2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI) | 978-1-6654-2923-8/22/$31.00 ©2022 IEEE | DOI: 10.1109/ISBI52829.2022.9761687

PERIPAPILLARY ATROPHY SEGMENTATION BASED ON ASM LOSS

Mengxuan Li[⋆] *He Zhao*[⋆] *Jie Xu*[†] *Huiqi Li*[⋆]

[⋆] Beijing Institute of Technology
[†] Beijing Tongren Hospital

ABSTRACT

Peripapillary atrophy (PPA) is a clinical finding that reflects atrophy of the retinal layer and retinal pigment epithelium and it is a symptom related to many diseases. The shape and region area indicate the pathologic severity of myopia and glaucoma. So it is important to segment PPA area to analyze the progression of these diseases. The shape of PPA area is mostly crescent. In order to combine the prior knowledge, a PPA segmentation method is proposed with a novel active shape model (ASM) loss. The shape constraint is introduced to improve the segmentation accuracy. It is realized by the proposed ASM loss module, which contains three parts, θ predictor, **b** predictor and an affine transformation module. Our approach is evaluated on a clinical dataset. Extensive experiments demonstrate that our method provides good performance both qualitatively and quantitatively.

Index Terms— Active shape model (ASM), Peripapillary atrophy (PPA), Segmentation

1. INTRODUCTION

Peripapillary atrophy (PPA) is also called the optic disc crescent, and it is a symptom related to many diseases such as myopia, glaucoma, and intracranial tumors [1]. PPA can provide an alarm for implementing an intervention to slow down the disease progression at an early stage [2]. Therefore, monitoring PPA area is very helpful for screening some eye diseases. PPA area can be observed through non-invasive fundus images, but manual detection and quantification of PPA in retinal fundus images are tedious and time-consuming. Based on the above reasons, it is very necessary to propose a fully automatic PPA segmentation algorithm.

Some methods of PPA segmentation have been proposed. Lu et al. [3] proposed a method that combines scanning filter, thresholding, region growing as well as a modified Chan-Vese model with a shape constraint to segment and quantify optic disk (OD) and PPA. Li et al. [4] utilized multiple evenly oriented radial lines to detect OD and PPA boundary simultaneously, which profile PPA region. Chai et al. [5] proposed a PPA segmentation using a multi-task fully convolutional network.

Corresponding author: Huiqi li (huiqili@bit.edu.cn).

Convolutional neural network (CNN) combined with shape constraint has been utilized to further improve medical image segmentation. Oktay et al. [6] proposed a generic training strategy that incorporates anatomical prior knowledge into CNN through a new regularization model using an encoding scheme. Fan et al. [7] proposed a method which incorporates active shape model (ASM)-derived information for the segmentation of the intra-cochlear anatomy in head CT images. Tilborghs et al. [8] performed left ventricle and myocardial segmentation by regression of pose and shape parameters derived from a statistical shape model.

In this paper, we propose a PPA segmentation method based on ASM loss. The contributions of our approach can be summarized as follows. (1) Unlike previous works that need to detect the PPA and OD areas simultaneously, we directly segment the PPA area. (2) Because the PPA area is mostly crescent-shaped, we are inspired by ASM method [9] and propose the ASM loss which is a shape constrain by cascading ASM loss module after the baseline segmentation network. (3) We have carried out extensive experiments on a private clinical dataset of 50 fundus images to evaluate our approach. The experimental results show that our approach achieves superior performance compared with other methods.

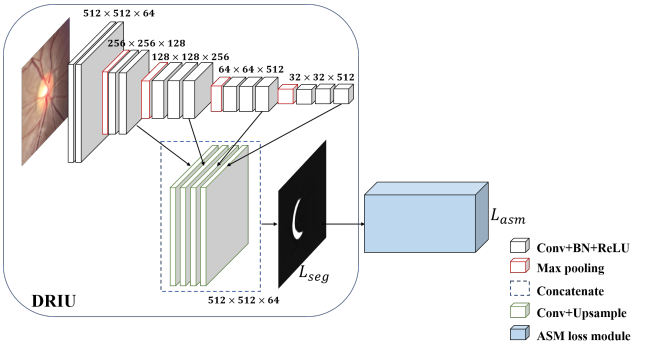
2. METHODOLOGY

We propose a novel PPA segmentation method learning with ASM loss that adds shape constraints by cascading a module after the baseline segmentation network. In this section, we will first introduce the active shape model (ASM) which leads to the proposed ASM loss and describe the whole method we use to segment PPA.

