IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 28 June 2024, accepted 23 July 2024, date of publication 29 July 2024, date of current version 7 August 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3435552

RESEARCH ARTICLE

Combined Optic Disc and Optic Cup Segmentation Network Based on Adversarial Learning

YONG LIU[1,2], JIN WU[1], (Member, IEEE), YUANPEI ZHU[3], AND XUEZHI ZHOU[2]
[1]School of Information Science and Engineering, Wuhan University of Science and Technology, Wuhan 430081, China
[2]School of Medical Engineering, Xinxiang Medical University, Xinxiang 453003, China
[3]School of Physics and Electronic Engineering, Xinxiang University, Xinxiang 453003, China

Corresponding author: Yong Liu (141001@xxmu.edu.cn)

This work was supported in part by Henan Province Key Research and Development and Promotion Projects under Grant 232102310009, and in part by the National Natural Science Foundation of China under Grant 82302298.

ABSTRACT Glaucoma is a group of diseases characterized by progressive optic nerve damage, ultimately resulting in irreversible visual impairment. Early diagnosis through color fundus photography, including measurement of the vertical cup-to-disk ratio (CDR), can help prevent vision loss. The normal range of CDR values is usually 0.3-0.5, and if it exceeds 0.6, then there may be some problems. However, asymmetrical thinning at the edges of the bottom-superior temporal-nasal region and large gaps in datasets pose challenges for existing automatic segmentation methods. To address these challenges, this paper proposes a joint segmentation method for the optic disc (OD) and optic cup (OC) based on an adversarial network, incorporating new monitoring functions to guide the network optimization process. The effectiveness and stability of this framework were evaluated using two public performance datasets of retinal fundus images, namely Drishti-GS and REFUGE. On the Drishti-GS dataset, our method achieved a score of 0.850/0.964/0.086, while on the REFUGE dataset, it obtained a score of 0.887/0.975/0.061. These results indicate the effectiveness of our approach.

INDEX TERMS Adversarial learning, deep learning, color fundus photography, glaucoma.

I. INTRODUCTION

Glaucoma is the most common blinding disease, and can cause optic nerve damage, visual field damage and irreversible vision loss, bringing great inconvenience to patients' live and work [1]. Owing to limited medical resources, analysis of the optic nerve head (ONH) can help prevent the occurrence and development of glaucoma. This can lead to differences in the diagnosis of glaucoma among different doctors. In addition, manual diagnosis methods are time-consuming, expensive, and not suitable for large-scale sample screening. With the increasing number of patients, there is an urgent need for an automatic segmentation method that can help clinicians improve their work efficiency.

The morphology of the optic disc (OD), optic cup (OC) and cup-to-disk ratio (CDR) are the main indices used to evaluate ONH, and the CDR value could indicate the possibility of glaucoma. The CDR value is defined as the ratio of the diameter of the optic disc to the diameter of the optic cup in the vertical direction, therefore, its accuracy depends on the accuracy of OD and OC segmentation.

In figure.1, we can know the general position of the nerve papilla in the fundus image. The yellow circle indicates the optic disc area and the blue circle indicates the optic cup area. VCD represents the diameter of the vertical optic cup area, and VDD represents the diameter of the optic disc in the disposal direction. This ratio represents the vertical cup-to-plate ratio. The difference between the cup-to-plate ratio of glaucoma patients and normal conditions can be clearly observed in (c) and (d). Under normal circumstances, the value of the cup-to-plate ratio is less than 0.6, and the difference between the left and right eyes is less than 0.2.

At present, many methods have been applied to OD and OC automatic task segmentation. The most commonly used methods are edge detection, region segmentation, and threshold methods (including color and contrast).

The associate editor coordinating the review of this manuscript and approving it for publication was Mohammad Zia Ur Rahman.

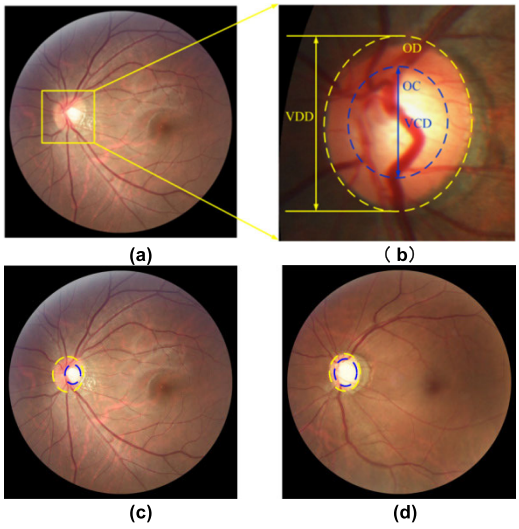

FIGURE 1. Schematic diagram of the position of the optic nerve papilla in the color fundus image. (a) is the optic nerve head topography image, (b) is enlarged structure of the optic papilla, (c) is a color fundus image without glaucoma and (d) is a color image of a patient with glaucoma.

For example, Aquino et al. [2] proposed a method using morphology and edge detection techniques to obtain circular OD boundaries by using circular transformations. If we can capture the pixel change law of the image boundary and near the boundary simultaneously, this will be very helpful for the segmentation result. Lu [3] proposed a method based on circular transformation. Chakravarty and Sivaswamy [4] extracted the boundaries of optic disc and cup based on the edge detection method of conditional random field. There are also some methods that use manual labeling of visual features to help us carry out segmentation tasks, such as stereo pair features [5], combines advanced feature extraction and attention mechanisms with spatial processing [6] and based on multi-channel and spatial attention mechanisms [7]. However, with the increase of model parameters, higher requirements will be put forward for training time and hardware. Although these methods have made progress in the automatic segmentation the optic cup from the optic disc, they are easily affected by training datasets, pathological changes, blood vessels near the optic disk and other factors, and the boundary between the OD and OC is fuzzy.

With the development of science and technology, especially computers and neural networks [8], deep learning technology has made remarkable achievements in the medical field and has achieved good results in the joint segmentation task of OD and OC [9], [10], [11]. Deep learning networks no longer rely on manually marked features, and after several rounds of iterative calculation, they can obtain the feature information of the training image, and can apply them to subsequent segmentation tasks, and the segmentation effect is better than that of traditional methods. Wang et al. [12] proposed a method to capture OD and OC boundaries with estimated ellipses on the basis of deep learning, which obtained better results under less supervision. Maninis et al. [13] proposed a DRIU network to segment OD and related blood vessels, which is an improved full convolutional network [14] with the advantages of VGG16 [15]. The complexity of the calculation can be reduced using certain sampling techniques. Zilly et al. [16] adopted entropy-based techniques to achieve this goal. Ding et al. [17] proposed a high-order attentional mechanism to obtain global context information and applied this method in medical image segmentation tasks. In these deep learning-based approaches, each pixel of images must be classified. For fundus color images, pixels can be divided into the background, OD, and OC. This approach often requires manual labeling to monitor network training. In supervised learning based on manual labeling, the quality of the segmentation results is affected by the manually labeled data. Most training schemes assume that all the datasets have the same distribution. However, there are significant differences between each dataset, including data distribution and picture quality, which results in the poor generalization ability of network models trained with a single datasets [18].

