

AUTOMATIC ANGLE-CLOSURE GLAUCOMA SCREENING BASED ON THE LOCALIZATION OF SCLERAL SPUR IN ANTERIOR SEGMENT OCT

Panming Li¹, Le Geng¹, Weifang Zhu¹, Fei Shi¹, Xinjian Chen^{1,2†}

¹School of Electronic and Information Engineering, Soochow University, Suzhou, China

²State Key Laboratory of Radiation Medicine and Protection, Soochow University, Suzhou, China

ABSTRACT

As one of the major types of glaucoma, closed-angle glaucoma is the leading cause of irreversible blindness in the world. The ability of Anterior Segment Optical Coherence Tomography (AS-OCT) to obtain high-resolution cross-sectional images of the entire anterior chamber in a single image makes it an important tool for glaucoma diagnosis. In this paper, we propose a practical and efficient system based on deep learning to accurately classify anterior chamber angle (ACA) closure by using the location of scleral spur (SS) points. First, the localization problem is reformulated as a pixel-wise regression task. A fully convolutional deep neural network is optimized to predict the probability that each pixel belongs to the SS points, and the numerical coordinates are obtained by the maximum likelihood estimation theory. Second, the ACA region centered on the detected SS is cropped as the input of the classification model. The single model applied for classification is SE-ResNet18 and optimized with focal loss. In the AGE Challenge 2019[1], our proposed method obtained superior performance for angle-closure glaucoma screening.

Index Terms— AS-OCT, glaucoma screening, scleral spur localization, angle-closure classification

1. INTRODUCTION

Glaucoma is the one of the leading cause of irreversible blindness worldwide. Primary angle-closure glaucoma (PACG) is one of the main types of glaucoma, which is a major form of glaucoma in Asia[1]. PACG is a progressive disease that can lead to acute angle closure, chronic glaucomatous optic nerve damage and blindness if left untreated. Active health examination is necessary for early detection and prevention. Anterior Segment Optical Coherence Tomography (AS-OCT) is an effective imaging modality for visually identifying the anterior segment structure and observing the condition of anterior chamber angle (ACA), which makes it an important diagnostic tool for glaucoma[2]. It is of great importance to develop automatic algorithms to make angle-closure glaucoma screening more convenient and stable.

Many automatic methods have been developed for this task in recent years. Tian et al. presented an algorithm to assess the anterior chamber angle based on the automatically detected Schwalbes line for High-Definition OCT (HD-OCT)[3]. Nongpiur et al. provided a classification algorithm based on stepwise logistic regression that used a combination of 6 clinical parameters obtained from a single horizontal AS-OCT scan[4]. Xu et al. localized the ACA region, and then extracted multiscale HOG features and visual features directly to classify the glaucoma subtype as open angle or angle-closure[5]. Xu et al. then presented a reconstruction-based method, called Similarity Weighted Linear Reconstruction (SWLR), for glaucoma classification from AS-OCT images containing the ACA[6]. Fu et al. put forward a data-driven approach to integrate AS-OCT segmentation, measurement, and screening. Fu et al. then proposed a multi-context deep network (MCDN) architecture to learn discriminative representations on particular regions of different scales. In 2019, Fu et al. detected the ACA region by using sliding windows, and then the proposed multilevel deep network (MLDN) combined global and local level representations for angle-closure detection in AS-OCT images[7].

Inspired by Fu et al.[7], we propose a practical and efficient system based on deep learning for angle-closure glaucoma screening in AS-OCT images. We firstly localize the scleral spur (SS) points, which is the important anatomical marker to provide the position of ACA. Different from the method based on sliding window regression framework proposed in [10], we reformulate the localization problem as a pixel-wise regression task. A fully convolutional deep neural network is optimized to predict the probability that each pixel belongs to the SS points, and the numerical coordinates are obtained based on the maximum likelihood estimation theory. According to clinical knowledge, the region of ACA, determined by the localization result, is cropped as the input of angle-closure classification. The only one model applied for classification is SE-ResNet18 and optimized with focal loss. In the AGE Challenge 2019, our proposed approach obtained superior performance for angle-closure glaucoma screening, with three commonly used classification metrics, including AUC, SEN and SPE, all reaching 1.00. The effectiveness of our proposed method is well demonstrated.

