AdarGCN: Adaptive Aggregation GCN for Few-Shot Learning
– Supplementary Material –

Jianhong Zhang[1*] Manli Zhang[1*] Zhiwu Lu[1,2†] Tao Xiang[3]
[1]Gaoling School of Artificial Intelligence, Renmin University of China
[2]Beijing Key Laboratory of Big Data Management and Analysis Methods
[3]University of Surrey, United Kingdom
jianhong@ruc.edu.cn manlizhang@ruc.edu.cn luzhiwu@ruc.edu.cn

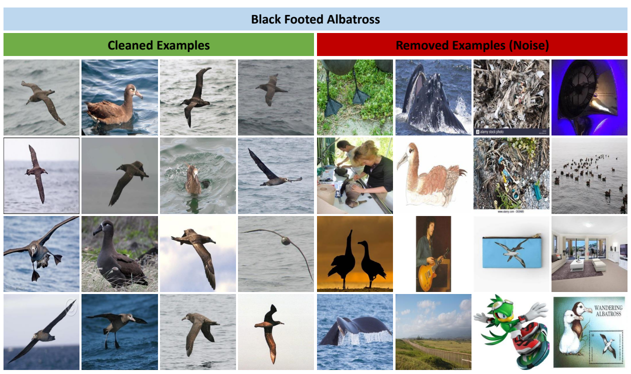

Figure 1. Illustration of the cleaned examples and the removed examples (i.e. noise) for the source class 'Black Footed Albatross'.

In this document, we provide more supporting results to show the effectiveness of our AdarGCN under both the new SSFSL and conventional FSL settings. Firstly, we illustrate the cleaned examples and the removed examples (i.e. noise) obtained by our AdarGCN-based label denoising (LDN) method under the new SSFSL setting. Secondly, we present the weight distributions of outlying examples obtained by our AdarGCN-based FSL method under the conventional FSL setting. These qualitative results (shown in Sections 1 and 2) suggest that our AdarGCN can effectively deal with both noisy and outlying images, explaining its superior performance under both FSL settings. Thirdly, we report some quantitative 5-way 1-shot results under the conventional FS-L setting, together with the 5-way 5-shot results in the main paper,to validate the effectiveness of our AdarGCN with d-ifferent 5-way settings.

1. Qualitative Results for Label Denoising

As mentioned in the main paper, out of the crawled images per source class, around 40% are noise; after LDN using our AdaGCN, this percent is reduced to around 10%. Some examples of the removed images are shown in Figure 1. We find that both the removed images and the cleaned images for the source class 'Black Footed Albatross' are generally recognized correctly by our AdarGCN-based LD-N method. This means that our AdarGCN can deal with the noisy images and thus achieve superior performance under the new SSFSL setting.

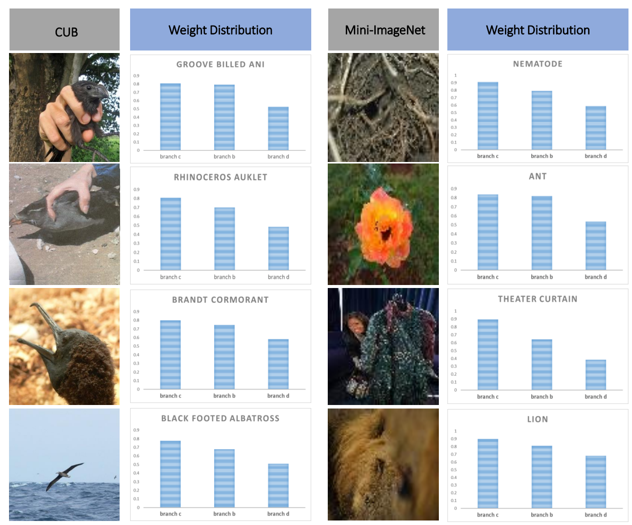

Figure 2. Illustration of the weight distributions of outlying examples over the CUB and mini-ImageNet datasets.

