Appendix for " Link-based Contrastive Learning for Single-shot Unsupervised Domain Adaptation "

Yue Zhang Mingyue Bin YuyangZhang Zhongyuan Wang Zhen Han Chao Liang*
School of Computer Science, Wuhan University
National Engineering Research Center for Multimedia Software(NERCMS)
Hubei Key Laboratory of Multimedia and Network Communication Engineering, Wuhan
{moozy924, bin_mingyue, 2021302131204, hanzhen, cliang}@whu.edu.cn, wzy_hope@163.com

A. The VSA Dataset

Data were collected in public areas, with prior notice given at the time of data collection, and consent obtained from individuals for the usage of their data. To protect personal privacy, facial features and eyes in the images were obscured. The related data and code will be made publicly available upon publication of the paper. To protect privacy, we will not provide the raw facial images. Instead, we will only release the facial features used in the experiments.

A.1. Construction Details

The VSA dataset is collected in a real-world office building with multiple surveillance cameras, as shown in Figure 1. We first collect the 720R@20fps video frames captured by 6 surveillance cameras (denoted by the red circles), then perform YOLOv3 [3] and SORT [1] to detect and trace persons' tracks. The ground truth person IDs are manually annotated on these tracks, on which we adopt RetinaFace [2] to detect faces from surveillance videos. These faces are treated as target samples in OSUDA. Meanwhile, we obtain the clear frontal faces from the enrollment gallery and take them as the source samples that inherently contain person IDs. Given that the OSUDA problem assumes the source and target domains share the same set of categories, we only retain the faces whose IDs are appeared in both domains to construct the VSA dataset. Since the faces in the same track are usually visually similar, we evenly select at most 10 faces in each track.[1] And to fit with the mainstream face recognition models, we resize all face images in both domains to 112 px × 112 px.

*Chao Liang is the corresponding author..

[1]The tracks inherently contain temporal information that can be utilized as weak constraints to establish in-domain links in the proposed LCL method. But in order to make a fair comparison with existing UDA/SFDA/FUDA methods that are not able to utilize such weak constraints, we remove the temporal information in the VSA dataset and treat each face as an independent sample.

Table 1. The statistics of the VSA dataset.

Domain	#Tracks	#IDs	#Images	#Images/ID		
				Max	Min	Avg.
Enrollment (source)	-	46	46	1	1	1
Surveillance (target)	516	46	3,803	188	19	83

A.2. Statistics and Examples

As shown in Figure 2, in the VSA dataset, each person has only one clear frontal portrait in the enrollment gallery (source domain), and numerous blurry angled faces, captured by surveillance cameras set up in high places, to be recognized (target domain). The overall statistics of VSA are presented in Table 1, where "#Images" denotes the total number of images in each domain, "#Images/ID" denotes the number of images per ID. It's worth noting that due to the privacy issue, there is only one image per ID available in the source domain, hence the VSA dataset is a typical case to study the OSUDA problem.

B. Additional Experimental Results

B.1. Hyperparameter Analysis

This section analyzes the hyperparameters of LCL, including α, β, λ and K on the VSA dataset, where we choose ArcFace as our backbone.

Figure 3 shows the dynamic performance with varying λ and K (actually K/N^s). We can observe that when K is fixed, the variation amplitude of adaptation accuracy is less than 1% with λ varying from 0 to 1, except when $K = 16N^s$, and when both K/N^s and λ equal 1. This shows that LCL is not sensitive to λ, and we choose $\lambda = 0.5$ as our default hyperparameter. And when λ is fixed, the performance first improves when K changes from N^s to $2N^s$ but then begins to drop when K/N^s is larger than 4. This may be due to directly setting K as N^s can hardly exclude

