



## Review

## Source-free unsupervised domain adaptation: A survey

Yuqi Fang<sup>a</sup>, Pew-Thian Yap<sup>a</sup>, Weili Lin<sup>a</sup>, Hongtu Zhu<sup>b</sup>, Mingxia Liu<sup>a,\*</sup><sup>a</sup> Department of Radiology and Biomedical Research Imaging Center, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, United States<sup>b</sup> Department of Biostatistics and Biomedical Research Imaging Center, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, United States

## ARTICLE INFO

## Keywords:

Domain adaptation

Source-free

Unsupervised learning

Survey

## ABSTRACT

Unsupervised domain adaptation (UDA) via deep learning has attracted appealing attention for tackling domain-shift problems caused by distribution discrepancy across different domains. Existing UDA approaches highly depend on the accessibility of source domain data, which is usually limited in practical scenarios due to privacy protection, data storage and transmission cost, and computation burden. To tackle this issue, many source-free unsupervised domain adaptation (SFUDA) methods have been proposed recently, which perform knowledge transfer from a pre-trained source model to the unlabeled target domain with source data inaccessible. A comprehensive review of these works on SFUDA is of great significance. In this paper, we provide a timely and systematic literature review of existing SFUDA approaches from a technical perspective. Specifically, we categorize current SFUDA studies into two groups, *i.e.*, white-box SFUDA and black-box SFUDA, and further divide them into finer subcategories based on different learning strategies they use. We also investigate the challenges of methods in each subcategory, discuss the advantages/disadvantages of white-box and black-box SFUDA methods, conclude the commonly used benchmark datasets, and summarize the popular techniques for improved generalizability of models learned without using source data. We finally discuss several promising future directions in this field.

## 1. Introduction

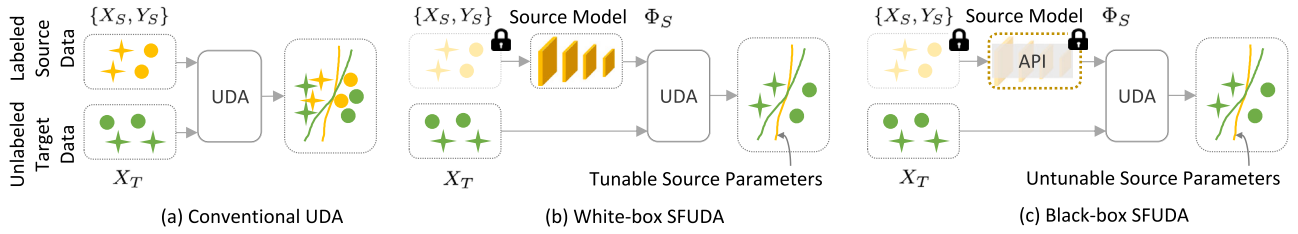
Deep learning has emerged as a promising technique and made remarkable progress in the field of computer vision (Hassaballah & Awad, 2020; Voulodimos, Doulamis, Doulamis, & Protopapadakis, 2018), medical data analysis (Litjens et al., 2017; Shen, Wu, & Suk, 2017), and natural language processing (Otter, Medina, & Kalita, 2020; Young, Hazarika, Poria, & Cambria, 2018). For problems with multiple domains (*e.g.*, different datasets or imaging sites), the typical learning process of a deep neural network is to transfer the model learned on a source domain to a target domain. However, performance degradation is often observed when there exists a distribution gap between source and target domains, which is termed “domain shift” problem (Li, Yang, Song, & Hospedales, 2017; Sankaranarayanan, Balaji, Jain, Lim, & Chellappa, 2018; Zhou, Liu, Qiao, Xiang, & Loy, 2022). To tackle this problem, various domain adaptation algorithms (Guan & Liu, 2021; Wang & Deng, 2018) have been proposed to perform knowledge transfer by reducing inter-domain distribution discrepancy. To avoid intensive burden of data annotation, unsupervised domain adaptation has achieved much progress (Dong, Cong, Sun, Fang, & Ding, 2021; Fang, Wang, Potter, & Liu, 2022; Ganin & Lempitsky, 2015; Saito, Watanabe, Ushiku, & Harada, 2018). As illustrated in Fig. 1(a), unsupervised domain adaptation aims to transfer knowledge from a labeled

source domain to a target domain without accessing any target label information.

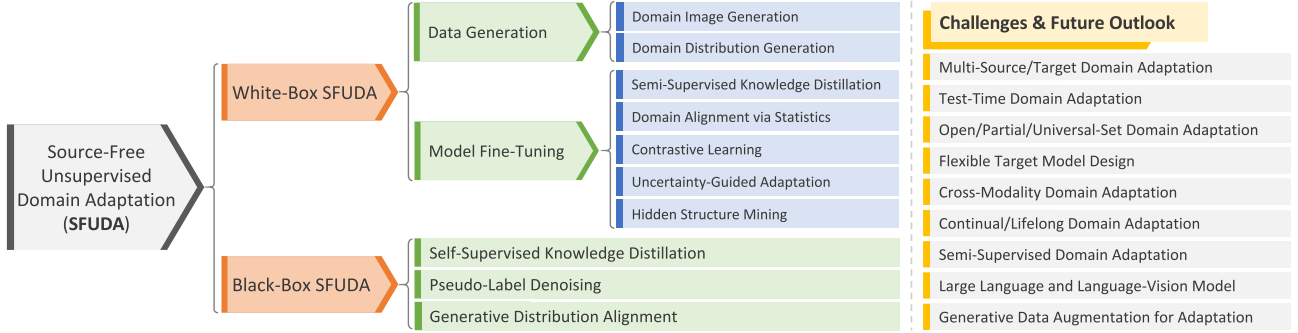
Existing deep learning studies on unsupervised domain adaptation highly depend on the accessibility of source data, which is usually limited in practical scenarios due to the following possible reasons. (1) *Data privacy protection*. Many source datasets containing confidential information, such as medical and facial data, are not available to the third parties due to privacy and security protection (Qu, Liu, Zhu, Nie, & Zhang, 2024; Yu, Xi, et al., 2023). (2) *Data storage and transmission cost*. The storage and transmission of large-scale source datasets, such as ImageNet (Deng et al., 2009), could bring much economic burden (Kumar, Lal, Patil, & Chakraborty, 2023; Xia, Zhao, & Ding, 2021; Zhao et al., 2023). (3) *Computation burden*. Training on extremely large source datasets requires high computational resources, which is not practical, especially in real-time deployment cases (Karim et al., 2023; Zhang et al., 2023). Thus, there is a high demand for source-free unsupervised domain adaptation (SFUDA) methods that transfer a pre-trained source model to the unlabeled target domain *without accessing any source data* (Huang, Guan, Xiao, & Lu, 2021; Kundu, Venkat, Babu, et al., 2020; Nayak, Mopuri, Jain, & Chakraborty, 2021; Yang, van de Weijer, Herranz, Jui, et al., 2021).

\* Corresponding author.

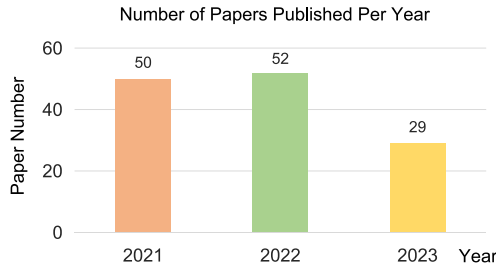
E-mail address: [mingxia.liu@med.unc.edu](mailto:mingxia.liu@med.unc.edu) (M. Liu).



**Fig. 1.** Illustration of (a) conventional unsupervised domain adaptation (UDA), (b) white-box source-free UDA (SFUDA), and (c) black-box SFUDA. Compared with (a) conventional UDA that relies on labeled source data  $\{X_S, Y_S\}$  and unlabeled target data  $X_T$ , (b, c) SFUDA performs knowledge transfer by directly leveraging a pre-trained source model  $\Phi_S$  and unlabeled target data  $X_T$ . The difference between (b) white-box SFUDA and (c) black-box SFUDA lies in whether the learnable parameters of the source model  $\Phi_S$  are accessible or not. API: application programming interface.



**Fig. 2.** Taxonomy of existing source-free unsupervised domain adaptation (SFUDA) methods and future outlook.



**Fig. 3.** Overview of the number of SFUDA studies (in terms of published years) that have been collected in this survey.

Many promising SFUDA algorithms have been developed recently to address problems in the fields of object recognition (Kundu, Bhambri, et al., 2022), semantic segmentation (Liu, Zhang, & Wang, 2021), image classification (Chen et al., 2022), object detection (Saltori, Lathuilière, Sebe, Ricci, & Galasso, 2020), face anti-spoofing (Liu, Chen, et al., 2022), etc. A comprehensive review of current studies on SFUDA and an outlook on future research directions are urgently needed. Liu, Zhang, Wang and Wang (2021) present a review on data-free knowledge transfer, where SFUDA only accounts for part of the review and the taxonomy of SFUDA is generally rough.

To fill the gap, in this paper, we provide a timely and thorough literature review of existing deep learning studies on source-free unsupervised domain adaptation. Our goal is to cover SFUDA studies of the past few years and provide a detailed and systematic SFUDA taxonomy. Specifically, we classify existing SFUDA approaches into two broad categories: (1) *white-box SFUDA* as shown in Fig. 1(b) and (2) *black-box SFUDA* as illustrated in Fig. 1(c). The difference between them lies in whether the model parameters of the pre-trained source model are available or not. Based on different learning strategies they use, we further subdivide the white-box and black-box SFUDA methods into finer categories, and the overall taxonomy is shown in Fig. 2. Moreover, we discuss the challenges and insights for methods in each category,

provide a comprehensive comparison between white-box and black-box SFUDA approaches, and summarize commonly used datasets in this field and popular techniques to improve model generalizability across different domains. Since SFUDA is still under vigorous development, we further discuss the main challenges and provide insights into potential future directions accordingly.

A recent paper (Yu, Li, Du, Zhu, & Shen, 2023) has reviewed source-free domain adaptation and its applications, and our survey significantly differs from this paper in the following aspects. (1) *Different coverage*. Our survey is with a more extensive coverage of SFUDA studies in terms of both quantity and timeliness, as shown in Fig. 3. In comparison, the previous review (Yu, Li, et al., 2023) included only 34, 39, and 3 papers in the years 2021, 2022, and 2023, respectively. (2) *Different problem settings*. This paper covers a wider spectrum of problem settings, including both white-box and black-box SFUDA, whereas the previous survey paper (Yu, Li, et al., 2023) focused primarily on white-box SFUDA. (3) *New research directions*. This survey offers a more comprehensive summary of newly arisen research problems and points out a broader range of future directions for SFUDA, such as large language models and generative artificial intelligence. The major contributions of this work are summarized as follows:

- Source-free unsupervised domain adaptation (SFUDA) has shown great significance but currently, there is a lack of review papers to summarize existing SFUDA studies. Our survey fills this gap, which helps researchers quickly grasp the fundamental concepts and methodologies in the field.
- Our survey divides current SFUDA methods into finer categories based on different learning strategies they use, and discusses their advantages and disadvantages in detail, which helps researchers choose the most suitable approach for their own work.
- We conclude the commonly used benchmark datasets in the SFUDA field, based on which the researchers can evaluate different methods for fair comparison. We also summarize the popular techniques for improved model generalizability, helping facilitate future works and spark new research ideas.

- We discuss promising future research directions in this field, which guides the trajectory of future investigations. And our paper includes an extensive list of references, serving as a valuable resource for researchers to explore specific aspects of the topic in more detail.

The rest of this survey is organized as follows. Sections 2 and 3 review existing white-box and black-box SFUDA methods, respectively. In Section 4, we compare white-box and black-box SFUDA and present useful strategies to improve model generalization. Section 5 discusses challenges of existing studies and future research directions. Finally, we conclude this paper in Section 6.

## 2. White-box source-free unsupervised domain adaptation

We denote the labeled source domain by  $D_S = \{X_S, Y_S\} \sim \{(x_s^i, y_s^i)\}_{i=1}^{n_s}, x_s^i \sim P_S\}$ . Here,  $X_S$  and  $Y_S$  represent source data and the corresponding label information, respectively;  $x_s^i$  represents  $i$ th source sample, following the marginal distribution  $P_S$ ;  $y_s^i$  denotes the label of  $x_s^i$ ; and  $n_s$  is the number of source data samples. Based on  $D_S$ , a well-trained source model  $\Phi_S$  can be obtained with supervised learning. We denote the unlabeled target domain by  $D_T = \{X_T\} \sim \{(x_t^i)\}_{i=1}^{n_t}, x_t^i \sim P_T\}$ , where  $X_T$  represents unlabeled target data,  $x_t^i$  denotes  $i$ th target sample, and  $P_T$  is the marginal distribution of  $x_t^i$  ( $P_T \neq P_S$ ). The goal of SFUDA is to learn a target model  $\Phi_T$  for predicting target label  $Y_T$  by only utilizing the pre-trained source model  $\Phi_S$  and unlabeled target data  $X_T$ . In the setting of white-box SFUDA, we aim to learn  $\Phi_T$ :  $\{\Phi_S, X_T\} \Rightarrow Y_T$ . Note that the source data  $\{X_S, Y_S\}$  cannot be accessed but the training parameters of the source model  $\Phi_S$  and unlabeled target data  $X_T$  are available. As shown in the upper middle of Fig. 2, existing white-box SFUDA studies can be divided into two categories: **Data Generation Method** and **Model Fine-Tuning Method**, with details elaborated as follows.

### 2.1. Data generation method

#### 2.1.1. Domain image generation

Many studies aim to generate source-like image data and achieve cross-domain adaptation by readily applying standard unsupervised domain adaptation techniques. Based on different image generation strategies, these studies can be divided into the following three subcategories: (1) batch normalization statistics transfer, (2) surrogate source data construction, and (3) generative adversarial network (GAN) based image generation.

(1) **Batch Normalization Statistics Transfer.** Considering that batch normalization (BN) stores the running mean and variance for a mini-batch of training data in each layer of a deep learning model, some studies (Hong, Zhang, & Chen, 2022; Hou & Zheng, 2020; Yang, Guo, Chen & Yuan, 2022) explicitly leverage such BN statistics for image style transfer, as illustrated in Fig. 4. For instance, Yang, Guo, et al. (2022) generate source-like images via a two-stage coarse-to-fine learning strategy. In the coarse image generation step, BN statistics stored in the source model are leveraged to preserve the style characteristics of source images and also maintain the content information of target data. In the fine image generation step, an image generator based on Fourier Transform is developed to remove ambiguous textural components of generated images and further improve image quality. With generated source-like images and given target images, a contrast distillation module and a compact consistency measurement module are designed to perform feature-level and output-level adaptation, respectively. Similarly, Hou and Zheng (2020) perform style transfer by matching BN statistics of generated source-style image features with those saved in the pre-trained source model for image translation. Hong et al. (2022) generate source-like images by designing a style-compensation transformation architecture guided by BN statistics stored in the source model and the generated reliable target pseudo-labels.

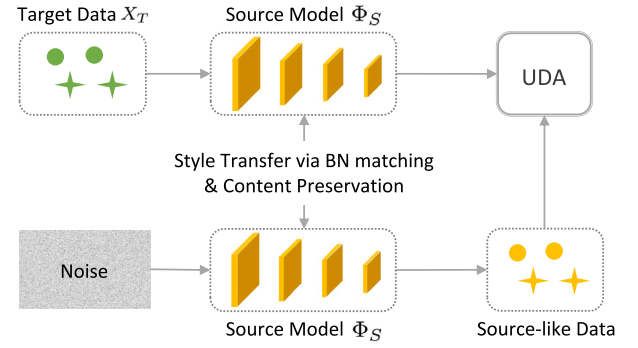


Fig. 4. Illustration of Batch Normalization Statistics Transfer methods for source image generation. By matching batch normalization (BN) statistics between the upper and lower branches, source-like data can be generated by preserving the target content but with source style. Unsupervised domain adaptation (UDA) is then performed between source-like data and target data.

