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

Juepeng Zheng, *Member, IEEE*, Guowen Li, Yibin Wen, Jinxiao Zhang, Runmin Dong, *Member, IEEE*, and Haohuan Fu, *Senior Member, IEEE*

***Abstract*—In this article, we first tackle a more realistic domain adaptation (DA) setting: source-free blending-target DA (SF-BTDA), where we *cannot* 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 coexistence of different label shifts in different targets, along with noisy target pseudolabels generated from the source model. In this article, we propose a new method called evidential graph contrastive alignment (EGCA) to decouple the blending-target domain and alleviate the effect of noisy target pseudolabels. First, to improve the quality of pseudo target labels, we propose a calibrated evidential learning (CEL) 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 the same class in the blending-target domain, which alleviates the coexistence 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 that have domain labels or source data in prior.**

***Index Terms*—Contrastive learning, domain adaptation (DA), evidential learning, source-free.**

I. Introduction

DESPITE the great success of deep learning in various fields [1], [2], [3], 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 statistical distribution, dimension, and context. Domain adaptation (DA) offers significant benefits in this scenario by adapting the pretrained models toward new domains with different distributions and properties from the original domain [4], [5], [6]. 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 the DA task, there are two main types of DA: single-target DA (STDA) [see Fig. 1(a)] and multitarget DA (MTDA) [see Fig. 1(c)]. While most existing DA works focus on STDA [7], [8], [9], MTDA is a more challenging task that simultaneously transfers one source domain to multiple target domains. For example, MTDA-ITA [10] aims to find a shared latent space common to all domains to allow adaptation from a single source to

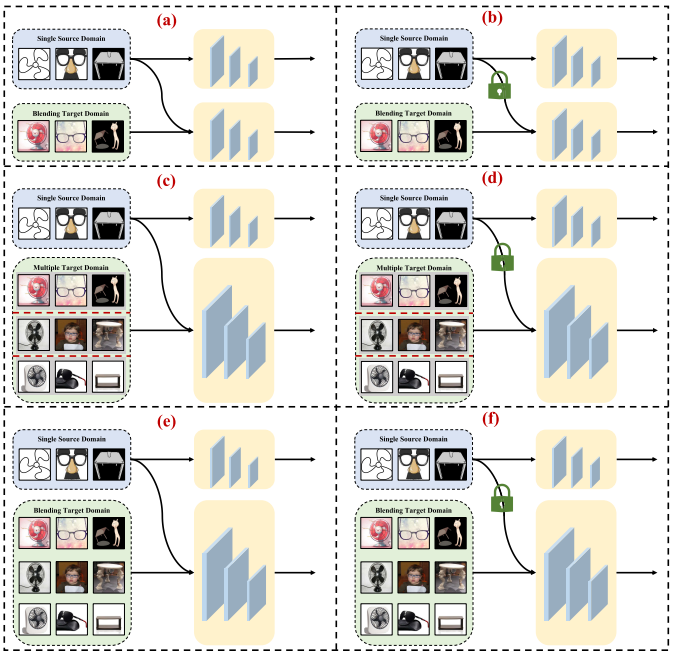

Fig. 1. Different DA settings. (a) STDA, a standard DA setting with access to labeled source data and source trained model. (b) SF-STDA, with only access to source trained model. (c) MTDA, with access to labeled source data, source trained model, and domain label in prior. (d) SF-MTDA with access to source trained model and domain label in prior. (e) BTDA with access to labeled source data and source trained model. (f) SF-BTDA, a new DA setting proposed in this article, with only access to source trained model.

Received 13 January 2025; revised 23 February 2025 and 30 May 2025; accepted 24 August 2025. Date of publication 8 September 2025; date of current version 8 January 2026. This work was supported in part by the Guangdong Science and Technology Program under Grant 2024B0101040005; in part by the National Natural Science Foundation of China under Grant T2125006 and Grant 42401415; in part by Shenzhen Science and Technology Program under Grant KCXFZ20240903093759004 and Grant KJZD20230923115106012; in part by the Fundamental Research Funds for the Central Universities, Sun Yat-sen University, under Project 24xkjc002; and in part by Jiangsu Innovation Capacity Building Program under Project BM2022028. *(Corresponding author: Haohuan Fu.)*

Juepeng Zheng is with the School of Artificial Intelligence, Sun Yat-sen University, Zhuhai 519082, China, and also with the National Supercomputing Center in Shenzhen, Shenzhen 518055, China (e-mail: zhengjp8@mail.sysu.edu.cn).

Guowen Li, Yibin Wen, and Runmin Dong are with the School of Artificial Intelligence, Sun Yat-sen University, Zhuhai 519082, China (e-mail: ligw8@mail2.sysu.edu.cn; wenyb5@mail2.sysu.edu.cn; dongrm3@mail.sysu.edu.cn).

Jinxiao Zhang is with the Ministry of Education Key Laboratory for Earth System Modeling and the Department of Earth System Science, Tsinghua University, Beijing 100190, China (e-mail: zhang-jx22@mails.tsinghua.edu.cn).

Haohuan Fu is with the Tsinghua Shenzhen International Graduate School, Tsinghua University, Shenzhen 518055, China, also with the National Supercomputing Center in Shenzhen, Shenzhen, China, and also with the Ministry of Education Key Laboratory for Earth System Modeling and the Department of Earth System Science, Tsinghua University, Beijing 100190, China (e-mail: haohuan@tsinghua.edu.cn).

Digital Object Identifier 10.1109/TNNLS.2025.3603224

2162-237X © 2025 IEEE. All rights reserved, including rights for text and data mining, and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission.
See https://www.ieee.org/publications/rights/index.html for more information.

Authorized licensed use limited to: SUN YAT-SEN UNIVERSITY. Downloaded on March 22,2026 at 15:35:20 UTC from IEEE Xplore. Restrictions apply.

TABLE I
COMPARISONS AMONG DIFFERENT DA SETTINGS

DA settings	Multiple targets	Domain labels	Source data
STDA	×	✓	✓
SF-STDA	×	✓	×
MTDA	✓	✓	✓
SF-MTDA	✓	✓	×
BTDA	✓	×	✓
SF-BTDA	✓	×	×

multiple target domains. HGAN [11] aims to learn a unified subspace common for all domains with a heterogeneous graph attention network for better semantic transfer in MTDA.

However, all aforementioned works have domain labels, which provide an important reference for deep learning models. However, 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 image collection device, street view, weather, lighting, time, and road conditions. 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. According to the authors' knowledge, there are only some works focusing on this mixed MTDA without domain labels, which is crucial for real-world applications [12], [13], [14]. Therefore, this article makes efforts on the more practical and challenging blending-target DA (BTDA) scenario [see Fig. 1(e)].

On the other hand, there is another crucial issue in real-world applications for DA, that is, whether the source data are available. Due to increasing privacy concerns (such as medical images, mobile applications, and surveillance videos), source-domain data cannot always be accessed. Recently, source-free DA (SFDA) has been introduced to transfer knowledge from a prior source domain to a new target domain without accessing the source data. As shown in Fig. 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 source-free STDA (SF-STDA) setting [see Fig. 1(b) and Table I] [15], [16], [17], while up to date, only ConMix [18] has discussed and designed new approach for source-free MTDA (SF-MTDA) [see Fig. 1(d)]. ConMix [18] introduces consistency between label-preserving augmentations and utilizes pseudo-label refinement methods to reduce noisy pseudo-labels for both SF-STDA and SF-MTDA settings.

Different from other DA scenarios, we proposed a new DA setting named source-free BTDA (SF-BTDA), where we have to transfer the knowledge from a single source to the blending-target domain without domain labels or the access of labeled source data [see Fig. 1(f)]. There are two major challenges in SF-BTDA.

1) *Challenge 1:* Facing a mixture of multiple target domains, the model is assumed to handle the coexistence of different label shifts in different targets with no prior information.
2) *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 and textures. Due to these two challenges, existing DA methods that count on utilizing the distribution shift between two specific domains become infeasible.

In this article, we first present the SF-BTDA setting and propose evidential graph contrastive alignment (EGCA) to better tackle the aforementioned two challenges in SF-BTDA scenarios. First, we calculate the domain distance and generate a pseudo-domain label through self-consistent clustering to address the coexistence of different label shifts in different targets with no prior information. Second, we propose calibrated evidential learning (CEL) to remove samples that are imbalanced in accuracy and certainty, and to iteratively improve both the accuracy and certainty of the resulting model. Finally, we propose graph contrastive alignment to pull samples that are from the same class both in intradomain and cross-domains through fully explore the advantages of pseudo-label information in the target domain and simultaneously reduce the negative transfer effect, which effectively tackles both *Challenges 1* and *2*. The contributions of this work are as follows.

1) We propose a novel DA task named SF-BTDA, where we have to transfer the knowledge from a single source to a blending-target domain without domain labels or the access of labeled source data.
2) We propose CEL 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.
3) We design graph contrastive learning with the domain distance matrix and confidence-uncertainty criterion, to minimize the distribution gap of samples of the same class in the blending-target domain.
4) We conduct a new benchmark for SF-BTDA setting and comprehensive experiments show that EGCA outperforms other methods with considerable gains and achieves comparable results compared with those that have domain labels or source data in prior.

The second contribution mainly addresses *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 coexistence of different label shifts in blended targets.

In the remaining sections of this article, a wide range of related work is introduced in Section II. We present a detailed description of our proposed EGCA in Section III. After that, a thorough experiment with promising results compared with other state-of-the-art approaches is detailed in Section IV. We further discuss our proposed EGCA in Section V and summarize it in Section VI.

Authorized licensed use limited to: SUN YAT-SEN UNIVERSITY. Downloaded on March 22,2026 at 15:35:20 UTC from IEEE Xplore. Restrictions apply.

II. Related Works

A. Domain Adaptation

DA aims to minimize domain shift, which refers to the distribution difference between the source and target domains. Most of them focus on STDA. Plenty of single target DA methods have been devised to different tasks including classification [7], [19], [20], [21], semantic segmentation [22], and object detection [23], [24]. To reduce domain discrepancies, there are different kinds of methods. One popular approach is motivated by generative adversarial networks (GANs) [25]. Adversarial networks are widely explored and try to minimize the domain shift by maximizing the confusion between source and target [26]. Alternatively, it is also common practice to adopt statistical measure matching methods to bridge the difference between source and target, including MMD [27], JMMD [28], and other variation metrics [29]. Although extensive works have been done on STDA, there is less research on MTDA, especially for BTDA, which is more similar to the real-world scenarios.

