Automatic Segmentation of Cortex and Nucleus in Anterior Segment OCT Images

Pengshuai Yin[1], Mingkui Tan[1], Huaqing Min[1], Yanwu Xu[2,4(✉)],
Guanghui Xu[1], Qingyao Wu[1], Yunfei Tong[2], Higashita Risa[3], and Jiang Liu[4]

[1] South China University of Technology, Guangzhou, China
[2] Guangzhou Shiyuan Electronic Technology Company Limited, Guangzhou, China
`ywxu@ieee.org`
[3] Tommy Corporation, Nagoya, Japan
[4] Cixi Institute of Biomedical Engineering, Chinese Academiy of Sciences,
Cixi, China

Abstract. We propose a pipeline for automatically segmenting cortex and nucleus in a 360-degree anterior segment optical coherence tomography (AS-OCT) image. The proposed pipeline consists of a U-shaped network followed by a shape template. The U-shaped network predicts a mask for cortex and nucleus. However, the boundary between cortex and nucleus is weak, so that the boundary of the prediction is an irregular shape and does not satisfy the physiological structure of nucleus. To address this problem, in the second step, we design a shape template according to the physiological structure of nucleus to refine the boundary. Our method integrates both appearance and structure information. The accuracy is measured by the normalized mean squared error (NMSE) between ground truth line and predicted line. We achieve NMSE 7.09/7.94 for nucleus top/bottom boundary and 2.49/2.43 for cortex top/bottom boundary.

Keywords: SS-OCT · AS-OCT · Image segmentation

1 Introduction

Anterior segment optical coherence tomography (AS-OCT) can assist the diagnosis of many eye diseases, such as glaucoma and cataract [2]. The measurement is made without contact and with low risk of infection. The role of AS-OCT in research and clinical care continues to accelerate [3].

AS-OCT nuclear density measurement is a repeatable and reliable objective cataract grading method, It is correlated with the Lens Opacity Classification System Version III (LOCS III) grading [5,8]. The nuclear density is got by delineates the lens nucleus and calculates the total average pixel intensity. If nucleus

---

P. Yin and M. Tan—Equally contribution to this work.

can be automatically segmented, cataract grading can be automatically acquired. To the best of our knowledge, no previous work focus on automatically segmenting the nucleus in AS-OCT images.

In this paper, we propose a pipeline to automatically segment cortex and nucleus in AS-OCT images. The proposed pipeline consists of a U-shaped network followed by a shape template. The U-shaped network predicts a preliminary mask for cortex and nucleus. However, the predicted boundary of nucleus is arbitrary because the boundary between cortex and nucleus is weak. To solve this problem, we design a shape template based on the physiological structure of nucleus to refine the boundary of nucleus. The basic idea of the refinement is to find a template in the training set to replace the boundary of the prediction. After the refinement, the boundary of nucleus satisfies the physiological structure of nucleus.

We summarize the contributions of this work as follows:

– We propose a simple and effective pipeline to segment cortex and nucleus by using a U-shaped network and a shape template. This method integrates both structure information and appearance information.
– We design a shape template that imitates the intrinsic concentric layers structure of nucleus. By using the template to refine the boundary of nucleus, the final prediction satisfies the physiological structure of nucleus.

2   Proposed Method

The proposed pipeline is shown in Fig. 1. First, we find the lens region using the Canny edge detector and divide the lens region into three sub-regions: capsule region, cortex region and nucleus region. We train a U-shaped network to predict a mask for each region. However, the output of the U-shaped network has no regular shape, especially for the nucleus which has a weak boundary. So we design a shape template to model the structure of nucleus and use the template to refine the boundary of nucleus.

2.1   Network Architecture

Motivated by [6]. We design a U-shaped network to predict a preliminary mask for capsule, cortex and nucleus. The U-shaped network can obtain a high-resolution mask with a clear boundary by using skip connections to restore the information loss caused by pooling layers.