(2) **Surrogate Source Data Construction.** To compensate for the inaccessible source domain, some studies (Ding, Sheng, Liang, Zheng, & He, 2023; Du et al., 2021; Tian, Ma, Zhang, Peng, & Xue, 2021; Yao, Guo, & Yang, 2021; Ye, Zhang, Ouyang, & Yuan, 2021) construct surrogate/proxy source data by selecting appropriate samples from the target domain directly, as illustrated in Fig. 5. For example, Tian, Ma, et al. (2021) construct pseudo source samples directly from the provided target samples under the guidance of a designed sample transport rule. The adaptation step and sample transport learning step are performed alternately to refine the approximated source domain and attain confident labels for target data, thus achieving effective cross-domain knowledge adaptation. Ding et al. (2023) build a category-balanced surrogate source domain using pseudo-labeled target samples based on a prototype similarity measurement. During model adaptation, intra-domain and inter-domain mixup regularizations are introduced to transfer label information from the proxy source domain to the target domain, as well as simultaneously eliminate negative effects caused by noisy labels. Ye et al. (2021) select target samples with high prediction confidence to construct a virtual source set that mimics source distribution. To align the target and virtual domains, they develop a weighted adversarial loss based on distribution and an uncertainty measurement to achieve cross-domain adaptation. Moreover, an uncertainty-aware self-training mechanism is proposed to iteratively produce the pseudo-labeled target set to further enhance adaptation performance. Du et al. (2021) construct a surrogate source domain by first selecting target samples near the source prototypes based on an entropy criterion, and then enlarging them by a mixup augmentation strategy (Zhang, Cisse, Dauphin, & Lopez-Paz, 2017). The adversarial training is then used to explicitly mitigate cross-domain distribution gap. Yao et al. (2021) simulate proxy source domain by freezing the source model and minimizing a supervised objective function for optimization. For the simulated source set, global fitting is enforced by a model gradient based equality constraint, which is optimized by an alternating direction method of multipliers algorithm (Boyd, Parikh, Chu, Peleato, Eckstein, et al., 2011).

(3) **GAN-based Image Generation.** Instead of approximating the source domain directly using existing target data, Kurmi, Subramanian, and Nambodiri (2021) simulate the source data by training a GAN-based generator, as illustrated in Fig. 6. Specifically, they first use a parametric conditional GAN to generate labeled proxy source data by treating the source classifier as an energy based function. Then, they learn feature patterns that are invariant across two domains via standard adversarial learning for further adaptation. Hou and Zheng (2021) also update an image generator framework but they aim to translate target images into the source-style ones instead of using the latent noise as in Kurmi et al. (2021). In their method, the knowledge adaptation

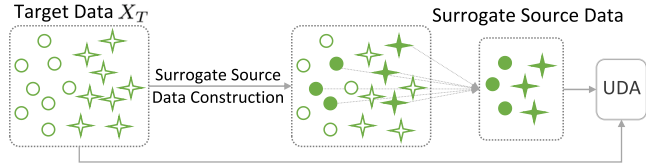


Fig. 5. Illustration of *Surrogate Source Data Construction* methods for source data generation. These methods first construct surrogate/proxy source data by selecting appropriate samples from the target domain and then perform standard unsupervised domain adaptation (UDA).

is achieved by training 1) a knowledge distillation loss that mitigates the difference between features of newly generated source-style images and those of target images, and 2) a relation-preserving loss that maintains channel-level relationship across different domains. Li, Jiao, Cao, Wong, Wu (2020) propose a *GAN-embedded generator* conditioned on a pre-defined label to generate target-style data. By incorporating real target samples, the learnable parameters of the generator and the adapted model can be updated in a collaborative manner. Moreover, two constraints, *i.e.*, weight regularization and clustering-based regularization, are utilized during model adaptation to preserve source knowledge and ensure high-confident target prediction, respectively.

### 2.1.2. Domain distribution generation

Instead of generating source-like images directly, some studies propose to align feature prototypes or feature distribution of source data (Ding et al., 2022; Qiu et al., 2021; Stan & Rostami, 2021a, 2021b; Tian, Zhang, Li & Xu, 2021) with those in the target domain. Specifically, Qiu et al. (2021) generate feature prototypes for each source category based on a conditional generator and produce pseudo-labels for the target data. The cross-domain prototype adaptation is achieved by aligning the features derived from pseudo-labeled target samples to source prototype with the same category label via contrastive learning. Tian, Zhang, et al. (2021) construct a virtual domain by simply sampling from an *approximated gaussian mixture model* (GMM) to mimic unseen source domain distribution. In terms of adaptation procedure, they reduce the distribution gap between the constructed virtual domain and the target domain via adversarial training, thus bypassing the inaccessible source domain. Their practice is based on the assumption that the feature prototype of each category can be mined from each row of the source classifier's weights (Chen, Liu, Kira, Wang & Huang, 2019). With the same assumption, Ding et al. (2022) leverage such source classifier weights and reliable target pseudo-labels derived by spherical k-means clustering to estimate source feature distribution. After that, proxy source data can be sampled from the estimated source distribution, and a conventional domain adaptation strategy (Kang, Jiang, Yang, & Hauptmann, 2019) is used to explicitly perform cross-domain feature distribution alignment. Stan and Rostami (2021a, 2021b) propose to first generate a prototypical distribution representing the source data in an embedding feature space via GMM, and then perform source-free adaptation by enforcing distribution alignment between source and target domains via sliced Wasserstein distance (Lee, Batra, Baig, & Ulbricht, 2019).

### 2.1.3. Challenges and insights

We classify existing domain image generation methods for SFUDA into three subcategories. We present the challenges of methods in each subcategory and our insights below.

- Among the above-mentioned three subcategories, the first one (*i.e.*, *batch normalization statistics transfer*) explicitly performs BN statistics matching between source and target domains for style transfer. Since the BN statistics of the source model are off-the-shelf, these methods are generally efficient and do not require complex model training. However, BN statistics mainly focus on keeping the style features

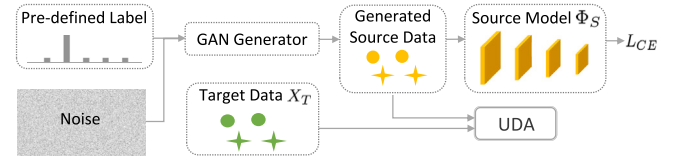


Fig. 6. Illustration of *Generative Adversarial Network (GAN) based Image Generation* methods for source data generation. Typically, a pre-defined label and random noise act as the inputs of a GAN-based generator. By utilizing the pre-trained source model, they synthesize source data for cross-domain adaptation.  $L_{CE}$ : Cross-entropy loss function.

while the content information cannot be well preserved. Therefore, this strategy is more applicable to scenarios where the contextual structure of images between source and target domains does not differ too much. It may not show good adaptation performance, *e.g.*, from a natural image to a cartoon image, since the content information has significant changes. Note that BN statistics transfer can also be used as a pre-processing step in source-free domain adaptation, and it can be combined with other strategies, *e.g.*, circular learning (Hong et al., 2022), for more effective knowledge transfer.

- Methods in the second subcategory (*i.e.*, *surrogate source data construction*) aim to approximate the proxy source domain using appropriate target samples directly, followed by conventional unsupervised domain adaptation. Their application is quite broad, including semantic segmentation (Ye et al., 2021), object recognition (Ding et al., 2023; Du et al., 2021; Yao et al., 2021), image classification (Tian, Ma, et al., 2021), and digital recognition (Du et al., 2021; Tian, Ma, et al., 2021). In general, methods in this group are straightforward and computation-efficient by avoiding introducing extra hyperparameters, which is different from generative models. However, because the proxy source samples are directly selected from the target domain, these generated source data may not effectively represent the original source domain. Moreover, how to effectively select informative target data for source data approximation is an important topic to be investigated. Some studies have proposed various strategies for target data selection based on entropy measurement (Ye et al., 2021), source prototype (Ding et al., 2023; Du et al., 2021), aggregated source decision boundary (Tian, Ma, et al., 2021), and equality constrained optimization (Yao et al., 2021). This is still an open but very interesting future direction. For multi-source settings, it is promising to study how to effectively utilize prediction results from different pre-trained source models for target data selection.
- Methods in the third category (*i.e.*, *GAN-based image generation*) typically synthesize images based on a generative model. Since the generator can model underlying complex distribution of source data with given random noise, GAN-based models generally create more diverse images compared with methods in the second category (*i.e.*, *surrogate source data construction*). However, these methods introduce additional frameworks and learnable parameters (*e.g.*, generators and discriminators), which may cost more computation resources. By comparing experimental results, we find the surrogate source data construction methods (Du et al., 2021; Yao et al., 2021) generally outperform the GAN-based generators (Kurmi et al., 2021; Li, Jiao, et al., 2020). The possible reason may be that the constructed source data in the former are closer to real data distributions, while those recovered in GAN-based methods usually suffer from a mode collapse problem (Ding et al., 2023) that leads to low-quality images. Note that the mode collapse problem can be partly mitigated by using a carefully tuned learning rate, manifold-guided training (Bang & Shim, 2021), and virtual mapping (Abusitta, Wahab, & Fung, 2021), which is worth exploring further.

Different from image generation methods (Section 2.1.1) that directly generate source/target-like images, the distribution generation methods (Section 2.1.2) generate feature prototype/distribution to



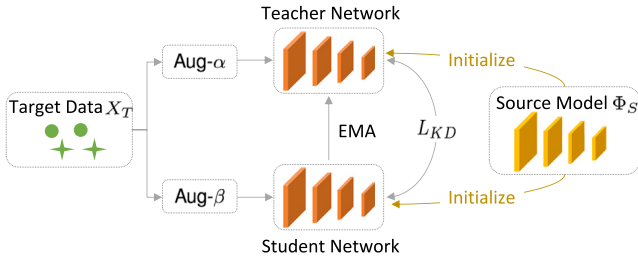


Fig. 7. Illustration of *Self-Supervised Knowledge Distillation* methods for source-free unsupervised domain adaptation. With target data from different augmentations as inputs, a teacher–student framework is utilized to exploit target features, where parameters of teacher network are usually exponential moving average (EMA) of those of student network. Aug- $\alpha$  and Aug- $\beta$  denote two data augmentation methods (e.g., flip, rotation, shift, noise addition, distortion, etc.), respectively.  $L_{KD}$ : Knowledge distillation loss function.

achieve cross-domain feature alignment. By comparing the reported experimental results, we find that the distribution generation approaches (Ding et al., 2022; Qiu et al., 2021; Tian, Zhang, et al., 2021) usually outperform the GAN-based image generation method (Li, Jiao, et al., 2020). And surrogate source data construction methods (Ding et al., 2023; Du et al., 2021) usually show superior performance compared with the distribution generation methods (Qiu et al., 2021; Tian, Zhang, et al., 2021). The underlying reason could be that the source distributions directly derived from the existing target data (Ding et al., 2023; Du et al., 2021) are more accurate and stable than the approximated ones (Qiu et al., 2021; Tian, Zhang, et al., 2021). How to drive the approximated source distribution to the real one can be further explored in the future.

## 2.2. Model fine-tuning method

Instead of generating source-like data for standard unsupervised domain adaptation, many studies attempt to fine-tune a pre-trained source model by exploiting unlabeled target data in a self-supervised training scheme. Based on different strategies for fine-tuning the source model, we divide existing studies into five subcategories: (1) self-supervised knowledge distillation, (2) domain alignment via statistics, (3) contrastive learning, (4) uncertainty-guided adaptation, and (5) hidden structure mining methods, as shown in Fig. 2. More details are introduced in the following.

### 2.2.1. Self-supervised knowledge distillation

Many studies (Chen et al., 2022; Liu & Yuan, 2022; Liu & Zhang, 2021; Raychaudhuri, Ta, Dutta, Lal, & Roy-Chowdhury, 2023; Tang, Shi, et al., 2021; VS, Oza, & Patel, 2023a, 2023b; VS, Oza, Sindagi, & Patel, 2022; VS, Valanarasu, & Patel, 2022; Xiong et al., 2021; Yang, Tang, et al., 2021; Yu et al., 2022) transfer knowledge learned from source data to the target model via knowledge distillation in a self-supervised manner, as illustrated in Fig. 7. In these works, most of them (Chen et al., 2022; Liu & Yuan, 2022; Liu & Zhang, 2021; Xiong et al., 2021; Yang, Tang, et al., 2021) achieve source-free domain adaptation via a *mean-teacher scheme* for knowledge transfer (Tarvainen & Valpola, 2017), where the target model not only learns from unseen target domain but also well preserves source model information. For instance, Liu and Yuan (2022) propose a self-supervised distillation scheme for automatic polyp detection. By means of keeping output consistency of weak and strong augmented polyp images, source knowledge is implicitly transferred to the target model with a mean teacher strategy (Tarvainen & Valpola, 2017). Besides, a diversification flow paradigm is designed to gradually eliminate the style sensitivity among different domains, further enhancing model robustness towards style diversification. Yang, Tang, et al. (2021) also propose a self-supervised mean-teacher approach for knowledge distillation, with a

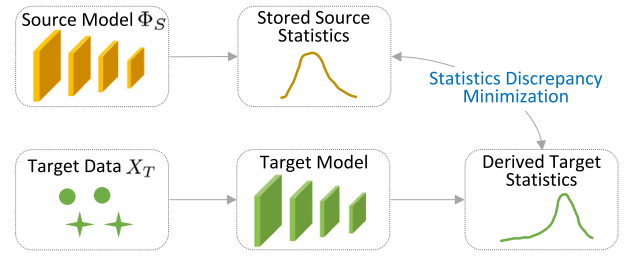


Fig. 8. Illustration of *Domain Alignment via Statistics* methods for source-free unsupervised domain adaptation. The corresponding methods leverage batch statistics stored in the pre-trained source model to approximate the distribution of inaccessible source data, and then perform cross-domain adaptation by reducing distribution discrepancy between source and target domains.

Transformer (Dosovitskiy et al., 2020) embedded. This helps the target model focus on object regions rather than less informative background in an image, thus improving model generalizability. Assuming that both source and target images are generated from a domain-invariant space by adding noise perturbations on each specific domain, Xiong et al. (2021) establish a *super target domain* via augmenting perturbations based on original target domain. The super and original target domains are fed into a mean-teacher framework, with three consistency regularization terms (w.r.t. image, instance, and class-wise) introduced for domain alignment. Chen et al. (2022) first divide the target data into clean and noisy subsets guided by a computation loss and regard them as labeled and unlabeled examples, and then utilize the mean teacher technique to self-generate pseudo-labels for the unlabeled target data for domain adaptation.

Instead of utilizing the conventional one-teacher one-student paradigm, Liu and Zhang (2021) construct a *multi-teacher multi-student* framework, where each teacher/student network is initialized using a public network pre-trained on a single dataset. Here, a graph is constructed to model the similarity among samples, and such relationships predicted by the teacher networks is used to supervise the student networks via a mean-teacher technique. Rather than leverage the mean-teacher paradigm that averages student's weights, Yu et al. (2022) propose to distill knowledge from teacher to student networks by style and structure regularizations, as well as physical prior constraints. Instead of employing a teacher–student network as the studies mentioned above, Tang, Shi, et al. (2021) achieve data-free adaptation through *gradual knowledge distillation*. Specifically, they first generate pseudo-labels via a constructed neighborhood geometry, and then use pseudo-labels obtained from the latest epoch to supervise the current training epoch for knowledge transfer.

### 2.2.2. Domain alignment via statistics

Many studies (Eastwood, Mason, Williams, & Schölkopf, 2021; Fan et al., 2022; Ishii & Sugiyama, 2021; Klingner, Termöhlen, Ritterbach, & Fingscheidt, 2022; Liu, Xing, Yang, El Fakhri & and Woo, 2021; Paul, Khurana, & Aggarwal, 2021; Zhang, Ye, Xiong, Li, & Li, 2021) leverage batch statistics stored in the pre-trained source model to approximate the distribution of inaccessible source data, and then perform cross-domain adaptation by reducing distribution discrepancy between source and target domains, as demonstrated in Fig. 8. For example, Ishii and Sugiyama (2021) approximate feature distribution of inaccessible source data by using *batch normalization statistics* (mean and variance) saved in the pre-trained source model. Then, Kullback–Leibler (KL) divergence is utilized to minimize the distributional discrepancy between source and target domains, thus achieving domain-level alignment. Inspired by Chang, You, Seo, Kwak, and Han (2019), Li, Wang, Shi, Liu, and Hou (2016), Paul et al. (2021) update the mean and variance of BatchNorm (Ioffe & Szegedy, 2015) or InstanceNorm (Ulyanov, Vedaldi, & Lempitsky, 2016) of the pre-trained

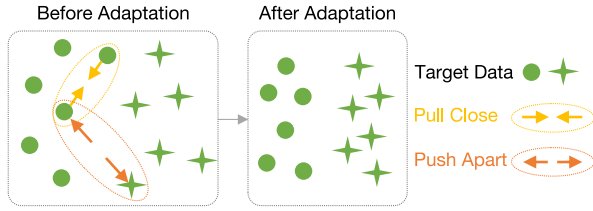


Fig. 9. Illustration of *Contrastive Learning* methods for source-free unsupervised domain adaptation. These methods exploit discriminative representations among unlabeled target data by pulling instances of similar categories closer and pushing instances of different categories away in feature space.

model based on unseen target data. Not limited to matching low-order batch-wise statistics (e.g., mean and variance), Liu, Xing, et al. (2021) additionally incorporate *high-order batch-wise statistics*, such as scale and shift parameters, to explicitly keep cross-domain consistency. Moreover, they quantify each channel's transferability based on its inter-domain divergence and assume that the channels with lower divergence contribute more to domain adaptation. Fan et al. (2022) propose to *align domain statistics adaptively* by modulating a learnable blending factor. By minimizing the total objective function, each BN layer can dynamically obtain its own optimal factor, which controls the contribution of each domain to BN statistics estimation. The methods mentioned above are all based on Gaussian-based statistics domain alignment, while Eastwood et al. (2021) attempt to align *histogram-based statistics* of the marginal feature distributions of the target domain with those stored in the pre-trained source model, thus well extending adaptation to non-Gaussian distribution scenarios.

