Unsupervised Domain Adaptation for Tubular Structure Segmentation Across Different Anatomical Sources

Yuxiang An
yuan5699@uni.sydney.edu.au
Dongnan Liu
dongnan.liu@sydney.edu.au
Weidong Cai
tom.cai@sydney.edu.au

School of Computer Science
The University of Sydney
Sydney, Australia

Abstract

Unsupervised domain adaptation (UDA) aims to boost the generalization ability of deep learning networks by leveraging the unlabeled data, which is a commonly used approach in medical image analysis. However, existing UDA methods only focus on bridging the gap between the same anatomical sites, which leads to the adaptation between different anatomical sites under-explored. UDA under multi-anatomical source settings is more challenging, due to 1) the gray-scale distribution and structural distribution of images of different anatomical sources of tubular structures vary greatly, and 2) The prior knowledge contained in the labels of different anatomical sources is inconsistent. In this paper, we propose an unsupervised domain adaptation method for segmenting tubular structures across different anatomical sources. Specifically, we treat different anatomical sites as different sources. Our method first reduces the domain gap by automatically adjusting the gray-scale distribution of images in different domains using a gray-scale transformation network. Secondly, we introduce target domain noise when training the segmentation network to further improve the segmentation accuracy of the framework in the target domain. In addition, we also design a mask adjustment module for modifying the masks of the raw data in both domains to make the model pay more attention to the common features of the segmentation targets, which further improves the generalization ability of the model. Experimental results on four public datasets at two anatomical sources (eyes and cells) demonstrate the superiority of our proposed method compared with existing state-of-the-art approaches.

1 Introduction

Segmentation of tubular structures such as blood vessels and extracellular membranes is essential in clinical medical image analysis [19, 20, 28]. Vessel segmentation can help determine the location of vascular lesions such as calcification sites [15], and can also provide auxiliary information for bioinformatic recognition such as retinal recognition [7]. Neuronal membrane segmentation is an important step in neural circuit reconstruction [5].

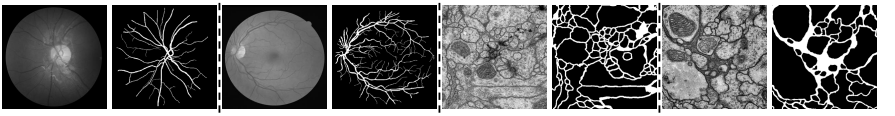

Figure 1: Examples of images and labels from four different sources: (a) the fundus dataset CHASE [9]; (b) the fundus dataset DRIVE [26]; (c) the electron microscopy cell dataset ISBI [2]; and (d) the electron microscopy cell dataset VNC [11]. In each set of examples, the left side is the image and the right side is the label.

Deep learning-based models have been widely used with good results for vessel segmentation in fundus images and neuronal membrane segmentation in electron microscopy images [13, 25], but these methods only perform well on the current dataset. These tasks on tubular structure segmentation require accurate labels during model training which is very time-consuming and error-prone. Therefore, it is necessary to enhance the model's generalization performance for the segmentation across different types of tissues. To address the challenge of limited labeling, unsupervised domain adaptation (UDA) methods that transfer knowledge from the labeled source domain to the unlabeled target domain have been intensively studied and applied to medical image analysis [18].

Recently, many approaches have been proposed to further improve the performance of UDA methods for tubular structure segmentation in medical images. FFO [6] can improve the segmentation accuracy of retinal vessels by fusing features and output space. SFUDA [1] improves the segmentation accuracy of neuronal membranes by acquiring prior knowledge of tubular structures. Although they have achieved great success, they still face the following challenges: first, they only consider domain adaptation for the same anatomical source. The scarcity of medical image data labels makes it important to utilize segmentation data across anatomical sources for segmentation tasks. The significant difference in gray-scale distribution and physiological structural distribution across anatomical sources greatly increases the challenge of cross-anatomical domain adaptation. Secondly, there is background noise that varies greatly between domains across anatomical sources. In addition, the medical prior knowledge contained in the labels of different anatomical sources is not entirely consistent and may be potentially conflicting. Facing the above challenges, we proposed a novel unsupervised domain adaptive tubular structure segmentation framework across anatomical sources. We validated the effectiveness of our framework on two anatomical sites, the eye and the cell. Figure 1 presents the differences in images and labels from different anatomical sources.

The major contributions of our work are summarized as follows: **(1)** We introduced a gray-scale transformation network (GTN) to adjust the gray-scale distribution of different domains to enhance the structural features thereby reducing the gap between different anatomical sources (domains). **(2)** We designed a feature-level noise interference strategy in the segmentation network (SNNI) to reduce background noise that varies greatly between different domains. **(3)** To minimize the discrepancy between the medical prior information of different anatomical sources, we designed a mask adjustment module (MAM) for modifying the tubular structure masks of different anatomical sources. **(4)** To the best of our knowledge, our method is the first to solve the task of UDA segmentation between different anatomical sources. Our method outperforms various UDA methods on the tubular segmentation across different tissues, including fundus, and neurons.

