Generative Adversarial Networks for Pre-training of Medical Image Segmentation Networks

Kun Chen (kchen17@fudan.edu.cn)
Fudan University School of Basic Medical Sciences https://orcid.org/0000-0001-9172-4562
Manning Wang
Fudan University School of Basic Medical Sciences
Zhijian Song
Fudan University School of Basic Medical Sciences

Research

Keywords: Generative Adversarial Networks, Segmentation networks, Medical image 29 segmentation, Preliminary training

Posted Date: January 30th, 2020

DOI: https://doi.org/10.21203/rs.2.22269/v1

Generative Adversarial Networks for Pre-training of Medical Image Segmentation Networks

Kun Chen[a,b], Manning Wang[a,b,*], Zhijian Song[a,b,*]
[a] Digital Medical Research Center, School of Basic Medical Sciences, Fudan University, Shanghai, CO 200032, China
[b] Shanghai Key Laboratory of Medical Imaging Computing and Computer Assisted Intervention, Fudan University, Shanghai, CO 200032, China

Abstract.

Background: Deep neural networks have been widely used in medical image segmentation and have achieved state-of-the-art performance in many tasks. However, different from the segmentation of natural images or video frames, the manual segmentation of anatomical structures in medical images needs high expertise so the scale of labeled training data is very small, which is a major obstacle for the improvement of deep neural networks performance in medical image segmentation.

Methods: In this paper, we proposed a new end-to-end generation-segmentation framework by integrating Generative Adversarial Networks (GAN) and a segmentation network and train them simultaneously. The novelty is that during the training of the GAN, the intermediate synthetic images generated by the generator of the GAN are used to pre-train the segmentation network. As the advances of the training of the GAN, the synthetic images evolve gradually from being very coarse to containing more realistic textures, and these images help train the segmentation network gradually. After the training of GAN, the segmentation network is then fine-tuned by training with the real labeled images.

Results: We evaluated the proposed framework on four different datasets, including 2D cardiac dataset and lung dataset, 3D prostate dataset and liver dataset. Compared with original U-net and CE-Net, our framework can achieve better segmentation performance. Our framework also can get better

* Corresponding author:
E-mail address: mnwang@fudan.edu.cn (Manning Wang)
zjsong@fudan.edu.cn (Zhijian Song)

segmentation results than U-net on small datasets. In addition, our framework is more effective than the usual data augmentation methods.

Conclusions: The proposed framework can be used as a pre-train method of segmentation network, which helps to get a better segmentation result. Our method can solve the shortcomings of current data augmentation methods to some extent.

Keywords: Generative Adversarial Networks, Segmentation networks, Medical image segmentation, Preliminary training

Background

Deep learning algorithm represented by convolutional neural network[1] has achieved great success in the field of medical image analysis, including medical image segmentation[2]. The performance of deep learning algorithms largely depends on the amount of labeled training data, and a training set with millions of annotated images is very common in natural image or video segmentation tasks, such as ImageNet Large Scale Visual Recognition Competition (ILSVRC)[3]. However, in the medical image segmentation task, accurate manual segmentation of anatomical structures needs high expertise, and it is very expensive and time-consuming to obtain training data with high quality manual segmentation. As a result, the training dataset of medical image segmentation tasks are very small, which usually consists of hundreds of images, and this is a major obstacle to improve the performance of deep learning algorithms in medical image segmentation.

The lack of high-quality annotated data will lead to over-fitting and insufficient generalization ability of deep learning models. At present, three kinds of directions are often

taken to solve this problem. One is to design new network structures to achieve better performance. Ronneberger et al.[4] proposed the U-net, which aggregated lower layer feature maps to the higher layer and achieved good results in biomedical image segmentation. It has become a backbone for most medical image segmentation tasks. Nie et al.[5] proposed a cascade framework, which first used a 3D U-net for coarse segmentation, and then a cascade of Convolutional Neural Networks (CNNs) for fine-grained segmentation. Wang et al.[6] presented a CNN-based interactive segmentation method (DeepIGeoS) with that only requires few interactions to obtain accurate segmentation results with high efficiency. Wang et al.[7] proposed a two-stage segmentation framework including locating a bounding box of the organ at risk (OAR) and segmenting the OAR from a small volume within the bounding box. Moreover, many methods adopted the image patches as the input of their framework[8, 9], which tend to lose some global image information. Many variants of CNN have been proposed, but these methods cannot well solve the problem of over-fitting caused by small scale of training data.

The second direction is carrying out transfer learning[10] on the segmentation network[11]. Transfer learning focuses on storing knowledge gained while solving one problem and applying it to a different but related problem, which allows two different domains. Currently, transfer learning can help initialize the segmentation network with the weight obtained from the pre-training with ImageNet dataset[12] or other medical image domain[13]. However, the current research [14, 15] pointed out that transfer learning has little impact on the performance improvement of the segmentation network.

The third direction is to expand the quantity and diversity of training data by data augmentation. At present, the commonly used data augmentation methods in medical images are based on affine transformation [16] and statistical model, such as scaling, rotation, flipping, translation, and elastic deformation, Active Shape Model (ASM)[17], and Active Appearance Model (AAM)[18]. Tang et al.[19] proposed an augmentation strategy based on statistical shape model (SSM) and 3D thin plate spline (3D TPS), which modeled shape information with SSM and filled the texture of the simulated shapes with 3D TPS. However, these transformations do not account for changes due to different imaging sequences, let alone changes in the size, shape, location, and appearance of specific lesions. And these methods will result in strong similarity between the texture of the synthetic images and the real images and limit their capability of improving the generalization performance of the segmentation model. GAN [20] has made a breakthrough in cross-modal image synthesis and is also used in synthesizing medical images. Zhao et al. [21] synthesized Magnetic resonance (MR) images into high-quality Computed Tomography (CT) images using GAN, and then used the synthetic CT images and MR images to segment bone structures. Costa et al. [22, 23] synthesized a complete retinal image with a given vascular tree by using GAN. Mok and Chung [24] used Conditional Generative Adversarial Nets (cGAN) [25] to augment training images for brain tumor segmentation. The generator of cGAN was conditioned on a segmentation map and generated brain MR images in a coarse to fine manner. However, GAN is not widely used in the image synthesis for segmentation task. One reason is that the training of GAN is even more difficult than training other neural networks, and it is still very difficult to train a GAN that is able to generate realistic and diverse medical images. Therefore,