### 2.2.3. Contrastive learning

Many contrastive learning studies (Agarwal, Paudel, Zaech, & Van Gool, 2022; Huang et al., 2021; Liu, Chen, et al., 2022; Wang, Wu, et al., 2022; Xia et al., 2021; Zhao, Stanislawski, et al., 2022) perform data-free adaptation, which helps the target model capture discriminative representations among unlabeled target data. The main idea is to *pull instances of similar categories closer* and *push instances of different categories away* in feature space, as illustrated in Fig. 9.

For instance, Xia et al. (2021) first adaptively divide target instances into source-similar and source-dissimilar sets, and then design a *class-aware contrastive module* for cross-set distribution alignment. The idea is to enforce the compactness of target instances from the same category and reduce cross-domain discrepancy, thus prompting effective knowledge transfer from the source model to target data. Wang, Wu, et al. (2022) present a cross-domain contrastive learning paradigm, which aims to minimize the distance between an anchor instance from one domain and instances from other domains that share the same category as the anchor. Due to the unavailability of source data, they utilize source prototypical representations, i.e., weight vectors in the classifier layer of a pre-trained source model, for feature alignment across two domains. Huang et al. (2021) tackle the data-free domain adaptation by taking advantage of the *historical source hypothesis*. Specifically, they propose a historical contrastive instance discrimination strategy to learn from target samples by contrasting their learned embeddings generated by the currently adapted and historical models. And they also design a historical contrastive category discrimination strategy to weight pseudo-labels of target data to learn category-discriminative target representations, by calculating the *consistency between the current and historical model predictions*. The two discrimination strategies help exploit historical source knowledge, bypassing the dependence on source data. Inspired by Sohn (2016), Agarwal et al. (2022) introduce a pair-wise contrastive objective function to reduce intra-category distance and meanwhile increase inter-category distance based on generated target pseudo-labels. They also introduce robust source and target models by taking advantage of the generated adversarial instances, which facilitates robust transfer of source knowledge to the target domain.

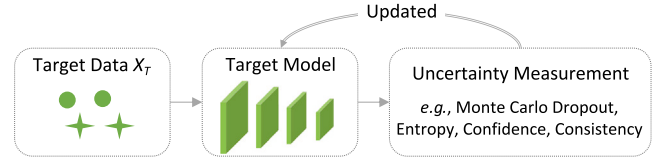


Fig. 10. Illustration of *Uncertainty-Guided Adaptation* methods for source-free unsupervised domain adaptation. These studies utilize uncertainty to guide target predictions, and such valuable information can be measured by Monte Carlo Dropout, entropy, etc.

### 2.2.4. Uncertainty-guided adaptation

Uncertainty can measure how well the target model fits the data distribution (Gawlikowski et al., 2023), and many studies (Chen, Liu, Jin, Dou, & Heng, 2021; Fleuret et al., 2021; Hegde et al., 2021; Lee & Lee, 2023; Li, Chen, et al., 2021; Pei et al., 2023; Roy et al., 2022; Xu, Lu, Wang, et al., 2022) utilize such information to guide target prediction in source-free adaptation settings (see Fig. 10).

For instance, Fleuret et al. (2021) estimate uncertainty based on differences between predicted outputs *with and without Dropout operation* (Srivastava, Hinton, Krizhevsky, Sutskever, & Salakhutdinov, 2014). By minimizing such differences, the prediction uncertainty on target data is reduced, meanwhile the learnable feature abstractor can be more robust to noise perturbations. Lee and Lee (2023) exploit aleatoric uncertainty by *encouraging intra-domain consistency* between target images and their augmented ones and enforcing *inter-domain feature distribution consistency*. Chen et al. (2021) introduce a prediction denoising approach for a cross-domain segmentation task. In this study, a key component is introducing *pixel-wise denoising via uncertainty evaluation using Monte Carlo Dropout* (Gal & Ghahramani, 2016; Kendall & Gal, 2017), which calculates the standard deviation of several stochastic outputs and keeps it under a manually-designed threshold. In this way, the noisy pseudo-labels can be filtered out, helping improve pseudo-label quality to achieve effective adaptation. Xu, Lu, Wang, et al. (2022) also propose an uncertainty-guided pseudo-labeling denoising scheme, but they use soft label correction instead of manually discarding unreliable data points. Specifically, they first identify mislabeled data points by utilizing a joint distribution matrix (Northcutt, Jiang, & Chuang, 2021; Xu et al., 2021), and then assign larger confident weights to those with higher certainty based on Monte Carlo Dropout. Combining target data and the corresponding rectified pseudo-labeling, a commonly used cross-entropy objective function can be leveraged for training the target model. Sharing the similar idea, Hegde et al. (2021) allocate lower weights for uncertain pseudo-labels, where the uncertainty is measured by prediction variance based on Monte Carlo Dropout (Gal & Ghahramani, 2016; Kendall & Gal, 2017). Considering that using Monte Carlo Dropout (Gal & Ghahramani, 2016) for uncertainty estimation requires manual hyperparameter adjustment (Gal, Hron, & Kendall, 2017), Roy et al. (2022) quantify source model's uncertainty using a *Laplace approximation* (MacKay, MacKay, et al., 2003; Tierney & Kadane, 1986). For model training, they assign smaller weights to those target samples that are farther away from source hypothesis (measured by uncertainty), avoiding misalignment of dissimilar samples. Pei et al. (2023) tackle the uncertainty issue from the perspective of improving *source model transferability*. Specifically, they estimate channel-aware transferability of the source model to target data based on an uncertainty distance, which measures the closeness between target instances and source distribution. With the aim of dynamically exploiting the source model and target data, the target model obtains the source knowledge from the transferable channels and neglects those less-transferable ones. Unlike previous studies, Li, Chen, et al. (2021) quantify uncertainty using self-entropy and propose a *self-entropy descent mechanism* to seek the optimal confidence threshold for robust pseudo-labeling of target data. They also leverage false negative mining and mosaic augmentation (Bochkovskiy, Wang, & Liao, 2020) to further eliminate the negative influence of noisy labels.

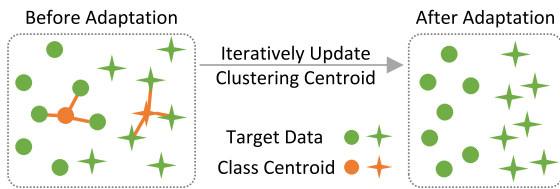


Fig. 11. Illustration of *Hidden Structure Mining* methods for source-free unsupervised domain adaptation. These methods take into consideration intrinsic feature structures of the target domain and iterate between target model refinement and clustering centroid update.

### 2.2.5. Hidden structure mining

Many studies (Lee, Jung, Yim, & Yoon, 2022; Li, Cao, & Chen, 2022; Liang, Hu, & Feng, 2020; Liang, Hu, Wang, He, & Feng, 2021; Lo, Oza, Chennupati, Galindo, & Patel, 2023; Tang, Yang, et al., 2021; Yang, van de Weijer, et al., 2021; Yang, Wang, van de Weijer, Herranz, & Jui, 2021; Yang, Wang, Wang, Jui, & van de Weijer, 2022) take into consideration intrinsic feature structures of the target domain and update the target model via clustering-aware pseudo-labeling. In Fig. 11, we illustrate the main idea of hidden structure mining methods.

For example, Yang, van de Weijer, et al. (2021) observe that target data can intrinsically form a *certain cluster structure* that can be used for domain adaptation. Specifically, they estimate affinity among target data by taking into account the neighborhood patterns captured from local-, reciprocal-, and expanded-neighbors. Source-free adaptation is achieved by encouraging consistent predictions for those with high affinity. Similarly, Yang, Wang, et al. (2021) also exploit *neighborhood structure information* of target data. They propose a local structure clustering strategy to encourage prediction consistency among  $k$ -nearest target features, thus pushing target data with semantically similar neighbors closer. Tang, Yang, et al. (2021) leverage *semantic constraints* hidden in *geometric structure* among target data to encourage robust clustering based on a cognition mechanism (Ashby, Maddox, et al., 2005). Source hypothesis transfer (SHOT) (Liang et al., 2020) and SHOT++ (Liang et al., 2021) attempt to mine the feature structure of the target domain, but they cannot fully exploit the meaningful context since the used self-supervised pseudo-labeling does not take into account each dimension's covariance in the feature space. To address this issue, Lee et al. (2022) utilize GMM in the target domain to obtain data structure and design a *joint model-data structure score* to concurrently exploit source and target knowledge. Li, Cao, and Chen (2022) utilize neighbor structure information from a new aspect by proposing a generic and model smoothness-assisted *Jacobian norm regularization* term, which is used to manipulate the consistency between each target instance and its neighbors. This Jacobian norm regularizer can be easily plugged into existing source-free domain adaptation frameworks for boosting performance.

Different from the above mentioned methods, some studies tackle source-free domain adaptation from other perspectives. Li, Du, et al. (2021) achieve data-free adaptation from an *adversarial-attack aspect*, which aims to generate adversarial target instances by adding diverse perturbations to attack the target model. Then, mutual information maximization is performed between representations extracted by the source and target model for the same target instance. The above two steps are performed alternatively, by which the domain-invariant source knowledge can be preserved and the rich target patterns can be well explored. Instead of exploring domain-invariant features for cross-domain knowledge transfer, Wang, Han, Gong and Yin (2022) mine *domain-invariant parameters* stored in the source model. They assume that only partial domain-invariant parameters of the source model contribute to domain adaptation, and their goal is to capture such parameters while penalizing the domain-specific ones. Liang, He, Sun, and Tan (2019) explore source-free adaptation from the perspective of

*minimum centroid shift*, with the aim of searching a subspace where target prototypes are mildly shifted from source prototypes. An alternating optimization scheme is leveraged for model convergence and target pseudo-label update. Inspired by maximum classifier discrepancy (Saito et al., 2018), Yang, Wang, Herranz, Jui, and van de Weijer (2023) introduce an *auxiliary bait classifier* for cross-domain feature alignment combined with the source anchor classifier. These two classifiers aim to collaboratively push uncertain target representations to the correct side of the source classifier boundary.

### 2.2.6. Challenges and insights

We classify existing fine-tuning methods for SFUDA into five subcategories. The challenges of methods in each subcategory and our insights are presented below.

- The methods in the first subcategory, i.e., *self-supervised knowledge distillation*, interpret source-free domain adaptation as a knowledge extraction and transfer process, aiming to learn domain-invariant feature representations. Most existing studies transfer source knowledge to the target model via a mean teacher strategy (Tarvainen & Valpola, 2017), where teacher weights are an exponential moving average of student weights. Hence, model parameters of both teacher and student networks are tightly coupled, which may lead to a performance bottleneck. A possible solution is to introduce a dual-student framework and let one student learn features flexibly, which may disentangle teacher-student weights to some extent (Ke, Wang, Yan, Ren, & Lau, 2019).
- The second subcategory, i.e., *domain alignment via statistics*, leverages batch statistics stored in a pre-trained source model to approximate the distribution of inaccessible source data. Compared with other categories, these statistics-based methods are lightweight and prone to generalize to other tasks, since they require only a few update steps of batch-wise statistics parameters and are potentially applicable to real-time deployment (Klingner et al., 2022). However, they are not suitable for problems that use deep network architectures without batch normalization layers. Additionally, only using the batch normalization technique for adaptation may not be sufficient, especially when the cross-domain distribution gap is large. To tackle this issue, current studies usually incorporate batch statistics with some other techniques to facilitate more effective cross-domain adaptation, such as consistent regularization (Zhang et al., 2021), confidence-filtered pseudo labeling (Paul et al., 2021), bottom-up feature restoration (Eastwood et al., 2021), and cross-domain statistics controlling (Fan et al., 2022).
- The methods in the third subcategory, i.e., *contrastive learning*, aim to bring similar-class samples closer and push dissimilar-class samples apart based on generated target pseudo-labels. Therefore, if the pseudo-labels contain much noise, these methods may suffer from substantial performance degradation. Moreover, a memory bank is usually required to store the similarity relationship between current and historical feature representations of target data, which could bring memory burdens. It is interesting to investigate the storage- and transmission-efficient contrastive learning strategies in source-free settings. In addition, several recent studies (Chen, Kornblith, Norouzi, & Hinton, 2020a; He, Fan, Wu, Xie, & Girshick, 2020) have shown that data pair construction is crucial for effective contrastive learning. One solution is utilizing contrastive information between target data and their augmented ones. Previous studies (Chen, Kornblith, Norouzi, & Hinton, 2020b) often use either strong or weak transformations for data augmentation, where strong augmentations mostly distort the structures of original images (e.g., shape distortion) while weak augmentations usually limit transformations to preserve the images' structures (e.g., flip). Here we propose to dynamically mix strong and weak augmentation of target data, which may help learn more robust representations.



- The methods in fourth subcategory, *i.e.*, *uncertainty-guided adaptation*, focus on reducing prediction uncertainty of target data. Many studies (Chen et al., 2021; Xu, Lu, Wang, et al., 2022) use Monte Carlo Dropout for uncertainty estimation, but this technique requires specialized network design and model training, bringing troublesome hyperparameter tuning (Gal et al., 2017). A recent study (Chen et al., 2021) points out that their method can only handle problems with minor domain shift, and performs poorly on problems with severe domain shift. It is interesting to explore this challenging problem in the future.
- The last subcategory, *i.e.*, *hidden structure mining*, considers intrinsic clustering structure of the target domain, assuming that geometric structure of target data may provide informative context (Tang, Yang, et al., 2021). The advantage of these methods is that no auxiliary frameworks are required, and thus, they can be easily incorporated into other adaptation frameworks. However, these methods have at least three disadvantages. (1) Most existing studies need to iterate between feature clustering and model update, which may hinder training efficiency and cause a memory burden. (2) These methods may be infeasible to handle extremely large-scale datasets due to the difficulty of saving global latent feature embeddings of the whole dataset (Chen et al., 2020c). (3) Most studies construct target geometric structures in Euclidean space, which may not be suitable for problems with non-Euclidean data such as graphs. Thus, how to improve training efficiency and deal with the large-size dataset and mining geometry information of non-Euclidean data deserve further research.

From the application perspective, computation-efficient approaches are more applicable for pixel-wise semantic segmentation tasks, which require higher resources compared with classification tasks. And those memory-intensive approaches such as contrastive learning may be not suitable for semantic segmentation. Moreover, it is worth noting that data generation methods detailed in Section 2.1 can be used in conjunction with the model fine-tuning methods described in this section. For instance, one can first generate a virtual source domain by selecting appropriate target samples, and thus a standard unsupervised domain adaptation framework could be applied. To further exploit target information, we then take account of geometric structure of target samples and generate corresponding target pseudo-labels to fine-tune the target model. These two steps can be optimized iteratively, helping generate more representative source domain to refine the target model.

### 3. Black-box source-free unsupervised domain adaptation

Different from white-box methods, *in the setting of black-box source-free domain adaptation*, both the source data  $\{X_S, Y_S\}$  and detailed parameters of the source model  $\Phi_S$  are not accessible. Only the hard or soft model predictions of the target data  $X_T$  from the source model  $\Phi_S$ , as well as  $X_T$ , are leveraged for domain adaptation. That is, we aim to learn a target model  $\Phi_T: \{\Phi_S(X_T), X_T\} \Rightarrow Y_T$ . Depending on the utilization of the black-box predictor, the existing black-box SFUDA studies can be mainly divided into three categories: **Self-Supervised Knowledge Distillation**, **Pseudo-Label Denoising**, and **Generative Distribution Alignment** methods, with detailed as follows.

#### 3.1. Self-supervised knowledge distillation

Some studies (Liang, Hu, Feng, & He, 2022; Liang, Hu, He, & Feng, 2021; Liu et al., 2022a, 2022b; Peng, Ding, Lyu, Sun, & Chen, 2022; Xu, Yang, Wu, et al., 2022) construct a *teacher-student-style network* architecture with knowledge distillation to translate the source knowledge to the target domain in a self-supervised manner. For instance, Liang et al. (2022), Liang, Hu, He, and Feng (2021) enforce output consistency between a source model (*i.e.*, teacher) and a customized target model

(*i.e.*, student) via a self-distillation loss. Specifically, a memory bank is first constructed to store the prediction of each target sample based on the black-box source model. This source model then acts as a teacher to maintain an exponential moving averaging of source and target prediction following (Kim, Ji, Yoon, & Hwang, 2021; Laine & Aila, 2016). Additionally, structural regularization on the target domain is further incorporated during adaptation for more effective knowledge distillation. Similarly, Liu et al. (2022a, 2022b) employ an *exponential mixup decay* scheme to explicitly keep prediction consistency of source and target domains, thus gradually capturing target-specific feature representations and obtaining the target pseudo-labels. Xu, Yang, Wu, et al. (2022) extend the teacher-student paradigm from image analysis to more challenging video analysis, where not only spatial features but also temporal information are taken into consideration during domain adaptation. For knowledge distillation, the target model is regarded as a student, which aims to learn similar predictions generated by a teacher (*i.e.*, source) model. The teacher model is meanwhile updated to maintain an exponential moving averaging prediction. Instead of distilling knowledge between source and target domains, Peng et al. (2022) transfer knowledge between the target network and its subnetwork in a mutual way, where the subnetwork is a slimmer version generated from the original target network following Yang et al. (2020). And target features are extracted by leveraging multi-resolution input images, helping improve the generalization ability of the target network. Moreover, a novel data augmentation strategy, called frequency MixUp, is proposed to emphasize task-related regions-of-interests while simultaneously reducing background interference.

