An Efficient Lens Structures Segmentation Method on AS-OCT Images

Guiping Cao[1] Wei Zhao[1] Risa Higashita[2] Jiang Liu[3,4] Wan Chen[5] Jin Yuan[5] Yubing Zhang[1] Ming Yang[1]

***Abstract*— Lens structures segmentation on anterior segment optical coherence tomography (AS-OCT) images is a fundamental task for cataract grading analysis. In this paper, in order to reduce the computational cost while keeping the segmentation accuracy, we propose an efficient segmentation method for lens structures segmentation. At first, we adopt an efficient semantic segmentation network in the work, and used it to extract the lens area image instead of the conventional object detection method, and then used it once again to segment the lens structures. Finally, we introduce the curve fitting processing (CFP) on the segmentation results. Experiment results show that our method has good performance on accuracy and processing speed, and could be applied to CASIA II device for practical applications.**

I. INTRODUCTION

Currently, the Lens Opacities Classification System III (LOCS III) is widely used in the world for cataract grading, which is a standard evaluation protocol that ophthalmologists compare observed images with standard slit lamp images subjectively [1]. Slit lamp images are two-dimensional (2D) only providing limited structures information. Anterior segment optical coherence tomography images (AS-OCT), which are three-dimensional (3D) images providing richer information of the lens internal structures, are thought to be more potential in cataract grading applications [2].

Automatic and accurate lens structures segmentation on AS-OCT images is the basic step of density measurement and the biological parameters calculation of anterior segment, which is significant for quantitative grading of cataract [3]. As shown in Fig.1 A, AS-OCT images can simultaneously show various different types of soft tissue such as lens, cornea and iris. In our previous work [4], we propose a simple but effective pipeline to segment lens structures. The pipeline consists of two steps: the first step is to extract the lens area from original AS-OCT images, and the second step is to segment lens structures based on lens area images, which are obtained from the first step.

Advantages of the lens area extraction are to reduce the noise interference of redundant tissue, so that the computational cost of the lens structures segmentation is reduced simultaneously. Our previous work [4] uses canny edge detector to detect the lens area boundaries with poor accuracy and robustness, which may lead to the wrong segmentation results of lens structures.

The lens structures segmentation on 3D sequence of AS-OCT images has challenges both on accuracy and speed. The work [4] adopts a U-shaped network followed by a shape template to segment nucleus structures of the lens automatically, which has good accuracy on small amount of the data. However, the shape template is unable to cover all kinds of lens structures with different people. In our work [5], we propose a guide-based model (G-MNet) to exploit edge information from multi-scale guided map, and achieve the state-of-the-art segmentation accuracy of lens structures. The main drawback is that the complex network structure makes the speed of algorithms can not meet the actual application requirements.

To improve the speed of the lens structures segmentation on AS-OCT images while keeping the accuracy, we propose an efficient method for automatic lens structures segmentation. The main contributions of this work are as follows:

- For the efficient segmentation algorithm, we adopt ShuffleSeg [6] network to realize excellent performance on the accuracy and speed.
- To reduce the interference of noise and other redundant structures, we treat the lens area detection problem as a segmentation task, which achieves higher area extraction accuracy and robustness.
- Considering that the borderlines of lens structures are smooth and approximately symmetrical, we propose a curve fitting processing method on segmentation results to improve the segmentation accuracy.

II. MATHOD

As shown in Fig.1, we firstly feed an original AS-OCT image into the pipeline to extract the lens area image by the efficient ShuffleSeg network [6]. And then, taking the extracted lens area image as input, we get the lens structures segmentation result by another segmentation network, which is also based on ShuffleSeg network. Finally, we perform curve fitting processing on the boundary line of the segmentation mask to obtain the final segmentation result.

A. ShuffleSeg Network

Based on encoder-decoder framework, we construct a ShuffleSeg network (shown in Fig.2) to segment the lens area and lens structures. The encoder of ShuffleSeg uses ShuffleNet unit (shown in Fig.3) in three convolution stages, which are composed of 3, 7 and 3 ShuffleNet units respectively. ShuffleNet unit is a residual bottleneck module with

[1]G. Cao, W. Zhao, Y. Zhang and M. Yang are with the CVTE Research, Guangzhou, China. {caoguiping, zhaowei, zhangyubing, yangming}@cvte.com

[2]R. Higashita is with Tomey Corporation, Japan. lisahigashita@gmail.com

[3,4]J. Liu is with Southern University of Science and Technology and Cixi Institute of Biomedical Engineering, Chinese Academiy of Sciences, China. liuj@sustech.edu.cn

