Black-box Domain Adaptative Cell Segmentation via Multi-source Distillation

Xingguang Wang[1], Zhongyu Li[1(✉)], Xiangde Luo[2], Jing Wan[1], Jianwei Zhu[1], Ziqi Yang[1], Meng Yang[3], and Cunbao Xu[4(✉)]

[1] School of Software Engineering, Xian Jiaotong University, Xi'an, China
zhongyuli@xjtu.edu.cn
[2] School of Mechanical and Electrical Engineering, University of Electronic Science and Technology of China, Chengdu, China
[3] Hunan Frontline Medical Technology Co., Ltd., Changsha, China
[4] Department of Pathology, Quanzhou First Hospital Affiliated to Fujian Medical University, Quanzhou, China
937340447@qq.com

Abstract. Cell segmentation plays a critical role in diagnosing various cancers. Although deep learning techniques have been widely investigated, the enormous types and diverse appearances of histopathological cells still pose significant challenges for clinical applications. Moreover, data protection policies in different clinical centers and hospitals limit the training of data-dependent deep models. In this paper, we present a novel framework for cross-tissue domain adaptative cell segmentation without access both source domain data and model parameters, namely Multi-source Black-box Domain Adaptation (MBDA). Given the target domain data, our framework can achieve the cell segmentation based on knowledge distillation, by only using the outputs of models trained on multiple source domain data. Considering the domain shift cross different pathological tissues, predictions from the source models may not be reliable, where the noise labels can limit the training of the target model. To address this issue, we propose two practical approaches for weighting knowledge from the multi-source model predictions and filtering out noisy predictions. First, we assign pixel-level weights to the outputs of source models to reduce uncertainty during knowledge distillation. Second, we design a pseudo-label cutout and selection strategy for these predictions to facilitate the knowledge distillation from local cells to global pathological images. Experimental results on four types of pathological tissues demonstrate that our proposed black-box domain adaptation approach can achieve comparable and even better performance in comparison with state-of-the-art white-box approaches. The code and dataset are released at: https://github.com/NeuronXJTU/MBDA-CellSeg.

Keywords: Multi-source domain adaptation · Black-box model · Cell segmentation · Knowledge distillation

H. Greenspan et al. (Eds.): MICCAI 2023, LNCS 14220, pp. 749–758, 2023.
https://doi.org/10.1007/978-3-031-43907-0_71

1 Introduction

Semantic segmentation plays a vital role in pathological image analysis. It can help people conduct cell counting, cell morphology analysis, and tissue analysis, which reduces human labor [19]. However, data acquisition for medical images poses unique challenges due to privacy concerns and the high cost of manual annotation. Moreover, pathological images from different tissues or cancer types often show significant domain shifts, which hamper the generalization of models trained on one dataset to others. Due to the abovementioned challenges, some researchers have proposed various white-box domain adaptation methods to address these issues.

Recently, [8,16] propose to use generative adversarial networks to align the distributions of source and target domains and generate source-domain look-alike outputs for target images. Source-free domain adaptation methods have been also widely investigated due to the privacy protection. [3,5,14] explore how to implicitly align target domain data with the model trained on the source domain without accessing the source domain data. There are also many studies on multi-source white-box domain adaptation. Ahmed et al. [1] propose a novel algorithm which automatically identifies the optimal blend of source models to generate the target model by optimizing a specifically designed unsupervised loss. Li et al. [13] extend the above work to semantic segmentation and proposed a method named model-invariant feature learning, which takes full advantage of the diverse characteristics of the source-domain models.

Nonetheless, several recent investigations have demonstrated that the domain adaptation methods for source-free white-box models still present a privacy risk due to the potential leakage of model parameters [4]. Such privacy breaches may detrimental to the privacy protection policies of hospitals. Moreover, the target domain uses the same neural network as the source domain, which is not desirable for low-resource target users like hospitals [15]. We thus present a more challenging task of relying solely on black-box models from vendors to avoid parameter leakage. In clinical applications, various vendors can offer output interfaces for different pathological images. While black-box models are proficient in specific domains, their performances greatly degrade when the target domain is updated with new pathology slices. Therefore, how to leverage the existing knowledge of black-box models to effectively train new models for the target domain without accessing the source domain data remains a critical challenge.