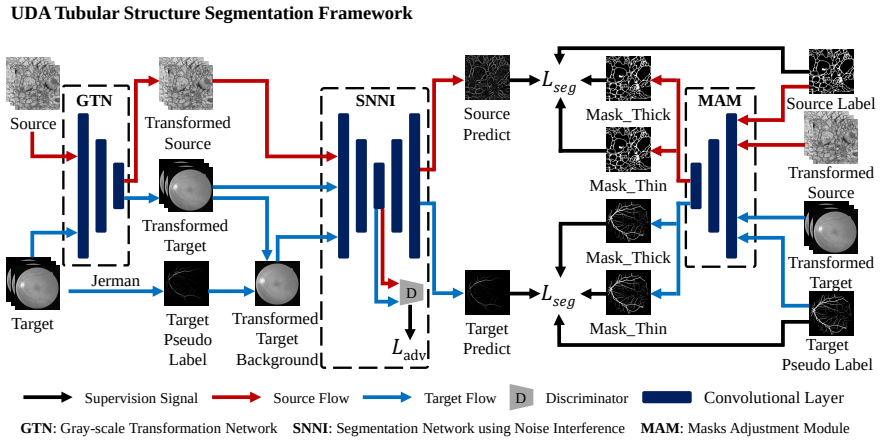

Figure 2: The framework of our proposed unsupervised domain adaptation method for segmentation of tubular structures across anatomical sources.

2 Related Work

Unsupervised domain adaptation methods can effectively improve the generalization ability and robustness of the model. Both feature alignment and image alignment can effectively reduce domain gaps. Feature alignment focuses on domain-invariant features. Image alignment reduces domain differences by transforming the styles of images from two domains. Domain Adversarial Neural Network (DANN) [10] based on the idea of adversarial learning is widely used in UDA methods due to its simple network structure and effectiveness. Many DANN-based network structures are used for medical image analysis [29]. AADG [22] uses a domain generalization method based on data manipulation to reduce the domain gap to achieve retinal vessel segmentation. SCAN [17] leverages the retinal layer structure to facilitate domain alignment to achieve optical coherence tomography (OCT) fluid segmentation. SFUDA [1] uses tubular structural features to promote model focus on domain-invariant features to achieve neuron membrane segmentation. The above-mentioned unsupervised domain adaptation methods have achieved good transfer results in their respective anatomical sources. However, differences in gray-scale distribution in different anatomical sources and inconsistencies in medical prior knowledge limit the applicability of existing UDA methods in different anatomical sources.

3 Methdology

As shown in Figure 2, our framework consists of three parts. The first part is a gray-scale transformation network (GTN) based on gray-scale distribution and structural distribution characteristics, which reduces domain gaps through image alignment. The input for GTN is the original image X, the enhanced image $X_{enhanced}$, and the mask $mask$. The enhanced images $X_{enhanced}$ are obtained following Frangi method [8] and Jerman method [17], which are further explained in Section 3.2. The mask $mask$ is used to provide a region of interest (ROI). The ROI in the fundus image is the part of the eyeball, and the ROI in the cell image is the entire image. The output of GTN is transformed image $X_{transformed}$. The second part is

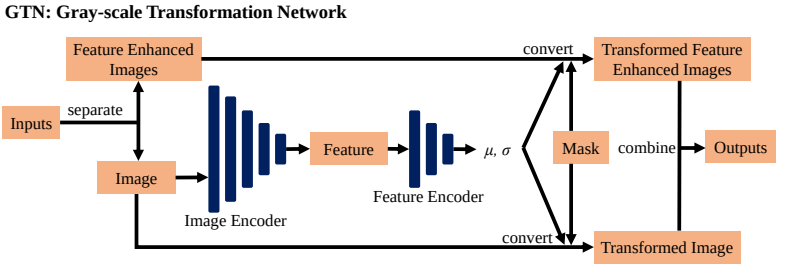

Figure 3: The architecture of our proposed gray-scale transformation network.

an image segmentation network using noise interference (SNNI) for feature alignment. The source domain flow in the segmentation network contains target domain noise. The pseudo label y_p of the target domain is obtained by the Jerman method [17]. The transformed background image of the target domain X_{tb} is obtained from the pseudo label y_p and the image of the target domain after gray-scale transformation $X_{transformed}$. The input for SNNI is the transformed source domain image X_s, the transformed target domain image X_t, and the transformed target domain background image X_{tb}. The output of SNNI is the predicted images of the source domain $\hat{y}_s$ and the target domain $\hat{y}_t$. The third part is the mask adjustment module (MAM) used to reduce the difference in prior knowledge of different anatomical sources. The input for MAM is the transformed image $X_{transformed}$ and label y or pseudo label y_p. The output of MAM includes two masks, $\hat{y}_{thick}$ and $\hat{y}_{thin}$.

