

Evidential Graph Contrastive Alignment for Source-Free Blending-Target Domain Adaptation

Juepeng Zheng[✉], Yibin Wen

School of Artificial Intelligence, Sun Yat-Sen University
zhengjp8@mail.sysu.edu.cn, wenyb5@mail2.sysu.edu.cn

Abstract—In this paper, we firstly tackle a more realistic Domain Adaptation (DA) setting: Source-Free Blending-Target Domain Adaptation (SF-BTDA), where we *can not* access to source domain data while facing mixed multiple target domains without any domain labels in prior. Compared to existing DA scenarios, SF-BTDA generally faces the co-existence of different label shifts in different targets, along with noisy target pseudo labels generated from the source model. In this paper, we propose a new method called Evidential Graph Contrastive Alignment (EGCA) to decouple the blending target domain and alleviate the effect from noisy target pseudo labels. First, to improve the quality of pseudo target labels, we propose a Calibrated Evidential Learning module to iteratively improve both the accuracy and certainty of the resulting model and adaptively generate high-quality pseudo target labels. Second, we design a Graph Contrastive Learning with the domain distance matrix and confidence-uncertainty criterion, to minimize the distribution gap of samples of a same class in the blended target domains, which alleviates the co-existence of different label shifts in blended targets. We conduct a new benchmark based on three standard DA datasets and EGCA outperforms other methods with considerable gains and achieves comparable results compared with those have domain labels or source data in prior.

Index Terms—source-free domain adaptation, blending target, evidential learning, contrastive learning

I. INTRODUCTION

Despite the great success of deep learning in various fields [1], it remains a challenge to achieve a generalized deep learning model that can perform well on unseen data, especially given the vast diversity of real-world data and problems. Performance degradation occurs due to various differences between the training data and new test data, for aspects such as distribution, dimension, context, etc. Domain adaptation (DA) offers significant benefits in this scenario by adapting the pre-trained models towards new domains with different distributions and properties from the original domain. Moreover, DA provides great potential for efficiency improvement by reducing the need for retraining on new data from scratch, which has been widely applied in various fields and real-world cases.

According to the number of target domains in DA tasks [2], [3], there are two main types of DA: Single-Target DA (STDA) (see Figure 1(a)) and Multi-Target DA (MTDA) (see Figure 1(c)). While most existing DA works focus on STDA [4]–[6], MTDA is a more challenging task that simultaneously transfer one source domain to multiple target domains. For example, MTDA-ITA [7] aims to find a shared latent space common

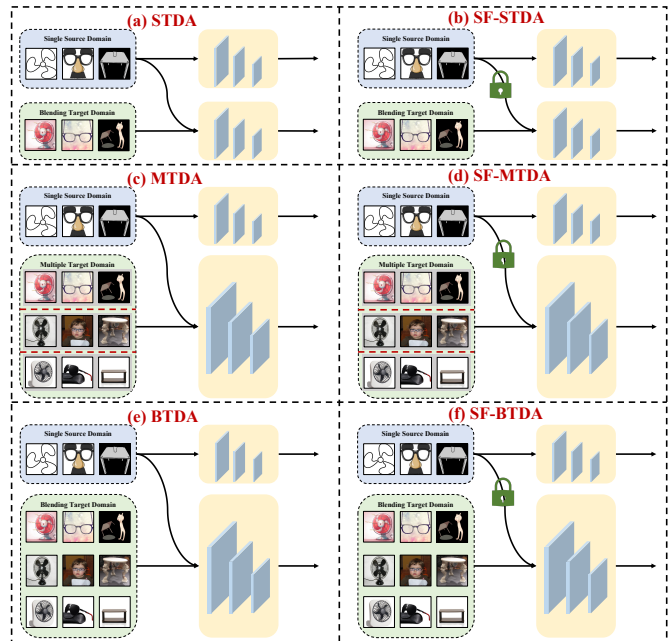


Fig. 1. Different DA settings. (a) Single Target Domain Adaptation (STDA). (b) Source-Free STDA (SF-STDA). (c) Multiple Target Domain Adaptation (MTDA). (d) Source-Free MTDA (SF-MTDA). (e) Blending-Target Domain Adaptation (BTDA). (f) Source-Free BTDA (SF-BTDA), a new DA setting proposed in this paper, with only access to source trained model.

to all domains to allow adaptation from a single source to multiple target domains. However, all aforementioned works have domain labels, which provides an important reference for deep learning models. Real-world tasks are much more complex. For instance, in object recognition for autonomous driving technology, various changes may introduce new domains into the task, such as sensor device, street view, weather, lighting, time and so on. Therefore, DA tasks in real-world applications often do not know their domain labels in prior, and the data from various target domains are mixed together. To our knowledge, there are only some works focusing on this mixed multi-target DA without domain labels, which is crucial for real-world applications [8]–[11]. Therefore, this paper makes efforts on the more practical and challenging Blending-Target Domain Adaptation (BTDA) scenario (see Figure 1(e)).

