RSA: Reducing Semantic Shift from Aggressive Augmentations for Self-supervised Learning Supplementary

Anonymous Author(s)
Affiliation
Address
email

A   Adaptability study

Table 1: Performance comparison with linear classification for 200 epochs. The mean and standard deviation are computed over three trails.

Method	CIFAR-100	STL-10	ImageNet-100
SimSiam	56.2±1.2	85.7±0.3	82.6
Simsiam + RSA	**63.8±0.4**	**89.9±0.5**	**84.0**

To evaluate the adaptability, we build our proposed RSA on SimSiam [1], termed SimSiam + RSA. Specifically, we change the same similarity function from mean square error to negative cosine similarity and adopt the same settings for all the hype-parameters mentioned in Section 4.1. We conduct experiments on CIFAR-100 [3], STL-10 [2], and ImageNet-100 [4] for 200 epochs.

As illustrated in Table 1, we observe that the proposed RSA achieves better transfer performance than SimSiam on three datasets. For instance, RSA significantly raise the accuracy of linear probing from 56.2% to 63.8% (+7.6%) on the CIFAR-100 dataset. Therefore, our proposed method does not rely on the momentum network, demonstrating the adaptability of RSA.

References

[1] Xinlei Chen and Kaiming He. Exploring simple siamese representation learning. In *CVPR*, pages 15750–15758, 2021.

[2] Adam Coates, Andrew Y. Ng, and Honglak Lee. An analysis of single-layer networks in unsupervised feature learning. In *AISTATS*, pages 215–223, 2011.

[3] Alex Krizhevsky, Geoffrey Hinton, et al. Learning multiple layers of features from tiny images. Technical report, 2009.

[4] Yonglong Tian, Dilip Krishnan, and Phillip Isola. Contrastive multiview coding. In *ECCV*, pages 776–794, 2020.