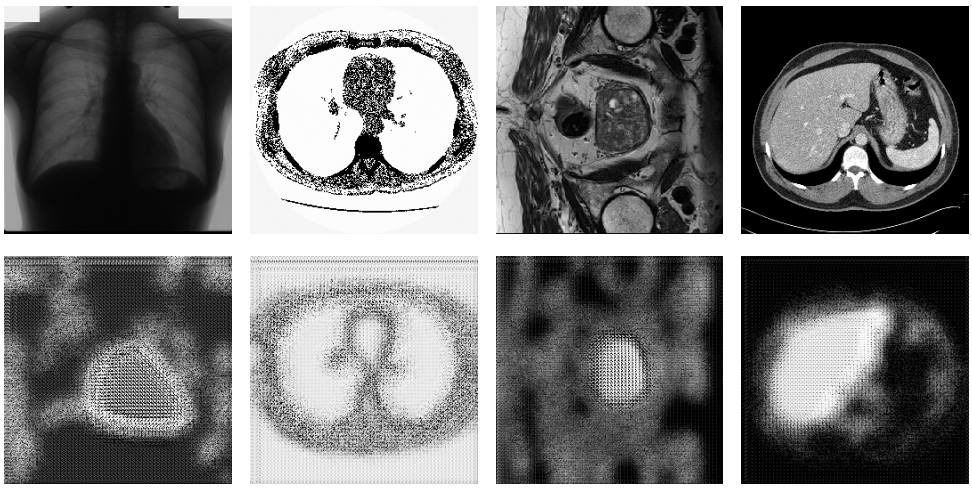
training with synthetic images and real images can`t improve the performance of segmentation network.

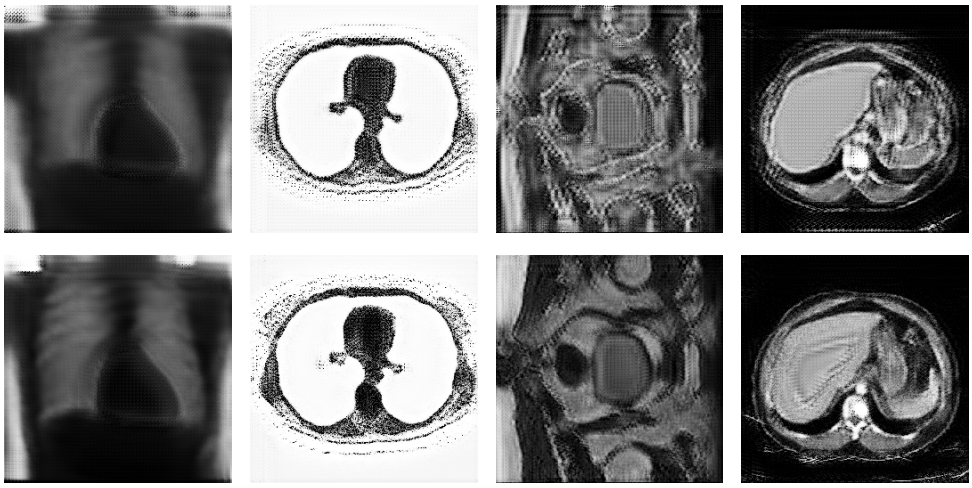
In this paper, we propose a new generation-segmentation framework by integrating a GAN and a segmentation network to improve segmentation accuracy with small training datasets. In the proposed network, the GAN and the segmentation network are first jointly trained, where the intermediate synthetic images during the training of GAN are used for the training of the segmentation network. Then the segmentation network is fine-tuned by using real training images. In the experiments, we test our framework on three public medical datasets and the results show that joint pre-training with GAN significantly improve the performance of the segmentation network.

Results

Synthesis results

With the continuous iteration of the proposed framework, the image generated by generator of GAN becomes increasingly fine. Fig. 1 shows the variation of generated image in four different datasets with respect to the number of iterations.





Cardiac LUNA Prostate Liver

Fig. 1 Generated images by the generator of GAN in different iterations for different datasets. The first row represents the original images, and the 2nd, 3rd and fourth rows represent the generated images after 1, 50, and 100 epochs, respectively.

Segmentation results

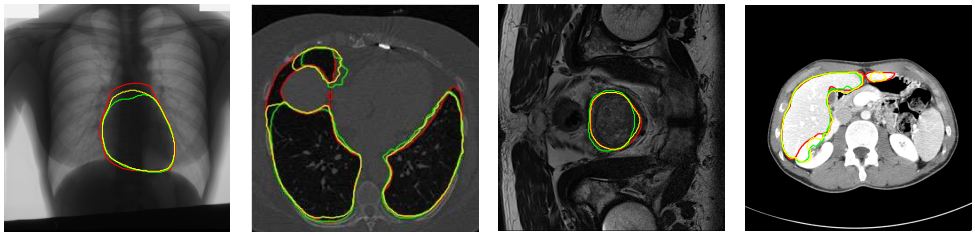
To illustrate the advantages of our framework in medical image segmentation, we compared it with the method of only using the U-net or CE-Net.