On the other hand, there is another crucial issue in real-world applications for DA, that is whether the source data

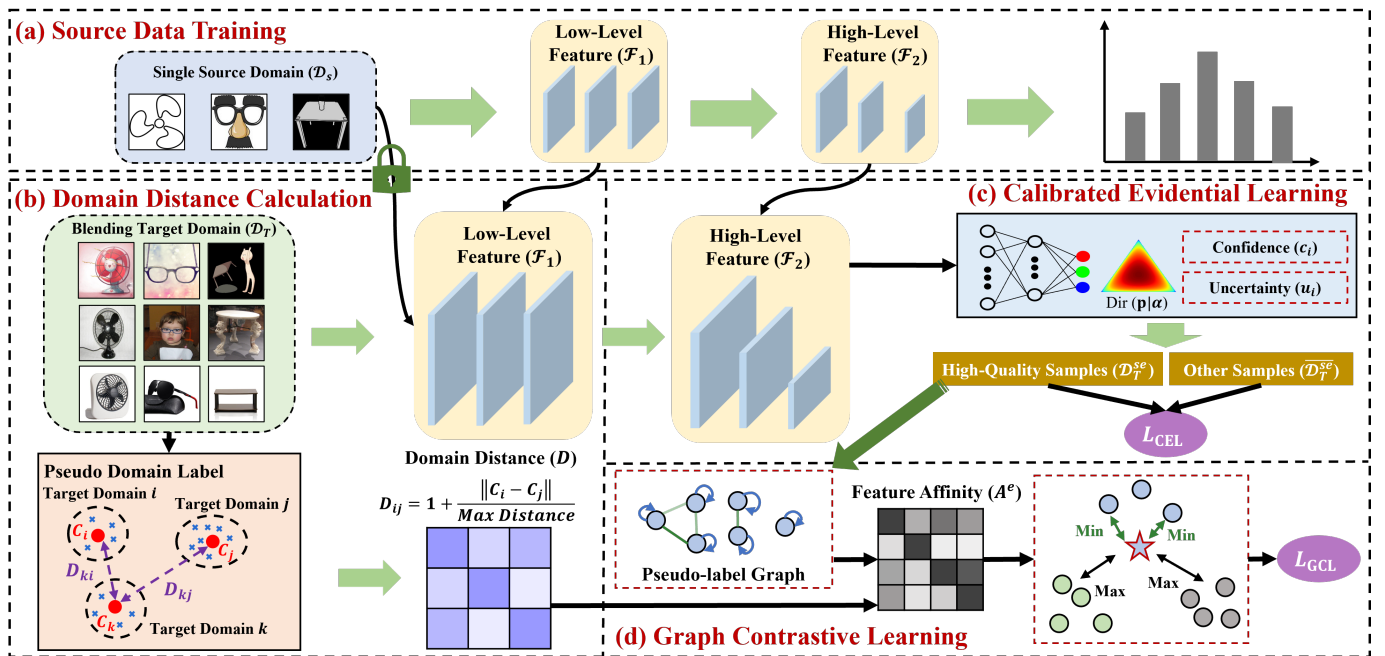


Fig. 2. The framework for our proposed Evidential Graph Contrastive Alignment (EGCA). (a) **Source Data Training**: We train the model through single source domain and we are not access to the source domain data. (b) **Domain Distance Calculation**: We generate the mixture domain distance through low-level features and original image textures in an unsupervised way. (c) **Calibrated Evidential Learning (CEL)**: We propose a calibrated evidential learning module to iteratively improve both the accuracy and certainty of the resulting model and adaptively generate high-quality pseudo target label (addressing **Challenge 1**). (d) **Graph Contrastive Learning (GCL)**: We design a graph contrastive learning with the domain distance matrix and confidence-uncertainty criterion, to minimize the distribution gap of samples of a same class in the blended target domains, which alleviates the co-existence of different label shifts in blended targets (addressing **Challenge 2**).

is available. Owing to increasing privacy concerns (such as medical images, mobile applications, surveillance videos, etc.), source-domain data can not be always accessed. Recently, Source-Free Domain Adaptation (SFDA) is introduced to transfer knowledge from a prior source domain to a new target domain without accessing the source data. As shown in Figure 1, we could define different SFDA settings according to the number of target domains in the DA task. However, most existing SFDA methods focus on SF-STDA setting (see Figure 1(b)) [12], [13], while only [14] have discussed and designed new approach for SF-MTDA (see Figure 1(d)).