[†] Corresponding author

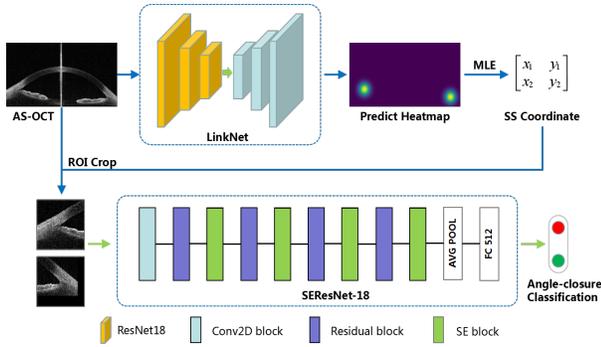


Fig. 1. Illustration of the proposed system for the localization of scleral spur and classification of angle closure.

2. METHOD

The overview of the proposed angle-closure glaucoma screening system is shown in Fig. 1. A U-shape fully convolutional neural network is trained to generate the possibility map for scleral spur detection. Then, a method based on maximum likelihood estimation is applied to obtain the numerical coordinates. According to several medical references, the anterior chamber angle (ACA) region provides key risk indicators for angle closure classification. Therefore, in the classification section, the ACA patch, centered on the detected scleral spur, is cropped as the region of interest (ROI). The classification model outputs the screening result based on the input ACA patch.

2.1. Scleral Spur Localization

Key points localization is a popular topic in computer vision and deep learning[8][9]. In the recent development of deep learning, numerous of algorithms for key points localization, especially in human pose estimation, have been proposed. The localization algorithms implemented by supervised neural network mainly include two forms: one is to directly use the coordinates as the ground truth; the other is to generate heatmap depended on the coordinate position as the ground truth, such as Gaussian map. Compared with the former, the latter reduces the complexity of the problem and is easier to converge in the training process. In addition, the method based on 2D heatmap ground truth can make use of the fully convolutional network for training and prediction. For this, we first generate the ground truth map $G(x, y)$ for each image. Given the coordinate (x_c, y_c) of the scleral spur, $G(x, y)$ is defined as follows:

$$G(x, y) = A \exp \left[-\frac{(x - x_c)^2 + (y - y_c)^2}{2\sigma^2} \right] \quad (1)$$

where A is a parameter used to normalize the values in $G(x, y)$ to $[0, 1]$. σ is the variance governing the spread of

ground truth map across the image domain. With the generated heatmap as ground truth, the localization problem is converted to a probability regression problem.

We employ the LinkNet[10], a typically light U-shape architecture, to learn the transformation from AS-OCT image to the probability map. The encoder part is the pre-trained ResNet-18, which retains the first four extracting blocks without average pooling layer and fully connected layers. Compared with U-net, LinkNet replaces the concatenation with addition operation to realize the skip connection, which will reduce computation and accelerate the training progress. To train the pixel-wise regression network, we utilize the mean square error (MSE) loss to calculate the difference between ground truth and prediction. We apply random data augmentation before training, including adjusting brightness, contrast and sharpness, in order to improve the generalization capabilities of the model. The enhancement factors are all following the log-normal distribution.

There are several methods to convert the output heatmap to the numerical coordinate. The most common way is to get the position of the peak value in the map. However, it is not robust enough and will lead to quantization error of coordinate rounding. In most human pose estimation researches, an offset in the direction from the highest response to the second highest response is used to obtain the final location. Here, we consider the pixel value of the output as the probability density, which ideally is a 2D Gaussian probability density. A method based on maximum likelihood estimation (MLE) theory is developed to obtain the coordinate (\hat{x}_c, \hat{y}_c) from the output heatmap. The formula is defined as follows:

$$\begin{aligned} \hat{x}_c &= \frac{\sum_{i \in C} x_i p_i}{\sum_{i \in C} p_i} \\ \hat{y}_c &= \frac{\sum_{i \in C} y_i p_i}{\sum_{i \in C} p_i} \end{aligned} \quad (2)$$

where C indicates the set containing pixels whose values are higher than the half of the maximum in the heatmap. (x_i, y_i) is corresponding coordinate, and p_i is the value of the i th pixel in the set. The results based on the weighted average operation have less error than finding the peak directly.

2.2. Angle Closure Classification

In the angle closure classification section, the coordinate of scleral spur is utilized to localize the ACA region, which provides key risk indicators for angle closure classification. We crop a 128×128 patch with scleral spur centered as the input of the classification network, shown in Fig. 2. To tolerate the minor localization error, Gaussian noise is added to the real coordinates when cropping patches during training process. The operation also works as a part of data augmentation to make the model more generalized.

