Distinguishing Differences Matters: Focal Contrastive Network for Peripheral Anterior Synechiae Recognition

Yifan Yang[1,2], Huihui Fang[3], Qing Du[1], Fei Li[4], Xiulan Zhang[4], Mingkui Tan[1,2,5(✉)], and Yanwu Xu[3(✉)]

[1] South China University of Technology, Guangzhou, China
mingkuitan@scut.edu.cn
[2] Pazhou Lab, Guangzhou, China
[3] Intelligent Healthcare Unit, Baidu, Beijing, China
ywxu@ieee.org
[4] State Key Laboratory of Ophthalmology, Zhongshan Ophthalmic Center, Sun Yat-sen University, Guangzhou, China
[5] Key Laboratory of Big Data and Intelligent Robot, Ministry of Education, Guangzhou, China

Abstract. We address the problem of Peripheral Anterior Synechiae (PAS) recognition, which aids clinicians in better understanding the progression of the type of irreversible angle-closure glaucoma. Clinical identification of PAS requires indentation gonioscopy, which is patient-contacting and time-consuming. Thus, we aim to design an automatic deep-learning-based method for PAS recognition based on non-contacting anterior segment optical coherence tomography (AS-OCT). However, modeling structural differences between tissues, which is the key for clinical PAS recognition, is especially challenging for deep learning methods. Moreover, the class imbalance issue and the tiny region of interest (ROI) hinder the learning process. To address these issues, we propose a novel Focal Contrastive Network (FC-Net), which contains a Focal Contrastive Module (FCM) and a Focal Contrastive (FC) loss to model the structural differences of tissues, and facilitate the learning of hard samples and minor class. Meanwhile, to weaken the impact of irrelevant structure, we introduce a zoom-in head to localize the tiny ROI. Extensive experiments on two AS-OCT datasets show that our proposed FC-Net yields 2.3%–8% gains on the PAS recognition performance regarding AUC, compared with the baseline models using different backbones. The code is available at https://github.com/YifYang993/FC-Net.

Keywords: PAS recognition · Contrastive loss · AS-OCT · Glaucoma

1 Introduction

Peripheral Anterior Synechiae (PAS), an eye condition in which the iris adheres to the cornea, causes closure of the anterior chamber angle [1,2] and further increases the risk of angle-closure glaucoma [3]. Clinically, experts identify PAS by measuring whether

Y. Yang, H. Fang and Q. Du—Authors contributed equally.

M. de Bruijne et al. (Eds.): MICCAI 2021, LNCS 12908, pp. 24–33, 2021.
https://doi.org/10.1007/978-3-030-87237-3_3

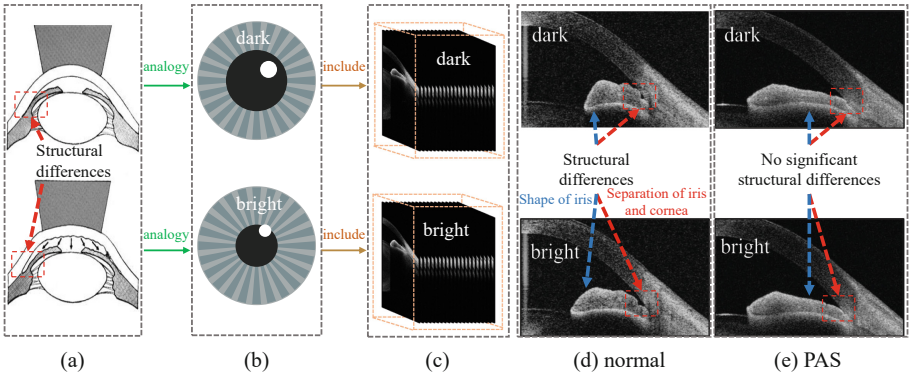

Fig. 1. A description of the motivation for the proposed FC-Net. (a) The structural differences (the dashed boxes) of the eye observed by indentation gonioscopy [9], (b) the eye in different light conditions, (c) the paired AS-OCT sequence, (d) the normal case with the existence of structural differences, (e) the PAS case without a significant structural difference. (Color figure online)

