Quantitative analysis of functional filtering bleb size using Mask R-CNN

Tao Wang[1#], Lei Zhong[1#], Jing Yuan[2#], Ting Wang[3], Shiyi Yin[1], Yi Sun[1,3], Xing Liu[3], Xun Liu[4], Shiqi Ling[1]

[1]Department of Ophthalmology, Third Affiliated Hospital of Sun Yat-Sen University, Guangzhou, China; [2]Eye Center, Renmin Hospital of Wuhan University, Wuhan, China; [3]State Key Laboratory of Ophthalmology, Zhongshan Ophthalmic Center, Sun Yat-sen University, Guangzhou, China; [4]Department of Medical Big Data, Third Affiliated Hospital of Sun Yat-Sen University, Guangzhou, China

Contributions: (I) Conception and design: T Wang, X Liu, X Liu, S Ling; (II) Administrative support: X Liu, X Liu, S Ling; (III) Provision of study materials or patients: T Wang, L Zhong, J Yuan, T Wang, S Yin, X Liu, X Liu, S Ling; (IV) Collection and assembly of data: T Wang, L Zhong, J Yuan, T Wang, S Yin, Y Sun; (V) Data analysis and interpretation: All authors; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Xing Liu. State Key Laboratory of Ophthalmology, Zhongshan Ophthalmic Center, Sun Yat-sen University, Guangzhou, China. Email: liuxing@mail.sysu.edu.cn; Xun Liu. Department of Medical Big Data, Third Affiliated Hospital of Sun Yat-Sen University, Guangzhou, China. Email: naturestyle@163.com; Shiqi Ling. Department of Ophthalmology, Third Affiliated Hospital of Sun Yat-Sen University, Guangzhou, China. Email: lingshiqi123@163.com.

Background: Deep learning has had a large effect on medical fields, including ophthalmology. The goal of this study was to quantitatively analyze the functional filtering bleb size with Mask R-CNN.

Methods: This observational study employed eighty-three images of post-trabeculectomy functional filtering blebs. The images were divided into training and test groups and scored according to the Indiana Bleb Appearance Grading Scale (IBAGS) system. Then, 70 images from the training group were used to train an automatic detection system based on Mask R-CNN and perform a quantitative analysis of the function bleb size. Thirteen images from the test group were used to evaluate the model. During the training process, left and right image-flipping algorithms were used for data augmentation. Finally, the correlation between the functional filtering bleb area and the intraocular pressure (IOP) was analyzed.

Results: The 83 functional filtering blebs have similar morphological features. According to IBAGS, the functional filtering blebs have a high incidence of E1/E2, H1/H2, and V0/V1. Our Mask R-CNN-based model using the selected parameters achieves good results on the training group after a 200-epoch training process. All the Intersection over Union (IoU) scores exceeded 93% on the test group. The Spearman correlation coefficient between the area of functional filtering blebs and the IOP value was -0.757 ($P<0.05$).

Conclusions: Deep learning is a powerful tool for quantitatively analyzing the functional filtering bleb size. This technique is suitable for use in monitoring post-trabeculectomy filtering blebs in the future.

Keywords: Glaucoma; trabeculectomy; filtering bleb; deep learning; Mask R-CNN

Introduction

With the advancement of powerful image processing and machine learning techniques, computer-aided diagnosis (CAD) will lead to substantial changes in many medical fields, including ophthalmology. Deep learning has been recently applied to perform automated detection of diabetic retinopathy from fundus photographs, grade nuclear cataracts, and segment foveal microvasculature (1-12).

Glaucoma is the leading cause of irreversible blindness globally and is estimated to be quite common in 3.5%

Page 2 of 9

Wang et al. Quantitative analysis functional filtering bleb using Mask R-CNN

of the population aged between 40 and 80 years (13). Trabeculectomy remains the classical surgical procedure for glaucoma when the maximal results of pressure-reducing medical or laser treatments fail to achieve the target intraocular pressure (IOP). Many clinical investigations have shown that the presence of filtration blebs is correlated with IOP after trabeculectomy. Creation and maintenance of these functional filtering blebs is considered to be a key aspect in a favorable trabeculectomy outcome. Filtering blebs can be assessed using slit-lamp examinations and color photography (14-16), ultrasound biomicroscopy (UBM) (17,18), and optical coherence tomography (OCT) (19-24). An objective and accurate evaluation of the morphology and structure of filtration blebs helps in monitoring and better maintaining the function of the filtration blebs.