#### 3.2. Pseudo-label denoising

Some studies (Luo et al., 2021; Zhang, Zhang, Jia, & Zhang, 2021) tackle domain shift by carefully denoising unreliable target pseudo-labels. For example, Zhang et al. (2021) combat noisy pseudo-labels via *noise rate estimation*, which first preserves more training samples at the start of the training process following Arpit et al. (2017) and then gradually filters out the noisy ones based on their loss values as training proceeds. The pseudo-labels are iteratively refined according to a category-dependent sampling strategy, encouraging the model to capture more diverse representations to improve model generalization ability. Different from Zhang et al. (2021) that only select part of reliable target data during model training, Luo et al. (2021) take into account all target data and rectify noisy pseudo-labels from a *negative learning* aspect. Specifically, their approach assigns complementary ground-truth labels for each target sample, helping alleviate error accumulation for noisy prediction. Moreover, a maximum squares objective function is utilized as confidence regularization to prevent the target model from being trapped in easy sample training. Yang, Peng, et al. (2022) incorporate *pseudo-label denoising* and *self-supervised knowledge distillation* into a unified framework. Specifically, domain knowledge is first distilled from the trained source predictor to warm up the target model by an exponential moving averaging scheme. The unlabeled target domain is then split into two subsets (*i.e.*, easy and hard groups) according to their adaptation difficulty (Arazo, Ortego, Albert, O'Connor, & McGuinness, 2019), and the MixMatch strategy (Berthelot et al., 2019) is leveraged to progressively exploit all target representations. In this way, the noise accumulation is further suppressed, thereby improving the efficacy of pseudo-label denoising.

#### 3.3. Generative distribution alignment

Different from the above methods, some studies perform distribution alignment across domains in a generative way. For instance, Yeh, Yang, Yuen, and Harada (2021) perform domain adaptation by maximizing the lower bound in variational inference. Specifically, they construct a *generation path* and an *inference path*, where the generation



path produces a prior feature distribution derived from predicted category labels, and the inference path approximates a posterior feature distribution based on each target instance. The latent distribution alignment can be achieved by maximizing the evidence lower bound in variational inference for cross-domain adaptation. Similarly, Yang, Yeh, Harada, and Yuen (2021) also build the generation and inference paths, but they achieve adaptation via minimizing the upper bound of the prediction error of target data in variational inference. Zhang and Wu (2023) achieve source-free adaptation by first building multiple source models and then generating a *virtual intermediate surrogate domain* to select target samples with minimum inconsistency predicted by the source models. The knowledge transfer is achieved by feature distribution alignment between the virtual surrogate domain and the target domain based on a maximum mean discrepancy (Zhang & Wu, 2020).

### 3.4. Challenges and insights

In this section, we classify existing black-box SFUDA methods into three categories based on how they utilize the noisy target predictions. The challenges of each category and our insights are presented below.

- The first category, *i.e.*, *self-supervised knowledge distillation*, aims to gradually transfer source knowledge to a customized target model by enforcing output consistency between a teacher (source) and a student (target) network. This learning strategy has also been used in white-box SFUDA (see Section 2.2.1). The difference is that model weights of student networks are accessible in white-box SFUDA methods, but not in black-box SFUDA. In black-box SFUDA, instead of leveraging any parameter details, the teacher network is only updated by source predictions and historical target predictions. The two items are typically weighted by a momentum factor, which helps dynamically adjust their contributions. Self-supervised knowledge distillation has shown promising performance in object recognition (Liang et al., 2022), semantic segmentation (Liu et al., 2022a), and video action recognition tasks (Xu, Yang, Wu, et al., 2022).
- The methods in the second category, *i.e.*, *pseudo-label denoising*, tackle black-box SFUDA from the perspective of noisy label rectification. It has shown that a pseudo-label denoising approach (Zhang et al., 2021) has inferior performance than the self-supervised knowledge distillation method (Liang et al., 2022). The reason may be that the former (Zhang et al., 2021) only focuses on noisy prediction itself while neglecting target data structure that is well considered in the latter (Liang et al., 2022). Considering that pseudo-label denoising methods can tackle unbalanced label noise via noise rate estimation, combining pseudo-label denoising with self-supervised knowledge distillation strategies will be a promising future direction, especially in class-imbalance scenarios. Moreover, if the black-box predictor only provides one-hot hard predictions instead of probability predictions, the utility of methods in this subcategory will be greatly reduced. The reason is that the noise rate cannot be well estimated in practice, *e.g.*, there is nearly no difference between the output of  $[0.45, 0.55]$  and that of  $[0.05, 0.95]$  because they produce the same output (*i.e.*,  $[0, 1]$ ).
- The third category, *i.e.*, *generative distribution alignment*, attempts to perform domain adaptation by minimizing feature distribution discrepancy across the source and target domains. Since source distribution is inaccessible in black-box models, some generative approaches are utilized to generate such reference distribution for target data to align with, including variational autoencoder (Yang, Yeh, et al., 2021; Yeh et al., 2021) and surrogate source domain construction (Zhang & Wu, 2023). These methods are more suitable for recognition/classification tasks, but less suitable for semantic segmentation tasks. For example, generating surrogate feature distribution of an object (*e.g.*, car) is usually easier than that of a semantic scene (*e.g.*, cityscape), since the latter contains different objects and thus the pixel-wise neighborhood relationship is difficult to model in practice.

For black-box SFUDA methods, most of them (Liu et al., 2022a, 2022b; Luo et al., 2021; Tian, Ma, et al., 2021; Wang, Zhang, Yan, Shao, & Li, 2023; Xu, Yang, Wu, et al., 2022; Yan & Guo, 2022; Yang, Yeh, et al., 2021) employ soft label prediction (*i.e.*, probability) of the unlabeled target data for adaptation, while a few studies (Wu et al., 2021; Yang, Peng, et al., 2022; Zhang et al., 2021) use hard label prediction (*i.e.*, one-hot label). Moreover, some works (Liang et al., 2022; Liang, Hu, He, & Feng, 2021; Mummadi et al., 2021; Xiao, Wei, Liu, & Lin, 2023) utilize both soft and hard labels, and they consistently find that using soft labels can produce superior adaptation performance. This finding is not surprising as employing soft labels enables the model to consider distinct label assignments in a weighted manner, which provides richer knowledge from the pre-trained source model. Besides, the use of soft labels can help alleviate dramatic oscillations during target model construction compared with using hard labels, facilitating more stable model fine-tuning (Yan & Guo, 2022). In the future, we can further explore scenarios where only hard labels are provided (*e.g.*, when source models are encapsulated as cloud services), which can enhance the practicability of SFUDA in privacy-related fields.

Besides the three strategies proposed above, it is also crucial to build a *general and robust black-box source model*, with which the target predictions tend to be more accurate. To achieve that, one possible solution is that the source data owner(s) could perform augmentation on their side (*e.g.*, adding some perturbation) before constructing the source model, which may eliminate style discrepancy between two domains, thus improving the generalizability of the source model. Another solution is using soft probability labels instead of hard one-hot labels (*e.g.*,  $[0, 1]$ ) for model training, which prevents the source model from being over-confident and helps enhance its generalizability. Compared to white-box methods, there are relatively few black-box SFUDA methods as well as benchmark datasets, which need to be further explored.

## 4. Discussion

In this section, we first compare the white-box and black-box SFUDA methods and then summarize several useful strategies to improve model generalizability. We also list datasets commonly used in the field in Table 1.

### 4.1. Comparison of white-box and black-box SFUDA

By comparing existing white-box and black-box SFUDA methods, we have the following interesting observations.

- Compared with black-box SFUDA which cannot access any source parameters, white-box SFUDA is capable of mining more source knowledge (*e.g.*, batch statistics) that facilitates more effective domain adaptation.
- White-box SFUDA methods may suffer from data privacy leakage problems (Zhang et al., 2021). For instance, Yin et al. (2020) reveal that raw data can be recovered based on source image distribution via a deep inversion technique. Using a membership inference attack strategy (Hu et al., 2022; Nasr, Shokri, & Houmansadr, 2019), it is possible to infer whether a given sample exists or not in training dataset, thereby revealing private information. The black-box SFUDA can help protect data privacy because only application programming interface (API) is accessible while detailed model weights are withheld, but it may suffer from performance degradation of cross-domain adaptation.
- Most white-box SFUDA methods assume that model architecture is shared between source and target domains, while the black-box SFUDA methods try to design task-specific target models for knowledge transfer. Such flexible model design in black-box SFUDA methods is very useful for target users with low computation resources, since they can design more efficient and lightweight target models for domain adaptation.

**Table 1**

The commonly used datasets for evaluating the performance of source-free unsupervised domain adaptation (SFUDA) approaches.

| Dataset  | Domain# | Instance# | Category# | Description  |
|--|---------|-----------|-----------|--|
| <b>Digit Recognition</b>   |         |           |           |  |
| Digits-Five (Peng et al., 2019)  | 5       | 215,695   | 10        | MNIST, SVHN, USPS, MNIST-M, Synthetic Digits           |
| <b>Semantic Segmentation</b>   |         |           |           |  |
| Segmentation datasets  | 4       | 45,766    | –         | GTA5, Cityscapes, SYNTHIA, NTHU                        |
| <b>Object Recognition</b>  |         |           |           |  |
| Office-31 (Saenko, Kulis, Fritz, & Darrell, 2010)                      | 3       | 4652      | 31        | Amazon, Webcam, DSLR                                   |
| Office-Home (Venkateswara, Eusebio, Chakraborty, & Panchanathan, 2017) | 4       | 15,500    | 65        | Artistic, Clip Art, Product, Real-World                |
| VisDA (Peng et al., 2017)  | 2       | 280,000   | 12        | Synthetic and real images                              |
| Office-Caltech-10 (Gong, Shi, Sha, & Grauman, 2012)                    | 4       | 2533      | 10        | Amazon, DSLR, Webcam, Caltech10                        |
| ImageCLEF-DA (Caputo et al., 2014)                                     | 4       | 2400      | 12        | Caltech-256, ImageNet ILSVRC2012, PASCAL VOC2012, Bing |
| PACS (Li et al., 2017)   | 4       | 9991      | 7         | Art painting, Cartoon, Photo, Sketch                   |
| DomainNet (Peng et al., 2019)  | 6       | 600,000   | 345       | Clipart, Infograph, Painting, Quickdraw, Real, Sketch  |
| MiniDomainNet (Zhou, Yang, Qiao, & Xiang, 2021)                        | 4       | 140,000   | 126       | Clipart, Painting, Real, Sketch                        |
| PointDA-10 (Qin, You, Wang, Kuo, & Fu, 2019)                           | 3       | 33,067    | 10        | ModelNet, ShapeNet, ScanNet                            |
| <b>Face Anti-Spoofing</b>  |         |           |           |  |
| Face datasets  | 4       | 7130      | –         | Replay-Attack, OULU-NPU, CASIA-MFSD, MSU-MFSD          |
| <b>LiDAR Detection</b>   |         |           |           |  |
| LiDAR datasets   | 3       | 158,510   | –         | Waymo, KITTI, nuScenes                                 |
| <b>Video Action Recognition</b>  |         |           |           |  |
| UCF-HMDB <sub>full</sub> (Chen, Kira, et al., 2019)                    | 2       | 3209      | 12        | UCF101, HMDB51   |
| Sports-DA (Xu et al., 2023)  | 3       | 40,718    | 23        | UCF10, Sports-1M, Kinetics                             |
| Daily-DA (Xu et al., 2023)   | 4       | 18,949    | 8         | ARID, HMDB51, Moments-in-Time, Kinetics                |
| <b>Traffic Sign Recognition</b>  |         |           |           |  |
| Sign datasets  | 2       | 151,839   | 43        | Syn.Signs, GTSRB                                       |
| <b>Image Classification</b>  |         |           |           |  |
| VLCS (Fang, Xu, & Rockmore, 2013)                                      | 4       | 10,729    | 5         | Caltech101, LabelMe, SUN09, VOC2007                    |
| <b>Medical Data</b>  |         |           |           |  |
| BraTS2018 (Menze et al., 2014)   | 2       | 285       | –         | Cross-disease, cross-modality (T1, T1ce, T2, FLAIR)    |
| MMWHS (Zhuang & Shen, 2016)  | 2       | 40        | –         | Cross-modality (MRI, CT)                               |
| Brain skull stripping (Li, Dan, et al., 2022)                          | 3       | 35        | –         | NFBS, ADNI, dHCP                                       |
| Polyp segmentation   | 4       | 2718      | –         | CVC-ClinicDB, Abnormal Symptoms, ETIS-Larib, EndoScene |
| MI Classification (Zhang & Wu, 2023)                                   | 4       | 528       | 2/4       | MI2-2, MI2-4, MI2015, AlexMI                           |
| Prostate segmentation  | 2       | 682       | –         | NCI-ISBI 2013 Challenge, PROMISE12 challenge           |
| Optic disc&cup segmentation  | 3       | 660       | –         | REFUGE, RIMONE-r3, Drishti-GS                          |
| Autism diagnosis   | 4       | 411       | 2         | NYU, USM, UCLA, and UM of ABIDE dataset                |

• Black-box SFUDA methods neither require data synthesis nor model fine-tuning, which helps to accelerate the convergence process of model training. In contrast, white-box methods are usually computationally intensive and time-consuming. For instance, it is reported that the computational cost of a black-box SFUDA method (Zhang & Wu, 2023) is 0.83s while that of two competing white-box methods are 3.17s (Liang et al., 2020) and 22.43s (Ahmed, Raychaudhuri, Paul, Oymak, & Roy-Chowdhury, 2021), respectively, reflecting the computation efficiency of black-box SFUDA.

In summary, when using white-box and black-box SFUDA methods, we have to make a trade-off between obtaining better performance, protecting confidential information, and reducing computational and memory costs.

#### 4.2. Useful strategies for improved generalizability

To facilitate research practice in this field, we summarize several useful techniques that could be used to improve the generalizability of learning models for source-free unsupervised domain adaptation.

##### 4.2.1. Entropy minimization loss

Most SFUDA methods utilize an entropy minimization loss (Grandvalet & Bengio, 2004) to reduce uncertainty of model predictions (Bateson, Kervadec, Dolz, Lombaert, & Ayed, 2022; Bateson, Kervadec, Dolz, Lombaert, & Ben Ayed, 2020; Fleuret et al., 2021; Hou & Zheng, 2020; Kothandaraman, Chandra, & Manocha, 2021; Kundu, Kulkarni, Singh, Jampani, & Babu, 2021; Liang et al., 2020; Liu, Li, An & Chen, 2022; Liu, Xing, et al., 2021; Liu et al., 2022a, 2022b; Wang, Shelhamer, Liu, Olshausen, & Darrell, 2020; Xu, Yang, Cao, et al., 2022). This simple yet effective strategy encourages the model to generate one-hot predictions for more confident learning.

##### 4.2.2. Diversity enforcing loss

To prevent predicted labels from collapsing to categories with larger number of samples, many studies leverage a diversity enforcing loss to encourage diverse predictions over target domain (Dong, Fang, Liu, Sun, & Liu, 2021; Huang, Yang, Zhang, & Xu, 2022; Liang et al., 2020; Liang, Hu, Feng & He, 2021; Liu, Li, An & Chen, 2022; Roy et al., 2022; Tian, Peng, & Ma, 2022; Wang, Liang, & Zhang, 2021). The usual practice is to maximize the entropy of empirical label distribution over the batch-wise average of model predictions.

##### 4.2.3. Label smoothing technique

In source-free adaptation studies, a pre-trained source model is generally obtained via training on labeled source data before adaptation stages. Currently, many studies use a label smoothing technique (Müller, Kornblith, & Hinton, 2019; Szegedy, Vanhoucke, Ioffe, Shlens, & Wojna, 2016) to produce a robust source model (Chen & Ma, 2022; Ding et al., 2023; Li, Lu, Zuo & Zhang, 2022; Liang et al., 2020; Wang, Han, Gong & Yin, 2022; Yang, van de Weijer, et al., 2021). This technique aims to transform original training labels from hard labels (e.g., 1) to soft labels (e.g., 0.95), which prevents the source model from being over-confident, helping enhance its generalization ability. Also, the experiments have shown that label smoothing can encourage closer representations of training samples from the same category (Müller et al., 2019). With a more general and robust source model, it is likely to boost adaptation performance on target domain.

