Informatics in Medicine Unlocked 9 (2017) 44–57

Contents lists available at ScienceDirect

Informatics in Medicine Unlocked

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

Detection of microaneurysms in retinal images using an ensemble classifier

M.M. Habib [a,*], R.A. Welikala [a], A. Hoppe [a], C.G. Owen [b], A.R. Rudnicka [b], S.A. Barman [a]

[a] School of Computer Science and Mathematics, Faculty of Science, Engineering and Computing, Kingston University, London, UK
[b] Population Health Research Institute, St. George's, University of London, United Kingdom

ARTICLE INFO

Keywords:
Image processing
Medical image analysis
Retinal imaging
Microaneurysm detection
Tree ensemble
Diabetic retinopathy

ABSTRACT

This paper introduces, and reports on the performance of, a novel combination of algorithms for automated microaneurysm (MA) detection in retinal images. The presence of MAs in retinal images is a pathognomonic sign of Diabetic Retinopathy (DR) which is one of the leading causes of blindness amongst the working age population. An extensive survey of the literature is presented and current techniques in the field are summarised. The proposed technique first detects an initial set of candidates using a Gaussian Matched Filter and then classifies this set to reduce the number of false positives. A Tree Ensemble classifier is used with a set of 70 features (the most commons features in the literature). A new set of 32 MA groundtruth images (with a total of 256 labelled MAs) based on images from the MESSIDOR dataset is introduced as a public dataset for benchmarking MA detection algorithms. We evaluate our algorithm on this dataset as well as another public dataset (DIARETDB1 v2.1) and compare it against the best available alternative. Results show that the proposed classifier is superior in terms of eliminating false positive MA detection from the initial set of candidates. The proposed method achieves an ROC score of 0.415 compared to 0.2636 achieved by the best available technique. Furthermore, results show that the classifier model maintains consistent performance across datasets, illustrating the generalisability of the classifier and that overfitting does not occur.

1. Introduction

Retinal Image Analysis (RIA) is an active area of research due to its application in screening programs for Diabetic Retinopathy (DR) – one of the leading causes of blindness in the developed world. During the screening process, fundus images of the retina are captured for the purpose of detection of diabetic retinopathy. The presence of microaneurysms (MAs) in retinal images is an early indicator of DR (Fig. 1). The automated detection of MAs from retinal images can aid in screening programs for DR diagnosis. Several algorithms have been proposed for the detection of MA, however, MA detection is still a challenging problem due to the variance in appearance of MAs in retinal images [1].

Through our review of MA detection in the literature, we have identified three main stages in MA detection algorithms: 1) preprocessing 2) MA candidate detection and 3) candidate classification. *Preprocessing* corrects non-uniform illumination in retinal images and enhances the contrast of MAs in the image. *MA candidate detection* seeks to detect an initial set of candidate regions where MAs are likely to exist. *MA candidate classification* applies machine learning techniques in order to improve the specificity of the algorithm by filtering out false positives from the candidate detection phase. Some of the proposed methods in the literature are unsupervised methods, which means they do not require the third classification stage [1–7]. A summary of MA candidate detection algorithms presented in the literature is listed in Table 1. For each algorithm the table describes image type, the initial candidates method, the classifier used, and the reported performance for each classifier. Most of the literature has differences in the method used to evaluate their algorithms or the dataset used, which makes it difficult to compare any 2 algorithms together. One of the earliest proposed techniques for MA detection was applied to fluorescein angiograms [8]. A Gaussian matched filter was used to detect the initial set of candidates. Finally, each initial candidate was classified as either a true candidate or a spurious one using some features, producing the final classification result. Cree [9] applied Spencer's technique [8] to multiple longitudinal florescence images in order to detect the 'MA turnover' – the appearance or disappearance of MA objects over time.

More recent techniques have tackled the problem of MA detection in colour fundus images. The main reason for this is that colour images are more common in screening programs and are also non-invasive to capture, unlike fluorescein images. The following methods are all based on MA detection in colour fundus images.

A number of techniques have adapted Spencer's approach in terms of

* Corresponding author.
E-mail addresses: m.habib@kingston.ac.uk (M.M. Habib), s.barman@kingston.ac.uk (S.A. Barman).

http://dx.doi.org/10.1016/j.imu.2017.05.006
Received 9 March 2017; Received in revised form 12 May 2017; Accepted 13 May 2017
Available online 17 May 2017

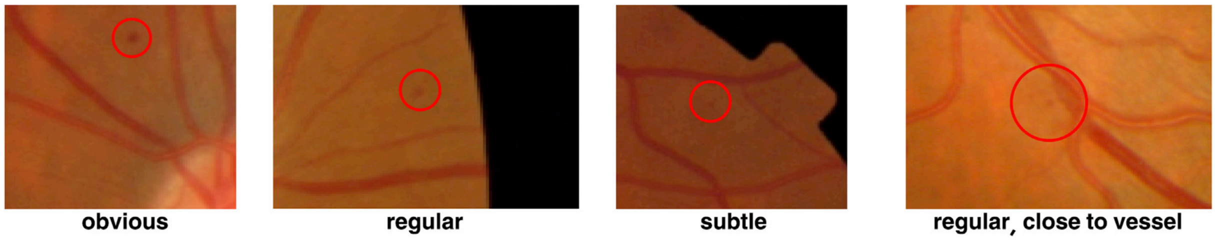

Fig. 1. Examples of various microaneurysms with varying appearances and locations.

