



Multi-Target Domain Adaptation with Prompt Learning for Medical Image Segmentation

Yili Lin¹, Dong Nie³, Yuting Liu², Ming Yang², Daoqiang Zhang¹,
and Xuyun Wen¹(✉)

¹ Department of Computer Science and Technology,
Nanjing University of Aeronautics and Astronautics, Nanjing, China
wenzuyun@nuaa.edu.cn

² Department of Radiology,
Children's Hospital of Nanjing Medical University, Nanjing, China

³ Alibaba Inc, El Monte, USA

Abstract. Domain shift is a big challenge when deploying deep learning models in real-world applications due to various data distributions. The recent advances of domain adaptation mainly come from explicitly learning domain invariant features (e.g., by adversarial learning, metric learning and self-training). While they cannot be easily extended to multi-domains due to the diverse domain knowledge. In this paper, we present a novel multi-target domain adaptation (MTDA) algorithm, i.e., prompt-DA, through implicit feature adaptation for medical image segmentation. In particular, we build a feature transfer module by simply obtaining the domain-specific prompts and utilizing them to generate the domain-aware image features via a specially designed simple feature fusion module. Moreover, the proposed prompt-DA is compatible with the previous DA methods (e.g., adversarial learning based) and the performance can be continuously improved. The proposed method is evaluated on two challenging domain-shift datasets, i.e., the Iseg2019 (domain shift in infant MRI of different ages), and the BraTS2018 dataset (domain shift between high-grade and low-grade gliomas). Experimental results indicate our proposed method achieves state-of-the-art performance in both cases, and also demonstrates the effectiveness of the proposed prompt-DA. The experiments with adversarial learning DA show our proposed prompt-DA can go well with other DA methods. Our code is available at <https://github.com/MurasakiLin/prompt-DA>.

Keywords: Domain Adaptation · Prompt Learning · Segmentation

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/978-3-031-43907-0_68.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023
H. Greenspan et al. (Eds.): MICCAI 2023, LNCS 14220, pp. 717–727, 2023.
https://doi.org/10.1007/978-3-031-43907-0_68

1 Introduction

Deep learning has brought medical image segmentation into the era of data-driven approaches, and has made significant progress in this field [1, 2], i.e., the segmentation accuracy has improved considerably. In spite of the huge success, the deployment of trained segmentation models is often severely impacted by a distribution shift between the training (or labeled) and test (or unlabeled) data since the segmentation performance will deteriorate greatly in such situations. Domain shift is typically caused by various factors, including differences in acquisition protocols (e.g., age, gender, the severity of the disease and so on).

Domain adaptation (DA) has been proposed and investigated to combat distribution shift in medical image segmentation. Many researchers proposed using adversarial learning to tackle distribution shift problems [3–7]. These methods mainly use the game between the domain classifier and the feature extractor to learn domain-invariant features. However, they easily suffer from the balance between feature alignment and discrimination ability of the model. Some recent researchers begin to explore self-training based DA algorithms, which generate pseudo labels for the ‘other’ domain samples to fulfill self-training [8–11]. While it is very difficult to ensure the quality of pseudo labels in the ‘other’ domain and is also hard to build capable models with noise labels. However, most of these methods cannot well handle the situation when the domains are very diverse, since it is very challenging to learn domain-invariant features when each domain contains domain-specific knowledge. Also, the domain information itself is well utilized in the DA algorithms.

To tackle the aforementioned issues, we propose utilizing prompt learning to take full advantage of domain information. Prompt learning [12, 13] is a recently emergent strategy to extend the same natural language processing (NLP) model to different tasks without re-training. Prompt learning models can autonomously tune themselves for different tasks by transferring domain knowledge introduced through prompts, and they can usually demonstrate better generalization ability across many downstream tasks. very few works have attempted to apply prompt learning to the computer vision field, and have achieved promising results. [14] introduced prompt tuning as an efficient and effective alternative to full fine-tuning for large-scale Transformer models. [15] exploited prompt learning to fulfill domain generalization in image classification tasks. The prompts in these models are generated and used in the very early stage of the models, which prevents the smooth combination with other domain adaptation methods.