In this paper, we present a novel source-free domain adaptation framework for cross-tissue cell segmentation without accessing both source domain data and model parameters, which can seamlessly integrate heterogeneous models from different source domains into any cell segmentation network with high generality. To the best of our knowledge, this is the first study on the exploration of multi-source black-box domain adaptation for cross-tissue cell segmentation. In this setting, conventional multi-source ensemble methods are not applicable due to the unavailability of model parameters, and simply aggregating the black-box outputs would introduce a considerable amount of noise, which can be detrimental to the training of the target domain model. Therefore, we develop two

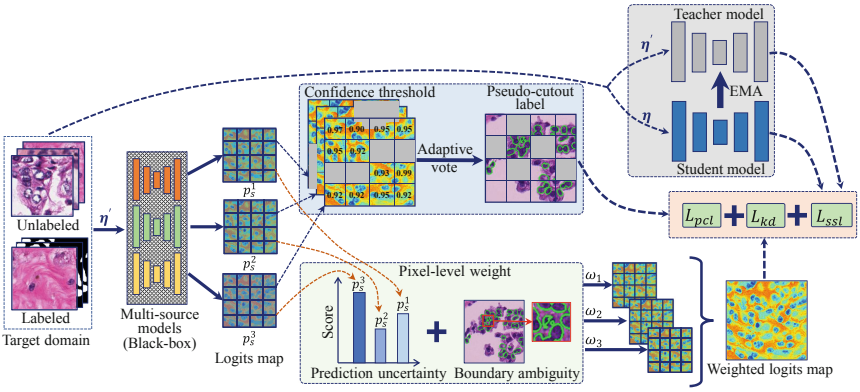

Fig. 1. Overview of our purposed framework, where logits maps denote the raw predictions from source models and ω denotes pixel-level weight for each prediction. The semi-supervised loss, denoted as L_{ssl}, encompasses the supervised loss, consistency loss, and maximize mutual information loss.

strategies within this new framework to address this issue. Firstly, we propose a pixel-level multi-source domain weighting method, which reduces source domain noise by knowledge weighting. This method effectively addresses two significant challenges encountered in the analysis of cellular images, namely, the uncertainty in source domain output and the ambiguity in cell boundary semantics. Secondly, we also take into account the structured information from cells to images, which may be overlooked during distillation, and design an adaptive knowledge voting strategy. This strategy enables us to ignore low-confidence regions, similar to Cutout [6], but with selective masking of pixels, which effectively balances the trade-off between exploiting similarities and preserving differences of different domains. As a result, we refer to the labels generated through the voting strategy as pseudo-cutout labels.

2 Method

Overview: Figure 1 shows a binary cell segmentation task with three source models trained on different tissues and a target model, i.e., the student model in Fig. 1. We only use the source models' predictions on the target data for knowledge transfer without accessing the source data and parameters. The η and η' indicate that different perturbations are added to the target images. Subsequently, we feed the perturbed images into the source domain predictor to generate the corresponding raw segmentation outputs. These outputs are then processed by two main components of our framework: a pixel-level weighting method that takes into account the prediction uncertainty and cell boundary ambiguity, and an adaptive knowledge voter that utilizes confidence gates and a dynamic ensemble strategy. These components we designed are to extract reliable knowledge from the predictions of source domain models and reduce noise

during distillation. Finally, we obtain a weighted logit for knowledge distillation from pixel level and a high-confidence pseudo-cutout label for further structured distillation from cell to global pathological image.

