iJRASET

International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 7 Issue: IX Month of publication: September 2019

DOI: http://doi.org/10.22214/ijraset.2019.9135

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.177
Volume 7 Issue IX, Sep 2019- Available at www.ijraset.com

Few-Shot Learning under Domain Shift using Adversarial Domain Adaptation

Akhil Dixit[1], Arush Agarwal[2]
[1]Electronics and Communication, Delhi Technological University, Bawana ,Delhi
[2]Electronics and Communication, Netaji Subhas University of Technology, Dwarka,Delhi

Abstract: Few-shot learning aims to tackle the limitations, of classical techniques, of needing large data sets for learning an effective model. Generally, in few shot learning we have a meta-learning stage (learning to learn) which shall be followed by the few examples to learn upon. Considered as a hallmark of human intelligence, the community has recently witnessed several contributions on this topic, in particular through meta-learning, where a model learns how to learn an effective model for few-shot learning. The main idea is to acquire prior knowledge from a set of training tasks, which is then used to perform (few-shot) test tasks. Most work assumes that both training and test tasks are drawn from the same distribution, and a large amount of labelled data is available in the training tasks. This is a very strong assumption which restricts the usage of meta-learning strategies in the real world where ample training tasks following the same distribution as test tasks may not be available. In this paper we shall see a possible approach for the above problem. In this approach, we use Model Agnostic MetaLearning by Finn et al. to address the problem of few shot learning and an adversarial approach to tackle to the problem of domain shift in the target dataset.

Keywords: Machine Learning, Few-Shot Learning, Meta Learning, Domain-Adaptation

I. INTRODUCTION

Few-Shot Learning aims to learn a prediction model from very limited amount of labelled data (Lake et al, 2015) [1]. Specifically, given a K-shot, N-class data for a classification task, the aim is to learn a multi-class classification model for N− classes, with K-labelled training examples for each class. Here K is usually a small number (e.g. 1, or 5). Considered as one of the hallmarks of human intelligence (Lake et al., 2011), this topic has received considerable interest in recent years (Lake et al., 2015; Koch et al., Vinyals 2015; et al., 2016; Finn et al., 2017). Modern techniques solve this problem through metalearning, using an episodic learning paradigm. The main idea is to use a labelled training dataset to effectively acquire prior knowledge, such that this knowledge can be transferred to novel tasks where few-shot learning is to be performed. Different from traditional transfer learning (Pan et al., 2010; Yosinski et al., 2014), here few-shot tasks are simulated using the labelled training data through episodes, in order to acquire prior knowledge that is specifically tailored for performing few-shot tasks. For example, given a set of labelled training data with a finite label space Y^{train}, the episodic paradigm is used to acquire prior knowledge which is stored in a model. Each episode is generated id from an unknown task distribution τ_{train}. This model is then used to do a novel few shot classification task which is drawn from an unknown task distribution τ_{test}. The test task comprises small amount of labelled data with a finite label space Y_{test} and the sets Y_{train} and Y_{test} are (possibly) mutually exclusive. Using this labelled data, and acquired prior knowledge, the goal is to predict the labels of all unlabelled instances in the test task.

The ability to learn from few examples has been seen in humans. Also, the human brain has shown the ability to adapt to different domains with this quick pace in learning. These attributes are highly desirable in a machine learning model as well. Few-shot learning aims to tackle the limitations, of classical techniques, of needing large data sets for learning an effective model. Generally, in few shot learning we have a meta-learning stage (learning to learn) which shall be followed by the few examples to learn upon. Domain Shift refers to the scenario when the distribution of data during model learning differs from the distribution of data during testing. Under domain shift, we see the inability of classical techniques to adapt to the test data distribution, since they tend to overfit on the training data distribution. Addressing these issues is a significant step in creating more robust Artificial Intelligence models. For instance, as a source of motivation, we can consider the example depicted in the following diagram (Fig. 1). In this approach, we use Model Agnostic Metalearning by Finn et al. to address the problem of few shot learning and an adversarial approach to tackle to the problem of domain shift in the target dataset. Model Agnostic Meta-learning provides us with a general framework for meta-learning stage for few shot learning. Adversarial Domain Discriminator is incorporated in this framework to make the model immune to Domain Shift.

