Devil is in Channels: Contrastive Single Domain Generalization for Medical Image Segmentation

Shishuai Hu, Zehui Liao, and Yong Xia(✉)

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

Abstract. Deep learning-based medical image segmentation models suffer from performance degradation when deployed to a new healthcare center. To address this issue, unsupervised domain adaptation and multi-source domain generalization methods have been proposed, which, however, are less favorable for clinical practice due to the cost of acquiring target-domain data and the privacy concerns associated with redistributing the data from multiple source domains. In this paper, we propose a **C**hannel-level **C**ontrastive **S**ingle **D**omain **G**eneralization (**C^{2}SDG**) model for medical image segmentation. In C^{2}SDG, the shallower features of each image and its style-augmented counterpart are extracted and used for contrastive training, resulting in the disentangled style representations and structure representations. The segmentation is performed based solely on the structure representations. Our method is novel in the contrastive perspective that enables channel-wise feature disentanglement using a single source domain. We evaluated C^{2}SDG against six SDG methods on a multi-domain joint optic cup and optic disc segmentation benchmark. Our results suggest the effectiveness of each module in C^{2}SDG and also indicate that C^{2}SDG outperforms the baseline and all competing methods with a large margin. The code is available at https://github.com/ShishuaiHu/CCSDG.

Keywords: Single domain generalization · Medical image segmentation · Contrastive learning · Feature disentanglement

1 Introduction

It has been widely recognized that the success of supervised learning approaches, such as deep learning, relies on the i.i.d. assumption for both training and test samples [11]. This assumption, however, is less likely to be held on medical image segmentation tasks due to the imaging distribution discrepancy caused by

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/978-3-031-43901-8_2.

H. Greenspan et al. (Eds.): MICCAI 2023, LNCS 14223, pp. 14–23, 2023.
https://doi.org/10.1007/978-3-031-43901-8_2

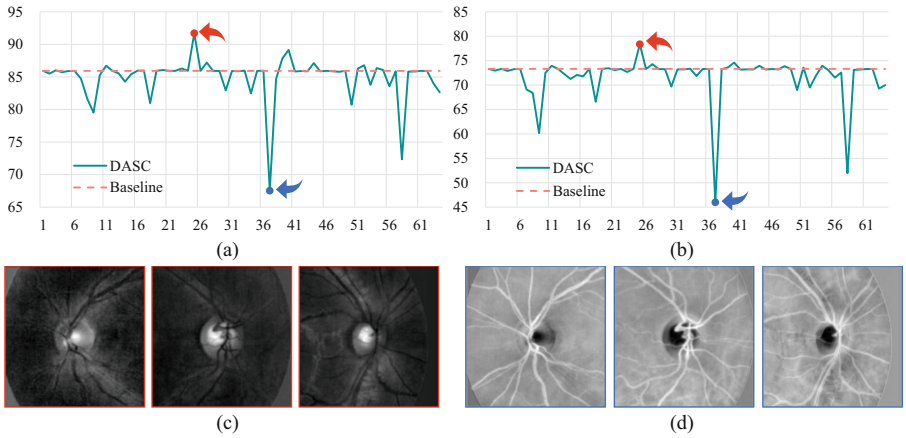

Fig. 1. Average OD (a) and OC (b) segmentation performance (Dice%) obtained on unseen target domain (BASE2) versus removed channel of shallow features. The Dice scores obtained before and after dropping a channel are denoted by 'Baseline' and 'DASC', respectively. The 24th channel (c) and 36th channel (d) obtained on three target-domain images are visualized.

non-uniform characteristics of the imaging equipment, inconsistent skills of the operators, and even compromise with factors such as patient radiation exposure and imaging time [14]. Therefore, the imaging distribution discrepancy across different healthcare centers renders a major hurdle that prevents deep learning-based medical image segmentation models from clinical deployment [7,18].

