



# Adversarial Consistency for Single Domain Generalization in Medical Image Segmentation

Yanwu Xu<sup>1</sup>(✉), Shaoan Xie<sup>3</sup>, Maxwell Reynolds<sup>1</sup>, Matthew Ragoza<sup>1</sup>,  
Mingming Gong<sup>2</sup>, and Kayhan Batmanghelich<sup>1</sup>

<sup>1</sup> Department of Biomedical Informatics, University of Pittsburgh, Pittsburgh, USA  
yanwuxu@pitt.edu

<sup>2</sup> School of Mathematics and Statistics, The University of Melbourne, Melbourne, Australia

<sup>3</sup> Department of Philosophy, Carnegie Mellon University, Pittsburgh, USA

**Abstract.** An organ segmentation method that can generalize to unseen contrasts and scanner settings can significantly reduce the need for retraining of deep learning models. Domain Generalization (DG) aims to achieve this goal. However, most DG methods for segmentation require training data from multiple domains during training. We propose a novel adversarial domain generalization method for organ segmentation trained on data from a *single* domain. We synthesize the new domains via learning an adversarial domain synthesizer (ADS) and presume that the synthetic domains cover a large enough area of plausible distributions so that unseen domains can be interpolated from synthetic domains. We propose a mutual information regularizer to enforce the semantic consistency between images from the synthetic domains, which can be estimated by patch-level contrastive learning. We evaluate our method for various organ segmentation for unseen modalities, scanning protocols, and scanner sites.

**Keywords:** Medical image segmentation · Single domain generalization · Adversarial training · Mutual information

## 1 Introduction

Deep Learning-based methods for the segmentation of medical images hold state-of-the-art performance across various organs and anatomies [13, 18, 24]. The independent and identically distributed (*i.i.d.*) is the underlying assumption of most of those methods. However, the difference in the image acquisition, such as scanning protocol and image modality, introduces domain shifts, rendering the assumption impractical. Domain Adaptation [2, 6, 7, 20] (DA) and Multi-source Domain generalization [11] (MDG) aim to alleviate the domain shift

---

M. Gong and K. Batmanghelich—Equal Contribution.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2022  
L. Wang (Eds.): MICCAI 2022, LNCS 13437, pp. 671–681, 2022.  
[https://doi.org/10.1007/978-3-031-16449-1\\_64](https://doi.org/10.1007/978-3-031-16449-1_64)

issue. However, those approaches are not data-efficient as they either require access to test distribution (i.e., DA) or need multiple labeled source domains during training (i.e., MDG). In this paper, we focus on *single-source* domain generalization (SDG), aiming to train a generalizable deep model on only one source domain.

The *SDG* does not require access to the test distribution or labeled data from multiple sources during the training. As a result, it reduces annotation costs and avoids repetitive adaptation for each new domain. In the literature, various SDG methods have been proposed that are based on augmentation of input image [15, 22] and meta learning [17]. The meta-learning techniques tend to be extremely slow during inference time. The augmentation methods synthesize new images using random initialization of convolution filter [15, 22]. However, those methods cannot avoid over-fitting to a regular pattern of synthetic data. Thus, we propose synthesizing the new domains via learning an adversarial framework.

We propose synthesizing the new domains via learning an adversarial domain synthesizer (ADS). The intuition is that the synthetic domains cover a large enough area of plausible distributions so that unseen domains can be interpolated from synthetic domains. Specifically, we design the synthesizer with a random style module, enabling ADS to synthesize random textures during adversarial training. Without a constraint, adversarial training may change the image semantics, making the synthetic domains irrelevant. To remedy this problem, we propose to keep the underlying semantic information between the source image and the synthetic image via a mutual information regularizer. As estimating mutual information is hard for high dimensional data, we utilize the patch-level contrastive loss [14] as a surrogate to maximize the mutual information between the original and synthesized images.