©IJRASET: All Rights are Reserved

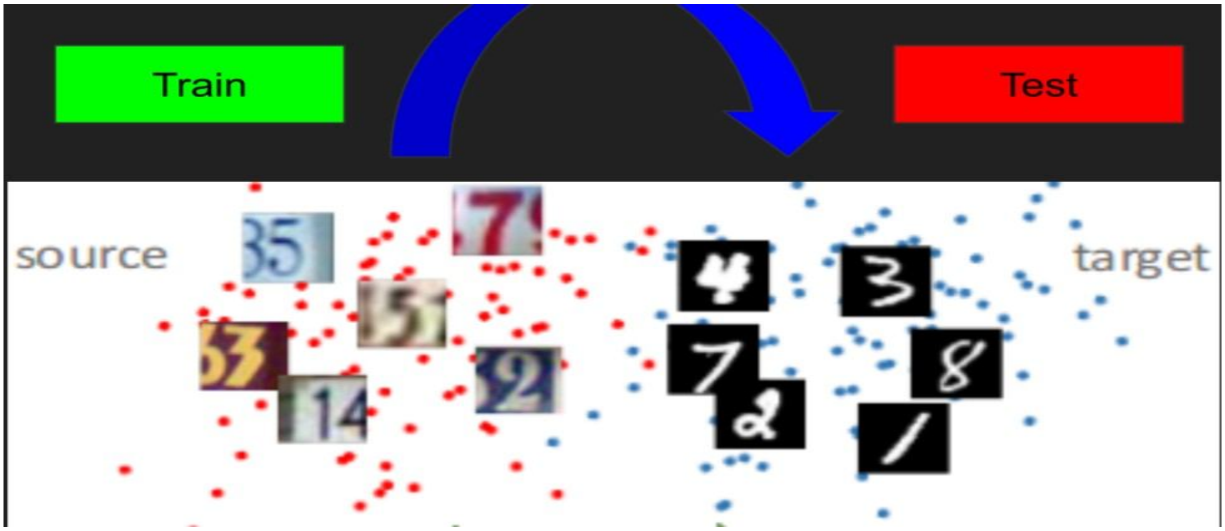


Fig. 1 Domain Adaptation

II. PROBLEM BACKGROUND

A. Few-Shot Learning

The idea of few shot learning arises based on the human ability to learn concepts based on handful of data. The human intelligence is structured to recognize objects from a few examples; such as a child can recognize a lion after seeing only a handful of images; or can learn new skills after only a few trials. But unlike human intelligence, in Artificial Intelligence, the latest state of the art algorithms have evolved that learn complex features using only a vast amount of data.

Few Shot learning aims at solving these issues, i.e. learning based on only a few past experiences and to adapt more quickly. In this work we use the method of Meta-Learning to do Few shot Learning.

B. Meta-Learning

Meta Learning is a field that aims at using metadata to make automatic learning more flexible for different learning problems. It, basically, strives to improve the performance of learning algorithms on diverse problems. Thus, an alternate term “learning to learn” is used for it as it tries to learn the learning algorithm itself with respect to a task. Every existing algorithm is suited for a particular task as it requires the problem to satisfy some prior assumptions/bias. An algorithm may perform better on one domain as compared to others. Meta Learning mitigates this barrier across domains for a learning algorithm so that it could become universal.

C. Adversarial-Learning

Adversarial Learning deals with a subject(generator) learning a realistic model in presence of an adversary(discriminator). The generator generates data from the learned model, hoping it to mimic the data drawn from the true distribution, in order to fool the discriminator. On the other hand, the discriminator tries to identify the counterfeit data. The generator is compelled to learn an implicit distribution as close to the true one so as to confuse the discriminator. Thus, this game between the two leads to an equilibrium where the generator produces such data that the discriminator is unable to distinguish it from real data. At this point, the generator is said to be adversially trained. Recently, adversarial learning methods have proven to be a promising approach to generate samples across various domains which precipitated in their proposed use in unsupervised domain adaptation tasks.

D. Domain-Adaptation

The problem of collecting large amount of labelled data is both expensive and time-consuming. Also, standard datasets are available but every specific task needs to be tested on specific unlabelled data, i.e. there is a domain shift between the train and test data even in these cases.

This problem of Domain Shift; i.e. the distribution of data in the target domain (say, real images) is different than in the source domain (say, synthetic images) which is handled using Domain Adaptation, a field associated with Transfer Learning. It focuses on learning a model based on some source data distribution that performs well for a similar but distinct target data distribution.

III. LITERATURE REVIEW AND RELATED WORK

Several separate works has been done on Few-Shot Learning, Meta Learning and Domain Shift, but none has tried combining all the three together to solve the issue of Domain Adaptation. Hence, we tried reviewing these separate literature's to understand the best approaches in practice for the three and use them as a base for our approach.

A. Model-Agnostic Meta-Learning(MAML)

In this work, the author tries to present a model agnostic framework, i.e. a model that is applicable to a number of different learning problems such as regression, classification and even reinforcement learning. It presents a meta learning, learning to learn, approach which uses gradient descent procedure and can be applied to any model that uses this. The key feature is that it makes the model sensitive so that it can learn about the task in very few gradient descent steps.

Meta Learning is a learning Algorithm that produce the parameter of a model or predictor M that will be able to learn from the few examples it provided. It trains a model with good initial parameters which can give optimal performance on a new task in a few gradient steps that is calculated based on small amount of data from new task, i.e. it is easy to fine tune for a new task because of high sensitivity of the model.

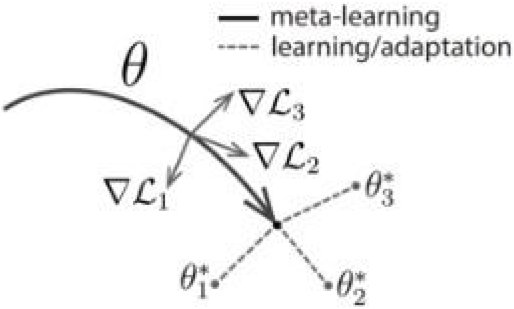
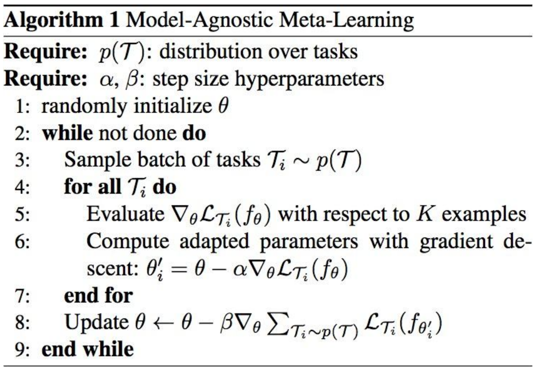


