Computers in Biology and Medicine 166 (2023) 107464

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

Semi-supervised peripapillary atrophy segmentation with shape constraint

Mengxuan Li [a], Weihang Zhang [a], Ruixiao Yang [a], Jie Xu [b], He Zhao [a], Huiqi Li [a,*]

[a] Beijing Institute of Technology, Beijing 100081, China
[b] Beijing Institute of Ophthalmology, Beijing Tongren Hospital, Capital Medical University, Beijing 100005, China

ARTICLE INFO

Keywords:
Peripapillary atrophy segmentation
Semi-supervised
Shape constraint
Active shape model
Mean teacher model

ABSTRACT

Peripapillary atrophy (PPA) is a clinical abnormality related to many eye diseases, such as myopia and glaucoma. The shape and area of PPA are essential indicators of disease progression. PPA segmentation is a challenging task due to blurry edge and limited labeled data. In this paper, we propose a novel semi-supervised PPA segmentation method enhanced by prior knowledge. In order to learn shape information in the network, a novel shape constraint module is proposed to restrict the PPA appearance based on active shape model. To further leverage large amount of unlabeled data, a Siamese-like model updated by exponential moving average is introduced to provide pseudo labels. The pseudo labels are further refined by region connectivity correction. Extensive experiments on a clinical dataset demonstrate that our proposed PPA segmentation method provides good qualitative and quantitative performance.

1. Introduction

Peripapillary atrophy (PPA) is also called the optic disc crescent, which is a clinical finding associated with chorioretinal thinning and disruption of the retinal pigment epithelium (RPE) in the area surrounding optic disc (OD) [1]. The presence of PPA is a symptom related to two ocular diseases, myopia and glaucoma [2], which may cause visual impairment even blindness [3,4]. The appearance of PPA can provide an alarm for medical treatment at an early stage. Clinically, to observe PPA, non-invasive fundus images are commonly used. However, manually measuring PPA parameters such as PPA's size is time-consuming because it needs to outline PPA region precisely from retinal images. Therefore, proposing an automatic method for PPA segmentation is necessary, especially in population-based clinical studies.

Recently, there have been some investigations into PPA segmentation which can be divided into traditional methods and deep learning-based methods. Traditional methods perform object segmentation based on image features such as color, intensity, texture and shape. Lu et al. [5] proposed a method based on a hybrid image processing technique to segment OD and PPA which combines several traditional techniques, including scanning filter, thresholding, region growing, and a modified Chan-Vese model [6] with a shape constraint. Furthermore, they programmed a retinal imaging tool set for PPA and OD assessment (PANDORA) [7]. The process includes segmentation of OD region by edge detection and ellipse fitting methods first, then segmentation of PPA outside the OD by sequence operations in the same way as the method in Ref. [5]. However, the above two methods are highly sensitive to initialization, which fail to obtain accurate boundaries for weak PPA. Li et al. [8] utilized evenly oriented radial lines to detect PPA boundary and used least square ellipse fitting to add shape constraints. However, the complicated illuminance around OD or overlapping blood vessels might affect the performance of the method.

With the rapid development of deep learning, the research on PPA segmentation has gradually shifted from traditional methods to deep learning-based methods. Chai et al. [9] proposed a multi-task fully convolutional network model to extract OD and 'PPA plus OD' (PPAOD) jointly, in which edge continuity of the object area is considered. The PPA segmentation result is obtained by subtracting OD from PPAOD. The edge continuity of this method considers the OD and PPAOD areas but does not consider the PPA area separately, so the edge continuity of the obtained PPA segmentation result is not satisfactory. In method proposed in Ref. [10], the boundary of PPA is directly considered, in which a boundary guidance convolution network is proposed to segment PPA area. These two deep learning methods mainly consider the boundary information but the shape features of PPA are not included yet.

Shape is an essential feature in image segmentation. Incorporating prior shape knowledge into traditional image segmentation algorithms has been proven to be helpful for obtaining more accurate and reasonable results. Deep learning methods combined with prior shape knowledge attract more and more researchers' attention. Shape constraints are often used as post-processing operations on deep learning segmentation results. In Refs. [11,12], conditional random field (CRF) [13] is used to refine the rough segmentation results. In method of

* Corresponding author.
E-mail addresses: zhaohe@bit.edu.cn (H. Zhao), huiqili@bit.edu.cn (H. Li).

https://doi.org/10.1016/j.compbiomed.2023.107464
Received 23 March 2023; Received in revised form 14 August 2023; Accepted 4 September 2023
Available online 7 September 2023

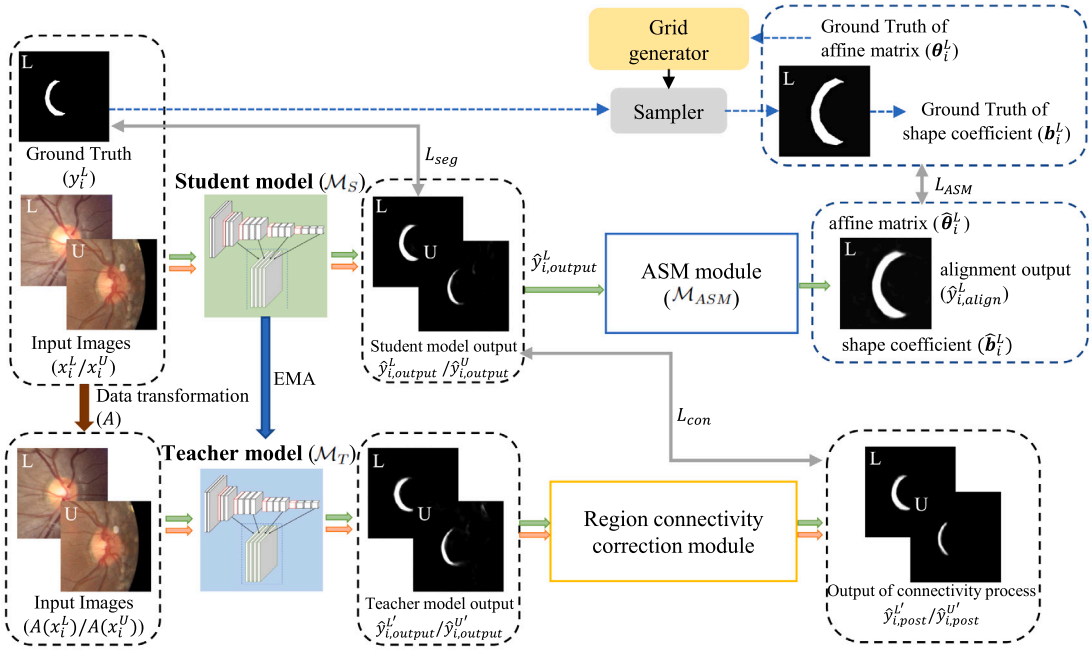

Fig. 1. Diagram of the proposed semi-supervised PPA segmentation approach. The upper left corner of the image 'L' represents labeled data and 'U' represents unlabeled data. Our proposed model contains a Siamese-like structure, including a student model and a teacher model. An ASM module is cascaded after the student model which is used to learn the shape parameters and further calculate the shape constraint loss. The region connectivity correction module is cascaded after the teacher model which is used to make the pseudo labels more accurate.

Ref. [14], the weighted fuzzy active shape model is used as post-processing to complete the segmentation of the bounding box. In Ref. [15], the phase-based active contour (PBAC) model [16] is further used to improve the rough segmentation results predicted by the convolutional network. In Ref. [17], denoising autoencoders can improve the anatomical plausibility of unfeasible segmentation masks as a post-processing step. Although better shape can be obtained by using the post-processing method, the post-processing operation is a multi-step operation, which is time-consuming and complex to obtain the final segmentation result. There are also some works incorporating shape constraints into convolutional neural network. The milestone work is ACNN [18], which learns the non-linear compact representation of anatomy by a stacked convolutional autoencoder. The auto encoder–decoder architecture or its variation for shape constraint is then utilized in various areas such as pediatric MR [19] and cardiac MRI [20]. However, the features encoded by the autoencoder have no corresponding physical interpretation.

In order to increase the interpretability of shape constraints, in our preliminary work published in a conference paper [21], a PPA segmentation method with shape constraint is proposed by introducing a shape-based loss to the convolution neural network, in which the shape information learned by the network has actual physical meaning. However, this method is a fully supervised method, and the number of pixel-level annotations in the actual segmentation task is limited. Lack of annotated samples limits the network's performance. Pixel-wise annotation is time-consuming especially when the boundary of PPA in fundus images is not clear such as in young patients' data. In this case, semi-supervised PPA segmentation is worth investigating since unlabeled data are more accessible. Recently, there have been some investigations on semi-supervised medical image segmentation. Due to the robustness and efficiency, mean teacher [22] has been employed into the medical image segmentation tasks successfully [23–26].

