



Cross-domain additive learning of new knowledge rather than replacement

Jiahao Liu¹ · Ge Jiao¹

Received: 1 July 2023 / Revised: 10 January 2024 / Accepted: 27 May 2024 / Published online: 7 June 2024
© Korean Society of Medical and Biological Engineering 2024

Abstract

In medical clinical scenarios for reasons such as patient privacy, information protection and data migration, when domain adaptation is needed for real scenarios, the source-domain data is often inaccessible and only the pre-trained source model on the source-domain is available. Existing solutions for this type of problem tend to forget the rich task experience previously learned on the source domain after adapting, which means that the model simply overfits the target-domain data when adapting and does not learn robust features that facilitate real task decisions. We address this problem by exploring the particular application of source-free domain adaptation in medical image segmentation and propose a two-stage additive source-free adaptation framework. We generalize the domain-invariant features by constraining the core pathological structure and semantic consistency between different perspectives. And we reduce the segmentation generated by locating and filtering elements that may have errors through Monte-Carlo uncertainty estimation. We conduct comparison experiments with some other methods on a cross-device polyp segmentation and a cross-modal brain tumor segmentation dataset, the results in both the target and source domains verify that the proposed method can effectively solve the domain offset problem and the model retains its dominance on the source domain after learning new knowledge of the target domain. This work provides valuable exploration for achieving additive learning on the target and source domains in the absence of source data and offers new ideas and methods for adaptation research in the field of medical image segmentation.

Keywords Domain adaptation · Image segmentation · Domain alignment · Uncertainty estimate

1 Introduction

Deep learning has achieved great success in mainstream computer vision in recent years, and these experiences provide a reference solution for the field of medical image analysis [1–5]. In real clinical scenarios, medical imaging is often acquired using different modalities, scanners, and protocols from different locations and populations that may have different characteristics, and these data suffer from severe domain shifts (inconsistent data distribution in the source and target domains) which may lead to performance degradation of pre-designed methods [6, 7]. Simultaneously, providing intensive professional annotations by high-level physicians is time-consuming and labor-intensive, and real patient data is subject to privacy protections and security

regulations resulting in limited access. The lack of annotation also leads to the fact that training a new target domain model is not only overhead but also difficult to implement. The above mentioned problems of domain shift and lack of annotation make it difficult for traditional vision processing solutions to work directly in the field of medical image analysis. How to adapt the trained source model with a small amount of target domain data in real clinical scenarios with drastic domain changes is an urgent problem.

Unsupervised Domain Adaptation (UDA) maps source and target domains to common feature space and then aligns differences in the distribution of two features so that model adapts itself to the target domain samples and obtains close or even consistent performance with source domain through similar representations [6–8]. These approaches require labeled source datasets and well-trained models to learn the source knowledge in domain adaptation training. As shown in Fig 1, compared to UDA, where source and target domain samples are directly aligned, Source-Free Domain Adaptation (SFDA) adjusts the parameters of a pre-trained

✉ Ge Jiao
jiaoge@126.com

¹ College of Computer Science, Hengyang Normal University, Hengyang 421008, China

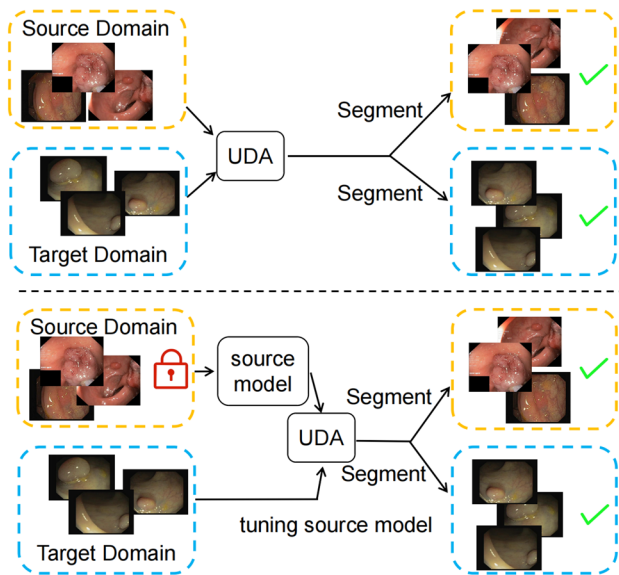


Fig. 1 Explanation of the difference between domain adaptation and source-free domain adaptation. Domain adaptation methods utilize both source domain data and target domain data to design feature alignment and other adaptation methods. In practical clinical scenarios, accessing source domain data may be infeasible due to issues such as privacy and security concerns. Source-free domain adaptation, on the other hand, is a solution that addresses this limitation. It involves tuning pre-trained source model parameters using alternative methods without accessing the source domain data, aiming to improve the model's performance on the target domain

source model using only the target domain samples. This adjustment aims to reduce the feature differences between the source and target domains [9, 10].