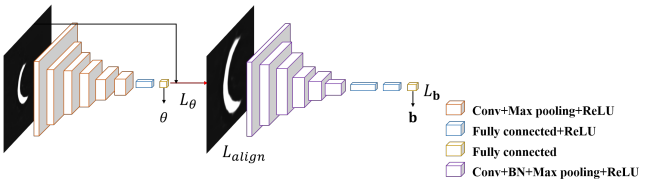
2.1. Active Shape Model (ASM)

ASM is a point-based distribution model. The structure of the image can be represented by a series of dots. This means that ASM can be used to extract feature points of objects and can be used as a form of representing object features.

In the theory of ASM, an image is selected as the standard reference at the beginning, and all the data including the standard one is sampled by n pairs of sampling points $\{(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)\}$. Next, each image is aligned



(a) The general structure of segmentation network



(b) Structure diagram of ASM loss module

Fig. 1. The overview framework for PPA segmentation.

to standard reference by an alignment operation performed with translation, rotation and scaling in order to remove the non-shape factors. The calculating equation is as follows:

$$\begin{bmatrix} x^{'} \\ y^{'} \end{bmatrix} = s \begin{bmatrix} cos(\alpha) & sin(\alpha) \\ -sin(\alpha) & cos(\alpha) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} x_m \\ y_m \end{bmatrix} \rightarrow \begin{bmatrix} x_{std} \\ y_{std} \end{bmatrix} \quad (1)$$

$$\begin{bmatrix} x_m \\ y_m \end{bmatrix} = \begin{bmatrix} \bar{x} - W/2 \\ \bar{y} - H/2 \end{bmatrix} \quad (2)$$

where $x = [x_1, x_2, ..., x_n]$, $y = [y_1, y_2, ..., y_n]$ represent the horizontal and vertical coordinates set of n samples respectively, $x^{'}$ and $y^{'}$ represent the set of the horizontal and vertical coordinates of the aligned sampling points, and the subscript *std* represents the standard reference. The parameters of the alignment operation can be calculated by the least square method.

Next, the shape model is established. The shape model is obtained by PCA on the shape descriptor set of all training data to perform dimension reduction. The expression of the shape descriptor is shown in (3).

$$\mathbf{C} = [x_1^{'}, x_2^{'}, ..., x_n^{'}, y_1^{'}, y_2^{'}, ..., y_n^{'}] \quad (3)$$

After dimensionality reduction, we obtain k maximum eigenvectors $\mathbf{P}$ and eigenvalues $\mathbf{\Lambda}$, where $\mathbf{P}= [p_1, p_2, ..., p_k]$, k is the dimension retained after dimensionality reduction, $\mathbf{\Lambda}= diag(\lambda_1, \lambda_2, ...\lambda_k)$.

The shape coefficient $\mathbf{b}$ is obtained by the following projection on k shape modes. By truncating each b_i in $\mathbf{b}$ within $\pm a\sqrt{\lambda_i}$.

$$\mathbf{b} = [p_1, p_2, ..., p_k]^T(\mathbf{C} - \bar{\mathbf{C}}) = \mathbf{P}^T(\mathbf{C} - \bar{\mathbf{C}}) \quad (4)$$

The corresponding constrained shape can be obtained through the inverse operation of the above formula:

$$\mathbf{C} = \bar{\mathbf{C}} + \mathbf{Pb} \quad (5)$$

The results obtained through the above processing will be used as the ground truth in Sec. 2.2.

2.2. ASMSeg

Fig. 1 shows the architecture of our proposed method called ASMSeg, which involves the DRIU [10] as baseline and introduces ASM loss by cascading an ASM loss module. The ASM loss module contains two encoders and an affine transformation operator as shown in Fig. 1(b).

