Queen's University
Belfast

Automatic Analysis of Digital Retinal Images for Glaucoma Detection

Document Version:
Peer reviewed version

Queen's University Belfast - Research Portal:
Link to publication record in Queen's University Belfast Research Portal

Automatic Analysis of Digital Retinal Images for Glaucoma Detection

Alexandre Guerre   **Jesús Martínez del Rincón**   **Paul Miller**
The Institute of Electronics, Communications and Information Technology
Queen's University Belfast, BT3 9DT
{a.guerre,j.martinez-del-rincon,p.miller}@qub.ac.uk

Augusto Azuara Blanco
Centre for Vision and Vascular Science
Queen's University Belfast, BT12 6BA
a.azuara-blanco@qub.ac.uk

Abstract

In this paper we propose a novel automated glaucoma detection framework for mass-screening that operates on inexpensive retinal cameras. The proposed methodology is based on the assumption that discriminative features for glaucoma diagnosis can be extracted from the optical nerve head structures, such as the cup-to-disc ratio or the neuro-retinal rim variation. After automatically segmenting the cup and optical disc, these features are feed into a machine learning classifier. Experiments were performed using two different datasets and from the obtained results the proposed technique provides better performance than approaches based on appearance. A main advantage of our approach is that it only requires a few training samples to provide high accuracy over several different glaucoma stages.

1   Introduction

Glaucoma is one of the most common causes of preventable blindness [1] and official population projections and epidemiological prevalence surveys predict that the number of glaucoma cases will increase by a third in the next twenty years [2]. Glaucoma, and those at risk of suffering from glaucoma, constitute a major part of the workload of secondary care eye services [3]. However, patient overload is not the only problem. Currently, referrals for suspected glaucoma are usually initiated by a community optometrist and then assessed at hospital by trained ophthalmologists. The reported diagnostic accuracy for detecting glaucoma by optometrists is suboptimal: only 20-30% of these referrals actually have glaucoma, and 45% of patients are discharged after their first visit [4]. This illustrates the inefficiency of current glaucoma detection methods and causes avoidable distress and worry to patients and carers. Interventions for optometrists, such as glaucoma training [5] or agreed guidelines [6], do not appear to affect the rates of false positive referrals. Even definitive glaucoma diagnosis, carried out by ophthalmologists, are not exempt from drawbacks: clinical optic nerve assessment is limited by subjectivity and reliance on examiner experience, while new diagnostic techniques for assessment of the structural changes at the optic nerve head (ONH) and retinal nerve fibre layer (RNFL) are expensive and therefore not widely available.

In this context, automatic detection methods are highly valuable for early glaucoma diagnosis [1], especially considering that glaucoma can be treated effectively if detected at an early stage. We propose a method based on the automatic analysis of the eye fundus, which brings together the expertise of human practitioners and the cost-effective advantages of computers. Given that digital fundus cameras are relatively inexpensive and are already widely available in optometrists' and hospital eye services, our system could potentially be deployed as a systematic screening programme for glaucoma.

Our method differs from other state-of-art systems in the usage of geometric parameters of the ONH structures that change in case of glaucoma disease: optic disc diameter, optic disc area, cup diameter, rim area, mean cup depth, etc. These features are extracted from the automatic segmentation of the structures and used for training a machine learning classifier which provides the final decision given a new fundus image. The usage of these features, traditionally employed in the manual analysis, has some competitive advantages regarding appearance based methods: they are less dependent on the camera model, they require a lower order of magnitude in the number of training images -dozens instead of
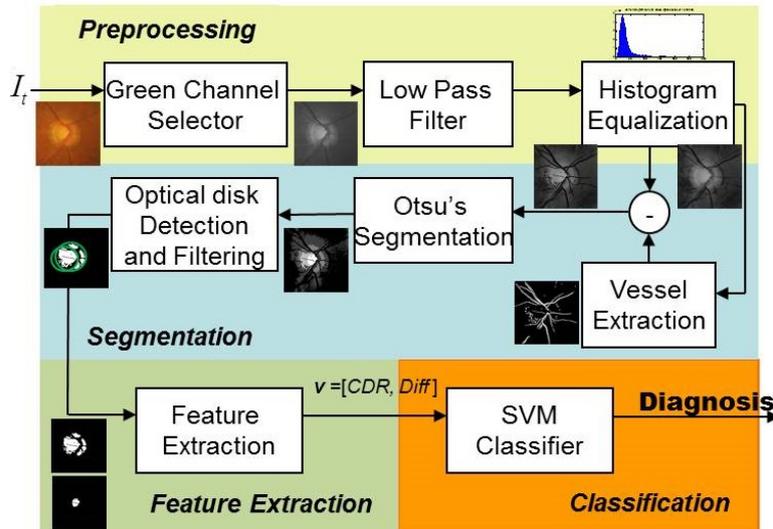
Figure 1: Block diagram of the proposed approach