Considering that PPAs are often crescent-shaped, especially for teenagers. In this paper, inspired by the success of mean teacher model, we propose a semi-supervised PPA segmentation with shape constraints. A novel shape constraint loss is proposed inspired by the active shape model (ASM) [27]. Instead of using the ASM in pre-/post-processing, our method can be trained in an end-to-end manner using the proposed loss function with a new ASM module. Moreover, the unlabeled data are leveraged by the Siamese-like structure with region connectivity correction. The main contributions of this article can be summarized as follows:

1. A novel semi-supervised method is first proposed for PPA segmentation. Due to the limited amount of labeled data, the semi-supervised method can make full use of unlabeled images to handle the problem of insufficient labeled data and further improve performance of the network.
2. Shape constraint is added to the network. The specific implementation is to cascade the proposed ASM module in the network to learn the shape parameters and further calculate the shape constraint loss.
3. In order to make the pseudo labels more accurate, a region connectivity correction module is applied to ensure that the prior information of PPA area is a consistent region in the semi-supervised segmentation.
4. Comprehensive experiments are carried out to evaluate our approach. The experimental results show that it achieves superior performance compared with other methods.

2. Methods

We propose a novel semi-supervised PPA segmentation method incorporating shape constraint and prior information. Given the network $\mathcal{M}_\phi(x)$ and the training dataset $D = L \cup U$, where $L = \left\{x_i^L, y_{i,GT}^L\right\}$ is the labeled data pair and $U = \left\{x_i^U\right\}$ is the unlabeled data without ground truth. The semi-supervised segmentation method tries to learn the network ϕ by optimizing the following loss function:

$$\mathcal{L} = \sum_i \mathcal{L}_s(\mathcal{M}_\phi(x_i^L), y_{i,GT}^L) + \omega \mathcal{L}_u(\phi, L, U) \quad (1)$$

where $\mathcal{L}_s$ and $\mathcal{L}_u$ indicate supervised loss and unsupervised loss, respectively. ω is a regularization weight.

Fig. 1 illustrates the structure of the proposed PPA segmentation method. Our model $\mathcal{M}$ contains a Siamese-like structure, including a student model $\mathcal{M}_S$ and a teacher model $\mathcal{M}_T$. Since most PPAs are

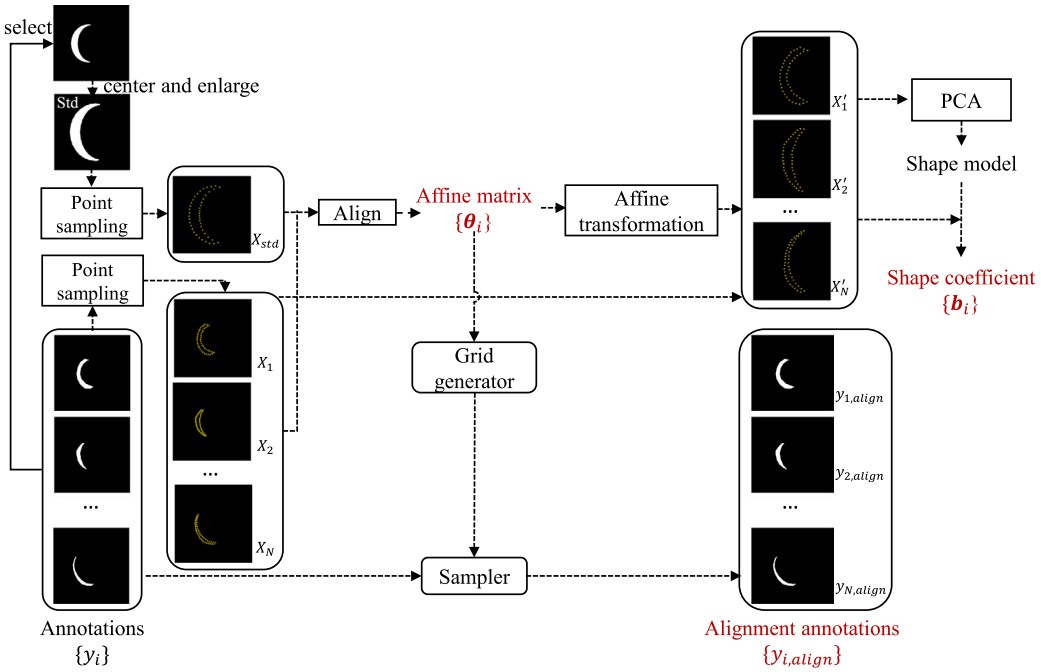

Fig. 2. Operations of generating ground truth about shape information. The upper left corner of the image 'Std' represents the standard image. The red mark is the ground truth used in the subsequent calculation of the shape constraint loss.

crescent shape, in order to learn the shape information in the process of network training, an ASM module is proposed to calculate the shape parameters which are used to restrict the PPA appearance by the proposed shape constraint loss. For labeled data pair L, the ground truth $y^L_{i,GT}$ contains four parts y^L_i, θ^L_i, $y^L_{i,align}$ and $\boldsymbol{b}^L_i$, where y^L_i is the original image annotation, $\boldsymbol{\theta}^L_i$ is the affine matrix, $y^L_{i,align}$ is the output after affine transformation, and $\boldsymbol{b}^L_i$ is the shape coefficient. For labeled and unlabeled images (x^L_i and x^U_i), the student model outputs the predictions ($\hat{y}^L_{i,output}$ and $\hat{y}^U_{i,output}$). The ASM module only inputs the prediction result $\hat{y}^L_{i,output}$ of labeled data, and outputs shape information ($\hat{\boldsymbol{\theta}}^L_i, \hat{y}^L_{i,align}, \hat{\boldsymbol{b}}^L_i$). The pseudo label $\hat{y}^{L'/U'}_{i,output}$ for $x^{L/U}_i$ is obtained by the teacher model. In order to make the pseudo label more accurate, the output of the teacher model is performed by a region connectivity correction module to introduce prior information that PPA is a single connected region, and the processed pseudo label is denoted by $\hat{y}^{L'/U'}_{i,post}$. The supervised loss $\mathcal{L}_s$ includes shape constraint loss $\mathcal{L}_{ASM}$ and segmentation loss $\mathcal{L}_{seg}$. The shape constraint loss is calculated based on the outputs of ASM module ($\hat{\boldsymbol{\theta}}^L_i, \hat{y}^L_{i,align}, \hat{\boldsymbol{b}}^L_i$) and the ground truth of shape parameters ($\boldsymbol{\theta}^L_i, y^L_{i,align}, \boldsymbol{b}^L_i$). Segmentation loss is calculated based on the student outputs $\hat{y}^L_{i,output}$ and the ground truth annotations y^L_i. The unsupervised loss $\mathcal{L}_u = \mathcal{L}_{con}$ is calculated between $\hat{y}^{L/U}_{i,output}$ and $\hat{y}^{L'/U'}_{i,post}$.

2.1. ASM module

The ASM module aims to propose a shape constraint loss to leverage the prior knowledge of PPA, which calculates the affine matrix and the shape coefficient of ASM [27]. In our method, the shape of PPA is represented by a set of points that can be used as a form of representing object features. For training data, we process the labeled annotations to obtain the ground truth of the shape information ($\boldsymbol{\theta}_i, y_{i,align}, \boldsymbol{b}_i$). The operation steps are shown in Fig. 2.

Firstly, all data are sampled by n pairs of landmarks to represent the shape of them. The j-th point of the i-th image is represented as $\{lx_{ij}, ly_{ij}\}$, and the shape of image i can be represented by a matrix composed of n pairs of landmark points, i.e., $\boldsymbol{X}_i = [lx_{i0}, ly_{i0}, \ldots, lx_{ik}, ly_{ik}, \ldots, lx_{i(n-1)}, ly_{i(n-1)}]^T$.

Secondly, select an image with the smallest angle between the horizontal axis and the line connecting image center and centroid of PPA area in the ground truth of training set as the standard reference image, and a point distribution model is derived to get its shape representation $\boldsymbol{X}_{std}$.