Fig. 2 Visualization of MAML Approach

It uses an episodic (task based) learning framework where it samples a task T_i and initialize the local meta-learner parameter as the global and learns a model M_i with parameters on the training set of that task (support set) and evaluate this model on the test set of that task (query set) to get a loss function L_i which is used to update the meta-learner parameter using gradient descent. This approach leads to generalization of the model and the resultant model after several iterations through meta train data results in a model that can be fine-tuned and can adapt to different tasks easily in a few iterations requiring few examples.



```
Algorithm 1 Model-Agnostic Meta-Learning
Require: $p(\mathcal{T})$: distribution over tasks
Require: $\alpha$, $\beta$: step size hyperparameters
randomly initialize $\theta$
while not done do
    Sample batch of tasks $\mathcal{T}_i \sim p(\mathcal{T})$
    for all $\mathcal{T}_i$ do
        Evaluate $\nabla_\theta \mathcal{L}_{\mathcal{T}_i}(f_\theta)$ with respect to $K$ examples
        Compute adapted parameters with gradient descent: $\theta_i' = \theta - \alpha \nabla_\theta \mathcal{L}_{\mathcal{T}_i}(f_\theta)$
    end for
    Update $\theta \leftarrow \theta - \beta \nabla_\theta \sum_{\mathcal{T}_i \sim p(\mathcal{T})} \mathcal{L}_{\mathcal{T}_i}(f_{\theta_i'})$
end while
```

Fig. 3 Algorithm of MAML Approach

B. Adversarial Discriminative Domain Adaptation (ADDA)

Adversarial Domain Adaptation method tries to minimize the domain discrepancy distance between the source and target data distribution through an adversarial loss function. The method is based on Generative Adversarial Network structure which pita a generator and a discriminator against each other. The generator tries to model the target distribution in such a way that samples from it fools the discriminator, which in turn tries to distinguish between the two distributions. After this process one must not be able to distinguish between the target and source data distributions as both mapped to a common feature space.

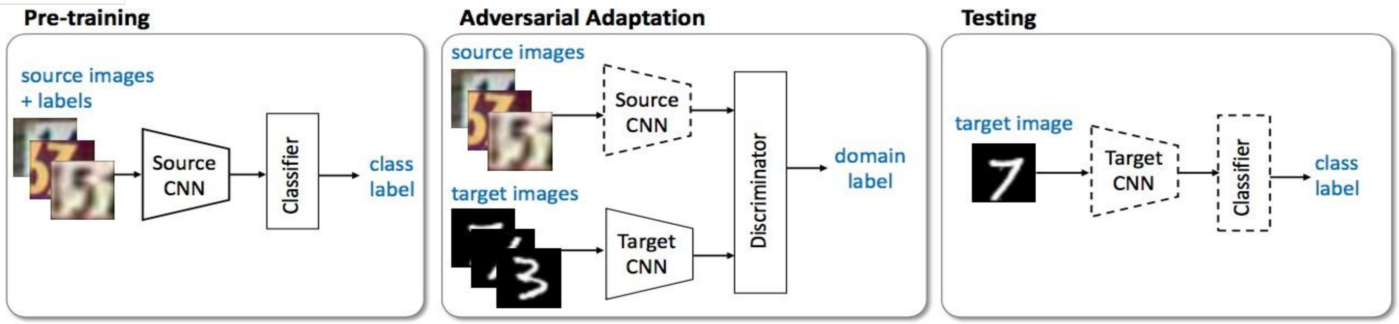


Fig. 4 Graphical view of ADDA ; the dashed line encoder is fixed and solid line encoder are being trained in the respective steps

ADDA first learns a discriminative representation by training the source encoder CNN using labelled source data. Next we implement the GAN on source and target distribution, where we learn the target distribution by keeping the source distribution fixed, such as the discriminator cannot distinguish between the two domains. Finally, the target CNN and source encoder together classifies the target images.

C. Meta-Learning Domain Adaptation(MLDA)

MLDA approaches to do domain shifted few shot learning using Mode Agnostic meta Learning with Adversarial Domain Adaptation. An episodic learning paradigm is used here. But in every episode (task) two events are happening:

1) First a prediction model is being built for the source domain.

2) It learns a feature mapping which establishes domain adaptation between the two domains.

For Few shot Learning Prototypical Networks are used where each instance of an episode is mapped to a d-dimension embedding space $X^{episode}$→ R^d. Now for each class a mean vector embedding is calculated as:

$$c_n = \frac{1}{S_n^{support}} \sum_{(x_i, y_i) \in S_n^{support}} F(x_i)$$

where, Support is the set all training instances of an episode. Next we find the probability distribution of each class calculated using the distance between the embeddings of a query instance and a class prototype and assign the instances of query set accordingly.

Second, we reach source target distribution invariant using ADDA approach, i.e. by using the GAN loss(being done in each episode).

D. Bayesian Model-Agnostic Meta-Learning(BMAML)

BMAML tries to solve the problem of meta-learning with a Bayesian approach. A Bayesian solution for this problem is extremely desired as learning from few examples induces a lot of uncertainty, and a Bayesian model greatly helps in accurately evaluating that uncertainty. Furthermore, it also enables us to have active learning or improve learning by Bayesian ensembling. It combines a non parametric variational inference with gradient based meta learning.

Therefore, Bayesian MAML algorithm is a significant step towards robust meta-learning.

