Glaucoma Assessment, Optic Disc and Optic Cup Segmentation on Retinal Images using Deep Learning

Andres Diaz-Pinto(✉)[1] and Valery Naranjo[1]

Instituto de Investigación e Innovación en Bioingeniería, I3B, Universitat Politècnica de València, Camino de Vera s/n, 46022 Valencia, Spain.
andiapin@upv.es

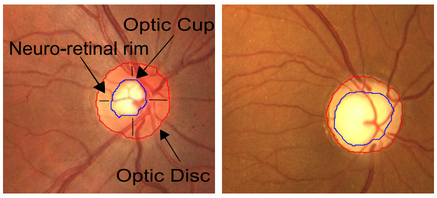
Abstract. Glaucoma is an irreversible neuro-degenerative eye disease and one of the major causes of permanent blindness around the globe. Since it does not show initial symptoms, early diagnosis is important to limit its progression. This paper presents an approach to classify glaucomatous retinal images based on fine-tuning four ImageNet-trained CNNs and using an ensemble setting. Additionally, automatic Optic Disc and Cup segmentation by using a U-Net architecture are also performed. Only publicly available retinal images were used to train, validate and test the presented methods. A Dice index of 0.91 and 0.78 were obtained for the Optic Disc and Optic Cup segmentation and an AUC of 0.94 was obtained for the classification task using the REFUGE validation data. This paper is presented for the REFUGE challenge hosted at the MICCAI 2018 conference in conjunction with OMIA workshop.

Keywords: Glaucoma Assessment, Retinal Image, Fine-tuning, REFUGE challenge

1 Introduction

Glaucoma is an irreversible eye disease and it is considered the second leading cause of blindness globally [1]. It is mainly characterised by optic nerve fibre loss and that is given by the increased intraocular pressure (IOP) and/or loss of blood flow to the optic nerve. In a fundus image, the optic nerve head or optic disc can be visually separated into two zones, a bright and central zone called optic cup and a peripheral part called neuro-retinal rim. See Fig. 1(a).

While the optic disc (OD) and cup are present in all individuals, an abnormal size of the cup with respect to the optic disc is a characteristic of a glaucomatous eye, as it is shown in Fig. 1(b). A deep understanding of the anatomy of the optic disc is crucial for glaucoma understanding. For that reason, different approaches have been developed towards optic disc analysis for Glaucoma assessment using retinal images [2,3]. For instance, in a state-of-the-art method developed by Chen et al. [4], they used cropped images to train and evaluate their system and obtaining an area under the ROC curve of 0.831 on a database of 650 images.



(a) (b)

Fig. 1. Digital fundus images cropped around the optic disc. (a) Main structures of a healthy optic disc and (b) glaucomatous optic disc.

In the paper published by Bock et al. [5], they proposed a data-driven method. This method is not based on accurate measurements of geometric optic nerve head structures such as the CDR. Instead, they used the idea of "Eigenimages" to extract features that are later classified by a Support Vector Machine (SVM). They evaluated their algorithm on 575 images randomly selected from the Erlangen Glaucoma Registry (EGR), obtaining a competitive AUC of 0.88. However, the images used in their work are private and their method cannot be compared with the presented in this paper.

Important limitations of the methods that are based on handcrafted characteristics (CDR, Area Cup/Disc ratio (ACDR), vessel kinks and ISNT rule) is the significant disagreement in estimating them even between expert human graders. For that reason, new algorithms have been focused on automatic feature extraction such as the data-driven methods [5] and convolutional neural networks (CNNs).

