Computer Methods and Programs in Biomedicine 161 (2018) 197–207

Contents lists available at ScienceDirect

Computer Methods and Programs in Biomedicine

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

Retinal artery/vein classification using genetic-search feature selection

Fan Huang[a], Behdad Dashtbozorg[a], Tao Tan[a,c], Bart M. ter Haar Romeny[b,a,*]

[a] Department of Biomedical Engineering, Eindhoven University of Technology, Eindhoven, The Netherlands
[b] Department of Biomedical and Information Engineering, Northeastern University, Shenyang, China
[c] Mammography, ScreenPoint Medical, Nijmegen, The Netherlands

ARTICLE INFO

Article history:
Received 22 November 2017
Revised 9 March 2018
Accepted 17 April 2018

Keywords:
Fundus image
Artery/vein classification
Genetic search feature selection

ABSTRACT

Background and objectives: The automatic classification of retinal blood vessels into artery and vein (A/V) is still a challenging task in retinal image analysis. Recent works on A/V classification mainly focus on the graph analysis of the retinal vasculature, which exploits the connectivity of vessels to improve the classification performance. While they have overlooked the importance of pixel-wise classification to the final classification results. This paper shows that a complicated feature set is efficient for vessel centerline pixels classification.

Methods: We extract enormous amount of features for vessel centerline pixels, and apply a genetic-search based feature selection technique to obtain the optimal feature subset for A/V classification.

Results: The proposed method achieves an accuracy of 90.2%, the sensitivity of 89.6%, the specificity of 91.3% on the INSPIRE dataset. It shows that our method, using only the information of centerline pixels, gives a comparable performance as the techniques which use complicated graph analysis. In addition, the results on the images acquired by different fundus cameras show that our framework is capable for discriminating vessels independent of the imaging device characteristics, image resolution and image quality.

Conclusion: The complicated feature set is essential for A/V classification, especially on the individual vessels where graph-based methods receive limitations. And it could provide a higher entry to the graph-analysis to achieve a better A/V labeling.

1. Introduction

Many systemic diseases including diabetes and hypertension cause blood vessels change (becoming tortuous, narrowing etc.) and even leakage may occur, leading to serious complications like blurry vision and hand/feet tingling and pain [1]. A retinal image provides direct access to vascular abnormalities and enables further quantitative analysis on the retinal vasculature. The study on retinal arteries and veins has received much attention in the field of retinal image analysis, since many artery-vein related biomarkers have been found significantly associated with the progress of diseases. In diabetic retinopathy (DR), the narrowing on arterioles and the widening on venules are observed, which result in a lower arteriolar-to-venular diameter ratio (AVR) of DR patients [2–4]. In hypertension, decrease on generalized arteriolar diameter is associated to the increased blood pressure level [5]. Additional measurements such as tortuosity (generalized vascular curvature) [6,7], vessel branching angle [8,9] and junction exponents [9,10] have received more and more interest. It is important to note that these clinical relevant features behave differently on arteries and veins respectively under pathological conditions. For instance, the arterial vessel wall is more elastic and thinner than the venous wall, thus abnormal arteries are usually more tortuous than veins [11]. Therefore, quantitative biomarkers extracted from arteries and veins separately might reveal more information for diseases progress rather than examining them together.

Due to the fact that high resolution fundus imaging is mostly low cost and fast, retinal screening programs usually produce huge amounts of data for analysis. It is then unrealistic to let human observers manually label the arteries and veins. Therefore, developing a fully automatic artery/vein (A/V) classification system is a prerequisite for automated large-scale retinal image analysis.

* Corresponding author at: Department of Biomedical Engineering, Eindhoven University of Technology, Eindhoven, The Netherlands.
E-mail address: b.m.terhaarromeny@tue.nl (B.M. ter Haar Romeny).