There are two available evaluation systems for filtering blebs, neither of which requires complex, sophisticated or expensive equipment: the Indiana Bleb Appearance Grading Scale (IBAGS) and the Moorfields Bleb Grading System (MBGS) (14-16). Although there are some differences between these systems, both consider area, height, and vascularity to be the key filtering bleb characteristics. However, a method for accurately assessing the morphologic characteristics of filtering blebs is lacking.

Convolutional neural networks (CNNs) have shown strong abilities in image recognition, object detection and segmentation tasks and have proved to be good tools for judging border characteristics and colors. A region-based CNN (R-CNN) has been used to apply CNNs to object detection tasks. The Mask R-CNN model, which can both segment instances and perform classification—is one of the latest, the most effective, and beneficial in-depth learning models, and it can be used even with some small datasets (25).

Although several studies using deep learning technology have been conducted on visual field perimetry in glaucoma patients and the detection of glaucoma in the fundus (3,5), deep learning technology has not yet been used to evaluate filtering blebs after trabeculectomy.

In this study, we focus on quantitatively evaluating the functional filtering bleb area with Mask R-CNN technology.

Methods

This study is an observational study. We collected color photos of functional filtering blebs. By training a deep learning model, the computer can learn to automatically quantify the size of the functional filtering blebs. Then, we analyze the correlation between size of the filtering bleb and the IOP. All the research was conducted in accordance with the tenets of the Declaration of Helsinki and with the approval of the Ethics Committee of the Third Affiliated Hospital of Sun Yat-sen University.

Inclusion criteria for functional filtering blebs

All included blebs met the following criteria:
- (I) The eyes had primary glaucoma.
- (II) Only one trabeculectomy had been performed.
- (III) An apparently functional filtering bleb was present.
- (IV) IOP measuring less than 21 mmHg without anti-glaucoma medication. (The IOP was measured with Goldmann applanation tonometry)
- (V) At least 1 year had elapsed after trabeculectomy.
- (VI) Trabeculectomy was performed in the upper part of the eye.
- (VII) Patients had not undergone any other internal eye surgery.
- (VIII) A negative Seidel test was performed.

Collection of filtering bleb images

All filtering bleb images were acquired in the following manner. Using a digital slit lamp biomicroscopy camera system, the patient looks at the underlying optotype, exposing all of the filter bleb above, dispersing the source, at a 10 objective. We selected the accurate and clear images containing a complete filter bleb for inclusion in this study.

Grading the functional filtering blebs

All the functional filtering blebs included in this research were evaluated by a researcher (Tao W) with the IBAGS system (14,16). The IBAGS is one of the commonly used clinical grading systems and has demonstrated good interobserver agreement. These grading systems include measurements of the bleb area, height, and vascularity, which are recognized as indicators of surgical success.

Quantitative analysis of the area of functional filtering blebs based on Mask R-CNN