The main contributions of this work can be summarized as follows: 1) We propose an adversarial framework for single domain generalization of medical image segmentation. 2) We redesign the network structure of synthesizing new domains. 3) To constrain the adversarial training, we propose a regularization method for synthetic images. To evaluate our model, we conduct experiments of single domain generalization of medical image segmentation on cross-modality image segmentation (CT  $\rightarrow$  MRI), -imaging protocol, and -organizations.

## 2 Related Works

*Unsupervised Domain Adaptation and Domain Generalization.* Unsupervised Domain Adaptation (UDA) is a proposed strategy for addressing domain shift between the training data and testing data of deployed applications. In the UDA framework, the model has access to labelled data from a source domain and unlabelled data from a target domain during training. Prior work on UDA can be classified as distribution alignment [2, 6] or self-supervised learning [4].

Multi-domain generalization (MDG) is an alternative framework where the goal is to learn domain-invariant features from multiple labelled source domains. Recent works [11, 12] give a theoretical proof that domain-invariant features can be learned using MDG. The common goal of UDA and MDG is to learn domain-invariant features for downstream tasks (i.e. classification, segmentation and object detection) by training a model on data from multiple domains. Therefore, UDA and MDG both require access to data from multiple domains during training. For single domain generalization, M-ADA [17] and L2D [21] propose an adversarial training framework for SDG learning. M-ADA proposes a meta-scheme method to find the adversarial perturbation, which calculates the gradient direction for each specific input and is extremely slow for real-life applications. While similar to RandConv [22] and GIN [15], M-ADA also proposes to synthesize new domain but with an adversarial training. However, M-ADA constrains the perturbation in latent space and cannot guarantee the semantic consistency for segmentation task, i.e. object boundary and shape, which is especially crucial for medical image segmentation.

*Data Augmentation.* Data Augmentation is a low-level data sampling technique, which gains a free performance improvement with use of human prior. To further improve the generalizability of models, AdaTransform [20] proposes to augment the data by maximizing the entropy of the data. AdvBias [3] also apply the adversarial technique for learning the bias-field and deformation field respectively. Cutout [5] randomly samples masks and crops out part of images via the masks. Mixup [23] simply mixes the images and also forces the prediction of mixed images to be the interpolation of ground truth labels.

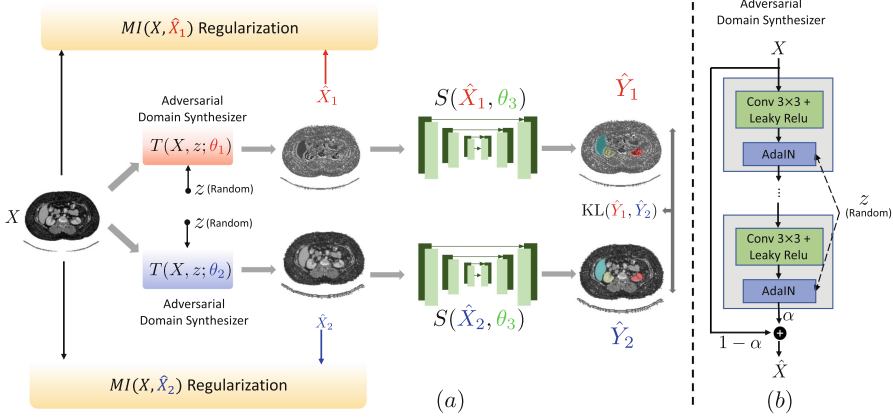
### 3 Proposed Method

In the single domain generalization setting, we aim at training a segmentation model  $S$  on a single source domain  $\mathcal{S} = \{x^i, y^i\}_{i=1}^N$  such that the learned model can generalize well to multiple target domains  $\mathcal{T}_k = \{x_k^j, y_k^j\}_{j=1}^M$ , where  $k$  is the domain indicator. The target domains are not accessible during training.

