Deep Learning based Retinal Image Analysis for evaluation of Glaucoma

Sharath M Shankaranarayana[1], Balamurali Murugesan[1,2], Kaushik Sarveswaran[2,3], Keerthi Ram[2], Kaushik Mitra[1], and Mohanasankar Sivaprakasam[1,2]

[1] Indian Institute of Technology Madras (IITM), India,
[2] Healthcare Technology Innovation Centre, IITM, India,
[3] Indian Institute of Information Technology, Design and Manufacturing (IIITD&M) Kancheepuram, India

Abstract. Glaucoma is one of the commonly occurring ocular diseases which could potentially lead to permanent loss in vision. Color fundus imaging is one of the most common modalities used for imaging the retina. Using the color fundus images, the assessment of glaucoma is performed by the examination of the optic nerve head and the region around it. For the computer aided diagnosis of glaucoma, we propose a deep learning based framework for three major tasks - localization of optic nerve head, image based classification for glaucoma, semantic segmentation of optic disc and cup. In our approach, we transform the fundus images with cropped region of interest using various operations. We propose a novel method for glaucoma classification in which we employ multiple pretrained deep networks and retrain the networks for glaucoma classification.During the prediction stage, we perform ensembling of multiple classifiers, each of which are trained with different networks and different kinds of input images. For the task of semantic segmentation, we employ a fully convolutional network for joint segmentation of optic disc and cup. On the validation set of the REFUGE challenge, we obtain an area under the curve (AUC) 0.9635 and reference sensitivity of 0.975, for the task of classification and dice scores for 0.8964 for optic disc and 0.7744 for optic cup and cup-to-disc ratio (CDR) error of 0.0974 for the task of semantic segmentation.

Keywords: Deep Learning, Fully Convolutional Networks, Glaucoma, Optic Disc and Cup Segmentation, Adversarial Networks

1 Introduction

Glaucoma is a sight-threatening ocular disorder which affects a large number of people worldwide. Assessment of glaucoma is predominantly performed using color fundus images by examining the optic nerve head(ONH) and the regions around it. The ONH region consists of two prominent parts- a bright circular structure called Optic Disc (OD) and an embedded region inside the OD called

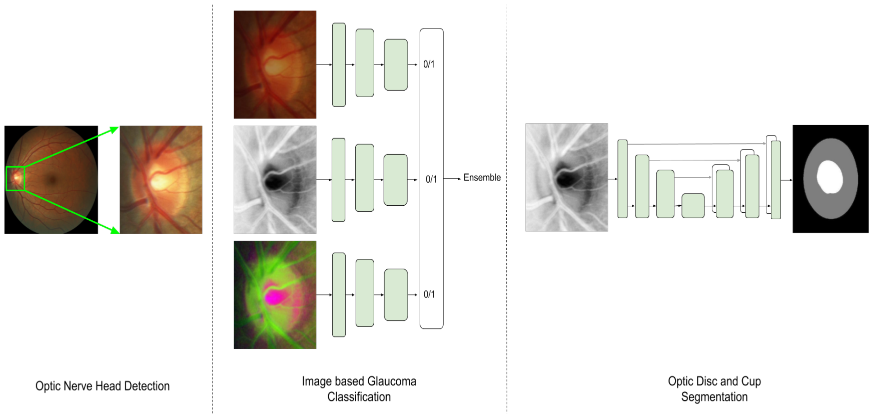

Fig. 1. Our proposed deep learning based pipeline for computer aided diagnosis of glaucoma

Optic Cup (OC). There have been several works on the assessment of ONH region for the analysis of glaucoma which usually follow the pipeline of performing segmentation of optic cup and disc and then deriving features such as cup-to-disc ratio (CDR). Recently, some works have employed deep networks for the task of glaucoma detection directly from the fundus image [1], [2]. For the task of segmentation of optic disc and cup, the work [6] provides a survey of different techniques for segmentation of optic disc and cup. Many methods based on morphological techniques [7] and deformable energy based models [8][9] and graph cuts [10] have been proposed. Recently, upon the advent of deep learning, the U-net [11] like fully convolutional architectures have been used for many kinds of semantic segmentation task. For the case of glaucoma, recently [3] proposed the use CNNs where filters are learned in a greedy fashion and the image is passed through the CNN to get pixelwise predictions for disc and cup segmentation. We recently proposed an end to end fully convolutional network for the task of joint optic disc and cup segmentation [5]. The work also explored the use of adversarial training for segmentation task.

In this work, we propose an automated pipeline for the task of analysis of retinal images for the detection of glaucoma. We propose methods using deep learning separately for the three important tasks in glaucoma analysis namely- ONH detection, image based glaucoma grading and segmentation of optic disc and cup.