Several different approaches have been applied to segmentation models to overcome the generalization problem between different datasets [20], [21], [22]. For example, in order to transmit weak class tag information, Hong et al. [19] used an attention network based on encoder-decoder mechanism. Additional methods were also introduced in the OD and OC joint segmentation tasks. Spatial adversarial networks have demonstrated a positive effect in recent semantic segmentation. Wang et al. [18] proposed a patch-based model, which changes the discriminator in adversarial network into a patch discriminator. They also introduced morphological perception segmentation loss in the OD and OC joint segmentation tasks. This model has shown good segmentation effect on several datasets, including Drishti-GS [24], RIM-ONE-r3 [25], and REFUGE [26]. Some researchers have found that the GAN network is unstable during the training process. Kadambi et al. [23] used WGAN [27] to carry out the training process and refined the details in the training process. Although these methods have made significant progress in the existing datasets, they cannot provide reliable support for the screening and diagnosis of glaucoma because of the different data distributions in the application process. This prompted us to explore a network that could provide reliable performance on different datasets and satisfy the joint OD and OC segmentation task. The inclusion of OC in the OD region is a physiological structural feature that is often ignored, and this relationship can be beneficial to the segmentation results. The above methods can be summarized in Table 1.

In this work, a convolution network model based on adversarial learning is proposed for simultaneous OD and OC joint segmentation of fundus images. It is worth noting that OC are always present within OD regions, and their boundaries are difficult to define. The proposed method involves using an extraction network to identify the region of interest, which primarily includes the region of the firm order and the region

of the apparent cup. Subsequently, the extracted image is transformed using polar coordinates to increase the proportion of optic disc and optic cup regions in the entire image, thereby enhancing the accuracy of subsequent segmentation. The segmentation process employs adversarial learning networks, which are adjusted in terms of network structure and loss function to improve overall performance. To evaluate the effectiveness of the proposed method, experiments were conducted on two public datasets (Drishti-GS and REFUGE). The results demonstrate that the model achieves the desired outcomes.

TABLE 1. Advantages and disadvantages of the above methods.

Methods	Advantages	Limitations	Techniques
Traditional Methods	Simple algorithms Low hardware requirements	Poor accuracy	Edge detection Region segmentation Threshold methods
Convolutional neural network Methods	Less human intervention	Labeled datasets Higher hardware requirements	FCN VGG Spatial attention
GAN Methods	High accuracy Good generalization ability	Higher hardware requirements	Spatial adversarial networks Patch GAN

This work's primary contributions are as follows:

(1) We explored supervised adversarial networks to improve the adaptability of different segmentation networks to different datasets and to improve the generalization ability of networks.

(2) A method of polar coordinate transformation is proposed to transform the image so that the accuracy of the final segmentation task can be improved.

(3) We evaluated our model on two public fundus image datasets, and achieved good results in OD and OC joint segmentation tasks.

The remainder of this paper is organized as follows. We review related techniques in Section II, and Section III introduces the proposed method in detail. The evaluation and results are presented in Section IV. Finally, we discuss the results and draw conclusions in Section V and Section VI.

II. RELATED WORKS

Nowadays, many researchers are engaged in research on task segmentation of OD and OC, and many methods are effective. These methods rely heavily on visual features of artificial markers for segmentation, such as image gradient information, features of stereoscopic image pairs, local texture features and superpixel based classifiers. The boundary between OC and OD is often difficult to distinguish, so OC segmentation is more difficult, and it relies more on manually annotated features. In recent years, it has been found that the joint segmentation of OD and OC can improve the performance of segmentation networks [28].

The optic disc, also known as the optic nerve papilla, is located at the posterior pole of the eyeball, approximately 3 millimeters nasally and 1.5 millimeters in diameter. The optic disc nerve plays an essential role in the fundus examination. Early methods of extracting optic disc boundaries relied on the use of templates. Lowell et al. [29] utilized image gradient changes to segment optic disc boundaries and incorporated the active contour method. Since both optic discs and optic cups have ellipsoid shapes, methods based on circular transformation technology have also been used [2], [3]. To enhance the robustness of the model, Fu et al [9] incorporated local texture features in a multidimensional space. Pixel classification methods have achieved notable results in current semantic segmentation tasks, particularly in the field of medical image segmentation. By converting the boundary segmentation task into a pixel classification problem, researchers have found it to be more conducive to solving the problem. Cheng et al. [30] used a superpixel classifier to segment the optic disc and optic cup, and manually produced visual features to improve detection accuracy. Abramoff et al. introduced parallax values extracted from stereo image pairs to identify the optic disc and background [5]. Although these methods have shown good results, they all rely on manual annotation information, and are therefore more susceptible to image quality and pathological changes.

OC Segmentation: An important indicator for diagnosis glaucoma is the optic cup, situated in the center of the optic disc within a brighter oval depression. The anterior movement of the optic cup obstructs the optic disc, leading to glaucoma. Under normal circumstances, the optic cup is less than 1/3 the size of the optic disc, but the proportion of the optic cup is larger patients with glaucoma. Wong et al. proposed a level-set algorithm to automatically segment the boundary of OC [31]. Later, the information about blood vessel curvature in retinal images has been shown to be beneficial for the segmentation of OC [32]. Due to the natural distortion of fundus blood vessels near the optic disc (OC) boundary, the accuracy of OC segmentation based on the information of blood vessel distortion information is not satisfactory. In addition, Cheng et al. introduced the method of pixel classifier method into the OC segmentation task [30]. More and more useful methods are being introduced to OC segmentation tasks [33], [34]. All of the aforementioned methods depend on manually labeled visual features, primarily focusing on the contrast information between the edge of the optic nerve and the optic cup.

Joint OD and OC Segmentation: Optic disc and optic cup are closely related in physiological structure, and the optic cup is contained in the optic disc, which means that the pixels belonging to the optic cup also belong to the optic disc. The joint segmentation of the optic disc can obtain high accuracy in calculating the value of CDR [35]. Joshi et al. divided OD and OC step by step [36]. Zheng et al. integrated the prior graph cut framework into OD and OC segmentation [37]. The above methods are based on the fact that any pixel in the fundus image only belongs to one part, such as background,

OD or OC. That is to say, they believe that OD and OC are independent, which is in contradiction with the actual physiological structure.

Recently, advancements have been made in domain adaptive technology, particularly in the field of medical image analysis. This technique is based on adversarial networks and explores the shared feature space between the source and target domains, enabling feature correspondence between the two domains. The network is trained using these ideas, and then applied to the target domain. These methods can generate realistic images in another domain without using paired training sets. One such method is Cycle-GAN, but it requires additional constraints to guide the unsupervised style change process. For instance, Sevastopolsky [38] used two segmentation networks stacked behind the cycle-GAN to achieve enhanced shape consistency. In [16], adversarial learning of semantic perception was used to prevent semantic distortion during conversion. In [39], a generative adversarial network was constructed to enhance segmentation consistency. However, these methods fail to consider the spatial correlations between the target domain and its neighborhood. Therefore, we propose a network that combines the joint segmentation task of OC and OD to improve the accuracy of the results.