hundred or thousands- and they allow detection of different levels of glaucoma even in the presence of other ocular pathologies.

1.1 State of the Art

Initial attempts to automatise the glaucoma diagnosis were based on quantitative parameters that can help to make the qualitative assessment more objective, reproducible and lead to a reduction of the observer variability. However, these methods were based on manually annotated ONH images [7, 20], or they lacked robustness and reliability. In response, researchers moved away from the extraction of geometrical features, which rely on good ONH segmentation, towards frameworks based on the pixel appearance of the whole retinal image [8, 16]. These methodologies were inspired by a pipeline previously used for face and object recognition [24] and have the advantage of not needing the segmentation of the ONH. Recently Bock et al. achieved 73% sensitivity and 85% specificity in the detection of glaucoma using a fully-automated analysis of monoscopic photographs [8] based on appearance. However, this is computationally expensive, their classification depends on the camera or machinery involved and they require hundreds of positive and negative glaucoma samples for retraining. The structure of the retina and how glaucoma affects its appearance is much more subtle than the differences between faces or objects for which these frameworks were originally designed. This makes it difficult for systems only based on appearance to detect glaucoma, specially in the early stages.

In the last few years there have been great advances, not only in medical image processing, but also in fundus cameras able to provide high resolution and low-noise retinal images. As a consequence, new studies have been performed showing high accuracy on the segmentation of the ONH characteristics such as disc area, disc diameter or the well established cup-to-disc ratio. Most successful approaches are based on ONH models able to automatically adjust to the image [14, 16, 17], although this implies additional training for model generation. In general, most of these approaches have only been validated on healthy eyes under assumptions that are not valid for glaucomatous eyes. Other approaches, tested on glaucomatous examples, have been validated against human segmentation, but they were not evaluated for diagnosis since they were neither input to a classifier, nor compared against pixel based approaches [12, 13, 15].

In this paper, we propose a computerised glaucoma diagnosis system which relies on the automatic extraction of high level geometrical features related to ONH structures. As a main advantage, it only requires a few training samples and provides high accuracy even in the presence of different stages of glaucoma. The approach is compared against a state-of-art methodology based on pixel appearance.

2 Methodology

Our image processing framework is structured in a standard 4-stage pipeline as depicted in Figure 1: (i) preprocessing, (ii) image-based segmentation, (iii) feature extraction and (iv) classification.

2.1 Preprocessing

Image normalisation is required to correct for variations caused by acquisition and illumination conditions. For this purpose, only the green channel is selected, as it has been shown as the most robust against variations [8]. After that, a low-pass filter [21] is applied to reduce the fine grain noise. Finally, histogram equalisation [21] is applied to ensure consistancy across images, Fig. 2(b).