2 Methods

The overall pipeline of our proposed approach is shown in Fig. 1. Given a color fundus image, optic nerve head is detected and cropped. The cropped image is classified as glaucoma or not using an ensemble of deep networks. Along with

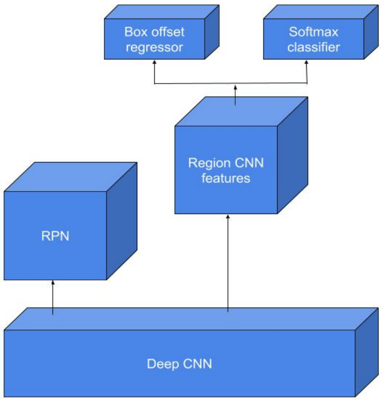

Fig. 2. Block diagram of Faster R-CNN.

this, semantic segmentation for optic disc and cup is carried out using fully convolutional networks (FCN).

2.1 Optic Nerve Head Detection

Optic nerve head detection is the primary step in glaucoma diagnosis. Localizing the optic nerve head helps in obtaining discriminative information, thus improving segmentation and classification performance. Faster R-CNN [4] is state of the art in object localization and classification. It consists of two stages: Region proposal network (RPN) and Fast R-CNN detector. An RPN takes a color fundus image as input and outputs a set of rectangular object proposals, each with an objectness score. The Fast R-CNN network then takes the region proposals and performs a region of interest (ROI) pooling operation and subsequently non-maximum suppression (NMS) to select only proposals with high confidence. A sample block diagram for Faster R-CNN can be seen in Fig. 2.

2.2 Processing of cropped image

The next important step in our pipeline is the processing of the cropped regions obtained from the ONH detection step. We perform three different kinds of processing on the original image. For the first kind of processing, we select a standard image and normalize all the cropped images with respect to the standard image for eah of the color channels. For the second kind of preprocessing, we first convert the normalized RGB image into HSV image and then scale it from $0 - 1$. We propose a novel scheme for retinal fundus images for the third kind of preprocessing, we take only the green channel of the normalized image

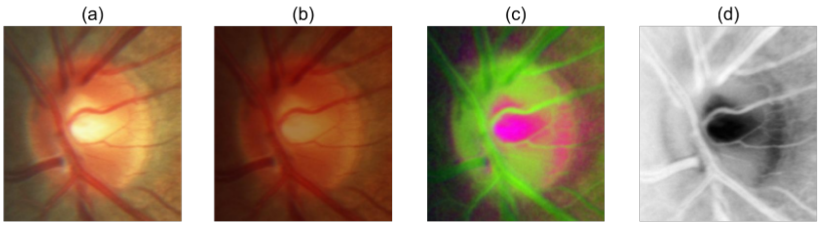

Fig. 3. Processed image crops- (a) Original Image, (b) Normalized RGB Image, (c) Normalized HSV Image, (d) Inverted Green Channel Image

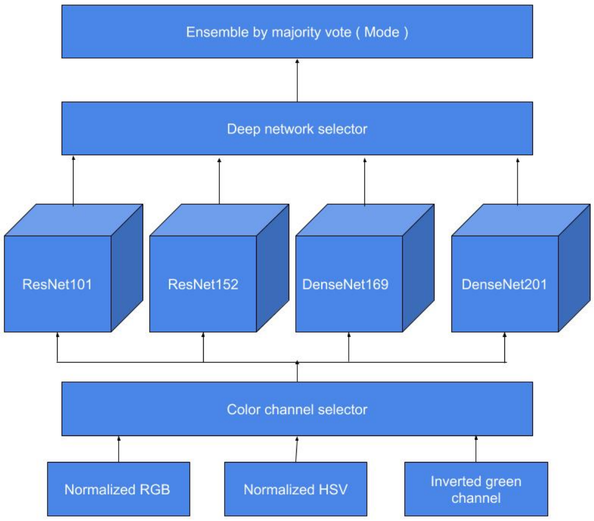

Fig. 4. Block diagram of Classifcation network

and then invert it. We see that this form of preprocessing enhances the discriminative ability of optic cup from disc. Comparison of different processing techniques along with the original image is shown in Fig. 3.

