Deep Learning-Based Glaucoma Diagnostic Assistance System on Mobile Devices

Bin Zhou, Yan Jiang, Ningyi Zhang, Yijian Fu, Yugen Yi, and Qiangqiang Zhou(✉)

School of Software, Jiangxi Normal University, Nanchang, China
{zbb,202040100754,zny,yijianfu,yiyg510,qiang}@jxnu.edu.cn

Abstract. Glaucoma is an irreversible, chronic eye disease for which there is no effective treatment. It is caused by elevated intraocular pressure that damages the optic nerve. There are no symptoms in the early stages of glaucoma, so early diagnosis is crucial to prevent blindness. In this paper, an accurate, reliable, and rapid glaucoma diagnosis method based on computer-aided detection (CAD) technology is presented. The effective use of CAD can significantly reduce the workload and burden of clinical doctors. The CAD system designed in this paper is successfully deployed on the Android platform and provides a user-friendly interface for clinical use. Experimental results demonstrate that the proposed strategy may precisely segment the optic cup and disc in retinal fundus images. Additionally, this research establishes the viability of using deep learning models on mobile devices to partition the optic cup and disc in fundus images.

Keywords: Glaucoma detection · U-net · Diagnostic Assistance System

1 Introduction

In persons aged 50 and older, glaucoma is the second most common cause of blindness and the fourth most common cause of moderate to severe vision loss. Statistics show that there were more than 76 million glaucoma sufferers globally in 2020, and that figure is projected to increase to 95.4 million by 2030 [1]. Glaucoma causes irreversible vision loss. Blindness risk can be decreased by around 50% with an early and precise diagnosis, followed by the proper treatment [2]. Therefore, there is a critical need for precise and effective screening methods for glaucoma early diagnosis.

There are various methods for detecting glaucoma, including intraocular pressure testing, fundus examination, visual acuity testing, and optical coherence tomography. Among them, intraocular pressure testing is the most common and simplest method, which determines the presence of glaucoma by measuring the intraocular pressure. Fundus examination can detect pathological changes in the eye, diagnose the degree of glaucoma, and whether there are other eye lesions. Optical coherence tomography is a high-tech method for detecting glaucoma, which can provide more accurate information about the structure and function of the eye. Regardless of the method used, glaucoma detection requires regular screening, especially for high-risk populations like the elderly, individuals with a family history of glaucoma, and diabetes patients. This can help ensure eye health and prevent the development and progression of eye disease.

Y. Cao and X. Shao (Eds.): DIONE 2023, LNICST 515, pp. 63–75, 2025.
https://doi.org/10.1007/978-3-031-80713-8_5

2 Related Works

In the course of the development of retinal image analysis, the segmentation of the optic cup(OC) and optic disc(OD) has consistently been an attractive study area. Numerous strategies have been presented forth, which can be roughly categorized into deep learning-based segmentation algorithms and conventional segmentation algorithms.

2.1 Machine Learning-Based Segmentation Methods

The evaluation of fundus images utilizes a variety of traditional machine learning segmentation algorithms to assist ophthalmologists in effectively diagnosing retinal issues at an early stage. These algorithms often base on retinal images and use edge detection, thresholding, morphological operations, and other image processing methods to separate the OD and OC. The segmentation of the OD and OC commonly employs the following conventional segmentation algorithms: Threshold-based segmentation [3] is a simple and effective segmentation method that divides the image into the OD and OC by setting a threshold on the image's grayscale values. Morphological-based segmentation methods [4] separate the OD and OC with morphological operations, such as erosion and dilation. Segmentation methods based on the watershed algorithm [5] use the watershed transform to separate the OD and OC. Projection-based segmentation methods [6] separate the optic cup and optic disc by using projections, such as the polar projection and polar angle projection. Above methods preprocess the image by obtaining the region of interest ROI and then segment the OD and OC by the proposed machine learning method as shown in Fig. 1.

