DOMAIN-AGNOSTIC META-LEARNING FOR CROSS-DOMAIN FEW-SHOT CLASSIFICATION

Wei-Yu Lee[⋆] *Jheng-Yu Wang*[⋆] *Yu-Chiang Frank Wang*[†]

[⋆]Technology and Research Corporate Division, MOXA Inc., Taipei, Taiwan
[†]Department of Electrical Engineering, National Taiwan University, Taipei, Taiwan

ABSTRACT

Few-shot classification requires one to classify instances of novel classes, given only a few examples of each class. Although promising meta-learning methods have been proposed recently, there is no guarantee that existing solutions would generalize to novel classes from an *unseen* domain. In this paper, we tackle the challenging task of *cross-domain few-shot classification* and propose *Domain-Agnostic Meta-Learning (DAML)* algorithm. Our DAML, serving as an optimization strategy, learns to adapt the model to novel classes in both seen and unseen domains by data sampled from multiple domains with desirable task settings. In our experiments, we apply DAML on three popular metric-based models under cross-domain settings. Experiments on several benchmarks (mini-ImageNet, CUB, Cars, Places, Plantae and META-DATASET) show that DAML significantly improves the generalization ability of learning models, and addresses cross-domain few-shot classification with promising results.

Index Terms— Meta-learning, Few-shot classification

1. INTRODUCTION

Few-shot classification [1, 2] aims to classify instances of novel classes, given only a few examples of each class. Recent meta-learning algorithms have been developed to solve few-shot classification problems. Among those approaches, metric-learning based models [3, 4, 5, 6] learn to project instances into an embedding space and then classify them by the associated similarities. On the other hand, optimization based models, such as model-agnostic meta-learning (MAML), learn the good initial parameters that are able to effectively adapt to a novel task [7, 8]. While these methods have achieved promising results, they typically assume a shared data distribution between training and test (novel) classes. Hence, when the novel classes come from an unseen domain, there is no guarantee that existing few-shot learning (FSL) solutions would still generalize [9].

Recently, [9, 10, 11] address the task of *cross-domain* few-shot learning (CD-FSL). In their settings, the test classes

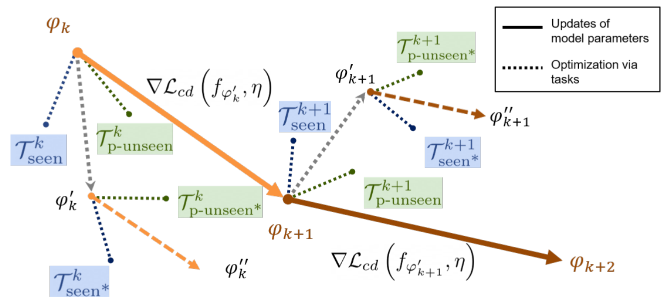


Fig. 1. Illustration of DAML. DAML jointly observes tasks $\mathcal{T}_{\text{seen}}$ and $\mathcal{T}_{\text{p-unseen}}$ from seen and pseudo-unseen domains. By minimizing $\mathcal{L}_{\mathcal{T}_{\text{seen}}}$ and $\mathcal{L}_{\mathcal{T}_{\text{p-unseen}}}$, the model learns domain-agnostic initial parameters φ, which would adapt to novel classes in unseen domains during meta-testing.

are from data domain different from that of training classes. One research branch in solving CD-FSL aims at learning a robust feature backbone from multiple source domains during meta-training, and expects generalization to unseen domains during meta-testing [11, 12, 13, 14]. Another research branch performs adaptation or finetuning on the unseen domain during meta-testing [15, 9, 10]; however, their meta-training is performed on a single source domain and does not guarantee the domain generalization ability.

