Proceedings of Machine Learning Research 176:1–8, 2022 NeurIPS 2021 Competition and Demonstration Track

Supplementary Material for Lessons learned from the NeurIPS 2021 MetaDL challenge: Backbone fine-tuning without episodic meta-learning dominates for few-shot learning image classification

Adrian El Baz* eb.adrian8@gmail.com
Ihsan Ullah ihsan2131@gmail.com
Edesio Alcobaça e.alcobaca@gmail.com
André C. P. L. F. Carvalho andre@icmc.usp.br
Hong Chen h-chen20@mails.tsinghua.edu.cn
Fabio Ferreira ferreira@cs.uni-freiburg.de
Henry Gouk henry.gouk@ed.ac.uk
Chaoyu Guan guancy19@mails.tsinghua.edu.cn
Isabelle Guyon guyon@chalearn.org
Timothy Hospedales t.hospedales@ed.ac.uk
Shell Hu shell.hu@samsung.com
Mike Huisman m.huisman@liacs.leidenuniv.nl
Frank Hutter fh@cs.uni-freiburg.de
Zhengying Liu zhengying.liu@inria.fr
Felix Mohr felix.mohr@unisabana.edu.co
Ekrem Öztürk ozturk@informatik.uni-freiburg.de
Jan N. van Rijn j.n.van.rijn@liacs.leidenuniv.nl
Haozhe Sun haozhe.sun@universite-paris-saclay.fr
Xin Wang xin_wang@tsinghua.edu.cn
Wenwu Zhu wwzhu@tsinghua.edu.cn

Editors: Douwe Kiela, Marco Ciccone, Barbara Caputo

Appendix A. Datasets

Table 1 summarizes the datasets. Figure 1 shows sample images per dataset.

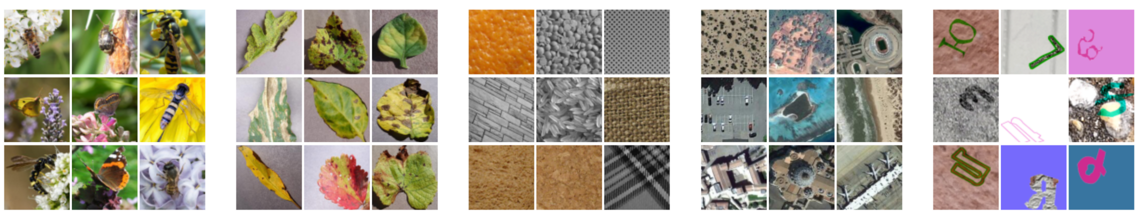
Appendix B. Related Work

While deep neural networks are capable of achieving performance superior to humans on various tasks (Krizhevsky et al., 2012; Mnih et al., 2015; He et al., 2015), they are notorious for requiring large amounts of data and processing power, restricting their success to domains where such resources are available. Humans, on the other hand, are more efficient learners as they can effectively draw on their prior knowledge and learning experience

* The two first authors are principal challenge organizer and dataset preparer; the other authors are in alphabetical order.

Table 1: MetaDL 2021 datasets summary

Domain	Dataset	Competition Phase	Categories	Images	Source
Ecology	Plankton	Feedback	91	3,640	Heidi M. Sosik (2015)
	Insects	Final	114	4,560	Serret et al. (2019)
Bio-medicine	Multiderma	Feedback	51	2,040	Sun et al. (2016)
	Plant Village	Final	37	1,480	Hughes and Salathé (2015) Geetharamani and Arun Pandian (2019)
Manufacturing	Texture DTD	Feedback	47	1,880	Cimpoi et al. (2014)
	Textures	Final	64	2,560	Fritz et al. (2004) Mallikarjuna et al. (2006) Kylberg (2011) Lazebnik et al. (2005)
Remote sensing	Mini RSICB	Feedback	45	1,800	Li et al. (2020)
	Mini RESISC	Final	45	1,800	Cheng et al. (2017)
OCR	OmniPrint-MD-mix	Feedback	706	28,240	Sun et al. (2021)
	OmniPrint-MD-5-bis	Final	706	28,240	Sun et al. (2021)