2.2 Automatic segmentation

After the glaucoma specific preprocessing, the ONH structures needs to be segmented in order to extract the features. Our automatic segmentation methods aims to segment the disc and the cup. First, retinal vessels are located accurately using the Isotropic Undecimated Wavelet Transform (IUWT) and edge location refinement [9], Fig. 2(c). The resulting image is used as a mask to remove blood vessels and facilitate the segmentation of the different image regions, Fig. 2(d). After that, an iterative and multilevel variation of the Otsu's adaptive thresholding [22] is applied, which allows us to identify several different image regions [25]. Given the composition of retinal images, four thresholds are applied to segment the first and second brightest regions, corresponding to the cup and the rim respectively, Fig. 2(e). The morphological operator open is applied to filter noise without changing the feature size.
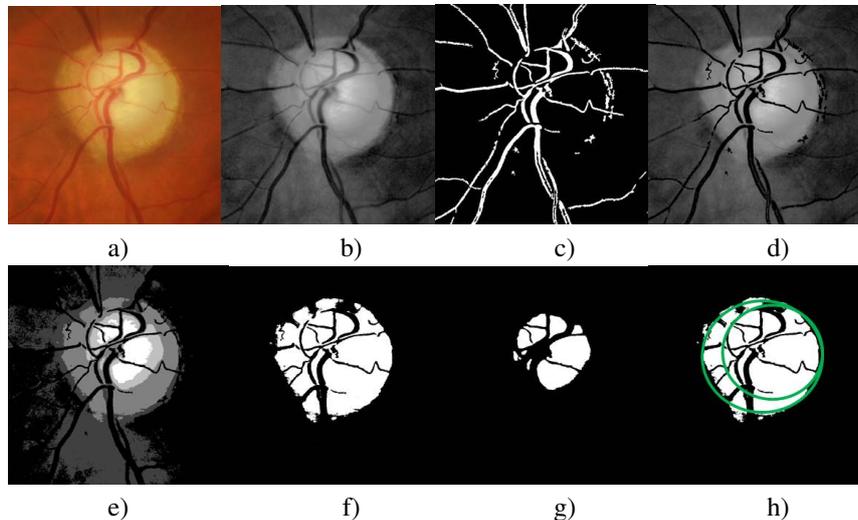

Figure 2: Segmentation process. From left to right, from top to bottom: a)original image, b) preprocessed image, c)Vessel mask, d) Vessel subtracted image, e) segmented image, f) segmented rim, g) segmented cup, h) optical discs candidates.

Although the algorithm gives an accurate segmentation, other image regions can be falsely detected as belonging to these ONH primary structures, especially in the presence of other retinal anomalies. In order to filter those false positives, an optical disc detection algorithm is employed. This detector applies a combination of the Circular Hough Transform with a scale invariant kernel operator, as described in [19], to detect circles within the retinal image. The primary goal is not to provide the geometrical parameters of the ONH, since its assumption about circularity may not be well-matched to the real shape of the ONH and may distort the feature values. Instead, the optical disc candidates are used for removing all those segmented areas outside the detected circles, thus filtering those false positives included by the region segmentation. More than one candidate is allowed, since the goal is not to uniquely identify the optical disc center, but to filter wrongly segmented pixels outside the ONH, Fig. 2(h).

2.3 Feature extraction

We hypothesise that geometric features measured from the segmentation of the disc and rim are of greater value than appearance features in detecting glaucoma. To this aim, two features are extracted and used in our framework:

Cup-to-disc ratio (CDR): The ratio of the vertical diameters of the inner cup and the outer disc rim is commonly used as an indicator of glaucoma likelihood or disease progression [10]. In our pipeline, the ratio is calculated by localising the highest and the lowest pixel in the vertical axis for both the rim and the cup segmented regions (see Fig. 3).
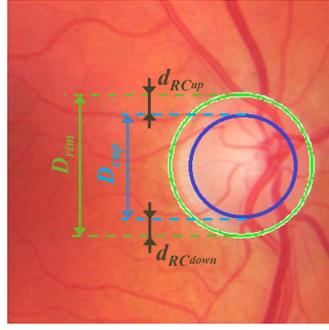
$$CDR = D_{cup}/D_{rim} \tag{1}$$

Figure 3: Feature extraction variables

Neuro-retinal rim width variation: The relative width of the neuro-retinal rim at different angular locations is known to differ between normal and glaucomatous discs. Normal subjects have a characteristic distribution, being widest at the inferior part of the disc, followed by decreasing width at the superior, nasal and temporal locations. Glaucomatous eyes typically do not follow such a pattern, which is commonly known as the "ISN'T rule" [11]. In our system, the upper and lower widths of the rim are calculated as the distance between the highest point of the rim and the highest point of the cup and the distance between the lowest point of the rim and the lowest point of the cup respectively, see Fig. 3. Then, the feature is implemented as the difference between these vertical distances.

$$Diff = d_{RC_{up}} - d_{RC_{down}} \tag{2}$$

The above features are the most common features used by ophthalmologists and therefore likely to provide discriminative information for classification.