the adhesion between iris and cornea could be disrupted (*i.e.,* structural differences, as shown in Fig. 1(a)) using indentation gonioscopy [4]. However, this diagnostic approach is patient-contacting. Anterior segment optical coherence tomography (AS-OCT), with the characteristics of easy-to-perform and non-contact, has become a critical tool to evaluate the anterior segment of an eye [5]. Thus, we seek to design an AS-OCT based algorithm to detect the PAS automatically.

Currently, there are a few deep-learning-based research endeavors pertaining to AS-OCT based automatic recognition of PAS, whereas several attempts have been made to recognize glaucoma related disease. Fu *et al.* combined clinical parameters and multi-scale features to improve the accuracy of a angle-closure screening method [6]. Hao *et al.* proposed a coarse-to-fine method to localize anterior chamber angle regions and then classified them into primary open-angle glaucoma (POAG), primary angle-closure suspect (PACS) and primary angle-closure glaucoma (PACG) [7]. Recently, Hao *et al.* proposed a novel MSDA block and multi-loss function to classify AS-OCT images into POAG, PACG and synechiae [8]. Although these deep learning methods can improve the performance of automatic recognition of glaucoma related diseases, they lack information regarding the structural differences of an eye (see Figs. 1(a), 1(d) and 1(e)), which is critical in identifying PAS in clinic [3]. Moreover, these methods ignore the class imbalance problem, which is unavoidable due to the difficulty in collecting positive data (*e.g.,* In the two AS-OCT datasets used in the experiment, the PAS samples accounted for 7.5% and 30.4%, respectively).

In view of the shortcomings of the above methods, we propose a novel framework to recognize PAS by considering distinguishing the structural differences of an eye, which requires a paired data representing conditions during indentation gonioscopy and a strategy to measure the differences between the paired data. Inspired by the diagnostic processes of doctors, PAS is identified if the adhesion of iris and cornea cannot be separated using indentation gonioscopy [4]. We introduce a medical prior [3], which states that the structural differences of the eye observed by indentation gonioscopy (Fig. 1 (a))

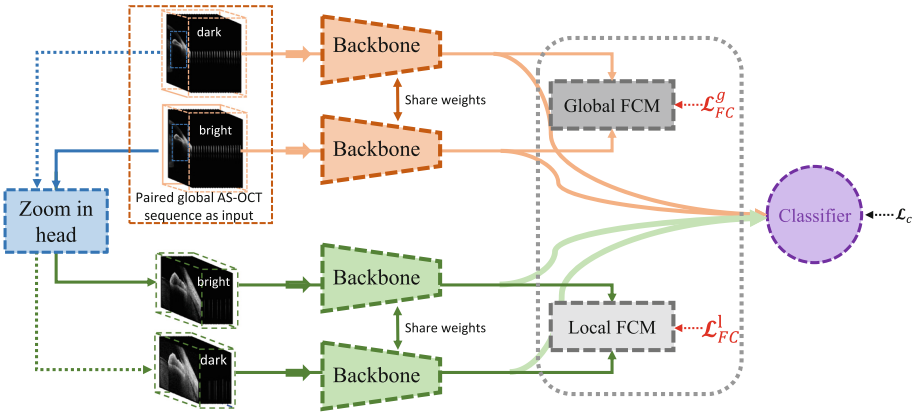

Fig. 2. Our proposed framework takes paired AS-OCT sequence under bright and dark conditions as input. The zoom-in head crops the discriminative area. Differences of the paired AS-OCT sequence are distinguished by means of the focal contrastive module in global and local views.

is analogized to the structural difference of an eye in different light conditions (Fig. 1 (b)), which can be observed from the paired AS-OCT sequence (Fig. 1 (c)). In normal cases, bright conditions cause pupillary constriction, which results in separation of iris and cornea (see the red dashed boxes in Fig. 1 (d)). For patients with PAS, the adhesion of iris and cornea prevents the separation (Fig. 1 (e)).

