Reliable Label Bootstrapping for Semi-Supervised Learning - Supplementary material

I. Ablation study for label propagation

We analyze the importance on the final semi-supervised performance of using diffusion and good unsupervised features for label propagation. We compare in Table I the performance achieved when using nearest neighbours compared to diffusion and two different unsupervised pre-training strategies (AND [1] and RotNet [2]). The results show the importance of estimating good features and proper label propagation strategy.

II. ReMixMatch training

We report in Table II the error rates for ReMixMatch, with and without our proposed ReLaB method, when training for 256 epochs instead of the original 1024 [3]. Although longer training is always beneficial, we observe convergence to a reasonable performance in 256 epochs. In the paper we adopt the 256 configuration as it substantially reduces training time.

III. Ablation study for Pseudo-labeling

We study the effect on the pseudo-labeling algorithm in [4] when using an unsupervised initialization with RotNet [2] and freezing all the layers up to the last convolutional block to avoid fitting label noise of the reliable extended set $\mathcal{D}_r$. Unsupervised initialization and early layers freezing is also adopted in [5] to improve pseudo-labeling. We show in Table III that both strategies contribute to better pseudo-labeling performance.

IV. Visualization of the bootstrapped samples

Figure 1 displays the capacity of our reliable sample selection to select an extended clean subset for the semi-supervised algorithm. The first row displays the seed samples; the middle row display a random subset of the samples labeled using label propagation on self-supervised features; the last row displays the reliable samples we select to extend the label set. Images with a red border have a noisy label. The label noise is reduced in the reliable extended set. The figure is best viewed on a computer.

TABLE I

Benefits of diffusion over a nearest neighbor (NN) for label propagation. We report error rates over the same labeled subset and use ReLaB + Pseudo-labeling.

Dataset # labeled samples	CIFAR-10 40	CIFAR-100 400
RotNet [2] + NN	25.64	69.76
AND [1] + NN	40.70	69.67
RotNet [2] + Diffusion	16.49	65.35
AND [1] + Diffusion	**14.21**	**57.09**

TABLE II

Error rate of short and long training of the ReMixMatch algorithm [3]. We report mean and standard deviation over 3 runs.

Dataset # labeled samples	CIFAR-10 250	CIFAR-10 40	CIFAR-100 400
ReLaB	No	Yes	Yes
ReMixMatch - 256	7.8 ± 0.83	10.04 ± 4.58	48.59 ± 0.7
ReMixMatch - 1024	6.24 ± 0.34	9.04 ± 4.37	47.24 ± 0.68

TABLE III

Ablation study on the error rate of pseudo-labeling (PL) algorithm in [4] combined with ReLaB when using unsupervised initialization and layer freezing (LF).

Dataset # labeled samples	CIFAR-10 40	CIFAR-100 400
ReLaB + PL [4]	22.12	58.17
ReLaB + PL (RotNet [2])	15.72	59.04
ReLaB + PL (RotNet [2] + LF)	**14.21**	**57.09**

References

[1] H. Jiabo, D. Qi, G. Shaogang, and Z. Xiatian, "Unsupervised Deep Learning by Neighbourhood Discovery," in *International Conference on Machine Learning (ICML)*, 2019.

[2] S. Gidaris, P. Singh, and N. Komodakis, "Unsupervised Representation Learning by Predicting Image Rotations," in *International Conference on Learning Representations (ICLR)*, 2018.

[3] D. Berthelot, N. Carlini, E. Cubuk, A. Kurakin, K. Sohn, H. Zhang, and C. Raffel, "ReMixMatch: Semi-Supervised Learning with Distribution Matching and Augmentation Anchoring," in *International Conference on Learning Representations (ICLR)*, 2020.

[4] E. Arazo, D. Ortego, P. Albert, N. O'Connor, and K. McGuinness, "Pseudo-Labeling and Confirmation Bias in Deep Semi-Supervised Learning," *arXiv: 1908.02983*, 2019.

[5] S.-A. Rebuffi, S. Ehrhardt, K. Han, A. Vedaldi, and A. Zisserman, "Semi-Supervised Learning with Scarce Annotations," *arXiv: 1905.08845*, 2019.

Fig. 1. Qualitative example of label propagation and reliable sample selection in CIFAR-10 with four seed samples per class. Best viewed on a computer.