SCIENCE CHINA
Information Sciences

• LETTER •

May 2026, Vol. 69, Iss. 5, 159101:1–159101:2
https://doi.org/10.1007/s11432-025-4783-y

Prototype-guided diffusion alignment for few-shot unsupervised domain adaptation

Heyang SUN, Chuanxing GENG & Songcan CHEN*

MIIT Key Laboratory of Pattern Analysis and Machine Intelligence, College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics (NUAA), Nanjing 211106, China

Received 29 May 2025/Revised 27 November 2025/Accepted 23 December 2025/Published online 23 March 2026

Citation Sun H Y, Geng C X, Chen S C. Prototype-guided diffusion alignment for few-shot unsupervised domain adaptation. Sci China Inf Sci, 2026, 69(5): 159101, https://doi.org/10.1007/s11432-025-4783-y

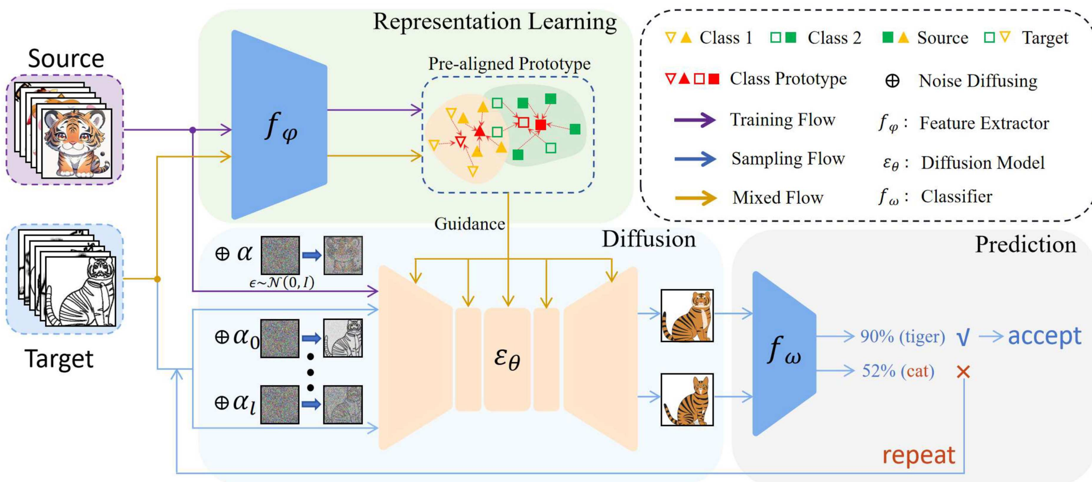
Deep neural network models achieve great and impressive successes on a variety of machine learning tasks. However, their performances always become frustrating once applied in the real-world scenarios where the test and training set no longer meet the independent and identically distributed (i.i.d.) hypothesis [1]. To cope with this problem, unsupervised domain adaptation (UDA) [1] as a popular paradigm is proposed to boost the prediction of the unlabeled target domain by transferring the knowledge from the labeled source domain.

Existing UDA methods usually rely on access to fully-annotated source data, which can hardly be guaranteed in some practical applications with sparse annotations due to data privacy protection. In these resource-constrained scenarios, it is difficult to train effective UDA models, let alone some large-scale models for domain adaptation. Therefore, few-shot unsupervised domain adaptation (FUDA) issues are gradually attracting attention and research [2–4]. Existing FUDA methods persistently strive to learn an effective discriminative space through various approaches such as self-supervised learning, thereby implicitly achieving distribution alignment between source and target domains. Although these approaches demonstrate promising progress by employing diverse auxiliary representation tasks to learn discriminative latent spaces for cross-domain knowledge transfer, they overlook a critical limitation: latent space learning is inherently an information selection process. In few-shot settings, such methods can hardly ensure that the latent space retains sufficient task-relevant semantic information for effective discrimination. In other words, due to the absence of sufficient supervisory information, the network prioritizes learning simple latent patterns, often at the cost of losing important discriminative semantics, which can lead to disastrous negative transfer.

To address this challenge, drawing inspiration from diffusion models [5], we propose a prototype-guided diffusion alignment (PGDA) method for FUDA, which aligns the source and target domain distributions in the original feature space while leveraging richer semantic information. By doing so, PGDA effectively mitigates negative transfer induced by unreliable latent representations. Specifically, we construct a diffusion alignment module that realizes domain alignment by transferring inputs into the noise distribution and then transforming the noise distribution back into the training distribution, where we embed the class prototypes as condition guidance to ensure the precision of the distribution transformation and the consistency of discriminative information. For the purpose of learning such effective guidance, we employ a representation learning module that selects high-confidence source domain samples to enrich the discriminative representations for prototype learning and performs coarse-grained cross-domain contrastive learning on the clustered target domain prototypes to bring the class prototypes closer together while avoiding negative transfer. Additionally, during the prediction phase, we design a confidence-based strategy to dynamically adjust the scale of the noise alignment, further ensuring accurate knowledge transfer. Extensive experiments on various benchmark datasets demonstrate the superiority of the proposed method. Please refer to Appendixes A–C for more details about motivation and related studies.

