Student Becomes Decathlon Master in Retinal Vessel Segmentation via Dual-Teacher Multi-target Domain Adaptation

Linkai Peng[1], Li Lin[1,2], Pujin Cheng[1], Huaqing He[1], and Xiaoying Tang[1,3(✉)]

[1] Department of Electrical and Electronic Engineering, Southern University
of Science and Technology, Shenzhen, China
tangxy@sustech.edu.cn
[2] Department of Electrical and Electronic Engineering,
The University of Hong Kong, Hong Kong SAR, China
[3] Jiaxing Research Institute, Southern University of Science and Technology,
Jiaxing, China
tangxy@sustech.edu.cn

Abstract. Unsupervised domain adaptation has been proposed recently to tackle the so-called domain shift between training data and test data with different distributions. However, most of them only focus on single-target domain adaptation and cannot be applied to the scenario with multiple target domains. In this paper, we propose RVms, a novel unsupervised multi-target domain adaptation approach to segment retinal vessels (RVs) from multimodal and multicenter retinal images. RVms mainly consists of a style augmentation and transfer (SAT) module and a dual-teacher knowledge distillation (DTKD) module. SAT augments and clusters images into source-similar domains and source-dissimilar domains via Bézier and Fourier transformations. DTKD utilizes the augmented and transformed data to train two teachers, one for source-similar domains and the other for source-dissimilar domains. Afterwards, knowledge distillation is performed to iteratively distill different domain knowledge from teachers to a generic student. The local relative intensity transformation is employed to characterize RVs in a domain invariant manner and promote the generalizability of teachers and student models. Moreover, we construct a new multimodal and multicenter vascular segmentation dataset from existing publicly-available datasets, which can be used to benchmark various domain adaptation and domain generalization methods. Through extensive experiments, RVms is found to be very close to the target-trained Oracle in terms of segmenting the RVs, largely outperforming other state-of-the-art methods.

Keywords: Multi-target domain adaptation · Dual teacher · Knowledge distillation · Style transfer · Retinal vessel segmentation

L. Peng and L. Lin—Contributed equally to this work.

1   Introduction

Multimodal ophthalmic images can be effectively employed to extract retinal structures, identify biomarkers, and diagnose diseases. For example, fundus images can depict critical anatomical structures such as the macula, optic disc, and retinal vessels (RVs) [3,11]. Optical coherence tomography angiography (OCTA) can efficiently and accurately generate volumetric angiography images [14,20]. These two modalities can both deliver precise representations of vascular structures within the retina, making them popular for the diagnoses of eye-related diseases. These years, advance in the imaging technologies brings new ophthalmic modalities including optical coherence tomography (OCT) [10], widefield fundus [4], and photoacoustic images [9]. These new modalities also provide important information of RVs and are likely to enhance the disease diagnosis accuracy. In retinal disease analysis, RV segmentation is a very important pre-requisite. Manual delineation is the most accurate yet highly labor-intensive and time-consuming manner, especially for RV segmentation across multiple modalities. One plausible solution is to manually label RVs on images of a single modality and then transfer the labels to other modalities of interest via advanced image processing and deep learning techniques. However, researches have shown that deep learning models trained on one domain generally perform poorly when tested on another domain with different data distribution, because of domain shift [22]. This indicates that segmentation models trained on existing modalities will have low generalizability on new modalities. Furthermore, there also exist variations within images of the same modality because of different equipments and imaging settings, posing more challenges to a generalized RV segmentation model for images with cross-center or cross-modality domain shift.

