



# ISDNet: Importance Guided Semi-supervised Adversarial Learning for Medical Image Segmentation

Qingtian Ning<sup>1</sup>, Xu Zhao<sup>1</sup>(✉), and Dahong Qian<sup>2</sup>

<sup>1</sup> Department of Automation, Shanghai Jiao Tong University, Shanghai, China  
zhaoxu@sjtu.edu.cn

<sup>2</sup> Institute of Medical Robotics, Shanghai Jiao Tong University, Shanghai, China

**Abstract.** Recent deep neural networks have achieved great success in medical image segmentation. However, massive labeled training data should be provided during network training, which is time consuming with intensive labor work and even requires expertise knowledge. To address such challenge, inspired by typical GANs, we propose a novel end-to-end semi-supervised adversarial learning framework for medical image segmentation, called “Importance guided Semi-supervised Deep Networks” (ISDNet). While most existing works based on GANs use a classifier discriminator to achieve adversarial learning, we combine a fully convolutional discriminator and a classifier discriminator to fulfill better adversarial learning and self-taught learning. Specifically, we propose an importance weight network combined with our FCN-based confidence network, which can assist segmentation network to learn better local and global information. Extensive experiments are conducted on the LASC 2013 and the LiTS 2017 datasets to demonstrate the effectiveness of our approach.

**Keywords:** Medical image segmentation · Semi-supervised · GAN

## 1 Introduction

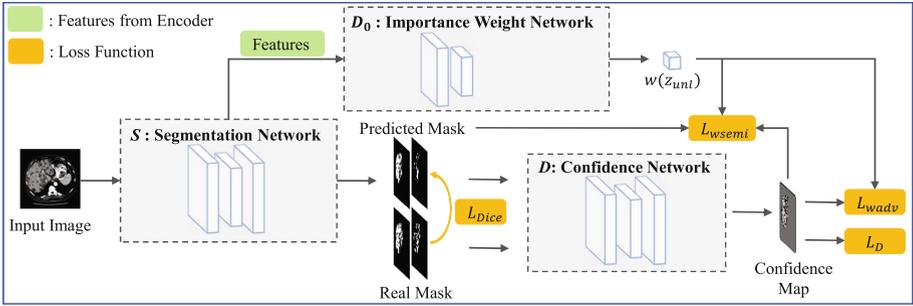
Medical image segmentation is a fundamental yet challenging problem in medical image analysis, which aims to segment organs or pathological area (e.g. left atrium or liver lesion) in medical images. Recently, a typical Fully Convolutional Network (FCN) [20] based methods of encoder-decoder structure have achieved considerable success in semantic image segmentation, such as U-Net [21] and deeplabv3+ [5]. Although CNN-based approaches have made great progress, they often require massive labeled training data. Medical image tasks are usually more difficult than natural image tasks because of little training data and high cost manual annotations, etc. In addition, most medical datasets are still constrained in limited size and application scenarios. Hence, it is a complex and expensive procedure to obtain large-scale medical labeled data.

In order to ease the effort of acquiring high-quality data, unsupervised learning is an alternative way to utilize massive unlabeled data. However, due to lacking the concept of classes, unsupervised learning methods have not attained convincing performance in segmentation. So semi/weakly-supervised methods have drawn much attention from many researchers in the community. These methods generally demand additional annotations or data, such as the image-level class label [18], box level [16, 17], point level [19], or scribble level [1].

Recently, Generative Adversarial Networks (GANs) [22] have achieved a wide range of success due to their ability to generate high-quality realistic images [2, 6, 11, 12]. It attracts extensive attention with successful applications in super-resolution [14], domain adaptation [8, 15], zero-shot learning [13], etc. A classic GAN consists of two sub-networks (i.e., generator and discriminator) that play a min-max game in the training procedure where generator produces a sample of the target data distribution, while discriminator aims to differentiate between real and fake data repeatedly. The generator is then optimized to generate more realistic samples that are more close to target data distribution. Recently, several works have applied the GANs framework in semantic segmentation. Luc *et al.* [23] propose to apply a classifier discriminator to assist the training process for semantic segmentation in a fully-supervised way. But this method has not achieved distinguished results over the baseline scheme and fails to tackle unlabeled data for semi-supervised setting. Frid-Adar *et al.* [9] employ GANs to generate synthetic medical images to train segmentation network. Nie *et al.* [7] use sample attention mechanism to improve the network training. However, those jobs utilize the classification networks as their discriminators in fully-supervised ways, so those discriminators only can capture the global information of generated masks and ground truth masks without considering the local details, which is more important for the task of semantic segmentation.