Mask R-CNN, is an extension of Faster R-CNN that includes an added network branch for predicting an object mask that functions in parallel with the existing branch
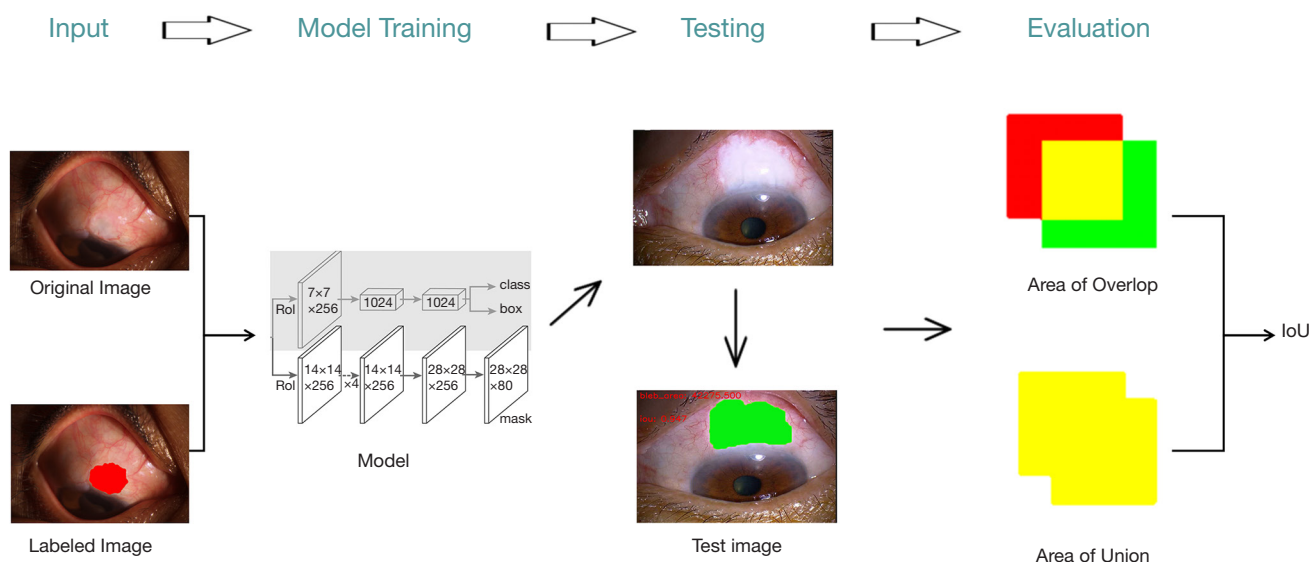
Figure 1 Framework of the quantitative analysis of functional filtering bleb size based on Mask R-CNN.
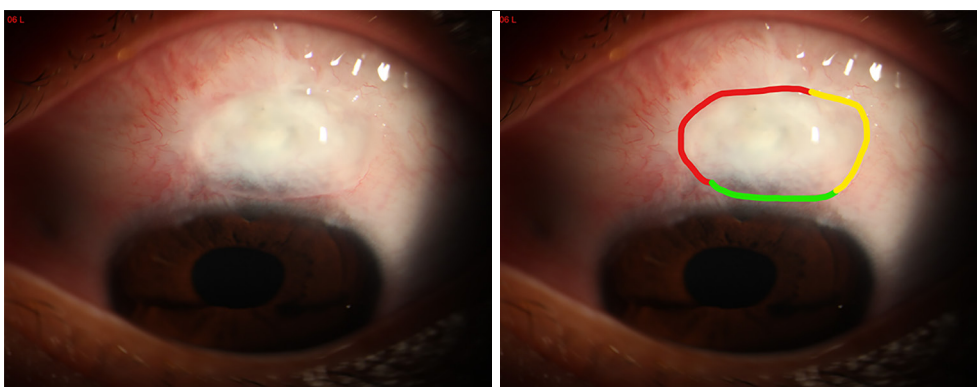

Figure 2 The labeling of a functional filtering bleb in a photograph. The IBAGS score of this bleb was H2E2V1. The red line represents the boundary of the outer edge of the rising filter bleb, the yellow line represents the medial edge of the vascular zone, and the green line represents the directional change of the microvascular network.

for bounding box regression and classification. Mask R-CNN is an efficient approach for detecting objects in an image while simultaneously generating a high-quality segmentation mask for each detected instance. In this study, we quantitatively analyzed the functional filtering bleb area using a Mask R-CNN model. The framework of our main method is illustrated in *Figure 1*.

Labeling the boundaries of the functional filtration bleb
The bleb area represents the horizontal dimension of the filtering bleb. The pixel values of the detected functional filtering blebs were used to calculate the size of each

filtering bleb. The boundaries of the filtering blebs were labeled according to the following standards (*Figure 2*):

❖ The outer edge of the obviously visible filtering bleb, rising around the scleral flap, with no scars;
❖ The medial edge of the vascular zone;
❖ The termination or directional change of the microvascular network.

Data set and model training
The training images of functional filtering bleb boundaries were labeled by two glaucoma experts. When the reviewers' opinions differed, the boundary was marked according to

```
# Learning rate and momentum
# The paper uses lr=0.02, but we found that to cause weights to explode often
# The learning rate in the paper is 0.02, but here we use 0.002 to avoid gradient explosion
LEARNING_RATE=0.002
LEARNING_MOMENTUM=0.9
# Weight decay regularization
WEIGHT_DECAY=0.0001
```

Figure 3 The key parameter settings of the model training process.

the senior glaucoma expert suggestion.