Existing SFDA methods can be roughly categorized into two types. The first type achieves domain alignment through virtual domain generation. Yang et al. [11] extract source domain knowledge from the source model to pattern the generative model of the source data to perform style transformation on the target domain data, and then using methods such as contrast learning and noise label filtering to perform target domain adaptation. Tian et al. [12] generate virtual domain samples based on a pre-trained source model in feature space using an approximate Gaussian mixture model, allowing the virtual domain to maintain a similar distribution to the source domain without accessing the source data. Qiu et al. [13] train a prototype generator by exploring the classification boundary information of the source model through contrast learning. However, these methods not only require additional computational costs but also necessitate the design of dedicated generation schemes for different domains, which can be a cumbersome process. The second type focuses on inter-domain feature alignment. These methods leverage techniques such as knowledge distillation and statistical consistency to achieve robust model parameterization against domain shifts [10]. Bateson et al. [14] guides the

weakly labeled target sample and the domain-invariant prior on the segmented region based on minimizing the unlabeled entropy learning defined on the target domain data. Kim et al. [15] selected reliable samples with self-entropy criterion and defined them as class prototypes. Self-supervised learning is then performed by assigning pseudo-labels to each target sample based on the similarity scores of the class prototypes. These methods are free from the constraints of domain generation and can achieve simple and generalized domain distribution adaptation by guiding the feature distribution for alignment through prior knowledge.

The above inter-domain feature alignment-based methods have made some progress without accessing the source data. However, on the one hand, adaptation learning over a long time on a small number of target domain samples leads to a bias of the model to fit these only samples, making it difficult for the model to maintain its original dominance over the source domain during the adaptation process. On the other hand, the feature signals obtained by these methods inevitably contain redundant, erroneous, and harmful noise signals due to the lack of explicit supervision. Indiscriminate use of these signals to update the source model parameters may lead to the model learning in a bad direction or even to a complete collapse. With these limitations in mind, this letter aims to investigate persistent adaptation over the target domain in the presence of effective supervision.

We propose a two-stage SFDA framework for additive source-free domain adaptation, aiming to achieve more robust and simplified target domain adaptation for medical image segmentation tasks. Inspired by consistency learning, in the first stage, we freeze the decoder part of model and generalize the encoding capability through aligning style and content consistency between rotated and cropped images. In the second stage, we utilize the model fine-tuned in the first stage with the frozen encoder part and guide the target domain samples with their enhanced knowledge distillation using uncertainty maps. The main contributions of this letter can be summarized as follows:

- We investigate a more realistic and challenging task to achieve continuous learning on both the target and source domains without accessing the source data. The proposed method can effectively deal with the problem of cross-distribution domain bias and the problem of unavailability of source domain data due to privacy and security protection in the field of medical image segmentation.
- We propose a two-stage approach to adapt the encoder and decoder of the model separately. The framework uses multi-view feature styles and content consistency in the encoder adaptation stage to generalize the feature extraction capability of the encoder of the segmentation model, and reduces errors in the decoder reconstruction of the segmentation results by finding and eliminating poten-

tially erroneous feature elements through uncertainty estimation in the decoder adaptation stage.

- We conduct fair comparison experiments with current state-of-the-art methods in cross-device polyp segmentation and cross-modal brain tumor segmentation application scenarios, respectively. We validate the effectiveness of the proposed method on the target domain to demonstrate that the proposed method can be well adapted to different target domain offsets. Further, we report the performance of the post-adaptation method on a test set in the source domain to demonstrate that the proposed method is knowledge-retentive rather than knowledge-replacing.

2 Method

2.1 Overview

We propose a two-stage adaptation framework to enable additive source-free domain adaptation by adjusting the model to learn domain-invariant features. It consists of an encoder adaptation stage that learns joint style and content invariance and a decoder adaptation stage that reduces feature uncertainty.

Let us define the source data $D_s = \{x_s, y_s\}$ and the target data $D_t = \{x_t\}$, where x_s is the source image, y_s is the corresponding source label, and y_t is the target image. x^{crop} and x^{rot} are the cropped and rotated enhanced images, respectively, and z^{crop} , z^{rot} , z_t are the intermediate vectors encoded by the encoder. p is the predicted probability map. M_s is the

pre-trained source model. Our goal aims to improve the performance of the adjusted model M_t on the target domain by adjusting M_s in conjunction with the target data D_t without accessing the source data D_s .

2.2 Encoder adaptation stage

Medical imaging exposes different anatomical structures of organs or tissues, and these images exhibit significant style information differences in terms of texture, contrast, saturation, and other visual attributes across different devices or modalities. These style variances exacerbate the domain shift issues along these cross-domain imaging pathways, leading to performance degradation of models in cross-scenario settings [11, 16]. Taking these inherent properties into consideration, in the encoder adaptation stage (fixed decoder, adjusted encoder), we decompose the high-level semantic representation space learned by the encoder into content representation and style representation. We then use style matching to enforce style consistency across the overall representation distribution between the base branch and the two branches based on cropping and rotation. The specific process is shown in the Fig 2.

Specifically, for the enhanced sample feature z model from the base feature is transformed (same range cropping or same angle rotation) to obtain the original feature corresponding to the field of view for the same type of transformation to constrain the style content consistency.

