DRISHTI-GS: RETINAL IMAGE DATASET FOR OPTIC NERVE HEAD(ONH) SEGMENTATION

Jayanthi Sivaswamy[†] S.R. Krishnadas[⋆] Gopal Dutt Joshi[†] Madhulika Jain[†] Ujjwal[†] Syed Tabish A.[†]

[†] IIIT, Hyderabad,India
[⋆]Arvind Eye Hospital, Madurai,India

ABSTRACT

Optic nerve head (ONH) segmentation problem has been of interest for automated glaucoma assessment. Although various segmentation methods have been proposed in the recent past, it is difficult to evaluate and compare the performance of individual methods due to a lack of a benchmark dataset. The problem of segmentation involves segmentation of optic disk and cup region within ONH region. Available datasets do not incorporate challenges present in this problem. In this data paper, we present a comprehensive dataset of retinal images which include both normal and glaucomatous eyes and manual segmentations from multiple human experts. Both area and boundary-based evaluation measures are presented to evaluate a method on various aspects relevant to the problem of glaucoma assessment.

Index Terms— Glaucoma, Dataset, Optic Nerve Head, Cup, Optic Disk

1. INTRODUCTION

Among several retinal abnormalities, glaucoma is one of the leading causes of irreversible vision loss in the world [1]. It involves a progressive degeneration of the optic nerve fibers. Structural changes in the ONH, which consists of optic disk (OD) and cup region, is one of the standard way to assess the amount of glaucomatous damage. These changes lead to *cupping* which is an enlargement of the cup with respect to the OD and is captured using cup-to-disk diameter ratio (CDR). A high value of CDR indicates the presence of glaucoma in an eye. Figure 1 shows the marked difference between a normal and a glaucomatous eye which is reflected in the corresponding CDR values. Different imaging modalities such as optical coherence tomography, heidelberg retina tomograph, fundus imaging, etc are used to assess glaucoma. Of these, color fundus (retinal) imaging is a low-cost, non-invasive and most preferred imaging modality for mass screening of retinal diseases. A retinal image is a true color, two dimensional projection of the retina. In order to quantify the amount of cupping and to estimate the indicative parameters of glaucoma, any automated analysis method has to accurately segment OD and the cup region from retinal image. Irregular disk shape, ill-
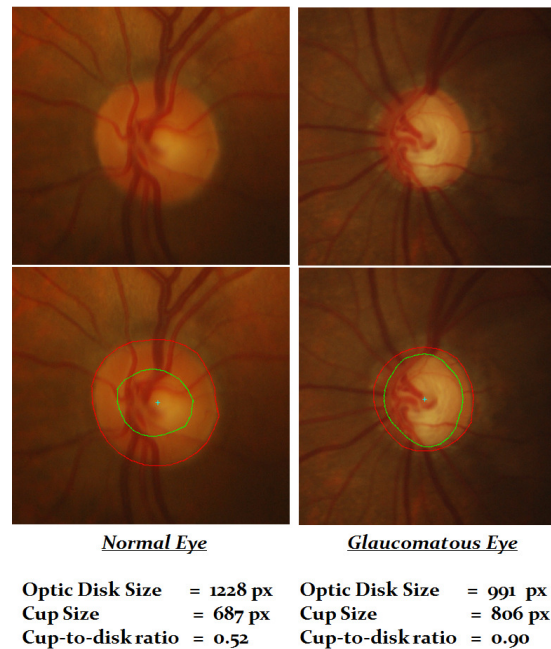

	Normal Eye		Glaucomatous Eye	
Optic Disk Size	= 1228 px	Optic Disk Size	= 991 px	
Cup Size	= 687 px	Cup Size	= 806 px	
Cup-to-disk ratio	= 0.52	Cup-to-disk ratio	= 0.90	

Fig. 1. Human expert marked boundary for optic disk (in red) and cup (in green) in dataset. First row: Cropped original retinal images; Second row: Marked boundaries; and Third row: Parameters derived from the marked boundary for normal and glaucomatous subject, respectively. The + sign indicates the location of disk center derived from the disk segmentation.

defined boundaries of OD region and variable imaging conditions make the problem of OD segmentation challenging. A number of methods have been proposed for OD and cup region segmentation [2] [3] [4] [5]. Techniques for OD segmentation are fairly mature and are capable of providing reliable OD boundaries. Recently, two standard datasets have been made public to compare methods and to establish state of art in OD segmentation [6][7]. In contrast, cup segmentation is still a challenging task as depth is the best marker (which is lost due to 2D projection in retinal image) and very few methods have been proposed. There is also high inter-observer variability seen in manual markings of the cup boundary by human experts. To the best of our knowledge, there is no ex-

isting public dataset which caters to cup segmentation problem. Hence, availability of a public dataset that provides cup and OD markings is highly desired.

