Class Bias-Aware Source-Free Domain Adaptation with Multi-Center Contrastive Learning

Hua Tong[1], Siya Yao[1*], Kaibo Zhou[1], and Xiaoyu Sean Lu[2]
[1] College of Information and Electronic Engineering, Zhejiang Gongshang University, Hangzhou, China
[2] School of Cyber Science and Engineering, Nanjing University of Science and Technology, Nanjing, China
Email: 24020090070@pop.zjgsu.edu.cn, yaosiya@mail.zjgsu.edu.cn, 25020090049@pop.zjgsu.edu.cn, xiaoyu.lu@njust.edu.cn

***Abstract*—Source-free Domain Adaptation (SFDA) transfers knowledge from a pre-trained source model to an unlabeled target domain with a different data distribution, without accessing source data. A key challenge in SFDA methods is the transfer difficulty across classes, leading to class bias and pseudo-label noise. This issue is pronounced for hard-to-transfer classes, where feature discrepancies result in unreliable pseudo-labels. Consequently, the pseudo-label distribution becomes biased: easily transferable classes dominate, while harder classes suffer from scarce or noisy labels, degrading model performance. These challenges highlight the need for robust methods addressing class-wise adaptation complexity and pseudo-label instability. Therefore, we propose a Class-Difficulty-aware and Minority-class Enhancement framework (CDME). First, we model intra-class compactness based on target domain feature pairs to quantify per-class transfer difficulty, and adjust pseudo-label confidence to suppress overconfidence in hard-to-transfer classes. Second, we introduce a dynamic pseudo-label strategy with multi-center prototype contrastive learning. This enhances intra-class stability and boosts pseudo-label reliability. Extensive results show our method achieves stable and competitive performance across multiple SFDA tasks. Code is available at https://github.com/Huatong9074/CDME.**

***Index Terms*—Source-free domain adaptation, Class-aware learning, Pseudo-label self-training, Intra-class compactness, Multi-prototype contrastive learning**

I. Introduction

Deep neural networks (DNNs) have achieved significant success in visual tasks, such as image classification [1], object detection, and semantic segmentation [2]. However, their generalization ability degrades under distribution shifts, known as domain shift. To address this, unsupervised domain adaptation (UDA) [3] reduces the distribution gap between source and target domains without requiring target labels.

However, most UDA methods rely on the assumption that both source-domain data and target-domain data are accessible during training. In many real-world scenarios, however, source-domain data cannot be shared due to privacy, storage, or copyright constraints, which has led to the emergence of Source-Free Domain Adaptation (SFDA) [4]. SFDA relies only on a pre-trained source model without accessing source data, making it practical in privacy-sensitive scenarios. Most SFDA methods adopt a pseudo-label-based self-training framework, enabling iterative adaptation and improving pseudo-label quality and feature alignment.

Nevertheless, SFDA faces more challenges, among which pseudo-label bias is particularly critical. In SFDA, pseudo labels are the only supervision, and their bias accumulates during training, causing performance degradation. Due to varying transfer difficulty across classes, some classes are easier to adapt, while others are more challenging. These differences in transfer difficulty primarily arise from the inherent feature variations of the classes and the discrepancy between the source and target domains. For some classes (e.g., those with stable appearances or clear textures), they are more easily recognized in the target domain. leading to high-confidence predictions [5] and forming "majority classes". In contrast, classes that are more sensitive to domain shifts struggle to obtain reliable predictions, resulting in fewer pseudo labels and becoming "minority classes". This bias drives optimization toward majority classes, suppressing minority representations and causing them to be weakened or forgotten.

As training progresses, the decision boundary increasingly biases toward majority classes, leading to negative transfer [6]. Fig. 1 illustrates domain shift and class bias on VisDA-C and compares previous methods with ours. In SFDA, some classes suffer from higher transfer difficulty due to feature discrepancies, leading to underrepresentation during training. We define these as minority classes. To address this, we propose a Class-Difficulty-aware Minority-class Enhancement framework (CDME) to improve pseudo-label quality and learn balanced features. First, we design a class-difficulty metric based on intra-class compactness to automatically adjust class-specific temperature settings in similarity normalization and pseudo-label generation, thereby controlling the degree of pseudo-label smoothing for different classes. Then, we introduce a multi-center prototype contrastive learning mechanism to construct dynamic pseudo labels, which strengthens intra-class structure and inter-class separation, and enhances the ability of multi-center prototypes to model complex intra-class distributions while alleviating pseudo-label bias. Our method goes beyond a simple combination of modules, and is built upon a unified optimization framework centered on category-level difficulty. Our contributions are summarized as follows:

*Corresponding author. This work was supported by the Youth Project of the National Natural Science Foundation of China under Grant 62406282 and 62302223, Zhejiang Provincial Natural Science Foundation of China under Grant LMS26F030018, and Fundamental Research Funds for the Provincial Universities of Zhejiang under Grant 3090JYN9924001G-016.