The process of style and content consistency constraint is shown in the following equation:

Encoder Adaptation Stage

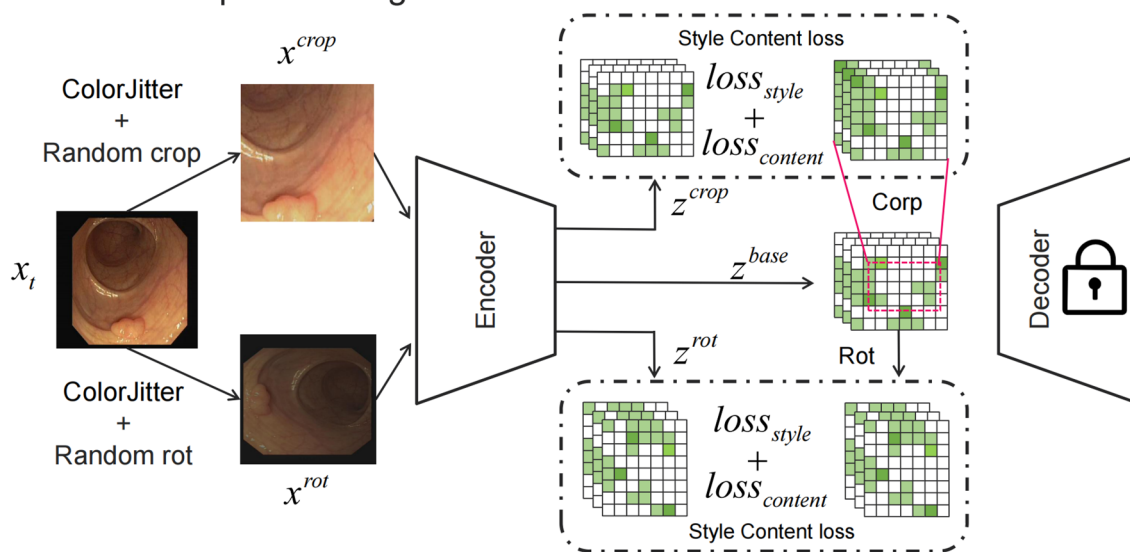


Fig. 2 The proposed encoder adaptation stage. The decoder is frozen first to improve the robustness and anti-interference capability of the encoder by learning styles and content consistency for features from different viewpoints

$$loss_{style} = \left(\frac{\sum_{i=1}^n z_i^{base}}{n} - \frac{\sum_{i=1}^n z_i^{aug}}{n} \right)^2 + \frac{\sum_{i=1}^n (z_i^{base} - z_{mean}^{base})^2}{n} + \frac{\sum_{i=1}^n (z_i^{aug} - z_{mean}^{aug})^2}{n} \quad (1)$$

$$loss_{content} = \frac{\sum_{i=1}^n (z_i^{base} - z_i^{aug})^2}{n} \quad (2)$$

where n is the number of samples in a batch, $z^{aug} = (z^{crop}, z^{rot})$, and z_{mean}^{aug} is actually the mean of the sample of the corresponding augmentation perspective in a batch. In the specific implementation, the alignment of z^{aug} from different perspectives with the base branch z^{base} is calculated separately.

2.3 Decoder adaptation stage

The high-level semantic representations encoded by the encoder may include sub-elements with low confidence and incorrect categorization. These errors will be gradually amplified by the upsampled decoder, resulting in regionally false-positive segmentation results [17, 18]. We can capture these low-confidence sub-elements by estimating the model's uncertainty on the samples. Then, we can jointly constrain the decoder's reconstruction process through two

mechanisms: filtering out low-confidence elements and enforcing consistency across different augmented perspectives of the same sample. The specific process is shown in the Fig 3.

Specifically, we perform Monte-Carlo uncertainty estimation [19] on the prediction probability maps p^{base} and p^{aug} from the base perspective and the perspective augmented by ColorJitter in the decoder adaptation stage (fixed encoder, adjusted decoder) to obtain the uncertainty maps U^{base} and U^{aug} in the prediction probability maps and the average probability prediction maps p^{base} and p^{aug} . Pixels with low confidence in U^{base} and U^{aug} are filtered by a threshold γ , and the filtered uncertainty maps are used to weight the average prediction maps. Finally, knowledge distillation is used to align the weighted probability prediction maps.

The uncertainty map is calculated as follows:

$$\Delta_U = \frac{\sum_{i=1}^n \sum_{j=1}^t (p_{ij} - p_{ij}^{mean})^2}{n} \quad (3)$$

$$U = \begin{cases} \Delta_U & \Delta_U \geq \gamma \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

where γ is a hyperparameter used to adjust the filtering ratio, t is the number of Monte-Carlo uncertainty estimation iterations, and p_j^{mean} represents the average of the results of t iterations.

