DA-M2Det: An Iris Classification Network for UBM Images

Luquan Wang1,3, Junxing Lao2,3, Lingfeng Yang2,3, Yaguang Zeng2,3 and Yong Chen1

1 School of Mechatronic Engineering and Automation, Foshan University, Foshan, Guangdong, 528225, China
2 School of Physics and Optoelectronics, Foshan University, Foshan, Guangdong, 528225, China
3 Guangdong-Hong Kong-Macao Joint Laboratory for Intelligent Micro-Nano Optoelectronic Technology, Foshan University, Foshan, Guangdong, 528225, China
E-mail: 17363994096@163.com

Abstract. Primary angle closure glaucoma (PACG) is primarily diagnosed by ophthalmologists through morphological analysis of the iris in ultrasonic biomicroscopy (UBM). In recent years, deep convolutional neural networks (CNNs) show potential for quick category definition in eye disease. According to the characteristics of iris in UBM images, we proposed a network (DenseNet and Attention gate) DA-M2Det to automatic classification iris morphology. Firstly, in the framework of M2Det network, We used the backbone of DenseNet to replace the VGG backbone of M2Det, better extraction of basic feature layers. Secondly, three scales of attention gate (AG) was added to the Thinned U-shape Module (TUM), enable the network to pay more attention to the iris region. Finally, we use the retraining method to further improve the accuracy of iris classification. The classification results of VGG-16, M2Det, ResNet-50 and DA-M2Det networks are compared experimentally. The results show that, in three different iris shapes (including arch, flat and depression), DA-M2Det achieves an average classification accuracy of 85%, which is higher than that of the other three networks. Experimental results show that DA-M2Det can accurately classify irises into three categories, assisting ophthalmologists to quickly diagnose the cause of glaucoma and accurately perform clinical treatment thereby.

1. Introduction
Glaucoma is one of the three major diseases that cause blindness in humans, according to the angle between the cornea and the iris, the anterior chamber angle (ACA) can be divided into PACG and primary open angle glaucoma (POAG) two common subtypes. Among them, the incidence of PACG is gradually increasing in Asia, especially China. It is estimated that 9 million people have obvious ACG, and more than 28 million people are susceptible to primary ACG[1]. Therefore, for glaucoma patients, it is necessary to accurately distinguish their etiology and take corresponding treatment.

PACG is due to the abnormal iris configuration, the ACA is mechanically blocked by the surrounding iris tissue, resulting in the obstruction of aqueous humor outflow and the increase of intraocular pressure. Therefore, it is of great significance to study the differences in iris morphology to understand the pathogenesis of PACG. The morphology of iris generally includes arch, flat and depression. Traditional gonioscopy has a certain degree of invasiveness, requires surface anesthesia, and may cause discomfort to the patient. With the development of anterior segment imaging technology, anterior segment optical coherence tomography (AS-OCT) and UBM have been derived, which makes the research on the anatomical risk factors of PACG gradually deepen. AS-OCT is a non-invasive optical imaging
diagnostic technology, which can obtain ACA high-resolution images. AS-OCT allows good quantitative and qualitative analysis of images, which can be used to determine ACA and better classify them [2-5]. However, its light cannot penetrate the iris, so the morphology of the iris cannot be observed well.

In order to further accurately identify the etiology of glaucoma patients, ophthalmologists usually observe the UBM images of patients. UBM is a B-type high-frequency ultrasonic diagnostic instrument, whose resolution can reach the level of ordinary optical microscope, and can provide detailed two-dimensional gray-scale images of anterior segment structure in vivo. UBM has been clinically studied and applied to a certain extent. Boonsong et al. proposed to take UBM image data as part of training and detect Plateau Iris from AS-OCT images through deep learning [6]. Wang et al. evaluated iris cysts and ciliary cysts under UBM [7], and described the narrowing or closure of atrial angle caused by cyst. Cataract and intraocular lens implantation can also be reflected on UBM images [1, 8]. According to literature search, there are few automatic detection algorithms for iris morphology. Therefore, the DA-M2Det target detection algorithm [9] in neural network method is used to detect the shape of iris for the first time in this paper. The deep learning algorithm makes use of multi-scale feature and channel attention module to obtain the feature information with the most distinguishing feature degree. We noticed that DenseNet pre-training weight can be used backbone of DA-M2Det network, enhance iris feature extraction by network, and in TUM module, attention gate is added to enable the network to better extract iris position features. In order to further explore the potential of the network for iris classification, this paper uses retraining for multiple training to achieve higher accuracy.

