Domain Specific Convolution and High Frequency Reconstruction based Unsupervised Domain Adaptation for Medical Image Segmentation

Shishuai Hu[1,*], Zehui Liao[1,*], and Yong Xia[1,2,3(✉)]

[1] National Engineering Laboratory for Integrated Aero-Space-Ground-Ocean Big Data Application Technology, School of Computer Science and Engineering, Northwestern Polytechnical University, Xi'an 710072, China
`yxia@nwpu.edu.cn`
[2] Ningbo Institute of Northwestern Polytechnical University, Ningbo 315048, China
[3] Research & Development Institute of Northwestern Polytechnical University in Shenzhen, Shenzhen 518057, China

Abstract. Although deep learning models have achieved remarkable success in medical image segmentation, the domain shift issue caused mainly by the highly variable quality of medical images is a major hurdle that prevents these models from being deployed for real clinical practices, since no one can predict the performance of a 'well-trained' model on a set of unseen clinical data. Previously, many methods have been proposed based on, for instance, CycleGAN or the Fourier transform to address this issue, which, however, suffer from either an inadequate ability to preserve anatomical structures or unexpectedly introduced artifacts. In this paper, we propose a multi-source-domain unsupervised domain adaptation (UDA) method called **Do**main specific **C**onvolution and high frequency **R**econstruction (**DoCR**) for medical image segmentation. We design an auxiliary high frequency reconstruction (HFR) task to facilitate UDA, and hence avoid the interference of the artifacts generated by the low-frequency component replacement. We also construct the domain specific convolution (DSC) module to boost the segmentation model's ability to domain-invariant features extraction. We evaluate DoCR on a benchmark fundus image dataset. Our results indicate that the proposed DoCR achieves superior performance over other UDA methods in multi-domain joint optic cup and optic disc segmentation. Code is available at: https://github.com/ShishuaiHu/DoCR.

Keywords: Unsupervised domain adaptation · Medical image segmentation · Domain specific convolution · High frequency reconstruction.

1 Introduction

Recent years have witnessed the tremendous success of deep learning in medical image segmentation [9,4,23,11,20]. Like any supervised learning approach,

---

* Equal contribution.

deep learning models reply their success on the assumption that training images and test ones share similar, if not identical, distributions. This assumption, however, is less likely to be held on medical image segmentation tasks, since the range of image quality and structure visibility can be considerable, depending on characteristics of the imaging equipment, skill of the operator, and compromises with factors such as patient radiation exposure and imaging time [15]. To relieve the performance degradation on test samples caused by the distribution discrepancy, tremendous research endeavors have been devoted to unsupervised domain adaptation (UDA), the setting where labeled training data is available on a source domain or source domains, but the goal is to have good performance on a target domain with only unlabeled data [16,24,21,14,2,17,18,5].

Currently, data-level UDA methods [16,24,7,21,13] attempt to address the distribution discrepancy issue by transforming source-domain images to the target domain using image synthesis. CycleGAN-based UDA methods [16,24,7] impose cycle consistency constraints between the source-domain image and corresponding synthetic target-domain image, and hence can be trained on the unpaired source- and target-domain images. Although they can transform image style well, these methods suffer from the poor preservation of structural information, which is unacceptable for segmentation applications. To preserve the anatomical structures in images, Fourier transform-based UDA methods [21,13] translate source-domain images to the target domain by replacing the low frequency components of the images in both domains, based on the assumption that the style information is embedded in low frequency components and structural information is embedded in high frequency components. Despite their prevalence, these approaches may unexpectedly introduce artifacts to synthetic images, since low frequency component replacement works like an ideal filter that introduces ringing artifacts in the filtered image. Moreover, it is less likely to estimate the cut-off between low and high frequency components. To address this issue, we suggest using high frequency image reconstruction (HFR), which is insusceptible to artifacts, to substitute for the low frequency component replacement. Similar to a high-pass filter, HFR filters out the low frequency components where most style information locates and hence improves the performance of UDA.

