Visualizing and Understanding Inherent Image Features in CNN-based Glaucoma Detection

Dhaval Vaghjiani
The University of Western Australia
Perth, Australia
Email: dhaval.vaghjiani@gmail.com

Sajib Saha
Australian e-Health Research Centre
Commonwealth Scientific and Industrial Research Organisation
Perth, Australia
Email: Sajib.Saha@csiro.au

Yann Connan
École Internationale des Sciences du Traitement de L'information
Paris, France
Email: connanyann@eisti.eu

Shaun Frost
Australian e-Health Research Centre
Commonwealth Scientific and Industrial Research Organisation
Perth, Australia
Email: Shaun.Frost@csiro.au

Yogesan Kanagasingam
Australian e-Health Research Centre
Commonwealth Scientific and Industrial Research Organisation
Perth, Australia
Email: Yogi.Kanagasingam@csiro.au

***Abstract*— Convolutional neural network (CNN)-based methods have achieved state-of-the-art performance in glaucoma detection. Despite this, these methods are often criticized for offering no opportunity to understand how classification decisions are made. In this paper, we develop an innovative visualization strategy that allows the inherent image features contributing to glaucoma detection at different CNN layers to be understood. We also develop a set of interpretable notions to better comprehend the contributing image features involved in the disease detection process. Extensive experiments are conducted on publicly available glaucoma datasets. Results show that the optic cup is the most influential ocular component for glaucoma detection (overall Intersection over Union (IoU) score of 0.18), followed by the neuro-retinal rim (NR) with IoU score 0.17. With an overall IoU score of 0.16 vessels in the photograph also play a considerable role in the disease detection.**

Keywords— Deep Learning, CNN, Glaucoma, Color Fundus Image, Optic Disc

I. INTRODUCTION

Glaucoma is an eye disease that affects more than 66 million people worldwide and is the second leading cause of blindness [1, 2]. Color fundus photographs (CFPs) are widely used in clinical practice for the diagnosis of glaucoma. Manual assessment of CFPs is a time-consuming, tedious and highly subjective task [3]. Automated methods could help reduce costs and enable large-scale screening programs by providing quick and consistent predictions [4]. Early work on automated analysis of fundus images for glaucoma assessment mainly focused on segmenting the optic disc and cup using rule-based machine learning techniques [5, 6, 7, 8]. In recent years, there has been an increasing interest in applying deep learning-based techniques to the automated diagnosis of glaucoma [9, 10, 11]. Whilst CNN-based models have achieved state-of-the-art performance for glaucoma detection, they are often criticized for their "black box" nature; it is not apparent how they reach the predictions [12]. In this paper we develop a CNN visualization strategy so that the inherent image features contributing to the detection of glaucoma can be understood. We also develop a set of interpretable notions to better comprehend the visualizations.

II. METHODOLOGY

A. Training CNN model for Glaucoma Detection

A VGG16 model [11] has been trained to perform glaucoma detection using cropped fundus images. The VGG16 architecture is chosen because of its simple sequential nature and reasonably high accuracy compared to more complex CNN architectures such as InceptionV3 [13], ResNet50 [13] and Xception [14] in glaucoma detection [11]. Transfer learning [13] has been used to fine-tune the weights of the model. Stochastic gradient descent with learning rate 0.0001 and momentum 0.9 has been used as the optimizer, and binary cross-entropy has been used as the loss function during training [11]. 80% of the images are used for training and the remaining 20% are used for testing.

B. Visualizing the CNN model's features

To gain an understanding of which image features are influencing the CNN's predictions for glaucoma detection, we have adopted the visualization strategy proposed by Zeiler et al. in [15]. We attach a deconvolutional network ('deconvnet' for short) to each convnet block of the VGG16 architecture (Fig. 2). Unlike in [15], in this work the reconstructed feature maps produced by deconvnets are binarized using Otsu's method [16].

C. Interpretable Notions to Summarize Inherent Image Features

We have proposed a set of interpretable notions so that the information available in the feature maps can be summarized in an interpretable manner [17]. We have defined five interpretable notions representing five mutually exclusive ocular components: 1) vessels within the optic cup (OC), 2) vessels within the neuro-retinal rim (NR), 3) vessels outside of the optic disc (OD), 4) the optic cup (excluding vessels) and 5) the neuro-retinal rim (excluding vessels), as shown in Figure 1.