(a) Insects (b) Plant Village (c) Textures (d) Mini RESISC (e) OmniPrint-MD-mix

(f) Plankton (g) Multiderma (h) Texture DTD (i) Mini RSICB (j) OmniPrint-MD-5-bis

Figure 1: NeurIPS 2021 meta-learning challenge datasets sample images

(Jankowski et al., 2011). Improving the learning efficiency of deep neural networks is being extensively studied within the area of few-shot learning (Wang et al., 2020; Bendre et al., 2020; Lu et al., 2020). We discuss the two main paradigms that are used to address this.

Meta-learning Meta-learning (Naik and Mammone, 1992; Thrun, 1998; Schmidhuber, 1987; Brazdil et al., 2022) aims to learn, from previous learning experiences, how to learn (Vanschoren, 2018; Hospedales et al., 2021; Huisman et al., 2021b). Matching networks aim to learn a good embedding such that a nearest-neighbour classifier can be effective

(Vinyals et al., 2016). Prototypical networks build on this technique by comparing inputs to class prototypes instead of instances (Snell et al., 2017). Relation networks replace the distance metric with a neural network (Sung et al., 2018). Model-based approaches, such as MANNs (Santoro et al., 2016), Meta-Nets (Munkhdalai and Yu, 2017), TURTLE (Huisman et al., 2021a) and SNAIL (Mishra et al., 2018), embed a given dataset into an activation state and use this state to make predictions for new data points. Optimization-based approaches use optimization, such as gradient descent, to learn new tasks. One of the most popular techniques from this approach is MAML (Finn et al., 2017), which aims to learn good initialization hyperparameters from which new tasks can be learned in a few gradient update steps. This work has been the inspiration for many follow-up works, such as Meta-SGD (Li et al., 2017), which also learns suitable learning rates, Reptile (Nichol et al., 2018), which is a first-order variant on MAML, and LEO (Rusu et al., 2019), whose goal is to learn the initialization hyperparameters in a low-dimensional latent space.

Transfer learning Transfer learning (Weiss et al., 2016; Tan et al., 2018; Pan and Yang, 2009) aims to transfer knowledge from a source task or domain (or set thereof), where a large amount of data may be present, to a target domain, where data may be sparse. One popular transfer learning approach in deep learning is to pre-train a network on a given source domain (e.g., ImageNet (Krizhevsky et al., 2012)), followed by fine-tuning parts (such as only the output layer) of the network on the target domain (Huang et al., 2013; Oquab et al., 2014). In this case, the knowledge transfer is parameter-based. Many other forms of transfer also exist, such as mapping-based, instance-based, and adversarial-based transfer (Tan et al., 2018).

Recent works illustrate that simple pre-training and fine-tuning can outperform more complicated meta-learning techniques (Chen et al., 2019; Tian et al., 2020) which raises the question of whether a good embedding is enough for achieving good few-shot learning performance. However, this could also indicate that the few-shot benchmarks such as Mini-ImageNet (Vinyals et al., 2016; Ravi and Larochelle, 2017), TieredImageNet (Ren et al., 2018), and CUB (Wah et al., 2011) are not challenging enough because test examples come from the same dataset as the one used for training.

Related competitions and benchmarks This competition is part of an established series of competitions such as the AutoML competition series (Guyon et al., 2019), the AutoDL competition series (Liu et al., 2021), the AutoCV competition series and the MetaDL competition series (El Baz et al., 2021). This competition is an extension to our previous hosted competition in the MetaDL series (El Baz et al., 2021). It challenges participants with a more challenging set of datasets, that were specifically designed for this challenge.

The Open Algorithm Selection Competition (OASC) is a competition that is closely related to meta-learning (Lindauer et al., 2017, 2019). In that competition, for a given dataset, an appropriate algorithm needs to be selected. While several machine learning datasets are present in the competition, it focuses also on algorithm selection beyond machine learning (e.g., MIP and SAT).