To further improve the efficiency of the system, the classification section contains only one network, which is a

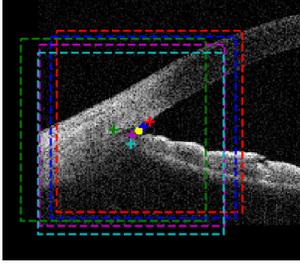


Fig. 2. ROI cropped randomly around ground truth. The SS ground truth is marked by yellow point. ROI centers are marked by colorful crosses.

modified SE-ResNet18, shown in Fig. 1. Experiments in Reference[11] show that integrating SE block into the different position of residual blocks performs similarly well. Here, we insert the SE block between each residual layers of pre-trained ResNet18, without destroying the original residual architecture.

Data imbalance is another difficulty in achieving accurate classification. To overcome this problem, the focal loss is employed as the cost function during training period, which is defined as Equation (3). Here, y is the ground truth of each angle, and \hat{y} is the predicted result. The α we determined is 6.0 and γ is 2.0.

$$L_{focal} = -\alpha y(1 - \hat{y})^\gamma \lg(\hat{y}) - (1 - y)\hat{y}^\gamma \lg(1 - \hat{y}) \quad (3)$$

According to the relevant medical knowledge, the closure status of the left and right anterior chambers in the same AS-OCT are extremely correlated, which is also confirmed in the official training data label of AGE 2019 challenge. Based on such rule, voting mechanism is introduced in the final test, which efficiently increase the accuracy of classification.

3. EXPERIMENTAL RESULTS

3.1. Data description and implementation details

The proposed system is estimated on a public challenge dataset, which is supplied for the challenge named Angle closure Glaucoma Evaluation Challenge(AGE). The dataset provided by AGE contains 4800 annotated AS-OCT images, which are divided equally for training, validation and testing. Classification of angle width was based on gonioscopy by glaucoma experts. Scleral spur localization was determined by the mean of 4 independent annotations from a group of ophthalmologists, followed by a manual adjustment from a senior glaucoma expert. Each RGB image has a size of 2130×998 captured by CASIA I. We find that the blue channel of each image has higher level noise than other two channels. Therefore, we remove the blue channel and convert the raw image as a gray image. In order to ensure the receptive field is sufficient and reduce the computation,

Table 1. Localization results of several baselines and our model on the online validation set. Bold numbers indicate the best performance.

Model	EDE[pixel]
FCN	21.19
RCNN	15.46
U-Net	14.72
Proposed Method	13.14

the modified OCT images are resized to 512×256 as the network inputs. We choose the Average Euclidean Distance Error(EDE) to estimate the localization results. The metrics used for estimating the screening results include area under curve(AUC), sensitivity (SEN) and specificity (SPE).

We implement our screening system using the publicly available Pytorch library. The loss for both localization and classification model is minimized by the Adam optimizer with an initial learning rate 0.0002. Two models included in the system are trained separately on Tesla K40 GPU(12GB). During inference section, it takes only about 200ms and 0.6GB GPU memory to output the final screening result for a single 512×256 AS-OCT image, which indicates that it is easy to deploy the screening system.

This study was approved by the ethical review committee of Zhongshan Ophthalmic Center GuangzhouChina and conducted in accordance with the tenets of the Helsinki Declaration. Written informed consent was obtained from each subject.

3.2. Results

Since the challenge has not released the label of test set, we first evaluate our system offline on the self-defined validation dataset, which is split from the training set. We compare different threshold level to determine the pixel set from coordinate calculation. The metric results show that when the threshold is about half of the maximum value in heatmap, the localization error is the smallest. When estimating the model online, we compare our proposed localization model with several baselines, including U-net, FCN and RCNN. In Table 1, from the quantitative results, our model outperforms others in Average Euclidean Distance. Finally, we achieve a result with Average Euclidean Distance about 13.14 pixel on online validation set (in 2130×998 AS-OCT).

For classification, on offline validation set, common metrics including area under curve, sensitivity and specificity are all 1.00. Due to the models great generalization, these metrics are still 1.00 on the official online validation set. The ablation experiment results for the classification model are shown in Table 2.

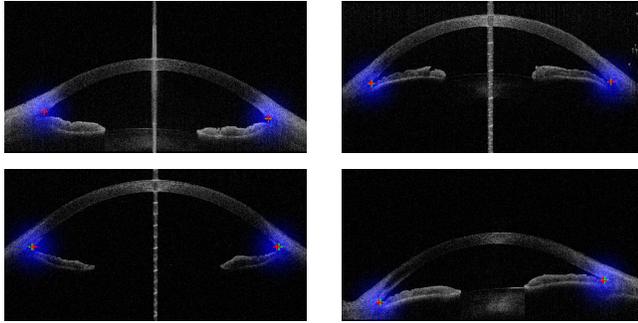


Fig. 3. Samples of localization results. Green cross is the label and red cross is the predicted result. The first column contains two open angles AS-OCT, and those in the second column are angle-closure.

