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

Ziyu Jiang¹, Tianlong Chen², Ting Chen³, Zhangyang Wang³
¹Texas A&M University, ²University of Texas at Austin, ³Google 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