III. METHOD

Figure. 2 and figure. 3 show the overall structure diagram of OD and OC joint segmentation proposed by us. It includes two parts: polar coordinate transformation for cropped fundu images and adversarial learning network. Firstly, we use the ROI extraction network to extract the OD region in the color fundus image [40]. As the proportion of the cut image of OD part is small, the network segmentation performance is not good after they are used as the training set. We will carry out polar coordinate transformation on the cut image, so as to improve the proportion of OD region in the whole image.

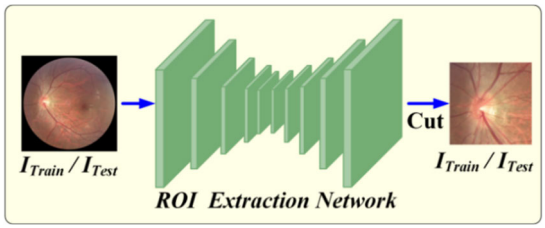

FIGURE 2. Overview of our ROI Extraction framework. ROI regions (I_{Train}; I_{Test}) are firstly extracted from the training section (I_{Train}) and the test section (I_{Test}).

The network can get more useful features from them. Finally, the transformed images are used as the inputs of adversarial learning network to obtain the segmentation result.

A. ROI EXTRACTION

Increasing the proportion of optic disc and optic cup in the whole image can help improve the accuracy of segmentation. For this purpose, we use the ROI extraction framework to manipulate the original image and crop it in the center of the optic disc. Specifically, the extraction network is a kind of U-Net network. We use the cut image block and the corresponding OD label for training, so that we can cut out the sub-image we are interested in with the OD center. In this case, the size of our clipped image is 480 × 480.

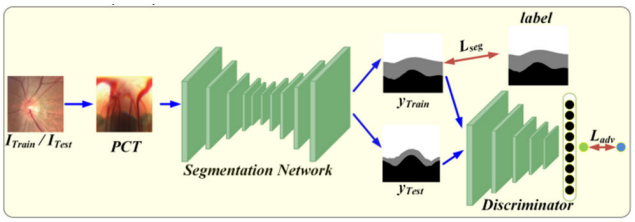

FIGURE 3. Segmentation process based on adversarial learning Network. *PCT* stands for the image after polar coordinate transformation, both the training part and the test part require this step. The discriminator can reduce the difference between training set and test set prediction. The entire network is monitored by two functions called L_{seg} and L_{adv}. L_{seg} is calculated by using the predicted value and the training labeled datasets (y_{Train}). And L_{adv} is calculated using the prediction of the unlabeled test datasets (y_{Test}).

B. SEGMENTATION NETWORK ARCHITECTURE

We used an improved network to get better performance for OD and OC joint split tasks.

Segmentor: We used MobileNetV2 [14] to replace xcexception [13] in the original DeepLabv3+ architecture [13], which reduces the overall network computation and improves network performance. See figure 4. We used an original convolutional network and seven inverted residual blocks of the MobileNetV2 in the down-sampling process. The stride of the first four convolution layers is set as the initial value, and the stride of the remaining three layers are set as 1. During the down-sampling process, a total of eight down-sampling operations are performed. In order to collect as many feature maps as possible, we use ASPP (Atrous Spatial Pyramid Pooling) [13]. Finally, the probability graph of OD and OC is generated according to the multi-label Settings in [1].

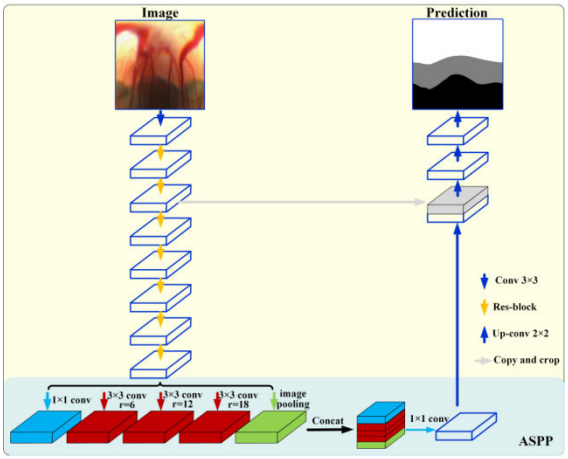

FIGURE 4. The proposed segmentation network architecture. It includes a down-sampling part, an up-sampling part and a skip connection part. And it includes the ASPP module.

Joint Morphology Loss: We used new functions to guide the learning of the network. A dice coefficient loss LDL and a smoothness loss LSL are make up this compound loss function. We can express it as:

$$L_{seg} = \lambda_1 L_{Dice}(p^d, y^d) + \lambda_2 L_{Dice}(p^c, y^c) + \lambda_3 \left[L_{Smooth}(p^d, y^d) + L_{Smooth}(p^c, y^c) \right] \quad (1)$$

where p^d, y^d, represents OD prediction probability graph and binary true mask after polar coordinate transformation; p^c and y^c represent the prediction probability graph of OC and the binary ground true mask after polar transformation; by adjusting λ, the weight of each loss function can be changed.

The dice coefficient loss [41] measures the overlap between the prediction and ground truth, and is written as

$$L_{Dice}(p, y) = 1 - \frac{2\sum_{i\in\Omega} p_i \cdot y_i}{\sum_{i\in\Omega} p^2 + \sum_{i\in\Omega} y^2} \quad (2)$$

where, Ω represents all pixels in the image after polar coordinate transformation, p represents the predicted probability graph, and y represents the ground truth mask after polar coordinate transformation. The smoothness loss is written as:

$$L_{Smooth}(p, y) = \sum_{i\in\Omega} \sum_{j\in N_4(i)} B_{i,j} \times y_i \times |p_i - p_j| \quad (3)$$

$$B_{i,j} = \begin{cases} 1 & if\ y_i = y_j \\ 0 & otherwise \end{cases} \quad (4)$$

The smoothness of the counter image can be improved by reducing the variation between adjacent pixels. Where $N_4(i)$ denote the four-connected neighbors of pixel i, p denote the prediction and y denote ground truth.

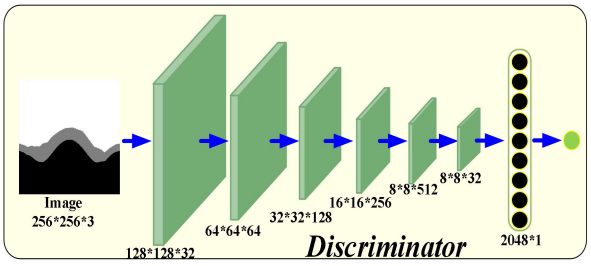

FIGURE 5. Network architecture of the discriminator.

Discriminator: In this work, we use discriminator [38], [42] to identify the results of generator, and obtain better segmentation results through the process of adversarial-learning. This discriminator can pay more attention to local features. As shown in figure 5, we implement the discriminator through a full convolutional network. The discriminator consists of a six-layer convolutional network, and the number of channels in each layer is 32; 64; 128; 256; 512; 32, respectively. with a kernel size of 4 × 4 and a stride of 2. The final layer is activated with the sigmoid function and the remaining four layers are LeakyRelu. Each patch is classified into real (1) or fake (0) through the discriminator.