We used Dice similarity coefficient (DSC) to quantitatively measure the accuracy of the segmentation results. In the 2D cardiac dataset and LUNA, we calculated the average DSC of all test images. In the 3D prostate dataset and the liver dataset, the 2D segmentation results were stacked into 3D and then average 3D DSC was calculated. The quantitative comparison of DSC is shown in Table 1.

Table 1 The DSC (%) of the proposed method on four datasets.

Datasets	Cardiac dataset	LUNA dataset	Prostate dataset	Liver dataset
U-net	94.44±0.07	96.87±0.01	82.88±1.15	93.11±0.24
Ours U-net	95.05±0.08	97.96±0.001	83.91±1.72	94.35±0.13
CE-Net	95.05±0.02	96.20[26]	75.49±2.01	93.95±0.24
Ours CE	95.45±0.00	97.69±0.17	83.79+1.31	94.15±0.15

From Table 1, we can see that the proposed framework consistently improves the segmentation result over U-net, and the increase of average DSC is 0.61%, 1.09%, 1.03% and 1.24% on the cardiac dataset, the LUNA dataset, the prostate dataset and the liver dataset, respectively. Compared with CE-Net, our proposed framework increased 0.40%, 1.49%, 8.30% and 0.20%. An illustrative case of each dataset is shown in Fig. 2.



(a) (b) (c) (d)

Fig. 2 Segmentation results in different datasets. (a) cardiac segmentation results, (b) LUNA segmentation results, (c) prostate segmentation results, (d) liver segmentation results, where red is the gold standard, green is the U-net segmentation results, and yellow is our segmentation results.

The impact of data size

In this part, we experimented on the impact of the number of training images on the segmentation results. We did experiments by using U-net with the 2D cardiac dataset and the 3D prostate dataset. In the cardiac segmentation dataset, we used 50, 100 and 200 images as training sets and test the segmentation network on the same 47 images. In the prostate segmentation dataset, we used 8 groups (154 slices), 16 groups (311 slices) and 24 groups (447 slices) as the training sets and test the segmentation network on the same 8 groups (155 slices). The DSC results of the heart dataset and the prostate dataset are shown in Table 2 and Table 3, respectively.

Table 2 The DSC results of the cardiac dataset with different numbers of training images.

Training image number	50	100	200
U-net	93.21%	93.80%	94.44%
Ours	94.16%	94.85%	95.05%

Table 3 The DSC results of the prostate dataset with different numbers of training images.

Training image number	8	16	24
U-net	74.87%	79.27%	82.88%
Ours	79.33%	82.26%	83.91%

As shown in Table 2 and Table 3, on the one hand, the segmentation result of our framework is better than that of U-net when the number of training images is the same. On the other hand, our framework requires less training data compared to the U-net to achieve the same segmentation accuracy. The segmentation result of the proposed framework trained with 100 chest radiographs is similar to that of the U-net trained with 200 images. In the prostate data, the proposed framework trained with 16 groups achieved higher segmentation accuracy than the U-net trained with 24 groups. These results show that the proposed framework can solve the problem of insufficient training data or over-fitting to some extent and obtain a better segmentation result in small dataset.

Comparison with other data augmentation methods

In this section, in order to compare the effects of different data augmentation methods, U-net was used to conduct experiments in two-dimensional cardiac dataset. For fair comparison, we synthesized 200 samples for training with different augmentation methods.

We compared to another three data augmentation methods. First, we adopted random elastic deformation, which is a traditional augmentation method, on the original images to create 200 new synthetic images. Then, GAN was used to generate another 200 new synthetic

145 images. The 200 synthetic images by GAN and 200 original images were mixed as the input
146 of the segmentation network, which was called GAN augmentation method-I. The 200 new
147 images by GAN were used for pre-training the segmentation network, and then the 200 real
148 images were used to refine the segmentation network, which was called GAN augmentation
149 method-II.

150 **Table 4 The DSC results of the prostate dataset in different augmentation methods.**

Methods	DSC
No augmentation	94.44%
Traditional augmentation	94.59%
GAN augmentation method-I	94.55%
GAN augmentation method-II	94.68%
Ours	95.05%

151 From Table 4, we can see that the DSCs of U-net segmentation results are increased by
152 using every data augmentation method. Specifically, the segmentation results of U-net
153 increased 0.15%, 0.11%, 0.24%, 0.61% with traditional augmentation, GAN augmentation
154 method-I, GAN augmentation method-II and our framework, respectively. The segmentation
155 results using our framework improved more than those using the other three data
156 augmentation methods.

157 Discussion

158 In this paper, we reported a novel pre-training strategy based on GAN. Our method is to
159 concatenate the GAN with the segmentation network. We examined the ability of our
160 proposed framework in four different datasets. With the continuous iteration of the GAN part,
161 the image generated by generator of GAN changes from rough to fine, which was shown in
162 Fig.1.

We adopted two different segmentation networks, U-net and CE-Net, to show the inclusiveness and universality of our proposed framework. From Table 1, we can find that the U-net and CE-Net segmentation results incorporated into our framework were improved to different degrees than the original U-net and CE-Net segmentation results. At the same time，the results also showed that our framework had good compatibility with different segmentation networks. The generated image changes from rough to fine, which leads to the task of segmentation network from simple to difficult, which is conducive to the convergence of network segmentation.

Our proposed framework also has good performance in medical image segmentation tasks with small data sets. In the first training stage, the segmentation network is trained by using a series of generated images with different degree of refinement, which is equivalent to the data augmentation step.

Compared with the common data augmentation methods, our framework also has batter performance. We can see from Table 4, traditional data augmentation and GAN data augmentation methods can only improve the segmentation result slightly. Because the traditional data augmentation method increases the number of training images by adopting a random elastic deformation, and the resulted images are not in conformity with the real situation of deformation, which widens the gap between the inputs of segmentation network. On the other hand, the existing problem of GAN data augmentation method is the blurry of synthesis image. In our framework, we don't pursue to generate images that are realistic enough. GAN of our framework generates a series of synthetic images from coarse to fine to pre-train the segmentation network, and successfully decomposes the segmentation task into

simple to difficult progressive tasks, which is conducive to the convergence and optimization of the segmentation network.