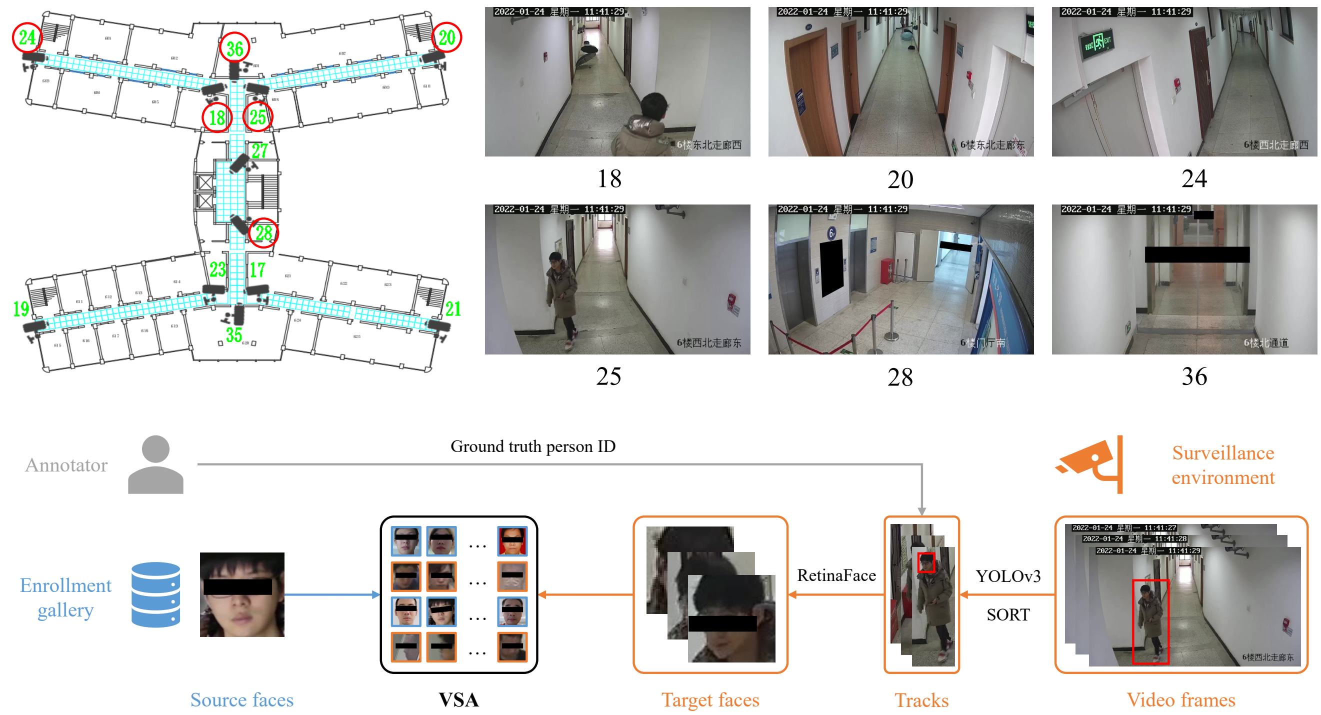

Figure 1. The construction process of the VSA dataset.

the outliers from each cluster, while a too-large K tends to separate semantically similar samples into different clusters. We choose $K = 2N^s$ as our default hyperparameter.

We then fix λ = 0.9 and $K = 2N^s$ to analyze the dynamic performance with varying α and β, and the results are illustrated in Figure 4. We can observe that LCL is less sensitive to the change of α compared with the change of β, and that LCL gains the best result when α = 0.5 and β = 5. When β = 5, the performance is relatively robust to the varying α, and the amplitude of performance variation is smaller than 0.5%. And when α = 0.5, the performance significantly improves as β goes up.

B.2. Full Experimental Results with Mean±Std

This section presents the full mean±std results of the 5-time repeated experiments. Tables 3-6 correspond to Tables 1 - 4 in the main text, respectively. And Table 7 corresponds to Table 5 in the main text.

B.3. Cross-Domain Image Retrieval

In this section, we present cross-domain image retrieval results on the VSA dataset. Given a query feature $\boldsymbol{f}_i^t$ in the target domain, we measure the cosine similarity between $\boldsymbol{f}_i^t$ and all source features $[\boldsymbol{f}_i^s]_{i=1}^{N^s}$. The most similar source images are returned as the top retrieval results. Table 2 shows the image retrieval results, where "R@1" and "R@2" denote the recall among the top-1, and top-2 retrieved images, respectively. We can observe that LCL outperforms the best baseline method, PCS, by over 7.6% in R@1 and R@2 on average.

Figure 5 shows some image retrieval examples using LCL and PCS, where both methods adopt PartialFC as the backbone for a fair comparison. We can observe that PCS fails to retrieve the same identity across domains when the query faces are blurry, occluded, or angled. In contrast, LCL successfully retrieves the correct faces even though the query images are blurry, angled or occluded. These results further demonstrate the effectiveness of the proposed LCL in the real-world OSUDA problem.

C. Limitations

Our goal is to learn discriminative features through in-domain contrastive learning and domain-invariant features through cross-domain contrastive learning, both crucial for domain adaptation. To achieve this, we propose link-based contrastive learning under two assumptions: (1) Samples in the same cluster tend to share the same label, and (2) Bidirectionally matched clusters across domains tend to be semantically consistent.

