PROCEEDINGS OF SPIE

SPIEDigitalLibrary.org/conference-proceedings-of-spie

SGCNet: a scale-aware and global context network for linear lesion segmentation in MCSL fundus images of high myopia

Luo, Gaohui, Zhu, Qianlong, Chen, Xinjian, Shi, Fei, Fan, Ying, et al.

Gaohui Luo, Qianlong Zhu, Xinjian Chen, Fei Shi, Ying Fan, Jiamin Xie, Qiuying Chen, Lingjiao Pan, Weifang Zhu, "SGCNet: a scale-aware and global context network for linear lesion segmentation in MCSL fundus images of high myopia," Proc. SPIE 11596, Medical Imaging 2021: Image Processing, 115963B (15 February 2021); doi: 10.1117/12.2579889

SPIE.

Event: SPIE Medical Imaging, 2021, Online Only

SGCNet: A scale-aware and global context network for linear lesion segmentation in MCSL fundus images of high myopia

Gaohui Luo[1,#], Qianlong Zhu[1,#], Xinjian Chen[1,2], Fei Shi[1], Ying Fan[3], Jiaming Xie[3], Qiuying Chen[3], Lingjiao Pan[4], Weifang Zhu[1,*]

[1]School of Electronics and Information Engineering, Soochow University, Suzhou, Jiangsu Province, 215006, China
[2] State Key Laboratory of Radiation Medicine and Protection, Soochow University, Suzhou, 215123, China
[3] Shanghai General Hospital, Shanghai 200080, China
[4] School of Electrical and Information Engineering, Jiangsu University of Technology, Changzhou, Jiangsu Province, 213000, China

ABSTRACT

At present, high myopia has become a hot spot for eye diseases worldwide because of its increasing prevalence. Linear lesion is an important clinical signal in the pathological changes of high myopia. ICGA is considered to be the "Ground Truth" for the diagnosis of linear lesions, but it is invasive and may cause adverse reactions such as allergy, dizziness, and even shock in some patients. Therefore, it is urgent to find a non-invasive imaging modality to replace ICGA for the diagnosis of linear lesions. Multi-color scanning laser (MCSL) imaging is a non-invasive imaging technique that can reveal linear lesion more richly than other non-invasive imaging technique such as color fundus imaging and red-free fundus imaging and some other invasive one such as fundus fluorescein angiography (FFA). To our best knowledge, there are no studies focusing on the linear lesion segmentation based on MCSL images. In this paper, we propose a new U-shape based segmentation network with multi-scale and global context fusion (SGCF) block named as SGCNet to segment the linear lesion in MCSL images. The features with multi-scales and global context information extracted by SGCF block are fused by learnable parameters to obtain richer high-level features. Four-fold cross validation was adopted to evaluate the performance of the proposed method on 86 MCSL images from 57 high myopia patients. The IoU coefficient, Dice coefficient, Sensitivity coefficient and Specialty are 0.494 ± 0.109, 0.654 ± 0.104, 0.676 ± 0.131 and 0.998 ± 0.002, respectively. Experiment results indicate the effectiveness of the proposed network.

KEYWORDS: multi-color scanning laser (MCSL) imaging, linear lesion segmentation.

1. INTRODUCTION

In many developed countries, pathological myopia is the main cause of blindness [1]. T. Tokoro et al. [2] proposed that high myopia with visual dysfunction can be defined as pathological myopia. According to [3, 4], linear lesions (indicated by the yellow arrows in Fig. 1) are important clinical signs for evaluating the development of high myopia to pathological myopia. At present, indocyanine green angiography (ICGA) (Fig.1 (a) and (c)) is considered to be the "Ground Truth" for the clinical diagnosis of linear lesion, but it requires injection of contrast agent indocyanine green (ICG) that may cause adverse reactions such as allergy, dizziness, and even shock. Therefore, it is urgent to find a non-invasive imaging modality to replace ICGA for the diagnosis of linear lesions. Multi-color scanning laser (MCSL) imaging (shown in Fig. 1 (b) and (d))[1] is a non-invasive imaging technique which can reveal linear lesions more richly than other non-invasive modality such as color fundus imaging and red-free fundus imaging and some other invasive one such as fundus fluorescein angiography (FFA). Therefore, we try to explore whether MCSL can replace ICGA as a non-invasive imaging technology for the diagnosis of linear lesions. In order to achieve this goal, the segmentation of linear