Meta-dataset is another notable benchmark used for few-shot learning. It is a collection of 10 datasets that are commonly used in few-shot learning (Triantafillou et al., 2020). Our competition setup with various datasets is partly inspired by this initiative.

References

Nihar Bendre, Hugo Terashima Marín, and Peyman Najafirad. Learning from few samples: A survey. *arXiv preprint arXiv:2007.15484*, 2020.

Pavel Brazdil, Jan N. van Rijn, Carlos Soares, and Joaquin Vanschoren. *Metalearning: Applications to Automated Machine Learning and Data Mining*. Springer, 2nd edition, 2022.

Wei-Yu Chen, Yen-Cheng Liu, Zsolt Kira, Yu-Chiang Frank Wang, and Jia-Bin Huang. A closer look at few-shot classification. In *International Conference on Learning Representations*, ICLR'19, 2019.

Gong Cheng, Junwei Han, and Xiaoqiang Lu. Remote sensing image scene classification: Benchmark and state of the art. *Proceedings of the IEEE*, 105(10):1865–1883, 2017.

Mircea Cimpoi, Subhransu Maji, Iasonas Kokkinos, Sammy Mohamed, and Andrea Vedaldi. Describing textures in the wild. In *2014 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2014, Columbus, OH, USA, June 23-28, 2014*, pages 3606–3613. IEEE Computer Society, 2014.

Adrian El Baz, Isabelle Guyon, Zhengying Liu, Jan N. van Rijn, Sebastien Treguer, and Joaquin Vanschoren. Advances in MetaDL: AAAI 2021 challenge and workshop. In *AAAI Workshop on Meta-Learning and MetaDL Challenge*, volume 140 of *Proceedings of Machine Learning Research*, pages 1–16. PMLR, 2021.

Chelsea Finn, Pieter Abbeel, and Sergey Levine. Model-agnostic meta-learning for fast adaptation of deep networks. In *Proceedings of the 34th International Conference on Machine Learning-Volume 70*, pages 1126–1135. JMLR. org, 2017.

Mario Fritz, E. Hayman, B. Caputo, and J. Eklundh. The kth-tips database, 2004. URL https://www.csc.kth.se/cvap/databases/kth-tips/index.html.

G. Geetharamani and J. Arun Pandian. Identification of plant leaf diseases using a nine-layer deep convolutional neural network. *Computers & Electrical Engineering*, 76:323–338, 2019.

Isabelle Guyon, Lisheng Sun-Hosoya, Marc Boullé, Hugo Jair Escalante, Sergio Escalera, Zhengying Liu, Damir Jajetic, Bisakha Ray, Mehreen Saeed, Michèle Sebag, Alexander Statnikov, Wei-Wei Tu, and Evelyne Viegas. *Analysis of the AutoML Challenge Series 2015–2018*, pages 177–219. Springer International Publishing, Cham, 2019.

Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In *Proceedings of the IEEE international conference on computer vision*, pages 1026–1034, 2015.

Emily F. Brownlee Heidi M. Sosik, Emily E. Peacock. Annotated plankton images - data set for developing and evaluating classification methods., 2015. URL https://hdl.handle.net/10.1575/1912/7341.

Timothy Hospedales, Antreas Antoniou, Paul Micaelli, and Amos Storkey. Meta-learning in neural networks: A survey. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2021.

Jui-Ting Huang, Jinyu Li, Dong Yu, Li Deng, and Yifan Gong. Cross-language knowledge transfer using multilingual deep neural network with shared hidden layers. In *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 7304–7308. IEEE, 2013.

David P. Hughes and Marcel Salathé. An open access repository of images on plant health to enable the development of mobile disease diagnostics through machine learning and crowdsourcing. *CoRR*, abs/1511.08060, 2015.

Mike Huisman, Aske Plaat, and Jan N van Rijn. Stateless neural meta-learning using second-order gradients. *arXiv preprint arXiv:2104.10527*, 2021a.

Mike Huisman, Jan N Van Rijn, and Aske Plaat. A survey of deep meta-learning. *Artificial Intelligence Review*, 54(6):4483–4541, 2021b.