To address this issue, domain adaptation has been explored, among which unsupervised domain adaptation (UDA) has gained the most favor in recent years. For instance, Javanmardi et al. [12] combined U-net [21] with a domain discriminator and introduced adversarial training based on DAAN [6]. Wang et al. [28] adopted this design to joint optic disc and cup segmentation from fundus images. These two methods focus on domain shift with small variations (i.e., cross-center domain shift). To tackle domain shift with large variations (i.e., cross-modality domain shift), Cai et al. [2] incorporated CycleGAN [32] into its own network and employed a shape consistency loss to ensure a correct translation of semantics in medical images. Dou et al. [5] designed a cross-modality UDA framework for cardiac MR and CT image segmentation by performing DA only at low-level layers, under the assumption that domain shift mainly lies in low-level characteristics. Zhang et al. [30] presented a Noise Adaptation Generative Adversarial Network for RV segmentation and utilized a style discriminator enforcing the translated images to have the same noise patterns as those in the target domain. Peng et al. [19] employed disentangled representation learning to disentangle images into content space and style space to extract domain-invariant features. These approaches are nevertheless designed to handle only a specific type of domain shift and may have limitations when applied to multimodal and multicenter ($m\&m$) ophthalmic images.

We here propose a novel unsupervised multi-target domain adaptation (MTD A) approach, RVms, to segment RVs from $m\&m$ retinal images. RVms mainly consists of two components: a style augmentation and transfer (SAT) module and a dual-teacher knowledge distillation (DTKD) module. SAT uses Bézier transformation to augment source images into diverse styles. By adjusting the interpolation rate, source-style images are gradually transferred to target styles via Fourier Transform. We then cluster images into source-similar domains D_{sim} and source-dissimilar domains D_{dis} according to the vessel-background relative intensity information. DTKD utilizes the augmented and transformed data to train two teachers, one for D_{sim} and the other for D_{dis}. Then knowledge distillation [8] is performed to iteratively distill different domain knowledge from teachers to a generic student. In addition, the local relative intensity transformation (LRIT) is employed to characterize RVs in a domain invariant manner and to improve the generalizability of teachers and student models.

The main contributions of this work are four-fold: (1) We propose a novel MTDA framework for RV segmentation from $m\&m$ retinal images. To the best of our knowledge, this is the first work that explores unsupervised MTDA for medical image segmentation across both modalities and centers. (2) MTDA features itself with style augmentation, style transfer, and dual-teacher knowledge distillation, which effectively tackles both cross-center and cross-modality domain shift. (3) Extensive comparison experiments are conducted, successfully identifying the proposed pipeline's superiority over representative state-of-the-art (SOTA) methods. (4) We construct a new $m\&m$ vascular segmentation dataset from existing publicly-available datasets, which can be used to benchmark various domain adaptation and domain generalization methods. Our code and dataset are publicly available at https://github.com/lkpengcs/RVms.

2   Methodology

The proposed RVms framework is shown in Fig. 1, which consists of a SAT module and a DTKD module.

2.1   Definition

We define the source domain as $S = \{x_i^s, y_i^s\}_{i=1}^{N^s}$, where x_i^s is the i-th input image, y_i^s is the corresponding RV segmentation label, and N^s is the total number of source domain images. The target domains are denoted as $T = \{T_1, T_2, \ldots T_n\}$, where $T_t = \{x_j^t, y_j^t\}_{j=1}^{N^t}$, x_j^t is the j-th image of the t-th target domain, y_j^t is the corresponding RV segmentation label, and N^t is the total number of images from the t-th target domain. From SAT, we obtain images in source-similar domains D_{sim} and source-dissimilar domains D_{dis}. Then in DTKD, a source-similar teacher T_{sim} and a source-dissimilar teacher T_{dis} are separately trained. Finally, we employ T_{sim} and T_{dis} to distill knowledge to a generic student S_g.

Fig. 1. Schematic demonstration of the architecture of our RVms framework. The upper part represents the LRIT module and the lower part represents the DTKD module. $P(\cdot)$ is the prediction of teachers and $Q(\cdot)$ is the prediction of the student.