We propose an adversarial training-based method to synthesize data from new domains  $\hat{\mathcal{T}}_l = \{\hat{x}_l^i, \hat{y}_l^i\}_{i=1}^N$  from the single source domain, where  $l$  is the synthetic domain indicator. An assumption of our method is that the space of the collections of the unseen domains belongs to the collections of synthetic domains, i.e.  $\cup_k \mathcal{T}_k \subset \cup_l \hat{\mathcal{T}}_l$ . Specifically, the unseen domains can be generated by interpolating the synthetic domains. Our adversarial framework can enlarge the diversity the synthetic domains and guarantee sufficient coverage of the unseen domains. To ensure the synthetic images adhere to the semantics of the source images, we also apply a regularizer that promotes the mutual information between synthetic images and source images.

Our method consists of four modules: two adversarial domain synthesizers parameterized as  $T(\cdot; \theta_1)$  and  $T(\cdot; \theta_2)$ , a segmentation network parameterized as

$S(\cdot; \theta_3)$  and the mutual information regularizer  $MI(\cdot, \cdot; \theta_4)$ . We train the models with a two-step min-max procedure that involves supervised, consistency, adversarial, and regularization loss terms. A schematic illustration of our proposed model is shown in Fig. 1.



**Fig. 1.** Schematic of the proposed method. (a) shows the completed model structure, which consists of our adversarial domain synthesizer (ADS), the mutual information regularization between input and synthetic image, the segmentation network and the KL consistency loss between two predictions. (b) is the detailed structure of our proposed ADS. (Color figure online)

### 3.1 Adversarial Domain Synthesizer

To maximize the effectiveness of the synthetic domains, we utilize adversarial training to learn an Adversarial Domain Synthesizer (ADS)  $T(X, z; \theta)$ , which takes as input a source image and random noise sample  $z$  and outputs a synthetic image  $\hat{X}$ . We assume that  $T$  only changes the texture of the source images and not their segmentation annotation. As a result, the segmentation network  $S$  should generate the same segmentation mask for all  $\hat{X}$  from a given  $X$ . For the adversarial training,  $T$  competes with  $S$  to generate the adversarial domain images. We construct the adversarial consistency via two differently parameterized  $T$ , and we can obtain two randomly synthesized images  $\hat{X}_1 = T(X, z; \theta_1)$ ,  $\hat{X}_2 = T(X, z; \theta_2)$ . The purpose of random variable  $z$  will be explained below. The adversarial competition can be formulated as follows:

$$\ell_{Con}(\hat{X}_1, \hat{X}_2) = KL(S(\hat{X}_1; \theta_3) || S(\hat{X}_2; \theta_3)), \quad (1)$$

where KL divergence is measured between the two softmax outputs. In Eq. 1, the two synthesizers  $T$  are trained to maximize  $\ell_{Consistency}$  and  $S$  is trained to minimize it.

*ADS Network Architecture.* The goal of the ADS  $T$  is to generate the synthetic domains with hardest perturbation for  $S$ . Meanwhile, the semantics of the images, such as the boundary and shape of the organs, should keep consistent with source images and the changes are limited to the texture of the tissues. To achieve the effect mentioned above, we proposed a modified version of Global Intensity Non-Linear Augmentation module (GIN) [15]. GIN utilizes a shallow CNN structure with randomly initialized weights for the convolution filters plus a mixup with ratio of  $\alpha \sim U(0, 1)$  to generate multi-modalities from source domain. Unlike GIN, in our method, the synthesizer  $T$  is learned adversarially, so we cannot initialize  $T$  randomly to generalize multiple modalities. Thus, without randomness from the parameters of  $T$ , we can only interpolate the domains with the mixup parameter, which limits the variance of synthetic domains. To alleviate this issue, we incorporate the adaptive instance normalization (AdaIN) block [8] to introduce randomness to the generated domain. The design of the proposed  $T$  is shown in Fig. 1 (b) (Fig. 2).