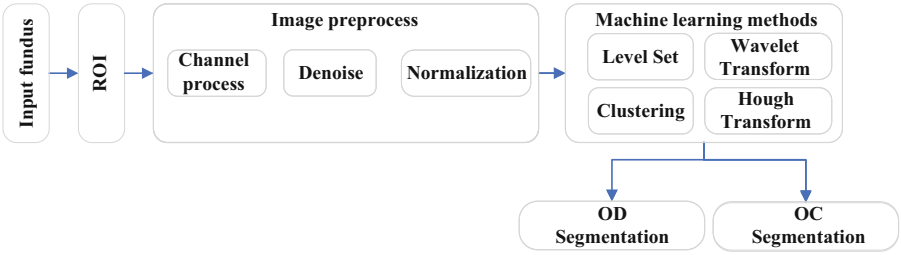

Fig. 1. Segmentation process of machine learning

2.2 Deep Learning Segmentation Methods

In contrast to traditional machine learning techniques that need manual feature extraction, deep learning-based systems may automatically identify characteristics from data. The approaches are hence increasingly advantageous as computer vision technology advances continuously. Additionally, deep learning models are typically trained using plenty of data and significant computational resources, which results in models with great accuracy and strong generalization capabilities.

The end-to-end feature optimization carried out automatically by deep learning algorithms helps them perform better. A deep learning-based technique to segment the OD and OC in retinal fundus images was carried out by Chen et al. [7]. They made use of a deep convolutional neural network (CNN) that has been trained on a sizable annotated dataset. When compared to conventional image processing techniques, the method produced results with improved accuracy and robustness. A fully convolutional network (FCN)-based technique to segment the OD and OC in retinal fundus images was put forth by Yao et al. [8]. They trained the model using a dataset of annotated images of the retinal fundus using an enhanced U-Net architecture. For the purpose of trying to segment the OD and OC in optical coherence tomography (OCT) images, Wei et al. [9] suggested a deep learning-based technique. Gharbi et al. [10] trained a multi-task deep neural network to segment the OD and OC in retinal fundus images. Conclusively belonging to the above, deep learning methods have been proficiently applied to medical image segmentation tasks and have achieved excellent performance (Fig. 2).

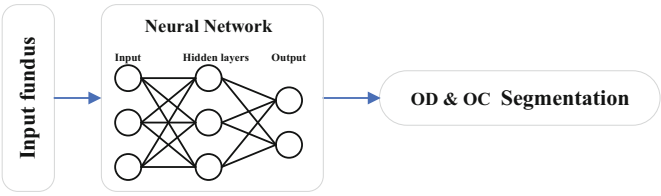

Fig. 2. Segmentation process using deep learning neural networks

3 U-Net-Related Segmentation Methods

3.1 U-net

The U-Net [11] model is a convolutional neural network (CNN)-based image segmentation method whose model structure resembles a U-shape, as seen in Fig. 3. A down-sampling encoder and an up-sampling decoder make up the model structure. While minimizing the size of the feature maps, the encoder extracts features from the image using CNNs and max-pooling layers. The decoder, which returns the size of the feature maps to their original size, is a deconvolutional neural network. The final feature maps are then semantically segregated using convolutional layers. The connection between the U-Net model's up-sampling and down-sampling components is its primary innovation. In the U-Net, the input to the decoder is the feature map from the corresponding layer of the encoder and a more detailed feature map from the same layer. While segmenting data using high-level features, this connection approach can maintain low-level features. U-Net has been successfully used to address a variety of medical picture segmentation issues and performs well in OD and OC segmentation tasks.

