Supplementary Material for the Paper: Class Relationship Embedded Learning for Source-Free Unsupervised Domain Adaptation

Yixin Zhang[1,2] Zilei Wang[*2] Weinan He[2]
[1] Institute of Artificial Intelligence, Hefei Comprehensive National Science Center
[2] University of Science and Technology of China
{zhyx12, zlwang}@ustc.edu.cn, hwn2018@mail.ustc.edu.cn

Contents

A. More Related Work **1**
A.1. Class Relationship 1
A.2. Contrastive Learning 1

B. Experiments **2**
B.1. Training Details 2
B.2. Detailed Results Beyond Average Accuracy . 2
B.3. Other Ways of Utilizing $\mathbf{A}^s$ 2
B.4. Compatibility with Non-source-free DA . . . 3
B.5. About the Class Relationship Prior 3
B.6. Setting of λ_{inter}. 3

C. Limitations **3**

A. More Related Work

A.1. Class Relationship

Numerous UDA methods have explored the prediction space to enhance feature learning [6, 28, 38], while most of them only focus on the intra-class relationship. Some methods [18, 24] propose to utilize the inter-class term, but their primary objective is to suppress it and produce more confident predictions. It is implemented by directly minimizing the inter-class term [18] or resorting to adversarial training between two classifiers [24]. Different from them, we use the source prior class similarity as coefficient for the inter-class term instead of a fixed value. It is beneficial to accurately represent the similarity between two samples. The experimental results show that our method can surpass them even without source supervision during target training.

In SFUDA, class relationships can be implicitly learned through pseudo labeling [14, 25] or prediction consistency of neighborhoods [52–54]. CAiDA [8] focus on multi-source-free domain adaptation, and propose to preserve the consistency of the interclass relationship by aligning soft label distributions across domains. It relies on multiple source models and needs to construct prototype-level class distribution through pseudo label. Differently, we propose to embed the source class relationship in the sample-level similarity computation, which is a finer grained constraint. The results under multi-source-free DA show the superiority of our method even without domain labels.

A.2. Contrastive Learning

Contrastive Learning has shown remarkable advantages in self-supervised learning [4, 9, 11, 13, 30, 32]. The contrastive loss measures the similarity of representation pairs and attempts to distinguish between positive and negative pairs. MoCo [11] maintains a queue of previously processed embeddings as negative memory bank. SimCLR [4] shows that large batch size and strong data augmentations has a comparable performance to the memory-based approaches. We adopt a similar architecture to MoCo [11] to perform contrastive learning.

In Unsupervised Domain Adaptation (UDA), most approaches [16, 19, 55] use contrastive loss on the basis of class-wise prototypes with a sample selection strategy. Only a few methods [20, 37, 41] use the instance discrimination based contrastive loss. CDS [20] employs contrastive learning in a pre-train step before proceeding to a domain alignment stage. CLDA [41] suggests that the classifier can be used as a contrastive projection head [4]. ContrastMix [37] conducts temporal contrastive self-supervised learning over the graph representations. Among these methods, none of them explicitly consider embedding class relationship into similarity computation as our method.

DaC [57] proposes a *Divide and Contrast* paradigm where the target samples are divided into source-like and target-specific ones, and learned through contrastive loss. For the source-like samples, they conduct class-wise contrastive learning where positive samples are class centroids. For the target-specific samples, local structure is considered where positive samples are nearest neighbors and another view. They design a momentum updated memory bank to provides source-like centroids and target-specific features.

*Corresponding author

Our method is actually different from DaC in four aspects: 1) Our main idea is to utilize source class relationship, and DaC is meant to learn source-like and target-specific samples differently. 2) Our method is more concise, we do not need to maintain a memory bank storing all features. 3) Our contrastive loss is based on probabilities, while DaC is based on original features. 4) We do not consider the nearest samples as positive samples, but show the complementarity between our method and local neighbor based method AaD [54]. Finally, our method can achieve better performance than DaC, for example, ours 75.9% *vs*. DaC 72.8% in Office-Home and ours 89.6% *vs*. DaC 87.3% in VisDA.

B. Experiments

B.1. Training Details

For our proposed method, we use two kinds of augmentation to generate two different views, namely weak augmentation and strong augmentation. For the weak augmentation, we use random resize, random crop, and random horizontal flip. For strong augmentation, we further add RandAugment [7] by randomly sampling two augmentations from a transformation set, including color, brightness, contrast adjustments, rotation, polarization, etc.