Thirdly, the alignment operation is performed to compute the transformation parameters. The alignment transformation can be represented by a scaling magnitude parameter s_i, a rotation parameter β_i, and a translation parameter matrix $\boldsymbol{t}_i$. After $\boldsymbol{X}_i$ is aligned, it can be expressed as $\boldsymbol{X}'_i$ and its specific matrix is expressed as $T(s_i, \beta_i)[\boldsymbol{X}_i] + \boldsymbol{t}_i$. The so-called alignment is to transform the shape to make it as close as possible to the reference shape $\boldsymbol{X}_{std}$, as shown in formula (2).

$$\min_{s_i,\beta_i} \left\{ (\boldsymbol{X}_{std} - T(s_i,\beta_i)[\boldsymbol{X}_i] - \boldsymbol{t}_i)^T (\boldsymbol{X}_{std} - T(s_i,\beta_i)[\boldsymbol{X}_i] - \boldsymbol{t}_i) \right\} \tag{2}$$

where

$$T(s_i,\beta_i)\begin{bmatrix} lx_{ik} \\ ly_{ik} \end{bmatrix} = \begin{bmatrix} s_i\cos\beta_i & s_i\sin\beta_i \\ -s_i\sin\beta_i & s_i\cos\beta_i \end{bmatrix}\begin{bmatrix} lx_{ik} \\ ly_{ik} \end{bmatrix} \tag{3}$$

$$\boldsymbol{t}_i = [t_{xi}, t_{yi}]^T \tag{4}$$

$t_{xi} = \frac{1}{n}\sum_{k=1}^{n} lx_{ik} - W/2$ and $t_{yi} = \frac{1}{n}\sum_{k=1}^{n} ly_{ik} - H/2$. H and W are the height and width of the image, respectively.

The affine transformation parameter $\boldsymbol{\theta}_i$ used in the ASM module is defined as:

$$\boldsymbol{\theta}_i = \begin{bmatrix} s_i\cos\beta_i & s_i\sin\beta_i & t_{xi}/W \\ -s_i\sin\beta_i & s_i\cos\beta_i & t_{yi}/H \end{bmatrix} \tag{5}$$

Fourthly, the affine transformation parameters are employed for all shape representations $\{\boldsymbol{X}_i\}$ and the original annotations to obtain the aligned shape representations $\{\boldsymbol{X}'_i\}$ and the aligned annotations $\{y_{i,align}\}$.

Next, principal component analysis (PCA) is performed on the aligned shape representations to obtain a shape model. According to PCA, k maximum eigenvectors $\boldsymbol{P} = [p_1, p_2, \ldots, p_k]$ and eigenvalues $\boldsymbol{\Lambda} = diag(\lambda_1, \lambda_2, \ldots, \lambda_k)$ can be obtained.

Finally, the shape model is used to calculate the shape coefficient for all data. The final shape coefficient $\boldsymbol{b}_i$ of image i is obtained by:

$$\boldsymbol{b}_i = [b_{i0}, \ldots, b_{ik}]^T = \boldsymbol{P}^T(\boldsymbol{X}'_i - \bar{\boldsymbol{X}}') \tag{6}$$

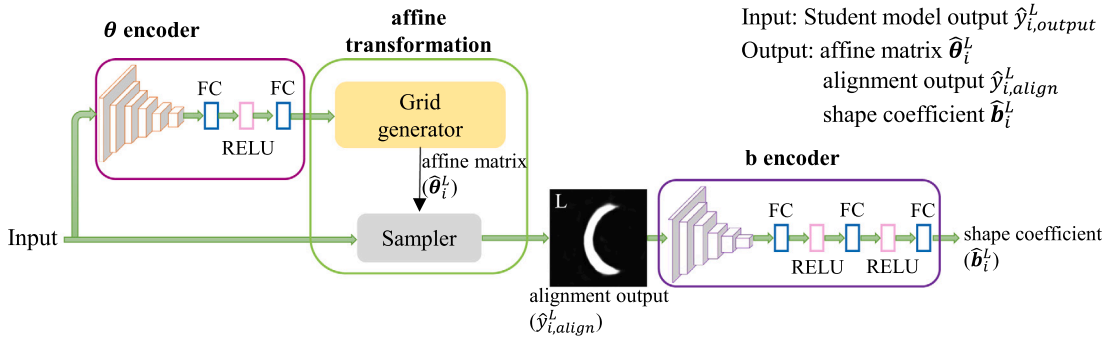

Fig. 3. The structure of the ASM module.

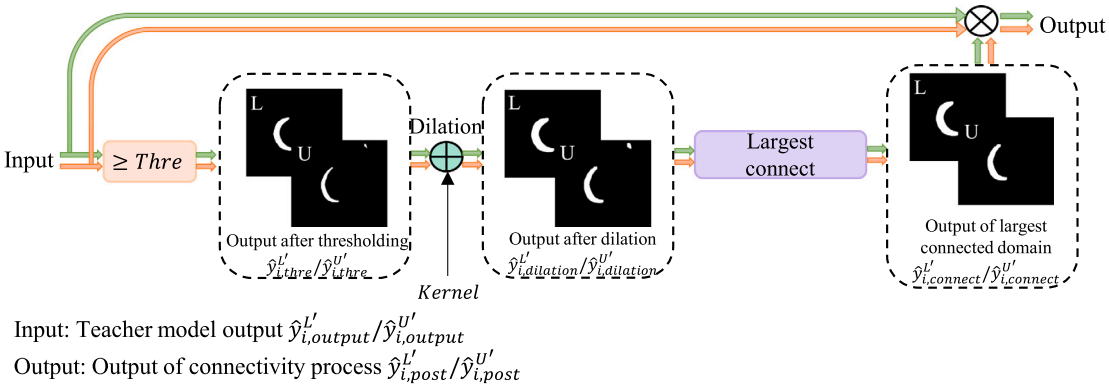

Fig. 4. The structure of the region connectivity correction module.

where $\bar{\boldsymbol{X}}'$ is the mean shape of $\{\boldsymbol{X}'_i\}_{i=1}^N$. The coefficient b_{ij} is further truncated within the range of $\left[-3\sqrt{\lambda_j}, 3\sqrt{\lambda_j}\right]$ to alleviate the outliers.

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

$$\boldsymbol{X}'_i = \bar{\boldsymbol{X}}' + \boldsymbol{P}\boldsymbol{b}_i \tag{7}$$

The ASM module used for calculating alignment and shape parameters consists of three parts, two encoders and an affine transformation operator, as shown in Fig. 3. The first encoder is named as 'θ predictor' which learns the parameters of the affine matrix $\boldsymbol{\theta}_i$. It consists of convolutional and fully connected layers. The affine transformation operator aligns the segmentation results according to the predicted affine matrix, and outputs the aligned segmentation results $y_{i,align}$. The affine transformation operator includes two parts: grid generator and sampler [28]. The grid generator is used to process the predicted affine matrix to create a sampling grid, which is a set of points where the input feature is sampled to produce the transformed output. The feature and the sampling grid are taken as inputs to the sampler, producing the output feature sampled from the input as the grid points. The second encoder called 'b predictor' encrypts the aligned result into shape coefficient $\boldsymbol{b}_i$. It also consists of convolutional and fully connected layers. Two predicted parameters $\boldsymbol{\theta}_i$ and $\boldsymbol{b}_i$ together with the alignment output $y_{i,align}$ are used for calculating the shape constraint loss.

2.2. Semi-supervised pipeline

Inspired by the mean teacher model [23], our model $\mathcal{M}$ is a Siamese-like structure including a student model and a teacher model which share the same model structure. In this paper, deep retinal image understanding (DRIU) [29] is introduced as the backbone of the student model and teacher model. The DRIU is a deep neural network specially designed for retinal image segmentation which can simultaneously segment the OD and vessels in the retina. In order to enable the network to learn shape information, an ASM module is added to optimize the model training, which is used to calculate the shape parameters and the shape constraint loss. During the training process, only the parameters of student model are obtained through learning, and the parameters of the teacher model are obtained by the exponential moving average (EMA) of the student model, so the ASM module is only cascaded after the student model. Both labeled and unlabeled data are input into the student model and the teacher model, but only the prediction results of the labeled part in student model are input into the ASM module. This is because the ASM module is used to calculate the shape constraint loss, and only labeled data have shape labels. The output of the teacher model is treated as the pseudo label, which contains many errors. In order to obtain more accurate pseudo labels, the output of the teacher model is processed by the region connectivity correction module, which uses the prior information that the PPA is a single connected region. In this part, we will first introduce the mean teacher model, then introduce the region connectivity correction module.