Different from other DA scenarios, we propose a new DA setting named Source-Free Blending-Target Domain Adaptation (SF-BTDA), where we have to transfer the knowledge from a single source to blending-target domain without domain labels nor the access of labeled source data (see Figure 1(f)). There are two major challenges in SF-BTDA: **Challenge 1**: facing a mixture of multiple target domains, the model is assumed to handle the co-existence of different label shifts in different targets with no prior information; **Challenge 2**: as we are only accessed to the source model, precisely alleviating the negative effect from noisy target pseudo labels is difficult because of the significantly enriched styles. Due to these two challenges, existing DA methods that count on utilizing distribution shifts between two specific domains become infeasible.

In this paper, we first present SF-BTDA setting and propose Evidential Graph Contrastive Alignment (EGCA) to better

tackle the aforementioned two challenges in SF-BTDA scenarios. The contributions of this work are as following:

- We propose a novel DA task named SF-BTDA, where we have to transfer the knowledge from a single source to blending-target domain without domain labels nor the access of labeled source data.
- We propose Calibrated Evidential Learning to iteratively improve both the accuracy and certainty of the resulting model and adaptively select high-quality pseudo target labels according to the balanced accuracy and certainty.
- We design Graph Contrastive Learning with the domain distance matrix and confidence-uncertainty criterion, to minimize the distribution gap of samples of a same class in the blended-target domain.

The second contribution mainly address the **Challenge 2**, improving the quality of pseudo target labels in the blending target domain to reduce the effect of the enriched styles and textures in the mixture of multiple target domains. The third contribution effectively tackles the **Challenge 1**, forcefully pulling samples that are from the same class to alleviate the co-existence of different label shifts in blended targets.

II. METHODOLOGY

A. Problem Setting and Notations

In contrast to MTDA, BTDA is established on a mixture of target domains $\mathcal{D}_T = \left\{ \mathcal{D}_t^j \right\}_{j=1}^k = \{(\mathbf{x}_i^t)\}_{i=1}^{n_T}$, where k is

the number of target domains and $n_T = \sum_{j=1}^k n_t^j$ denotes the total quantity of images in the target domain. Unlike MTDA scenario, the proportions of different target domains in the mixed datasets $\{w^j\}_{j=1}^k$ are unknown. In SF-BTDA setting, we consider labeled source domain data with n_s , the samples as $\mathcal{D}_s = \{(\mathbf{x}_i^s, \mathbf{y}_i^s)\}_{i=1}^{n_s}$, in which the \mathbf{y}_i^s is the corresponding label of \mathbf{x}_i^s , is only available during model pretraining. Our approach is based on any backbone, which we split into three parts: a shallow feature extractor \mathcal{F}_1 , a deep feature extractor \mathcal{F}_2 and a classifier \mathcal{G} . If we directly adopt existing SF-STDA algorithms and consider the mixed target domains as one target domain in a brute-force way, the training objective will facilitate domain-invariant representations to align the whole blending-target domain \mathcal{D}_T rather than k target domains $\{\mathcal{D}_t^j\}_{j=1}^k$. Because of the discrepancy among sub-subjects from k distributions, adaptation for the whole blending-target domain directly may lead to severe category misalignment.

B. Domain Distance Calculation

As introduced in Sec. I, SF-BTDA issue faces a mixture of multiple target domains, the model is assumed to handle the co-existence of different label shifts in different targets with no prior information (**Challenge 1**). Therefore, some samples in the same class may appear different styles and textures because of co-existence multiple targets with significantly enriched styles and textures. Under this consideration, we calculate the domain distance to better guide the contrastive learning (see Sec. II-D) to pull the samples in the same class that may appear different styles and textures. As we are not access to the specific domain label for the blending-target domain, we separate the target domains into k domains and assign pseudo domain labels for them using an unsupervised way (*i.e.*, k -means). Notably, we adopt the extracted shallow features from DNNs ($\mathcal{F}_1(\mathbf{x}^t)$) with the dimension of $56 \times 56 \times 256$ and simply concatenate them with original image textures (\mathbf{x}^t) to achieve pseudo domain assignment for better satisfying the specific image style as well as the universal feature style.

We calculate the domain distance D_{ij} between two samples \mathbf{x}_i^t and \mathbf{x}_j^t through Eq. (1), in which C_i and C_j are the corresponding centroids for \mathbf{x}_i^t and \mathbf{x}_j^t . The centroid for an target image (\mathbf{x}_i^t) denotes it belongs to the pseudo domain label. The *Max Distance* denotes we normalize the domain distance among different target domain centroids. If \mathbf{x}_i^t and \mathbf{x}_j^t are from the same domain, the domain distance will be 1.0. Here, we utilize the distance of domain centroids to represent the samples from different target domains for simplicity. We update the domain distance after each training epoch.

$$D_{ij} = 1 + \frac{\|C_i - C_j\|}{\text{Max Distance}} \quad (1)$$

C. Calibrated Evidential Learning (CEL)

In the meantime, as we are also not access to the source data, we could only achieve pseudo target labels for the blending-target images using the model trained from labeled source images. Because of **Challenge 2**, directly adopting pseudo

labels for the target domain may ruin and mislead the model training. To alleviate the negative effect from noisy target labels, we propose calibrated evidential learning to iteratively improve both the accuracy and certainty of the resulting model, and adaptively select high-quality pseudo target labels.