[5]W. Chen and J. Yuan are with Zhongshan Ophthalmic Center, Sun Yat-sen University. yeah-cw@126.com yuanjincornea@126.com

978-1-7281-1990-8/20/$31.00 ©2020 IEEE

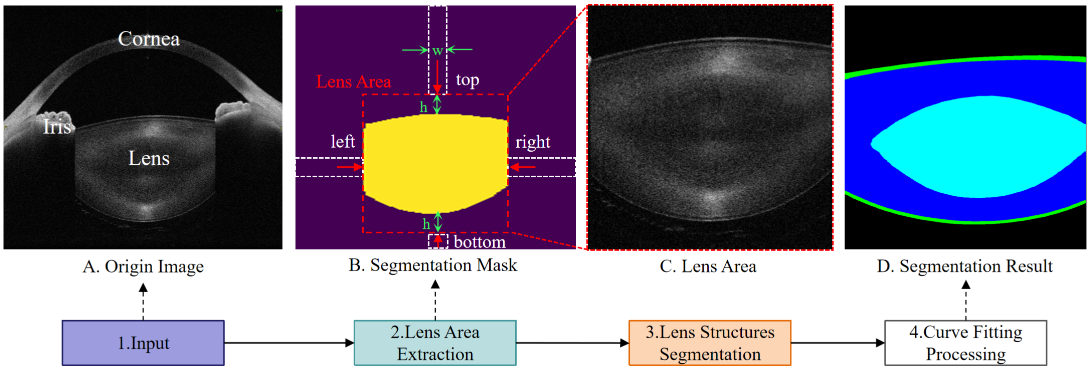

Fig. 1. The proposed pipeline: 1. Original AS-OCT image. 2. Lens area extraction. Segmentation mask is a binary image of lens area. 3. Segmentation of lens structures. 4. The curve fitting processing on the segmentation results (green: capsule, blue: cortex, cyan: nucleus).

average pooling (AVG Pool) of kernel size 3×3 as residual connection, and uses depthwise separable convolution (DWC) and grouped convolution (GC) to reduce the computational cost and maintain good representation capability. The decoder of ShuffleSeg is based on skip connections to benefit from higher resolution feature maps to improve accuracy, and the upsampling is achieved by transposed convolution.

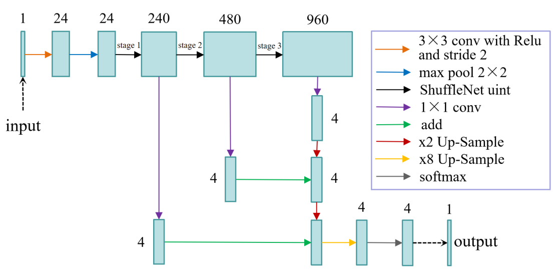

Fig. 2. The architecture of ShuffleSeg network, and the number (like 480) indicates the data channels.

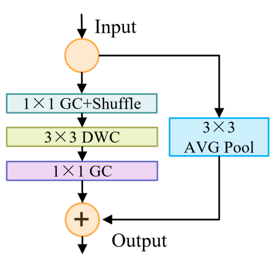

Fig. 3. The architecture of ShuffleNet unit.

B. Lens Area Extraction and Lens Structures Segmentation

We first extract the lens area images from original AS-OCT images by the segmentation method, and then feed it into the segmentation network to realize the lens structures segmentation.

To reduce the interference of noise and other excess structures, we treat the object detection problem of lens area extraction as a segmentation task. Taking the original AS-OCT image as input, we adopt the ShuffleSeg network to get the lens area segmentation mask, which is a binary image and to be resized into the same size with the original AS-OCT image. As shown in Fig.1 B and C, the coordinates (left, right, top and bottom) of the lens area boundaries are searched from the segmentation mask, and the lens area image is extracted by the coordinates from the original AS-OCT image.

To get the lens area boundary coordinates, we search the boundaries by detecting non-zero pixel value coordinates from the binary image, and the red arrows in Fig.1 B indicate the search direction. In order to improve the robustness of boundary searching algorithm and reduce the noise pixels influence, we update the center pixel value by using the mode pixel value of its neighboring pixels, which is in the vertical direction of the search direction. As shown in Fig.1 B, the width w of the white dotted rectangle represents the number of neighboring pixels. Meanwhile, to ensure that the extracted lens area image contains integral lens structures, we extend the top and bottom boundary coordinates with height h.