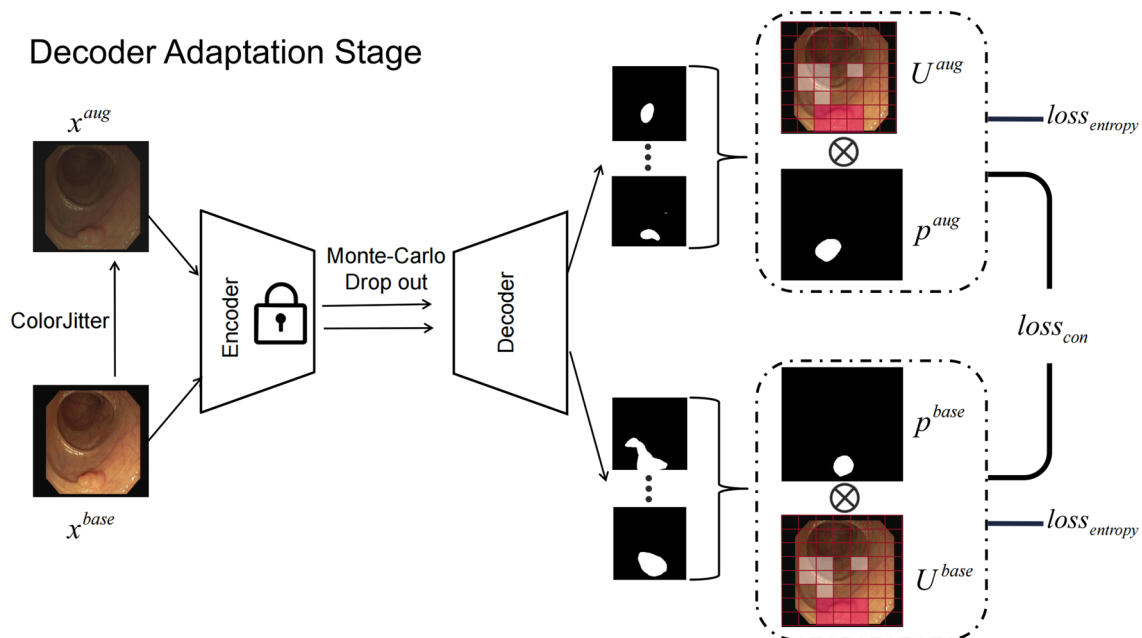


Fig. 3 The proposed decoder adaptation stage. At the end of encoder self-adaptation, the encoder is frozen and then the decoder is finetuned. The uncertainty of the samples is calculated by Monte-Carlo uncertainty estimation, and then the features are filtered by threshold

filtering and feature weighting to reduce the uncertainty elements in the features and ensure that the decoder is optimized in the right direction

The calculation process of consistency distillation is as follows:

$$loss_{con} = \frac{\sum_{i=1}^n (p_i^{base} * U_i^{base} - p_i^{aug} * U_i^{aug})^2}{n} \quad (5)$$

$$loss_{entropy} = - \sum_{i=1}^n p_i \times \log p_i \quad (6)$$

2.4 Optimization

The optimization process of the proposed method is to first perform encoder adaptation, and then decoder adaptation based on the adaptation results, which are optimized according to $loss_{stage1}$ and $loss_{stage2}$ respectively. During the encoder adaptation stage, the decoder is frozen, and the adaptation is only for the encoder. Similarly, during the decoder adaptation stage, the encoder is frozen, and the adaptation is only for the decoder.

$$loss_{stage1} = \alpha \times loss_{style} + (1 - \alpha) \times loss_{content} \quad (7)$$

$$loss_{stage2} = \beta \times loss_{con} + (1 - \beta) \times loss_{entropy} \quad (8)$$

where α and β are corresponding weight hyperparameters.

3 Experiment

3.1 Experiment setting

We utilize DeeplabV3+ [20] as our base segmentation model for conducting source-free domain adaptation experiments and implemented the entire framework using PyTorch. Following the workflow of previous works [11, 14], we divide the data into source domain data and target domain data. The source domain data is used to pretrain the source model, and then the corresponding domain adaptation methods were employed to optimize the model on the target domain. We used the SGD optimizer and applied basic data augmentation using Colorjitter. For methods with open-source code, we conduct experiments using the provided code, while for methods without open-source implementations, we strictly follow the descriptions in their papers to construct the corresponding pipelines.

3.2 Dataset

We perform fair comparison experiments with current state-of-the-art methods on two publicly available medical image segmentation datasets. We follow previous methods

to segment the datasets [11]. And for each task we perform a 3-fold cross-validation.

Cross-device polyp segmentation: The publicly available colonoscopy datasets EndoScene [21] and ETIS-Larib [22] are used for cross-device adaptation. The EndoScene dataset includes 912 images from 36 patients. It is collected by Olympus Q160AL and Q165L, additional II video processors. We set the EndoScene dataset [21] as the source domain and follow the standard setting for polyp segmentation as described by [21] with a ratio of 3:1:1 for the training set validation set and test set, respectively. We set the ETIS-Larib dataset [22] as the target domain, which is composed of Pentax 90i series, EPKi 7000 video processor collected 196 frames. We randomly set the training set to 4:1 with the test set. To facilitate training and testing, we resized all images to 256x256 dimensions.

Cross-modal brain tumor segmentation: The Multimodal Brain Tumor Segmentation Challenge 2018 dataset [23] is a dataset providing multimodal 3D brain MRI and ground truth segmentation for each case including 4 MRI modalities (T1, T1c, T2 and FLAIR). We refer to the pipeline of previous work [24] to partition the source and target domains of the samples in a 1:1 ratio from the data of 285 HGG patients on this dataset, and then further randomly partition the training and test sets in a 4:1 ratio on the corresponding domains. We perform experiments in both flair and T2 modals of MRI imaging, where the size of each axial slice is adjusted to 192×168 .

3.3 Comparison with the state-of-the-art methods

To verify the effectiveness of the proposed method, we conduct fair comparison experiments with some of the most popular methods in the same environment. We compare the experimental results with some state-of-the-art methods on the cross-device polyp segmentation task and the cross-modal brain tumor segmentation task, respectively. These methods are described below:

- AdaEnt [25]: This method combines domain-invariant prior with entropy loss minimization to guide segmentation. It learns an analogical prior through an auxiliary network and integrates it in the overall loss function in the form of Kullback–Leibler (KL) divergence.
- AUGCO [26]: This method utilizes pixel-level prediction consistency of the model, automatically generates views of each target image, and utilizes model confidence to identify reliable pixel predictions. It selectively self-trains these images.
- SFDA [15]: This method employs a pre-trained model from the source domain and progressively updates the target model in a self-learning manner. It assigns pseudo-

labels to each target sample using reliable samples selected based on the self-entropy criterion.