To address this issue, unsupervised domain adaptation (UDA) [8,17] and multi-source domain generalization (MSDG) [10,16] have been studied. UDA needs access to the data from source domain(s) and unlabeled target domain, while MSDG needs access to the data from multiple source domains. In clinical practice, both settings are difficult to achieve, considering the cost of acquiring target-domain data and the privacy concerns associated with redistributing the data from multiple source domains [9,22].

By contrast, single domain generalization (SDG) [2,13,15,19,22,23] is a more practical setting, under which only the labeled data from one source domain are used to train the segmentation model, which is thereafter applied to the unseen target-domain data. The difficulty of SDG is that, due to the existence of imaging distribution discrepancy, the trained segmentation model is prone to overfit the source-domain data but generalizes poorly on target-domain data. An intuitive solution is to increase the diversity of training data by performing data augmentation at the image-level [13,15,19,21]. This solution has recently been demonstrated to be less effective than a more comprehensive one, *i.e.*, conducting domain adaptation on both image- and feature-levels [2,8,12]. As a more comprehensive solution, Dual-Norm [23] first augments source-domain images into 'source-similar' images with similar intensities and 'source-dissimilar' images

with inverse intensities, and then processes these two sets of images using different batch normalization layers in the segmentation model. Although achieving promising performance in cross-modality CT and MR image segmentation, Dual-Norm may not perform well under the cross-center SDG setting, where the source- and target-domain data are acquired at different healthcare centers, instead of using different imaging modalities. In this case, the 'source-dissimilar' images with inverse intensities do not really exist, and it remains challenging to determine the way to generate both 'source-similar' and 'source-dissimilar' images [1,4]. To address this challenge, we suggest resolving 'similar' and 'dissimilar' from the perspective of contrastive learning. Given a source image and its style-augmented counterpart, only the structure representations between them are 'similar', whereas their style representations should be 'dissimilar'. Based on contrastive learning, we can disentangle and then discard the style representations, which are structure-irrelevant, using images from only a single domain.

Specifically, to disentangle the style representations, we train a segmentation model, *i.e.*, the baseline, using single domain data and assess the impact of the features extracted by the first convolutional layer on the segmentation performance, since shallower features are believed to hold more style-sensitive information [8,18]. A typical example was given in Fig. 1(a) and (b), where the green line is the average Dice score obtained on the target domain (the BASE2 dataset) versus the index of the feature channel that has been dropped. It reveals that, in most cases, removing a feature does not affect the model performance, indicating that the removed feature is redundant. For instance, the performance even increases slightly after removing the 24th channel. This observation is consistent with the conclusion that there exists a sub-network that can achieve comparable performance [6]. On the contrary, it also shows that some features, such as the 36th channel, are extremely critical. Removing this feature results in a significant performance drop. We visualize the 24th and 36th channels obtained on three target-domain images in Fig. 1(c) and (d), respectively. It shows that the 36th channel is relatively 'clean' and most structures are visible on it, whereas the 24th channel contains a lot of 'shadows'. The poor quality of the 24th channel can be attributed to the fact that the styles of source- and target-domain images are different and the style representation ability learned on source-domain images cannot generalize well on target-domain images. Therefore, we suggest that the 24th channel is more style-sensitive, whereas the 36th channel contains more structure information. This phenomenon demonstrates that 'the devil is in channels'. Fortunately, contrastive learning provides us a promising way to identify and expel those style-sensitive 'devil' channels from the extracted image features.

In this paper, we incorporate contrastive feature disentanglement into a segmentation backbone and thus propose a novel SDG method called **C**hannel-level **C**ontrastive **S**ingle **D**omain **G**eneralization (**C^{2}SDG**) for joint optic cup (OC) and optic disc (OD) segmentation on fundus images. In C^{2}SDG, the shallower features of each image and its style-augmented counterpart are extracted and used for contrastive training, resulting in the disentangled style representations