In this paper, we introduce a domain prompt learning method (prompt-DA) to tackle distribution shift in multi-target domains. Different from the recent prompt learning methods, we generate domain-specific prompts in the encoding feature space instead of the image space. As a consequence, it can improve the quality of the domain prompts, more importantly, we can easily consolidate the prompt learning with the other DA methods, for instance, adversarial learning based DA. In addition, we propose a specially designed fusion module to reinforce the respective characteristics of the encoder features and domain-

specific prompts, and thus generate domain-aware features. As a way to prove the prompt-DA is compatible with other DAs, a very simple adversarial learning module is jointly adopted in our method to further enhance the model’s generalization ability (we denote this model as comb-DA). We evaluate our proposed method on two multi-domain datasets: 1). the infant brain MRI dataset for cross-age segmentation; 2). the BraTS2018 dataset for cross-grade tumor segmentation. Experiments show our proposed method outperforms state-of-the-art methods. Moreover, ablation study demonstrates the effectiveness of the proposed domain prompt learning and the feature fusion module. Our claim about the successful combination of prompt learning with adversarial learning is also well-supported by experiments.

2 Methodology

Our proposed prompt-DA network consists of three main components as depicted in Fig. 1(a): a typical encoder-decoder network (e.g., UNet) serving the segmentation baseline, a prompt learning network to learn domain-specific prompts, and a fusion module aggregating the image features and domain-specific prompts to build domain-aware feature representation, where the fused features are fed into the decoder. It is worth noting that our proposed method is compatible with other DA algorithms, and thus we can add an optional extra DA module to further optimize the domain generalization ability, in this paper, we choose an adversarial learning based DA as an example since it is the mostly used DA methods in medical image segmentation (as introduced in Sect. 1).

There are various encoder-decoder segmentation networks, many of which are well known. As a result, we donot introduce the details of the encoder-decoder and just choose two typical networks to work as the segmentation backbone, that is, 3D-UNet [16] (convolution based and 3D) and TransUNet [17] (transformer based and 2D). In the following subsections, our focus will be on domain-specific prompt generation and domain-aware feature learning with feature fusion.

2.1 Learning Domain-Specific Prompts

In our designed prompt learning based DA method, it is essential to learn domain-specific prompts. Moreover, the quality of generated prompts directly determines the domain-aware features. Therefore, we specially designed a prompt generation module to learn domain-specific prompts which mainly consists of two components, i.e., a classifier and a prompt generator.

Our approach incorporates domain-specific information into the prompts to guide the model in adapting to the target domain. To achieve this, we introduce a classifier $h(x)$ that distinguishes the domain (denoted as \hat{d}) of the input image, shown in Eq. 1.

$$\hat{d} = h(x) \quad (1)$$

where x is the image or abstracted features from the encoder.

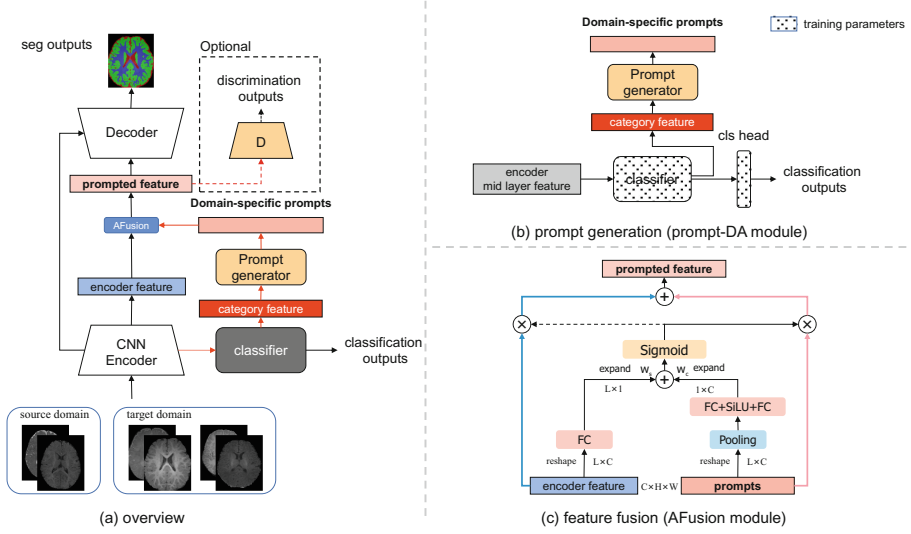


Fig. 1. a). Overview of the proposed prompt-DA network; b). The prompt generation module to learn domain-specific prompt; c). The feature fusion to learn domain-aware features.

