

# PROCEEDINGS OF SPIE

[SPIDigitalLibrary.org/conference-proceedings-of-spie](https://SPIDigitalLibrary.org/conference-proceedings-of-spie)

## A generative adversarial framework for capillary non-perfusion regions segmentation in fundus fluorescein angiograms

Cai, Mulin, Xiang, Dehui, Pan, Shengxue, Shi, Fei, Zhu, Weifang, et al.

Mulin Cai, Dehui Xiang, Shengxue Pan, Fei Shi, Weifang Zhu, Xinjian Chen, Bei Tian, "A generative adversarial framework for capillary non-perfusion regions segmentation in fundus fluorescein angiograms," Proc. SPIE 11596, Medical Imaging 2021: Image Processing, 115962D (15 February 2021); doi: 10.1117/12.2581045

**SPIE.**

Event: SPIE Medical Imaging, 2021, Online Only

# A Generative Adversarial Framework for Capillary Non-perfusion Regions Segmentation in Fundus Fluorescein Angiograms

Mulin Cai<sup>1</sup>, Dehui Xiang<sup>1\*</sup>, Shengxue Pan<sup>1</sup>, Fei shi<sup>1</sup>, Weifang Zhu<sup>1</sup>, Xinjian Chen<sup>1,2</sup>

<sup>1</sup>School of Electronics and Information Engineering, Soochow University, Suzhou, 215006, China

<sup>2</sup>State Key Laboratory of Radiation Medicine and Protection, Soochow University, Suzhou, 215123, China

## ABSTRACT

Retinal capillary non-perfusion (CNP) is one of diabetic retinal vascular diseases. As the capillaries are occluded, blood stops flowing to certain regions of the retina, resulting in the formation of non-perfused regions. Accurate determination of the area and change of CNP is of great significance in clinical judgment of the extent of vascular obstruction and selection of treatment methods. This paper proposes a novel generative adversarial framework, and realize the segmentation of non-perfusion regions in fundus fluorescein angiography images. The generator G of GANs is trained to produce “real” images; while an adversarially trained discriminator D is trained to do as well as possible at detecting “fakes” images from the generator. In this paper, a U-shaped network is used as the discriminator. Our method is validated using on 138 clinical fundus fluorescein angiography images. Experimental results show that our method achieves more accurate segmentation results than that of state-of-the-art approaches.

**KEYWORDS:** Capillary non-perfusion regions, image segmentation, generative adversarial networks

## 1. INTRODUCTION

Retinal capillary non-perfusion (CNP) is one of diabetic retinal vascular diseases. In the non-perfusion area of the retina, the macula of the fundus is in a state of ischemia and insufficient oxygen, and new blood vessels are prone to occur near the non-perfusion area. Because the permeability of the neovascular wall is extremely unstable, exudation and hemorrhage are likely to occur. Vitreous hemorrhage can cause blindness in severe cases. Therefore, accurate determination of the range and variation of CNP has important reference significance for determining the extent of vascular obstruction and selecting the treatment method in the clinical.

There are huge differences in the shape, distribution, and size of the non-perfusion regions between patients. The gray values between the normal regions and the non-perfusion regions are also the similar. Fig. 1 shows fundus fluorescein angiography images from two different patients, (a) and (c) are the original image, (b) and (d) are the corresponding ground truth. In Fig. 1 (a), the non-perfusion regions have different shapes and sizes, and distribute throughout the image. However, in Fig. 1 (c), most of them are large regions with perfusion and mainly distribute in the upper half of the image.

There are many methods for detecting and segmenting diabetic retinal lesions, but there are few studies on non-perfusion areas based on FFA images. Piotr Jasiobedzki<sup>[1]</sup> used the position of the main blood vessel to subdivide the image into many original regions, and estimated the extent of non-perfusion of each region by measuring the texture characteristics through mathematical morphology. The extent of perfusion information merges adjacent areas. Kwong EMT<sup>[5]</sup> used the variational texture segmentation model and a rapid optimization strategy to segment the image into CNP and non-CNP. Jasiobedzki<sup>[4]</sup> used homomorphic filtering to perform light correction and healthy capillary detection on FFA images; Jayanthi Sivaswamy<sup>[3]</sup> modeled CNP as a valley function, and proposed a new method for