Evidential Deep Learning (EDL) [15] offers a principled way to formulate the uncertainty estimation, which overcomes the shortcoming of softmax-based DNNs. Assuming that the class probability follows a prior Dirichlet distribution [16], we can acquire the predictive uncertainty u for each sample: $u = M/S$, in which M is the number of classes, and $S = \sum_{m=1}^M \alpha_m$ is the total evidence. α_m is the non-negative network prediction output and can be expressed $\alpha_m = \mathbf{e}_m + 1$, where $\mathbf{e}(i) = g(\mathcal{F}_2(\mathcal{F}_1(\mathbf{x}(i))))$ and g is the evidence function.

Although the evidential uncertainty from EDL can be directly learned in the target domain samples, the uncertainty may not be well calibrated to address our target pseudo label learning in SF-BTDA scenario. According to existing model calibration literature [17], a better calibrated model is supposed to be uncertain in its predictions when being inaccurate, while be confident about accurate ones. In this section, we propose to calibrate the EDL model by considering the relationship between the accuracy and uncertainty for the blending-target domain under source-free setting.

In particular, we propose Calibrated Evidential Learning (CEL) to minimize the following sum of samples that are accurate but uncertainty, or inaccurate but certain, using the logarithm constraint between confidence c_i and uncertainty u_i :

$$L_{\text{CEL}} = -\lambda_t \sum_{\mathbf{x}_i \in \mathcal{D}_T^{\text{se}}} c_i \log(1 - u_i) - (1 - \lambda_t) \sum_{\mathbf{x}_i \in \overline{\mathcal{D}_T^{\text{se}}}} (1 - c_i) \log(u_i) \quad (2)$$

in which u_i denotes the evidential uncertainty of an input sample \mathbf{x}_i and c_i is the associated maximum class probability. $\mathcal{D}_T^{\text{se}}$ and $\overline{\mathcal{D}_T^{\text{se}}}$ are the selected high-quality target samples and the remainder in the target domain (low-quality target samples), respectively. That is, $|\mathcal{D}_T| = |\mathcal{D}_T^{\text{se}}| + |\overline{\mathcal{D}_T^{\text{se}}}|$.

Different from other high-quality selection schemes in SFDA setting, we utilize the evidential uncertainty and prediction confidence to measure the label reliability simultaneously. That is, the high-quality target pseudo labels both satisfy that $c_i > \eta_c$ and $u_i < \eta_u$, where η_c and η_u are two selection thresholds that can be adaptively estimated in each mini-batch. To this end, we do not need to set fixed hyper-parameters for η_c and η_u .

$$\eta_c = \frac{1}{B} \sum_{i=1}^{i=B} c_i; \quad \eta_u = \frac{1}{B} \sum_{i=1}^{i=B} u_i \quad (3)$$

The first term of L_{CEL} in Eq. (2) tries to give low uncertainty ($u_i \rightarrow 0$) when the model makes accurate prediction ($\mathbf{x}_i \in \mathcal{D}_T^{\text{se}}, c_i \rightarrow 1$) for the selected high-quality target samples, while the second term of L_{CEL} aims to give high uncertainty ($u_i \rightarrow 1$) when the model makes inaccurate prediction

($\mathbf{x}_i \in \overline{\mathcal{D}}_T^{sc}, c_i \rightarrow 0$) for the low-quality target samples. Notably, γ_t is the annealing factor that will be exponentially increasing from γ_0 to 1.0 during training.

To this end, through our proposed calibrated evidential learning, the model tries to make high accuracy and low uncertainty for the high-quality target domain samples, while make a restraint for low-quality target domain samples, which addresses the **Challenge 2** introduced in Sec. I. This strategy helps us to better achieve the knowledge transfer from the source model to the blending-target domain.

D. Graph Contrastive Learning (GCL)

The original intention for contrastive learning is to push each image far from others and pull the augmented images closer with the original one in the feature space. However, it is not compatible with the classification task which requires to cluster images at the class level, especially that we have some high-quality target samples generated from Sec. II-C. Furthermore, according to the description of **Challenge 1**, some samples in the same class may appear different styles and textures because of co-existence multiple targets with significantly enriched styles and textures. To this end, we design the domain distance calculated from Sec. II-B in graph contrastive learning to pay more attention for abovementioned samples, which could solve the conflict between contrastive learning and the classification task.