Knowledge Distillation by Weighted Logits Map: We denote $\mathcal{D}_S^N = \{X_s, Y_s\}_N$ as a collection of N source domains and $\mathcal{D}_T = \{X_t^i, Y_t^j\}$ as single target domain, where the number of labeled instances $Y_t^j \ll X_t^i$. We are only provided with black-box models $\{f_s^n\}_{n=1}^N$ trained on multiple source domains $\{x_s^i, y_s^i\}_{n=1}^N$ for knowledge transfer. The parameters $\{\theta_s^n\}_{n=1}^N$ of these source domain predictors are not allowed to participate in gradient backpropagation as a result of the privacy policy. Thus, our ultimate objective is to derive a novel student model $f_t : X_t \to Y_t$ that is relevant to the source domain task. Accordingly, direct knowledge transfer using the output of the source domain predictor may lead to feature bias in the student model due to the unavoidable covariance [20] between the target and source domains. Inspired by [21], we incorporate prediction uncertainty and cell boundary impurity to establish pixel-level weights for multi-source outputs. We assume that *k-square-neighbors* of a pixel as a cell region, i.e., for a logits map with height H and width W, we define the region as follow:

$$\mathcal{N}_k\{(i,j) \mid (i,j) \in (H,W)\} = \{(u,v) \mid |u-i| \leq k, |v-j| \leq k\} \tag{1}$$

where (i,j) denotes centre of region, and k denotes the size of *k-square-neighbors*.

Firstly, we develop a pixel-level predictive uncertainty algorithm to aid in assessing the correlation between multiple source domains and the target domain. For a given target image $x_t \in X_t^i$, we initially feed it into the source predictors $\{f_s^n\}_{n=1}^N$ to obtain their respective prediction $\{p_s^n\}_{n=1}^N$. To leverage the rich semantic information from the source domain predictor predictions, we utilize predictive entropy of the softmax outputs to measure the prediction uncertainty scores. In the semantic segmentation scenario of C-classes classification, we define the pixel-level uncertainty score $\mathcal{U}_n^{(i,j)}$ as follow:

$$\mathcal{U}_n^{(i,j)} = -\sum_{c=1}^{C} O_s^{n(i,j,c)} \log O_s^{n(i,j,c)} \tag{2}$$

where O_s^n denotes softmax output,i.e.,$O_s^n = softmax(p_s^n)$ from nth source predictor.

Due to the unique characteristics of cell morphology, merely relying on uncertainty information is insufficient to produce high-quality ensemble logits map that accurately capture the relevance between the source and target domains. The target pseudo-label for the nth predictor f_s^n can be obtained by applying the softmax function to the output and selecting the category with the highest probability score, i.e., $\widehat{Y}_t = \arg\max_{c \in \{1,\ldots,C\}}(softmax(p_s^n))$. Then according to C-classes classification tasks, we divide the cell region into C subsets, $\mathcal{N}_k^c(i,j) = \{(u,v) \in \mathcal{N}_k(i,j) \mid \widehat{Y}_t = c\}$. After that, we determine the degree of

impurity in an area of interest by analyzing the statistics of the boundary region, which represents the level of semantic information ambiguity. Specifically, the number of different objects within the area is considered a proxy for its impurity level, with higher counts indicating higher impurity.The boundary impurity $\mathcal{P}^{(i,j)}$ can be calculated as:

$$\mathcal{P}_n^{(i,j)} = -\sum_{c=1}^{C} \frac{|\mathcal{N}_k^c(i,j)|}{|\mathcal{N}_k(i,j)|} \log \frac{|\mathcal{N}_k^c(i,j)|}{|\mathcal{N}_k(i,j)|} \tag{3}$$

where $|\cdot|$ denotes the number of pixels in the area.