In this paper, we propose a domain-agnostic meta-learning (DAML) algorithm for CD-FSL. Our DAML can be viewed as an optimization-based meta-learning algorithm. By jointly learning from data across multiple seen domains with few-shot adaptation, novel classes in unseen domains can be recognized accordingly. To be more specific, our DAML concurrently renders generalization and adaptation abilities to a learner by guiding the updated parameters towards domain-agnostic gradient direction. Once the learning process is complete, the preferable domain-agnostic initial parameters are obtained for the learning models to adapt to few-shot tasks in the unseen domain.

It is worth noting that, the core idea of this paper lies in the proposed novel learning scheme, *but not* particular network architecture or loss function. We uniquely introduce domain generalization into existing FSL solutions. In our exper-

[⋆]Indicates equal contribution.

iments, we apply our DAML on three metric-learning based models, and conduct experiments on 6 benchmarks (mini-ImageNet [16], CUB [17], Cars [18], Places [19], Plantae [20], and META-DATASET [21]). We show that our algorithm clearly improves the generalization ability of existing FSL models in recognizing novel classes of unseen domains.

The contributions of this paper are highlighted as follows:

- We present a novel optimization-based meta-learning strategy, allowing the leaning models to efficiently and effectively adapt to novel classes in unseen domains.
- Instead of proposing novel network architectures or losses, DAML presents a unique training scheme of learning from few-shot task sampled across multiple seen domains with domain generalization guarantees.
- Applicable to existing FSL solutions like metric-based methods , we provide insights to the task of CD-FSL and confirm that a proper meta-learning scheme would be a more preferable solution than a large training dataset.

2. PROPOSED METHOD

2.1. Optimization-Based Few-Shot Classification

A standard few-shot classification task $\mathcal{T}$ is comprised of support set S and query set Q. A learner is required to classify instances in $Q = \{(\mathcal{X}_q, \mathcal{Y}_q)\}$ given a small amount of labeled data in $S = \{(\mathcal{X}_s, \mathcal{Y}_s)\}$, where Q and S have the same label space. The learning process often consists of two stages: *meta-training* and *meta-testing*. During meta-training, a dataset $\mathcal{D}$ with C classes is provided to train a properly initialized learner for adaptation. During meta-testing, the learner would adapt to novel classes.

However, traditional FSL models assume training and test classes share the same data domain, which might not be practical. To handle novel classes in an unseen domain, we present a novel learning strategy called domain-agnostic meta-learning (DAML). Inspired by MAML [7], we aim at learning a set of domain-agnostic initial parameters φ for adaptation. Our DAML updates φ using $S \in \mathcal{T}$ sampled across multiple seen domains to acquire the domain generalization and adaptation ability and prevent the learning biases towards any single domain.

2.2. Domain-Agnostic Meta-Learning (DAML)

Serving as an optimization-based approach, DAML learns a set of *domain-agnostic* initial parameters that can adapt to tasks from arbitrary unseen domains. The core idea of DAML is that the updated learning model are required to perform well on tasks drawn not only from the seen domain but also from the *pseudo*-unseen domain (i.e., cross domain). By introducing data of seen and pseudo-unseen domains during meta-training, our DAML pursues a set of domain-agnostic initial parameters; this is realized by simultaneously updating the model through a few gradient steps on different domains. Thus, the derived model parameters would be domain-agnostic, and could adapt to novel classes in unseen domains.

One can easily apply DAML on existing metric-based FSL models. Note that the metric-based models typically project instances into an embedding space, and then perform recognition by a meta-learned metric function. Specifically, the prediction is made by: $\hat{\mathcal{Y}}_q = M(\mathcal{Y}_s, E(\mathcal{X}_s), E(\mathcal{X}_q))$, where E and M are feature extractor and metric function, respectively. In order to train the metric-based models that can adapt to unseen domain, we design the training scheme as follows. As shown in Fig. 1, the optimization process is based on tasks drawn from the seen and pseudo-unseen domains, i.e., $\mathcal{T}_{\text{seen}} \in \mathcal{D}_{\text{seen}}$ and $\mathcal{T}_{\text{p-unseen}} \in \mathcal{D}_{\text{p-unseen}}$, rather than the standard S and Q (drawn from the same domain) as used in MAML. Specifically, at each iteration, we first update the model parameters using $\mathcal{T}_{\text{seen}}$ and $\mathcal{T}_{\text{p-unseen}}$: $\varphi_i' = \varphi - \gamma \nabla_\varphi \mathcal{L}_{cd}\left(f_\varphi, \eta\right)$, where γ is the learning rate and $\mathcal{L}_{cd}$ is cross-domain loss:

$$\mathcal{L}_{cd}\left(f_\varphi, \eta\right) = (1-\eta)\mathcal{L}_{\mathcal{T}_{\text{seen}}}\left(f_\varphi\right) + \eta \mathcal{L}_{\mathcal{T}_{\text{p-unseen}}}\left(f_\varphi\right), \tag{1}$$

where η strikes the balance between seen and pseudo-unseen domains. Since the tasks drawn from different domains might exhibit various characteristics, leading to various degrees of difficulty, we do not use a fixed value of η. Instead, we have η updated based on the observed difficulties between seen and pseudo-unseen data. Specifically,

$$\eta\left(f_\varphi\right) = \mathcal{L}_{\mathcal{T}_{\text{p-unseen}}}\left(f_\varphi\right) / \left[\mathcal{L}_{\mathcal{T}_{\text{seen}}}\left(f_\varphi\right) + \mathcal{L}_{\mathcal{T}_{\text{p-unseen}}}\left(f_\varphi\right)\right]. \tag{2}$$

Thus, when the tasks of pseudo-unseen domain are more difficult than tasks of seen domain, the learning objective will give higher weights on pseudo-unseen tasks and vice versa.

Now, the updated φ_i' is expected to not only address $\mathcal{T}_{\text{seen}}$ but also generalize to $\mathcal{T}_{\text{p-unseen}}$. For learning a proper initial model parameters φ, the model parameter φ would be updated as: $\varphi \leftarrow \varphi - \alpha \nabla_\varphi \mathcal{L}_{cd}\left(f_{\varphi'}, \eta\right)$, where α indicates the learning rate. Note that there could be multiple pseudo-unseen domains $\{\mathcal{D}^1_{\text{p-unseen}}, ..., \mathcal{D}^d_{\text{p-unseen}}\}$. We randomly sample one domain and draw new tasks at each optimization step.

During meta-testing, we replace $\mathcal{T}_{\text{p-unseen}}$ with $\mathcal{T}_{\text{unseen}}$ and update φ to adapt to the unseen domain. Note that the loss of $\mathcal{T}_{\text{seen}}$ is also considered to make sure that the model would generalize to both domains.

3. EXPERIMENTS

3.1. Experimental Settings and Implementation

We compare our proposed DAML with MAML and three metric-learning based FSL models: MatchingNet [4], RelationNet [5], and GNN [6]. For fair comparisons, we follow

Table 1. CD-FSL on CUB, Cars, Places, and Plantae. All methods are trained on mini-ImageNet together with three other datasets (except the one listed in each column as the unseen domain for testing). The left/right digits in each entry denote the accuracy on seen (mini-ImageNet) and unseen domains, respectively. Note that [13] did not report results on seen domains.