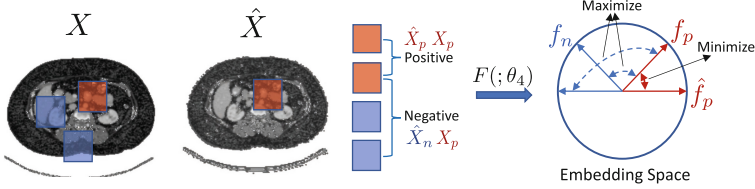


Fig. 2. Mutual information maximization.

### 3.2 MI Regularization

Without proper regularization,  $T$  can change semantics of the image arbitrarily. To ensure that the boundary and shape are shared underlying semantics between the synthetic image and the source image, we can maximize the mutual information between two images to keep the semantics. Thus, we propose to add the mutual information maximization (MIM) constraint to the synthesized images as maximizing  $MI(X, \hat{X})$ , where  $\theta_4$  is the parameters of the mutual information estimator. Exact calculation of mutual information is computationally prohibitive. Inspired by the success of contrastive learning in image-to-image translation [16], we use patch-based contrastive loss as a surrogate for MI. Specifically, we utilize contrastive learning on the image patch level to maximize the MI between the source images and the transformed images, which exactly satisfies our goal. Thus, we can adapt [16] to our model and reformulate it as follows:

$$\ell_{MI}(X, \hat{X}) = \log \frac{\exp(\hat{f}_p \cdot f_p / \tau)}{\exp(\hat{f}_p \cdot f_p / \tau) + \sum_{n \neq p} \exp(f_p \cdot f_n / \tau)}. \quad (2)$$

In Eq. 2, the  $\ell_{MI}$  is maximized and the numerator is the inner product of two corresponding patch features with exponential re-scaling, the  $\hat{f}_p$  and  $f_p$  denote the features extracted by  $F(\hat{X}_p; \theta_4)$ ,  $F(X_p; \theta_4)$ , where  $\hat{X}_p, X_p$  represent the patches of the synthesized image and the source image at the corresponding location respectively and they are the positive pair for contrastive learning. For the negative pair, we can see the same operation of the  $f_p$  and  $f_q$  in the denominator. In Eq. 2, we are trying to maximize the semantic correlation between the synthesized image and the input via pushing similarity of patches at the corresponding location and dissimilarity at different locations.

### 3.3 Model Optimization

As mentioned above, the segmentation masks are kept consistent from the source domain to synthesized domains. Thus we also have the supervised loss as below:

$$\ell_{Sup}(\hat{X}, Y) = \text{KL}(S(\hat{X}; \theta_3) || Y). \quad (3)$$

Our two-step optimization for the proposed model can be summarized as:

$$\begin{aligned} \text{update } \theta_3 : \min_{\theta_3} \mathbb{E}_{X, Y \sim P_{XY}} \ell_{Sup}(\hat{X}_1, Y) + \ell_{Sup}(\hat{X}_2, Y) + \ell_{Cons}(\hat{X}_1, \hat{X}_2) \\ \text{update } \theta_1, \theta_2, \theta_4 : \max_{\theta_1, \theta_2, \theta_4} \mathbb{E}_{X \sim P_X, z \sim P_z} \ell_{Con}(\hat{X}_1, \hat{X}_2) + \ell_{MI}(X, \hat{X}_1) + \ell_{MI}(X, \hat{X}_2). \end{aligned}$$

## 4 Experiments

We evaluate our model on two experimental settings. For the first setting, we test the generalizability of the model on an unseen modality by using abdominal CT scans from [1] and MRI scans from [9] as the source domain and target domain respectively. For the second setting, we aim to test the generalizability of our model on unseen data distribution shift caused by different scanner machine settings. We obtain the data from six different organizations (RUNMC, BMC, HCRUDb, UCL, BIDMC, HK) and adopt the following cross-validation procedure: we select the data from one organization as the source domain for training and hold out the rest as the target domains for testing. Furthermore, we conduct a detailed ablation study to analyze the effectiveness of each component of the proposed method.

