Supplementary for SLADE: A Self-Training Framework For Distance Metric Learning

Jiali Duan[1]* Yen-Liang Lin[2] Son Tran[2] Larry S. Davis[2] C.-C. Jay Kuo[1]
[1] University of Southern California [2] Amazon
{jialidua,jckuo}@usc.edu {yenliang,sontran,lrrydav}@amazon.com

1. t-SNE visualization

A common setting in semi-supervised learning is to treat a small random fraction (e.g., 10%) of an existing (labeled) dataset as the labeled set and the rest (e.g., 90%) as the unlabeled set ([5, 3, 2, 1, 4, 6, 7]). The labeled and unlabeled data, in this case, come from the same distribution, which makes the underlying tasks such as label propagation, sample interpolation, or pseudo label prediction easier. In practice, the label and unlabeled data sets might not be created by the same source and therefore would have distribution differences (e.g., MS COCO vs. ImageNet, CUB-200 vs. NABirds, In-shop vs. Fashion200K, etc). It is unclear how techniques such as label propagation, sample augmentation ([1, 4, 6, 7], etc) can be extended to this setting which we face in this work.

The following set of figures give an illustration of the distinction between these two settings. In figures 1 and 2, we show the t-SNE visualization using ImageNet embedding for labeled (e.g., CUB) and unlabeled data sets (e.g., NABirds). As can be seen from these figures, there is a significant gap between the labeled and unlabeled data sets. Note that we did not not supply the t-SNE algorithm with any label information to minimize visualization biases, and data from labeled set and unlabeled sets were sampled uniformly.

In figures 3 and 4, we show the t-SNE visualization using the embedding from our teacher model (after fine-tuning on the labeled data and would be used to predict pseudo labels for unlabeled data). The distribution gap became smaller, but it is still noticeable and would affect the accuracy of pseudo label prediction. Nevertheless, our student model is still able to utilize the noisy pseudo labels to improve the final retrieval performance as shown in the experimental section.

* Work done during an internship at Amazon.

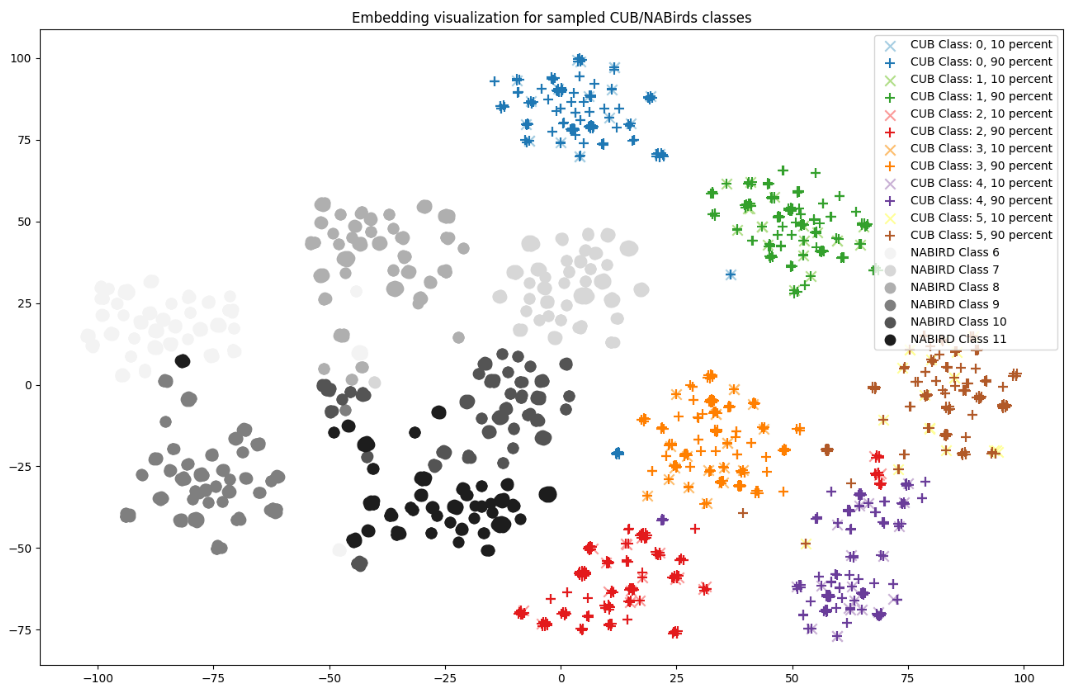

Figure 1. t-SNE visualization for CUB-200 (Colored) and NABirds (Grayed) using ImageNet embedding. A traditional setting would use 10% ($\times$) as the labeled and 90% ($+$) as the unlabeled sets, which share the same distribution. In our setting, CUB-200 (Colored) is the labeled set and NABirds (Grayed) is the unlabeled set. The distribution gap is quite significant. Best viewed electronically.