Backbone		CUB	Cars	Places	Plantae
MAML	[7]	69.72 / 54.25 %	69.54 / 42.48 %	67.52 / 62.39 %	70.26 / 44.57 %
MatchingNet	[4]	72.98 / 51.92 %	73.17 / 39.87 %	70.58 / 61.82 %	72.84 / 47.29 %
	[11]	73.51 / 61.41 %	73.23 / 43.08 %	70.82 / 64.99 %	72.76 / 48.32 %
	Ours	**74.58 / 62.48 %**	**74.77 / 50.43 %**	**72.47 / 66.04 %**	**73.08 / 51.49 %**
RelationNet	[5]	74.65 / 62.13 %	76.09 / 40.64 %	73.85 / 64.34 %	75.91 / 46.29 %
	[11]	76.12 / 64.99 %	**76.80** / 43.44 %	**75.07** / 67.35 %	**76.58** / 50.39 %
	[13]	- / 62.71 %	- / 41.05 %	- / 66.08 %	- / 48.78 %
	Ours	**76.57 / 67.05 %**	75.44 / **51.35 %**	74.77 / **68.97** %	75.75 / **55.97 %**
GNN	[6]	83.67 / 69.26 %	84.37 / 48.91 %	80.87 / 72.59 %	83.60 / 58.36 %
	[11]	84.81 / 73.11 %	84.10 / 49.88 %	82.76 / 77.05 %	85.37 / 58.84 %
	Ours	**85.46 / 75.05 %**	**85.45 / 56.85 %**	**83.45 / 77.27** %	**85.88 / 61.06 %**

Table 2. CD-FSL on META-DATASET. Note that our DAML is trained on multiple intermediate-size source domains (mini-ImageNet together with Cars, Places and Plantae), while others methods are trained on ImageNet. All models are evaluated on unseen datasets as listed in each column. Avg. rank denotes the average performance ranking across datasets for each method.

Model	Training set size/class	Omniglot	Aircraft	Texture	Draw	Flower	Traffic Sign	COCO	Avg. rank
k-NN	∼100 GB / 712	37.07	46.81	66.36	32.06	83.10	44.59	30.38	6.3
Finetune	∼100 GB / 712	60.85	**68.69**	69.05	42.60	85.51	66.79	34.86	3.3
RelationNet [5]	∼100 GB / 712	45.35	40.73	52.97	43.30	68.76	33.67	29.15	6.6
Relation-DAML	**∼51.5 GB / 445**	61.51	47.36	**69.12**	51.27	84.04	**68.56**	48.32	**2.7**
ProtoNet [3]	∼100 GB / 712	59.98	53.10	66.56	48.96	85.27	47.12	41.00	4.1
Proto-MAML [21]	∼100 GB / 712	63.37	55.95	66.49	51.52	**87.15**	48.83	43.74	2.9
Proto-DAML	**∼51.5 GB / 445**	**72.68**	59.67	67.10	**58.51**	83.47	67.50	**51.15**	**2.1**

the experimental settings of [11] to consider the leave-one-out scheme for CD-FSL. That is, we use the mini-ImageNet [16] as seen domain, and three of the four datasets (i.e., CUB [17], Cars [18], Places [19], and Plantae [20]) as pseudo-unseen domains during meta-training. The remaining one dataset serves as unseen domain during meta-testing. In addition, we test our DAML on META-DATASET [21], a newly proposed large-scale challenging dataset taking into account the class imbalance issue within a few-shot task.

We utilize the source code released by [11] to conduct experiments. As [11], we adopt the ResNet-10 [22] as our backbone feature extractor and pre-train it on the training set of mini-ImageNet. During meta-testing, we split S into disjoint sub-support and sub-query set to form a $\mathcal{T}$ for the metric-based models to adapt to unseen domain. For instance, in a N-way K-shot experiment, we randomly separate s shot from S to make a task (i.e., N-way $(K-s)$-shot) for adaptation, and then evaluate the updated models on Q.

3.2. Evaluation

3.2.1. Quantitative evaluation

We compare our DAML with four different previous methods on the 5-way 5-shot tasks on unseen domain. In Table 1, each column indicates the data domain unseen during meta-training. We train our model on the seen domain (mini-ImageNet) and the rest of three pseudo-unseen domains. Table 1 shows that our DAML performs favorably against the state-of-the-arts on both seen and unseen domain, suggesting that the models learn a set of proper initial parameters that can effectively adapt to tasks from arbitrary domains.

