

Annotation-Free Cardiac Vessel Segmentation via Knowledge Transfer from Retinal Images

Fei Yu¹, Jie Zhao², Yanjun Gong³, Zhi Wang³, Yuxi Li³, Fan Yang³, Bin Dong^{4,1,2}, Quanzheng Li^{5,2}, and Li Zhang^{1,2}(✉)

¹ Center for Data Science, Peking University, Beijing 100871, China;
zhangli_pku@pku.edu.cn

² Center for Data Science in Health and Medicine, Peking University,
Beijing 100871, China;

³ Department of Cardiology, Peking University First Hospital, Beijing 100034, China;

⁴ Beijing International Center for Mathematical Research (BICMR), Peking
University, Beijing 100871, China;

⁵ MGH/BWH Center for Clinical Data Science, Boston, MA 02115, USA.

Abstract. Segmenting coronary arteries is challenging, as classic unsupervised methods fail to produce satisfactory results and modern supervised learning (deep learning) requires manual annotation which is often time-consuming and can some time be infeasible. To solve this problem, we propose a knowledge transfer based shape-consistent generative adversarial network (SC-GAN), which is an annotation-free approach that uses the knowledge from publicly available annotated fundus dataset to segment coronary arteries. The proposed network is trained in an end-to-end fashion, generating and segmenting synthetic images that maintain the background of coronary angiography and preserve the vascular structures of retinal vessels and coronary arteries. We train and evaluate the proposed model on a dataset of 1092 digital subtraction angiography images, and experiments demonstrate the supreme accuracy of the proposed method on coronary arteries segmentation.

Keywords: Coronary Artery Segmentation · Knowledge Transfer · Generative Adversarial Network · Deep Learning.

1 Introduction

Quantitative measurement of coronary arteries in medical images is important for the diagnosis, prevention and therapeutic evaluation of related diseases including hypertension, myocardial infarction, and coronary atherosclerotic disease. In the diagnosis of coronary diseases, digital subtraction angiography (DSA) has been widely used and considered the "gold standard". To quantitatively segment blood vessels in DSA, researchers have developed automated methods including region growing, level sets, and Hessian analysis [1, 11]. However, due to the complexity of the vascular morphology, these unsupervised methods are difficult to obtain a clinically satisfactory segmentation of coronary

arteries. On the other hand, supervised learning such as deep neural networks (DNNs) can produce better segmentation results but relies heavily on pixel-level image annotation, which is often expensive, time-consuming, and even impossible to access especially for coronary artery segmentation.

To solve the problem of the lack of DSA vessel annotation, we aim to apply a strategy to transfer the knowledge of retinal vessel segmentation to the coronary artery segmentation. The researchers have established and validated a number of public retinal vessel segmentation datasets, including DRIVE[12], STARE[4], and RITE[5]. Due to the significant differences between their anatomical regions, traditional transfer methods are not suitable. Therefore, we present a novel knowledge transfer based adversarial model containing three parts called generator, discriminator, and segmentor. Training of the model contains three major steps: 1) Frangi vessel analysis [2] is used to segment the coronary artery in the DSA images roughly. 2) The adversarial training between generator and discriminator allows the model to fuse the fundus image and the DSA image. 3) A synthetic label is then created by computing the union of rough coronary artery segmentation and retinal vessel annotation. Moreover, a shape-consistent scheme is used to ensure the shape consistency of synthetic images and synthetic annotations. The fused image and corresponding synthetic label are used to train the segmentor. The supreme accuracy demonstrates the effectiveness of our methods, which improves the accuracy of coronary artery segmentation by using the knowledge of fundus segmentation without additional manual annotation of the DSA images. The ideas of knowledge transfer and data fusion in this paper have many other application scenarios, including cell segmentation, neural segmentation, and airway segmentation. As long as the object structures in the two datasets are similar, we can use the knowledge from the annotated dataset to guide the analysis of the unannotated one.

Our work relates closely to the recent rise of knowledge transfer techniques. In the field of natural image analysis, Domain-Adversarial Neural Network (DANN) transfers the feature distribution to solve the domain-shift problem [3]. CycleGAN introduces a cycle consistency loss and achieves unpaired image-to-image translation [17]. AdaptSegNet adopts adversarial learning in the output space and receives favorably accuracy and visual quality [13]. In the field of medical image analysis, many studies have been dedicated to exploring cross-modality translation with GAN [10]. Using synthetic data to overcome insufficient labeled data is also an active research area. For example, using synthetic data as augmented training data can help lesion segmentation [7] and cardiovascular volumes segmentation [16]. These methods show that knowledge transfer is effective for the same anatomical region, and our work contributes by accomplishing knowledge transfer between two different anatomical regions.