- AdaMI [14]: This method minimizes the unlabeled entropy loss defined on the target domain data to further guide weak labels of target samples and domain-invariant prior on segmentation regions.
- SMG [11]: In the generation stage, this method achieves inverse-source class images using a pre-trained source model and statistical information of mutually Fourier-transformed. In the adaptation stage, it transfers relational knowledge using domain distillation loss and reduces domain discrepancy through domain contrastive loss in a self-supervised paradigm.

In Table 1 the data reported is the performance of these methods on the test set of target domain and in Table 2 the data is the performance of these methods on the test set of source domain. The Source Only method in Table 1 is the result obtained by training on the source domain data through a single-step transfer learning method. Specifically, we use ResNet101 pre-trained on ImageNet as initialization backbone, and then use source domain data for fine-tuning the whole model. Similarly, Target Only use the ResNet101 pre-trained on ImageNet as initialization backbone, and then use target domain data to fine-tuning the whole model. As shown in Table 1, the performance of Source Only tested on target domain test set can be treated as a pass mark for source domain-free adaptive method for this task. While Target Only is trained directly on target domain, the performance on target domain test set can be regarded as the upper limit of source-free adaptation method for this task. In the cross-device polyp segmentation task, the performance of model can generally be improved compared to using domain-adaptive methods and not using any domain-adaptive methods. For example, using methods such as AdaEnt, SFDA, and AdaMI can improve the model’s Dice score on target domain to around

Table 2 Compare with State-of-The-Art methods on the source domain test set. The Source Model represents the performance of the source model on the source domain test set

Method	Cross-device polyp segmentation		Cross-modal brain tumor segmentation	
	Dice	Miou	Dice	Miou
Source Model	87.04	78.82	85.67	76.31
AdaEnt [25]	83.17	73.97	80.93	70.3
AUGCO [26]	85.8	76.9	83.82	73.61
SFDA [15]	85.1	76.16	83.12	72.84
AdaMI [14]	83.41	74.1	81.2	71.02
SMG [11]	85.24	76.33	82.84	72.35
ours	86.16	77.42	84.26	74.9

70.12, and these methods can also improve performance in a cross-modal brain tumor segmentation task (Dice scores 68.31 and 67.64). The limitation of these methods to further improve their performance may be that they only perform simple adaptation of output layer features of the model without in-depth consideration of the consistency of different viewpoints of same representation and bias caused by the uncertainty of model’s prediction. Whereas methods such as AUGCO and SMG consider more intermediate layer features for adaptation, these methods are either very sensitive to changes in viewpoints or require complex hyperparameters for tuning. In contrast, the Dice Score on the tasks of polyp segmentation and brain tumor segmentation using the proposed method reaches 73.7 and 70.61, respectively. This is mainly attributed to the fact that the proposed method uses a two-phase adaptive approach to learn domain-invariant representations and uncertainty-reducing feature factors, respectively. These operations allow the model to keep the model learning discriminative and robust features in the presence of complex

Table 1 Compare with State-of-The-Art methods on the target domain test set. Source Only indicates the performance of the source model trained on the source domain and tested on the target domain

Method	Cross-device polyp segmentation					Cross-modal brain tumor segmentation				
	Dice	Specificity	Sensitivity	Acc	Miou	Dice	Specificity	Sensitivity	Acc	Miou
Source Only	68.03	98.94	74.26	96.12	59.22	62.75	98.25	74.26	94.19	56.47
Target Only	79.32	99.53	82.97	99.04	76.81	77.46	99.76	82.9	99.13	73.2
AdaEnt [25]	69.84	99.37	75.9	97.09	61.43	66.13	99.19	72.66	95.56	56.89
AUGCO [26]	69.2	98.88	75.71	97.14	61.2	66.4	98.7	72.11	95.74	57.21
SFDA [15]	70.12	98.9	77.65	98.07	61.21	68.31	98.57	76.41	97.01	59.6
AdaMI [14]	70.58	98.72	76.11	97.92	61.87	67.64	98.8	75.56	96.22	58.4
SMG [11]	71.66	98.84	78.8	97.36	62.13	68.23	99.37	76.12	97.12	59.24
ours	73.7	98.91	80.74	98.28	64.4	70.61	99.5	78.32	98.04	61.42

(without any domain adaptation) Target Only represents the performance of the source model trained on the target domain and tested on the target domain

changes, while selecting features with higher confidence for final decision making and segmentation.

At the same time, to evaluate the persistence of the proposed method in the source domain more objectively, we also report the performance of the adaptation model on the source domain test set to prove that the knowledge learned by the proposed method is retention-addition rather than replacement-forgetting. It can be seen from the Table 2 that after the model completes the adaptation process, the model shows different degrees of performance degradation on the test set of the source domain, for example, AdaMI, AdaEnt and other methods have Dice Scores of only 83.41 and 83.17 (87.04 for the source model) in the adaptation model. Although these methods achieve performance improvement on the target domain after adaptation, they show significant performance degradation on the source domain, which means that these methods forget the rich experience of previous learning after learning new knowledge, and some beneficial weight parameters are replaced with new knowledge. This phenomenon can also be interpreted as the overfitting of the model to the target domain sample. In contrast, using proposed method the model can almost maintain its original advantage in the source domain (Dice Score 86.16) while achieving the best performance in the target domain (Dice Score 73.7 and 70.61). This shows that the features learned by proposed method are more generalized and robust than others.