2.4 Classification

Although in the past these features have been used directly for diagnosis by applying a set of rules, mimicking what optometrists do manually, this is subjective, depending on the experience of the expert and the demography of the population. It also assumes that features are perfectly extracted, which is not always true due to segmentation difficulties. On the contrary, by feeding the above feature set into a robust classifier, more complex rules can be automatically inferred and deviations produced by segmentation failures accounted for. Therefore, in our implementation, a SVM classifier with linear kernel is used [23]. This linear classifier determines a maximum-margin and soft hyperplane that best separates the considered classes. Data is normalised and transformed via the linear radial basis kernel.

The decision to use a linear classifier is supported by the literature [16, 8], where linear classifiers have reported excellent results for glaucoma diagnosis. The choice of SVM over more traditional approaches such as nearest neighbours, regression, neural networks and discriminant analysis is supported by their reported advantages [23]: they do not require regularity in the data so it can be applied to data following an unknown distribution, it delivers a unique solution since the optimality problem is convex contrary to neural networks, it can be easily extended to non-linear nonparametric problems by replacing the linear kernel, it scales well to high dimensional data, and the trade-off between complexity and error can be controlled explicitly.

Although more complex classification pipelines could be applied in our framework, such as the double schema proposed in [8] or non linear kernels [16], it is not the scope of our paper to state the best classification technique but to prove the validity of geometrical features. Therefore, SVM is the classifier providing the best framework for comparison with other state-of-art methodologies without compromising future improvements of the system.

3 Experimental Results

Two different datasets have been used to validate the experiments and ensure that the conclusions are not dependant on the fundus camera. The first dataset was captured with a stereoscopic camera Kowa nonmyd WX. Only one of the two images provided was used since our goal is to evaluate monoscopic systems for screening, given their broader availability. The dataset is composed of 29 samples, 14 healthy eyes and 15 glaucomatous ones. The second dataset is a standard set, publicly available [18], which facilitates future comparison of our methodology with others. It contains 26 samples, 8 healthy and 18 glaucomatous discs. Both datasets contain different degrees of glaucoma, from very early stages

Table 1: Results over the 2 datasets. Four first rows show the state-of-art appearance features, while three middle rows show different variation of our framework and last two rows the combination of our pipeline with appearance features

Method	Dataset 1						Dataset 2					
	Acc	Sens	Spec	Prec	Recl	Fmes	Acc	Sens	Spec	Prec	Recl	Fmes
Intensity + PCA [8]	0.59	0.60	0.57	0.6	0.6	0.6	0.48	0.55	0.33	0.65	0.55	0.59
FFT + PCA [8]	0.45	0.40	0.50	0.46	0.40	0.43	0.38	0.45	0.22	0.56	0.45	0.50
Spline + PCA [8]	0.59	0.60	0.57	0.60	0.60	0.60	0.59	0.65	0.44	0.72	0.65	0.68
All appearance [8]	0.55	0.53	0.57	0.57	0.53	0.55	0.46	0.50	0.33	0.63	0.50	0.55
CDR (no circle detect.)	0.79	0.73	0.85	0.85	0.73	0.79	0.46	0.50	0.38	0.62	0.50	0.55
CDR	**0.89**	**0.93**	**0.85**	**0.88**	**0.93**	**0.90**	0.59	0.56	0.63	0.75	0.56	0.64
CDR + Diff	0.82	0.87	0.77	0.81	0.87	0.84	**0.71**	**0.69**	**0.75**	**0.85**	**0.69**	**0.76**
Intensity + PCA + CDR	0.69	0.73	0.64	0.69	0.73	0.71	0.50	0.56	0.38	0.64	0.56	0.60
All Appearance + CDR	0.62	0.60	0.64	0.64	0.60	0.62	0.46	0.56	0.25	0.60	0.56	0.58

to severe cases, as well as other disorders, such as hypermetrope, haemorrhages or peripapillary atrophy, that can make diagnosis difficult.