With the recent advance of adversarial training and feature normalization, feature-level UDA methods [2,12,19,14] boost their ability to domain-invariant feature extraction. Among them, feature disentangling based models [14] attempt to decouple domain-variant and domain-invariant features under the guidance of the game theory. Most of these models contain a pair of alternately trained feature extractor and domain discriminator. The feature extractor is trained to fool the discriminator, and the discriminator is expected to identify the domain in which the features are extracted. After training, these models can well disentangle the discriminative features for domain classification. However, the equilibrium between the feature extractor and domain discriminator is hard to maintain, and hence it is difficult to train these models well enough to improve UDA. To avoid this issue, domain-specific batch normalization (DSBN) was developed to improve the ability to domain-invariant feature extraction [2,12]. Al-
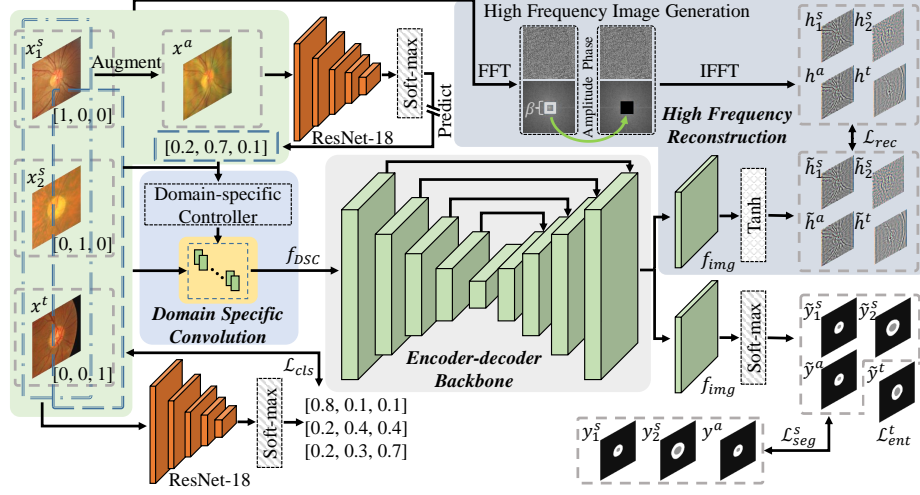
Fig. 1. Overview of DoCR for medical image segmentation. FFT: Fast Fourier Transform; IFFT: Inverse FFT. The box in yellow represents the domain-specific convolutional head.

though DSBN can reduce the distribution discrepancy between source and target domain images in the feature space, its feature normalization ability is limited by the few learnable parameters and batch-level operations [19]. Therefore, we designed a domain-specific convolution (DSC) module based on dynamic convolutions [22,8,6] to extract domain-insensitive features.

In this paper, we incorporate both HFR and DSC into an encoder-decoder backbone, and thus propose a multi-source-domain UDA method called **Do**main specific **C**onvolution and high frequency **R**econstruction (**DoCR**) for medical image segmentation, with a particular focus on addressing the issue of training-test data distribution discrepancy caused by image quality variation. This method has been evaluated against other UDA methods on the task of the joint optic cup (OC) and optic disc (OD) segmentation on fundus images. Our main contributions are three-fold: (1) we use HFR as an auxiliary task to facilitate UDA in medical image segmentation, and hence avoid the interference of the artifacts generated by the low-frequency component replacement; (2) we design the DSC module to improve the segmentation model's ability to domain-invariant features extraction; and (3) as indicated by our results on OC/OD segmentation, the proposed DoCR model achieves superior performance over other state-of-the-art UDA methods.

2   Method

2.1   Problem Definition and Method Overview

Let a set of K source domains be represented as $\mathcal{D}^s = \{(x_{ki}^s, y_{ki}^s)_{i=1}^{N_k}\}_{k=1}^K$, where x_{ki}^s is the i-th image in the k-th source domain, and y_{ki}^s is the segmentation mask of x_{ki}^s. The target domain with unlabeled images can be denoted as $\mathcal{D}_u^t = (x_i^t)_{i=1}^{N_u}$. Our goal is to train a segmentation model $F_\theta : x \to y$ on $\mathcal{D}^s$ and $\mathcal{D}_u^t$, which can perform well on the target domain $\mathcal{D}^t$.