B. Multitarget DA and Blending-Target DA

Different from STDA, MTDA is more flexible because the model can be adapted to multiple target domains with different characteristics, which requires the model to generalize better on target data without annotations [10], [30]. These requirements introduce new challenges, especially the computational requirements and catastrophic forgetting of previously learned targets. Wang et al. [31] combine federated learning with multitarget DA and propose an effective DualAdapt method to reduce the training cost significantly. MTDA-DTM [32] allows the object detector to be incrementally updated without degrading performance on data from previously learned target domains. However, BTDA has received little attention. This approach aims to provide solutions for situations where there are no explicit domain labels and samples from different target domains are mixed together, which is more consistent with reality [33]. The differences between STDA, MTDA, and mixed MTDA is shown in Fig. 1. To the best of authors' knowledge, there are only few works related to this field. For instance, DADA [13] aims to tackle domain-agnostic learning by disentangling the domain-invariant features from both domain-specific and class-irrelevant features simultaneously. CGCT [14] uses curriculum learning and co-teaching methods to obtain more reliable pseudo-labels. In addition, open compound DA (OCDA) also contains blending-target domain [34], [35], while it contains an extra "unseen" domain, which is beyond the scope of this article.

C. Source-Free DA

Because of the limit of privacy constraints, SFDA is becoming more and more popular, which only provides a well-trained source model without any access to source data during the adaptation process. There are two mainstreams for SFDA methods. On the one hand, some methods focus on reconstructing the pseudo source-domain samples in the feature space [36], [37] On the other hand, some methods exploit pseudo target labels from the source model and adopt self-training or noisy learning so that the model is well-fit to the target domain [38]. For example, SHOT [39] adopts entropy minimization and information maximization through pseudo-labeling learning to transfer the knowledge from a trained classifier to target features. G-SFDA [40] is proposed to refine pseudo-labels through encouraging consistent label predictions between local neighbor samples. To date, most existing SFDA methods focus on the SF-STDA setting, while only Kumar et al. [18] have discussed and designed a new approach for SF-MTDA. ConMix [18] introduces consistency between label-preserving augmentations and utilizes pseudo-label refinement methods to reduce noisy pseudo-labels for both SF-STDA and SF-MTDA settings.

However, existing literature has not exploited a new DA setting named SF-BTDA, where we have to transfer the knowledge from a single source to a blending-target domain without domain labels or the access of labeled source data. As introduced in Section I, there are two major challenges in SF-BTDA.

1) *Challenge 1:* Facing a mixture of multiple target domains, the model is assumed to handle the coexistence of different label shifts in different targets with no prior information.
2) *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 and textures. In this article, we propose EGCA to better address both two major challenges for SF-BTDA.

III. Methodology

In this section, we first propose the problem setting for SF-BTDA. Then, we introduce our domain distance matrix to prevent the negative effect from mixture of multiple target domains. Next, we elaborate how to select high-quality target samples from noisy pseudo-label in CEL module (addressing *Challenge 2*). Finally, we adopt graph contrastive learning with the domain distance matrix and confidence-uncertainty criterion, to minimize the distribution gap of samples of the same class in the blending-target domain (addressing *Challenge 1*).

A. Problem Setting and Notations

In contrast to MTDA, BTDA is established on a mixture of target domains $\mathcal{D}_T = \{\mathcal{D}_t^j\}_{j=1}^k = \{(\mathbf{x}_i^t)\}_{i=1}^{n_T}$, 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 the MTDA scenario, the proportions of different target domains in the mixed datasets $\{w^j\}_{j=1}^k$ are unknown. In the SF-BTDA setting, we consider the 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 $\mathbf{x}_i^s$, are only available during model pretraining. Our proposed 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 $\mathcal{G}$.

As we have to transfer the knowledge from a single source to blending-target domain without domain labels or the access

Authorized licensed use limited to: SUN YAT-SEN UNIVERSITY. Downloaded on March 22,2026 at 15:35:20 UTC from IEEE Xplore. Restrictions apply.

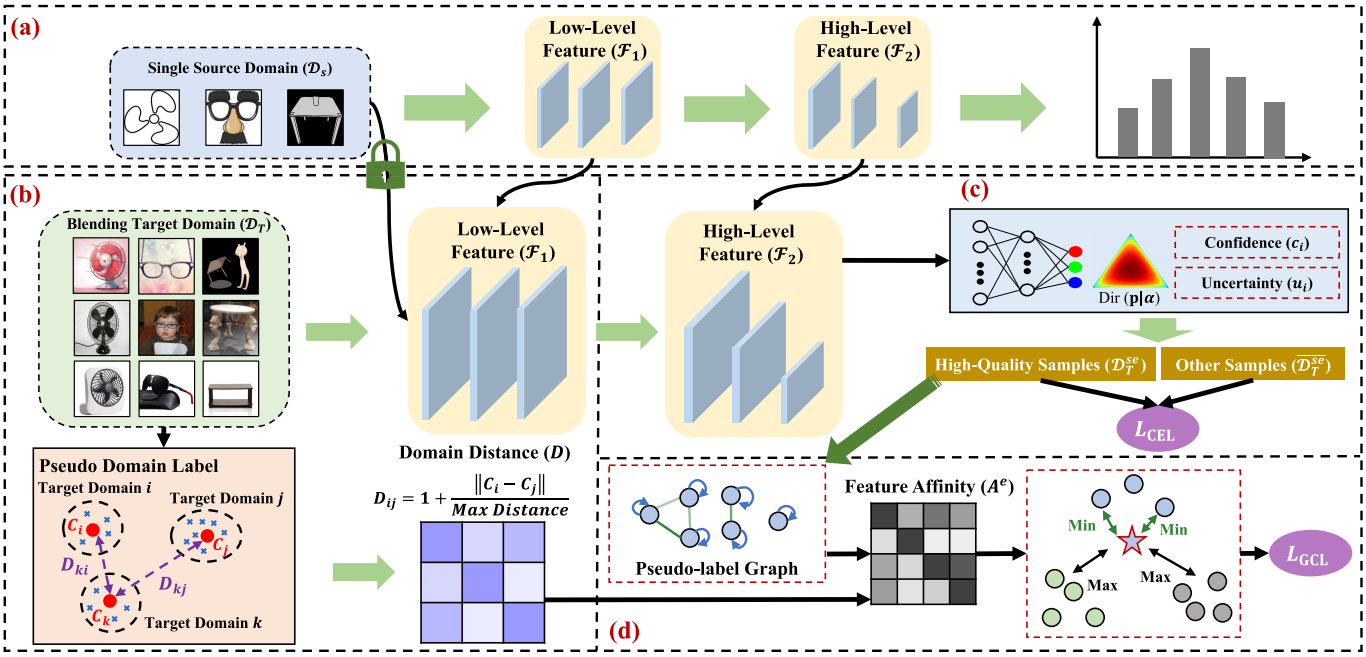

Fig. 2. Framework for our proposed EGCA. (a) *Source data training*: we train the model through a single source domain, and we do not have access to the source-domain data. (b) *Domain distance calculation*: we generate the mixture domain distance through low-level features and original image textures in an unsupervised way. (c) CEL: we propose a CEL module to iteratively improve both the accuracy and certainty of the resulting model and adaptively generate high-quality pseudo target label (addressing *Challenge 2*). (d) *Graph contrastive learning*: we design a graph contrastive learning with the domain distance matrix and confidence-uncertainty criterion, to minimize the distribution gap of samples of the same class in the blending-target domain, which alleviates the coexistence of different label shifts in blended targets (addressing *Challenge 1*).

of labeled source data, 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 subsubjects from k distributions, adaptation for the whole blending-target domain directly may lead to severe category misalignment and negative transfer effects. Compared to MTDA, it is a key message in the SF-BTDA scenario that we are unavailable to any information about the domain label of each image in the target dataset in advance. Therefore, the challenges of the SF-BTDA scenario are twofold. On the one hand, the mixture of multiple target domains would naturally introduce significantly enriched styles and textures. On the other hand, the model is assumed to handle the coexistence of different label shifts in different targets with no prior information. Due to these two challenges, most existing DA methods that count on utilizing the distribution shift between two specific domains become infeasible.

To address these problems in the SF-BTDA issue, we propose EGCA method (Fig. 2). First, we calculate the domain distance and generate pseudo-domain label through self-consistent clustering to address the coexistence of different label shifts in different targets with no prior information (see Section III-B). Second, we propose a CEL module, to remove samples that are imbalanced in accuracy and certainty, and to iteratively improve both the accuracy and certainty of the resulting model (see Section III-C). Finally, based on the samples that perform well in both accuracy and certainty, we further propose graph contrastive alignment to minimize the distribution gap of samples of the same class in the blending-target domain, and between the source and target domains (see Section III-D).

B. Domain Distance Calculation

As introduced in Section I, the SF-BTDA issue faces a mixture of multiple target domains, and the model is assumed to handle the coexistence of different label shifts in different targets with no prior information (*Challenge 1*). Therefore, some samples in the same class may appear with different styles and textures because of the coexistence of multiple targets with significantly enriched styles and textures. Under this consideration, we calculate the domain distance to better guide the contrastive learning (see Section III-D) to pull the samples in the same class that may appear in different styles and textures. As we do not have 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). Therefore, given that original image textures ($\mathbf{x}^t$) may represent the specific image style, while shallow features from DNNs [$\mathcal{F}_1(\mathbf{x}^t)$] may represent the universal feature style [41], we combine these two types of features to achieve better pseudo-domain label assignment. Notably, we simply concatenate the low-level features $\mathcal{F}_1(\mathbf{x}_i)$ with the original image x_i and we adopt the low-level features that generated from conv2_x with the dimension of $56 \times 56 \times 256$.

Notably, we simply concatenate the low-level feature with the original image x_i and we adopt the low-level features that

Authorized licensed use limited to: SUN YAT-SEN UNIVERSITY. Downloaded on March 22,2026 at 15:35:20 UTC from IEEE Xplore. Restrictions apply.

were generated from conv2_x with the dimension of $56 \times 56 \times 256$.