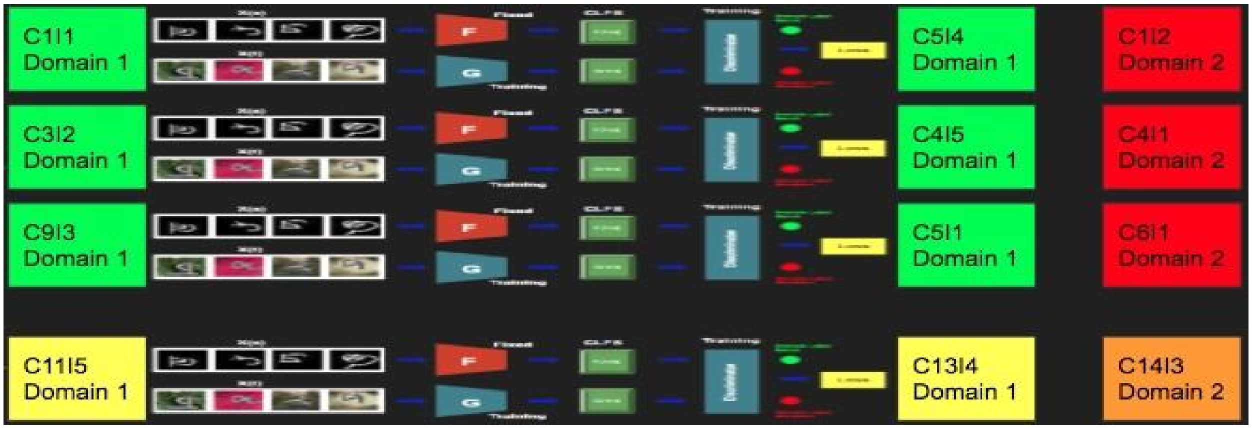


Fig. 5 A simple representation of MLDA Approach. In each episode; 1) Learn a classifier, 2) Do domain adaptation using GAN loss

IV. PROPOSED APPROACH

During episodic meta learning in MLDA, examples from two different domains are fed to the learner. Training examples come from one domain while test examples come from the other. As MLDA performs domain shift (across different domains used for training and testing in each episode) during each episode while doing episodic learning, it becomes computationally too expensive. In order to mitigate this, we decided to extract the part related to domain shift out of the iteration loop of episodes. This would imply that we would have to focus on shifting across domains only once at the beginning of the learning phase. As a result of this segregation, the problem gets bifurcated into two sequential phases:
Phase 1: Adversarial Domain Agnostic Encoding Learning(ADAEL)
Phase 2: Latent Model-Agnostic Meta-Learning(LMAML)

A. Phase 1: Adversarial Domain Agnostic Encoding Learning(ADAEL)

In this phase, we adversially learn an encoder that is domain-agnostic. For this, we define the concept of Common Latent Feature Space(CLFS). Let us train a convolutional neural network(CNN) F on the tuple X, y corresponding to Domain 1 (Omniglot) which was used to sample training images (X corresponds to the set of images and y denotes the corresponding set of class labels; C is the classifier used at the end of F to predict labels). The space of feature of the last layer of F is our CLFS. Our task in this phase is to learn an encoder for domain other than Domain 1 that would transform images belonging to that domain to feature encodings that lie in our CLFS. Once the images from distinct domains are brought down to the same CLFS, the identity of their domain becomes irrelevant. As these latent encodings are used in second phase, it becomes domain agnostic. For obtaining image encodings in CLFS, we learn an encoder corresponding to each source domain (apart from Domain 1). We adversially learn this encoder using F. This process is repeated for all source domains used in the learning phase at the end of which we would have encoders for each domain. In our case, test images are sampled from Domain 2 (Omniglot-M). So, our objective is to learn a neural network G that maps images of Domain 2 to CLFS, thereby, eliminating domain identity. The architecture used for this phase is shown in Figure 6.
Keeping the weights of F fixed with the values that were learnt (as a CNN) during feature encoding process, we learn the weights of G in an adversarial manner by treating G as the generator. The weights of G are randomly initialized. In each epoch (consisting of K pairs of images, one for F and other for G), we pass X(s) and X(t), both from same class (in our case, same character), to F and G respectively. X(s) belongs to Domain 1 and X(t) is of Domain 2. Let X(t) be the image of Domain 2. X(t) is fed to G that learns some feature encoding G(X(t)) which is forwarded to the discriminator. Similarly, F computes latent encoding F (X(s)) from X(s). Both F(X(s)) and G(X(t)) are passed as inputs to the discriminator whose task is to differentiate them on the basis of domain labels. The discriminator is a neural network that predicts the probability(extent) of domain.