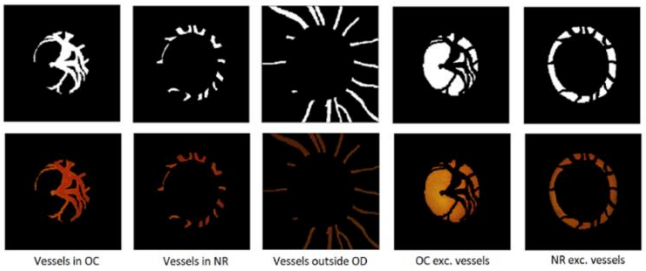

Fig. 1. Illustration of our five ocular components. Top row: binary annotation masks of each ocular components for an image in the Drishti-GS dataset. Bottom row: binary masks applied to the image.

We compute the structural similarity of the binarized deconvnet outputs with the five different ocular components using Intersection over Union (IoU) scores, as defined below.

The IoU score for image X, feature f and component c is defined as:

$$IoU_{X,f,c} = \frac{|D_f(X) \cap M_c(X)|}{|D_f(X) \cup M_c(X)|},$$

where $D_f(X)$ is the binarized deconvnet output of image X for filter f, and $M_c(X)$ is the binary annotation mask of component c for image X (in $M_c(X)$, a pixel p is equal to 1 if p belongs to component c, and 0 otherwise).

A semi-automated system was developed to generate the ocular component binary masks.

III. Experimental Analysis

A. Dataset

A total of 1817 fundus images from six publicly available datasets have been used in this study: ACRIMA [11], Drishti-GS [18], HRF [19], RIM-ONE r2 [20], sjchoi86 HRF [21] and DRIONS-DB [22]. To remain consistent and in-line with other studies in the literature, all images are manually cropped to keep only the optic disc and its surrounding region [11].

B. Experiments and Results

1) Reproducing state-of-the-art performance in glaucoma detection using VGG16

Figure 2(a) shows the training and validation accuracy of the VGG16 model. We have obtained an accuracy of 93% with sensitivity and specificity of 92% and 94% respectively, which is consistent with the findings in the literature [11].

2) Investigating the relative importance of the optic disc (OD) and surrounding region for glaucoma detection

It is generally believed that glaucoma primarily affects the optic disc and its surrounding region [11]. This experiment has been conducted to determine the relative importance of the optic disc compared to its surrounding region for glaucoma detection. We have trained and validated the VGG16 model independently under two different scenarios. In the first case, we have used the cropped images without introducing any image pre-processing. In the second case, the images are pre-processed to mask out the regions outside the optic disc. For case-I we have obtained an accuracy of 92.7% and for case-II we have obtained an accuracy of 91.2%. The area under the ROC curve (AUC) scores for case-I and case-II are 97.4% and 96.5% respectively (see Figure 2(b)).

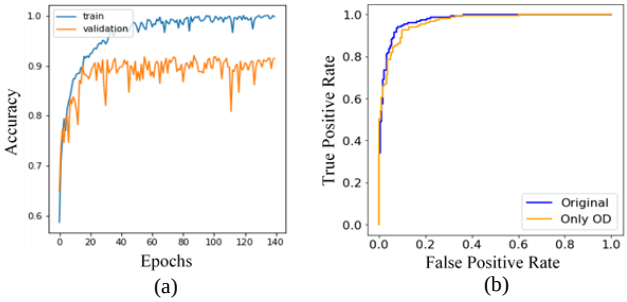

Fig. 2. (a) Training and validation accuracy of VGG16 in glaucoma detection. (b) ROC curves showing the accuracy of VGG16 for two different cases.

3) Visualizing CNN features

In this experiment we aim to visualize the inherent image features that contribute to the CNN's predictions at different layers, using the methods described in section II.B. Figure 3 shows some sample deconvnet visualizations for filters that are found to exhibit greatest variability in activation strength.