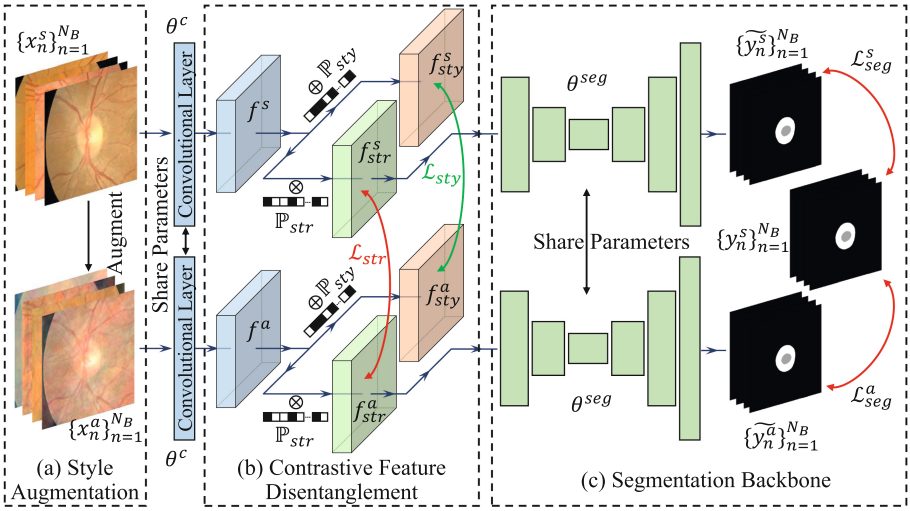

Fig. 2. Diagram of our C^{2}SDG. The rectangles in blue and green represent the convolutional layer and the segmentation backbone, respectively. The cubes represent different features. The projectors with parameters θ^p in (b) are omitted for simplicity. (Color figure online)

and structure representations. The segmentation is performed based solely on the structure representations. This method has been evaluated against other SDG methods on a public dataset and improved performance has been achieved. Our main contributions are three-fold: (1) we propose a novel contrastive perspective for SDG, enabling contrastive feature disentanglement using the data from only a single domain; (2) we disentangle the style representations and structure representations explicitly and channel-wisely, and then diminish the impact of style-sensitive 'devil' channels; and (3) our C^{2}SDG outperforms the baseline and six state-of-the-art SDG methods on the joint OC/OD segmentation benchmark.

2 Method

2.1 Problem Definition and Method Overview

Let the source domain be denoted by $\mathcal{D}^s = \{x_i^s, y_i^s\}_{i=1}^{N_s}$, where x_i^s is the i-th source domain image, and y_i^s is its segmentation mask. Our goal is to train a segmentation model $F_\theta : x \to y$ on $\mathcal{D}^s$, which can generalize well to an unseen target domain $\mathcal{D}^t = \{x_i^t\}_{i=1}^{N_t}$. The proposed C^{2}SDG mainly consists of a segmentation backbone, a style augmentation (StyleAug) module, and a contrastive feature disentanglement (CFD) module. For each image x^s, the StyleAug module generates its style-augmented counterpart x^a, which shares the same structure but different style to x^s. Then a convolutional layer extracts high-dimensional representations f^s and f^a from x^s and x^a. After that, f^s and f^a are fed to

the CFD module to perform contrastive training, resulting in the disentangled style representations f_{sty} and structure representations f_{str}. The segmentation backbone only takes f_{str} as its input and generates the segmentation prediction $\widetilde{y}$. Note that, although we take a U-shape network [5] as the backbone for this study, both StyleAug and CFD modules are modularly designed and can be incorporated into most segmentation backbones. The diagram of our C^{2}SDG is shown in Fig. 2. We now delve into its details.

2.2 Style Augmentation

Given a batch of source domain data $\{x_n^s\}_{n=1}^{N_B}$, we adopt a series of style-related data augmentation approaches, *i.e.*, gamma correction and noise addition in BigAug [21], and Bezier curve transformation in SLAug [15], to generate $\{x_n^{BA}\}_{n=1}^{N_B}$ and $\{x_n^{SL}\}_{n=1}^{N_B}$.