*Corresponding author: Weifang Zhu, E-mail: wfzhu@suda.edu.cn.

#indicates these authors contributed equally to this work.

Medical Imaging 2021: Image Processing, edited by Ivana Išgum, Bennett A. Landman, Proc. of SPIE Vol. 11596, 115963B · © 2021 SPIE · CCC code: 1605-7422/21/$21 · doi: 10.1117/12.2579889

lesion in MCSL images is necessary and important. As can be seen from Fig. 1, the automatic linear lesion segmentation in MCSL fundus images is a big challenge due to two aspects: (1) Compared with ICGA images, linear lesions are less obvious in MCSL images and is greatly disturbed by the background. (2) The shapes of linear lesions are irregular and their bounders are very blurred.

Jiang et al. [5] proposed an improved conditional generative adversarial (cGAN) network for linear lesion segmentation in ICGA images and achieved good performance. But there still are some drawbacks such as the high complexity of the network model and too many hyperparamters. Feng et al. [6] designed the CPFNet with a SAPF block to segment the linear lesion in ICGA images. However, the CPFNet only fused multi-scale information without acquiring global context information, which plays an extremely important role in the field of image segmentation. Global context information can represent the relationship between each pixel and all pixels, so that it can help the network segment the target more accurately. Cao et al. [7] designed the global context (GC) block, which is lightweight and can effectively model the global context. To the best of our knowledge, linear lesion segmentation in MCSL images has not been reported yet. In this paper, we proposed a new U-shape based segmentation network with a multi-scale and global context fusion (SGCF) block named as SGCNet to segment the linear lesion in MCSL images.