By assigning lower weights to the pixels with high uncertainty and boundary ambiguity, we can obtain pixel-level weight scores $\mathcal{W}^n$ for each p_s^n, i.e.,

$$\mathcal{W}^n = -\log\left(\frac{\exp\left(\mathcal{U}_n \odot \mathcal{P}_n\right)}{\sum_{n=1}^{N} \exp\left(\mathcal{U}_n \odot \mathcal{P}_n\right)}\right) \tag{4}$$

where $\odot$ denotes element-wise matrix multiplication. According to the pixel-level weight, we will obtain an ensemble logits map $\mathcal{M} = \sum_{n=1}^{N} \mathcal{W}^n \cdot p_s^n$. And the object of the knowledge distillation is a classical regularization term [9]:

$$\mathcal{L}_{kd}(f_t; X_t, \mathcal{M}) = \mathbb{E}_{x_t \in X_t} \mathcal{D}_{kl}(\mathcal{M} \mid\mid f_t(x_t)) \tag{5}$$

where D_{kl} denotes the Kullback-Leibler (KL) divergence loss.

Adaptive Pseudo-Cutout Label: As previously mentioned, the outputs from the source domain black-box predictors have been adjusted by the pixel-level weight. However, they are still noisy and only pixel-level information is considered while ignoring structured information in the knowledge distillation process. Thus, we utilize the output of the black-box predictor on the target domain to produce an adaptive pseudo-cutout label, which will be employed to further regularize the knowledge distillation process. We have revised the method in [7] to generate high-quality pseudo labels that resemble the Cutout augmentation technique. For softmax outputs $\{O_s^n\}_{n=1}^N$ from N source predictors, we first set a threshold α to filter low-confidence pixels. To handle pixels with normalized probability values below the threshold, we employ a Cutout-like operation and discard these pixels. Subsequently, we apply an adaptive voting strategy to the N source domain outputs. Initially, during the training of the target model, if at least one source domain output exceeds the threshold, we consider the pixel as a positive or negative sample, which facilitates rapid knowledge acquisition by the model. As the training progresses, we gradually tighten the voting strategy and only retain regional pixels that have received adequate votes. The strategy can be summarised as follow:

$$\mathcal{V}_n^{((i,j) \mid (i,j) \in (H,W))} = \begin{cases} 1, & O_s^n(i,j) > \alpha, \\ 0, & \text{otherwise.} \end{cases} \tag{6}$$

where α is empirically set as 0.9.

Then we will aggregate the voting scores, i.e., $V^{(i,j)} = \sum_{n=1}^{N} \mathcal{V}_n^{(i,j)}$ and determine whether to retain each pixel using an adaptive vote gate $G \in \{1, 2, 3, etc.\}$. By filtering with a threshold and integrating the voting strategy, we generate high-confidence pseudo-labels that remain effective even when the source and target domains exhibit covariance. Finally, we define the ensemble result as a pseudo-cutout label $\hat{P}_s$ and employ consistency regularization as below:

$$\mathcal{L}_{pcl}(f_t; X_t, \hat{P}_s) = \mathbb{E}_{x_t \in X_t} l_{ce}(\hat{P}_s \mid\mid f_t(x_t)) \tag{7}$$

where l_{ce} denotes cross-entropy loss function.

Loss Functions: Finally, we incorporate global structural information about the predicted outcome of the target domain into both distillation and semi-supervised learning. To mitigate the noise effect of the source domain predictors, we introduce maximize mutual information targets to facilitate discrete representation learning by the network. We define $E(p) = -\sum_i p_i \log p_i$ as conditional entropy. The object can be described as follow:

$$\begin{aligned} \mathcal{L}_{mmi}(f_t; X_t) &= H(Y_t) - H(Y_t|X_t) \\ &= E\left(\mathbb{E}_{x_t \in X_t} f_t(x_t)\right) - \mathbb{E}_{x_t \in X_t} E\left(f_t(x_t)\right), \end{aligned} \tag{8}$$

where the increasing $H(Y_t)$ and the decreasing $H(Y_t|X_t)$ help to balances class separation and classifier complexity [15].