---

\* Corresponding author: Dehui Xiang, E-mail: xiangdehui@suda.edu.cn

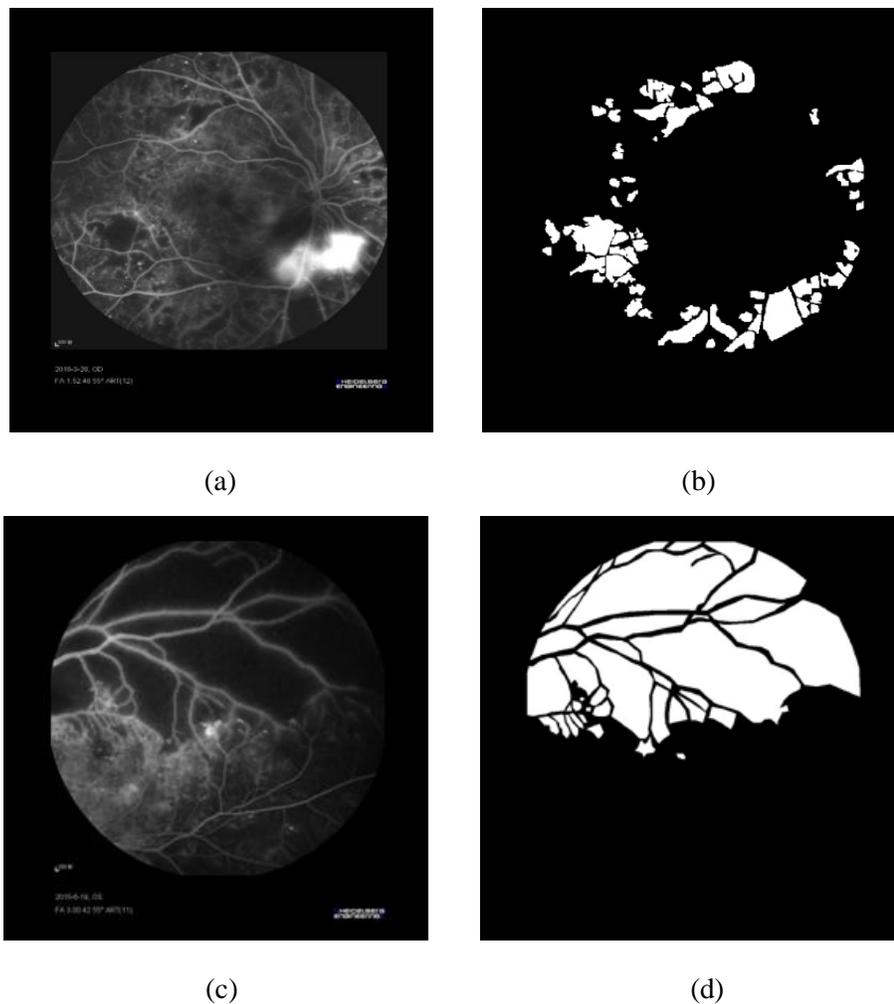


Fig.1. Illustration of the challenges in CNP area segmentation. (a) and (c) are the original image, (b) and (d) are the corresponding ground truth.

valley function detection based on extreme pyramids. For the obtained valley points, the CNP region is segmented using a variance-based region growth scheme. Yalin Zheng<sup>[2]</sup> proposed an unsupervised texture segmentation technology, using a total mutation energy minimization algorithm. The segmentation results of this algorithm are trained on a set of texture features and become the candidate regions for further refinement of the supervised integrated classifier.

## 2. METHODS

We introduce the proposed method in the following three parts: the method of data augmentation, the structure of the proposed deep network and the loss function.

### 2.1 Data Augmentation

138 clinical fundus fluorescein angiography images for training in our experiment. We horizontally flip, shift and randomly rotate the slices to make data augmentation for each patient.

## 2.2 Architecture

### A. Segment Generator

The generator of this new framework is a U-net. The encoder path has 5 convolution blocks, and each convolution block contains a set of  $3 \times 3$  convolutional layers and a BatchNormalization layer. After the first four convolution blocks, a maxpooling layer extracts the deep features. The decoder path has 5 decoding blocks, and each decoding block first upsamples the feature map. Then, the feature will be fused through the  $3 \times 3$  convolutional layer. At the end of the decoder, a  $1 \times 1$  convolutional layer is used to output a feature map with the same size as the original image and 1 channel, and then the prediction result is obtained through sigmoid. The encoder path and the decoder path are skip-connected.