We use the same network architecture as SHOT [25] *i.e.*, the final part of the network is: *fully connected layer - Batch Normalization [17] - fully connected layer with weight normalization [40]*. This is also used in our experiments in Transformer based backbone. We further adopt the same learning rate scheduler $\eta = \eta_0 \cdot (1 + 10 \cdot p)^{-0.75}$ as [10, 28], where p is the training progress changing from 0 to 1. The initial learning rate for Office-31 and Office-Home is set to 1e-3 for all layers, except for the last two newly added fc layers, where we apply 1e-2. Learning rates are set 10 times smaller for VisDA. For DomainNet, the initial learning rate is set to 5e-4. We train 40 epochs for Office-31 and Office-Home, 25 epochs for DomainNet, and 15 epochs for VisDA.

We fixed source pre-trained classifier following SHOT [25]. In the official implementation of AaD [54], the classifier is not fixed. We find that fixing the classifier in AaD will decrease some performance. However, when combining AaD with our method, fixing classifier or not does not effect the final performance. Thus, we fix classifier in all experiments.

Following the protocols in SHOT [25], the source domain consists of 25 classes (the first 25 in alphabet order) but the target domain contains 65 classes including unknown samples for an open-set scenario. For a partial-set scenario, the source domain consists of 65 classes, but the target domain contains the same 25 classes.

B.2. Detailed Results Beyond Average Accuracy

In some benchmarks, we only report the average performance due to the limitation of space. Here we give the detailed results. Table 3 shows the class-wise accuracy in VisDA under single-source unsupervised domain adaptation setting. Table 4 shows the results of partial-set and open-set domain adaptation on Office-Home. Table 5 shows the results of ViT-B backbone on Office-Home under single-source unsupervised domain adaptation setting.

B.3. Other Ways of Utilizing $\mathbf{A}^s$

Given the source class similarity matrix $\boldsymbol{A}^s$, an natural way of utilizing it is to obtain target class similarity matrix $\boldsymbol{A}^t$ by target prototypes and then enforce consistency between $\boldsymbol{A}^s$ and $\boldsymbol{A}^t$. To get $\boldsymbol{A}^t$, we first obtain the pseudo labels following SHOT [25]. For the target prototype, we tried batch-level or EMA-updated global-level prototypes. There are no obvious differences in performance. The loss can be presented as

$$\ell_{con} = ||\boldsymbol{A}^s - \boldsymbol{A}^t||_F^2 \tag{1}$$

where $||M||_F$ represents the Frobenius norm of M.

Another way is to embed $\boldsymbol{A}^s$ into the cross-entropy loss of FixMatch [42] as weights. Specifically, the weighted FixMatch loss is presented as

$$\ell_{weighted\_fm}(\boldsymbol{p}) = -\sum_{i=1}^{C} \boldsymbol{A}^s_{l^*,i} log(p_i) \tag{2}$$

where $\boldsymbol{p}$ is the predicted probability of strongly augmented view. C is the number of classes. l^* is the pseudo label class obtained from clustering based strategy in SHOT [25].

The result is shown in Table 1. It can be seen that consistency between $\boldsymbol{A}^s$ and $\boldsymbol{A}^t$ can bring improvements over SHOT-IM baseline, but is not as effective as FixMatch. Using weighted FixMatch by $\boldsymbol{A}^s$ can bring very limited improvement. This validates the importance of our proposed contrastive loss, which can fully exploit the source class similarity.

Table 1. Comparison with other ways using $\boldsymbol{A}^s$ in a class-wise manner. The experiments are conducted on Office-Home under SUDA setting.

Method	Avg
SHOT-IM	70.5
+ FixMatch	72.9
+ Prototype Similarity Consistency	72.3
+ Weighted FixMatch	73.1
+CR-CACo	74.8

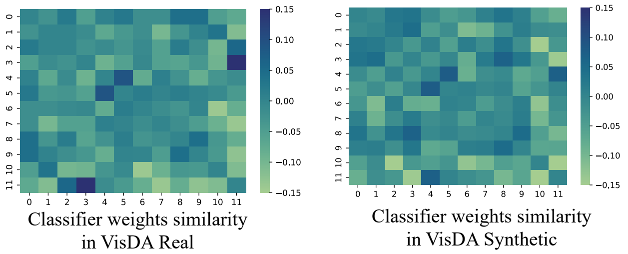

Figure 1. Classifier weights similarity of different domains. The diagonal elements with original value 1 are changed to 0 for better visualization.

