Frontiers of Information Technology & Electronic Engineering
www.jzus.zju.edu.cn; engineering.cae.cn; www.springerlink.com
ISSN 2095-9184 (print); ISSN 2095-9230 (online)
E-mail: jzus@zju.edu.cn

Dynamic parameterized learning for unsupervised domain adaptation*

Runhua JIANG[1,2], Yahong HAN[‡1,2]

[1]College of Intelligence and Computing, Tianjin University, Tianjin 300350, China

[2]Tianjin Key Lab of Machine Learning, Tianjin University, Tianjin 300350, China

E-mail: ddghjikle1@gmail.com; yahong@tju.edu.cn

Received Dec. 9, 2022; Revision accepted Apr. 13, 2023; Crosschecked Oct. 26, 2023

Abstract: Unsupervised domain adaptation enables neural networks to transfer from a labeled source domain to an unlabeled target domain by learning domain-invariant representations. Recent approaches achieve this by directly matching the marginal distributions of these two domains. Most of them, however, ignore exploration of the dynamic trade-off between domain alignment and semantic discrimination learning, thus rendering them susceptible to the problems of negative transfer and outlier samples. To address these issues, we introduce the dynamic parameterized learning framework. First, by exploring domain-level semantic knowledge, the dynamic alignment parameter is proposed, to adaptively adjust the optimization steps of domain alignment and semantic discrimination learning. Besides, for obtaining semantic-discriminative and domain-invariant representations, we propose to align training trajectories on both source and target domains. Comprehensive experiments are conducted to validate the effectiveness of the proposed methods, and extensive comparisons are conducted on seven datasets of three visual tasks to demonstrate their practicability.

Key words: Unsupervised domain adaptation; Optimization steps; Domain alignment; Semantic discrimination
https://doi.org/10.1631/FITEE.2200631 **CLC number:** TP391

1 Introduction

Deep neural networks have achieved remarkable success in diverse computer vision tasks (Li S et al., 2021a; Zhang MY et al., 2021; Wu KH et al., 2023). Unfortunately, most of them rely on access to massive well-labeled training data. However, collecting a large number of well-labeled samples is time-consuming and laborious. Neural networks tend to have poor generalization on new datasets due to the distribution discrepancy issue (Li S et al., 2021a). As a result, unsupervised domain adaptation (UDA) methods have been widely studied.

The mainstream methodology of UDA methods is to learn domain-invariant representations by optimizing alignment loss functions. For example, Zhang YC et al. (2019) proposed to learn domain-invariant representations by minimizing the multi-kernel maximum mean discrepancy metric. Long et al. (2017) enhanced transferability of learned features by aligning the joint distributions of multiple domain-specific layers. Subsequently, several research works have found that semantic discriminability is critical. Specifically, Dai et al. (2021) proposed to regularize the geodesic path among intermediate, source, and target domains, thus obtaining strong class discrimination ability. Wang XM et al. (2019) introduced two types of attention mechanisms to highlight discriminative and transferable samples. By emphasizing domain-invariant

‡ Corresponding author
* Project supported by the National Natural Science Foundation of China (No. 61932009) and the Starry Night Science Fund of Zhejiang University Shanghai Institute for Advanced Study, China
ORCID: Runhua JIANG, https://orcid.org/0000-0003-2402-8684; Yahong HAN, https://orcid.org/0000-0003-2768-1398
© Zhejiang University Press 2023

or semantic discrimination, these loss functions can effectively enhance the representation ability of neural networks. However, independently considering these two characteristics causes networks susceptible to negative transfer and outlier samples.

In fact, for various domain adaptation (DA) scenarios, the degrees of alignment and discrimination should be different. Fig. 1 provides an intuitive example. It can be seen that other than bikes, this image contains components of other categories. Excessive alignment will emphasize components with small distribution discrepancy, e.g., backpacks or helmets, while excessive pursuit of discriminability will make neural networks overfit to the real-world domain. Motivated by this observation, some methods have been introduced to find dynamic balance. For instance, some works (Liu et al., 2021; Yang SQ et al., 2021; Balgi and Dukkipati, 2022) achieved this consideration by abstracting UDA as the combination of two related tasks: supervised classification of labeled source data and discriminative clustering of unlabeled target data. Furthermore, Tian et al. (2022) proposed to regularize the relationship between domain alignment and discrimination by class-level weights, while Xiao and Zhang (2021) introduced a sample-weighting mechanism. However, because samples randomly drawn from source and target domains cannot cover all valuable categories, these methods are still sensitive to outlier samples in a mini-batch.

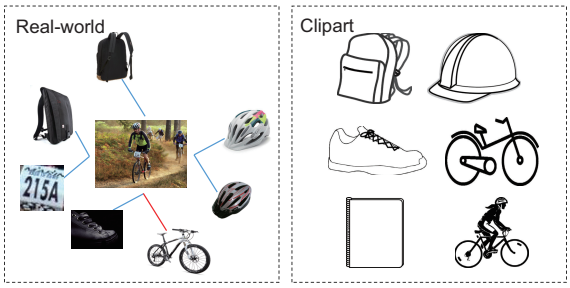

Fig. 1 Taking an image of a bike in the real-world domain as an example, there exist components of other categories. Excessive alignment will force neural networks to concentrate on backpacks or helmets, which have smaller cross-domain discrepancies than other components

To address this issue, we introduce dynamic parameterized learning (DPL) framework. On one hand, the dynamic alignment parameter (DAP) is calculated by recurrently estimating the consistency between class distributions of source and target domains. During optimization, the processes of domain alignment and discrimination learning are dynamically adjusted by taking DAPs as the weights of loss functions. On the other hand, by considering that these two processes cannot be absolutely separated, we propose the previously contrastive loss (PCL) and adaptive learning parameter (ALP) to constrain the optimization trajectories (Cazenavette et al., 2022) on source and target domains. In each iteration, PCL and ALP first estimate the variations between the previous and current parameters and then enforce the parameter trajectories on the target domain to match those on the source domain. With the above methods, UDA models are able to adaptively adjust the learning processes, thus becoming robust to the outlier samples and negative transfer.

The main contributions of this work can be concluded as follows:

1. We introduce the DPL framework to address the issues of outlier samples and improper trade-off between domain alignment and semantic discrimination learning.

2. In this framework, DAP is proposed to dynamically adjust the degrees of domain alignment and semantic discrimination learning. We also propose PCL and ALP to enable learning of cross-domain discriminative representations.

3. Comprehensive experimental analysis is provided to demonstrate the effectiveness of the proposed methods. Extensive comparisons are conducted on seven datasets of regression, classification, and person re-identification (Re-ID) tasks to validate the practicability of the proposed methods.

2 Related works

UDA has been extensively studied in various visual tasks (Han YH et al., 2021; Li S et al., 2021a; Wu AM et al., 2022). Among various tasks, domain adaptation regression is a basic but challenging task due to its continuous solution space. Chen XY et al. (2021) found that regression methods are not robust to feature scaling. Therefore, they introduced a geometrical loss function, termed the representation subspace distance (RSD) loss, to learn transferable representations. Different from regression, classification generates discrete labels by normalizing multi-class probabilities. Most of existing methods follow

the methodology of aligning intermediate representations of convolution neural networks. As a result, these features should contain both class discriminability and domain transferability. For discriminability, Li MX et al. (2020) took attention mechanism into account in enhanced transport distance (ETD) and then proposed the attention-aware transport distance and entropy-based regularization. Han ZY et al. (2022) demonstrated that neural networks always include nontransferable parameters, which should not be always updated for domains with small distribution discrepancies. For transferability, the domain-adversarial neural network (DANN) (Ganin et al., 2016) achieves UDA classification by both generative adversarial network and adversarial learning strategies. Dai et al. (2021) found that forming an intermediate domain can greatly enhance the representative ability of learned features. According to these research works, it can be concluded that enhancing semantic discriminability or domain transferability of deep representations can improve generalization performance.

However, how to balance semantic discriminability and domain transferability for effective UDA is still underexplored. Liu et al. (2021) demonstrated that cross-domain alignment should be achieved based on underlying discriminative features of source and target domains, rather than all features. Therefore, they introduced the cycle self-training (CST) to learn the underlying discriminative features of source and target domains by self-supervised learning. In contrast, other works (Yang SQ et al., 2021; Balgi and Dukkipati, 2022) considered the domain adaptation task as supervised learning on the source domain and unsupervised learning on the target domain. Both of them first learned discriminative representations from the source domain and then fine-tuned these representations on the target domain by adopting clustering schemes. Moreover, Li WK and Chen (2022) demonstrated that the target risk is bounded by both intra-domain structure and between-domain discrepancy. Wang XM et al. (2019) proposed transferable attention for domain adaptation (TADA), which consists of local attention generated by region-level domain discriminators and global attention generated by image-level discriminators. Although the motivations of this present work and the above methods are similar, there exist two important differences. First, the proposed DPL achieves the above motivation from the model optimization perspective. DPL dynamically weights the optimization steps on the source and target domains, while these methods generate sample-wise, class-level, or subdomain-level weights for random samples. Second, DPL is a flexible framework and thus can be used to solve not only DA classification but also regression and Re-ID.

3 Proposed methods

To introduce the proposed methods, we formally define the labeled source domain $D_\mathrm{s} = (x_\mathrm{s}, y_\mathrm{s})$ and the unlabeled target domain $D_\mathrm{t} = (x_\mathrm{t})$. At each iteration, a neural network with encoder $f(\cdot)$ and projection head $g(\cdot)$ takes x_s and x_t as the inputs and the corresponding y_s as supervision. Without losing generality, we assume the loss function of training the neural network as follows:

$$L = L_\mathrm{D} + \alpha L_\mathrm{A}, \tag{1}$$

where the discrimination loss L_D takes y_s and $g(f(x_\mathrm{s}))$ to supervise semantic discrimination learning, and the alignment loss L_A aims at enhancing domain transferability. α is a predefined hyperparameter for L_A.