Inevitably, these assumptions have their limitations: (1) The clustering results are prone to be affected by outliers. (2) The bidirectional matching mechanism is not guaranteed to generate a matching result for every cluster. As a result, the performance of the proposed LCL is likely to be

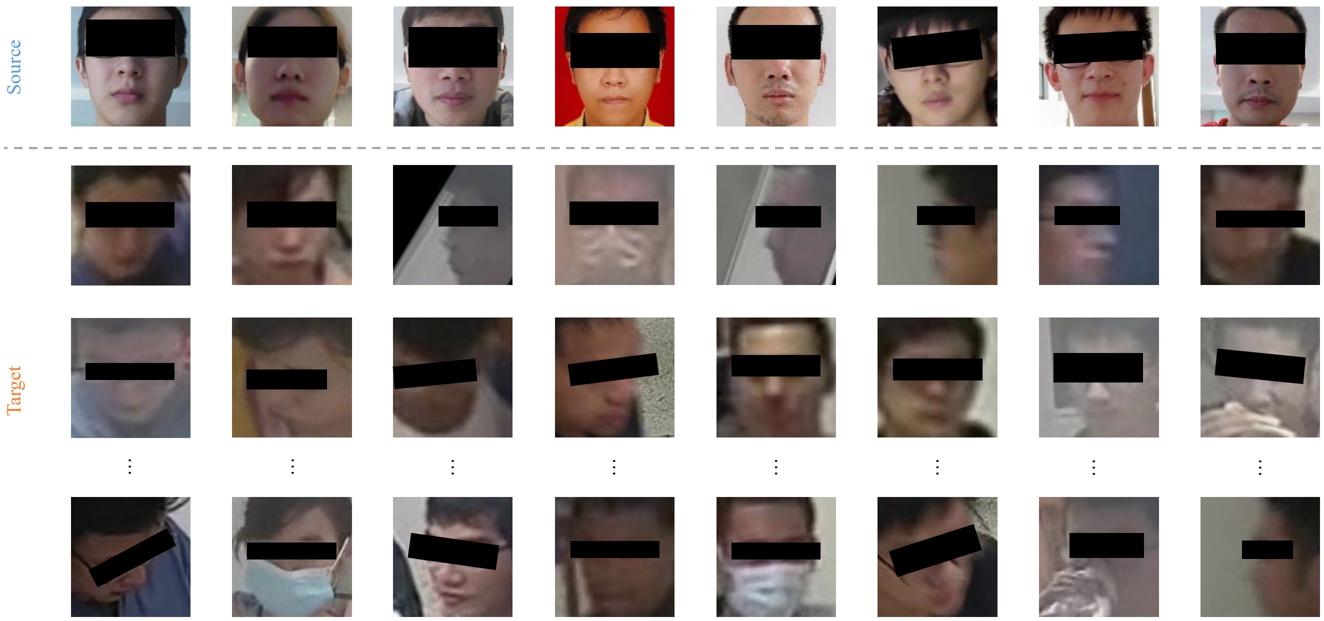

Figure 2. Examples of the VSA dataset, where images in the same column represent the same person.

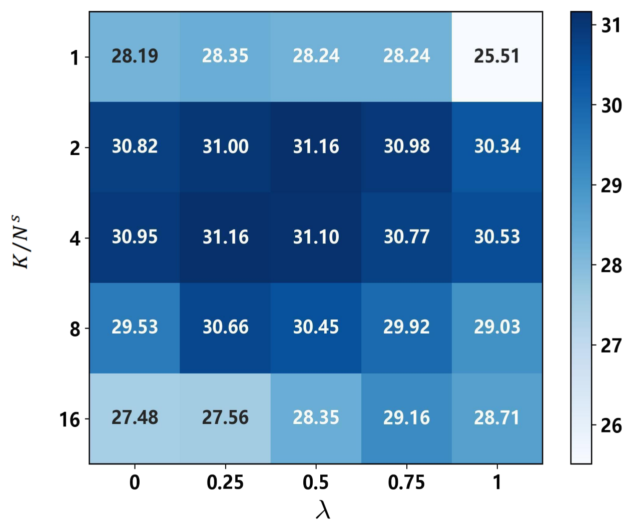

Figure 3. Dynamic performance of λ and K on VSA.

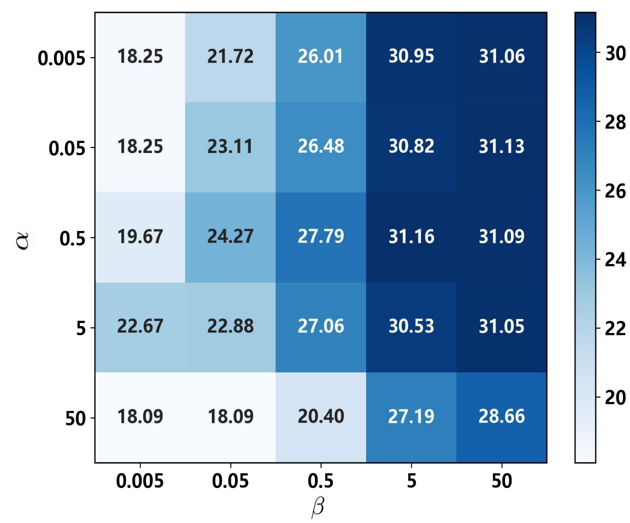

Figure 4. Dynamic performance of α and β on VSA.

affected by low-quality clustering and bidirectional matching results.

