BMD: A General Class-balanced Multicentric Dynamic Prototype Strategy for Source-free Domain Adaptation

Sanqing Qu[1], Guang Chen[1(✉)], Jing Zhang[2], Zhijun Li[3],
Wei He[4], and Dacheng Tao[2,5]

[1] Tongji University {2011444, guangchen}@tongji.edu.cn
[2] The University of Sydney jing.zhang1@sydney.edu.au
[3] University of Science and Technology of China zjli@ieee.org
[4] University of Science and Technology Beijing weihe@ieee.org
[5] JD Explore Academy dacheng.tao@gmail.com

Abstract. Source-free Domain Adaptation (SFDA) aims to adapt a pre-trained source model to the unlabeled target domain without accessing the well-labeled source data, which is a much more practical setting due to the data privacy, security, and transmission issues. To make up for the absence of source data, most existing methods introduced feature prototype based pseudo-labeling strategies to realize self-training model adaptation. However, feature prototypes are obtained by instance-level predictions based feature clustering, which is category-biased and tends to result in noisy labels since the visual domain gaps between source and target are usually different between categories. In addition, we found that a monocentric feature prototype may be ineffective to represent each category and introduce negative transfer, especially for those hard-transfer data. To address these issues, we propose a general class-**B**alanced **M**ulticentric **D**ynamic prototype (BMD) strategy for the SFDA task. Specifically, for each target category, we first introduce a global inter-class balanced sampling strategy to aggregate potential representative target samples. Then, we design an intra-class multicentric clustering strategy to achieve more robust and representative prototypes generation. In contrast to existing strategies that update the pseudo label at a fixed training period, we further introduce a dynamic pseudo labeling strategy to incorporate network update information during model adaptation. Extensive experiments show that the proposed model-agnostic BMD strategy significantly improves representative SFDA methods to yield new state-of-the-art results. The code is available at https://github.com/ispc-lab/BMD.

Keywords: Domain Adaptation, Source-free, Class-balanced Sampling, Multicentric Prototype Pseudo-labeling

1 Introduction

Deep neural networks have achieved remarkable success in various visual tasks at the expense of massive data collections and annotations [10,17,12,57,4] but
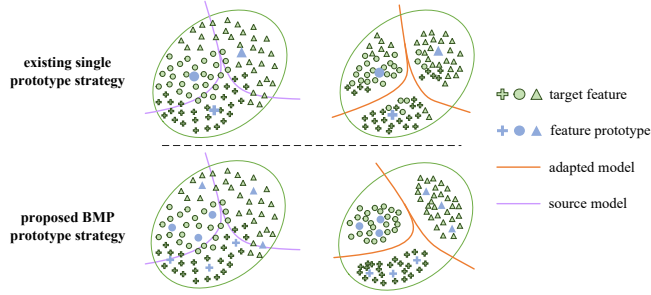
Fig. 1. Comparison between existing prototype strategy (left) and our BMD prototype strategy (right). During SFDA model adaptation, existing prototype strategies are monocentric and class-biased, which often lead to negative transfers for those hard-transfer instances, while our class-balanced multicentric dynamic prototype (BMD) strategy can effectively address the issue.

still often generalized poorly to the unseen new domains due to the inter-domain discrepancy. To reduce the annotation burden when dealing with new domain data, unsupervised domain adaptation (UDA) methods have been developed by aligning the well-labeled source data and the unlabeled target data distribution, which have achieved promising results in object recognition [52,13,46,63], object detection [5,55,37], and semantic segmentation [64,65,61,60,14].

However, most existing UDA methods require to access the source and target data simultaneously during model adaptation, which is often impractical due to the concerns about data privacy, data security and data transmission efficiency. Therefore, current frontiers have emerged a few works [24,44,29,26,27,59,58,1] seeking to realize source-free domain adaptation (SFDA), where only a source pretrained model is available. To make up for the absence of source data, existing methods can be divided into two main categories: GAN based methods [24,44,29] and self-training based methods [26,27,59,58,1]. For those self-training methods, pseudo-labeling strategies based on feature prototypes are popular and offer promising results. [26,27,1] introduce a weighted k-means clustering based feature prototype generation strategy. However, existing strategies are implemented with instance-level prediction results, which are category-biased and tend to introducing noisy labels, since the visual (e.g. scale, appearance, etc) domain gaps between source and target are usually different between categories [64,65]. In addition, we argue that due to the domain gap, a rough monocentric feature prototype for each category could not effectively represent the target data and would introduce negative transfer, especially for those hard-transfer data.

In light of the above issues, in this paper, we focus on existing self-training based SFDA methods and propose a general class-**B**alanced **M**ulticentric **D**ynamic (BMD) prototype strategy. Specifically, to avoid the gradual dominance of easy-transfer classes on prototype generation, for each target category we first introduce a novel inter-class balanced sampling strategy to aggregate the potential and representative data samples. Even though we can obtain category balanced

feature prototype with above strategy, it is still inferior for those hard transfer data samples. Therefore, we then introduce an intra-class multicentric clustering strategy to generate multiple feature prototypes for each category to assign more robust and precise pseudo labels. In addition, we conjecture that existing strategies that update the pseudo label bank at a fixed training period, may not effectively exploit the dynamic information of network optimization. Thus, we further introduce a dynamic pseudo-labeling strategy to incorporate network update information during model adaptation. We compare our BMD strategy with existing methods in Fig. 1.

To evaluate the effectiveness and generality of our model agnostic strategy, we have applied our strategy to four existing representative methods (SHOT [26], SHOT++ [27], G-SFDA [59] and NRC [58]). Extensive experiments on four benchmark datasets (VisDA-C [40], Office-Home [53], Office-31 [45] and PointDA-10 [43]) show that our BMD strategy significantly improves these methods to yield new state-of-the-art performance.

Our contribution can be summarized as follows:

- We propose a general class-balanced multicentric dynamic prototype strategy BMD for SFDA tasks that is model-agnostic and can be applied to existing self-training based SFDA methods.
- To avoid the gradual dominance of easy-transfer classes on prototype generation, we propose a novel inter-class balanced sampling strategy to aggregate potential and representative data samples.
- To reduce the noisy labels for those hard-transfer data samples, we introduce an intra-class multicentric prototype strategy for each category to assign more robust and precise pseudo labels.
- We conducted extensive experiments to evaluate the effectiveness of our BMD strategy. The results show that the proposed strategy can significantly boost existing methods, e.g., improving SHOT [26] from 82.9% to 85.8% on VisDA-C and NRC [58] from 52.6% to 57.0% on PointDA-10.