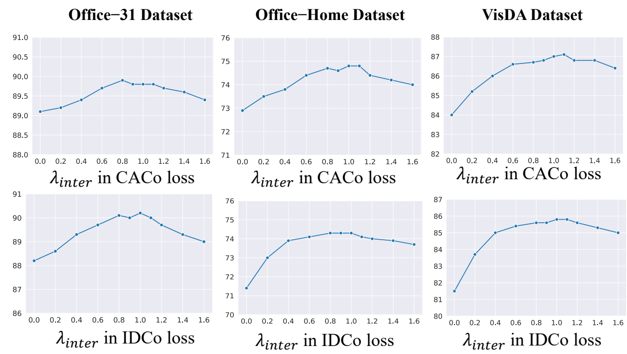

Figure 2. λ_{inter} in different losses and datasets.

B.4. Compatibility with Non-source-free DA

We further evaluate the effectiveness of our method in non-source-free DA. As described before, our method should be combined with other methods that can produce confident samples. Here we choose the output space based domain adversarial methods [28, 47]. The distribution across domains are encouraged to be aligned through domain adversarial training, and the target samples will become more confident as the source samples.

For Office-Home dataset, we use ResNet-50 backbone, and for VisDA dataset, we use ResNet-101 backbone. Instead of using clustering based pseudo label strategy as in SHOT [25], we directly generate pseudo label from the prediction of weakly augmented views for simplicity.

Unlike SFUDA where the $\boldsymbol{A}^s$ is fixed after given the source pre-trained model, here we use the learnable classifier weights to generate $\boldsymbol{A}^s$. This is reasonable since the learning rate of classifier is ten times larger than feature extractor, and the class similarity can be quickly learned.

Results are shown in Table 2. It can be seen that our method can bring consistent improvements.

Table 2. **Compatibility with non-source-free DA** works on Office-Home and VisDA under SUDA setting.

Method	Office-Home	VisDA
CDAN [28]	65.8	73.9
CDAN + ours	**73.0** **(+7.2)**	**82.8** **(+8.9)**
ToAlign [47]	72.0	80.1
ToAlign + ours	**75.1** **(+3.1)**	**88.7** **(+8.6)**

B.5. About the Class Relationship Prior

Besides intuitive rationality, we visualized the cosine similarity matrix of classifier weights trained by the data from different domains. We found that similar classes in one domain are more likely to be similar in other domains. In particular, we provide the results of VisDA in Figure 1 here.

B.6. Setting of λ_{inter}.

$\lambda_{inter} = 1.0$ is always set in both losses and all datasets (L590-591). Here we provide the results of adjusting λ_{inter} in Figure 2. It can be seen that for two losses and different datasets, our method is stable for $\lambda_{inter} \in [0.6, 1.2]$.

C. Limitations

The limitations of our proposed approach are summarized as follows:

- Our proposed CR-CACo loss must be used in conjunction with methods that can generate high-confidence samples. And our method should be combined with other methods to achieve better performance.

- While in some datasets (*i.e.* VisDA with 10 classes) which have enough samples and are relatively simple, our method can achieve performance that is almost as good as Oracle's, it still has a larger gap with Oracle's performance in other more difficult datasets (*i.e.* DomainNet).

References

[1] Abhijit Bendale and Terrance E Boult. Towards open set deep networks. In *CVPR*, 2016. 4

[2] Zhangjie Cao, Mingsheng Long, Jianmin Wang, and Michael I Jordan. Partial transfer learning with selective adversarial networks. In *CVPR*, 2018. 4

[3] Zhangjie Cao, Kaichao You, Mingsheng Long, Jianmin Wang, and Qiang Yang. Learning to transfer examples for partial domain adaptation. In *CVPR*, 2019. 4

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

[5] Tong Chu, Yahao Liu, Jinhong Deng, Wen Li, and Lixin Duan. Denoised maximum classifier discrepancy for source-free unsupervised domain adaptation. In *AAAI*, 2022. 4

Table 3. **Single-Source Unsupervised Domain Adaptation (SUDA)** on VisDA benchmarks. * indicates results from released code. **(+x.x)** indicates absolute improvements over SHOT-IM and AaD respectively.