Despite this, these assumptions help achieve our goal: (1) The clustering assumption enables unsupervised in-domain links, and (2) Bidirectional matching enforces reliable cross-domain links. Table 7 confirms the contribution of each component, demonstrating the rationality of the above assumptions.

In the future, we will generalize the binary hard links to probabilistic soft links to improve the robustness of the method.

References

[1] Alex Bewley, ZongYuan Ge, Lionel Ott, Fabio Tozeto Ramos, and Ben Upcroft. Simple online and realtime tracking. In *IEEE International Conference on Image Processing*, pages 3464–3468. IEEE, 2016. 1

[2] Jiankang Deng, Jia Guo, Evangelos Ververas, Irene Kotsia, and Stefanos Zafeiriou. Retinaface: Single-shot multi-level face localisation in the wild. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 5202–5211. IEEE, 2020. 1

[3] Joseph Redmon and Ali Farhadi. Yolov3: An incremental improvement. *arXiv*, 2018. 1

Figure 5. Image retrieval examples on the VSA dataset using PCS and LCL. The top-8 retrieval results are arranged in descending order of cosine similarity from left to right, where green boxes indicate the correct faces in the source domain.

Table 2. Image retrieval results on the VSA dataset.

Method	**SphereFace**		**CosFace**		**ArcFace**		**VPL**		**AdaFace**		**PartialFC**		**Avg.**	
	R@1	**R@2**	**R@1**	**R@2**	**R@1**	**R@2**	**R@1**	**R@2**	**R@1**	**R@2**	**R@1**	**R@2**	**R@1**	**R@2**
Direct	16.54	22.32	28.14	37.73	17.62	24.14	16.17	21.75	24.64	32.00	25.64	34.03	21.46	28.66
PCS	25.53	33.68	30.66	41.10	25.93	32.90	22.77	30.19	26.74	35.58	36.71	44.28	28.06	36.29
UCDIR	21.59	28.00	35.81	45.75	21.80	29.74	19.62	25.56	30.06	37.97	37.21	45.12	27.68	35.36
SDAT	2.60	4.65	3.42	6.02	3.21	5.10	3.18	5.86	3.47	6.02	3.63	6.29	3.25	5.66
CAF-A	3.84	6.21	3.47	6.07	3.68	6.73	2.05	3.50	1.81	3.68	1.42	3.21	2.71	4.90
CAF-D	1.57	3.71	1.35	2.86	1.98	4.18	2.13	4.58	2.27	4.37	1.87	4.01	1.86	3.95
ICON	2.10	4.40	2.20	4.60	1.90	4.20	2.60	4.50	2.00	4.30	2.30	4.20	2.18	4.36
ProtoOT	2.45	4.68	2.76	5.31	2.73	5.18	3.55	6.26	2.81	5.10	2.45	5.36	2.79	5.32
MUCDA	2.39	4.52	2.47	4.85	2.10	4.13	2.85	4.92	2.23	4.44	2.30	4.08	2.39	4.49
SHOT	7.10	11.25	9.99	17.17	8.52	12.88	7.47	10.70	4.92	8.55	17.12	23.64	9.19	14.03
NRC	13.28	18.88	21.85	29.61	13.99	19.83	13.36	18.25	19.04	26.06	18.62	25.27	16.69	22.98
AaD	13.23	18.70	20.85	28.45	14.28	20.06	13.70	18.83	19.30	25.95	19.09	25.80	16.74	22.97
SFDA-DE	11.10	15.94	21.90	28.40	11.96	16.86	11.81	16.17	22.43	28.93	21.01	28.27	16.70	22.43
BMD	14.41	21.11	21.98	30.08	15.46	21.67	13.83	19.06	21.69	29.50	20.83	28.00	18.03	24.90
GPUE	19.75	26.77	25.27	34.84	18.80	25.95	15.99	23.40	27.08	35.95	28.16	37.63	22.51	30.76
SFADA	11.54	16.67	15.28	20.75	11.52	16.04	13.96	18.06	4.73	7.73	11.20	14.46	11.37	15.62
TPDS	8.69	12.48	10.37	25.83	11.78	15.46	13.54	19.69	6.29	11.28	10.70	13.75	10.23	18.08
LCL	**28.77**	**35.50**	**43.44**	**53.54**	**30.74**	**37.76**	**25.35**	**30.71**	**42.41**	**48.86**	**44.33**	**54.40**	**35.84**	**43.46**

Table 3. Acc. (%) on VisDA-2017.