In this paper, we propose a semi-supervised semantic segmentation algorithm (called ISDNet) based on GANs in order to alleviate the demand for large-scale medical labeled data. Inspired by [10] and [8], our network consists of three sub-networks: (1) segmentation network, (2) confidence network and (3) importance weight network. In this work, two semi-supervised loss are conducted to leverage the unlabeled data. First, Our FCN-based confidence network can generate confidence maps, which can guide our segmentation network in a self training strategy. The confidence maps provide us the trustworthy regions in the segmented label map, which can be selected to generate proxy label for unlabeled data. Second, we apply the adversarial loss on unlabeled data in the supervised setting, which encourages the model to predict segmentation outputs of unlabeled data close to the ground truth distributions. Then, our importance weight network can identify the importance score of unlabeled samples, which represents the probability of the sample come from the labeled data distribution. Finally, we integrate the importance score into semi-supervised loss and obtain our two weighted semi-supervised loss.

In sum, our main contributions include: (i) we develop an adversarial framework, which improves semantic segmentation performance without adding any



**Fig. 1.** Illustration of the architecture of our proposed ISDNet, which consists of a segmentation network, a confidence network and an importance weight network.

computational cost during inference; (ii) we propose a semi-supervised framework to improve the segmentation accuracy with unlabeled data; (iii) we combine a fully convolutional discriminator and a classifier discriminator to facilitate the semi-supervised learning, which can better use unlabeled data. To demonstrate the effectiveness of our proposed approach, ablation studies are conducted on the LASC 2013 dataset. Overall, our proposed approach brings 1.8% improvement on the LASC 2013 dataset and 11.4% improvement on the LiTS 2017 dataset.

## 2 Our Approach

The proposed ISDNet consists of three subnetworks, i.e., (1) segmentation network (denoted as  $S$ ), (2) confidence network (denoted as  $D$ ) and (3) importance weight network (denoted as  $D_0$ ). The architecture of our proposed framework is presented in Fig. 1.

In order to facilitate elaboration, we first give the definitions of terminologies used throughout the paper. Given an input medical image  $X_n$  of dimension  $H \times W \times 1$  and its one-hot encoded ground truth label  $Y_n$ , we denote the segmentation network as  $S(\cdot)$ , the features as  $Z(X_n)$  extracted from the encoder of our segmentation network and the predicted probability map as  $\hat{Y}_n = S(X_n)$  of size  $H \times W \times C$ , where  $C$  is the category number. We denote our fully convolutional discriminator as  $D(\cdot)$  which takes class probability maps as the input (the output from segmentation network  $S(X_n)$  or ground truth label maps  $Y_n$ ) and then outputs a confidence map of size  $H \times W \times 1$ . Confidence map generated by  $D$  scores each pixel  $p$ , which represents whether that pixel is sampled from the ground truth label ( $p = 1$ ) or the segmentation network ( $p = 0$ ). The importance weight network denoted as  $D_0(\cdot)$ , takes  $Z(X_n)$  (the features of labeled sample or unlabeled sample) as the input and outputs an importance weight. The importance weight represents whether that sample is sampled from the labeled data ( $X_n = 1$ ) or the unlabeled data ( $X_n = 0$ ).

## 2.1 Segmentation Network

In ISDNet as shown in Fig. 1, we employ a simplified 2D U-Net (SU-Net) as our segmentation network, but any end-to-end segmentation network can be applied, such as FCN and deeplabv3+, etc. In this paper, we halve the number of all convolutional layers in original U-Net [21], and use leaky ReLU and group normalization [3] instead of ReLU and batch normalization to balance the performance and memory cost.

**Total Loss for Segmentation Network.** We train the segmentation network by minimizing a multi-task loss function

$$L_{seg} = L_{Dice} + \lambda_{adv}L_{wadv} + \lambda_{semi}L_{wsemi}, \quad (1)$$

where  $L_{Dice}$ ,  $L_{wadv}$  and  $L_{wsemi}$  denote the multi-class dice loss, adversarial loss and semi-supervised loss, respectively. In (1),  $\lambda_{adv}$  and  $\lambda_{semi}$  are two weights for minimizing the proposed multi-task loss function.