We calculate the domain distance D_{ij} between two samples $\mathbf{x}_i^t$ and $\mathbf{x}_j^t$ through (1), in which C_i and C_j are the corresponding centroids for samples ($\mathbf{x}_i^t$) from the target domain i ($\mathcal{D}_t^i$) and samples ($\mathbf{x}_j^t$) from the target domain j ($\mathcal{D}_t^j$). Notably, i and j are the two nominal variables. The centroid for a target image (x_i^t) denotes that it belongs to the pseudo-domain label. The *Max Distance* denotes that we normalize the domain distance among different target domain centroids. For example, when there are three target domains ($\mathcal{D}_t^1$, $\mathcal{D}_t^2$, and $\mathcal{D}_t^3$) in the blended target ($\mathcal{D}_T$), the *Max Distance* denotes the max value among $\|C_1 - C_2\|$, $\|C_1 - C_3\|$, and $\|C_2 - C_3\|$. In this article, we utilize the distance of the domain centroids to represent the samples from different target domains for simplicity

$$D_{ij} = 1.0 + \frac{\|C_i - C_j\|}{\text{Max Distance}}. \tag{1}$$

To avoid $D_{ij} = 0$ when these two samples come from the same target domain, we add 1.0 in (1). That is, if $\mathbf{x}_i^t$ and $\mathbf{x}_j^t$ are from the same domain, the domain distance will be 1.0 instead of 0. In particular, after each training epoch, we recluster target samples into k domains and reassign their pseudo-domain labels, and then, we update the specific target-domain centroids and update the domain distance D_{ij}.

C. Calibrated Evidential Learning

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* introduced in Section I, 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 CEL to iteratively improve both the accuracy and certainty of the resulting model, and adaptively select high-quality pseudo target labels.

Evidential deep learning (EDL) [42], [43] offers a principled way to formulate both the uncertainty estimation and multiclass classification, which overcomes the shortcomings of softmax-based DNNs. Assuming that the class probability follows a prior Dirichlet distribution, the cross-entropy loss to be minimized for the learning evidence $\mathbf{e}(i)$ from a given sample $\mathbf{x}(i)$ can be formulated as:

$$L_{\text{EDL}}^{(i)}\left(\mathbf{y}^{(i)}, \mathbf{e}^{(i)}\right) = \sum_{m=1}^{M} y_k^{(i)} \left(\log S^{(i)} - \log\left(e^{(i)_k+1}\right)\right) \tag{2}$$

in which M is the number of classes, and $\mathbf{y}(i)$ is an one-hot K-dimensional label for the sample $\mathbf{x}(i)$. $\mathbf{e}(i) = g(f(\mathbf{x}(i)))$, where f is the feature extractor and g is the evidence function. S is the total evidence that can be defined as $S = \sum_{m=1}^{M} \boldsymbol{\alpha}_m$, in which $\boldsymbol{\alpha}_m$ is the nonnegative network prediction output and can be expressed as $\boldsymbol{\alpha}_m = \mathbf{e}_m + 1$, where $\mathbf{e}(i) = g(\mathcal{F}_2(\mathcal{F}_1(\mathbf{x}(i))))$ and g is the evidence function. According to the evidence theory of Dempster–Shafer theory (DST) [44], the discriminative probability of mth class is $p_m = \boldsymbol{\alpha}_k/S$ and the predictive uncertainty u can be determined as $u = M/S$.

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 the SF-BTDA scenario. According to existing model calibration literature [45], 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 attempt to calibrate the EDL model by considering the relationship between the accuracy and uncertainty for the blending-target domain under a source-free setting.

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

$$\begin{aligned} L_{\text{CEL}} = &-\gamma_t \sum_{\mathbf{x}_i \in \mathcal{D}_T^{\text{se}}} c_i \log(1 - u_i) \\ &- (1 - \gamma_t) \sum_{\mathbf{x}_i \in \overline{\mathcal{D}_T^{\text{se}}}} (1 - c_i) \log(u_i) \end{aligned} \tag{3}$$

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 the SFDA setting [46], [47], [48], [49], 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 the two selection thresholds that can be adaptively estimated in each mini-batch. To this end, we do not need to set fixed hyperparameters for η_c and η_u

$$\eta_c = \frac{1}{B} \sum_{i=1}^{i=B} c_i; \quad \eta_u = \frac{1}{B} \sum_{i=1}^{i=B} u_i. \tag{4}$$

The first term of L_{CEL} in (3) tries to give low uncertainty ($u_i \to 0$) when the model makes accurate prediction ($\mathbf{x}_i \in \mathcal{D}_T^{\text{se}}, c_i \to 1$) for the selected high-quality target samples, while the second term of L_{CEL} aims to give high uncertainty ($u_i \to 1$) when the model makes inaccurate prediction ($\mathbf{x}_i \in \overline{\mathcal{D}_T^{\text{se}}}, c_i \to 0$) for the low-quality target samples. Notably, γ_t is the annealing factor defined as

$$\gamma_t = \gamma_0 \exp\left\{-\left(\ln \frac{\gamma_0}{\mathcal{T}}\right) t\right\} \tag{5}$$

in which t is the training epoch and $\mathcal{T}$ is the total number of training epochs. γ_0 is a small positive constant and monotonically increasing.

The reason of the annealing weight γ is that the dominant periods of accurate and inaccurate predictions in model training are different. In the early training stages, the inaccurate predictions are the dominant cases so that the IC loss [the second term in L_{CEL} in (3)] should be more penalized, while in the late training stages, the accurate predictions are the dominant so that the AU loss [the first term in L_{CEL} in (3)] should be more penalized. Since the annealing weight factor γ_t dynamically balances the two terms in training, γ_0 is a small positive constant ($\gamma_0 = 0.001$) and the factor γ_t will be exponentially increasing from γ_0 to 1.0 as the training epoch t increasing to $\mathcal{T}$.

Authorized licensed use limited to: SUN YAT-SEN UNIVERSITY. Downloaded on March 22,2026 at 15:35:20 UTC from IEEE Xplore. Restrictions apply.

To this end, through our proposed CEL, 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 Section I. This strategy helps us to better achieve the knowledge transfer from the source model to the blending-target domain.

D. Graph Contrastive Learning

Contrastive learning tries to learn a universal prior information for downstream tasks [50], [51]. According to existing contrastive learning frameworks [50], [52], [53], [54], the self-supervised contrastive loss InfoNCE [55] is

$$L_{\mathrm{CON}} = -\sum_{i\in\mathcal{I}} \log \frac{\exp\left(z_i \cdot z_{j(i)}/\tau\right)}{\sum_{q\in Q(i)} \exp\left(z_i \cdot z_q/\tau\right)} \tag{6}$$

in which $i \in \mathcal{I} \equiv \{1 \ldots 2N\}$ is the index of an arbitrary augmented sample and $j(i)$ is the index of the other augmented sample originating from the same sample. z is the projection of $f(\mathbf{x})$, and τ is a scalar temperature parameter. $Q(i) \equiv I\backslash\{i\}$. The index i is named the anchor, index $j(i)$ is named the positive pair, and the other $2(N-1)$ indices ($\{k \in Q(i)\backslash\{j(i)\}\}$) are named the negative pairs. Note that for each anchor i, there is 1 positive pair and $2N-2$ negative pairs. The denominator has a total of $2N-1$ terms (the positive and negative).

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 Section III-C. Furthermore, according to the description of *Challenge 1*, some samples in the same class may appear in different styles and textures because of the coexistence of multiple targets with significantly enriched styles and textures. To this end, we design the domain distance calculated from Section III-B in graph contrastive learning to pay more attention to the above-mentioned samples, which could solve the conflict between contrastive learning and the classification task.

Since Section III-C has generated high-quality labels for the target domain, we could fully explore their advantages in the contrastive learning. Notably, our contrastive alignment module aims to learn representations guided by a pseudo-label graph. Here, 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:

$$a_{ij} = \begin{cases} 1, & \text{if } j = i, \\ 1, & \text{if } z_i \text{ and } z_j \text{ are from same category,} \\ & \text{and } \mathbf{x}_i \in \mathcal{D}_T^{\mathrm{se}} \text{ and } \mathbf{x}_j \in \mathcal{D}_T^{\mathrm{se}}, \\ 0, & \text{otherwise.} \end{cases} \tag{7}$$

Notably, $\mathbf{x}_i \in \mathcal{D}_T^{\mathrm{se}}$ and $\mathbf{x}_j \in \mathcal{D}_T^{\mathrm{se}}$ denote 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 alleviate this problem, we adopt the confidence-uncertainty evaluation produced from Section III-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

$$a_{ij}^e = \begin{cases} a_{ij} * I_{ij} * D_{ij}, & \text{if } i \neq j, \\ a_{ij}, & \text{otherwise.} \end{cases} \tag{8}$$

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

$$L_{\mathrm{GCL}} = -\sum_{i\in\mathcal{I}} \frac{1}{1+|\mathcal{P}(i)|} \left\{ \log \frac{\exp\left(z_i \cdot z_{j(i)}/\tau\right)}{\sum_{q\in Q(i)} \exp\left(z_i \cdot z_q/\tau\right)} - \sum_{p\in\mathcal{P}(i)} \log \frac{a_{ij}^e \cdot \exp\left(z_i \cdot z_p/\tau\right)}{\sum_{q\in Q(i)} \exp\left(z_i \cdot z_q/\tau\right)} \right\} \tag{9}$$

in which $\mathcal{P}(i)$ denotes the 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 [63]. As our setting is SF-BTDA, it is noteworthy that our graph contrastive learning module is conducted on the blended targets due to its difficulties from classification confusions because of the coexistence 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 graph contrastive learning loss L_{GCL}

$$L = L_{\mathrm{CEL}} + \beta L_{\mathrm{GCL}} \tag{10}$$

where β is a tradeoff parameter balancing two different loss terms between L_{CEL} and L_{GCL}.

To enhance clarity, we summarize the training and infer ence procedures of EGCA in Algorithm 1. The algorithm outlines the key steps, including domain distance calculation, CEL, and graph contrastive alignment, as described in Sections III-B–III-E.

IV. Experiments

A. Datasets

We conduct experiments on three standard DA benchmarks including *Office-31* [64], *Office-Home* [65], and *DomainNet* [66]. *Office-31* [64] consists of 4652 images from three domains: DSLR (D), Amazon (A), and Webcam (W). *Office-Home* [65] is a more challenging dataset, which consists of 15500 images in total from 65 categories of everyday objects. There are four significantly different domains: artistic images (Ar), clip-art images (Cl), product images (Pr), and real-world images (Rw). *DomainNet* [66] is the most challenging

Authorized licensed use limited to: SUN YAT-SEN UNIVERSITY. Downloaded on March 22,2026 at 15:35:20 UTC from IEEE Xplore. Restrictions apply.