2.2 Style Augmentation and Transfer

Style Augmentation. Retinal images are either grayscale or can be converted to grayscale and the content across different images is generally similar; the difference mainly lies in the intensity and the clarity of tiny vessels. Besides, the style in different modalities is also a key factor that leads to domain shift, especially for small-variation domain shift. Thus, we follow the work of [31] and utilize a non-linear transformation via the monotonic and smooth Bézier Curve function. It is a one-to-one mapping that assigns each pixel a new and unique value. We generate the Bézier Curve from two end points (P_0 and P_3) and two control points (P_1 and P_2), which is defined as

$$B(t) = (1-t)^3 P_0 + 3(1-t)^2 t P_1 + 3(1-t)t^2 P_2 + t^3 P_3, \ t \in [0,1], \qquad (1)$$

where t is a fractional value along the length of the line. We set $P_0 = (0,0)$ and $P_3 = (1,1)$ to get source-similar augmentations and the opposite to get source-dissimilar augmentations. The x-axis coordinates and y-axis coordinates of P_1 and P_2 are randomly selected from the interval $(0,1)$.

Style Transfer. In addition to grayscale differences, different modalities also differ in the image style, such as vessel shape and existence of other anatomical landmarks. To imitate the diverse appearance of different target domains, we adopt Fourier Transform to progressively translate the style of source domain to those of target domains. Given a source image x_i^s and a randomly selected target image x_j^t from target domain T_t, we first perform Fourier Transform on both images to get amplitude spectrums $\mathcal{A}_s, \mathcal{A}_t$ and phase spectrums $\mathcal{P}_s, \mathcal{P}_t$ [16,29]. Then we use a binary mask $\mathcal{M} = \mathbb{1}_{(h,w)\in[-\alpha H:\alpha H, -\alpha W:\alpha W]}$ to extract the central regions of $\mathcal{A}_s$
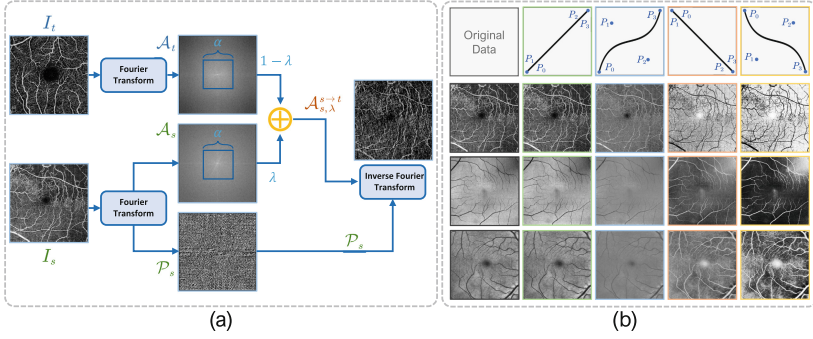
Fig. 2. Schematic demonstration of our proposed SAT module. Panel (a) is the style transfer module and Panel (b) is the style augmentation module.

and $\mathcal{A}_t$ and combine them with an interpolation rate λ. In this way the contributions of $\mathcal{A}_s$ and $\mathcal{A}_t$ to the synthesized image can be adjusted.

$$\mathcal{A}_{s,\lambda}^{s\rightarrow t} = ((1-\lambda)\mathcal{A}_s + \lambda\mathcal{A}_t) * \mathcal{M} + \mathcal{A}_s * (1-\mathcal{M}). \tag{2}$$

Finally, we perform inverse Fourier Transform to obtain an image with content from source domain and style from both source and target domains (Eq. 3). Detailed illustration of SAT is shown in Fig. 2.

$$x_{i,\lambda}^{s\rightarrow t} = \mathcal{F}^{-1}\left(\mathcal{A}_{s,\lambda}^{s\rightarrow t}, \mathcal{P}_s\right). \tag{3}$$