Additionally, to fully utilize the style diversity inside single domain data, we also adopt low-frequency components replacement [20] within a batch of source domain images. Specifically, We reverse $\{x_n^s\}_{n=1}^{N_B}$ to match x_n^s with x_r^s, where $r = N_B + 1 - n$ to ensure x_r^s provides a different reference style. Then we transform x_n^s and x_r^s to the frequency domain and exchange their low-frequency components $Low(Amp(x^s); \beta)$ in the amplitude map, where β is the cut-off ratio between low and high-frequency components and is randomly selected from (0.05, 0.15]. After that, we recover all low-frequency exchanged images to generate $\{x_n^{FR}\}_{n=1}^{N_B}$.

The style-augmented images batch $\{x_n^a\}_{n=1}^{N_B}$ is set to $\{x_n^{BA}\}_{n=1}^{N_B}$, $\{x_n^{SL}\}_{n=1}^{N_B}$, and $\{x_n^{FR}\}_{n=1}^{N_B}$ in turn to perform contrastive training and segmentation.

2.3 Contrastive Feature Disentanglement

Given x^s and x^a, we use a convolutional layer with parameter θ^c to generate their shallow features f^s and f^a, which are 64-channel feature maps for this study.

Then we use a channel mask prompt $\mathbb{P} \in \mathbb{R}^{2\times 64}$ to disentangle each shallow feature map f into style representation f_{sty} and structure representation f_{str} explicitly channel-wisely

$$\begin{cases} f_{sty} = f \times \mathbb{P}_{sty} = f \times SM(\frac{\mathbb{P}}{\tau})_1 \\ f_{str} = f \times \mathbb{P}_{str} = f \times SM(\frac{\mathbb{P}}{\tau})_2, \end{cases} \tag{1}$$

where $SM(\cdot)$ is a softmax function, the subscript i denotes i-th channel, and $\tau = 0.1$ is a temperature factor that encourages $\mathbb{P}_{sty}$ and $\mathbb{P}_{str}$ to be binary-element vectors, *i.e.*, approximately belonging to $\{0, 1\}^{64}$.

After channel-wise feature disentanglement, we have $\{f_{sty}^s, f_{str}^s\}$ from x^s and $\{f_{sty}^a, f_{str}^a\}$ from x^a. It is expected that (a) f_{sty}^s and f_{sty}^a are different since we want to identify them as the style-sensitive 'devil' channels, and (b) f_{str}^s and f_{str}^a are the same since we want to identify them as the style-irrelevant channels

and x^s and x^a share the same structure. Therefore, we design two contrastive loss functions $\mathcal{L}_{sty}$ and $\mathcal{L}_{str}$

$$\begin{cases} \mathcal{L}_{str} = \sum |Proj(f^s_{str}) - Proj(f^a_{str})| \\ \mathcal{L}_{sty} = -\sum |Proj(f^s_{sty}) - Proj(f^a_{sty})|, \end{cases} \tag{2}$$

where the $Proj(\cdot)$ with parameters θ^p reduces the dimension of f_{str} and f_{sty}.

Only f^s_{str} and f^a_{str} are fed to the segmentation backbone with parameters θ^{seg} to generate the segmentation predictions $\widetilde{y^s}$ and $\widetilde{y^a}$.

2.4 Training and Inference

Training. For the segmentation task, we treat OC/OD segmentation as two binary segmentation tasks and adopt the binary cross-entropy loss as our objective

$$\mathcal{L}_{ce}(y, \widetilde{y}) = -(\widetilde{y} \log y + (1 - \widetilde{y}) \log (1 - y)) \tag{3}$$

where y represents the segmentation ground truth and $\widetilde{y}$ is the prediction. The total segmentation loss can be calculated as

$$\mathcal{L}_{seg} = \mathcal{L}^s_{seg} + \mathcal{L}^a_{seg} = \mathcal{L}_{ce}(y^s, \widetilde{y^s}) + \mathcal{L}_{ce}(y^s, \widetilde{y^a}). \tag{4}$$

During training, we alternately minimize $\mathcal{L}_{seg}$ to optimize $\{\theta^c, \mathbb{P}, \theta^{seg}\}$, and minimize $\mathcal{L}_{str} + \mathcal{L}_{sty}$ to optimize $\{\mathbb{P}, \theta^p\}$.