2 Methods

In this section, we show two models based on knowledge transfer. The first one is the GAN model with a constraint of shape consistency (SC-GAN) proposed in

this work. Then, to verify the necessity of fusion, we show another simpler model. It adopts from Mixup[15], which is to train the U-Net model by computing an average of the fundus image and the DSA image (Add U-Net).

2.1 SC-GAN

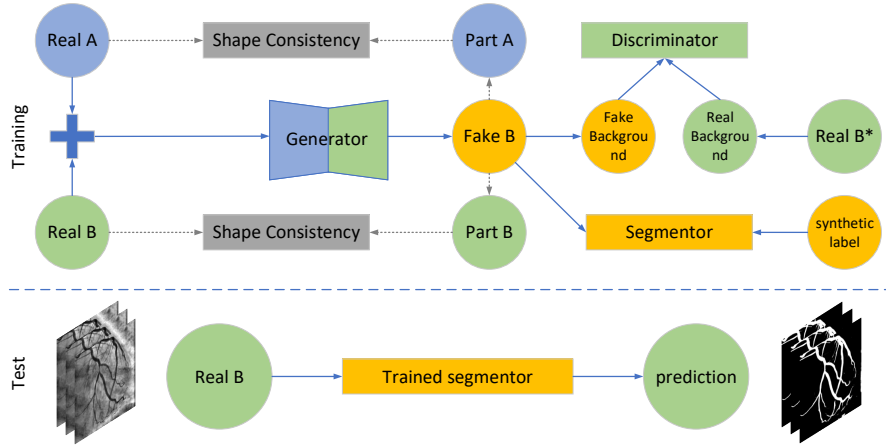


Fig. 1. The illustration of the proposed SC-GAN, an end-to-end approach for coronary artery segmentation requiring no new manual annotations.

Figure 1 shows the training and test processes of the proposed model. During the training process, the generator, discriminator, and segmentor are trained simultaneously. In the test process, only the trained segmentor is needed to segment coronary arteries, which requires much less memory and inference time than training. In the previously reported knowledge transfer between different modalities within the same anatomy, the foreground and the background can be reasonably registered. In our task, however, the foreground and background of the images are completely mismatched. So we have designed a shape-consistent scheme that allows the generator and discriminator to complete the knowledge transfer of foreground and background respectively. In the following subsections, we will discuss the model’s architectures and the objective functions in detail.

Generator: As shown in Figure 1, the generator uses U-Net as its network backbone. The input of the generator is an average of the fundus image (Real A) and the DSA image (Real B). The output is a synthetic image (Fake B) with the same dimensions as the input. To ensure that the Fake B has both retinal vessels and coronary arteries, we extract the two vascular regions (Part A and Part B) in Fake B using manual annotation of retinal vessels and Frangi segmentation of DSA images. A shape-consistent loss (l_1 loss) is then used to regularize the

content of the vessel regions in Fake B to be consistent with the corresponding regions in the original images,

$$\mathcal{L}_{shape_A}(G) = \mathbb{E}_{A \sim p_{data}(A), B \sim p_{data}(B)}[\|label_A * (G(A + B) - A)\|_1] \quad (1)$$

$$\mathcal{L}_{shape_B}(G) = \mathbb{E}_{A \sim p_{data}(A), B \sim p_{data}(B)}[\|label_B * (G(A + B) - B)\|_1] \quad (2)$$

where A represents the fundus image, B represents the DSA image, and $label_A$ and $label_B$ are the retinal vessel annotation and the Frangi segmentation results, respectively.

Discriminator: We expect that the background of Fake B is sufficiently similar to the background of a DSA image, so we first use the annotations and Frangi analysis results to mask out the vascular regions and extract the background in the generated and real DSA images,

$$Fake_{bg} = \neg(label_A \cup label_B) * G(A + B) \quad (3)$$

$$Real_{bg} = \neg(label_A \cup label_B) * B^* \quad (4)$$

Where B^* (Real B^* in Figure 1) represents a randomly chosen DSA image before the injection of the contrast medium, which guarantees a vessel-free background. For the structure of the discriminator, we adopt PatchGAN [17]. The adversarial loss between the generator and the discriminator can be expressed as,

$$\begin{aligned} \mathcal{L}_{GAN}(G, D) = & \mathbb{E}_{B^* \sim p_{data}(B^*)}[\log D(Real_{bg})] \\ & + \mathbb{E}_{A \sim p_{data}(A), B \sim p_{data}(B)}[\log(1 - D(Fake_{bg}))] \end{aligned} \quad (5)$$

Segmentor: The main structure of the segmentor is also a U-Net. We use MultiLabelSoftMarginLoss [9] as the objective function of the segmentor:

$$\mathcal{L}_{seg}(S) = -(y(i) \log[\frac{\exp(\hat{y}(i))}{1 + \exp(\hat{y}(i))}] + (1 - y(i)) \log[\frac{1}{1 + \exp(\hat{y}(i))}]) \quad (6)$$

where \hat{y} is the prediction and y the synthetic label.