2   Related Work

2.1   Unsupervised Domain Adaptation

In pursuit of transfering knowledge from a different but well-labeled source dataset to an unlabeled but relevant target dataset, unsupervised domain adaptation (UDA) has received considerable interests in recent years. Existing methods can be broadly classified into three categories: discrepancy based, reconstruction based, and adversarial based. Discrepancy based methods usually introduce a divergence criterion to measure the distance between the source and target data distributions, and then achieve model adaptation by minimizing the corresponding criterion, e.g. the maximum mean discrepancy (MMD) [31], the wasserstein metric [6], and the contrastive domain discrepancy [21]. Reconstruction based methods [15,2,35] typically introduce an auxiliary image reconstruction task that guides the network to extract domain-invariant features for model adaptation.

Inspired by GAN, there are also approaches [13,30,47] that introduce domain discriminators to learn domain-invariant features in an adversarial manner. Despite of effectiveness, these methods require access to the source data, which is often impractical due to data privacy or security concerns.

2.2   Source-free Domain Adaptation

Currently, there have been several works [24,44,29,26,27,59,58,1] attempting to realize source-free domain adaptation, where only a pre-trained source model and unlabeled target data are available. In these approaches, [24,44,29] introduce generative networks to generate pseudo-data similar to sources or targets, which are difficult and inefficient. Instead, self-training based on feature prototype could be a promising direction [26,27,49,1]. The most relevant papers to our BMD are SHOT [26] and SHOT++ [27], which introduce a weighted k-means clustering algorithm to generate feature prototype and then assign pseudo-labels based on prototype matching. However, their feature prototype generation process is category-biased and prone to introducing negative transfer for those hard-transfer data. In contrast, our BMD can obtain more robust and stable feature prototypes by introducing inter-class balanced sampling and intra-class multi-centric prototypes generation. Our strategy may share some similarities in other fields approaches [16,41]. However, our methods are fundamentally different. Unlike these works which apply the labeled data to train a network to recognize the multimodal classes by introducing the multicentric prototypes, we explore the multicentric idea to assign preciser pseudo labels for those unlabeled data, especially those hard-transfer data, to achieve source-free model adaptation.

2.3   Imbalanced Learning

Massive studies [62] have been proposed for long-tail vision tasks to maintain diversity and balance predictions for minority categories. One straightforward idea is to perform category-balanced sampling [48,20,54] using prior knowledge about the category distributions. However, these strategies are not applicable to the UDA task because we do not have access to knowledge of the target distribution. For DA methods, [7,8] introduce a nuclear-norm regularization item to achieve balanced learning, and [64,65] design a self-training framework to realize class-balance segmentation adaptation. In the absence of labeled data, DeepCluster [3], one of the best self-supervised and class-balanced learning methods, generates pseudo-labels via k-means clustering and utilizes them to re-train the current model. Considering domain shift and absence of source data and leveraging advantages of existing methods, we design a novel and general class-balanced multicentric prototype strategy for SFDA.

3   Preliminary

In this paper, we consider the K-way object (2D images and 3D point cloud) recognition task. In the conventional UDA task, we are given two domain data,

the labeled source domain with n_s samples as $\mathcal{D}_s = \{(x_s^i, y_s^i)\}_{i=1}^{n_s}$ where $x_s^i \in \mathcal{X}_s$, $y_s^i \in \mathcal{Y}_s$ and y_s^i is the corresponding label of data sample x_s^i, and the unlabeled target domain with n_t samples as $\mathcal{D}_t = \{(x_t^i)\}_{i=1}^{n_t}$ where $x_t^i \in \mathcal{X}_t$. The goal of UDA is to predict the labels $\{y_t^i\}_{i=1}^{n_t}$ of $\mathcal{D}_t$ where $y_t^i \in \mathcal{Y}_t$ with the well-labeled source domain $\mathcal{D}_s$. It is commonly assumed that the data space of $\mathcal{D}_s$ and $\mathcal{D}_t$ are distinct but label space are identical, i.e., $\mathcal{X}_s \neq \mathcal{X}_t$, $\mathcal{Y}_s = \mathcal{Y}_t$. But under the SFDA setting, the $\mathcal{D}_s$ is inaccessible and it is replaced by the source model. Assume that the source model f_s has been well-trained and it consists of two parts: a feature extractor $g_s \colon \mathcal{X}_s \to \mathbb{R}^d$ and a classifier $h_s \colon \mathbb{R}^d \to \mathbb{R}^K$, i.e., $f_s(x) = h_s(g_s(x))$. Here d is the dimension of the extracted feature. Therefore, the goal for SFDA is to learn the target model $f_t \colon \mathcal{X}_t \to \mathcal{Y}_t$ with only access to the source model f_s and the unlabeled target domain $\mathcal{D}_t$.

To transfer the knowledge from the pre-trained source model, feature prototype based pseudo-labeling strategy is a promising direction. Inspired by Deep-Cluster [3], existing feature prototype based pseudo-labeling strategies [26,27,1] first attain the prototype c_k for each class similar to weighted k-means clustering as follows:

$$c_k = \frac{\sum_{x_t \in \mathcal{X}_t} \delta_k(\hat{f}_t(x_t)) \hat{g}_t(x_t)}{\sum_{x_t \in \mathcal{X}_t} \delta_k(\hat{f}_t(x_t))}, \tag{1}$$

where $\hat{f}_t = \hat{g}_t \circ \hat{h}_t$ denotes the previously learned target model and $\delta_k(\hat{f}_t(x_t))$ denotes the softmax probability of target instance x_t belonging to the k-th class. Then one can obtain the pseudo label $\hat{y}_t$ of target x_t via the nearest prototype classifier as:

$$\hat{y}_t = \arg\min_k D_f(\hat{g}_t(x_t), c_k), \tag{2}$$

where $D_f(a, b)$ measures the distance between a and b. One may iterate above process to obtain more stable prototype and pseudo labels like:

$$c_k = \frac{\sum_{x_t \in \mathcal{X}_t} \mathbb{1}(\hat{y}_t = k) \, \hat{g}_t(x_t)}{\sum_{x_t \in \mathcal{X}_t} \mathbb{1}(\hat{y}_t = k)},$$
$$\hat{y}_t = \arg\min_k D_f(\hat{g}_t(x_t), c_k). \tag{3}$$

where $\mathbb{1}(\cdot)$ is an indicator function. Thereafter, based on the obtained pseudo labels, one can realize self-training based model adaptation with the categorical cross-entropy (CE) loss as follows:

$$\mathcal{L}_{st} = -\frac{1}{n_t} \sum_{i=1}^{n_t} \sum_{k=1}^{K} \mathbb{1}_{[k=\hat{y}_t]} \log \delta_k(f_t(x_t^i)). \tag{4}$$