The proposed DoCR model has a U-shape architecture [4], equipped with a DSC module, an HFR head, and a segmentation head (see Fig. 1). The DSC module is used as the first convolutional layer to remove as much domain information as possible from the input images x and extract domain-insensitive features f_{DSC}. Then the encoder-decoder backbone takes f_{DSC} as its input and produces further processed features f_{img}. Based on f_{img}, the HFR head reconstructs the high frequency components h of x, and the segmentation head generates the segmentation prediction $\widetilde{y}$. We now delve into the details of each part.

2.2   DSC Module

The DSC module consists of a DSC head and a domain-specific controller. The DSC head has a light-weighted design, containing a 3×3 convolutional layer, a ReLU layer, and a batch normalization layer. It takes the C-channel input image x as input and outputs a 32-channel domain-insensitive feature map f_{DSC}, which will be fed to the encoder-decoder backbone. Consequently, there are totally $C \times 32 \times 3 \times 3 + 32$ parameters in this head, denoted by ω_D. These parameters are dynamically generated by the controller F_C, which is also a convolutional layer, conditioned on the domain code $\mathscr{D}$ of the input images x.

Since there are K source domains and a target domain, $\mathscr{D}$ is a one-hot $K+1$ dimension vector. To relieve the insufficient optimization issue caused by one-hot $\mathscr{D}$, we conduct low frequency component replacement [21] to increase the diversity of source-domain data. Specifically, the low frequency components in a source-domain image can be replaced by the same number of low frequency components in either a source-domain image or a target-domain image. Then, we train ResNet-18 as a domain predictor using original images and their one-hot domain codes, and then use the trained ResNet-18 to predict the domain code $\mathscr{D}^a$ of each augmented image x^a (see the bottom left part of Fig. 1).

2.3   Encoder-decoder Backbone

We tailor ResNet-34 as the encoder by replacing its average pooling layer and fully connected (FC) layer with a ReLU layer and setting the channel number of the first convolutional layer to 32. The encoder contains a convolutional block and four residual blocks, which gradually convert the input feature map f_{DSC} into a lower resolution and more abstract one. Symmetrically, the decoder is

also composed of five convolutional blocks, which gradually upsample the feature map to its original resolution. Skip connections are made from the first convolutional block and first three residual blocks in the encoder to the corresponding locations in the decoder. In each of the first four decoder blocks, the feature map is processed by a 2×2 transposed convolutional layer with a stride of 2 and a 1×1 convolutional layer with a stride of 1. After that, the feature map is concatenated with the feature map skipped from the encoder, and then fed to a ReLU layer and a batch normalization layer. The last decoder block only upsamples the feature map using a transposed convolutional layer and a batch normalization layer and produces a 32-channel feature map f_{img} with the same size $H \times W$ as the input image.

2.4 HFR and Segmentation

The HFR head F_{RC} contains a convolutional layer and a tanh layer to reconstruct the high frequency image $\widetilde{h}$ based on the input feature map f_{img}. This can be expressed as

$$\widetilde{h} = Tanh(F_{RC}(f_{img}; \theta_{RC})) \tag{1}$$

where θ_{RC} is the parameters of F_{RC}, and $Tanh(\cdot)$ represents the tanh layer. The ground truth of high frequency image h is generated by setting low frequency components to zero in the frequency domain (see the top right part of Fig. 1). The size of the replaced area is $\beta \times (H \times W)$, where β is a hyper-parameter.

The segmentation head F_S is composed of a convolutional layer and a soft-max layer to convert f_{img} to segmentation map $\widetilde{y}$. This can be expressed as

$$\widetilde{y} = SM(F_S(f_{img}; \theta_S)) \tag{2}$$

where θ_S is the parameters of F_S, and $SM(\cdot)$ is a soft-max layer.

2.5 Training and Inference

Training. We pretrain the domain predictor for 20 epochs using the one-hot domain code and corresponding images without augmentation. The objective of the domain prediction task is the cross-entropy loss

$$\mathcal{L}_{cls} = -\sum_{k=1}^{K+1} d_k \log(d_k^p) \tag{3}$$

where d_k is the domain label, and d_k^p is the soft-max probability of belonging to the k-th domain.