Method	SF	plane	bcycl	bus	car	horse	knife	mcycl	person	plant	sktbrd	train	truck	Avg
BCDM (AAAI'20) [24]	✗	95.1	87.6	81.2	73.2	92.7	95.4	86.9	82.5	95.1	84.8	88.1	39.5	83.4
MCC (ECCV'20) [18]	✗	88.7	80.3	80.5	71.5	90.1	93.2	85.0	71.6	89.4	73.8	85.0	36.9	78.8
FixBi (CVPR'21) [31]	✗	96.1	87.8	**90.5**	90.3	96.8	95.3	92.8	**88.7**	97.2	94.2	90.9	25.7	87.2
FAA (ICCV'21) [15]	✗	91.6	80.5	81.5	70.7	89.6	81.0	87.5	79.9	87.1	86.4	81.0	75.1	82.7
SHOT (ICML'20) [25]	✓	94.3	88.5	80.1	57.3	93.1	94.9	80.7	80.3	91.5	89.1	86.3	58.2	82.9
CPGA (IJCAI'21) [34]	✓	94.8	83.6	79.7	65.1	92.5	94.7	90.1	82.4	88.8	88.0	88.9	60.1	84.1
VDM-DA (TCSVT'21) [44]	✓	96.9	90.0	80.0	64.4	96.8	96.4	86.7	83.3	96.2	87.9	89.8	54.7	85.3
A^2Net (ICCV'21) [48]	✓	94.0	87.8	85.6	66.8	93.7	95.1	85.8	81.2	91.6	88.2	86.5	56.0	84.3
HCL(NeurIPS'21) [14]	✓	93.3	85.4	80.7	68.5	91.0	88.1	86.0	78.6	86.6	88.8	80.0	**74.7**	83.5
NRC (NeurIPS'21) [52]	✓	96.8	**91.3**	82.4	62.4	96.2	95.9	86.1	80.6	94.8	94.1	90.4	59.7	85.9
SHOT++ (TPAMI'21) [26]	✓	97.7	88.4	90.2	86.3	**97.9**	**98.6**	92.9	84.1	97.1	92.2	**93.6**	28.8	87.3
D-MCD (AAAI'22) [5]	✓	97.0	88.0	90.0	81.5	95.6	98.0	86.2	**88.7**	94.6	92.7	83.7	53.1	87.5
DIPE (CVPR'22) [45]	✓	95.2	87.6	78.8	55.9	93.9	95.0	84.1	81.7	92.1	88.9	85.4	58.0	83.1
Sub-Sup (ECCV'22) [21]	✓	-	-	-	-	-	-	-	-	-	-	-	-	88.2
BMD (ECCV'22) [35]	✓	96.9	87.8	90.1	**91.3**	97.8	97.8	90.6	84.4	96.9	94.3	90.9	45.9	88.7
CoWA-JMDS (ICML'22) [23]	✓	96.1	87.8	**90.5**	90.3	96.8	95.3	92.8	**88.7**	97.2	94.2	90.9	25.7	87.2
Feat-Mixup (ICML'22) [22]	✓	-	-	-	-	-	-	-	-	-	-	-	-	87.8
DaC (NeurIPS'22) [57]	✓	96.6	86.8	86.4	78.4	96.4	96.2	93.6	83.8	96.8	95.1	89.6	50.0	87.3
SHOT-IM (ICML'20) [25]	✓	93.7	86.4	78.7	50.7	91.0	93.5	79.0	78.3	89.2	85.4	87.9	51.1	80.4
Ours + SHOT-IM	✓	**98.1**	91.0	86.9	84.8	97.5	96.5	91.4	83.1	97.0	**95.8**	92.8	54.1	89.1 **(+8.7)**
AaD (NeurIPS) [54]	✓	97.4	90.5	80.8	76.2	97.3	96.1	89.8	82.9	95.5	93.0	92.0	64.7	88.0
Ours + AaD	✓	**98.1**	90.3	87.2	88.7	97.6	96.3	**93.7**	84.5	**97.5**	95.3	91.3	55.0	**89.6** **(+1.6)**

Table 4. **Partial-set** and **open-set Domain Adaptation** (**PDA** and **ODA**) on Office-Home. * indicates results from released code. **(+x.x)** indicates absolute improvements over SHOT-IM and AaD respectively.