4 BMD Strategy

4.1 Inter-class Balanced Prototype

The visual domain gaps between source and target are typically different between categories, resulting in relatively higher prediction confidence scores for
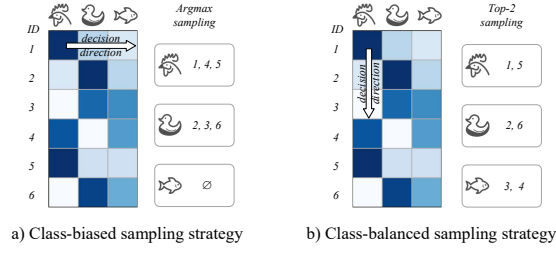
Fig. 2. A toy example compares the existing class-biased strategy (left) with BMD class-balanced strategy (right). To better understand the difference, we illustrate the decision direction of these two sampling strategies. In the presence of large domain gaps, existing strategy is prone to aggregating class-biased data instances.

those easy-transfer classes in target domain [64,65]. Therefore, we conjecture that existing strategies are category-biased and tend to generating noisy labels for those hard data. To avoid the gradual dominance of easy-transfer classes on prototype generation, we propose a novel global inter-class balanced sampling strategy to aggregate those potential data samples. Different from existing methods that decide whether to sample instances based on the instance-level prediction results, we formulate this as a multiple instance learning (MIL) problem [11,25]. In MIL, individual samples are grouped in two bags, i.e., positive and negative bags. A positive bag contains at least one positive instance and a negative bag contains no positive instance. For a specific class k, we treat the target domain $\mathcal{D}_t$ as a combination of a positive bag and a negative bag, where each data instance x_t is represented by a feature vector $\hat{g}_t(x_t)$ and a classification result $p(x_t) = \delta(\hat{f}_t(x_t))$. Thus, the feature prototype for the k-th class is a representative of the positive bag. Since the *top* instances are most likely to be positive, we then aggregate the top-M $\delta_k(\hat{f}_t(x_t))$ scores represented instances along all target domain $\mathcal{D}_t$ for the k-th class as potential instances. After that we can average them to build the class-balanced feature prototype c_k and assign the pseudo label $\hat{y}_t$ as:

$$\mathcal{M}_k = \underset{\substack{x_t \in \mathcal{X}_t \\ |\mathcal{M}_k| = M}}{\arg\max} \delta_k(\hat{f}_t(x_t)),$$

$$c_k = \frac{1}{M} \sum_{i \in \mathcal{M}_k} \hat{g}_t(x_t^i), \qquad (5)$$

$$\hat{y}_t = \underset{k}{\arg\min} D_f(\hat{g}_t(x_t), c_k).$$

where $M = \max\{1, \lfloor \frac{n_t}{r \times K} \rfloor\}$, r is a hyperparameter denoting the top-M selection ratio, and K is the number of object classes in the target domain. For simplicity, we refer to this class-balanced sampling based feature prototype pseudo-labeling strategy as BP. It is worth noting that our BP strategy is not based on local instance-level prediction results to decide whether to sample instances, but rather to select the top-M most likely instances to construct feature prototypes

from a global perspective. Therefore, we argue that our strategy is inter-class balanced. We compare our class-balanced strategy with existing class-biased strategy in Fig. 2 with a toy example. We may iterate this process like existing methods to obtain more stable prototypes and pseudo labels as:

$$
\begin{aligned}
\mathcal{M}_k &= \arg\max_{\substack{x_t \in \mathcal{X}_t \\ |\mathcal{M}_k|=M}} \frac{\exp\left(\hat{g}_t(x_t) \cdot c_k\right)}{\sum_{j=1}^{K} \exp\left(\hat{g}_t(x_t) \cdot c_j\right)}, \\
c_k &= \frac{1}{M} \sum_{i \in \mathcal{M}_k} \hat{g}_t(x_t^i), \\
\hat{y}_t &= \arg\min_{k} D_f(\hat{g}_t(x_t), c_k).
\end{aligned}
\tag{6}
$$

4.2   Intra-class Multicentric Prototype

Even though with the above strategy we can obtain class-balanced feature prototype and assign more robust pseudo labels, we found that a coarse monocentric feature prototype may be not effectively represent those ambiguous data and even introduce negative transfer. In contrast to [60] that introduces uncertainty for pseudo labels to mitigate the negative transfer caused by monocentric prototype, in this paper, we aim to assign more robust and precise label for each instance. Therefore, we propose an intra-class multicentric prototype strategy for each category to obtain more robust and precise pseudo label. We compare our multicentric strategy with existing monocentric strategy in Fig. 3.

Clustering is an essential data analysis technique for grouping unlabeled data in unsupervised learning [18]. In our implementations, assume the sampled data instances for k-th class as $\mathcal{X}_t^k$, the predefined multiple feature prototype number is S. We represent each data instance x_t as the extracted $\hat{g}_t(x_t)$, and apply the classical k-means [33] algorithm to realize intra-class clustering. Then for simplicity, we directly denote the S cluster centroids $\{c_k^i\}_{i=1}^S$ as intra-class multiple feature prototypes for the k-th class. After obtaining multiple feature prototypes

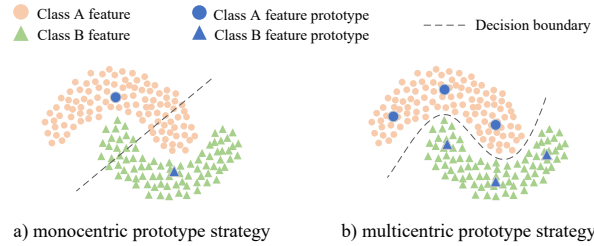a) monocentric prototype strategy   b) multicentric prototype strategy

Fig. 3. Comparison between existing monocentric prototype strategy (left) and our proposed multicentric prototype strategy (right). We can conclude that with multicentric prototype strategy we would obtain more robust and precise decision boundaries during pseudo labels generation.

for all categories, we can assign the pseudo-labels as follows:

$$\hat{y}_t = \arg\max_k \frac{\max\limits_{1 \le i \le S}(\exp(\hat{g}_t(x_t) \cdot c_k^i))}{\sum_{j=1}^{K} \max\limits_{1 \le i \le S}(\exp(\hat{g}_t(x_t) \cdot c_j^i))}, \tag{7}$$

where c_k^i is the i-th feature prototype of class k. For simplicity, we refer to this class-balanced sampling based multicentric pseudo-labeling strategy as BMP. We may also iterate this process like before to obtain more stable prototype and pseudo labels as:

$$\mathcal{M}_k = \arg\max_{\substack{x_t \in \mathcal{X}_t \\ |\mathcal{M}_k|=M}} \frac{\max\limits_{1 \le i \le S}(\exp(\hat{g}_t(x_t) \cdot c_k^i))}{\sum_{j=1}^{K} \max\limits_{1 \le i \le S}(\exp(\hat{g}_t(x_t) \cdot c_j^i))},$$

$$\{c_k^i\}_{i=1}^S = K\!\!\underset{n \in \mathcal{M}_k}{means}(\hat{g}_t(x_t^n)), \tag{8}$$

$$\hat{y}_t = \arg\max_k \frac{\max\limits_{1 \le i \le S}(\exp(\hat{g}_t(x_t) \cdot c_k^i))}{\sum_{j=1}^{K} \max\limits_{1 \le i \le S}(\exp(\hat{g}_t(x_t) \cdot c_j^i))}.$$