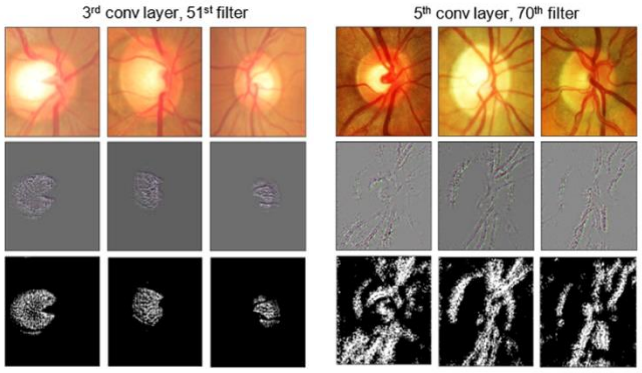

Fig. 3. Sample visualizations of inherent image features. The 1st, 2nd and 3rd rows show respectively the cropped fundus images, deconvnet outputs and binarized deconvnet outputs.

4) Interpreting the Inherent Image Features

The five mutually exclusive ocular components and the IoU score proposed in section II.C are used in this experiment to interpret the inherent image features generated by the methods of section II.B. For this experiment the images from the Drishti-GS dataset are used. Figure 4 shows the mean IoU scores in different layers. The relative importance of the components remains reasonably consistent across the different layers, with the optic cup (exc. vessels) and neuro-retinal rim (exc. vessels) showing the highest influence, followed by the vessels outside the optic disc, vessels in the optic cup and vessels in the neuro-retinal rim. These results confirm the hypothesis that most of the important information for detecting glaucoma is contained inside the optic disc.

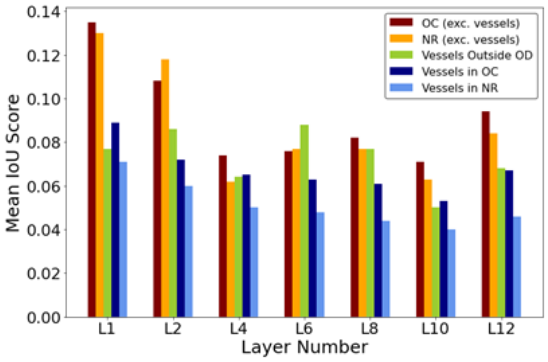

Fig. 4. Plot of mean IoU scores, aggregated by ocular components and layers. Only selected layers are shown.

IV. Discussions and Conclusion

In this study, we have investigated inherent image features in CNN-based glaucoma detection. We have adapted the visualization strategy of Zeiler et al. [15] for this specific context. Another alternative to Zeiler et al.'s method for CNN visualization could be the Grad-cam [23]. However, Zeiler et al.'s method offers greater opportunity to discover inherent image features and their progression within the CNN.

Numerous experiments on publicly available glaucoma datasets show that with an IoU score of 0.18 (including both vessels and non-vessels), optic cup features contribute the most towards glaucoma detection. With an IoU score of 0.17 (including both vessels and non-vessels) the neuro-retinal rim is the second most influential ocular component. The overall IoU score of 0.16 of the vessel components indicate that vasculature in the photograph also plays a considerable role.

Future work could involve applying our method to a larger and more diverse set of images.

REFERENCES

[1] E. Prokofyeva and E. Zrenner, "Epidemiology of major eye diseases leading to blindness in Europe: a literature review," Ophthalmic research, vol. 47, no. 4, pp. 171-188, 2012.

[2] A. Giangiacomo and A. L. Coleman, "The epidemiology of glaucoma," Glaucoma, Springer, Berlin Heidelberg, 2009.

[3] H.D. Jampel, D. Friedman, H. Quigley, S. Vitale, R. Miller, F. Knezevich, Y. Ding, "Agreement among glaucoma specialists in assessing progressive disc changes from photographs in open-angle glaucoma patients," American journal of ophthalmology, vol. 147, no. 1, pp. 39-44, 2009.

[4] S.K. Saha, B. Fernando, J. Cuadros, D. Xiao, and Y. Kanagasingam, "Automated quality assessment of colour fundus images for diabetic retinopathy screening in telemedicine," Journal of digital imaging, vol. 31, no. 6, pp. 869-78, 2018.