Partial-set DA	SF	Ar→Cl	Ar→Pr	Ar→Re	Cl→Ar	Cl→Pr	Cl→Re	Pr→Ar	Pr→Cl	Pr→Re	Re→Ar	Re→Cl	Re→Pr	Avg.
ResNet-50 [12]	✗	46.3	67.5	75.9	59.1	59.9	62.7	58.2	41.8	74.9	67.4	48.2	74.2	61.3
IWAN (CVPR'18) [56]	✗	53.9	54.5	78.1	61.3	48.0	63.3	54.2	52.0	81.3	76.5	56.8	82.9	63.6
SAN (CVPR'18) [2]	✗	44.4	68.7	74.6	67.5	65.0	77.8	59.8	44.7	80.1	72.2	50.2	78.7	65.3
ETN (CVPR'19) [3]	✗	59.2	77.0	79.5	62.9	65.7	75.0	68.3	55.4	84.4	75.7	57.7	84.5	70.5
SAFN (ICCV'19) [49]	✗	58.9	76.3	81.4	70.4	73.0	77.8	72.4	55.3	80.4	75.8	60.4	79.9	71.8
Source model only	✓	45.2	70.4	81.0	56.2	60.8	66.2	60.9	40.1	76.2	70.8	48.5	77.3	62.8
SHOT (ICML'20) [25]	✓	64.8	85.2	92.7	76.3	77.6	88.8	79.7	64.3	89.5	80.6	66.4	85.8	79.3
SHOT+HCL (NeurIPS'21) [14]	✓	66.9	**85.5**	92.5	78.3	77.2	87.1	78.3	65.1	90.7	82.4	68.7	88.4	80.1
CoWA-JMDS (ICML'22) [23]	✓	69.6	93.2	92.3	78.9	81.3	92.1	79.8	71.7	90.0	**83.8**	72.2	**93.7**	**83.2**
SHOT-IM (ICML'20) [25]	✓	57.9	83.6	88.8	72.4	74.0	79.0	76.1	60.6	90.1	81.9	68.3	88.5	76.8
Ours +SHOT-IM	✓	68.6	85.1	90.9	80.1	79.4	86.3	79.2	66.1	90.5	82.2	69.5	89.3	80.6 (+3.8)
AaD* (NeurIPS'22) [54]	✓	67.0	83.5	93.1	80.5	76.0	87.6	78.1	65.6	90.2	83.5	64.3	87.3	79.7
Ours +AaD	✓	69.0	85.5	**93.2**	**83.3**	**82.2**	90.2	**82.1**	66.8	**91.9**	83.7	**69.7**	90.7	82.4 (+2.7)
Open-set DA	**SF**	**Ar→Cl**	**Ar→Pr**	**Ar→Re**	**Cl→Ar**	**Cl→Pr**	**Cl→Re**	**Pr→Ar**	**Pr→Cl**	**Pr→Re**	**Re→Ar**	**Re→Cl**	**Re→Pr**	**Avg.**
ResNet [12]	✗	53.4	52.7	51.9	69.3	61.8	74.1	61.4	64.0	70.0	78.7	71.0	74.9	65.3
ATI-λ (ICCV'17) [33]	✗	55.2	52.6	53.5	69.1	63.5	74.1	61.7	64.5	70.7	79.2	72.9	75.8	66.1
OSBP (ECCV'18) [39]	✗	56.7	51.5	49.2	67.5	65.5	74.0	62.5	64.8	69.3	80.6	74.7	71.5	65.7
OpenMax (CVPR'16) [1]	✗	56.5	52.9	53.7	69.1	64.8	74.5	64.1	64.0	71.2	80.3	73.0	76.9	66.7
STA (CVPR'19) [27]	✗	58.1	53.1	54.4	**71.6**	69.3	81.9	63.4	**65.2**	74.9	**85.0**	**75.8**	80.8	69.5
Source model only	✓	36.3	54.8	69.1	33.8	44.4	49.2	36.8	29.2	56.8	51.4	35.1	62.3	46.6
SHOT (ICML'22) [25]	✓	64.5	**80.4**	84.7	63.1	75.4	81.2	65.3	59.3	83.3	69.6	64.6	82.3	72.8
SHOT+HCL (NeurIPS'21) [14]	✓	64.2	78.3	83.0	61.1	72.2	79.6	65.5	59.3	80.6	80.1	72.0	82.8	73.2
U-SFAN (ECCV'22) [36]	✓	62.9	77.9	84.0	67.9	74.6	79.6	68.8	61.3	83.3	76.0	63.9	82.3	73.5
CoWA-JMDS (ICML'22) [23]	✓	63.3	79.2	**85.4**	67.6	**83.6**	**82.0**	66.9	56.9	81.1	68.5	57.9	**85.9**	73.2
SHOT-IM (ICML'20) [25]	✓	62.5	77.8	83.9	60.9	73.4	79.4	64.7	58.7	83.1	69.1	62.0	82.1	71.5
Ours +SHOT-IM	✓	65.2	76.6	80.2	66.2	75.3	77.8	**70.4**	61.8	79.3	71.1	61.1	78.3	73.2(+1.7)
AaD* (NeurIPS'22) [54]	✓	63.7	77.3	80.4	66.0	72.6	77.6	69.1	62.5	79.8	71.8	62.3	78.6	71.8
Ours +AaD	✓	**65.8**	79.9	84.9	65.5	75.6	77.6	68.2	60.9	**83.5**	71.0	64.2	82.3	**73.3**(+1.5)