3.1 Gray-scale Transformation Network

Due to the different collection methods and collection instruments of datasets for different anatomical sources, as well as the inconsistent structures of different parts, there is a huge gap in the gray-scale distribution and structural distribution of different anatomy images. Such issue makes existing UDA method cannot work, such as the one based on style transfer [23].

To solve this problem, we propose a gray-scale transformation network (GTN) to normalize the source and target images into a new space, by deprecating the domain-specific structural information of these two domains. As shown in Figure 3, our data contains original images X and their corresponding enhanced images $X_{enhanced}$. The input of the Image Encoder is the original image X. After multiple downsampling, the size of the feature map is obtained as $b \times 512 \times 32 \times 32$, where b represents the batch size. Next, the feature map is modified to a feature with size $b \times 2048 \times 1 \times 1$ through convolutional operations. Then, the underlying feature is passed through the multiple linear layers and activation layers to obtain the expected mean eigenvalue μ and standard deviation eigenvalue σ. By removing the influence of pixels outside the mask *mask* on gray-scale information and distribution can enable the GTN to focus more on the pixels within the ROI. Since masks from different sources are different, we only calculate the inside pixels of the mask *mask* when calculating μ and σ of the original image. The transformed images are defined as follows:

$$X_{\text{transformed}} = \begin{cases} \left(\frac{X - \frac{1}{\sum_i \text{mask}_i} \sum_i \text{mask}_i \cdot X_i}{\sqrt{\frac{1}{\sum_i \text{mask}_i} \sum_i \text{mask}_i \cdot (X - \frac{1}{\sum_i \text{mask}_i} \sum_i \text{mask}_i \cdot X_i)^2}} \right) \cdot \sigma + \mu, & \text{mask}_i = 1 \\ 0, & \text{mask}_i = 0 \end{cases}, \quad (1)$$

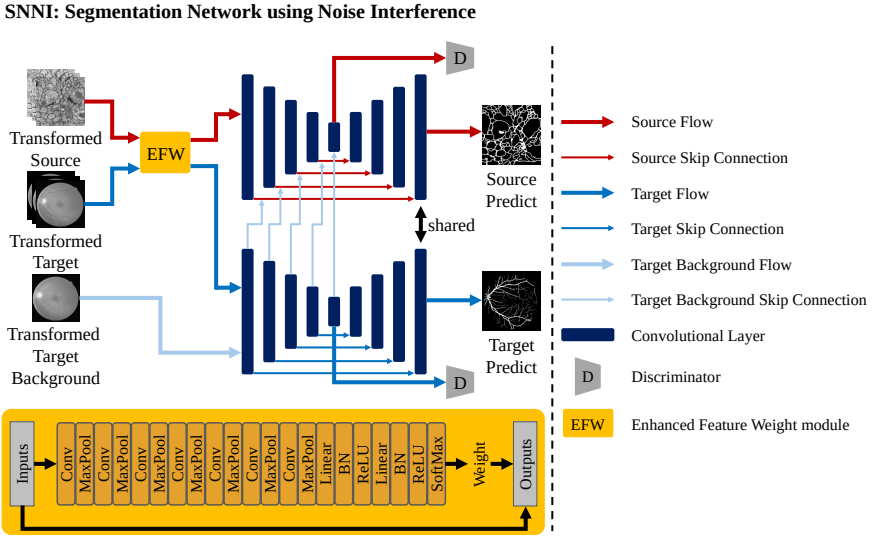

Figure 4: The architecture of our proposed segmentation network using noise interference. We utilize the EFW module [1] to optimize the enhanced image.

where i represents each pixel. To ensure the consistency between $X_{enhanced}$ and X, we use μ and σ to perform gray-scale transformation on both the enhanced images $X_{enhanced}$ and the original image X and combine them as the network output.

3.2 Segmentation Network using Noise Interference

SFUDA [1] demonstrated that tubular structural features as prior knowledge perform well on small dataset segmentation tasks. However, they failed to consider that the accuracy of prior knowledge of tubular structures in different anatomical sources varies greatly and that the prior knowledge is not entirely correct. This phenomenon causes the model to potentially learn incorrect features using incomplete prior knowledge. Therefore, we let the model learn background information adversarially through the background noise of the target domain, reducing the error caused by the model's exclusive focus on incomplete foreground information. The extracellular structures of neurons and retinal vessels are tubular structures. Therefore, we use the Frangi feature extraction method [8] that considers the curvature of the tubular structure and the Jerman feature extraction method [17] that considers different contrasts in different parts of the tubular structure to extract features to obtain the enhanced images $X_{enhanced}$. To address significant gaps in background noise across domains, we introduced noise interference in the target domain during the training process of the source stream.