Category	Method	Venue	Acc.
w/o. DA	Direct	Baseline	40.20 ±8.54
UDA/FUDA	PCS	CVPR 2021	56.22 ±3.86
	UCDIR	ECCV 2022	45.60 ±2.92
	SDAT	ICML 2022	49.67 ±8.71
	CAF-A	TKDE 2023	34.55 ±5.63
	CAF-D	TKDE 2023	50.28 ±11.27
	ICON	NeurIPS 2023	50.46 ±2.77
	ProtoOT	AAAI 2024	34.38 ±4.66
	MCUDA	TOMM 2024	49.77 ±7.20
SFDA	SHOT	ICML 2020	59.67 ±6.47
	NRC	NeurIPS 2021	38.94 ±5.37
	AaD	NeurIPS 2022	39.81 ±5.35
	SFDA-DE	CVPR 2022	59.23 ±7.05
	BMD	ECCV 2022	61.11 ±1.19
	GPUE	CVPR 2023	55.92 ±7.56
	SFADA	PR 2023	49.24 ±7.01
	TPDS	IJCV 2024	61.31 ±2.98
OSUDA	LCL	Ours	**63.09 ±3.63**

Table 4. Adaptation accuracy (%) on Office-31.

Method	A→D	A→W	D→A	D→W	W→A	W→D	Avg.
Direct	55.26 ±3.87	54.42 ±5.37	53.46 ±2.99	77.41 ±2.02	50.49 ±3.72	75.42 ±3.88	61.08 ±2.07
PCS	65.98 ±5.80	63.34 ±8.14	62.62 ±2.99	**86.44 ±2.51**	61.72 ±3.24	84.74 ±2.76	70.81 ±2.65
UCDIR	34.26 ±2.63	35.95 ±5.38	38.48 ±3.40	54.49 ±5.02	34.15 ±1.03	45.70 ±7.63	40.50 ±1.46
SDAT	39.60 ±3.15	32.88 ±2.71	39.36 ±0.71	63.60 ±1.81	32.72 ±1.84	56.54 ±4.11	44.12 ±0.18
CAF-A	57.47 ±3.25	58.31 ±7.61	54.30 ±4.01	83.22 ±4.92	50.37 ±5.14	79.52 ±3.68	63.87 ±2.28
CAF-D	49.72 ±3.11	53.13 ±4.54	52.47 ±3.27	79.22 ±3.20	47.77 ±3.18	77.75 ±5.50	59.98 ±0.87
ICON	63.67 ±1.07	68.87 ±0.62	66.26 ±0.27	85.96 ±1.56	**67.68 ±0.46**	78.58 ±2.04	71.80 ±0.63
ProtoOT	38.96 ±4.38	39.12 ±1.70	39.63 ±2.73	62.87 ±2.78	33.28 ±3.39	57.79 ±3.30	45.28 ±0.49
MCUDA	40.22 ±3.40	33.42 ±2.62	39.36 ±0.71	63.30 ±1.81	32.72 ±1.84	56.54 ±4.11	44.31 ±1.22
SHOT	54.54 ±6.28	51.09 ±5.11	54.86 ±1.54	73.33 ±1.93	51.35 ±3.35	70.72 ±2.50	59.32 ±1.43
NRC	51.73 ±3.93	47.52 ±3.12	46.86 ±1.20	72.05 ±2.33	43.29 ±3.04	73.25 ±3.21	55.78 ±0.78
AaD	51.73 ±3.93	47.52 ±3.12	47.18 ±1.25	72.05 ±2.33	43.59 ±2.87	73.25 ±3.21	55.89 ±0.83
SFDA-DE	60.84 ±4.35	61.54 ±4.40	57.01 ±2.51	78.39 ±1.43	52.86 ±2.70	75.82 ±1.62	64.41 ±1.69
BMD	51.57 ±1.13	48.38 ±3.62	53.49 ±1.97	72.43 ±1.50	52.69 ±2.15	74.05 ±1.76	58.77 ±0.88
GPUE	61.29 ±4.48	58.59 ±2.82	54.53 ±2.19	78.93 ±5.28	50.54 ±3.87	75.50 ±4.29	63.23 ±1.49
SFADA	54.86 ±7.17	55.25 ±5.00	56.35 ±2.57	75.02 ±3.34	51.55 ±2.26	75.46 ±3.94	61.42 ±1.45
TPDS	57.47 ±1.50	57.59 ±1.02	52.33 ±0.77	71.13 ±1.07	51.64 ±1.49	75.34 ±2.10	61.92 ±0.47
LCL	**69.28 ±5.36**	**68.96 ±3.46**	**68.14 ±3.97**	84.10 ±2.19	65.46 ±5.78	**90.12 ±3.31**	**74.34 ±0.90**

Table 5. Adaptation accuracy (%) comparison on Office-Home.