To distinguish the differences between the paired AS-OCT sequence for PAS recognition, we design a Focal Contrastive Network (FC-Net) with a focal contrastive (FC) loss. The network consists of two modules including a zoom-in head module and a Focal Contrastive Module (FCM). The zoom-in head module is designed to address the difficulty in capturing the tiny region, where the iris and cornea intersect. The main idea of FCM is to distinguish differences of features from two light conditions, and to guide the network to focus on hard samples and minor class by the designed FC loss. The major contributions of this paper are threefold. (1) We introduce a medical prior of clinical PAS recognition to the deep-learning-based method, and experimentally demonstrate that this medical prior knowledge can enhance the performances. (2) We design a novel focal contrastive module with a focal contrastive loss, which can distinguish the differences between the paired AS-OCT images, as well as alleviate the class imbalance problem and focus on the hard samples. (3) Our proposed FC-Net can be implemented with different backbones, and the promising experimental results show that the FC-Net has strong generalizability and can be applied to 3D data such as video or CT volume for difference measurement.

2 Methodology

Dividing an eye into 12 paired sectors (Fig. 3 (a1)–(a2)), a paired sectors (*e.g.*,2–3 o'clock, shown as the yellow fan area) of the eye under dark and bright conditions is represented by a paired AS-OCT sequence, *i.e.*,2×21 AS-OCT images (Fig. 1 (c)) [2]. We aim to classify the 12 paired sectors of an eye into normal or PAS, respectively.

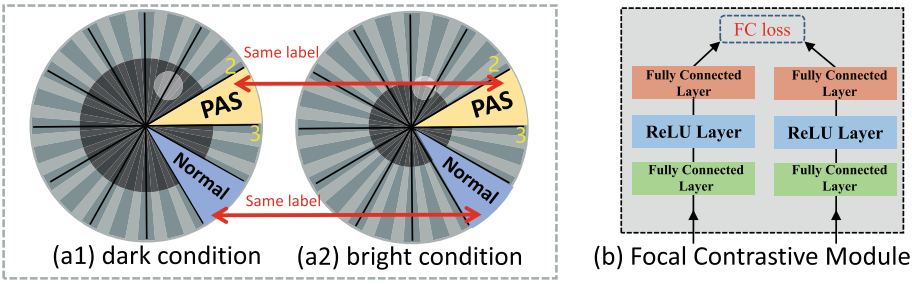

Fig. 3. (a1–a2) Illustration of the paired sectors of an eye in different light conditions. Colored paired sectors shares same label and is represented by paired AS-OCT sequence. (b) Overview of the focal contrastive module (FCM). (Color figure online)

The pipeline of our FC-Net is illustrated in Fig. 2 given a paired AS-OCT sequence $\{\mathcal{X}_d, \mathcal{X}_b\}$ that represents a paired sectors of an eye from different light conditions. First, a zoom-in head module is implemented to extract the discriminative area by cropping $\{\mathcal{X}_d, \mathcal{X}_b\}$ into local version $\{\mathcal{X}_d^l, \mathcal{X}_b^l\}$. After the dual stream backbone, the learned features $\{\mathcal{H}_d, \mathcal{H}_b\}$, and $\{\mathcal{H}_d^l, \mathcal{H}_b^l\}$ are sent to the focal contrastive module(FCM) to distinguish the differences between two light conditions ($C(\mathcal{H}_d, \mathcal{H}_b) \rightarrow \mathcal{L}_{FC}^g, C(\mathcal{H}_d^l, \mathcal{H}_b^l) \rightarrow \mathcal{L}_{FC}^l$), generating corresponding global focal contrastive (FC) loss $\mathcal{L}_{FC}^g$ and local FC loss $\mathcal{L}_{FC}^l$ which measure the difference between the paired AS-OCT sequence in global and local views. Note that the FC loss in FCM introduces two factors to guide the contrastive module to focus on hard samples and minor class simultaneously. Finally, the supervised information of FC loss ($\mathcal{L}_{FC}^l$ and $\mathcal{L}_{FC}^g$) and classification loss ($\mathcal{L}_c$) is fused by the decay learning strategy, which iteratively decreases the weight of FC loss. The overall training procedure of PAS recognition aims to minimize the following objective function:

$$\mathcal{L}_{all} = \mathcal{L}_c + decay(\mathcal{L}_{FC}),\ \mathcal{L}_{FC} = \mathcal{L}_{FC}^g + \mu\mathcal{L}_{FC}^l \tag{1}$$

where *decay* is our proposed decay learning strategy to balance between $\mathcal{L}_c$ and $\mathcal{L}_{FC}$, and μ is a given ratio between local FC loss and global FC loss. The details of our FC-Net are illustrated as follows.

2.1 Focal Contrastive Neural Network

The focal contrastive network (FC-Net) contains a zoom-in head module for cropping the AS-OCT sequence, a dual-stream backbone for learning representation of the paired AS-OCT sequence and a focal contrastive module for measuring the differences between the dark and bright conditions.

Zoom-in head: Clinically, PAS is identified by the adhesion of iris and cornea [4], yet the adhesion region is tiny comparing with an eye. Hence, we introduce a localization method termed the zoom-in head to crop the tiny ROI of the AS-OCT sequence, based on an observation that the horizontal centerline of the ROI is close to the horizontal line with max sum of pixels. Consider a matrix X representing an image, the zoom-in head

module traverses each row of X to search the row with the largest sum of pixels (*i.e.,* $\max_i\{\sum_j X_{ij}\}$), then we set the row as the horizontal centerline of the ROI. With the centerline, we crop X at a fixed size to obtain the ROI. The original and the cropped sequence are simultaneously learned by the dual-stream backbone.

Dual Stream Backbone: Motivated by an observation that the AS-OCT sequence in different light conditions differ on both global scale (*e.g.,* change in lens size) and local scale (*i.e.,* adhesion and separation of the conjunction of iris and cornea), we design a dual-stream neural network that includes global and local information of the paired data. The backbone aims to extract features of the sequence from multiple levels in the two light conditions. Specifically, the global and local sequence ($\mathcal{X}_d$ and $\mathcal{X}_d^l$) captured in the dark condition are fed into a global dark stream (*i.e.,*$\mathcal{X}_d \rightarrow \mathcal{H}_d$) and a local dark stream (*i.e.,*$\mathcal{X}_d^l \rightarrow \mathcal{H}_d^l$), and their counterparts ($\mathcal{X}_b$ and $\mathcal{X}_b^l$) work in the same manner. Note that the backbones of the same scale share weights to benefit from the corresponding representation learning and to reduce computational complexity.

Focal Contrastive Module: Modeling the difference between the dark and bright sequence is the core procedure for PAS recognition, which is achieved by our proposed FCM, as shown in Fig. 3 (b). Taking the global branch as an example, the feature vectors $\mathcal{H}_d^g$, $\mathcal{H}_b^g$ are fed into FCM to generate representation vectors of corresponding AS-OCT sequence (*i.e.,*$\mathcal{C}_d$, and $\mathcal{C}_b$). Then, the difference, *i.e.,*Euclidean distance d between $\mathcal{C}_d$ and $\mathcal{C}_b$, is selected by the mining strategy in [10] and then measured by the FC loss.