Since Sec. II-C has generated high-quality labels for target domain, we could fully explore their advantages in contrastive learning. Notably, our contrastive alignment module aims to learn representations guided by a pseudo-label graph. We build the pseudo-label graph by constructing a similarity matrix (A). To this end, each element a_{ij} in A has the following format:

$$a_{ij} = \begin{cases} 1 & \text{if } j = i, \\ 1 & \text{if } z_i \text{ and } z_j \text{ are from same category,} \\ & \text{and } \mathbf{x}_i \in \mathcal{D}_T^{se} \text{ and } \mathbf{x}_j \in \mathcal{D}_T^{se}, \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

Notably $\mathbf{x}_i \in \mathcal{D}_T^{se}$ and $\mathbf{x}_j \in \mathcal{D}_T^{se}$ denotes that \mathbf{x}_i and \mathbf{x}_j are selected from the high-quality target domain samples. Although high-quality pseudo label-based supervised contrastive learning is theoretically functional, directly using these pseudo labels may be problematic because they might be noisy labels, too. To alliviate this problem, we adopt the confidence-uncertainty evaluation produced from Sec. II-C to strengthen the robustness and reduce the negative effect on noisy pseudo labels. Actually, the pseudo-label graph (A) serves as the target to train an embedding graph (A^e), which is defined as:

$$a_{ij}^e = \begin{cases} a_{ij} * I_{ij} * D_{ij} & \text{if } i \neq j, \\ a_{ij} & \text{otherwise} \end{cases} \quad (5)$$

in which $I_{ij} = c_i * c_j * (1 - u_i) * (1 - u_j)$ denotes that samples with high confidence and low uncertainty will have higher weights in our evidential contrastive alignment

module. D_{ij} can be calculated from Eq. (1), which could make our contrastive learning module to pay more attention on the samples of a same class that belonged to different target domains (addressing **Challenge 1**). To this end, the loss of L_{GCL} can be formulated as:

$$L_{GCL} = - \sum_{i \in \mathcal{I}} \frac{1}{1 + |\mathcal{P}(i)|} \left\{ \log \frac{\exp(z_i \cdot z_{j(i)}/\tau)}{\sum_{q \in Q(i)} \exp(z_i \cdot z_q/\tau)} \right. \\ \left. - \sum_{p \in \mathcal{P}(i)} \log \frac{a_{ij}^e \cdot \exp(z_i \cdot z_p/\tau)}{\sum_{q \in Q(i)} \exp(z_i \cdot z_q/\tau)} \right\} \quad (6)$$

in which $\mathcal{P}(i)$ denotes indices of the views from other images of the same class in the selected high-quality target domain samples. $|\mathcal{P}(i)|$ denotes its number and $|\mathcal{P}(i)| + 1$ denotes all positive pairs. τ is the temperature parameter. Other notations could refer to [27]. As our setting is SF-BTDA, it is noteworthy that our graph contrastive learning module conducts on the blending target owing to its difficulties from classification confusions because of co-existence of different label shifts (addressing **Challenge 1**).

E. Overall Objective

The model is jointly optimized with two loss terms, including CEL loss L_{CEL} and GCL loss L_{GCL} :

$$L = L_{CEL} + \beta L_{GCL} \quad (7)$$

where β is a trade-off parameter balancing two different loss terms between L_{CEL} , and L_{GCL} .

III. EXPERIMENTS

A. Datasets

We conduct experiments on three standard domain adaptation benchmarks. **Office-31** [28] consists of 4,652 images from three domains: DSLR (D), Amazon (A), and Webcam (W). **OfficeHome** [29] is a more challenging dataset, which consists of 15,500 images in total from 65 categories of everyday objects. There are four significantly different domain: Artistic images (Ar), Clip-Art images (Cl), Product images (Pr), and Real-World images (Rw). **DomainNet** [30] is the most challenging and very large scale DA benchmark, which has six different domains: Clipart (Cli), Infograph (Inf), Painting (Pai), Quickdraw (Qui), Real (Rea) and Sketch (Ske). It has around 0.6 million images, including both train and test images, and has 345 different object categories.

B. Training Details

Our methods were implemented based on the PyTorch. We adopt ResNet-50 [18] for **Office-31** and **OfficeHome**, and ResNet-101 for **DomainNet**, both of which pretrained on the ImageNet dataset. Whatever module trained from scratch, its learning rate was set to be 10 times that of the lower layers. We adopt mini-batch stochastic gradient descent (SGD) with momentum of 0.95 using the learning rate and progressive training strategies. We set β in Eq (7) as 1.0.

State-of-the-art. ResNet [18] is the baseline backbone without any domain adaptation tricks. We list some methods that

TABLE I

OVERALL ACCURACY (%) ON **OFFICE-31** AND **OFFICEHOME** FOR FOR STATE-OF-THE-ART METHODS. ALL METHODS USE THE RESNET-50 AS THE BACKBONE. THE RESULTS ON EACH DATASET ARE THE AVERAGE OF THREE/FOUR LEAVE-ONE-DOMAIN-OUT CASES.