To show the generalization capability of DAML, we also evaluate our DAML on the newly proposed benchmark META-DATASET [21] and make comparisons to the reported models trained on ImageNet [23]. We test all models on the same unseen datasets of META-DATASET (Fungi [24] is already unavailable). Table 2 shows that our DAML outperforms other metric-learning based methods and significantly improves generalization. It is worth noting that the size of our training data is only about half of the ImageNet. Moreover, our training data contains only a subset of labels in ImageNet, and thus *no* auxiliary label information are used during meta-training.

In addition, we compare our DAML to the newly proposed CD-FSL algorithm [12] with the same 5-way 5-shot classification on unseen domains. Again, we use the same 4 datasets to train our DAML, while other competitors use larger training datasets (ImageNet [23], Aircraft [26], CIFAR100 [27], Texture [28], Traffic sign [29], Omniglot [2], UCF101 [30] and Flower [31], except for the testing

Table 3. Additional comparisons to recent CD-FSL methods. Following [12], all models are evaluated on 600 5-way 5-shot tasks with 10 query images. Note that our DAML utilizes only multiple intermediate-size source domains with a subset of label information in the training data used in other models.

Model	Training set size	Omniglot	Aircraft	Texture	Flower
FEAT [25]	>100 GB	89.90	34.90	52.00	76.70
ProtoNet [3]	>100 GB	92.70	36.05	52.00	79.60
Proto-MAML [21]	>100 GB	83.00	33.10	45.30	69.40
DoA [12]	>100 GB	95.10	39.00	56.50	84.10
DoA-Ch [12]	>100 GB	95.20	39.20	55.00	84.30
Proto-DAML	**~51.5 GB**	**95.60**	**41.11**	**57.22**	**85.88**

dataset of interest). As listed in Table 3, even with a significantly smaller training dataset and less label information, our DAML still outperforms other state-of-the-art FSL models in unseen domains. Since we *do* not observe auxiliary label information during training, the above comparisons supports our proposed learning strategy.

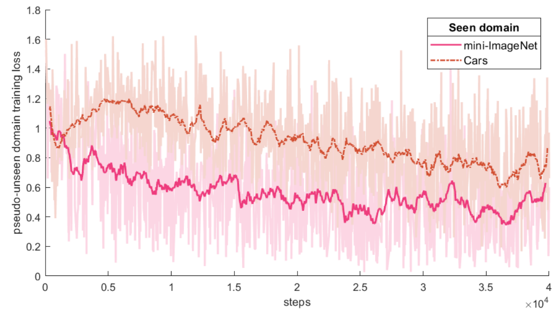
3.2.2. Ablation studies

We analyze how data diversity in the seen domain affects generalization ability of the learning models. To inspect the effect of source datasets on the learning models, we take RelationNet as example backbone and train our DAML on different seen datasets, i.e. mini-ImageNet vs. Cars, with the same pseudo-unseen domains. According to Table 4(a) and Fig. 2(a), we observe that the model trained on mini-ImageNet significantly outperforms the model trained on Cars. We attribute the improved generalization to the greater data variability of mini-ImageNet, favoring a learner to obtain the domain-agnostic initial parameters and to adapt to unseen domain. By contrast, Cars is a relatively fine-grained dataset, thus presenting constrained learning space for a learner.

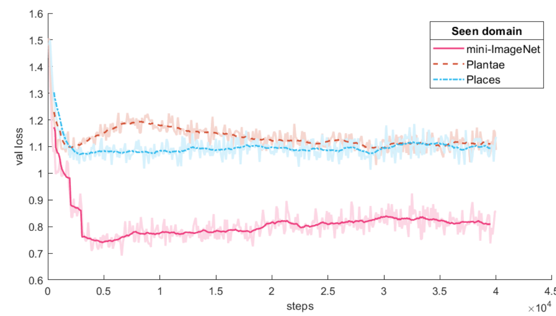
In addition, we use the same source datasets (mini-ImageNet, Places, and Plantae) and swap the seen and pseudo-unseen domains to validate the importance of the seen domain in our algorithm. From Table 4(b) and Fig. 2(b), we again confirm that the use of mini-ImageNet as seen domain would be more desirable for CD-FSL.