[6] Safa Cicek and Stefano Soatto. Unsupervised domain adaptation via regularized conditional alignment. In *ICCV*, 2019. 1

[7] Ekin D Cubuk, Barret Zoph, Jonathon Shlens, and Quoc V Le. Randaugment: Practical automated data augmentation with a reduced search space. In *CVPR Workshops*, 2020. 2

[8] Jiahua Dong, Zhen Fang, Anjin Liu, Gan Sun, and Tongliang Liu. Confident anchor-induced multi-source free domain

Table 5. **Single-Source Unsupervised DA (SUDA)** on Office-Home with ViT-B backbone.

Method	SF	Office-Home												
		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
ViT-B	✗	54.68	83.04	87.15	77.30	83.42	85.54	74.41	50.90	87.22	79.56	53.79	88.80	75.48
CDTrans (ICLR'22) [50]	✗	68.8	85.0	86.9	81.5	87.1	87.3	79.6	63.3	88.2	82.0	66.0	90.6	80.5
TVT (WACV'23) [51]	✗	74.89	86.82	89.47	82.78	87.95	88.27	79.81	71.94	90.13	85.46	74.62	90.56	83.56
DOT-B (ACMMM'22) [29]	✗	73.1	89.1	90.1	85.5	89.4	89.6	83.2	72.1	90.4	84.4	72.9	91.5	84.3
SSRT (CVPR'22) [43]	✗	75.17	88.98	91.09	85.13	88.29	89.95	85.04	74.23	91.26	85.70	78.58	91.78	85.43
BCAT-DTF (Arxiv'22) [46]	✗	75.3	90.0	92.9	88.6	90.3	92.7	87.4	73.7	92.5	86.7	75.4	93.5	86.6
Source model	✓	61.7	84.9	87.9	79.5	85.9	86.5	77.5	58.1	88.0	81.1	59.4	88.4	78.2
SHOT-IM	✓	73.6	88.6	90.8	84.4	89.9	88.8	82.4	70	90	85.4	72.5	91.9	84.0
SHOT-IM+CR-CACo	✓	78.2	89.3	92.2	87.8	92.0	90.9	86.9	75.9	91.8	88.0	78.8	92.4	87.0(+3.0)
SHOT-IM+Ours	✓	79.9	90.3	92.9	88.6	92.7	91.5	88.3	77.6	92.6	89.5	81.9	93.6	**88.3**(+4.3)

adaptation. In *NeurIPS*, 2021. 1

[9] Debidatta Dwibedi, Yusuf Aytar, Jonathan Tompson, Pierre Sermanet, and Andrew Zisserman. With a Little Help From My Friends: Nearest-Neighbor Contrastive Learning of Visual Representations. In *ICCV*, 2021. 1

[10] Yaroslav Ganin and Victor Lempitsky. Unsupervised domain adaptation by backpropagation. In *ICML*, 2015. 2

[11] Kaiming He, Haoqi Fan, Yuxin Wu, Saining Xie, and Ross Girshick. Momentum contrast for unsupervised visual representation learning. In *CVPR*, 2020. 1

[12] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *CVPR*, 2016. 4

[13] R Devon Hjelm, Alex Fedorov, Samuel Lavoie-Marchildon, Karan Grewal, Phil Bachman, Adam Trischler, and Yoshua Bengio. Learning deep representations by mutual information estimation and maximization. In *ICLR*, 2019. 1

[14] Jiaxing Huang, Dayan Guan, Aoran Xiao, and Shijian Lu. Model adaptation: Historical contrastive learning for unsupervised domain adaptation without source data. In *NeurIPS*, 2021. 1, 4

[15] Jiaxing Huang, Dayan Guan, Aoran Xiao, and Shijian Lu. RDA: Robust domain adaptation via fourier adversarial attacking. In *ICCV*, 2021. 4

[16] Jiaxing Huang, Dayan Guan, Aoran Xiao, Shijian Lu, and Ling Shao. Category contrast for unsupervised domain adaptation in visual tasks. In *CVPR*, 2022. 1