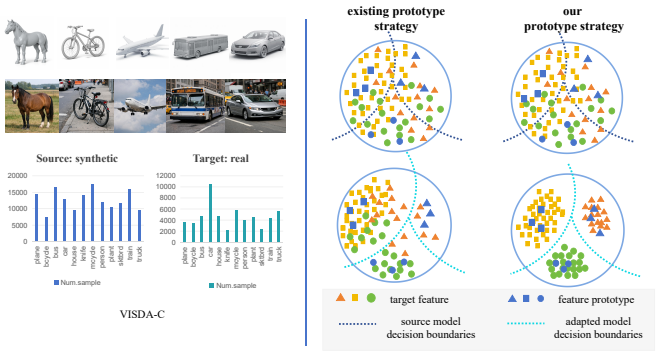

Fig. 1: (Left) Illustration of domain shift and class imbalance on the VisDA-C dataset. (Right) Comparison between the adaptation results of previous prototype-based training strategies and our proposed strategy.

- We propose a Class-Difficulty-aware and Minority-class Enhancement (CDME) framework for SFDA tasks, which can be seamlessly integrated into existing self-training-based methods.
- We introduce a class-difficulty-aware temperature adjustment mechanism that dynamically calibrates pseudo-label confidence for hard-to-transfer classes, and we develop a multi-center prototype contrastive learning strategy to improve intra-class distribution modeling and minority-class representation.
- We conduct extensive experiments on multiple benchmark datasets, demonstrating consistent performance improvements on VisDA (84.6% to 87.4%), Office-Home (71.2% to 72.0%), and Office-31 (87.7% to 88.7%).

II. Related Work

Source-free domain adaptation (SFDA) aims to achieve cross-domain adaptation without access to source data. Existing methods address the lack of supervision through pseudo-label self-training, prototype learning, and dynamic pseudo-label strategies.

Pseudo-label self-training and consistency learning. The consistency loss by Sajjadi et al. [7] lays the foundation for consistency-based learning. Source Hypothesis Transfer (SHOT) [8] represents a typical SFDA approach, combining prototype-based pseudo labeling with self-training. These methods generate pseudo labels using weak augmentations and enforce consistency under strong augmentations. However, approaches such as SHOT and NRC adopt fixed or globally adjusted temperatures, lacking class-wise adaptability and thus limiting performance under varying transfer difficulty.

Prototype learning and multi-center structure modeling. Previous work introduced multi-prototype representations to capture intra-class variations and improve pseudo-label reliability [9]. Balanced Multicentric Dynamic prototypes (BMD) [10] maintains multiple prototype centers per class with iterative updates for robust representations. Compared to single-center methods, BMD reduces prototype drift and improves pseudo-label stability, but it neglects intra-class compactness and transfer difficulty. As a result, pseudo-label confidence fails to capture class-level differences, leading to instability under hard class shifts.

Dynamic pseudo-labeling. Dynamic pseudo-labeling is widely used in SFDA to iteratively refine pseudo labels, improving stability and reliability. Methods like those by Lu et al. [11] and energy-based methods improve label quality by modeling uncertainty and dynamically selecting labels. However, these methods focus on noisy labels but overlook the weakening of minority classes, causing instability under class bias and domain shift.

These methods improve pseudo-label quality and feature representation but face challenges due to noise and hard-to-transfer classes. Unlike prior methods, our approach introduces a class-difficulty-aware mechanism to adjust pseudo-label confidence and uses multi-center prototype contrastive learning to enhance minority class representations, improving adaptability and robustness.

III. Method

In this section, we introduce the SFDA setting and notations, followed by our framework for robust and generalizable SFDA. Fig. 2 shows the overall framework. We first propose a class-difficulty-aware method based on intra-class compactness, then adopt a dynamic pseudo-label strategy with multi-center prototype contrastive learning to mitigate noise, class-difficulty variation, and complex intra-class structure.

A. Preliminaries and notations

In this paper, we focus on source-free domain adaptation (SFDA), a special case of traditional UDA. We consider two datasets: a labeled source domain dataset, denoted as $D_S = \left\{\left(x_s^i, y_s^i\right)\right\}_{i=1}^{n_s}$, where $x_s^i \in X_s$ is the feature and $y_s^i \in Y_s$ is the corresponding label, and an unlabeled target domain dataset, denoted as $D_T = \left\{\left(x_t^i\right)\right\}_{i=1}^{n_t}$, where $x_t^i \in X_t$ is the feature without label. The goal of UDA is to use the source domain (including both data and model) to predict the labels Y_T in the target domain D_T, where the data distributions in the source D_S and target domains D_T are different. However, in the SFDA setting, the source domain dataset D_S is inaccessible and there is only the source model to complete the target domain adaptation task.