In this data paper, we present a reference dataset of 50 training and 51 testing images. For each image, manual segmentations are collected for both OD and cup region from four different human experts with varying clinical experience. Markings were collected with a dedicated marking tool which provides a fully deformable circle to enable capturing different OD and cup shapes, along with any localized shape changes such as *notching*. In the next section, we summarize the existing datasets for ONH segmentation and present a detailed description of the presented dataset.

2. EXISTING ONH SEGMENTATION DATASETS

There are relatively few public datasets for glaucoma assessment as against a large number of datasets for diabetic retinopathy [8] [9] [10] and for vessel segmentation [11] [12]. Public datasets like Rim-one [7] and Drions-db [6] cater only to OD segmentation. Rim-one has 169 images (majority of them normal) with 5 manual segmentations for each one. Drions-db has 110 images with 2 manual markings for each image. OD and cup segmentation are both important parts of ONH segmentation and together they form a base for glaucoma assessment. A reference dataset to evaluate cup segmentation methods was announced in [13]. However, it is not available for free, public access.

Our dataset is released with an aim to address these issues and stimulate research efforts in relevant problems in ONH segmentation.

3. DATASET DESCRIPTION

The dataset DRISHTI-GS consists of a total of 101 images. It is divided into 50 training and 51 testing images. Ground truth is provided for the training set whose details are provided in Sub-section 3.3.

3.1. Image Acquisition

All images were collected at Aravind eye hospital, Madurai from visitors to the hospital, with their consent. Galucoma patient selection was done by clinical investigators based on clinical findings during examination. Selected patients were 40-80 years of age, with roughly equal number of males and females. Patients undergoing routine refraction test and were not glaucomatous were chosen to represent the *normal* class.

All images were taken with the eyes dilated using the following data collection protocol: centered on OD with a Field-of-View of 30-degrees and of dimension 2896×1944 pixels and PNG uncompressed image format. Apart from these, no other imaging constraints were imposed on the acquisition process. For each image, ground truth was collected from four
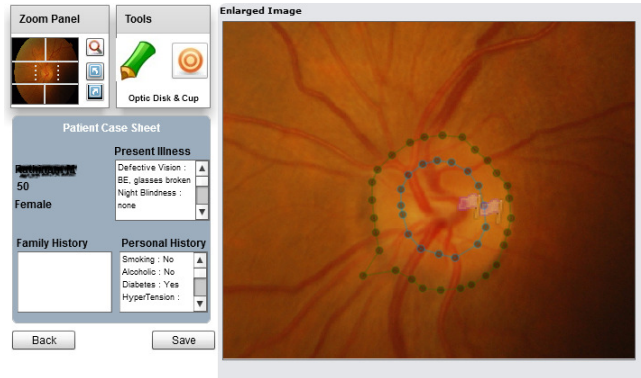

Fig. 2. Screenshot of the marking tool. Patient ID is smudged (shown in black) for anonymity.

glaucoma experts with experience of 3, 5, 9, and 20 years, respectively to capture inter-observer variance in marking. Bad quality images in terms of poor contrast, positioning of OD region, etc. were discarded. In this dataset release, fundus region (image region having retinal structures) has been extracted from the original image by eliminating the surrounding non-fundus black region, to obtain an image of approximately 2047×1760 pixels.

3.2. Ground Truth Collection

A dedicated marking tool was developed to get boundary marking on images from human experts. This tool allows precise boundary marking for even irregular shapes of OD and cup regions. This was favored over approximating the boundary with some parametric shape such as ellipse, as local deformations in the cup are possible and need to be captured in the marking. A deformable circle with several free-to-move control points was provided to the user. This helps the user to first position the circle close to the approximate region boundary and next do a better fit to the region by moving the individual control points. Thus, an accurate shape marking is possible in two steps: rough localization of the circle followed by local deformations to the circle. Figure 2 shows a screen shot of the tool wherein one point in the OD boundary is shown in pre-final position.

The ground truth is released only for the 50 training images. A web page is designed to allow interested research groups to register, download the test set and submit segmentation results (for OD and cup) on the same. The submitted results will be analyzed using evaluation measures presented in Section 4 and made publicly available on DRISHTI-GS web page (see section 6 for link).