We adopt the classical and effective mean-teacher framework as a baseline for semi-supervised learning and update the teacher model parameters by exponential moving average. Also, we apply two different perturbations (η, η') to the target domain data and feed them into the student model and the mean-teacher model respectively. The consistency loss of unsupervised learning can be defined as below:

$$\mathcal{L}_{cons}(\theta_t, \theta_t') = \mathbb{E}_{x_t \in X_t}\left[||f_t(x_t, \theta_t', \eta') - f_t(x_t, \theta_t, \eta)||^2\right] \tag{9}$$

Finally, we get the overall objective:

$$\mathcal{L}_{all} = \mathcal{L}_{kd} + \mathcal{L}_{pcl} + \mathcal{L}_{cons} - \mathcal{L}_{mmi} + \mathcal{L}_{sup} \tag{10}$$

where $\mathcal{L}_{sup}$ denotes the ordinary cross-entropy loss for supervised learning and we set the weight of each loss function to 1 in the training.

3 Experiments

Dataset and Setting: We collect four pathology image datasets to validate our proposed approach. Firstly, we acquire 50 images from a cohort of patients with Triple Negative Breast Cancer (TNBC), which is released by Naylor et al [18]. Hou et al. [10] publish a dataset of nucleus segmentation containing 5,060 segmented slides from 10 TCGA cancer types. In this work, we use 98 images from

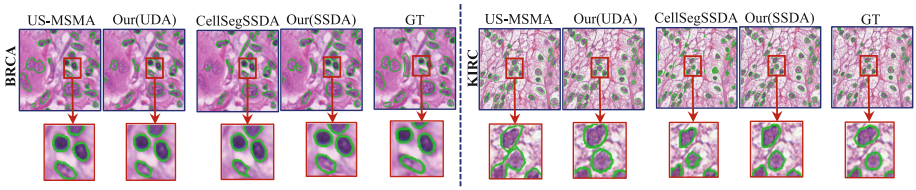

Fig. 2. Visualized segmentation on the BRCA and KIRC target domains respectively.

invasive carcinoma of the breast (BRCA). We have also included 463 images of Kidney Renal Clear cell carcinoma (KIRC) in our dataset, which are made publicly available by Irshad et al [11]. Awan et al. [2] publicly release a dataset containing tissue slide images and associated clinical data on colorectal cancer (CRC), from which we randomly select 200 patches for our study. In our experiments, we transfer knowledge from three black-box models trained on different source domains to a new target domain model (e.g.,from CRC, TNBC, KIRC to BRCA). The backbone network for the student model and source domain black-box predictors employ the widely adopted residual U-Net [12], which is commonly used for medical image segmentation. For each source domain network, we conduct full-supervision training on the corresponding source domain data and directly evaluate its performance on target domain data. The upper performance metrics (Source-only upper) are shown in the Table 1. To ensure the reliability of the results, we use the same data for training, validation, and testing, which account for 80%, 10%, and 10% of the original data respectively. For the target domain network, we use unsupervised and semi-supervised as our task settings respectively. In semi-supervised domain adaptation, we only use 10% of the target domain data as labeled data.

Experimental Results: To validate our method, we compare it with the following approaches: (1) CellSegSSDA [8], an adversarial learning based semi-supervised domain adaptation approach. (2) US-MSMA [13], a multi-source model domain aggregation network. (3) SFDA-DPL [5], a source-free unsupervised domain adaptation approach. (4) BBUDA [17], an unsupervised black-box model domain adaptation framework. A point worth noting is that most of the methods we compared with are white-box methods, which means they can obtain more information from the source domain than us. For single-source domain adaptation approach, CellsegSSDA and SFDA-DPL, we employ two strategies to ensure the fairness of the experiments: (1) single-source, i.e. performing adaptation on each single source, where we select the best results to display in the Table 1; (2) source-combined, i.e. all source domains are combined into a traditional single source. The Table 1 and Fig. 2 demonstrate that our proposed method exhibits superior performance, even when compared to these white-box methods, surpassing them in various evaluation metrics and visualization results. In addition, the experimental results also show that simply combining multiple

Table 1. Quantitative comparison with unsupervised and semi-supervised domain adaptation methods under 3 segmentation metrics.