Fig. 3. Samples from three standard DA benchmark datasets used in our experiments. (a) Office-31 dataset, featuring images from Amazon, DSLR, and Webcam domains. (b) Office-Home dataset, showcasing diverse categories including art, clipart, product, and real-world domains. (c) DomainNet dataset, presenting a wide range of domains such as clipart, infograph, painting, quickdraw, real, and sketch.

Algorithm 1 Training and Inference Procedures of EGCA

Input: Blending-target domain data $\mathcal{D}^T$, source model components $\{\mathcal{F}_1, \mathcal{F}_2, \mathcal{G}\}$, number of pseudo-domains k, trade-off weight β, total epochs T

Output: Adapted model parameters θ^*

TRAINING PHASE:

Compute domain-distance matrix D_{ij} via Eq.(1) and assign pseudo-domain labels by k-means.

Generate pseudo-labels for all $x \in \mathcal{D}^T$ using current $\{\mathcal{F}_1, \mathcal{F}_2, \mathcal{G}\}$.

for $t = 1$ T **do**

  Update high-quality sample set $\mathcal{D}_T^{\text{se}}$ and $\overline{\mathcal{D}_T^{\text{se}}}$ via calibrated evidential selection.

  Compute **Calibrated Evidential Loss** L_{CEL} in Eq. (3).

  Compute **Graph Contrastive Alignment Loss** L_{GCL} using D_{ij} and confidence/uncertainty in Eq. (9).

  Update parameters: $\theta \leftarrow \theta - \eta \nabla_\theta \big(L_{\text{CEL}} + \beta L_{\text{GCL}}\big)$.

end for

INFERENCE PHASE:

Given test sample x_0, extract $z_1 = \mathcal{F}_1(x_0)$, $z_2 = \mathcal{F}_2(z_1)$.

Compute logits $\alpha = \mathcal{G}(z_2)$ and class-probabilities $p_m = \alpha_m / \sum_n \alpha_n$.

OUTPUT: predicted label $\hat{y} = \arg\max_m p_m$.

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. In these three DA datasets, we select only one domain as the source domain while the others as the blending-target domains. Fig. 3 shows some examples of these three datasets.

B. Training Details

Our methods were implemented based on the PyTorch [67]. We adopt ResNet-50 [56] for *Office-31* and *Office-Home*, and ResNet-101 for *DomainNet*, both of which pretrained on the ImageNet dataset [68]. Whatever module was trained from scratch, its learning rate was set to be ten times that of the lower layers. We adopt mini-batch stochastic gradient descent (SGD) with a momentum of 0.95 using the learning rate and progressive training strategies. We set β in (10) and γ_0 in (5) as 1.0 and 0.001, respectively. We set τ in (6) and (9) as 0.1 following previous studies [63]. In our experiments, we select the same value of k as the true number of target domains in each dataset to report, with $k = 2$, $k = 3$, and $k = 5$ for *Office-31*, *Office-Home*, and *DomainNet*, respectively.

1) State of the Art: ResNet [56] is the baseline backbone without any DA tricks. We list some methods that have been proposed for the STDA scenario, such as DAN [27], DANN [7], RTN [8], JAN [28], SE [9], MCD [61], CDAN [57], and MCC [62]. We also compare our method with the existing three MTDA methods (i.e., MTDA-ITA [10], DCL [30], and DCGCT [14]) and five BTDA methods (i.e., AMEAN [12], DADA [13], CGCT [14], DML [31], and MCDA [58]). Furthermore, we compare three source-free DA methods, including G-SFDA [40], GPUE [59], and SHOT++ [60]. However, these SFDA methods mainly focus on SF-STDA settings. Finally, we compare our performance with ConMix [18], which is the only one that concentrates on SF-MTDA scenarios with domain labels in prior, while without access to the labels of source data. It is noteworthy that those methods designed for single target DA tasks (such as DANN [7], CDAN [57], G-SFDA [40], and GPUE [59]) consider the whole blended target domain as the single target domain and are lack of capacity to address blended target DA scenarios. Notably, the bold numbers in Tables II and III denote the highest accuracy without source data. The underline numbers denote the highest accuracy without domain labels.

C. Results

1) Office-31: Our EGCA overpasses the comparison methods without domain labels or source data on all transfer tasks, with at least 1.5% improvement. The EGCA even outperforms all comparison methods only without source data. It is remarkable that EGCA promotes the accuracy on transfer tasks of **A** and **D** compared to ConMix with domain labels in prior.

2) Office-Home: The accuracies on *Office-Home* for the SF-BTDA setting are shown in Table II. We can observe

Authorized licensed use limited to: SUN YAT-SEN UNIVERSITY. Downloaded on March 22,2026 at 15:35:20 UTC from IEEE Xplore. Restrictions apply.

TABLE II

OVERALL ACCURACY (%) ON OFFICE-31 AND OFFICE-HOME 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				Office-Home				
				A	D	W	Avg	Ar	Cl	Pr	Rw	Avg
ResNet-50 [56]	×	×	×	76.3	68.7	67.0	70.7	62.5	61.2	55.1	61.8	60.1
DAN [27]	×	×	✓	79.5	80.3	81.2	80.3	58.4	58.1	52.9	62.1	57.9
DANN [7]	×	×	✓	78.2	72.2	69.8	73.4	58.4	58.1	52.9	62.1	57.9
CDAN [57]	×	×	✓	93.6	80.5	81.3	85.1	59.5	61.0	54.7	62.9	59.5
MTDA-ITA [10]	✓	✓	✓	87.9	83.7	84.0	85.2	64.6	66.4	59.2	67.1	64.3
DCL [30]	✓	✓	✓	92.6	82.5	84.7	86.6	63.0	63.0	60.0	67.0	63.3
DCGCT [14]	✓	✓	✓	93.4	86.0	87.1	88.8	70.5	71.6	66.0	71.2	69.8
AMEAN [12]	✓	×	✓	90.1	77.0	73.4	80.2	64.3	65.5	59.5	66.7	64.0
CGCT [14]	✓	×	✓	93.9	85.1	85.6	88.2	67.4	68.1	61.6	68.7	66.5
DML [31]	✓	×	✓	94.9	85.3	86.2	88.8	67.0	70.6	62.9	69.0	67.4
MCDA [58]	✓	×	✓	92.4	87.7	88.8	89.6	71.7	72.8	68.0	71.7	71.1
G-SFDA [40]	×	×	×	90.6	78.9	77.4	82.3	70.3	74.9	66.8	70.0	70.5
GPUE [59]	×	×	×	78.8	72.4	60.7	70.6	50.5	45.5	46.4	52.9	48.8
SHOT++ [60]	×	×	×	93.3	79.6	78.4	83.8	70.6	74.2	66.2	68.6	69.9
ConMix [18]	✓	✓	×	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 III

OVERALL ACCURACY (%) ON DOMAINNET 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
ResNet-101 [56]	×	×	×	25.6	16.8	25.8	9.2	20.6	22.3	20.1
SE [9]	×	×	✓	21.3	8.5	14.5	13.8	16.0	19.7	15.6
MCD [61]	×	×	✓	25.1	19.1	27.0	10.4	20.2	22.5	20.7
CDAN [57]	×	×	✓	31.6	27.1	31.8	12.5	33.2	35.8	28.7
MCC [62]	×	×	✓	33.6	30.0	32.4	13.5	28.0	35.3	28.8
DCL [30]	✓	✓	✓	35.1	31.4	37.0	20.5	35.4	41.0	33.4
DCGCT [14]	✓	✓	✓	37.0	32.2	37.3	19.3	39.8	40.8	34.4
DADA [13]	✓	×	✓	26.4	20.0	26.5	12.9	20.7	22.8	21.6
CGCT [14]	✓	×	✓	36.1	33.3	35.0	10.0	39.6	39.7	32.3
DML [31]	✓	×	✓	32.0	25.4	29.4	12.7	31.5	36.4	27.9
MCDA [58]	✓	×	✓	37.5	37.3	36.6	17.8	36.1	41.4	34.5
G-SFDA [40]	×	×	×	34.2	26.7	32.7	11.7	18.7	32.0	26.0
GPUE [59]	×	×	×	23.4	22.7	31.4	16.5	16.9	24.8	22.6
SHOT++ [60]	×	×	×	32.9	28.8	32.1	13.9	17.9	34.9	26.7
ConMix [18]	✓	✓	×	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**

TABLE IV

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

DomainNet	Cli	Inf	Pai	Qui	Rea	Ske	Avg
ResNet-101 [56]	25.6	16.8	25.8	9.2	20.6	22.3	20.1
+ L_{CEL}	32.6	30.1	35.3	12.5	30.3	36.8	29.6
+ L_{CEL} + L_{CON}	37.1	**33.2**	**38.1**	14.6	34.0	39.2	32.7
+ L_{CEL} + L_{GCL}	**39.2**	**32.6**	37.3	**16.0**	**35.5**	**39.8**	**33.4**

TABLE V

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**

that our ECGA achieves the highest average accuracy (70.6%) compared to methods without domain labels nor source data. Especially, EGCA is the best model for the task of **Cl** among methods without domain labels and outperforms other methods without source data for tasks of **Pr** and **Rw**.

3) DomainNet: As for the most difficult dataset, our proposed EGCA improves the accuracy with a considerably larger rooms (with at least +7.2%) compared to those without domain labels or source data. In addition, it is encouraging that EGCA yields the highest accuracy without source data. Furthermore, EGCA is slightly lower (with only −0.6%) than MCDA, which has source data without domain labels, which suggests that EGCA can fully explore high-quality pseudo target-domain labels and EGCA for effective DA.

D. Ablation Studies

To tooth apart the separate contributions of the strategies in EGCA, we denote by L_{CEL} and L_{GCL} the module of CEL and semi-supervised contrastive learning. As listed in Fig. 4, L_{CON} achieves 2.5% and 3.8% improvement for *Office-31* and *Office-Home*, respectively. In addition, L_{GCL} further increase 1.1% and 1.4% for average accuracy. Table IV lists the ablation studies for *DomainNet* for our proposed ECGA. As listed in Table IV, L_{CEL} achieves +9.5% improvement compared

Authorized licensed use limited to: SUN YAT-SEN UNIVERSITY. Downloaded on March 22,2026 at 15:35:20 UTC from IEEE Xplore. Restrictions apply.