The first encoder named as θ predictor learns the parameters of the affine matrix which is described in Eq. (6). The meaning of the parameters in Eq. (6) is the same as in Eq. (1), which is the concatenation of two matrices in Eq. (1).

$$\theta = \begin{bmatrix} s \cdot cos(\alpha) & s \cdot sin(\alpha) & x_m \\ -s \cdot sin(\alpha) & s \cdot cos(\alpha) & y_m \end{bmatrix} \quad (6)$$

After the affine matrix is obtained, it is applied to the output of baseline segmentation, so as to obtain the alignment result of removing the influence of angle, size and position. The second encoder called $\mathbf{b}$ predictor encrypts the aligned result into shape coefficient $\mathbf{b}$, where the meaning of $\mathbf{b}$ is the same as Eq. (4).

The differences between the encoder predictions of θ and $\mathbf{b}$ and their ground truths for the labelled segmentation shape, both obtained using Eq. (6) and (4), are calculated separately as the loss terms. Finally, the proposed ASM loss can be formulated as:

$$L_{ASM} = L_\theta + 10 \times L_{\mathbf{b}} + L_{align} \quad (7)$$

where L_θ and $L_{\mathbf{b}}$ are loss terms updating the encoders. In addition, we also include an align loss L_{align} to restrict the alignment, which can further improve the accuracy of parameters in the affine matrix. Among them, $L_{\mathbf{b}}$ must be multiplied by the hyperparameters to make the above three items all at the same magnitude. Empirically, the hyperparameter is selected as 10.

The calculation formulas of L_θ, L_{align} and $L_{\mathbf{b}}$ are shown in Eq. (8)-(10), where θ, Y_{align} and $\mathbf{b}$ represent the ground truth of affine matrix, alignment result and shape coefficient respectively, and the parameters with $\hat{}$ represent the predicted result.

$$L_\theta = \sum_{i=1}^{n} (\theta_i - \hat{\theta}_i)^2 \quad (8)$$

$$L_{align} = 1 - \frac{2|Y_{align} \cap \hat{Y}_{align}|}{|Y_{align}| + |\hat{Y}_{align}|} \quad (9)$$

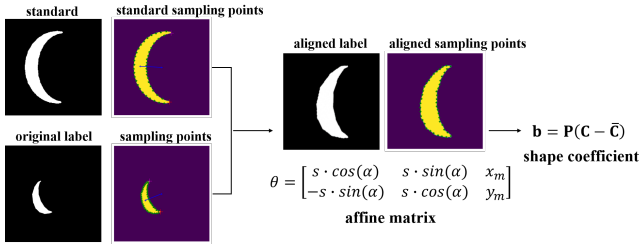


Fig. 2. The flow chart of ground truth generation process.

$$L_{\mathbf{b}} = \sum_{i=1}^{n} (b_i - \hat{b}_i)^2 \tag{10}$$

The final loss function is shown as follows. It contains two parts, segmentation loss $\mathcal{L}_{seg}$ and ASM loss L_{ASM}:

$$L = L_{seg} + L_{ASM} \tag{11}$$

$$L_{seg} = 1 - \frac{2|Y \cap \hat{Y}|}{|Y| + |\hat{Y}|} \tag{12}$$

where Y and $\hat{Y}$ represent the ground truth and the segmentation result. The input of the ASM loss module is Y and the operations in it are all derivable, so ASM loss will impose constraints on baseline segmentation network (DRIU) in backpropagation.

3. EXPERIMENTS AND RESULTS

In this section, we first introduce the dataset and pre-processing steps. The implementation details and the evaluation metrics are given. After that, we compare our method with other state-of-the-art methods.

3.1. Dataset description and Pre-processing

The dataset we use is provided by Beijing Tongren Hospital, which contains 200 clinical data. This dataset is collected from primary school students in grades one to six with a crescent PPA shape.

We randomly select 50 images as testing set, and the rest as training set. For all the data, we first normalize them to the right eye and extract the region of interest (ROI), and then resize them to 512×512. ROI extraction first uses the method of [11] to locate the center of OD, and then crop the ROI at 0.4 times the height of the fundus image.