2.2.1. Mean teacher model

The mean teacher model consists of two components: student model and teacher model, where the architecture of student model is replicated into the teacher model. For the same input data, different data transformations are used to process the data and then they are input to the student model and teacher model respectively to obtain the predicted outputs $\hat{y}^{L/U}_{output}$ and $\hat{y}^{L'/U'}_{output}$. In the loss function, the supervised loss only includes the segmentation loss $\mathcal{L}_{seg}$, and the unsupervised loss is the consistency loss between $\hat{y}^{L/U}_{output}$ and $\hat{y}^{L'/U'}_{output}$. According to the loss function, the student model is trained by the gradient descent optimizer while the teacher model is optimized as the EMA of the student parameters. Specifically, assuming the student model is parameterized by ϕ and the teacher model is parameterized by ϕ', the weights updating formula of the teacher model at training step t is expressed in Eq. (8).

$$\phi'_t = \alpha\phi'_{t-1} + (1-\alpha)\phi_t \tag{8}$$

where α is the coefficient of EMA decay to control the updating rate, ϕ_t is the parameter of the student model at training step t, and ϕ'_{t-1} is the parameter of the teacher model at training step $t-1$.

2.2.2. Region connectivity correction module

Since there is no restriction on unlabeled data, the PPA segmentation result from model $\mathcal{M}_T$ might gain some undesirable discrete areas. This is mainly caused by the disturbances in the image, such as the white fiber layer and abnormal bright region. Moreover, the appearance of PPA is crescent and a consistent region. Based on this prior knowledge, we propose a region connectivity correction module to refine the pseudo label generated by $\mathcal{M}_T$, as shown in Fig. 4. It is achieved by generating a mask to filter out the area with low probability.

The output of $\mathcal{M}_T$ is first thresholded by a constant value to remove the predictions with low probability. After thresholding, the output may occur with minor breaks. If taking the largest connected region as the output, it will be far from ground truth. In order to avoid the above situation, the dilation operation is used to process the output after thresholding and then take the largest connected region. Due to the expansion of the region after dilation will introduce additional distractions, the current output cannot be directly used as the pseudo label. Therefore, the final pseudo label is obtained by multiplying the current output with the output of $\mathcal{M}_T$.

2.3. Loss function

Our loss function consists of three parts: the segmentation loss $\mathcal{L}_{seg}$ and shape constraint loss $\mathcal{L}_{ASM}$ on labeled data as well as the consistency loss $\mathcal{L}_{con}$ on both labeled and unlabeled data.

Segmentation loss: The segmentation loss calculates the difference between the PPA segmentation map and the ground truth, which is formulated by:

$$\mathcal{L}_{seg} = \frac{1}{M}\sum_{i=1}^{M}(1 - \frac{2y_i^L \hat{y}_{i,output}^L}{y_i^L + \hat{y}_{i,output}^L}) \tag{9}$$

where y_i^L and $\hat{y}_{i,output}^L$ represent the ground truth and the segmentation result for image i. M is the number of labeled images.

Shape constraint loss: The proposed ASM loss can be formulated as:

$$\mathcal{L}_{ASM} = \mathcal{L}_\theta + \gamma\mathcal{L}_b + \mathcal{L}_{align} \tag{10}$$

where $\mathcal{L}_\theta$ and $\mathcal{L}_b$ are loss terms updating the encoders, and γ is a hyperparameter. In addition, we also include an align loss $\mathcal{L}_{align}$ to constrain the alignment, which can further improve the accuracy of parameters in the affine matrix.

The calculation formulas of L_θ, L_{align} and L_b are defined as:

$$\mathcal{L}_\theta = \frac{1}{M}\sum_{i=1}^{M}(\boldsymbol{\theta}_i^L - \hat{\boldsymbol{\theta}}_i^L)^2 \tag{11}$$

$$\mathcal{L}_{align} = \frac{1}{M}\sum_{i=1}^{M}(1 - \frac{2y_{i,align}^L \hat{y}_{i,align}^L}{y_{i,align}^L + \hat{y}_{i,align}^L}) \tag{12}$$

$$\mathcal{L}_b = \frac{1}{M}\sum_{i=1}^{M}(\boldsymbol{b}_i^L - \hat{\boldsymbol{b}}_i^L)^2 \tag{13}$$

where $\boldsymbol{\theta}_i^L$, $y_{i,align}^L$ and $\boldsymbol{b}_i^L$ represent the ground truth of affine matrix, alignment result and shape coefficient for image i respectively, and $\hat{(\cdot)}$ represent the predicted results.

Consistency loss: The consistency loss is calculated by the pseudo label generated by $\mathcal{M}_T$ and the output of the student model, as shown in Eq. (14).

$$\mathcal{L}_{con} = \frac{1}{M+K}\sum_{i=1}^{M+K}[-\hat{y}_{i,post}^{L'/U'} log(\hat{y}_{i,output}^{L/U}) - (1-\hat{y}_{i,post}^{L'/U'}) log(1-\hat{y}_{i,output}^{L/U})] \tag{14}$$

where $\hat{y}_{output}^{L/U}$ and $\hat{y}_{post}^{L'/U'}$ respectively represent the segmentation result of student model and the post-processing output of the teacher model. K is the number of unlabeled images.

Finally, the overall loss function to optimize $\mathcal{M}_S$ can be formulated as:

$$\mathcal{L} = \mathcal{L}_{seg} + \mathcal{L}_{ASM} + \omega\mathcal{L}_{con} \tag{15}$$

where ω is a hyperparameter for the proportion of consistency loss $\mathcal{L}_{con}$.

The pseudo-code of our proposed method is shown in Algorithm 1.

Algorithm 1 Our proposed semi-supervised segmentation algorithm

Input:

Labeled training set $D^L = \{(x_i^L, y_i^L, \boldsymbol{\theta}_i^L, y_{i,align}^L, \boldsymbol{b}_i^L)\}_{i=1}^M$

Unlabeled training set $D^U = \{x_i^U\}_{i=1}^K$

Variable definitions:

$\mathcal{M}_S(\cdot, \phi)$ →student model

$\mathcal{M}_T(\cdot, \phi')$ →teacher model

$\mathcal{M}_{ASM}(\cdot, \phi_{ASM})$ →ASM module

$\mathcal{M}_{S,ASM}(\cdot, \phi, \phi_{ASM})$ →cascade ASM module behind student model

$A(\cdot)$ →data transformation.

$\omega(\cdot)$ → consistency weight ramp-up function

$Thre$ → manually selected threshold

Output:

Student model $\mathcal{M}_S$

for k in [1,num_epochs] **do**

for $(x^L, y^L) \in D^L$ *and* $x^U \in D^U$ **do**

$\hat{y}_{output}^L, \hat{\boldsymbol{\theta}}^L, \hat{y}_{align}^L, \hat{\boldsymbol{b}}^L \leftarrow \mathcal{M}_{S,ASM}(x^L, \phi, \phi_{ASM})$

$\hat{y}_{output}^U \leftarrow \mathcal{M}_S(x^U, \phi)$

$\hat{y}_{output}^{L'} \leftarrow \mathcal{M}_T(A(x^L), \phi')$

$\hat{y}_{output}^{U'} \leftarrow \mathcal{M}_T(A(x^U), \phi')$

$\hat{y}_{thre}^{L'/U'} = \hat{y}_{output}^{L'/U'} > Thre$

$\hat{y}_{dilation}^{L'/U'} = Dilation(\hat{y}_{thre}^{L'/U'}, kernel)$

$\hat{y}_{connect}^{L'/U'} = Largest\ connected(\hat{y}_{dilation}^{L'/U'})$

$\hat{y}_{post}^{L'/U'} = \hat{y}_{connect}^{L'/U'} \cdot \hat{y}_{output}^{L'/U'}$

$\mathcal{L} = \mathcal{L}_{seg} + \mathcal{L}_{ASM} + \omega(k)\mathcal{L}_{con}$

Update ϕ *and* ϕ_{ASM} with regard to $\mathcal{L}$

$\phi' = \alpha\phi' + (1-\alpha)\phi$

end for

end for

3. Experimental results

3.1. Dataset description

Two datasets are used to validate the proposed method which are provided by Beijing Tongren Hospital. All images are collected from primary school students from grade one to six. Dataset 1 has 200 labeled images. Dataset 2 contains 407 unlabeled images. The Institutional Review Board (IRB) number of this study is TRECKY2017-061.

We randomly select 50 images from the labeled dataset (Dataset 1) as the testing set, and the rest as the training set. The dataset used for fully supervised manner in the follow-up experiments only uses Dataset 1 and keeps the testing set unchanged.

The data used in the experiments focus on the region of interest (ROI) in retinal images. The ROI is extracted using the method in Ref. [30] to locate the center of OD followed by a resizing operation to 512×512, and normalized to the right eye. Since the original image is centered on the macula, it can be normalized to the right eye according to the central position of OD.