To optimize the parameters, we adopt cross-entropy loss to train the classifier, as shown in Eq. 2.

$$L_{cls}(\hat{d}, d) = L_{ce}(\hat{d}, d) = - \sum_{i=1}^C d^{(i)} \log \hat{d}^{(i)} \quad (2)$$

where \hat{d} is the predicted domain information, and d is the ground truth domain information.

Prompt Generation: Instead of directly using \hat{d} as the category information, we fed the second-to-last layer’s features (i.e., z) of the classifier, namely, $g(z)$. In particular, the $g(z)$ is a multi-layer-perception, as defined in Eq. 3.

$$g(z) = \phi_3(\phi_2(\phi_1(z))) \quad (3)$$

where ϕ can be a Conv+BN+ReLU sequence. Note this module does not change the size of the feature map, instead, it transforms the extracted category features into domain-specific prompts.

2.2 Learning Domain-Aware Representation by Fusion

The learned prompt captures clearly about a certain domain and the features from the encoder describe the semantics as well as spatial information for the images. We can combine them to adapt the image features to domain-aware representations.

Basically, suppose we have an image denoted as I , and the prompt encodings for the domain knowledge is $g(e(I))$ (where $e(I)$ is the features from a shallow layer), $E(I)$ is the encoder features for this image. Then the domain-aware features (i.e., F) are extracted by a fusion module as Eq. 4.

$$F = \psi(g(e(I)), E(I)) \quad (4)$$

As the learned prompt and encoder feature capture quite different aspects of the input data, we cannot achieve good effect by simply using addition, multiplication or concatenation to serve as the fusion function ψ . Specifically, while the encoder feature emphasizes spatial information for image segmentation, the prompt feature highlights inter-channel information for domain-related characteristics. To account for these differences, we propose a simple attention-based fusion (denoted as AFusion) module to smoothly aggregate the information. This module computes channel-wise and spatial-wise weights separately to enhance both the channel and spatial characteristics of the input. Figure 1(c) illustrates the structure of our proposed module.

Our module utilizes both channel and spatial branches to obtain weights for two input sources. The spatial branch compresses the encoder feature in the channel dimension using an FC layer to obtain spatial weights. Meanwhile, the channel branch uses global average pooling and two FC layers to compress the prompt and obtain channel weights. We utilize FC layers for compression and rescaling, denoted as f_{cp} and f_{re} respectively. The spatial and channel weights are computed according to Eq. 5.

$$\begin{aligned} \mathcal{W}_s &= f_{cp}(E(I)), \\ \mathcal{W}_c &= f_{re}(f_{cp}(avgpool(g(e(I)))))) \end{aligned} \quad (5)$$

Afterward, we combine the weights from the spatial and channel dimensions to obtain a token that can learn both high-level and low-level features from both the encoder feature and the prompt, which guides the fusion of the two features. The process is illustrated as follows:

$$\begin{aligned} \mathcal{W} &= sigmoid(\mathcal{W}_c + \mathcal{W}_s), \\ \mathcal{F}_{out} &= g(e(I)) * \mathcal{W} + E(I) * (1 - \mathcal{W}) \end{aligned} \quad (6)$$

This module introduces only a few parameters, yet it can effectively improve the quality of the prompted domain-aware features after feature fusion. In the experimental section (i.e., Sect. 3.3), we conducted relevant experiments to verify that this module can indeed improve the performance of our prompt-DA method.

2.3 Adversarial Learning to Enhance the Generalization Ability

As aforementioned, our proposed prompt-DA is fully compatible with other DA algorithms. We thus use adversarial learning, which is widely adopted in medical image DA, to work as an optional component in our network to continuously enhance the domain adaptation ability.