3.4 Ablation experiments

We perform a series of ablation experiments to verify some other details in the overall framework.

Loss function curves: We plot the loss function curves of the proposed method in the encoder and decoder adaptation phases. As shown in Fig. 4, since the entire network is initialized using the parameters of the source model, which means that the network already has a certain amount of discriminative ability, the loss function starts to decrease from a relatively low point at the beginning of the encoder adaptation phase. After about 100 epochs of training, the model converges in the encoder adaptation phase. At the beginning of the decoder adaptation phase, the loss starts to decrease from around 0.31, and the loss converges to around 0.26 after 50 epochs of training. These results validate that combining the consistency and uncertainty estimates in the two complex augmented viewpoint features can effectively improve the performance of the model.

Backbone: To explore the robustness of the proposed method and the influence of different architectures on the model, we conduct ablation experiments on some popular backbone networks [27–32]. The experiment uses the same configuration, replacing only the backbone network portion of the model. It can be seen from the Fig. 5 that

Table 3 Ablation experiments contributed by each module. Baseline is the effect of the basic model without any adaptive adjustment on the target domain test set

Method	Cross-device polyp segmentation		Cross-modal brain tumor segmentation	
	Dice	Miou	Dice	Miou
Baseline	68.03	59.22	62.75	56.47
Baseline+ $loss_{stage1}$	70.78	61.61	66.27	58.33
Baseline+ $loss_{stage1}$ + $loss_{stage2}$	73.7	64.42	70.61	61.42

Table 4 The ablation of Monte-Carlo uncertainty maps

Method	Cross-device polyp segmentation	Cross-modal brain tumor segmentation
w/o Monte-Carlo	71.64	62.87
Monte-Carlo	73.1	63.56

backbones such as Shufflenet [29] and InceptionV3 [31] perform the worst, which may be due to the complexity of the design of these architectures, which leads to their poor performance in some special scenarios, compared to simpler architectures such as EfficientNet [30] and ResNet [27]. After comprehensive consideration, we finally use ResNet as the basic feature extractor.

Ablation of Monte-Carlo uncertainty maps: We perform ablation experiments on decoder adaptation stage to demonstrate the necessity of Monte-Carlo uncertainty maps. The proposed decoder adaptation stage computes Monte-Carlo uncertainty map for the original and enhanced views separately, and then unifies high-confidence pixel computation consistency among different views in uncertainty map by distillation learning. From Table 4, it can be seen that using Monte-Carlo uncertainty map (Monte-Carlo) has a significant advantage over not using Monte-Carlo (w/o Monte-Carlo) in both Dice Score and Miou. These results validate the effectiveness of Monte-Carlo uncertainty map in this passive-domain adaptive scenario for medical imaging.

Loss-function weighted proportion: We design experiments to explore the proportions between different loss functions. An excessively large value of α will cause the model to pay too much attention to the style of the sample and ignore the content modeling. α that is too small causes the model to pay too much attention to the content of the sample and neglect to generalize the features of different styles. Similarly, β is used to regulate the degree of optimization between consistency learning and entropy learning. The results are shown in the Fig. 6, and after comprehensive

Fig. 4 Loss function decline curve. The encoder adaptation phase and the decoder adaptation phase are conducted in 100 epoch and 50 epoch, respectively, and the whole training is conducted in two phases

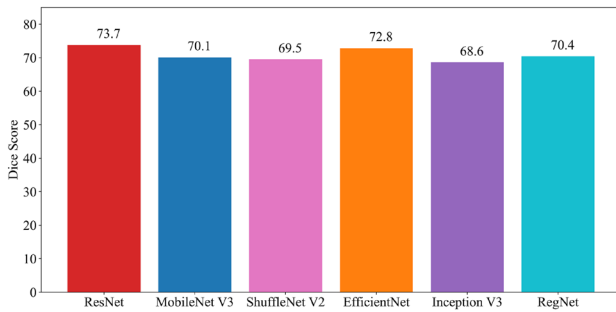
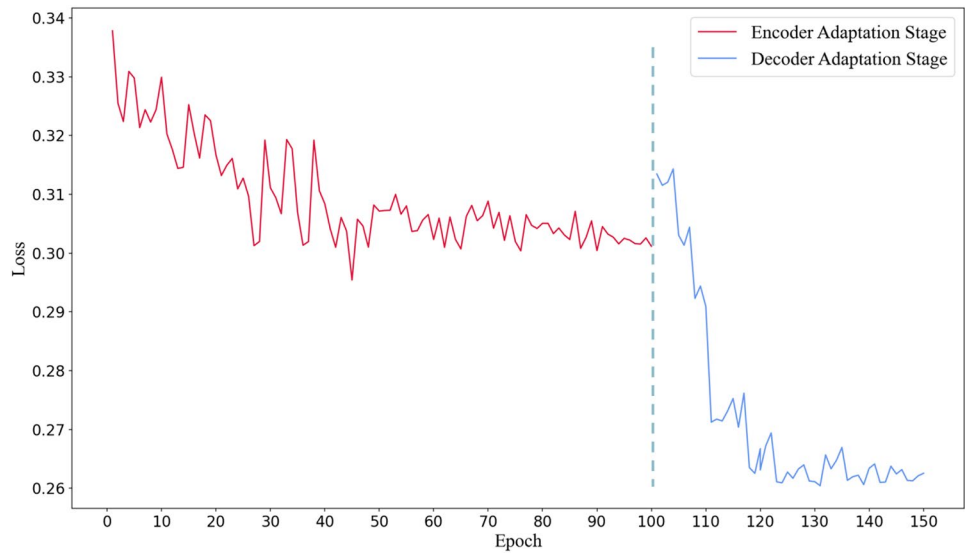