2.2 Designed Loss Function

Focal Contrastive Loss: The novel FC loss is designed to measure the difference of the paired AS-OCT sequence and to alleviate the class imbalance problem. Inspired by contrastive loss [11] and focal loss [12], the FC loss is designed as follows.

$$\mathcal{L}_{FC} = \frac{S}{m}\sum_{i=1}^{m}\left\{\alpha^{\mathcal{P}}\sum_{y_i=1}\beta_i^{\mathcal{P}}d_i^2 + \alpha^{\mathcal{N}}\sum_{y_i=0}\beta_i^{\mathcal{N}}\left([M-d_i]_+\right)^2\right\} \tag{2}$$

where m denotes the number of training samples, S is a scale factor that determines the largest scale of the FC loss, and M is a fixed margin for similarity separation, which is set as 2 following [11]. d is the Euclidean distance that measures similarity between a data pair (*e.g.,*paired AS-OCT sequence from different light conditions). β and α are a focal factor and a balance factor following the focal loss style. β is formulated as:

$$\beta^{\mathcal{P}} = [\text{sigmoid}\,(d)]^2,\ \beta^{\mathcal{N}} = [1-\text{sigmoid}\,(d)]^2 \tag{3}$$

where $\mathcal{P}$ and $\mathcal{N}$ denote PAS and normal samples(*i.e.,* similar pairs and dissimilar pairs).

β measures the amount of information that a data pair contains. Starting from scaling non-negative value d to range $[0.5, 1)$ by a sigmoid function, we then measure the information of each pair according to how far the distance d is from the expected distance by Eq. 3 (*e.g.,* the expected distance of a pair is near 1, thus a hard dissimilar pair with a close distance contains much information). We specify a property of the focal factor. When a similar pair is a hard one and leads to sigmoid(d) near 1, the corresponding $\beta^{\mathcal{P}}$ approaches 1. In this way, the loss for the hard similar pair is up-weighted.

Table 1. The statistics of AS-I and AS-II. Note that a sample in the table denotes a paired sequence (Fig. 1 (c)) consisting of 2×21 AS-OCT images, except for the last column, which represents the number of eyes.

Dataset	Set	Category		#Total	#Eye
		#Normal	#PAS		
AS-I	Train	212	64	276	23
	Test	202	62	264	22
AS-II	Train	2109	159	2268	189
	Test	2099	157	2256	188

α balances the impact of the similar and dissimilar pairs. Since class imbalance problem leads the training to be dominated by the dissimilar pairs, we design $\alpha^{\mathcal{P}}$ and $\alpha^{\mathcal{N}}$ subject to $\alpha^{\mathcal{P}} + \alpha^{\mathcal{N}} = 1$. Specifically, for a normal pair that occupies the majority, $\alpha^{\mathcal{N}}$ is set to less than 0.5 and hence down-weights the impact of the major class.

Decay Learning Strategy: To use the information obtained from the structural differences, we design a decay learning strategy to fuse the above FC loss and a classification loss. In our network, the α-balanced focal loss [12–14] is used as the classification loss. Specifically, recall that the overall loss is formulated as in Eq. 1, where *decay* controls the impact of FC loss. We down-weight the ratio of FC loss by:

$$decay = 1 - \left(\frac{T}{T_{\max}}\right) \tag{4}$$

where T denotes the current training epoch, and $T_{\max}$ denotes the total epochs. During the training procedure, the decay learning strategy aims to first learn an embedding by distinguishing the difference via the FC loss and then gradually focus on the classification by making use of the difference. Thus, in our designed decay strategy, as $T \rightarrow T_{max}$, *decay* goes to 0 and the impact of FC loss is down-weighted.

3 Experimental Results

Datasets: Our experiments utilize AS-I and AS-II, which are captured from 377 eyes using a CASIA II machine and from 45 eyes using a CASIA I machine, respectively. Each eye is collected in bright and dark conditions to obtain bright and dark volumes, and contains 2×128 AS-OCT images with 800×1876 pixels in AS-II, with 800×1000 in AS-I. Differences between the AS-OCT images in these two datasets lie in noise, iris and crystalline lens, which lead to variation in prediction performance. Following [15], we crop 128 images of an AS-OCT volume into 256 left/right images. Then, 256 images are divided into 12 AS-OCT sequence, each of which denotes a sector of an eye under dark or bright light condition. The statistics of the category distribution, as well as the data distribution in training and testing processes, are shown in Table 1. Note that we carefully checked the datasets to ensure that there was no patient overlap between the training and testing sets.