Table 2. Classification results of different methods on online validation set. Bold numbers indicate the best performance.

Method	AUC	SEN	SPE
ResNet18	0.998	0.943	0.996
ResNet18-Focal	0.997	0.981	0.983
ResNet18-Focal-Vote	0.999	0.997	0.993
ResNet18-Focal-Vote-ROI	0.999	1.000	0.998
Proposed Method	1.000	1.000	1.000

4. CONCLUSION

In this work, we propose an automatic system with two deep learning networks for localization of scleral spur and classification of angle closure. Inspired by human pose estimation, we convert the scleral spur localization problem as a regression problem. When obtaining the coordinate from predicted heatmap, a method based on MLE theory is proposed to reduce the transform error. In classification section, we utilize the result of localization to crop a ACA patch as the input of classification model. The closure status at both sides of AS-OCT have high correlation, so a vote mechanism is applied to reduce the classification error. The experiment results indicate that our method performs well on AGE challenge dataset, and it is time-saving and requires little GPU memory. Therefore, it can provide efficient and reliable screening results for ophthalmologist.

5. REFERENCES

[1] Huazhu Fu, Fei Li, Jos Ignacio Orlando, Hrvoje Bogunovi, Xu Sun, Jingan Liao, Yanwu Xu, Shaochong Zhang, and Xiulan Zhang, “Age: Angle closure glaucoma evaluation challenge,” 2019.

[2] Reetika Sharma, Ajay Sharma, Tarun Arora, Sourabh Sharma, Amit Sobti, Bhaskar Jha, Neha Chaturvedi, and Tanuj Dada, “Application of anterior segment optical

coherence tomography in glaucoma,” *Survey of ophthalmology*, vol. 59, no. 3, pp. 311–327, 2014.

[3] Jing Tian, Pina Marziliano, Mani Baskaran, Hong-Tym Wong, and Tin Aung, “Automatic anterior chamber angle assessment for hd-oct images,” *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 11, pp. 3242–3249, 2011.

[4] Monisha E Nongpiur, Benjamin A Haaland, David S Friedman, Shamira A Perera, Mingguang He, Li-Lian Foo, Mani Baskaran, Lisandro M Sakata, Tien Y Wong, and Tin Aung, “Classification algorithms based on anterior segment optical coherence tomography measurements for detection of angle closure,” *Ophthalmology*, vol. 120, no. 1, pp. 48–54, 2013.

[5] Yanwu Xu, Jiang Liu, Jun Cheng, Beng Hai Lee, Damon Wing Kee Wong, Mani Baskaran, Shamira Perera, and Tin Aung, “Automated anterior chamber angle localization and glaucoma type classification in oct images,” in *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2013, pp. 7380–7383.

[6] Yanwu Xu, Jiang Liu, Damon Wing Kee Wong, Mani Baskaran, Shamira A Perera, and Tin Aung, “Similarity-weighted linear reconstruction of anterior chamber angles for glaucoma classification,” in *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*. IEEE, 2016, pp. 693–697.

[7] Huazhu Fu, Yanwu Xu, Stephen Lin, Damon Wing Kee Wong, Mani Baskaran, Meenakshi Mahesh, Tin Aung, and Jiang Liu, “Angle-closure detection in anterior segment oct based on multilevel deep network,” *IEEE transactions on cybernetics*, 2019.

[8] Alejandro Newell, Kaiyu Yang, and Jia Deng, “Stacked hourglass networks for human pose estimation,” in *European conference on computer vision*. Springer, 2016, pp. 483–499.

[9] Yilun Chen, Zhicheng Wang, Yuxiang Peng, Zhiqiang Zhang, Gang Yu, and Jian Sun, “Cascaded pyramid network for multi-person pose estimation,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 7103–7112.

[10] Abhishek Chaurasia and Eugenio Culurciello, “Linknet: Exploiting encoder representations for efficient semantic segmentation,” in *2017 IEEE Visual Communications and Image Processing (VCIP)*. IEEE, 2017, pp. 1–4.

[11] Jie Hu, Li Shen, and Gang Sun, “Squeeze-and-excitation networks,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 7132–7141.