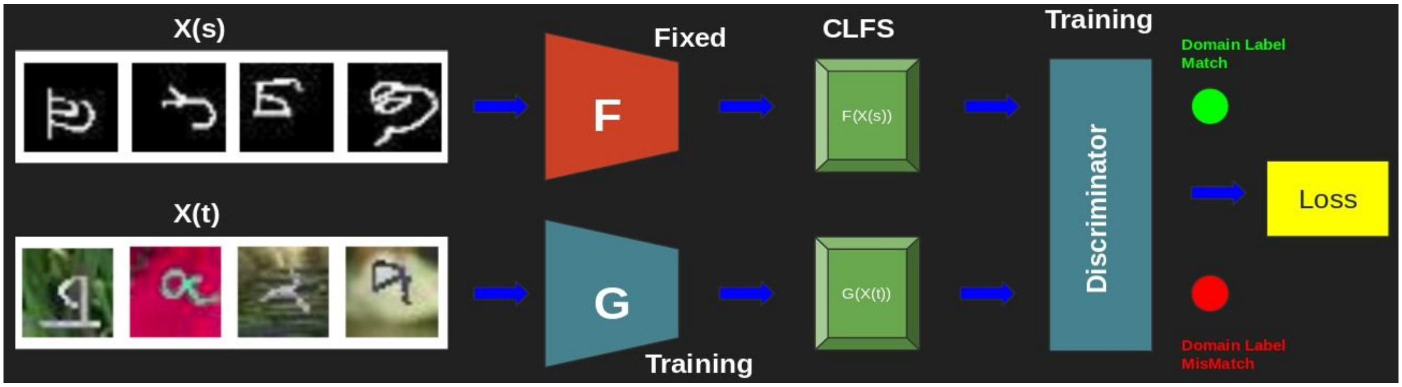


Fig. 6 Learning G adversarially from fixed F

Labels corresponding to input feature F (X(s)) and G(X(t)) being same(D(G(X(t))) is close to 1 if the domain labels are perceived to be same; close to 0, otherwise). the predicted probability D(G(X(t))) is close to 0 due to domains labels mismatch, the discriminator is able to identify the encodings as being of images from different domains. The goal is to learn G such that the feature G(X(t)) produce D(G(X(t))) close to 1 which corresponds to being assigned same domain label as F (X)) when passed through the discriminator. When this happens, the discriminator is no longer able to identify the background domain of the feature. It perceives the images to be coming from the same domain corresponding to the CLFS. We say that the generator(G) is trained when the discriminator appears to predict domain label match with probability 1 At the end of the training phase (Fig. 7), both F (X(s)) and G(X(t)) lie in the CLFS.

The loss function that **F** minimized as a CNN is given in equation 1. Equation 2 and 3 present the adversarial losses that are alternatively maximized and minimized by the discriminator and the generator respectively while training. We make use of this learnt **G** to map images to an element in **CLFS** which is used as an input for LMAML described in the next section.

$$min_{F,C}L_{cls}(X_s,Y_s) = -E(x,y)\sim(X,Y)\sum_{K=1}^{K=y} K \log C(F(xs)) - \S$$

$$min_D L_{advD}(X_s,X_t,F,tt) = -E_{xs}\sim X_s[\log D(F(xs))] - E_{xt}\sim X_t[\log(1 - D(tt(x_t)))] - \S$$

$$min_G L_{advM}(X_s,X_t,D) = E_{xt}\sim X_t[\log(1 - D(tt(x_t)))] - \S$$

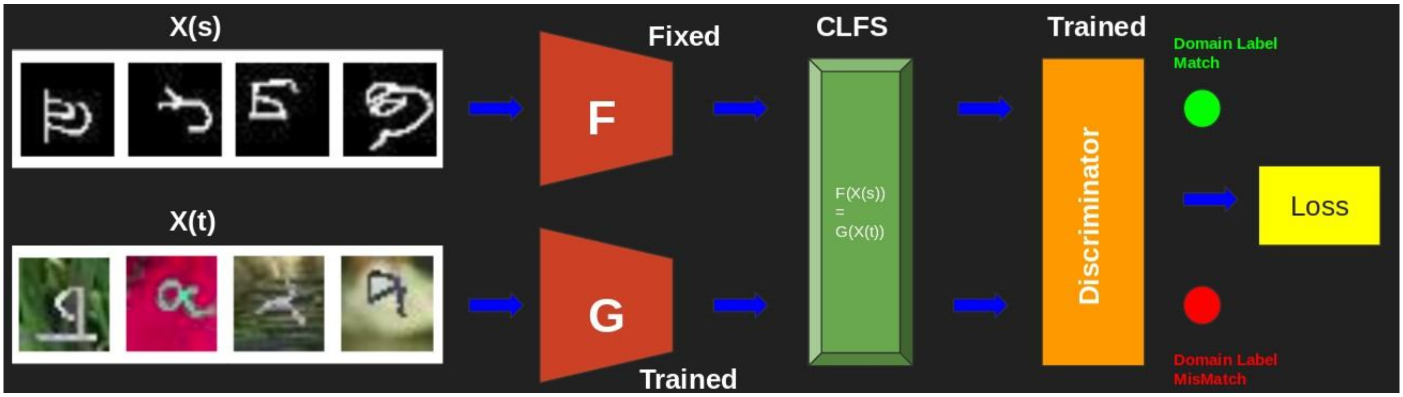


Fig. 7 Trained G and CLFS at the end of training

B. *Phase 2: Latent Model-Agnostic Meta-Learning(LMAML)*

After learning G from the above phase, we pass every image X(t) (of every class) in Domain 2 to G to get corresponding feature encoding G(X(t)). There is a dictionary D maintained that gets indexed by this feature encoding and the value stored at that index is the class label of the corresponding original image X(t). In the same way, all the images X(s) (of every class) of Domain 1 are fed to F to get their respective encoding F (X(s)). As both the encodings belong to the same latent space **CLFS**, the same dictionary D can be used to store the feature F (X(s)) and the class labels of their matching images. This is a one-time process. Now, we have a dataset consisting of feature encodings and their corresponding labels all sharing the same domain which corresponds to CLFS. We can now use a MAML-type architecture that takes feature encodings instead of images for meta-learning. These inputs will be sampled from the dataset D, Using the concept of adversarial domain shift, we have converted the entire dataset of images, affiliated to separate domains, to new images (represented by respective feature encodings) that make up dataset D all associated with the same domain (corresponding to CLFS). Phase 1 eliminates the domain identity of images bringing them to the same platform of CLFS. As depicted the structure of meta-learner is modified to accommodate feature encodings (Fig instead of their corresponding images (Fig. 8) for episodic meta learning. Some notations used in the diagrams are described as follows. Here, CxIy represent yth belonging to class x. n-way, k-shot learning (shown in legend in Fig.10) translates to a training set containing images from n different classes, each class having k examples. F(CxIy) represents the feature in CLFS of belonging to class x.

