SCOMatch: Alleviating Overtrusting in Open-set Semi-supervised Learning

Zerun Wang[1], Liuyu Xiang[2], Lang Huang[1], Jiafeng Mao[1], Ling Xiao[1], and Toshihiko Yamasaki[1]

[1] The University of Tokyo
`{ze_wang, langhuang, ling, yamasaki}@cvm.t.u-tokyo.ac.jp`
`mao@hal.t.u-tokyo.ac.jp`
[2] Beijing University of Posts and Telecommunications
`xiangly@bupt.edu.cn`

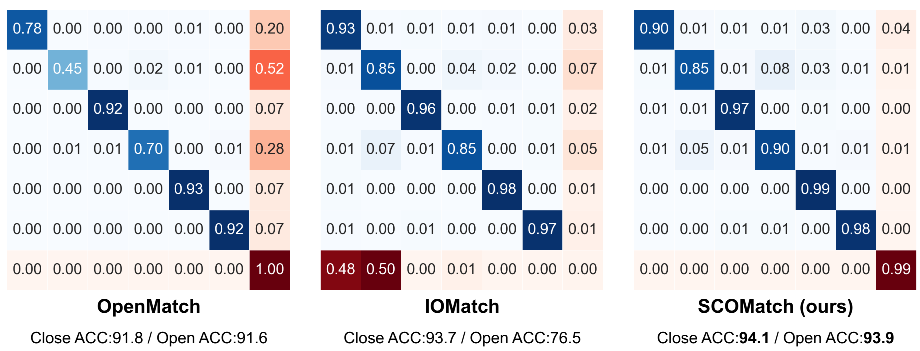

Fig. 5: The quantitative confusion matrixes of three methods. The results indicate that our performance is better considering both close-set and open-set accuracy.

We provide a detailed description of our implementation and show more analysis. Our code will be available at `https://github.com/komejisatori/SCOMatch`.

1 Additonal analysis

Analysis of the confusion matrix. We provide the confusion matrixes and corresponding accuracy of the three methods mentioned in Figure 1 of the main paper. The results in Figure 5 show that compared with SCOMatch, 1) OpenMatch [5] generates more false-negative ID samples. Thus, resulting in fewer pseudo-labels for SSL as they will be filtered. 2) IOMatch [4] generates more false-positive ID samples. Therefore, there will be more OOD samples joining the SSL incorrectly. The corresponding accuracy in the figure shows that both situations affect the SSL performance. The models are trained on CIFAR-10 with six ID classes and 50 labeled images per ID class.

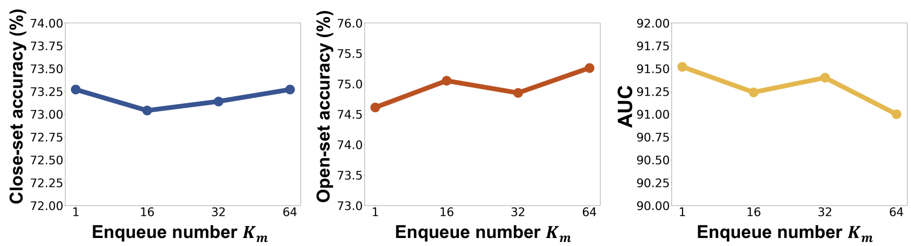

Fig. 6: Analysis of the enqueue number of the OOD memory queue.

Analysis of the enqueue number. We conduct additional ablation studies on the enqueue number K_m of the OOD memory queue. The results in Figure 6 indicate that the choice of K_m does not significantly affect the accuracy. In our experiments, we selected $K_m = 1$ as the default value. The analysis is conducted on CIFAR-100 with 55 ID classes and 25 labeled images per ID class.
Discussion of efficiency. Our SCOMatch is end-to-end trainable and does not require additional head branches. It only brings one extra dimension to the original classification head. Meanwhile, the number of sampled OODs per iteration is equal to a training batch, which only consumes about 0.9 GB of extra GPU memory. On the same device (NVIDIA A100), the training speed of SCOMatch (0.35 ms/iter) is comparable with the recent work IOMatch (0.32 ms/iter).

2 Datasets.

Following prior methods, we use the image size of 28×28 for MNIST [3], 32×32 for CIFAR-10 [2], CIFAR-100 [2], and TinyImageNet [1]. We use the image size of 224×224 for IN-30 [1].

3 Evaluation metrics.

We mainly reported three metrics in the paper: close-set classification accuracy, open-set classification accuracy, and Area under the ROC Curve (AUC).
Close-set classification accuracy. For evaluating the close-set classification accuracy, we remove the OOD samples in the test data regarding the class split of each dataset. Then we use the output of $\mathcal{C}^K$ as the prediction result.
Open-set classification accuracy. For evaluating the open-set classification accuracy, we use the entire test dataset and the output of the entire $\mathcal{C}$ as the prediction result.
Area under the ROC Curve. For evaluating the open-set classification accuracy, we use the entire test dataset and treat all the ID classes as one single class. For the confidence score, we use the predicted probability of the OOD class in $\mathcal{C}$ (*i.e.* the last dimension of $\mathcal{C}$).

Table 1: Hyper-parameters used in SCOMatch.

	Hyper-parameter	Description	MNIST	CIFAR-10/100	TinyImageNet	IN30
SSL (FixMatch)	τ	SSL confidence threshold	0.95			
	λ	Unsupervised loss weight	1			
	β	Momentum	0.9			
	α	EMA weight	0.999			
	Weight decay	Weight decay	0.0005			
	lr	Learning rate	0.001	0.03		
	B	Labeled data batch size	50	64		
Ours	K_m	No. of enqueue samples per iter	1			
	N_m	OOD memory queue length	8 $\times$ number of ID classes			
	τ^{min}	Min value of τ^{ood}	0.8	0.75		
	τ^{ood}	Confidence threshold for OOD class	0.8	CPL based threshold		

4 Implementation details.

Environment. All experiments are conducted with one NVIDIA A100 GPU. We use Python 3.6, PyTorch 1.10.0 with CUDA 11.1.
Hyper-parameters. Like many recent OSSL methods, SCOMatch is based on the SSL framework of FixMatch [6]. Therefore, we follow FixMatch's setting for SSL hyper-parameters and the data augmentation methods. Table 1 summarizes the hyper-parameters. For other OSSL methods mentioned in the paper, we use their original code and hyper-parameter settings if available.

References

1. Deng, J., Dong, W., Socher, R., Li, L.J., Li, K., Fei-Fei, L.: Imagenet: A large-scale hierarchical image database. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 248–255 (2009)
2. Krizhevsky, A., Hinton, G., et al.: Learning multiple layers of features from tiny images (2009)
3. LeCun, Y.: The mnist database of handwritten digits. http://yann. lecun. com/exdb/mnist/ (1998)
4. Li, Z., Qi, L., Shi, Y., Gao, Y.: Iomatch: Simplifying open-set semi-supervised learning with joint inliers and outliers utilization. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 15870–15879 (2023)
5. Saito, K., Kim, D., Saenko, K.: Openmatch: Open-set semi-supervised learning with open-set consistency regularization. Advances in Neural Information Processing Systems **34**, 25956–25967 (2021)
6. Sohn, K., Berthelot, D., Carlini, N., Zhang, Z., Zhang, H., Raffel, C.A., Cubuk, E.D., Kurakin, A., Li, C.L.: Fixmatch: Simplifying semi-supervised learning with consistency and confidence. Advances in Neural Information Processing Systems **33**, 596–608 (2020)