As shown in Figure 4, our segmentation network using noise interference (SNNI) is established based on [1], which achieves good performance on UDA electron microscopy (EM) tubular segmentation. The input of the network is the original image, enhanced images, and the background image of the target domain, all processed by GTN. Since the Jerman method tends to perform better than the Frangi method, we obtain the background image of the target domain X_{tb} through the image $X_{transformed}$ and the pseudo label y_p obtained by the Jerman

method. y_p is defined as $y_p = Jerman(y)$. The EFW module converts the target domain image from three layers to two layers, while the segmentation network's underlying feature from the target domain image is inputted into the discriminator. The background image of the target domain is directly fed into the segmentation network, yielding five layers of features through downsampling. The source domain image is converted from three layers to two layers through the EFW module and is then put into the SNNI, and the underlying feature is fed into the discriminator. Additionally, we add the five-layer features downsampled from the background image of the target domain to the five-layer features downsampled from the source domain and then participate in the skip connection process of the source domain. This allows the model to optimize segmentation features under the interference of target domain background noise.

we use adversarial learning and self-supervised learning in the target domain. The segmentation loss of the target domain should not have the same weight as the segmentation loss of the source domain due to the inaccuracy of pseudo labels. In addition, to further improve the segmentation accuracy, we provide weight to the domain loss to reduce the negative impact of the recognition of domain differences on segmentation feature extraction. The loss of the segmentation network is defined as follows:

$$L = L_s^{Seg}(y,\hat{y}_s) + \rho L_t^{Seg}(y_p,\hat{y}_t) + \lambda GRL\left(L_s^D(y_{D\_s},\hat{y}_{D\_s}) + L_t^D(y_{D\_t},\hat{y}_{D\_t})\right), \tag{2}$$

where L_s^{Seg} and L_t^{Seg} are the segmentation losses of the source domain and target domain. L_s^D and L_t^D are the domain losses of the source domain and target domain. $\hat{y}_s$ and $\hat{y}_t$ are the predicted images of the source domain and target domain. y_{D_s} and $\hat{y}_{D\_s}$ are the domain label and domain prediction of source domain. y_{D_t} and $\hat{y}_{D\_t}$ are the domain label and domain prediction of target domain. ρ and λ are respectively used to control the impact of target segmentation loss and domain loss on the overall loss. All losses are based on binary cross-entropy loss.

3.3 Masks Adjustment Module

The labels of different anatomical sources contain professional doctors' medical knowledge of the current anatomical source, and the same medical prior information is not completely suitable for different anatomical sources. This medical prior information specific to the current anatomical source is not conducive to UDA segmentation tasks. Based on the idea of adversarial learning, we design the mask adjustment module (MAM) on the source domain and target domain. MAM automatically modifies the annotation y of the original data in the source domain to allow the model to pay more attention to the image itself while removing unhelpful medical prior information. Since the pseudo label y_p for the target domain is not completely correct, the MAM also modifies the pseudo label y_p of the original data in the target domain to reduce the sensitivity of the model to pseudo label error information. There are great differences in the thickness of tubular structures in different anatomical sources. To address such cross-domain differences, we propose to optimize the segmentation performance of the overall algorithm framework by adjusting the thickness of the tubular structure.

As shown in Figure 5, the input of MAM is the image $X_{transformed}$. The image is passed through an encoder similar to the U-Net downsampling structure to obtain two thresholds. One threshold is used to thicken the tubular structure mask, and the other threshold is used to thin the tubular structure mask. Using the idea of adversarial learning to simultaneously

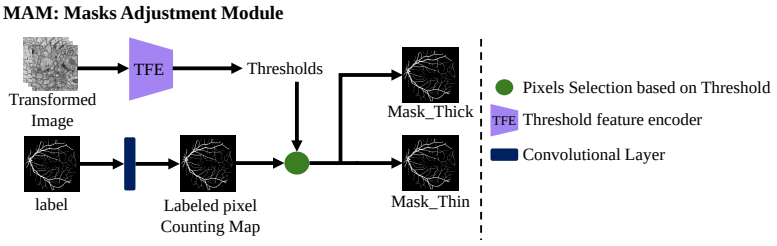

Figure 5: The architecture of our proposed masks adjustment module.

thicken and thin the mask can prevent the generated masks from being too thick or too thin. The thickness of tubular structures varies across different domains. Thickening and thinning the mask for each domain can further mitigate domain differences. Due to the possibility that pixels with similar features may belong to different classes in different domains during cross-domain tasks, the thicker mask $\hat{y}_{thick}$ enables the network to ignore potentially controversial pixels at boundary locations in the cross-domain task, thus better discern the overall difference between the foreground and background. And the thinner mask $\hat{y}_{thin}$ allows the network to focus more precisely on common features shared by all tubular structures, such as the centerline feature.