The source model has been fully trained and consists of two components: a feature extractor $g_s : X_s \to \mathbb{R}^d$ and a classifier $h_s : \mathbb{R}^d \to \mathbb{R}^K$, i.e., $f_s(x) = h_s(g_s(x))$, where d denotes the dimension of the extracted features. SFDA learns a target model $f_t : X_t \to Y_T$ by leveraging the source model f_s and the unlabeled target domain D_T.

B. SFDA Basic Framework

Many self-training methods are built upon SHOT. SFDA methods rely on pseudo labels for training. Pseudo labels are generated based on similarity to class prototypes, a simple and general approach. An inter-class balanced sampling strategy [12] aggregates samples to construct class prototypes c_k and

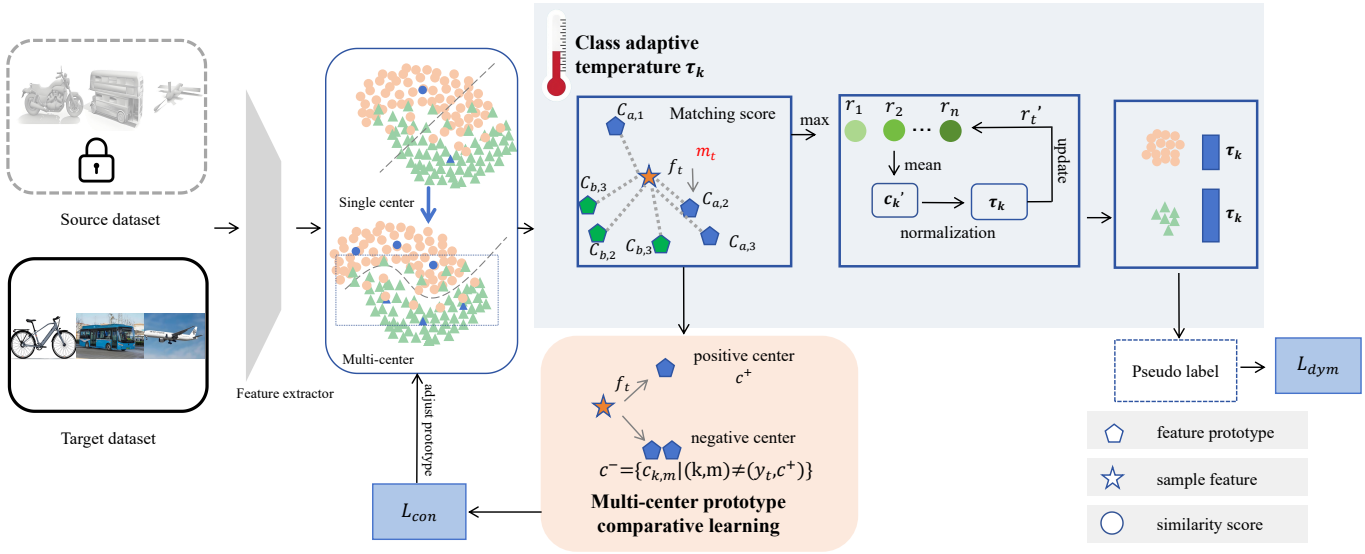

Fig. 2: Overview of CDME: A Class-Difficulty-aware and Minority-class Enhancement framework for SFDA. The feature extractor transforms source domain data to the target domain, adapting it to a multi-center setting. The sample-to-center matching score calculates intra-class compactness, then dynamically adjusts the class-specific temperature to generate pseudo-labels. Finally, multi-center prototype contrastive learning is applied to guide samples toward the positive prototype center.

assign pseudo labels $\hat{y}_t = \arg\max_k D_f(\hat{g}_T(x_t), c_k)$, where each sample x_t is represented by a feature vector $\hat{g}_T(x_t)$, c_k is the prototype of class k, and D_f is a function that calculates the similarity between a sample and the class prototype.

However, existing methods suffer from class bias and introduce noisy labels for hard samples. his is partly due to ignoring varying transfer difficulty across classes, leading to poor adaptability. Moreover, coarse prototypes fail to capture domain shift, leading to blurred decision boundaries. To address these issues, we propose a class-difficulty-aware and minority-class enhancement framework for SFDA. Our method will be described in detail in the following sections.