2.3   Dual-Teacher Knowledge Distillation

Local Relative Intensity Transformation. In our situation, despite that retinal vessels are in different modalities and multi-scaled, they always have a consistent relationship with the background in grayscale intensity. This relationship can be used to depict vessels in a domain-invariant manner. We employ LRIT to extract RV features by taking advantage of this vessel-background relationship. Inspired by [18,24], each pixel serves as an anchor point and the intensity values of the adjacent eight pixels are compared to generate a new value for the anchor point through the formula in Eq. 4. Adjacent pixels in all four directions (i.e., up, down, left, and right) are compared separately, resulting in four transformed images.

$$
\begin{aligned}
V_{new}(a) &= \sum_{i=1}^{8} c\left(V(a) - V(n_i)\right) \times 2^i, \\
c(x) &= \begin{cases} 1, & \text{if } x > 0 \\ 0, & \text{otherwise}, \end{cases}
\end{aligned}
\tag{4}
$$

where $V(\cdot)$ is the intensity value and n_i is the i-th adjacent pixel of the anchor point. For pixels near edges, we use edge padding to ensure there are eight neighboring pixels in each direction. The four transformed images are shuffled and concatenated as external channels for both teachers and student models.

Knowledge Distillation. The dual teachers T_{sim} and T_{dis} are respectively trained with augmented and style-transferred images in D_{sim} and D_{dis}. The source images, style augmented images, and style transferred images are used as inputs and the corresponding labels are used for supervision via Dice losses $\mathcal{L}_{seg}^{sim}$ and $\mathcal{L}_{seg}^{dis}$. When the two teachers are trained to converge (after τ epochs), we start the knowledge distillation process. The domain-generic student model S_g is supervised with the outputs of T_{sim} and T_{dis}, each being responsible for one type of the target domains (source-similar or source-dissimilar). Specifically, we adopt a cross-entropy loss $\mathcal{L}_{kd}$ to minimize the distribution differences between the corresponding outputs of the teachers and student models (i.e., P_{sim} and Q_{sim}, P_{dis} and Q_{dis}). The groundtruth is also used to supervise the training on the source domain through a Dice loss $\mathcal{L}_{seg}^s$. The total loss function is

$$\mathcal{L}_{DTKD} = \begin{cases} \mathcal{L}_{seg}^{sim} + \mathcal{L}_{seg}^{dis}, & \text{epoch} \leqslant \tau \\ \mathcal{L}_{kd} + \mathcal{L}_{seg}^s + \mathcal{L}_{seg}^{sim} + \mathcal{L}_{seg}^{dis}, & \text{epoch} > \tau. \end{cases} \tag{5}$$

3   Experiments and Results

3.1   Dataset

We construct a new dataset named mmRV, consisting of five domains from five publicly available datasets [1,4,13,17,25]. The details are shown in Table A1 of the appendix. For OCTA-500, we resize each image to 384×384 and discard samples with severe quality issues. For fundus images, we crop out the micro-vascular region surrounding the macula to avoid interference from other unrelated structures such as the optic disc and only focus on vessels; the capillaries near the macula are mostly multi-scaled and difficult to segment [15]. We also apply Contrast Limited Adaptive Histogram Equalization to fundus images as preprocessing. For PRIME-FP20, we augment each image for four times.

3.2   Experimental Setting

We train our RVms framework on the newly-constructed mmRV dataset. We conduct two sets of MTDA experiments respectively using DRIVE and OCTA-500 as the source domain because the image modalities in those two datasets are the most commonly used in clinical practice and those two datasets are most well-annotated. All compared methods and RVms are implemented with Pytorch using NVIDIA RTX 3090 GPUs. For both teachers and student models in RVms, we use the Adam optimizer with a learning rate of 1×10^{-3}. The number of epochs τ for training the dual teachers is set to be 200 and then we co-train the teachers with the student for another 400 epochs. α is set to 0.2 and λ is randomly selected in $(0, 1)$. During testing, the target domain images are directly inputted to S_g to get the corresponding predictions. We use ResNet34 [7] with ImageNet pretrained initialization as the encoder in a modified U-net [21] architecture.