Specially, inspired by the adversarial DA in [18], we adopt the classic GAN loss to train the discriminator and prompt generator (Note the adversarial loss, L_{adv} , for the generator will only be propagated to the prompt generator).

2.4 Total Loss

To optimize the segmentation backbone network, we use a combined loss function, L_{seg} , that incorporates both dice loss [19] and cross-entropy loss with a balance factor.

By summing the above-introduced losses, the total loss to train the segmentation network can be defined by Eq. 7.

$$L_{total} = L_{seg} + \lambda_{cls}L_{cls} + \lambda_{adv}L_{adv}, \quad (7)$$

where λ is the scaling factor to balance the losses.

2.5 Implementation Details

We use basic 3D-UNet [16] or TransUnet [17] as our segmentation network. We use a fully convolutional neural network consisting of four convolutional layers with 3×3 kernels and stride of 1 as the domain classifier, with each convolution layer followed by a ReLU parameterized by 0.2. We used three convolutional layers with ReLU activation function as the Prompt Generator and constructed a Discriminator with a similar structure to the Classifier.

We adopt Adam as the optimizer and set the learning rate to 0.0002 and 0.002 for the segmentation and domain classifier, respectively. The learning rate will be decayed by 0.1 every quarter of the training process.

3 Experiments and Results

3.1 Datasets

Our proposed method was evaluated using two medical image segmentation DA datasets. The first dataset, i.e., cross-age infant segmentation [20], was used for cross-age infant brain image segmentation, while the second dataset, i.e., Brats2018 [21], was used for HGG to LGG domain adaptation.

The first dataset is for infant brain segmentation (white matter, gray matter and cerebrospinal fluid). To build the cross-age dataset, we take advantage 10 brain MRIs of 6-month-old from iSeg2019 [20], and also build 3-month-old and 12-month-old in-house datasets. In this dataset, we collect 11 brain MRI for both the 3-month-old and 12-month-old infants. We take the 6-month-old data as the source domain, the 3-month-old and 12-month-old as the target domains.

The 2nd dataset is for brain tumor segmentation (enhancing tumor, peritumoral edema and necrotic and non-enhancing tumor core), which has 285 MRI samples (210 HGG and 75 LGG). We take HGG as the source domain and LGG as the target domain.

Table 1. Comparison with SOTA DA methods on the infant brain segmentation task. The evaluation metric shown is DICE.

Method	6 month				12 month				3 month			
	WM	GM	CSF	avg.	WM	GM	CSF	avg.	WM	GM	CSF	avg.
no-DA	82.47	88.57	93.84	88.29	68.58	72.37	76.45	72.47	69.29	61.92	62.84	64.68
nn-UNet	83.44	88.97	94.76	89.06	65.66	73.23	66.74	68.54	55.54	63.67	67.19	62.13
ADDA	80.88	87.36	92.96	87.07	69.98	75.12	75.78	73.62	70.02	68.13	62.94	67.03
CyCADA	81.12	87.89	93.06	87.36	70.12	75.24	77.13	74.16	70.12	70.54	62.91	67.86
SIFA	81.71	87.87	92.98	87.52	70.37	76.98	77.02	74.79	69.89	71.12	63.01	68.01
ADR	81.69	87.94	93.01	87.55	71.81	77.02	76.65	75.10	70.16	72.04	62.98	68.39
ours	81.77	88.01	93.04	87.61	75.03	80.03	78.74	77.93	70.59	74.51	63.18	69.43

3.2 Comparison with State-of-the-Art (SOTA) Method

We compared our method with four SOTA methods: ADDA [18], CyCADA [22], SIFA [23] and ADR [24]. We directly use the code from the corresponding papers. For fair comparison, we have replaced the backbone of these models with the same we used in our approach. The quantitative comparison results of cross-age infant brain segmentation is presented in Table 1, and due to space limitations, we put the experimental results of the brain tumor segmentation task in Table 1 of Supplementary Material, Sec.3.

As observed, our method demonstrates very good DA ability on the cross-age infant segmentation task, which improves about 5.46 DICE and 4.75 DICE on 12-month-old and 3-month-old datasets, respectively. When compared to the four selected SOTA DA methods, we also show superior transfer performance in all the target domains. Specially, we outperform other SOTA methods by at least 2.83 DICE and 1.04 DICE on the 12-month-old and 3-month-old tasks.