Several limitations of this work should be noted. First, our proposed framework is based on 2D images. Even in 3D data, our strategy was to segment the goal area slice by slice and then a series of 2D results were reconstructed into 3D results to compare with other methods. Our future work is to integrate 3D GAN and 3D segmentation network into our framework. Another limitation is that our medical segmentation accuracy does not outperform state-of-the-art segmentation algorithms. On the one hand, we did not use the state-of-the-art segmentation networks, but used the common U-net structure. which was recognized as the backbone of medical image segmentation. We believe that our framework will get better segmentation results integrating those state-of-the-art segmentation networks. On the other hand, our focus is not whether the segmentation accuracy of our framework can exceed the state-of art segmentation networks, but whether the segmentation performance of the current segmentation networks can be effectively improved after using our proposed framework.

Conclusion

In this paper, we proposed a new end-to-end generation-segmentation framework by integrating a GAN and a segmentation network. During the training of GAN, the intermediate synthetic images generated by GAN are used to pre-train the segmentation network. After the GAN training is completed, we fine-tune the segmentation network by training it with the real image data. The proposed framework was validated on 2D cardiac dataset, 2D LUNA dataset, 3D prostate dataset and 3D liver dataset. The experiment results showed that our framework achieved higher segmentation accuracy than the only U-net and CE-Net, which

were used as the segmentation network in our framework. At the same time. Our framework had good compatibility and can incorporate other segmentation networks besides U-net and CE-Net. Our proposed framework also has good performance in medical image segmentation tasks with small data sets. Besides, compared with the common data augmentation methods, our framework also has good segmentation performance.

Methods

In this paper, we propose an end-to-end generation-segmentation framework. As shown in Fig.3, our framework includes two parts, the data synthesis part and the segmentation part. The data synthesis part is a GAN, and the segmentation part can be any segmentation network. The innovation of this paper is that during the process of GAN training, the synthetic image generated by the generator are used for the pre-training of the segmentation network. When the GAN training is completed, the pre-trained segmentation network is then refined with real training data.

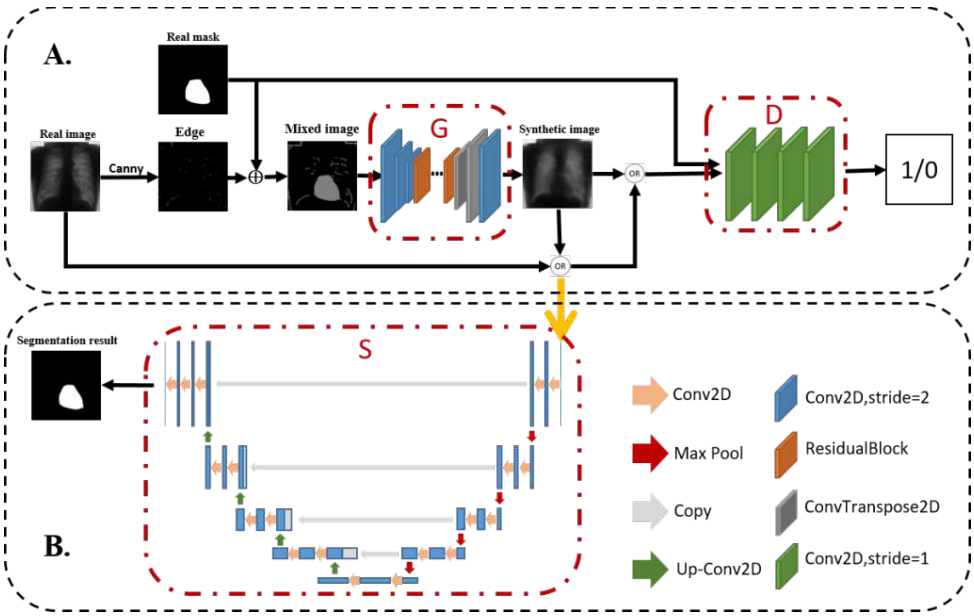


Fig. 3 The structure of our proposed framework. (A)The data synthesis part. G represents the Generator of GAN, D represents the Discriminator of GAN. (B)The segmentation part. S represents the segmentation network, such as U-net.

The training of our framework is divided into two stages. In the first stage, the GAN and the segmentation network were optimized simultaneously. In each iteration, synthetic images were generated by the GAN from inputted mixed images, which are obtained by overlaying mask images and edge images. Then the synthetic images and the corresponding masks are used to train the segmentation network. The weights of GAN and the segmentation network are updated simultaneously. In early iterations, the synthetic images mainly contain the edge information and the location information of the segmented target, and it is relatively easy to use these synthetic images to train the segmentation network. With the continuous optimization of the GAN, the synthetic images become more and more like the real images. Therefore, the GAN generates a series of synthetic images from coarse to fine to train the segmentation network, and successfully decompose the segmentation task into progressive tasks from simple to difficult, which is conducive to the convergence and optimization of the segmentation network.

Synthesis part

In this study, we borrow the idea of pix2pix [27] and used the conditional GAN as the synthesis part. The conditional GAN achieve image-to-image translation, which synthesizes a new image from an input image instead of random noise. The data synthesis part is a conditional GAN composed of a generator and a discriminator. The generator consists of 4 convolutional layers, 9 residual blocks and 2 deconvolution layers. The use of the residual block can make the generator deeper and prevent the gradient from disappearing. The discriminator is a simple two-channel Fully convolutional network composed of 4

convolutional layers (composed of convolution, batch normalization and activation function). Real or synthetic image and mask pairs are used as input to the discriminator. When the input is a real image-mask pair, the output is a full 1 image; when the input is a synthetic image-mask pair, the output is a full 0 image. The discriminator is optimized continuously to better distinguish between the real and the synthetic images, and ensure the similarity and diversity of synthetic data.