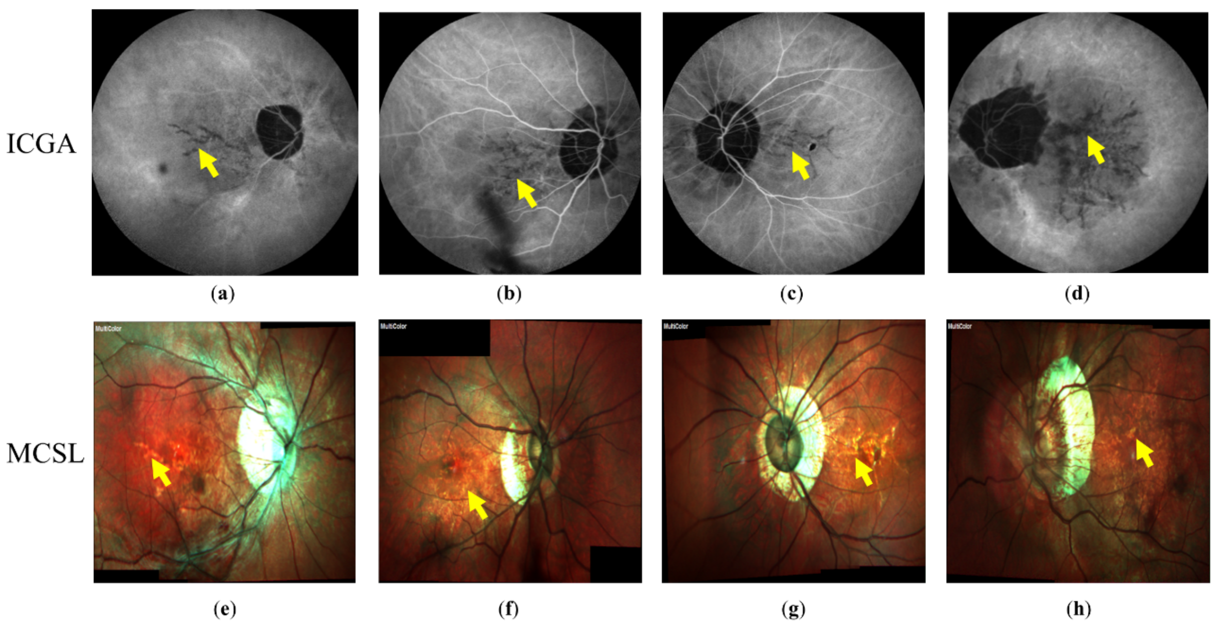

Fig. 1. ICGA images and MCSL images. (a), (b), (c) and (d) ICGA image with linear lesions. (e), (f), (g) and (h) are the corresponding MCSL images for (a), (b), (c) and (d) respectively. The yellow arrow indicates linear lesions.

2. METHODS

We introduce the proposed method in the following three parts: the structure of the proposed SGCNet network, the multi-scale and global context fusion (SGCF) block and the loss function.

2.1 Network structure of SGCNet

In recent years, U-shape network has achieved excellent performance in the field of medical image segmentation. Specifically, Attention U-Net [8] with attention gate (AGs) can implicitly learn to suppress irrelevant regions in an input image and highlight salient features which are useful for a specific task. Our baseline network is improved based on Attention U-Net. We add SE block [9] to the encoder and decoder blocks of Attention U-Net, which can improve the utilization rate of feature channels. As shown in Fig. 2, we have added the SGCF block proposed in this paper on this baseline to form SGCNet.

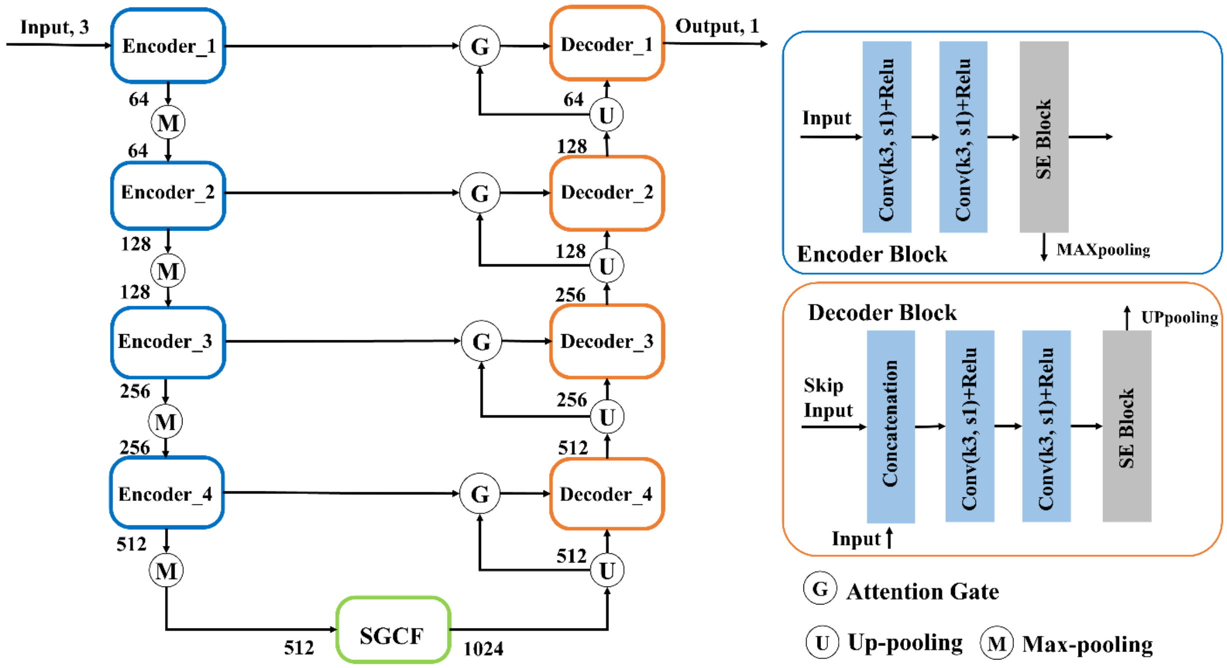

Fig 2. An overview of the SGCNet.