Table 2. Performance comparison based on two datasets. The suffix "*-FC" represents FC-Net with "*" as the backbone, "*-B(D)" denotes methods taking bright(dark) sequence as input only, and "*-C" indicates baselines concatenating features extracted from paired AS-OCT sequence.

Method	AS-I		AS-II	
	AUC	F1	AUC	F1
SMA-Net-D [8]	0.8675	0.6400	0.9271	0.4013
SMA-Net-B [8]	0.8521	0.5856	0.9024	0.5086
S3D-D [16]	0.8084	0.5492	0.8581	0.3402
S3D-B [16]	0.8204	0.5665	0.8532	0.3303
S3D-FC(ours)	**0.8986**	**0.6853**	**0.9381**	**0.5428**
I3D-D [17]	0.8713	0.6345	0.9360	0.4945
I3D-B [17]	0.8403	0.6341	0.9340	0.5056
I3D-FC(Ours)	**0.8986**	**0.6906**	**0.9599**	**0.5646**
NL I3D-D [18]	0.8907	0.6667	0.9082	0.4103
NL I3D-B [18]	0.8897	0.6212	0.9162	0.3657
NL I3D-C(Ours)	0.9016	**0.6667**	0.9359	0.4217
NL I3D-FC(Ours)	**0.9342**	0.6526	**0.9401**	**0.4778**

Implementation Details: We implement FC-Net via Pytorch. During training, we use an SGD optimizer with learning rate $= 10^{-3}$ and weight decay $= 10^{-4}$ on NVIDIA TITAN X GPU. The learning rate is decayed by 0.01 at the 120-th and 160-th epochs. We implement the decay learning strategy to trade off between classification loss ($\mathcal{L}_C$) and FC loss ($\mathcal{L}_{FC}$). For the trade-off parameters, we set $\mu = 0.5$, $M = 2$, and $\alpha = 0.25$ on all datasets, and set $S = 16$ and $S = 24$ for AS-I and AS-II, respectively.

Baselines: We compare our FC-Net with 4 3D deep learning models: S3D [16], I3D [17], NL I3D [18], and SMA-Net [8]. For fair comparisons, we evaluate each model on both dark and bright AS-OCT sequence, which are termed as "*-D" and "*-B". In addition, "*-C" denotes a variation of our proposed method that simply concatenates features of dark and bright sequence, while our proposed method "*-FC" considers the difference between both AS-OCT sequence. Note that "*" denotes different backbones.

Metrics: We use AUC and F_1-score as the metrics. Both metrics are robust with respect to the class imbalance problem. The AUC shows desirable properties in the binary classification task [19], and the F_1-score can trade off between precision and recall [20].

3.1 Overall Performance

Table 2 shows the PAS recognition results of different methods under different light conditions for two datasets. Specifically, on AS-II, when using the AS-OCT sequence captured only in dark or bright condition, the best performance of the methods with SMA-Net, S3D, I3D and NL I3D as backbones is the $AUC = 0.9360$ obtained by the I3D-D method. When adding FCM to I3D, the corresponding performance is improved by 2.39%. Similarly, on AS-I, using the proposed FCM, the PAS recognition performance is increased by 4.35% over the NL I3D-D method, which exhibits the best performance among the methods when using only one light condition. Besides, as illustrated

Table 3. Ablation studies performed on AS-I and AS-II in terms of AUC, where ZIH means zoom-in head, FCM indicates the focal contrastive module with FC loss, and DSL denotes decay learning strategy.

Component			Dataset	
ZIH	FCM	DLS	AS-I	AS-II
			0.8395	0.8784
✓			0.9003	0.9425
✓	✓		0.9177	0.9540
✓	✓	✓	**0.9342**	**0.9599**

Table 4. Effect of α in FC loss on the performance of NL I3D-FC in terms of AUC (on AS-I).