After that, the augmented image, the source image, and the target image batches are fed to the model randomly. For the reconstruction task, the objective is the $L1$ loss

$$\mathcal{L}_{rec} = \sum \left| h - \widetilde{h} \right| \tag{4}$$

For the segmentation task, the cross-entropy loss is adopted as the objective for the supervised source domain data, and entropy loss [21] is used for the unlabeled target domain images

$$\mathcal{L}_{seg} = -(y^s \log \widetilde{y}^s + (1 - y^s) \log(1 - \widetilde{y}^s)) + \lambda_{ent}((-\widetilde{y}^t \log_2 \widetilde{y}^t)^2 + 0.001^2)^2 \quad (5)$$

where y^s and $\widetilde{y}^s$ represent the ground truth and prediction of source domain image, $\widetilde{y}^t$ is the prediction of target domain image, and λ_{ent} is a hyper-parameter, which is set to 0.005 as suggested in [21].

The total loss can be expressed as

$$\mathcal{L} = I(x^a \notin x)\mathcal{L}_{cls} + \mathcal{L}_{seg} + \lambda_{rec}\mathcal{L}_{rec} \quad (6)$$

where $I(\cdot)$ is an indicator function, and λ_{rec} is a hyper-parameter.
Inference. During inference, given a test image and its domain code, the DSC block produces a 32-channel feature map f_{DSC}, which is then fed to the backbone network and converted into a 32-channel feature map f_{img}. Next, the segmentation head takes f_{img} as input and generates the segmentation prediction $\widetilde{y}$.

3 Experiments and Results

Materials and Evaluation Metrics. To reduce the annotator bias [10] among different datasets, we use a multi-domain joint OC/OD segmentation dataset annotated by the same group of ophthalmologists, namely RIGA+ [1,3], and only use rater 1's annotations to train and evaluate all the algorithms. It contains 195 labeled data from BinRushed, 95 labeled data from Magrabia, 454 labeled, and 717 unlabeled data from the MESSIDOR database (including data collected from 3 medical centers, *i.e.*, BASE1, BASE2, and BASE3). The details of the RIGA+ were summarized in Table 1. Dice similarity coefficient ($D_{disc}(\%)$, $D_{cup}(\%)$) is adopted as evaluation metrics.

Table 1. Details of the dataset used for this study. The 20% test cases in the target domain were selected randomly.

Domain	Dataset	Labeled Images (Training/Test)	Unlabeled Images
Source	BinRushed	195 (195/0)	0
	Magrabia	95 (95/0)	0
Target	BASE1	173 (138/35)	227
Target	BASE2	148 (118/30)	238
Target	BASE3	133 (106/27)	252

Implementation Details. Each input image was normalized by subtracting its mean and dividing by its standard deviation. For all experiments, the mini-batch size was set to 8, and all the images were center-cropped [17] and resized to 512×512. The SGD algorithm with a momentum of 0.99 and an initial learning
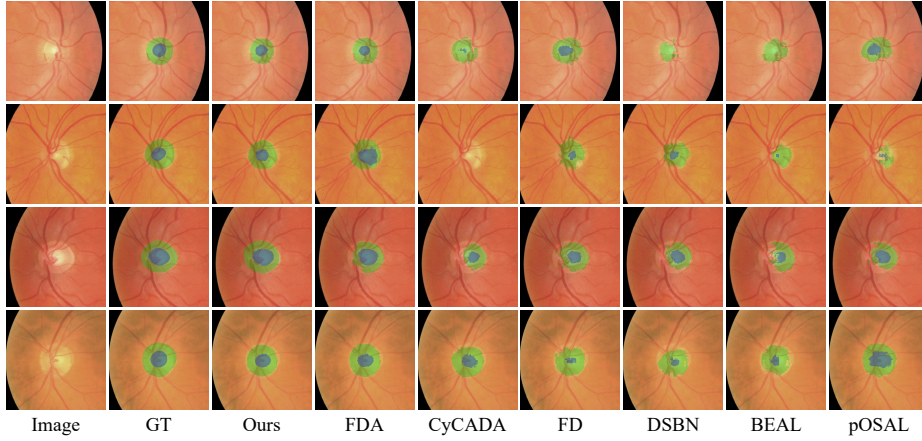
Fig. 2. Visualization of segmentation masks predicted by our DoCR and six competing methods, together with ground truth.