After the lens area images extracted, we feed these images into another ShuffleSeg network, in which the segmentation class is set to four (capsule, cortex, nucleus and background), to get lens structures segmentation result.

C. Curve Fitting Processing

We find that the borderlines of lens structures are smooth and approximately symmetrical, and this structural features may have some advantages on the segmentation accuracy improving. Inspired by this phenomenon, we extract the boundary points of capsule, cortex and nucleus from the lens structures segmentation results and perform curve fitting processing on each boundary line. The curve fitting processing is realized by lagrange interpolation with borderline points $p = \{x, y\}$:

$$\hat{y} = \sum_{j=0}^{n}(y_i \prod_{i=0, i \neq j}^{n} \frac{x - x_i}{x_j - x_i}) \tag{1}$$

$\hat{y}$ is obtained by lagrange interpolation. i and j are the pixel index. n is the number of the points used for interpolation, and the default value is 5.

III. EXPERIMENTS

Dataset. AS-OCT image data is acquired by CASIA II device of Tomey Corporation, Japan. The original AS-OCT image size is 2130×1864, and the width and height of the pixel size are approximately 0.015 *mm* and 0.008 *mm* respectively. The dataset of the lens area extraction and the lens structures segmentation is labeled by experienced ophthalmologists respectively. The dataset contains 2298 images. Referencing [5], we randomly select 1711 images for training and 587 images for testing, and the images of training set and the testing set are from different people respectively.

Implementation details. For the lens area segmentation, we resize the original AS-OCT images into 120×240. We set the *w* to 40 pixel, and set *h* to 100 pixel to search the boundary coordinates.

For lens structures segmentation, we resize the lens area images into 256×256 for segmentation network. For speed test, we take the average segmentation time of 1000 images as the algorithm running time in the environment: Core i-5-8250U 1.6GHz, RAM 16.0GB, without GPU.

Evaluation metrics. For lens area extraction, the accuracy is measured by the average distance error (ABDE) of the left and right boundary coordinates. ABDE is defined as:

$$ABDE = \frac{1}{n}\sum_{i=1}^{n} |x_i - x'_i| \tag{2}$$

n is the number of images. i reprents the ith image. x_i and x'_i denote the left or right border coordinates of the lens area in predicted result and ground truth respectively.

Following the previous work [5] in AS-OCT images segmentation, we employ the normalized mean squared error (NMSE) to measure the segmentation accuracy of lens structures. NMSE (defined in eq.3) is calculated by the ground truth $S_{gt} = \{x_i, y_i\}$ and predicted result $S_p = \{\hat{x}_i, \hat{y}_i\}$, where S_{gt} and S_p are the coordinate position sets of the boundary points, and n_p denotes the number of the annotation boundary points.

$$NMSE = \frac{1}{n_p}\sum_{i=1}^{n_p} \sqrt{(\hat{x}_i - x_i)^2 + (\hat{y}_i - y_i)^2} \tag{3}$$

Results. **Lens area extraction:** TABLE I compares our method with common detection network yolov2 [7], canny edge detection [4] and threshold techniques [8]. Experiment results show that our lens area extraction method has the best performance in both mean and standard deviation of the ABDE. Our approach reduces errors of the ABDE over 3.5 pixels than yolov2 and 38 pixles than canny detection method. In terms of algorithm speed, compared with yolov2 and canny, our method can save 28.13% and 96.17% of the time respectively.

Lens structures segmentation: We compare the performance of our method with several state-of-the-art segmentation methods in TABLE II. Although the mean and standard deviation of NMSE with our proposed method is slightly lower than our previous work [5], it is better than other methods. In addition, our method has an absolute advantage in algorithm speed that it can save 94.86% of the time compared with the method [5].

TABLE I
ABDE AND SPEED OF LEFT AND RIGHT BORDER FOR LENS AREA EXTRACTION (ABDE UNIT:PIXEL, TIME UNIT:MS)

Method	Left	Right	Time
Yolov2 [7]	6.15 ± 5.48	5.88 ± 5.11	173.62
Canny [4]	40.35 ± 49.06	43.06 ± 48.52	3254.79
Threshold [8]	8.69 ± 34.79	8.39 ± 31.79	**4.63**
Ours	**2.38 ± 2.65**	**2.39 ± 3.02**	124.78

TABLE II
NMSE AND SPEED OF LENS STRUCTURES SEGMENTATION (NMSE UNIT:PIXEL, TIME UNIT:MS)