Method	Multiple Targets	Domain Labels	Source Data	Office-31				OfficeHome				
				A	D	W	Avg	Ar	Cl	Pr	Rw	Avg
ResNet-50 [18]	×	×	×	76.3	68.7	67.0	70.7	62.5	61.2	55.1	61.8	60.1
DAN [19]	×	×	✓	79.5	80.3	81.2	80.3	58.4	58.1	52.9	62.1	57.9
DANN [4]	×	×	✓	78.2	72.2	69.8	73.4	58.4	58.1	52.9	62.1	57.9
CDAN [20]	×	×	✓	93.6	80.5	81.3	85.1	59.5	61.0	54.7	62.9	59.5
MTDA-ITA [7]	✓	✓	✓	87.9	83.7	84.0	85.2	64.6	66.4	59.2	67.1	64.3
DCL [21]	✓	✓	✓	92.6	82.5	84.7	86.6	63.0	63.0	60.0	67.0	63.3
DCGCT [10]	✓	✓	✓	93.4	86.0	87.1	88.8	70.5	71.6	66.0	71.2	69.8
AMEAN [8]	✓	×	✓	90.1	77.0	73.4	80.2	64.3	65.5	59.5	66.7	64.0
CGCT [10]	✓	×	✓	93.9	85.1	85.6	88.2	67.4	68.1	61.6	68.7	66.5
DML [22]	✓	×	✓	<u>94.9</u>	85.3	86.2	88.8	67.0	70.6	62.9	69.0	67.4
MCDA [23]	✓	×	✓	92.4	<u>87.7</u>	88.8	89.6	71.7	72.8	68.0	71.7	71.1
G-SFDA [12]	×	×	×	90.6	78.9	77.4	82.3	70.3	74.9	66.8	70.0	70.5
GPUE [24]	×	×	×	78.8	72.4	60.7	70.6	50.5	45.5	46.4	52.9	48.8
SHOT++ [13]	×	×	×	93.3	79.6	78.4	83.8	70.6	74.2	66.2	68.6	69.9
ConMix [14]	✓	✓	×	92.4	81.8	80.4	84.9	75.6	81.4	71.4	73.4	75.4
EGCA (ours)	✓	×	×	93.5	84.4	81.3	86.4	76.5	82.2	72.2	73.6	76.1

TABLE II

OVERALL ACCURACY (%) ON **DOMAINNET** FOR FOR STATE-OF-THE-ART METHODS. ALL METHODS USE THE RESNET-101 AS THE BACKBONE. THE RESULTS ON EACH DATASET ARE THE AVERAGE OF SIX LEAVE-ONE-DOMAIN-OUT CASES.

Method	Multiple Targets	Domain Labels	Source Data	Cli	Inf	Pai	Qui	Rea	Ske	Avg
SE [5]	×	×	✓	21.3	8.5	14.5	13.8	16.0	19.7	15.6
MCD [25]	×	×	✓	25.1	19.1	27.0	10.4	20.2	22.5	20.7
CDAN [20]	×	×	✓	31.6	27.1	31.8	12.5	33.2	35.8	28.7
MCC [26]	×	×	✓	33.6	30.0	32.4	13.5	28.0	35.3	28.8
DCL [21]	✓	✓	✓	35.1	31.4	37.0	20.5	35.4	41.0	33.4
DCGCT [10]	✓	✓	✓	37.0	32.2	37.3	19.3	39.8	40.8	34.4
DADA [9]	✓	×	✓	26.4	20.0	26.5	12.9	20.7	22.8	21.6
CGCT [10]	✓	×	✓	36.1	33.3	35.0	10.0	39.6	39.7	32.3
DML [22]	✓	×	✓	32.0	25.4	29.4	12.7	31.5	36.4	27.9
MCDA [23]	✓	×	✓	37.5	<u>37.3</u>	36.6	17.8	36.1	41.4	34.5
G-SFDA [12]	×	×	×	34.2	26.7	32.7	11.7	18.7	32.0	26.0
GPUE [24]	×	×	×	23.4	22.7	31.4	16.5	16.9	24.8	22.6
SHOT++ [13]	×	×	×	32.9	28.8	32.1	13.9	17.9	34.9	26.7
ConMix [14]	✓	✓	×	41.8	29.2	39.9	17.5	32.7	41.2	33.7
EGCA (ours)	✓	×	×	42.2	32.6	40.3	18.0	35.5	41.8	34.6

proposed for STDA scenario, such as DAN [19], DANN [4], RTN [31], SE [5], MCD [25], CDAN [20] and MCC [26]. We also compare our method with existing three MTDA methods (*i.e.*, MTDA-ITA [7], DCL [21], DCGCT [10]) and five BTDA methods (*i.e.*, AMEAN [8], DADA [9], CGCT [10], DML [22] and MCDA [23]). Furthermore, we compare three source-free DA methods, including G-SFDA [12], GPUE [24] and SHOT++ [13]. However, these SFDA methods mainly focus on SF-STDA settings. Finally, we compare our performance with ConMix [14], which is the only one that concentrates on SF-MTDA scenarios with domain labels in prior while without the access to the labels of source data.

