Semi-Supervised Domain Adaptation with Source Label Adaptation
Supplementary Material

In this supplementary material, we provide our detailed implementation results. The link to the code to reproduce our main results on *Office-Home* and *DomainNet* datasets is available at `https://github.com/chu0802/SLA`.

1. Implementation Detail

Our proposed framework, Source Label Adaptation (SLA) invovles cooperation with other state-of-the-art algorithms. We take MME and CDAC as our backbone models, named MME + SLA and CDAC + SLA, respectively. For **MME + SLA**, we use the official implementation at `https://github.com/VisionLearningGroup/SSDA_MME` to obtain the MME loss. For **CDAC + SLA**, we use the official implementation at `https://github.com/lijichang/CVPR2021-SSDA` to obtain the CDAC loss. We follow the suggestions in both papers to select all hyper-parameters across different datasets.

2. Experiment Detail

For each sub-task on *DomainNet* and *Office-Home* datasets, we run three times with different seeds and take the average to obtain the value. This sections provides the average values and the standard deviations of our experiments. Table 2 and Table 3 shows the detailed statistics for one-shot and three-shot Semi-Supervised Domain Adaptation (SSDA) on *Office-Home* dataset, respectively. Table 4 shows the statistics for one-shot and three-shot SSDA on *DomainNet* dataset.

3. Reproducibility Issue for MCL

MCL [1] is a state-of-the-art algorithm for SSDA, which performs consistency learning at three different levels and achieve great results. In our study, we also try to couple the MCL loss with our SLA framework. We follow the official implementation at `https://github.com/chester256/MCL` to reproduce the experiments. However, when reproducing the results on 3-shot *Office-Home* dataset. We found that it is generally hard to reach the reported numbers provided in their original paper. We address the issue by first reproducing MCL five times with different seeds using totally the same code in above. The detailed statistics are shown in Table 5. We then run another three trials for MCL and MCL + SLA by fixing the seed for the generator in the DataLoader. This step is to compare the two approaches in a much more fair manner. The link to our modified code is available in *MCL-SLA.txt*, and the results are shown in Table 6. As we stated in the main paper, though after applying SLA, we can generally do better than our reproducing MCL, we are still not able to achieve the reported values in the original work.

Method	MME	MME + SLA	CDAC	CDAC + SLA
1-shot	86.57	87.39	87.56	**88.59**
3-shot	88.31	88.64	88.07	**88.91**

Table 1. Avg. Accuracy (%) on *Office31* w/ ResNet34.

4. More examples for the adapted labels

We have shown the top-3 probabilities of the average adapted source label for Backpack (class 1) and Knives (class 30). We provide the whole probability distributions of the adapted labels on 6 representive classes in Figure 1. Compared with the original source labels, which are one-hot encoded, our adapted labels can be much closer to the ideal labels.

5. Results on *Office-31*

We show the results on *Office31* in Table 1, which again confirms the competitive advantages of SLA. Note that we reproduce the results for MME and CDAC on ResNet34 because we select it as our backbone model throughout the work.

References

[1] Zizheng Yan, Yushuang Wu, Guanbin Li, Yipeng Qin, Xiaoguang Han, and Shuguang Cui. Multi-level consistency learning for semi-supervised domain adaptation. *arXiv preprint arXiv:2205.04066*, 2022. 1, 3

Stats	A→C	A→P	A→R	C→A	C→P	C→R	P→A	P→C	P→R	R→A	R→C	R→P	**Mean**
							MME + SLA						
avg.	62.1%	76.3%	78.6%	67.5%	77.1%	75.1%	66.7%	59.9%	80.0%	72.9%	64.1%	83.8%	72.0%
std.	0.231	0.243	0.129	0.208	0.378	0.033	0.104	0.366	0.033	0.080	0.306	0.032	0.179
							CDAC + SLA						
avg.	63.0%	78.0%	79.2%	66.9%	77.6%	77.0%	67.3%	61.8%	80.6%	72.7%	66.1%	84.6%	72.9%
std.	0.431	0.873	0.133	0.111	0.653	0.200	0.404	0.324	0.066	0.489	0.270	0.117	0.339

Table 2. Results on *Office-Home* dataset for 1-shot Semi-Supervised Domain Adaptation with ResNet34.

Stats	A→C	A→P	A→R	C→A	C→P	C→R	P→A	P→C	P→R	R→A	R→C	R→P	**Mean**
							MME + SLA						
avg.	65.9%	81.1%	80.5%	69.2%	81.9%	79.4%	69.7%	67.4%	81.9%	74.7%	68.4%	87.4%	75.6%
std.	0.119	0.135	0.082	0.279	0.033	0.286	0.084	0.085	0.060	0.329	0.115	0.179	0.149
							CDAC + SLA						
avg.	67.3%	82.6%	81.4%	69.2%	82.1%	80.1%	70.1%	69.3%	82.5%	73.9%	70.1%	87.1%	76.3%
std.	0.295	0.186	0.060	0.411	0.233	0.178	0.128	0.119	0.181	0.436	0.426	0.073	0.227

Table 3. Results on *Office-Home* dataset for 3-shot Semi-Supervised Domain Adaptation with ResNet34.