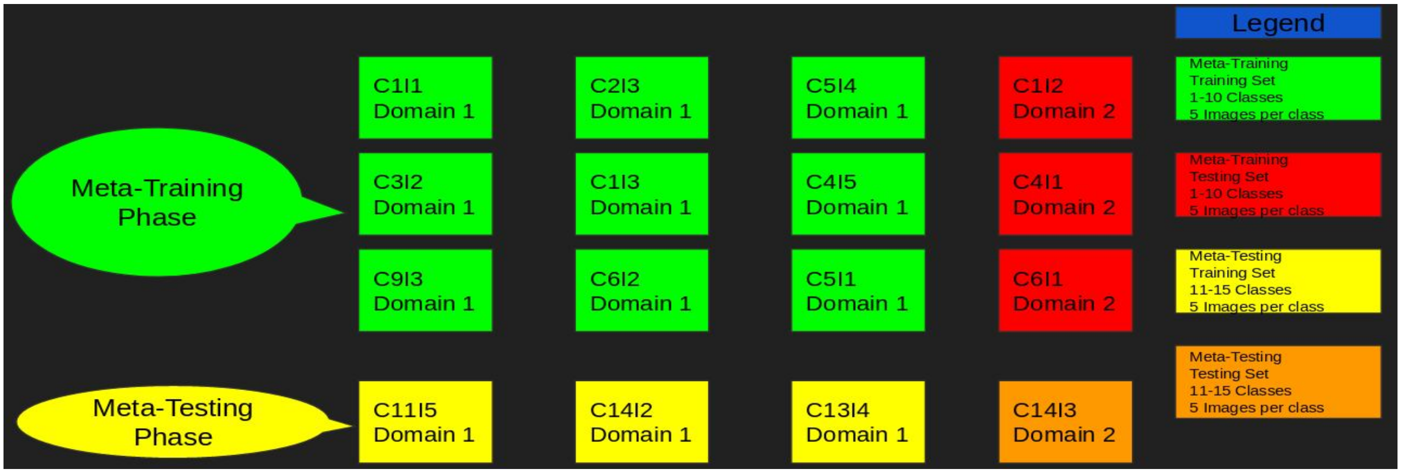


Fig. 8 Episodic learning unoriginal MAML

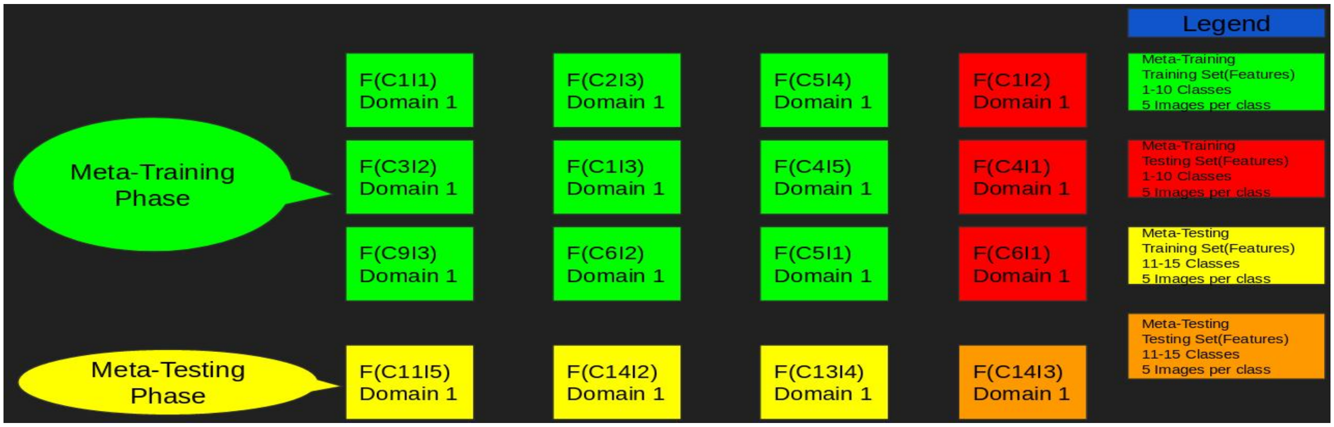


Fig. 9 Episodic learning in LMAML

V. DATASET AND IMPLEMENTATION RESULTS

A. Omniglot

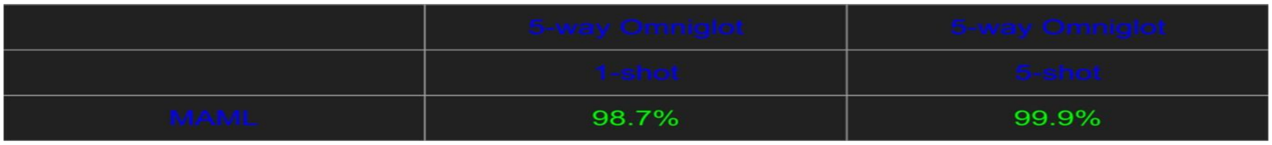
The dataset consists of 1623 handwritten characters from 50 different languages. They have been drawn by 20 different persons (different pen strokes). The dataset has been expanded using 45° rotations and lateral flips. All the images in this dataset have white background.

B. Omniglot-M

This dataset is created by blending images from Omniglot with random backgrounds from BSDS500. This leads to a variation in the domain of this dataset from the domain of Omniglot dataset.