[17] Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *ICML*, 2015. 2

[18] Ying Jin, Ximei Wang, Mingsheng Long, and Jianmin Wang. Minimum class confusion for versatile domain adaptation. In *ECCV*, 2020. 1, 4

[19] Guoliang Kang, Lu Jiang, Yi Yang, and Alexander G Hauptmann. Contrastive adaptation network for unsupervised domain adaptation. In *CVPR*, 2019. 1

[20] Donghyun Kim, Kuniaki Saito, Tae-Hyun Oh, Bryan A Plummer, Stan Sclaroff, and Kate Saenko. Cds: Cross-domain self-supervised pre-training. In *ICCV*, 2021. 1

[21] Jogendra Nath Kundu, Suvaansh Bhambri, Akshay Kulkarni, Hiran Sarkar, Varun Jampani, and R Venkatesh Babu. Concurrent subsidiary supervision for unsupervised source-free domain adaptation. In *ECCV*, 2022. 4

[22] Jogendra Nath Kundu, Akshay R Kulkarni, Suvaansh Bhambri, Deepesh Mehta, Shreyas Anand Kulkarni, Varun Jampani, and Venkatesh Babu Radhakrishnan. Balancing discriminability and transferability for source-free domain adaptation. In *ICML*, 2022. 4

[23] Jonghyun Lee, Dahuin Jung, Junho Yim, and Sungroh Yoon. Confidence score for source-free unsupervised domain adaptation. In *ICML*, 2022. 4

[24] Shuang Li, Fangrui Lv, Binhui Xie, Chi Harold Liu, Jian Liang, and Chen Qin. Bi-classifier determinacy maximization for unsupervised domain adaptation. In *AAAI*, 2021. 1, 4

[25] Jian Liang, Dapeng Hu, and Jiashi Feng. Do we really need to access the source data? source hypothesis transfer for unsupervised domain adaptation. In *ICML*, 2020. 1, 2, 3, 4

[26] Jian Liang, Dapeng Hu, Yunbo Wang, Ran He, and Jiashi Feng. Source data-absent unsupervised domain adaptation through hypothesis transfer and labeling transfer. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021. 4

[27] Hong Liu, Zhangjie Cao, Mingsheng Long, Jianmin Wang, and Qiang Yang. Separate to adapt: Open set domain adaptation via progressive separation. In *CVPR*, 2019. 4

[28] Mingsheng Long, Zhangjie Cao, Jianmin Wang, and Michael I Jordan. Conditional adversarial domain adaptation. In *NeurIPS*, 2018. 1, 2, 3

[29] Wenxuan Ma, Jinming Zhang, Shuang Li, Chi Harold Liu, Yulin Wang, and Wei Li. Making the best of both worlds: A domain-oriented transformer for unsupervised domain adaptation. In *ACMMM*, 2022. 5

[30] Ishan Misra and Laurens van der Maaten. Self-supervised learning of pretext-invariant representations. In *CVPR*, 2020. 1

[31] Jaemin Na, Heechul Jung, Hyung Jin Chang, and Wonjun Hwang. FixBi: Bridging domain spaces for unsupervised domain adaptation. In *CVPR*, 2021. 4

[32] Aaron van den Oord, Yazhe Li, and Oriol Vinyals. Representation learning with contrastive predictive coding. *arXiv preprint arXiv:1807.03748*, 2018. 1

[33] Pau Panareda Busto and Juergen Gall. Open set domain adaptation. In *ICCV*, 2017. 4

[34] Zhen Qiu, Yifan Zhang, Hongbin Lin, Shuaicheng Niu, Yanxia Liu, Qing Du, and Mingkui Tan. Source-free domain

adaptation via avatar prototype generation and adaptation. In *IJCAI*, 2021. 4

[35] Sanqing Qu, Guang Chen, Jing Zhang, Zhijun Li, Wei He, and Dacheng Tao. Bmd: A general class-balanced multicentric dynamic prototype strategy for source-free domain adaptation. In *ECCV*, 2022. 4

[36] Subhankar Roy, Martin Trapp, Andrea Pilzer, Juho Kannala, Nicu Sebe, Elisa Ricci, and Arno Solin. Uncertainty-guided source-free domain adaptation. In *ECCV*, 2022. 4

[37] Aadarsh Sahoo, Rutav Shah, Rameswar Panda, Kate Saenko, and Abir Das. Contrast and mix: Temporal contrastive video domain adaptation with background mixing. In *NeurIPS*, 2021. 1