Due to the lack of a publicly available filtering bleb image data set, the included filtering bleb images were divided into two parts: one part of the filtering bleb images was randomly selected and used as the training images, and the remaining part, which was not used during model training, was used as independent test data to evaluate the performance of the trained model. During model training, we applied the multi-task loss approach to optimize the filtering bleb localization, classification and segmentation loss (25). The Mask R-CNN loss function is as follows:

$$L = L_{class} + L_{box} + L_{mask} .\qquad [1]$$

Due to the small scale of training data, to improve the model generalization ability, we applied data augmentation to enlarge the training data, including left and right flip, random scale and crop, rotate, and brightness (26).

In the training stage, we first trained the head [the Region Proposal Network (RPN), classifier and mask heads of the network] for 40 epochs with a learning rate of 0.002 and stochastic gradient descent (SGD) as the optimizer. Then we trained the Resnet block for 100 epochs using the same learning rate and optimizer, and finally we trained the whole network (all layers) for another 200 epochs, while the learning rate and optimizer remained the same.

The parameter settings for the model training process

The key parameters are shown in the *Figure 3*. Detailed parameter settings are at http://cdn.amegroups.cn/static/application/f7c546885d37faf2cb4faf9923f4eed8/ATM-2019-MAIR-17-supplementary.pdf.

Model evaluation

We evaluated the model in two ways.

Intersection over Union (IoU)

IoU is commonly used to evaluate the accuracy of object detection algorithms (25). The overlapping area between the two regions of the model prediction result and the expert label result (intersection) is divided by the total area (the union) of the two regions. The obtained ratio can be used to evaluate model performance: the higher the degree of coincidence between the model-predicted area and the expert-label area, the closer the IoU ratio is to 1, and the better the performance is.

Correlate the functional filtering bleb size and IOP

A functional filtration bleb has been shown to be correlated with IOP. To further evaluate our Mask R-CNN model, we correlated the size of the functional filtering bleb automatically generated by the computer with the IOP.

Statistical analysis

The statistical analyses were performed with IBM SPSS software version 24. The Kolmogorov–Smirnov test method first tests whether the data are normally distributed. When the data have a normal distribution, the Kolmogorov–Smirnov uses the Pearson correlation analysis; otherwise, it uses the Spearman correlation analysis. The correlation coefficient was used to evaluate the correlation between the two categorical variables. A two-tailed test was used for all measurements, and a P value of <0.05 was considered significant for the measured variables. Data were expressed as the mean ± standard deviation for metric values and as a frequency (percentage) for categorical variables.

Results

A total of 83 of the collected functional filtering bleb images met the requirements to create a filtering bleb image dataset.

For model training, we randomly sampled 70 labeled images as training data (Training Group), and used the remaining 13 images as independent test data (Test Group); the test images were reserved for model effect evaluation and were not involved in the model training process.

We graded both the training group and the test group

Table 1 Bleb clinical characteristics of training group and test group in IBAGS

Parameters	IBAGS												
	Extension				Height				Vascularity				
	E_0	E_1	E_2	E_3	H_0	H_1	H_2	H_3	V_0	V_1	V_2	V_3	V_4
Training group (n=70)	0	21	43	6	3	21	38	8	28	40	2	0	0
Test group (n=13)	0	4	9	0	0	4	7	2	6	7	0	0	0
P	0.543>0.05				0.873>0.05				0.781>0.05				