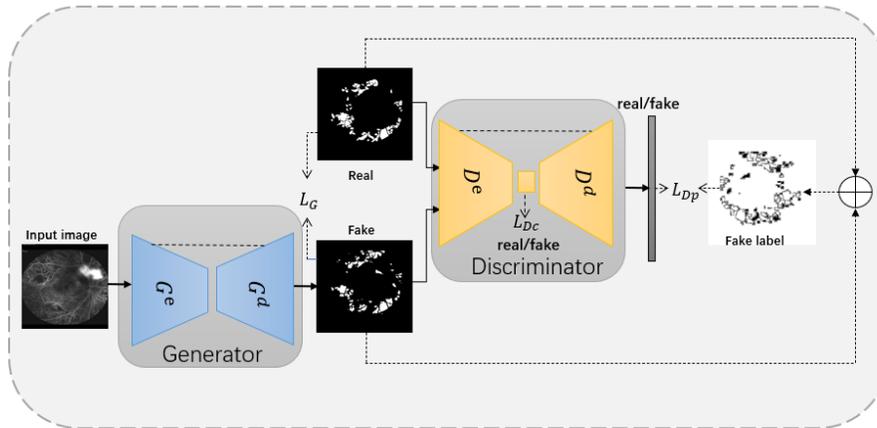


Fig.2. Framework of proposed method for CNP area segmentation.

### B. Pixel Discriminator

The discriminator is an encoder-decoder network shown in Fig.3. The encoder path is a binary classifier. This path has 8 convolutional layers, and the maxpooling layers are used after the first 7 convolutional layers. The output of a 1 channel feature map is computed through the  $3 \times 3$  convolutional layer. The decoder path is up-sampled from the fifth convolutional layer of the encoder, and then passes through a  $3 \times 3$  convolutional layer, and finally produces a result image with the same size and channel as the original image.

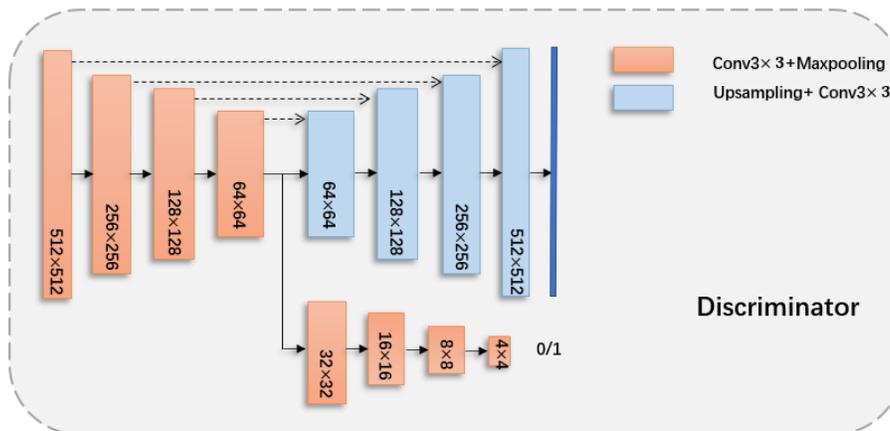


Fig. 3. Framework of proposed Pixel Discriminator.

### 2.3 Loss function

We define our loss function as:

$$L = \alpha L_{bce}(G(X_G), Y_G) + \beta L_{DC} + \gamma L_{DP} \quad (1)$$

$$L_D = L_{DC} + L_{DP} \quad (2)$$

where  $X_G$  is the input of the generator,  $Y_G$  is the segmentation label.  $L_D$  is the loss of discriminator.  $L_{DC}$  and  $L_{DP}$  are the classification loss and pixel loss of the discriminator, respectively.  $\alpha$ ,  $\beta$  and  $\gamma$  are hyperparameters. In this paper, take 10,1,1, respectively.

$$L_{bce} = -\sum_{i=1}^{H \times W} (y_i \log \hat{y}_i + (1 - y_i) \log(1 - \hat{y}_i)) \quad (3)$$

$$L_{DC} = L_{mes}(D_e(X_D), Y_{DC}) \quad (4)$$

$$L_{DP} = L_{mse}(D_a(X_D), Y_{DP}) \quad (5)$$

$$L_{mse} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (6)$$

where  $L_{mse}$  refers to the mean square error loss.  $L_{bce}$  refers to the binary cross entropy loss.  $H$  and  $W$  are the width and height of the input image respectively.  $X_D$  is the input of the discriminator.  $Y_{DC}$  is 1 if the input of discriminator is "real" image and is 0 if input of discriminator is generated image. If the input of discriminator is "real" image,  $Y_{DP}$  will be a mask that all the values are 1. On the contrary, when inputting generated image,  $Y_{DP}$  is the result of segment label xor generated image. It means that we judge all pixel values of the original image to be 'real' and judge the pixel to be 'real' if it is predicted correctly by generator. The formula is as follows:

$$Y_{DP} = \begin{cases} 1 & \text{"real" image} \\ G(X_G) \oplus Y_G & \text{generated image} \end{cases} \quad (7)$$

## 3. RESULTS

### 3.1 Datasets

The proposed method is evaluated on 138 fundus fluorescein angiography images obtained from different patients with CNP. All the image size is  $512 \times 512 \times 3$ , and we preprocess the image into a single-channel grayscale image. In our experiments, we divide 138 images into 107 training images and 31 testing images.