2.3 Image Based Glaucoma Classification

After preprocessing, images with three color space will be obtained and is resized to 224x224. Each color space is classified by an ensemble of state of the art ImageNet [12] classifiers like ResNet101, ResNet152, DenseNet169, and DenseNet201. Comparing to the normal convolutional networks, ResNet [13] and DenseNet [14] have improved the performance of classification tasks by introducing residual and dense connections respectively. The classifier network takes in a normalized image as input and predicts output label. The networks are initialized

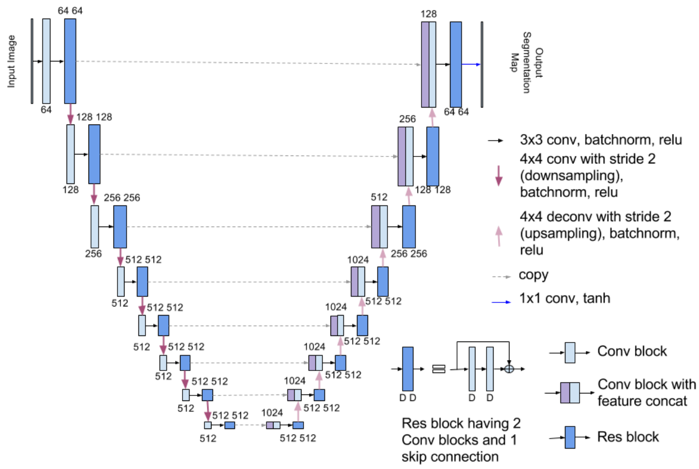

Fig. 5. Segmentation network

with ImageNet pre-trained weights. The weights of the networks are finetuned using gradients obtained from negative log-likelihood loss. The overall layout of the architecture can be seen in Fig . 4.

2.4 Semantic Segmentation of Optic Disc and Cup

For the task of joint optic disc and cup segmentation, we employ the network proposed by our group called ResUnet [5], shown in figure Fig. 5. The network is a fully convolutional network (FCN) with both long and short skip connections. In this task, we learn a mapping from the processed image containing the region of interest to the corresponding segmentation map with the segmentation map consisting of three classes- optic disc, cup and the background. Let x denote the processed image and y denote the corresponding segmentation map, our goal here is to learn mapping $G : x \to y$, where the function G is learned using a base network. The base network is a FCN which is trained to minimize multi-class cross entropy loss $L_{mce}(G)$ given by

$$L_{mce}(G) = \mathbb{E}_{x,y\sim p_{data}(x,y)}[-\sum_{i}^{N} y_i \log\left(G_i(x)\right)] \tag{1}$$

where i represents the pixel index with total of N pixels.

2.5 Datasets

ORIGA dataset contains 650 retinal fundus images with 482 normal cases and 168 confirmed glaucoma. This is used for training the detector, classifier, and segmentation network. REFUGE dataset consists of total 1200 color fundus images. The dataset is split 1:1:1 into 3 subsets equally for training, offline validation and online test, stratified to have equal glaucoma presence percentage. Training set has a total of 400 color fundus image with 40 glaucoma and 360 normal cases. The images provided have pixel-wise annotations for optic-disc and cup segmentation. Testing set consists of 800 fundus images in which 400 is used for off-site and the remaining 400 for on-site validation.

3 Experiments

3.1 Optic Nerve Head Detection

- *Train* : Faster RCNN is trained with resnet101 backbone for 25000 steps. The dimensions of the images are 360 x 572. Batch size is fixed to 1. The anchor scale and aspect ratios are $[0.25, 0.5, 1.0, 2.0]$, $[0.5, 1.0, 2.0]$ respectively. NMS iou threshold is set to 0.7.
- *Test* : The threshold of the confidence score is kept at 0.5. With the assumption, only one optic nerve head can be present in a single image, when multiple bounding boxes are predicted, the one with the highest confidence score is considered.

3.2 Image Based Glaucoma Classification

- *Train* : 4 models with 3 color spaces contribute 12 different models. Each one is trained separately with ORIGA dataset. Settings of the classifier are batch size (4), loss function (NLLLoss), final layer activation function (logsoftmax), optimizer (SGD), learning rate (0.001), momentum (0.9) and epochs (25).
- *Test* : In an ensemble of classifiers, the mode is taken across the available models. The class which is contributing to the higher mode is taken. If the class is glaucoma, then maximum confidence score of the glaucoma class group is taken. If the class is non-glaucoma, then minimum confidence score of the non-glaucoma group is taken.

3.3 Semantic Segmentation of Optic Disc and Cup

- *Train* : The model was trained from scratch with initialization from a Gaussian distribution with mean 0 and standard deviation 0.02 and run for 200 epochs. The initial learning rate was kept to be 10^{-4} and halved every 50 epochs.
- *Test* : The network produces a segmentation map for OD and OC. In order to constrain the segmentation maps as a single connected component, we employ a morphological opening and remove spurious regions. We then fit an ellipse to both the OD and OC segmentation maps.