$\alpha^{\mathcal{N}}$,$\alpha^{\mathcal{P}}$	0.25, 0.75	0.5, 0.5	0.75, 0.25
AUC	**0.9342**	0.9107	0.8753

in the table, our proposed FC-Net can be added to different backbones, and outperforms the one stream counterpart by 2.3%–8% in terms of AUC. Such results demonstrate the superiority as well as the generalizability of our FC-Net. Thus, these experimental results show that distinguishing the differences between dark and bright sequence with our proposed FCM and FC loss boosts the performance of PAS recognition.

To verify the effect of only adding AS-OCT sequence of two illumination conditions, we use the NL I3D backbone to extract the features of the paired sequence and concatenate them (NL I3D-C) for PAS recognition. From Table 2, NL I3D-C achieves 2.77% and 1.97% gains compared with NL I3D-D and NL I3D-B in terms of AUC on AS-II, respectively. A similar conclusion can be drawn from AS-I, which demonstrates the effectiveness of the clinical prior, i.e., the differences of an eye observed by clinical indentation gonioscopy are analogized to the difference of a paired AS-OCT sequence captured under dark and bright conditions.

3.2 Ablation Studies and Parameter Discussion

In this section, we discuss the ablation studies (Table 3) and the selection of α (Table 4). We choose I3D as the backbone for AS-II, and NL I3D is set as the backbone for AS-I. In Table 3, the method in the first line simply concatenates the features of the AS-OCT sequence from different light conditions. The table demonstrates that the model with the zoom-in head module (ZIH) achieves 6.41% and 6.08% gains regarding AUC on AS-II and AS-I compared with that without ZIH, since ZIH can introduce the discriminative information by cropping the AS-OCT sequence. Based on ZIH, the focal contrastive module with FC loss (FCM) further increases AUC by 1.15% and 1.74% on AS-II and AS-I respectively, by modeling the structural differences of tissues. Moreover, the decay learning strategy (DLS) also boosts the model performance owint to its characteristic of gradually down-weighting FC loss that focuses on modeling structural differences instead of classification. In addition, we introduce parameter α in FC loss to

re-weight the minority class. To further evaluate the impact of α, we discuss its sensitivity at different values while retaining other parameter settings. According to Table 4, when we reduce $\alpha^{\mathcal{N}}$ from 0.75 to 0.25, the minor class in Eq. 2 is up-weighted and the re-weighting procedure leads to improvement of the experimental results.

4 Conclusion

In this paper, we demonstrated that the paired AS-OCT sequence captured from dark and bright conditions can provide fruitful information for PAS recognition. Additionally, we proposed a FC-Net to model the difference of the paired data, which contains an FCM and an FC loss to distinguish the difference of tissues accompanied by up-weighting hard samples and minor class, and a zoom-in head to crop the discriminative area of the AS-OCT sequence. Our FC-Net outperforms the state-of-the-art methods on both AS-II and AS-I. The ablation study shows that the modules designed in this paper are meaningful and effective. Moreover, the proposed FCM with the FC loss can be applied to more tasks aiming at identification of differences between paired 3D data such as video and CT volume, which we leave to our future work.

Acknowledgements. This work was partially supported by Key Realm R&D Program of Guangzhou (202007030007), National Natural Science Foundation of China (NSFC) 62072190, Program for Guangdong Introducing Innovative and Enterpreneurial Teams 2017ZT07X183, Fundamental Research Funds for the Central Universities D2191240, Guangdong Natural Science Foundation Doctoral Research Project (2018A030310365), International Cooperation open Project of State Key Laboratory of Subtropical Building Science, South China University of Technology (2019ZA02).