Method	Ar→Cl	Ar→Pr	Ar→Rw	Cl→Ar	Cl→Pr	Cl→Rw	Pr→Ar	Pr→Cl	Pr→Rw	Rw→Ar	Rw→Cl	Rw→Pr	Avg.
Direct	18.25 ±0.31	30.42 ±2.81	36.65 ±2.49	24.72 ±2.56	27.02 ±2.32	29.11 ±1.69	31.13 ±1.04	22.81 ±1.04	44.76 ±1.94	35.58 ±3.72	23.00 ±2.61	41.90 ±2.19	30.45 ±1.01
PCS	26.62 ±1.84	42.62 ±2.86	49.36 ±3.66	33.68 ±2.67	40.68 ±2.98	41.84 ±2.34	38.47 ±2.50	30.09 ±2.80	56.97 ±3.70	44.25 ±4.21	29.86 ±3.52	54.47 ±1.49	40.74 ±1.42
UCDIR	15.64 ±1.42	24.55 ±2.41	30.20 ±1.84	17.12 ±4.22	22.68 ±4.03	22.63 ±4.71	24.94 ±2.77	18.14 ±1.05	33.10 ±2.24	28.68 ±2.51	18.86 ±1.08	37.03 ±1.10	24.47 ±1.55
SDAT	13.40 ±0.75	19.92 ±1.69	25.62 ±1.91	15.13 ±2.87	15.89 ±1.54	16.89 ±2.94	22.17 ±1.86	15.46 ±1.00	31.62 ±1.81	29.55 ±2.65	16.49 ±1.45	31.23 ±2.51	21.11 ±1.03
CAF-A	15.02 ±1.77	32.12 ±1.63	40.13 ±2.98	17.87 ±2.58	22.18 ±2.56	24.46 ±4.89	26.44 ±1.65	17.32 ±2.09	42.47 ±3.57	34.12 ±4.23	20.43 ±1.71	47.00 ±4.25	28.30 ±1.62
CAF-D	15.85 ±0.92	26.48 ±5.25	33.92 ±4.17	18.67 ±2.48	20.26 ±1.56	22.31 ±2.44	27.22 ±2.79	19.39 ±2.54	40.69 ±3.06	35.02 ±6.53	20.75 ±1.67	42.04 ±3.65	26.88 ±1.72
ICON	23.64 ±2.88	33.30 ±3.66	45.82 ±5.77	30.82 ±3.19	28.10 ±3.43	31.68 ±2.68	31.70 ±1.68	28.46 ±2.03	53.38 ±3.82	44.85 ±7.14	28.98 ±1.76	49.58 ±3.34	36.44 ±1.53
ProtoOT	14.66 ±1.09	22.52 ±3.80	27.52 ±3.42	17.42 ±3.53	21.51 ±1.96	22.63 ±1.17	24.05 ±2.31	16.55 ±1.17	32.24 ±4.58	29.17 ±1.83	17.25 ±0.52	34.35 ±1.45	23.32 ±1.28
MUCDA	16.20 ±1.39	20.92 ±1.98	27.22 ±1.72	15.53 ±3.21	18.89 ±1.82	19.09 ±4.55	21.21 ±1.06	17.46 ±4.11	30.94 ±1.07	31.35 ±4.28	16.61 ±1.91	31.95 ±3.30	22.28 ±0.58
SHOT	21.53 ±2.17	31.28 ±1.93	37.21 ±2.63	24.23 ±2.88	26.91 ±1.71	28.73 ±2.96	32.44 ±0.62	24.64 ±1.47	45.12 ±3.18	37.42 ±2.43	26.46 ±1.86	45.85 ±2.45	31.82 ±1.38
NRC	19.33 ±1.15	29.02 ±1.74	25.56 ±2.59	22.37 ±1.83	24.56 ±1.25	26.77 ±1.50	30.71 ±1.88	23.36 ±1.34	42.36 ±1.75	35.12 ±2.62	24.32 ±1.67	40.64 ±2.13	29.51 ±0.82
AaD	19.33 ±1.15	29.02 ±1.74	35.56 ±2.59	22.37 ±1.83	24.56 ±1.24	26.77 ±1.50	30.71 ±1.88	23.36 ±1.34	42.37 ±1.75	35.12 ±2.62	24.32 ±1.67	40.64 ±2.13	29.51 ±0.82
SFDA-DE	24.19 ±1.86	34.07 ±2.30	40.33 ±2.85	26.25 ±2.48	29.11 ±2.05	30.96 ±1.89	34.07 ±2.61	29.08 ±0.86	47.72 ±1.41	38.48 ±3.25	27.99 ±2.82	46.57 ±3.84	34.07 ±0.98
BMD	21.91 ±1.44	34.78 ±2.97	41.12 ±3.31	25.92 ±3.40	30.52 ±1.79	32.07 ±1.40	34.72 ±2.35	26.65 ±1.65	47.97 ±1.72	37.67 ±2.74	26.79 ±1.68	46.47 ±1.85	33.88 ±1.09
GPUE	19.99 ±2.71	33.03 ±3.57	40.84 ±3.52	24.11 ±2.47	27.06 ±1.99	28.07 ±2.26	35.48 ±1.16	23.56 ±2.15	47.16 ±1.31	41.14 ±3.37	25.55 ±2.23	47.90 ±1.20	32.82 ±1.20
SFADA	19.26 ±2.83	28.67 ±4.62	35.71 ±4.26	24.81 ±6.14	23.27 ±5.77	27.65 ±3.58	34.53 ±3.38	25.45 ±3.65	45.76 ±3.21	37.55 ±5.00	25.52 ±3.98	44.47 ±4.16	31.05 ±1.76
TPDS	25.13 ±1.54	33.71 ±2.41	40.99 ±1.98	26.66 ±2.92	28.45 ±3.30	31.64 ±2.31	34.91 ±1.10	29.90 ±2.49	43.14 ±4.03	38.70 ±3.61	26.85 ±3.29	45.17 ±3.02	33.77 ±0.98
LCL	**32.55 ±3.41**	**49.57 ±3.85**	**55.91 ±4.89**	**38.88 ±2.82**	**46.18 ±4.11**	**46.95 ±2.61**	**48.17 ±4.52**	**36.43 ±1.79**	**61.03 ±4.39**	**51.77 ±3.59**	**37.58 ±2.50**	**61.99 ±3.62**	**47.25 ±1.71**