4.3   Dynamic Pseudo Label

Existing strategies update the pseudo label bank at a fixed training period, which may not effectively exploit the updated network during optimization. Thus, we further explore a dynamic pseudo-labeling strategy to improve the target model performance. At the beginning of each epoch, we first update the multiple feature prototype for each class and the corresponding pseudo labels for each instance from a global perspective. And then, for each iteration step we update the feature prototypes as the exponentially moving average (EMA) [50] of the cluster centroids in mini-batches. Specifically, we obtain the dynamic pseudo labels $\hat{y}_t^d$ and update the feature prototypes as follows:

$$\hat{y}_t^d = \frac{\max\limits_{1 \le i \le S}(\exp(\hat{g}_t(x_t) \cdot c_k^i))}{\sum_{j=1}^{K} \max\limits_{1 \le i \le S}(\exp(\hat{g}_t(x_t) \cdot c_j^i))},$$

$$p_k^i(x_t^n) = \frac{\exp(\hat{g}_t(x_t^n) \cdot c_k^i)}{\sum_{j=1}^{K} \sum_{s=1}^{S} \exp(\hat{g}_t(x_t^n) \cdot c_j^s)}, \tag{9}$$

$$\hat{c}_k^i = \frac{\sum_{n=1}^{N} \hat{g}_t(x_t^n) \cdot p_k^i(x_t^n)}{\sum_{n=1}^{N} p_k^i(x_t^n)},$$

$$c_k^i \leftarrow \lambda c_k^i + (1-\lambda)\hat{c}_k^i.$$

where $p_k^i(x_t)$ denotes the similarity of instance x_t with existing feature prototypes, $\hat{c}_k^i$ represents the i-th feature prototype of class k calculated with current training minibatch, and λ is the momentum coefficient of EMA which we set to

0.9999. With the obtained dynamic pseudo labels $\hat{y}_t^d$, instead of using a standard cross-entropy loss, following [60], we adopt a more robust variant, symmetric cross-entropy loss (SCE) [56] to further enhance the noise-tolerance. Formally, the dynamic pseudo label loss is defined as follows:

$$\mathcal{L}_{dym} = -\frac{1}{N}\sum_{i=1}^{N}\sum_{k=1}^{K}\hat{y}_{t,k}^d \log \delta_k(f_t(x_t^i))- \\ \frac{1}{N}\sum_{i=1}^{N}\sum_{k=1}^{K}\delta_k(f_t(x_t^i))\log \hat{y}_{t,k}^d, \tag{10}$$

However, the dynamic pseudo label did not take into account the potential for a large domain shift and may lead to less informative class prototypes, especially when many samples are misclassified, as it is updated based on the features of the local minibatch. Therefore, we combine the static pseudo label based self-training loss with the dynamic loss to achieve more stable results as :

$$\mathcal{L}_{bmd} = \alpha\mathcal{L}_{st} + \beta\mathcal{L}_{dym}. \tag{11}$$

where α and β are hyper-parameters to balance the two losses. Overall, we denote the combination of the class-balanced multicentric pseudo-labeling and dynamic pseudo labeling strategies as BMD.

5   Experiment

5.1   Experimental Setup

We conduct extensive experiments to evaluate the effectiveness of our BMD strategy covering several popular benchmarks and representative methods below.

Datasets We evaluate our BMD strategy on three 2D image and one 3D point cloud recognition benchmarks. **Office-31** [45] is a standard benchmark that contains three domains (Amazon (**A**), DSLR (**D**), and Webcam (**W**)) and each domains contains 31 object classes under the office environment. **Office-Home** [53] is a challenging medium-sized benchmark that contains 4 domains (Real (**Rw**), Clipart (**Cl**), Art (**Ar**) and Product (**Pr**)) with 65 classes and a total of 15,500 images. **VisDA-C** [40] is a more challenging large-scale benchmark, which focus on 12-class synthetic-to-real object recognition tasks. Its source domain contains about 152k synthetic 3D object images while the target domain consists of 55k real object images sampled from Microsoft CoCo [28]. **PointDA-10** [43] is the first 3D dataset designed for domain adaptation on point cloud, which contains three domains (ModelNet-10, ShapeNet-10 and ScanNet-10). There are about 27.7k training and 5.1k testing frame point clouds.

Table 1. Classification accuracies (%) on small-sized Office-31 dataset with ResNet-50 as backbone. SF denotes source-free.

Method	Venue	SF	A→D	A→W	D→A	D→W	W→A	W→D	**Avg**
DANN [13]	JMLR 2016	✗	79.7	82.0	68.2	96.9	67.4	99.1	82.2
CDAN [30]	NeurIPS 2018	✗	92.9	94.1	71.0	98.6	69.3	100.0	87.7
MDD [63]	ICML 2019	✗	93.5	94.5	74.6	98.4	72.2	100.0	88.9
GVB-GD [9]	CVPR 2020	✗	95.0	94.8	73.4	98.7	73.7	100.0	89.3
SHOT [26]	ICML 2020		94.0	90.1	74.7	98.4	74.3	99.9	88.6
SHOT w/ BMD	ours	✓	95.6	93.0	75.6	97.5	75.0	99.8	**89.4**
SHOT++ [27]	TPAMI 2021		94.3	90.4	76.2	98.7	75.8	99.9	89.2
SHOT++ w/ BMD	ours	✓	96.2	94.2	76.0	98.0	76.0	100.0	**90.1**

Table 2. Accuracies (%) on medium-sized Office-Home dataset with ResNet-50 as backbone.[* using our reproduced performance]

Methods	Venue	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**
CDAN [30]	NeurIPS 2018	✗	50.7	70.6	76.0	57.6	70.0	70.0	57.4	50.9	77.3	70.9	56.7	81.6	65.8
CDAN+BNM [7]	CVPR 2020	✗	56.2	73.7	79.0	63.1	73.6	74.0	62.4	54.8	80.7	72.4	58.9	83.5	69.4
GVB-GD [9]	CVPR 2020	✗	57.0	74.7	79.8	64.6	74.1	74.6	65.2	55.1	81.0	74.6	59.7	84.3	70.4
Fixbi [36]	CVPR 2021	✗	58.1	77.3	80.4	67.7	79.5	78.1	65.8	57.9	81.7	76.4	62.9	86.7	72.7
SHOT* [26]	ICML 2020		54.3	78.1	80.3	68.3	79.1	80.1	68.7	54.1	82.0	73.1	57.0	83.0	71.5
SHOT w/ BMD	ours	✓	55.9	77.8	80.8	69.7	79.3	79.9	69.6	56.6	82.6	73.3	59.5	85.1	**72.5**
G-SFDA* [59]	ICCV 2021		55.2	77.6	80.1	67.7	75.6	79.1	66.3	54.8	81.6	72.5	58.1	84.0	71.0
G-SFDA w/ BMD	ours	✓	56.0	78.2	80.4	69.1	79.0	79.4	67.5	55.8	82.4	73.7	58.7	83.8	**72.0**
SHOT++* [27]	TPAMI 2021		55.9	79.1	81.8	69.9	81.3	81.0	70.3	56.2	83.6	72.9	59.0	84.3	72.9
SHOT++ w/ BMD	ours	✓	58.1	79.7	82.6	69.3	81.0	80.7	70.8	57.6	83.6	74.0	60.0	85.9	**73.6**