Objective Function: The segmentation network is constantly generating images in an attempt to fool the discriminator, whose goal is to find the generated fake data. By constantly updating the parameters of the segmentation network and discriminator, better segmentation results can be obtained. We use formula 5 as the objective function of the discriminator:

$$L_D = -\sum_{m,n} zlog(D(S(I_{cS}))) + (1 - z)log(1 - D(S(I_{cT}))) \quad (5)$$

where $z = 1$ if the patch prediction is from the training set, and $z = 0$ if the patch prediction is from the testing set. In the process of segmentation, we use the supervision function to ensure the reliability of the segmentation results, which includes two parts and:

$$L_S = L_{seg}(I_{cS}) + L_{adv}(I_{cT}) \quad (6)$$

$$L_{adv}(I_{cT}) = -\sum_{m,n} log(D(S(I_{cT}))) \quad (7)$$

We used two public datasets, their training parts contain the original images and the corresponding manual annotation images, in the optimization process of the network, we can use the joint morphological perception to segment the loss L_{seg}.

C. POLAR TRANSFORMATION FOR FUNDUS IMAGE

In order to improve the performance of our network, we use the method of polar coordinate transformation. The image after transformation will make OD and OC have a certain constraint relationship in space, which accords with the actual situation of physiological structure. If we set the center of the disk as $O(u_0, v_0)$, then any pixel $P(u, v)$ in the original color fundus image corresponds to pixel $P'(\theta, r)$ in the transformed image. Figure. 6 (C), if r is used to represent the radius to the origin p, θ is used to represent the azimuth. We can make polar coordinates correspond to Cartesian coordinates:

$$\begin{cases} u = r\cos\theta \\ v = r\sin\theta \end{cases} \Leftrightarrow \begin{cases} r = \sqrt{u^2 + v^2} \\ \theta = \tan^{-1} v/u \end{cases} \quad (8)$$

Select the radius R of the circular region in the original image to correspond to the height of the transformed image, and 2PI to correspond to the width of the transformed image. After polar coordinate transformation, it can bring outstanding improvement to the network, which is mainly manifested in:

1)Geometric constraint: In physiological structure, OD and OC have a certain structural relationship. They tend to present an elliptical structure, and OC is included in OD. As shown in (b). After polar coordinate transformation, OC, OD and background three parts show a regular layer structure, as shown in (d). Such a relationship is conducive to image segmentation [43] and [44].

2) Data enhancement: We can do data enhancement in different ways. If we move the center $O(u_0, v_0)$ of the circle, it's equivalent to the polar transformation of the different regions. When we change the radius R, we're going to transform it with a different coefficient.

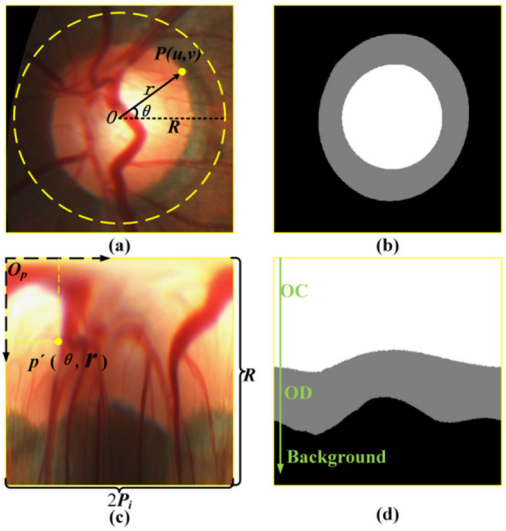

FIGURE 6. A schematic diagram of polar coordinate transformation, where the image is transformed from (a) to (c). (a) to (c) represents the polar coordinate transformation of the organization image, where the transformation area selects the circular area with radius R in the original image, as indicated by the circle in the figure. (b) to (c) marks the transformation of GT, where the white area represents the optic cup, the gray area represents the optic disc, and the black area represents the background.

3) Increase OC ratio: Due to the low proportion of OC region in the original color fundus image, the network error will be large in the training process, which is easy to produce overfitting. Even if we use the ROI extraction technology, this result will be greatly improved. As shown in (b) and (d), if the method of polar coordinate transformation is used, the proportion of OC can be increased from about 4% to about 23%. This is conducive to the performance improvement and reduction of overfitting in the process of model training.

The Python tool contains library functions commonly used in neural network learning today, so we use it to build our network model. During the training phase, we introduced Stochastic Gradient Descent (SGD) to guide the training of the network. We will start with a learning rate of 0.0001 and gradually decrease, with momentum of 0.9. In order to make the transformed image size 480 × 480 pixels, we set the transform radius R to 480 and plot the direction Angle into 480 different bins. The output image is the probability graph of OD and OC. We set the threshold to 0.5 and then get the binary mask from the probability plot. We used the same scheme in the [10] and [36] to generate the final segmentation result.

Dataset: We verified our algorithm on three public datasets: DRISHTI-GS[1] dataset and REFUGE[2] dataset. The DRISHTI-GS dataset is a publicly available dataset dedicated to the study and evaluation of the optic nerve head (ONH) segmentation problem. The enrolled patients ranged in age from 40 to 80 years old, with roughly equal numbers of men and women. The dataset consists of 101 high-quality retinal images, divided into 50 training images and 51 test images, all derived from clinical data at the Madurai Aravan Eye Hospital. Centered on OD, the field of view is 30 degrees, the size is 2896 × 1944 pixels, and the format is an uncompressed PNG image. Each image was labeled by four glaucoma specialists with three, five, nine and 20 years of experience to capture differences in labeling between observers. The REFUGE dataset is a public data set released in MICCAI 2018 for the Retinal fundus Glaucoma Challenge to advance the development of automated assessment technologies for retinal diseases such as glaucoma. The dataset contained 1200 fundus images, of which 120 were of glaucoma patients and the rest were of non-glaucoma patients. All images are stored in JPEG format and each color channel is 8 bits. There are two image resolutions: 2124 × 2056 pixels and 1634 × 1634 pixels. All images were accompanied by detailed disc and cup labeling information and were divided into three subsets for easy experimental evaluation. Each subset contains 400 fundus images. The proportion of glaucoma and non-glaucoma cases also remained consistent, i.e., each subset contained 10% of glaucoma cases and 90% of non-glaucoma cases.

Implementation Detail: Training strategy: We used the same method in [16] to optimize the segmentation network and discriminator. The image I_{cS} in the training set and the image I_{cT} in the test set are passed to the network for each iteration. Then the parameters of the network are optimized according to the functions L_S and L_D. This process is repeated over and over again.