3.3. Ground truth Format

The ground truth for each image comprises of three different pieces of information which are briefly summarized below:
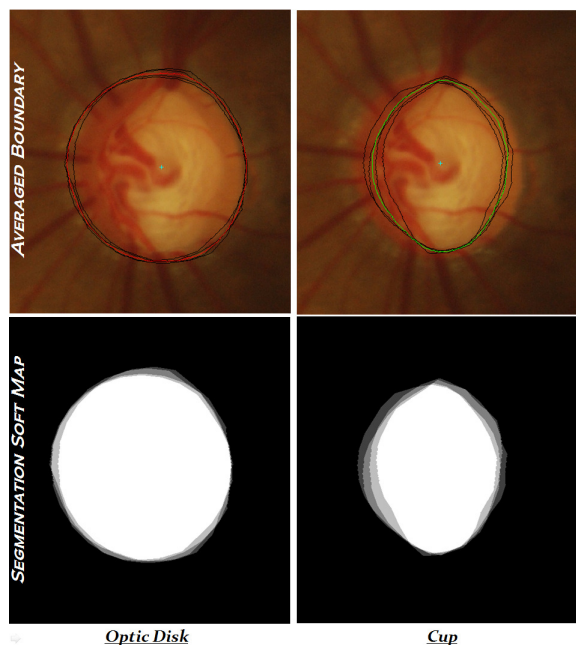
Fig. 3. Sample markings and soft segmentation maps. Top row: 4 expert markings (black) for OD (left) and cup (right) and averaged OD (red) and cup (green) boundaries. Bottom row: segmentation soft maps.

- *Region Boundary:* Average boundaries for both OD and cup region are derived from the manual marked boundaries. The image region is divided into 80 equal angular sectors with respect to the disk center and the average boundary is obtained by averaging manual markings in the respective sector. A total of Pixel locations [x,y] corresponding to the averaged boundary of OD and cup are provided in two separate text files named as *ImageName_RegionTypeAvgBoundary.txt*. Disk center location is stored in a separate file named as *ImageName_diskCenter.txt*. Top row of Figure 3 shows four manual markings in black color and obtained averaged boundary is shown in red for OD and green for cup.

- *Segmentation Soft map:* Manual segmentations obtained from four experts are fused together to generate a soft map $B(p)$ for both OD and cup region. B takes on values from 0-1 with 0 and 1 indicating no agreement and total consensus, respectively on p being a boundary point. Similarly, $B \geq 0.75$ indicates agreement among at least three experts. For each image, segmentation soft maps are provided for both OD and cup region in .PNG image format. Bottom row of Figure 3 shows sample segmentation soft maps for OD and cup region.

- *Cup-to-Disk Ratio:* CDR values corresponding to the markings of the 4 experts are provided in a separate text file named as *ImageName_cdrValues.txt*. This information allows evaluation of the CDR estimation accuracy of an algorithm against each individual expert.

4. EVALUATION MEASURES

Three different measures are proposed for the quantitative evaluation of a method. These are accuracies of a) region segmentation, b) boundary localization and c) CDR estimation.

4.1. Region Segmentation

Segmentation result is evaluated against binary ground truth and a single score is awarded to the method. The binary ground truth segmentation for an image can be obtained by thresholding the given segmentation soft map. In this paper, we present results against ground truth obtained with a value of threshold 0.75 . For a given segmentation, precision and recall values are computed as follows:

$$Precision = \frac{tp}{tp + fp} \qquad Recall = \frac{tp}{tp + fn}$$

where tp = count of true positive, fp = count of false positive and fn= count of false negative pixels. Next, the F-score (F), which is the harmonic mean of precision and recall is computed. It is defined as:

$$F = 2 \frac{Precision \cdot Recall}{Precision + Recall}$$

This is also known as the F_1 score, because recall and precision are evenly weighted. Maximum possible value of F-score is 1.

4.2. Boundary localization

The distance between computed region boundary and ground truth is a metric for boundary localization. Let C_g and C_o be the ground truth and computed boundary by a method. The distance (D, in pixels) between two curves is defined as :

$$D = \frac{1}{N} \sum_{n=0}^{N-1} \sqrt{(d_g^n)^2 - (d_o^n)^2}$$

where, d_g^n and d_o^n are the distance from disk center to points on C_g and C_o, respectively in the angular direction indexed by n. N is taken to be 24 in our evaluation.

4.3. CDR Estimation

CDR is the ratio of vertical diameters of the OD and cup regions. It ranges from 0-1 and represents the extent of glaucomatous damage in an eye. The CDR estimation error is the metric used to assess a segmentation method's success in deriving clinically relevant disk parameters.