Table 6. Adaptation accuracy (%) comparison on DomainNet.

Method	Cl→Pa	Cl→Re	Cl→Sk	Pa→Cl	Pa→Re	Pa→Sk	Re→Cl	Re→Pa	Re→Sk	Sk→Cl	Sk→Pa	Sk→Re	Avg.
Direct	11.24 ±1.79	17.45 ±1.87	9.11 ±1.15	10.26 ±1.28	28.88 ±2.52	10.24 ±1.11	14.27 ±1.01	26.79 ±0.78	14.18 ±0.88	9.97 ±0.96	13.45 ±2.03	18.22 ±1.88	15.34 ±0.59
PCS	15.67 ±2.34	25.78 ±3.87	10.41 ±1.60	14.37 ±1.50	37.87 ±4.79	10.76 ±0.89	17.42 ±1.58	30.62 ±1.70	14.18 ±0.87	14.16 ±0.87	19.88 ±1.72	28.78 ±2.67	19.99 ±0.49
UCDIR	12.02 ±2.06	15.39 ±2.16	13.95 ±2.25	15.22 ±3.00	23.97 ±2.43	20.59 ±1.76	18.78 ±1.96	33.85 ±1.62	28.19 ±0.81	15.57 ±0.81	20.64 ±2.19	25.07 ±1.92	21.10 ±0.71
SDAT	8.53 ±1.11	14.36 ±1.39	7.36 ±0.58	10.28 ±1.52	23.51 ±2.58	8.30 ±1.18	14.54 ±0.91	22.72 ±1.78	11.53 ±1.14	9.44 ±0.88	9.40 ±1.46	13.95 ±1.01	12.83 ±0.56
CAF-A	10.35 ±3.00	20.97 ±3.47	8.28 ±0.98	12.15 ±1.67	30.89 ±4.64	9.97 ±2.09	20.00 ±2.59	26.48 ±2.73	15.70 ±1.97	10.29 ±0.73	11.00 ±1.14	19.11 ±1.96	16.27 ±1.17
CAF-D	10.61 ±2.77	16.18 ±2.86	6.37 ±1.08	11.24 ±1.17	25.30 ±3.14	7.92 ±1.64	15.58 ±3.08	25.69 ±3.71	11.55 ±2.98	9.30 ±0.90	10.64 ±1.16	15.96 ±1.27	13.86 ±1.39
ICON	8.78 ±2.73	13.98 ±1.49	6.24 ±1.53	9.34 ±1.95	25.62 ±2.81	6.80 ±1.70	15.27 ±1.25	21.36 ±3.40	9.66 ±2.10	7.70 ±1.90	6.51 ±1.41	11.18 ±1.04	11.87 ±0.68
ProtoOT	8.26 ±0.92	13.57 ±0.94	9.10 ±1.28	8.19 ±1.61	25.84 ±2.55	12.12 ±1.33	11.59 ±0.69	26.25 ±2.41	14.87 ±1.25	9.84 ±1.34	14.24 ±1.54	18.57 ±2.17	14.37 ±0.28
MUCDA	11.75 ±2.33	15.55 ±3.33	8.89 ±2.14	11.33 ±1.41	27.65 ±5.48	9.98 ±2.10	14.80 ±4.41	25.10 ±3.56	15.38 ±2.27	10.71 ±1.39	11.99 ±1.64	16.64 ±3.30	14.98 ±0.67
SHOT	10.17 ±2.44	16.49 ±3.71	7.74 ±1.22	11.44 ±1.00	29.66 ±3.35	9.29 ±1.67	16.60 ±1.45	27.62 ±1.76	14.59 ±0.80	8.53 ±1.04	12.35 ±1.66	18.60 ±2.70	15.26 ±0.67
NRC	10.32 ±1.73	16.59 ±2.18	10.16 ±1.01	14.41 ±0.97	25.42 ±5.64	12.95 ±0.92	20.67 ±0.41	26.80 ±1.30	19.04 ±0.62	11.36 ±1.49	13.07 ±1.90	17.68 ±1.79	16.54 ±0.86
AaD	10.31 ±1.73	16.55 ±2.19	10.16 ±1.01	14.41 ±0.97	27.36 ±2.68	12.93 ±0.93	20.69 ±0.38	26.74 ±1.31	19.00 ±0.59	11.40 ±1.53	13.03 ±1.92	17.58 ±1.80	16.68 ±0.57
SFDA-DE	10.42 ±1.55	17.81 ±2.23	11.27 ±1.19	16.26 ±0.41	30.72 ±2.51	14.88 ±1.07	23.11 ±0.78	30.67 ±0.66	22.36 ±0.96	12.02 ±1.30	12.84 ±1.73	17.91 ±1.37	18.35 ±0.47
BMD	11.29 ±2.28	21.31 ±3.28	10.63 ±0.99	16.39 ±0.83	34.91 ±3.38	14.05 ±1.43	22.77 ±0.98	29.52 ±1.37	20.35 ±0.66	13.24 ±1.85	14.93 ±2.73	21.83 ±3.47	19.27 ±0.66
GPUE	13.28 ±1.80	18.24 ±1.46	13.43 ±1.75	18.05 ±1.93	31.38 ±3.71	19.52 ±2.41	26.16 ±1.70	35.10 ±1.06	27.50 ±1.22	13.95 ±1.50	17.39 ±2.15	21.13 ±2.04	21.26 ±0.70
SFADA	5.76 ±1.11	11.34 ±1.72	5.10 ±1.00	8.46 ±0.98	18.37 ±2.69	6.42 ±1.50	13.07 ±0.89	18.77 ±1.67	8.41 ±1.03	7.19 ±1.59	8.67 ±2.92	12.90 ±2.33	10.37 ±0.43
TPDS	10.65 ±1.86	16.39 ±3.35	8.02 ±1.26	11.96 ±1.43	31.36 ±3.69	9.21 ±1.81	16.14 ±2.06	25.84 ±1.57	15.05 ±1.34	9.53 ±2.24	11.77 ±1.82	19.72 ±2.58	15.55 ±0.43
LCL	**20.44 ±3.61**	**28.42 ±2.92**	**15.78 ±2.82**	**23.38 ±1.58**	**44.21 ±5.43**	**21.77 ±1.91**	**31.29 ±0.62**	**44.24 ±1.82**	**28.19 ±1.41**	**20.12 ±2.37**	**25.84 ±4.44**	**28.16 ±2.45**	**27.65 ±0.63**