applying morphological vessel removal followed by a Gaussian matched filter. Hipwell [10] performed a modification of Cree [9] in order to apply the algorithm to colour fundus images. Streeter [11] used a method based on Cree [9]. However, during the classification phase, 80 features are extracted and Linear Discriminant Analysis (LDA) was used to perform the classification. Feature Selection was performed to filter the features down to 16 features. Feature Selection is a process to reduce redundant features in order to reduce computational time and decrease chances of overfitting. Another Spencer-based approach was introduced in Fleming [12]. This technique introduced a novel region-growing step based on gradient values, rather than a simple threshold. In addition a paraboloid was fitted to each candidate using a parameter optimization process. The paraboloid parameters are used to compute many of the features used in the candidate classification phase. Instead of using a single Gaussian matched filter, Zhang [13] applied multiple Gaussian filters at multiple scales and computed the maximum response to produce a probability map of the likelihood of presence of MA candidates. This probability map was then thresholded to produce the initial set of MA candidates. Finally a rule-based classifier using 30 features was used to perform the final classification. Li [4] used an unsupervised method based on a Multi-orientation Sum of Matched Filter (MSMF). This filter is a modification of the classical Gaussian Matched filter. This modified filter is anisotropic in nature and is applied in multiple directions. Hence, this filter is better at suppressing responses to blood vessels than the Gaussian Matched filter. Wu [14] modified the MSMF filter to take into account the varying size of MAs.

Sánchez used a mixture model-based clustering technique to detect the initial MA candidate regions [7]. The technique fits three normal distribution histograms to the retinal image histogram. These histograms correspond to foreground, background and outliers. The foreground histogram pixels are considered as the initial set of MA candidate regions. Finally, logistic regression was used to classify each MA region as belonging to either a foreground or background region. Quellec [15] based his technique on wavelet transforms applied in different sub-bands of the colour image.

A double-ring filter was used in Mizutani [16] to detect the initial candidates. The filter used the property that MAs are dark circular regions within a brighter region to detect the MA candidates. It consists of an inner ring and an outer ring. A given pixel is considered to be a MA pixel if the average intensity of the inner ring is smaller than the average intensity of the outer ring. After the initial candidates are detected, classification is performed using 12 extracted features and an Artificial Neural Network (ANN).

Initial candidates were detected using a simple thresholding in Giancardo [5,6]. A novel Radon-based transform was used to extract the features of the initial candidates and a Support Vector Machine (SVM) classifier was used to perform the final classification. An initial set of 31 features were computed for classification. The dimensionality of the features was reduced to 10 dimensions using Principle Component Analysis (PCA), and this reduced representation was used to perform the classification. A reduced dimension for the features reduces the risk of overfitting and also makes the classification more computationally efficient.

Sinthanayothin [17] used a 'moat operator' to enhance red lesions in the image and then these regions were segmented. Vessel regions were then removed to produce the final set of candidates. Note that this method detected both MAs and haemorrhages. The moat operator was not defined in the paper and we were unable to find the exact definition in the literature.

AbdelAzeem [18] used a Hessian matrix in order to detect the initial MA candidate set. A rule based classifier is then used to detect false MA detections. The rule is simply based on the candidate 'energy'. The exact definition of the computed 'energy' was not mentioned in the paper, however, it is likely to be the same definition as in Fleming [12]. Inoue [19] relied on a Hessian matrix in order to detect the initial candidates and an Artificial Neural Network (ANN) was used to classify the features. A group of 126 features were fed into the ANN for classification. However this group of features was reduced using Principle Component Analysis (PCA) in order to reduce computational complexity and avoid overfitting. Moreover, Srivastava [20] used the eigenvalues of the hessian matrix in order to detect the initial candidates. Recently, Adal has used a hessian matrix in order to detect the initial set of MA candidates. A combination of SURF, Radon and scale-space features were extracted from the initial candidates. Multiple classifiers (Support Vector Machines, K-Nearest-Neighbours, Naive Bayes and Random Forest) were also experimented with in this technique.

An adaptation of Spencer [8] and Frame [22] is presented in Niemeijer [22]. Two main contributions were added: A pixel based classification system for the initial candidate detection phase and an extended set of features used for pixel classification.

A unique method was introduced in Lazar [1,2] since it is an unsupervised technique that does not require any training or classification steps. Moreover the reported results of this technique are comparable to other supervised methods, which make it a promising method. The essence of this technique is to discriminate between vessels and MAs by using a 1D scanline at different directions for each pixel. While a MA will have local minima in all directions of the rotated scanlines, a vessel will have only one minima corresponding to when the scanline is perpendicular to the vessel. Hence, using this property, a probability map is produced at each pixel and then simple thresholding is applied to produce the final set of candidates.

Garcia [30] compared the accuracy of four neural network variants: Multilayer Perceptron (MP), Radial Basis Function (RBF), Support Vector Machine (SVM) and Majority Voting (MV). The initial candidates were detected using a local thresholding technique based on the mean pixels of the entire image compared to mean intensity in a small window around a pixel. According to their experiments, the RBF was suggested as the preferred classifier among all 4. An interesting approach that relies on visual dictionaries was presented in Rocha [24]. The use of visual dictionaries (bag of words) makes this approach more generalizable since it does not rely on specific features during the classification. Therefore, the same approach can be used to perform detection of lesions other than MAs as well. The disadvantage of this is that it requires a larger training set. Haloi [29] recently applied deep neural networks to detect MAs in colour images. Deep neural networks have gained popularity in the field of computer vision in the recent years since they do not require manual feature engineering (selection of features). Moreover, algorithms based on deep learning have produced results that out-perform other state-of-