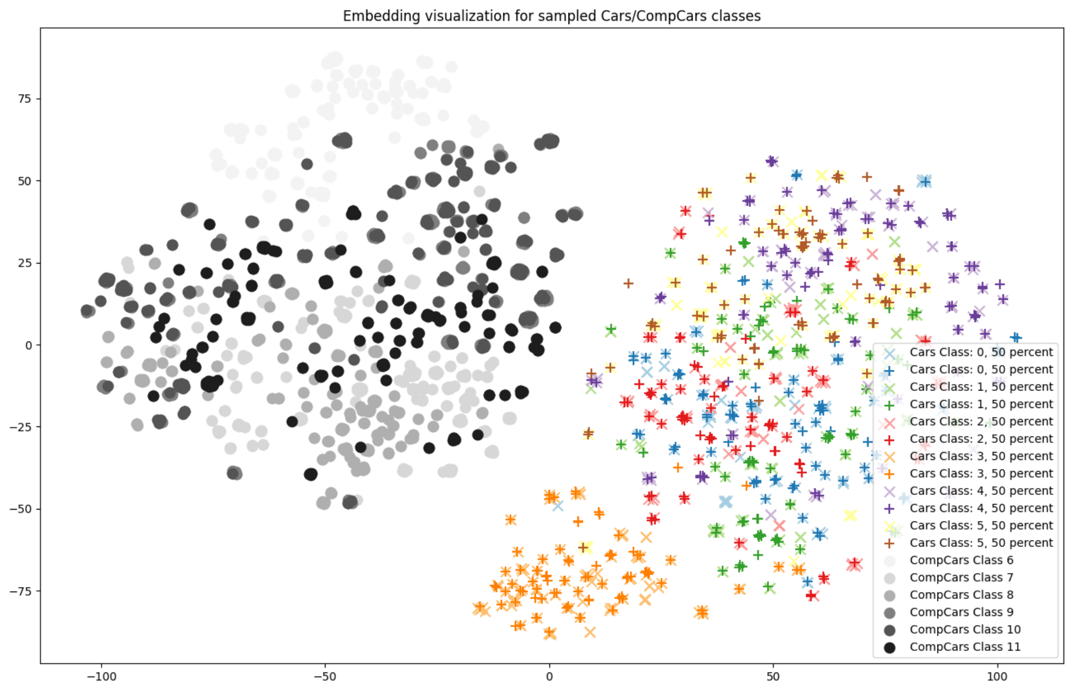

Figure 2. t-SNE visualization for Cars-196 (Colored) and CompCars (Greyed) using ImageNet embedding. See figure 1 for detailed explanations.