2.2 Multi-scale and global context fusion (SGCF) block

As shown in Fig. 3, the proposed SGCF block consists of two parts: global context block and information fusion module. The global context block is designed based on Non-local Net [10], which greatly reduces the amount of calculation and parameters. The information fusion block uses self-attention mechanism [11] to fuse the input feature information of three different scales. The specific structure of SGCF block is described as follows.

Firstly, the input features are obtained by the global context block for long distance dependence. Secondly, three different expansive convolutions (three different dilated rates, d=1, d=2, and d=4) are used to obtain multi-scale context. Then the feature information with global context and multi-scale context is obtained by the feature fusion block. Finally, global and multi-scale context feature information and the input feature information are fused by learnable parameter weights.

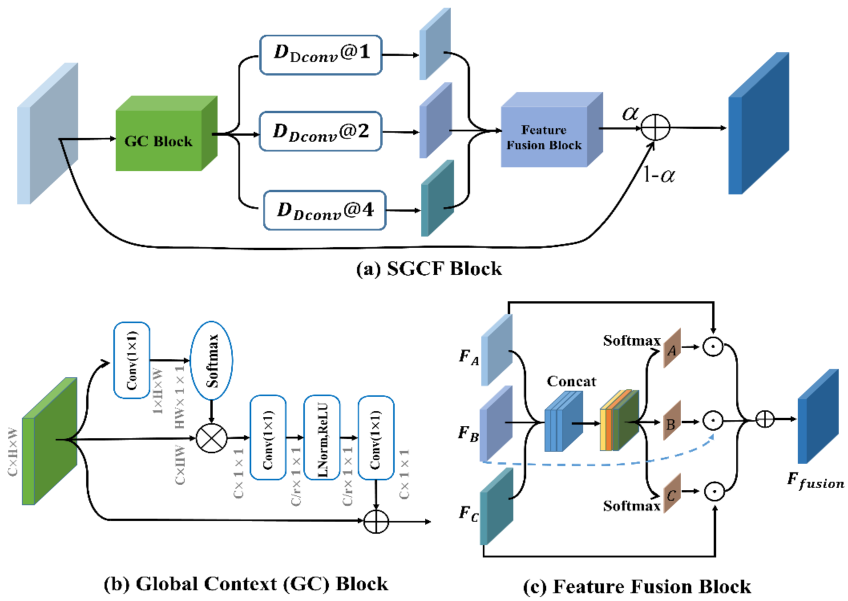

Fig 3. Multi-scale and global context fusion (SGCF) block. ⊗ denotes matrix multiplication, ⊕ denotes broadcast element-wise addition, and ⊙ denotes element-wise product operations.

2.3 Loss function

As can be seen from Fig.1, linear lesions usually occur in the small regions between macular and optic disk, which may lead to an imbalance problem between the target and background. To solve this problem, Dice loss is used in this paper. The binary cross entropy (BCE) is usually used in neural networks as a loss function for single class segmentation tasks. We finally employ a joint loss L_{total} consisting of Dice loss L_{dice} and binary cross-entropy loss L_{bce}, which is described as:

$$L_{total} = L_{bce} + L_{dice} \tag{1}$$

$$L_{bce} = -\frac{1}{n}\sum_{i}^{n}\left[t_i \ln o_i + (1-t_i)\ln(1-o_i)\right] \tag{2}$$

$$L_{dice} = 1 - \frac{1}{n}\sum_{i=1}^{n}\frac{2t_i o_i}{t_i^2 + o_i^2} \tag{3}$$

Where t represent the target pixel value in the ground truth, o represents the value output by the network, n represents the total number of pixels in the image, and i represents the ith pixel.