Different variations of our methodology were tested, using only the cup-to-disc ratio, the rim width variation or a combination of both. The circle detector and filter, used to reduce segmentation errors outside the ONH was also evaluated. All the parameters were setup experimentally and kept identical for all experiments and datasets in order to compare the methods in equal conditions and to avoid overfitting to specific cases. In order to validate our method and extract pertinent conclusions regarding the comparison between geometrical and appearance features, different appearance based methodologies -pixel values, fft coefficients B-spline coefficients and a combination of all- were implemented following the description, setup and conclusions by Block et al. [8]. All experiments were performed using leave-one-out cross validation.

Results are shown in Table 1 in terms of accuracy (Acc), sensitivity (Sens), specificity (Spec), precision (Prec), recall (Recl) and F measurement (Fmes). It can be seen how our framework provides accurate glaucoma diagnosis. The extraction and usage of geometrical features seems to provide superior diagnosis accuracy under realistic conditions: when the number of training images is small, they perform much better than appearance based. This explains the significant decrease in performance of appearance based feature compared to other results reported in the literature [16, 8], where hundred of examples were available. Since those features depend heavily on the camera setting, they require retraining for every camera model and therefore they are difficult to deploy in the real world.

Other conclusions can be drawn from these results. The optical disc detection and filtering gives a significant improvement in the final classification. The rim variation feature does not always provide a significant increase in accuracy, especially in the first dataset where the high resolution allows a perfect segmentation of disc and cup, but it plays a significant role for cheaper cameras. The second dataset complexity, with a much lower resolution, is reflected in the final performance of all the tested methods. Finally, by adding geometrical features to the appearance feature vector, results appear to invariably improve, which shows the potential of combining both methodologies with the potential of fully exploiting the advantages of both techniques.

4   Conclusions

In this paper, a method for glaucoma diagnosis, based on ONH segmentation of retinal images, is proposed. Our framework is able to accurately extract the cup and the rim of the optical disc to extract high level geometrical features. The obtained values are then used as input to a machine learning classifier, responsible for detecting glaucoma given a new retinal image.

Our approach has been designed for glaucoma screening in real world conditions. Experiments on varied datasets were performed to evaluate our schema with different cameras and resolutions, and both colour and black and white images. The proposed method achieved high accuracy rates overperforming state-of-art methodologies in real conditions, when small training sets are available. The experiments also validated the usage of geometrical features for glaucoma detection and as a complement to appearance based methods. As future work, a diagnosis study will be performed to ensure the validity of our conclusion in a larger scale and the potential of our framework for glaucoma screening and diagnosis in real life.

References

[1] Burr J, Mowatt G, Siddiqui MAR, Herandez R et al. (2007). The clinical and cost effectiveness of screening for open angle glaucoma: a systematic review and economic evaluation. *Health Technology Assessment*, 11(41).

[2] Tuck MW, Crick RP (2003). The projected increase in glaucoma due to an ageing population. *Ophthalmic and Physiological Optics*, 23(2):175-179.

[3] Harrison RJ, Wild JM, Hobley AJ (1988). Referral patterns to an ophthalmic outpatient clinic by general practitioners and ophthalmic opticians and the role of these professionals in screening for ocular disease. *Br Med J*, 297:1162-7.

[4] Bowling B, Chen SD, Salmon JF (2005). Outcomes of referrals by community optometrists to a hospital glaucoma service. *Br J Ophthalmol*, 89:1102-4.

[5] Patel UD, Murdoch IE, Theodossiades J (1920). Glaucoma detection in the community: does ongoing training of optometrists have a lasting effect?. *Eye*, 591-4.