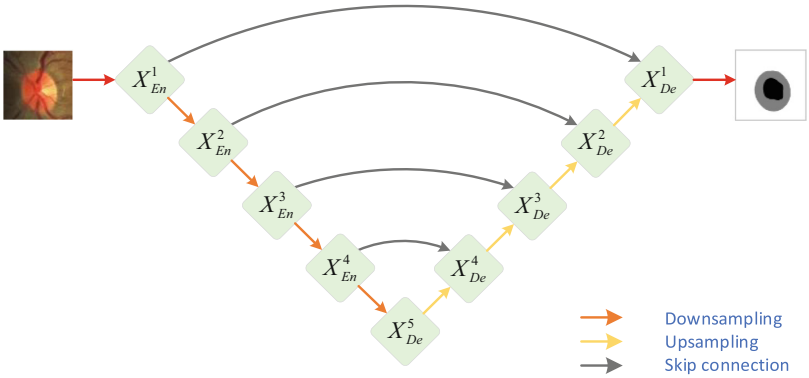

Fig. 3. U-Net architecture diagram

3.2 U-Net++

As seen in Fig. 4, the U-Net++ [12] is an enhanced version of the U-Net concept. In contrast to the U-Net model, U-Net++ initially employs a multi-level up-sampling technique to improve the model's capacity for semantic segmentation. Secondly, the model's efficiency is improved through feature reuse, that is, features between different layers can be reused multiple times, reducing the computational cost of the model. Finally, U-Net++ adds a global information pathway to the model's structure, enabling the model to simultaneously utilize global and local information, thereby improving its accuracy. U-Net++ shows positive performance in many image segmentation tasks, especially in handling more complex and larger images, where the efficiency and accuracy of U-Net++ have been significantly improved.

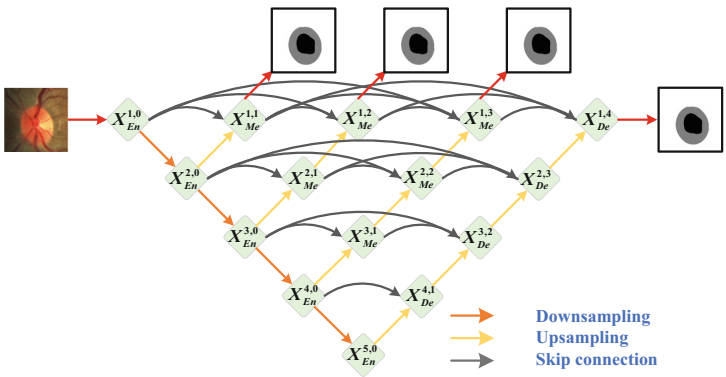

Fig. 4. U-Net++ model architecture diagram

3.3 U-Net3+

Development over U-Net+ +, U-Net3+ [13] has a similar U-shape structure to Fig. 5. It is an encoder-decoder construction that has an expanding and contracting route. The expanding path up-samples the feature maps to restore the entire resolution of the output segmentation, while the contracting path down-samples the input cup and disc to extract high-level features. Compared to U-Net++, U-Net3 + uses a more advanced technique, namely, dense connectivity. In traditional convolutional neural networks, feature maps are only transmitted from one layer to the next, while in U-Net3+, each layer is connected to every other layer in the same block, resulting in a denser feature representation. In addition, U-Net3+ also uses full-scale deep supervision information, generating a side output for each layer in the decoder through a convolutional layer, up-sampling, and activation function, which is compared with the true label.

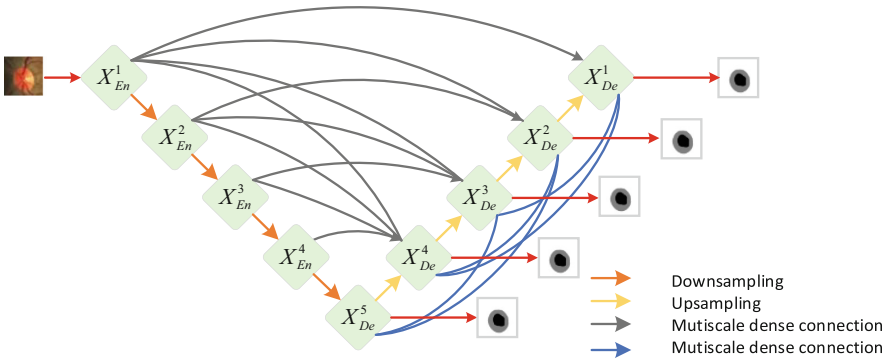

Fig. 5. U-Net3 + model architecture diagram

4 Glaucoma Diagnostic Assistance System

4.1 Hardware Configuration

In this article, the hardware environment is split between desktop and mobile devices. The desktop is utilized for model testing and training, as well as for transforming the learned model into a deployable version for mobile devices. Through appropriate library functions, the mobile device invokes the trained model for inference.