Table 1
Results of our semi-supervised method under different numbers of labeled data. ↑ means that the higher the indicator, the better, and ↓ means that the lower the indicator, the better.

	L/U	Dice (%) ↑	F1 (%) ↑	Pre (%) ↑	Rec (%) ↑	Acc (%) ↑	IoU (%) ↑	ASSD↓	mAP (%) ↑
Supervised	150/0	80.05	81.42	81.04	81.81	99.06	68.67	5.17	79.29
Ours	25/125	65.45	66.05	68.00	64.26	98.33	49.32	10.67	66.23
	50/100	75.50	76.52	75.23	77.86	98.79	61.98	6.52	79.52
	75/75	79.02	79.76	77.40	82.29	98.95	66.34	5.59	81.80
	100/50	79.79	80.61	78.97	82.32	99.00	67.52	5.19	82.26
	125/25	**80.87**	**81.86**	**79.31**	**84.57**	**99.05**	**69.29**	**4.73**	**83.45**

[1] "L/U" denotes the number of labeled images versus unlabeled images.

Table 2
Results of our semi-supervised method under different numbers of unlabeled data.

L/U	Dice(%) ↑	F1(%) ↑	Pre(%) ↑	Rec(%) ↑	Acc(%) ↑	IoU(%) ↑	ASSD↓	mAP(%) ↑
150/50	80.81	81.98	80.86	83.13	99.08	69.46	5.33	83.73
150/100	81.00	82.07	80.04	**84.22**	99.07	69.60	4.89	83.98
150/150	81.20	82.24	80.48	84.09	99.08	69.84	4.83	83.98
150/300	81.43	82.39	80.78	84.07	99.09	70.06	4.83	83.96
150/407	**81.53**	**82.57**	**81.50**	83.67	**99.11**	**70.32**	**4.72**	**84.00**

[1] "L/U" denotes the number of labeled images versus unlabeled images.

Table 3
Results of the baseline method (MT) and our method with different student and teacher models.

		Dice (%) ↑	F1 (%) ↑	Pre (%) ↑	Rec (%) ↑	Acc (%) ↑	IoU (%) ↑	ASSD↓	mAP (%) ↑
DRIU [29]	Baseline	79.68	80.97	80.28	81.70	99.03	68.02	5.47	80.88
	Ours	**81.53**	**82.57**	**81.50**	**83.67**	**99.11**	**70.32**	**4.72**	**84.00**
U-Net [33]	Baseline	79.45	80.74	**80.49**	80.99	99.03	67.70	5.41	81.00
	Ours	**80.31**	**81.39**	79.87	**82.96**	**99.04**	**68.62**	**4.84**	**82.31**
PSPNet [34]	Baseline	80.54	81.75	**81.72**	81.78	99.08	69.13	5.21	81.95
	Ours	**80.95**	**82.23**	81.26	**83.21**	**99.09**	**69.82**	**4.87**	**84.42**

3.2. Implementation details

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 next 100 epochs is linearly decreased to 0. All experiments are optimized using Adam optimizer with the batch size of 12. Moreover, when processing the labeled data to generate the shape ground truth, the number of sampling points n is 44. The information retention rate in PCA is selected as 99%, and the length of shape coefficient vector is 12. The coefficient of EMA $\alpha = 0.999$ in Eq. (8). The hyperparameter γ in Eq. (10) is selected as 10. The consistency weight ω in Eq. (15) is dynamically changed over time as a function of:

$$\omega = \begin{cases} 0 & t \le 10 \\ \exp(-5 \times (1 - t/t_{max})^2) & 10 < t \le t_{max} \\ 1 & t > t_{max} \end{cases} \tag{16}$$

where t and t_{max} denote the current and the ramp-up maximum epoch, and t_{max} is set to 100 in this experiment. Besides, the threshold is set to 0.3 and the kernel size is selected as 5×5 in Algorithm 1.

3.3. Evaluation metrics

Dice, F1-score (F1), Precision (Pre), Recall (Rec), Accuracy (Acc), IoU, Average symmetric surface distance (ASSD) [31] and Mean Average Precision (mAP) [32] are adopted to evaluate PPA segmentation performance. Both Dice and IoU are metrics to measure the similarity between the segmentation results and the ground truth. mAP is used by the PASCAL Visual Object Classes (VOC) Challenge to evaluate the segmentation results. ASSD is the average of all the distances from points on the boundary of the segmentation region to the boundary of ground truth. Our results are the average score of three runs. The threshold used to obtain the final segmentation result is determined based on the best value of F1-score.

3.4. Experiments with different ratios of labeled and unlabeled data

In order to verify the effectiveness of the proposed semi-supervised pipeline, we compare it with the original fully supervised method. The dataset used in this part of experiment is Dataset 1, and the experimental results are shown in Table 1. The fully supervised model is a DRIU network trained with all labeled data. When training the semi-supervised model, the total amount of training data is kept constant and the experiments are carried out under the condition of different ratios of labeled data and unlabeled data. It can be observed that the performance is increased by using more labeled data for training. When only 1/3 of the labeled data are used for training, the mAP obtained exceeds that of the fully supervised method. The ASSD metric also outperforms fully supervised method when 2/3 of the labeled data are used in the training process. When trained with 5/6 labeled data, almost all indicators surpass the fully supervised method. The results demonstrate that our semi-supervised method is effective, which can achieve better results with less labeled data.

We also conduct experiments to explore the contribution of unlabeled data. In this experiment, the amount of labeled data is kept constant, and the amount of unlabeled data is changed. The labeled data comes from Dataset 1, and the unlabeled data comes from Dataset 2. The experimental results are shown in Table 2. It can be observed that the more unlabeled data, the better performance of the network. In order to make full use of all available data, in the subsequent semi-supervised experiments, Dataset 1 is fully used as labeled data and Dataset 2 is fully used as unlabeled data.

3.5. Results on different model structures

To demonstrate the generality of the semi-supervised pipeline, different student and teacher structures are compared including DRIU [29], U-Net [33], PSPNet [34] respectively. The experimental results are shown in Table 3. 'Baseline' in the table indicates that the semi-supervised scheme is mean teacher model, and 'Ours' indicates that the

Table 4
Quantitative results of ablation study.

Mean teacher	Shape constraint	Region connectivity correction	Dice (%) ↑	F1 (%) ↑	Pre(%) ↑	Rec (%) ↑	Acc (%) ↑	IoU (%) ↑	ASSD↓	mAP (%) ↑
✓			79.68	80.97	80.28	81.70	99.03	68.02	5.47	80.88
✓	✓		81.27	82.41	**82.55**	82.28	**99.11**	70.08	5.12	83.30
✓		✓	79.81	80.99	80.20	81.81	99.03	68.06	5.22	81.87
✓	✓	✓	**81.53**	**82.57**	81.50	**83.67**	**99.11**	**70.32**	**4.72**	**84.00**

Table 5
Complexity analysis of ablation study.

Mean teacher	Shape constraint	Region connectivity correction	FLOPs (G)	Params (M)	Train-time (s/epoch)
✓			85.16	14.94	24.23
✓	✓		86.91	17.58	25.39
✓		✓	85.16	14.94	36.83
✓	✓	✓	86.91	17.58	37.19

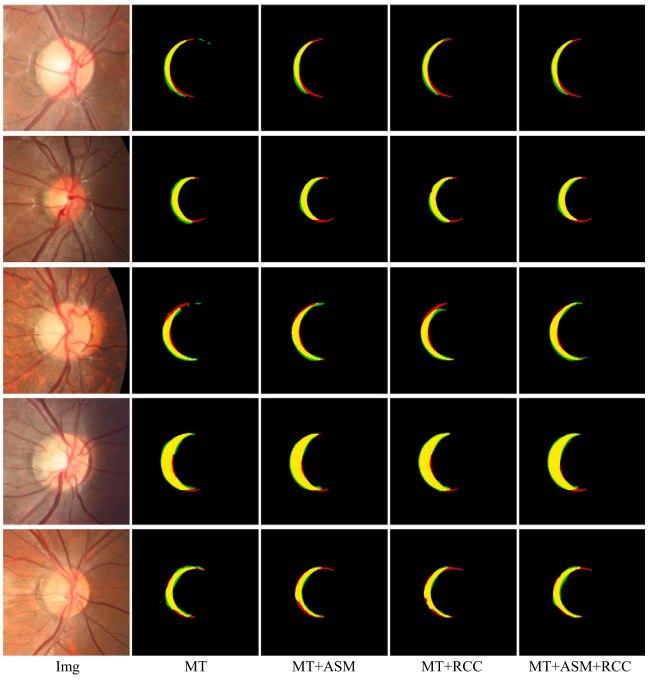

Fig. 5. Visual results of ablation experiments on shape constraint. MT is the abbreviation of mean teacher. ASM refers to the ASM module here, and RCC is the abbreviation of region connectivity correction module. The red is the ground truth, the green is the segmentation result, and the yellow is the overlap area.