Baselines We inject our BMD strategy to four existing SFDA methods to verify its versatility. **SHOT** [26] proposes to freeze the source classifier and fine-tunes the source features extraction module by maximizing the mutual information and feature prototype based pseudo labels. **SHOT++** [27] extends the SHOT [26] by introducing self-supervised learning for fine-tuning the feature extraction module and employing semi-supervised learning strategy to further improve the target domain performance. Different from SHOT and SHOT++ that introduce pseudo labeling strategy to realize model adaptation, **G-SFDA** [59] and **NRC** [58] explore the local neighborhood structure of the target data in feature space to realize model adaptation. Although these two approaches do not introduce pseudo labeling, we find that our BMD strategy still fits seamlessly with these methods and consistently improves their performance.

Implementation Details For a fair comparison, we adopt the same network architecture and training recipe with baselines. Specifically, we adopt the ResNet-50 [17] pretrained on ImageNet [10] as backbone for Office-31 and Office-Home benchmarks, and the ResNet-101 for VisDA-C benchmark. As for PointDA-10, we utilize the PointNet [42] with local node aggregation network proposed in [43] as feature extraction backbone. To prepare the pretrained source model, following SHOT and NRC, we utilize the label smoothing [34] to increase the discrim-

inability of the source model and facilitate the following target data alignment. During target model adaptation, to achieve source and target domain alignment, we fix the target classifier $h_t = h_s$ and update only the target feature extractor g_t initialized from g_s. Following previous methods, we apply the SGD optimizer with momentum 0.9 and the Adam optimizer for PointDA-10. The batch size is set to 64 for all benchmark datasets. We set the learning rate to 1e-2 for Office-31 and OfficeHome, 1e-3 for VisDA-C, and 1e-6 for PointDA-10. We train 30 epochs for all 2D image datasets and 50 epochs for PointDA. We set the hyperparamter r to 3 for all datasets, and $S = 4$ for Office-Home and VisDA-C, $S = 2$ for PointDA-10 and Office-31. We set $\alpha = 2$ and $\beta = 0.5$ for VisDA-C, $\alpha = 0.3$ and $\beta = 0.1$ for Office-31 and Office-Home, and $\alpha = 1.0$ and $\beta = 0.1$ PointDA-10. All experiments are conducted on a RTX-3090 GPU with PyTorch-1.7.

5.2   Results

2D Image Recognition We first evaluate the effectiveness of our strategy with existing methods on three 2D image recognition datasets. The results are summarized in Table 1-3, the top part illustrates results for the traditional UDA methods with access to source data during model adaptation, and the bottom part presents results for the SFDA methods. As shown in Table 1, on **Office-31**, our BMD strategy can consistently improve SHOT and SHOT++ to yield new state-of-the-art performance, especially on the challenging A → D task, our BMD strategy can improve SHOT from 90.1% to 93.0% and SHOT++ from 90.4% to 94.2 %, respectively. As excepted in Table 2, on the medium-sized **Office-Home**, our BMD strategy can also consistently improve existing state-of-the-art methods. Specifically, by injecting BMD strategy, we can improve SHOT from 71.5% to 72.5%, G-SFDA from 71.0% to 72.0%, SHOT++ from 72.9% to 73.6%. For the large-scale synthetic-to-real **VisDA-C** dataset in Table 3, our BMD strategy can also significantly improve existing methods by a large margin, especially, we can improve SHOT from 82.9% to 85.7%, G-SFDA from 84.8% to 86.5%, NRC from 85.9% to 86.9% and SHOT++ from 87.3% to 88.7%. With our BMD strategy, SHOT++ can even achieve a performance comparable to the target supervised approach (88.7% vs 89.6%). In addition, on VisDA-C, we can find that with our strategy above methods can achieve more class-balanced performance. Especially for the challenging class 'truck', we can significantly improve SHOT from 58.2% to 70.8%, G-SFDA from 44.8% to 59.7%, and SHOT++ from 28.8% to 45.9%. We also report the standard deviation σ of the accuracy achieved by our method. The σ of SHOT w/BMD on Office-31 is 0.07, while the σ of SHOT w/BMD on VisDA-C is 0.11, showing that the improvement of using BMD in SHOT is significant. Beyond closet-set SFDA, we further evaluate BMD with SHOT on two other DA scenarios, multi-source [38] and multi-target [39]. Due to space limitations, we will present these experiments in the Appendix.

3D Point Cloud Recognition In addition to 2D images, to verify the generality of our BMD strategy, we also conducted experiments on the 3D point

Table 3. Per-class accuracy (%) on large-scale VisDA-C validation set with ResNet-101 as backbone.

Methods	Venue	SF	plane	bcycl	bus	car	horse	knife	mcycl	person	plant	sktbrd	train	truck	Avg
CDAN [30]	NeurIPS 2018	✗	85.2	66.9	83.0	50.8	84.2	74.9	88.1	74.5	83.4	76.0	81.9	38.0	73.9
SWD [22]	CVPR 2019	✗	90.8	82.5	81.7	70.5	91.7	69.5	86.3	77.5	87.4	63.6	85.6	29.2	76.4
MCC [19]	ECCV 2020	✗	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
STAR [32]	CVPR 2020	✗	95.0	84.0	84.6	73.0	91.6	91.8	85.9	78.4	94.4	84.7	87.0	42.2	82.7
FixBi [36]	CVPR 2021	✗	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
SHOT [26]	ICML 2020		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
SHOT w/ BMD	ours	✓	96.2	87.8	81.4	61.7	95.0	97.5	87.9	82.9	92.6	88.8	87.4	70.8	**85.8**
G-SFDA [59]	ICCV 2021		95.9	88.1	85.4	72.5	96.1	93.7	88.5	80.6	92.3	92.2	87.6	44.8	84.8
G-SFDA w/ BMD	ours	✓	95.9	87.5	83.9	75.7	96.5	96.6	91.4	81.8	95.9	88.4	85.1	59.7	**86.5**
NRC [58]	NeurIPS 2021		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
NRC w/ BMD	ours	✓	96.7	87.2	85.0	75.6	96.8	97.0	91.6	84.9	94.7	89.0	88.6	55.6	**86.9**
SHOT++ [27]	TPAMI 2021		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
SHOT++ w/ BMD	ours	✓	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**
Target-Supervised	–	–	97.0	86.6	84.3	88.7	96.3	94.4	92.0	89.4	95.5	91.8	90.7	68.7	89.6

Table 4. Accuracies (%) on PointDA-10 dataset with PointNet [42] as backbone.