3. RESULTS

3.1 Datasets

The collection and analysis of image data were approved by the Institutional Review Board of Shanghai General Hospital and adhered to the tenets of the Declaration of Helsinki. Totally 86 MCSL images from 57 high myopia patients were included in our experiments, which were collected and authorized from Shanghai General Hospital. The ground truth of the linear lesion is manually labeled under the supervision of the experienced ophthalmologists. The heights of the original MCSL images are fixed at 496, while the widths vary from 596 to 960 because of different view angles. To facilitate image processing, all images are resampled to (496, 496). Online data augmentation was applied during training, including rotation, flipping and scaling. Four-fold cross validation was adopted to evaluate the performance of the proposed method. All data were randomly split into four parts according to the subjects, which contain 22, 22, 21 and 21 images.

3.2 Implementation details

All the experiments were implemented with PyTorch on a Linux server. The networks were trained on a single NVIDIA Tesla K40m GPU with 10 GB RAM. The initial learning rate is set to 1e-3 with the SGD optimizer and Poly strategy. Batch size is set as 4.

3.3 Experimental results

We choose the following three evaluation metrics to evaluate our network: Jaccard index, Dice coefficient, Sensitivity and Specialty. Paired Wilcoxon signed-rank tests (significance level $\alpha_H = 0.05$) are applied to compare medians of the segmentation results between different methods.

Table 1 the mean Jaccard, Dice, Sensitivity and Specialty compared with different methods

Method	**Result**			
	Jaccard	**Dice**	**Sensitivity**	**Specialty**
UNet	0.452±0.121	0.614±0.125	0.636±0.002	0.998±0.002
Attention-UNet	0.469±0.112	0.630±0.108	0.658±0.148	0.998±0.002
TiramisuNet	0.324±0.157	0.467±0.186	0.532±0.225	0.997±0.003
CENet	0.483±0.117	0.642±0.114	0.665±0.142	0.998±0.002
PSPNet	0.341±0.086	0.503±0.095	0.490±0.115	0.997±0.002
CPFNet	0.475±0.109	0.636±0.103	0.670±0.134	0.998±0.002
Baseline	0.479±0.106	0.641±0.099	0.656±0.122	0.998±0.002
Baseline +GC [7]	0.482±0.111	0.643±0.108	0.663±0.143	0.998±0.002
Baseline +SAPF [6]	0.482±0.119	0.641±0.115	**0.678±0.134**	0.997±0.003
SGCNet	**0.491±0.106**	**0.651±0.098**	0.676±0.131	**0.998±0.002**

As can be seen from Table 1, our baseline is better than U-Net and Attention U-Net on Jaccard, Dice, Sensitivity and Specialty indicators. When Multi-scale and global context fusion (SGCF) block is added, the performance is further improved, and the indicators of Jaccard, Dice, Sensitivity and Specialty reach 0.491, 0.651, 0.678 and 0.998 respectively. The SGCNet significantly outperforms the Baseline on metrics of Jaccard, Dice and Sensitivity with p-values < 0.05. The results indicate that the proposed SGCF block can fuse multi-scale and global context feature information effectively and improve the performance of linear lesion segmentation in MCSL images significantly. The results also indicate that the proposed SGCF block has a better effect than SAPF block [6] and GC block [7] on linear lesion segmentation in MCSL images.