Therefore, the final objective function of our proposed model is,

$$\mathcal{L}(G, D, S) = \mathcal{L}_{GAN}(G, D) + \mathcal{L}_{seg}(S) + \lambda \mathcal{L}_{shape_A}(G) + \mu \mathcal{L}_{shape_B}(G) \quad (7)$$

where λ and μ control the relative importance of the objectives. During training, we set $\lambda = 100, \mu = 50$.

The model uses instance normalization [14] instead of batch normalization [6], the generator uses ReLU and the discriminator uses LeakyReLU as activations.

2.2 Add U-Net

Figure 2 shows the overall structure of Add U-Net. Referring to Mixup, the model takes the average of the fundus image (Real A) and the DSA image (Real B). The manual annotation of retinal vessels and the Frangi analysis results of the DSA image are combined to obtain the label of the added image. We then train a U-Net with such added images and annotations. An independent DSA dataset is used to evaluate the trained U-Net.

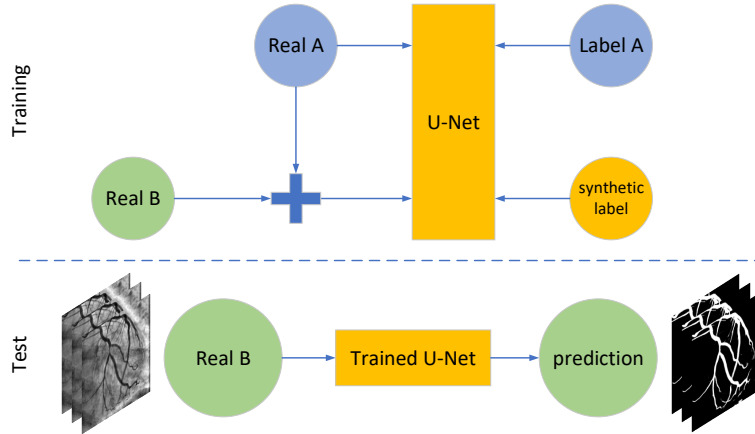


Fig. 2. The illustration of Add U-Net, where the input is an average of fundus photography and DSA image.

3 Experiments

In this section, to evaluate the effectiveness of our proposed SC-GAN, we compare the segmentation of four methods: **1) Frangi Algorithm** Multi-scale Frangi vessel analysis is used to segment coronary arteries. **2) Classic U-Net.** A U-Net model is trained using Frangi Algorithm results as the learning targets. **3) Add U-Net.** A U-Net model is trained using an average of fundus photographs and DSA images. **4) SC-GAN** The proposed shape-consistent GAN. We also compare the synthetic results of SC-GAN and Cycle-GAN [17]. The data and experiment details are presented below.

Data: We use the DRIVE [12] dataset as the source domain of knowledge transfer. The DRIVE dataset includes 40 fundus images with manually annotated retinal vessels. We also collect 1092 coronary angiographies (DSA) with no annotations as the target domain of knowledge transfer. Several preprocessing approaches are performed on the fundus images, including color to grayscale transform, median-filtering and contrast-limited adaptive histogram equalization [18]. Finally, we resize all images into the same size of 512×512 and randomly choose 256×256 patches as inputs of the models.

Experiment details: In all experiments, 50% of DSA images are randomly selected as training set, 20% are validation set and 30% are test set. Meanwhile, we use the Adam solver [8] with a learning rate of $2e-4$. After training for 50 epochs, we decrease the learning rate linearly for 50 epochs till 0.

4 Results

In this section, we briefly report the evaluation results in two aspects: 1) Images synthesis and 2) Images segmentation.

Images synthesis: Figure 3 shows some examples of the fundus, DSA, and synthetic image patches. Compared to the results of Cycle-GAN [17], the synthetic images from our proposed SC-GAN have more realistic DSA background and also preserve the vascular structures corresponding to the labels (see the columns (c) and (d) in Figure 3).