The desktop device is configured with an Intel® Core™ i9-10900K Processor CPU, 32G of RAM, and an Nvidia GeForce RTX 3090 GPU. Using the PyCharm integrated development environment, and Python programming language with version 3.8.16. Other toolboxes used include torch = 2.0.1, cuda = 11.8, cudnn = 8.3, and CUDA Driver Version = 526.47.

The mobile device used is a Redmi Note9 with MIUI12.5.9 operating system version, Android version, 128GB storage space, 8GB of RAM, and a MediaTek Dimensity 800U processor.

4.2 Dataset

DRIONS-DB [16] and Drishti-GS1 [17] are two open datasets that are used in this research. Each of the 110 600 × 400-resolution color fundus photos in the DRIONS-DB contains annotation data for two sets of optic discs. The first set is utilized in our lab for training and testing. The first 60 photographs are used for training, while the final 50 images are used for testing. There are a total of 101 retinal images in Drishti-GS1 with a resolution of 2896 × 1944, including 31 images of healthy eyes and 70 images with glaucoma. This dataset has two sets of photos: a training set of 50 images and a test set of 51 images. Four ophthalmologists, each with a different level of clinical competence, manually name each image's optic disc and cup, which are then shown as grayscale image boundaries with various number values. In this experiment, we take the intersection of the annotations from three experts as the ground truth, which is generated by thresholding the grayscale values at 65 to produce a binary mask image (Table 1).

Table 1. The information of DRIONS-DB and Drishti-GS datasets

Dataset	Origin Resolution	Resolution after crop	Number of images	Training images	Testing images	Number after augment
DRIONS-DB	600 × 400	192 × 192	110	60	50	1260
DRISHTI-GS	2896 × 1944	640 × 640	101	50	51	1050

4.3 Performance Indicators

When performing segmentation tasks, the DICE and Jaccard assessment indicators are frequently used to assess how closely the segmentation results are aligned with predictions [14].

4.3.1 DICE Index

The DICE index, also known as the Sørensen-Dice coefficient, is calculated as:

$$Dice(A, B) = \frac{2TP}{2TP + FP + FN} \tag{1}$$

where TP denotes the number of samples that the model properly identified as positive; False Positive (FP) stands for the proportion of samples that the model mistakenly predicts to be positive, and False Negative (FN) for the proportion of samples that the model incorrectly expects to be negative. The DICE index has a value between 0 and 1, and the closer it is to 1, the more accurately the model predicts the outcome in comparison to the actual outcome.

4.3.2 Jaccard Coefficient

Jaccard index [15], also known as Intersection over Union (IoU), is calculated as follows:

$$Jaccard\,(A, B) = \frac{TP}{TP + FP + FN} \tag{2}$$

where the definitions of *TP*, *FP*, and *FN* are the same as previously. The Jaccard index likewise has a value between 0 and 1, and the closer it is to 1, the closer the anticipated outcome of the model is to the actual result.

The DICE index is more skewed toward lower values than the Jaccard index, which is how the two metrics differ from one another. The DICE index and the Jaccard index may produce identical findings in some instances, while in others, the DICE index may be more sensitive to minute variations.