##### 4.2.4. Model regularization

Many regularization terms are utilized in existing SFUDA methods by incorporating some prior knowledge. For instance, an early learning regularization (Arpit et al., 2017; Qiu et al., 2021; Xu, Wei, Zhang, & Nie, 2022) is used to prevent the model from over-fitting to label noise. A stability regularization (Li, Jiao, et al., 2020; Xiong, Ye, Zhang, Gan, & Liu, 2022; Yan, Guo, & Yang, 2021; Yang, Kuo, & Hsu, 2022)

is leveraged to prevent parameters of the target model to deviate from those of the source model. A local smoothness regularization (Li, Jiao, et al., 2020; Ma et al., 2021) is used to encourage output consistency between the target model and its noise-perturbed counterpart, helping improve robustness of the target model. A mixup regularization (Ding et al., 2023; Guan, Sun, Liu, & Zhou, 2022; Kundu, Kulkarni, et al., 2022; Liang et al., 2022; Peng et al., 2022) is used to enforce prediction consistency between original and augmented data, which can mitigate the negative influence of noisy labels.

#### 4.2.5. Confidence thresholding

Many studies leverage pseudo-labeling to train the target model in a self-supervised way. Instead of utilizing a manually-designed threshold to identify reliable/confident pseudo-labels, a commonly used strategy is automatically learning the confidence threshold for reliable pseudo-label selection (Li, Chen, Luo, He & Tan, 2022). To further tackle the class-imbalance problem, some studies (Fleuret et al., 2021; Kim, Cho, Han, Panda & Hong, 2021; Prabhu, Khare, Kartik, & Hoffman, 2021; Xu, Lu, Wang, et al., 2022; Yang, Kuo, & Hsu, 2022) propose to learn the dynamic threshold for each category, which provides a fair chance for categories with limited samples to generate pseudo-labels for self-training.

## 5. Future outlook

### 5.1. Multi-source/target domain adaptation

To utilize diverse and rich information of multiple domains, many studies (Ahmed et al., 2021; Dong, Fang, et al., 2021; Han et al., 2023; Jeon, Lee, & Kang, 2021; Li, Lu, Zuo & Zhang, 2023; Li & Wu, 2023; Li, Zhang, Wang, He, & Wu, 2023; Shen, Bu, & Wornell, 2023; Yeh, Meng, & Harada, 2023) propose multi-source data-free adaptation to transfer source knowledge to the target domain. Tian, Ma, et al. (2021) introduce a sample transport learning method, but the proposed model is shallow, and thus cannot handle highly nonlinear feature extraction. To tackle this problem, several deep learning models (Ahmed et al., 2021; Dong, Fang, et al., 2021) are proposed. However, they often ignore the fact that the generated target pseudo-labels may be noisy, which may cause training bias when matching target domains with large domain gaps. The key to solving problems with multi-source domains is quantifying the transferability of different source models and utilizing their complementary information to promote domain adaptation. Even several strategies are proposed (e.g., information-theoretic bound optimization Shen et al., 2023, Bayesian source importance estimation Han et al., 2023, weighted net confidence Li, Zhang, et al., 2023, transformer-based learning Li & Wu, 2023), more explorations are encouraged to address the problem of negative transfer during cross-domain knowledge transfer.

A few studies (Feng, You, et al., 2021; Kang et al., 2022; Peng, Huang, Zhu, & Saenko, 2019; Qin, Yang, Gao, Hu, & Shen, 2022; Song, Ma, Zhang, & Zhang, 2020) incorporate federated learning into domain adaptation scenarios. Federated learning (Bonawitz et al., 2019, 2017; Li, Sahu, Talwalkar, & Smith, 2020) is a decentralized scheme to facilitate collaborative learning among multiple distributed clients without sharing training data or model parameters. The constraint that prevents data and parameter transmission across different source domains is not required in multi-source-free domain adaptation. For instance, *federated adversarial domain adaptation (FADA)* introduced by Peng, Huang, Zhu, and Saenko (2019) is among the first attempts to propose the concept of federated domain adaptation, which employs a dynamic attention mechanism to transfer knowledge from multi-source domains to an unlabeled target domain. In this method, each source model needs to synchronize with the target domain after each training batch, resulting in huge computation costs and potential risk of privacy leakage (Zhu, Liu, & Han, 2019). To tackle this problem, Feng, You, et al. (2021)

introduces a consensus focus schema that greatly improves communication efficiency for decentralized domain adaptation. Moreover, Song et al. (2020) utilize a homomorphic encryption approach for privacy protection, and Qin et al. (2022) introduce a flexible uncertainty-aware strategy for reliable source selection. However, current federated learning studies usually produce a common model for all clients without considering heterogeneity of data distribution of different clients. Therefore, the common model cannot adapt to each client adaptively, which may affect adaptation performance. It would be very interesting to investigate personalized federated learning (Tan, Yu, Cui, & Yang, 2023), with which current or new clients can easily adapt to their own local dataset by performing a few optimization steps. Besides, all the methods mentioned above require labeled data from multiple sources to train a federated model, inevitably increasing annotation costs. Therefore, approaches that effectively exploit unlabeled data from multi-source domains in a decentralized way are urgently needed.

On the other hand, Yao et al. (2022) and Shenaj et al. (2023) propose several federated *multi-target* domain adaptation strategies for transferring knowledge of a labeled source server to multiple unlabeled target clients. More advanced techniques for federated multi-target domain adaptation are highly desirable, by considering computation and communication cost, annotation burden, and privacy protection of different target domains.

### 5.2. Test-time domain adaptation

Most SFUDA approaches require pre-collected unlabeled target data for model training, termed “training-time adaptation”. Test-time adaptation (Boudiaf, Mueller, Ben Ayed, & Bertinetto, 2022; Iwasawa & Matsuo, 2021; Karani, Erdil, Chaitanya, & Konukoglu, 2021; Wang et al., 2020; Yazdanpanah & Moradi, 2022) has been investigated by adapting the source model to the target domain during inference procedure only. The advantages of test-time adaptation are mainly twofold: (1) The adaptation process does not need iterative training, which greatly improves computational efficiency, so the model can be easily deployed in an online manner. (2) Without relying on target training data, test-time adaptation is expected to be well generalized to diverse target domains. Even current studies have made promising achievements, there are still some problems worth exploring, listed as follows. Some studies (Ma et al., 2022; Wang et al., 2020; You, Li, & Zhao, 2021) need to access batch-sized ( $>1$ ) target samples during inference, which cannot handle scenarios where target samples arrive one-by-one sequentially. Two studies (He, Carass, Zuo, Dewey, & Prince, 2020, 2021) perform image-wise adaptation rather than batch-wise adaptation, but they cannot deal with cases with large distribution shift. It is interesting to explore how to handle the scenarios where test instances come from continuous changeable domains in the future. Additionally, the solutions on how to adaptively exploit test data can be further explored (Yang, Chen, et al., 2022), such as adjusting model weights dynamically based on sample discrepancy across domains.

### 5.3. Open/partial/universal-set domain adaptation

This survey focuses on *close-set* source-free domain adaptation, where the label space of source and target domains is consistent. Practical scenarios are much more complicated when the category shift issue occurs across different domains. There are three *non-close-set* scenarios: (1) open-set ( $C_s \subset C_t$ ) problems, (2) partial-set ( $C_s \supset C_t$ ) problems, and (3) universal-set ( $C_s \setminus C_t \neq \emptyset$  and  $C_t \setminus C_s \neq \emptyset$ ,  $C_s \subset C_t$ ,  $C_s \supset C_t$ ) problems, where  $C_s$  and  $C_t$  denote the category label set for source and target domains, respectively.

Currently, only a few studies (Kundu, Venkat, Babu, et al., 2020; Kundu, Venkat, Revanur, Babu, et al., 2020; Zhao, Zhong, Luo, Lee, & Sebe, 2022) attempt to handle the category shift problem in source-free adaptation scenarios, including theory-driven bound regularization (Fang, Lu, Liu, Xuan, & Zhang, 2020), out-of-distribution data



construction (Kundu, Venkat, Babu, et al., 2020; Kundu, Venkat, Revanur, et al., 2020), neighborhood clustering learning (Saito, Kim, Sclaroff, & Saenko, 2020), uncertainty-based progressive learning (Luo, Wang, Chen, Huang, & Baktashmotlagh, 2023), and mutual information maximization (Liang, Hu, Feng & He, 2021). The main idea behind these studies is to recognize out-of-source distribution samples and improve generalization ability of the source model. However, the model performance of existing studies is not quite satisfactory due to the inaccessibility of valuable category-gap knowledge. One possible solution is to adaptively learn a threshold instead of using a fixed one to determine the acceptance/rejection of each target sample as a “known” category via some similarity measurement. Moreover, some strategies used in non-source-free domain adaptation can also be borrowed, such as distribution weighted combining rule (Xu, Chen, Zuo, Yan, & Lin, 2018), category-invariant representation learning (Lekhtman, Ziser, & Reichart, 2021), one-vs-all learning scheme (Saito & Saenko, 2021), and global-local optimization (Feng, Chen, He, Pan, & Zhou, 2021).

#### 5.4. Flexible target model design

For black-box SFUDA methods, due to unavailability of structure and parameters of target model, one usually has to manually design a target model. For instance, Liu et al. (2022a) choose a U-Net based framework as the target model for segmentation. However, such manually designed architectures may be not suitable when adapting to the target domain. It is expected that the automatic design of target models, e.g., using neural architecture search (NAS) (Elsken, Metzen, & Hutter, 2019; Lu et al., 2021; Wistuba, Rawat, & Pedapati, 2019), helps improve the learning performance. Considering that NAS has recently become a popular strategy for searching proper network architectures in deep learning, we can integrate it into SFUDA scenarios to find more proper and efficient target models. And how to balance the search space and search cost of network parameters can be further investigated. Moreover, hyperparameters used in NAS (e.g., optimizer strategy, weight decay regularization) should be carefully considered since they also have a significant impact on network performance (Ren et al., 2021).

#### 5.5. Cross-modality domain adaptation

Existing studies mainly focus on one single modality for domain adaptation, while a few studies perform cross-modality adaptation in source-free settings (Ahmed, Lohit, Peng, Jones, & Roy-Chowdhury, 2022; Hong et al., 2022). For instance, for medical data analysis, the acquisition expense of computed tomography (CT) scans is generally less than that of magnetic resonance imaging (MRI) scans, hence it may greatly reduce annotation cost for a segmentation task when transferring a source model trained on CT images to MRI scans (Hong et al., 2022). Moreover, in computer vision field, it would be promising to investigate cross-modality adaptation in the future, e.g., image→video, which aims to achieve video recognition based on the source model trained on the image dataset. Also, how to effectively integrate multi-modality (e.g., image, sound, text, and video) data for domain adaptation in a source-free way is an interesting but not yet widely studied problem.

#### 5.6. Continual/lifelong domain adaptation

Most current studies focus on improving adaptation performance on the target domain while neglecting the performance on source domain, running the risk of catastrophic forgetting problems (Kemker, McClure, Abitino, Hayes, & Kanan, 2018). To address this issue, several solutions have been developed from different aspects, such as domain expansion (Zhang, Li, Tang, Ogunbona, et al., 2020), historical contrastive learning (Huang et al., 2021), domain attention regularization (Yang, Wang, et al., 2021), and model perturbation (Jing, Zhen, Li, & Snoek,

2022), while there is still massive room for performance improvement. Inspired by continual/lifelong learning (Lange et al., 2021; Kirkpatrick et al., 2017; Li & Hoiem, 2017; Lopez-Paz & Ranzato, 2017), continual domain adaptation has recently made great progress by investigating gradient regularization (Tang, Su, Chen & Ouyang, 2021), iterative neuron restoration (Wang, Fink, Van Gool & Dai, 2022), buffer sample mixture (Taufique, Jahan, & Savakis, 2021), etc. The continual domain adaptation in source-free settings for mitigation of catastrophic forgetting remains an underdeveloped topic that can be further explored in the future.

#### 5.7. Semi-supervised domain adaptation

Source-free domain adaptation in semi-supervised settings (i.e., with a few labeled target data involved for model training) has also been explored in recent years (Chidlovskii, Clinchant, & Csurka, 2016; Kothandaraman et al., 2021; Nelakurthi, Maciejewski, & He, 2018). It usually performs semi-supervised adaptation with the help of active learning (Kothandaraman et al., 2022; Wang, Han, Zhang, & Yin, 2022), model memorization revelation (Yang, Ma, & Yuen, 2022), and consistency and diversity learning (Wang, Zhuo, Cui, & Wang, 2021). There is still a lot of space for improvement with a limited number of labeled target samples, e.g., by fine-tuning the current source-free adaptation frameworks, but this is not the focus of this survey.

#### 5.8. Large language and language-vision model

Several recent studies on SFUDA (Hu et al., 2024; Tanwisuth, Zhang, Zheng, He, & Zhou, 2023; Zara et al., 2023) have leveraged large language and language-vision models to improve target adaptation, which usually have millions or billions of parameters. Specifically, Zara et al. (2023) propose to first initialize the source model with a popular large model called Contrastive Language-Image Pretraining (CLIP) (Radford et al., 2021), and then introduce two domain-specific adapters to distill source knowledge to the target domain, requiring only a small set of learnable parameters. Hu et al. (2024) propose to exclusively fine-tune the layer-normalization (Ba, Kiros, & Hinton, 2016) weights of the pre-trained source model while keeping other weights frozen. Moreover, Tanwisuth et al. (2023) perform source-free adaptation by fine-tuning the large source model based on a parameter-efficient prompt tuning strategy (Lester, Al-Rfou, & Constant, 2021). This strategy introduces a few numbers of tunable parameters within input embedding space, eliminating the necessity to fine-tune the entire source model.

It is noted that large model based SFUDA still lies in its early stage, and its potential for the future appears promising. Given that the large models (trained with extensive datasets) can capture comprehensive and versatile feature representations from the source domain, they are expected to show great improvements in target inference. It is interesting to explore effective methods to distill valuable knowledge from large source models to unseen target domains in the future, and techniques like spot-adaptive knowledge distillation (Song, Chen, Ye, & Song, 2022), self-paced learning (Li, Zhou, et al., 2023), and soft label dynamic tuning (Rao et al., 2023), may present promising directions. Furthermore, considering existing studies mainly focus on the language field, future works can integrate SFUDA with other large foundation models in the fields of computer vision, medicine, and genomics.

#### 5.9. Generative data augmentation for domain adaptation

In source-free domain adaptation, many studies (Kundu, Kulkarni, et al., 2022; Liang et al., 2022) employ data augmentation to generate synthetic data that resemble real data in source and target domains. These studies mainly use the MixUp strategy (Zhang et al., 2017) for augmentation, which may cause limited diversity in synthetic data as

they rely only on existing data. Given that generative artificial intelligence can potentially generate a wider range of diverse samples, we can incorporate more advanced generative data augmentation techniques (e.g., differential privacy Faisal, Mohammed, Leung, & Wang, 2022; Sun, van Soest, & Dumontier, 2023, diffusion model Ho, Jain, & Abbeel, 2020) in SFUDA to further improve target inference.

Specifically, we can generate more diverse target samples based on unlabeled target data, which may produce a more general and robust feature encoder, helping improve the generalizability of the target model. Moreover, we can create more data for minority classes, which helps mitigate data imbalance issues and prevent overfitting during target model training. In the future, there is potential for integrating advanced generative AI techniques within more source-free scenarios.

## 6. Conclusion

In this paper, we provide a comprehensive review of recent progress in source-free unsupervised domain adaptation (SFUDA). We classify existing SFUDA studies into white-box and black-box groups, and each group is further categorized based on different learning strategies. The challenges of methods in each category and our insights are provided. We then compare white-box and black-box SFUDA methods, discuss effective techniques for improving adaptation performance, and summarize commonly used datasets. We finally discuss promising future research directions. It is worth noting that the research topic of source-free unsupervised domain adaptation is still in its early stages, and we hope this survey can spark new ideas and attract more researchers to advance this high-impact research field.

## CRedit authorship contribution statement

**Yuqi Fang:** Methodology, Software, Validation, Visualization, Writing – original draft. **Pew-Thian Yap:** Writing – review & editing. **Weili Lin:** Writing – review & editing. **Hongtu Zhu:** Writing – review & editing. **Mingxia Liu:** Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Writing – review & editing.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

No data was used for the research described in the article.

## Acknowledgment

This work was supported in part by National Institutes of Health (NIH), United States grants RF1AG073297, R01EB035160, and RF1AG082938.