In this study, the input to the GAN is a mixed image, which is an overlay of the segmentation mask of the real image and the edges extracted from the real image by canny edge detector. We train the GAN to synthesize images similar to the real image from the mixed image.

When we optimize the generator of GAN, on the one hand, the synthetic image $\mathrm{G}(x^*)$ and real mask x pair should be able to fool the discriminator D; on the other hand, the synthetic image $\mathrm{G}(x^*)$ should be as similar as possible to the real image y. As in the training of generator G, we use the following loss function:

$$L_G = \mathbb{E}_{x,x^*}[logD(x, G(x^*))] + \lambda L1(y, G(x^*)) \quad (1)$$

where $L1$ represents L1 distance, and λ is a constant.

When we optimize the discriminator of GAN, D needs to distinguish mask-real image pair or mask-synthetic image pair through FCN. We use the following loss function for training the discriminator as in most GAN training:

$$L_D = \mathbb{E}_{x,y}[logD(x, y)] + \mathbb{E}_{x,G(x^*)}\big[log\big(1 - D(x, G(x^*))\big)\big] \quad (2)$$

Segmentation part

We use U-net and CE-Net [26] as the segmentation network, and it can be replaced by any other segmentation networks. The U-net of the segmentation part consists of a contracting path (right side) and an expansive path (left side). It utilizes skip connections at the same resolution to concatenate the lower fine feature maps to the higher coarse feature maps to improve the accuracy in biomedical image segmentation. In CE-Net, dense atrous convolution (DAC block) and residual multi-kernel pooling (RMP block) were integrated into encoder-decoder framework to reduce information loss caused by pooling and convolution layers. CE-Net improved the performance of several 2D medical image segmentation tasks.

The training of the segmentation network S is divided into two stages. In the first stage, only the synthetic images with different levels of detail are used to coarsely train the segmentation network, and we can obtain the initialization weights of the segmentation network and the predicted segmentation result of the synthetic image. In the second stage, only real images are used to fine-tune the weights of the segmentation network, and the predicted segmentation results of the real images are obtained. We calculated the L1 distance of the predicted result and mask as the loss function of segmentation network, as shown below:

$$L_S = L1\big(x, S(I)\big) \tag{3}$$

Where I represents the input of segmentation network, the synthetic image in the first stage and the real image in the second stage.

In the second stage, GAN is abandoned and only real labeled images were used to train the segmentation network. After the first stage, the segmentation network can capture the location information and approximate edge information of the target, so the learning rate can be set to a small value in second stage. The second stage is to fine-tune the weight of the segmentation network and optimize the segmentation results at the boundary of the target.

Experiments

Datasets and preprocessing

Our goal was to develop a general approach to improve the accuracy of medical image segmentation instead of improving the performance in a specific task, so we evaluated the proposed framework on four different datasets.

The first dataset was the publicly available JSRT dataset [28] that contains 247 chest radiographs. The manual segmentations of five anatomical structures (right lung, left lung, heart, right clavicle, left clavicle) for all the 247 JSRT radiographs are provided in the SCR dataset [29]. The image size is 1024×1024. In our experiment, the 247 chest radiographs and the corresponding manual segmentations of heart were used, and 200 of them were used as the training set and the other 47 were used as the test set.

The second dataset was the lung structure in 2D CT images from the Lung Nodule Analysis (LUNA) competition (*https://www.kaggle.com/kmader/finding-lungs-in-ct-data/data/*), which contains 267 images and corresponding segmentation maps. The image size is 512×512. The whole dataset was divided into 200 images for training and 67 images for testing.

The third dataset was the prostate dataset provided by Medical Segmentation Decathlon challenge (*https://decathlon-10.grand-challenge.org*). It consists of 32 transverse T2-weighted scans with resolution of $0.6\text{mm} \times 0.6\text{mm} \times 4\text{mm}$ and the manual segmentation of the whole prostate. Our framework was based on 2D images, so the 3D data were segmented slice by slice. The size of each slice was 320×320. In total, 24 groups (442 slices) were used for training and 8 groups (155 slices) for testing.

The fourth dataset was the first batch liver CT data provided by Chaos challenge (*https://chaos.grand-challenge.org*), which contains 10 groups of 3D data. The size of each slice was 512×512. We used 5 groups (483 slices) for training and 5 groups (497 slices) for testing.

In order to use the same framework for all the experiments, we used bicubic interpolation to down-sample all the data to a resolution of 256×256. For each dataset, the intensities of all images were rescaled into $[-1,1]$.

Experiment settings

We used the Canny edge detection function provided in OpenCV, which needed two threshold parameters, the minimum threshold and the maximum threshold. They were set 20 and 60 for the heart dataset, 200 and 600 for the prostate dataset, and 120 and 360 for the liver dataset.

For the post-processing, conditional random field (CRF) was adopted to optimize the predicted segmentation results. We adopted PyTorch to implement the proposed framework and ran it on a NVIDIA GeForce GTX 2080Ti GPU. For fair comparison, we trained U-net and CE-Net for 200 epochs as baseline. The segmentation network of proposed framework was trained by the generated images for 100 epochs and the real images for 200 epochs.

