Semi-supervised Long-tailed Recognition using Alternate Sampling

Bo Liu
UC, San Diego
boliu@ucsd.edu

Haoxiang Li
Wormpex AI Research
lhxustcer@gmail.com

Hao Kang
Wormpex AI Research
haokheseri@gmail.com

Nuno Vasconcelos
UC, San Diego
nuno@ece.ucsd.edu

Gang Hua
Wormpex AI Research
ganghua@gmail.com

1. iNaturalist2018-SSLT

Dataset. We further curate a benchmark for semi-supervised long-tailed recognition based on iNaturalist 2018 [2]. iNaturalist 2018 is a long-tailed dataset sampled from natural distribution. We follow the distribution in both of the labeled and unlabeled subset. More specifically, Samples in each class is randomly down-sampled one-fifth of the total number as labeled data, and the remains are assigned as unsupervised subset. Classes with less than 2 labeled samples are eliminated. In result, iNaturalist2018-SSLT contains 8080 classes, with labeled samples from 200 to 2, and the unsupervised subset is 4 times larger.

Classes are divided into three splits based on the number of labeled samples: many-shot ($[100, +\infty)$), medium-shot ($[10, 100)$), and few-shot ($[2, 10)$). It is a extremely long-tailed dataset, with 134 many-shot classes, 1220 medium-shot classes, and 7010 few-shot classes.

Results. Results are shown in Table 1. Our method is the only one that improves the overall performance upon baseline. iNaturalist2018-SSLT is different from our other benchmarks in the amount of few-shot classes. It has a very long tail taking up 87% of the label space. This makes the dataset especially hard when combined with unsupervised data.

With the inferior quality of predictions, we see significant drop of Pseudo-Label method in many-shot split. In fact, Pseudo-Label decreases the accuracy of baselines in all splits. Our method mitigates this problem, and improve the few-shot performance. Given the fact that most classes are in few-shot split, our method is the only one that increase the overall performance.

Comparison among benchmarks. From CIFAR-10-SSLT to ImageNet-SSLT and iNaturalist2018-SSLT, the datasets have more and more classes and few-shot classes. In result, they are more and more challenging. This challenge makes Pseudo-Label method ineffective. From CIFAR-10-SSLT to ImageNet-SSLT, the shortcoming first appears in many-shot splits. On ImageNet-SSLT, Pseudo-Label improves the few-shot performance with a sacrifice of many-shot performance. Our method is more robust to this difficulty. It keeps the many-shot performance while improves the few-shot performance. On iNaturalist2018-SSLT, the Pseudo-Label improvement on few-shot split also disappears, and the drop on many-shot is big. Our method, however, can still improves the few-shot performance and control the drop of many-shot compared to the baseline.

All of these results show that semi-supervised long-tailed recognition is a challenging problem. Given the fact that this problem follows the natural workflow of data collecting, we believe it deserves more attention in the literature.

References

[1] Bingyi Kang, Saining Xie, Marcus Rohrbach, Zhicheng Yan, Albert Gordo, Jiashi Feng, and Yannis Kalantidis. Decoupling representation and classifier for long-tailed recognition. In *Eighth International Conference on Learning Representations (ICLR)*, 2020. 2

[2] 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 *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 8769–8778, 2018. 1

Table 1. Results(Accuracy in $\%$) on iNaturalist2018-SSLT. ResNet-50 are used for all methods. For many-shot $t > 100$, for medium-shot $t \in (10, 100]$, and for few-shot $t \leq 10$, where t is the number of labeled samples.

Method	Overall	Many-Shot	Medium-Shot	Few-Shot
Decoupling [1]	27.9	54.1	41.7	24.8
Pseudo-Label + Decoupling	26.3	39.9	35.8	24.3
Ours	28.4	49.5	38.7	26.1