Fig. 5 Comparison of different backbones. All experiments are performed using the same configuration changing only the type of backbone

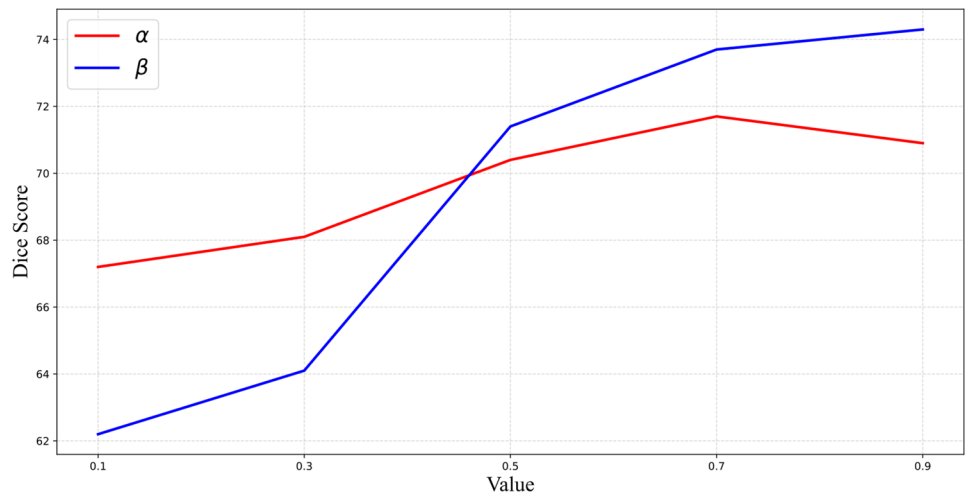
consideration, we set the values of α and β to 0.7 and 0.9, respectively.

Contribution of each module: Table 3 reports the experimental results of the proposed method decoupling, verifying that the proposed components work together and benefit the overall framework. It can be seen from the table that the proposed two-stage domain adaptation fine-tuning is carried out sequentially on the basis of baseline, and the performance of the model is significantly improved (Dice Score from 68.03 to 73.7, 62.75 to 70.61).

4 Conclusion

This letter summarizes the limitations of domain feature alignment methods in adaptation learning and proposes a new two-stage additive SFDA framework to address these

Fig. 6 Weighted ratio of different losses



issues. The proposed method is extensively evaluated on two medical image segmentation tasks: cross-device polyp segmentation adaptation and cross-modal brain tumor segmentation adaptation, achieving significant results that validate the effectiveness and potential applications of the framework. Overall, this work provides valuable exploration for achieving additive learning on the target and source domains in the absence of source data and offers new ideas and methods for adaptation research in the field of medical image segmentation.

Funding This work is supported by the Hunan Provincial Natural Science Foundation of China under Grant 2021JJ50074, in part by the Scientific Research Fund of Hunan Provincial Education Department under Grant 19B082, in part by the Science and Technology Plan Project of Hunan Province under Grant 2016TP1020, and in part by the 14th Five-Year Plan Key Disciplines and Application-oriented Special Disciplines of Hunan Province under Grant Xiangjiaotong [2022] 351.

Declarations

Conflict of interest The authors declare that they have no Conflict of interest.