### 4.1 Training Configuration

We use Efficient-b2 [19] as the backbone for the segmentation network  $S$  and modify it with UNet-style skip connection. The synthesizer network  $T$  has 4 convolutional blocks, each block consisting of  $3 \times 3$  convolutional kernel with the channel size of 2, Leaky-ReLU and AdaIN. For the MIM model, we use the encoder part of the generator with a fully connected layer as the feature

extractor  $F$  from [16]. We choose Adam optimizer [10] for all of the models with initial learning rate  $3 \times 10^{-4}$  and  $\beta = [0.5, 0.999]$ . The learning rate will linearly decay to zero at the end of the training. The total training epochs are 2,000. We compare our method to Cutout, AdvBias, RandConv and GIN [15] under the same settings.

## 4.2 Data Preprocessing and Evaluation Metrics

Before training and testing, we first resize all of the 3D volumes at the axial plane to the size of  $192 \times 192$ . For the CT modality, we clip the intensity into the range of  $[-275, 125]$ . For the MRI modality, the top 99.5% of intensity values are cut out. We also apply all variations of predefined augmentations to the 3D volumes before we slice them into 2D images. These augmentations include random contrast via gamma transformation, Gaussian noise addition, affine transformation, 3D elastic transformation and intensity normalization of zero mean and unit variance. All of these predefined augmentations are performed with MONAI. More details of the data preprocessing and augmentation steps can be found in Appendix A.

To evaluate single source domain generalization, we do not access the target testing dataset during training. Thus, we split all source datasets into training and evaluation partitions with proportions 70% and 30%, respectively. For the final testing on the target datasets, we pick the saved model checkpoint which has the best predictive performance on the source evaluation split. We use the Dice score (DS) for evaluating the performance of the different methods.

## 4.3 Experimental Results and Empirical Analysis

We summarize our results in Table 1. Across three different single domain generalization settings, our proposed method achieves the best performance compared with the baseline models. Notably, compared with [15], the overall performance on Cardiac bSSFP-LGE is worse than their reported results, although we have tried our best to replicate the baseline results under the suggested settings of [15]. We also visualize the segmentation results in the Fig. 3. Our method succeeds in some difficult samples, where the partial spleen is missing or mis-classified in the Abdominal CT-MRI, or missing semantic structure in Cardiac bSSFP-LGE. We also observe the coarse segmentation mask in Prostate Cross-Centers for the baseline methods. The above quantitative and qualitative results show the effectiveness of our adversarial training framework. more visual results are shown in Appendix B.

## 4.4 Ablation Study

We conduct extra experiments to study the effectiveness of each component in our model. In Table 2, we compare the performance results of our model without adversarial training, our model without the mutual information regularizer,