rate lr_0 of 0.01 was adopted as the optimizer. The learning rate was decayed according to the polynomial rule $lr = lr_0 \times (1 - t/T)^{0.9}$, where t is the current epoch and T is the maximum epoch, which was set to 100. The hyper-parameters λ_{rec} and β were set to 0.1 and 0.01 according to the ablation experiments. All experiments were implemented using the PyTorch framework. It takes about 5 hours to train our DoCR model on one NVIDIA 2080Ti GPU.

Table 2. Performance of our DoCR and nine competing methods in OC/OD segmentation. The best results except for 'Intra-Domain' are highlighted with **bold**.

Methods		Target Domain		
		BASE1	BASE2	BASE3
	Intra-Domain	(95.87, <u>87.27</u>)	(95.92, <u>88.83</u>)	(95.73, <u>88.42</u>)
Baseline	w/o DA	(94.30, 84.01)	(94.58, 79.22)	(93.82, 81.40)
	Self-Training	(95.66, 85.47)	(94.97, 78.69)	(94.98, 82.76)
Data Level	FDA [21]	(**95.91**, 85.69)	(96.22, 83.19)	(95.93, 84.36)
	CyCADA [7]	(95.04, 83.50)	(93.86, 80.64)	(94.77, 82.55)
Feature Level	FD [14]	(95.58, 84.29)	(95.26, 82.72)	(95.40, 85.26)
	DSBN [2]	(95.60, 85.30)	(94.19, 82.43)	(95.08, 85.14)
Decision Level	BEAL [17]	(94.96, 83.60)	(95.29, 80.12)	(94.78, 82.86)
	pOSAL [18]	(93.96, 83.07)	(95.36, 84.36)	(94.73, 84.36)
	DoCR (Ours)	(95.89, **85.72**)	(**96.39**, **86.17**)	(**96.19**, **85.96**)

Comparative Experiments. We compare our DoCR with three baselines: 'Intra-Domain' setting (*i.e.*, training and test on the data from the same target domain), 'w/o DA' (*i.e.*, training on the data from the source domains and test on

the target domain data), and 'Self-Training' (*i.e.*, using source domain data and pseudo labeled target domain data for iterative training and test on the target domain data); two data level UDA methods, including (1) FDA: a Fourier transform based method [21], and (2) CyCADA: a CycleGAN based method [7]; two feature level UDA methods, including (1) FD: a feature disentanglement based approach [14], and (2) DSBN: a feature normalization based method [2]; and two decision level UDA methods, including (1) pOSAL: an output adversarial training based method [18], and (2) BEAL: a boundary and entropy adversarial training based method [17]. Each target domain is chosen as the test domain alternately to simulate the deployment scenario, *i.e.*, only images from one target domain can be accessed per time.

The results of DoCR and its competitors were given in Table 2. It can be seen that the overall performance of our DoCR is not only higher than the baseline methods but also better than that of the other UDA approaches. We also visualize the segmentation maps predicted by our DoCR and other six competing methods, as shown in Fig. 2. It reveals that our DoCR can produce the most accurate segmentation map compared to the ground truth.

Table 3. Performance of DoCR, its four variants, and two baseline methods.

Methods		Target Domain		
		BASE1	BASE2	BASE3
Baseline	w/o DA	(94.30, 84.01)	(94.58, 79.22)	(93.82, 81.40)
	Self-Training	(95.66, 85.47)	(94.97, 78.69)	(94.98, 82.76)
Analysis of DSC	HFR	(95.81, 85.03)	(95.85, 84.88)	(95.43, 85.67)
	HFR + Multi-Input-Head	(57.92, 36.74)	(52.96, 23.67)	(14.82, 00.00)
	HFR + DSC (DoCR)	(**95.89, 85.72**)	(**96.39, 86.17**)	(**96.19, 85.96**)
Analysis of HFR	DSC (Self-Training)	(95.94, 84.65)	(95.46, 82.82)	(95.92, 84.39)
	DSC + Image Rec.	(**96.24**, 85.45)	(95.90, 82.51)	(95.59, 84.54)
	DSC + HFR (DoCR)	(95.89, **85.72**)	(**96.39, 86.17**)	(**96.19, 85.96**)