We can use a flowchart to summarize the whole process of network hyperparameter updating:

Algorithm 1 Stochastic Gradient Descent Training of Segmentation Network Based on Adversarial Learning. The Number of Steps to Apply to the Discriminator, k, Is a Hyperparameter. We Used $k = 10$ in Our Experiments

for number of training iterations **do**

for k steps **do**

- Samples from Segmentation network $\left\{I_{cS}^1, I_{cS}^2, \cdots, I_{cS}^N\right\}$
- Samples from GT $\left\{I_{cT}^1, I_{cT}^2, \cdots, I_{cT}^N\right\}$
- Update the discriminator by:

$$\nabla_d \frac{1}{N} \sum_{i=1}^{N} \left[-\sum_{m,n} z log(D(S(I_{cS}^i))) + (1-z) log(1 - D(S(I_{cT}^i))) \right]$$

end for

- Samples from Segmentation network $\left\{I_{cS}^1, I_{cS}^2, \cdots, I_{cS}^N\right\}$
- Samples from GT $\left\{I_{cT}^1, I_{cT}^2, \cdots, I_{cT}^N\right\}$
- Update Segmentation network:

$$\nabla_s \frac{1}{N} \sum_{i=1}^{N} \left[L_{seg}(I_{cS}^i) + L_{adv}(I_{cT}^i) \right]$$

end for

The gradient-based updates can use any standard gradient-based learning rule. We used momentum in our experiments.

[1] http://cvit.iiit.ac.in/projects/mip/drishti- gs/mip- dataset2/Home.php
[2] https://refuge.grand-challenge.org

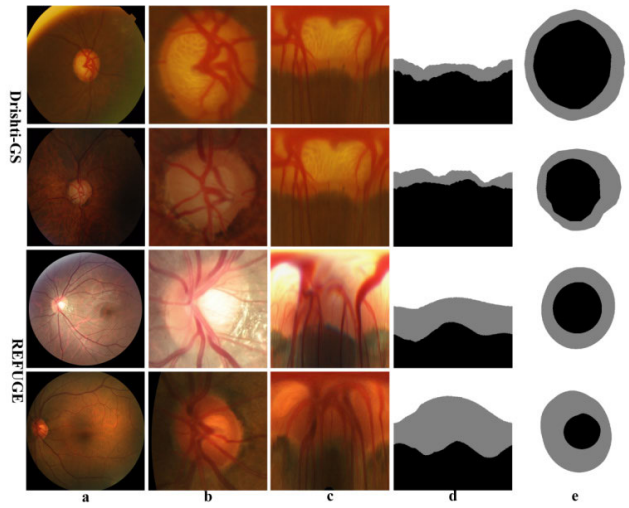

FIGURE 7. The visual examples of polar coordinate transformation on Drishti-GS and REFUGE datasets. The top three lines are from the Drishti-GS dataset, and the bottom three lines are from REFUGE datasets, Where a) fundus image, b) ROI region c) the polar coordinate transformation corresponding to the ROI region, d) the polar transformation corresponding to the label, d) the label for the ROI region.

The method is verified on an Intel i5-9400, 32G RAM, Nvidia 2080Ti GPU and Windows10 platform computer, which is implemented in Python language and the back-end is Tensorflow. First, we crop the original image to 480×480 size with the optic disc as the center, and pass this region to the network as the ROI region for the next operation. According to the previous experience, too small datasets tend to cause overfitting of the network. Simple data enhancement, such as translation, rotation and other operations, can not improve the network overfitting problem.

To address this issue, polar coordinate transformation is utilized to enhance the performance of the network. The process is depicted in figure 7. During the training phase, the Adam optimizer iteratively optimizes the network's parameters. The initial learning rate is set at 1e-3, which is adjusted in subsequent iterations. The discriminator network is also continuously optimized during training, with a simulation learning rate set at 2e-5 and 1e-5, respectively. The learning rate is reduced by 10 times every 4 learning periods, for a total of 400 learning periods. After obtaining the predicted mask, morphological operation is performed as a post-processing step to refine the segmentation mask.

Evaluation Metrics: We use the following indicators to evaluate our segmentation results, Dice coefficient (Dice), Jaccard (Jac) and vertical cup-to-disc ratio (CDR). The criteria are defined as:

$$Dice = \frac{2 \times N_{tp}}{2 \times N_{tp} + N_{fp} + N_{fn}} \tag{9}$$

$$Jaccard = \frac{N_{tp}}{N_{tp} + N_{fp} + N_{fn}} \tag{10}$$

$$\delta = \left| CDR_p - CDR_g \right|, \quad CDR = \frac{VD_{cup}}{VD_{disc}} \tag{11}$$

In the formula, N_{tp}, N_{fp}, and N_{fn} represent true positive, false positive and false negative respectively. VD_{cup} and VD_{disc} represent the vertical diameters of OC and OD, respectively. Their ratios represent the cup-plate ratio, and CDR_p and CDR_g represent the predicted and actual ratios. The error between the predicted value and the actual value is denoted by δ, and the smaller the value, the better prediction result.

IV. RESULTS

Experiments and Results: To verify the validity of the proposed method, we conducted experiments on two publicly available datasets, Drishti-GS and REFUGE.

According to the experimental results, our approach has demonstrated superior performance on the two publicly available datasets. As shown in Table 2, our method outperformed other methods in the OD and OC joint segmentation task. Specifically, on the Drishti-GS dataset, our method achieved a score of 0.850/0.964/0.086, while on the REFUGE dataset, it obtained a score of 0.887/0.975/0.061. These results indicate the effectiveness of our approach.

Figure 8 presents the segmentation results on the Drishti-GS and REFUGE datasets. Based on these results, we can conclude that our method's segmentation results (column C) are more accurate and closer to the gold standard compared to U-Net (column D). This indicates that our approach can provide more reliable technical support to clinicians.

In the experiment, we also calculated the CDR value to measure the role of the proposed method in glaucoma screening. The results of ORIGA and REFUGE dataset are shown in columns 4 and 8 of Table 1. From the glaucoma screening results, we have the following observations: Most FCN-based segmentation networks rely heavily on pixel-level labels, resulting in unsatisfactory results. However, GNS network has some advantages in network structure, which can collect high latitude features and finally make the results better. As shown in figure. 9, the ROC values on Drishti-GS (AUC=0.8590) and REFUGE datasets (AUC=0.8788).

Ablation Experiments: Ablation experiments were conducted on the Drishti-GS datasets. The result achieved by different components of the model is shown in Table 3. For the sake of presentation, we take the U-net framework as the baseline. When GAN network is combined with baseline, a better dice score can be obtained in optic disc segmentation, with an increase of 4.65%, but the optic cup segmentation get a lower score.

V. DISCUSSION

Glaucoma diagnosis primarily depends on retinal images, particularly the optic disc and cup region in the center of the image. However, existing methods suffer from limitations such as differences in labels, datasets, and physiological structure between the optic disc and cup. To address these issues, we propose a generative adversarial network. Our approach offers several improvements over existing methods:

TABLE 2. Results of joint OD and OC segmentation on different datasets.