References

- Chen Y, Yang X-H, Wei Z, Heidari AA, Zheng N, Li Z, Chen H, Hu H, Zhou Q, Guan Q. Generative adversarial networks in medical image augmentation: a review. *Comput Biol Med.* 2022;144: 105382.
- Malhotra P, Gupta S, Koundal D, Zaguia A, Enbeyle W. et al. Deep neural networks for medical image segmentation. *J Healthcare Eng* 2022; vol. 2022.
- Yu H, Yang LT, Zhang Q, Armstrong D, Deen MJ. Convolutional neural networks for medical image analysis: state-of-the-art, comparisons, improvement and perspectives. *Neurocomputing.* 2021;444:92–110.
- Chen X, Wang X, Zhang K, Fung K-M, Thai TC, Moore K, Mannel RS, Liu H, Zheng B, Qiu Y. Recent advances and clinical applications of deep learning in medical image analysis. *Med Image Anal.* 2022;79: 102444.
- Ma X, Niu Y, Gu L, Wang Y, Zhao Y, Bailey J, Lu F. Understanding adversarial attacks on deep learning based medical image analysis systems. *Pattern Recogn.* 2021;110: 107332.
- Yan W, Wang Y, Gu S, Huang L, Yan F, Xia L, Tao Q. The domain shift problem of medical image segmentation and vendor-adaptation by unet-gan. In: *Medical image computing and computer assisted intervention-MICCAI, 22nd international conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part II* 22. Springer. 2019:623–31.
- Na J, Jung H, Chang HJ, Hwang W. Fixbi: Bridging domain spaces for unsupervised domain adaptation. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition.* 2021. pp. 1094–1103.
- Zhao S, Yue X, Zhang S, Li B, Zhao H, Wu B, Krishna R, Gonzalez JE, Sangiovanni-Vincentelli AL, Seshia SA, et al. A review of single-source deep unsupervised visual domain adaptation. *IEEE Trans Neural Netw Learn Syst.* 2020;33(2):473–93.
- Fang Y, Yap P-T, Lin W, Zhu H, Liu M. Source-free unsupervised domain adaptation: a survey. *arXiv preprint arXiv:2301.00265,* 2022.
- Yu Z, Li J, Du Z, Zhu L, Shen HT. A comprehensive survey on source-free domain adaptation. *arXiv preprint arXiv:2302.11803,* 2023.
- Yang C, Guo X, Chen Z, Yuan Y. Source free domain adaptation for medical image segmentation with fourier style mining. *Med Image Anal.* 2022;79: 102457.
- Tian J, Zhang J, Li W, Xu D. Vdm-da: virtual domain modeling for source data-free domain adaptation. *IEEE Trans Circuits Syst Video Technol.* 2021;32(6):3749–60.
- Qiu Z, Zhang Y, Lin H, Niu S, Liu Y, Du Q, Tan M. Source-free domain adaptation via avatar prototype generation and adaptation. *arXiv preprint arXiv:2106.15326,* 2021.
- Bateson M, Kervadec H, Dolz J, Lombaert H, Ayed IB. Source-free domain adaptation for image segmentation. *Med Image Anal.* 2022;82: 102617.
- Kim Y, Cho D, Han K, Panda P, Hong S. Domain adaptation without source data. *IEEE Trans Artif Intell.* 2021;2(6):508–18.
- Su Z, Yao K, Yang X, Huang K, Wang Q, Sun J. Rethinking data augmentation for single-source domain generalization in medical image segmentation. *Proc AAAI Conf Artif Intell.* 2023;37(2):2366–74.
- Ju L, Wang X, Wang L, Mahapatra D, Zhao X, Zhou Q, Liu T, Ge Z. Improving medical images classification with label noise using dual-uncertainty estimation. *IEEE Trans Med Imaging.* 2022;41(6):1533–46.
- Loquercio A, Segu M, Scaramuzza D. A general framework for uncertainty estimation in deep learning. *IEEE Robot Autom Lett.* 2020;5(2):3153–60.
- Abdar M, Salari S, Qahremani S, Lam H-K, Karray F, Hussain S, Khosravi A, Acharya UR, Makarenkov V, Nahavandi S. Uncertaintyfusenet: robust uncertainty-aware hierarchical feature fusion model with ensemble Monte Carlo dropout for covid-19 detection. *Inf Fusion.* 2023;90:364–81.
- Chen L-C, Zhu Y, Papandreou G, Schroff F, Adam H. Encoder-decoder with atrous separable convolution for semantic image segmentation. In: *Proceedings of the European conference on computer vision (ECCV).* 2018. pp. 801–818.
- Vázquez D, Bernal J, Sánchez FJ, Fernández-Esparrach G, López AM, Romero A, Drozdal M, Courville A, et al. A benchmark for endoluminal scene segmentation of colonoscopy images. *J Healthcare Eng.* 2017;2017:4037190.
- Silva J, Histace A, Romain O, Dray X, Granado B. Toward embedded detection of polyps in WCE images for early diagnosis of colorectal cancer. *Int J Comput Assist Radiol Surg.* 2014;9:283–93.
- Bakas S, Reyes M, Jakab A, Bauer S, Rempfler M, Crimi A, Shinohara RT, Berger C, Ha SM, Rozycki M. 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,* 2018.
- Wang Y, Cheng J, Chen Y, Shao S, Zhu L, Wu Z, Liu T, Zhu H. Fvp: fourier visual prompting for source-free unsupervised domain adaptation of medical image segmentation. *arXiv preprint arXiv:2304.13672,* 2023.
- Bateson M, Kervadec H, Dolz J, Lombaert H, Ben Ayed I. Source-relaxed domain adaptation for image segmentation. In: *Medical image computing and computer assisted intervention-MICCAI, 23rd international conference, Lima, Peru, October 4–8, 2020, Proceedings, Part I* 23. Springer. 2020. pp. 490–9.
- Prabhu V, Khare S, Kartik D, Hoffman J. Augco: augmentation consistency-guided self-training for source-free domain adaptive semantic segmentation. *arXiv preprint arXiv:2107.10140,* 2021.

27. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2016. pp. 770–778.
28. Howard A, Sandler M, Chu G, Chen L-C, Chen B, Tan M, Wang W, Zhu Y, Pang R, Vasudevan V. et al. Searching for mobilenet3. In: Proceedings of the IEEE/CVF international conference on computer vision 2019. pp. 1314–1324.
29. Ma N, Zhang X, Zheng H-T, Sun J. Shufflenet v2: Practical guidelines for efficient cnn architecture design. In: Proceedings of the European conference on computer vision (ECCV). 2018. pp. 116–131.
30. Tan M, Le Q. Efficientnet: rethinking model scaling for convolutional neural networks. In: International conference on machine learning. PMLR. 2019. pp. 6105–6114.
31. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. Rethinking the inception architecture for computer vision. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2016. pp. 2818–2826.
32. Radosavovic I, Kosaraju RP, Girshick R, He K, Dollár P. Designing network design spaces. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2020. pp. 10 428–10 436.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.