4.3.3 Cup-to-Disc Ratio (CDR)

The CDR simplifies the ratio of the optical cup and disc diameters. CDR measurements can be utilized to get a fast glaucoma diagnosis. A healthy person's fundus picture typically has a CDR value of 0.58 or less. The horizontal cup-to-disc ratio (HCDR), vertical cup-to-disc ratio (VCDR), and area cup-to-disc ratio (ACDR) are the three basic approaches for computing CDR values. The vertical diameter ratio of the optic disc and cup is one of them, and it is calculated as:

$$VCDR = \frac{V_{\text{cup}}}{V_{\text{disc}}} \tag{3}$$

where V_{disc} and V_{cup} are the vertical diameters of the OD and OC, respectively. HCDR is the ratio of the horizontal diameters of the OD and OC, which is calculated as:

$$HCDR = \frac{H_{\text{cup}}}{H_{\text{disc}}} \tag{4}$$

where H_{disc} and H_{cup} are the horizontal diameters of the OD and OC, respectively. ACDR is the ratio of the areas of the optic cup and optic disc, which is calculated as:

$$ACDR = \frac{A_{\text{cup}}}{A_{\text{disc}}} \tag{5}$$

where A_{disc} and A_{cup} are the areas of the optic disc and optic cup, respectively. The positions of H_{disc}, H_{cup}, V_{disc}, V_{cup}, A_{disc}, and A_{cup} in the color fundus photographs are shown in Fig. 6.

4.4 Algorithm Description

The OD and OC segmentation task based U-Net model in Algorithm 1.

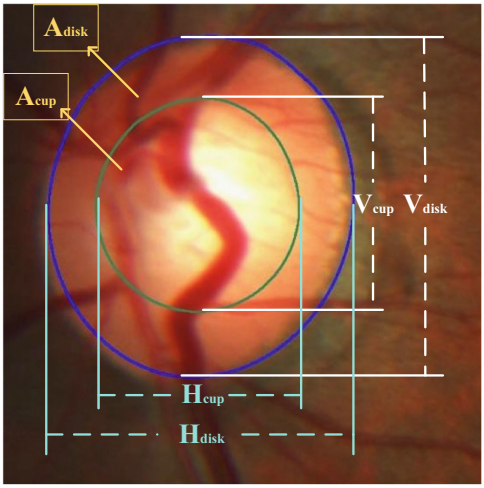

Fig. 6. Horizontal and vertical diameters of the optic disc and cup in color fundus photographs

Algorithm 1 OD and OC Segmentation whit U-Net

Input: Labeled training data $\mathrm{D}^l = \{(x_i, y_i)\}_{i=1}^{M}$, $U - Net/U - Net3+$

Output: Fine segmentation of OD and OC $Mask_F$

function COARSE SEGMENTATION*(D^l, U-Net)*
    Train *U-Net* on D^l with Dice loss $Loss_Dice$
    Obtain Coarse segmentation $Mask_C$ of OD
    Apply CHT on $Mask_C$ then crop to get ROI
    return cropped data D^c
end function
function FINE SEGMENTATION *(D^c, U-Net)*
    Train *U-Net* on D^c with hybrid loss $Loss_H$
    Obtain Fine segmentation $Mask_F$ of OD and OC
    return $Mask_F$
end function

4.5 Model Training

In the OD and OC segmentation challenge based on color fundus images, only a tiny region surrounding the OD is helpful for glaucoma screening and diagnosis. As seen in Fig. 7, this research suggests a two-stage method for segmenting the optic cup and disk in fundus images, starting with coarse segmentation before proceeding to fine segmentation. Combining a priori understanding that OD and OC both have prominent and cyclic elements. The U-Net model is first used to provide the coarse segmentation output for the fundus picture. The circular Hough transform (CHT) is then used to calculate the

disk's center and radius. The image's region of interest (ROI) is then determined by establishing the proper threshold. Finally, U-Net and U-Net3+ are employed for precise ROI-based segmentation. The resulting optic cup and disk diameters are then used to get a CDR glaucoma diagnosis.