The ground truth includes four parts, the PPA labeled by experienced ophthalmologists, the affine matrix, the aligned area of PPA region and the shape coefficient. The method to obtain the ground truth refers to Sec. 2.1. The flow chart of ground truth is shown in Fig. 2.

3.2. Implementation Details and Metrics

The code of our method is implemented using PyTorch. For all experiments, 200 epochs are trained. The first 100 epochs have a fixed learning rate of 0.0001, and the learning rate for the last 100 epochs are linearly decreased, and finally decreased to 0. During training, all experiments are optimized using Adam optimizer with the batch size of 8. And when processing the data to generate the ground truth, the information retention rate in PCA is selected as 99%, so the final shape coefficient vector length is 12.

We use dice coefficient, precision (P), recall (R), accuracy (ACC), IOU and mAP to evaluate PPA segmentation performance. Since this segmentation task has only one target area of PPA, mAP is equivalent to AP. Both Dice and IOU are metrics to measure the similarity between the segmentation results and the ground truth.

3.3. Comparison with other methods

We compare our method with DRIU [10], ACNN [6], Li et al. [4], and ASM post-processing [9]. In the comparison experiment, the regularization model in ACNN has the same structure as the **b** predictor of our ASM loss module in order to maintain fairness. The standard image, point sampling method and ASM model in the post-processing method are all the same as our method.

Table 1. Quantitative results compared with other methods

	DRIU [10]	ACNN [6]	Li et al. [4]	ASM post processing [9]	Ours w/o L_{align}	Ours
Dice	0.8006	0.8056	0.6370	0.8032	0.8071	**0.8104**
P	0.8084	0.8068	0.6319	**0.8256**	0.8125	0.8030
R	0.8049	0.8159	0.6813	0.7946	0.8122	**0.8267**
ACC	0.9906	0.9907	0.9793	0.9909	0.9908	**0.9909**
IOU	0.6729	0.6793	0.4859	0.6769	0.6810	**0.6859**
mAP	0.7745	0.8257	-	-	0.7993	**0.8358**

Tab. 1 shows the quantitative results and Fig. 3 displays the visual comparison between our method and others. The method of Li et al. and ASM post-processing method can output smoother results. This is because the method of Li et al. uses the ellipse constraint, and the ASM post-processing method reconstructs the shape coefficient. However, due to the complicated brightness transformation around the OD, the method of Li et al. visually has a large deviation from the ground truth in the recognition of the boundary. Compared with the proposed method, ASM post-processing method has two disadvantages. Firstly, the noise of segmentation map (e.g., the output of DRIU) will have a great negative impact on the post-processing result. Secondly, it is time-consuming compared with end-to-end segmentation methods due to the sampling operations. On the other hand, our method improves the performance by the proposed ASM module, especially the

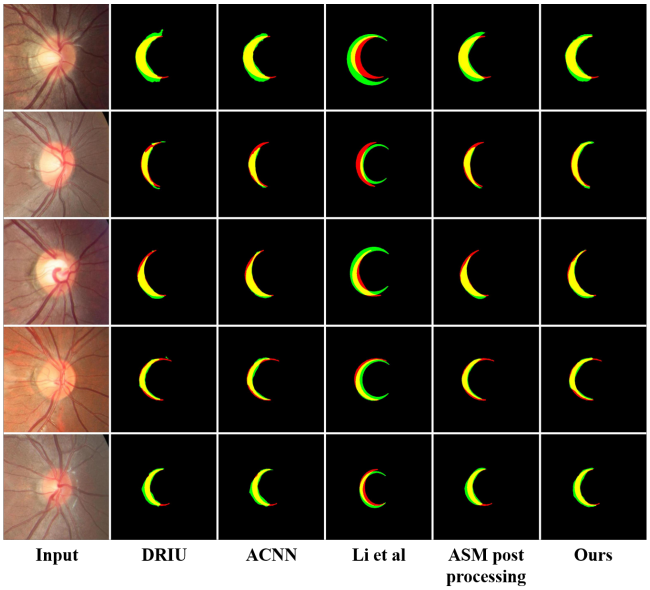


Fig. 3. Visual results compared with other methods. In the result, the red is the ground truth, the green is the segmentation result, and the yellow is the overlap area.

variant with the alignment component. With the help of alignment, accurate parameters of affine matrix can be obtained by the image-level constraints.