Inference. Given a test image x^t, its shallow feature map f^t can be extracted by the first convolutional layer. Based on f^t, the optimized channel mask prompt $\mathbb{P}$ can separate it into f^t_{sty} and f^t_{str}. Only f^t_{str} is fed to the segmentation backbone to generate the segmentation prediction $\widetilde{y^t}$.

3 Experiments and Results

Materials and Evaluation Metrics. The multi-domain joint OC/OD segmentation dataset RIGA+ [1,4,8] was used for this study. It contains annotated fundus images from five domains, including 195 images from BinRushed, 95 images from Magrabia, 173 images from BASE1, 148 images from BASE2, and 133 images from BASE3. Each image was annotated by six raters, and only the first rater's annotations were used in our experiments. We chose BinRushed and Magrabia, respectively, as the source domain to train the segmentation model, and evaluated the model on the other three (target) domains. We adopted the Dice Similarity Coefficient (D, %) to measure the segmentation performance.

Implementation Details. The images were center-cropped and normalized by subtracting the mean and dividing by the standard deviation. The input batch contains eight images of size 512×512. The U-shape segmentation network, whose encoder is a modified ResNet-34, was adopted as the segmentation backbone of our C^{2}SDG and all competing methods for a fair comparison. The

Table 1. Average performance of three trials of our C^{2}SDG and six competing methods in joint OC/OD segmentation using BinRushed (row 2–row 9) and Magrabia (row 10–row 17) as source domain, respectively. Their standard deviations are reported as subscripts. The performance of 'Intra-Domain' and 'w/o SDG' is displayed for reference. The best results except for 'Intra-Domain' are highlighted in blue.

Methods	BASE1		BASE2		BASE3		Average	
	D_{OD}	D_{OC}	D_{OD}	D_{OC}	D_{OD}	D_{OC}	D_{OD}	D_{OC}
Intra-Domain	$94.71_{0.07}$	$84.07_{0.35}$	$94.84_{0.18}$	$86.32_{0.14}$	$95.40_{0.05}$	$87.34_{0.11}$	94.98	85.91
w/o SDG	$91.82_{0.54}$	$77.71_{0.88}$	$79.78_{2.10}$	$65.18_{3.24}$	$88.83_{2.15}$	$75.29_{3.23}$	86.81	72.73
BigAug [23]	$94.01_{0.34}$	$81.51_{0.58}$	$85.81_{0.68}$	$71.12_{1.64}$	$92.19_{0.51}$	$79.75_{1.44}$	90.67	77.46
CISDG [13]	$93.56_{0.13}$	$81.00_{1.01}$	$94.38_{0.23}$	$83.79_{0.58}$	$93.87_{0.03}$	$83.75_{0.89}$	93.93	82.85
ADS [19]	$94.07_{0.29}$	$79.60_{5.06}$	$94.29_{0.38}$	$81.17_{3.72}$	$93.64_{0.28}$	$81.08_{4.97}$	94.00	80.62
MaxStyle [2]	$94.28_{0.14}$	$82.61_{0.67}$	$86.65_{0.76}$	$74.71_{2.07}$	$92.36_{0.39}$	$82.33_{1.24}$	91.09	79.88
SLAug [15]	$95.28_{0.12}$	$83.31_{1.10}$	$95.49_{0.16}$	$81.36_{2.51}$	$95.57_{0.06}$	$84.38_{1.39}$	95.45	83.02
Dual-Norm [23]	$94.57_{0.10}$	$81.81_{0.76}$	$93.67_{0.11}$	$79.16_{1.80}$	$94.82_{0.28}$	$83.67_{0.60}$	94.35	81.55
Ours	$95.73_{0.08}$	$86.13_{0.07}$	$95.73_{0.09}$	$86.82_{0.58}$	$95.45_{0.04}$	$86.77_{0.19}$	95.64	86.57
w/o SDG	$89.98_{0.54}$	$77.21_{1.15}$	$85.32_{1.79}$	$73.51_{0.67}$	$90.03_{0.27}$	$80.71_{0.63}$	88.44	77.15
BigAug [23]	$92.32_{0.13}$	$79.68_{0.38}$	$88.24_{0.82}$	$76.69_{0.37}$	$91.35_{0.14}$	$81.43_{0.78}$	90.64	79.27
CISDG [13]	$89.67_{0.76}$	$75.39_{3.22}$	$87.97_{1.04}$	$76.44_{3.48}$	$89.91_{0.64}$	$81.35_{2.81}$	89.18	77.73
ADS [19]	$90.75_{2.42}$	$77.78_{4.23}$	$90.37_{2.07}$	$79.60_{3.34}$	$90.34_{2.93}$	$79.99_{4.02}$	90.48	79.12
MaxStyle [2]	$91.63_{0.12}$	$78.74_{1.95}$	$90.61_{0.45}$	$80.12_{0.90}$	$91.22_{0.07}$	$81.90_{1.14}$	91.15	80.25
SLAug [15]	$93.08_{0.17}$	$80.70_{0.35}$	$92.70_{0.12}$	$80.15_{0.43}$	$92.23_{0.16}$	$80.89_{0.14}$	92.67	80.58
Dual-Norm [23]	$92.35_{0.37}$	$79.02_{0.39}$	$91.23_{0.29}$	$80.06_{0.26}$	$92.09_{0.28}$	$79.87_{0.25}$	91.89	79.65
Ours	$94.78_{0.03}$	$84.94_{0.36}$	$95.16_{0.09}$	$85.68_{0.28}$	$95.00_{0.09}$	$85.98_{0.29}$	94.98	85.53