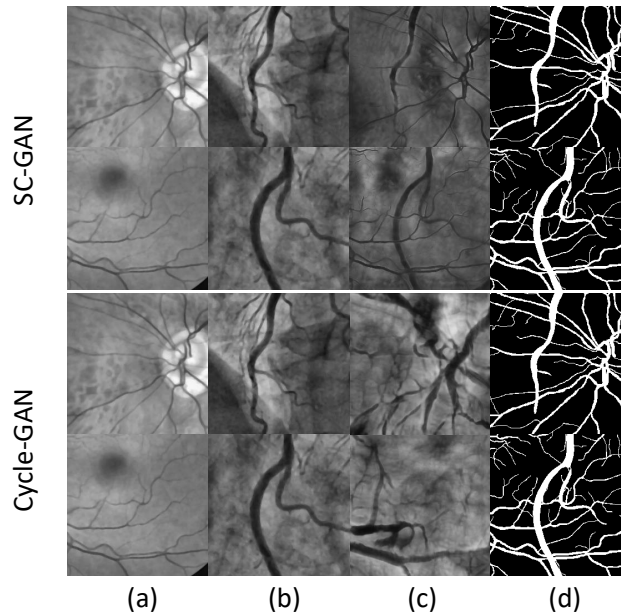


Fig. 3. Comparison of SC-GAN and Cycle-GAN. (a) Fundus patches, (b) DSA patches, (c) synthetic images, (d) synthetic labels.

Images segmentation: We annotate 30% of the DSA dataset (328 out of 1092 images) and evaluate our proposed model on it. Table 4 compares the performance of different methods. As the baseline method of this article, the Frangi algorithm has a Dice score of 0.636 ± 0.046 . If the result of the Frangi algorithm is used as an annotation to train a U-Net (Classic U-Net), the Dice score reduces to 0.589 ± 0.049 . Both Add U-Net and SC-GAN have higher Dice scores (0.742 ± 0.048 and 0.824 ± 0.026). And SC-GAN also outperforms the other methods in terms of accuracy and recall. Figure 4 shows some typical examples in the test set. Columns (d-e) show better results than columns (b-c), indicating that

knowledge transfer effectively enhances the identification of small blood vessels. By comparing the results of Add U-Net and SC-GAN, we can also find that GAN is better than an average in terms of the quality of knowledge transfer.

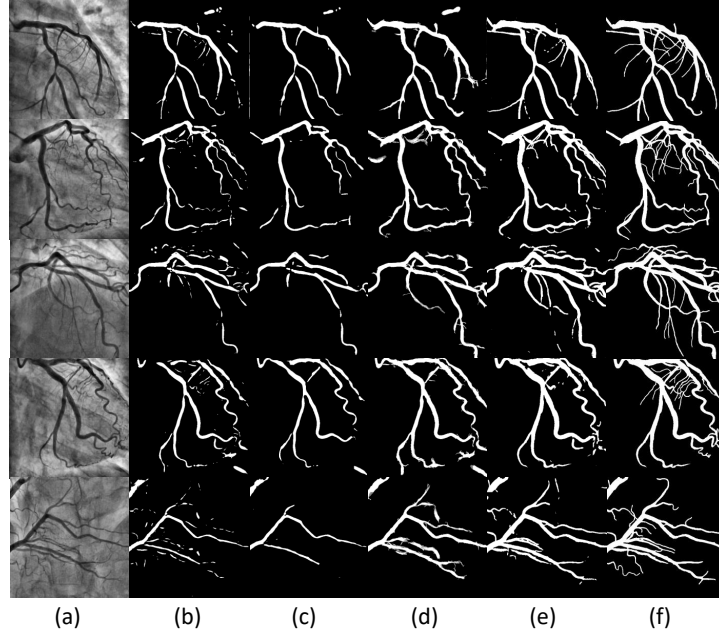


Fig. 4. Examples of different vessel segmentation methods. (a) Original images, (b) Frangi Algorithm, (c) Classic U-Net, (d) Add U-Net, (e) Proposed SC-GAN, (f) Ground truth.

Table 1. Quantitative performance of different vessel segmentation methods.

Methods	Frangi Algorithm	Classic U-Net	Add U-Net	Proposed
Accuracy	0.927±0.014	0.921±0.015	0.940±0.012	0.953±0.009
Precision	0.943±0.057	0.975±0.052	0.864±0.078	0.820±0.031
Recall	0.481±0.048	0.423±0.047	0.653±0.050	0.829±0.039
Dice Coefficient	0.636±0.046	0.589±0.049	0.742±0.048	0.824±0.026

5 Discussion

In this paper, we proposed a shape-consistent GAN model (SC-GAN) for coronary artery segmentation, which was able to transfer the knowledge of the seg-

mentation on public fundus dataset to an unlabeled DSA dataset. Experimental results demonstrated that SC-GAN obtained an obvious superior performance on coronary arteries segmentation. Despite the promising results, our method has several limitations and requires further investigation: 1. How well does the method perform on other datasets; 2. Although the segmentor is light-weighted in test, training SC-GAN is much more complex than in a classic supervised deep model. In future work, we will further test the proposed SC-GAN on other application scenarios and simplify the training process.