ASM module and the region connectivity correction module are added on the basis of mean teacher model. It can be noted from the table that no matter what structures are adopted, our proposed method improves all metrics except Pre. It can be seen from the experimental results that the proposed semi-supervised scheme is a generalized method and can be extended to other network structures.

3.6. Ablation study

In this part, we verify the effectiveness of the proposed ASM module and region connectivity correction module. Table 4 displays the results of ablation study. When only adding the ASM module, all metrics are better than the original mean teacher model, especially mAP increases more than 2%. The proposed shape constraint loss makes the student model incorporate with the shape information in the training process, resulting in a precise shape appearance. As shown in Fig. 5, the results generated by the model with shape constraint are with more smooth boundary and crescent-like appearance. When only adding the region connectivity correction module, although most of the indicators are the same as the original mean teacher model, it has improved in both mAP and ASSD. The reduction of ASSD means that the segmentation boundary is closer to the ground truth boundary, which can reduce the influence of noise region. It can also be seen from Fig. 5 that the connectivity of the result obtained after adding the region connectivity module is better. When both the ASM module and the region connectivity correction module are added to the baseline with mean teacher model, the mAP and ASSD can be significantly improved.

We carry out the complexity analysis of ablation study. The results are shown in Table 5. After adding the ASM module, although the number of parameters (Params) and floating point operations (FLOPs) increase, the training time remains basically the same. After adding the region connectivity correction module, there is no change in the number of Params and FLOPs, but the training time will increase due to the calculation of the maximum connected region. Since the ASM module and region connectivity correction module are used only in the training phase to facilitate the learning of DRIU, and the final segmentation results can be obtained directly by using the DRIU network during testing, the increase in time and space costs only affects the training phase.

3.7. Comparison with existing methods

3.7.1. Comparison with other semi-supervised methods

We compare our method with the latest semi-supervised segmentation methods, including mean teacher model [23], RLSSS [35], DTC [36], DTML [37], MC-Net [38], RMT_VAT [39] and SS-Net [40]. The mean teacher model has been introduced in Section 2.2.1. The specific operation of RLSSS is that the student model learns from the pseudo labels generated by the teacher model, and then the parameters of the teacher model are updated according to the labeled data performance of the student model [35]. Both DTC and DTML adopt a dual-task framework by predicting a pixel-wise segmentation map and a geometry-aware level set representation of the target simultaneously and calculating the consistency association between the two predicted results. MC-Net includes multiple decoder outputs, and makes consistency constraints between the probability output of one decoder and the soft pseudo labels of other decoders. RMT_VAT is composed of mean teacher model and a regularization-driven strategy with virtual adversarial training and entropy minimization. SS-Net explores pixel-level smoothness by applying adversarial noises and inter-class separation by shrinking each class distribution. These semi-supervised methods are all tested on our datasets. It should be noted that the ASM module and the region connectivity correction module are

Table 6
Segmentation results of different semi-supervised methods.

	Dice (%) ↑	F1 (%) ↑	Pre (%) ↑	Rec (%) ↑	Acc (%) ↑	IoU (%) ↑	ASSD↓	mAP (%) ↑
MT [23]	79.68	80.97	80.28	81.70	99.03	68.02	5.47	80.88
RLSSS [35]	80.60	81.74	79.64	83.95	99.05	69.11	5.00	78.67
DTC [36]	80.51	81.96	80.28	81.70	99.09	69.44	5.54	76.61
DTML [37]	80.79	82.28	81.75	82.83	99.10	69.90	5.00	79.67
MC-Net [38]	80.74	82.19	**82.16**	82.23	99.10	69.77	5.38	76.45
RMT_VAT [39]	81.19	82.34	80.10	**84.71**	99.08	69.98	4.74	81.19
SS-Net [40]	81.35	82.49	81.27	83.75	99.10	70.20	5.10	81.35
Ours	**81.53**	**82.57**	81.50	83.66	**99.11**	**70.32**	**4.72**	**84.00**

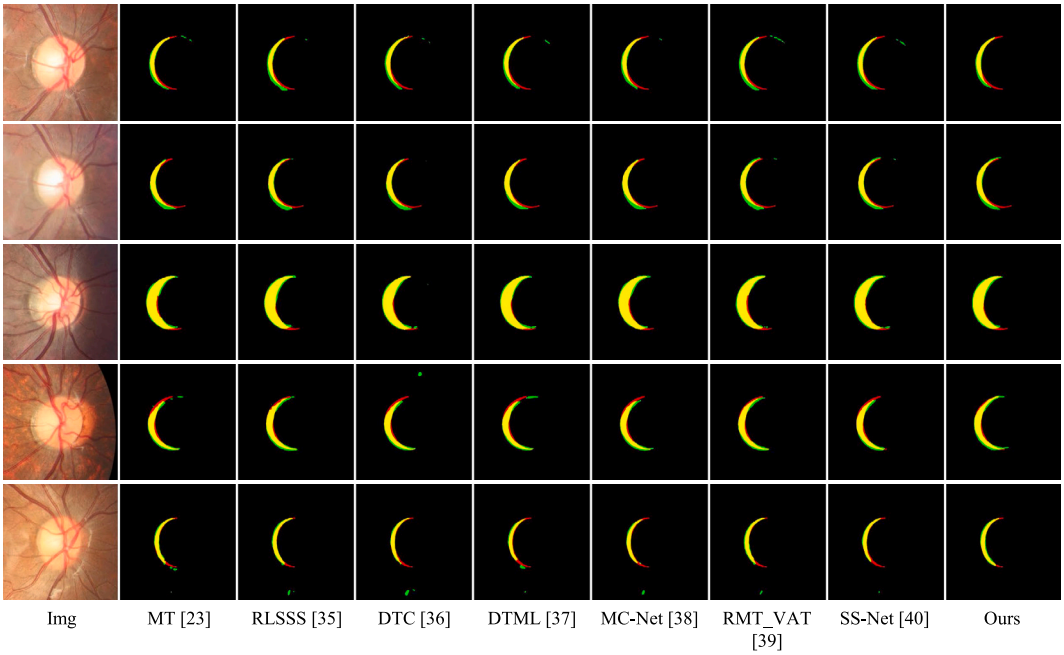

Fig. 6. Segmentation results of different semi-supervised methods on our PPA dataset.

Table 7
Segmentation results of different post-processing methods.

	Dice (%) ↑	F1 (%) ↑	Pre (%) ↑	Rec (%) ↑	Acc (%) ↑	IoU (%) ↑	ASSD↓	mAP↑
MT+RCC	79.81	80.99	80.20	81.81	99.03	68.06	5.22	81.87
MT+RCC+Post_CRF	78.30	79.61	**85.70**	74.33	99.04	66.13	6.32	–
MT+RCC+Post_ASM	79.94	81.23	81.68	80.80	99.06	68.40	5.35	–
MT+RCC+Post_DAE [17]	81.15	82.21	79.12	**85.42**	99.07	69.79	4.76	–
Ours	**81.53**	**82.57**	81.50	83.66	**99.11**	**70.32**	**4.72**	**84.00**

[1] The '–' indicates that the prediction result cannot calculate mAP.

not involved in this experiment, since the above methods are different in network structures. To compare with our proposed method fairly, DRIU [29] is employed as the segmentation backbone for all methods.

Comparison results of different semi-supervised segmentation methods are shown in Table 6. Among them, DTC and DTML methods adopt a dual-task model, and add shape constraints by calculating the signed distance maps. RMT_VAT and SS-Net methods constrain the pixel level smoothness to improve network performance. It can be observed from the table that our method is slightly lower than DTML and MC-Net in precision, and lower in recall when compared with RLSSS, RMT_VAT and SS-Net. Precision and recall are often mutually constrained, and only one of the indicators is higher in the above comparison methods. Our method achieves state-of-the-art performance when evaluated on comprehensive metrics such as Dice and F1-score, and obtains the closest segmentation boundary to the ground truth. Fig. 6 shows the visual comparison between our method and others. It can be seen that our model gets the most accurate results in these samples with smoother PPA boundary and more connective segmentation.