Method	Capsule	Cortex	Nucleus	Time
FCN-VGG16 [9]	3.08±4.84	3.34±3.14	11.03±4.08	349.07
DeepLabV2 [10]	3.97±4.08	6.18±4.31	10.88±8.04	2325.46
PSPNet-Res34 [11]	1.37±0.96	1.73±0.75	8.20±3.97	496.22
M-Net [12]	1.37±2.62	1.60±0.93	7.93±3.65	1992.41
U-Net [13]	1.32±1.14	1.49±1.20	8.54±4.11	2159.68
U-shaped Net [4]	——	——	18.10±6.73	2940.36
G-MNet [5]	**0.57 ± 0.29**	**0.97 ± 0.60**	**7.45 ± 3.24**	3558.88
Ours (No CFP)	0.96 ± 0.59	1.45±0.68	8.23±3.33	**183.09**
Ours	0.79 ± 0.66	1.43 ± 1.38	7.95 ± 3.26	184.62

IV. DISCUSSION

Lens area extraction: The drawback of canny is that it is easily interfered by noise pixels (the standard deviation of ABDE is close to 50 pixel). Although the threshold technique [8] only takes 4.63ms, it is susceptible to noise and does not use spatial position information, which leads to low accuracy and extremely poor algorithm stability. For object detection method [7], as shown in Fig.4 B, the main problem is that the misdetected structures such as iris may lead to the incorrect segmentation of the lens edge structures.

In AS-OCT images, the left and right boundaries of the lens area are very clear, and its binary segmentation results generally do not include other interfering tissues. Based on the characteristics mentioned above, we first transform the detection problem into a segmentation task, and then obtain the boundary coordinates from the segmented mask. Experiment results show that our method has advantages in accuracy, robustness and speed, which means searching for the lens boundary on the binary segmentation result is more accurate than detecting the boundary on the original AS-OCT images directly.

Lens structures segmentation: Segmentation accuracy and speed are both important, and the latter is even more significant in 3D sequence AS-OCT images. It is very meaningful that the segmentation speed can be improved while keeping the segmentation accuracy.

In terms of accuracy, [11]–[13] show similar performance that they exploit the multi-scale information through various network structures. The G-MNet [5] gets higher-quality segmentation results by incorporating guided filter into CNNs to learn better features. However, as the most time-consuming method, G-MNet needs to construct and calculate multiple

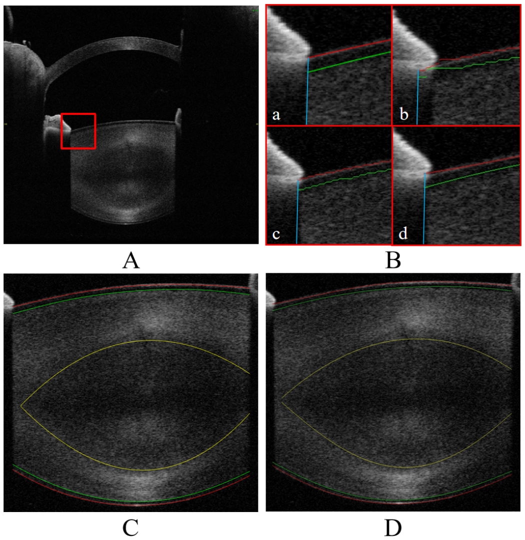

Fig. 4. One example of lens area extraction and CFP of the lens structures segmentation results (LSSR). A: Original AS-OCT images. B: LSSR of the red box (RB). a: Ground truth. b: Lens area is extracted by [7], which is extracted incorrectly with noise tissues and LSSR is wrong. c: Lens area is extracted by our method, and LSSR is better than b. d: Lens area is extracted by our method, and LSSR with CFP is better than b and c. C: Ground truth of A. D: LSSR of our method with CFP.

filters and multi-scale images to learn more features, which will greatly increase the burden of computing resources.

Inspired by the FCN [9] and ShuffleNet [14] network, the ShuffleSeg network improves segmentation accuracy by adopting the residual bottleneck modules in encoding part and the skip connection in decoding part. To improve the segmentation speed, ShuffleNet unit (Fig.3) uses DWC, GC and channel shuffling to reduce the computational cost, while maintaining good representation capability. Experiments show that our method achieves the excellent segmentation accuracy and speed, and the method can be easily applied to other medical images.