Acknowledgments. This work was supported by National Key R&D Program of China (No. 2018YFC0910700); The National Natural Science Foundation of China (NSFC) under Grants 81801778, 11831002; Beijing Natural Science Foundation (Z180001).

References

1. Adams, R., Bischof, L.: Seeded region growing. *IEEE Transactions on pattern analysis and machine intelligence* **16**(6), 641–647 (1994)
2. Frangi, A.F., Niessen, W.J., Vincken, K.L., Viergever, M.A.: Multiscale vessel enhancement filtering. In: *International conference on medical image computing and computer-assisted intervention*. pp. 130–137. Springer (1998)
3. Ganin, Y., Ustinova, E., Ajakan, H., Germain, P., Larochelle, H., Laviolette, F., Marchand, M., Lempitsky, V.: Domain-adversarial training of neural networks. *The Journal of Machine Learning Research* **17**(1), 2096–2030 (2016)
4. Hoover, A., Goldbaum, M.: Locating the optic nerve in a retinal image using the fuzzy convergence of the blood vessels. *IEEE transactions on medical imaging* **22**(8), 951–958 (2003)
5. Hu, Q., Abramoff, M.D., Garvin, M.K.: Automated separation of binary overlapping trees in low-contrast color retinal images. In: *International conference on medical image computing and computer-assisted intervention*. pp. 436–443. Springer (2013)
6. Ioffe, S., Szegedy, C.: Batch normalization: Accelerating deep network training by reducing internal covariate shift. In: *International Conference on Machine Learning*. pp. 448–456 (2015)
7. Kamnitsas, K., Baumgartner, C., Ledig, C., Newcombe, V., Simpson, J., Kane, A., Menon, D., Nori, A., Criminisi, A., Rueckert, D., et al.: Unsupervised domain adaptation in brain lesion segmentation with adversarial networks. In: *International conference on information processing in medical imaging*. pp. 597–609. Springer (2017)
8. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980* (2014)
9. Lapin, M., Hein, M., Schiele, B.: Analysis and optimization of loss functions for multiclass, top-k, and multilabel classification. *IEEE transactions on pattern analysis and machine intelligence* **40**(7), 1533–1554 (2018)
10. Nie, D., Trullo, R., Lian, J., Petitjean, C., Ruan, S., Wang, Q., Shen, D.: Medical image synthesis with context-aware generative adversarial networks. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. pp. 417–425. Springer (2017)

11. Osher, S., Sethian, J.A.: Fronts propagating with curvature-dependent speed: algorithms based on hamilton-jacobi formulations. *Journal of computational physics* **79**(1), 12–49 (1988)
12. Staal, J., Abràmoff, M.D., Niemeijer, M., Viergever, M.A., Van Ginneken, B.: Ridge-based vessel segmentation in color images of the retina. *IEEE transactions on medical imaging* **23**(4), 501–509 (2004)
13. Tsai, Y.H., Hung, W.C., Schuler, S., Sohn, K., Yang, M.H., Chandraker, M.: Learning to adapt structured output space for semantic segmentation. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. pp. 7472–7481 (2018)
14. Ulyanov, D., Vedaldi, A., Lempitsky, V.: Instance normalization: The missing ingredient for fast stylization. *arXiv preprint arXiv:1607.08022* (2016)
15. Zhang, H., Cisse, M., Dauphin, Y.N., Lopez-Paz, D.: mixup: Beyond empirical risk minimization. *arXiv preprint arXiv:1710.09412* (2017)
16. Zhang, Z., Yang, L., Zheng, Y.: Translating and segmenting multimodal medical volumes with cycle-and shape-consistency generative adversarial network. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. pp. 9242–9251 (2018)
17. Zhu, J.Y., Park, T., Isola, P., Efros, A.A.: Unpaired image-to-image translation using cycle-consistent adversarial networks. In: *Proceedings of the IEEE International Conference on Computer Vision*. pp. 2223–2232 (2017)
18. Zuiderveld, K.: Contrast limited adaptive histogram equalization. In: *Graphics gems IV*. pp. 474–485. Academic Press Professional, Inc. (1994)