**Multi-class Dice Loss.** To overcome the class imbalance problem, we propose to use a weighted multi-class dice loss as the segmentation loss

$$L_{Dice} = 1 - 2 \sum_{c=1}^C \frac{w^c \hat{Y}_n^c Y_n^c}{w^c (\hat{Y}_n^c + Y_n^c)}, \quad (2)$$

where  $\hat{Y}_n^c$  denotes the predicted probability belonging to class  $c$  (i.e. background, liver, or liver lesion),  $Y_n^c$  denotes the ground truth probability, and  $w^c$  denotes a class dependent weighting factor. Empirically, we set the weights to be 0.2 for background, 1.2 for liver, and 2.2 for liver lesion. But for left atrium (LA) segmentation, we use normal two-class dice loss.

## 2.2 Confidence Network

Different from using CNN-based discriminator, we propose to use a FCN-based discriminator called confidence network to generate more detailed adversarial information in local region. Hence, we combine adversarial learning in our work to further optimize the segmentation network.

**Adversarial Loss of the Confidence Network.** To train the confidence network, we minimize the binary cross-entropy loss  $L_D$  using

$$L_D = L_{bce}(D(Y_n)^{h,w}, 1) + L_{bce}(D(S(X_n))^{h,w}, 0), \quad (3)$$

where  $X_n$  and  $Y_n$  represent the input data and its corresponding ground truth label map (one-hot encoding scheme), respectively. In addition,  $D(S(X_n))^{(h,w)}$  is the confidence map of  $X_n$  at location  $(h, w)$ , and  $D(Y_n)^{(h,w)}$  is defined similarly.

**Adversarial Loss of the Segmentation Network.** In the conventional GANs, generators and discriminators play a min-max game. Hence, there is

another loss from  $D$  working as adversarial Loss to further improve segmentation network. It enforces higher-order consistency between ground truth label and predicted masks. We utilize the loss  $L_{adv}$  based on a fully convolutional discriminator network

$$L_{adv} = L_{bce}(D(S(X_n))^{h,w}, 1). \quad (4)$$

With this loss, we further improve the ability of segmentation network to fool the discriminator by maximizing the probability of the predicted results being more close to the ground truth distribution.

### 2.3 Importance Weight Network

**Training with Unlabeled Data.** In this work, we concern more about how to use unlabeled data to improve the performance of segmentation network. Since there is no corresponding ground truth label for unlabeled data,  $L_{Dice}$  no longer works. But the adversarial loss  $L_{adv}$  is still applicable as it only requires the discriminator network.

In addition, since our confidence network could provide local confidence information, we propose a self-taught learning framework with our trained discriminator  $D$  for unlabeled data. The main idea is that our confidence network can generate a confidence map  $CM = D(S(X_n))^{h,w}$  to indicate us which regions of the predicted results are sufficiently close to the ground truth distribution. Then we process the confidence map with a threshold  $T_{semi}$  to obtain the confident regions. In this way, we can exploit these confident regions to filter the segmentation results of unlabeled data to improve the segmentation network. The semi-supervised loss is defined as

$$L_{semi} = 1 - 2 \sum_{c=1}^C \frac{w^c [CM > T_{semi}] \hat{Y}_n^c \bar{Y}_n^c}{w^c [CM > T_{semi}] (\hat{Y}_n^c + \bar{Y}_n^c)}, \quad (5)$$

where  $\bar{Y}_n$  is the one-hot encoding of  $\arg \max(\hat{Y}_n)$ .  $[\cdot]$  is the indicator function.

**Sample Weights Learning.** Furthermore, we employ another discriminator named  $D_0$ . It can output the probability that an unlabeled sample belongs to labeled data distribution on the image level. Specifically, the importance weight network is similar to the original GANs with min-max loss

$$L_{adv} = L_{bce}(D_0(Z(X_l)), 1) + L_{bce}(D_0(Z(X_{unl})), 0), \quad (6)$$

where  $Z(\cdot)$  is the feature extractor (also is the encoder of our SU-Net) for labeled data  $X_l$  and unlabeled data  $X_{unl}$  respectively, and  $D_0$  is a binary classifier with all the labeled data labeled as 1 and all the unlabeled data labeled as 0.