3.1 Dynamic alignment parameter

As discussed earlier, optimizing Eq. (1) with fixed α easily causes imbalance between domain alignment and semantic discrimination learning. Therefore, previous methods (Xiao and Zhang, 2021; Tian et al., 2022) proposed to weight each sample or category according to its contributions. However, they are still sensitive to outlier samples in a mini-batch. In addition, computing domain discrepancy based on limited samples tends to be inaccurate. Therefore, we propose to use domain-level knowledge for achieving dynamic balance. In general, there are two kinds of methods for modeling domain-level knowledge (Pan YW et al., 2019; Tanwisuth et al., 2021). The first is to form average latent features for all categories in each domain (Pan YW et al., 2019), while the other is to learn parametric representations (Tanwisuth et al., 2021). Specifically, Tanwisuth et al. (2021) demonstrated that with optimization, parameters in the mapping head are encouraged to form domain-invariant prototypes. Therefore, we can obtain prototypes for source and

target domains by optimizing two mapping heads, i.e., $g_{\mathrm{s}}(\cdot)$ and $g_{\mathrm{t}}(\cdot)$. In each iteration, they respectively handle features of source or target domain. After being optimized, parameters in $g_{\mathrm{s}}(\cdot)$ and $g_{\mathrm{t}}(\cdot)$, termed as θ_{s} and θ_{t} respectively, are encouraged to form prototypes for each domain. Therefore, the dynamic balance can be achieved by estimating the similarity between these prototypes. When prototypes of different domains are similar, the representation space within neural networks is robust to the domain discrepancy. Thus, neural networks should focus more on discrimination learning. Otherwise, optimization contribution of domain alignment should be emphasized. To this end, we define DAP as follows:

$$\mathrm{DAP} = \frac{\lambda}{N}\sum_{n=1}^{N}|\cos(\theta_{\mathrm{s}}^{n}, \theta_{\mathrm{t}}^{n})|, \tag{2}$$

where N is the number of the parameters θ, $\cos(\cdot)$ is the cosine distance, and λ is a hyper-parameter for regularizing the range of the alignment loss function. By taking DPL to replace α in Eq. (1), the optimization of domain alignment and discrimination learning can be adaptively controlled.

3.2 Cross-domain discriminability

As Fig. 2 presents, in each iteration, neural networks are optimized with dynamic balance between domain alignment and discrimination learning. However, domain alignment and discrimination learning are two closely related processes. Therefore, we further consider jointly enhancing domain transferability and semantic discriminability for effective DA. It is straightforward to find that aligning class-level representations can achieve this purpose (Tian et al., 2022). However, estimating the transferability for each class is still challenging due to the unavailability of target labels. Therefore, we propose to achieve the above consideration from the optimization perspective.

The core of our method involves constraining the optimization trajectories on the source and target domains to be synchronized, such that semantic knowledge within learned representations is forced to be transferable. Specifically, as discussed in previous works (Liu et al., 2021; Yang SQ et al., 2021; Balgi and Dukkipati, 2022), domain adaptation can be seen as two processes: supervised learning on the source domain and unsupervised learning on the target domain. During optimization of objective functions such as Eq. (1), the decision boundary on the source domain is formed according to the source label, while the boundary of the target domain is obtained by minimizing the statistical distance between the source and target samples. By matching the optimization trajectories of the source and target domains, these two boundaries are hoped to share similar discrimination. Thus, semantic discriminability of the learned representations can be maintained as highly as possible, while transferability is assured by cooperative learning. Accordingly, here we propose PCL based on contrastive learning (Li QB et al., 2021). For supervising the optimization steps on the target domain, PCL is defined as follows:

$$L_{\mathrm{PCL}} = -\log\frac{\exp(\cos(f(x_{\mathrm{s}}), f(x_{\mathrm{t}}))/\tau)}{\exp(\cos(f(x_{\mathrm{s}}), f(x_{\mathrm{t}}))/\tau)+\exp(\cos(f(x_{\mathrm{t}}), f_{\mathrm{p}}(x_{\mathrm{t}}))/\tau)}, \tag{3}$$

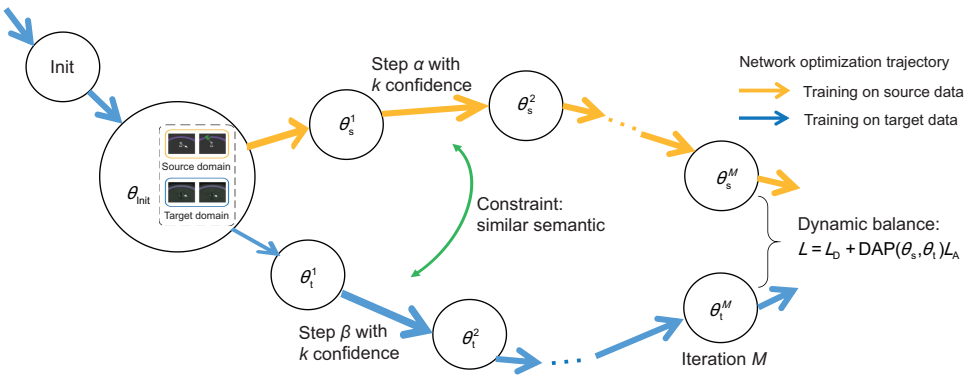

Fig. 2 Illustration of the proposed methods. In each iteration, optimization steps (Cazenavette et al., 2022) on the source and target domains are constrained to be semantically similar. Meanwhile, loss functions of domain alignment and discrimination learning are dynamically balanced by the proposed dynamic parameterized learning

where $f_{\mathrm{p}}(\cdot)$ denotes the previous encoder and $\cos(\cdot)$ is the cosine similarity. τ is a temperature parameter that equals 0.5. By replacing x_{t} in the second term of the denominator as x_{s}, PCL can be used to supervise the optimization steps on the source domain. With similar methodology of contrastive learning, optimizing the proposed PCL forces the encoder $f(\cdot)$ to decrease domain-specific information in the latent space and maximize the semantic consistency in representations learned by the current and previous parameters.

We hope that the proposed PCL can be dynamically optimized. Therefore, we further propose ALP as the confidence of PCL. This is achieved by measuring the representative ability of learned features (Yang YC and Soatto, 2020; Wang W et al., 2022). If these features can effectively represent the distributions of the input samples, values of the ALP become large, thus encouraging cross-domain transferability. Otherwise, ALP suppresses the performance of the PCL. Concretely, given x_{s} and $f(x_{\mathrm{s}})$, frequency distribution of the input image x_{s} is obtained by Fourier transform and compared with distributions of $f(x_{\mathrm{s}})$. Formally, the above operations can be defined as follows:

$$\mathrm{ALP_s} = \frac{\gamma}{N}\sum_{n=1}^{N}\cos(\mathrm{Pool}(\mathrm{FFT}(x_{\mathrm{s}}^{n})), f(x_{\mathrm{s}}^{n})), \quad (4)$$

where $\mathrm{Pool}(\cdot)$ and $\mathrm{FFT}(\cdot)$ denote the max-pooling operation and Fourier transform, respectively. γ is a predefined parameter for regularizing the range of ALP. We can also obtain $\mathrm{ALP_t}$ by replacing x_{s} by x_{t}.

3.3 Overall training objective

The overall training objective of this work includes discrimination loss L_{D}, alignment loss L_{A} with the proposed DAP, and the proposed PCL with ALP. Formally, we propose to optimize the following function:

$$\begin{aligned} L =& L_{\mathrm{D}} + \mathrm{DPL}\cdot L_{\mathrm{A}} + \mathrm{ALP_s}\cdot L_{\text{PCL-s}} \\ &+ \mathrm{ALP_t}\cdot L_{\text{PCL-t}}, \end{aligned} \quad (5)$$

where the last two terms in the right-hand side are computed to match and supervise the optimization steps on the source and target domains. As the above objective presents, the proposed framework can be used with different alignment loss L_{A} and various discrimination loss L_{D} to solve different tasks.

4 Experiments

In this section, the proposed methods are evaluated and compared with several state-of-the-art UDA methods on three tasks. Thus, we first introduce these three tasks and implementation details. Then, variations of the proposed methods are verified to validate their effectiveness. Finally, the proposed methods are compared with existing methods to demonstrate their practicability.

4.1 Tasks and datasets

4.1.1 UDA regression

UDA regression models take images as inputs and generate constants to describe the visual characteristics of the input images. In this work, datasets dSprites (Higgins et al., 2016) and MPI3D (Gondal et al., 2019) are taken to make comparisons. dSprites is a two-dimensional (2D) synthetic dataset for deep representation learning. It consists of three domains: Color (C), Noisy (N), and Scream (S). These three domains contain 737 280 images, which have five factors of variations. Following Chen XY et al. (2021), three factors are used for the regression experiments: scale, position X, and position Y, whose possible values are taken from [0.5, 1], [0, 1], and [0, 1], respectively. MPI3D is a simulation-to-real dataset of three-dimensional (3D) objects. There are three domains in this dataset: Toy (T), Realistic (RC), and Real (RL). Each of these domains has 1 036 800 images with seven factors. According to previous works (Gondal et al., 2019; Chen XY et al., 2021), the values of the horizontal and vertical axes are evaluated for regression. For comparing model performance, the mean absolute error (MAE) distance is taken as the metric.

4.1.2 UDA classification

For the classification task, the popular Office-Home (Venkateswara et al., 2017), Office-31 (Saenko et al., 2010), and DomainNet (Peng et al., 2019) datasets are used. The Office-Home dataset contains 15 500 images of 65 categories. These images can be split into four distinct domains: Artistic images (Ar), Clip art (Cl), Product images (Pr), and Real-World images (Rw). Therefore, 12 adaptation tasks are formed to evaluate the proposed methods. Office-31 is a real-world benchmark dataset for DA

classification. In this dataset, 4110 images of 31 categories can be split into three domains: Amazon (A), Webcam (W), and DSLR (D). As a result, six adaptation tasks are set up to make comparisons.