Method	Drishti-GS			Method	REFUGE		
	$Dice_{cup}$	$Dice_{disc}$	δ		$Dice_{cup}$	$Dice_{disc}$	δ
Eff-s net	0.834	0.963	0.118	Unet	0.862	0.953	-
TD-GAN [45]	0.747	0.924	0.117	Robust	0.873	0.967	-
Hoffman et al. [46]	0.851	0.959	0.093	DeepLabv3+	0.859	0.972	-
Javanmardi et al. [40]	0.849	0.961	0.091	CE-Net	0.883	0.974	-
M-Net	0.802	**0.967**	-	CUHKMED[26]	0.883	0.960	-
AGNet	0.844	0.963	-	TDSNET[47]	0.885	0.949	-
CCNet	0.881	0.848	-	U^2-Net	0.886	0.961	
Ours	**0.853**	0.964	**0.086**	Ours	**0.887**	**0.975**	**0.061**

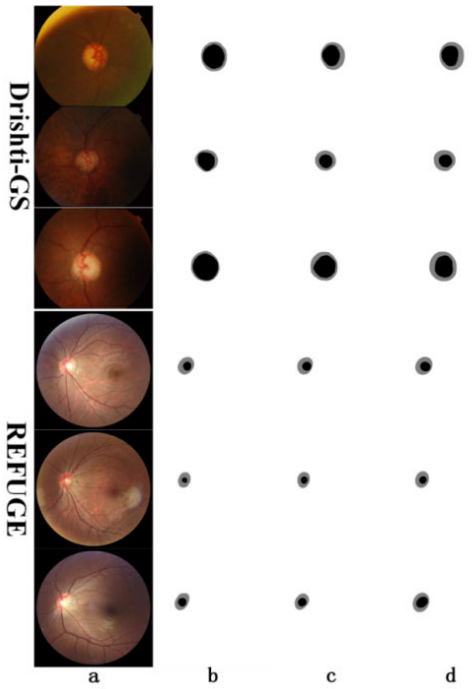

FIGURE 8. A visual example of optic disc and cup segmentation on the Drishti-GS and REFUGE datasets, where the top three rows of results are from the Drishti-GS dataset and the bottom three rows of results are from the REFUGE dataset. From left to right, each column represents: a) original fundus image, b) ground truth, c) prediction results of our method, d) U-Net prediction results.

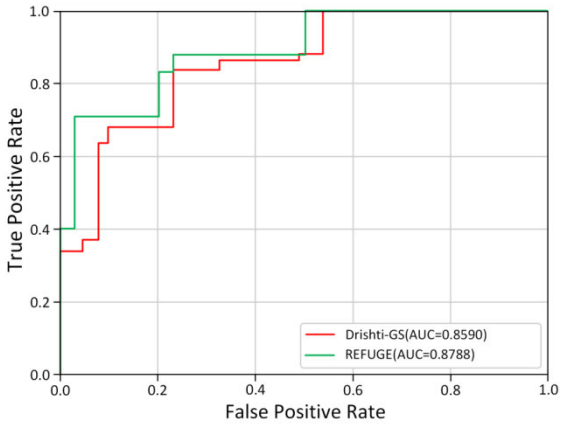

FIGURE 9. The ROC curves with AUC scores for glaucoma screening based on the segmentation results on Drishti-GS and REFUGE datasets.

TABLE 3. Effect of different blocks of the frame on the Drishti-GS datasets.

Model	$Dice_{cup}$	$Dice_{disc}$	δ
Baseline	0.847	0.943	0.097
Baseline+GAN	0.840	0.954	0.0956
PT+Baseline+GAN	**0.853**	**0.964**	**0.086**

1. Running Time: We reduced the size of the original images during training and used an NVIDIA 2080Ti GPU to compute 400 iterations, which took approximately 9 hours. Our model can generate a prediction image in approximately 0.9 seconds, which is faster than existing methods.

2. Improvements to the proposed approach: During our training, we iterated 400 times, reducing the learning rate by a factor of 10 every four learning cycles and recording the weights every 10 rounds. However, we found that this led to an underestimation of CDR values for fundus images, particularly in the glaucoma group.

3. Methodology: In contrast to existing methods, our proposed approach uses a compound feature loss to calculate the difference between the predicted image and the gold standard. This facilitates the segmentation network and discriminator to capture hierarchical features of long- and short-distance spatial relationships between pixels. Additionally, the segmentation network and discriminator share loss functions, making the training of the segmentation stable from end to end.

We know from figure 10 that when the choice of R value is different, the resolution of the corresponding converted image is also different. For example, when R=800, the resolution of the converted image is 800 by 800 pixels, and the field of view is wider than R=400. However, too large input image will affect the training speed of the network, and does not significantly improve the resolution of the network. After comparison, we chose R=480, which is very network friendly.

Although the proposed model performs well in the above two data sets, the test results need to be improved when different dataset are used, such as on the HRF dataset [48], the experimental results are shown in Table 4. There is a gap between the performance of the proposed framework and the

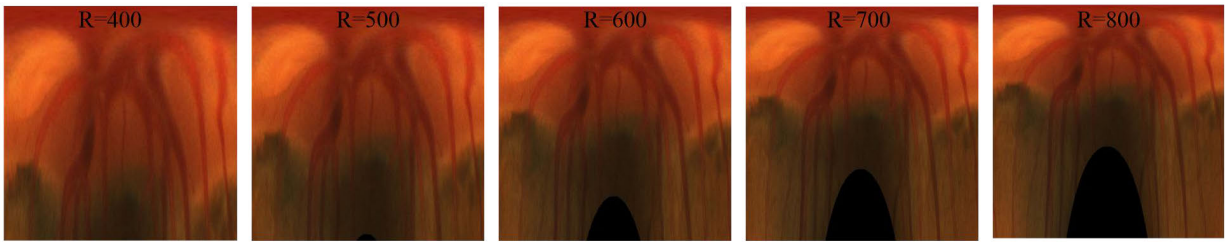

FIGURE 10. Example of polar coordinate transformation, when choosing different R-values for the same image. From left to right, R=400, 500, 600, 700, 800.

TABLE 4. Performance of the proposed model on the HRF datasets.

DATASETS	METHOD	ACC	SP	SE	PRE	IOU	F1
HRF	MULTI-CHANNEL GENERATIVE ADVERSARIAL NETWORK	0.9982	0.9986	0.9406	0.9552	0.9003	0.9420
	OURS	0.9840	0.9952	0.9184	0.9279	0.8986	0.9163

results in the references. In the next work we will explore the validation of model performance across different data sets.

VI. CONCLUSION

In this study, we utilized an enhanced generative adversarial network to segment the optic disc (OD) and optic cup (OC) in color fundus data. By transforming the joint segmentation task into a multi-label segmentation task, we successfully achieved the separation of these two structures. Our proposed framework comprises two components: a segmentation network and a discriminator. The segmentation network is designed to learn the conditional distribution between the fundus image and its corresponding label, while the discriminator is responsible for distinguishing between the source of the image-label pair. To ensure that the OD and OC information is well balanced in the image, we first employed an EOI network to isolate the region of interest, followed by a polar coordinate transformation for optimal performance. Our experimental results on two separate datasets indicate the viability of our approach. This method has the potential to aid clinicians in diagnosis and we plan to evaluate its performance on additional public datasets in future work. Our approach also holds valuable insights for other related fields.

REFERENCES

[1] Y.-C. Tham, X. Li, T. Y. Wong, H. A. Quigley, T. Aung, and C.-Y. Cheng, "Global prevalence of glaucoma and projections of glaucoma burden through 2040," *Ophthalmology*, vol. 121, no. 11, pp. 2081–2090, Nov. 2014, doi: 10.1016/j.ophtha.2014.05.013.