C. Results

The **bold** numbers in Table I and II denote the highest accuracy without source data. The underline numbers denote the highest accuracy without domain labels. According to Table I and II, our EGCA overpasses these comparison methods without domain labels nor source data on all transfer tasks.

Notably, MCDA and CGCT must use source data in prior. However, our EGCA *does not* use the source data nor domain labels. Therefore, it is surprising that EGCA outperforms these methods for **Office-Home** and **DomainNet**. In addition, EGCA neither has source data nor domain labels, but it even outperforms ConMix that adopts domain labels during training on all datasets. Therefore, EGCA outperforms other methods with considerable gains compared with those that have domain labels or source data in prior.

IV. DISCUSSION

A. Ablation studies

As listed in Fig. 3, L_{CON} achieves 2.5% and 3.8% improvement for **Office-31** and **OfficeHome**, respectively. In addition, L_{GCL} further increase 1.1% and 1.4% for average accuracy. In addition, Table III lists ablation results for different high-quality selection schemes. Therefore, we both utilize the evidential uncertainty and prediction confidence to measure

TABLE III
ACCURACY ON DIFFERENT SAMPLE SELECTION SCHEMES.

Method	Office-31	Office-Home	DomainNet
only η_u	85.9	74.3	32.8
only η_c	85.6	75.5	32.2
both η_u and η_c	86.4	76.1	34.6

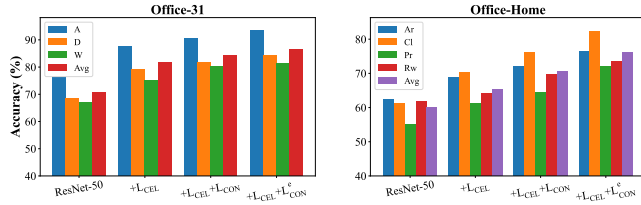


Fig. 3. Overall Accuracy on **Office-31** and **OfficeHome** for ablation studies.

the label reliability simultaneously (see more details in Sec. II-C).

B. Do We Need Domain Label or Source Data?

According to Table I, source data is a crucial impact on **Office-31**, while domain label is not the necessary information during adaption. On the other hand, our proposed EGCA achieves highest accuracy compared with those methods with source data or domain label on **Office-Home**. As for **DomainNet**, source data and domain label are not the necessary conditions. In conclusion, **we may not need domain label nor source data in SF-BTDA scenarios.**

V. CONCLUSIONS

In this paper, we are the first to propose a new DA settings (*i.e.*, SF-BTDA), which is a more practical and challenging task where we can not access to source domain data while facing mixed multiple target domains without any domain labels in prior. Towards to address this scenario, we propose Evidential Graph Contrastive Alignment (EGCA) and conduct a new benchmark for SF-BTDA. EGCA effectively addressing two major challenges in SF-BTDA: (1) facing a mixture of multiple target domains, the model is assumed to handle the co-existence of different label shifts in different targets with no prior information; and (2) as we are only accessed to the source model, precisely alleviating the negative effect from noisy target pseudo labels is difficult because of the significantly enriched styles and textures. We conduct extensive experiments on three DA benchmarks, and empirical results show that EGCA outperforms state-of-the-art SF-STDA methods with considerable gains and achieves comparable results compared with those that have domain labels or source data.

ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China (42401415) and Jiangsu Innovation Capacity Building Program (Project BM2022028).