4.1.3 UDA person Re-ID

In this work, three datasets of UDA Re-ID are used to make comparisons, as done by Dai et al. (2021). To be specific, the Market-1501 dataset, which contains over 32 000 images, is first taken (Zheng et al., 2015). Each identity in this dataset has multiple images under different cameras, leading to this dataset being quite challenging. Then, DukeMTMC-ReID (Ristani et al., 2016) is adopted as a different domain. This dataset contains more than 2×10^6 frames of 1080 pixels, which are taken from 60-FPS (FPS is short for frame per second) videos captured by eight cameras. Finally, the real-world dataset MSMT17 (Wei LH et al., 2018) is used. This dataset contains 4101 identities and 126 441 bounding boxes. Following Dai et al. (2021), four adaptation tasks are formed: DukeMTMC-ReID → Market-1501, Market-1501 → DukeMTMC-ReID, Market-1501 → MSMT17, and DukeMTMC-ReID → MSMT17. Popular metrics such as mean average precision (mAP) and Rank-1/5/10 (R1/5/10) of cumulative match characteristic (CMC) are used to evaluate the performances.

4.2 Implementation details

4.2.1 UDA regression

Following Chen XY et al. (2021), all images are resized to 224×224 pixels and data augmentation is not used. During training, the batch size is fixed as 128. The learning rates of $f(\cdot)$ and $g(\cdot)$ are set to 0.1 and 1.0, respectively. For updating the network parameters, mini-batch stochastic gradient descent (SGD) with a momentum of 0.95 and weight decay of 0.000 05 is adopted. During the experiments, pretrained ResNet-18 (He et al., 2016) and two fully connected layers are taken as models. For optimization, the discrimination loss L_{D} and alignment loss L_{A} are defined as the MAE and RSD functions, respectively (Chen XY et al., 2021). All experiments are implemented on a single GeForce RTX 2080Ti GPU and trained about 20 000 iterations.

4.2.2 UDA classification

Following standard protocol for UDA classification (Li S et al., 2021a; Xiao and Zhang, 2021), we use all the labeled source data and unlabeled target data as training samples and evaluate on the unlabeled target data. To fairly compare with existing methods, we use pretrained ResNet-50 or ResNet-101 (He et al., 2016) as the encoder. In the experiments, the discrimination loss and alignment loss are defined the same as those in previous research (Li S et al., 2021a). All input images are cropped to 224×224 pixel dimensions. The mini-batch SGD optimizer with momentum of 0.9 and the learning rate strategy described by Ganin and Lempitsky (2015) is used to learn the parameters. All experiments are conducted on a single GeForce RTX 2080Ti GPU with 20 000 iterations in total.

4.2.3 UDA person Re-ID

For all UDA Re-ID experiments, the input images are resized to 256×128, and common data augmentation operations are applied. During training, each batch consists of 64 source images with 16 identities and 64 target images with 16 pseudo identities. The initial learning rate is set to 3.5e-4, which will be divided by 10 at the 20th and 40th epochs, respectively. The Adam optimizer with the weight decay of 5×10^{-4} and momentum of 0.9 is adopted in our experiments. Similar to Dai et al. (2021), ResNet-50 pretrained on ImageNet and domain-specific batch normalization layers are adopted as the backbone network. In this work, all experiments of this task are implemented on two GeForce RTX 3090 GPUs. During testing, there are no postprocessing techniques.

4.3 Experimental analysis

In this subsection, we present several experiments to demonstrate the effectiveness of the proposed methods. Due to the regression task being more sensitive than the other two tasks, all experiments are conducted on the dSprites dataset. For comprehensive validation, we first conduct an ablation study based on Eq. (5). Then, we compare the variants of the proposed DAP. Finally, the benefits of matching the optimization trajectories on the source and target domains are discussed.

4.3.1 Ablation study

As Eq. (5) indicates, DAP serves as the coefficient of the alignment loss. The proposed PCL and ALP supervise the optimization trajectories on the source or target domain. Therefore, the ablation study is conducted by removing one of these contributions. The corresponding results are presented in Table 1, from which we can find that removing these contributions always degrades the final performance. First, by comparing the top and bottom rows, it can be observed that by removing the proposed DAP, neural networks are easily affected by inaccurate balance, leading to suboptimal performance. Similarly, without supervising and matching of the optimization trajectories on the source or target domain, the model performance tends to be worse. In addition, the results in the second and fourth rows indicate that the source trajectory has a more important influence than the target trajectory. This verifies the motivation of cross-domain discriminability. Last but not least, the proposed ALP, which adaptively controls the performance of the PCL, is beneficial for the overall performance.

Even with the above results, it is still not demonstrated that DPL can address the issue of outlier samples. Therefore, we present three experiments to validate this ability. Specifically, DPL is trained with outlier samples generated by three methods: (1) perturbing source labels with random noise; (2) perturbing source images with random noise; (3) replacing the source or target images with samples from the other domain. As can be seen from Figs. 3a–3c, when the source labels are perturbed by noise randomly sampled from [0, 0.1], [0, 0.5], and [0.5, 1], the proposed DPL can achieve similar MAEs. Similarly, it is robust to outlier images caused by random noise. Last but not least, when the source or target images are replaced, the proposed DPL still tends to work well on S→C and S→N, indicating that the proposed methods can adaptively adjust domain alignment and discrimination learning.

4.3.2 Analysis of dynamic alignment parameter

For further validating the effectiveness of the proposed DAP, we change the hyper-parameter λ in Eq. (2). The corresponding results are presented in Fig. 4a. For analysis, we further provide the performance after replacing DAP by a constant of 0.001 and $\cos(\cdot)$ by the Kullback–Leibler (KL) distance. As shown in Table 2, the performance in Fig. 4a always outperforms these results. From Fig. 4a, it is obvious that with small λ, the values of DAP become small, and thus the effectiveness of the alignment loss is reduced, leading to inadequate alignment and worse performance (Xiao and Zhang, 2021). By comparing the results of the N → C and S → N scenarios, it can be found that the values of λ do not seriously influence the overall performance. Therefore, it is demonstrated that even without suitable λ, the proposed DAP can still coordinate with alignment loss functions to achieve dynamic trade-off and effective representations. In addition, it can be found that with suitable λ, methods with DAP can remarkably

Table 2 Experimental results related to the proposed dynamic alignment parameter

Method	Mean average error		
	C→S	N→C	S→N
Constant of 0.001	0.79	0.15	0.12
KL with $\lambda = 0.1$	0.84	0.16	0.11

For comparison, we replace dynamic parameterized learning by the constant of 0.001. Meanwhile, $\cos(\cdot)$ is replaced by the Kullback–Leibler (KL) divergence distance

Table 1 Ablation study of the proposed methods

Method	DAP	$L_{\text{PCL-s}}$	ALP_s	$L_{\text{PCL-t}}$	ALP_t	Mean average error		
						C→N	C→S	S→N
w/o DAP		✓	✓	✓	✓	0.27	0.78	0.10
w/o source trajectory	✓			✓	✓	0.32	0.84	0.07
w/o ALP source	✓	✓		✓	✓	1.03	0.85	0.14
w/o target trajectory	✓	✓	✓			0.39	0.58	0.07
w/o ALP target	✓	✓	✓	✓		1.00	0.47	0.26
Ours	✓	✓	✓	✓	✓	0.24	0.46	0.07

"✓" indicates that the corresponding contributions are used. For simplicity, the inputs of these contributions are abbreviated. w/o: without

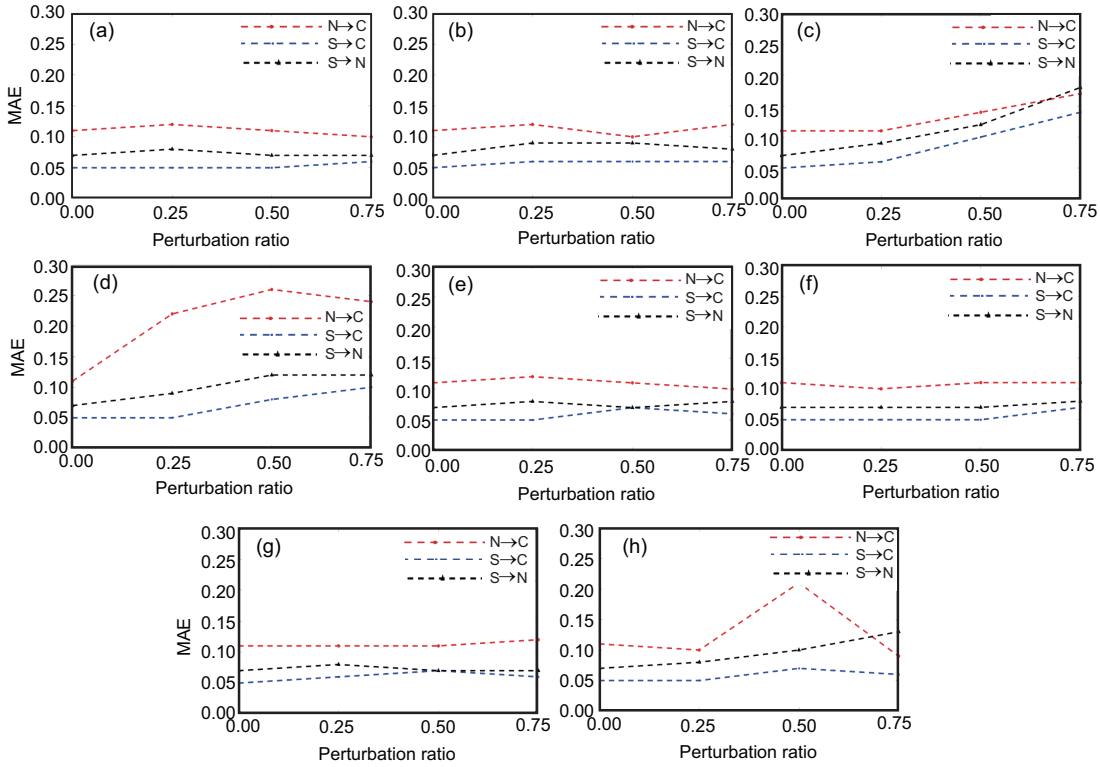

Fig. 3 Experimental results of whether dynamic parameterized learning can address the issue of outlier samples: (a–c): perturbing source labels by noise randomly sampled from [0, 0.1], [0, 0.5], [0.5, 1]; (e–g): perturbing source images by noise randomly sampled from [0, 0.1], [0, 0.5], [0.5, 1]; (d, h): replacing source or target images by samples from the other domain. Values in the horizontal axis denote the ratio of the number of perturbed samples to that of all the samples