C. Class-Difficulty and Intra-Class Compactness

In traditional SFDA methods, a single prototype [13] represents each class. However, this fails to capture intra-class diversity in target domains. BMD introduces multi-center prototypes to capture intra-class variations. While multi-center prototypes can mitigate class bias and noisy label issues, they do not account for transfer difficulty differences between classes and struggle with varying transfer difficulty in complex domains. To address this, we propose a class-difficulty-aware method based on intra-class compactness, which estimates class difficulty from feature distributions in the target domain.

First, we calculate intra-class compactness for each class. For a feature f_t of a target domain sample x_t, we compute its matching score $r_{t,k,m} = \langle f_t, c_{k,m} \rangle$ with each center, where $c_{k,m}$ is the m-th prototype within class k, and $\langle f_t, c_{k,m} \rangle$ is the inner product between f_t and $c_{k,m}$. We then select the center with the highest similarity as the nearest neighbor center m_t for the sample:

$$m_t = \arg\max_m r_{t,k,m} \tag{1}$$

Class compactness is defined as the average similarity to the nearest prototype centers:

$$c_k = \frac{1}{N_k} \sum_{t=1}^{N_k} \max_m r_{t,k,m} \tag{2}$$

N_k denotes the number of samples in class k. Transfer difficulty is closely related to intra-class compactness. High compactness leads to clustered features and easier transfer, while low compactness results in dispersed features and harder transfer. Thus, majority classes benefit from higher compactness and more samples, while minority classes suffer from low compactness and are more vulnerable to pseudo-label noise.

Next, we dynamically adjust the temperature for each class. Specifically, we normalize the intra-class compactness c_k via $\hat{c}_k = \frac{c_k - \min(c)}{\max(c) - \min(c)}$, and then apply a linear mapping to obtain class-specific coefficients:

$$\tau_k = \tau_{\min} + (1 - \hat{c}_k)\,(\tau_{\max} - \tau_{\min}) \tag{3}$$

Let $\tau_{\min}$ be the minimum temperature coefficient, representing higher learning difficulty when class transfer is hardest, and $\tau_{\max}$ be the maximum temperature coefficient, representing lower learning difficulty when class transfer is easier. We set $\tau_{\min}$ to 0.6 and $\tau_{\max}$ to 1.2 in our experiments, with a range for $\tau_{\min}$ between 0 and 1, and for $\tau_{\max}$ between 1 and 2. To ensure stability, we apply percentile clipping [14] to suppress extreme variations in τ. The temperature coefficient τ_k dynamically adjusts class-wise learning difficulty. Classes with higher compactness have lower temperatures and receive more attention, while those with lower compactness have higher temperatures and receive less attention. Pseudo labels

are updated via this dynamic temperature. The similarity $r_{t,k,m}$ is adjusted as follows:

$$r'_{t,k,m} = \frac{r_{t,k,m}}{\tau_k} \tag{4}$$

The predicted probability $\hat{p}_t(k)$ is computed from the adjusted similarity.

$$\hat{p}_t(k) = \frac{\exp\left(r'_{t,k}\right)}{\sum_{j=1}^{K} \exp\left(r'_{t,j}\right)} \tag{5}$$

Then, the class with the highest probability is used as the pseudo-label:

$$\hat{y}_t = \arg\max_k \hat{p}_t(k) \tag{6}$$

Thus, the model's predicted class probability $\hat{p}_t(k)$ is obtained. Finally, the pseudo-label loss function is given as follows:

$$L_{dym} = -\sum_{k=1}^{K} 1[\hat{y}_t = k] \log \hat{p}_t(k) \tag{7}$$

This method distinguishes easy and hard classes and adjusts pseudo labels via class-specific τ. It reduces bias in hard classes and improves pseudo-label stability under class imbalance.

D. Multi-center Prototype Contrastive Learning

In the previous section, we introduced a class-difficulty-aware method based on intra-class compactness. However, distribution discrepancies between domains remain a challenge. To address this, we propose a dynamic pseudo-label strategy with multi-center prototype contrastive learning, which improves adaptability by dynamically adjusting class prototypes.