Method	Venue	SF	M→SC	M→SH	SC→M	SC→SH	SH→M	SH→SC	Avg.
ADDA [52]	CVPR 2017	✗	30.5	61.0	48.9	51.1	40.4	29.3	43.5
MCD [46]	CVPR 2018	✗	31.0	62.0	46.8	59.3	41.4	31.3	45.3
PointDAN [43]	NeurIPS 2019	✗	33.0	64.2	49.1	64.1	47.6	33.9	48.7
VDM [51]	Arxiv 2021	✓	30.9	58.4	45.3	61.8	61.0	40.8	49.7
SHOT [26]	ICML 2020		31.8	62.1	67.6	56.9	75.8	24.3	53.1
SHOT w/ BMD	ours	✓	32.8	66.1	75.0	62.0	81.5	24.4	**57.0**
NRC [58]	NeurIPS 2021		25.8	64.8	70.1	68.1	59.8	26.9	52.6
NRC w/ BMD	ours	✓	33.8	66.7	70.8	62.6	83.4	24.8	**57.0**

cloud **PointDA-10** dataset. As shown in Table 4, our BMD strategy can also significantly improve existing methods by a large margin. Specifically, by injecting BMD strategy, we can improve SHOT from 53.1% to 57.0% and NRC from 52.6% to 57.0%. Especially on the challenging task, SH → M, we can improve SHOT from 75.8% to 81.5% and NRC from 59.8% to 83.4%.

5.3   Performance Analysis

Ablation Study As we presented before, the core components of our BMD strategy are inter-class balanced sampling and intra-class multicentric prototype based pseudo label generation. In order to incorporate the network dynamic optimization information, we further introduce the EMA based dynamic pseudo label strategy. To study the advantage of each part of our BMD strategy, we conduct the ablation study on Office-Home, VisDA-C and PointDA-10 with SHOT, the results are summarized in Table 5. To verify the superiority of our BMD strategy, we also introduce two existing strategies, the naive argmax based pseudo label [23] strategy, and the monocentric prototype based pseudo

Table 5. Ablation study on three UDA datasets.

Methods/Datasets	Office-Home	VisDA-C	PointDA-10
Source-model	59.6	46.6	39.4
SHOT w/ naive PL [23]	70.3	82.9	51.0
SHOT w/ mono PL [26]	71.5	82.9	53.1
SHOT w/ BP (ours)	72.0	83.8	55.0
SHOT w/ BMP (ours)	72.5	84.7	56.4
SHOT w/ BMD (ours)	72.5	85.7	57.0

Table 6. Statistics of class-wise performance on VisDA-C.

Methods	Acc avg $\mu \uparrow$	Acc std $\sigma \downarrow$	Acc cv $c_v \downarrow$
SHOT [26]	82.9	12.857	0.155
SHOT w/ BMD	85.8	10.127	**0.118**
G-SFDA [59]	84.8	14.279	0.168
G-SFDA w/ BMD	86.5	10.766	**0.124**
SHOT++ [27]	87.3	19.027	0.218
SHOT++ w/ BMD	88.7	14.146	**0.159**

label [26] strategy. As expected, the results show that the simple BP strategy can outperform existing strategies, which indicates the importance of the class-balanced sampling strategy for pseudo label generation. When we incorporate the intra-class multicentric prototype strategy with BP strategy, i.e. the BMP, the performance is further significantly boosted. We attribute this to the fact that the MP strategy introduce more fine-grained feature prototypes for each class, which allows the model to assign more accurate pseudo-labels for those hard-transfer data. As for the dynamic strategy, we find that it is not as effective as BP and BMP on the Office-Home and PointDA datasets. We suspect it may be due to the relatively small size of the datasets, thus the EMA-based dynamic feature prototypes cannot effectively utilize the information during training. Due to space limitations, we presented more ablation experiments in the Appendix.

Does our strategy really achieve more class-balanced results? To verify whether our BMD strategy is really helpful in achieving the class-balanced results, in this part, we introduce the coefficient of variation (also known as the relative standard deviation) c_v as metric to evaluate the inter-class balance performance, which is a standardized measure of dispersion of a probability distribution or frequency distribution. Formally, the coefficient of variation is defined as $c_v = \frac{\sigma}{\mu}$, where σ and μ are the standard deviation and expected mean of the data distribution, respectively. We conduct experiments on VisDA-C dataset, the results are summarized in Table 6. As shown in this table, for all methods by injecting our BMD strategy we can arrive higher accuracy mean μ, lower standard deviation σ and lower coefficient of variation c_v, which demonstrates that our BMD strategy indeed facilitates existing methods to achieve more class-balanced performance. To verify the robustness, we also conducted the c_v experiments on PointDA, on task SC→M, SHOT gets 0.291, SHOT w/BMD is 0.186; NRC gets 0.464, NRC w/BMD is 0.301. These results further demonstrate that our BMD strategy can improve the existing methods to achieve class-balanced results.

Visualization To demonstrate the superiority of our BMD strategy, we present the t-SNE feature and confusion matrix on PointDA-10 (SC→M), and the pseudo and predicted label accuracy curves on VisDA-C in Fig. 4. From Fig. 4a, we can see that after model adaptation by SHOT w/ BMD, the target features are more compactly clustered. The confusion matrix in Fig. 4b demonstrates
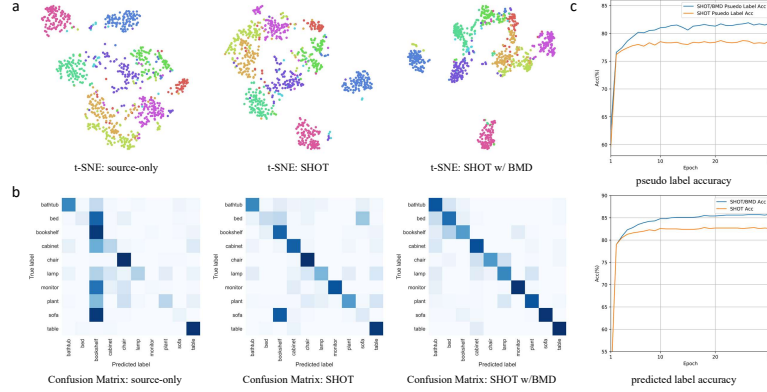
Fig. 4. a: The t-SNE visualization of target features for source model, SHOT, and SHOT w/ BMD on PointDA-10 (SC→M). **b**: The Confusion Matrix visualization for source model, SHOT, and SHOT w/BMD on PointDA-10 (SC→M). **c**: The pseudo and predicted label accuracy curves for SHOT, and SHOT w/BMD on VisDA-C.

that our BMD strategy can achieve more class-balanced accuracy compared to the vanilla SHOT. In particular, when there is severe class bias in the source model, e.g., for the hard-transfer *'sofa'* class, SHOT cannot achieve good model adaptation due to the severe domain gap, while BMD strategy can overcome this well by our inter-class balanced sampling. The accuracy curves in Fig. 4c further support that our BMD strategy can facilitate existing methods to achieve more superior pseudo labels and predicted labels during model adaptation.