Our network mainly includes two modules: encoding module and decoding module. The encoding module consists of six blocks. Each block contains two or three convolutional layers, and each convolutional layer is followed by a rectified linear unit (ReLU) and a 2×2 max pooling operation with stride 2. The decoding module is also composed of six blocks. Each block consists of a concatenation with the corresponding feature maps from the contracting path, a spatial upsampling of the feature maps with a factor of 2 followed by two convolutions and
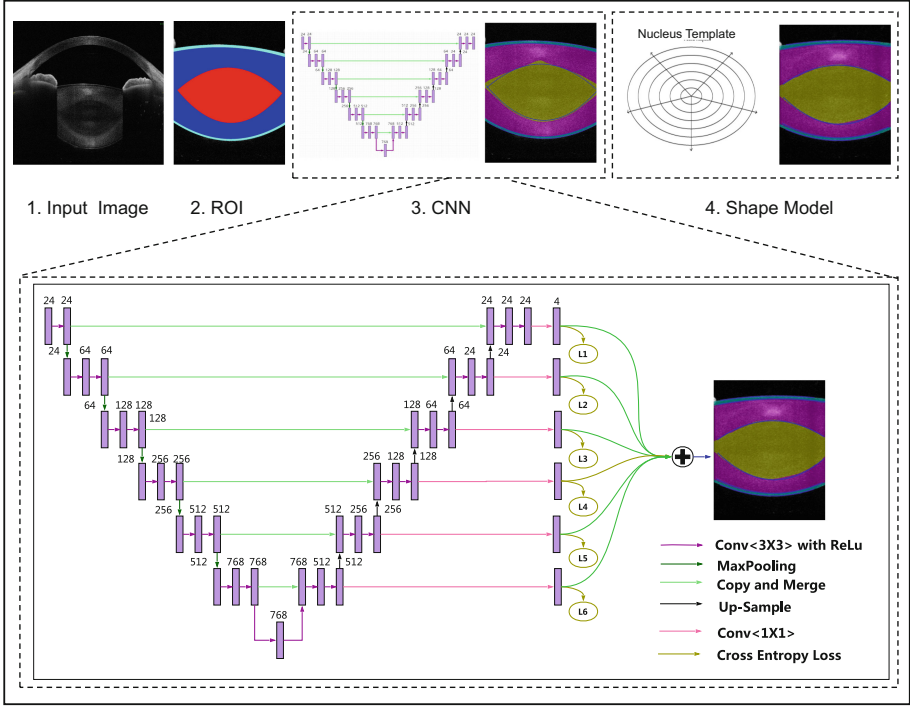
Fig. 1. The proposed pipeline: 1, AS-OCT image. 2, Region of interest (ROI). We locate the lens and divide the lens into three sub-regions according to the ground truth annotations (blue lines), pink: cortex region, yellow: nucleus region, cyan: capsule region. 3, U-shaped network with side-output predicts a mask for each region. 4, Shape template refines the boundary of nucleus. (Color figure online)

a ReLU layer. At the final layer, a 1×1 convolution maps each 24-dim feature vector to the desired number of classes (here 4).

Motivated by [1], we add a side-output layer which acts as a classifier that produces a companion local output for early layers and also integrates different level information. Cross-entropy loss is used for each side-output layer. There are M side-outputs in the network. The loss function of the side-output layer is given as:

$$L_{side-output} = \frac{1}{M} \sum_{m=1}^{M} L_{cross-entropy}(y, y'), \tag{1}$$

$L_{cross-entropy}$ is the cross entropy loss:

$$L_{cross-entropy} = -\sum_{i} (y'_i \log(y_i)), \tag{2}$$

y_i is the predicted probability value for class i and y_i' is the true probability for that class. The overall structure of the U-shaped network with side-outputs is shown in Fig. 1.