projector in our CFD module contains a convolutional layer followed by a batch normalization layer, a max pooling layer, and a fully connected layer to convert f_{sty} and f_{str} to 1024-dimensional vectors. The SGD algorithm with a momentum of 0.99 was adopted as the optimizer. The initial learning rate was set to $lr_0 = 0.01$ and decayed according to $lr = lr_0 \times (1 - e/E)^{0.9}$, where e is the current epoch and $E = 100$ is the maximum epoch. All experiments were implemented using the PyTorch framework and performed with one NVIDIA 2080Ti GPU.

Comparative Experiments. We compared our C^{2}SDG with two baselines, including 'Intra-Domain' (*i.e.*, training and testing on the data from the same target domain using 3-fold cross-validation) and 'w/o SDG' (*i.e.*, training on the source domain and testing on the target domain), and six SDG methods, including BigAug [21], CISDG [13], ADS [19], MaxStyle [2], SLAug [15], and Dual-Norm [23]. In each experiment, only one source domain is used for training, ensuring that only the data from a single source domain can be accessed during training. For a fair comparison, all competing methods are re-implemented using the same backbone as our C^{2}SDG based on their published code and paper. The results of C^{2}SDG and its competitors were given in Table 1. It shows that C^{2}SDG

improves the performance of 'w/o SDG' with a large margin and outperforms all competing SDG methods. We also visualize the segmentation predictions generated by our C^{2}SDG and six competing methods in Fig. 3. It reveals that our C^{2}SDG can produce the most accurate segmentation map.

Ablation Analysis. To evaluate the effectiveness of low-frequency components replacement (FR) in StyleAug and CFD, we conducted ablation experiments using BinRushed and Magrabia as the source domain, respectively. The average performance is shown in Table 2. The performance of using both BigAug and SLAug is displayed as 'Baseline'. It reveals that both FR and CFD contribute to performance gains.

