Automatic classification of multispectral eye fundus images using deep learning

Francisco J Burgos-Fernández, Buntheng Ly, Fernando Díaz-Doutón, Meritxell Vilaseca, Jaume Pujol, Maxime Sermesant

Francisco J Burgos-Fernández, Buntheng Ly, Fernando Díaz-Doutón, Meritxell Vilaseca, Jaume Pujol, et al.. Automatic classification of multispectral eye fundus images using deep learning. Libro de Resúmenes RNO2021, Nov 2021, Online, Spain. hal-03513023

HAL Id: hal-03513023
https://hal.inria.fr/hal-03513023

XIII Spanish National Meeting on Optics, 22-24 November 2021

Automatic classification of multispectral eye fundus images using deep learning

Francisco J. Burgos-Fernández[1,2], Buntheng Ly[2], Fernando Díaz-Doutón[1], Meritxell Vilaseca[1], Jaume Pujol[1] and Maxime Sermesant[2,3]

[1] Centre for Sensors, Instruments and Systems Development, Universitat Politècnica de Catalunya (Rambla Sant Nebridi 10, 08222, Terrassa, Spain)
[2] National Institute for Research in Computer Science and Automation (Inria), Université Côte d'Azur, Epione team (2004 route des lucioles, 06902, Sophia Antipolis, France)
[3] Electrophysiology and Heart Modelling Institute (IHU, Liryc), Université de Bordeaux (Avenue du Haut Lévêque, 33604, Pessac, France)

Contact: francisco.javier.burgos@upc.edu

The evaluation of the eye fundus from retinographies could be challenging, especially at early stages where the changes are very subtle or when the ocular media is not transparent enough. This combined with the intrinsic metamerism of colour fundus cameras constrain the diagnostic capabilities of these instruments. However, spectral imaging overcomes this limitation since it offers high and extended spectral sampling, even towards the near infrared (NIR), providing information from layers behind the retina[1,2]. The subsequent image processing and analysis have traditionally applied techniques such as intensity thresholds, principal component analysis or probability distribution models[3] which are different from the mechanisms followed for physicians during clinical judgement. In order to leverage recent progress in deep learning for computer vision, artificial neural networks have been recently developed for eye fundus assessments with promising results[4,5]. Therefore, the goal of this work was to apply deep learning techniques to the classification of multispectral images from healthy and diseased eyes fundus obtained by means of a fast visible (VIS) and extended infrared multispectral fundus camera (400 nm – 1300 nm) with high spectral and spatial resolution, including the relatively unexplored range beyond 900 nm[6].

The study was conducted at the Institute of Ocular Microsurgery (Barcelona, Spain) and the University Vision Centre of the Universitat Politècnica de Catalunya (Terrassa, Spain). From all the database, 68 healthy and 68 diseased eyes from 89 subjects (63% females and 37% males) with a mean ± standard deviation in age of 54.5 ± 16.0 years have been considered. Ethical committee approval was obtained and all patients provided written informed consent before any examination. The Declaration of Helsinki tenets of 1975 (as revised in Tokyo in 2004) were followed throughout the study. The multispectral fundus camera[6] comprises 15 spectral bands divided in two detections arms, one for the VIS-NIR range (400 nm – 950 nm) and another only for the NIR range (960 nm – 1300 nm). The deep learning network was originally developed by Ly et al.[7] at Inria (Sophia-Antipolis, France) for sustained ventricular arrhythmia prediction using cardiac computed tomography images and it was recently adapted to work with multispectral eye fundus images. It consists of a conditional variational autoencoder (CVAE), which is divided in an encoder and a decoder branch, and a classifier model. In the encoder of the CVAE, the eye fundus images are combined with a mask (Fig. 1) followed by a series of regular convolution layers to generate a low dimensional space that feeds the classifier. From the low dimensional features as inputs, the decoder applies regular and transpose convolution layers to reconstruct the original eye fundus images, concatenating them again with the corresponding masks. In short, the main advantage of the autoencoder architecture is the comparison between the encoder-decoder outputs with the original data since the resulting error is backpropagated through the network at each iteration to improve its performance, while also optimising the classification performance. It enables to generate a low dimensional representation of the data that is discriminating between healthy and diseased patients. The dataset was divided in training/validation (90% data) and test (10% data) subsets.