## References

- Abusitta, A., Wahab, O. A., & Fung, B. C. (2021). VirtualGAN: Reducing mode collapse in generative adversarial networks using virtual mapping. In *2021 International Joint Conference on Neural Networks* (pp. 1–6). IEEE.
- Agarwal, P., Paudel, D. P., Zaech, J. N., & Van Gool, L. (2022). Unsupervised robust domain adaptation without source data. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision* (pp. 2009–2018).
- Ahmed, S. M., Lohit, S., Peng, K. C., Jones, M. J., & Roy-Chowdhury, A. K. (2022). Cross-modal knowledge transfer without task-relevant source data. In *European Conference on Computer Vision* (pp. 111–127). Springer.
- Ahmed, S. M., Raychaudhuri, D. S., Paul, S., Oymak, S., & Roy-Chowdhury, A. K. (2021). Unsupervised multi-source domain adaptation without access to source data. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 10103–10112).
- Arazo, E., Ortego, D., Albert, P., O'Connor, N., & McGuinness, K. (2019). Unsupervised label noise modeling and loss correction. In *International Conference on Machine Learning* (pp. 312–321). PMLR.
- Arpit, D., Jastrzebski, S., Ballas, N., Krueger, D., Bengio, E., Kanwal, M. S., et al. (2017). A closer look at memorization in deep networks. In *International Conference on Machine Learning* (pp. 233–242). PMLR.
- Ashby, F. G., Maddox, W. T., et al. (2005). Human category learning. *Annual Review of Psychology*, 56(1), 149–178.
- Ba, J. L., Kiros, J. R., & Hinton, G. E. (2016). Layer normalization. arXiv preprint arXiv:1607.06450.
- Bang, D., & Shim, H. (2021). MGGAN: Solving mode collapse using manifold-guided training. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 2347–2356).
- Bateson, M., Kervadec, H., Dolz, J., Lombaert, H., & Ayed, I. B. (2022). Source-free domain adaptation for image segmentation. *Medical Image Analysis*, 82, Article 102617.
- Bateson, M., Kervadec, H., Dolz, J., Lombaert, H., & Ben Ayed, I. (2020). Source-relaxed domain adaptation for image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 490–499). Springer.
- Berthelot, D., Carlini, N., Goodfellow, I., Papernot, N., Oliver, A., & Raffel, C. A. (2019). MixMatch: A holistic approach to semi-supervised learning. *Advances in Neural Information Processing Systems*, 32.
- Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M. (2020). YOLOv4: Optimal speed and accuracy of object detection. arXiv preprint arXiv:2004.10934.
- Bonawitz, K., Eichner, H., Grieskamp, W., Huba, D., Ingerman, A., Ivanov, V., et al. (2019). Towards federated learning at scale: System design. *Proceedings of Machine Learning and Systems*, 1, 374–388.
- Bonawitz, K., Ivanov, V., Kreuter, B., Marcedone, A., McMahan, H. B., Patel, S., et al. (2017). Practical secure aggregation for privacy-preserving machine learning. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security* (pp. 1175–1191).
- Boudiaf, M., Mueller, R., Ben Ayed, I., & Bertinetto, L. (2022). Parameter-free online test-time adaptation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 8344–8353).
- Boyd, S., Parikh, N., Chu, E., Peleato, B., Eckstein, J., et al. (2011). Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends® in Machine Learning*, 3(1), 1–122.
- Caputo, B., Müller, H., Martinez-Gomez, J., Villegas, M., Acar, B., Patricia, N., et al. (2014). ImageCLEF 2014: Overview and analysis of the results. In *International Conference of the Cross-Language Evaluation Forum for European Languages* (pp. 192–211). Springer.
- Chang, W. G., You, T., Seo, S., Kwak, S., & Han, B. (2019). Domain-specific batch normalization for unsupervised domain adaptation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 7354–7362).
- Chen, M. H., Kira, Z., AlRegib, G., Yoo, J., Chen, R., & Zheng, J. (2019). Temporal attentive alignment for large-scale video domain adaptation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 6321–6330).
- Chen, T., Kornblith, S., Norouzi, M., & Hinton, G. (2020a). A simple framework for contrastive learning of visual representations. In *International Conference on Machine Learning* (pp. 1597–1607). PMLR.
- Chen, T., Kornblith, S., Norouzi, M., & Hinton, G. (2020b). A simple framework for contrastive learning of visual representations. In *International Conference on Machine Learning* (pp. 1597–1607). PMLR.
- Chen, W., Lin, L., Yang, S., Xie, D., Pu, S., & Zhuang, Y. (2022). Self-supervised noisy label learning for source-free unsupervised domain adaptation. In *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 10185–10192). IEEE.
- Chen, C., Liu, Q., Jin, Y., Dou, Q., & Heng, P. A. (2021). Source-free domain adaptive fundus image segmentation with denoised pseudo-labeling. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 225–235). Springer.
- Chen, W.-Y., Liu, Y.-C., Kira, Z., Wang, Y.-C. F., & Huang, J.-B. (2019). A closer look at few-shot classification. arXiv preprint arXiv:1904.04232.
- Chen, P., & Ma, A. J. (2022). Source-free temporal attentive domain adaptation for video action recognition. In *Proceedings of the 2022 International Conference on Multimedia Retrieval* (pp. 489–497).
- Chen, W., Pu, S., Xie, D., Yang, S., Guo, Y., & Lin, L. (2020c). Unsupervised image classification for deep representation learning. In *European Conference on Computer Vision* (pp. 430–446). Springer.
- Chidlovskii, B., Clinchant, S., & Csurka, G. (2016). Domain adaptation in the absence of source domain data. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 451–460).
- De Lange, M., Aljundi, R., Masana, M., Parisot, S., Jia, X., Leonardis, A., et al. (2021). A continual learning survey: Defying forgetting in classification tasks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(7), 3366–3385.
- Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009). ImageNet: A large-scale hierarchical image database. In *2009 IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 248–255). IEEE.

- Ding, Y., Sheng, L., Liang, J., Zheng, A., & He, R. (2023). ProxyMix: Proxy-based mixup training with label refinery for source-free domain adaptation. *Neural Networks*, 167, 92–103.
- Ding, N., Xu, Y., Tang, Y., Xu, C., Wang, Y., & Tao, D. (2022). Source-free domain adaptation via distribution estimation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 7212–7222).
- Dong, J., Cong, Y., Sun, G., Fang, Z., & Ding, Z. (2021). Where and how to transfer: Knowledge aggregation-induced transferability perception for unsupervised domain adaptation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 46(3), 1664–1681.
- Dong, J., Fang, Z., Liu, A., Sun, G., & Liu, T. (2021). Confident anchor-induced multi-source free domain adaptation. *Advances in Neural Information Processing Systems*, 34, 2848–2860.
- Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., et al. (2020). An image is worth 16×16 words: Transformers for image recognition at scale. arXiv preprint arXiv:2010.11929.
- Du, Y., Yang, H., Chen, M., Jiang, J., Luo, H., & Wang, C. (2021). Generation, augmentation, and alignment: A pseudo-source domain based method for source-free domain adaptation. arXiv preprint arXiv:2109.04015.
- Eastwood, C., Mason, I., Williams, C. K., & Schölkopf, B. (2021). Source-free adaptation to measurement shift via bottom-up feature restoration. arXiv preprint arXiv: 2107.05446.
- Elsken, T., Metzen, J. H., & Hutter, F. (2019). Neural architecture search: A survey. *Journal of Machine Learning Research*, 20(1), 1997–2017.
- Faisal, F., Mohammed, N., Leung, C. K., & Wang, Y. (2022). Generating privacy preserving synthetic medical data. In *2022 IEEE 9th International Conference on Data Science and Advanced Analytics* (pp. 1–10). IEEE.
- Fan, J., Zhu, H., Jiang, X., Meng, L., Chen, C., Fu, C., et al. (2022). Unsupervised domain adaptation by statistics alignment for deep sleep staging networks. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 30, 205–216.
- Fang, Z., Lu, J., Liu, F., Xuan, J., & Zhang, G. (2020). Open set domain adaptation: Theoretical bound and algorithm. *IEEE Transactions on Neural Networks and Learning Systems*, 32(10), 4309–4322.
- Fang, Y., Wang, M., Potter, G. G., & Liu, M. (2022). Unsupervised cross-domain functional MRI adaptation for automated major depressive disorder identification. *Medical Image Analysis*, Article 102707.
- Fang, C., Xu, Y., & Rockmore, D. N. (2013). Unbiased metric learning: On the utilization of multiple datasets and web images for softening bias. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 1657–1664).
- Feng, Y., Chen, J., He, S., Pan, T., & Zhou, Z. (2021). Globally localized multisource domain adaptation for cross-domain fault diagnosis with category shift. *IEEE Transactions on Neural Networks and Learning Systems*, 34(6), 3082–3096.
- Feng, H., You, Z., Chen, M., Zhang, T., Zhu, M., Wu, F., et al. (2021). KD3A: Unsupervised multi-source decentralized domain adaptation via knowledge distillation. In *International Conference on Machine Learning* (pp. 3274–3283).
- Fleuret, F., et al. (2021). Uncertainty reduction for model adaptation in semantic segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 9613–9623).
- Gal, Y., & Ghahramani, Z. (2016). Dropout as a bayesian approximation: Representing model uncertainty in deep learning. In *International Conference on Machine Learning* (pp. 1050–1059). PMLR.
- Gal, Y., Hron, J., & Kendall, A. (2017). Concrete dropout. *Advances in Neural Information Processing Systems*, 30.
- Ganin, Y., & Lempitsky, V. (2015). Unsupervised domain adaptation by back-propagation. In *International Conference on Machine Learning* (pp. 1180–1189). PMLR.
- Gawlikowski, J., Tassi, C. R. N., Ali, M., Lee, J., Humt, M., Feng, J., et al. (2023). A survey of uncertainty in deep neural networks. *Artificial Intelligence Review*, 1–77.
- Gong, B., Shi, Y., Sha, F., & Grauman, K. (2012). Geodesic flow kernel for unsupervised domain adaptation. In *2012 IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 2066–2073). IEEE.
- Grandvalet, Y., & Bengio, Y. (2004). Semi-supervised learning by entropy minimization. *Advances in Neural Information Processing Systems*, 17.
- Guan, H., & Liu, M. (2021). Domain adaptation for medical image analysis: A survey. *IEEE Transactions on Biomedical Engineering*, 69(3), 1173–1185.
- Guan, X., Sun, H., Liu, N., & Zhou, H. (2022). Polycentric clustering and structural regularization for source-free unsupervised domain adaptation. arXiv preprint arXiv: 2210.07463.
- Han, Z., Zhang, Z., Wang, F., He, R., Su, W., Xi, X., et al. (2023). Discriminability and transferability estimation: A Bayesian source importance estimation approach for multi-source-free domain adaptation. In *Proceedings of the AAAI Conference on Artificial Intelligence*: vol. 37, (no. 6), (pp. 7811–7820).
- Hassaballah, M., & Awad, A. I. (2020). *Deep Learning in Computer Vision: Principles and Applications*. CRC Press.
- He, Y., Carass, A., Zuo, L., Dewey, B. E., & Prince, J. L. (2020). Self domain adapted network. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 437–446). Springer.
- He, Y., Carass, A., Zuo, L., Dewey, B. E., & Prince, J. L. (2021). Autoencoder based self-supervised test-time adaptation for medical image analysis. *Medical Image Analysis*, 72, Article 102136.
- He, K., Fan, H., Wu, Y., Xie, S., & Girshick, R. (2020). Momentum contrast for unsupervised visual representation learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 9729–9738).
- Hegde, D., Sindagi, V., Kilic, V., Cooper, A. B., Foster, M., & Patel, V. (2021). Uncertainty-aware mean teacher for source-free unsupervised domain adaptive 3D object detection. arXiv preprint arXiv:2109.14651.
- Ho, J., Jain, A., & Abbeel, P. (2020). Denoising diffusion probabilistic models. *Advances in Neural Information Processing Systems*, 33, 6840–6851.
- Hong, J., Zhang, Y. D., & Chen, W. (2022). Source-free unsupervised domain adaptation for cross-modality abdominal multi-organ segmentation. *Knowledge-Based Systems*, Article 109155.
- Hou, Y., & Zheng, L. (2020). Source free domain adaptation with image translation. arXiv preprint arXiv:2008.07514.
- Hou, Y., & Zheng, L. (2021). Visualizing adapted knowledge in domain transfer. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 13824–13833).
- Hu, H., Salic, Z., Sun, L., Dobbie, G., Yu, P. S., & Zhang, X. (2022). Membership inference attacks on machine learning: A survey. *ACM Computing Surveys*, 54(11s), 1–37.
- Hu, X., Zhang, K., Xia, L., Chen, A., Luo, J., Sun, Y., et al. (2024). ReCLIP: Refine contrastive language image pre-training with source free domain adaptation. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision* (pp. 2994–3003).
- Huang, J., Guan, D., Xiao, A., & Lu, S. (2021). Model adaptation: Historical contrastive learning for unsupervised domain adaptation without source data. *Advances in Neural Information Processing Systems*, 34, 3635–3649.
- Huang, Y., Yang, X., Zhang, J., & Xu, C. (2022). Relative alignment network for source-free multimodal video domain adaptation. In *Proceedings of the 30th ACM International Conference on Multimedia* (pp. 1652–1660).
- Ioffe, S., & Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *International Conference on Machine Learning* (pp. 448–456). PMLR.
- Ishii, M., & Sugiyama, M. (2021). Source-free domain adaptation via distributional alignment by matching batch normalization statistics. arXiv preprint arXiv:2101.10842.
- Iwasawa, Y., & Matsuo, Y. (2021). Test-time classifier adjustment module for model-agnostic domain generalization. *Advances in Neural Information Processing Systems*, 34, 2427–2440.
- Jeon, H., Lee, S., & Kang, U. (2021). Unsupervised multi-source domain adaptation with no observable source data. *PLoS One*, 16(7), Article e0253415.
- Jing, M., Zhen, X., Li, J., & Snoek, C. (2022). Variational model perturbation for source-free domain adaptation. *Advances in Neural Information Processing Systems*, 35, 17173–17187.
- Kang, Y., He, Y., Luo, J., Fan, T., Liu, Y., & Yang, Q. (2022). Privacy-preserving federated adversarial domain adaptation over feature groups for interpretability. *IEEE Transactions on Big Data*.
- Kang, G., Jiang, L., Yang, Y., & Hauptmann, A. G. (2019). Contrastive adaptation network for unsupervised domain adaptation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 4893–4902).
- Karani, N., Erdil, E., Chaitanya, K., & Konukoglu, E. (2021). Test-time adaptable neural networks for robust medical image segmentation. *Medical Image Analysis*, 68, Article 101907.
- Karim, N., Mithun, N. C., Rajvanshi, A., Chiu, H. p., Samarasekera, S., & Rahnavard, N. (2023). C-SFDA: A curriculum learning aided self-training framework for efficient source free domain adaptation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 24120–24131).
- Ke, Z., Wang, D., Yan, Q., Ren, J., & Lau, R. W. (2019). Dual student: Breaking the limits of the teacher in semi-supervised learning. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 6728–6736).
- Kemker, R., McClure, M., Abitino, A., Hayes, T., & Kanan, C. (2018). Measuring catastrophic forgetting in neural networks. In *Proceedings of the AAAI Conference on Artificial Intelligence*: vol. 32, (no. 1).
- Kendall, A., & Gal, Y. (2017). What uncertainties do we need in bayesian deep learning for computer vision? *Advances in Neural Information Processing Systems*, 30.
- Kim, Y., Cho, D., Han, K., Panda, P., & Hong, S. (2021). Domain adaptation without source data. *IEEE Transactions on Artificial Intelligence*, 2(6), 508–518.
- Kim, K., Ji, B., Yoon, D., & Hwang, S. (2021). Self-knowledge distillation with progressive refinement of targets. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 6567–6576).
- Kirkpatrick, J., Pascanu, R., Rabinowitz, N., Veness, J., Desjardins, G., Rusu, A. A., et al. (2017). Overcoming catastrophic forgetting in neural networks. *Proceedings of the National Academy of Sciences*, 114(13), 3521–3526.
- Klingner, M., Termöhlen, J.-A., Ritterbach, J., & Fingscheidt, T. (2022). Unsupervised batchnorm adaptation (UBNA): A domain adaptation method for semantic segmentation without using source domain representations. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision* (pp. 210–220).