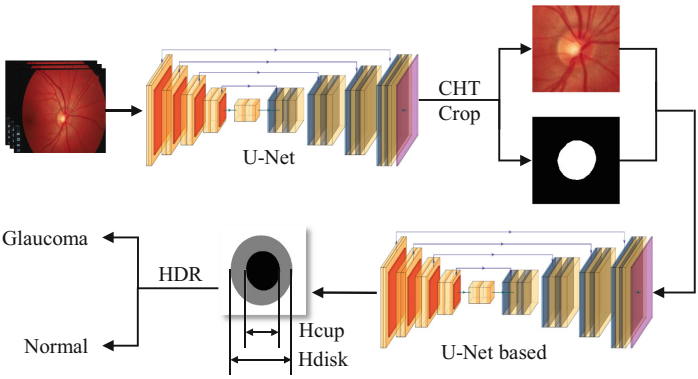

Fig. 7. The diagnosis process of the glaucoma CAD system

On the DRIONS-DB test set, the U-Net trained using the previously described method produced a Dice coefficient of 0.9566, and on the Drishti-GS test set, disk and cup segmentation Dice coefficients of 0.9589 and 0.8405, respectively. On the DRIONS-DB test set, the U-Net3+ achieved a Dice coefficient of 0.9627, and on the Drishti-GS test set, disk and cup segmentation Dice coefficients of 0.9679 and 0.8692, respectively (Table 2).

Table 2. Segmentation results of U-Net and U-Net3 + on the two datasets

Dice	DRIONS-DB	Drishti-GS	
	OD	OD	OC
U-Net	0.9566	0.9589	0.8405
U-Net3+	0.9627	0.9679	0.8692

4.6 Deployment on Mobile Devices

The diagnostic system designed in this paper first includes the function of selecting the model and dataset. As shown in Fig. 8(a), two models, U-Net and U-Net3 +, can be selected from the tab bar at the top. As shown in Fig. 8(b), two datasets can be selected from the drop-down list below the page, namely DRIONS-DB, and Drishti-GS.

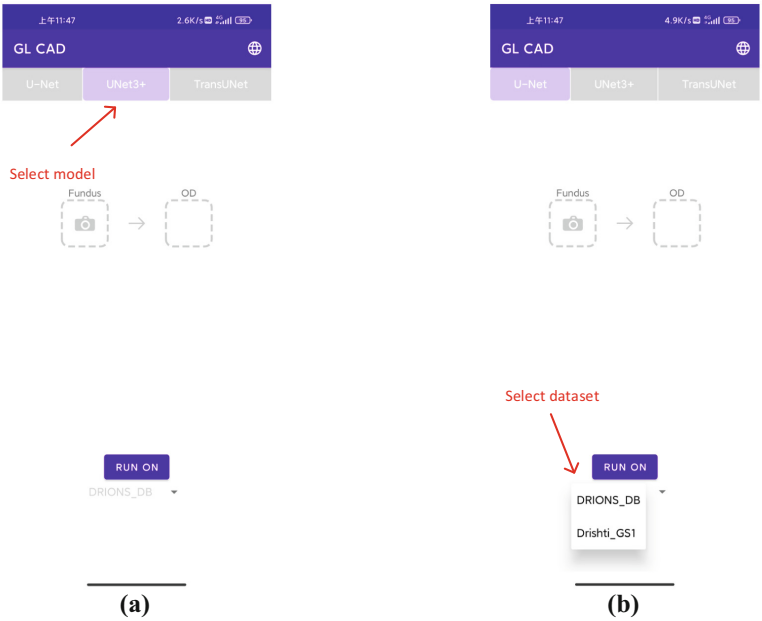

Fig. 8. Dataset and model selection of CAD system. (a) Select model(b) Select data

While the Drishti-GS dataset includes labels for both the optic disc and the optic cup, the DRIONS-DB collection only includes color fundus pictures and labels for the optic disc. So, for each of the three segmentation tasks listed above, U-Net and U-Net3+ trained six models that are stored in the.pth format. Then, the.pth files can be converted into.pt files that can be called by Java using the torch.jit.trace() method.