When transferring to a single target domain in the brain tumor segmentation task, our proposed DA solution improves about 3.09 DICE in the target LGG domain. Also, the proposed method shows considerable improvements over ADDA and CyCADA, but very subtle improvements to the SIFA and ADR methods (although ADR shows a small advantage on the Whole category).

We also visualize the segmentation results on a typical test sample of the infant brain dataset in Fig. 2, which once again demonstrates the advantage of our method in some detailed regions.

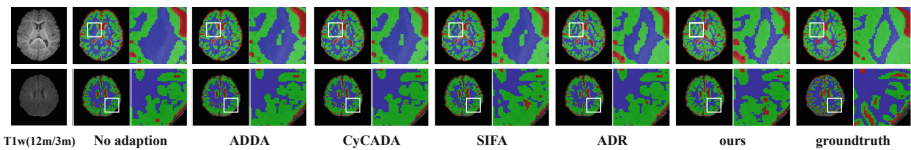
**Fig. 2.** Visualization of segmentation maps (details) for all the comparison methods.

Table 2. Ablation study about prompt-DA, adv-DA and comb-DA.

model	experiment	6 month				12 month				3 month			
		WM	GM	CSF	avg.	WM	GM	CSF	avg.	WM	GM	CSF	avg.
3D-Unet	no-DA	82.47	88.57	93.84	88.29	68.58	72.37	76.45	72.47	69.29	61.92	62.84	64.68
	adv-DA	80.88	87.36	92.96	87.07	69.98	75.12	75.78	73.62	70.02	68.13	62.94	67.03
	prompt-DA	81.57	87.90	93.06	87.51	71.3	77.82	77.16	75.06	70.69	69.51	62.83	67.68
	comb-DA	81.77	88.01	93.04	87.61	75.03	80.03	78.74	77.93	70.59	74.51	63.18	69.43
TransUnet	no-DA	73.24	81.12	84.19	79.52	66.04	70.12	54.94	63.72	39.70	59.49	59.25	52.81
	adv-DA	72.76	80.72	82.98	78.82	66.72	70.39	55.21	64.11	39.89	60.02	59.89	53.27
	prompt-DA	73.01	80.31	83.21	78.84	67.41	71.01	55.41	64.61	40.17	60.22	60.09	53.49
	comb-DA	72.98	80.59	83.61	79.06	70.25	72.57	57.04	66.62	42.61	61.03	61.57	55.07

3.3 Ablation Study

Prompt-DA vs. adv-DA: Since the performance reported in Table 1 is achieved with the method combining prompt-DA and adv-DA, we carry out more studies to investigate: 1). Does prompt-DA itself shows the transfer ability? 2). Is prompt-DA compatible with adv-DA?

The corresponding experiments are conducted on the infant brain dataset and experimental results are shown in Table 2. To make the table more readable, we denote: *no-DA* means only training the segmentation network without any DA strategies; *adv-DA* presents only using adversarial learning based DA; *prompt-DA* is the proposed prompt learning based DA and *comb-DA* is our final DA algorithm which combines both *adv-DA* and *prompt-DA*.

As observed in Table 2, both adv-DA and prompt-DA can improve the transfer performance on all the target domains. When looking into details, the proposed prompt-DA can improve more (1.44 DICE and 0.65 DICE respectively) compared to the adv-DA on both 12-month-old and 3-month-old with 3D-UNet segmentation backbone. When combined together (i.e., comb-DA), the performance can be further improved by a considerable margin, 2.87 DICE and 1.75 DICE on 12-month-old and 3-month-old respectively, compared to prompt-DA. With TransUNet segmentation backbone, we can find the similar phenomenon. To this end, we can draw conclusions that 1). Prompt-DA itself is beneficial to improve the transfer ability; 2). prompt-DA is quite compatible with adv-DA.