Ablation Analysis. To evaluate the effectiveness of HFR and DSC, we conducted a series of ablation experiments, as shown in Table 3. The results of 'w/o DA', 'Self-Training', 'HFR', 'DSC', and 'DoCR' reveal that both HFR and DSC contribute to the final results. We compare the performances of reconstructing the high frequency image and reconstructing the image itself. The results are presented in 'Analysis of HFR' part of Table 3. It can be seen that reconstructing the image itself can not bring much performance gain and the OD segmentation results on 'BASE2' even get worse. It can be attributed to the fact that reconstructing the image itself forces the network to recover the details of the image, most of which are redundant for the segmentation task. Different from that, high frequency image reconstruction facilitates the network filter out domain-sensitive low frequency features and hence improves the segmentation performance on the target domain data.

We also compare the results of adopting dynamic convolution (*i.e.*, HFR + DSC) and using multi-input-head (*i.e.*, HFR + Multi-Input-Head). It can be seen that using multi-input-head is hardly to perform segmentation since the target domain-specific parameters are mainly optimized under the supervision of reconstruction. Even though, adopting dynamic convolution can achieve improved performance than HFR, owing to its flexible generated parameters and diverse domain codes.

4 Conclusion

In this paper, we propose a DoCR based unsupervised domain adaptation model for medical image segmentation. It is composed of a DSC module, an encoder-decoder backbone, an HFR head, and a segmentation head. The DSC module improves the model's domain-invariant feature extraction ability by adopting domain-specific parameters generated conditioned on the domain code. The HFR head performs reconstruction under the supervision of high frequency image, and hence forces the model to behave like a high-pass filter to filter out low frequency domain-sensitive features. Experimental results on the multi-domain joint OC/OD segmentation task suggest the proposed DoCR can achieve superior performance over the baselines and other state-of-the-art UDA methods.

Acknowledgment

References

1. Almazroa, A., Alodhayb, S., Osman, E., Ramadan, E., Hummadi, M., Dlaim, M., Alkatee, M., Raahemifar, K., Lakshminarayanan, V.: Retinal fundus images for glaucoma analysis: the riga dataset. In: Medical Imaging 2018: Imaging Informatics for Healthcare, Research, and Applications. vol. 10579, p. 105790B. International Society for Optics and Photonics (2018)
2. Chang, W.G., You, T., Seo, S., Kwak, S., Han, B.: Domain-specific batch normalization for unsupervised domain adaptation. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 7354–7362 (2019)
3. Decencière, E., Zhang, X., Cazuguel, G., Lay, B., Cochener, B., Trone, C., Gain, P., Ordonez, R., Massin, P., Erginay, A., et al.: Feedback on a publicly distributed image database: the messidor database. Image Analysis & Stereology **33**(3), 231–234 (2014)
4. Falk, T., Mai, D., Bensch, R., Çiçek, Ö., Abdulkadir, A., Marrakchi, Y., Böhm, A., Deubner, J., Jäckel, Z., Seiwald, K., et al.: U-Net: Deep learning for cell counting, detection, and morphometry. Nature Methods **16**(1), 67–70 (Jan 2019). https://doi.org/10.1038/s41592-018-0261-2

5. Guan, H., Liu, M.: Domain adaptation for medical image analysis: a survey. IEEE Transactions on Biomedical Engineering **69**(3), 1173–1185. https://doi.org/10.1109/TBME.2021.3117407

6. Han, Y., Huang, G., Song, S., Yang, L., Wang, H., Wang, Y.: Dynamic Neural Networks: A Survey. arXiv:2102.04906 [cs] (Feb 2021)

7. Hoffman, J., Tzeng, E., Park, T., Zhu, J.Y., Isola, P., Saenko, K., Efros, A., Darrell, T.: Cycada: Cycle-consistent adversarial domain adaptation. In: International Conference on Machine Learning (ICML). pp. 1989–1998. PMLR (2018)

