Computers in Biology and Medicine ∎ (∎∎∎∎) ∎∎∎–∎∎∎

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/cbm

Automated retinal image quality assessment on the UK Biobank dataset for epidemiological studies

R.A. Welikala [a,*], M.M. Fraz [b], P.J. Foster [c], P.H. Whincup [d], A.R. Rudnicka [d], C.G. Owen [d], D.P. Strachan [d], S.A. Barman [a],

on behalf of the UK Biobank Eye and Vision Consortium [1]

[a] School of Computing and Information Systems, Kingston University, Surrey KT1 2EE, United Kingdom
[b] School of Electrical Engineering and Computer Science, National University of Sciences and Technology, Islamabad 44000, Pakistan
[c] NIHR Biomedical Research Centre, Moorfields Eye Hospital, London, EC1V 2PD and UCL Institute of Ophthalmology, London EC1V 9EL, United Kingdom
[d] Population Health Research Institute, St. George's, University of London, London SW17 0RE, United Kingdom

ARTICLE INFO

Article history:
Received 13 November 2015
Accepted 30 January 2016

Keywords:
Retinal image
Image quality
Vessel segmentation
Large retinal datasets
UK Biobank
Epidemiological studies

ABSTRACT

Morphological changes in the retinal vascular network are associated with future risk of many systemic and vascular diseases. However, uncertainty over the presence and nature of some of these associations exists. Analysis of data from large population based studies will help to resolve these uncertainties. The QUARTZ (QUantitative Analysis of Retinal vessel Topology and siZe) retinal image analysis system allows automated processing of large numbers of retinal images. However, an image quality assessment module is needed to achieve full automation. In this paper, we propose such an algorithm, which uses the segmented vessel map to determine the suitability of retinal images for use in the creation of vessel morphometric data suitable for epidemiological studies. This includes an effective 3-dimensional feature set and support vector machine classification. A random subset of 800 retinal images from UK Biobank (a large prospective study of 500,000 middle aged adults; where 68,151 underwent retinal imaging) was used to examine the performance of the image quality algorithm. The algorithm achieved a sensitivity of 95.33% and a specificity of 91.13% for the detection of inadequate images. The strong performance of this image quality algorithm will make rapid automated analysis of vascular morphometry feasible on the entire UK Biobank dataset (and other large retinal datasets), with minimal operator involvement, and at low cost.

1. Introduction

Examination of the blood vessel structure in retinal images offers an opportunity to directly and non-invasively observe the blood circulatory system. The morphological characteristics of retinal vessels (e.g. vessel calibre, tortuosity) have been prospectively associated with cardiovascular and systemic disease [1–4]. Approximately 20 million people in the UK have a long term health condition and its increasing prevalence is a major challenge for the healthcare system [5]. Cardiovascular disease alone accounts for nearly 200,000 deaths in the UK per year [6], with coronary heart disease, stroke and heart failure accounting for most of these deaths. Early detection and prevention of disease outcome is key. Accurate assessment of retinal vessel morphology may be an important biomarker of vascular health, which might predict those at high risk of disease [7].

UK Biobank includes probably the world's largest retinal image repository (nearly 136,000 retinal images) in a middle-aged population-based cohort study. The depth and breadth of health data recorded make it a powerful research resource for improving the prevention, diagnosis and treatment of a wide range of serious illnesses [8]. Considerable potential exists in using this retinal dataset to discover biomarkers for the identification of high risk patients. Retinal vessel morphology may provide such a measure, which could be used to identify those at high risk, particularly of vascular related disease. However, the extraction of quantitative measures from the vessel morphology in large datasets are needed to establish the presence (or absence) of associations, and this has been prohibitive to date given the considerable amount of manual

* Corresponding author.
E-mail address: R.Welikala@kingston.ac.uk (R.A. Welikala).
[1] Members of the UK Biobank Eye and Vision Consortium are listed before References.