TABLE VI

COMPARISONS OF ACCURACY GAP AMONG DIFFERENT METHODS ON OFFICE-31 AND OFFICE-HOME. THE RED NUMBERS DENOTE THAT OUR PROPOSED EGCA OUTPERFORMS OTHER METHODS, WHILE THE BLUE NUMBERS DENOTE THAT OTHER METHODS OUTPERFORM OUR PROPOSED EGCA

Method		Office-31				Office-Home				
		A	D	W	Avg	Ar	Cl	Pr	Rw	Avg
without domain labels nor source data	ResNet-50 [56]	+16.9	+14.7	+12.3	+14.6	+22.9	+31.6	+23.0	+19.3	+24.2
	G-SFDA [40]	+2.6	+4.5	+2.0	+3.0	+0.2	-0.7	+0.4	+0.6	+0.1
	GPUE [59]	+14.4	+11.0	+18.6	+14.7	+20.0	+28.7	+20.8	+17.7	+21.8
	SHOT++ [60]	-0.1	+3.8	+0.9	+1.5	-0.1	+0.0	+1.0	+2.0	+0.7
only without domain labels	DAN [27]	+13.7	+3.1	-1.9	+5.0	+14.9	+17.6	+18.7	+13.9	+16.3
	DANN [7]	+12.4	+0.9	-3.9	+3.1	+12.1	+16.1	+14.3	+8.5	+12.8
	CDAN [57]	-0.4	+2.9	-2.0	+0.2	+11.0	+13.2	+12.5	+7.7	+11.1
	AMEAN [12]	+3.1	+6.4	+5.9	+5.1	+6.2	+8.7	+7.7	+3.9	+6.6
	CGCT [14]	-0.7	-1.7	-6.3	-2.9	+3.1	+6.1	+5.6	+1.9	+4.2
	DML [31]	-1.7	-1.9	-6.9	-3.5	+3.5	+3.6	+4.3	+1.6	+3.3
	MCDA [58]	+0.8	-4.3	-9.5	-4.3	-1.2	+1.4	-0.8	-1.1	-0.4
only without source data	ConMix [18]	+0.8	+1.6	-1.1	+0.4	-1.4	-2.2	+0.8	+1.3	-0.4
both with domain labels and source data	MTDA-ITA [10]	+5.3	-0.3	-4.7	+0.1	+5.9	+7.8	+8.0	+3.5	+6.3
	DCL [30]	+0.6	+0.9	-5.4	-1.3	+7.5	+11.2	+7.2	+3.6	+7.4
	DCGCT [14]	-0.2	-2.6	-7.8	-3.5	+0.0	+2.6	+1.2	-0.6	+0.8

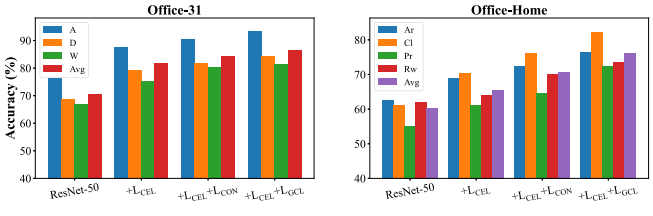

Fig. 4. Overall accuracy (%) on *Office-31* and *Office-Home* for ablation studies.

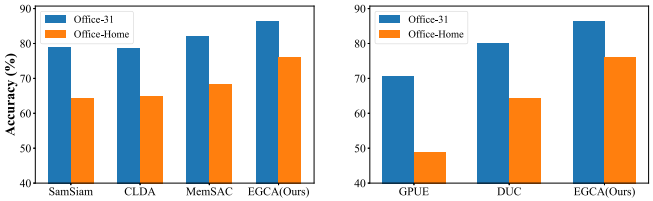

Fig. 5. Comparisons among other contrastive learning methods and other uncertainty methods with our EGCA.

to ResNet-101 [56]. In addition, L_{CON} increases +3.1% for average accuracy compared to L_{CEL}. Furthermore, L_{GCL} further improves the average accuracy with slight gains of +0.7%.

In addition, Table V lists ablation results for different high-quality selection schemes. Therefore, we both utilize the evidential uncertainty and prediction confidence to measure the label reliability simultaneously (see more details in Section III-C). Furthermore, we compare different contrastive learning methods (such as SamSiam [69], CLDA [70], and MemSAC [71]) and uncertainty methods (such as GPUE [59] and DUC [72]) in Fig. 5. The experimental results indicate that the evidential contrastive learning and the CEL in our proposed EGCA perform the best among other state-of-the-art methods.

As introduced in Section III-B, we simply concatenate the low-level feature with the original image x_i. We adopt the low-level features that were generated from conv2_x with the dimension of $56 \times 56 \times 256$. Fig. 6(a) shows that after adding the original image x_i, the performance has gains for *Office-31* and *Office-Home* by using shallow features ($\mathcal{F}_1(\mathbf{x}^t)$). Also, Fig. 6(b) shows the accuracy of pseudo-labels for the target data (introduced in Sections III-C and III-D) with gradually improving along with the training epochs, indicating that our model is becoming more and more confident and certainty during the model training. Fig. 6(c) shows that increasing the number of k for single target DA, the accuracy may only have a very slight reduction. This experiment indicates that our proposed method is also robust to address the scenario where the number of target domains is only one (i.e., single target DA). However, this article focuses on SF-BTDA instead of single-target DA.

V. DISCUSSION

A. Do We Really Need Domain Label or Source Data?

We calculate the accuracy gap using the following equation between our proposed EGCA and other methods in Table VI (*Office-31* and *Office-Home* and Table VII (*DomainNet*):

$$\text{Accuracy Gap} = \text{Acc}_{\text{EGCA}} - \text{Acc}_{\text{Others}}. \tag{11}$$

The following sections list the detailed analysis for these three datasets, including *Office-31*, *Office-Home*, and *DomainNet*.

1) Office-31: According to Table VI, we can observe that EGCA outperforms all methods in all transfer tasks without domain labels or source data, except $\mathbf{A} \rightarrow \{\mathbf{D}, \mathbf{W}\}$ for SHOT++ [60], with only −0.1% degradation. As for other methods, only without domain labels and with source data, EGCA outperforms some of them, with $+0.2\% \sim +5.0\%$ improvement. Also, EGCA performs worse than some BTDA methods, such as CGCT [14], DML [31], and MCDA [58], with $-2.9\% \sim -4.3\%$ degradation. In addition, EGCA performs better than ConMix with +0.4%, which has domain labels but cannot access to source data in prior. Furthermore, EGCA outperforms MTDA-ITA [10]. However, EGCA still has a certain gap with other methods, both with domain labels and source data, such as DCL [30] (−1.3%) and DCGCT [14] (−3.5%). To this end, source data are not the necessary condition for *Office-31* in SF-BTDA scenarios, while the domain labels are kind of important.

Authorized licensed use limited to: SUN YAT-SEN UNIVERSITY. Downloaded on March 22,2026 at 15:35:20 UTC from IEEE Xplore. Restrictions apply.

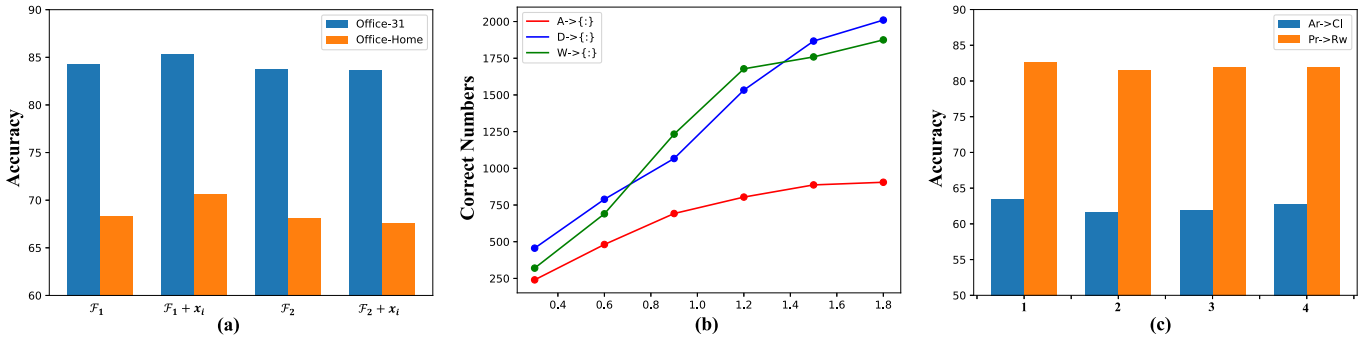

Fig. 6. More ablation studies in this article. (a) Average performance of *Office-31* and *Office-Home* that if adding the original image x_i as introduced in Section III-B. (b) Accuracy of pseudo-labels for the target data for *Office-31* dataset introduced in Sections III-C and III-D. (c) Performance under increasing k for two single target DA tasks in *Office-Home* dataset.