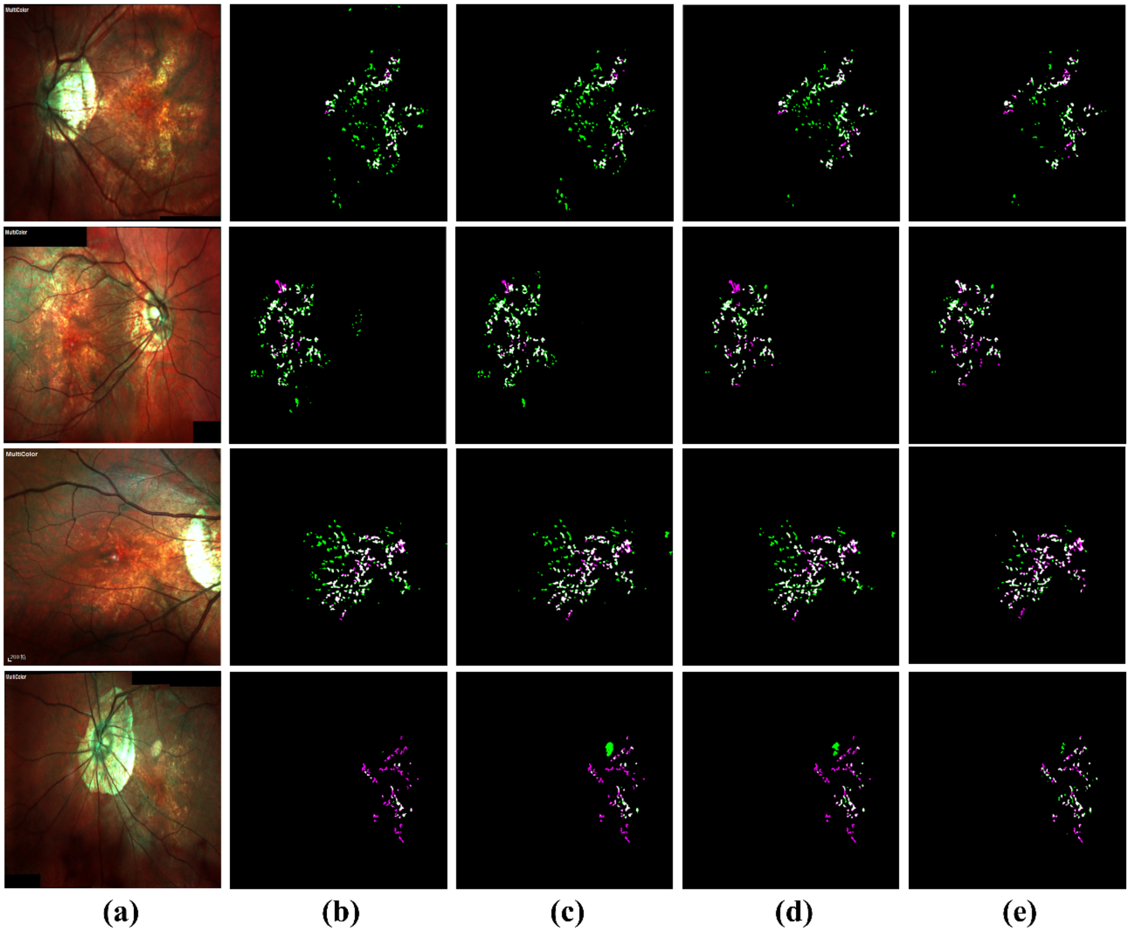

Fig 4. Linear lesion segmentation results. (a)Raw image. (b)UNet. (c)Attention UNet. (d)Baseline. (e)SGCNet. The white region is the overlap of ground truth and network prediction, and lime and magenta regions represent the ground truth and network prediction respectively.

4. CONCLUSIONS

In this paper, we propose a new SGCNet network for the linear lesion segmentation in MCSL images, which can improve the utilization of feature and extract the integrated multi-scale and global context feature information. Experiments indicate the effectiveness of the proposed network.