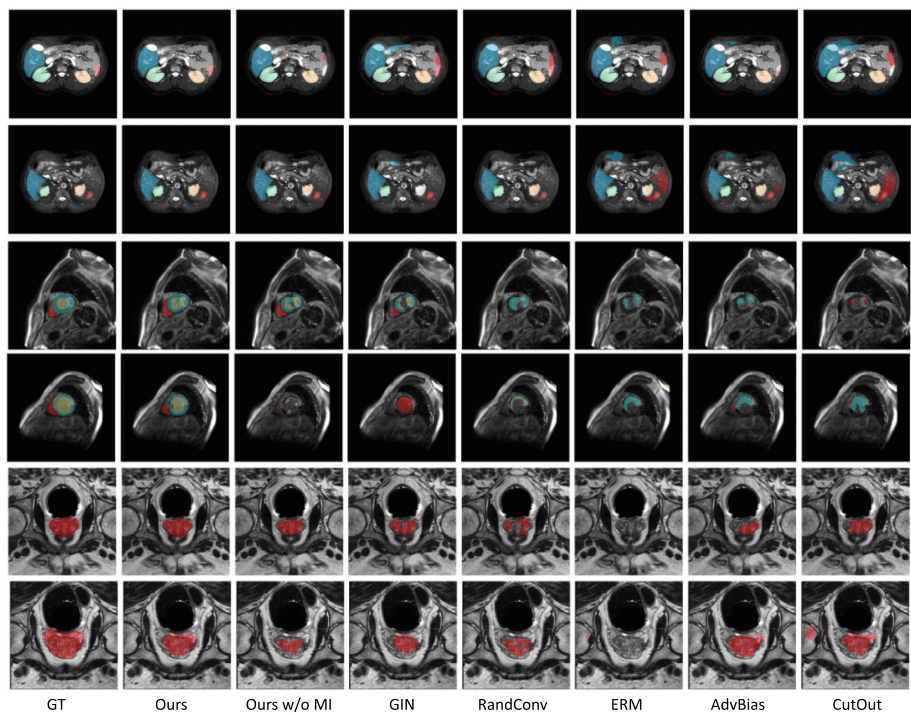


Fig. 3. Visualization of results.

Table 1. Results on three single source domain generalization settings.

| Method   | Abdominal CT-MRI |              |              |              |              | Cardiac bSSFP-LGE |              |              |              | Prostate Cross-center |
|----------|------------------|--------------|--------------|--------------|--------------|-------------------|--------------|--------------|--------------|-----------------------|
|          | Liver            | R-Kidney     | L-Kidney     | Spleen       | Average      | L-ventricle       | Myocardium   | R-ventricle  | Average      | Average               |
| ERM      | 73.35            | 75.92        | 74.20        | <b>82.93</b> | 76.60        | 51.19             | 72.7         | 70.40        | 64.76        | 51.57                 |
| Cutout   | 76.57            | 80.06        | 77.96        | 78.90        | 78.37        | 65.13             | 77.58        | 70.44        | 71.05        | 58.79                 |
| AdvBias  | 74.97            | 86.16        | 78.39        | 72.45        | 77.99        | 60.10             | 77.52        | 72.29        | 69.97        | 60.47                 |
| RandConv | 78.07            | 83.31        | 80.69        | 80.23        | 80.58        | 67.48             | 85.39        | 80.46        | 77.78        | 68.14                 |
| GIN      | 83.16            | 85.99        | 82.17        | 82.22        | 83.39        | 71.22             | 84.10        | <b>82.06</b> | 79.12        | 70.15                 |
| Ours     | <b>85.24</b>     | <b>89.87</b> | <b>86.92</b> | 81.69        | <b>85.93</b> | <b>73.37</b>      | <b>86.10</b> | 81.45        | <b>80.31</b> | <b>71.42</b>          |

and our completed model. This experiment is conducted in the Abdominal CT-MRI under the same setups as described above. The model without adversarial training still achieves better results than the model without mutual information regularization by a small gap. However, we can see that both adversarial training and the MI regularizer can cooperate with each other to achieve superior performance.



**Table 2.** Ablation on Abdominal CT  $\rightarrow$  MRI.

| Abdominal CT-MRI     |              |              |              |              |              |
|----------------------|--------------|--------------|--------------|--------------|--------------|
| Method               | Liver        | R-kidney     | L-kidney     | Spleen       | Average      |
| Ours w/o adversarial | 84.59        | 85.52        | <b>87.79</b> | 79.21        | 84.28        |
| Ours w/o MI          | 83.23        | 85.62        | 83.71        | <b>82.79</b> | 83.59        |
| Ours                 | <b>85.24</b> | <b>89.87</b> | 86.92        | 81.69        | <b>85.93</b> |