Curve Fitting Processing: In TABLE II, compared with no CFP method, our method with CFP can efficiently improve the lens structures segmentation accuracy. From Fig.4 B (b, c and d), we also can see that the curve fitting processing has advantages to smooth sawtooth shapes and improve accuracy of the segmentation results. However, there are a few images not suitable to process with CFP. For example, the disappearance of some structures (Fig.5 B) and incorrect region segmentation (Fig.5 C) will make that the curve fitting process cannot be performed successfully.

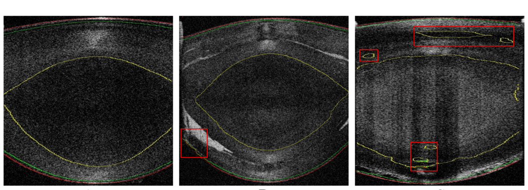

Fig. 5. Lens structures segmentation results with different quality. A: Normal. B: Missing lens posterior capsule structure (shown in the red box). C: Incorrect segmentation result of the lens nucleus.

V. CONCLUSIONS

In this paper, we propose an efficient method that it can reduce the computational cost while keeping the segmentation accuracy in the lens structures segmentation on AS-OCT images. We adopt efficient ShuffleSeg network to realize lens area extraction and lens structures segmentation. Compared with existing methods, our method has good precision and excellent speed advantages and could be convenient applied to the CASIA II device for practical applications. In the future, we will use more data to verify the effectiveness of the method, and explore the advantages of different curve fitting processing algorithms on segmentation results.

ACKNOWLEDGMENT

This work was supported by National Key R&D Program of China #2017YFC0112402.

REFERENCES

[1] L. Huiqi, L. Joo Hwee, L. Jiang, M. Paul, T. Ava Grace, W. Jie Jin, and W. Tien Yin, "A computer-aided diagnosis system of nuclear cataract," *IEEE transactions on bio-medical engineering*, vol. 57, no. 7, p. 1690, 2010.

[2] M. Doors, T. T. Berendschot, J. de Brabander, C. A. Webers, and R. M. Nuijts, "Value of optical coherence tomography for anterior segment surgery," *Journal of Cataract & Refractive Surgery*, vol. 36, no. 7, pp. 1213–1229, 2010.

[3] A. L. Wong, C. K. Leung, R. Weinreb, A. K. Cheng, C. Y. L. Cheung, P. T. Lam, C.-P. Pang, and D. S. C. Lam, "Quantitative assessment of lens opacities with anterior segment optical coherence tomography," *British Journal of Ophthalmology*, vol. 93, no. 1, pp. 61–65, 2009.

[4] P. Yin, M. Tan, H. Min, Y. Xu, G. Xu, Q. Wu, Y. Tong, H. Risa, and J. Liu, "Automatic segmentation of cortex and nucleus in anterior segment oct images," in *Computational Pathology and Ophthalmic Medical Image Analysis*. Springer, 2018, pp. 269–276.

[5] S. Zhang, Y. Yan, P. Yin, Z. Qiu, W. Zhao, G. Cao, W. Chen, J. Yuan, R. Higashita, Q. Wu *et al.*, "Guided m-net for high-resolution biomedical image segmentation with weak boundaries," in *International Workshop on Ophthalmic Medical Image Analysis*. Springer, 2019, pp. 43–51.

[6] M. Gamal, M. Siam, and M. Abdel-Razek, "Shuffleseg: Real-time semantic segmentation network," *arXiv preprint arXiv:1803.03816*, 2018.

[7] J. Redmon and A. Farhadi, "Yolo9000: better, faster, stronger," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 7263–7271.

[8] S. S. Al-Amri, N. V. Kalyankar *et al.*, "Image segmentation by using threshold techniques," *arXiv preprint arXiv:1005.4020*, 2010.

[9] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 3431–3440.

[10] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs," *IEEE transactions on pattern analysis and machine intelligence*, vol. 40, no. 4, pp. 834–848, 2017.

[11] H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia, "Pyramid scene parsing network," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 2881–2890.

[12] H. Fu, J. Cheng, Y. Xu, D. W. K. Wong, J. Liu, and X. Cao, "Joint optic disc and cup segmentation based on multi-label deep network and polar transformation," *IEEE transactions on medical imaging*, vol. 37, no. 7, pp. 1597–1605, 2018.

[13] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.

[14] X. Zhang, X. Zhou, M. Lin, and J. Sun, "Shufflenet: An extremely efficient convolutional neural network for mobile devices," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 6848–6856.