Table 1. Quantitative evaluations of different methods. **Bold** and underlined numbers respectively denote the best and second-best results.

	Source domain: OCTA (OCTA-500)											
Modalities	Fundus image (DRIVE)		OCTA (ROSE)		OCT (OCTA-500)		Fundus image (HRF)		UWF fundus (PRIME-FP20)		Average	
Metrics	Dice ↑	HD ↓	Dice ↑	HD ↓	Dice ↑	HD ↓	Dice ↑	HD ↓	Dice ↑	HD ↓	Dice ↑	HD ↓
Source only	2.18	178.06	52.35	25.52	5.70	62.89	0.43	189.60	1.33	180.42	12.40	127.30
Dofe	24.03	27.91	48.52	21.21	34.41	21.76	24.67	25.63	30.37	24.75	32.40	24.25
ADVENT	1.46	52.53	52.86	12.80	4.91	44.48	0.37	44.16	1.01	20.11	12.12	34.81
Multi-Dis	1.45	39.65	58.98	10.29	5.07	30.77	0.31	46.91	1.03	52.09	13.37	35.94
T_{sim}	5.35	36.18	53.04	23.27	2.92	32.89	2.90	30.75	5.07	37.08	13.85	39.27
T_{dis}	**72.77**	**10.77**	5.68	19.52	**77.76**	15.60	72.90	12.12	76.49	12.57	49.73	**14.16**
Ours	72.64	17.72	**60.80**	13.49	72.81	24.70	74.03	12.97	74.54	21.09	**70.96**	17.99
Oracle	67.72	9.05	71.43	12.14	82.68	11.13	72.09	12.67	77.80	9.33	74.34	10.86
	Source domain: fundus image (DRIVE)											
Modalities	OCTA (OCTA-500)		OCTA (ROSE)		OCT (OCTA-500)		Fundus image (HRF)		UWF fundus (PRIME-FP20)		Average	
Metrics	Dice ↑	HD ↓	Dice ↑	HD ↓	Dice ↑	HD ↓	Dice ↑	HD ↓	Dice ↑	HD ↓	Dice ↑	HD ↓
Source only	11.39	21.81	17.69	10.24	59.24	19.77	58.77	**10.38**	64.87	14.92	42.39	**15.42**
Dofe	30.94	22.72	38.39	11.48	62.69	19.13	60.28	12.86	65.46	13.88	51.55	16.01
ADVENT	22.38	47.40	20.36	33.88	47.32	42.66	53.15	34.59	52.86	42.68	39.21	40.24
Multi-Dis	32.75	21.30	33.73	11.16	52.42	35.79	59.46	28.22	60.62	38.54	47.80	27.00
T_{sim}	12.05	19.58	14.32	23.08	59.11	21.19	59.15	15.37	64.96	17.69	41.91	19.38
T_{dis}	64.44	18.51	**63.81**	16.13	11.01	**14.63**	6.87	25.28	10.08	24.69	31.24	19.84
Ours	**68.93**	15.67	61.31	16.05	64.35	18.93	63.18	13.50	**69.44**	14.04	**65.44**	15.63
Oracle	88.35	4.21	71.43	12.14	82.68	11.13	72.09	12.67	77.80	9.33	78.47	9.89
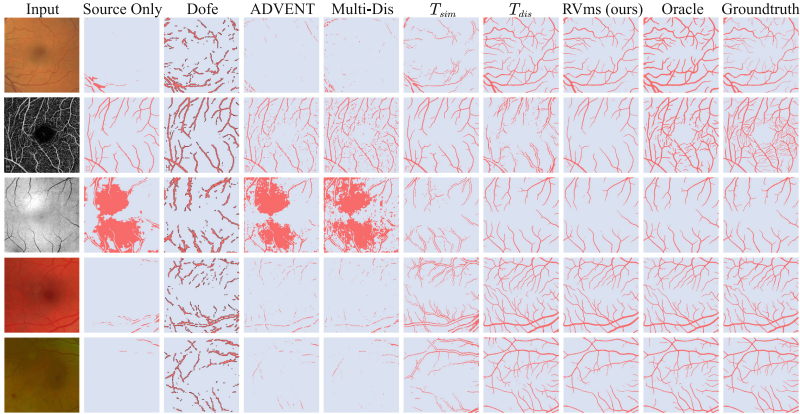
Fig. 3. Representative RV segmentation results using the OCTA images from OCTA-500 as the source domain. The target domains are Fundus Image (DRIVE), OCTA (ROSE), OCT (OCTA-500), Fundus Image (HRF), and UWF Fundus (PRIME-FP20) from top to bottom.