6   Conclusion

In this paper, we present a general class-balanced multicentric dynamic (BMD) prototype strategy for source-free domain adaptation, which is model agnostic and can be applied to existing self-training based SFDA methods. Specifically, our BMD strategy consists of a novel inter-class balanced sampling strategy, an intra-class multicentric prototype strategy, and a dynamic feature prototype based pseudo-labeling strategy. We have injected our strategy into four existing representative methods and conducted experiments on both 2D images and 3D point cloud datasets. The results demonstrate that our BMD strategy can consistently and significantly boost existing methods to yield new state-of-the-art performance. For future work, we will extend our BMD strategy to those source-free dense prediction tasks.

References

1. Ahmed, S.M., Raychaudhuri, D.S., Paul, S., Oymak, S., Roy-Chowdhury, A.K.: Unsupervised multi-source domain adaptation without access to source data. In: CVPR (2021) 2, 4, 5
2. Bousmalis, K., Trigeorgis, G., Silberman, N., Krishnan, D., Erhan, D.: Domain separation networks. In: NeurIPS (2016) 3
3. Caron, M., Bojanowski, P., Joulin, A., Douze, M.: Deep clustering for unsupervised learning of visual features. In: ECCV (2018) 4, 5
4. Chen, G., Chen, K., Zhang, L., Zhang, L., Knoll, A.: Vcanet: Vanishing-point-guided context-aware network for small road object detection. Automotive Innovation (2021) 1
5. Chen, Y., Li, W., Sakaridis, C., Dai, D., Van Gool, L.: Domain adaptive faster r-cnn for object detection in the wild. In: CVPR (2018) 2
6. Courty, N., Flamary, R., Tuia, D., Rakotomamonjy, A.: Optimal transport for domain adaptation. IEEE TPAMI **39**(9), 1853–1865 (2016) 3
7. Cui, S., Wang, S., Zhuo, J., Li, L., Huang, Q., Tian, Q.: Towards discriminability and diversity: Batch nuclear-norm maximization under label insufficient situations. In: CVPR (2020) 4, 10
8. Cui, S., Wang, S., Zhuo, J., Li, L., Huang, Q., Tian, Q.: Fast batch nuclear-norm maximization and minimization for robust domain adaptation. arXiv preprint arXiv:2107.06154 (2021) 4
9. Cui, S., Wang, S., Zhuo, J., Su, C., Huang, Q., Tian, Q.: Gradually vanishing bridge for adversarial domain adaptation. In: CVPR (2020) 10
10. Deng, J., Dong, W., Socher, R., Li, L.J., Li, K., Fei-Fei, L.: Imagenet: A large-scale hierarchical image database. In: CVPR (2009) 1, 10
11. Dietterich, T.G., Lathrop, R.H., Lozano-Pérez, T.: Solving the multiple instance problem with axis-parallel rectangles. Artificial intelligence **89**(1-2), 31–71 (1997) 6
12. Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., et al.: An image is worth 16x16 words: Transformers for image recognition at scale. In: ICLR (2020) 1
13. Ganin, Y., Ustinova, E., Ajakan, H., Germain, P., Larochelle, H., Laviolette, F., Marchand, M., Lempitsky, V.: Domain-adversarial training of neural networks. JMLR **17**(1), 2096–2030 (2016) 2, 4, 10
14. Gao, L., Zhang, J., Zhang, L., Tao, D.: Dsp: Dual soft-paste for unsupervised domain adaptive semantic segmentation. In: ACM Multimedia (2021) 2
15. Ghifary, M., Kleijn, W.B., Zhang, M., Balduzzi, D., Li, W.: Deep reconstruction-classification networks for unsupervised domain adaptation. In: ECCV (2016) 3
16. Han, J., Luo, P., Wang, X.: Deep self-learning from noisy labels. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 5138–5147 (2019) 4
17. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: CVPR (2016) 1, 10
18. Jain, A.K., Murty, M.N., Flynn, P.J.: Data clustering: a review. ACM computing surveys (CSUR) **31**(3), 264–323 (1999) 7
19. Jin, Y., Wang, X., Long, M., Wang, J.: Minimum class confusion for versatile domain adaptation. In: ECCV (2020) 12
20. Kang, B., Xie, S., Rohrbach, M., Yan, Z., Gordo, A., Feng, J., Kalantidis, Y.: Decoupling representation and classifier for long-tailed recognition. In: ICLR (2020) 4