For instance, Chen et al. [6] proposed and trained from scratch a CNN architecture that contains six layers: four convolutional layers and two fully-connected layers; to automatically classify glaucomatous fundus images. They performed the experiments on two private databases: ORIGA-(light) which contains 650 images and SCES which contains 1676 images, achieving an AUC of 0.831 and 0.887 respectively. For ORIGA database, they trained their CNN architecture by randomly selecting 99 images, and using the remaining 551 images for test. For SCES database, they used the 650 images from ORIGA database for training, and all the 1676 images of SCES database for test. The main disadvantage is the unbalanced data. The ORIGA database is comprised of 168 glaucomatous and 482 normal fundus images and the SCES database contains 1676 fundus images of which only 46 are glaucomatous images. Another limitation of this work is that the obtained results are difficult to reproduce because the ORIGA and SCES databases are not publicly available.

A study conducted by Alghamdi et al. [7] makes use of eight databases (four public and four private databases) to detect optic disc abnormality. They developed a new approach using two CNNs: one CNN was trained to first classify the optic disc region and the other CNN to classify the optic disc region into normal, suspicious and abnormal classes. However, the four public databases (DRIVE, STARE, DIARETDB1, and MESSIDOR) used in the work of Alghamdi et al.

cannot be used for glaucoma classification. They were taken for different purposes to glaucoma classification. This means those images do not have any glaucoma sign or do not have glaucoma annotations. The glaucoma labelled databases used in their work are private and, for that reason, it is difficult to reproduce the results presented in their work.

In this paper, the results of four different ImageNet-trained CNN architectures on an ensemble setting were used for glaucoma assessment using single retinal fundus images. Additionally, we used the U-Net network architecture presented in [8] to segment the Optic Disc and Optic Cup. Contrary to most of the works in the literature, this work presents both classification and segmentation methods using images from six public databases. To the best of the author's knowledge, there are no works in the literature that address the problem of classification and segmentation with the same quantity of images as used in this work.

2 Materials and Methods

2.1 Materials

Six public databases were used in this work: the HRF database [9], composed by 15 glaucoma and 15 non-glaucoma images, the Drishti-GS1 database [10], composed by 101 glaucoma-labelled images, the third release of RIM-ONE [11] which consists of 159 stereo retinal fundus images with optic disc and optic cup annotations and labels for glaucoma and non-glaucoma images (74 glaucoma and 85 non-glaucoma). The ORIGA-light database [12] with 650 glaucoma-labelled images and the DRIONS-DB database [13] composed of 110 images with optic disc and optic cup annotations. All these databases are shown in detail in Table 1.

For all the experiments carried out in this work, the open source Deep Learning library Keras [14] and NVIDIA Titan Xp GPU were used.

Table 1. List of all the publicly available databases. GT stands for Ground Truth

Database	Glaucoma	Non-Glaucoma	OD/Cup GT	**Total**
DRIONS-DB [13]	-	-	Yes	**110**
Drishti-GS1 [10]	70	31	Yes	**101**
HRF [9]	15	15	No	**30**
ORIGA-Light [12]	168	482	No	**650**
RIM-ONE v3 [11]	74	85	Yes	**159**
REFUGE	40	360	Yes (train data)	**1200**
Total	**367**	**973**		**2250**

2.2 Glaucoma Assessment

In this section, the four ImageNet-trained CNN architectures used for glaucoma classification are described. These architectures are the state-of-the-art in image classification, object detection, and localisation. Moreover, their pre-trained versions are publicly available to use for applications different from their initial task.

2.2.1 VGG19 The VGG19 network architecture was published in 2014 by Simonyan et al. [15]. The architecture of this network is based on the same model and characterized by their simplicity. This architecture was the basis of their ImageNet 2014 submission, with which they secured the first and the second places in the localisation and classification tracks.

2.2.2 GoogLeNet This CNN was first introduced by Szegedy et al. [16]. With this architecture, the authors made the submission to the ImageNet 2014 and secured the first place in the object detection track. Based on the Inception module, they developed an architecture significantly more complex and deeper than all previous CNN architectures.

2.2.3 Microsoft ResNet Microsoft ResNet is the CNN architecture proposed by the Microsoft Research Asia team (MSRA) to the ImageNet competition in 2015. With this architecture, the MSRA not only won the competition, but also set new records in classification, detection, and localisation. Unlike traditional sequential network architectures such as AlexNet [17] and VGG, ResNet is an "exotic architecture" that relies on residual blocks.