We perform a convolution operation on the source domain labels y and the target domain pseudo-labels y_p separately to obtain two pixel count maps representing the number of positive samples surrounding each pixel. Small convolution kernels may not be well-suited for the curvature of tubular structures, while large kernels may be less sensitive to the edge information of vessels [16, 30]. Therefore, we adopt a 5×5 kernel size for the convolution. Subsequently, we apply two thresholds (coarsening and thinning) on the positive pixel count map to derive coarsening and thinning masks. The loss computed in each domain includes two masks and one label. The label is used to calculate the segmentation loss, and the thickening mask $\hat{y}_{thick}$ and thinning mask $\hat{y}_{thin}$ are used to fine-tune the segmentation loss. Therefore, when we calculate the loss, we need to reduce the weight of the mask to reduce the impact on the correct segmentation result. The segmentation loss of the source domain L_s^{Seg} and the segmentation loss of the target domain L_t^{Seg} are as follows:

$$L_s^{Seg} = L_s(y, \hat{y}_s) + \eta(L_s(y, \hat{y}_{thick\_s}) + L_s(y, \hat{y}_{thin\_s})), \tag{3}$$

$$L_t^{Seg} = L_t(y_p, \hat{y}_t) + \eta(L_t(y_p, \hat{y}_{thick\_t}) + L_t(y_p, \hat{y}_{thin\_t})), \tag{4}$$

where η is controllable parameter. $\hat{y}_{thick\_s}$ and $\hat{y}_{thin\_s}$ are the two masks of source domain. $\hat{y}_{thick\_t}$ and $\hat{y}_{thin\_t}$ are the two masks of target domain. y is the lable of the source domain. y_p is the pseudo label of the target domain. L represents binary cross-entropy loss.

4 Experiments

4.1 Datasets and Implementation Details

The experiment aims to segment tubular structures across anatomical sources. The experiment includes four publicly available datasets from two anatomical sources. These consist

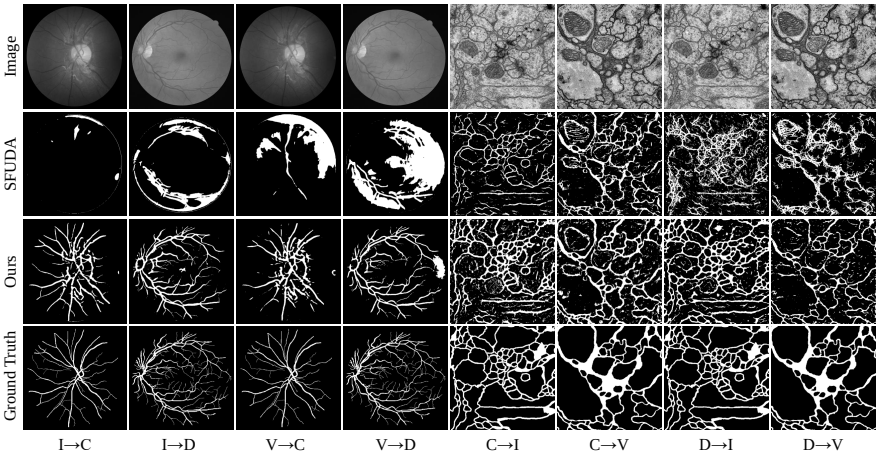

Figure 6: Visualization examples of comparative experiments. I: ISBI, V: VNC, C: CHASE, D: DRIVE.

of two cell datasets: ISBI 2012 EM Segmentation Challenge [2] (ISBI) and Neural Tissue Segmented Anisotropy ssTEM dataset [10] (VNC), and two eye datasets: Digital Retinal Images for Vessel Extraction [26] (DRIVE) and a subset of retinal images of multiethnic children from the Child Heart and Health Study in England [9] (CHASE). The ISBI dataset has 30 2D data with a size of 512×512, of which 20 are used for training and 10 are used for testing. The VNC dataset has 20 2D data with a size of 1024×1024. We cut it into 80 images of 512×512, 60 for training and 20 for testing. The DRIVE dataset has 40 2D data with a size of 584×565, of which 20 are used for training and 20 are used for testing. The CHASE dataset has 28 2D data with a size of 999×960, 20 for training and 8 for testing. Since the input sizes of different fields in the model should be the same, we used the cubic interpolation method to adjust the data size of the four datasets to 512×512. Due to the imbalance in the number of samples used for training in different datasets, we increased each of the four datasets used for training to 100 by rotating and flipping.

We used Adam optimizer. The initial learning rate is set to 0.001 with a weight decay rate of 0.0001. Furthermore, the learning rate decreases by a factor of 0.9 every 10 epochs. ρ, λ, η are set to 0.8, 0.1, and 0.2 respectively. Our model is trained for a total of 50 epochs, and the evaluation metric employed is the Dice similarity coefficient. Our framework leverages GTN, SNNI, and MAM during the training phase, while only GTN and SNNI are used during the testing phase.