2. Qualitative Results for Conventional FSL

Figure 2 shows the weight distributions of outlying examples obtained by our AdarGCN under the conventional FSL setting. It can be clearly observed that the weight of branch **c** is forced to be significantly larger than those of the other two branches (i.e. **b**, **d**) for the outlying examples so that their negative effect can be effectively limited. This also explains why our AdarGCN can achieve state-of-the-

*Equal Contribution
†Corresponding Author

Models	GCN	*mini*ImageNet
MatchingNet [12]	no	43.60
ProtoNet [10]	no	46.61
MAML [4]	no	48.70
Relation Net [11]	no	50.44
TPN [8]	no	53.75
R2-D2 [2]	no	51.80
IMP [1]	no	49.60
Baseline++ [3]	no	48.24
MetaOptNet [7]	no	53.23
GCN [9]	yes	50.33
wDAE-GNN [5]	yes	51.02
EGCN [6]	yes	53.24
AdarGCN (ours)	yes	**55.71**

Table 1. Comparative 5-way 1-shot results for conventional FSL.

art results under the conventional FSL setting. Note that the problem of outlying images is really very hard to solve, because it tends to occur in a variety of ways (e.g. occlusion, very small scale, and so on). However, this challenging problem is shown to be well solved by our AdarGCN, due to its adaptive aggregation mechanism.

3. More Quantitative Results

Since only the 5-way 5-shot protocol is considered in the main paper for fair comparison to the latest GCN-based FSL methods, we also provide the 5-way 1-shot results for conventional FSL in Table 1 to make more extensive comparison. It can be seen that our AdarGCN FSL method yields 2–5% improvements over the latest GCN-based FSL methods [9, 6, 5] and 2–7% improvements over the state-of-the-art FSL baselines [8, 2, 1, 3, 7], validating the effectiveness of our AdarGCN module under one-shot setting.

References

[1] Kelsey R. Allen, Evan Shelhamer, Hanul Shin, and Joshua B. Tenenbaum. Infinite mixture prototypes for few-shot learning. In *ICML*, pages 232–241, 2019.

[2] Luca Bertinetto, Joao F Henriques, Philip HS Torr, and Andrea Vedaldi. Meta-learning with differentiable closed-form solvers. In *ICLR*, 2019.

[3] 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.

[4] Chelsea Finn, Pieter Abbeel, and Sergey Levine. Model-agnostic meta-learning for fast adaptation of deep networks. In *ICML*, pages 1126–1135, 2017.

[5] Spyros Gidaris and Nikos Komodakis. Generating classification weights with GNN denoising autoencoders for few-shot learning. In *CVPR*, pages 21–30, 2019.

[6] Jongmin Kim, Taesup Kim, Sungwoong Kim, and Chang D. Yoo. Edge-labeling graph neural network for few-shot learning. In *CVPR*, pages 11–20, 2019.

[7] Kwonjoon Lee, Subhransu Maji, Avinash Ravichandran, and Stefano Soatto. Meta-learning with differentiable convex optimization. In *CVPR*, pages 10657–10665, 2019.

[8] Yanbin Liu, Juho Lee, Minseop Park, Saehoon Kim, Eunho Yang, Sung Ju Hwang, and Yi Yang. Learning to propagate labels: Transductive propagation network for few-shot learning. In *ICLR*, 2019.

[9] Victor Garcia Satorras and Joan Bruna. Few-shot learning with graph neural networks. In *ICLR*, 2018.

[10] Jake Snell, Kevin Swersky, and Richard Zemel. Prototypical networks for few-shot learning. In *Advances in Neural Information Processing Systems*, pages 4077–4087, 2017.

[11] 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*, pages 1199–1208, 2018.

[12] Oriol Vinyals, Charles Blundell, Timothy Lillicrap, koray kavukcuoglu, and Daan Wierstra. Matching networks for one shot learning. In *Advances in Neural Information Processing Systems*, pages 3630–3638, 2016.