We generate multiple prototypes for each class, $c_k^1, c_k^2, \ldots, c_k^S$, where S denotes the number of prototypes per class. These prototypes represent different clusters in the feature space. We use K-means to cluster target samples and obtain S prototypes per class. The formula for calculating the S-th prototype is as follows:

$$\{\mathbf{c}_{k,1}, \mathbf{c}_{k,2}, \ldots, \mathbf{c}_{k,S}\} = \mathrm{KMeans}\big(\{\hat{g}_t(x_i) \mid \hat{y}_i = k\}\big) \tag{8}$$

These prototypes serve as representative feature points that capture local structure and support contrastive learning. In contrastive learning, we separate positive and negative centers. For each sample, the prototype with the highest similarity is selected as the positive center c^+, and the remaining prototypes form the negative set $c^- = \{c_{k,m} \mid (k,m) \neq (y_t, c^+)\}$. The pseudo-label is defined based on the contrastive objective:

$$\hat{y}_t = \arg\max_k \frac{\exp\left(r^+_{t,k}\right)}{\sum_{j=1}^{K} \exp\left(r^+_{t,j}\right)} \tag{9}$$

$r^+_{t,k}$ denotes the similarity with the positive center c^+_k, and $r^-_{t,k}$ denotes the similarity with the negative center c^-_k. For each sample, define a contrastive loss L_{con} that pulls features toward the positive prototype and pushes them away from others. The loss is defined as:

$$L_{con} = -\sum_{k=1}^{K} \log \frac{\exp\left(r^+_{t,k}\right)}{\sum_{j=1}^{K} \exp\left(r^-_{t,j}\right)} \tag{10}$$

To improve pseudo-label quality and reduce domain discrepancy, we combine the dynamic pseudo-label loss and contrastive loss with a weighted sum.The total loss is defined as:

$$L_{total} = \lambda_1 L_{dym} + \lambda_2 L_{con} \tag{11}$$

λ_1 and λ_2 are balancing hyperparameters. Combining pseudo-label and contrastive losses improves pseudo-label reliability, enhancing robustness.

IV. Experiments

Datasets. We conduct experiments on three datasets: Office-Home [23], with 12 tasks across four domains, 65 classes, and about 15,500 images; VisDA-2017 [24], a synthetic-to-real dataset with 12 classes, containing 152K synthetic source images, 55K real target images, and 72K test images; and Office-31 [25], with 4,652 images from three domains (Amazon, DSLR, Webcam) and 31 classes.

Implementation Details. For a fair comparison, we use the same network architecture and training setup as existing methods. We adopt ResNet-50 and ResNet-101 [29], pre-trained on ImageNet [30], as backbones. The SGD optimizer with 0.9 momentum is applied, and all benchmark batch sizes are set to 64. The learning rate for Office-31 and OfficeHome is 1×10^{-2}, with training for 50 epochs and r set to 3. The dynamic pseudo-label loss is 0.5, and the prototype contrastive loss is 0.2, with VisDA learning rate set to 1×10^{-3}. All experiments are conducted on a GeForce RTX-3090 GPU with PyTorch-3.10.8.

Experiments Analysis. We evaluate our method on VisDA-2017, Office-Home, and Office-31, reporting classification accuracy after entropy adaptation and comparing with UDA and SFDA baselines. We apply our method to SHOT, SHOT++, and NRC, comparing results with and without our approach. Bolded results highlight key benchmarks.

SHOT [8] introduces source hypothesis transfer, adapting the model through information maximization and self-supervised pseudo-labeling. SHOT++ [21] extends SHOT with task-based self-supervised learning and a semi-supervised strategy for improved target domain performance. NRC [22] enforces consistency in feature predictions using local neighborhood structures. While NRC doesn't use pseudo-labeling, it still benefits from combining with our CDME method.

Table I shows that on the VisDA-2017 dataset, our method achieves the highest class accuracy in most categories. It improves SHOT's accuracy from 82.3% to 87.0%, SHOT++ from 85.8% to 87.3%, and NRC from 85.9% to 88.0%. For the "truck" class, our method improves accuracy with minimal decrease in SHOT++, consistent with BMD. Overall, our method outperforms BMD in total accuracy. On VisDA-2017, where class transfer difficulty varies, our method reduces

Table I Per-class accuracy (%) on large-scale visda validation set with resnet-101 as backbone