Analysis of CFD. Our CFD is modularly designed and can be incorporated into other SDG methods. We inserted our CFD to ADS [19] and SLAug [15], respectively. The performance of these two approaches and their variants, denoted as C^2-ADS and C^2-SLAug, was shown in Table 3. It reveals that our CFD module can boost their ability to disentangle structure representations and improve the segmentation performance on the target domain effectively. We also adopted 'Ours w/o CFD' as 'Baseline' and compared the channel-level contrastive feature disentanglement strategy with the adversarial training strategy and channel-level

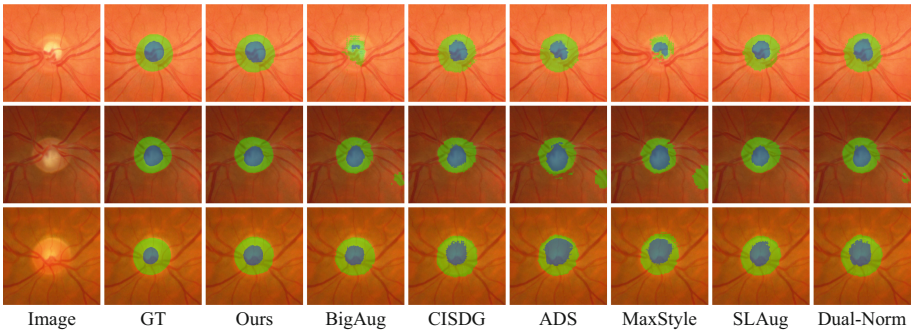

Fig. 3. Visualization of segmentation masks predicted by our C^{2}SDG and six competing methods, together with ground truth.

Table 2. Average performance of our C^{2}SDG and three variants.

Methods	Average	
	D_{OD}	D_{OC}
Baseline	94.13	81.62
w/o FR	95.07	84.90
w/o CFD	95.07	84.83
Ours	95.31	86.05

Table 3. Average performance of ADS, SLAug, and their two variants.

Methods	Average	
	D_{OD}	D_{OC}
ADS [19]	92.24	79.87
C^2-ADS	93.76	81.35
SLAug [15]	94.06	81.80
C^2-SLAug	94.24	83.68

Table 4. Average performance of using contrastive and other strategies.

Methods	Average	
	D_{OD}	D_{OC}
Baseline	95.07	84.83
Dropout	95.14	84.95
Adversarial	90.27	78.47
Ours	95.31	86.05

dropout (see Table 4). It shows that the adversarial training strategy fails to perform channel-level feature disentanglement, due to the limited training data [3] for SDG. Nonetheless, our channel-level contrastive learning strategy achieves the best performance compared to other strategies, further confirming the effectiveness of our CFD module.

4 Conclusion

In this paper, we propose a novel SDG method called C^2SDG for medical image segmentation. In C^2SDG, the StyleAug module generates style-augmented counterpart of each source domain image and enables contrastive learning, the CFD module performs channel-level style and structure representations disentanglement via optimizing a channel prompt $\mathbb{P}$, and the segmentation is performed based solely on structure representations. Our results on a multi-domain joint OC/OD segmentation benchmark indicate the effectiveness of StyleAug and CFD and also suggest that our C^2SDG outperforms the baselines and six completing SDG methods with a large margin.

Acknowledgement. This work was supported in part by the National Natural Science Foundation of China under Grant 62171377, in part by the Key Technologies Research and Development Program under Grant 2022YFC2009903/2022YFC2009900, in part by the Key Research and Development Program of Shaanxi Province, China, under Grant 2022GY-084, in part by the Innovation Foundation for Doctor Dissertation of Northwestern Polytechnical University under Grant CX2023016.

References

1. Almazroa, A., et al.: 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. Chen, C., Li, Z., Ouyang, C., Sinclair, M., Bai, W., Rueckert, D.: MaxStyle: adversarial style composition for robust medical image segmentation. In: Wang, L., Dou, Q., Fletcher, P.T., Speidel, S., Li, S. (eds.) MICCAI 2022. LNCS, vol. 13435, pp. 151–161. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-16443-9_15
3. Clarysse, J., Hörrmann, J., Yang, F.: Why adversarial training can hurt robust accuracy. In: International Conference on Learning Representations (ICLR) (2023). https://openreview.net/forum?id=-CA8yFkPc7O
4. Decencière, E., et al.: Feedback on a publicly distributed image database: the Messidor database. Image Anal. Stereol. **33**(3), 231–234 (2014)
5. Falk, T., et al.: U-net: deep learning for cell counting, detection, and morphometry. Nat. Methods **16**(1), 67–70 (2019)
6. Frankle, J., Carbin, M.: The lottery ticket hypothesis: finding sparse, trainable neural networks. In: International Conference on Learning Representations (ICLR) (2018)
7. Guan, H., Liu, M.: Domain adaptation for medical image analysis: a survey. IEEE Trans. Biomed. Eng. **69**(3), 1173–1185 (2021)