8. Hu, S., Liao, Z., Zhang, J., Xia, Y.: Domain and content adaptive convolution for domain generalization in medical image segmentation. arXiv preprint arXiv:2109.05676 (2021)

9. Isensee, F., Jaeger, P.F., Kohl, S.A.A., Petersen, J., Maier-Hein, K.H.: nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation. Nature Methods **18**(2), 203–211 (Dec 2020). https://doi.org/10.1038/s41592-020-01008-z

10. Liao, Z., Hu, S., Xie, Y., Xia, Y.: Modeling human preference and stochastic error for medical image segmentation with multiple annotators. arXiv preprint arXiv:2111.13410 (2021)

11. Litjens, G., Kooi, T., Bejnordi, B.E., Setio, A.A.A., Ciompi, F., Ghafoorian, M., van der Laak, J.A., van Ginneken, B., Sánchez, C.I.: A survey on deep learning in medical image analysis. Medical Image Analysis **42**, 60–88 (2017). https://doi.org/https://doi.org/10.1016/j.media.2017.07.005

12. Liu, Q., Dou, Q., Yu, L., Heng, P.A.: MS-Net: Multi-Site Network for Improving Prostate Segmentation With Heterogeneous MRI Data. IEEE Transactions on Medical Imaging **39**(9), 2713–2724 (Sep 2020). https://doi.org/10.1109/TMI.2020.2974574

13. Liu, Q., Chen, C., Qin, J., Dou, Q., Heng, P.A.: FedDG: Federated Domain Generalization on Medical Image Segmentation via Episodic Learning in Continuous Frequency Space. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 1013–1023 (Jun 2021)

14. Shin, S.Y., Lee, S., Summers, R.M.: Unsupervised domain adaptation for small bowel segmentation using disentangled representation. In: Medical Image Computing and Computer Assisted Intervention (MICCAI). pp. 282–292. Springer (2021)

15. Sprawls, P.: Image characteristics and quality. In: Physical principles of medical imaging, pp. 1–16. Aspen Gaithersburg (1993)

16. Wang, R., Zheng, G.: Cycmis: Cycle-consistent cross-domain medical image segmentation via diverse image augmentation. Medical Image Analysis **76**, 102328 (2022)

17. Wang, S., Yu, L., Li, K., Yang, X., Fu, C.W., Heng, P.A.: Boundary and entropy-driven adversarial learning for fundus image segmentation. In: Medical Image Computing and Computer Assisted Intervention (MICCAI). pp. 102–110 (2019)

18. Wang, S., Yu, L., Yang, X., Fu, C.W., Heng, P.A.: Patch-based output space adversarial learning for joint optic disc and cup segmentation. IEEE Transactions on Medical Imaging **38**(11), 2485–2495 (2019)

19. Xiao, J., Yu, L., Xing, L., Yuille, A., Zhou, Y.: Dualnorm-unet: Incorporating global and local statistics for robust medical image segmentation. arXiv preprint arXiv:2103.15858 (2021)

20. Xie, X., Niu, J., Liu, X., Chen, Z., Tang, S., Yu, S.: A survey on incorporating domain knowledge into deep learning for medical image analysis. Medical Image Analysis **69**, 101985 (Apr 2021). https://doi.org/10.1016/j.media.2021.101985

21. Yang, Y., Soatto, S.: FDA: Fourier Domain Adaptation for Semantic Segmentation. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 4085–4095 (Jun 2020)
22. Zhang, J., Xie, Y., Xia, Y., Shen, C.: DoDNet: Learning To Segment Multi-Organ and Tumors From Multiple Partially Labeled Datasets. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 1195–1204 (Jun 2021)
23. Zhou, Z., Siddiquee, M.M.R., Tajbakhsh, N., Liang, J.: UNet++: Redesigning Skip Connections to Exploit Multiscale Features in Image Segmentation. IEEE Transactions on Medical Imaging **39**(6), 1856–1867 (Jun 2020). https://doi.org/10.1109/TMI.2019.2959609
24. Zhu, J.Y., Park, T., Isola, P., Efros, A.A.: Unpaired image-to-image translation using cycle-consistent adversarial networks. In: International Conference on Computer Vision (ICCV). pp. 2223–2232 (2017)