4. CONCLUSION

In this paper, we propose a method of adding ASM loss to the segmentation network to improve segmentation accuracy by introducing shape constraints. This loss is added by cascading the ASM loss module after the segmentation network. The method we proposed has achieved excellent results in both quantitative and qualitative analysis.

5. COMPLIANCE WITH ETHICAL STANDARDS

The research database comes from Beijing Tongren Hospital and written informed consent is obtained from all participants prior to their entry into the database.

6. ACKNOWLEDGMENTS

The research work is supported by the National Natural Science Foundation of China (NSFC) (Grant No. 82072007) and China Postdoctoral Science Foundation (No. 2020M680387).

7. REFERENCES

[1] Jiao Sun, Jialin Wang, and Yanling Wang, "Research progress on morphology and associations of peripapillary atrophy," *Chinese Journal of Ocular Fundus Diseases*, pp. 510–513, 2019.

[2] Botong Wu, Sijie Ren, Jing Li, Xinwei Sun, Shi-Ming Li, and Yizhou Wang, "Forecasting irreversible disease via progression learning," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, pp. 8117–8125.

[3] Cheng-Kai Lu, Tong Boon Tang, F Murray Alan, Augustinus Laude, and Baljean Dhillon, "Automatic parapapillary atrophy shape detection and quantification in colour fundus images," in *2010 Biomedical Circuits and Systems Conference (BioCAS)*. IEEE, 2010, pp. 86–89.

[4] Hanxiang Li, Huiqi Li, Jieliang Kang, Yunlong Feng, and Jie Xu, "Automatic detection of parapapillary atrophy and its association with children myopia," *Computer methods and programs in biomedicine*, vol. 183, pp. 105090, 2020.

[5] Yidong Chai, Hongyan Liu, and Jie Xu, "A new convolutional neural network model for peripapillary atrophy area segmentation from retinal fundus images," *Applied Soft Computing*, vol. 86, pp. 105890, 2020.

[6] Ozan Oktay, Enzo Ferrante, Konstantinos Kamnitsas, Mattias Heinrich, Wenjia Bai, Jose Caballero, Stuart A Cook, Antonio De Marvao, Timothy Dawes, Declan P O'Regan, et al., "Anatomically constrained neural networks (acnns): application to cardiac image enhancement and segmentation," *IEEE transactions on medical imaging*, vol. 37, no. 2, pp. 384–395, 2017.

[7] Yubo Fan, Dongqing Zhang, Jianing Wang, Jack H Noble, and Benoit M Dawant, "Combining model-and deep-learning-based methods for the accurate and robust segmentation of the intra-cochlear anatomy in clinical head ct images," in *Medical Imaging 2020: Image Processing*. International Society for Optics and Photonics, 2020, vol. 11313, p. 113131D.

[8] Sofie Tilborghs, Tom Dresselaers, Piet Claus, Jan Bogaert, and Frederik Maes, "Shape constrained cnn for cardiac mr segmentation with simultaneous prediction of shape and pose parameters," in *International Workshop on Statistical Atlases and Computational Models of the Heart*. Springer, 2020, pp. 127–136.

[9] Timothy F Cootes, Christopher J Taylor, David H Cooper, and Jim Graham, "Active shape models-their training and application," *Computer vision and image understanding*, vol. 61, no. 1, pp. 38–59, 1995.

[10] Kevis-Kokitsi Maninis, Jordi Pont-Tuset, Pablo Arbeláez, and Luc Van Gool, "Deep retinal image understanding," in *International conference on medical image computing and computer-assisted intervention*. Springer, 2016, pp. 140–148.

[11] Huiqi Li and Opas Chutatape, "Automatic location of optic disk in retinal images," in *Proceedings 2001 International Conference on Image Processing (Cat. No. 01CH37205)*. IEEE, 2001, vol. 2, pp. 837–840.