TABLE VII

COMPARISONS OF ACCURACY GAP AMONG DIFFERENT METHODS ON DOMAINNET. THE RED NUMBERS DENOTE THAT OUR PROPOSED EGCA OUTPERFORMS OTHER METHODS, WHILE THE BLUE NUMBERS DENOTE THAT OTHER METHODS OUTPERFORM OUR PROPOSED EGCA

Method		Cli	Inf	Pai	Qui	Rea	Ske	Avg
without domain labels nor source data	ResNet-101 [56]	+13.6	+15.8	+11.5	+6.8	+14.9	+17.5	+13.4
	G-SFDA [40]	+5.0	+5.9	+4.6	+4.3	+16.8	+7.8	+7.4
	GPUE [59]	+15.8	+9.9	+5.9	-0.5	+18.6	+15.0	+10.8
	SHOT++ [60]	+6.3	+3.8	+5.2	+2.1	+17.6	+4.9	+6.7
only without domain label	SE [9]	+17.9	+24.1	+22.8	+2.2	+19.5	+20.1	+17.8
	MCD [61]	+14.1	+13.5	+10.3	+5.6	+15.3	+17.3	+12.7
	CDAN [57]	+7.6	+5.5	+5.5	+3.5	+2.3	+4.0	+4.7
	MCC [62]	+5.6	+2.6	+4.9	+2.5	+7.5	+4.5	+4.6
	DADA [13]	+12.8	+12.6	+10.8	+3.1	+14.8	+17.0	+11.9
	CGCT [14]	+3.1	-0.7	+2.3	+6.0	-4.1	+0.1	+1.1
	DML [31]	+7.2	+7.2	+7.9	+3.3	+4.0	+3.4	+5.5
	MCDA [58]	+1.7	-4.7	+0.7	-1.8	-0.6	-1.6	-1.1
only without source data	ConMix [18]	-2.6	+3.4	-2.6	-1.5	+2.8	-1.4	-0.3
both with domain labels and source data	DCL [30]	+4.1	+1.2	+0.3	-4.5	+0.1	-1.2	+0.0
	DCGCT [14]	+2.2	+0.4	+0.0	-3.3	-4.3	-1.0	-1.0

2) Office-Home: From Table VI, we can conclude that EGCA outperforms all methods in all transfer tasks without domain labels or source data, except **Ar** $\rightarrow$ {**Cl, Pr, Rw**} for SHOT++ [60] and **Cl** $\rightarrow$ {**Ar, Pr, Rw**} for G-SFDA [40], with only -0.1% and -0.7% degradation, respectively. In addition, for other methods only without domain labels and with source data, EGCA outperforms all of them with respect to average accuracy, with $+3.3\% \sim +16.3\%$ improvement, except MCDA [58], with only -0.4% gap. Also, EGCA has comparable performance with ConMix (with only -0.4%), which has domain labels but cannot access to source data in prior. Furthermore, EGCA outperforms all methods both with domain labels and source data with considerable improvements ($+0.8 \sim +6.3\%$). Therefore, for *Office-Home*, source data and domain labels are not the necessary conditions in SF-BTDA scenarios.

3) DomainNet: According to Table VII, we can observe that EGCA outperforms all methods in all transfer tasks without domain labels or source data, except when *Qui* as the source domain for GPUE [59], with only -0.5% degradation. As the same as *Office-Home*, for other methods only without domain labels and with source data, EGCA significantly outperforms all of them, with $+1.1\% \sim +17.8\%$ improvement, except MCDA [58] (only with -1.1% degradation). Also, EGCA only performs slightly worse than ConMix with -0.3%, which has domain labels but cannot access to source data in prior. In addition, although other methods both with domain labels and source data outperform EGCA in average accuracy, EGCA considerably outperforms them for **Cli** $\rightarrow$ {**Inf, Pai, Qui, Rea, Ske**}, with $+4.1\%$ and $+2.2\%$ for DCL [30] and DCGCT [14], respectively. Therefore, in most cases, source data and domain labels are not the necessary condition for *DomainNet* in SF-BTDA scenarios.

All in all, although our proposed EGCA *does not* have domain labels in prior, our proposed EGCA also outperforms most of the existing methods that have domain labels for SF-BTDA scenarios for all datasets. Also, without source data, our proposed EGCA outperforms most of the methods that can access source data in prior for *Office-Home* and *DomainNet*, but EGCA performs considerably worse than existing methods with source data for *Office-31*.

Furthermore, Table VIII displays overall accuracy on different datasets with different k. We observe that choosing k as the number of subtargets may not lead to superior performance. While the k is larger, the accuracy only has slight changes. Therefore, we select the same value of k as the true number of target domains in each dataset to report. In addition, our

Authorized licensed use limited to: SUN YAT-SEN UNIVERSITY. Downloaded on March 22,2026 at 15:35:20 UTC from IEEE Xplore. Restrictions apply.

TABLE VIII

OVERALL ACCURACY (%) ON OFFICE-31, OFFICE-HOME, AND DOMAINNET WITH DIFFERENT k. * DENOTES THE REPORTED RESULT FOR OUR PROPOSED EGCA IN TABLES II AND III

k	2	3	4	5	6
Office-31	86.4*	86.1	**86.9**	85.7	86.6
Office-Home	74.4	76.1*	**77.0**	76.1	76.6
DomainNet	24.2	32.9	33.3	**34.6***	33.8

TABLE IX

ECE RESULTS FOR OFFICE-31. SMALL ECE INDICATES THAT THE MODEL IS BETTER CALIBRATED. ALL METHODS USE THE RESNET-50 AS THE BACKBONE

Method	A	D	W
EGCA (w/o L_{CEL})	0.042	0.105	0.117
EGCA (full)	**0.040**	**0.084**	**0.109**

TABLE X

AVERAGE OVERALL ACCURACY (%) ON OFFICE-31 USING DIFFERENT SOURCE MODELS FOR DIFFERENT METHODS

Method	ResNet-18	ResNet-34	ResNet-50	Avg
Source Only	61.3	67.9	70.7	66.6
DANN [7]	65.7	71.6	73.4	70.2
MTDA-ITA [10]	66.9	72.7	85.2	74.9
AMEAN [12]	63.2	70.1	80.2	71.2
SHOT++ [60]	70.1	76.7	83.8	76.9
ECGA (ours)	**74.5**	**82.0**	**86.4**	**81.0**

proposed ECGA is also obviously suitable for the MTDA setting, significantly outperforms other SOTA methods, such as DCDCT [14] and DCL [30].

B. ECE Results

Although our proposed CEL module can improve the performance on the SF-BTDA task (as shown in Table IX), we further delve into the question whether if the performance gain of CEL results from better calibrating a classification model. Therefore, we adopt the widely used expected calibration error (ECE) [73] to evaluate the model calibration performance of our full method EGCA (full) and its variant without CEL loss L_{CEL}. We evaluate ECE through partitioning predictions into M equally spaced bins (similar to the reliability diagrams) and taking a weighted average of the bins' difference between accuracy and confidence, which can be formulated as

$$\text{ECE} = \sum_{m=1}^{M} \frac{|B_m|}{n} |\text{acc}(B_m) - \text{conf}(B_m)| \tag{12}$$

in which n is the number of samples. The difference between acc and conf for a given bin represents the calibration gap. Quantitative results are reported in Table IX. It shows that L_{CEL} can reduce the ECE values in all transfer scenarios in *Office-31*. This validates our claim that the proposed L_{CEL} could calibrate the SF-BTDA model.

While CEL improves pseudo-label quality iteratively, its effectiveness may depend on the performance of the initial model in generating reasonably good pseudo-labels. Therefore, we discuss the robustness of the method under different initial models. Table X lists the average overall accuracy in the *Office-31* dataset using different source models (i.e., ResNet-18, ResNet-34, and ResNet-50) for different methods. The experimental results indicate that our proposed EGCA performs considerable accuracy gains compared to different initial models, which could prove the robustness under varying initial model qualities.

TABLE XI

OVERALL ACCURACY (%) ON OFFICE-31, OFFICE-HOME, AND DOMAINNET WITH DIFFERENT PARAMETER SELECTIONS FOR η_c AND η_u

Parameters	η_c	η_u	Office-31	Office-Home	DomainNet
Fixed Parameters	0.4	0.4	82.3	65.1	25.9
	0.5	0.4	77.4	70.5	31.0
	0.5	0.5	84.2	68.9	30.3
	0.6	0.6	83.9	68.3	30.1
	0.7	0.5	84.0	69.3	32.7
	0.7	0.6	80.3	66.2	29.8
Adaptive Parameters			**85.3**	**70.6**	**33.4**

C. Different Parameter Selections of η_c and η_u

We compare different parameter selection methods for η_c and η_u that are introduced in Section III-C. Here, we first display the performance of fixed parameters for η_c and η_u in Table XI. We can find that adaptive parameters can perform better than fixed parameters. Different from other high-quality selection schemes in the SFDA setting [46], [47], [48], [49], 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 $c_u > \eta_u$, 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 hyperparameters for η_c and η_u.

D. Theoretical Understanding

Inspired by the generalized label shift theorem in [74], in SF-BTDA setting, for any classifier $\hat{Y} = (h \circ g)(X)$, the blended target error rate could be formulated as $\|(1/K)\sum_j^K \epsilon_{T_j}\| \leq \text{BER}_{f_s}(\hat{Y}\|Y) + 2(c-1)\Delta_{\text{BTCE}}(\hat{Y})$. $\text{BER}_{f_s}(\hat{Y}\|Y)$ is the classification performance only related to the source model. Δ_{BTCE} measures the conditional distribution discrepancy of each class between the source and each target. In this sense, we only need to minimize the $\Delta_{\text{BTCE}}(\hat{Y})$. Therefore, the most important objective is to improve the quality of pseudo-labels, which may lead to category misalignment and negative transfer effects.

In addition, following the DA theory proposed in [4], the upper boundary of the expected error $\mathcal{E}_{\mathcal{T}}(h)$ of a hypothesis h on the target domain, can be expressed as the sum of three terms ($\mathcal{E}_{\mathcal{T}}(h) \leq \mathcal{E}_{\mathcal{S}}(h) + (1/2)d_{\mathcal{H}\Delta\mathcal{H}}(\mathcal{S},\mathcal{T}) + \lambda$): 1) expected error of h on the source domain, $\mathcal{E}_{\mathcal{S}}(h)$; 2) $\mathcal{H}\Delta\mathcal{H}$-distance $d_{\mathcal{H}\Delta\mathcal{H}}(\mathcal{S},\mathcal{T})$, measuring domain shift as the discrepancy between the disagreement of two hypotheses $h, h' \in \mathcal{H}\Delta\mathcal{H}$; and 3) the error λ of the ideal joint hypothesis h^* on both source and target domains. Fig. 7(a) and (b) visualizes the values of $\mathcal{H}\Delta\mathcal{H}$-distance (measuring domain shift as the discrepancy between the disagreement of two hypotheses $h, h' \in \mathcal{H}\Delta\mathcal{H}$) and λ (the ideal joint hypothesis h^* on both source and target domains), and demonstrate that

Authorized licensed use limited to: SUN YAT-SEN UNIVERSITY. Downloaded on March 22,2026 at 15:35:20 UTC from IEEE Xplore. Restrictions apply.

