "A Bayesian approach to unsupervised one-shot learning of object categories," *in Proc. the Ninth IEEE International Conference on Computer Vision*, France, 2003, vol. 2, pp. 1134–1141.

[8] W. Liu, X. Chang, Y. Yan, Y. Yang, and A. G. Hauptmann, "Few-shot text and image classifification via analogical transfer learning," *ACM Trans. Intell. Syst. Technol.*, vol. 9, pp. 1–20, 2018.

[9] L. Gui, Y.-X. Wang, and M. Hebert, "Few-shot hash learning for image retrieval," in *Proc. the 2017 IEEE International Conference on Computer Vision Workshops*, Venice, Italy, 2017, pp. 1228–1237.

[10] S. Qiao, C. Liu, and A. L. Yuille, "Few-shot image recognition by predicting parameters from activations," in *Proc. the IEEE Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, 2018, pp. 7229–7238.

[11] A. Shaban, S. Bansal, Z. Liu, I. Essa, and B. Boots, "One-shot learning for semantic segmentation," arXiv preprint, arXiv:1709.03410, 2017.

[12] B. Liu, X. Yu, A. Yu, P. Zhang, G. Wan, and R. Wang, "Deep few-shot learning for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, pp. 2290–2304, 2018

[13] I. Ashrafifi, M. Mohammad, A. S. Mauree, and K. M. Habibullah, "Attention guided relation network for few-shot image classifification," in *Proc. the 2019 7th International Conference on Computer and Communications Management*, Bangkok, Thailand, 2019, pp. 177–180.

[14] L. Gui, Y.-X. Wang, and M. Hebert, "Few-shot hash learning for image retrieval," in *Proc. the 2017 IEEE International Conference on Computer Vision Workshops*, Venice, Italy, 2017, pp. 1228–1237.

[15] R. Singh, V. Bharti, V. Purohit, A. Kumar, A. K. Singh, and S. K. Singh, "MetaMed: Few-shot medical image classification using gradient-based meta-learning," *Pattern Recognit.*, vol. 120, 108111. 2021.

[16] Y. Yu and N. Bian, "An intrusion detection method using few-shot learning," *IEEE Access*, vol. 8, pp. 49730–49740, 2020.

[17] K. Zhao, X. Jin, and Y. Wang, "Survey on few-shot learning," *J. Softw.* vol. 32, pp. 349–369, 2021.

[18] Y. Wang, Q. Yao, J. Kwok, and L. M. Ni, "Generalizing from a few examples: A survey on few-shot learning," arXiv preprint, arXiv:1904.05046, 2020. doi: http://arxiv.org/abs/1904.05046

[19] J. Snell, K. Swersky, and R. S. Zemel, "Prototypical networks for few-shot learning," arXiv preprint, arXiv:1703.05175, 2012. doi: http://arxiv.org/abs/1703.05175

[20] J. Kim, T. H. Oh, S. Lee, F. Pan, and I. S. Kweon, "Variational prototyping-encoder: One-shot learning with prototypical images," in *Proc. the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 9462–9470.

[21] M. Assran and M. Rabbat, "On the convergence of Nesterov's accelerated gradient method in stochastic settings," in *Proc. ICML*, 2020.

[22] Y. Wang, R. Girshick, M. Hebert, and B. Hariharan, "Low-shot learning from imaginary data," in *Proc. the 2018 IEEE/CVF*

Conference on Computer Vision and Pattern Recognition, Salt Lake City, USA, 2018, pp. 7278–7286.

[23] H. Zhang, J. Zhang, and P. Koniusz, "Few-shot learning via saliency-guided hallucination of samples," in *Proc. the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Long Beach, CA, USA, 2019, pp. 15–20.

[24] Y. Liu, H. Zhang, W. Zhang, G. Lu, Q. Tian, and N. Ling, "Few-shot image classification: Current status and research trends," *Electronics,* no. 11, 1752. 2022. https://doi.org/10.3390/electronics11111752

[25] E. Triantafillou, T. Zhu, and V. Dumoulin *et al.*, "Meta-dataset: A dataset of datasets for learning to learn from few examples," in *Proc. ICLR*, 2020.

[26] B. Ermis, G. Zappella, Cedric, and C. Archambeau, "Towards robust episodic meta-learning," in *Proc. the Thirty-Seventh Conference on Uncertainty in Artificial Intelligence*, 2021, pp. 1342–1351.

[27] B. M. Lake, R. Salakhutdinov, and J. B. Tenenbaum, "The Omniglot challenge: A 3-year progress report," *Current Opinion in Behavioral Sciences*, vol. 29, pp. 97–104, 2019.

[28] Y. Netzer, T. Wang, A. Coates, A. Bissacco, B. Wu, and A. Y. Ng, "Reading digits in natural images with unsupervised feature learning," in *Proc. NIPS Workshop on Deep Learning and Unsupervised Feature Learning*, 2011.

[29] K. Killamsetty, C. Li, C. Zhao, R. Iyer, and F. Chen. "A reweighted meta learning framework for robust few shot learning," arXiv preprint, arXiv:2011.06782, 2020.