References

1. Foster, P.J., Aung, T., et al.: Defining "occludable" angles in population surveys: drainage angle width, peripheral anterior synechiae, and glaucomatous optic neuropathy in East Asian people. Brit. J. Ophthalmol. **88**(4), 486–490 (2004)
2. Lee, J.Y., Kim, Y.Y., Jung, H.R.: Distribution and characteristics of peripheral anterior synechiae in primary angle-closure glaucoma. Korean J. Ophthalmol. **20**(2), 104–108 (2006)
3. Lai, I., Mak, H., Lai, G., Yu, M., Lam, D.S., Leung, C.: Anterior chamber angle imaging with swept-source optical coherence tomography: measuring peripheral anterior synechia in glaucoma. Ophthalmology **120**(6), 1144–1149 (2013)
4. Forbes, M.: Gonioscopy with corneal indentation: a method for distinguishing between appositional closure and synechial closure. Arch. Ophthalmol. **76**(4), 488–492 (1966)
5. Ang, M., Baskaran, M., et al.: Anterior segment optical coherence tomography. Prog. Retinal Eye Res. **66**, 132–156 (2018)
6. Fu, H., Xu, Y., et al.: Multi-context deep network for angle-closure glaucoma screening in anterior segment OCT. In: Frangi, A., Schnabel, J., Davatzikos, C., Alberola-Lopez, C., Fichtinger, G. (eds) Medical Image Computing and Computer Assisted Intervention, vol. 11071, pp. 356–363. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-00934-2_40
7. Hao, H., Zhao, Y., Fu, H., Shang, Q., Li, F., Zhang, X., Liu, J.: Anterior chamber angles classification in anterior segment oct images via multi-scale regions convolutional neural networks. In: International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 849–852 (2019)

8. Hao, H.: Open-appositional-synechial anterior chamber angle classification in AS-OCT sequences. In: Martel, A.L., et al. (eds) Medical Image Computing and Computer Assisted Intervention, vol. 12265, pp. pp. 715–724. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-59722-1_69
9. Friedman, D.S., He, M.: Anterior chamber angle assessment techniques. Surv. Ophthalmol. **53**(3), 250–273 (2008)
10. Wang, X., Han, X., Huang,W., Dong, D., Scott, M.R.: Multi-similarity loss with general pair weighting for deep metric learning. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 5022–5030 (2019)
11. Hadsell, R., Chopra, S., LeCun, Y.: Dimensionality reduction by learning an invariant mapping. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 2, pp. 1735–1742 (2006)
12. Lin, T.Y., Goyal, P., Girshick, R., He, K., Dollar, P.: Focal loss for dense object detection. IEEE Trans. Pattern Anal. Mach. Intell. **42**(2), 318–327 (2020)
13. Zhang, Y., Wei, Y., et al.: Collaborative unsupervised domain adaptation for medical image diagnosis. IEEE Trans. Image Process. **29**, 7834–7844 (2020)
14. Zhang, Y., Chen, H., et al.: From whole slide imaging to microscopy: deep microscopy adaptation network for histopathology cancer image classification. In: Shen, D., et al. (eds) Medical Image Computing and Computer Assisted Intervention, vol. 11764, pp. 360–368. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-32239-7_40
15. Fu, H., Xu, Y., et al.: Segmentation and quantification for angle-closure glaucoma assessment in anterior segment OCT. IEEE Trans. Med. Imaging **36**, 1930–1938 (2017)
16. Xie, S., Sun, C., Huang, J., Tu, Z., Murphy, K.: Rethinking spatiotemporal feature learning: speed-accuracy trade-offs in video classification. In: The European Conference on Computer Vision, pp. 318–335 (2018)
17. Carreira, J., Zisserman, A.: Quo vadis, action recognition? A new model and the kinetics dataset. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 4724–4733 (2017)
18. Wang, X., Girshick, R., Gupta, A., He, K.: Non-local neural networks. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 7794–7803 (2018)
19. Bradley, A.P.: The use of the area under the ROC curve in the evaluation of machine learning algorithms. Pattern Recogn. **30**(7), 1145–1159 (1997)
20. Powers, D.: Evaluation: from precision, recall and F-factor to ROC, informedness, markedness and correlation. Mach. Learn. Technol. 2 (2008)