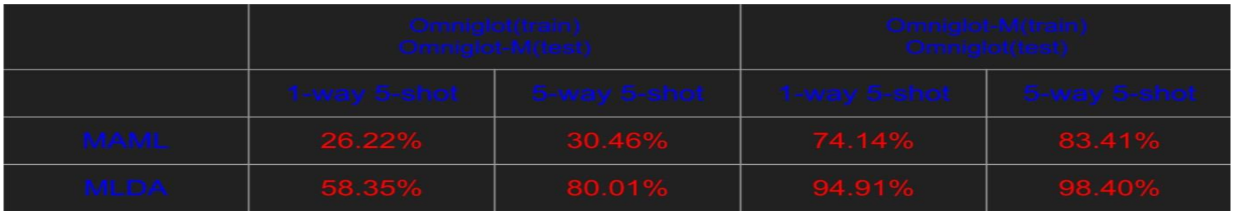
C. Implementation Results

Code of MAML was run on Omniglot dataset. We compared the implementation results of MAML (that we executed) with the results available for the case of domain adaptation in MLDA and MAML [2]. The results were not so good which in turn motivated us to address the issue of domain shift in meta-learning as there was a huge scope of improvement.



	5-way Omniglot	5-way Omniglot
	1-shot	5-shot
MAML	98.7%	99.9%

Fig. 10 Classification accuracy when MAML is run without Domain shift



	Omniglot(train)
	Omniglot-M(test)
			Omniglot-M(train)
			Omniglot(test)	

	1-way 5-shot	5-way 5-shot	1-way 5-shot	5-way 5-shot
MAML	26.22%	30.46%	74.14%	83.41%
MLDA	58.35%	80.01%	94.91%	98.40%

Fig. 11 Comparison of MAML and MLDA under Domain shift

VI. ADVANTAGES AND DISADVANTAGES OF APPROACH

One of the foremost advantage of our approach is that data from multiple domains could be used for the learning process as phase 1 removes their domain identity. Thus, it would appear as if all the images are sampled from the same domain. Our approach is time-efficient as compared to existing Domain Adaptation approaches due to the extraction of time consuming domain shift part from the episode loops. Now, the domain shift has to be performed only at the start of the learning process.

Addition of images corresponding to a specific domain is synonymous to training a neural network affiliated to that domain which converts images associated with the domain into CLFS.

The primary drawback of our approach is the issue of convergence of the adversarial training that we use in phase 1. It would not always converge (if it does) to a global optimum. Finding an ideal neural network architecture (no. of layers and no. of neurons in each layer) of the encoders (e.g. F,G) is very difficult. There is an extra storage requirement for maintaining a dictionary D (discussed in Section 5.2) that maps feature encodings to class labels. As the total no. of images(N) in our entire dataset increases, the size of this dictionary D grows linearly(O(N)) with it.

VII. FUTUTRE SCOPE AND DEVELOPMENT

There are two variations of our novel approach that could be further explored. Both of them have disparate phase 1 while phase 2 remains the same as our novel approach. They are briefly described in the following sections below:

A. Alternating Adversarial Domain Agnostic Encoding Learning(AADAEL)

In case of our approach, we keep F fixed to the values it learned as a CNN while we keep on updating G during adversarial learning. That trains G to reduce an image in the CLFS which is nothing but the feature encoding space corresponding to Domain 1(encoder F). In this alternative, we do not keep F fixed. Instead, we update both F and G alternatively in each adversarial step, i.e., in one step, the game proceeds between F and discriminator and in the next step, the game is played by G and the discriminator (this steps are looped over until convergence). As a result of this alternate trainings, the CLFS would not be the latent space corresponding to the feature encodings of Domain 1(encoder F) instead it would be some new latent space that is intermediate between that of latent spaces of Domain 1(encoder F) and Domain 2(encoder G).