Source		CRC&TNBC&KIRC			CRC&TNBC&BRCA		
Target		BRCA			KIRC		
Standards	Methods	Dice	HD95	ASSD	Dice	HD95	ASSD
Source-only	Source(upper)	0.6991	41.9604	10.8780	0.7001	34.5575	6.7822
Single-source(upper)	SFDA-DPL [5]	0.6327	43.8113	11.6313	0.6383	26.3252	4.6023
	BBUDA [17]	0.6620	39.6950	11.3911	0.6836	42.9875	7.4398
Source-Combined	SFDA-DPL [5]	0.6828	46.5393	12.1484	0.6446	25.4274	4.2998
	BBUDA [17]	0.6729	41.8879	11.5375	0.6895	46.7358	8.7463
Multi-source	US-MSMA [13]	0.7334	**37.1309**	8.7817	0.7161	**18.7093**	**3.0187**
Multi-source	Our(UDA)	**0.7351**	39.4103	**8.7014**	**0.7281**	30.9221	6.2080
Single-source(upper)	CellSegSSDA [8]	0.6852	45.2595	9.9133	0.6937	58.7221	12.5176
Source-Combined	CellSegSSDA [8]	0.7202	43.9251	**8.0944**	0.6699	55.1768	10.3623
Multi-source	Our(SSDA)	**0.7565**	**39.0552**	9.3346	**0.7443**	**31.7582**	**6.0873**
fully-supervised upper bounds		0.7721	35.1449	7.2848	0.7540	23.53767	4.1882

Table 2. Ablation study of three modules in our proposed method.

CRC&KIRC&BRCA to TNBC					
WL	PCL	MMI	Dice	HD95	ASSD
×	×	×	0.6708	56.9111	16.3837
✓	×	×	0.6822	**54.3386**	14.9817
✓	✓	×	0.6890	57.0889	12.9512
✓	✓	✓	**0.7075**	58.8798	**10.7247**

source data into a traditional single source will result in performance degradation in some cases, which also proves the importance of studying multi-source domain adaptation methods.

Ablation Study: To evaluate the impact of our proposed methods of weighted logits(WL), pseudo-cutout label(PCL) and maximize mutual information(MMI) on the model performance, we conduct an ablation study. We compare the baseline model with the models that added these three methods separately. We chose CRC, KIRC and BRCA as our source domains, and TNBC as our target domain. The results of these experiments, presented in the Table 2, show that our proposed modules are indeed useful.

4 Conclusion

Our proposed multi-source black-box domain adaptation method achieves competitive performance by solely relying on the source domain outputs, without the need for access to the source domain data or models, thus avoiding information leakage from the source domain. Additionally, the method does not assume

the same architecture across domains, allowing us to learn lightweight target models from large source models, improving learning efficiency. We demonstrate the effectiveness of our method on multiple public datasets and believe it can be readily applied to other domains and adaptation scenarios. Moving forward, we plan to integrate our approach with active learning methods to enhance annotation efficiency in the semi-supervised setting. By leveraging multi-source domain knowledge, we aim to improve the reliability of the target model and enable more efficient annotation for better model performance.

Acknowledgements. This work is partially supported by the National Natural Science Foundation of China under grant No. 61902310 and the Natural Science Basic Research Program of Shaanxi, China under grant 2020JQ030.