Table 7. Ablation study on Office-Home.

Method	Ar→Cl	Ar→Pr	Ar→Rw	Cl→Ar	Cl→Pr	Cl→Rw	Pr→Ar	Pr→Cl	Pr→Rw	Rw→Ar	Rw→Cl	Rw→Pr	Avg.
None (Direct)	18.25 ±0.31	30.42 ±2.81	36.65 ±2.49	24.72 ±2.56	27.02 ±2.32	29.11 ±1.69	31.13 ±1.04	22.81 ±1.04	44.76 ±1.94	35.58 ±3.72	23.00 ±2.61	41.90 ±2.19	30.45 ±1.01
+ $\mathcal{L}_{\text{PC}}$	24.91 ±1.89	39.52 ±1.73	42.89 ±2.49	30.75 ±2.91	32.45 ±3.45	35.06 ±3.17	35.69 ±1.48	28.87 ±0.95	50.83 ±2.38	41.65 ±3.81	30.85 ±2.27	52.07 ±2.46	37.13 ±1.26
+ $\mathcal{L}_{\text{ILC}}$	30.82 ±1.90	44.60 ±2.98	47.84 ±3.15	34.65 ±2.70	38.39 ±3.05	39.09 ±3.11	41.17 ±2.16	35.13 ±1.77	55.32 ±2.49	48.21 ±2.57	36.87 ±2.23	57.30 ±2.41	42.45 ±1.29
+ $\mathcal{L}_{\text{CLC}}$ (w/o. RM)	30.44 ±1.30	47.30 ±4.17	52.78 ±3.59	35.43 ±3.84	40.95 ±4.17	43.26 ±2.79	45.56 ±4.02	34.89 ±1.81	59.07 ±2.44	50.78 ±3.29	37.38 ±1.58	59.95 ±2.83	44.82 ±1.69
+ $\mathcal{L}_{\text{CLC}}$ (w. RM)	**32.55 ±3.41**	**49.57 ±3.85**	**55.91 ±4.89**	**38.88 ±2.82**	**46.18 ±4.11**	**46.95 ±2.61**	**48.17 ±4.52**	**36.43 ±1.79**	**61.03 ±4.39**	**51.77 ±3.59**	**37.58 ±2.50**	**61.99 ±3.62**	**47.25 ±1.71**