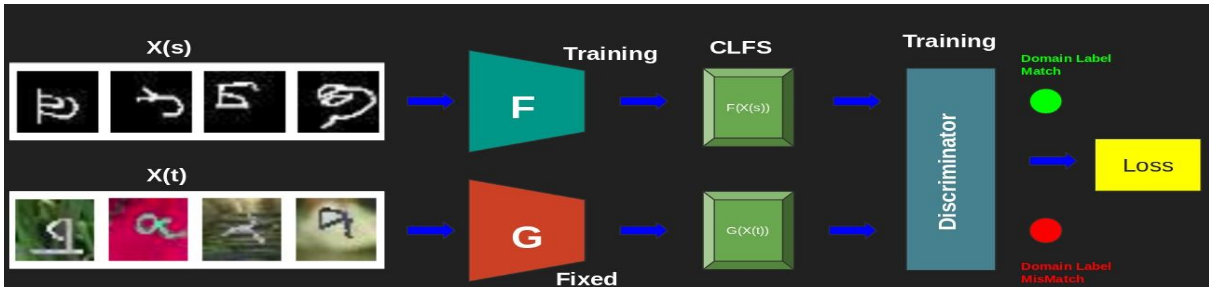


Fig. 11 Alternatively training F and G during adversarial learning

B. Adversarial Encoding-Decoding Learning Approach(AEDLA)

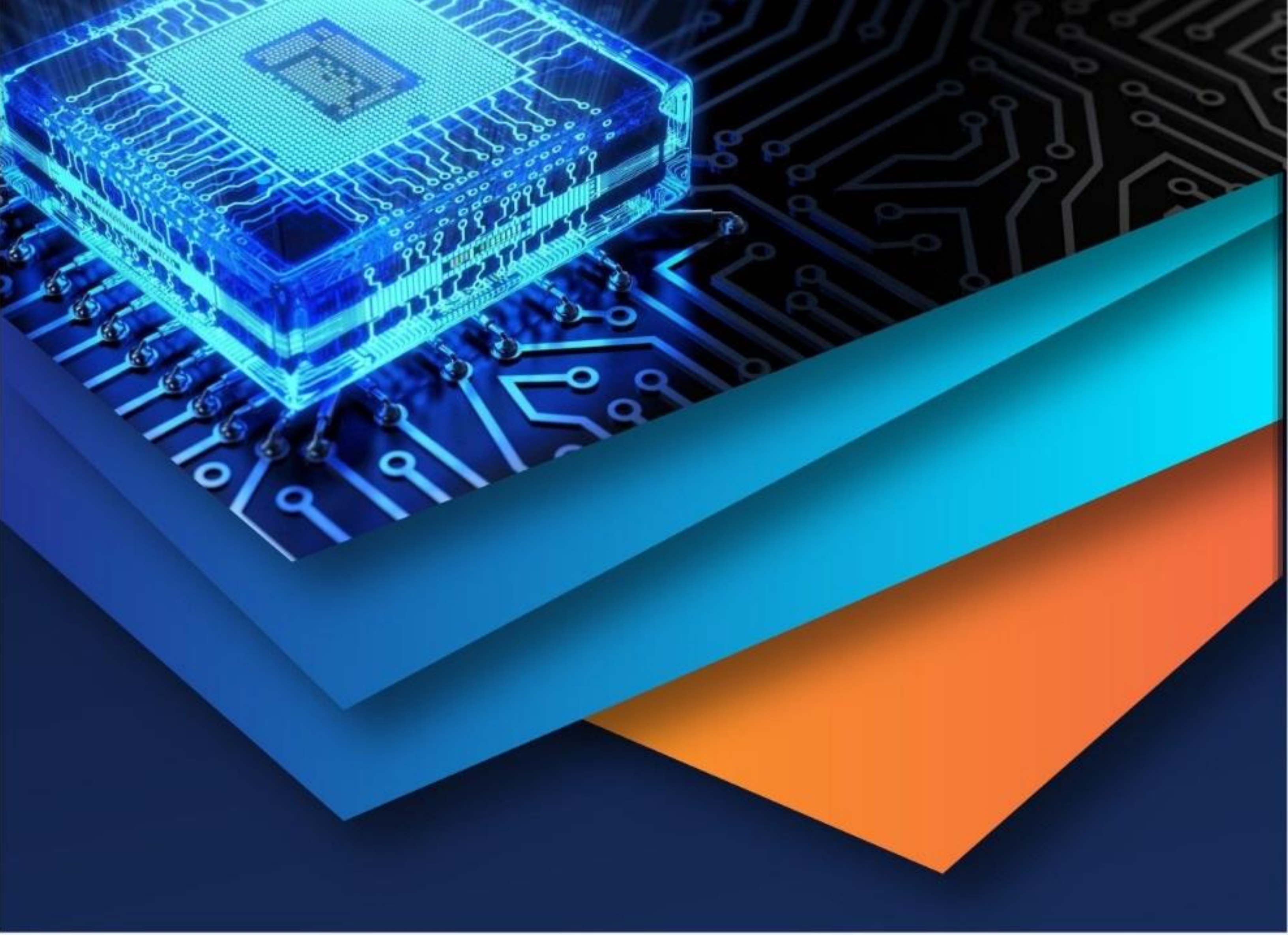
As we strive in our approach to bring the feature encodings of images (from two distinct domains) to a common latent space CLFS, there exists an alternate approach which deals with images instead of feature encodings during meta-learning process in phase 2. So, encoders F and G are learned in the same way but an extra neural network(decoder) F $^{-1}$ that does an inverse mapping (from feature encoding space of Domain 1 to an image in Domain 1; basically inverse of encoder F) is also maintained. An image(X(t)) in Domain 2 is converted via G to an element(G(X(t))) in CLFS (feature space of Domain 1). Now, decoder F $^{-1}$ is applied to this feature encoding G(X(t)) to get the corresponding image representation in Domain 1 for an image in Domain 2. Basically, we are transforming an image from one domain to another. In the second phase, we could now use original MAML for doing meta-learning as the images have been converted so as to belong to Domain 1.

VIII. CONCLUSION

To conclude, we have studied about different approaches to tackle to problem of few-shot learning under Domain shift, and finally proposed a solution which comprises of two phases: Adversarial Domain Adaptation and Latent MAML. With these approaches multiple Domains can be interlinked and be brought to a common Latent Feature space. The proposed approach avoids episodic learning of the Adversarial Domain Adaptation which should result in a computationally efficient model, but it might not converge and it is also storage inefficient as we are required to keep a mapping dictionary.

REFERENCES

[1] Sergey Levine Chelsea Finn, Pieter Abbeel. Model-agnostic meta-learning for fast adaptation of deep networks. International Conference on Machine Learning, 2017.
[2] Doyen Sahoo et al. Meta-learning with domain adaptation for few-shot learning under domain shift, 2018.
[3] Eric Tzeng et al. Adversarial discriminative domain adaptation. IEEE Conference on Computer Vision and Pattern Recognition, 2017.
[4] Taesup Kim et al. Bayesian model-agnostic meta-learning. NIPS 2018, 2018.



10.22214/IJRASET

INDEX COPERNICUS
45.98

IMPACT FACTOR:
7.129

TOGETHER WE REACH THE GOAL
IMPACT FACTOR:
7.429

INTERNATIONAL JOURNAL
FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)