2. Method

2.1. Preprocessing

In the experiment, the information display in the upper left corner and the gray bar display in the right of the original image were cut, and the original 1024×576 was cut to 830×576, so as to remove useless information in the image and retain the iris area completely. We also screened the original data set in the experiment. Filtering out the distorted, inaccurate scanning position and different scanning area in the collected images... The final cut images were shown on the right of figure 1.

![Figure 1. UBM image, the left is the original image, and the right is the preprocessed image.](image)

2.2. DA-M2Det Network

2.2.1. Overall Structure. In order to accurately detect and classify the iris, we selected M2Det network. M2Det is a single-shot type network proposed by Zhao et al. in 2019, which is mainly composed of three modules Feature Fusion Module (FFM), TUM and Scale-wise Feature Aggregation Module (SFAM). Based on the M2Det network, we replaced VGG backbone of M2Det to DenseNet [10] backbone, which deepened the depth of the network. In the TUM module, the AG is added to form the Att-TUM module, which further improves the extraction accuracy of the model for different scale features. The overall block diagram of the model is shown in figure 2 below.
In figure 2, BasDense is the pre-training module of DenseNet-121. Att-TUM outputs from shallow to deep to the scale feature aggregation module of SFAM, and then outputs 6 scales (40×40 to 1×1) Features to the SFAM layer.

2.2.2. Basic Feature Extraction Module. M2det network selects VGG-16[11] as the backbone. VGG-16 is a relatively early network model. Due to its shallow network depth, the network cannot better extract more abstract image features. M2det also uses ResNet-101[12] as the basic feature extraction. After adding residual blocks, ResNet-101 network can make the network deeper without degradation, improved M2Det classification results. This paper selects DenseNet-121 as the backbone. DenseNet realizes feature reuse. It takes all the previous features as the later input to form denseblock. Due to the superposition mode of denseblock, it not only ensure the depth of the network and extract rich feature layer information, but also increases the information flow transmission of features and allows the network to learn more complex and abstract image features. Therefore, DenseNet-121 is selected as the backbone of the DA-M2Det. At the same time, in order to speed up the network training, the pre training weight of DenseNet-121 of IMGNET is used.

Figure 3. Structure of Att-TUM and AG.

a is the u-net structure in Att-TUM, and AG module is added in the connection between encoding
and decoding. AG module is shown as b, where Relu and Sigmoid are activation functions, Resampler is resampling the attention coefficient for trilinear interpolation, and c1 is the result after the AG module.

2.2.3. Att-TUM Module. The core idea of M2Det network is in the design of TUM module. It accepts the feature layer extracted from the basic feature module, passes through a shallow U-Net[13] network module, then adds the output of the last layer to the basic feature module, and then passes through the same u-net module. Each u-net module will output 6 feature layers of different scales (40×40 to 1×1), thus obtaining features of different scales from shallow to deep, because the input of each layer is added to the basic feature module Therefore, the TUM output combines the information flow before and after, ensuring the dissemination of characteristic information. This paper considers adding the attention mechanism AG[14] to the TUM module. The added Att-TUM module is shown in Figure 3 below. The AG module was added to the feature maps with sizes of 40×40, 20×20, and 10×10, respectively. For feature maps smaller than 10, the effect of the AG module is not obvious because of the lack of expression information, so it is not added[15]. The output of c1 is added with the feature graph of the corresponding encoding part, and then enter the next layer of convolution. The AG module ensures that in each feature transmission process, the network will pay more attention on the target area, so as to better detect the position and shape of iris.