- Kothandaraman, D., Chandra, R., & Manocha, D. (2021). SS-SFDA: Self-supervised source-free domain adaptation for road segmentation in hazardous environments. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 3049–3059).
- Kothandaraman, D., Shekhar, S., Sancheti, A., Ghuhan, M., Shukla, T., & Manocha, D. (2022). DistillAdapt: Source-free active visual domain adaptation. arXiv preprint arXiv:2205.12840.
- Kumar, V., Lal, R., Patil, H., & Chakraborty, A. (2023). Conmix for source-free single and multi-target domain adaptation. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision* (pp. 4178–4188).
- Kundu, J. N., Bhambri, S., Kulkarni, A., Sarkar, H., Jampani, V., & Babu, R. V. (2022). Concurrent subsidiary supervision for unsupervised source-free domain adaptation. In *European Conference on Computer Vision* (pp. 177–194). Springer.
- Kundu, J. N., Kulkarni, A. R., Bhambri, S., Mehta, D., Kulkarni, S. A., Jampani, V., et al. (2022). Balancing discriminability and transferability for source-free domain adaptation. In *International Conference on Machine Learning* (pp. 11710–11728). PMLR.
- Kundu, J. N., Kulkarni, A., Singh, A., Jampani, V., & Babu, R. V. (2021). Generalize then adapt: Source-free domain adaptive semantic segmentation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 7046–7056).
- Kundu, J. N., Venkat, N., Babu, R. V., et al. (2020). Universal source-free domain adaptation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 4544–4553).
- Kundu, J. N., Venkat, N., Revanur, A., Babu, R. V., et al. (2020). Towards inheritable models for open-set domain adaptation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 12376–12385).
- Kurmi, V. K., Subramanian, V. K., & Nambodiri, V. P. (2021). Domain impression: A source data free domain adaptation method. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision* (pp. 615–625).
- Laine, S., & Aila, T. (2016). Temporal ensembling for semi-supervised learning. arXiv preprint arXiv:1610.02242.
- Lee, C. Y., Batra, T., Baig, M. H., & Ulbricht, D. (2019). Sliced wasserstein discrepancy for unsupervised domain adaptation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 10285–10295).
- Lee, J., Jung, D., Yim, J., & Yoon, S. (2022). Confidence score for source-free unsupervised domain adaptation. In *International Conference on Machine Learning* (pp. 12365–12377). PMLR.
- Lee, J., & Lee, G. (2023). Feature alignment by uncertainty and self-training for source-free unsupervised domain adaptation. *Neural Networks*, 161, 682–692.
- Lekhtman, E., Ziser, Y., & Reichart, R. (2021). DILBERT: Customized pre-training for domain adaptation with category shift, with an application to aspect extraction. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing* (pp. 219–230).
- Lester, B., Al-Rfou, R., & Constant, N. (2021). The power of scale for parameter-efficient prompt tuning. arXiv preprint arXiv:2104.08691.
- Li, W., Cao, M., & Chen, S. (2022). Jacobian norm for unsupervised source-free domain adaptation. arXiv preprint arXiv:2204.03467.
- Li, C., Chen, W., Luo, X., He, Y., & Tan, Y. (2022). Adaptive pseudo labeling for source-free domain adaptation in medical image segmentation. In *ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing* (pp. 1091–1095). IEEE.
- Li, X., Chen, W., Xie, D., Yang, S., Yuan, P., Pu, S., et al. (2021). A free lunch for unsupervised domain adaptive object detection without source data. In *Proceedings of the AAAI Conference on Artificial Intelligence*: vol. 35, (no. 10), (pp. 8474–8481).
- Li, Y., Dan, R., Wang, S., Cao, Y., Luo, X., Tan, C., et al. (2022). Source-free domain adaptation for multi-site and lifespan brain skull stripping. arXiv preprint arXiv:2203.04299.
- Li, J., Du, Z., Zhu, L., Ding, Z., Lu, K., & Shen, H. T. (2021). Divergence-agnostic unsupervised domain adaptation by adversarial attacks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(11), 8196–8211.
- Li, Z., & Hoiem, D. (2017). Learning without forgetting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(12), 2935–2947.
- Li, R., Jiao, Q., Cao, W., Wong, H.-S., & Wu, S. (2020). Model adaptation: Unsupervised domain adaptation without source data. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 9641–9650).
- Li, K., Lu, J., Zuo, H., & Zhang, G. (2022). Source-free multi-domain adaptation with generally auxiliary model training. In *2022 International Joint Conference on Neural Networks* (pp. 1–8). IEEE.
- Li, K., Lu, J., Zuo, H., & Zhang, G. (2023). Source-free multi-domain adaptation with fuzzy rule-based deep neural networks. *IEEE Transactions on Fuzzy Systems*.
- Li, T., Sahu, A. K., Talwalkar, A., & Smith, V. (2020). Federated learning: Challenges, methods, and future directions. *IEEE Signal Processing Magazine*, 37(3), 50–60.
- Li, Y., Wang, N., Shi, J., Liu, J., & Hou, X. (2016). Revisiting batch normalization for practical domain adaptation. arXiv preprint arXiv:1603.04779.
- Li, G., & Wu, C. (2023). Transformer-based multi-source domain adaptation without source data. In *2023 International Joint Conference on Neural Networks* (pp. 1–8). IEEE.
- Li, D., Yang, Y., Song, Y. Z., & Hospedales, T. M. (2017). Deeper, broader and artier domain generalization. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 5542–5550).
- Li, G., Zhang, Q., Wang, P., He, R., & Wu, C. (2023). Target-discriminability-induced multi-source-free domain adaptation. In *2023 IEEE International Conference on Image Processing* (pp. 76–80). IEEE.
- Li, J., Zhou, S., Li, L., Wang, H., Bu, J., & Yu, Z. (2023). Dynamic data-free knowledge distillation by easy-to-hard learning strategy. *Information Sciences*, 642, Article 119202.
- Liang, J., He, R., Sun, Z., & Tan, T. (2019). Distant supervised centroid shift: A simple and efficient approach to visual domain adaptation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 2975–2984).
- Liang, J., Hu, D., & Feng, J. (2020). Do we really need to access the source data? Source hypothesis transfer for unsupervised domain adaptation. In *International Conference on Machine Learning* (pp. 6028–6039). PMLR.
- Liang, J., Hu, D., Feng, J., & He, R. (2021). UMAD: Universal model adaptation under domain and category shift. arXiv preprint arXiv:2112.08553.
- Liang, J., Hu, D., Feng, J., & He, R. (2022). DINE: Domain adaptation from single and multiple black-box predictors. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 8003–8013).
- Liang, J., Hu, D., He, R., & Feng, J. (2021). Distill and fine-tune: Effective adaptation from a black-box source model. arXiv preprint arXiv:2104.01539.
- Liang, J., Hu, D., Wang, Y., He, R., & Feng, J. (2021). Source data-absent unsupervised domain adaptation through hypothesis transfer and labeling transfer. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.
- Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., et al. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42, 60–88.
- Liu, Y., Chen, Y., Dai, W., Gou, M., Huang, C. T., & Xiong, H. (2022). Source-free domain adaptation with contrastive domain alignment and self-supervised exploration for face anti-spoofing. In *European Conference on Computer Vision* (pp. 511–528). Springer.
- Liu, J., Li, X., An, S., & Chen, Z. (2022). Source-free unsupervised domain adaptation for blind image quality assessment. arXiv preprint arXiv:2207.08124.
- Liu, X., Xing, F., Yang, C., El Fakhri, G., & Woo, J. (2021). Adapting off-the-shelf source segmenter for target medical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 549–559). Springer.
- Liu, X., Yoo, C., Xing, F., Kuo, C. C. J., El Fakhri, G., Kang, J. W., et al. (2022a). Unsupervised black-box model domain adaptation for brain tumor segmentation. *Frontiers in Neuroscience*, 341.
- Liu, X., Yoo, C., Xing, F., Kuo, C. C. J., El Fakhri, G., Kang, J. W., et al. (2022b). Unsupervised domain adaptation for segmentation with black-box source model. In *Medical Imaging 2022: Image Processing*: vol. 12032, (pp. 255–260). SPIE.
- Liu, X., & Yuan, Y. (2022). A source-free domain adaptive polyp detection framework with style diversification flow. *IEEE Transactions on Medical Imaging*, 41(7), 1897–1908.
- Liu, X., & Zhang, S. (2021). Graph consistency based mean-teaching for unsupervised domain adaptive person re-identification. arXiv preprint arXiv:2105.04776.
- Liu, Y., Zhang, W., & Wang, J. (2021). Source-free domain adaptation for semantic segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 1215–1224).
- Liu, Y., Zhang, W., Wang, J., & Wang, J. (2021). Data-free knowledge transfer: A survey. arXiv preprint arXiv:2112.15278.
- Lo, S. Y., Oza, P., Chennupati, S., Galindo, A., & Patel, V. M. (2023). Spatio-temporal pixel-level contrastive learning-based source-free domain adaptation for video semantic segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 10534–10543).
- Lopez-Paz, D., & Ranzato, M. (2017). Gradient episodic memory for continual learning. *Advances in Neural Information Processing Systems*, 30.
- Lu, Z., Sreekumar, G., Goodman, E., Banzhaf, W., Deb, K., & Boddeti, V. N. (2021). Neural architecture transfer. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(9), 2971–2989.
- Luo, X., Chen, W., Tan, Y., Li, C., He, Y., & Jia, X. (2021). Exploiting negative learning for implicit pseudo label rectification in source-free domain adaptive semantic segmentation. arXiv preprint arXiv:2106.12123.
- Luo, Y., Wang, Z., Chen, Z., Huang, Z., & Baktashmotlagh, M. (2023). Source-free progressive graph learning for open-set domain adaptation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.
- Ma, N., Bu, J., Lu, L., Wen, J., Zhang, Z., Zhou, S., et al. (2021). Semi-supervised hypothesis transfer for source-free domain adaptation. arXiv preprint arXiv:2107.06735.
- Ma, W., Chen, C., Zheng, S., Qin, J., Zhang, H., & Dou, Q. (2022). Test-time adaptation with calibration of medical image classification nets for label distribution shift. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 313–323). Springer.
- MacKay, D. J., Mac Kay, D. J., et al. (2003). *Information Theory, Inference and Learning Algorithms*. Cambridge University Press.

- Menze, B. H., Jakab, A., Bauer, S., Kalpathy-Cramer, J., Farahani, K., Kirby, J., et al. (2014). The multimodal brain tumor image segmentation benchmark (BRATS). *IEEE Transactions on Medical Imaging*, 34(10), 1993–2024.
- Müller, R., Kornblith, S., & Hinton, G. E. (2019). When does label smoothing help? *Advances in Neural Information Processing Systems*, 32.
- Mummadi, C. K., Huttmacher, R., Rambach, K., Levinkov, E., Brox, T., & Metzen, J. H. (2021). Test-time adaptation to distribution shift by confidence maximization and input transformation. *arXiv preprint arXiv:2106.14999*.
- Nasr, M., Shokri, R., & Houmansadr, A. (2019). Comprehensive privacy analysis of deep learning: Passive and active white-box inference attacks against centralized and federated learning. In *2019 IEEE Symposium on Security and Privacy* (pp. 739–753). IEEE.
- Nayak, G. K., Mopuri, K. R., Jain, S., & Chakraborty, A. (2021). Mining data impressions from deep models as substitute for the unavailable training data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(11), 8465–8481.
- Nelakurthi, A. R., Maciejewski, R., & He, J. (2018). Source free domain adaptation using an off-the-shelf classifier. In *2018 IEEE International Conference on Big Data (Big Data)* (pp. 140–145). IEEE.
- Northcutt, C., Jiang, L., & Chuang, I. (2021). Confident learning: Estimating uncertainty in dataset labels. *Journal of Artificial Intelligence Research*, 70, 1373–1411.
- Otter, D. W., Medina, J. R., & Kalita, J. K. (2020). A survey of the usages of deep learning for natural language processing. *IEEE Transactions on Neural Networks and Learning Systems*, 32(2), 604–624.
- Paul, S., Khurana, A., & Aggarwal, G. (2021). Unsupervised adaptation of semantic segmentation models without source data. *arXiv preprint arXiv:2112.02359*.
- Pei, J., Jiang, Z., Men, A., Chen, L., Liu, Y., & Chen, Q. (2023). Uncertainty-induced transferability representation for source-free unsupervised domain adaptation. *IEEE Transactions on Image Processing*, 32, 2033–2048.
- Peng, X., Bai, Q., Xia, X., Huang, Z., Saenko, K., & Wang, B. (2019). Moment matching for multi-source domain adaptation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 1406–1415).
- Peng, Q., Ding, Z., Lyu, L., Sun, L., & Chen, C. (2022). Toward better target representation for source-free and black-box domain adaptation. *arXiv preprint arXiv:2208.10531*.
- Peng, X., Huang, Z., Zhu, Y., & Saenko, K. (2019). Federated adversarial domain adaptation. *arXiv preprint arXiv:1911.02054*.
- Peng, X., Usman, B., Kaushik, N., Hoffman, J., Wang, D., & Saenko, K. (2017). VisDA: The visual domain adaptation challenge. *arXiv preprint arXiv:1710.06924*.
- Prabhu, V., Khare, S., Kartik, D., & Hoffman, J. (2021). S4T: Source-free domain adaptation for semantic segmentation via self-supervised selective self-training. *arXiv preprint arXiv:2107.10140*.
- Qin, Z., Yang, L., Gao, F., Hu, Q., & Shen, C. (2022). Uncertainty-aware aggregation for federated open set domain adaptation. *IEEE Transactions on Neural Networks and Learning Systems*, 1–15.
- Qin, C., You, H., Wang, L., Kuo, C. C. J., & Fu, Y. (2019). PointDAN: A multi-scale 3D domain adaption network for point cloud representation. *Advances in Neural Information Processing Systems*, 32.
- Qiu, Z., Zhang, Y., Lin, H., Niu, S., Liu, Y., Du, Q., et al. (2021). Source-free domain adaptation via avatar prototype generation and adaptation. *arXiv preprint arXiv:2106.15326*.
- Qu, X., Liu, L., Zhu, L., Nie, L., & Zhang, H. (2024). Source-free style-diversity adversarial domain adaptation with privacy-preservation for person re-identification. *Knowledge-Based Systems*, 283, Article 111150.
- Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., et al. (2021). Learning transferable visual models from natural language supervision. In *International Conference on Machine Learning* (pp. 8748–8763). PMLR.
- Rao, J., Meng, X., Ding, L., Qi, S., Liu, X., Zhang, M., et al. (2023). Parameter-efficient and student-friendly knowledge distillation. *IEEE Transactions on Multimedia*, 26, 4230–4241.
- Raychaudhuri, D. S., Ta, C. K., Dutta, A., Lal, R., & Roy-Chowdhury, A. K. (2023). Prior-guided source-free domain adaptation for human pose estimation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 14996–15006).
- Ren, P., Xiao, Y., Chang, X., Huang, P. Y., Li, Z., Chen, X., et al. (2021). A comprehensive survey of neural architecture search: Challenges and solutions. *ACM Computing Surveys*, 54(4), 1–34.
- Roy, S., Trapp, M., Pilzer, A., Kannala, J., Sebe, N., Ricci, E., et al. (2022). Uncertainty-guided source-free domain adaptation. In *European Conference on Computer Vision* (pp. 537–555). Springer.
- Saenko, K., Kulis, B., Fritz, M., & Darrell, T. (2010). Adapting visual category models to new domains. In *European Conference on Computer Vision* (pp. 213–226). Springer.
- Saito, K., Kim, D., Sclaroff, S., & Saenko, K. (2020). Universal domain adaptation through self supervision. *Advances in Neural Information Processing Systems*, 33, 16282–16292.
- Saito, K., & Saenko, K. (2021). OVANet: One-vs-all network for universal domain adaptation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 9000–9009).
- Saito, K., Watanabe, K., Ushiku, Y., & Harada, T. (2018). Maximum classifier discrepancy for unsupervised domain adaptation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 3723–3732).
- Saltoni, C., Lathuilière, S., Sebe, N., Ricci, E., & Galasso, F. (2020). SF-UDA<sup>3D</sup>: Source-free unsupervised domain adaptation for LIDAR-based 3D object detection. In *2020 International Conference on 3D Vision* (pp. 771–780). IEEE.
- Sankaranarayanan, S., Balaji, Y., Jain, A., Lim, S. N., & Chellappa, R. (2018). Learning from synthetic data: Addressing domain shift for semantic segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 3752–3761).
- Shen, M., Bu, Y., & Wornell, G. W. (2023). On balancing bias and variance in unsupervised multi-source-free domain adaptation. In *International Conference on Machine Learning* (pp. 30976–30991). PMLR.
- Shen, D., Wu, G., & Suk, H. I. (2017). Deep learning in medical image analysis. *Annual Review of Biomedical Engineering*, 19, 221.
- Shenaj, D., Fanì, E., Toldo, M., Caldarella, D., Tavera, A., Michieli, U., et al. (2023). Learning across domains and devices: Style-driven source-free domain adaptation in clustered federated learning. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision* (pp. 444–454).
- Sohn, K. (2016). Improved deep metric learning with multi-class N-pair loss objective. *Advances in Neural Information Processing Systems*, 29.
- Song, J., Chen, Y., Ye, J., & Song, M. (2022). Spot-adaptive knowledge distillation. *IEEE Transactions on Image Processing*, 31, 3359–3370.
- Song, L., Ma, C., Zhang, G., & Zhang, Y. (2020). Privacy-preserving unsupervised domain adaptation in federated setting. *IEEE Access*, 8, 143233–143240.
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, 15(1), 1929–1958.
- Stan, S., & Rostami, M. (2021a). Privacy preserving domain adaptation for semantic segmentation of medical images. *arXiv preprint arXiv:2101.00522*.
- Stan, S., & Rostami, M. (2021b). Unsupervised model adaptation for continual semantic segmentation. In *Proceedings of the AAAI Conference on Artificial Intelligence: vol. 35, (no. 3)*, (pp. 2593–2601).
- Sun, C., van Soest, J., & Dumontier, M. (2023). Generating synthetic personal health data using conditional generative adversarial networks combining with differential privacy. *Journal of Biomedical Informatics*, Article 104404.
- Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 2818–2826).
- Tan, A. Z., Yu, H., Cui, L., & Yang, Q. (2023). Towards personalized federated learning. *IEEE Transactions on Neural Networks and Learning Systems*, 34(12), 9587–9603.
- Tang, S., Shi, Y., Ma, Z., Li, J., Lyu, J., Li, Q., et al. (2021). Model adaptation through hypothesis transfer with gradual knowledge distillation. In *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 5679–5685). IEEE.
- Tang, S., Su, P., Chen, D., & Ouyang, W. (2021). Gradient regularized contrastive learning for continual domain adaptation. In *Proceedings of the AAAI Conference on Artificial Intelligence: vol. 35, (no. 3)*, (pp. 2665–2673).
- Tang, S., Yang, Y., Ma, Z., Hendrich, N., Zeng, F., Ge, S. S., et al. (2021). Nearest neighborhood-based deep clustering for source data-absent unsupervised domain adaptation. *arXiv preprint arXiv:2107.12585*.
- Tanwisuth, K., Zhang, S., Zheng, H., He, P., & Zhou, M. (2023). POUF: Prompt-oriented unsupervised fine-tuning for large pre-trained models. *arXiv preprint arXiv:2305.00350*.
- Tarvainen, A., & Valpola, H. (2017). Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results. *Advances in Neural Information Processing Systems*, 30.
- Taufique, A. M. N., Jahan, C. S., & Savakis, A. (2021). ConDA: Continual unsupervised domain adaptation. *arXiv preprint arXiv:2103.11056*.
- Tian, Q., Ma, C., Zhang, F. Y., Peng, S., & Xue, H. (2021). Source-free unsupervised domain adaptation with sample transport learning. *Journal of Computer Science and Technology*, 36(3), 606–616.
- Tian, Q., Peng, S., & Ma, T. (2022). Source-free unsupervised domain adaptation with trusted pseudo samples. *ACM Transactions on Intelligent Systems and Technology*, 14(2), 1–17.
- Tian, J., Zhang, J., Li, W., & Xu, D. (2021). VDM-DA: Virtual domain modeling for source data-free domain adaptation. *IEEE Transactions on Circuits and Systems for Video Technology*, 32(6), 3749–3760.
- Tierney, L., & Kadane, J. B. (1986). Accurate approximations for posterior moments and marginal densities. *Journal of the American Statistical Association*, 81(393), 82–86.
- Ulyanov, D., Vedaldi, A., & Lempitsky, V. (2016). Instance normalization: The missing ingredient for fast stylization. *arXiv preprint arXiv:1607.08022*.
- Venkateswara, H., Eusebio, J., Chakraborty, S., & Panchanathan, S. (2017). Deep hashing network for unsupervised domain adaptation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 5018–5027).
- Voulodimos, A., Doulamis, N., Doulamis, A., & Protopapadakis, E. (2018). Deep learning for computer vision: A brief review. *Computational Intelligence and Neuroscience*, 2018.
- VS, V., Oza, P., & Patel, V. M. (2023a). Instance relation graph guided source-free domain adaptive object detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 3520–3530).