Methods	SF	plane	bcycl	bus	car	horse	knife	mcyl	person	plant	sktbrd	train	truck	Avg
CDAN [15]	×	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 [16]	×	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 [17]	×	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 [18]	×	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
Source Model	-	65.6	21.5	44.7	73.5	58.8	4.4	77.6	17.8	58.3	42.5	85.1	7.6	46.5
HCL [19]	✓	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
U-SFAN [20]	✓	94.9	87.4	78.0	56.4	93.8	95.1	80.5	79.9	90.1	90.1	85.3	60.4	82.7
SHOT [8]	✓	94.8	82.7	**83.1**	67.9	92.7	96.0	87.6	79.4	88.5	85.2	85.3	43.9	82.3
SHOT w/BMD	✓	96.3	87.4	78.9	63.8	95.0	**97.1**	88.4	83.7	93.2	**88.7**	86.1	**71.2**	85.8
SHOT w/CDME(ours)	✓	**96.9**	**88.8**	82.9	**75.3**	**97.0**	96.3	**91.0**	**85.1**	**95.6**	86.1	**86.4**	63.1	**87.0**
SHOT++ [21]	✓	96.8	86.3	86.7	75.3	96.0	98.2	89.0	**83.0**	95.4	90.9	87.5	44.2	85.8
SHOT++ w/BMD	✓	97.2	86.9	89.0	81.7	96.2	98.2	91.0	82.7	96.4	92.5	92.5	37.5	86.8
SHOT++ w/CDME(ours)	✓	**97.6**	**87.2**	**89.1**	**82.0**	**96.4**	**98.3**	**92.8**	82.7	**97.7**	**92.5**	**93.7**	**37.7**	**87.3**
NRC [22]	✓	97.2	**92.6**	84.4	60.7	96.8	97.0	**91.6**	84.9	94.7	89.0	88.6	55.6	85.9
NRC w/BMD	✓	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
NRC w/CDME(ours)	✓	**97.2**	89.3	**86.7**	**75.9**	**97.9**	**97.4**	91.4	**85.5**	**95.9**	**90.3**	**89.5**	**59.2**	**88.0**

Table II Accuracies (%) on small-sized Office-31 dataset with ResNet-50 as backbone

Method	SF	A→D	A→W	D→A	D→W	W→A	W→D	Avg
DANN [26]	×	79.7	82.0	68.2	96.9	67.4	99.1	82.2
CDAN [15]	×	92.9	82.7	71.0	97.0	67.8	99.8	85.2
Source Model	-	80.3	74.6	61.2	92.5	58.8	94.8	77.0
SHOT [8]	✓	93.2	90.1	72.6	97.7	67.4	99.2	86.7
SHOT w/BMD	✓	93.6	89.2	72.6	**98.2**	67.5	99.6	86.8
SHOT w/CDME(ours)	✓	**93.8**	**90.8**	**75.8**	97.0	**72.0**	**99.8**	**88.2**
SHOT++ [21]	✓	93.4	90.1	73.6	97.4	75.6	99.2	88.2
SHOT++ w/BMD	✓	93.4	90.1	73.6	97.7	75.5	99.6	88.3
SHOT++ w/CDME(ours)	✓	**93.8**	**90.1**	**74.2**	**97.7**	**75.7**	**99.6**	**88.5**
NRC [22]	✓	93.1	90.1	73.9	97.7	75.8	99.6	88.4
NRC w/BMD	✓	93.1	90.5	74.2	97.9	75.9	99.6	88.5
NRC w/CDME(ours)	✓	**96.1**	**92.2**	**74.8**	**98.2**	**75.9**	**99.8**	**89.5**

Table III Accuracies (%) on medium-sized Office-Home dataset with ResNet-50 as backbone

Method	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 [15]	×	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 [27]	×	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 [28]	×	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
Source Model	-	43.9	66.8	73.5	51.8	62.2	64.5	53.5	39.2	73.1	64.2	45.0	77.3	59.6
SHOT [8]	✓	53.4	75.7	79.3	66.1	75.2	75.4	64.7	50.7	**81.9**	72.5	**58.6**	84.0	69.8
SHOT w/BMD	✓	53.8	**76.3**	79.8	**66.5**	76.5	76.5	64.8	50.9	80.1	**72.8**	57.4	84.5	70.0
SHOT w/CDME(ours)	✓	**54.6**	74.4	**80.7**	66.0	**78.8**	**77.4**	**66.0**	**52.4**	81.5	70.6	58.1	**84.5**	**70.4**
SHOT++ [21]	✓	54.2	76.4	81.4	66.1	78.0	78.7	65.0	**55.1**	80.9	72.2	58.6	84.0	71.9
SHOT++ w/BMD	✓	56.7	78.9	**81.5**	66.8	78.6	78.9	66.7	54.9	83.1	73.5	58.7	84.5	71.9
SHOT++ w/CDME(ours)	✓	**57.8**	**79.0**	81.3	**66.8**	**78.9**	**78.9**	**67.0**	54.9	**83.7**	**73.5**	**58.7**	**84.8**	**72.1**
NRC [22]	✓	57.2	80.1	**82.1**	67.7	**78.9**	77.7	65.3	56.5	82.5	72.9	**59.0**	84.5	72.0
NRC w/BMD	✓	57.2	80.9	82.0	68.3	78.6	78.9	67.0	56.4	82.6	72.9	58.9	86.4	72.5
NRC w/CDME(ours)	✓	**58.9**	**81.2**	81.9	**68.5**	78.8	**78.9**	**67.0**	**57.1**	**82.9**	**74.9**	58.9	**87.0**	**73.0**

pseudo-label bias in hard classes, improving pseudo-label quality and stability.