## 5 Conclusion

We successfully trained a model on a single source domain and deploy it on unseen target domains for the purpose of *single* domain generalization (SDG). Because of the limited training data and the unpredictable target domain shift, the generalizability of segmentation models is restricted, especially for medical images. Therefore, we propose a novel adversarial training framework to improve the generalizability of the SDG segmentation model. We synthesize the new domains via learning an adversarial domain synthesizer (ADS) in the proposed method. We assume that the unseen domains can be interpolated from synthetic domains, and the adversarial synthetic domains can guarantee sufficient coverage. To constrain the semantic consistency between synthetic images and the corresponding source images, we propose a mutual information regularizer, which can be estimated by patch-level contrastive learning. We evaluate our method for various organ segmentation for unseen modalities, scanning protocols, and scanner sites. The proposed method shows a consistent improvement over the baseline methods.

**Acknowledge.** This work was partially supported by NIH Award Number 1R01HL141813-01, NSF 1839332 Tripod+X, SAP SE, and Pennsylvania Department of Health. We are grateful for the computational resources provided by Pittsburgh Super-Computing grant number TG-ASC170024. MG is supported by Australian Research Council Project DE210101624. KZ would like to acknowledge the support by the National Institutes of Health (NIH) under Contract R01HL159805, by the NSF-Convergence Accelerator Track-D award #2134901, and by the United States Air Force under Contract No. FA8650-17-C7715.

## References

1. (Author Name Not Available): Segmentation outside the cranial vault challenge (2015). <https://doi.org/10.7303/SYN3193805>, <https://repo-prod.prod.sagebase.org/repo/v1/doi/locate?id=syn3193805&type=ENTITY>
2. Bousmalis, K., Trigeorgis, G., Silberman, N., Krishnan, D., Erhan, D.: Domain separation networks. In: Lee, D., Sugiyama, M., Luxburg, U., Guyon, I., Garnett, R. (eds.) *Advances in Neural Information Processing Systems*, vol. 29. Curran Associates, Inc. (2016). <https://proceedings.neurips.cc/paper/2016/file/45fbc6d3e05ebd93369ce542e8f2322d-Paper.pdf>