Assume that in the case of the optimal classifier  $D_0$ , the output value of  $D_0$  is the probability that the sample comes from the labeled data distribution. If  $D^*(z) \approx 1$ , then the sample will highly likely originate from the labeled data distribution, since the features  $Z(X_l)$  are quite different from the features  $Z(X_{unl})$

and can be ideally separated from labeled data distribution by  $D_0$ . Then we should reduce the contribution of these samples because feature extractor has already been trained by some similar samples. On the other hand, if  $D^*(z)$  is small, it means feature extractor has not been trained by those samples. A larger importance weight should be applied to these samples to improve segmentation network. Hence, the sample weights function should be inversely related to  $D^*(z)$  and the importance weight function of the unlabeled samples is defined as

$$w(z) = 1 - D^*(z). \quad (7)$$

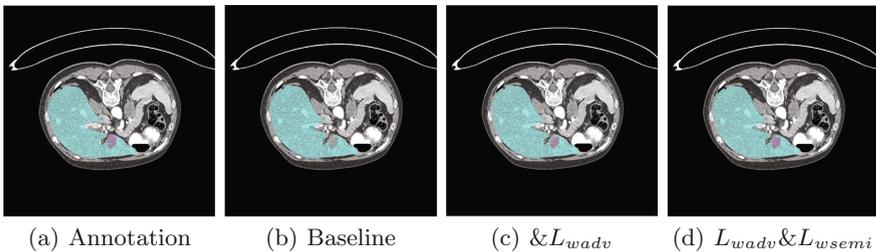
As can be seen that if  $D^*(z)$  is small,  $w(z)$  is large. Hence, the weights for unlabeled samples that are similar to the labeled data will be smaller than those are not similar to the labeled data. Our aim is to obtain the relative importance of unlabeled samples. The weight function can also be expressed as a function of density ratio between labeled and unlabeled features. If we apply the weights to  $D_0$ , then the Jensen-Shannon divergence between two densities can not be reduced from the theoretical results of the minimax game [8]. Hence, we utilize the weights to  $D$  to solve this issue. In this way,  $D_0$  is only used for obtaining the importance weights for unlabeled samples. Thus, we will not update the encoder with the gradient of  $D_0$ . So we can integrate the importance weight into our semi-supervised loss.

$$L_{wsemi} = w(z)L_{semi}. \quad (8)$$

And adversarial loss for unlabeled data can be modified as

$$L_{wadv} = w(z)L_{adv}. \quad (9)$$

But for labeled data,  $w(z)$  should be set to 1. By summing the above losses, the total loss to train the segmentation network can be defined as (1).



**Fig. 2.** Comparisons on the LiTS 2017 dataset using 1/2 labeled data. Green area and red area represent liver and lesion, respectively. It can be seen that our confidence network assists segmentation network discover parts of lesions while our baseline does not. Furthermore, our semi-supervised algorithm refines the segmentation results.

### 3 Experiments

**Evaluation Datasets and Metrics.** Experiments are conducted on two publicly available datasets to report the state-of-the-art performance of ISDNet. Ablation studies are conducted on LASC 2013 dataset [4]. We also conduct experiments on LiTS 2017 dataset to verify the validity of our method. For LASC 2013, we employ data augmentation and obtain 3K 2D images with size  $320 \times 320$  in total. Dice index for left atrium (LA) and running time are selected to compare with other state-of-the-art method. For LiTS 2017, we use two metrics to evaluate the segmented liver lesions, including dice per case and global dice.

**Implementation Details.** All experiments are built with Pytorch framework on a single NVIDIA 1080ti GPU. We use the Adam optimizer for both our segmentation network and two discriminators with the learning rate  $10^{-4}$ . For the hyper-parameters in the proposed method,  $\lambda_{adv}$  is set as 0.001 when trained with labeled and unlabeled data. We set  $\lambda_{semi}$  as 1 and  $T_{semi}$  as 0.1. For semi-supervised training, we randomly divide all dataset into labeled and unlabeled data. We initiate the semi-supervised learning after 5 epochs training with labeled data. In each epoch, we train both the segmentation network and two discriminator networks, only labeled data are used for training of the discriminator  $D$  while  $D_0$  demands the part of unlabeled data.

**Table 1.** Results on the LASC test set. We utilize 1/2 images as labeled data, the rest as unlabeled data for semi-supervised learning. *Running time* indicates mean inference time on CPU.