3.7.2. Comparison with shape constraint post-processing methods

In order to explore the effectiveness of shape constraints in our proposed method, we compare the segmentation results with methods using shape constraints as post-processing. In this case, the rough segmentation results before post-processing operation are obtained using the mean teacher model only with the region connectivity correction module. In this part, we compare the post-processing methods mentioned in Section 1 including CRF [13], ASM [27] and Post-DAE [17]. Quantitative results are shown in Table 7. It can be seen from the table that when using ASM and Post-DAE, the post-processing results of Dice are improved, but they are still lower than our method. The results verify the superiority of embedding ASM in loss in this paper. In terms of performance, our methods has reached the optimum, and our method can directly output the segmentation results end-to-end, while the post-processing method is a two-step operation. Compared with the post-processing methods, our method is more convenient. The visual results are shown in Fig. 7. From the figure, we can see that our method can make the final segmentation results smoother and play the role of shape constraint.

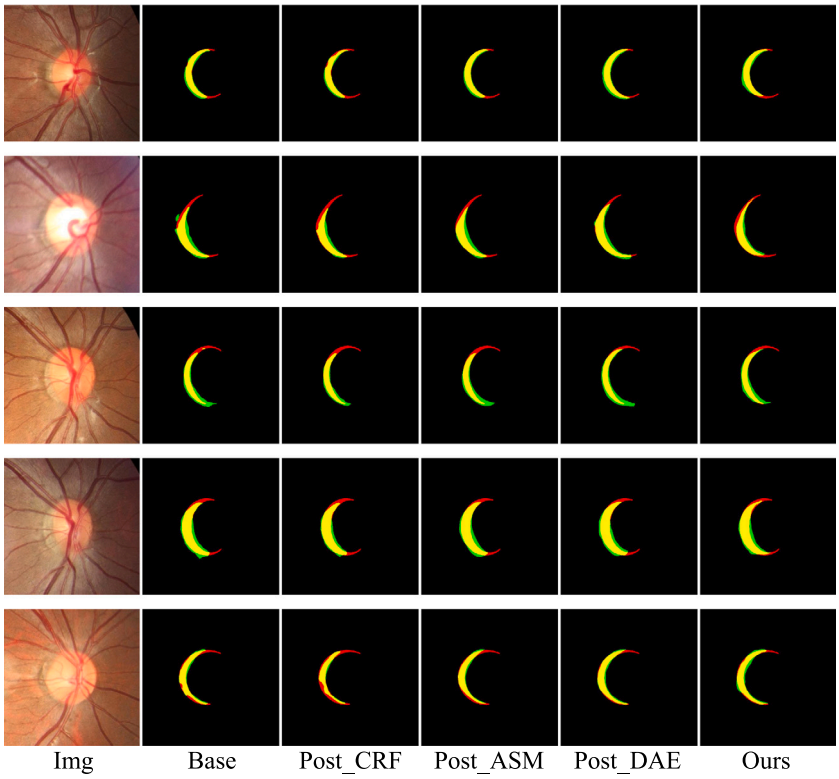

Fig. 7. Segmentation results of different post-processing methods.

Table 8
Quantitative results compared with other PPA segmentation methods.

		Dice (%) ↑	F1 (%) ↑	Pre (%) ↑	Rec (%) ↑	Acc (%) ↑	IoU (%) ↑	ASSD↓	mAP (%) ↑
Traditional method	EORL [8]	63.70	63.18	57.24	70.49	97.93	46.18	8.35	–
Supervised	MFCN [9]	76.49	78.40	80.67	76.25	98.94	64.47	9.93	74.58
	BGN [10]	79.99	81.52	80.18	82.92	99.05	68.81	5.18	81.59
	SGN [21]	81.03	82.28	80.99	83.61	99.09	69.89	4.89	83.93
Semi-supervised	Ours	**81.53**	**82.57**	**81.50**	**83.67**	**99.11**	**70.32**	**4.72**	**84.00**

[1] The '–' indicates that the prediction result cannot calculate mAP.

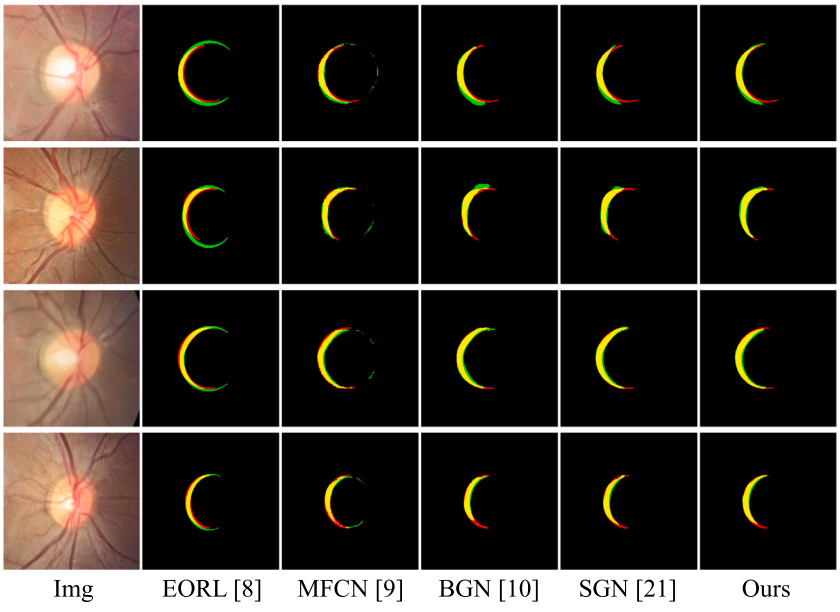

Fig. 8. Segmentation results of different PPA segmentation methods.

3.7.3. Comparison with methods on PPA segmentation

In this part, we compare our segmentation results with other state-of-the-art methods on PPA segmentation. It can be observed from Table 8 that the traditional method proposed in Ref. [8] achieves relatively lower performance than other deep learning methods. Compared with the fully supervised segmentation method in Refs. [9,10,21], our method improves on all metrics. It is important to note that the MFCN [9] is a multi-task method which segments OD and PPAOD simultaneously. The evaluation index mAP used in the original paper is the average result of calculating mAP for OD and PPAOD respectively. However, mAP in our paper is only calculated for the PPA segmentation results. If the evaluation index mAP in the original MFCN paper is used, it is 89.3%, which is close to the results shown in the paper, which proves the credibility of the recurring experiment.

Fig. 8 shows the visual comparison between our method and others. It can be observed that the method in Ref. [8] outputs smoother results. The reason is that it uses the ellipse constraint. However, due to the complicated brightness transformation around the OD, this method visually has a significant deviation from the ground truth of the boundary. The results obtained by MFCN are weak in edge continuity. The reason is that the edge continuity constraints are added to OD and PPAOD, but the edge continuity after subtraction is not guaranteed for PPA segmentation. In contrast, our method can directly segment the PPA, and the obtained results are closer to the ground truth.

4. Conclusion

In this paper, we propose a novel semi-supervised PPA segmentation method with shape constraints and connectivity information. The shape features can be achieved by the proposed shape constraint loss, which is performed using an ASM module. For unlabeled data, in order to ensure the pseudo label to be more accurate, the connectivity information is added by the proposed region connectivity correction module to guarantee that PPA is a consistent region in the semi-supervised pipeline. The proposed method achieves excellent performance quantitatively and qualitatively. In future work, we will further conduct quantitative analysis on PPA to explore its correlation with various eye diseases.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China (NSFC) under Grant 82072007, in part by the China Postdoctoral Science Foundation under Grant 2020M680387, and in part by the Beijing Institute of Technology Research Fund Program for Young Scholars, China.

References

[1] Varsha Manjunath, Heeral Shah, James G Fujimoto, Jay S Duker, Analysis of peripapillary atrophy using spectral domain optical coherence tomography, Ophthalmology 118 (3) (2011) 531–536.

[2] Jiao Sun, Jialin Wang, Yanling Wang, Research progress on morphology and associations of peripapillary atrophy, Chin. J. Ocular Fundus Dis. (2019) 510–513.

[3] Sharon Kingman, Glaucoma is second leading cause of blindness globally, Bull. World Health Organ. 82 (2004) 887–888.

[4] Lan Chang, Chen-Wei Pan, Kyoko Ohno-Matsui, Xiaoyu Lin, Gemmy CM Cheung, Gus Gazzard, Victor Koh, Haslina Hamzah, E Shyong Tai, Su Chi Lim, et al., Myopia-related fundus changes in Singapore adults with high myopia, Am. J. Ophthalmol. 155 (6) (2013) 991–999.

[5] Cheng-Kai Lu, Tong Boon Tang, F Murray Alan, Augustinus Laude, 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.