List of abbreviations

GAN: Generative Adversarial Networks; ILSVRC: ImageNet Large Scale Visual Recognition Competition; CNNs: Convolutional Neural Networks; OAR: Organ At Risk; ASM: Active Shape Model; AAM: Active Appearance Model; SSM: Statistical Shape Model; TPS: Thin Plate Spline; MR: Magnetic Resonance; CT: Computed Tomography; cGAN: Conditional Generative Adversarial Nets; DSC: Dice Similarity Coefficient; DAC: Dense Atrous Convolution; RMP: Residual Multi-kernel Pooling; LUNA: Lung Nodule Analysis.

Declarations

Ethics approval and consent to participate

Bioethics Commission for this study is not necessary.

Consent for publication

Not applicable.

Availability of data and materials

The datasets analyzed during the current study are available in the SCR database (http://www.isi.uu.nl/Research/Databases/SCR/), LUNA (https://www.kaggle.com/kmader/finding-lungs-in-ct-data/data/), Medical Segmentation Decathlon challenge (https://decathlon-10.grand-challenge.org) and Chaos challenge (https://chaos.grand-challenge.org).

Competing interests

The authors declare that they have no competing interests.

Funding

There was no funding for the presented research.

Authors' contributions

KC designed and conducted this study under the supervision of MW and ZS. All authors read and approved the final manuscript.

Acknowledgements

Not applicable.

References

1. Krizhevsky A, Sutskever I, Hinton GE, editors. Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems; 2012.
2. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, et al. A survey on deep learning in medical image analysis. Medical image analysis. 2017;42:60-88.
3. Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, et al. Imagenet large scale visual recognition challenge. International journal of computer vision. 2015;115(3):211-52.
4. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. Lect Notes Comput Sc. 2015;9351:234-41.
5. Nie D, Wang L, Trullo R, Li J, Yuan P, Xia J, et al. Segmentation of Craniomaxillofacial Bony Structures from MRI with a 3D Deep-Learning Based Cascade Framework. Mach Learn Med Imaging. 2017;10541:266-73.
6. Wang G, Zuluaga MA, Li W, Pratt R, Patel PA, Aertsen M, et al. DeepIGeoS: a deep interactive geodesic framework for medical image segmentation. IEEE transactions on pattern analysis and machine intelligence. 2018;41(7):1559-72.
7. Wang Y, Zhao L, Wang M, Song Z. Organ at Risk Segmentation in Head and Neck CT Images Using a Two-Stage Segmentation Framework Based on 3D U-Net. IEEE Access. 2019;7:144591-602.
8. Havaei M, Davy A, Warde-Farley D, Biard A, Courville A, Bengio Y, et al. Brain tumor segmentation with deep neural networks. Medical image analysis. 2017;35:18-31.
9. Zhang W, Li R, Deng H, Wang L, Lin W, Ji S, et al. Deep convolutional neural networks for multi-modality isointense infant brain image segmentation. NeuroImage. 2015;108:214-24.
10. Pan SJ, Yang Q. A survey on transfer learning. IEEE Transactions on knowledge and data engineering. 2009;22(10):1345-59.
11. de Bruijne M. Machine learning approaches in medical image analysis: From detection to diagnosis. Elsevier; 2016.

12. Shin H-C, Roth HR, Gao M, Lu L, Xu Z, Nogues I, et al. Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. Ieee T Med Imaging. 2016;35(5):1285-98.
13. Ghafoorian M, Mehrtash A, Kapur T, Karssemeijer N, Marchiori E, Pesteie M, et al., editors. Transfer Learning for Domain Adaptation in MRI: Application in Brain Lesion Segmentation2017; Cham: Springer International Publishing.
14. Raghu M, Zhang C, Kleinberg J, Bengio S, editors. Transfusion: Understanding transfer learning for medical imaging. Advances in Neural Information Processing Systems; 2019.
15. He K, Girshick R, Dollár P, editors. Rethinking imagenet pre-training. Proceedings of the IEEE International Conference on Computer Vision; 2019.
16. Shin HO, Blietz M, Frericks B, Baus S, Savellano D, Galanski M. Insertion of virtual pulmonary nodules in CT data of the chest: development of a software tool. Eur Radiol. 2006;16(11):2567-74.
17. Cootes TF, Hill A, Taylor CJ, Haslam J, editors. The use of active shape models for locating structures in medical images. Biennial International Conference on Information Processing in Medical Imaging; 1993: Springer.
18. Cootes TF, Edwards GJ, Taylor CJ, editors. Active appearance models. European conference on computer vision; 1998: Springer.
19. Tang Z, Chen K, Pan M, Wang M, Song Z. An Augmentation Strategy for Medical Image Processing Based on Statistical Shape Model and 3D Thin Plate Spline for Deep Learning. IEEE Access. 2019;7:133111-21.
20. Goodfellow IJ, Pouget-Abadie J, Mirza M, Bing X, Warde-Farley D, Ozair S, et al., editors. Generative Adversarial Nets. International Conference on Neural Information Processing Systems; 2014.
21. Zhao M, Wang L, Chen J, Nie D, Cong Y, Ahmad S, et al. Craniomaxillofacial Bony Structures Segmentation from MRI with Deep-Supervision Adversarial Learning. Med Image Comput Comput Assist Interv. 2018;11073:720-7.
22. Costa P, Galdran A, Meyer MI, Abràmoff MD, Niemeijer M, Mendonça AM, et al. Towards adversarial retinal image synthesis. arXiv preprint arXiv:170108974. 2017.
23. Costa P, Galdran A, Meyer MI, Niemeijer M, Abramoff M, Mendonca AM, et al. End-to-End Adversarial Retinal Image Synthesis. Ieee T Med Imaging. 2018;37(3):781-91.
24. Mok TC, Chung AC, editors. Learning data augmentation for brain tumor segmentation with coarse-to-fine generative adversarial networks. International MICCAI Brainlesion Workshop; 2018: Springer.
25. Mirza M, Osindero S. Conditional generative adversarial nets. arXiv preprint arXiv:14111784. 2014.
26. Gu Z, Cheng J, Fu H, Zhou K, Hao H, Zhao Y, et al. CE-Net: Context Encoder Network for 2D Medical Image Segmentation. Ieee T Med Imaging. 2019.
27. Isola P, Zhu JY, Zhou TH, Efros AA. Image-to-Image Translation with Conditional Adversarial Networks. Proc Cvpr Ieee. 2017:5967-76.
28. Shiraishi J, Katsuragawa S, Ikezoe J, Matsumoto T, Kobayashi T, Komatsu K-i, et al. Development of a digital image database for chest radiographs with and without a

lung nodule: receiver operating characteristic analysis of radiologists' detection of pulmonary nodules. American Journal of Roentgenology. 2000;174(1):71-4.