8. Hu, S., Liao, Z., Xia, Y.: Domain specific convolution and high frequency reconstruction based unsupervised domain adaptation for medical image segmentation. In: Wang, L., Dou, Q., Fletcher, P.T., Speidel, S., Li, S. (eds.) MICCAI 2022. LNCS, vol. 13437, pp. 650–659. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-16449-1_62
9. Hu, S., Liao, Z., Xia, Y.: ProSFDA: prompt learning based source-free domain adaptation for medical image segmentation. arXiv preprint arXiv:2211.11514 (2022)
10. Hu, S., Liao, Z., Zhang, J., Xia, Y.: Domain and content adaptive convolution based multi-source domain generalization for medical image segmentation. IEEE Trans. Med. Imaging **42**(1), 233–244 (2022)
11. Litjens, G., et al.: A survey on deep learning in medical image analysis. Med. Image Anal. **42**, 60–88 (2017). https://doi.org/10.1016/j.media.2017.07.005
12. Ma, H., Lin, X., Yu, Y.: I2F: a unified image-to-feature approach for domain adaptive semantic segmentation. IEEE Trans. Pattern Anal. Mach. Intell. (2022)
13. Ouyang, C., et al.: Causality-inspired single-source domain generalization for medical image segmentation. IEEE Trans. Med. Imaging **42**, 1095–1106 (2022)
14. Sprawls, P.: Image characteristics and quality. In: Physical Principles of Medical Imaging, pp. 1–16. Aspen Gaithersburg (1993)
15. Su, Z., Yao, K., Yang, X., Wang, Q., Sun, J., Huang, K.: Rethinking data augmentation for single-source domain generalization in medical image segmentation. In: AAAI Conference on Artificial Intelligence (AAAI) (2023)
16. Wang, J., et al.: Generalizing to unseen domains: a survey on domain generalization. IEEE Trans. Knowl. Data Eng. **35**, 8052–8072 (2022)
17. Wilson, G., Cook, D.J.: A survey of unsupervised deep domain adaptation. ACM Trans. Intell. Syst. Technol. **11**(5), 1–46 (2020)
18. 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. Med. Image Anal. **69**, 101985 (2021). https://doi.org/10.1016/j.media.2021.101985
19. Xu, Y., Xie, S., Reynolds, M., Ragoza, M., Gong, M., Batmanghelich, K.: Adversarial consistency for single domain generalization in medical image segmentation. In: Wang, L., Dou, Q., Fletcher, P.T., Speidel, S., Li, S. (eds.) MICCAI 2022. LNCS, vol. 13437, pp. 671–681. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-16449-1_64
20. Yang, Y., Soatto, S.: FDA: fourier domain adaptation for semantic segmentation. In: IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 4085–4095 (2020)
21. Zhang, L., et al.: Generalizing deep learning for medical image segmentation to unseen domains via deep stacked transformation. IEEE Trans. Med. Imaging **39**(7), 2531–2540 (2020)
22. Zhou, K., Liu, Z., Qiao, Y., Xiang, T., Loy, C.C.: Domain generalization: a survey. IEEE Trans. Pattern Anal. Mach. Intell. **45**, 4396–4415 (2022)
23. Zhou, Z., Qi, L., Yang, X., Ni, D., Shi, Y.: Generalizable cross-modality medical image segmentation via style augmentation and dual normalization. In: IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 20856–20865 (2022)