[30] O. Vinyals, C. Blundell, T. Lillicrap, K. Kavukcuoglu, and D. Wierstra, "Matching networks for one shot learning," arXiv preprint, arXiv:1606.04080, 2017. doi: http://arxiv.org/abs/1606.04080

[31] M. A. Shedivat, L. Li, E. Xing, and A. T. Walkar, "On data efficiency of meta-learning," in *Proc. International Conference on Artificial Intelligence and Statistics*, 2021, pp.1369–1377.

[32] C. Finn, P. Abbeel, and S. Levine, "Model-agnostic meta-learning for fast adaptation of deep networks," in *Proc. the 34th International Conference on Machine Learning*, Sydney, NSW, Australia, 2017, pp. 1156–1168.

[33] A. A. Rusu, D. Rao, J. Sygnowski, O. Vinyals, R. Pascanu, S. Osindero, and R. Hadsell, "Meta-learning with latent embedding optimization," arXiv preprint, arXiv:1807.05960, 2018.

[34] P. Chikontwe, S. Kim, and S. H. Park, "CAD: Co-adapting discriminative features for improved few-shot classification," in *Proc. the 2022 IEEE/CVF Conference on Computer Vision and PatternRecognition*, New Orleans, USA, 2022, pp. 19–20.

[35] N. Mishra, M. Rohaninejad, X. Chen, and P. Abbeel, "A simple neural attentive meta-learner," arXiv preprint, arXiv:1707.03141, 2018. doi: http://arxiv.org/abs/1707.03141

[36] W.-H. Chu, Y.-J. Li, J.-C. Chang, and Y.-C. F. Wang, "Spot and learn: A maximum-entropy patch sampler for few-shot image classifification," in *Proc. the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Long Beach, CA, USA, 2019; pp. 6244–6253.

[37] F. Sung, Y. Yang, L. Zhang, T. Xiang, P. H. S. Torr, and T. M. Hospedales, "Learning to compare: Relation network for few-shot learning," in *Proc. the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, USA, 2018, pp. 1199–1208.

[38] W. Li, J. Xu, J. Huo, L. Wang, Y. Gao, and J. Luo, "Distribution Consistency Based Covariance Metric Networks for Few-Shot Learning," in *Proc. the AAAI Conference on Artificial Intelligence*, 2019, vol. 33, no. 1, pp. 8642–8649.

[39] J. Zhu, J. J.-Jaccard, A. Singh, I. Welch, A. S. Harith, and S. Camtepe, "A few-shot meta-learning based Siamese neural network using entropy features for ransomware classification," *Computers & Security*, vol. 117, 102691, 2022.

[40] F. Hao, F. He, J. Cheng, L. Wang, J. Cao, and D. Tao, "Collect and select: Semantic alignment metric learning for few-shot learning," in *Proc. the 2019 IEEE/CVF International Conference on Computer Vision*, Seoul, Korea, 2019, pp. 8459–8468.

[41] P. Chikontwe, S. Kim, and S. H. Park, "CAD: Co-adapting discriminative features for improved few-shot classification," in *Proc. the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, New Orleans, LA, USA, June 2022, 19–20.

[42] M. Andrychowicz, M. Denil, S. Gomez, M. W. Hoffman, D. Pfau, T. Schaul, B. Shillingford, and N. de Freitas, "Learning to learn by gradient descent by gradient descent," *Adv. Neural Inf. Process. Syst*, vol. 29, 3981–3989, 2016.

[43] S. Ravi and H. Larochelle, "Optimization as a model for few-shot learning," in *Proc. the 5th International Conference on Learning Representations, ICLR 2017*, Toulon, France, 2017, pp. 24–26.

[44] A. Graves, G. Wayne, and I. Danihelka, "Neural turing machines," arxiv preprint, arxiv:1410.5401, 2014.

[45] A. Santoro, S. Bartunov, M. Botvinick, D. Wierstra, and T. Lillicrap, "Meta-learning with memory-augmented neural networks," in *Proc. the International Conference on Machine Learning*, New York, NY, USA, 2016, pp. 1842–1850.

[46] A. A. Rusu, D. Rao, J. Sygnowski, O. Vinyals, R. Pascanu, S. Osindero, and R. Hadsell, "Meta-learning with latent embedding optimization," arXiv preprint, arXiv:1807.05960, 2018.

[47] M. Garouani, A. Ahmad, M. Bouneffa, M. Hamlich, G. Bourguin, and A. Lewandowski, "Using meta-learning for automated algorithms selection and confifiguration: an experimental framework for industrial big data," *Journal of Big Data*, vol. 9, no. 1, pp. 1–20, 2022.

[48] J. Vanschoren, "Meta-learning: A survey," arXiv preprint, arXiv:1810.03548, 2018.

[49] P. Novais, "Meta-learning and the new challenges of machine learning," *International Journal of Intelligent Systems*, vol. 36, no. 11, pp. 6240–6272, 2021

[50] M. A. Shedivat, L. Li, E. Xing, and A. T. Walkar, "On data efficiency of meta-learning," in *Proc. International Conference on Artificial Intelligence and Statistics*, 2021, pp. 1369–1377.

[51] J. T. Ash, C. Zhang, A. Krishnamurthy, J. Langford, and A. Agarwal, "Deep batch active learning by diverse, uncertain gradient lower bounds," arXiv preprint, arXiv:1906.03671, 2019.