No Seidel test was included because a positive Seidel test result represents leakage of aqueous humor through the bleb surface; thus, blebs with positive Seidel tests were excluded.
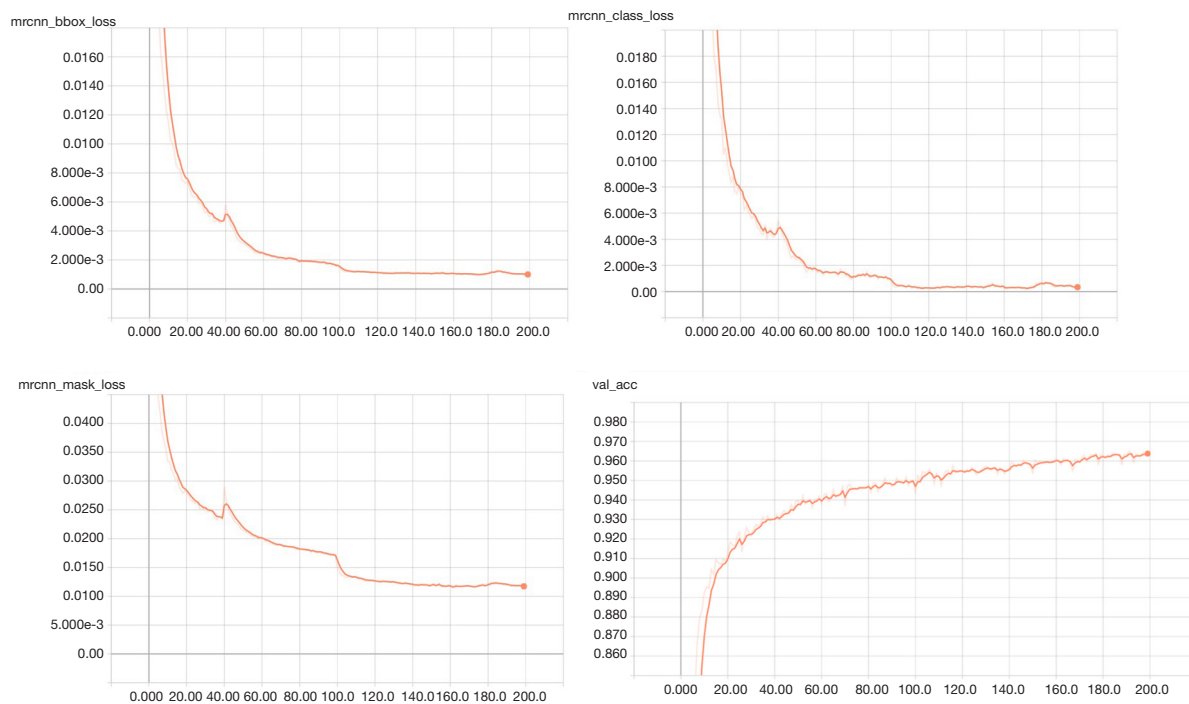
Figure 4 The loss graph of the training process and the classification accuracy.

based on their clinical characteristics according to IBAGS. All functional filtering blebs have similar morphological features. There was no significant difference between the two groups (*Table 1*).

Using the selected model hyperparameters for training, the model achieves good results after approximately 200 epochs. The loss graphs of the training process and the classification accuracy are shown in *Figure 4*. The model converges after approximately 200 training epochs.

The Precision-Recall curves of bounding box regression and segmentation are shown in *Figure 5*.

Model evaluation results

We applied the IoU algorithm to evaluate the model's performance on the test data. The IoU results of the model for the filtering bleb area range from 0.938 to 0.974, which means that the automatically recognized results of functional filtering blebs by the computer were very close to those labeled by the expert. The IoU values of the test group are shown in *Figure 6*, and one of the functional filtering blebs from the test group automatically quantitatively analyzed by the computer is shown in *Figure 7*.
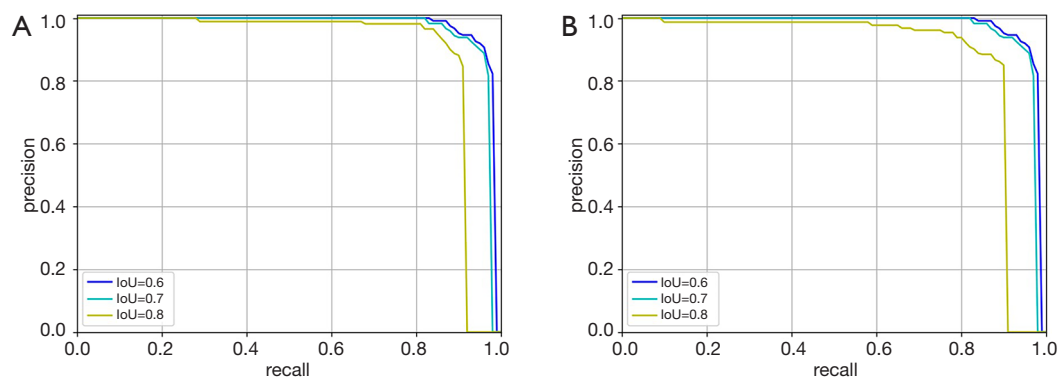
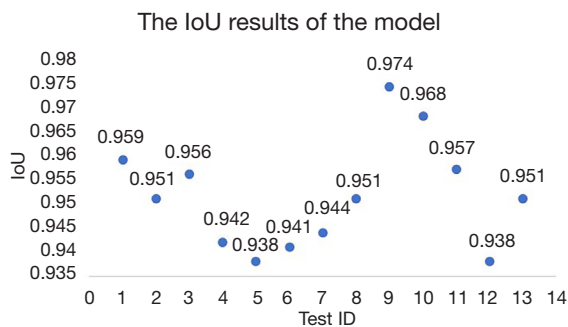**Figure 5** The Precision-Recall curves of bounding box regression and segmentation.