Fusion Strategy for Learning Domain-Aware Features: One of the key components of the prompt-DA is to learn domain-aware features through fusion. We have evaluated the effectiveness of our proposed feature fusion strategy in both 3D and 2D models. For comparison, we considered several other fusion strategies, including ‘add/mul’, which adds or multiplies the encoder feature and prompt directly, ‘conv’, which employs a single convolutional layer to process the concatenated features, and ‘rAFusion’, which utilizes a reverse version of the AFusion module, sending the prompt to the spatial branch and the encoder feature to the channel branch. The results of these experiments are presented in Table 3.

Table 3. Ablation study about fusion strategies to learn domain-aware features.

model	experiment	6 month				12 month				3 month			
		WM	GM	CSF	avg.	WM	GM	CSF	avg.	WM	GM	CSF	avg.
3D-Unet	no-DA	82.47	88.57	93.84	88.29	68.58	72.37	76.45	72.47	69.29	61.92	62.84	64.68
	add/mul	81.31	87.62	92.67	87.2	73.88	78.59	76.52	76.33	68.27	74.17	62.90	68.45
	conv	81.4	87.61	93.16	87.39	73.06	77.73	78.73	76.51	69.91	71.93	63.02	68.29
	rAFusion	81.72	88.17	93.33	87.74	74.82	79.75	77.84	77.47	69.42	74.54	62.98	68.98
	AFusion	81.77	88.01	93.04	87.61	75.03	80.03	78.74	77.93	70.59	74.51	63.18	69.43
TransUnet	no-DA	73.24	81.12	84.19	79.52	66.04	70.12	54.94	63.72	39.70	59.49	59.25	52.81
	add/mul	72.66	80.69	83.77	79.04	68.66	70.92	55.52	65.03	40.41	60.87	60.32	53.87
	conv	72.72	80.56	83.71	79.00	69.75	71.01	55.56	65.44	41.32	60.97	60.24	54.18
	rAFusion	72.96	80.72	83.74	79.14	69.80	72.44	56.18	66.14	42.12	61.01	60.50	54.54
	AFusion	72.98	80.59	83.61	79.06	70.25	72.57	57.04	66.62	42.61	61.03	61.57	55.07

Our experimental results demonstrate that the proposed AFusion module improves the model’s performance significantly, and it is effective for both 3D and 2D models.

4 Conclusion

In this paper, we propose a new DA paradigm, namely, prompt learning based DA. The proposed prompt-DA uses a classifier and a prompt generator to produce domain-specific information and then employs a fusion module (for encoder features and prompts) to learn domain-aware representation. We show the effectiveness of our proposed prompt-DA in transfer ability, and also we prove that the prompt-DA is smoothly compatible with the other DA algorithms. Experiments on two DA datasets with two different segmentation backbones demonstrate that our proposed method works well on DA problems.

Acknowledgements. This work was supported by the National Natural Science Foundation of China (No. 62001222), the China Postdoctoral Science Foundation funded project (No. 2021TQ0150 and No. 2021M701699).

References

1. Ronneberger, O., Fischer, P., Brox, T.: U-net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) MICCAI 2015. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28
2. Zhou, S., Nie, D., Adeli, E., Yin, J., Lian, J., Shen, D.: High-resolution encoder-decoder networks for low-contrast medical image segmentation. IEEE Trans. Image Process. **29**, 461–475 (2019)
3. Ouyang, C., Kamnitsas, K., Biffi, C., Duan, J., Rueckert, D.: Data efficient unsupervised domain adaptation for cross-modality image segmentation. In: Shen, D., et al. (eds.) MICCAI 2019. LNCS, vol. 11765, pp. 669–677. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-32245-8_74