3. Chen, C., et al.: Realistic adversarial data augmentation for MR image segmentation. In: Martel, A.L., et al. (eds.) MICCAI 2020. LNCS, vol. 12261, pp. 667–677. Springer, Cham (2020). [https://doi.org/10.1007/978-3-030-59710-8\\_65](https://doi.org/10.1007/978-3-030-59710-8_65)
4. Chen, Y., Wei, C., Kumar, A., Ma, T.: Self-training avoids using spurious features under domain shift. CoRR abs/2006.10032 (2020). <https://arxiv.org/abs/2006.10032>
5. Devries, T., Taylor, G.W.: Improved regularization of convolutional neural networks with cutout. CoRR abs/1708.04552 (2017). <http://arxiv.org/abs/1708.04552>
6. Ganin, Y., et al.: Domain-adversarial training of neural networks. J. Mach. Learn. Res. **17**(59), 1–35 (2016). <http://jmlr.org/papers/v17/15-239.html>
7. Hoffman, J., et al.: CyCADA: cycle-consistent adversarial domain adaptation. In: Dy, J., Krause, A. (eds.) Proceedings of the 35th International Conference on Machine Learning. Proceedings of Machine Learning Research, vol. 80, pp. 1989–1998. PMLR, 10–15 Jul 2018. <https://proceedings.mlr.press/v80/hoffman18a.html>
8. Huang, X., Belongie, S.J.: Arbitrary style transfer in real-time with adaptive instance normalization. CoRR abs/1703.06868 (2017). <http://arxiv.org/abs/1703.06868>
9. Kavur, A.E., et al.: CHAOS challenge - combined (CT-MR) healthy abdominal organ segmentation. Med. Image Anal. **69**, 101950 (2021). <https://doi.org/10.1016/j.media.2020.101950>, <http://www.sciencedirect.com/science/article/pii/S1361841520303145>
10. Kingma, D.P., Ba, J.: Adam: a method for stochastic optimization (2014). <http://arxiv.org/abs/1412.6980>, cite [arxiv:1412.6980](https://arxiv.org/abs/1412.6980)[Comment](https://arxiv.org/abs/1412.6980#comment): Published as a conference paper at the 3rd International Conference for Learning Representations, San Diego (2015)
11. Li, Y., Tian, X., Gong, M., Liu, Y., Liu, T., Zhang, K., Tao, D.: Deep domain generalization via conditional invariant adversarial networks. In: Proceedings of the European Conference on Computer Vision (ECCV), September 2018
12. Liu, Q., Dou, Q., Heng, P.: Shape-aware meta-learning for generalizing prostate MRI segmentation to unseen domains. CoRR abs/2007.02035 (2020). <https://arxiv.org/abs/2007.02035>
13. Milletari, F., Navab, N., Ahmadi, S.A.: V-net: fully convolutional neural networks for volumetric medical image segmentation. In: 2016 Fourth International Conference on 3D Vision (3DV), pp. 565–571 (2016)
14. van den Oord, A., Li, Y., Vinyals, O.: Representation learning with contrastive predictive coding. CoRR abs/1807.03748 (2018). <http://arxiv.org/abs/1807.03748>
15. Ouyang, C., et al.: Causality-inspired single-source domain generalization for medical image segmentation. CoRR abs/2111.12525 (2021). <https://arxiv.org/abs/2111.12525>
16. Park, T., Efros, A.A., Zhang, R., Zhu, J.-Y.: Contrastive learning for unpaired image-to-image translation. In: Vedaldi, A., Bischof, H., Brox, T., Frahm, J.-M. (eds.) ECCV 2020. LNCS, vol. 12354, pp. 319–345. Springer, Cham (2020). [https://doi.org/10.1007/978-3-030-58545-7\\_19](https://doi.org/10.1007/978-3-030-58545-7_19)
17. Qiao, F., Zhao, L., Peng, X.: Learning to learn single domain generalization. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 12556–12565 (2020)
18. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) MICCAI 2015. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). [https://doi.org/10.1007/978-3-319-24574-4\\_28](https://doi.org/10.1007/978-3-319-24574-4_28)

19. Tan, M., Le, Q.: EfficientNet: rethinking model scaling for convolutional neural networks. In: Chaudhuri, K., Salakhutdinov, R. (eds.) Proceedings of the 36th International Conference on Machine Learning. Proceedings of Machine Learning Research, vol. 97, pp. 6105–6114. PMLR, 09–15 June 2019. <https://proceedings.mlr.press/v97/tan19a.html>
20. Tang, Z., Peng, X., Li, T., Zhu, Y., Metaxas, D.: Adatransform: adaptive data transformation. In: 2019 IEEE/CVF International Conference on Computer Vision (ICCV), pp. 2998–3006 (2019). <https://doi.org/10.1109/ICCV.2019.00309>
21. Wang, Z., Luo, Y., Qiu, R., Huang, Z., Baktashmotlagh, M.: Learning to diversify for single domain generalization. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), pp. 834–843, October 2021
22. Xu, Z., Liu, D., Yang, J., Raffel, C., Niethammer, M.: Robust and generalizable visual representation learning via random convolutions. arXiv preprint [arXiv:2007.13003](https://arxiv.org/abs/2007.13003) (2020)
23. Zhang, H., Cisse, M., Dauphin, Y.N., Lopez-Paz, D.: mixup: Beyond empirical risk minimization. In: International Conference on Learning Representations (2018). <https://openreview.net/forum?id=r1Ddp1-Rb>
24. Zhu, Q., Du, B., Yan, P.: Boundary-weighted domain adaptive neural network for prostate MR image segmentation. CoRR abs/1902.08128 (2019). <http://arxiv.org/abs/1902.08128>