Figure 9 illustrates the glaucoma identification method on the Drishti-GS dataset and the optic disc segmentation process on the DRIONS-DB dataset. The DRIONS-DB dataset must first be chosen from the drop-down list at the bottom, followed by the ROI (resolution of 192 × 192) of the color fundus image that has to be segmented, and then the run button to produce the optic disc mask of the color fundus picture. The same techniques should be followed for glaucoma identification on the Drishti-GS dataset: The segmentation results of the optic cup and optic disc are obtained simultaneously by selecting the Drishti-GS dataset from the drop-down list in the first step, choosing the segmentation model in the second step, choosing the ROI (resolution of 640 × 640) of the color fundus image to be segmented in the third step, and clicking run. Calculate the cup-disc ratio's (CDR) horizontal diameter in the fifth step. Finally, contrast the estimated CDR with the glaucoma detection threshold. Glaucoma is identified in the color fundus imaging if the CDR is higher than 0.58.

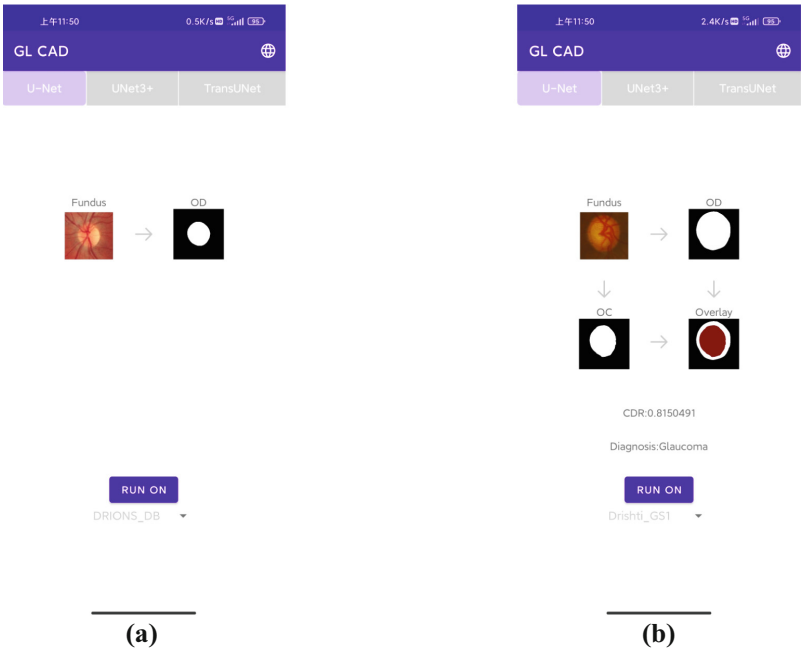

Fig. 9. Optic disc segmentation process. (a) Inferencing process for optic disc segmentation on DRIONS-DB dataset (b) Glaucoma detection process on Drishti-GS dataset

5 Conclusions

In our work, we propose a two-stage glaucoma-assisted diagnosis model deployed on mobile devices. Traditional U-Net and CHT are used to extract the ROI of OD in the first stage. The background noise in the training images is reduced using this process. In the second stage, to obtain CDR for glaucoma diagnosis, the U-Net and U-Net3 + are used for fine segmentation based on the ROI. Both datasets were used in the experiments, and both generated positive results. To assist physicians in making faster glaucoma diagnoses at an early stage, we deployed the model on a mobile device and designed an aesthetically pleasing and easy-to-use interface.

In the future, we'll seek to improve the segmentation process's resilience and dependability under increasingly difficult circumstances. And via further model optimization, we'll achieve better diagnostic results.

Acknowledgment. This work is supported in part by grants from the National Natural Science Foundation of China (No. 62262030 and No.62062040), the Outstanding Youth Project of Jiangxi Natural Science Foundation (No. 20212ACB212003), the Jiangxi Province Key Subject Academic and Technical Leader Funding Project (No. 20212BCJ23017), and Jiangxi Provincial Natural Science Foundation (No.20232BAB202021).