[2] A. Aquino, M. E. Gegúndez-Arias, and D. Marín, "Detecting the optic disc boundary in digital fundus images using morphological, edge detection, and feature extraction techniques," *IEEE Trans. Med. Imag.*, vol. 29, no. 11, pp. 1860–1869, Nov. 2010, doi: 10.1109/TMI.2010.2053042.

[3] S. Lu, "Accurate and efficient optic disc detection and segmentation by a circular transformation," *IEEE Trans. Med. Imag.*, vol. 30, no. 12, pp. 2126–2133, Dec. 2011, doi: 10.1109/TMI.2011.2164261.

[4] A. Chakravarty and J. Sivaswamy, "Joint optic disc and cup boundary extraction from monocular fundus images," *Comput. Methods Programs Biomed.*, vol. 147, pp. 51–61, Aug. 2017, doi: 10.1016/j.cmpb.2017.06.004.

[5] M. D. Abramoff, W. L. M. Alward, E. C. Greenlee, L. Shuba, C. Y. Kim, J. H. Fingert, and Y. H. Kwon, "Automated segmentation of the optic disc from stereo color photographs using physiologically plausible features," *Investigative Opthalmol. Vis. Sci.*, vol. 48, no. 4, p. 1665, Apr. 2007, doi: 10.1167/iovs.06-1081.

[6] M. Hayat, S. Aramvith, and T. Achakulvisut, "SEGSRNet for stereo-endoscopic image super-resolution and surgical instrument segmentation," 2024, *arXiv:2404.13330*.

[7] M. Hayat, S. Aramvith, and T. Achakulvisut, "Combined channel and spatial attention-based stereo endoscopic image super-resolution," in *Proc. IEEE Region 10 Conf. (TENCON)*, Chiang Mai, Thailand, Oct. 2023, pp. 920–925, doi: 10.1109/TENCON58879.2023.10322331.

[8] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*, 2012, pp. 1097–1105, doi: 10.1145/3065386.

[9] H. Fu, J. Cheng, Y. Xu, D. W. K. Wong, J. Liu, and X. Cao, "Joint optic disc and cup segmentation based on multi-label deep network and polar transformation," *IEEE Trans. Med. Imag.*, vol. 37, no. 7, pp. 1597–1605, Jul. 2018, doi: 10.1109/TMI.2018.2791488.

[10] Y. Jiang, L. Duan, J. Cheng, Z. Gu, H. Xia, H. Fu, C. Li, and J. Liu, "JointRCNN: A region-based convolutional neural network for optic disc and cup segmentation," *IEEE Trans. Biomed. Eng.*, vol. 67, no. 2, pp. 335–343, Feb. 2020, doi: 10.1109/TBME.2019.2913211.

[11] Z. Gu, J. Cheng, H. Fu, K. Zhou, H. Hao, Y. Zhao, T. Zhang, S. Gao, and J. Liu, "CE-Net: Context encoder network for 2D medical image segmentation," *IEEE Trans. Med. Imag.*, vol. 38, no. 10, pp. 2281–2292, Oct. 2019, doi: 10.1109/TMI.2019.2903562.

[12] Z. Wang, N. Dong, S. D. Rosario, M. Xu, P. Xie, and E. P. Xing, "Ellipse detection of optic disc-and-cup boundary in fundus images," in *Proc. IEEE 16th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2019, pp. 601–604.

[13] K. K. Maninis, J. Pont-Tuset, P. Arbeláez, and L. Van Gool, "Deep retinal image understanding," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Cham, Switzerland: Springer, 2016, pp. 140–148, doi: 10.1007/978-3-319-46723-8_17.

[14] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 3431–3440, doi: 10.1109/CVPR.2015.7298965.

[15] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*.

[16] J. Zilly, J. M. Buhmann, and D. Mahapatra, "Glaucoma detection using entropy sampling and ensemble learning for automatic optic cup and disc segmentation," *Computerized Med. Imag. Graph.*, vol. 55, pp. 28–41, Jan. 2017, doi: 10.1016/j.compmedimag.2016.07.012.

[17] F. Ding, G. Yang, J. Wu, D. Ding, J. Xv, G. Cheng, and X. Li, "High-order attention networks for medical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Cham, Switzerland: Springer, 2020, pp. 253–262, doi: 10.1007/978-3-030-59710-8_25.

[18] S. Wang, L. Yu, X. Yang, C.-W. Fu, and P.-A. Heng, "Patch-based output space adversarial learning for joint optic disc and cup segmentation," *IEEE Trans. Med. Imag.*, vol. 38, no. 11, pp. 2485–2495, Nov. 2019, doi: 10.1109/TMI.2019.2899910.

[19] S. Hong, J. Oh, H. Lee, and B. Han, "Learning transferrable knowledge for semantic segmentation with deep convolutional neural network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 3204–3212.

[20] J. Hoffman, D. Wang, F. Yu, and T. Darrell, "FCNs in the wild: Pixel-level adversarial and constraint-based adaptation," 2016, *arXiv:1612.02649*.

[21] Y.-H. Chen, W.-Y. Chen, Y.-T. Chen, B.-C. Tsai, Y.-C. F. Wang, and M. Sun, "No more discrimination: Cross city adaptation of road scene segmenters," 2017, *arXiv:1704.08509*.

[22] M. Wang and W. Deng, "Deep visual domain adaptation: A survey," *Neurocomputing*, vol. 312, pp. 135–153, Oct. 2018, doi: 10.1016/j.neucom.2018.05.083.

[23] S. Kadambi, Z. Wang, and E. Xing, "WGAN domain adaptation for the joint optic disc-and-cup segmentation in fundus images," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 15, no. 7, pp. 1205–1213, Jul. 2020, doi: 10.1007/s11548-020-02144-9.

[24] J. Sivaswamy, S. R. Krishnadas, G. D. Joshi, M. Jain, and A. U. S. Tabish, "Drishti-GS: Retinal image dataset for optic nerve head(ONH) segmentation," in *Proc. IEEE 11th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2014, pp. 53–56.

[25] F. Fumero, S. Alayon, J. L. Sanchez, J. Sigut, and M. Gonzalez-Hernandez, "RIM-ONE: An open retinal image database for optic nerve evaluation," in *Proc. 24th Int. Symp. Comput.-Based Med. Syst. (CBMS)*, Jun. 2011, pp. 1–6.

[26] J. I. Orlando et al., "REFUGE challenge: A unified framework for evaluating automated methods for glaucoma assessment from fundus photographs," *Med. Image Anal.*, vol. 59, Jan. 2020, Art. no. 101570, doi: 10.1016/j.media.2019.101570.

[27] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein GAN," 2017, *arXiv:1701.07875*.

[28] P. V. C. Hough, "Method and means for recognizing complex patterns," U.S. Patent 3 069 654, Dec. 3, 1962.

[29] J. Lowell, A. Hunter, D. Steel, A. Basu, R. Ryder, E. Fletcher, and L. Kennedy, "Optic nerve head segmentation," *IEEE Trans. Med. Imag.*, vol. 23, no. 2, pp. 256–264, Feb. 2004, doi: 10.1109/tmi.2003.823261.