4.2 Comparison Experiments

To verify the superiority of the proposed method, we conduct a comprehensive comparison with 9 unsupervised domain adaptation methods on four datasets, including DANN [10], UMDA-SNA [21], DCDA [23], SAM-UDA [4], ADANet [31], FFO [27], SFUDA [1], MIC [14], and LA-UDA [3]. DANN is a very popular basic architecture used in UDA methods based on the idea of adversarial learning. UMDA-SNA adds spatial neural attention (SNA) to the classic adversarial domain adaptation network to achieve domain adaptation. DCDA uses a disentangling representation style transfer module and a collaborative consis-

Table 1: Comparison experiments. I: ISBI, V: VNC, C: CHASE, D: DRIVE.

Method	I→C	I→D	V→C	V→D	C→I	C→V	D→I	D→V	Average
U-Net(No adaptation) [24]	7.79%	2.88%	16.56%	22.84%	36.13%	20.13%	36.79%	27.76%	21.36%
U-Net(Supervised) [24]	80.59%	80.87%	80.59%	80.87%	78.73%	88.30%	78.73%	88.30%	82.12%
DANN [10]	12.34%	13.75%	13.43%	19.27%	57.88%	62.81%	47.88%	44.84%	34.03%
UMDA-SNA [21]	15.87%	10.85%	13.75%	21.35%	42.36%	32.24%	37.70%	19.57%	24.21%
DCDA [23]	15.92%	14.17%	16.33%	20.16%	61.23%	63.06%	46.59%	45.28%	35.34%
SAM-UDA [9]	14.10%	10.78%	13.92%	16.89%	43.43%	39.95%	40.35%	22.83%	25.28%
ADANet [31]	11.43%	9.69%	14.31%	12.49%	65.40%	36.48%	41.69%	23.56%	26.88%
FFO [22]	19.73%	18.13%	14.80%	20.04%	48.57%	36.67%	46.90%	35.33%	30.02%
SFUDA [9]	13.79%	26.40%	14.80%	26.43%	54.44%	64.75%	54.15%	59.51%	39.28%
MIC [14]	15.42%	10.86%	14.85%	15.13%	61.85%	68.74%	59.14%	64.57%	38.82%
LA-UDA [8]	36.76%	39.70%	24.67%	26.34%	56.19%	44.15%	45.29%	47.80%	40.11%
Ours	**60.46%**	**67.11%**	**53.93%**	**61.68%**	**67.52%**	**70.84%**	**68.05%**	**69.94%**	**64.94%**

tency learning module to assist domain transfer. SAM-UDA improves the performance of the UDA method by focusing on the correlation between pixels through the self-attention mechanism. ADANet uses a discriminator based on atrous convolution to improve its discriminative ability and optimize segmentation performance. FFO uses two discriminators to achieve UDA retinal vessel segmentation by fusing features and output space. SFUDA utilizes priori knowledge of tubular structures to achieve neuronal membrane segmentation. MIC enhances global semantic information correlation using masked image consistency. LA-UDA accomplishes domain adaptation by training the augmentation module and classifier in two phases, respectively.

For a fair comparison, we use the same U-Net backbone network as the segmentation network in these methods. We show the upper and lower bounds of the U-Net metric without domain adaptation. Our framework achieves considerable performance improvements, outperforming mainstream unsupervised domain adaptation methods in terms of metrics. Figure 6 is the visualization of the comparison results between SFUDA as the baseline method and our method. The visualization of the complete comparison results are in the supplementary materials. Table 1 quantitatively shows the difference in indicators of different UDA methods, indicating that our method produces more accurate segmentation results. Experimental results show that existing unsupervised domain adaptation methods initially proposed for a single anatomical source do not perform well when applied to different anatomical sources. This is due to the significant differences in data characteristics across anatomical sources, and the failure of these methods to provide clear feature constraints and assistance for segmentation objectives.

4.3 Ablation Studies

We verify the effectiveness of the proposed modules in our framework through ablation experiments. We use the segmentation network in [9] as the baseline and attach the proposed modules to the framework respectively. It shows that our method improves segmentation accuracy through three modules. Table 2 shows the quantitative results of the ablation experiments. The results demonstrate the effectiveness of our proposed three modules. Through ablation experiments, we also found that there are cases where the performance of using GTN and MAM is not as high as using either of these two modules alone. We posit that utilizing the first and third modules encourages the model to refine its segmentation ability

Table 2: Ablation studies. I: ISBI, V: VNC, C: CHASE, D: DRIVE.

