Appendix: Improving Contrastive Learning on Imbalanced Seed Data via Open-World Sampling

Ziyu Jiang[1], Tianlong Chen[2], Ting Chen[3], Zhangyang Wang[2]
[1]Texas A&M University, [2]University of Texas at Austin, [3]Google Research, Brain Team
jiangziyu@tamu.edu, {tianlong.chen,atlaswang}@utexas.edu, iamtingchen@google.com

This appendix contains the following details that we could not include in the main paper due to space restrictions.

- **(Sec. A)** Details of the computing infrastructure.
- **(Sec. B)** Details of the employed hyperparameters.

A Details of computing infrastructure

Our codes are based on Pytorch [1], and all models are trained with NVIDIA A100 Tensor Core GPU.

B Details of hyper-parameter settings

B.1 Pre-training

We identically follow [2] for pre-training settings except the epochs number: we pre-train for 1000 epochs for all our experiments following [3] (Including the feature extractor).

B.2 Fine-tuning

For all fine-tuning, the optimizer is set as SGD with momentum of 0.9 and initial learning rate of 30 following [4]. When fine-tuning for *linear separability performance*, we train for 30 epochs and decrease the learning rate by 10 times at epochs 10 and 20. When fine-tuning for *few-shot performance*, we follow [5] fine-tuning from the first MLP projection layer. We train for 100 epochs with batch size 64. The initial lr is set as 0.02 and employing cosine learning rate decay without warm up [5].

References

[1] Adam Paszke, Sam Gross, Soumith Chintala, Gregory Chanan, Edward Yang, Zachary DeVito, Zeming Lin, Alban Desmaison, Luca Antiga, and Adam Lerer. Automatic differentiation in pytorch. 2017.

[2] Ziyu Jiang, Tianlong Chen, Bobak Mortazavi, and Zhangyang Wang. Self-damaging contrastive learning. In *International Conference on Machine Learning*, 2021.

[3] Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. A simple framework for contrastive learning of visual representations. In *International conference on machine learning*, pages 1597–1607. PMLR, 2020.

[4] Xinlei Chen, Haoqi Fan, Ross Girshick, and Kaiming He. Improved baselines with momentum contrastive learning. *arXiv preprint arXiv:2003.04297*, 2020.

[5] Ting Chen, Simon Kornblith, Kevin Swersky, Mohammad Norouzi, and Geoffrey Hinton. Big self-supervised models are strong semi-supervised learners. *arXiv preprint arXiv:2006.10029*, 2020.

35th Conference on Neural Information Processing Systems (NeurIPS 2021).