2.2.4 Xception Xception stands for Extreme Inception and is the name of the architecture proposed in [18]. It is an extension of the Inception architecture which replaces the standard Inception modules with depthwise separable convolutions. This type of convolution is also called "separable convolution" in frameworks such as TensorFlow and Keras.

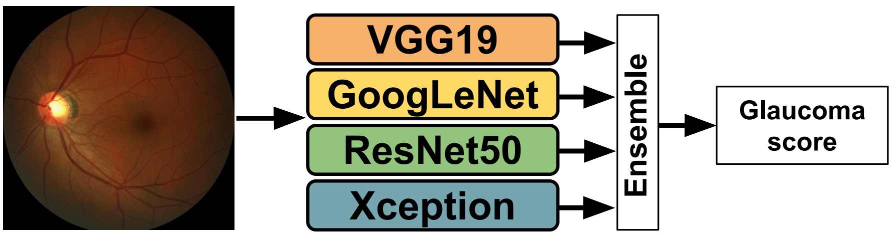


Fig. 2. Schema used for glaucoma classification

2.3 Optic Disc and Optic Cup Segmentation

The approach used in this work for Optic Disc and Optic Cup segmentation is also based on deep learning technique. Specifically, it is based on the U-Net architecture [19,8] showed in Fig. 3

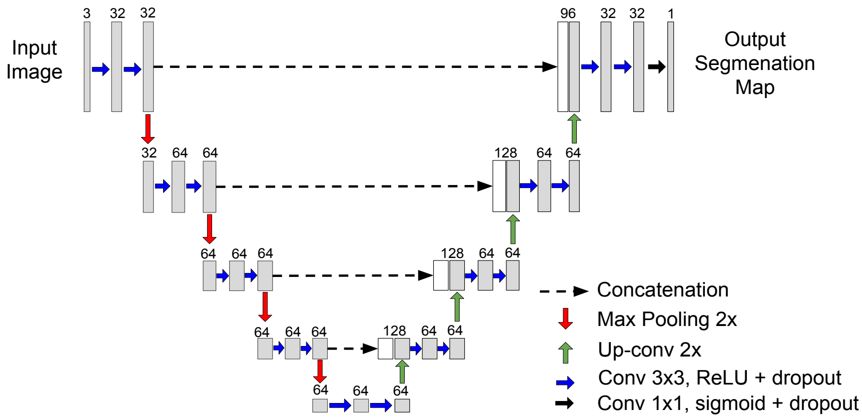


Fig. 3. U-Net architecture used for Optic disc and Optic Cup segmentation. The number on top of the box represents the number of channels.

This network was originally introduced as a Fully-convolutional neural network in which feature maps are depth-concatenated to layers that are upsampled from the bottleneck layer. In simple terms, this network receives as input an RGB image and output a probability map.

3 Results and Discussion

We fine-tuned four ImageNet-trained CNN architectures using data augmentation based on the glaucoma-labelled images (Drishti-GS1, HRF, ORIGA-light, RIM-ONE and the train set of the REFUGE database). We first re-scaled the images to 256x256px, normalize and generated smooth deformations using random factors for image rotations, image mirroring, shape deformation, vertical and horizontal flips.

Once these networks were independently fine-tuned, their results were merged together by using the average model ensemble. This technique combines the results of all these networks by using the average of the results obtained from each architecture.

It is important to highlight that before we fine-tuned the architectures presented above, we first balanced the data using the Synthetic Minority Oversampling Technique (SMOTE) [20] on the training set of the REFUGE database.

It reduces the bias on the prediction model towards the more common class (Normal).

3.1 Optic Disc and Optic Cup Segmentation

We trained a U-Net architecture for the Optic Disc and Optic Cup Segmentation task. Images from four publicly available databases (DRIONS-DB, Drishti-GS1, RIM-ONE v3 and REFUGE train set) were used to train, validate and test this system.