Method	I-C	I-D	V-C	V-D	C-I	C-V	D-I	D-V
Baseline	13.79%	26.40%	14.80%	26.43%	54.44%	64.75%	54.15%	59.51%
Baseline+GTN	51.92%	60.62%	36.47%	59.00%	53.26%	57.48%	56.42%	58.82%
Baseline+SNNI	56.60%	63.23%	50.79%	60.96%	54.79%	63.06%	62.36%	64.99%
Baseline+MAM	55.06%	34.30%	43.36%	36.24%	65.89%	57.01%	60.02%	62.03%
Baseline+GTN+SNNI	55.98%	64.38%	52.10%	61.29%	55.42%	60.95%	61.22%	65.73%
Baseline+GTN+MAM	34.28%	57.25%	27.67%	57.48%	64.29%	63.60%	64.47%	62.53%
Baseline+SNNI+MAM	59.35%	64.48%	53.32%	59.24%	64.54%	66.66%	63.04%	65.92%
Ours	**60.46%**	**67.11%**	**53.93%**	**61.68%**	**67.52%**	**70.84%**	**68.05%**	**69.94%**

within the source domain without enhancing the segmentation network, thereby diminishing the segmentation performance within the target domain.

5 Conclusion

In this work, we proposed a novel domain-adaptive tubular structure segmentation framework across anatomical sources. Compared with other UDA methods, our method has achieved great success. We designed GTN, SNNI, and MAM to face the task of segmenting tubular structures across anatomical sources. We achieved domain adaptation from both the image alignment and the feature alignment, that is, the differences in gray distribution and background noise in different anatomical sources. In addition, we further reduce the domain gap caused by inconsistent medical prior knowledge of masks from different anatomical sources. To the best of our knowledge, we are in the early stages of trying to implement unsupervised domain adaptation methods across anatomical sources on medical images. Due to the attractive performance of our method, it can be expanded for future applications in multi-anatomical segmentation tasks.

References

[1] Yuxiang An, Dongnan Liu, and Weidong Cai. Unsupervised domain adaptation for neuron membrane segmentation based on structural features. In *IEEE International Conference on Multimedia and Expo*, pages 948–953, 2023.

[2] Albert Cardona, Stephan Saalfeld, Stephan Preibisch, Benjamin Schmid, Anchi Cheng, Jim Pulokas, Pavel Tomancak, and Volker Hartenstein. An integrated micro-and macroarchitectural analysis of the drosophila brain by computer-assisted serial section electron microscopy. *PLoS Biology*, 8(10):e1000502, 2010.

[3] Julio Ivan Davila Carrazco, Pietro Morerio, Alessio Del Bue, and Vittorio Murino. Learnable data augmentation for one-shot unsupervised domain adaptation. In *British Machine Vision Conference*, 2023.

[4] Inseop Chung, Jayeon Yoo, and Nojun Kwak. Exploiting inter-pixel correlations in unsupervised domain adaptation for semantic segmentation. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 12–21, 2023.

[5] Dan Ciresan, Alessandro Giusti, Luca Gambardella, and Jürgen Schmidhuber. Deep neural networks segment neuronal membranes in electron microscopy images. *Advances in Neural Information Processing Systems*, 25, 2012.

[6] Wei Feng, Lie Ju, Lin Wang, Kaimin Song, Xin Zhao, and Zongyuan Ge. Unsupervised domain adaptation for medical image segmentation by selective entropy constraints and adaptive semantic alignment. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, pages 623–631, 2023.

[7] Isabel N Figueiredo, Susana Moura, Júlio S Neves, Luís Pinto, Sunil Kumar, Carlos M Oliveira, and João D Ramos. Automated retina identification based on multiscale elastic registration. *Computers in Biology and Medicine*, 79:130–143, 2016.

[8] Alejandro F Frangi, Wiro J Niessen, Koen L Vincken, and Max A Viergever. Multiscale vessel enhancement filtering. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 130–137, 1998.

[9] Muhammad Moazam Fraz, Paolo Remagnino, Andreas Hoppe, Bunyarit Uyyanonvara, Alicja R Rudnicka, Christopher G Owen, and Sarah A Barman. An ensemble classification-based approach applied to retinal blood vessel segmentation. *IEEE Transactions on Biomedical Engineering*, 59(9):2538–2548, 2012.

[10] Yaroslav Ganin, Evgeniya Ustinova, Hana Ajakan, Pascal Germain, Hugo Larochelle, François Laviolette, Mario March, and Victor Lempitsky. Domain-adversarial training of neural networks. *Journal of Machine Learning Research*, 17(59):1–35, 2016.

[11] Stephan Gerhard, Jan Funke, Julien Martel, Albert Cardona, and Richard Fetter. Segmented anisotropic sstem dataset of neural tissue. *Figshare*, 2013.

[12] Xingxin He, Zhun Zhong, Leyuan Fang, Min He, and Nicu Sebe. Structure-guided cross-attention network for cross-domain oct fluid segmentation. *IEEE Transactions on Image Processing*, 32:309–320, 2022.