Norbert Jankowski, Wlodzislaw Duch, and Krzysztof Grabczewski, editors. *Meta-Learning in Computational Intelligence*, volume 358 of *Studies in Computational Intelligence*. Springer, 2011.

Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. ImageNet Classification with Deep Convolutional Neural Networks. In *Advances in Neural Information Processing Systems 25*, NIPS'12, pages 1097–1105, 2012.

Gustaf Kylberg. The kylberg texture dataset v. 1.0. Technical Report 35, Centre for Image Analysis, Swedish University of Agricultural Sciences and Uppsala University, Uppsala, Sweden, 2011.

Svetlana Lazebnik, Cordelia Schmid, and Jean Ponce. A sparse texture representation using local affine regions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(8):1265–1278, 2005.

Haifeng Li, Xin Dou, Chao Tao, Zhixiang Wu, Jie Chen, Jian Peng, Min Deng, and Ling Zhao. Rsi-cb: A large-scale remote sensing image classification benchmark using crowdsourced data. *Sensors*, 20(6):1594, 2020.

Zhenguo Li, Fengwei Zhou, Fei Chen, and Hang Li. Meta-SGD: Learning to Learn Quickly for Few-Shot Learning. *arXiv preprint arXiv:1707.09835*, 2017.

Marius Lindauer, Jan N van Rijn, and Lars Kotthoff. Open algorithm selection challenge 2017: Setup and scenarios. In *Open Algorithm Selection Challenge 2017*, volume 79 of *PMLR*, pages 1–7, 2017.

Marius Lindauer, Jan N van Rijn, and Lars Kotthoff. The algorithm selection competitions 2015 and 2017. *Artificial Intelligence*, 272:86–100, 2019.

Zhengying Liu, Adrien Pavao, Zhen Xu, Sergio Escalera, Fabio Ferreira, Isabelle Guyon, Sirui Hong, Frank Hutter, Rongrong Ji, Julio C. S. Jacques Junior, Ge Li, Marius Lindauer, Zhipeng Luo, Meysam Madadi, Thomas Nierhoff, Kangning Niu, Chunguang Pan, Danny Stoll, Sebastien Treguer, Jin Wang, Peng Wang, Chenglin Wu, Youcheng Xiong, Arbër Zela, and Yang Zhang. Winning solutions and post-challenge analyses of the chalearn autodl challenge 2019. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(9):3108–3125, 2021.

Jiang Lu, Pinghua Gong, Jieping Ye, and Changshui Zhang. Learning from very few samples: A survey. *arXiv preprint arXiv:2009.02653*, 2020.

P. Mallikarjuna, Alireza Tavakoli Targhi, Mario Fritz, E. Hayman, B. Caputo, and J. Eklundh. The kth-tips 2 database, 2006. URL https://www.csc.kth.se/cvap/databases/kth-tips/index.html.

Nikhil Mishra, Mostafa Rohaninejad, Xi Chen, and Pieter Abbeel. A simple neural attentive meta-learner. In *International Conference on Learning Representations*, ICLR'18, 2018.

Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A Rusu, Joel Veness, Marc G Bellemare, Alex Graves, Martin Riedmiller, Andreas K Fidjeland, Georg Ostrovski, et al. Human-level control through deep reinforcement learning. *Nature*, 518(7540):529–533, 2015.

Tsendsuren Munkhdalai and Hong Yu. Meta networks. In *Proceedings of the 34th International Conference on Machine Learning*, ICML'17, pages 2554–2563. JLMR.org, 2017.

Devang K Naik and Richard J Mammone. Meta-neural networks that learn by learning. In *International Joint Conference on Neural Networks*, pages 437–442. IEEE, 1992.

Alex Nichol, Joshua Achiam, and John Schulman. On First-Order Meta-Learning Algorithms. *arXiv preprint arXiv:1803.02999*, 2018.

Maxime Oquab, Leon Bottou, Ivan Laptev, and Josef Sivic. Learning and transferring mid-level image representations using convolutional neural networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1717–1724, 2014.