3. Experiment

3.1. Datasets
The materials used to obtain UBM in this study were provided by the South Hospital of Southern Medical University. The study was approved by the ethics committee of Nanfang Hospital, Southern Medical University, and the images used were approved by all participants. The ground-truth data were labeled by two ophthalmologists with over 10 years of clinical experience. These images were acquired on the MD-300L Ultrasound Biomicroscope with a probe frequency of 50MHz, a detection depth of 5mm, and a resolution of 50μm. UBM images are randomly divided into training set, validation set, and test set with a ratio of 4:1:1.

3.2. Experimental Environment and Training
The network training and testing equipment is NVIDIA GeForce GTX2080Ti GPU, 2.90-GHz CPU and 32GB of running memory, the experimental operating system is Windows 10, and the main software packages are Python 3.6, TensorFlow 1.14, cuDNN 7.6.4 and CUDA 10.0. The network learning rate is set to 1e-5. If the loss of 3 epochs does not decrease, the learning rate is reduced to half. If the loss of 30 cycles does not decrease, stop early. The total training epochs is 160. At the same time, the weight of the DenseNet pre-training parameters of the first 80 epochs does not change. Make the first 80 epochs only train the non-basic feature network layer.

4. Result and Discussion
The experiment compares the detection and classification results of VGG-16, M2Det, ResNet-50 and DA-M2Det. The detection results of DA-M2Det on the three types of UBM images are shown in Figure 4. In order to further explore the performance of the network and improve the accuracy of the network for the classification of different iris images, the experiment tried to take the trained model as the pre-training weight and train again, that is retraining, so as to improve the accuracy. Although this will increase the total training time, it can be determined from the experimental results that this does increase the overall accuracy of the result classification. At the same time, in practical applications, this does not affect the prediction time of the model.
It can be seen from figure 4 that the DA-M2Det can accurately frame the area where the iris is located for all iris positions, indicating that the network can locate the area well. At the same time, in terms of category judgment, DA-M2Det can better detect depression and arching. In the flat image, the network in Figure 4(f) thinks that the image may be flat or depression, but in terms of probability, it thinks that the image is flat with a probability of 0.85, and obtains the correct result.

Table 1. Classification accuracy results of UBM after three training of four networks
Each network from left to right is the first, second and third retraining results

| Method | VGG-16 | ResNet-50 | M2Det | DA-M2Det |
|--------|--------|-----------|-------|----------|
| Depression | 0.00 | 0.19 | 0.250.69 | 0.81 | 0.88 | 0.55 | 0.62 | 0.68 | 0.81 | 0.88 | 0.94 |
| Arch | 0.94 | 0.99 | 0.92 | 0.90 | 0.90 | 0.90 | 0.90 | 0.94 | 0.92 | 0.91 | 0.89 |
| Flat | 0.00 | 0.00 | 0.070.38 | 0.57 | 0.600.69 | 0.68 | 0.750.60 | 0.68 | 0.72 |

Table 1 shows the classification accuracy of four networks for three types of iris morphology. Each network has been trained three times, so three results have been obtained for each classification. It can be seen that after retraining, most networks have improved the classification accuracy of all iris types. Among them, the average accuracy of DA-M2Det has increased from 0.77 to 0.85, reaching the highest accuracy, and three other networks have also improved to varying degrees in retraining. Finally, VGG-16, ResNet-50, M2Det achieved average accuracy of 0.41, 0.79 and 0.79. It is worth noting that during the first training of VGG-16, the accuracy of depression and flat are all 0, and all the predicted iris images are arched. This may be because the extraction of iris features by VGG-16 network is not complete, and the extracted iris image features cannot distinguish depression and flatness, so they are all mistaken for arched. DA-M2Det is very sensitive to the depression image at the beginning, and can finally achieve an accuracy of 0.94, which is much higher than M2Det and the other two networks, indicating that the depth of attention module and DenseNet can improve the accuracy of the network for depression and flat, so as to achieve the best results.