### 3.2 Parameter settings

All the image size is  $512 \times 512 \times 3$ , and we preprocess the image into a single-channel grayscale image. In our experiments, we divide 138 images into 110 training images and 30 testing images. We used Adam algorithm with an initial learning rate of 0.0002 to optimize the weight of the network in the training process. In our experiment, the batchsize was set to 2 and our model was trained for 100 epochs. The models are implemented based on the Keras framework.

### 3.3 Evaluation metrics

The segmentation results are compared to the ground truth according to four metrics: dice similarity coefficient (DSC), sensitivity (SEN), specificity (SPE) and ACC (accuracy).

The calculation of SEN, SPE, ACC is shown as follows:

$$SEN = \frac{TP}{TP+FN} \quad (8)$$

$$SPE = \frac{TN}{FP+TN} \quad (9)$$

$$ACC = \frac{TP+TN}{TP+FP+TN+FN} \quad (10)$$

where TP, TN, FP, FN represents the number of true positive, true negative, false positive, and false negative predictions, respectively.

DSC is a measure of similarity between two regions. Assuming that  $LA_{pred}$  and  $LA_{gt}$  represent the predicted results of this method and the lesion area in the ground truth, respectively, and  $|\cdot|$  means the size of the pixel set. DSC can be calculated as follows,

$$DSC = 2 \frac{|LA_{pred} \cap LA_{gt}|}{|LA_{pred}| + |LA_{gt}|} \quad (11)$$

### 3.4 Experimental results

In order to demonstrate the effectiveness of our approach, we compare our network with some other existing networks. First, we compare the proposed network to the classic segmentation network, FCN8 and Unet. We compare our architecture with the cGAN which is the baseline of our method. To demonstrate the advantages of our proposed method, we also compare it to the cGAN with ASPP block.

As can be seen from Table 1, cGAN achieves better segmentation results than the FCN8 and Unet. After the ASPP block is added, the performance is worst. For the proposed architecture, by using the U shape discriminator, the Dice coefficient reaches  $0.7538 \pm 0.1115$ .

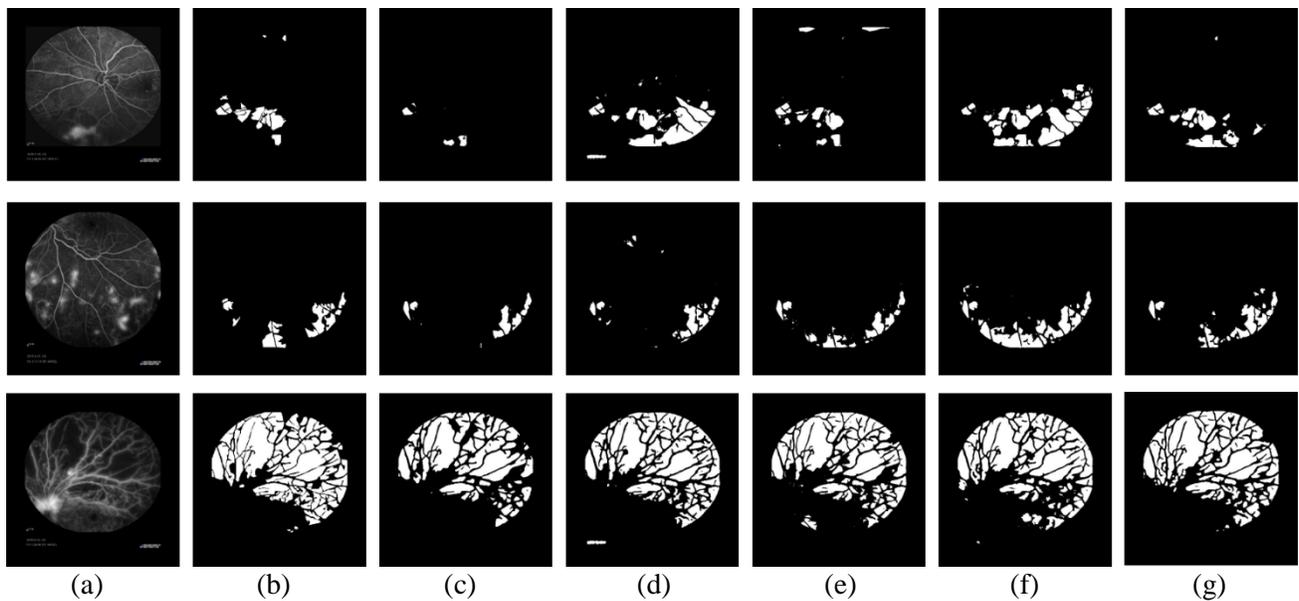


Fig. 4. Examples of segmentation results. (a)Original image. (b)Ground truth. (c)FCN. (d)Unet. (e)cGAN . (f) cGAN+ASPP, (g) our methods.