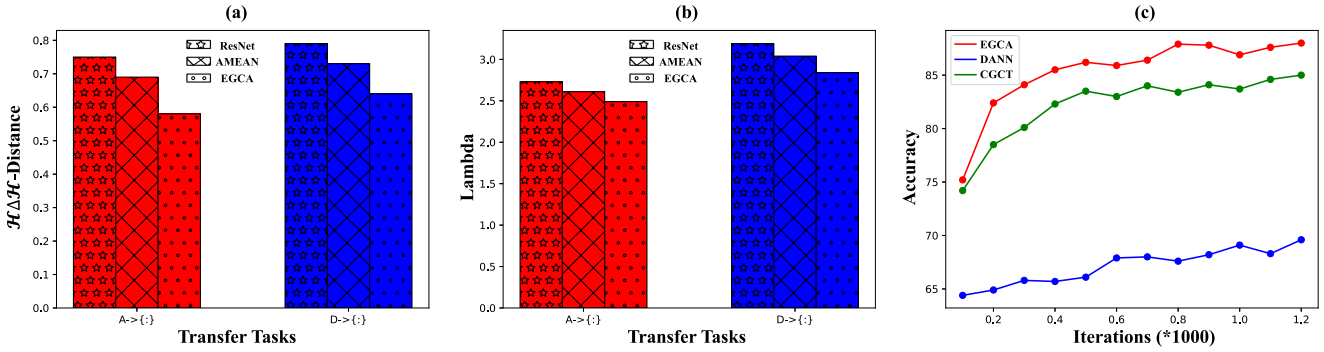

Fig. 7. Visualization of (a) $\mathcal{H}\Delta\mathcal{H}$-distance, (b) λ, and (c) convergence speed in *Office-31*.

TABLE XII
OVERALL ACCURACY (%) ON DIFFERENT SEMI-SUPERVISED CONTRASTIVE LEARNING FOR OFFICE-31, OFFICE-HOME, AND DOMAINNET

Different A	Office-31	Office-Home	DomainNet
None A	81.7	69.2	29.6
A	84.2	65.4	32.7
A^i	84.8	68.4	32.9
A^d	84.2	69.2	32.1
A^e	**85.3**	**70.6**	**33.4**

TABLE XIII
OVERALL ACCURACY (%) ON DIFFERENT β FOR OFFICE-31, OFFICE-HOME, AND DOMAINNET

Different β	Office-31	Office-Home	DomainNet
0.05	82.0	65.3	30.0
0.1	84.2	67.8	31.8
0.5	**85.5**	69.5	32.2
1.0	85.3	**70.6**	**33.4**
2.0	83.1	68.2	30.2

EGCA leads to a lower $\mathcal{H}\Delta\mathcal{H}$-distance as well as a lower λ, indicating a lower generalization error. Note that our EGCA balances the classification ability and transferable pattern. Furthermore, Fig. 7(c) indicates that EGCA can also accelerate the convergence process with a higher accuracy.

E. Sensitive Analysis

There are much less hyperparameters that need to be tuned. To show the sensitivity of domain distance matrix, we list different semi-supervised contrastive learning settings in Table XII. Notably, we keep the same experimental setting based on L_{CEL}. None A denotes the naive contrastive learning using (6). A denotes the semi-supervised contrastive learning using 7. A^i and A^d denote only evidential weights and only domain-distance weights, respectively. A^e denotes the evidential and domain-distance matrix. From Table XII, we can observe that A^e achieves the highest accuracy for these three public datasets.

We also analyze the behavior of our proposed EGCA when changing the tradeoff parameters (i.e., β), balancing three different loss terms between L_{CEL} and L_{GCL} on *Office-31*, *Office-Home*, and *DomainNet* in the SF-BTDA setting [as presented in (10)]. Notably, we set suitable hyperparameters in the validation dataset for DomainNet, while in the test dataset for other DA datasets, just like other DA works. From Table XIII, we select β as 1.0 in all our experiments.

VI. CONCLUSION

In this article, we are the first to propose a new DA setting (i.e., SF-BTDA), which is a more practical and challenging task where we cannot access to source-domain data while facing mixed multiple target domains without any domain labels in prior. To address this scenario, we propose 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 coexistence 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.

REFERENCES

[1] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 25, 2012, pp. 1–9.
[2] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 3431–3440.
[3] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2961–2969.
[4] S. Ben-David, J. Blitzer, K. Crammer, A. Kulesza, F. Pereira, and J. W. Vaughan, "A theory of learning from different domains," *Mach. Learn.*, vol. 79, nos. 1–2, pp. 151–175, May 2010.
[5] S. Cui, S. Wang, J. Zhuo, C. Su, Q. Huang, and Q. Tian, "Gradually vanishing bridge for adversarial domain adaptation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 12455–12464.
[6] W. Chen, Y. Wen, J. Zheng, J. Huang, and H. Fu, "BAN: A universal paradigm for cross-scene classification under noisy annotations from RGB and hyperspectral remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 63, 2025, Art. no. 5610213.
[7] Y. Ganin et al., "Domain-adversarial training of neural networks," *J. Mach. Learn. Res.*, vol. 17, no. 1, pp. 2030–2096, 2016.

Authorized licensed use limited to: SUN YAT-SEN UNIVERSITY. Downloaded on March 22,2026 at 15:35:20 UTC from IEEE Xplore. Restrictions apply.

[8] M. Long, H. Zhu, J. Wang, and M. I. Jordan, "Unsupervised domain adaptation with residual transfer networks," in *Proc. 30th Int. Conf. Neural Inf. Process. Syst.*, 2016, pp. 136–144.

[9] G. French, M. Mackiewicz, and M. Fisher, "Self-ensembling for visual domain adaptation," in *Proc. Int. Conf. Learn. Represent.*, 2017, pp. 1–15.

[10] B. Gholami, P. Sahu, O. Rudovic, K. Bousmalis, and V. Pavlovic, "Unsupervised multi-target domain adaptation: An information theoretic approach," *IEEE Trans. Image Process.*, vol. 29, pp. 3993–4002, 2020.

[11] X. Yang, C. Deng, T. Liu, and D. Tao, "Heterogeneous graph attention network for unsupervised multiple-target domain adaptation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 4, pp. 1992–2003, Apr. 2022.

[12] Z. Chen, J. Zhuang, X. Liang, and L. Lin, "Blending-target domain adaptation by adversarial meta-adaptation networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 2248–2257.

[13] X. Peng, Z. Huang, X. Sun, and K. Saenko, "Domain agnostic learning with disentangled representations," in *Proc. Int. Conf. Mach. Learn.*, 2019, pp. 5102–5112.

[14] S. Roy, E. Krivosheev, Z. Zhong, N. Sebe, and E. Ricci, "Curriculum graph co-teaching for multi-target domain adaptation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 5351–5360.

[15] J. N. Kundu, N. Venkat, and R. V. Babu, "Universal source-free domain adaptation," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2020, pp. 4544–4553.

[16] Y. Liu, W. Zhang, and J. Wang, "Source-free domain adaptation for semantic segmentation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 1215–1224.

[17] Y. Zhang, Z. Wang, and W. He, "Class relationship embedded learning for source-free unsupervised domain adaptation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2023, pp. 7619–7629.

[18] V. Kumar, R. Lal, H. Patil, and A. Chakraborty, "CoNMix for source-free single and multi-target domain adaptation," in *Proc. IEEE/CVF Winter Conf. Appl. Comput. Vis. (WACV)*, Jan. 2023, pp. 4167–4177.

[19] J. Zheng et al., "Open-set domain adaptation for scene classification using multi-adversarial learning," *ISPRS J. Photogramm. Remote Sens.*, vol. 208, pp. 245–260, Feb. 2024.

[20] Q. Li, Y. Wen, J. Zheng, Y. Zhang, and H. Fu, "HyUniDA: Breaking label set constraints for universal domain adaptation in cross-scene hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 62, 2024, Art. no. 5518415.

[21] Q. Li, Y. Zhang, J. Zheng, Y. Zhang, J. Huang, and H. Fu, "Boosting universal domain adaptation in remote sensing with dual-classifiers consistency discrimination and cross-domain feature mixup," *IEEE Trans. Geosci. Remote Sens.*, vol. 63, 2025, Art. no. 5515015.

[22] Y. Li, L. Yuan, and N. Vasconcelos, "Bidirectional learning for domain adaptation of semantic segmentation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 6936–6945.

[23] H.-K. Hsu et al., "Progressive domain adaptation for object detection," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, Mar. 2020, pp. 749–757.

[24] J. Zheng et al., "Cross-regional oil palm tree counting and detection via a multi-level attention domain adaptation network," *ISPRS J. Photogramm. Remote Sens.*, vol. 167, pp. 154–177, Sep. 2020.

[25] I. Goodfellow et al., "Generative adversarial nets," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 27, 2014, pp. 1–9.

[26] H. Xia, H. Zhao, and Z. Ding, "Adaptive adversarial network for source-free domain adaptation," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2021, pp. 9010–9019.

[27] M. Long, Y. Cao, J. Wang, and M. Jordan, "Learning transferable features with deep adaptation networks," in *Proc. 32nd Int. Conf. Mach. Learn.*, vol. 37, Jul. 2015, pp. 97–105.

[28] M. Long, H. Zhu, J. Wang, and M. I. Jordan, "Deep transfer learning with joint adaptation networks," in *Proc. Int. Conf. Mach. Learn.*, 2017, pp. 2208–2217.

[29] Y. Liang, S. Cao, J. Zheng, X. Zhang, J. Huang, and H. Fu, "Low saturation confidence distribution-based test-time adaptation for cross-domain remote sensing image classification," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 139, May 2025, Art. no. 104463.

[30] L. T. Nguyen-Meidine, A. Belal, M. Kiran, J. Dolz, L.-A. Blais-Morin, and E. Granger, "Unsupervised multi-target domain adaptation through knowledge distillation," in *Proc. IEEE/CVF Winter Conf. Appl. Comput. Vis.*, Jan. 2021, pp. 1339–1347.

[31] J. Wang, C. Zhong, C. Feng, Y. Zhang, J. Sun, and Y. Yokota, "Discriminative mutual learning for multi-target domain adaptation," in *Proc. 26th Int. Conf. Pattern Recognit. (ICPR)*, Aug. 2022, pp. 2900–2906.

[32] L. T. Nguyen-Meidine, M. Kiran, M. Pedersoli, J. Dolz, L.-A. Blais-Morin, and E. Granger, "Incremental multi-target domain adaptation for object detection with efficient domain transfer," *Pattern Recognit.*, vol. 129, Sep. 2022, Art. no. 108771.