Sinno Jialin Pan and Qiang Yang. A Survey on Transfer Learning. *IEEE Transactions on knowledge and data engineering*, 22(10):1345–1359, 2009.

Sachin Ravi and Hugo Larochelle. Optimization as a Model for Few-Shot Learning. In *International Conference on Learning Representations*, ICLR'17, 2017.

Mengye Ren, Sachin Ravi, Eleni Triantafillou, Jake Snell, Kevin Swersky, Josh B. Tenenbaum, Hugo Larochelle, and Richard S. Zemel. Meta-learning for semi-supervised few-shot classification. In *International Conference on Learning Representations*, ICLR'18, 2018.

Andrei A. Rusu, Dushyant Rao, Jakub Sygnowski, Oriol Vinyals, Razvan Pascanu, Simon Osindero, and Raia Hadsell. Meta-learning with latent embedding optimization. In *International Conference on Learning Representations*, ICLR'19, 2019.

Adam Santoro, Sergey Bartunov, Matthew Botvinick, Daan Wierstra, and Timothy Lillicrap. Meta-learning with Memory-augmented Neural Networks. In *Proceedings of the 33rd International Conference on International Conference on Machine Learning*, ICML'16, pages 1842–1850, 2016.

Jürgen Schmidhuber. Evolutionary principles in self-referential learning, or on learning how to learn: the meta-meta-... hook. Master's thesis, Technische Universität München, 1987.

Hortense Serret, Nicolas Deguines, Yikweon Jang, Grégoire Lois, and Romain Julliard. Data quality and participant engagement in citizen science: comparing two approaches for monitoring pollinators in france and south korea. *Citizen Science: Theory and Practice*, 4(1):22, 2019.

Jake Snell, Kevin Swersky, and Richard Zemel. Prototypical Networks for Few-shot Learning. In *Advances in Neural Information Processing Systems 30*, NIPS'17, pages 4077–4087. Curran Associates Inc., 2017.

Haozhe Sun, Wei-Wei Tu, and Isabelle M Guyon. Omniprint: A configurable printed character synthesizer. In *Proceedings of the Neural Information Processing Systems Track on Datasets and Benchmarks*, NIPS'21, 2021.

Xiaoxiao Sun, Jufeng Yang, Ming Sun, and Kai Wang. A benchmark for automatic visual classification of clinical skin disease images. In *European Conference on Computer Vision*, pages 206–222. Springer, 2016.

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 *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1199–1208. IEEE, 2018.

Chuanqi Tan, Fuchun Sun, Tao Kong, Wenchang Zhang, Chao Yang, and Chunfang Liu. A survey on deep transfer learning. In *International conference on artificial neural networks*, pages 270–279. Springer, 2018.

Sebastian Thrun. Lifelong Learning Algorithms. In *Learning to learn*, pages 181–209. Springer, 1998.

Yonglong Tian, Yue Wang, Dilip Krishnan, Joshua B Tenenbaum, and Phillip Isola. Rethinking few-shot image classification: a good embedding is all you need? *arXiv preprint arXiv:2003.11539*, 2020.

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 *International Conference on Learning Representations*, 2020.

Joaquin Vanschoren. Meta-Learning: A Survey. *arXiv preprint arXiv:1810.03548*, 2018.

Oriol Vinyals, Charles Blundell, Timothy Lillicrap, Koray Kavukcuoglu, and Daan Wierstra. Matching Networks for One Shot Learning. In *Advances in Neural Information Processing Systems 29*, NIPS'16, pages 3637–3645, 2016.

C. Wah, S. Branson, P. Welinder, P. Perona, and S. Belongie. The Caltech-UCSD Birds-200-2011 Dataset. Technical Report CNS-TR-2011-001, California Institute of Technology, 2011.

Yaqing Wang, Quanming Yao, James T Kwok, and Lionel M Ni. Generalizing from a few examples: A survey on few-shot learning. *ACM computing surveys (csur)*, 53(3):1–34, 2020.

Karl Weiss, Taghi M Khoshgoftaar, and DingDing Wang. A survey of transfer learning. *Journal of Big data*, 3(1):1–40, 2016.