There are still some shortcomings in this paper: (1) The quantity of the experimental dataset is insufficient, which only includes 86 MCSL images from 57 high myopia patients. The generalization of the proposed SGCNet network can be improved by increasing the amount of dataset, especially the quantity of data without linear lesions. (2) Because of the irregular shape and small area of linear lesion, it is very challenging to automatically segment linear lesion only in MCSL image. Therefore, we will try to combine the MCSL images and some other modality images such as optical coherence tomography (OCT) and infrared fundus images to automatically segment the linear lesion in the future work. As mentioned above, ICGA is invasive and a part of patients may suffer from allergic reactions. To solve this problem, our team focuses the study on the possibility evaluation for the replacement between non-invasive MCSL imaging and invasive ICGA imaging in linear lesion diagnosis and analysis. In our previous related research [5], an improved cGAN based framework was proposed to segment linear lesions in ICGA images. In research [12], we proposed a two-stage registration framework for the registration of ICGA and MCSL images. With these previous studies and the further improvement of the proposed work of linear lesion segmentation in MCSL images, we will evaluate the possibility of non-invasive linear lesion diagnosis in the near future.

5. ACKNOWLEDGEMENTS

This work was supported in part by the National Key Research and Development Program of China under Grant 2018YFA0701700, in part by the National Nature Science Foundation of China under Grant 61622114 and 62001196, and in part by the National Basic Research Program of China under Grant 2014CB748600.

6. REFERENCE

[1] Y. Tano, "Lix Edward Jackson memorial lecture - Pathologic myopia: Where are we now?," American Journal of Ophthalmology 134, 645-660 (2002).

[2] T. Tokoro, "On the definition of pathologic myopia in group studies," Acta Ophthalmologica 66, 107-108 (1988).

[3] K. Ohno-Matsui, T. Yoshida, S. Futagami, K. Yasuzumi, N. Shimada, A. Kojima, T. Tokoro, and M. Mochizuki, "Patchy atrophy and lacquer cracks predispose to the development of choroidal neovascularisation in pathological myopia," British Journal of Ophthalmology 87, 570-573 (2003).

[4] X. Xu, Y. Fang, K. Uramoto, N. Nagaoka, K. Shinohara, T. Yokoi, N. Tanaka, and K. Ohno-Matsui, "Clinical features of lacquer cracks in eyes with pathologic myopia," Retina 39, 1265-1277 (2019).

[5] H. Jiang, X. Chen, F. Shi, Y. Ma, D. Xiang, L. Ye, J. Su, Z. Li, Q. Chen, and Y. Hua, "Improved cGAN based linear lesion segmentation in high myopia ICGA images," Biomedical Optics Express 10(5), 2355-2366 (2019).

[6] S. Feng, H. Zhao, F. Shi, X. Cheng, M. Wang, Y. Ma, D. Xiang, W. Zhu, and X. Chen, "CPFNet: Context Pyramid Fusion Network for Medical Image Segmentation," IEEE Transaction Medical Imaging 39, 3008-3018 (2020).

[7] Y. Cao, J. Xu, S. Lin, F. Wei and H. Hu, "GCNet: Non-Local Networks Meet Squeeze-Excitation Networks and Beyond," 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW), 2019.

[8] O. Oktay, J. Schlemper, L. L. Folgoc, M. Lee, M. Heinrich, K. Misawa, K. Mori, S. McDonagh, N. Y. Hammerla, and B. Kainz, "Attention u-net: Learning where to look for the pancreas," arXiv:1804.03999 (2018).

[9] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2018.

[10] X, Wang, et al. "Non-local neural networks," Proceedings of the IEEE conference on computer vision and pattern recognition, 2018.

[11] Vaswani, Ashish, et al. "Attention is all you need." Advances in neural information processing systems 30, 5998-6008 (2017).

[12] Luo G, Chen X, Shi F, Peng Y, Xiang D, Chen Q, Xu X, Zhu W, Fan Y, “Multimodal affine registration for ICGA and MCSL fundus images of high myopia,” Biomedical Optics Express, 11(8), 4443-4457 (2020).