21. Kang, G., Jiang, L., Yang, Y., Hauptmann, A.G.: Contrastive adaptation network for unsupervised domain adaptation. In: CVPR (2019) 3
22. Lee, C.Y., Batra, T., Baig, M.H., Ulbricht, D.: Sliced wasserstein discrepancy for unsupervised domain adaptation. In: CVPR (2019) 12
23. Lee, D.H., et al.: Pseudo-label: The simple and efficient semi-supervised learning method for deep neural networks. In: Workshop on challenges in representation learning, ICML (2013) 12, 13
24. Li, R., Jiao, Q., Cao, W., Wong, H.S., Wu, S.: Model adaptation: Unsupervised domain adaptation without source data. In: CVPR (2020) 2, 4
25. Li, W., Vasconcelos, N.: Multiple instance learning for soft bags via top instances. In: CVPR (2015) 6
26. Liang, J., Hu, D., Feng, J.: Do we really need to access the source data? source hypothesis transfer for unsupervised domain adaptation. In: ICML (2020) 2, 3, 4, 5, 10, 12, 13
27. Liang, J., Hu, D., Wang, Y., He, R., Feng, J.: Source data-absent unsupervised domain adaptation through hypothesis transfer and labeling transfer. IEEE TPAMI (2021) 2, 3, 4, 5, 10, 12, 13
28. Lin, T.Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., Zitnick, C.L.: Microsoft coco: Common objects in context. In: ECCV (2014) 9
29. Liu, Y., Zhang, W., Wang, J.: Source-free domain adaptation for semantic segmentation. In: CVPR (2021) 2, 4
30. Long, M., Cao, Z., Wang, J., Jordan, M.I.: Conditional adversarial domain adaptation. In: NeurIPS (2018) 4, 10, 12
31. Long, M., Wang, J., Ding, G., Sun, J., Yu, P.S.: Transfer feature learning with joint distribution adaptation. In: ICCV (2013) 3
32. Lu, Z., Yang, Y., Zhu, X., Liu, C., Song, Y.Z., Xiang, T.: Stochastic classifiers for unsupervised domain adaptation. In: CVPR (2020) 12
33. MacQueen, J., et al.: Some methods for classification and analysis of multivariate observations. In: Proceedings of the fifth Berkeley symposium on mathematical statistics and probability. Oakland, CA, USA (1967) 7
34. Müller, R., Kornblith, S., Hinton, G.: When does label smoothing help? In: NeurIPS (2019) 10
35. Murez, Z., Kolouri, S., Kriegman, D., Ramamoorthi, R., Kim, K.: Image to image translation for domain adaptation. In: CVPR (2018) 3
36. Na, J., Jung, H., Chang, H.J., Hwang, W.: Fixbi: Bridging domain spaces for unsupervised domain adaptation. In: CVPR (2021) 10, 12
37. Oza, P., Sindagi, V.A., VS, V., Patel, V.M.: Unsupervised domain adaption of object detectors: A survey. arXiv preprint arXiv:2105.13502 (2021) 2
38. Peng, X., Bai, Q., Xia, X., Huang, Z., Saenko, K., Wang, B.: Moment matching for multi-source domain adaptation. In: ICCV (2019) 11
39. Peng, X., Huang, Z., Sun, X., Saenko, K.: Domain agnostic learning with disentangled representations. In: ICML. PMLR (2019) 11
40. Peng, X., Usman, B., Kaushik, N., Hoffman, J., Wang, D., Saenko, K.: Visda: The visual domain adaptation challenge. arXiv preprint arXiv:1710.06924 (2017) 3, 9
41. Prabhu, V., Kannan, A., Ravuri, M., Chaplain, M., Sontag, D., Amatriain, X.: Few-shot learning for dermatological disease diagnosis. In: Machine Learning for Healthcare Conference. pp. 532–552. PMLR (2019) 4
42. Qi, C.R., Su, H., Mo, K., Guibas, L.J.: Pointnet: Deep learning on point sets for 3d classification and segmentation. In: CVPR (2017) 10, 12
43. Qin, C., You, H., Wang, L., Kuo, C.C.J., Fu, Y.: Pointdan: A multi-scale 3d domain adaption network for point cloud representation. In: NeurIPS (2019) 3, 9, 10, 12

44. Qiu, Z., Zhang, Y., Lin, H., Niu, S., Liu, Y., Du, Q., Tan, M.: Source-free domain adaptation via avatar prototype generation and adaptation. In: IJCAI (2021) 2, 4
45. Saenko, K., Kulis, B., Fritz, M., Darrell, T.: Adapting visual category models to new domains. In: ECCV (2010) 3, 9
46. Saito, K., Watanabe, K., Ushiku, Y., Harada, T.: Maximum classifier discrepancy for unsupervised domain adaptation. In: CVPR (2018) 2, 12
47. Shu, R., Bui, H.H., Narui, H., Ermon, S.: A dirt-t approach to unsupervised domain adaptation. In: ICLR (2018) 4
48. Song, J., Shen, C., Yang, Y., Liu, Y., Song, M.: Transductive unbiased embedding for zero-shot learning. In: CVPR (2018) 4
49. Tanwisuth, K., Fan, X., Zheng, H., Zhang, S., Zhang, H., Chen, B., Zhou, M.: A prototype-oriented framework for unsupervised domain adaptation. In: NeurIPS (2021) 4
50. Tarvainen, A., Valpola, H.: Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results. In: NeurIPS (2017) 8
51. Tian, J., Zhang, J., Li, W., Xu, D.: Vdm-da: Virtual domain modeling for source data-free domain adaptation. arXiv preprint arXiv:2103.14357 (2021) 12
52. Tzeng, E., Hoffman, J., Saenko, K., Darrell, T.: Adversarial discriminative domain adaptation. In: CVPR (2017) 2, 12
53. Venkateswara, H., Eusebio, J., Chakraborty, S., Panchanathan, S.: Deep hashing network for unsupervised domain adaptation. In: CVPR (2017) 3, 9
54. Wang, T., Li, Y., Kang, B., Li, J., Liew, J., Tang, S., Hoi, S., Feng, J.: The devil is in classification: A simple framework for long-tail instance segmentation. In: ECCV (2020) 4
55. Wang, W., Cao, Y., Zhang, J., He, F., Zha, Z.J., Wen, Y., Tao, D.: Exploring sequence feature alignment for domain adaptive detection transformers. In: ACM Multimedia (2021) 2
56. Wang, Y., Ma, X., Chen, Z., Luo, Y., Yi, J., Bailey, J.: Symmetric cross entropy for robust learning with noisy labels. In: ICCV (2019) 9
57. Xu, Y., Zhang, Q., Zhang, J., Tao, D.: Vitae: Vision transformer advanced by exploring intrinsic inductive bias. NeurIPS (2021) 1
58. Yang, S., Wang, Y., van de Weijer, J., Herranz, L., Jui, S.: Exploiting the intrinsic neighborhood structure for source-free domain adaptation. In: NeurIPS (2021) 2, 3, 4, 10, 12
59. Yang, S., Wang, Y., van de Weijer, J., Herranz, L., Jui, S.: Generalized source-free domain adaptation. In: ICCV (2021) 2, 3, 4, 10, 12, 13
60. Zhang, P., Zhang, B., Zhang, T., Chen, D., Wang, Y., Wen, F.: Prototypical pseudo label denoising and target structure learning for domain adaptive semantic segmentation. In: CVPR (2021) 2, 7, 9
61. Zhang, Q., Zhang, J., Liu, W., Tao, D.: Category anchor-guided unsupervised domain adaptation for semantic segmentation. In: NeurIPS (2019) 2
62. Zhang, Y., Kang, B., Hooi, B., Yan, S., Feng, J.: Deep long-tailed learning: A survey. arXiv preprint arXiv:2110.04596 (2021) 4
63. Zhang, Y., Liu, T., Long, M., Jordan, M.: Bridging theory and algorithm for domain adaptation. In: ICML (2019) 2, 10
64. Zou, Y., Yu, Z., Kumar, B., Wang, J.: Unsupervised domain adaptation for semantic segmentation via class-balanced self-training. In: ECCV (2018) 2, 4, 6
65. Zou, Y., Yu, Z., Liu, X., Kumar, B., Wang, J.: Confidence regularized self-training. In: ICCV (2019) 2, 4, 6