29. Van Ginneken B, Stegmann MB, Loog M. Segmentation of anatomical structures in chest radiographs using supervised methods: a comparative study on a public database. Medical image analysis. 2006;10(1):19-40.

Figures

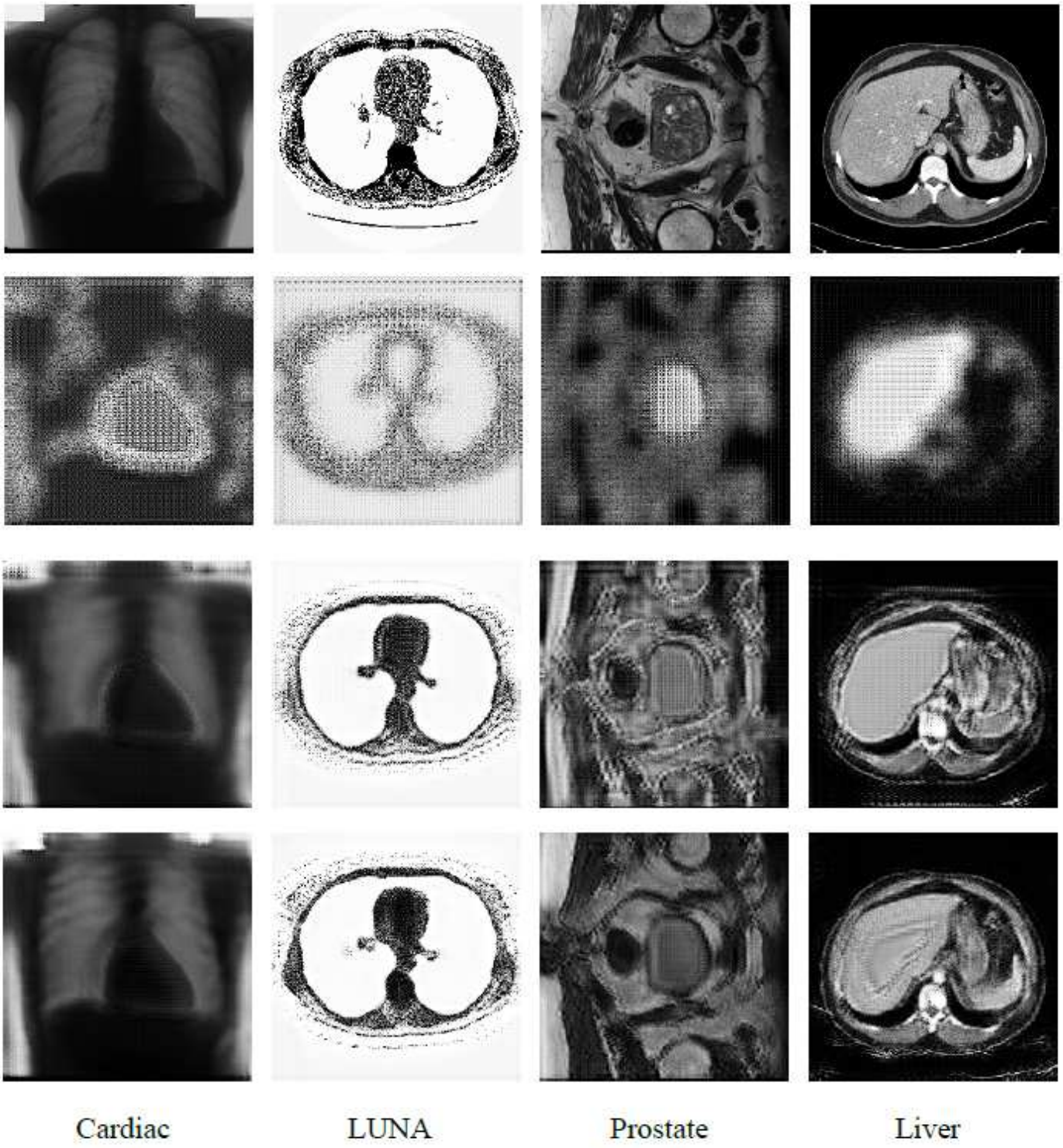


Figure 1

Generated images by the generator of GAN in different iterations for different datasets. The first row represents the original images, and the 2nd, 3rd and fourth rows represent the generated images after 1, 50, and 100 epochs, respectively.