3.3   Results

All methods are evaluated using two metrics, i.e., Dice[%] and 95% Hausdorff Distance (HD[px]), the results of which are tabulated in Table 1. We compare RVms with two recently-developed SOTA DA/MTDA models, namely ADVENT [26] and Multi-Dis [23]. Note that [26] is trained with mixed target domains. We also compare with a domain generalization method Dofe [27], which leaves one out as the target domain and uses multiple source-similar domains as the source domains. The results from the dual teachers T_{sim} and T_{dis} are also reported. **Source Only** means the model is trained with source domain data only. **Oracle** means the model is trained and tested on the specific target domain. Our method achieves Dice scores that are about 38% and 13% higher than the second best method when using OCTA (OCTA-500) and fundus image (DRIVE) as the source domain. It is evident that our proposed RVms delivers superior RV segmentation performance when encountering both cross-modality and cross-center domain shift. Besides, it achieves the highest average Dice score.

Representative visualization results are illustrated in Fig. 3 and Fig. A1 of the appendix. We observe that all compared methods fail in many cases while our framework is very close to the Oracle. Apparently, our framework achieves superior performance when tested on all target domains.

To evaluate the effectiveness of several key components in RVms, we conduct ablation studies. We compare with the proposed RVms without style augmentation (w/o sa), without style transfer (w/o st), without LRIT (w/o LRIT). We also compare with training a single model with images in D_{sim} and D_{dis} without knowledge distillation (w/o KD) and employing only one teacher model (w/o DT). The results are shown in Table 2 and Table A2 of the appendix. The performance degrades when removing any component in the framework, in terms of the average Dice score.

Table 2. Ablation analysis results for several key components in our proposed framework using OCTA (OCTA-500) as the source domain. **Bold** and underlined numbers respectively denote the best and second-best results.

	Source domain: OCTA (OCTA-500)										
Modalities	Fundus image (DRIVE)		OCTA (ROSE)		OCT (OCTA-500)		Fundus image (HRF)		UWF fundus (PRIME-FP20)		Average	
Metrics	Dice ↑	HD ↓	Dice ↑	HD ↓	Dice ↑	HD ↓	Dice ↑	HD ↓	Dice ↑	HD ↓	Dice ↑	HD ↓
Source only	2.18	178.06	52.35	25.52	5.70	62.89	0.43	189.60	1.33	180.42	12.40	127.30
w/o sa	62.74	24.82	48.25	26.51	60.47	30.50	60.48	30.89	60.17	29.91	58.42	26.73
w/o st	70.19	21.58	41.95	27.29	71.59	24.81	70.86	18.19	69.57	23.33	55.94	23.04
w/o LRIT	_71.86_	**13.35**	_56.65_	_19.93_	71.60	_18.84_	71.54	**10.28**	_73.83_	**15.17**	_69.09_	**15.31**
w/o KD	69.73	_14.78_	13.08	20.57	71.07	**18.83**	70.53	16.75	71.96	_18.54_	59.27	_17.89_
w/o DT	70.08	16.06	43.20	26.12	_72.11_	20.91	_72.21_	13.85	72.25	19.64	65.97	19.31
Ours	**72.64**	17.72	**60.80**	**13.49**	**72.81**	24.70	**74.03**	_12.97_	**74.54**	21.09	**70.96**	17.99
Oracle	67.72	9.05	71.43	12.14	82.68	11.13	72.09	12.67	77.80	9.33	74.34	10.86