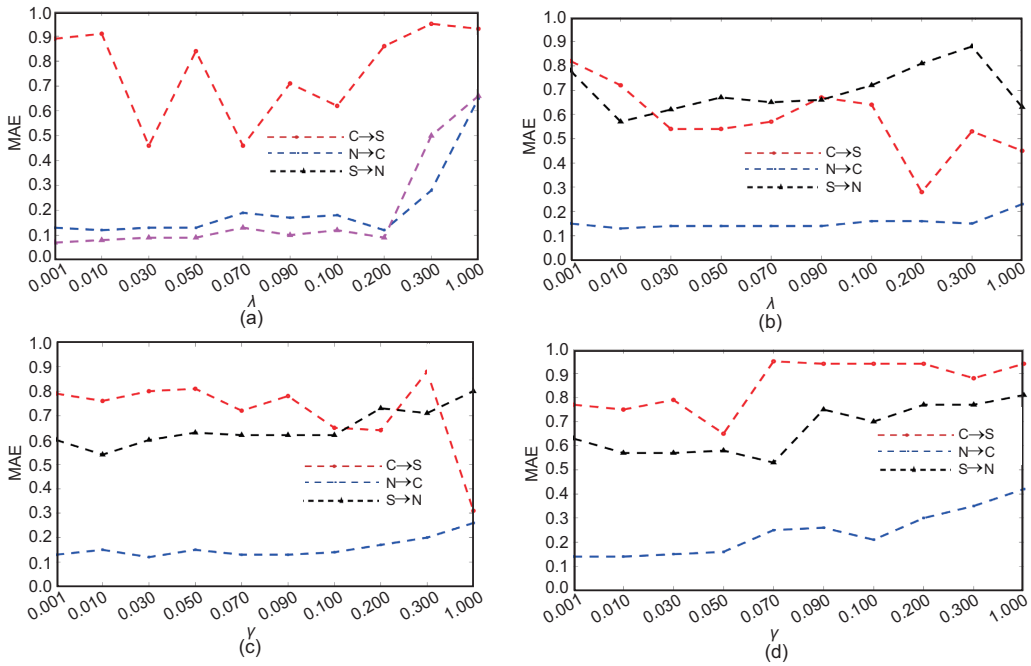

Fig. 4 Experimental results obtained by varying λ and γ: (a) the proposed DAP; (b) PCL with different weights; (c) influence of hyper-parameters in $\mathrm{ALP_s}$; (d) influence of hyper-parameters in $\mathrm{ALP_t}$. Note that in (c) and (d), only $L_{\text{PCL-s}}$ or $L_{\text{PCL-t}}$ is used with the proposed ALP. The proposed DPL is used only in (a)

outperform those with 0.001 as the hyper-parameter. Last but not least, by comparing Fig. 4a and Table 2, we conclude that compared with KL divergence, the cosine distance adopted in Eq. (2) is more effective. From the above results, it is demonstrated that the proposed DAP can effectively estimate and promote the balance between domain alignment and discrimination learning.

4.3.3 Effectiveness of matching optimization trajectories

Based on the above experiments, we consider the effectiveness of matching the optimization trajectories by PCL and ALP. To this end, we first give $\mathrm{PCL_s}$ with different constants as the hyper-parameters. From Fig. 4b, it can be found that these constant hyper-parameters always degrade the regression performance. By comparing Figs. 4b and 4c, we note that with the proposed ALP, it is not necessary to find the optimal hyper-parameter of $L_{\text{PCL-s}}$ by numerous experiments. After being equipped with the ALP, the proposed PCL can effectively constrain the optimization trajectory, leading to the best performance on most tasks. This phenomenon demonstrates that by matching the learning process on the source and target domains, models can achieve better adaptation performance. In addition, as the best performance on each adaptation task is achieved with different γ values, we conclude that the proposed PCL and ALP can be easily applied to various adaptation scenarios. In Fig. 4, we can further compare the influence of $\mathrm{ALP_s}$ and $\mathrm{ALP_t}$ by observing Figs. 4c and 4d. It is easy to find that with the same γ, methods trained with $\mathrm{ALP_s}$ always work better than those with $\mathrm{ALP_t}$. Therefore, the core consideration of PCL and ALP is demonstrated to be empirically reasonable, i.e., using the optimization trajectory on the source domain to guide the learning process on the target domain.

As PCL aims at matching the optimization trajectories on different domains, there exists a natural question about whether the steps of the matched trajectories should be the same. For addressing this issue, we implement $\mathrm{ALP_s}$ and $\mathrm{ALP_t}$ with different values of γ. As a result, the matched trajectories are controlled to various steps. Table 3 illustrates the results of the above experiments. First, when γ of $\mathrm{ALP_s}$ and $\mathrm{ALP_t}$ becomes different, the overall performance is drastically degraded. This validates the reason that steps of the matched trajectories should be similar. In addition, it is obvious that matching the short-term trajectories is more effective than aligning the long-term steps. Overall, we demonstrate that the steps of the matched trajectories should be similar for learning effective representations.

4.4 Comparison

4.4.1 UDA regression

The proposed methods are first compared with regression methods on both dSprites and MPI3D datasets. According to the methodology, these methods can be split into two categories: machine learning and deep representation learning. To name a few, transfer component analysis (TCA) (Pan SJ et al., 2011) and joint distribution optimal transport (JDOT) (Courty et al., 2017) are two classical machine learning methods. Both of them focus on importance weighting. For deep representation learning methods, RSD (Chen XY et al., 2021) is the latest approach, which concentrates on aligning orthogonal bases. The experimental results are presented in Table 4, from which we can find that our method achieves state-of-the-art performance. To be specific, on the dSprites dataset,

Table 3 Performance of matching different optimization trajectories on the source and target domains

γ^*	Mean average error											
	C→N				N→C				S→N			
	γ^{**}=0.001	0.01	0.1	1	0.001	0.01	0.1	1	0.001	0.01	0.1	1
0.001	0.25	0.24	0.22	0.59	0.13	0.12	0.28	0.52	0.10	0.10	0.11	0.61
0.01	0.26	0.27	0.35	0.69	0.12	0.13	0.32	0.42	0.11	0.11	0.11	0.07
0.1	0.31	0.30	0.31	0.57	0.17	0.13	0.14	0.42	0.11	0.10	0.14	0.36
1	1.03	0.78	1.03	0.93	0.26	0.23	0.22	0.54	0.20	0.21	0.18	0.30

For fair comparison, the proposed DPL is not used. * γ is used for computing $\mathrm{ALP_t}$; ** γ is used for computing $\mathrm{ALP_s}$

Table 4 Experimental results on the UDA regression task

Method	Mean average error on dSprites						
	C→N	C→S	N→C	N→S	S→C	S→N	Avg
ResNet-18 (He et al., 2016)	0.94	0.90	0.16	0.65	0.08	0.26	0.498
TCA (Pan SJ et al., 2011)	0.94	0.87	0.19	0.66	0.10	0.23	0.498
DAN (Long et al., 2015)	0.70	0.77	0.12	**0.50**	0.06	0.11	0.377
DANN (Ganin et al., 2016)	0.47	0.46	0.16	0.65	**0.05**	0.10	0.315
JDOT (Courty et al., 2017)	0.86	0.79	0.19	0.64	0.10	0.23	0.468
MCD (Saito et al., 2018)	0.81	0.81	0.17	0.65	0.07	0.19	0.450
AFN (Xu RJ et al., 2019)	1.00	0.96	0.16	0.62	0.08	0.32	0.523
RSD (Chen XY et al., 2021)	0.32	**0.35**	0.16	0.57	0.08	0.09	**0.258** (0.262)
Ours	**0.24**	0.46	**0.11**	0.62	**0.05**	**0.07**	**0.258**

Method	Mean average error on MPI3D						
	RL→RC	RL→T	RC→RL	RC→T	T→RL	T→RC	Avg
ResNet-18 (He et al., 2016)	0.17	0.44	0.19	0.45	0.51	0.50	0.377
TCA (Pan SJ et al., 2011)	0.17	0.42	0.19	0.42	0.50	0.50	0.373 (0.367)
DAN (Long et al., 2015)	0.12	0.35	0.12	0.27	0.40	0.41	0.278
DANN (Ganin et al., 2016)	0.09	0.24	**0.11**	0.41	0.48	0.37	0.283
JDOT (Courty et al., 2017)	0.16	0.41	0.16	0.41	0.47	0.47	0.353 (0.347)
MCD (Saito et al., 2018)	0.13	0.40	0.15	0.45	0.52	0.50	0.358
AFN (Xu RJ et al., 2019)	0.18	0.45	0.20	0.46	0.53	0.52	0.390
RSD (Chen XY et al., 2021)	0.10	0.23	**0.11**	0.17	0.41	0.42	0.242 (0.240)
Ours	**0.08**	**0.20**	0.14	**0.12**	**0.35**	**0.33**	**0.203**

Best results are in bold. The average values in the bracket are re-calculated in this study according to the results cited from the literature

the proposed methods outperform ResNet-18 on all adaptation scenarios. Compared with DANN, the proposed methods tend to achieve better performance on difficult tasks such as C→N and N→S. On the MPI3D dataset, our methods outperform RSD on most scenarios except RC→RL. In addition, since RSD can be seen as the baseline model of this work, it is demonstrated that adaptively learning semantic discriminability and dynamically aligning domain-invariant knowledge can improve the overall adaptation performance.