Methods	Data amount		
	1/2	Full	Running time (sec)
LTSI_VRG [24]	N/A	91.0	3100
UCL_1C [25]	N/A	93.8	1200
UCL_4C [25]	N/A	85.9	1200
OBS_2 [4]	N/A	90.8	N/A
Hung et al. [10]	91.0	N/A	1.1
Luc et al. [23]	90.0	93.6	1.1
Liu et al. [26]	N/A	94.0	23.7
AF-CNN-SP [27]	N/A	<b>95.1</b>	450
U-Net [21]	87.4	90.8	1.4
Baseline (SU-Net)	89.5	93.2	1.1
Baseline+ $L_{wadv}$	90.5	<b>94.0</b>	1.1
Baseline+ $L_{wadv}+L_{wsemi}$	<b>91.3</b>	N/A	<b>1.1</b>

### 3.1 LASC 2013

**Results and Analysis.** Table 1 shows the evaluation results on the LASC 2013 test dataset. We randomly sample 1/2 images as labeled data, the rest as unlabeled data. We compare the proposed algorithm against LTSI-VRG [24], UCL-1C [25], UCL-4C [25], U-Net [21], [26] and [27] to demonstrate that our baseline model (SU-Net) performs comparably with the state-of-the-art schemes. LTSI-VRG, UCL-1C, UCL-4C are the top three methods in Dice index, which are all based on multi-atlas. OBS-2 [4] is the result from human observer. [26] combines CNN and recurrent neural network (RNN) to achieve a Dice index of 0.94, but it will greatly increase the inference time (22 times slower than our SU-Net). [27] is based on a multi-view CNN with an adaptive fusion strategy and a new loss function, which is 400 times slower than ours and use 20 times more data than us. Table 1 shows that the adversarial loss brings consistent performance improvement (from 0.8% to 1.0%) over different amounts of training data. Incorporating the proposed semi-supervised learning scheme brings overall 1.8% improvement. For [10] and [23], we use our SU-Net to replace its original segmentation network for equal comparisons. Specially, our importance weight network brings 0.3% improvement compared to [10]. It means our network makes better use of unlabeled data to improve network performance. Apart from this, our confidence network brings improvement (from 0.4% to 0.5%) compared to a typical classifier discriminator [23].

**Table 2.** Hyper parameter analysis.

Data amount	$\lambda_{adv}$	$\lambda_{semi}$	$T_{semi}$	Dice
1/2	0.0001	0	N/A	89.3
1/2	0.001	0	N/A	90.5
1/2	0.005	0	N/A	88.7
1/2	0.001	5	0.1	89.2
1/2	0.001	1	0.1	91.3
1/2	0.001	0.5	0.1	90.6
1/2	0.001	1	0	90.1
1/2	0.001	1	0.1	91.3
1/2	0.001	1	0.5	90.9
1/2	0.001	1	1	90.3

**Hyper-parameter Analysis.** The proposed algorithm is governed by three hyper parameters:  $\lambda_{adv}$  and  $\lambda_{semi}$  for balancing the multi-task learning in (1), and  $T_{semi}$  used to control the sensitivity in the semi-supervised learning described in (5). Table 2 shows sensitivity analysis of hyper parameters on the LASC 2013 dataset under semi-supervised setting. Different from [10], we find that smaller  $\lambda_{adv}$  must be used for medical image tasks, this is because the content of natural images is richer and requires larger loss to guide network learning.

Then, we conduct the experiments with different values of  $T_{semi}$ . With higher  $T_{semi}$ , our algorithm will select regions, which are more close to the ground truth distribution. When  $T_{semi} = 0$ , all the pixel predictions in unlabeled images will be applied for semi-supervised training, which leads to performance degradation. Overall, the proposed model achieves the best results when  $T_{semi} = 0.1$ .

**Ablation Study.** We present ablation study of our proposed system in Table 3 on LASC test dataset. Our confidence network gains 0.5% and 0.4% improvement compared to a classifier discriminator with half and full data, respectively. Then, we apply the semi-supervised learning method without the adversarial loss. The results show that the adversarial procedure on the labeled data is necessary to our semi-supervised scheme. If the segmentation network does not participate in adversarial training, the confidence maps generated by the discriminator would be pointless. As shown in Table 3, our semi-supervised methods in ISDNet help to improve segmentation performance.

**Table 3.** Ablation study of the proposed method on the LASC dataset.  $\checkmark$  denotes the setting of corresponding column is employed. *CN* denotes confidence map.