Figure 4 shows some FFAs and the corresponding segmentation results. It shows that our method can segment the CNP region well. Although the CNP lesions in different images have different sizes, shapes and intensities, our method still can achieve good results. When compared to other methods, segmentation of the boundaries are more accurate.

Table 1 the mean DSC and SEN, SPE and ACC with different methods

	<b>DSC (%)</b>	<b>SEN (%)</b>	<b>SPE (%)</b>	<b>ACC (%)</b>
<b>FCN[3]</b>	70.55±14.65	63.71±19.25	<b>98.60±1.40</b>	94.82±3.02
<b>Unet[4]</b>	68.73±13.93	74.99±11.78	95.62±3.59	93.42±2.79
<b>cGAN [5]</b>	72.16±11.86	75.83±11.25	96.43±2.97	94.23±2.79
<b>cGAN+ASPP[6]</b>	69.27±16.19	77.95±11.73	93.16±3.95	94.81±4.44
<b>Our method</b>	<b>75.38±11.15</b>	<b>79.60±9.94</b>	96.72±2.78	<b>94.89±2.40</b>

#### 4. CONCLUSIONS

In this paper, a new approach for automatic segmentation of CNP is introduced. We propose a new generative adversarial framework, and it can discriminate each pixel of generated image to segment non-perfusion regions. The discriminator in this paper can provide richer information for the generator. Our method achieves more accurate segmentation results compared to state-of-the-art approaches.

#### 5. ACKNOWLEDGEMENTS

This work has been supported in part by the National Natural Science Foundation of China (NSFC) under Grant 61971298, and in part by the National Key R&D Program of China under Grant 2018YFA0701700.

#### 6. REFERENCE

- [1] Jasiobedzki, P. , Mcleod, D. , & Taylor, C. J. . (1991). Detection of non-perfused zones in retinal images. Computer-based Medical Systems, Fourth IEEE Symposium. IEEE.
- [2] Yalin, Z. , Ting, K. M. , Maccormick, I. J. C. , Beare, N. A. V. , Harding, S. P. , & Friedemann, P. . (2014). A comprehensive texture segmentation framework for segmentation of capillary non-perfusion regions in fundus fluorescein angiograms. Plos One, 9(4), e93624-.
- [3] Jayanthi Sivaswamy, Amit Agarwal, Mayank Chawla, Alka Rani, & Taraprasad Das. (2008). Extraction of Capillary Non-perfusion from Fundus Fluorescein Angiogram. Biomedical Engineering Systems and Technologies. Springer Berlin Heidelberg.
- [4] Rasta, S. H. , Nikfarjam, S. , & Javadzadeh, A. . (2015). Detection of retinal capillary nonperfusion in fundus fluorescein angiogram of diabetic retinopathy. Bioimpacts, 5(4), 183-190.
- [5] Kwong, EMT, Zheng, Y, MacCormick, IJ, Beare, NAV, & Harding, SP. (2012). Automated segmentation of capillary non-perfusion (cnp) regions in fundus fluorescein angiograms (fa) using a texture-based approach.
- [6] Goodfellow I J , Pouget-Abadie J , Mirza M , et al. Generative Adversarial Networks[J]. Advances in Neural Information Processing Systems, 2014, 3:2672-2680.
- [7] Phillip Isola, Jun-Yan Zhu , et al. "Image-to-Image Translation with Conditional Adversarial Networks."(2017)
- [8] Khosravan, Naji , et al. "PAN: Projective Adversarial Network for Medical Image Segmentation." (2019).
- [9] Chen, L. C. , et al. "DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs." IEEE Transactions on Pattern Analysis and Machine Intelligence 40.4(2018):834-848.
- [10] Long, Jonathan , E. Shelhamer , and T. Darrell . "Fully Convolutional Networks for Semantic Segmentation." IEEE Transactions on Pattern Analysis and Machine Intelligence 39.4(2015):640-651.
- [11] Ronneberger, Olaf , P. Fischer , and T. Brox . "U-Net: Convolutional Networks for Biomedical Image Segmentation." (2015).