4.4.2 UDA classification

Results of UDA classification are illustrated in Tables 5–9. From Table 5, it can be found that this work outperforms all the methods on average accuracy. This is attributed to the improvement on tasks such as Ar→Cl, Ar→Pr, Cl→Ar, and Cl→Pr. However, the proposed methods achieve comparable performance on other adaptation tasks. The above phenomenon demonstrates the benefits of the dynamic balance between domain alignment and discrimination learning. On the Office-31 dataset, it is easy to find that almost all the methods cannot perform well on difficult tasks such as D→A, while DPL achieves the best performance. In addition, the proposed methods can be extended to multi-source DA. As can be seen from Tables 8 and 9, DPL achieves the best performance on the multi-source DA cases. Compared with close-set DA, multi-source, open-set, and partial DA are more difficult cases. However, it is easy to find that under the open-set and partial-set settings, DPL is sensitive to the mismatch of label spaces. In our future research, this issue will be addressed by further considering the discrimination learning process of each class.

4.4.3 UDA person Re-ID

Results of UDA person Re-ID are summarized in Table 10. It can be found that the proposed methods achieve state-of-the-art performance. Specifically, compared with generative adversarial network (GAN) based transferring method such as pose disentanglement and adaptation network (PDA-Net) (Li YJ et al., 2019), the proposed methods tend to show better performance by considering the cross-domain discriminability between source and target domains. For the fine-tuning methods such as AD-Cluster (Zhai et al., 2020) and noise resistible mutual-training (NRMT) (Zhao F et al., 2020), the

Table 5 Experimental results on Office-Home

Method	Accuracy (%)												
	Ar→Cl	Ar→Pr	Ar→Rw	Cl→Ar	Cl→Pr	Cl→Rw	Pr→Ar	Pr→Cl	Pr→Rw	Rw→Ar	Rw→Cl	Rw→Pr	Avg
SymNets (Zhang YB et al., 2019)	47.7	72.9	78.5	64.2	71.3	**74.2**	64.2	48.8	79.5	**74.5**	52.6	82.7	67.6
TADA (Wang XM et al., 2019)	53.1	72.3	77.2	59.1	71.2	72.1	59.7	53.1	78.4	72.4	60.0	82.9	67.6
MDD (Zhang YC et al., 2019)	54.9	73.7	77.8	60.0	71.4	71.8	61.2	53.6	78.1	72.5	**60.2**	82.3	68.1
ALDA (Chen MH et al., 2020)	53.7	70.1	76.4	60.2	72.6	71.5	56.8	51.9	77.1	70.2	56.3	82.1	66.6
ETD (Li MX et al., 2020)	51.3	71.9	**85.7**	57.6	69.2	73.7	57.8	51.2	79.3	70.2	57.5	82.1	67.3
SymmNets-V2 (Zhang YB et al., 2020)	48.1	74.3	78.7	**64.6**	71.8	74.1	**64.4**	50.0	80.2	74.3	53.1	83.2	68.1
TransPar-DANN (Han ZY et al., 2022)	55.1	70.4	77.8	61.6	70.8	72.1	62.1	55.1	**80.8**	73.8	60.0	83.0	68.6
CKB (Luo YW and Ren, 2021)	54.7	74.4	77.1	63.7	72.2	71.8	64.1	51.7	78.4	73.1	58.0	82.4	68.5
TSA (Li S et al., 2021b)	53.6	75.1	78.3	64.4	**73.7**	72.5	62.3	49.4	77.5	72.2	58.8	82.1	68.3
DANNPE (Wei GQ et al., 2021)	54.7	72.8	78.5	62.3	71.1	73.1	61.0	53.0	80.0	72.8	56.5	**83.4**	68.3
DWL (Xiao and Zhang, 2021)	45.6	63.9	72.2	55.5	60.8	64.6	58.0	46.8	73.9	69.3	52.0	78.0	61.7
DIA (Lu et al., 2021)	51.7	70.3	74.4	58.5	68.4	68.2	56.1	**67.5**	75.3	66.5	56.9	81.1	66.2
AMRC (Jing et al., 2022)	49.8	69.9	75.5	57.1	69.8	70.2	56.9	50.5	77.1	69.6	55.8	81.2	65.3
CRSL (Xu YY and Yan, 2022)	51.4	70.7	76.8	53.3	65.5	66.1	55.0	46.7	74.1	67.7	51.7	81.0	63.3
Ours	**55.9**	**76.7**	80.0	64.5	73.3	73.5	62.9	51.9	78.3	73.1	59.6	82.3	**69.3**

Best results are in bold

Table 6 Experimental results on Office-31

Method	Accuracy (%)						
	A→W	D→W	W→D	A→D	D→A	W→A	Avg
MDD (Zhang YC et al., 2019)	94.5	98.4	**100.0**	93.5	74.6	72.2	88.9
SymNets (Zhang YB et al., 2019)	90.8	98.8	**100.0**	93.9	74.6	72.5	88.4
TADA (Wang XM et al., 2019)	94.3	98.7	99.8	91.6	72.9	73.0	88.4
ETD (Li MX et al., 2020)	92.1	**100.0**	**100.0**	88.0	71.0	67.8	86.2 (86.5)
ALDA (Chen MH et al., 2020)	**95.6**	97.7	**100.0**	94.0	72.2	72.5	88.7
SymmNets-V2 (Zhang YB et al., 2020)	94.2	98.8	**100.0**	93.5	74.4	73.4	89.1
TransPar-DANN (Han ZY et al., 2022)	92.3	98.9	**100.0**	89.8	74.9	72.3	88.0
DWL (Xiao and Zhang, 2021)	89.2	99.2	**100.0**	91.2	73.1	69.8	87.1
TSA (Li S et al., 2021b)	94.8	99.1	**100.0**	92.6	74.9	74.4	89.3
DANNPE (Wei GQ et al., 2021)	92.5	98.5	99.9	89.9	70.7	71.1	87.0 (87.1)
NRC (Yang SQ et al., 2021)	90.8	99.0	**100.0**	**96.0**	75.3	**75.0**	89.4
CUDA (Balgi and Dukkipati, 2022)	**95.6**	99.1	**100.0**	**96.0**	73.2	74.7	**89.8**
SUDA (Zhang JY et al., 2022)	90.8	98.7	**100.0**	91.2	72.2	71.4	87.4
AMRC (Jing et al., 2022)	92.7	97.7	**100.0**	91.2	69.8	69.1	86.8
SEMA (Zuo et al., 2022)	87.1	98.3	99.4	83.6	70.4	70.9	85.0
Ours	93.7	99.2	99.8	94.6	**75.4**	74.8	89.6

Best results are in bold. The average values in the bracket are re-calculated in this study according to the results cited from the literature

Table 7 Accuracy (%) on DomainNet for unsupervised domain adaptation

DAN	Accuracy (%)							JAN	Accuracy (%)						
	clp	inf	pnt	qdr	rel	skt	Avg		clp	inf	pnt	qdr	rel	skt	Avg
clp	–	9.1	23.4	16.2	37.9	29.7	23.2 (23.3)	clp	–	7.8	24.5	14.3	38.1	25.7	22.1
inf	17.2	–	15.6	4.4	24.8	13.5	15.1	inf	17.6	–	18.7	8.7	28.1	15.3	17.7
pnt	29.9	8.9	–	7.9	42.1	26.1	23.0	pnt	27.5	8.2	–	7.1	43.1	23.9	22.0
qdr	14.2	1.6	4.4	–	8.5	10.1	7.8	qdr	17.8	2.2	7.4	–	8.1	10.9	9.3
rel	37.4	11.5	33.3	10.1	–	26.4	23.7	rel	33.5	9.1	32.5	7.5	–	21.9	20.9
skt	39.1	8.8	28.2	13.9	36.2	–	25.2	skt	35.3	8.2	27.7	13.3	36.8	–	24.3
Avg	27.6	8.0	21.0	10.5	29.9	21.2	19.7	Avg	26.3	7.1	22.2	10.2	30.8	19.5	19.4

DANN	Accuracy (%)							RTN	Accuracy (%)						
	clp	inf	pnt	qdr	rel	skt	Avg		clp	inf	pnt	qdr	rel	skt	Avg
clp	–	9.1	23.2	13.7	37.6	28.6	22.4	clp	–	8.1	21.1	13.1	36.1	26.5	21.0
inf	17.9	–	16.4	2.1	27.8	13.3	15.5	inf	15.6	–	15.3	3.4	25.1	12.8	14.4
pnt	29.1	8.6	–	5.1	41.5	24.7	21.8	pnt	26.8	8.1	–	5.2	40.6	22.6	20.7
qdr	16.8	1.8	4.8	–	9.3	10.2	8.6	qdr	15.1	1.8	4.5	–	8.5	8.9	7.8
rel	36.5	11.4	33.9	5.9	–	24.5	22.4	rel	35.3	10.7	31.7	7.5	–	22.9	21.6
skt	37.9	8.2	26.3	12.2	35.3	–	24.0	skt	34.1	7.4	23.3	12.6	32.1	–	21.9
Avg	27.6	7.8	20.9	7.8	30.3	20.3	19.1	Avg	25.4	7.2	19.2	8.4	28.4 (28.5)	18.7	17.9