[6] Yandong Tang, Xiaomao Li, Axel von Freyberg, Gert Goch, Automatic segmentation of the papilla in a fundus image based on the cv model and a shape restraint, in: 18th International Conference on Pattern Recognition (ICPR'06), vol. 1, IEEE, 2006, pp. 183–186.

[7] Cheng-Kai Lu, Tong Boon Tang, Augustinus Laude, Baljean Dhillon, Alan F Murray, Parapapillary atrophy and optic disc region assessment (PANDORA): retinal imaging tool for assessment of the optic disc and parapapillary atrophy, J. Biomed. Opt. 17 (10) (2012) 106010.

[8] Hanxiang Li, Huiqi Li, Jieliang Kang, Yunlong Feng, Jie Xu, Automatic detection of parapapillary atrophy and its association with children myopia, Comput. Methods Programs Biomed. 183 (2020) 105090.

[9] Yidong Chai, Hongyan Liu, Jie Xu, A new convolutional neural network model for peripapillary atrophy area segmentation from retinal fundus images, Appl. Soft Comput. 86 (2020) 105890.

[10] Mengxuan Li, He Zhao, Jie Xu, Huiqi Li, Peripapillary atrophy segmentation with boundary guidance, in: Ophthalmic Medical Image Analysis: 8th International Workshop, OMIA 2021, Held in Conjunction with MICCAI 2021, Strasbourg, France, September 27, 2021, Proceedings 8, Springer, 2021, pp. 101–108.

[11] Kai Hu, Qinghai Gan, Yuan Zhang, Shuhua Deng, Fen Xiao, Wei Huang, Chunhong Cao, Xieping Gao, Brain tumor segmentation using multi-cascaded convolutional neural networks and conditional random field, IEEE Access 7 (2019) 92615–92629.

[12] Bhargav J Bhatkalkar, Dheeraj R Reddy, Srikanth Prabhu, Sulatha V Bhandary, Improving the performance of convolutional neural network for the segmentation of optic disc in fundus images using attention gates and conditional random fields, IEEE Access 8 (2020) 29299–29310.

[13] Philipp Krähenbühl, Vladlen Koltun, Efficient inference in fully connected crfs with gaussian edge potentials, Adv. Neural Inf. Process. Syst. 24 (2011).

[14] Pooneh R Tabrizi, Awais Mansoor, Juan J Cerrolaza, James Jago, Marius George Linguraru, Automatic kidney segmentation in 3D pediatric ultrasound images using deep neural networks and weighted fuzzy active shape model, in: 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), IEEE, 2018, pp. 1170–1173.

[15] Yuzhou Hu, Yi Guo, Yuanyuan Wang, Jinhua Yu, Jiawei Li, Shichong Zhou, Cai Chang, Automatic tumor segmentation in breast ultrasound images using a dilated fully convolutional network combined with an active contour model, Med. Phys. 46 (1) (2019) 215–228.

[16] Lingyun Cai, Yuanyuan Wang, A phase-based active contour model for segmentation of breast ultrasound images, in: 2013 6th International Conference on Biomedical Engineering and Informatics, IEEE, 2013, pp. 91–95.

[17] Agostina J Larrazabal, César Martínez, Ben Glocker, Enzo Ferrante, Post-DAE: anatomically plausible segmentation via post-processing with denoising autoencoders, IEEE Trans. Med. Imaging 39 (12) (2020) 3813–3820.

[18] 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 Trans. Med. Imaging 37 (2) (2017) 384–395.

[19] Arnaud Boutillon, Bhushan Borotikar, Valérie Burdin, Pierre-Henri Conze, Multi-structure bone segmentation in pediatric MR images with combined regularization from shape priors and adversarial network, Artif. Intell. Med. 132 (2022) 102364.

[20] Chen Chen, Carlo Biffi, Giacomo Tarroni, Steffen Petersen, Wenjia Bai, Daniel Rueckert, Learning shape priors for robust cardiac MR segmentation from multi-view images, in: Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part II 22, Springer, 2019, pp. 523–531.

[21] Mengxuan Li, He Zhao, Jie Xu, Huiqi Li, Peripapillary atrophy segmentation based on ASM loss, in: 2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI), IEEE, 2022, pp. 1–4.

[22] Antti Tarvainen, Harri Valpola, Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results, Adv. Neural Inf. Process. Syst. 30 (2017).

[23] Christian S. Perone, Julien Cohen-Adad, Deep semi-supervised segmentation with weight-averaged consistency targets, in: Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, Proceedings 4, Springer, 2018, pp. 12–19.

[24] Wenhui Cui, Yanlin Liu, Yuxing Li, Menghao Guo, Yiming Li, Xiuli Li, Tianle Wang, Xiangzhu Zeng, Chuyang Ye, Semi-supervised brain lesion segmentation with an adapted mean teacher model, in: Information Processing in Medical Imaging: 26th International Conference, IPMI 2019, Hong Kong, China, June 2–7, 2019, Proceedings 26, Springer, 2019, pp. 554–565.

[25] Xiaomeng Li, Lequan Yu, Hao Chen, Chi-Wing Fu, Lei Xing, Pheng-Ann Heng, Transformation-consistent self-ensembling model for semisupervised medical image segmentation, IEEE Trans. Neural Netw. Learn. Syst. 32 (2) (2020) 523–534.

[26] Sukesh Adiga Vasudeva, Jose Dolz, Herve Lombaert, Leveraging labeling representations in uncertainty-based semi-supervised segmentation, in: Medical Image Computing and Computer Assisted Intervention–MICCAI 2022: 25th International Conference, Singapore, September 18–22, 2022, Proceedings, Part VIII, Springer, 2022, pp. 265–275.

[27] Timothy F Cootes, Christopher J Taylor, David H Cooper, Jim Graham, Active shape models-their training and application, Comput. Vis. Image Underst. 61 (1) (1995) 38–59.

[28] Max Jaderberg, Karen Simonyan, Andrew Zisserman, et al., Spatial transformer networks, Adv. Neural Inf. Process. Syst. 28 (2015).

[29] Kevis-Kokitsi Maninis, Jordi Pont-Tuset, Pablo Arbeláez, Luc Van Gool, Deep retinal image understanding, in: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2016: 19th International Conference, Athens, Greece, October 17-21, 2016, Proceedings, Part II 19, Springer, 2016, pp. 140–148.

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

[31] Varduhi Yeghiazaryan, Irina Voiculescu, Family of boundary overlap metrics for the evaluation of medical image segmentation, J. Med. Imaging 5 (1) (2018) 015006.

[32] Mark Everingham, Luc Van Gool, Christopher KI Williams, John Winn, Andrew Zisserman, The pascal visual object classes (voc) challenge, Int. J. Comput. Vis. 88 (2009) 303–308.

[33] Olaf Ronneberger, Philipp Fischer, Thomas Brox, U-net: Convolutional networks for biomedical image segmentation, in: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18, Springer, 2015, pp. 234–241.

[34] Hengshuang Zhao, Jianping Shi, Xiaojuan Qi, Xiaogang Wang, Jiaya Jia, Pyramid scene parsing network, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 2881–2890.

[35] Xiangyun Zeng, Rian Huang, Yuming Zhong, Dong Sun, Chu Han, Di Lin, Dong Ni, Yi Wang, Reciprocal learning for semi-supervised segmentation, in: Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part II 24, Springer, 2021, pp. 352–361.

[36] Xiangde Luo, Jieneng Chen, Tao Song, Guotai Wang, Semi-supervised medical image segmentation through dual-task consistency, in: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 35, 2021, pp. 8801–8809.

[37] Yichi Zhang, Jicong Zhang, Dual-task mutual learning for semi-supervised medical image segmentation, in: Pattern Recognition and Computer Vision: 4th Chinese Conference, PRCV 2021, Beijing, China, October 29–November 1, 2021, Proceedings, Part III 4, Springer, 2021, pp. 548–559.

[38] Yicheng Wu, Zongyuan Ge, Donghao Zhang, Minfeng Xu, Lei Zhang, Yong Xia, Jianfei Cai, Mutual consistency learning for semi-supervised medical image segmentation, Med. Image Anal. 81 (2022) 102530.

[39] Qing Wang, Xiang Li, Mingzhi Chen, Lingna Chen, Junxi Chen, A regularization-driven mean teacher model based on semi-supervised learning for medical image segmentation, Phys. Med. Biol. 67 (17) (2022) 175010.

[40] Yicheng Wu, Zhonghua Wu, Qianyi Wu, Zongyuan Ge, Jianfei Cai, Exploring smoothness and class-separation for semi-supervised medical image segmentation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2022, pp. 34–43.