REFERENCES

- Juepeng Zheng, Shuai Yuan, Weijia Li, et al., "A review of individual tree crown detection and delineation from optical remote sensing images: Current progress and future," *IEEE Geosci. Remote Sens. Mag.*, 2024.
- Qingmei Li, Yibin Wen, et al., "Hyunida: Breaking label set constraints for universal domain adaptation in cross-scene hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, 2024.
- Wentang Chen, Yibin Wen, et al., "Ban: A universal paradigm for cross-scene classification under noisy annotations from rgb and hyperspectral remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, 2025.
- Yaroslav Ganin, Evgeniya Ustinova, et al., "Domain-adversarial training of neural networks," *JMLR*, vol. 17, no. 1, pp. 2096–2030, 2016.
- Geoff French, Michal Mackiewicz, and Mark Fisher, "Self-ensembling for visual domain adaptation," in *ICLR*, 2018.
- Juepeng Zheng, Yibin Wen, Mengxuan Chen, et al., "Open-set domain adaptation for scene classification using multi-adversarial learning," *ISPRS J. Photogramm. Remote Sens.*, vol. 208, pp. 245–260, 2024.
- Behnam Gholami et al., "Unsupervised multi-target domain adaptation: An information theoretic approach," *TIP*, vol. 29, pp. 3993–4002, 2020.
- Ziliang Chen et al., "Blending-target domain adaptation by adversarial meta-adaptation networks," in *CVPR*, 2019, pp. 2248–2257.
- Xingchao Peng, Zijun Huang, et al., "Domain agnostic learning with disentangled representations," in *ICML. PMLR*, 2019, pp. 5102–5112.
- Subhankar Roy et al., "Curriculum graph co-teaching for multi-target domain adaptation," in *CVPR*, 2021, pp. 5351–5360.
- Juepeng Zheng, Wenzhao Wu, Shuai Yuan, Yi Zhao, et al., "A two-stage adaptation network (tsan) for remote sensing scene classification in single-source-mixed-multiple-target domain adaptation (s^2m^2t da) scenarios," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–13, 2021.
- Shiqi Yang, Yaxing Wang, Joost Van De Weijer, et al., "Generalized source-free domain adaptation," in *ICCV*, 2021, pp. 8978–8987.
- Jian Liang, Dapeng Hu, Yunbo Wang, Ran He, and Jiashi Feng, "Source data-absent unsupervised domain adaptation through hypothesis transfer and labeling transfer," *TPAMI*, vol. 44, no. 11, pp. 8602–8617, 2021.
- Vikash Kumar, Rohit Lal, et al., "Conmix for source-free single and multi-target domain adaptation," in *WACV*, 2023, pp. 4178–4188.
- Alexander Amini, Wilko Schwarting, Ava Soleimany, et al., "Deep evidential regression," *NeurIPS*, vol. 33, pp. 14927–14937, 2020.
- Kari Sentz and Scott Ferson, "Combination of evidence in dempster-shafer theory," 2002.
- Ranganath Krishnan and Omesh Tickoo, "Improving model calibration with accuracy versus uncertainty optimization," *NeurIPS*, vol. 33, pp. 18237–18248, 2020.
- Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, "Deep residual learning for image recognition," in *CVPR*, 2016, pp. 770–778.
- Mingsheng Long, Yue Cao, et al., "Learning transferable features with deep adaptation networks," in *ICML. PMLR*, 2015, pp. 97–105.
- Mingsheng Long, Zhangjie Cao, Jianmin Wang, et al., "Conditional adversarial domain adaptation," in *NeurIPS*, 2018, pp. 1640–1650.
- Le Thanh Nguyen-Meidine, Atif Belal, Madhu Kiran, Jose Dolz, Louis-Antoine Blais-Morin, et al., "Unsupervised multi-target domain adaptation through knowledge distillation," in *WACV*, 2021, pp. 1339–1347.
- Jie Wang, Chaoliang Zhong, et al., "Discriminative mutual learning for multi-target domain adaptation," in *ICPR. IEEE*, 2022, pp. 2900–2906.
- Pengcheng Xu et al., "Class overwhelms: Mutual conditional blended-target domain adaptation," in *AAAI*, 2023, vol. 37, pp. 3036–3044.
- Mattia Litrico, Alessio Del Bue, and Pietro Morerio, "Guiding pseudo-labels with uncertainty estimation for source-free unsupervised domain adaptation," in *CVPR*, 2023, pp. 7640–7650.
- Kuniaki Saito, Kohei Watanabe, et al., "Maximum classifier discrepancy for unsupervised domain adaptation," in *CVPR*, 2018, pp. 3723–3732.
- Ying Jin, Ximei Wang, et al., "Minimum class confusion for versatile domain adaptation," in *ECCV*. Springer, 2020, pp. 464–480.
- Prannay Khosla, Piotr Teterwak, Chen Wang, et al., "Supervised contrastive learning," *NeurIPS*, vol. 33, pp. 18661–18673, 2020.
- Kate Saenko, Brian Kulis, Mario Fritz, et al., "Adapting visual category models to new domains," in *ECCV*. Springer, 2010, pp. 213–226.
- Hemant Venkateswara, Jose Eusebio, et al., "Deep hashing network for unsupervised domain adaptation," in *CVPR*, 2017, pp. 5018–5027.
- Xingchao Peng, Qinxun Bai, Xide Xia, et al., "Moment matching for multi-source domain adaptation," in *ICCV*, 2019, pp. 1406–1415.
- Mingsheng Long, Han Zhu, et al., "Unsupervised domain adaptation with residual transfer networks," in *NeurIPS*, 2016, pp. 136–144.