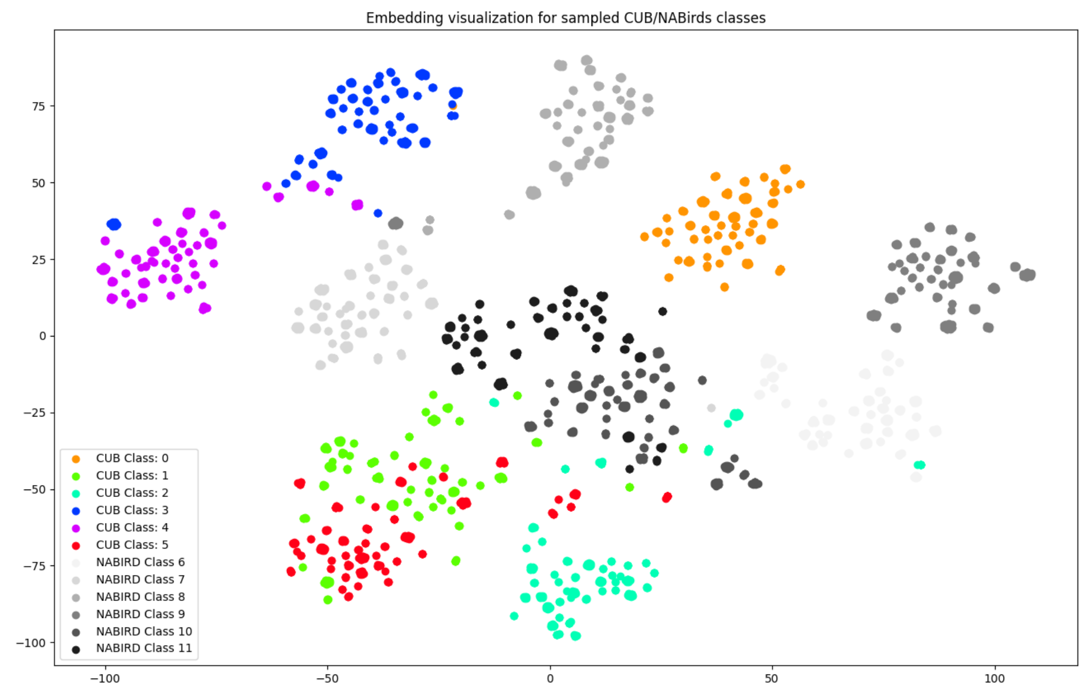

Figure 3. t-SNE visualization for CUB-200 (Colored, labeled) and NABirds (Grayed, unlabeled) using the embeddings from the teacher model (i.e., after supervised fine-tuning).

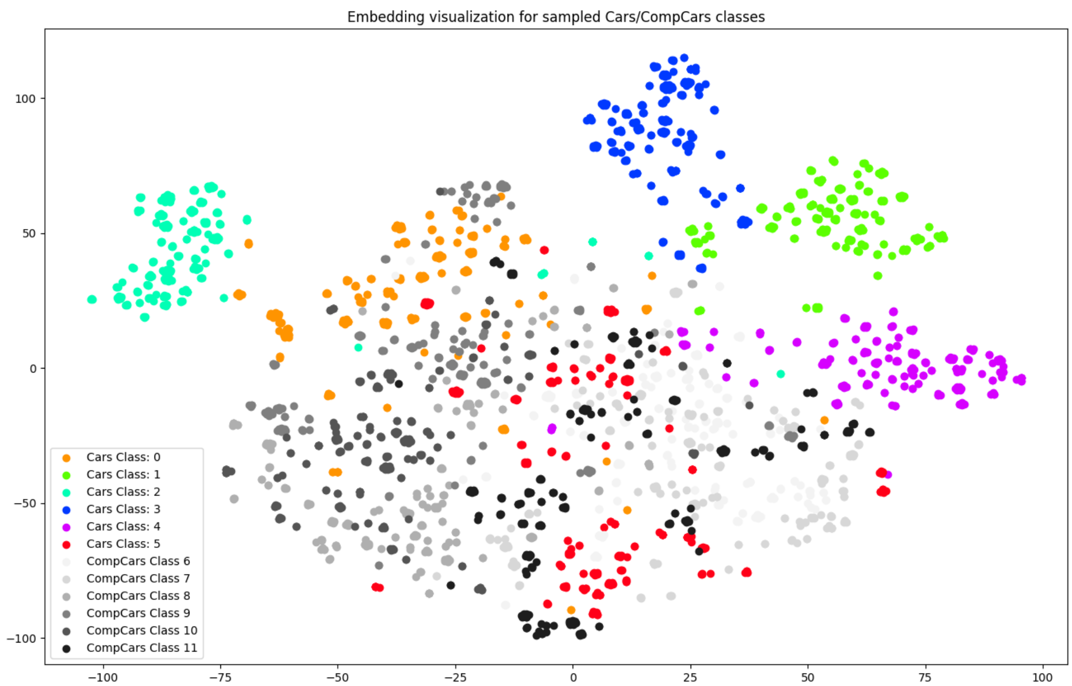

Figure 4. t-SNE visualization for Cars-196 (Colored, labeled) and CompCars (Greyed, unlabeled) using the embeddings from the teacher model

.

References

[1] Mathilde Caron, Ishan Misra, Julien Mairal, Priya Goyal, Piotr Bojanowski, and Armand Joulin. Unsupervised learning of visual features by contrasting cluster assignments. In *NeurIPS*, 2020. 1

[2] Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. A simple framework for contrastive learning of visual representations. In *ICML*, 2020. 1

[3] Ting Chen, Simon Kornblith, Kevin Swersky, Mohammad Norouzi, and Geoffrey Hinton. Big self-supervised models are strong semi-supervised learners. In *NeurIPS*, 2020. 1

[4] A. Iscen, G. Tolias, Y. Avrithis, and O. Chum. Label propagation for deep semi-supervised learning. In *CVPR*, 2019. 1

[5] Antti Tarvainen and Harri Valpola. Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results. In *NeurIPS*, 2017. 1

[6] Qin Wang, Wen Li, and Luc Van Gool. Semi-supervised learning by augmented distribution alignment. In *ICCV*, 2019. 1

[7] Xiaohua Zhai, Avital Oliver, Alexander Kolesnikov, and Lucas Beyer. S4l: Self-supervised semi-supervised learning. In *ICCV*, 2019. 1