2.2   Shape Template

As shown in Fig. 2 (left). The U-shaped network tends to misclassify the areas similar to nucleus as nucleus because of the weak boundary between nucleus and cortex. To solve the problem, we design a shape template based on the physiological structure of nucleus to refine the boundary of nucleus. The basic idea is to find the closest shape of nucleus in the training set to replace the boundary of prediction.
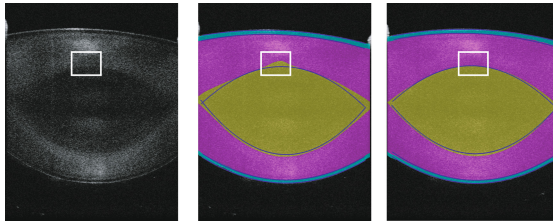

Fig. 2. Blue lines are ground truth annotations. Left: AS-OCT image, region in white rectangle shows weak contrast boundary. Mid: Prediction of ours U-shaped network, the U-shaped network misclassify the regions similar to nucleus as nucleus, shown in white rectangle. Right: Shape template refine the boundary of nucleus. (Color figure online)
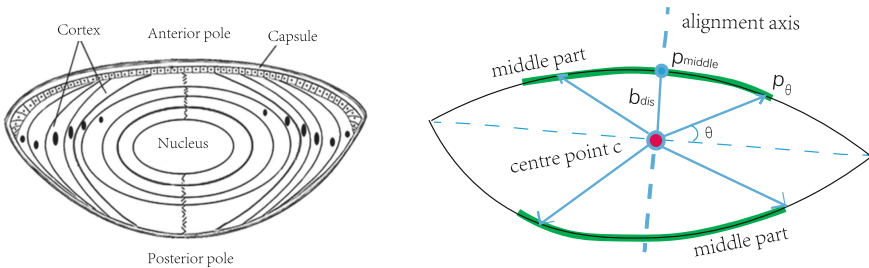

Fig. 3. Left: The structure of the lens, the lens fibres are arranged in concentric layers. Right: The shape template of nucleus, p_{middle} is the intersection between boundary and alignment axis, b_{dis} is the distance from center point c to point p_{middle}, the middle part of the template is shown in green. (Color figure online)

Lens fibers form the bulk of the lens. The lens fibers stretch lengthwise from the posterior to the anterior poles and are arranged in concentric layers rather like the layers of an onion when cutting horizontally [4], as shown in Fig. 3

(left). Motivated by the ray feature [7], the structure of concentric layers can be represented by a center point and the distance between the center point and the nearest point on the boundary, as shown in Fig. 3 (right). Different layers share the same center point and different distances to the center point.
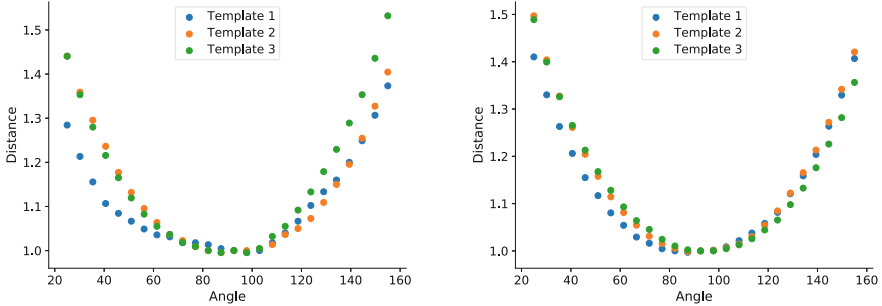

Fig. 4. Three nucleus templates of the top part (left) and the bottom part (right). Different templates are shown in different colors.

The boundary of nucleus is represented by n points: $S_i = \{x_i, y_i\}$. The center point of nucleus is defined as $c = (\sum \frac{x_i}{n}, \sum \frac{y_i}{n})$. Then the boundary of nucleus is encoded by c and the distance from c to the nearest point p_θ on the boundary of nucleus in direction θ:

$$f(I, c, \theta) = \frac{\|c - p_\theta\|}{z}, \tag{3}$$

where $z = \|c - p_{middle}\| = b_{dis}$ is a normalization factor. p_{middle} is the intersection of alignment axis and the boundary of nucleus as shown in Fig. 3 (right). p_{middle} is different in each image due to the eye movement between AS-OCT images mentioned in [9]. Normalization not only ease the rotational misalignment problem but also eliminates scale effects. We sample θ every 5 degrees. In this way all the shapes are encoded by K points $\{\theta_k, f(I, c, \theta_k)\}$.