4. Xie, X., Chen, J., Li, Y., Shen, L., Ma, K., Zheng, Y.: MI²GAN: generative adversarial network for medical image domain adaptation using mutual information constraint. In: Martel, A.L., et al. (eds.) MICCAI 2020. LNCS, vol. 12262, pp. 516–525. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-59713-9_50
5. Dou, Q., et al.: Pnp-adanet: plug-and-play adversarial domain adaptation network at unpaired cross-modality cardiac segmentation. *IEEE Access* **7**, 99 065–99 076 (2019)
6. Chen, C., Dou, Q., Chen, H., Qin, J., Heng, P.-A.: Synergistic image and feature adaptation: Towards cross-modality domain adaptation for medical image segmentation. In: *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, no. 01, pp. 865–872 (2019)
7. Cui, H., Yuwen, C., Jiang, L., Xia, Y., Zhang, Y.: Bidirectional cross-modality unsupervised domain adaptation using generative adversarial networks for cardiac image segmentation. *Comput. Biol. Med.* **136**, 104726 (2021)
8. Kumar, A., Ma, T., Liang, P.: Understanding self-training for gradual domain adaptation. In: *International Conference on Machine Learning*, pp. 5468–5479. PMLR (2020)
9. Sheikh, R., Schultz, T.: Unsupervised domain adaptation for medical image segmentation via self-training of early features. In: *International Conference on Medical Imaging with Deep Learning*, pp. 1096–1107. PMLR (2022)
10. Xie, Q., et al.: Unsupervised domain adaptation for medical image segmentation by disentanglement learning and self-training. *IEEE Trans. Med. Imaging* (2022)
11. Yang, C., Guo, X., Chen, Z., Yuan, Y.: Source free domain adaptation for medical image segmentation with Fourier style mining. *Med. Image Anal.* **79**, 102457 (2022)
12. Liu, X., Ji, K., Fu, Y., Du, Z., Yang, Z., Tang, J.: P-tuning v2: prompt tuning can be comparable to fine-tuning universally across scales and tasks. *arXiv preprint arXiv:2110.07602* (2021)
13. Zhou, K., Yang, J., Loy, C.C., Liu, Z.: Learning to prompt for vision-language models. *Int. J. Comput. Vision* **130**(9), 2337–2348 (2022)
14. Jia, M., et al.: Visual prompt tuning. In: Avidan, S., Brostow, G., Cissé, M., Farinella, G.M., Hassner, T. (eds.) *Computer Vision - ECCV 2022. Lecture Notes in Computer Science*, vol. 13693, pp. 709–727. Springer, Cham (2022)
15. Zheng, Z., Yue, X., Wang, K., You, Y.: Prompt vision transformer for domain generalization. *arXiv preprint arXiv:2208.08914* (2022)
16. Çiçek, Ö., Abdulkadir, A., Lienkamp, S.S., Brox, T., Ronneberger, O.: 3D U-net: learning dense volumetric segmentation from sparse annotation. In: Ourselin, S., Joskowicz, L., Sabuncu, M.R., Unal, G., Wells, W. (eds.) MICCAI 2016. LNCS, vol. 9901, pp. 424–432. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46723-8_49
17. Chen, J., et al.: Transunet: transformers make strong encoders for medical image segmentation. *arXiv preprint arXiv:2102.04306* (2021)
18. Tzeng, E., Hoffman, J., Saenko, K., Darrell, T.: Adversarial discriminative domain adaptation. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 7167–7176 (2017)
19. Sudre, C.H., Li, W., Vercauteren, T., Ourselin, S., Jorge Cardoso, M.: Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations. In: Cardoso, M.J., et al. (eds.) *DLMIA/ML-CDS -2017. LNCS*, vol. 10553, pp. 240–248. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-67558-9_28
20. Sun, Y., et al.: Multi-site infant brain segmentation algorithms: the ISEG-2019 challenge. *IEEE Trans. Med. Imaging* **40**(5), 1363–1376 (2021)

21. Bakas, S., et al.: Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the brats challenge. arXiv preprint [arXiv:1811.02629](https://arxiv.org/abs/1811.02629) (2018)
22. Hoffman, J., et al.: Cycada: cycle-consistent adversarial domain adaptation. In: International Conference on Machine Learning, pp. 1989–1998. PMLR (2018)
23. Chen, C., Dou, Q., Chen, H., Qin, J., Heng, P.A.: Unsupervised bidirectional cross-modality adaptation via deeply synergistic image and feature alignment for medical image segmentation. *IEEE Trans. Med. Imaging* **39**(7), 2494–2505 (2020)
24. Zeng, G., et al.: Semantic consistent unsupervised domain adaptation for cross-modality medical image segmentation. In: de Bruijne, M., et al. (eds.) *MICCAI 2021. LNCS*, vol. 12903, pp. 201–210. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-87199-4_19