We think the detection accuracy improvement of DA-M2Det is mainly brought by the AG module. On one hand, the backbone of DenseNet can effectively avoid network degradation such as gradient dispersion, so that the network can get better basic characteristics. On the other hand, we looked at the feature map after the AG module and found a higher activation value in the iris region. which indicates that DA-M2Det can get more accurate iris classification results.

5. Conclusion
This paper proposes a DA-M2Det network for UBM image iris classification, which used backbone of DenseNet replace the VGG backbone of M2Det, and add attention gate module into TUM, it makes the network better focus on the iris image information, and use retraining to improve the classification accuracy of the network. In order to verify the accuracy of the network, we compared the accuracy of DA-M2Det with VGG-16, ResNet-50 and M2Det. The experimental results show that DA-M2Det achieves the highest average accuracy of 85%. Therefore, DA-M2Det can automatically and accurately classify the iris image and frame the iris area, so as to provide a certain reference for ophthalmologists to distinguish the morphology of iris, and then make rapid diagnosis and clinical intervention for glaucoma.

Reference
[1] Hao H, Zhao Y, Fu H, et al. 2019 Anterior chamber angles classification in anterior segment OCT images via multi-scale regions convolutional neural networks[C] 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) IEEE 849-852.
[2] Fu H, Li F, Sun X, et al. 2020 AGE challenge: Angle Closure Glaucoma Evaluation in Anterior Segment Optical Coherence Tomography [J] Medical Image Analysis 66 101798.
[3] Li W, Chen, Q, Jiang, C, et al. 2021 Automatic Anterior Chamber Angle Classification Using Deep Learning System and Anterior Segment Optical Coherence Tomography Images [J] Transl Vis Sci Technol 10 19.
[4] Riva I, Michelelli, E, Oddone, F, et al. 2020 Anterior Chamber Angle Assessment Techniques: A Review [J] J Clin Med 9 3814
[5] Hao H, Zhao, Y, Yan, Q, et al. 2021 Angle-closure assessment in anterior segment OCT images via deep learning [J] Med Image Anal 69 101956.
[6] Wanichwecharungruang B, Kaothanthong, N, Pattanapongpaiboon, W, et al. 2021 Deep Learning for Anterior Segment Optical Coherence Tomography to Predict the Presence of Plateau Iris [J] Transl Vis Sci Technol 10 7.
[7] Wang B H, Yao, Y F. 2012 Effect of primary iris and ciliary body cyst on anterior chamber angle in patients with shallow anterior chamber [J] J Zhejiang Univ Sci B 13 723-730.
[8] Le C, Baroni, M, Vinnett, A, et al. 2021 Deep Learning Model for Accurate Automatic Determination of Phakic Status in Pediatric and Adult Ultrasound Biomicroscopy Images [J] Transl Vis Sci Technol 9 63.
[9] Zhao Q, Sheng, T, Wang, Y, et al. M2det: 2019 A single-shot object detector based on multi-level feature pyramid network[C] AAAI conference on artificial intelligence 33 9259-9266.
[10] Huang G, Liu, Z, Van Der Maaten, L, et al. 2017 Densely connected convolutional networks[C] IEEE conference on computer vision and pattern recognition 4700-4708.
[11] Simonyan K, Zisserman, A J a p a. 2014 Very deep convolutional networks for large-scale image recognition [J] arXiv preprint arXiv:1409.1556
[12] He K, Zhang, X, Ren, S, et al. 2016 Deep Residual Learning for Image Recognition [C] 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 770-778.
[13] Ronneberger O, Fischer, P, Brox, T. U-net: 2015 Convolutional networks for biomedical image segmentation[C] International Conference on Medical image computing and computer-assisted intervention 234-241.
[14] Schlemper J, Oktay, O, Schaap, M, et al. 2019 Attention gated networks: Learning to leverage salient regions in medical images [J] Med Image Anal 53 197-207.
[15] Pinheiro P H O, Lin, T-Y, Collobert, R, et al. 2016 Learning to Refine Object Segments[C] European conference on computer vision (ECCV) 75-91.