ADDA	Accuracy (%)							MCD	Accuracy (%)						
	clp	inf	pnt	qdr	rel	skt	Avg		clp	inf	pnt	qdr	rel	skt	Avg
clp	–	11.2	24.1	3.2	41.9	30.7	22.2	clp	–	14.2	26.1	1.6	45.0	33.8	24.1
inf	19.1	–	16.4	3.2	26.9	14.6	16.0	inf	23.6	–	21.2	1.5	36.7	18.0	20.2
pnt	31.2	9.5	–	8.4	39.1	25.4	22.7	pnt	34.4	14.8	–	1.9	50.5	28.4	26.0
qdr	15.7	2.6	5.4	–	9.9	11.9	9.1	qdr	15.0	3.0	7.0	–	11.5	10.2	9.3
rel	39.5	14.5	29.1	12.1	–	25.7	24.2	rel	42.6	19.6	42.6	2.2	–	29.3	27.2 (27.3)
skt	35.3	8.9	25.2	14.9	37.6	–	25.4 (24.4)	skt	41.2	13.7	27.6	3.8	34.8	–	24.2
Avg	28.2	9.3	$20.1^{20.0}$	8.4	31.1	21.7	19.8	Avg	31.4	13.1	24.9	2.2	35.7	23.9	21.9

TransPar	Accuracy (%)							DPL	Accuracy (%)						
	clp	inf	pnt	qdr	rel	skt	Avg		clp	inf	pnt	qdr	rel	skt	Avg
clp	–	26.2	41.7	24.8	53.1	57.6	40.7	clp	–	18.3	39.2	2.81	54.0	42.3	31.3
inf	20.5	–	19.4	4.7	22.7	20.5	17.6	inf	30.1	–	33.1	1.0	45.3	22.5	26.4
pnt	38.5	27.7	–	10.2	51.8	47.0	35.0	pnt	44.0	18.7	–	1.2	55.9	40.5	32.1
qdr	13.3	4.7	0.0	–	10.1	12.0	8.0	qdr	25.4	4.3	13.8	–	22.3	18.8	16.9
rel	56.2	44.0	55.5	17.4	–	55.8	45.8	rel	52.2	22.2	50.8	1.9	–	40.7	33.5
skt	45.2	25.9	38.1	20.4	39.4	–	33.8	skt	55.3	17.9	44.0	8.1	52.8	–	35.6
Avg	34.7	25.7	30.9	15.5	35.4	38.6	30.1	Avg	41.4	16.3	36.2	3.0	46.1	33.0	29.3

The column-wise domains are selected as the source domain, and the row-wise domains are selected as the target domain. The average values in the bracket are re-calculated in this study according to the results cited from the literature

performance improvement achieved by our methods demonstrates that excessively or inadequately aligning the source and target domains is not beneficial for the high-level object Re-ID task. By comparing the joint training methods (Zhong et al., 2019, 2021) and our methods, it can also be validated that the dynamic alignment between different domains can enhance the overall adaptation performance. Last but not least, compared with multi-source DA methods (Bai ZC et al., 2021), the performance improvement further validates that matching the optimization trajectories on different domains is beneficial for learning effective features. Overall, from the above experiments, we demonstrate that our methods achieve state-of-the-art performance on the UDA Re-ID task.

Table 8 Accuracy comparisons with multi-source domain adaptation methods on Office-31 dataset

Model	Accuracy (%)			
	R→D	R→W	R→A	Avg
D-CORAL (Sun et al., 2016)	99.2	98.9	69.2	89.1
DCTN (Xu RJ et al., 2018)	99.3	98.2	64.2	87.2
MADAN (Zhao SC et al., 2021)	99.4	98.4	63.9	87.2
Adv-Ensemble (Rakshit et al., 2019)	99.3	97.3	68.1	88.3 (88.2)
MFSAN (Zhu et al., 2019)	99.5	98.5	72.7	90.2
DAN (Long et al., 2015)	99.6	97.8	67.6	88.3
D-CORAL (Sun et al., 2016)	99.3	98.0	67.1	88.1
DSBN-M (Chang et al., 2019)	99.0	98.8	70.1	89.3
DSAN (Zhu et al., 2021)	99.1	98.6	72.4	90.0
DPL (ours)	**99.8**	**99.0**	**75.0**	**91.3**

R means other domains. The average value in the bracket is re-calculated in this study according to the results cited from the literature. Best results are in bold

Table 9 Accuracy comparisons with multi-source domain adaptation methods on Office-Home dataset

Model	Accuracy (%)				
	R→Ar	R→Cl	R→Pr	R→Rw	Avg
MFSAN (Zhu et al., 2019)	72.1	62.0	80.3	81.8	74.1
SImpAl (Venkat et al., 2020)	70.8	56.3	80.2	81.5	72.2
SHOT-Ens (Liang et al., 2020)	72.2	59.3	82.8	82.9	74.3
DECISION (Ahmed et al., 2021)	74.5	59.4	**84.4**	83.6	75.5
DAN (Long et al., 2015)	68.5	59.4	79.0	82.5	72.4
D-CORAL (Sun et al., 2016)	68.1	58.6	79.5	82.7	72.2
RevGrad (Ganin and Lempitsky, 2015)	68.4	59.1	79.5	82.7	72.4
SHOT (Liang et al., 2020)	72.1	57.2	83.4	81.4	73.5
DPL (ours)	**74.6**	**62.4**	84.1	**84.0**	**76.3**

R means other domains. Best results are in bold

5 Conclusions

In this work, we propose to explore the dynamic balance between semantic discrimination learning and domain alignment learning. We propose DPL to achieve task-adaptive balance by modeling domain-level semantic knowledge, which is more robust and effective than previous methods based on sample-level or class-wise knowledge. In addition, we propose PCL and ALP to adaptively match the optimization trajectories on source and target domains, thus obtaining cross-domain discriminative representations. Extensive experimental results are presented to validate the rationality and effectiveness of the proposed methods.

Contributors

Runhua JIANG designed the research and drafted the paper. Yahong HAN helped organize the paper. Runhua JIANG revised and finalized the paper.

Compliance with ethics guidelines

Yahong HAN is a corresponding expert of *Frontiers of Information Technology & Electronic Engineering*, and he was not involved with the peer review process of this paper. Runhua JIANG and Yahong HAN declare that they have no conflict of interest.

Data availability

The data that support the findings of this study are openly available in Transfer-Learning-Library at https://github.com/thuml/Transfer-Learning-Library.

References

Ahmed SM, Raychaudhuri DS, Paul S, et al., 2021. Unsupervised multi-source domain adaptation without access to source data. Proc IEEE/CVF Conf on Computer Vision and Pattern Recognition, p.10103-10112. https://doi.org/10.1109/CVPR46437.2021.00997

Bai Y, Wang C, Lou YH, et al., 2021. Hierarchical connectivity-centered clustering for unsupervised domain adaptation on person re-identification. *IEEE Trans Image Process*, 30:6715-6729. https://doi.org/10.1109/TIP.2021.3094140

Bai ZC, Wang ZG, Wang J, et al., 2021. Unsupervised multi-source domain adaptation for person re-identification. Proc IEEE/CVF Conf on Computer Vision and Pattern Recognition, p.12914-12923. https://doi.org/10.1109/CVPR46437.2021.01272

Table 10 Experimental results on UDA Re-ID

Method	DukeMTMC-ReID→Market-1501				Market-1501→DukeMTMC-ReID			
	mAP	R1	R5	R10	mAP	R1	R5	R10
ECN (Zhong et al., 2019)	43.0	75.1	87.6	91.6	40.4	63.3	75.8	80.4
PDA-Net (Li YJ et al., 2019)	47.6	75.2	86.3	90.2	45.1	63.2	77.0	82.5
SSG (Fu et al., 2019)	58.3	80.0	90.0	92.4	53.4	73.0	80.6	83.2
MMCL (Wang DK and Zhang, 2020)	60.4	84.4	92.8	95.0	51.4	72.4	82.9	85.0
ECN-GPP (Zhong et al., 2021)	63.8	84.1	92.8	95.4	54.4	74.0	83.7	87.4
AD-Cluster (Zhai et al., 2020)	68.3	86.7	94.4	96.5	54.1	72.6	82.5	85.5
NRMT (Zhao F et al., 2020)	71.7	87.8	94.6	96.5	62.2	77.8	86.9	89.5
HCC-UDPL (Bai Y et al., 2021)	69.8	87.1	94.7	96.7	59.9	74.9	86.0	89.6
PREST (Zhang H et al., 2021)	62.4	82.5	92.1	94.9	56.1	74.4	83.7	85.9
MSDA (Bai ZC et al., 2021)	68.4	85.9	94.3	96.5	55.8	72.9	84.4	88.1
ICE (Li YY et al., 2022)	73.8	**90.8**	95.8	97.2	**66.4**	**80.2**	**88.5**	91.6
HCN (Zhang Z et al., 2022)	70.2	90.2	–	–	57.3	78.9	–	–
MGML (Tao et al., 2022)	68.5	88.3	94.9	96.9	64.4	79.9	88.2	**93.1**
Ours	**75.6**	89.7	**96.0**	**97.7**	62.2	78.0	85.9	88.4