[5] 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," 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 2266-2269, IEEE, Vancouver, BC, Canada, 2008.

[6] G. D Joshi, J. Sivaswamy, and S. R. Krishnadas, "Optic disk and cup segmentation from monocular color retinal images for glaucoma assessment," IEEE transactions on medical imaging, vol. 30, no. 6, pp. 1192-1205, 2011.

[7] 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 transactions on Medical Imaging, vol. 32, no. 6, pp. 1019-1032, 2013.

[8] A. Diaz, S. Morales, V. Naranjo, P. Alcocer, and A. Lanzagorta, "Glaucoma diagnosis by means of optic cup feature analysis in color fundus images," 24th European Signal Processing Conference (EUSIPCO), pp. 2055-2059, IEEE, Hilton Budapest, Hungary, 2016.

[9] X. Chen, Y. Xu, D.W.K. Wong, T.Y. Wong, and J. Liu, "Glaucoma detection based on deep convolutional neural network," 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 715-718, IEEE, Milano, Italy (2015).

[10] B. Al-Bander, W. Al-Nuaimy, M.A. Al-Taee, Y. Zheng, "Automated glaucoma diagnosis using deep learning approach," 14th International Multi-Conference on Systems, Signals & Devices (SSD), pp. 207-210, IEEE, 2017.

[11] A. Diaz-Pinto, S. Morales, V. Naranjo, T. Köhler, J. M. Mossi, and A. Navea, "CNNs for automatic glaucoma assessment using fundus images: an extensive validation," Biomedical engineering online, vol. 18, no. 29, pp. 1-19, 2019.

[12] K. Wickstrøm, M. Kampffmeyer, R. Jenssen, "Uncertainty modeling and interpretability in convolutional neural networks for polyp segmentation," 2018 ieee 28th international workshop on machine learning for signal processing (mlsp), pp. 1-6. IEEE, 2018.

[13] S. Saha, M. Nassisi, M. Wang, S. Lindenberg, S. Sadda, and Z.J. Hu, "Automated detection and classification of early AMD biomarkers using deep learning," Scientific reports, vol. 9, no. 1, pp.1-9, 2019.

[14] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 1251-1258. IEEE, 2017.

[15] M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks," European conference on computer vision, pp. 818-833. Springer, Cham, 2014.

[16] N. Otsu, "A threshold selection method from gray-level histograms," IEEE transactions on systems, man, and cybernetics, vol. 9, no. 1, pp. 62-66, 1979.

[17] M.T. Ribeiro, S. Singh, C. Guestrin, "Why should i trust you? Explaining the predictions of any classifier," Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining, pp. 1135-1144. ACM, 2016.

[18] J. Sivaswamy, S. R. Krishnadas, G. D. Joshi, M. Jain, A. U. S. Tabish, "Drishti-GS: Retinal image dataset for optic nerve head (ONH) segmentation," 2014 IEEE 11th international symposium on biomedical imaging (ISBI), pp. 53-56. IEEE, 2014.

[19] A. Budai, R. Bock, A. Maier, J. Hornegger, G. Michelson, "Robust Vessel Segmentation in Fundus Images," International Journal of Biomedical Imaging, 2013, pp. 1-11, 2013.

[20] F. Fumero, S. Alayón, J. L. Sanchez, J. Sigut, and M. Gonzalez-Hernandez, "RIM-ONE: An open retinal image database for optic nerve evaluation," 2011 24th international symposium on computer-based medical systems (CBMS), pp. 1-6. IEEE, 2011.

[21] sjchoi86-HRF Database. Last accessed 1 Feb 2017.

[22] E.J. Carmona, M. Rincón, J. García-Feijoó, J.M. Martínez-de-la-Casa, "Identification of the optic nerve head with genetic algorithms," Artificial Intelligence in Medicine, vol. 43, no. 3, pp. 243-259, 2008.

[23] R. R Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, D. Batra, "Grad-cam: Visual explanations from deep networks via gradient-based localization," IEEE international conference on computer vision (ICCV), pp. 618-626, 2017.