The training procedure is shown in Algorithm 1. The purpose of training is to learn all possible shapes of nucleus. For each shape in the training set, we encode it into K points $\{\theta_k, f(I, c, \theta_k)\}$ and then cluster them into v templates. In the experiments, we only learn the middle part of shape and use quadratic curve fitting to get the other part because the shape of middle part is relatively stable. The middle part is divided into a top part and a bottom part as shown in Fig. 3 (right). For each part, we cluster the corresponding shapes into v' templates using K-means. Figure 4 shows three normalized templates of the top part and the bottom part.

The refinement procedure is shown in Algorithm 2, s is the boundary of nucleus predicted by the U-shaped network. We calculate the center point c', alignment axis and p'_{middle} of s. The next step is to find a template closest to s. For each template T_v, we finetune the template by multiply T_v by z, where

---

Algorithm 1. Training shape template

Input: Training shape $S_i = \{x_i, y_i\}$
Output: V templates $T_v = \{f_v(I, c, \theta)\}$
for *Each shape S_i in training set* **do**
    Center point: $c = (\sum \frac{x_i}{n}, \sum \frac{y_i}{n})$
    Calculate alignment axis and p_{middle} of S_i
    Encode S_i into K points $\{\theta_k, f_i(I, c, \theta_k)\}$
end
Extract the middle part of the shape
Using K-Means to get V templates $\{f_v(I, c, \theta)\}$

Algorithm 2. Fitting shape template

Input: V templates $T_v = \{f_v(I, c, \theta)\}$, unrefined shape $s = \{x_j, y_j\}$
Output: Refined shape s'
Calculate center point c', alignment axis and p'_{middle} of s
for *Each template T_v* **do**
    $T'_v = T_v * z$, where $z = \|c' - p'_{middle}\|$
    for *offset $\leftarrow \{-10, -9, ..., 9, 10\}$* **do**
        $B = T'_v +$offset
        Calculate the similarity between s and B
    end
end
Find the template $B' = B$ with maximum similarity
The refined shape $s' = B'$

$z = \|c' - p'_{middle}\|$. After the multiplication, p_{middle} of T_v is coincide with p'_{middle} of s. Then we add an offset to T_v. The positive offset means the template move to outer layer. For each transformed template B, we calculate the similarity between B and s. We find a most similar template B' to replace s. B and s can be represent by K points: $\{\theta_k, f_B(I, c, \theta_k)\}$ and $\{\theta_k, f_s(I, c, \theta_k)\}$. The similarity between B and s is defined as:

$$Similarity(B, s) = \Sigma_{k=1}^K | f_B(I, c, \theta_k) - f_s(I, c, \theta_k) |. \tag{4}$$

3 Performance Evaluation and Discussion

We acquire the data from CASIA-2000 anterior SS-OCT produced by Tomey Co. Ltd. The dataset contains 20 eyes from 10 people, 8 images per eye. We select 7 people (120 images) for training and 3 people (40 images) for testing. All the images are annotated by one experienced ophthalmologist. The accuracy is measured by the normalized mean squared error (NMSE) between the predicted shape $S_p = \{x_i, y_i\}$ and the ground truth $S_g = \{x_j, y_j\}$, it is defined as

$$NMSE = \frac{\sum_{n_g} \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}}{n_g}, \tag{5}$$

where n_g is the number of annotation points. The results are shown in Table 1. Nucleus top is the top boundary of nucleus. Nucleus bottom is the bottom boundary of nucleus. The same thing is conducted for the cortex.

The U-shaped network is trained from scratch. The initial learning rate is 0.001. The network is trained for 70 epochs. The origin image size is 2130×1864, we resize the image into 1024×1024. For shape template, we learn 40 templates for nucleus top and 40 templates for nucleus bottom from the training set. The offset range of template is set to $[-10, 10]$. As shown in Table 1, the U-shaped network is better than U-Net [6] and M-Net [1] because of the larger reception field. Side-output layer adds the supervision to mid layers and eases the difficulty to train the network. Side-output layer also integrates different scale information and obviously improve the results. However, multi-scale structure such as M-Net shows no improvement. The shape template refines the boundary of nucleus predicted by the U-shaped network with side-output. The entire refinement process can be seen as finding the most similar template in the training set to replace the boundary of nucleus. Using shape template to refine the nucleus boundary not only improves the performance but also makes the prediction consistent with the physiological structure of nucleus. More results are shown in Fig. 5.