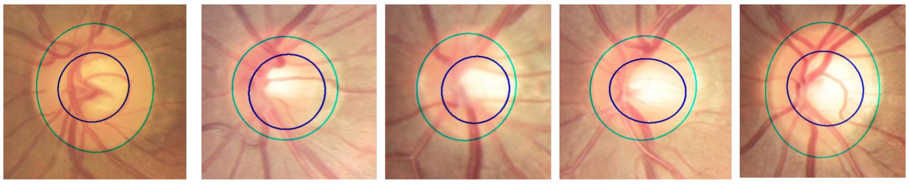

Fig. 6. Sample results on the REFUGE validation dataset. The green line indicates the dileneations for the optic disc and the dark blue line indicates dileneations for optic cup.

4 Results

For the task of ONH detection, the detector gave a precision of 1.0, recall of 1.0 and F1-score of 1.0. For the task of localized ONH classification, the ensemble of ResNet152 with the three color channels gave the AUC of 0.9635 and reference sensitivity of 0.975. For the task of semantic segmentation, the proposed model gave a dice score of 0.8964 for optic disc and 0.7744 for optic cup and cup-to-disc ratio (CDR) error of 0.0974. Sample results on the REFUGE dataset are shown in Fig. 6.

5 Conclusion

In this work, we proposed methods for the tasks of ONH detection, image based glaucoma classification and semantic segmentation of OD and OC. We also proposed the use of novel inverted green channel image for retinal image analyses, which gave superior performances on both the tasks of image classification and segmentation. We then proposed an ensemble based framework for image based glaucoma classification utilizing three kinds of processed images and also different deep networks. We obtain a high accuracy and AUC ROC for the task of glaucoma classification. We also employed FCNs for the task of OD and OC segmentation. Thus, our proposed framework serves as an automated pipeline for retinal image analysis for glaucoma.

6 Acknowledgement

We thank the organizers of the Retinal Fundus Glaucoma Challenge (REFUGE) (https : //refuge.grand challenge.org/) for hosting the challenge and kindly providing the dataset.

References

1. Chen, Xiangyu, et al. Automatic Feature Learning for Glaucoma Detection Based on Deep Learning. Medical Image Computing and Computer-Assisted Intervention-MICCAI, pp. 669-677, 2015.

2. Chen, X., Xu, Y., Wong, D.W.K., Wong, T.Y. and Liu, J., 2015, August. Glaucoma detection based on deep convolutional neural network. In Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE (pp. 715-718). IEEE.
3. J. Zilly, J. M. Buhmann, and D. Mahapatra, Glaucoma detection using entropy sampling and ensemble learning for automatic optic cup and disc segmentation, Computerized Medical Imaging and Graphics , vol. 55, pp. 2841, 2017
4. Girshick, R., 2015. Fast r-cnn. In Proceedings of the IEEE international conference on computer vision (pp. 1440-1448).
5. Shankaranarayana, Sharath M., Keerthi Ram, Kaushik Mitra, and Mohanasankar Sivaprakasam. Joint optic disc and cup segmentation using fully convolutional and adversarial networks. In Fetal, Infant and Ophthalmic Medical Image Analysis, pp. 168-176. Springer, Cham, 2017.
6. Almazroa, Ahmed, et al. Optic Disc and Optic Cup Segmentation Methodologies for Glaucoma Image Detection: A Survey, J. Ophthalmol, 2015.
7. A.Aquino, M. Gegundez-Arias, and D. Marin., Detecting the optic disc boundary in digital fundus images using morphological edge detection and feature extraction techniques, IEEE Trans. Med. Imag, vol. 20, no. 11, pp. 1860-1869, 2010.
8. J. Lowell, A. Hunter, D. Steel, A. Basu, R. Ryder, E. Fletcher et al., Optic nerve head segmentation, IEEE Trans. Med. Imag. , vol. 23, no. 2, pp. 256264, 2004
9. J. Xu, O. Chutatape, E. Sung, C. Zheng, and P. Chew, Optic disk feature extraction via modified deformable model technique for glaucoma analysis, Pattern Recognit. , vol. 40, no. 7, pp. 20632076, 2007
10. Y. Zheng, D. Stambolian, J. OBrien, and J. C. Gee, Optic disc and cup segmentation from color fundus photograph using graph cut with priors, in MICCAI 2013, 2013, pp. 7582
11. O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. In MICCAI, pages 234241. Springer, 2015
12. Deng, J., Dong, W., Socher, R., Li, L.J., Li, K. and Fei-Fei, L., 2009, June. Imagenet: A large-scale hierarchical image database. In Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on (pp. 248-255). Ieee.
13. He, K., Zhang, X., Ren, S. and Sun, J., 2016. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).
14. Huang, G., Liu, Z., Van Der Maaten, L. and Weinberger, K.Q., 2017, July. Densely Connected Convolutional Networks. In CVPR (Vol. 1, No. 2, p. 3).