[38] Kuniaki Saito, Kohei Watanabe, Yoshitaka Ushiku, and Tatsuya Harada. Maximum classifier discrepancy for unsupervised domain adaptation. In *CVPR*, 2018. 1

[39] Kuniaki Saito, Shohei Yamamoto, Yoshitaka Ushiku, and Tatsuya Harada. Open set domain adaptation by backpropagation. In *ECCV*, 2018. 4

[40] Tim Salimans and Durk P Kingma. Weight normalization: A simple reparameterization to accelerate training of deep neural networks. In *NeurIPS*, 2016. 2

[41] Ankit Singh. Clda: Contrastive learning for semi-supervised domain adaptation. In *NeurIPS*, 2021. 1

[42] Kihyuk Sohn, David Berthelot, Chun-Liang Li, Zizhao Zhang, Nicholas Carlini, Ekin D Cubuk, Alex Kurakin, Han Zhang, and Colin Raffel. Fixmatch: Simplifying semi-supervised learning with consistency and confidence. In *NeurIPS*, 2020. 2

[43] Tao Sun, Cheng Lu, Tianshuo Zhang, and Haibin Ling. Safe self-refinement for transformer-based domain adaptation. In *CVPR*, 2022. 5

[44] Jiayi Tian, Jing Zhang, Wen Li, and Dong Xu. VDM-DA: Virtual domain modeling for source data-free domain adaptation. *IEEE Transactions on Circuits and Systems for Video Technology*, 2021. 4

[45] Fan Wang, Zhongyi Han, Yongshun Gong, and Yilong Yin. Exploring domain-invariant parameters for source free domain adaptation. In *CVPR*, 2022. 4

[46] Xiyu Wang, Pengxin Guo, and Yu Zhang. Domain adaptation via bidirectional cross-attention transformer. *arXiv preprint arXiv:2201.05887*, 2022. 5

[47] Guoqiang Wei, Cuiling Lan, Wenjun Zeng, Zhizheng Zhang, and Zhibo Chen. Toalign: Task-oriented alignment for unsupervised domain adaptation. In *NeurIPS*, 2021. 3

[48] Haifeng Xia, Handong Zhao, and Zhengming Ding. Adaptive adversarial network for source-free domain adaptation. In *ICCV*, 2021. 4

[49] Ruijia Xu, Guanbin Li, Jihan Yang, and Liang Lin. Larger norm more transferable: An adaptive feature norm approach for unsupervised domain adaptation. In *ICCV*, 2019. 4

[50] Tongkun Xu, Weihua Chen, WANG Pichao, Fan Wang, Hao Li, and Rong Jin. Cdtrans: Cross-domain transformer for unsupervised domain adaptation. In *International Conference on Learning Representations*, 2021. 5

[51] Jinyu Yang, Jingjing Liu, Ning Xu, and Junzhou Huang. Tvt: Transferable vision transformer for unsupervised domain adaptation. In *WACV*, 2023. 5

[52] Shiqi Yang, Joost van de Weijer, Luis Herranz, Shangling Jui, et al. Exploiting the intrinsic neighborhood structure for source-free domain adaptation. In *NeurIPS*, 2021. 1, 4

[53] Shiqi Yang, Yaxing Wang, Joost van de Weijer, Luis Herranz, and Shangling Jui. Generalized source-free domain adaptation. In *ICCV*, 2021. 1

[54] Shiqi Yang, Yaxing Wang, Kai Wang, Shangling Jui, and Joost van de Weijer. Attracting and dispersing: A simple approach for source-free domain adaptation. In *NeurIPS*, 2022. 1, 2, 4

[55] Xiangyu Yue, Zangwei Zheng, Shanghang Zhang, Yang Gao, Trevor Darrell, Kurt Keutzer, and Alberto Sangiovanni Vincentelli. Prototypical cross-domain self-supervised learning for few-shot unsupervised domain adaptation. In *CVPR*, 2021. 1

[56] Jing Zhang, Zewei Ding, Wanqing Li, and Philip Ogunbona. Importance weighted adversarial nets for partial domain adaptation. In *CVPR*, 2018. 4

[57] Ziyi Zhang, Weikai Chen, Hui Cheng, Zhen Li, Siyuan Li, Liang Lin, and Guanbin Li. Divide and contrast: Source-free domain adaptation via adaptive contrastive learning. In *NeurIPS*, 2022. 1, 4