[30] J. Cheng, J. Liu, Y. Xu, F. Yin, D. W. K. Wong, N.-M. Tan, D. Tao, C.-Y. Cheng, T. Aung, and T. Y. Wong, "Superpixel classification based optic disc and optic cup segmentation for glaucoma screening," *IEEE Trans. Med. Imag.*, vol. 32, no. 6, pp. 1019–1032, Jun. 2013, doi: 10.1109/TMI.2013.2247770.

[31] D. W. K. Wong, J. Liu, J. H. Lim, X. Jia, F. Yin, H. Li, and T. Y. Wong, "Level-set based automatic cup-to-disc ratio determination using retinal fundus images in ARGALI," in *Proc. 30th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2008, p. 10, doi: 10.1109/IEMBS.2008.4649648.

[32] D. W. K. Wong, J. Liu, J. H. Lim, H. Li, and T. Y. Wong, "Automated detection of kinks from blood vessels for optic cup segmentation in retinal images," *Proc. SPIE*, vol. 7260, no. 6, pp. 964–970, 2009, doi: 10.1117/12.810784.

[33] Y. Xu, D. Xu, S. Lin, J. Liu, J. Cheng, C. Y. Cheung, T. Aung, and T. Y. Wong, "Sliding window and regression based cup detection in digital fundus images for glaucoma diagnosis," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Berlin, Germany: Springer, 2011, pp. 1–8, doi: 10.1007/978-3-642-23626-6_1.

[34] Y. Xu, L. Duan, S. Lin, X. Chen, D. Wong, T. Wong, and J. Liu, "Optic cup segmentation for glaucoma detection using low-rank superpixel representation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, 2014, pp. 1–8, doi: 10.1007/978-3-319-10404-1_98.

[35] Y. Xu, J. Liu, S. Lin, D. Xu, C. Y. Cheung, T. Aung, and T. Y. Wong, "Efficient optic cup detection from intra-image learning with retinal structure priors," in *Proc. MICCAI*, vol. 15, 2012, pp. 58–65, doi: 10.1007/978-3-642-33415-3_8.

[36] G. D. Joshi, J. Sivaswamy, and S. R. Krishnadas, "Optic disk and cup segmentation from monocular color retinal images for glaucoma assessment," *IEEE Trans. Med. Imag.*, vol. 30, no. 6, pp. 1192–1205, Jun. 2011, doi: 10.1109/TMI.2011.2106509.

[37] Y. Zheng, D. Stambolian, J. O'Brien, and J. Gee, "Optic disc and cup segmentation from color fundus photograph using graph cut with priors," in *Proc. MICCAI*, 2013, pp. 75–82, doi: 10.1007/978-3-642-40763-5_10.

[38] A. Sevastopolsky, "Optic disc and cup segmentation methods for glaucoma detection with modification of U-Net convolutional neural network," *Pattern Recognit. Image Anal.*, vol. 27, no. 3, pp. 618–624, Jul. 2017, doi: 10.1134/s1054661817030269.

[39] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. 18th Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, vol. 9351. Cham, Switzerland: Springer, 2015, pp. 234–241, doi: 10.1007/978-3-319-24574-4_28.

[40] R. O. Duda and P. E. Hart, "Use of the Hough transformation to detect lines and curves in pictures," *Commun. ACM*, vol. 15, no. 1, pp. 11–15, Jan. 1972, doi: 10.1145/361237.361242.

[41] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 7132–7141, doi: 10.1109/CVPR.2018.00745.

[42] A. Sevastopolsky, S. Drapak, K. Kiselev, B. M. Snyder, J. D. Keenan, and A. Georgievskaya, "Stack-U-Net: Refinement network for improved optic disc and cup image segmentation," *Proc. SPIE*, vol. 10949, Mar. 2019, Art. no. 1094928, doi: 10.1117/12.2511572.

[43] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted residuals and linear bottlenecks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 4510–4520, doi: 10.1109/CVPR.2018.00474.

[44] M. Tan, B. Chen, R. Pang, V. Vasudevan, M. Sandler, A. Howard, and Q. V. Le, "MnasNet: Platform-aware neural architecture search for mobile," 2018, *arXiv:1807.11626*.

[45] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," 2015, *arXiv:1502.03167*.

[46] S. Yu, D. Xiao, S. Frost, and Y. Kanagasingam, "Robust optic disc and cup segmentation with deep learning for glaucoma detection," *Computerized Med. Imag. Graph.*, vol. 74, pp. 61–71, Jun. 2019, doi: 10.1016/j.compmedimag.2019.02.005.

[47] X. Ren, S. Ahmad, L. Zhang, L. Xiang, D. Nie, F. Yang, Q. Wang, and D. Shen, "Task decomposition and synchronization for semantic biomedical image segmentation," *IEEE Trans. Image Process.*, vol. 29, pp. 7497–7510, 2020, doi: 10.1109/TIP.2020.3003735.

[48] M. K. Kar and M. K. Nath, "Efficient segmentation of vessels and disc simultaneously using multi-channel generative adversarial network," *Social Netw. Comput. Sci.*, vol. 5, no. 3, p. 288, Feb. 2024, doi: 10.1007/s42979-024-02610-0.

YONG LIU was born in Henan, China, in 1986. He received the bachelor's degree in electrical engineering and automation from Yanshan University, in 2009, and the master's degree in circuit and system from Henan Normal University, in 2013. He is currently pursuing the Ph.D. degree in control science and engineering with Wuhan University of Science and Technology.

His research interests include artificial intelligence, medical image processing, and intelligent control.

JIN WU (Member, IEEE) was born in Anhui, China, in 1967. She received the degree in electronic information engineering from Huazhong University of Science and Technology, in 1988, the master's degree in detection technology and automatic equipment from the University of Science and Technology Beijing, in 1997, and the Ph.D. degree in pattern recognition and intelligent systems Huazhong University of Science and Technology, in 2006.

She became the Director of the Department of Electronic and Information Engineering, Wuhan University of Science and Technology, in 1998, and the Associate Dean of the School of Information, Wuhan University of Science and Technology, in 2007. Her research interests include image processing, pattern recognition and intelligent systems, signal and information processing, and multimedia communication.

Dr. Wu is a member of the Professional Committee of Image and Video Processing and Communication of Chinese Society of Image and Graphics, and a Senior Member of Chinese Institute of Electronics and Chinese Optical Society.

YUANPEI ZHU was born in Henan, China, in 1989. She received the bachelor's degree in electronic information engineering and the master's degree in circuits and systems from Henan Normal University, in 2009 and 2010, respectively. Since 2013, she has been with Xinxiang University, Henan. Her research interests include medical image processing and artificial intelligence.

XUEZHI ZHOU was born in Henan, China, in 1993. He received the B.S. degree in electronic information engineering from Henan Normal University, Xinxiang, China, in 2014, and the Ph.D. degree in information and communication engineering from Xidian University, Xi'an, China, in 2020. His research interests include radiomics, medical image analysis, and bioinformatics.