$L_{wadv}$	$L_{wsemi}$	$L_{adv}&L_{semi}$	CN	Data amount	
				1/2	Full
				89.5	93.2
$\checkmark$			$\checkmark$	90.5	94.0
$\checkmark$				90.0	93.6
	$\checkmark$		$\checkmark$	89.1	N/A
$\checkmark$	$\checkmark$		$\checkmark$	91.3	N/A
		$\checkmark$	$\checkmark$	91.0	N/A

### 3.2 LiTS 2017

**Results and Visualization.** Furthermore, we extend the experiment on LiTS 2017 dataset. Figure 2 shows visual comparisons of the segmentation results on the LiTS 2017 validation dataset generated by our proposed method. It can be seen that no lesion is found in our baseline (Fig. 2(b)), but with the assistance of adversarial loss, segmentation network can detect parts of lesions (Fig. 2(c)). Further more, with our semi-supervised adversarial learning algorithm, segmentation network could segment a majority of lesions (Fig. 2(d)). Table 4 shows the liver lesion evaluation results on the LiTS 2017 test dataset with random sampled 1/2 images as labeled data. It can be seen that our methods have made the best dice per case of 50.6% with 10.5% gain and the best global dice of 72.5% with 11.4% gain.

**Table 4.** Results on the LiTS test set. *DPC* indicates dice per case. *GD* denotes global dice.

Methods	Data Amount	DPC	GD
Our baseline	Full	54.0	76.6
Baseline+ $L_{wadv}$	Full	56.8	76.9
Our baseline	1/2	40.1	61.1
Baseline+ $L_{wadv}$	1/2	45.9	64.0
Baseline+ $L_{wadv}$ + $L_{wsemi}$	1/2	50.6	72.5

## 4 Conclusion

In this work, we have presented a novel importance guided semi-supervised adversarial learning scheme (ISDNet) for medical image segmentation. Specifically, we train two discriminators to enhance the segmentation network with both labeled and unlabeled data to effectively address the insufficient labeled data problem. We combine FCN-based discriminator with CNN-based discriminator for our semi-supervised learning strategy. It can be seen that by integrating these components into our framework, the ISDNet has achieved significant improvement in terms of both accuracy and robustness.

**Acknowledgement.** This work is supported by: National Natural Science Foundation of China (61673269, 61273285).

## References

1. Lin, D., Dai, J., Jia, J., et al.: ScribbleSup: scribble-supervised convolutional networks for semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3159–3167 (2016)
2. Brock, A., Donahue, J., Simonyan, K.: Large scale GAN training for high fidelity natural image synthesis. arXiv preprint [arXiv:1809.11096](https://arxiv.org/abs/1809.11096) (2018)
3. Wu, Y., He, K.: Group normalization. In: Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y. (eds.) ECCV 2018. LNCS, vol. 11217, pp. 3–19. Springer, Cham (2018). [https://doi.org/10.1007/978-3-030-01261-8\\_1](https://doi.org/10.1007/978-3-030-01261-8_1)
4. Tobon-Gomez, C., Geers, A.J., Peters, J., et al.: Benchmark for algorithms segmenting the left atrium from 3D CT and MRI datasets. IEEE Trans. Med. Imaging **34**(7), 1460–1473 (2015)
5. Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F., Adam, H.: Encoder-decoder with atrous separable convolution for semantic image segmentation. In: Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y. (eds.) ECCV 2018. LNCS, vol. 11211, pp. 833–851. Springer, Cham (2018). [https://doi.org/10.1007/978-3-030-01234-2\\_49](https://doi.org/10.1007/978-3-030-01234-2_49)
6. Radford, A., Metz, L., Chintala, S.: Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv preprint [arXiv:1511.06434](https://arxiv.org/abs/1511.06434) (2015)