Method	Market-1501→MSMT17				DukeMTMC-ReID→MSMT17			
	mAP	R1	R5	R10	mAP	R1	R5	R10
ECN (Zhong et al., 2019)	8.5	25.3	36.3	42.1	10.2	30.2	41.5	46.8
SSG (Fu et al., 2019)	13.2	31.6	–	49.6	13.3	32.2	–	51.2
ECN-GPP (Zhong et al., 2021)	15.2	40.4	53.1	58.7	16.0	42.5	55.9	61.5
MMCL (Wang DK and Zhang, 2020)	15.1	40.8	51.8	56.7	16.2	43.6	54.3	58.9
NRMT (Zhao F et al., 2020)	19.8	43.7	56.5	62.2	20.6	45.2	57.8	63.3
CAIL (Luo CC et al., 2020)	20.4	43.7	56.1	61.9	24.3	51.7	64.0	68.9
MMT (Ge et al., 2020)	22.9	49.2	63.1	68.8	23.3	50.1	63.9	69.8
PREST (Zhang H et al., 2021)	15.9	37.8	51.8	57.8	18.5	43.8	57.5	63.6
MDL (Xu X and Zhang, 2021)	20.6	45.5	59.6	66.3	23.1	50.0	62.5	69.2
HGA (Zhang MY et al., 2021)	25.5	**55.1**	61.2	65.5	26.8	**58.6**	64.7	69.2
MSDA (Bai ZC et al., 2021)	20.7	46.9	60.0	65.0	21.3	47.7	60.8	66.4
I-CBAM (Wang W et al., 2022)	24.2	51.3	**65.1**	**70.8**	24.1	51.7	65.5	71.1
ICE (Li YY et al., 2022)	24.0	50.9	–	–	24.3	51.6	–	–
MGML (Tao et al., 2022)	22.2	47.3	60.3	66.5	23.0	49.1	61.5	66.2
Ours	**26.1**	52.0	63.7	68.2	**28.0**	55.2	**67.2**	**71.5**

Bests results are in bold

Balgi S, Dukkipati A, 2022. Contradistinguisher: a Vapnik's imperative to unsupervised domain adaptation. *IEEE Trans Patt Anal Mach Intell*, 44(9):4730-4747. https://doi.org/10.1109/TPAMI.2021.3071225

Cazenavette G, Wang TZ, Torralba A, et al., 2022. Dataset distillation by matching training trajectories. Proc IEEE/CVF Conf on Computer Vision and Pattern Recognition, p.4750-4759. https://doi.org/10.1109/CVPR52688.2022.01045

Chang WG, You T, Seo S, et al., 2019. Domain-specific batch normalization for unsupervised domain adaptation. Proc IEEE/CVF Conf on Computer Vision and Pattern Recognition, p.7354-7362. https://doi.org/10.1109/CVPR.2019.00753

Chen MH, Zhao S, Liu HF, et al., 2020. Adversarial-learned loss for domain adaptation. Proc 34th AAAI Conf on Artificial Intelligence, p.3521-3528. https://doi.org/10.1609/aaai.v34i04.5757

Chen XY, Wang SN, Wang JM, et al., 2021. Representation subspace distance for domain adaptation regression. Proc 38th Int Conf on Machine Learning, p.1749-1759.

Courty N, Flamary R, Habrard A, et al., 2017. Joint distribution optimal transportation for domain adaptation. Proc 31st Int Conf on Neural Information Processing Systems, p.3733-3742.

Dai YX, Liu J, Sun YF, et al., 2021. IDM: an intermediate domain module for domain adaptive person Re-ID. Proc IEEE/CVF Int Conf on Computer Vision, p.11864-11874. https://doi.org/10.1109/ICCV48922.2021.01165

Fu Y, Wei YC, Wang GS, et al., 2019. Self-similarity grouping: a simple unsupervised cross domain adaptation approach for person re-identification. Proc IEEE/CVF Int Conf on Computer Vision, p.6112-6121. https://doi.org/10.1109/ICCV.2019.00621

Ganin Y, Lempitsky V, 2015. Unsupervised domain adaptation by backpropagation. Proc 32nd Int Conf on Machine Learning, p.1180-1189.

Ganin Y, Ustinova E, Ajakan H, et al., 2016. Domain-adversarial training of neural networks. In: Csurka G, (Ed.), Domain Adaptation in Computer Vision Applications. Advances in Computer Vision and Pattern Recognition. Springer, Cham, p.189-209. https://doi.org/10.1007/978-3-319-58347-1_10

Ge Y, Chen DP, Li HS, 2020. Mutual mean-teaching: pseudo label refinery for unsupervised domain adaptation on

person re-identification. Proc 8th Int Conf on Learning Representations.

Gondal MW, Wuthrich M, Miladinovic D, et al., 2019. On the transfer of inductive bias from simulation to the real world: a new disentanglement dataset. Proc 33rd Int Conf on Neural Information Processing Systems, Article 1410.

Han YH, Wu AM, Zhu LC, et al., 2021. Visual commonsense reasoning with directional visual connections. *Front Inform Technol Electron Eng*, 22(5):625-637. https://doi.org/10.1631/FITEE.2000722

Han ZY, Sun HL, Yin YL, 2022. Learning transferable parameters for unsupervised domain adaptation. *IEEE Trans Image Process*, 31:6424-6439. https://doi.org/10.1109/TIP.2022.3184848

He KM, Zhang XY, Ren SQ, et al., 2016. Deep residual learning for image recognition. Proc IEEE Conf on Computer Vision and Pattern Recognition, p.770-778. https://doi.org/10.1109/CVPR.2016.90

Higgins I, Matthey L, Pal A, et al., 2016. beta-VAE: learning basic visual concepts with a constrained variational framework. Proc 5th Int Conf on Learning Representations.

Jing MM, Meng LC, Li JJ, et al., 2022. Adversarial mixup ratio confusion for unsupervised domain adaptation. *IEEE Trans Multimedia*, 25:2559-2572. https://doi.org/10.1109/TMM.2022.3148592

Li MX, Zhai YM, Luo YW, et al., 2020. Enhanced transport distance for unsupervised domain adaptation. Proc IEEE/CVF Conf on Computer Vision and Pattern Recognition, p.13936-13944. https://doi.org/10.1109/CVPR42600.2020.01395

Li QB, He BS, Song D, 2021. Model-contrastive federated learning. Proc IEEE/CVF Conf on Computer Vision and Pattern Recognitionn, p.10713-10722. https://doi.org/10.1109/CVPR46437.2021.01057

Li S, Xie MX, Lv FR, et al., 2021a. Semantic concentration for domain adaptation. Proc IEEE/CVF Int Conf on Computer Vision, p.9102-9111. https://doi.org/10.1109/ICCV48922.2021.00897

Li S, Xie MX, Gong KX, et al., 2021b. Transferable semantic augmentation for domain adaptation. Proc IEEE/CVF Conf on Computer Vision and Pattern Recognition, p.11516-11525. https://doi.org/10.1109/CVPR46437.2021.01057

Li WK, Chen SC, 2022. Partial domain adaptation without domain alignment. *IEEE Trans Patt Anal Mach Intell*, 45(7):8787-8797. https://doi.org/10.1109/TPAMI.2022.3228937

Li YJ, Lin CS, Lin YB, et al., 2019. Cross-dataset person re-identification via unsupervised pose disentanglement and adaptation. Proc IEEE/CVF Int Conf on Computer Vision, p.7919-7929. https://doi.org/10.1109/ICCV.2019.00801

Li YY, Yao HT, Xu CS, 2022. Intra-domain consistency enhancement for unsupervised person re-identification. *IEEE Trans Multimedia*, 24:415-425. https://doi.org/10.1109/TMM.2021.3052354

Liang J, Hu DP, Feng JS, 2020. Do we really need to access the source data? Source hypothesis transfer for unsupervised domain adaptation. Proc 37th Int Conf on Machine Learning, Article 560.

Liu H, Wang JM, Long MS, 2021. Cycle self-training for domain adaptation. Proc 35th Advances in Neural Information Processing Systems, p.22968-22981.

Long MS, Cao Y, Wang JM, et al., 2015. Learning transferable features with deep adaptation networks. Proc 32nd Int Conf on Machine Learning, p.97-105.

Long MS, Zhu H, Wang JM, et al., 2017. Deep transfer learning with joint adaptation networks. Proc 34th Int Conf on Machine Learning, p.2208-2217.

Lu YW, Li DS, Wang WJ, et al., 2021. Discriminative invariant alignment for unsupervised domain adaptation. *IEEE Trans Multimedia*, 24:1871-1882. https://doi.org/10.1109/TMM.2021.3073258

Luo CC, Song CF, Zhang ZX, 2020. Generalizing person re-identification by camera-aware invariance learning and cross-domain mixup. Proc 16th European Conf on Computer Vision, p.224-241. https://doi.org/10.1007/978-3-030-58555-6_14

Luo YW, Ren CX, 2021. Conditional bures metric for domain adaptation. Proc IEEE/CVF Conf on Computer Vision and Pattern Recognition, p.13989-13998. https://doi.org/10.1109/CVPR46437.2021.01377

Pan SJ, Tsang IW, Kwok JT, et al., 2011. Domain adaptation via transfer component analysis. *IEEE Trans Neur Netw*, 22(2):199-210. https://doi.org/10.1109/TNN.2010.2091281

Pan YW, Yao T, Li YH, et al., 2019. Transferrable prototypical networks for unsupervised domain adaptation. Proc IEEE/CVF Conf on Computer Vision and Pattern Recognition, p.2239-2247. https://doi.org/10.1109/CVPR.2019.00234

Peng XC, Bai QX, Xia XD, et al., 2019. Moment matching for multi-source domain adaptation. Proc IEEE/CVF Int Conf on Computer Vision, p.1406-1415. https://doi.org/10.1109/ICCV.2019.00149

Rakshit S, Banerjee B, Roig G, et al., 2019. Unsupervised multi-source domain adaptation driven by deep adversarial ensemble learning. Proc 41st German Conf on Pattern Recognition, p.485-498. https://doi.org/10.1007/978-3-030-33676-9_34

Ristani E, Solera F, Zou R, et al., 2016. Performance measures and a data set for multi-target, multi-camera tracking. European Conf on Computer Vision, p.17-35. https://doi.org/10.1007/978-3-319-48881-3_2

Saenko K, Kulis B, Fritz M, et al., 2010. Adapting visual category models to new domains. Proc 11th European Conf on Computer Vision, p.213-226. https://doi.org/10.1007/978-3-642-15561-1_16

Saito K, Watanabe K, Ushiku Y, et al., 2018. Maximum classifier discrepancy for unsupervised domain adaptation. Proc IEEE/CVF Conf on Computer Vision and Pattern Recognition, p.3723-3732. https://doi.org/10.1109/CVPR.2018.00392

Sun BC, Feng JS, Saenko K, 2016. Return of frustratingly easy domain adaptation. Proc 30th AAAI Conf on Artificial Intelligence, p.2058-2065.