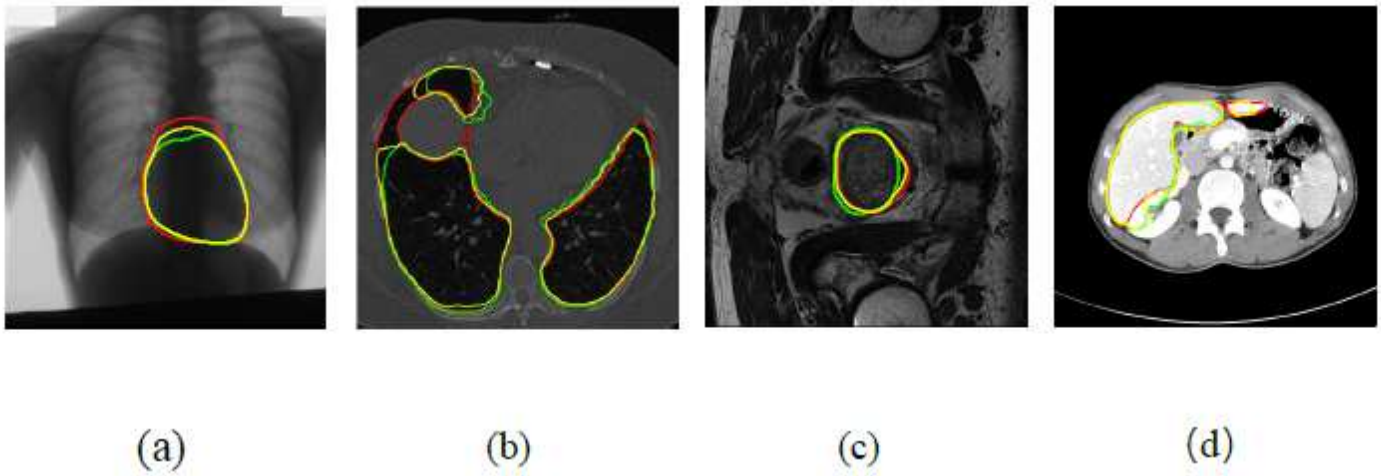


Figure 3

Segmentation results in different datasets. (a) cardiac segmentation results, (b) LUNA segmentation results, (c) prostate segmentation results, (d) liver segmentation results, where red is the gold standard, green is the U-net segmentation results, and yellow is our segmentation results.

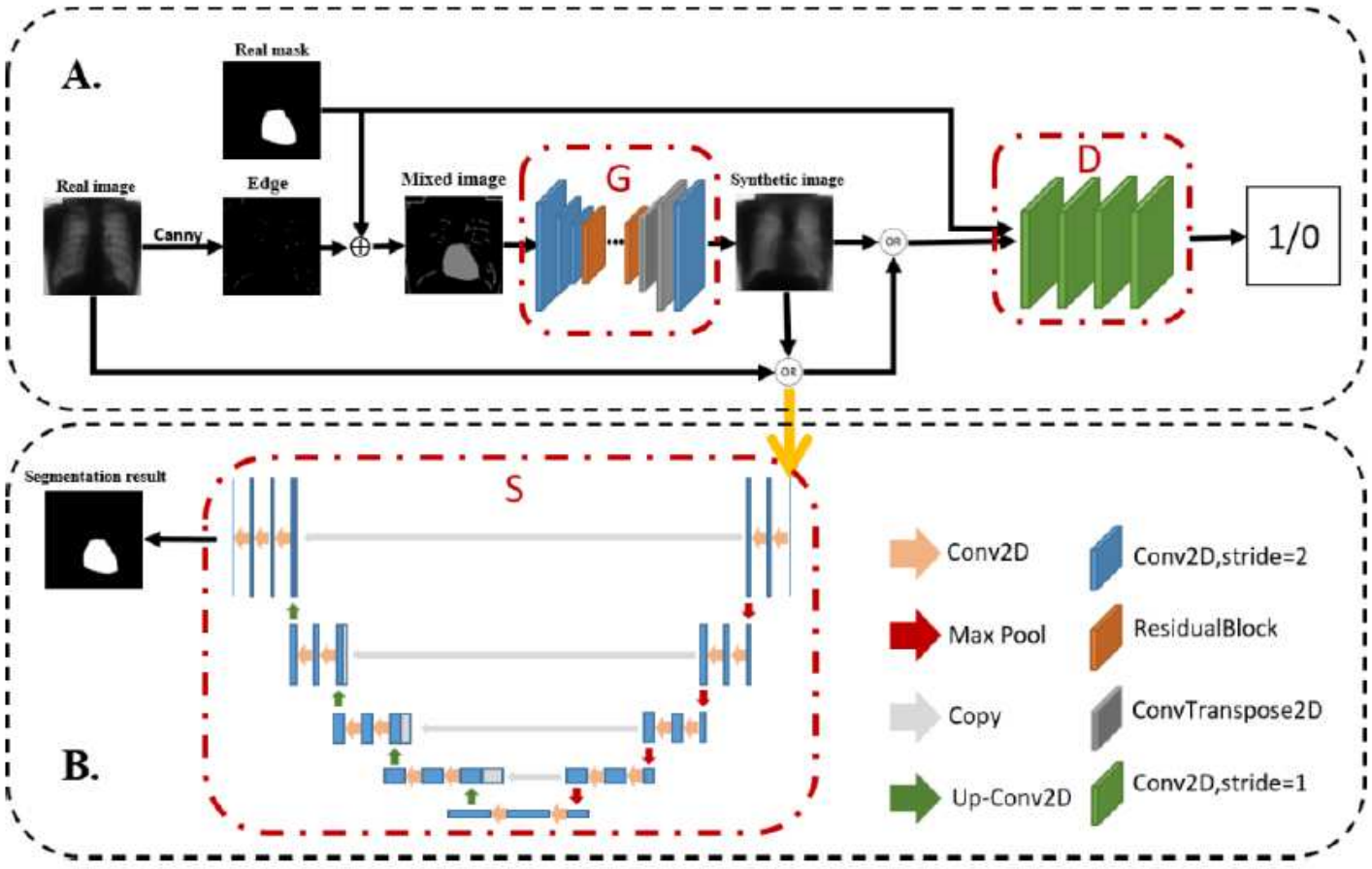


Figure 6

The structure of our proposed framework. (A)The data synthesis part. G represents the Generator of GAN, D represents the Discriminator of GAN. (B)The segmentation part. S represents the segmentation network, such as U-net.