As a pre-processing technique, we used Contrast Limited Adaptive Histogram Equalization (CLAHE) [21,22] for both Optic Disc and Optic Cup segmentation. A pipeline of our approach for Optic Disc and Optic Cup segmentation is showed in Fig. 4.

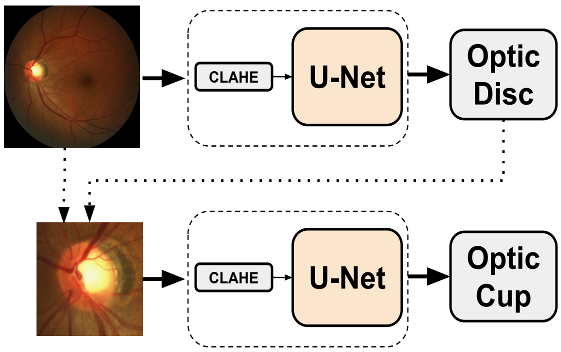


Fig. 4. Schema used for Optic disc and Optic Cup segmentation

As it is possible to see from the Fig. 4, we first segment the Optic Disc to obtain the Optic Cup mask. This means a two-stage process: we first segment the Optic Disc and used the mask to cop the image and segment the Optic Cup.

We trained the model to segment the Optic Disc by using all the databases with Optic Disc annotations (Drishti-GS, RIM-ONE v3 and the train set of the REFUGE database.) The same databases were used for training the model to segment the Optic Cup. Moreover, we do data augmentation to increase the number of images used for training. Data augmentation is essential to teach the network the desired invariance and robustness properties when only few training samples are available. We first re-scaled the images to 256x256px, normalize and generate smooth deformations using random factors for image rotations, image mirroring, and shape deformation.

Using only public databases, a Dice index of 0.91 and 0.78 were obtained for the Optic Disc and Optic Cup segmentation and an AUC of 0.94 was obtained for the classification task. DRIONS-DB, Drishti-GS1, REFUGE train dataset, and the RIM-ONE v3 were used for segmentation. HRF, ORIGA-light and the REFUGE train dataset were used for classification.

4 Conclusion

In this work, an automatic system based on fine-tuning four CNN architectures is presented. Results from all these architectures were merged together by doing model ensemble. Additionally to glaucoma assessment system, a U-Net network was trained to segment the Optic Disc and the Optic Cup. Using only public databases, a Dice index of 0.91 and 0.78 were obtained for the Optic Disc and Optic Cup segmentation and an AUC of 0.94 was obtained for the classification task. These results demonstrated the potential of these systems to be used in a clinical setting to help ophthalmologists in the glaucoma detection task.

References

1. World Health Organization: Bulletin of the World Health Organization, Volume 82, Number 11. `http://www.who.int/bulletin/volumes/82/11/en/infocus.pdf?ua=1` (2004) Accessed 2017-08-01.
2. Diaz-Pinto, A., Morales, S., Naranjo, V., Alcocer, P., Lanzagorta, A.: Glaucoma diagnosis by means of optic cup feature analysis in color fundus images. In: 24th European Signal Processing Conference (EUSIPCO). Volume 24. (2016) 2055–2059
3. Morales, S., Naranjo, V., Angulo, J., Alcañiz, M.: Automatic detection of optic disc based on PCA and mathematical morphology. IEEE Transactions on Medical Imaging. **32** (2013) 786–796
4. Chen, X., Xu, Y., Yan, S., Wong, D.W.K., Wong, T.Y., Liu, J.: Automatic Feature Learning for Glaucoma Detection Based on Deep Learning. In: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, Cham, Springer International Publishing (2015) 669–677
5. Bock, R., Meier, J., Nyúl, L.G., Hornegger, J., Michelson, G.: Glaucoma risk index: Automated glaucoma detection from color fundus images. Medical Image Analysis. **14** (2010) 471–481.
6. Chen, X., Xu, Y., Wong, D.W.K., Wong, T.Y., Liu, J.: Glaucoma detection based on deep convolutional neural network. In: 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). (Aug 2015) 715–718
7. Alghamdi, H.S., Tang, H.L., A.Waheeb, S., Peto, T.: Automatic optic disc abnormality detection in fundus images: A deep learning approach. In: OMIA3 (MICCAI 2016). (2016) 17–24
8. Sevastopolsky, A.: Optic disc and cup segmentation methods for glaucoma detection with modification of U-Net convolutional neural network. Pattern Recognition and Image Analysis **27**(3) (Jul 2017) 618–624
9. Köhler, T., Budai, A., Kraus, M.F., Odstrčilik, J., Michelson, G., Hornegger., J.: Automatic no-reference quality assessment for retinal fundus images using vessel segmentation. In: Proceedings of the 26th IEEE International Symposium on Computer-Based Medical Systems. (2013) 95–100