Tanwisuth K, Fan XJ, Zheng HJ, et al., 2021. A prototype-oriented framework for unsupervised domain adaptation. Proc 35th Conf on Neural Information Processing Systems, p.17194-17208.

Tao XF, Kong J, Jiang M, et al., 2022. Unsupervised domain adaptation by multi-loss gap minimization learning for person re-identification. *IEEE Trans Circ Syst Video Technol*, 32(7):4404-4416. https://doi.org/10.1109/TCSVT.2021.3135274

Tian Q, Zhu YN, Sun HY, et al., 2022. Unsupervised domain adaptation through dynamically aligning both the feature and label spaces. *IEEE Trans Circ Syst Video Technol*, 32(12):8562-8573. https://doi.org/10.1109/TCSVT.2022.3192135

Venkat N, Kundu JN, Singh DK, et al., 2020. Your classifier can secretly suffice multi-source domain adaptation. Proc 34th Int Conf on Neural Information Processing Systems, Article 390.

Venkateswara H, Eusebio J, Chakraborty S, et al., 2017. Deep hashing network for unsupervised domain adaptation. Proc IEEE Conf on Computer Vision and Pattern Recognition, p.5018-5027. https://doi.org/10.1109/CVPR.2017.572

Wang DK, Zhang SL, 2020. Unsupervised person re-identification via multi-label classification. Proc IEEE/CVF Conf on Computer Vision and Pattern Recognition, p.10981-10990. https://doi.org/10.1109/CVPR42600.2020.01099

Wang W, Zhao F, Liao S, et al., 2022. Attentive waveblock: complementarity-enhanced mutual networks for unsupervised domain adaptation in person re-identification and beyond. *IEEE Trans Image Process*, 31:1532-1544. https://doi.org/10.1109/TIP.2022.3140614

Wang XM, Li L, Ye WR, et al., 2019. Transferable attention for domain adaptation. Proc 33rd AAAI Conf on Artificial Intelligence, p.5345-5352. https://doi.org/10.1609/aaai.v33i01.33015345

Wei GQ, Lan CL, Zeng WJ, et al., 2021. MetaAlign: coordinating domain alignment and classification for unsupervised domain adaptation. Proc IEEE/CVF Conf on Computer Vision and Pattern Recognition, p.16643-16653. https://doi.org/10.1109/CVPR46437.2021.01637

Wei LH, Zhang SL, Gao W, et al., 2018. Person transfer GAN to bridge domain gap for person re-identification. Proc IEEE/CVF Conf on Computer Vision and Pattern Recognition, p.79-88. https://doi.org/10.1109/CVPR.2018.00016

Wu AM, Han YH, Zhu LC, et al., 2022. Instance-invariant domain adaptive object detection via progressive disentanglement. *IEEE Trans Patt Anal Mach Intell*, 44(8):4178-4193. https://doi.org/10.1109/TPAMI.2021.3060446

Wu KH, Jia F, Han YH, 2023. Domain-specific feature elimination: multi-source domain adaptation for image classification. *Front Comput Sci*, 17(4):174705. https://doi.org/10.1007/s11704-022-2146-x

Xiao N, Zhang L, 2021. Dynamic weighted learning for unsupervised domain adaptation. Proc IEEE/CVF Conf on Computer Vision and Pattern Recognition, p.15242-15251. https://doi.org/10.1109/CVPR46437.2021.01499

Xu RJ, Chen ZL, Zuo WM, et al., 2018. Deep cocktail network: multi-source unsupervised domain adaptation with category shift. Proc IEEE/CVF Conf on Computer Vision and Pattern Recognition, p.3964-3973. https://doi.org/10.1109/CVPR.2018.00417

Xu RJ, Li GB, Yang JH, et al., 2019. Larger norm more transferable: an adaptive feature norm approach for unsupervised domain adaptation. Proc IEEE/CVF Int Conf on Computer Vision, p.1426-1435. https://doi.org/10.1109/ICCV.2019.00151

Xu X, Zhang LY, 2021. Rectifying pseudo label by mutual disagreement learning for unsupervised domain adaptation person re-identification. Proc IEEE Int Conf on Multimedia and Expo, p.1-6. https://doi.org/10.1109/ICME51207.2021.9428307

Xu YY, Yan H, 2022. Cycle-reconstructive subspace learning with class discriminability for unsupervised domain adaptation. *Patt Recognit*, 129:108700. https://doi.org/10.1016/j.patcog.2022.108700

Yang SQ, van de Weijer J, Herranz L, et al., 2021. Exploiting the intrinsic neighborhood structure for source-free domain adaptation. Proc 35th Advances in Neural Information Processing Systems, p.29393-29405.

Yang YC, Soatto S, 2020. FDA: Fourier domain adaptation for semantic segmentation. Proc IEEE/CVF Conf on Computer Vision and Pattern Recognition, p.4085-4095. https://doi.org/10.1109/CVPR42600.2020.00414

Zhai YP, Lu SJ, Ye QX, et al., 2020. AD-Cluster: augmented discriminative clustering for domain adaptive person re-identification. Proc IEEE/CVF Conf on Computer Vision and Pattern Recognition, p.9021-9030. https://doi.org/10.1109/CVPR42600.2020.00904

Zhang H, Cao HH, Yang X, et al., 2021. Self-training with progressive representation enhancement for unsupervised cross-domain person re-identification. *IEEE Trans Image Process*, 30:5287-5298. https://doi.org/10.1109/TIP.2021.3082298

Zhang JY, Huang JX, Tian ZC, et al., 2022. Spectral unsupervised domain adaptation for visual recognition. Proc IEEE/CVF Conf on Computer Vision and Pattern Recognition, p.9829-9840. https://doi.org/10.1109/CVPR52688.2022.00960

Zhang MY, Liu K, Li YD, et al., 2021. Unsupervised domain adaptation for person re-identification via heterogeneous graph alignment. Proc 35th AAAI Conf on Artificial Intelligence, p.3360-3368. https://doi.org/10.1609/aaai.v35i4.16448

Zhang YB, Tang H, Jia K, et al., 2019. Domain-symmetric networks for adversarial domain adaptation. Proc IEEE/CVF Conf on Computer Vision and Pattern Recognition, p.5031-5040. https://doi.org/10.1109/CVPR.2019.00517

Zhang YB, Deng B, Tang H, et al., 2020. Unsupervised multi-class domain adaptation: theory, algorithms, and practice. *IEEE Trans Patt Anal Mach Intell*, 44(5):2775-2792. https://doi.org/10.1109/TPAMI.2020.3036956

Zhang YC, Liu TL, Long MS, et al., 2019. Bridging theory and algorithm for domain adaptation. Proc 36th Int Conf on Machine Learning, p.7404-7413.

Zhang Z, Wang YN, Liu S, et al., 2022. Cross-domain person re-identification using heterogeneous convolutional network. *IEEE Trans Circ Syst Video Technol*, 32(3):1160-1171. https://doi.org/10.1109/TCSVT.2021.3074745

Zhao F, Liao SC, Xie GS, et al., 2020. Unsupervised domain adaptation with noise resistible mutual-training for person re-identification. Proc 16th European Conf on Computer Vision, p.526-544. https://doi.org/10.1007/978-3-030-58621-8_31

Zhao SC, Li B, Xu PF, et al., 2021. MADAN: multi-source adversarial domain aggregation network for domain adaptation. *Int J Comput Vis*, 129(8):2399-2424. https://doi.org/10.1007/s11263-021-01479-3

Zheng L, Shen LY, Tian L, et al., 2015. Scalable person re-identification: a benchmark. Proc IEEE Int Conf on Computer Vision, p.1116-1124. https://doi.org/10.1109/ICCV.2015.133

Zhong Z, Zheng L, Luo ZM, et al., 2019. Invariance matters: exemplar memory for domain adaptive person re-identification. Proc IEEE/CVF Conf on Computer Vision and Pattern Recognition, p.598-607. https://doi.org/10.1109/CVPR.2019.00069

Zhong Z, Zheng L, Luo ZM, et al., 2021. Learning to adapt invariance in memory for person re-identification. *IEEE Trans Patt Anal Mach Intell*, 43(8):2723-2738. https://doi.org/10.1109/TPAMI.2020.2976933

Zhu YC, Zhuang FZ, Wang DQ, 2019. Aligning domain-specific distribution and classifier for cross-domain classification from multiple sources. Proc 33rd AAAI Conf on Artificial Intelligence, p.5989-5996. https://doi.org/10.1609/aaai.v33i01.33015989

Zhu YC, Zhuang FZ, Wang JD, et al., 2021. Deep subdomain adaptation network for image classification. *IEEE Trans Neur Netw Learn Syst*, 32(4):1713-1722. https://doi.org/10.1109/TNNLS.2020.2988928

Zuo YK, Yao HT, Zhuang LS, et al., 2022. Margin-based adversarial joint alignment domain adaptation. *IEEE Trans Circ Syst Video Technol*, 32(4):2057-2067. https://doi.org/10.1109/TCSVT.2021.3081729