References

1. Tham, Y.C., Li, X., Wong, T.Y., et al.: Global prevalence of glaucoma and projections of glaucoma burden through 2040: a systematic review and meta-analysis. Ophthalmology **121**(11), 2081–2090 (2014). https://doi.org/10.1016/j.ophtha.2014.05.013
2. Michelson, G., Hornegger, J., Wärntges, S., et al.: The papilla as screening parameter for early diagnosis of glaucoma. Deutsches Aerzteblatt International **105**(34–35), 583–588 (2008). https://doi.org/10.3238/arztebl.2008.0583
3. Jayadeva, B.V., Jayadeva, K.S.: Automated detection and segmentation of optic cup and disc in retinal fundus images. J. Med. Syst. **38**(10), 121 (2014). https://doi.org/10.1007/s10916-014-0121-5
4. Kabir, M.A., Uddin, M.G.: Automated optic disc and cup segmentation in retinal fundus images using mathematical morphology. J. Med. Syst. **39**(11), 173 (2015). https://doi.org/10.1007/s10916-015-0325-5
5. Bhuyan, M.K., Jana, P.K.: Optic cup and disc segmentation in retinal images using watershed algorithm. Int. J. Image Graph. Sig. Process. **8**(8), 24–32 (2016). https://doi.org/10.5815/ijigsp.2016.08.04
6. Agrawal, S.K., Phatak, A.V.: Optic cup and disc segmentation in retinal images using polar transformation. J. Med. Imaging Health Inform. **7**(2), 403–411 (2017). https://doi.org/10.1166/jmihi.2017.2075
7. Chen, X., Wang, Y., Chen, Y., Peng, Y.: Optic disc and cup segmentation in retinal fundus images using convolutional neural networks. PLoS ONE **11**(4), e0153061 (2016). https://doi.org/10.1371/journal.pone.0153061
8. Yao, Y., Wang, X., Liu, X.: Fully Convolutional Networks for Optic Disc and Cup Segmentation in Retinal Fundus Images. arXiv preprint arXiv:1703.06519 (2017)
9. Wei, L., Yu, X., Hu, Y., Chen, S.: Automated optic disc and cup segmentation in optical coherence tomography images using deep learning. J. Med. Syst. **42**(6), 218 (2018). https://doi.org/10.1007/s10916-018-0942-1
10. Gharbi, M., Dugelay, J.L., Kaminiarz, M.: Optic disc and cup segmentation in retinal fundus images using deep convolutional neural network. In: Proceedings of the 2016 IEEE International Conference on Imaging Systems and Techniques (IST), pp. 102–107. IEEE (2016)
11. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W., Frangi, A. (eds.) Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. MICCAI 2015. Lecture Notes in Computer Science, vol. 9351. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28
12. Zhou, Z., Rahman Siddiquee, M.M., Tajbakhsh, N., Liang, J.: UNet++: a nested u-net architecture for medical image segmentation. In: Stoyanov, D., et al. Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. DLMIA ML-CDS 2018 2018. Lecture Notes in Computer Science, vol. 11045. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-00889-5_1
13. Jha, D., Sahay, R.R., Sivaswamy, J.: U-Net3+: A Full-Scale Connected U-Net for Medical Image Segmentation. arXiv preprint arXiv:2004.08790 (2020)
14. Thakur, N., Juneja, M.: Survey on segmentation and classification approaches of optic cup and optic disc for diagnosis of glaucoma. Biomed. Signal Process. Control **42**, 162–189 (2018). https://doi.org/10.1016/j.bspc.2018.01.014
15. Jimenez, S., Gonzalez, F.A., Gelbukh, A.: Mathematical properties of soft cardinality: enhancing Jaccard, Dice, and cosine similarity measures with element-wise distance. Inf. Sci. **367–368**, 373–389 (2016). https://doi.org/10.1016/j.ins

16. Carmona, E.J., Rincón, M., García-Feijoó, J., et al.: Identification of the optic nerve head with genetic algorithms. Artif. Intell. Med. **43**(3), 243–259 (2008)
17. Sivaswamy, J., Krishnadas, S.R., Joshi, J.D., et al.: Drishti-GS: retinal image dataset for optic nerve head (onh) segmentation. In: 2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI), pp. 53–56. IEEE (2014)