[33] J. Zheng 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, 2022, Art. no. 5609213.

[34] Z. Liu et al., "Open compound domain adaptation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 12406–12415.

[35] R. Gong et al., "Cluster, split, fuse, and update: Meta-learning for open compound domain adaptive semantic segmentation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 8340–8350.

[36] Z. Qiu et al., "Source-free domain adaptation via avatar prototype generation and adaptation," in *Proc. 30th Int. Joint Conf. Artif. Intell.*, Aug. 2021, pp. 2921–2927.

[37] J. Tian, J. Zhang, W. Li, and D. Xu, "VDM-DA: Virtual domain modeling for source data-free domain adaptation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 32, no. 6, pp. 3749–3760, Jun. 2022.

[38] W. Chen, L. Lin, S. Yang, D. Xie, S. Pu, and Y. Zhuang, "Self-supervised noisy label learning for source-free unsupervised domain adaptation," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2022, pp. 10185–10192.

[39] J. Liang, D. Hu, and J. Feng, "Do we really need to access the source data? Source hypothesis transfer for unsupervised domain adaptation," in *Proc. Int. Conf. Mach. Learn.*, 2020, pp. 6028–6039.

[40] S. Yang, Y. Wang, J. Van De Weijer, L. Herranz, and S. Jui, "Generalized source-free domain adaptation," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2021, pp. 8978–8987.

[41] J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, "How transferable are features in deep neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 27, 2014, pp. 3320–3328.

[42] A. Amini, W. Schwarting, A. P. Soleimany, and D. Rus, "Deep evidential regression," in *Proc. Adv. Neural Inf. Process. Syst.*, 2019, pp. 14927–14937.

[43] W. Bao, Q. Yu, and Y. Kong, "Evidential deep learning for open set action recognition," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2021, pp. 13329–13338.

[44] K. Sentz and S. Ferson, "Combination of Evidence in Dempster–Shafer Theory," Sandia Nat. Laboratories, 2002.

[45] R. Krishnan and O. Tickoo, "Improving model calibration with accuracy versus uncertainty optimization," in *Proc. Adv. Neural Inf. Process. Syst.*, 2020, pp. 18237–18248.

[46] D. Guillory, V. Shankar, S. Ebrahimi, T. Darrell, and L. Schmidt, "Predicting with confidence on unseen distributions," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2021, pp. 1114–1124.

[47] J. Dong, Z. Fang, A. Liu, G. Sun, and T. Liu, "Confident anchor-induced multi-source free domain adaptation," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 34, 2021, pp. 2848–2860.

[48] Y. Wang, J. Peng, and Z. Zhang, "Uncertainty-aware pseudo label refinery for domain adaptive semantic segmentation," in *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, Oct. 2021, pp. 9092–9101.

[49] N. Karim, N. C. Mithun, A. Rajvanshi, H.-P. Chiu, S. Samarasekera, and N. Rahnavard, "C-SFDA: A curriculum learning aided self-training framework for efficient source free domain adaptation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2023, pp. 24120–24131.

[50] T. Chen, S. Kornblith, M. Norouzi, and G. E. Hinton, "A simple framework for contrastive learning of visual representations," in *Proc. 37th Int. Conf. Mach. Learn.*, vol. 119, 2020, pp. 1597–1607.

[51] K. He, H. Fan, Y. Wu, S. Xie, and R. Girshick, "Momentum contrast for unsupervised visual representation learning," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 9729–9738.

[52] R. D. Hjelm et al., "Learning deep representations by mutual information estimation and maximization," in *Proc. Int. Conf. Learn. Represent.*, 2018, pp. 1–24.

[53] O. Henaff, "Data-efficient image recognition with contrastive predictive coding," in *Proc. Int. Conf. Mach. Learn. (PMLR)*, 2020, pp. 4182–4192.

[54] Y. Tian, D. Krishnan, and P. Isola, "Contrastive multiview coding," in *Proc. Eur. Conf. Comput. Vis*. Springer, 2020, pp. 776–794.

Authorized licensed use limited to: SUN YAT-SEN UNIVERSITY. Downloaded on March 22,2026 at 15:35:20 UTC from IEEE Xplore. Restrictions apply.

[55] A. Van Den Oord, Y. Li, and O. Vinyals, "Representation learning with contrastive predictive coding," 2018, *arXiv:1807.03748*.
[56] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
[57] M. Long, Z. Cao, J. Wang, and M. I. Jordan, "Conditional adversarial domain adaptation," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, pp. 1640–1650.
[58] P. Xu, B. Wang, and C. X. Ling, "Class overwhelms: Mutual conditional blended-target domain adaptation," in *Proc. AAAI Conf. Artif. Intell.*, 2023, vol. 37, no. 3, pp. 3036–3044.
[59] M. Litrico, A. Del Bue, and P. Morerio, "Guiding pseudo-labels with uncertainty estimation for source-free unsupervised domain adaptation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2023, pp. 7640–7650.
[60] J. Liang, D. Hu, Y. Wang, R. He, and J. Feng, "Source data-absent unsupervised domain adaptation through hypothesis transfer and labeling transfer," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 11, pp. 8602–8617, Nov. 2022.
[61] K. Saito, K. Watanabe, Y. Ushiku, and T. Harada, "Maximum classifier discrepancy for unsupervised domain adaptation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 3723–3732.
[62] Y. Jin, X. Wang, M. Long, and J. Wang, "Minimum class confusion for versatile domain adaptation," in *Proc. Eur. Conf. Comput. Vis.*, 2020, pp. 464–480.
[63] P. Khosla et al., "Supervised contrastive learning," in *Proc. Adv. Neural Inf. Process. Syst.*, 2020, pp. 18661–18673.
[64] K. Saenko, B. Kulis, M. Fritz, and T. Darrell, "Adapting visual category models to new domains," in *Proc. 11th Eur. Conf. Comput. Vis.*, 2010, pp. 213–226.
[65] H. Venkateswara, J. Eusebio, S. Chakraborty, and S. Panchanathan, "Deep hashing network for unsupervised domain adaptation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 5018–5027.
[66] X. Peng, Q. Bai, X. Xia, Z. Huang, K. Saenko, and B. Wang, "Moment matching for multi-source domain adaptation," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 1406–1415.
[67] A. Paszke et al., "PyTorch: An imperative style, high-performance deep learning library," in *Proc. Adv. Neural Inf. Process. Syst.*, 2019, pp. 1–12.
[68] O. Russakovsky et al., "ImageNet large scale visual recognition challenge," *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, Dec. 2015.
[69] X. Chen and K. He, "Exploring simple Siamese representation learning," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 15750–15758.
[70] A. Singh, "CLDA: Contrastive learning for semi-supervised domain adaptation," in *Proc. NIPS*, 2021, pp. 5089–5101.
[71] T. Kalluri, A. Sharma, and M. Chandraker, "MemSAC: Memory augmented sample consistency for large scale domain adaptation," in *Proc. Eur. Conf. Comput. Vis.*, 2022, pp. 550–568.
[72] M. Xie, S. Li, R. Zhang, and C. H. Liu, "Dirichlet-based uncertainty calibration for active domain adaptation," in *Proc. Int. Conf. Learn. Represent.*, 2023, pp. 1–22.
[73] C. Guo, G. Pleiss, Y. Sun, and K. Q. Weinberger, "On calibration of modern neural networks," in *Proc. 34th Int. Conf. Mach. Learn.*, 2017, pp. 1321–1330.
[74] R. T. D. Combes, H. Zhao, Y.-X. Wang, and G. J. Gordon, "Domain adaptation with conditional distribution matching and generalized label shift," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 33, 2020, pp. 19276–19289.

Juepeng Zheng (Member, IEEE) received the bachelor's degree from the College of Surveying and Geoinformatics, Tongji University, Shanghai, China, in 2019, and the Ph.D. degree from the Department of Earth System Science, Tsinghua University, Beijing, China, in 2023.

He is currently an Assistant Professor with the School of Artificial Intelligence, Sun Yat-sen University, Zhuhai, China. He also works with the National Supercomputing Center in Shenzhen, Shenzhen, China. His research interests include remote sensing image understanding, high-performance computing, and deep learning.

Guowen Li received the B.S. degree from the School of Intelligent Systems Science and Engineering, Jinan University, Zhuhai, China, in 2023. He is currently pursuing the M.S. degree with the School of Artificial Intelligence, Sun Yat-sen University, Zhuhai.

His primary research interests include artificial intelligence for climate and transfer learning.

Yibin Wen received the B.Eng. degree from the School of Artificial Intelligence, Sun Yat-sen University, Zhuhai, China, in 2025, where he is currently pursuing the M.Eng. degree.

His research interests include intelligent interpretation and analysis of remote sensing imagery and domain adaptation.

Jinxiao Zhang received the bachelor's and master's degrees from Shandong University, Jinan, China, in 2019 and 2022, respectively. She is currently pursuing the Ph.D. degree in ecological science with the Department of Earth System Science, Tsinghua University, Beijing, China.

Her research interests include high-performance computing, remote sensing compression, and their applications.

Runmin Dong (Member, IEEE) received the Ph.D. degree from Tsinghua University, Beijing, China, in 2022.

She was a "Shuimu Scholar" Post-Doctoral Fellow at Tsinghua University, and held visiting positions at The University of Hong Kong, Hong Kong, and the Technical University of Munich, Munich, Germany. She is currently an Associate Professor with the School of Artificial Intelligence, Sun Yat-sen University, Zhuhai, China. Her research interests include computer vision and remote sensing applications.

Haohuan Fu (Senior Member, IEEE) received the Ph.D. degree in computing from Imperial College London, London, U.K., in 2009.

He is currently a Professor with the Ministry of Education Key Laboratory for Earth System Modeling and the Department of Earth System Science, Tsinghua University, Beijing, China. He is also the Deputy Director of the National Supercomputing Center, Wuxi, China. His research interests include design methodologies for highly efficient and highly scalable simulation applications that can take advantage of emerging multicore, many-core, and reconfigurable architectures, and make full utilization of current peta-flops and future exa-flops supercomputers; and intelligent data management, analysis, and data mining platforms that combine the statistics methods and machine learning technologies.

Authorized licensed use limited to: SUN YAT-SEN UNIVERSITY. Downloaded on March 22,2026 at 15:35:20 UTC from IEEE Xplore. Restrictions apply.