7. Nie, D., Gao, Y., Wang, L., Shen, D.: ASDNet: attention based semi-supervised deep networks for medical image segmentation. In: Frangi, A.F., Schnabel, J.A., Davatzikos, C., Alberola-López, C., Fichtinger, G. (eds.) MICCAI 2018. LNCS, vol. 11073, pp. 370–378. Springer, Cham (2018). [https://doi.org/10.1007/978-3-030-00937-3\\_43](https://doi.org/10.1007/978-3-030-00937-3_43)
8. Zhang, J., Ding, Z., Li, W., et al.: Importance weighted adversarial nets for partial domain adaptation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 8156–8164 (2018)
9. Frid-Adar, M., Diamant, I., Klang, E., et al.: GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification. *Neuro-computing* **321**, 321–331 (2018)
10. Hung, W.C., Tsai, Y.H., Liou, Y.T., et al.: Adversarial learning for semi-supervised semantic segmentation. arXiv preprint [arXiv:1802.07934](https://arxiv.org/abs/1802.07934) (2018)
11. Arjovsky, M., Chintala, S., Bottou, L.: Wasserstein GAN. arXiv preprint [arXiv:1701.07875](https://arxiv.org/abs/1701.07875) (2017)
12. Gulrajani, I., Ahmed, F., Arjovsky, M., et al.: Improved training of Wasserstein GANs. In: Advances in Neural Information Processing Systems, pp. 5767–5777 (2017)
13. Zhu, Y., Elhoseiny, M., Liu, B., et al.: A generative adversarial approach for zero-shot learning from noisy texts. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1004–1013 (2018)
14. Lai, W.S., Huang, J.B., Ahuja, N., et al.: Deep Laplacian pyramid networks for fast and accurate super-resolution. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 624–632 (2017)
15. Hu, L., Kan, M., Shan, S., et al.: Duplex generative adversarial network for unsupervised domain adaptation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1498–1507 (2018)
16. Dai, J., He, K., Sun, J.: BoxSup: exploiting bounding boxes to supervise convolutional networks for semantic segmentation. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 1635–1643 (2015)
17. Oh, S.J., Benenson, R., Khoreva, A., et al.: Exploiting saliency for object segmentation from image level labels. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 5038–5047. IEEE (2017)
18. Pathak, D., Krahenbuhl, P., Darrell, T.: Constrained convolutional neural networks for weakly supervised segmentation. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 1796–1804 (2015)
19. Bearman, A., Russakovsky, O., Ferrari, V., Fei-Fei, L.: What’s the point: semantic segmentation with point supervision. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.) ECCV 2016. LNCS, vol. 9911, pp. 549–565. Springer, Cham (2016). [https://doi.org/10.1007/978-3-319-46478-7\\_34](https://doi.org/10.1007/978-3-319-46478-7_34)
20. Long, J., Shelhamer, E., Darrell, T.: Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3431–3440 (2015)
21. 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)
22. Goodfellow, I., Pouget-Abadie, J., Mirza, M., et al.: Generative adversarial nets. In: Advances in Neural Information Processing Systems, pp. 2672–2680 (2014)
23. Luc, P., Couprie, C., Chintala, S., et al.: Semantic segmentation using adversarial networks. arXiv preprint [arXiv:1611.08408](https://arxiv.org/abs/1611.08408) (2016)

24. Sandoval, Z., Betancur, J., Dillenseger, J.-L.: Multi-atlas-based segmentation of the left atrium and pulmonary veins. In: Camara, O., Mansi, T., Pop, M., Rhode, K., Sermesant, M., Young, A. (eds.) STACOM 2013. LNCS, vol. 8330, pp. 24–30. Springer, Heidelberg (2014). [https://doi.org/10.1007/978-3-642-54268-8\\_3](https://doi.org/10.1007/978-3-642-54268-8_3)
25. Zuluaga, M.A., Cardoso, M.J., Modat, M., Ourselin, S.: Multi-atlas propagation whole heart segmentation from MRI and CTA using a local normalised correlation coefficient criterion. In: Ourselin, S., Rueckert, D., Smith, N. (eds.) FIMH 2013. LNCS, vol. 7945, pp. 174–181. Springer, Heidelberg (2013). [https://doi.org/10.1007/978-3-642-38899-6\\_21](https://doi.org/10.1007/978-3-642-38899-6_21)
26. Liu, X., Shen, Y., Zhang, S., et al.: Segmentation of left atrium through combination of deep convolutional and recurrent neural networks. *J. Med. Imaging Health Inform.* **8**(8), 1578–1584 (2018)
27. Mortazi, A., Karim, R., Rhode, K., Burt, J., Bagci, U.: *CardiacNET*: segmentation of left atrium and proximal pulmonary veins from MRI using multi-view CNN. In: Descoteaux, M., Maier-Hein, L., Franz, A., Jannin, P., Collins, D.L., Duchesne, S. (eds.) MICCAI 2017. LNCS, vol. 10434, pp. 377–385. Springer, Cham (2017). [https://doi.org/10.1007/978-3-319-66185-8\\_43](https://doi.org/10.1007/978-3-319-66185-8_43)