http://dx.doi.org/10.1016/j.compbiomed.2016.01.027
0010-4825/

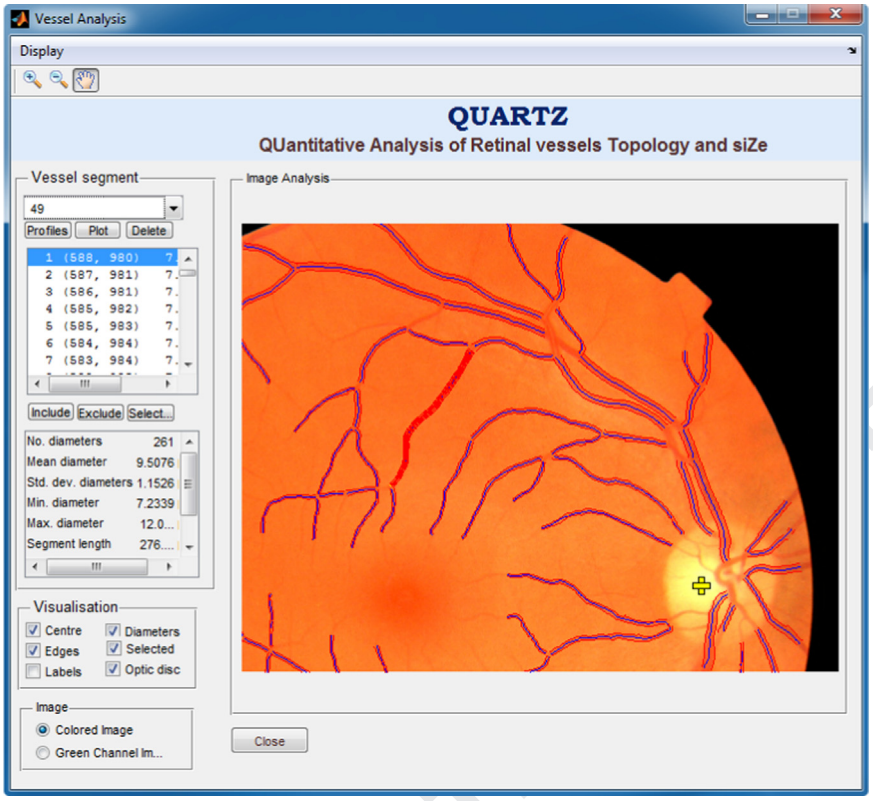

Fig. 1. QUARTZ vessel analysis screen.

operator involvement required. Therefore, automation is needed to process and analyse the large amount of images and extract useful, objective and quantitative information from vessel morphology [9].

Our research group have developed a retinal image analysis system called QUARTZ (QUantitative Analysis of Retinal vessel Topology and siZe), which is used for the automated processing of large numbers of retinal images and obtains quantitative measures of vessel morphology to be used in epidemiological studies [10]. QUARTZ was designed to be fully automated and includes vessel segmentation, measurements from retinal vessels, arteriole/venule (a/v) classification and optic disc localisation. The software also derives information from the whole retina and not simply concentric areas centred on the optic disc or manually selected vessels. The vessel analysis screen of the QUARTZ software is illustrated in Fig. 1.

QUARTZ is considered fully automated in respect to its ability to run on large datasets uninterrupted. However, if quantitative results are to be of high precision, then it is vital to identify and remove images of inadequate quality. Currently, this step of image quality assessment is performed manually. This takes 15 s per image, which totalled to approximately 67 h of manual processing time when QUARTZ was previously used to process 16,000 images [11] from another large population based study (the European Prospective Investigation on Cancer study in Norfolk; EPIC-Norfolk) [12]. VAMPIRE [13,14], SIRIUS [15], ARIA [16] and other notable vessel analysis software, also rely on manual image quality processing. Hence, the development of automated image quality assessment would be extremely beneficial in speeding up the process.

Retinal images can be poorly and unevenly illuminated, blurred and obstructed. This is caused by factors including camera exposure, focal plane error, poorly dilated/small pupils, eye lashes/blinking, lens artefacts, media opacity (e.g. cataracts, vitreous haemorrhages, asteroid hyalosis) and head/eye movement.

The criteria for an image to be classed as adequate differs for retinal images used to make a conventional diagnosis (e.g. detection of diabetic retinopathy) to those which are useful for extracting vessel morphometric data suitable for epidemiological studies. For the former, there must be good clarity of the entire image to ensure any signs of possible pathology are not missed. For the latter, the criteria are that image clarity must be sufficient to allow for the accurate vessel segmentation for at least a portion of the image. Useful information can still be extracted from well segmented sections of the vasculature, even if this only represents a portion of the vascular tree. From our experience most large retinal datasets used in epidemiological studies (EPIC-Norfolk, UK Biobank) contain large amounts of images of poor quality. Poorly/unevenly illuminated images are a particular problem which is expected when images are captured without the use of pharmacological mydriasis. Therefore this approach ensures that there is little wastage by making use of many of the poor images, extracting as much information as possible from these retinal datasets.

In this paper, an automated image quality assessment methodology is proposed which is designed to classify retinal images as inadequate or adequate for use in the creation of vessel morphometric data suitable for epidemiological studies. As far as we are aware, this paper presents the first automated methodology designed to tackle this important issue. The methodology is based on the assessment of the segmented vasculature, which involves the extraction of three global features (area, fragmentation and complexity) measured from the segmented vessel map and support vector machine (SVM) classification. This step completely removes the need for any manual processing, allowing QUARTZ to analyse the entire UK Biobank dataset (and other large retinal datasets) without manual intervention.

Please cite this article as: R.A. Welikala, et al., Automated retinal image quality assessment on the UK Biobank dataset for epidemiological studies, Comput. Biol. Med. (2016), http://dx.doi.org/10.1016/j.compbiomed.2016.01.027