Table 1. Accuracy of the segmentation

Method	Nucleus top	Nucleus bottom	Cortex top	Cortex bottom
U-Net [6]	12.4	12.8	3.19	2.73
M-Net [1]	15.75	13.04	3.03	2.82
U-shaped Net (ours)	8.57	9.83	**2.45**	2.74
Side output (ours)	8.28	8.96	2.49	**2.43**
Shape template (ours)	**7.09**	**7.94**		

Note: Shape template only refines the boundary of nucleus

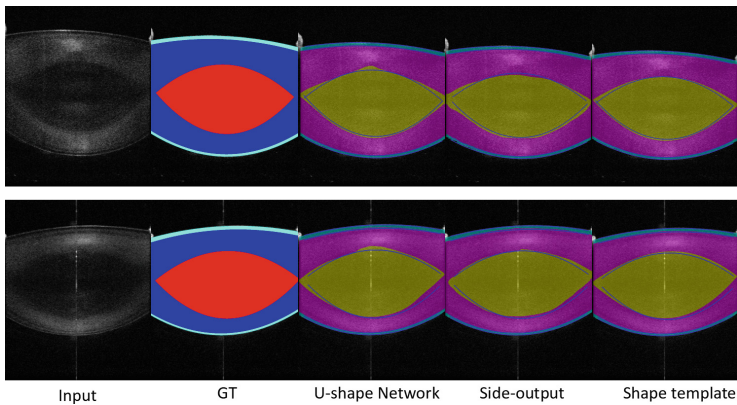Input   GT   U-shape Network   Side-output   Shape template

Fig. 5. Example results. From left to right: Input image, ground truth, prediction of the U-shape Network, prediction after side-output layer, prediction after the refinement. Blue lines are ground truth annotations. (Color figure online)

4 Conclusions

In this paper, we use a U-shaped network to produce a preliminary mask and we design a shape template to refine the mask. The experiments show the effectiveness of our method. After the refinement, the result satisfies the physiological structure of the lens.

References

1. Fu, H., Cheng, J., Xu, Y., Wong, D.W.K., Liu, J., Cao, X.: Joint optic disc and cup segmentation based on multi-label deep network and polar transformation. IEEE Trans. Med. Imag. **37**(7), 1597–1605 (2018)
2. Fu, H., et al.: Segmentation and quantification for angle-closure glaucoma assessment in anterior segment OCT. IEEE Trans. Med. Imag. **36**(9), 1930–1938 (2017)
3. Fujimoto, J., Huang, D.: Foreword: 25 years of optical coherence tomography. Invest. Ophthalmol. Vis. Sci. **57**(9), OCTi–OCTii (2016)
4. Kaufman, P.L., Levin, L.A., Adler, F.H., Alm, A.: Adler's Physiology of the Eye. Elsevier Health Sciences, New York (2011)
5. Panthier, C., Burgos, J., Rouger, H., Saad, A., Gatinel, D.: New objective lens density quantification method using swept-source optical coherence tomography technology: comparison with existing methods. J. Cataract Refract. Surg. **43**(12), 1575–1581 (2017)
6. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) MICCAI 2015. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28
7. Smith, K., Carleton, A., Lepetit, V.: Fast ray features for learning irregular shapes. In: IEEE 12th International Conference on Computer Vision, pp. 397–404. IEEE (2009)
8. Wong, A.L., et al.: Quantitative assessment of lens opacities with anterior segment optical coherence tomography. Br. J. Ophthalmol. **93**(1), 61–65 (2009)
9. Xu, Y., et al.: Axial alignment for anterior segment swept source optical coherence tomography via robust low-rank tensor recovery. In: Ourselin, S., Joskowicz, L., Sabuncu, M.R., Unal, G., Wells, W. (eds.) MICCAI 2016. LNCS, vol. 9902, pp. 441–449. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46726-9_51