10. Sivaswamy, J., Krishnadas, S., Joshi, G.D., Jain, M., Ujjwal, A., S.T.: Drishti-GS: Retinal image dataset for optic nerve head (ONH) segmentation. In: 2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI). (2014) 53–56.
11. Medina-Mesa, E., Gonzalez-Hernandez, M., Sigut, J., Fumero-Batista, F., Pena-Betancor, C., Alayon, S., de la Rosa., M.G.: Estimating the amount of hemoglobin in the neuroretinal rim using color images and OCT. Current Eye Research. **41**(6) (2015) 798–805
12. Zhang, Z., Yin, F.S., Liu, J., Wong, W.K., Tan, N.M., Lee, B.H., Cheng, J., Wong, T.Y.: ORIGA-light: An online retinal fundus image database for glaucoma analysis and research. In: 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology. (Aug 2010) 3065–3068
13. Carmona, E.J., Rincón, M., García-Feijoó, J., de-la Casa, J.M.M.: Identification of the optic nerve head with genetic algorithms. Artificial Intelligence in Medicine **43**(3) (2008) 243 – 259
14. Chollet, F., et al.: Keras. `https://github.com/fchollet/keras` (2015) Accessed: 2017-05-21.
15. Simonyan, K., Zisserman, A.: Very Deep Convolutional Networks for Large-Scale Image Recognition. ArXiv e-prints **abs/1409.1556** (September 2014)
16. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A.: Going deeper with convolutions. In: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). (June 2015) 1–9
17. Krizhevsky, A., Sutskever, I., Hinton, G.E.: ImageNet Classification with Deep Convolutional Neural Networks. In: Advances in Neural Information Processing Systems 25. Curran Associates, Inc. (2012) 1097–1105
18. Chollet, F.: Xception: Deep learning with depthwise separable convolutions. ArXiv e-prints **abs/1610.02357** (2016)
19. Ronneberger, O., Fischer, P., Brox, T.: U-Net: Convolutional Networks for Biomedical Image Segmentation. In Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F., eds.: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, Cham, Springer International Publishing (2015) 234–241
20. Chawla, N.V., Bowyer, K.W., Hall, L.O., Kegelmeyer, W.P.: SMOTE: Synthetic Minority Over-sampling Technique. ArXiv e-prints (June 2011)
21. Pizer, S.M., Johnston, R.E., Ericksen, J.P., Yankaskas, B.C., Muller, K.E.: Contrast-limited adaptive histogram equalization: speed and effectiveness. In: [1990] Proceedings of the First Conference on Visualization in Biomedical Computing. (May 1990) 337–345
22. Pizer, S.M., Amburn, E.P., Austin, J.D., Cromartie, R., Geselowitz, A., Greer, T., ter Haar Romeny, B., Zimmerman, J.B., Zuiderveld, K.: Adaptive histogram equalization and its variations. Computer Vision, Graphics, and Image Processing **39**(3) (1987) 355 – 368