[13] Fabian Hörst, Moritz Rempe, Lukas Heine, Constantin Seibold, Julius Keyl, Giulia Baldini, Selma Ugurel, Jens Siveke, Barbara Grünwald, Jan Egger, et al. Cellvit: Vision transformers for precise cell segmentation and classification. *Medical Image Analysis*, 94:103143, 2024.

[14] Lukas Hoyer, Dengxin Dai, Haoran Wang, and Luc Van Gool. Mic: Masked image consistency for context-enhanced domain adaptation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 11721–11732, 2023.

[15] Yi Huang, Wenjun Yan, Menghua Xia, Yi Guo, Guohui Zhou, and Yuanyuan Wang. Vessel membrane segmentation and calcification location in intravascular ultrasound images using a region detector and an effective selection strategy. *Computer Methods and Programs in Biomedicine*, 189:105339, 2020.

[16] David H Hubel and Torsten N Wiesel. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of Physiology*, 160(1):106, 1962.

[17] Tim Jerman, Franjo Pernuš, Boštjan Likar, and Žiga Špiclin. Enhancement of vascular structures in 3d and 2d angiographic images. *IEEE Transactions on Medical Imaging*, 35(9):2107–2118, 2016.

[18] Canran Li, Dongnan Liu, Haoran Li, Zheng Zhang, Guangming Lu, Xiaojun Chang, and Weidong Cai. Domain adaptive nuclei instance segmentation and classification via category-aware feature alignment and pseudo-labelling. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 715–724, 2022.

[19] Fengming Lin, Yan Xia, Nishant Ravikumar, Qiongyao Liu, Michael MacRaild, and Alejandro F Frangi. Adaptive semi-supervised segmentation of brain vessels with ambiguous labels. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 106–116. Springer, 2023.

[20] Dongnan Liu, Donghao Zhang, Yang Song, Chaoyi Zhang, Heng Huang, Mei Chen, and Weidong Cai. Large kernel refine fusion net for neuron membrane segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 2212–2220, 2018.

[21] Jinping Liu, Hui Liu, Subo Gong, Zhaohui Tang, Yongfang Xie, Huazhan Yin, and Jean Paul Niyoyita. Automated cardiac segmentation of cross-modal medical images using unsupervised multi-domain adaptation and spatial neural attention structure. *Medical Image Analysis*, 72:102135, 2021.

[22] Junyan Lyu, Yiqi Zhang, Yijin Huang, Li Lin, Pujin Cheng, and Xiaoying Tang. Aadg: automatic augmentation for domain generalization on retinal image segmentation. *IEEE Transactions on Medical Imaging*, 41(12):3699–3711, 2022.

[23] Linkai Peng, Li Lin, Pujin Cheng, Ziqi Huang, and Xiaoying Tang. Unsupervised domain adaptation for cross-modality retinal vessel segmentation via disentangling representation style transfer and collaborative consistency learning. In *IEEE International Symposium on Biomedical Imaging*, pages 1–5, 2022.

[24] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 234–241, 2015.

[25] Jihyoung Ryu, Mobeen Ur Rehman, Imran Fareed Nizami, and Kil To Chong. Segrnet: A deep learning framework with multi-scale feature fusion for robust retinal vessel segmentation. *Computers in Biology and Medicine*, 163:107132, 2023.

[26] Joes Staal, Michael D Abràmoff, Meindert Niemeijer, Max A Viergever, and Bram Van Ginneken. Ridge-based vessel segmentation in color images of the retina. *IEEE Transactions on Medical Imaging*, 23(4):501–509, 2004.

[27] Shengsheng Wang, Zihao Fu, Bilin Wang, and Yulong Hu. Fusing feature and output space for unsupervised domain adaptation on medical image segmentation. *International Journal of Imaging Systems and Technology*, 33(5):1672–1681, 2023.

[28] Yicheng Wu, Yong Xia, Yang Song, Donghao Zhang, Dongnan Liu, Chaoyi Zhang, and Weidong Cai. Vessel-net: Retinal vessel segmentation under multi-path supervision. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd*

International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part I 22, pages 264–272, 2019.

[29] Qianbi Yu, Dongnan Liu, Chaoyi Zhang, Xinwen Zhang, and Weidong Cai. Unsupervised domain adaptive fundus image segmentation with few labeled source data. In *British Machine Vision Conference*, 2022.

[30] Matthew D Zeiler and Rob Fergus. Visualizing and understanding convolutional networks. In *European Conference on Computer Vision*, pages 818–833, 2014.

[31] Ting Zhang, Zihang Gao, Zhaoying Liu, Syed Fawad Hussain, Muhammad Waqas, Zahid Halim, and Yujian Li. Infrared ship target segmentation based on adversarial domain adaptation. *Knowledge-Based Systems*, 265:110344, 2023.