	R → C		R → P		P → C		C → S		S → P		R → S		P → R		**Mean**	
Stats	1-shot	3-shot	1-shot	3-shot	1-shot	3-shot	1-shot	3-shot	1-shot	3-shot	1-shot	3-shot	1-shot	3-shot	**1-shot**	**3-shot**
									MME + SLA							
avg.	71.8%	73.3%	68.2%	70.1%	70.4%	72.7%	59.3%	63.4%	64.9%	67.3%	61.8%	63.9%	77.2%	79.6%	68.8%	70.0%
std.	0.217	0.231	0.082	0.135	0.244	0.207	0.361	0.238	0.129	0.097	0.148	0.083	0.213	0.203	0.199	0.171
									CDAC + SLA							
avg.	79.8%	81.6%	75.6%	76.0%	77.4%	80.3%	68.1%	71.2%	71.7%	73.5%	71.7%	73.5%	80.4%	82.5%	75.0%	76.9%
std.	0.224	0.363	0.079	0.122	0.231	0.213	0.713	0.198	0.326	0.235	0.135	0.099	0.387	0.174	0.299	0.201

Table 4. Results on *DomainNet* dataset for 1-shot and 3-shot Semi-Supervised Domain Adaptation with ResNet34.

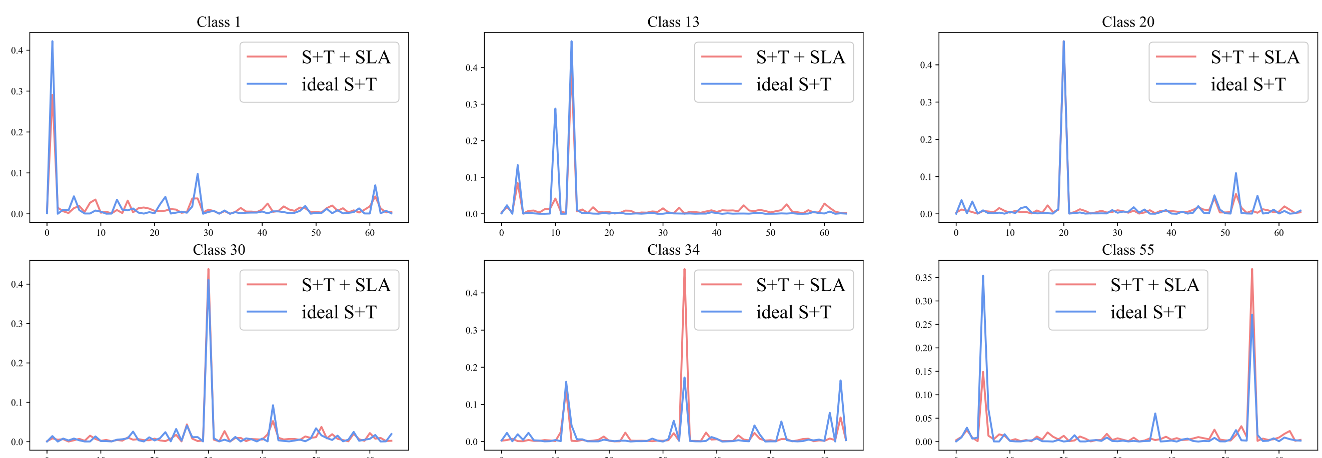

Figure 1. Average adapted source labels from the PPC and ideal S+T for a certain class (3-shot Office-Home A → C with ResNet34). x-axis: the classes, y-axis: the probability of the average adapted labels. We illustrate the average adapted source labels in S+T + SLA on six representative classes. Note that the original source labels should be one-hot encoded. The results show that the adapted labels could be much closer to the ideal labels.

Stats	A→C	A→P	A→R	C→A	C→P	C→R	P→A	P→C	P→R	R→A	R→C	R→P	**Mean**
avg.	63.5%	81.6%	80.7%	69.7%	82.4%	79.2%	70.6%	65.0%	82.7%	75.2%	67.8%	86.6%	75.4%
std.	0.678	0.647	0.476	0.648	1.033	0.506	0.311	0.823	0.151	0.269	0.847	0.301	0.558
min.	62.5%	80.7%	79.8%	68.9%	80.5%	78.3%	70.3%	63.8%	82.4%	74.8%	66.7%	86.3%	74.6%
max.	64.4%	82.4%	81.1%	70.7%	83.5%	79.7%	71.2%	66.3%	82.9%	75.5%	69.3%	87.1%	76.2%
reported	67.5%	83.9%	82.4%	71.4%	84.3%	81.6%	69.9%	68.0%	83.0%	75.3%	70.1%	88.1%	77.1%

Table 5. The detailed statistics of our reproducing results for MCL on 3-shot *Office-Home* dataset with ResNet34. We reproduce MCL five times with different seeds. **reported**: The reported numbers provided in the original paper [1].

Stats	A→C	A→P	A→R	C→A	C→P	C→R	P→A	P→C	P→R	R→A	R→C	R→P	**Mean**
							MCL*						
avg.	64.1%	81.6%	80.6%	70.3%	82.2%	79.2%	70.6%	64.0%	81.8%	75.3%	67.8%	86.6%	75.3%
std.	0.237	0.345	0.318	0.678	0.830	0.730	0.073	0.106	0.212	0.147	0.321	0.440	0.370
							MCL + SLA						
avg.	64.3%	81.6%	80.8%	70.2%	82.6%	79.4%	70.9%	64.2%	82.2%	75.5%	68.0%	86.8%	75.6%
std.	0.380	0.090	0.250	0.551	0.900	0.489	0.077	0.114	0.000	0.090	0.261	0.332	0.295
reported	67.5	83.9	82.4	71.4	84.3	81.6	69.9	68.0	83.0	75.3	70.1	88.1	77.1

Table 6. Results of MCL* and MCL + SLA with another 3 different seeds on 3-shot *Office-Home* dataset. *: Reproduced by ourselves. **reported**: The reported values in the original paper [1].