4   Conclusion

In this paper, we proposed and validated a novel framework for unsupervised multi-target domain adaptation in retinal vessel segmentation. We used style augmentation and style transfer to generate source-similar images and source-dissimilar images to improve the robustness of both teachers and student models. We also conducted knowledge distillation from dual teachers to a generic student, wherein a domain invariant method named LRIT was utilized to facilitate the training process. Another contribution of this work is that we constructed a new dataset called mmRV from several public datasets, which can be used as a new benchmark for DA and domain generalization. Through extensive experiments, our proposed RVms was found to largely outperform representative SOTA MTDA methods, in terms of RV segmentation from different modalities.

References

1. Budai, A., Bock, R., Maier, A., et al.: Robust vessel segmentation in fundus images. Int. J. Biomed. Imaging **2013**(6), 154860 (2013)
2. Cai, J., Zhang, Z., Cui, L., et al.: Towards cross-modal organ translation and segmentation: a cycle-and shape-consistent generative adversarial network. Med. Image Anal. **52**, 174–184 (2019)
3. Cheng, P., Lin, L., Huang, Y., Lyu, J., Tang, X.: I-secret: importance-guided fundus image enhancement via semi-supervised contrastive constraining. In: Medical Image Computing and Computer Assisted Intervention, vol. 12908, pp. 87–96. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-87237-3_9
4. Ding, L., Kuriyan, A.E., Ramchandran, R.S., et al.: Weakly-supervised vessel detection in ultra-widefield fundus photography via iterative multi-modal registration and learning. IEEE Trans. Med. Imaging **40**(10), 2748–2758 (2020)
5. Dou, Q., et al.: PnP-AdaNet: plug-and-play adversarial domain adaptation network at unpaired cross-modality cardiac segmentation. IEEE Access **7**, 99065–99076 (2019)
6. Ganin, Y., Ustinova, E., Ajakan, H., et al.: Domain-adversarial training of neural networks. J. Mach. Learn. Res. **17**(1), 2030–2096 (2016)
7. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 770–778 (2016)
8. Hinton, G., Vinyals, O., Dean, J., et al.: Distilling the knowledge in a neural network. arXiv preprint arXiv:1503.02531 (2015)
9. Hu, Z., Liu, Q., Paulus, Y.M.: New frontiers in retinal imaging. Int. J. Ophthalmic Res. **2**(3), 148–158 (2016)
10. Huang, D., Swanson, E.A., Lin, C.P., et al.: Optical coherence tomography. Science **254**(5035), 1178 (1991)

11. Huang, Y., Lin, L., Li, M., et al.: Automated hemorrhage detection from coarsely annotated fundus images in diabetic retinopathy. In: 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI), pp. 1369–1372. IEEE (2020)

12. Javanmardi, M., Tasdizen, T.: Domain adaptation for biomedical image segmentation using adversarial training. In: 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), pp. 554–558. IEEE (2018)

13. Li, M., Zhang, Y., Ji, Z., et al.: IPN-V2 and OCTA-500: methodology and dataset for retinal image segmentation. arXiv preprint arXiv:2012.07261 (2020)

14. Lin, L., Wang, Z., Wu, J., et al.: BSDA-Net: a boundary shape and distance aware joint learning framework for segmenting and classifying octa images. In: Medical Image Computing and Computer Assisted Intervention, vol. 12908, pp. 65–75. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-87237-3_7