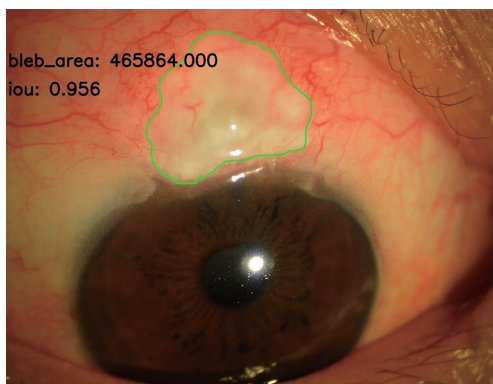**Figure 6** The IoU values of the test group.

Figure 7 One color image from the test group. The functional filtering bleb was automatically quantitatively analyzed by the computer. The green line shows the boundary of the bleb with a high IoU value (0.956).

The functional filtering bleb area and the IOP value in the test group were tested with the Kolmogorov-Smirnov normal distribution test. The results showed that the bleb area did not fit a normal distribution (P=0.023<0.05), but

the IOP value fit the normal distribution (P=0.09>0.05). Therefore, the correlations between bleb area and IOP value were analyzed using a nonparametric correlation test. Bivariate analyses showed a statistically significant correlation between the functional bleb size based on Mask R-CNN and IOP value with a Spearman rank correlation coefficient of r=−0.757 (P=0.003).

Discussion

In this study, we quantitatively analyzed the functional filtering bleb size with Mask R-CNN. To the best of our knowledge, this is the first study to quantitatively evaluate filtering blebs using deep learning in peer-reviewed literature.

The success of a trabeculectomy depends on the long-term presence of the functional filtering bleb. The wound healing process continues for an extended period post trabeculectomy. We termed the functional filtering bleb the "target bleb" for all the post-trabeculectomy targets interventions to form a functional bleb. The earlier that patients at risk of developing filtering scarring are identified, the earlier a proper intervention may result in the creation and maintenance of a functional filtering bleb (27,28). Many factors can affect IOP during the early stage, and a long-term stable functional filtering bleb is an important factor that affects the IOP (29). Therefore, we chose the functional filtering bleb at least 1 year after trabeculectomy with no use of any anti-glaucoma medication.

In general, there is a substantial association between the morphology and structure of functional blebs (30-32). A functional bleb is more likely avascular or has only mild vascularization, and has low or medium height in the morphology. It presents as a low reflection signal with more

microcapsules in UBM or OCT (17,19-21). Compared with UBM or OCT scans, photographic images of blebs are easier to obtain. The size of the bleb may be easier to analyze quantitatively from bleb morphologic features.

The development of CNN layers has fostered significant gains in the ability to classify images and detect objects from an image (1-12). In this study, based on our parameter settings, we used a Mask R-CNN model to automatically identify of functional filter blebs and perform quantitative analyses of filter bleb size. According to our parameter settings, the model converges after 200 training cycles and reaches good results. The test data show that the consistency of functional filtering bleb size identification between the trained model and a glaucoma expert exceeds 93%. A correlation between the IOP value and the size of the filtering bleb in the test group was performed, and a significant inverse correlation was found (P=0.023<0.05). The Spearman rank correlation coefficient was high, with a value of –0.757. This result provides some indirect insight into the feasibility and efficacy of our method. Quantitative analyses of the sizes of filtering blebs based on deep learning technology may become a powerful tool for post-trabeculectomy monitoring of filtering blebs.