4. CONCLUSION

We proposed a domain-agnostic meta-learning (DAML) algorithm for CD-FSL. Instead of designing new network architectures or losses, DAML presents an optimization-based learning scheme, aiming to initialize the model with domain-agnostic parameters, so that the model can efficiently and effectively adapt to novel tasks from both seen and unseen domains. DAML is realized across multiple source-domains with intermediate training set sizes, and exhibits promising few-shot classification and generalization abilities. Our extensive experiments confirm that DAML is applicable to existing metric-based models with improved generalization.



(a) Pseudo-unseen domain training loss



(b) Validation loss on mini-ImageNet

Fig. 2. Loss comparisons of (a) $\mathcal{L}_{\mathcal{T}_{\text{p-unseen}}}$ and (b) validation loss on mini-ImageNet with different seen datasets. Note that (a) verifies the use of mini-ImageNet as seen domain for better generalization and (b) further confirms that the model trained with seen mini-ImageNet shows better performance.

Table 4. Comparisons of using different seen datasets in DAML. (a) mini-ImageNet vs. Cars (with shared pseudo-unseen datasets), (b) different seen and pseudo-unseen pairs with the same source datasets.

5-way 5-shot		CUB
Seen	Pseudo-unseen	
mini-ImageNet	Places, Plantae	**66.86 %**
Cars	Places, Plantae	61.56 %

(a)

5-way 5-shot		CUB
Seen	Pseudo-unseen	
mini-ImageNet	Places, Plantae	**66.86 %**
Plantae	mini-ImageNet, Places	62.73 %
Places	mini-ImageNet, Plantae	61.88 %

(b)

5. REFERENCES

[1] Brenden Lake, Ruslan Salakhutdinov, Jason Gross, and Joshua Tenenbaum, "One shot learning of simple visual concepts," in *Proceedings of the annual meeting of the cognitive science society*, 2011.

[2] Brenden M Lake, Ruslan Salakhutdinov, and Joshua B Tenenbaum, "Human-level concept learning through probabilistic program induction," *Science*, 2015.

[3] Jake Snell, Kevin Swersky, and Richard Zemel, "Prototypical networks for few-shot learning," in *Advances in neural information processing systems*, 2017.

[4] Oriol Vinyals, Charles Blundell, Timothy Lillicrap, Daan Wierstra, et al., "Matching networks for one shot learning," in *Advances in neural information processing systems*, 2016.

[5] Flood Sung, Yongxin Yang, Li Zhang, Tao Xiang, Philip HS Torr, and Timothy M Hospedales, "Learning to compare: Relation network for few-shot learning," in *CVPR*, 2018.

[6] Victor Garcia and Joan Bruna, "Few-shot learning with graph neural networks," in *ICLR*, 2018.

[7] Chelsea Finn, Pieter Abbeel, and Sergey Levine, "Model-agnostic meta-learning for fast adaptation of deep networks," in *ICML*, 2017.

[8] Alex Nichol, Joshua Achiam, and John Schulman, "On first-order meta-learning algorithms," *arXiv preprint arXiv:1803.02999*, 2018.

[9] Wei-Yu Chen, Yen-Cheng Liu, Zsolt Kira, Yu-Chiang Frank Wang, and Jia-Bin Huang, "A closer look at few-shot classification," in *ICLR*, 2019.

[10] Yunhui Guo, Noel C Codella, Leonid Karlinsky, James V Codella, John R Smith, Kate Saenko, Tajana Rosing, and Rogerio Feris, "A broader study of cross-domain few-shot learning," in *ECCV*, 2020.

[11] Hung-Yu Tseng, Hsin-Ying Lee, Jia-Bin Huang, and Ming-Hsuan Yang, "Cross-domain few-shot classification via learned feature-wise transformation," in *ICLR*, 2020.