Table II shows results on the Office-Home benchmark, where our method outperforms at least three out of six SFDA tasks. Specifically, in the A→W task, our method significantly improves adaptation and surpasses traditional domain adaptation methods requiring source data. Despite the smaller dataset size, our method provides more accurate task estimates, demonstrating its broad applicability.

Table III shows task-wise and overall accuracy on Office-Home. Our method improves performance across SFDA models, increasing accuracy from 69.8% to 70.4% for SHOT, 71.9% to 72.1% for SHOT++, and 72.0% to 73.0% for NRC. Our method demonstrates strong generalization and effectively alleviates performance issues in hard-to-transfer classes.

Ablation Experiment. Our method consists of a class-difficulty-aware strategy and a dynamic pseudo-label strategy with multi-center prototype contrastive learning. We conduct ablation studies on Office-Home, VisDA-2017, and Office-31 under SHOT. As shown in Fig. 3, we incrementally add components to SHOT: first the BMD multi-center prototype strategy, then class-difficulty-aware pseudo-label loss and multi-center prototype contrastive loss. Specifically, we add basic pseudo-label loss, followed by the multi-center strategy with dynamic pseudo-label adjustment to evaluate their contributions.

V. Conclusions

We propose a Class-Difficulty-aware and Minority-class Enhancement (CDME) framework that adjusts pseudo-label confidence based on intra-class compactness to reduce errors in hard-to-transfer classes. A dynamic pseudo-label strategy with multi-center prototype contrastive learning further improves label reliability and intra-class consistency. Experiments demonstrate that CDME outperforms existing SFDA methods, particularly under class bias in the VisDA target domain, achieving better adaptability and overall performance.

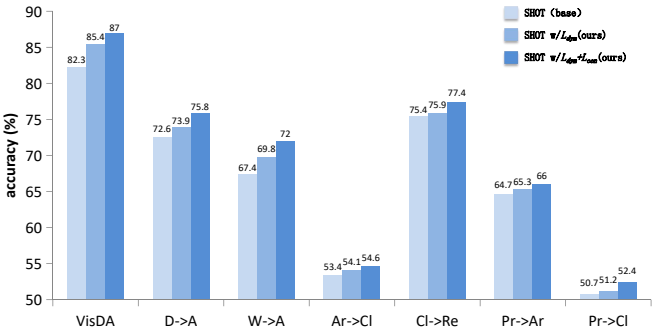

Fig. 3: Performance changes (%) on partial tasks of Office-31, Office-Home, and VisDA

References

[1] Z. Zhu, L. Fan, M. Pagnucco, and Y. Song, "Interpretable image classification via non-parametric part prototype learning," in *Proceedings of the Computer Vision and Pattern Recognition Conference*, 2025, pp. 9762–9771.

[2] G. Yang, Y. Wang, D. Shi, and Y. Wang, "Golden cudgel network for real-time semantic segmentation," in *Proceedings of the Computer Vision and Pattern Recognition Conference*, 2025, pp. 25 367–25 376.

[3] S. Liu, J. Lv, J. Kang, H. Zhang, Z. Liang, and S. He, "Modfinity: Unsupervised domain adaptation with multimodal information flow intertwining," in *Proceedings of the Computer Vision and Pattern Recognition Conference*, 2025, pp. 5092–5101.

[4] G. Xu, H. Guo, L. Yi, C. Ling, B. Wang, and G. Yi, "Revisiting source-free domain adaptation: a new perspective via uncertainty control," in *The Thirteenth International Conference on Learning Representations*, 2024.

[5] H. Liu, J. Wang, and M. Long, "Cycle self-training for domain adaptation," *Advances in Neural Information Processing Systems*, vol. 34, pp. 22 968–22 981, 2021.

[6] H. Li, R. Zhang, H. Yao, X. Zhang, Y. Hao, X. Song, S. Peng, Y. Zhao, C. Zhao, Y. Wu *et al.*, "Seen-da: Semantic entropy guided domain-aware attention for domain adaptive object detection," in *Proceedings of the Computer Vision and Pattern Recognition Conference*, 2025, pp. 25 465–25 475.

[7] S. Laine and T. Aila, "Temporal ensembling for semi-supervised learning," *arXiv preprint arXiv:1610.02242*, 2016.

[8] J. Liang, D. Hu, and J. Feng, "Do we really need to access the source data? source hypothesis transfer for unsupervised domain adaptation," in *International conference on machine learning*. PMLR, 2020, pp. 6028–6039.

[9] S. Wan, Y. Zhan, S. Chen, S. Pan, J. Yang, D. Tao, and C. Gong, "Boosting graph contrastive learning via adaptive sampling," *IEEE Transactions on Neural Networks and Learning Systems*, 2023.