However, this study has some limitations. The number of training images was insufficient. Furthermore, from its establishment to the point at which it maintains relative stability, the appearance and structure of a filtering bleb can change constantly. Even a functional filtering bleb should not be considered in a static state but rather in a dynamic change state. Because we focused on relatively stable function filtering blebs, the reliability and effectiveness of the method need to be further verified when applied to other filtering blebs. In future work, we will attempt to obtain additional image data through the Internet to increase the accuracy of our method and to verify it on other filtering bleb applications.

Conclusions

In summary, deep learning is a powerful tool for quantitatively analyzing the functional filtering bleb size. According to our parameter settings, this Mask R-CNN model is suitable for use in monitoring post-trabeculectomy filtering blebs in the future.

Acknowledgments

Footnote

Provenance and Peer Review: This article was commissioned by the Guest Editors (Haotian Lin and Limin Yu) for the series "Medical Artificial Intelligent Research" published in *Annals of Translational Medicine*. The article was sent for external peer review organized by the Guest Editors and the editorial office.

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at http://dx.doi.org/10.21037/atm.2020.03.135). The series "Medical Artificial Intelligent Research" was commissioned by the editorial office without any funding or sponsorship. All authors report other from Guangdong Kingpoint Data Technology Co., Ltd., during the conduct of the study.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. All the research was conducted in accordance with the tenets of the Declaration of Helsinki and with the approval of the Ethics Committee of the Third Affiliated Hospital of Sun Yat-sen University.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: https://creativecommons.org/licenses/by-nc-nd/4.0/.

References

1. Abràmoff MD, Lou Y, Erginay A, et al. Improved Automated Detection of Diabetic Retinopathy on a

Page 8 of 9

Wang et al. Quantitative analysis functional filtering bleb using Mask R-CNN

Publicly Available Dataset Through Integration of Deep Learning. Invest Ophthalmol Vis Sci 2016;57:5200-6.

2. Lee CS, Baughman DM, Lee AY. Deep learning is effective for the classification of OCT images of normal versus Age-related Macular Degeneration. Ophthalmol Retina 2017;1:322-7.

3. Asaoka R, Murata H, Iwase A, et al. Detecting preperimetric glaucoma with standard automated perimetry using a deep learning classifier. Ophthalmology 2016;123:1974-80.

4. Rampasek L, Goldenberg A. Learning from Everyday Images Enables Expert-like Diagnosis of Retinal Diseases. Cell 2018;172:893-5.

5. Bajwa MN, Malik MI, Siddiqui SA, et al. Two-stage framework for optic disc localization and glaucoma classification in retinal fundus images using deep learning. BMC Med Inform Decis Mak 2019;19:136.

6. Gao X, Lin S, Wong TY. Automatic Feature Learning to Grade Nuclear Cataracts Based on Deep Learning. IEEE Trans Biomed Eng 2015;62:2693-701.

7. Prentašic P, Heisler M, Mammo Z, et al. Segmentation of the foveal microvasculature using deep learning networks. J Biomed Opt 2016;21:75008.

8. Zheng C, Johnson TV, Garg A, et al. Artificial intelligence in glaucoma. Curr Opin Ophthalmol 2019;30:97-103.

9. Wei Zhang, Jie Zhong, Shijun Yang, et al. Automated identification and grading system of diabetic retinopathy using deep neural networks. Knowledge-Based Systems 2019;175:12-25.

10. Saha S, Nassisi M, Wang M, et al. Automated detection and classification of early AMD biomarkers using deep learning. Sci Rep 2019;9:10990.

11. Gibson E, Li W, Sudre C, et al. NiftyNet: a deep-learning platform for medical imaging. Comput Methods Programs Biomed 2018;158:113-22.

12. Tham YC, Li X, Wong TY, et al. Global prevalence of glaucoma and projections of glaucoma burden through 2040: a systematic review and meta-analysis. Ophthalmology 2014;121:2081-90.