[12] Yongseok Choi, Junyoung Park, Subin Yi, and Dong-Yeon Cho, "Domain-agnostic few-shot classification by learning disparate modulators," *arXiv preprint arXiv:1909.04999*, 2019.

[13] Jiamei Sun, Sebastian Lapuschkin, Wojciech Samek, Yunqing Zhao, Ngai-Man Cheung, and Alexander Binder, "Explanation-guided training for cross-domain few-shot classification," *arXiv preprint arXiv:2007.08790*, 2020.

[14] Yixiong Zou, Shanghang Zhang, José MF Moura, JianPeng Yu, and Yonghong Tian, "Revisiting mid-level patterns for distant-domain few-shot recognition," *arXiv preprint arXiv:2008.03128*, 2020.

[15] John Cai and Sheng Mei Shen, "Cross-domain few-shot learning with meta fine-tuning," *arXiv preprint arXiv:2005.10544*, 2020.

[16] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *CVPR*, 2009.

[17] P. Welinder, S. Branson, T. Mita, C. Wah, F. Schroff, S. Belongie, and P. Perona, "Caltech-UCSD Birds 200," Tech. Rep. CNS-TR-2010-001, California Institute of Technology, 2010.

[18] Jonathan Krause, Michael Stark, Jia Deng, and Li Fei-Fei, "3d object representations for fine-grained categorization," in *ICCV Workshops*, 2013.

[19] Bolei Zhou, Agata Lapedriza, Aditya Khosla, Aude Oliva, and Antonio Torralba, "Places: A 10 million image database for scene recognition," *IEEE transactions on pattern analysis and machine intelligence*, 2017.

[20] Grant Van Horn, Oisin Mac Aodha, Yang Song, Yin Cui, Chen Sun, Alex Shepard, Hartwig Adam, Pietro Perona, and Serge Belongie, "The inaturalist species classification and detection dataset," in *CVPR*, 2018.

[21] Eleni Triantafillou, Tyler Zhu, Vincent Dumoulin, Pascal Lamblin, Utku Evci, Kelvin Xu, Ross Goroshin, Carles Gelada, Kevin Swersky, Pierre-Antoine Manzagol, and Hugo Larochelle, "Meta-dataset: A dataset of datasets for learning to learn from few examples," in *ICLR*, 2020.

[22] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, "Deep residual learning for image recognition," in *CVPR*, 2016.

[23] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, et al., "Imagenet large scale visual recognition challenge," *International journal of computer vision*, 2015.

[24] B Schroeder and Y Cui, "Fgvcx fungi classification challenge 2018," 2018.

[25] Han-Jia Ye, Hexiang Hu, De-Chuan Zhan, and Fei Sha, "Learning embedding adaptation for few-shot learning," *arXiv preprint arXiv:1812.03664*, 2018.

[26] Subhransu Maji, Esa Rahtu, Juho Kannala, Matthew Blaschko, and Andrea Vedaldi, "Fine-grained visual classification of aircraft," *arXiv preprint arXiv:1306.5151*, 2013.

[27] Alex Krizhevsky, Geoffrey Hinton, et al., "Learning multiple layers of features from tiny images," 2009.

[28] Mircea Cimpoi, Subhransu Maji, Iasonas Kokkinos, Sammy Mohamed, and Andrea Vedaldi, "Describing textures in the wild," in *CVPR*, 2014.

[29] Johannes Stallkamp, Marc Schlipsing, Jan Salmen, and Christian Igel, "Man vs. computer: Benchmarking machine learning algorithms for traffic sign recognition," *Neural networks*, 2012.

[30] Khurram Soomro, Amir Roshan Zamir, and Mubarak Shah, "Ucf101: A dataset of 101 human actions classes from videos in the wild," *arXiv preprint arXiv:1212.0402*, 2012.

[31] Maria-Elena Nilsback and Andrew Zisserman, "Automated flower classification over a large number of classes," in *2008 Sixth Indian Conference on Computer Vision, Graphics & Image Processing*, 2008.