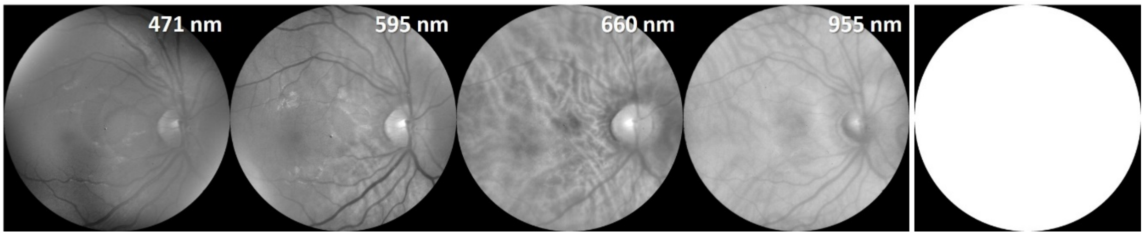

Figure 1: Four spectral eye fundus images (left) and the mask applied (right).

The CVAE was applied to three approaches: (1) using the 12 images of the VIS-NIR arm (image size = 1757 x 1757), (2) using the 15 images from the VIS-NIR and NIR arms resizing the images from the latter (image sizes = 1757 x 1757), (3) using the 15 images from the VIS-NIR and NIR arms with the original size (image sizes = 1757 x 1757 and 386 x 386, respectively). The best classification accuracies for the three approaches when applied to the test subset were (1) 92.86%, (2) 100.00%, and (3) 100.00%. The loss parameter has also to be considered since it quantifies the error of the performance of the algorithm; for the three approaches it was (1) 0.188, (2) 0.185, and (3) 0.187. From these results, it can be concluded that the combination of the 15 spectral images from the VIS-NIR and NIR led to the highest accuracies and lowest losses, even the differences with respect to the other approach were very small. Although the use of the wavelengths of the NIR arm improved the performance of the network, only considering those from the VIS-NIR also produced an excellent classification. The image resizing used in the approach (2) involved a technically less complex implementation since the CVAE remained unchanged, while for the approach (3) the encoder and decoder had to be split in two branches due to the different image size. Future work will be focused on increasing the dataset and adapting the network to detect specific eye fundus pathologies such as age-related macular degeneration, glaucoma or diabetic retinopathy.

ACKNOWLEDGEMENTS: This project has received funding from the European Union's Horizon 2020 research and innovation programme under Marie Skłodowska-Curie grant agreement No. 801342 (Tecniospring INDUSTRY) and the Government of Catalonia's Agency for Business Competitiveness (ACCIÓ).

DISCLAIMER: This work only expresses the opinion of the authors and neither the European Union nor ACCIÓ are liable for the use made of the information provided.

---

[1] C. Zimmer, D. Kahn, R. Clayton, P. Dugel, and K. B. Freund; *Retina Today*, **9**, 94 (2014).

[2] L. Everdell, I. B. Styles, A. Calcagni, J. Gibson, J. Hebden, and E. Claridge; *Review of Scientific Instruments*, **81**, 1 (2010).

[3] V. S. Mary, E. B. Rajsingh, and G. R. Naik; *IEEE Access*, **4**, 4327 (2016).

[4] W. Zhang, Y. Dai, M. Liu, Y. Chen, J. Zhong, and Z. Yi; *Applied Intelligence*, 1 (2021).

[5] X. Li, M. Jia, M. T. Islam, L. Yu, and L. Xing; *IEEE Transactions on Medical Imaging*, **39**, 4023 (2020).

[6] T. Alterini, F. Díaz-Doutón, F. J. Burgos-Fernández, L. González, C. Mateo, and M. Vilaseca; *Journal of Biomedical Optics*, **24**, 096007-1 (2019).

[7] B. Ly, S. Finsterbach, M. Nuñez-García, H. Cochet, and M. Sermesant; Proceedings of *Functional Imaging and Modeling of the Heart*, 461 (2021).