13. Picht G, Grehn F. Classification of filtering blebs in trabeculectomy: biomicroscopy and functionality. Curr Opin Ophthalmol 1998;9:2-8.

14. Cantor LB, Mantravadi A, WuDunn D, et al. Morphologic classification of filtering blebs after glaucoma filtration surgery: the Indiana Bleb Appearance Grading Scale. J Glaucoma 2003;12:266-71.

15. Wells AP, Crowston JG, Marks J, et al. A pilot study of a system for grading of drainage blebs after glaucoma surgery. J Glaucoma 2004;13:454-60.

16. Wells AP, Ashraff NN, Hall RC, et al. Comparison of two clinical Bleb grading systems. Ophthalmology 2006;113:77-83.

17. Yamamoto T, Sakuma T, Kitazawa Y. An ultrasound biomicroscopic study of filtering blebs after mitomycin C trabeculectomy. Ophthalmology 1995;102:1770-6.

18. El Salhy AA, Elseht RM, Al Maria AF, et al. Functional evaluation of the filtering bleb by ultrasound biomicroscopy after trabeculectomy with mitomycin C. Int J Ophthalmol 2018;11:245-50.

19. Leung CK, Yick DW, Kwong YY, et al. Analysis of bleb morphology after trabeculectomy with Visante anterior segment optical coherence tomography. Br J Ophthalmol 2007;91:340-4.

20. Singh M, Chew PT, Friedman DS, et al. Imaging of trabeculectomy blebs using anterior segment optical coherence tomography. Ophthalmology 2007;114:47-53.

21. Tominaga A, Miki A, Yamazaki Y, et al. The assessment of the filtering bleb function with anterior segment optical coherence tomography. J Glaucoma 2010;19:551-5.

22. Napoli PE, Zucca I, Fossarello M. Qualitative and quantitative analysis of filtering blebs with optical coherence tomography. Can J Ophthalmol 2014;49:210-6.

23. Cvenkel B, Kopitar AN, Ihan A. Correlation between filtering bleb morphology, expression of inflammatory marker HLA-DR by ocular surface, and outcome of trabeculectomy. J Glaucoma 2013;22:15-20.

24. Kasaragod D, Fukuda S, Ueno Y, et al. Objective Evaluation of Functionality of Filtering Bleb Based on Polarization-Sensitive Optical Coherence Tomography. Invest Ophthalmol Vis Sci 2016;57:2305-10.

25. He K, Gkioxari G, Dollar P, et al. Mask R-CNN. IEEE Trans Pattern Anal Mach Intell 2020;42:386-97.

26. Hussain Z, Gimenez F, Yi D, et al. Differential Data Augmentation Techniques for Medical Imaging Classification Tasks. AMIA Annu Symp Proc 2018;2017:979-84.

27. Narita A, Morizane Y, Miyake T, et al. Characteristics of early filtering blebs that predict successful trabeculectomy identified via three-dimensional anterior segment optical coherence tomography. Br J Ophthalmol 2018;102:796-801.

28. Tsutsumi-Kuroda U, Kojima S, Fukushima A, et al. Early bleb parameters as long-term prognostic factors for surgical success: a retrospective observational study using three-dimensional anterior-segment optical coherence

tomography. BMC Ophthalmol 2019;19:155.

29. Narita A, Morizane Y, Miyake T, et al. Characteristics of successful filtering blebs at 1 year after trabeculectomy using swept-source three-dimensional anterior segment optical coherence tomography. Jpn J Ophthalmol 2017;61:253-9.

30. Wen JC, Stinnett SS, Asrani S. Comparison of Anterior Segment Optical Coherence Tomography Bleb Grading, Moorfields Bleb Grading System, and Intraocular Pressure After Trabeculectomy. J Glaucoma 2017;26:403-8.

31. Sacu S, Rainer G, Findl O, et al. Correlation between the early morphological appearance of filtering blebs and outcome of trabeculectomy with mitomycin C. J Glaucoma 2003;12:430-5.

32. Seo JH, Kim YA, Park KH, et al. Evaluation of Functional Filtering Bleb Using Optical Coherence Tomography Angiography. Transl Vis Sci Technol 2019;8:14.