- VS, V., Oza, P., & Patel, V. M. (2023b). Towards online domain adaptive object detection. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision* (pp. 478–488).
- VS, V., Oza, P., Sindagi, V. A., & Patel, V. M. (2022). Mixture of teacher experts for source-free domain adaptive object detection. In *2022 IEEE International Conference on Image Processing* (pp. 3606–3610). IEEE.
- VS, V., Valanarasu, J. M. J., & Patel, V. M. (2022). Target and task specific source-free domain adaptive image segmentation. arXiv preprint arXiv:2203.15792.
- Wang, M., & Deng, W. (2018). Deep visual domain adaptation: A survey. *Neurocomputing*, 312, 135–153.
- Wang, Q., Fink, O., Van Gool, L., & Dai, D. (2022). Continual test-time domain adaptation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 7201–7211).
- Wang, F., Han, Z., Gong, Y., & Yin, Y. (2022). Exploring domain-invariant parameters for source free domain adaptation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 7151–7160).
- Wang, F., Han, Z., Zhang, Z., & Yin, Y. (2022). Active source free domain adaptation. arXiv preprint arXiv:2205.10711.
- Wang, Y., Liang, J., & Zhang, Z. (2021). Give me your trained model: Domain adaptive semantic segmentation without source data. arXiv preprint arXiv:2106.11653.
- Wang, D., Shelhamer, E., Liu, S., Olshausen, B., & Darrell, T. (2020). Tent: Fully test-time adaptation by entropy minimization. arXiv preprint arXiv:2006.10726.
- Wang, R., Wu, Z., Weng, Z., Chen, J., Qi, G.-J., & Jiang, Y.-G. (2022). Cross-domain contrastive learning for unsupervised domain adaptation. *IEEE Transactions on Multimedia*, 25, 1665–1673.
- Wang, S., Zhang, D., Yan, Z., Shao, S., & Li, R. (2023). Black-box source-free domain adaptation via two-stage knowledge distillation. arXiv preprint arXiv:2305.07881.
- Wang, X., Zhuo, J., Cui, S., & Wang, S. (2021). Learning invariant representation with consistency and diversity for semi-supervised source hypothesis transfer. arXiv preprint arXiv:2107.03008.
- Wistuba, M., Rawat, A., & Pedapati, T. (2019). A survey on neural architecture search. arXiv preprint arXiv:1905.01392.
- Wu, K., Shi, Y., Han, Y., Shao, Y., Li, B., & Tian, Q. (2021). Domain adaptation without model transferring. arXiv preprint arXiv:2107.10174.
- Xia, H., Zhao, H., & Ding, Z. (2021). Adaptive adversarial network for source-free domain adaptation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 9010–9019).
- Xiao, Y., Wei, P., Liu, C., & Lin, L. (2023). Adversarially robust source-free domain adaptation with relaxed adversarial training. In *2023 IEEE International Conference on Multimedia and Expo* (pp. 2681–2686). IEEE.
- Xiong, L., Ye, M., Zhang, D., Gan, Y., Li, X., & Zhu, Y. (2021). Source data-free domain adaptation of object detector through domain-specific perturbation. *International Journal of Intelligent Systems*, 36(8), 3746–3766.
- Xiong, L., Ye, M., Zhang, D., Gan, Y., & Liu, Y. (2022). Source data-free domain adaptation for a faster R-CNN. *Pattern Recognition*, 124, Article 108436.
- Xu, R., Chen, Z., Zuo, W., Yan, J., & Lin, L. (2018). Deep cocktail network: Multi-source unsupervised domain adaptation with category shift. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 3964–3973).
- Xu, Z., Lu, D., Wang, Y., Luo, J., Jayender, J., Ma, K., et al. (2021). Noisy labels are treasure: Mean-teacher-assisted confident learning for hepatic vessel segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 3–13). Springer.
- Xu, Z., Lu, D., Wang, Y., Luo, J., Wei, D., Zheng, Y., et al. (2022). Denoising for relaxing: Unsupervised domain adaptive fundus image segmentation without source data. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 214–224). Springer.
- Xu, Z., Wei, W., Zhang, L., & Nie, J. (2022). Source-free domain adaptation for cross-scene hyperspectral image classification. In *IGARSS 2022-2022 IEEE International Geoscience and Remote Sensing Symposium* (pp. 3576–3579). IEEE.
- Xu, Y., Yang, J., Cao, H., Wu, K., Wu, M., & Chen, Z. (2022). Source-free video domain adaptation by learning temporal consistency for action recognition. In *European Conference on Computer Vision* (pp. 147–164). Springer.
- Xu, Y., Yang, J., Cao, H., Wu, K., Wu, M., Li, Z., et al. (2023). Multi-source video domain adaptation with temporal attentive moment alignment network. *IEEE Transactions on Circuits and Systems for Video Technology*, 33(8), 3860–3871.
- Xu, Y., Yang, J., Wu, M., Li, X., Xie, L., & Chen, Z. (2022). EXTERN: Leveraging endo-temporal regularization for black-box video domain adaptation. arXiv preprint arXiv:2208.05187.
- Yan, H., & Guo, Y. (2022). Dual moving average pseudo-labeling for source-free inductive domain adaptation. arXiv preprint arXiv:2212.08187.
- Yan, H., Guo, Y., & Yang, C. (2021). Augmented self-labeling for source-free unsupervised domain adaptation. In *NeurIPS 2021 Workshop on Distribution Shifts: Connecting Methods and Applications*.
- Yang, H., Chen, C., Jiang, M., Liu, Q., Cao, J., Heng, P. A., et al. (2022). DLTTA: Dynamic learning rate for test-time adaptation on cross-domain medical images. *IEEE Transactions on Medical Imaging*, 41(12), 3575–3586.
- Yang, S., van de Weijer, J., Herranz, L., Jui, S., et al. (2021). Exploiting the intrinsic neighborhood structure for source-free domain adaptation. *Advances in Neural Information Processing Systems*, 34, 29393–29405.
- Yang, C., Guo, X., Chen, Z., & Yuan, Y. (2022). Source free domain adaptation for medical image segmentation with fourier style mining. *Medical Image Analysis*, 79, Article 102457.
- Yang, C. Y., Kuo, Y. J., & Hsu, C. T. (2022). Source free domain adaptation for semantic segmentation via distribution transfer and adaptive class-balanced self-training. In *2022 IEEE International Conference on Multimedia and Expo* (pp. 1–6). IEEE.
- Yang, B., Ma, A. J., & Yuen, P. C. (2022). Revealing task-relevant model memorization for source-protected unsupervised domain adaptation. *IEEE Transactions on Information Forensics and Security*, 17, 716–731.
- Yang, J., Peng, X., Wang, K., Zhu, Z., Feng, J., Xie, L., et al. (2022). Divide to adapt: Mitigating confirmation bias for domain adaptation of black-box predictors. arXiv preprint arXiv:2205.14467.
- Yang, G., Tang, H., Zhong, Z., Ding, M., Shao, L., Sebe, N., et al. (2021). Transformer-based source-free domain adaptation. arXiv preprint arXiv:2105.14138.
- Yang, S., Wang, Y., van de Weijer, J., Herranz, L., & Jui, S. (2021). Generalized source-free domain adaptation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 8978–8987).
- Yang, S., Wang, Y., Herranz, L., Jui, S., & van de Weijer, J. (2023). Casting a BAIT for offline and online source-free domain adaptation. *Computer Vision and Image Understanding*, Article 103747.
- Yang, S., Wang, Y., Wang, K., Jui, S., & van de Weijer, J. (2022). Attracting and dispersing: A simple approach for source-free domain adaptation. CoRR.
- Yang, B., Yeh, H. W., Harada, T., & Yuen, P. C. (2021). Model-induced generalization error bound for information-theoretic representation learning in source-data-free unsupervised domain adaptation. *IEEE Transactions on Image Processing*, 31, 419–432.
- Yang, T., Zhu, S., Chen, C., Yan, S., Zhang, M., & Willis, A. (2020). MutualNet: Adaptive convnet via mutual learning from network width and resolution. In *European Conference on Computer Vision* (pp. 299–315). Springer.
- Yao, C. H., Gong, B., Qi, H., Cui, Y., Zhu, Y., & Yang, M. H. (2022). Federated multi-target domain adaptation. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision* (pp. 1424–1433).
- Yao, H., Guo, Y., & Yang, C. (2021). Source-free unsupervised domain adaptation with surrogate data generation. In *Proceedings of NeurIPS 2021 Workshop on Distribution Shifts: Connecting Methods and Applications*.
- Yazdanpanah, M., & Moradi, P. (2022). Visual domain bridge: A source-free domain adaptation for cross-domain few-shot learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 2868–2877).
- Ye, M., Zhang, J., Ouyang, J., & Yuan, D. (2021). Source data-free unsupervised domain adaptation for semantic segmentation. In *Proceedings of the 29th ACM International Conference on Multimedia* (pp. 2233–2242).
- Yeh, H. W., Meng, Q., & Harada, T. (2023). Misalignment-free relation aggregation for multi-source-free domain adaptation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 4313–4322).
- Yeh, H. W., Yang, B., Yuen, P. C., & Harada, T. (2021). SoFA: Source-data-free feature alignment for unsupervised domain adaptation. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision* (pp. 474–483).
- Yin, H., Molchanov, P., Alvarez, J. M., Li, Z., Mallya, A., Hoiem, D., et al. (2020). Dreaming to distill: Data-free knowledge transfer via DeepInversion. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 8715–8724).
- You, F., Li, J., & Zhao, Z. (2021). Test-time batch statistics calibration for covariate shift. arXiv preprint arXiv:2110.04065.
- Young, T., Hazarika, D., Poria, S., & Cambria, E. (2018). Recent trends in deep learning based natural language processing. *IEEE Computational Intelligence Magazine*, 13(3), 55–75.
- Yu, H., Huang, J., Liu, Y., Zhu, Q., Zhou, M., & Zhao, F. (2022). Source-free domain adaptation for real-world image dehazing. In *Proceedings of the 30th ACM International Conference on Multimedia* (pp. 6645–6654).
- Yu, Z., Li, J., Du, Z., Zhu, L., & Shen, H. T. (2023). A comprehensive survey on source-free domain adaptation. arXiv preprint arXiv:2302.11803.
- Yu, Q., Xi, N., Yuan, J., Zhou, Z., Dang, K., & Ding, X. (2023). Source-free domain adaptation for medical image segmentation via prototype-anchored feature alignment and contrastive learning. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 3–12). Springer.
- Zara, G., Conti, A., Roy, S., Lathuilière, S., Rota, P., & Ricci, E. (2023). The unreasonable effectiveness of large language-vision models for source-free video domain adaptation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 10307–10317).
- Zhang, H., Cisse, M., Dauphin, Y. N., & Lopez-Paz, D. (2017). Mixup: Beyond empirical risk minimization. arXiv preprint arXiv:1710.09412.
- Zhang, J., Li, W., Tang, C., Ogunbona, P., et al. (2020). Unsupervised domain expansion from multiple sources. arXiv preprint arXiv:2005.12544.
- Zhang, Y., Ren, Z., Feng, K., Yu, K., Beer, M., & Liu, Z. (2023). Universal source-free domain adaptation method for cross-domain fault diagnosis of machines. *Mechanical Systems and Signal Processing*, 191, Article 110159.



- Zhang, W., & Wu, D. (2020). Discriminative joint probability maximum mean discrepancy (DJP-MMD) for domain adaptation. In *2020 International Joint Conference on Neural Networks* (pp. 1–8). IEEE.
- Zhang, W., & Wu, D. (2023). Lightweight source-free transfer for privacy-preserving motor imagery classification. *IEEE Transactions on Cognitive and Developmental Systems*, 15(2).
- Zhang, D., Ye, M., Xiong, L., Li, S., & Li, X. (2021). Source-style transferred mean teacher for source-data free object detection. In *ACM Multimedia Asia* (pp. 1–8).
- Zhang, H., Zhang, Y., Jia, K., & Zhang, L. (2021). Unsupervised domain adaptation of black-box source models. *arXiv preprint arXiv:2101.02839*.
- Zhao, Y., Feng, S., Li, C., Song, R., Liang, D., & Chen, X. (2023). Source-free domain adaptation for privacy-preserving seizure prediction. *IEEE Transactions on Industrial Informatics*, 20(2), 2787–2798.
- Zhao, X., Stanislawski, R., Gardoni, P., Sulowicz, M., Glowacz, A., Krolczyk, G., et al. (2022). Adaptive contrastive learning with label consistency for source data free unsupervised domain adaptation. *Sensors*, 22(11), 4238.
- Zhao, Y., Zhong, Z., Luo, Z., Lee, G. H., & Sebe, N. (2022). Source-free open compound domain adaptation in semantic segmentation. *IEEE Transactions on Circuits and Systems for Video Technology*, 32(10), 7019–7032.
- Zhou, K., Liu, Z., Qiao, Y., Xiang, T., & Loy, C. C. (2022). Domain generalization: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(4), 4396–4415.
- Zhou, K., Yang, Y., Qiao, Y., & Xiang, T. (2021). Domain adaptive ensemble learning. *IEEE Transactions on Image Processing*, 30, 8008–8018.
- Zhu, L., Liu, Z., & Han, S. (2019). Deep leakage from gradients. *Advances in Neural Information Processing Systems*, 32.
- Zhuang, X., & Shen, J. (2016). Multi-scale patch and multi-modality atlases for whole heart segmentation of MRI. *Medical Image Analysis*, 31, 77–87.