[10] H. Sun, Y. Zhang, L. Xu, S. Jin, P. Luo, C. Qian, W. Liu, and Y. Chen, "Unsupervised continual domain shift learning with multi-prototype modeling," in *Proceedings of the Computer Vision and Pattern Recognition Conference*, 2025, pp. 10 131–10 141.

[11] Z. Wang, H. Zhu, B. Huang, Z. Wang, W. Lu, N. Chen, and Y. Wang, "M-msseu: source-free domain adaptation for multi-modal stroke lesion segmentation using shadowed sets and evidential uncertainty," *Health Information Science and Systems*, vol. 11, no. 1, p. 46, 2023.

[12] J.-X. Shi, T. Wei, Y. Xiang, and Y.-F. Li, "How re-sampling helps for long-tail learning?" *Advances in Neural Information Processing Systems*, vol. 36, pp. 75 669–75 687, 2023.

[13] S. Shimizu, S. Yada, L. Raithel, and E. Aramaki, "Improving self-training with prototypical learning for source-free domain adaptation on clinical text," in *Proceedings of the 23rd Workshop on Biomedical Natural Language Processing*, 2024, pp. 1–13.

[14] I. Merad and S. Gaïffas, "Robust stochastic optimization via gradient quantile clipping," *arXiv preprint arXiv:2309.17316*, 2023.

[15] M. Long, Z. Cao, J. Wang, and M. I. Jordan, "Conditional adversarial domain adaptation," *Advances in neural information processing systems*, vol. 31, 2018.

[16] C.-Y. Lee, T. Batra, M. H. Baig, and D. Ulbricht, "Sliced wasserstein discrepancy for unsupervised domain adaptation," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019, pp. 10 285–10 295.

[17] Y. Jin, X. Wang, M. Long, and J. Wang, "Minimum class confusion for versatile domain adaptation," in *European conference on computer vision*. Springer, 2020, pp. 464–480.

[18] Z. Lu, Y. Yang, X. Zhu, C. Liu, Y.-Z. Song, and T. Xiang, "Stochastic classifiers for unsupervised domain adaptation," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 9111–9120.

[19] J. Huang, D. Guan, A. Xiao, and S. Lu, "Model adaptation: Historical contrastive learning for unsupervised domain adaptation without source data," *Advances in neural information processing systems*, vol. 34, pp. 3635–3649, 2021.

[20] S. Roy, M. Trapp, A. Pilzer, J. Kannala, N. Sebe, E. Ricci, and A. Solin, "Uncertainty-guided source-free domain adaptation," in *European conference on computer vision*. Springer, 2022, pp. 537–555.

[21] J. Liang, D. Hu, Y. Wang, R. He, and J. Feng, "Source data-absent unsupervised domain adaptation through hypothesis transfer and labeling transfer," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 11, pp. 8602–8617, 2021.

[22] S. Yang, J. Van de Weijer, L. Herranz, S. Jui *et al.*, "Exploiting the intrinsic neighborhood structure for source-free domain adaptation," *Advances in neural information processing systems*, vol. 34, pp. 29 393–29 405, 2021.

[23] H. Venkateswara, J. Eusebio, S. Chakraborty, and S. Panchanathan, "Deep hashing network for unsupervised domain adaptation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 5018–5027.

[24] X. Peng, B. Usman, N. Kaushik, J. Hoffman, D. Wang, and K. Saenko, "Visda: The visual domain adaptation challenge," *arXiv preprint arXiv:1710.06924*, 2017.

[25] J. Hoffman, E. Tzeng, T. Park, J.-Y. Zhu, P. Isola, K. Saenko, A. Efros, and T. Darrell, "Cycada: Cycle-consistent adversarial domain adaptation," in *International conference on machine learning*. PMLR, 2018, pp. 1989–1998.

[26] Y. Ganin, E. Ustinova, H. Ajakan, P. Germain, H. Larochelle, F. Laviolette, M. March, and V. Lempitsky, "Domain-adversarial training of neural networks," *Journal of machine learning research*, vol. 17, no. 59, pp. 1–35, 2016.

[27] S. Cui, S. Wang, J. Zhuo, L. Li, Q. Huang, and Q. Tian, "Towards discriminability and diversity: Batch nuclear-norm maximization under label insufficient situations," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 3941–3950.

[28] S. Cui, S. Wang, J. Zhuo, C. Su, Q. Huang, and Q. Tian, "Gradually vanishing bridge for adversarial domain adaptation," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 12 455–12 464.

[29] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[30] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE conference on computer vision and pattern recognition*. IEEE, 2009, pp. 248–255.