References

1. Ahmed, S.M., Raychaudhuri, D.S., Paul, S., Oymak, S., Roy-Chowdhury, A.K.: 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 (2021)
2. Awan, R., et al.: Glandular morphometrics for objective grading of colorectal adenocarcinoma histology images. Sci. Rep. **7**(1), 16852 (2017)
3. Bateson, M., Kervadec, H., Dolz, J., Lombaert, H., Ben Ayed, I.: Source-relaxed domain adaptation for image segmentation. In: Martel, A.L., et al. (eds.) MICCAI 2020. LNCS, vol. 12261, pp. 490–499. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-59710-8_48
4. Carlini, N., et al.: Extracting training data from large language models. In: USENIX Security Symposium, vol. 6 (2021)
5. Chen, C., Liu, Q., Jin, Y., Dou, Q., Heng, P.-A.: Source-free domain adaptive fundus image segmentation with denoised pseudo-labeling. In: de Bruijne, M., et al. (eds.) MICCAI 2021. LNCS, vol. 12905, pp. 225–235. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-87240-3_22
6. DeVries, T., Taylor, G.W.: Improved regularization of convolutional neural networks with cutout. arXiv preprint arXiv:1708.04552 (2017)
7. Feng, H., et al.: Kd3a: unsupervised multi-source decentralized domain adaptation via knowledge distillation. In: ICML, pp. 3274–3283 (2021)
8. Haq, M.M., Huang, J.: Adversarial domain adaptation for cell segmentation. In: Medical Imaging with Deep Learning, pp. 277–287. PMLR (2020)
9. Hinton, G., Vinyals, O., Dean, J.: Distilling the knowledge in a neural network. arXiv preprint arXiv:1503.02531 (2015)
10. Hou, L., et al.: Dataset of segmented nuclei in hematoxylin and eosin stained histopathology images of ten cancer types. Sci. Data **7**(1), 185 (2020)
11. Irshad, H., et al.: Crowdsourcing image annotation for nucleus detection and segmentation in computational pathology: evaluating experts, automated methods, and the crowd. In: Pacific Symposium on Biocomputing Co-chairs, pp. 294–305. World Scientific (2014)
12. Kerfoot, E., Clough, J., Oksuz, I., Lee, J., King, A.P., Schnabel, J.A.: Left-ventricle quantification using residual U-Net. In: Pop, M., et al. (eds.) STACOM 2018. LNCS, vol. 11395, pp. 371–380. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-12029-0_40

13. Li, Z., Togo, R., Ogawa, T., Haseyama, M.: Union-set multi-source model adaptation for semantic segmentation. In: Computer Vision-ECCV 2022: 17th European Conference, Tel Aviv, Israel, 23–27 October 2022, Proceedings, Part XXIX, pp. 579–595. Springer, Heidelberg (2022). https://doi.org/10.1007/978-3-031-19818-2_33
14. Liang, J., Hu, D., Feng, J.: 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 (2020)
15. Liang, J., Hu, D., Feng, J., He, R.: 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 (2022)
16. Liu, X., et al.: Adversarial unsupervised domain adaptation with conditional and label shift: Infer, align and iterate. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 10367–10376 (2021)
17. Liu, X., et al.: Unsupervised black-box model domain adaptation for brain tumor segmentation. Front. Neurosci. **16**, 837646 (2022)
18. Naylor, P., Laé, M., Reyal, F., Walter, T.: Segmentation of nuclei in histopathology images by deep regression of the distance map. IEEE Trans. Med. Imaging **38**(2), 448–459 (2018)
19. Scherr, T., Löffler, K., Böhland, M., Mikut, R.: Cell segmentation and tracking using cnn-based distance predictions and a graph-based matching strategy. Plos One **15**(12), e0243219 (2020)
20. Shimodaira, H.: Improving predictive inference under covariate shift by weighting the log-likelihood function. J. Stat. Plan. Infer. **90**(2), 227–244 (2000)
21. Xie, B., Yuan, L., Li, S., Liu, C.H., Cheng, X.: Towards fewer annotations: active learning via region impurity and prediction uncertainty for domain adaptive semantic segmentation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 8068–8078 (2022)