https://doi.org/10.1016/j.cmpb.2018.04.016

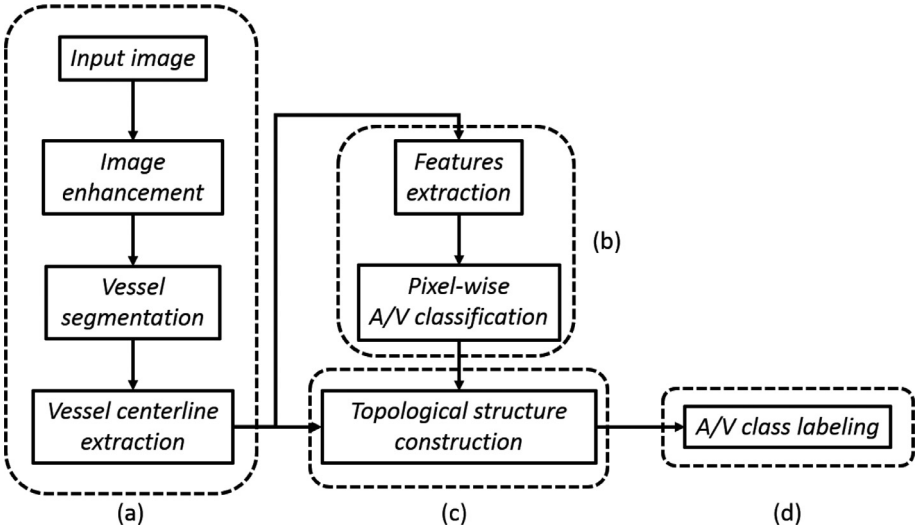

Fig. 1. The general pipeline for an automatic A/V classification program, where (a): image preprocessing, (b): pixel-wise A/V classification, (c): vessel topological structure construction and (d): A/V label assignment.

Several automatic A/V classification systems have been proposed in literature [12–16]. In summary, most of these methods consist of 4 main modules: (1) image pre-processing, (2) pixel-wise A/V probability assignment, (3) vessel topological structure construction and (4) A/V label determination (see Fig. 1). In the pre-processing step, images are enhanced by image processing techniques such as luminosity normalization and histogram equalization which correct for the illumination and background inhomogeneity. Afterwards, the retinal vessels are segmented yielding a vessel binary map. In the pixel-wise A/V classification module, intensity-based features are extracted for all vessel centerline pixels. Using these features, a supervised or unsupervised machine learning technique is exploited to assign a probability (between 0 and 1) to these pixels. After that, a topological structure of the vascular network is built in order to extract the connectivity relation between each individual segment. It improves the result of pixel-wise classification, because arteries only cross veins but not themselves and vessels connected to each other must be of the same type. Finally, the A/V label of each vessel segment, or even a full vessel tree is determined by using both the local and contextual information.

In the last few years, publications on A/V classification mainly focus on constructing the topological structure of vessels using graph theory. Joshi et al. used Dijkstra's graph search algorithm to connect vessel segments as subtrees and clustered them into arterial and venous classes [13]. Dashtbozorg et al. applied graph analysis on individual vessel segments and determined the type of vessels by combining the graph label and the pixel-wise A/V label [14]. Hu et al. constructed the vascular structure by a graph-based and a meta-heuristic algorithm [15]. Estrada et al. incorporated domain-specific features with a topology framework to construct a global likelihood model for A/V classification [17].

Exploiting vessel contextual information for A/V classification is novel, while good pixel-wise classification is also a crucial entry step. Because even if the graph analysis was perfect, an incorrect local A/V probability might still result in wrong A/V label estimation and further affect the corrected labeling of the whole vessel tree. The recent frameworks proposed in literature still use the information extracted by a small amount of features for supervised/un-supervised classification. Joshi et al. extracted only 4 features, Dashtbozorg et al. used 19 features (after feature selection), Niemeijer et al. [12] and Hu et al. [15] used 31 features, Mirsharif et al. [18] used 8 features (after feature selection) and Xu et al. [19] used 21 features for pixel-wise classification. Additionally, the category of features used in these works is limited, where only the local intensity values on multiple color channels (e.g. RGB, HSB and CIExyz) are used.

In this paper, we show that a more complicated feature set is more efficient in the discrimination of artery and vein. We developed a novel framework for pixel-wise A/V classification, which extracts features of different categories for vessel centerline pixels. An advanced feature selection technique, named genetic-search feature selection, is applied to obtain the optimal subset of features for classification. Then this framework was validated on five retinal image databases, including two public datasets and three clinical datasets.

2. Method

Arteries carry oxyhemoglobin which transports oxygen molecules from respiratory organs (e.g. lungs) to the rest of body (e.g. tissue), while veins carry deoxygenated hemoglobin which without the bound oxygen. Oxyhemoglobin is visually brighter and deoxygenated hemoglobin is darker. Therefore, on retinal images arteries are mostly brighter than veins, which makes the pixel intensities of vessels become very important features for discrimination. In the proposed method, we examined many intensity based features such as red, green, blue, hue, saturation, brightness etc. for every vessel centerline pixel. While if we simply feed the color intensities (such as RGB or HSB) of a pixel to a machine learning classifier, the classification results are usually disappointing. It turns out that our brain must take into account more information than considering only the local intensities to determine the vessel type. In addition, the fundus cameras installed in eye clinic are different from each other in the sense of field-of-view, image resolution, imaging flashlight and the embedded post-processing techniques. A well-trained system might work perfectly on the images from one dataset, but it may fail on the ones from others.

In this paper, we developed a framework to address this issue, which extracts a large amount of features for each vessel pixels, followed by a feature selection algorithm. The methodology starts by enhancing the contrast and correcting the luminosity variation of imported image. Afterwards, we apply vessel segmentation technique to obtain the vessel binary map as well as centerlines. For each centerline pixel we extract in total 455 features containing the categories of local intensity, contextual intensity, global intensity and spatial information. Then we use genetic-search