Methods. The method aligns feature representations in the original feature space with more comprehensive semantic information, thereby mitigating the negative transfer caused by inadequate latent representation, which consists of class-prototype learning and diffusion alignment, as shown in Figure 1(a). Specifically, class prototype learning employs a pre-trained feature extractor to obtain sample features, which are then stored in a memory bank via momentum updating. Category prototypes for the source and target domains are computed by selecting high-confidence samples based on cosine similarity. Building on this, prototype learning applies contrastive learning within each domain to encourage the features of each sample extracted by the model to align more closely with their respective class prototypes. Additionally, a spatial proximity learning loss is used to minimize the distance between each sample's corresponding class prototype and the class prototypes from the reverse domain, thereby capturing semantic invariance and spatial similarity. Diffusion alignment leverages the learned class prototypes to guide the sampling of target domain samples toward the source domain, achieving alignment between domains. Meanwhile, a pre-trained classifier serves as a supervisor, accepting high-confidence samples while applying larger noise steps to resample those with low confidence. Our proposed method performs alignment in the original feature space, effectively mitigating the semantic loss that may occur when aligning in the latent space

* Corresponding author (email: s.chen@nuaa.edu.cn)



(a)

Method	A→D	A→W	D→A	D→W	W→A	W→D	Avg
Source Only	31.3/49.0	19.6/43.3	41.3/55.7	58.7/81.8	39.9/49.8	61.9/81.7	42.1/60.2
CDS	52.6/65.1	55.2/68.8	65.7/71.2	76.6/88.1	59.7/71.0	73.3/87.3	63.9/75.3
PCS	60.2/78.2	69.8/82.9	<u>76.1</u>/76.4	90.6/94.1	71.2/76.3	91.8/96.0	76.6/84.0
C-VisDiT	<u>74.1</u>/**82.9**	<u>72.3</u>/<u>86.0</u>	75.7/<u>76.5</u>	**93.2**/**95.0**	<u>76.4</u>/<u>76.9</u>	<u>94.2</u>/**97.0**	<u>81.0</u>/<u>85.7</u>
Ours	**75.9**/<u>82.3</u>	**76.3**/**88.2**	**77.7**/**79.3**	<u>91.7</u>/<u>94.9</u>	**79.1**/**79.9**	**94.7**/<u>96.6</u>	**82.8**/**86.6**

(b)

Figure 1 (Color online) (a) Illustration of PGDA. The feature extractor f_φ is trained by the prototype learning to guide the knowledge transfer of the alignment module. The diffusion model ϵ_θ with U-Net architecture trained on the source domain is responsible for transforming the noise-added target domain distribution into the source domain distribution under the guidance of class prototypes, which is supervised by the classifier f_ω pre-trained on the source domain, accepting high-confidence samples while resampling low-confidence samples with an increased noise scale. (b) Adaptation accuracy (%) on 1-shot/3-shots labeled source per class on Office-31. The best and second-best results are highlighted in bold and underlined, respectively.

under label-scarce scenarios. Please refer to Appendix D for more detailed methods.

Experiment. We evaluate the performance of PGDA on datasets with various scales including the small-sized Office-31, the medium-sized Office-Home, and the large-sized VisDA-C and DomainNet. We conduct the source only method as the baseline, which is training on the labeled source images and classifying the target images. Then we compare our method with the related UDA and FUDA methods. Figure 1(b) presents the comparison of our proposed method with baselines on Office-31. It can be observed that our method demonstrates clear advantages over existing few-shot domain adaptation approaches, achieving optimal or sub-optimal results on most tasks. This validates the benefits of performing alignment in the original feature space under label-scarce scenarios. Additionally, further experiments on various datasets, ablation studies, parameter sensitivity analysis, and implementation details are provided in Appendix E.

Conclusion. In this study, we propose a PGDA method for FUDA, aiming to address the negative transfer problem caused by the loss of task-related semantic information during the process of latent space learning under sparsely labeled scenarios. The proposed PGDA method leverages class prototypes to guide the diffusion model in aligning the distribution discrepancy between the source and target domains across the entire space. In addition, PGDA employs a confidence-based strategy to adaptively select the noise scale, ensuring that each sample is aligned while preserving semantic consistency and reducing computational overhead. In addition, we offer a theoretical upper bound on the alignment error, offering insights into the design motivation of our method. Finally, extensive experiments on various datasets demonstrate the superiority of the proposed method.

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant Nos. 62376126, 62076124, 62576167) and Fundamental Research Funds for the Central Universities (Grant No. NS2024058).

Supporting information Appendixes A–F. The supporting information is available online at info.scichina.com and link.springer.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

References

1 Zhuang F, Qi Z, Duan K, et al. A comprehensive survey on transfer learning. Proc IEEE, 2021, 109: 43–76
2 Kim D, Saito K, Oh T H, et al. Cross-domain self-supervised learning for domain adaptation with few source labels. 2020. ArXiv:2003.08264
3 Yue X, Zheng Z, Zhang S, et al. Prototypical cross-domain self-supervised learning for few-shot unsupervised domain adaptation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021. 13834–13844
4 Xiong Y, Chen H, Lin Z, et al. Confidence-based visual dispersal for few-shot unsupervised domain adaptation. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2023. 11621–11631
5 Ho J, Salimans T. Classifier-free diffusion guidance. In: Proceedings of the NeurIPS 2021 Workshop on Deep Generative Models and Downstream Applications, 2021