[6] Vernon SA, Ghosh G (2001). Do locally agreed guidelines for optometrists concerning the referral of glaucoma suspects influence referral practice?. *Eye*, 15:458-63

[7] Schultz RO, Radius RL, Hartz AJ, Brown DB, Eytan ON, Ogawa GS, Kuhn E, Simons KB (1995). Screening for glaucoma with stereo disc photography. *J Glaucoma*, 4:177-82.

[8] Bock R, Meier J, Nyl LG, Hornegger J, Michelson G (2010). Glaucoma risk index: Automated glaucoma detection from color fundus images. *Med Image Anal*, 14:471-81.

[9] Bankhead P, Scholfield CN, McGeown JG, Curtis TM (2012). Fast retinal vessel detection and measurement using wavelets and edge location refinement. *PLoS ONE*, 7(3):1-12.

[10] Armaly MF, Sayegh RE (1969). The cup/disc ratio. The findings of tonometry and tonography in the normal eye. *Arch Opthalmol*, 82:191-6.

[11] Jonas JB, Gusek GC, Naumann GO (1988). Optic disc morphometry in chronic primary open-angle glaucoma. I. Morphometric intrapapillary characteristics. *Graefes Arch Clin Exp Ophthalmol*, 226(6):522-30.

[12] Hatanaka Y, Noudo A, Muramatsu C, Sawada A, Hara T, et al. (2010). Automatic Measurement of Vertical Cup-to-Disc Ratio on Retinal Fundus Images. *Lecture Notes in Computer Science*, 6165:64-72.

[13] Liu J, Wong DWK, Lim JH, Li H, Tan NM, Zhang Z, Wong TY, Lavanya R (2009). ARGALI: An Automatic Cup-to-Disc Ratio Measurement System for Glaucoma Analysis Using Level-set Image Processing. *International Conference on Biomedical Engineering IFMBE Proceedings*, 23:559-562

[14] Fondn I, Nez F, Tirado M, Jimnez S, Alemany P, Abbas Q, Serrano C, Acha B (2012). Automatic Cup-to-Disc Ratio Estimation Using Active Contours and Color Clustering in Fundus Images for Glaucoma Diagnosis, Image Analysis and Recognition. *Lecture Notes in Computer Science*, 7325:390-399

[15] Narasimhan K, Vijayarekha, K (2011). An Efficient Automated System For Glaucoma Detection Using Fundus Image. *Journal of Theoretical and Applied Information Technology*, 33:1

[16] Mookiah MRK, Acharya UR, Lim CM, Petznick A, Suri JS (2012). Data mining technique for automated diagnosis of glaucoma using higher order spectra and wavelet energy features, *Knowledge-Based Systems*, 33:73-82.

[17] Xu J, Chutatape O, Sung E, Zheng C, Chewteckuan P (2007). Optic disk feature extraction via modified deformable model technique for glaucoma analysis. *Pattern Recognition*, 2063-2076.

[18] Spaeth G (2014). An online resource for ophtalmologists, physicians, medical students and optometrists. Online: http://www.optic-disc.org/ [Last access: January 2014]

[19] Atherton TJ, Kerbyson DJ (1999). Size invariant circle detection. *Image and Vision Computing*, 17(11):795-803.

[20] Wollstein G, Garway-Heath DF, Hitchings RA (1998). Identification of early glaucoma cases with the scanning laser ophthalmoscope. *Ophthalmology*, 105(8):15571563.

[21] Acharya R (2005). *Image Processing: Principles and Applications*, Wiley-Interscience.

[22] Otsu N (1979). A threshold selection method from gray-level histograms. *IEEE SMCB*, 9(1):6266.

[23] Auria L, Moro R (2008). Support Vector Machines (SVM) as a technique for solvency analysis. *Discussion papers*, 811, DIW Berlin.

[24] Turk M, Pentland A (1991). Eigenfaces for recognition. *J. Cognit. Neurosci.*, 3(1):7186.

[25] Sezgin M, Sankur B (2004). Survey over image thresholding techniques and quantitative performance evaluation. *J. Electronic Imaging*, 13(1):146-165.