Table 1. F-score and Boundary(B) Localization Performance

Set	F Score		B Localization (px)	
	OD	Cup	OD	Cup
Train	0.96	0.77	8.61	24.24
Test	0.96	0.79	8.93	25.48

Table 2. Error in CDR Estimation (mean/std)

	Train	Test
Expert-1	0.15/0.12	0.18/0.14
Expert-2	0.13/0.10	0.17/0.11
Expert-3	0.10/0.10	0.13/0.12
Expert-4	0.11/0.11	0.14/0.12
Avg	0.12/0.02	0.16/0.02

5. EXPERIMENTAL RESULTS

The methods for OD [5] and cup [3] segmentation were re-evaluated on DRISHTI-GS. Table 1 presents the F-score and boundary distance obtained for both OD and cup regions. Table 2 presents the CDR estimation error from the obtained segmentation results. The obtained results are consistent with the results reported in [5][3] on a larger dataset.

6. CONCLUSION AND DISCUSSION

This paper presents a dataset of retinal images for OD and cup segmentation along with ground truth. It is intended to foster active research in this direction and contribute towards developing robust and mature solutions to glaucoma assessment. The dataset is available at the url: *http://cvit.iiit.ac.in/projects/mip/drishti-gs/mip-dataset2/Home.php* Users can submit their results on the test images. The results of evaluation will be posted on the above link.

7. ACKNOWLEDGEMENT

8. REFERENCES

[1] W.H.O, "Prevention of blindness and deafness. global initiative for the elimination of avoidable blindness," *WHO document WHO/PBL/97.61 Rev2*, 2000.

[2] Y. Hatanaka, A. Noudo, C. Muramatsu, A. Sawada, T. Hara, T. Yamamoto, and H. Fujita, "Automatic measurement of vertical cup-to-disc ratio on retinal fundus images," in *Medical Biometrics*, pp. 64–72. 2010.

[3] G D Joshi, J Sivaswamy, and SR Krishnadas, "Depth discontinuity-based cup segmentation from multiview color retinal images," *IEEE TBME*, vol. 59, no. 6, pp. 1523–1531, 2012.

[4] DWK Wong, J Liu, JH Lim, X. Jia, F. Yin, H. Li, and TY Wong, "Level-set based automatic cup-to-disc ratio determination using retinal fundus images in argali," in *IEEE EMBS*, 2008, pp. 2266–2269.

[5] G D Joshi, J Sivaswamy, and SR Krishnadas, "Optic disk and cup segmentation from monocular color retinal images for glaucoma assessment," *IEEE TMI*, vol. 30, no. 6, pp. 1192–1205, 2011.

[6] E J Carmona, M. Rincón, J García-Feijoó, and J. Martínez de-la Casa, "Identification of the optic nerve head with genetic algorithms," *AIM*, vol. 43, no. 3, pp. 243–259, 2008.

[7] F. Fumero, S. Alayón, J L Sanchez, J. Sigut, and M Gonzalez-Hernandez, "Rim-one: An open retinal image database for optic nerve evaluation," in *IEEE CBMS*, 2011, pp. 1–6.

[8] Jun. 2011 [Online] MESSIDOR, "Available : http://messidor.crihan.fr/index-en.php," .

[9] T. Kauppi, V. Kalesnykiene, JK. Kamarainen, L. Lensu, I. Sorri, A. Raninen, R. Voutilainen, H. Uusitaloand H. Kälviäinen, and J. Pietilä, "The diaretdb1 diabetic retinopathy database and evaluation protocol.," in *BMVC*, 2007, pp. 1–10.

[10] L. Giancardo, F. Meriaudeau, T P Karnowski, Y. Li, S. Garg, K W Tobin, and E. Chaum, "Exudate-based diabetic macular edema detection in fundus images using publicly available datasets," *MIA*, vol. 16, no. 1, pp. 216–226, 2012.

[11] Lee B and Adam H, "Drusen detection in a retinal image using multi-level analysis," in *MICCAI*, pp. 618–625. 2003.

[12] J. Staal, M D Abràmoff, M. Niemeijer, M A Viergever, and B. van Ginneken, "Ridge-based vessel segmentation in color images of the retina," *IEEE TMI*, vol. 23, no. 4, pp. 501–509, 2004.

[13] Z. Zhang, F S Yin, J. Liu, W K Wong, N M Tan, B H Lee, J Cheng, and T Y Wong, "Origa¡ sup¿-light¡/sup¿: An online retinal fundus image database for glaucoma analysis and research," in *IEEE EMBC*, 2010, pp. 3065–3068.