15. Lin, L., Wu, J., Cheng, P., Wang, K., Tang, X.: BLU-GAN: bi-directional ConvLSTM U-net with generative adversarial training for retinal vessel segmentation. In: Intelligent Computing and Block Chain, vol. 1385, pp. 3–13. Springer, Cham (2020). https://doi.org/10.1007/978-981-16-1160-5_1

16. Liu, Q., Chen, C., Qin, J., et al.: FedDG: federated domain generalization on medical image segmentation via episodic learning in continuous frequency space. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 1013–1023 (2021)

17. Ma, Y., Hao, H., Xie, J., et al.: Rose: a retinal oct-angiography vessel segmentation dataset and new model. IEEE Trans. Med. Imaging **40**(3), 928–939 (2020)

18. Ojala, T., Pietikainen, M., Maenpaa, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans. Pattern Anal. Mach. Intell. **24**(7), 971–987 (2002)

19. Peng, L., Lin, L., Cheng, P., Huang, Z., Tang, X.: Unsupervised domain adaptation for cross-modality retinal vessel segmentation via disentangling representation style transfer and collaborative consistency learning. arXiv preprint arXiv:2201.04812 (2022)

20. Peng, L., Lin, L., Cheng, P., et al.: Fargo: a joint framework for FAZ and RV segmentation from octa images. In: Fu, H., Garvin, M.K., MacGillivray, T., Xu, Y., Zheng, Y. (eds.) Ophthalmic Medical Image Analysis, vol. 12970, pp. 42–51. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-87000-3_5

21. Ronneberger, O., Fischer, P., Brox, T.: U-net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W., Frangi, A. (eds.) Medical Image Computing and Computer-Assisted Intervention, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28

22. Saenko, K., Kulis, B., Fritz, M., Darrell, T.: Adapting visual category models to new domains. In: Daniilidis, K., Maragos, P., Paragios, N. (eds.) Computer Vision, vol. 6314, pp. 213–226. Springer, Cham (2010). https://doi.org/10.1007/978-3-642-15561-1_16

23. Saporta, A., Vu, T.H., Cord, M., Pérez, P.: Multi-target adversarial frameworks for domain adaptation in semantic segmentation. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 9072–9081 (2021)

24. Shi, T., Boutry, N., Xu, Y., Géraud, T.: Local intensity order transformation for robust curvilinear object segmentation. IEEE Trans. Image Processing **31**, 2557–2569 (2022)

25. Staal, J., Abràmoff, M.D., Niemeijer, M., et al.: Ridge-based vessel segmentation in color images of the retina. IEEE Trans. Med. Imaging **23**(4), 501–509 (2004)

26. Vu, T.H., Jain, H., et al.: Advent: adversarial entropy minimization for domain adaptation in semantic segmentation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 2517–2526 (2019)

27. Wang, S., Yu, L., Li, K., et al.: DoFE: domain-oriented feature embedding for generalizable fundus image segmentation on unseen datasets. IEEE Trans. Med. Imaging **39**(12), 4237–4248 (2020)

28. Wang, S., Yu, L., Yang, X., et al.: Patch-based output space adversarial learning for joint optic disc and cup segmentation. IEEE Trans. Med. Imaging **38**(11), 2485–2495 (2019)

29. Yang, Y., Soatto, S.: FDA: fourier domain adaptation for semantic segmentation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 4085–4095 (2020)

30. Zhang, T., Cheng, J., et al.: Noise adaptation generative adversarial network for medical image analysis. IEEE Trans. Med. Imaging **39**(4), 1149–1159 (2019)

31. Zhou, Z., Sodha, V., Rahman Siddiquee, M.M., et al.: Models genesis: generic autodidactic models for 3D medical image analysis. In: Medical Image Computing and Computer Assisted Intervention, vol. 11767, pp. 384–393. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-32251-9_42

32. Zhu, J.Y., Park, T., Isola, P., Efros, A.A.: Unpaired image-to-image translation using cycle-consistent adversarial networks. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 2223–2232 (2017)