Feature selection for detection of new vessels on the optic disc

Keith A. Goatman[1]
k.a.goatman@abdn.ac.uk
Alan D. Fleming[1]
a.d.fleming@abdn.ac.uk
John A. Olson[2]
john.olson@nhs.net
Peter F. Sharp[1]
p.sharp@abdn.ac.uk

[1] University of Aberdeen
Foresterhill,
Aberdeen, AB25 2ZD, UK.

[2] Diabetic Retinal Screening Service,
David Anderson Building,
Foresterhill,
Aberdeen, AB25 2ZP, UK.

Abstract

The development of new vessels on the retina of people with diabetes is rare, but is likely to lead to severe visual impairment. This paper investigates the selection of suitable image features for the automatic detection of new vessels on the optic disc. The features are chosen based on their discrimination capability (tested using the non-parametric Wilcoxon rank sum and Ansari-Bradley dispersion tests) and absence of correlation with other features (tested using the Kendall Tau coefficient). Classification was performed using a support vector machine. The system was trained and tested by cross-validation using 38 images with new vessels and 71 normal images without new vessels. Fourteen features were selected, giving an area under the receiver operator characteristic curve of 0.911 for detecting images with new vessels on the disc. The method could have a useful role as part of an automated retinopathy analysis system.

1 Introduction

Diabetic retinopathy causes several different retinal lesions. Usually the first sign of retinopathy is the microaneurysm (MA). These appear in the photograph as small red dots. As the disease progresses capillaries may begin to leak, forming exudates, bright yellow/white lipid deposits. Larger, dark red blot haemorrhages may also form at this stage. As the disease progresses to its proliferative stage, ischaemia can trigger abnormal vessel changes, such as venous beading (VB), intra-retinal microvascular abnormalities (IRMA) and new vessel growth. New vessels are classified according to their position, either new vessels at the disc (NVD) if they occur on or within one optic disc diameter of the disc, or new vessels elsewhere (NVE). Although the prevalence of new vessels is low, typically 0.4% of the screening population [3], the associated risk of rapid vision loss mean it must be detected reliably. Figure 1 shows some examples of normal and abnormal optic disc vessels.

There has been little work automating new vessel detection. However, Jelinek *et al.*, in a study of 27 images, examined vessel characteristics in fluorescein angiograms in an attempt

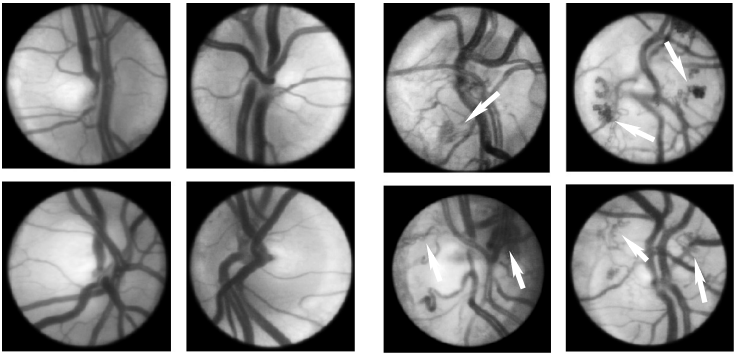

Figure 1: Examples of (a) normal discs and (b) discs containing abnormal vessels, indicated by the white arrows. New vessels are likely to bleed and lead to retinal detachment.

to predict proliferative disease [1]. This paper describes a method for detecting new vessels on the optic disc in standard screening photographs.

2 Method

A total of 109 colour retinal images were collected from two screening programmes and a hospital eye clinic. Thirty-eight of the images contained new vessels. Two experienced graders annotated vessels they considered abnormal; the logical AND of their annotations was the reference standard.

2.1 Small vessel detection

Several authors have described methods for segmenting normal retinal vessels outside of the optic disc, but there has been little interest in the disc vessels themselves (indeed the disc is often masked out altogether). The method below combines a watershed transform and ridge strength to detect the small and often tortuous disc vessels.

The image was first inverted and filtered with a Gaussian function (standard deviation equal to 2 pixels) to prevent over-segmentation. The binary watershed lines were thinned, such that only the pixels at vessel bifurcations have more than two neighbours. Candidate segments were separated by removing the pixels at bifurcations. The watershed transform generates closed regions, not all of which coincide with vessels. To remove the non-vessel segments the ridge strength, *kappa*, was calculated as [2]

$$\kappa = \frac{L_x^2 L_{yy} + L_y^2 L_{xx} - 2L_x L_y L_{xy}}{(L_x^2 + L_y^2)^{3/2}}, \tag{1}$$

where L is the Gaussian filtered image, the standard deviation determining the ridge scale. The subscripts indicate partial derivatives, for example L_x is the first partial derivative of L with respect to x and L_{xx} is the second partial derivative with respect to x. κ will be positive for the vessel centre line ridges and negative in the valleys between vessels.

2.2 Classification and feature selection

Fifteen features were calculated for each segment, based on characteristics human observers use to recognise abnormal vessels. They were, briefly: (1) segment length, (2) gradient (the mean Sobel gradient magnitude along the segment), (3) the Sobel gradient variance along the segment, (4) segment direction, (5) tortuosity 1 (sum of absolute changes in direction), (6) tortuosity 2 (maximum difference in angular direction along segment), (7) tortuosity 3 (mean change in angular direction per pixel), (8) grey level (mean segment grey level), (9) grey level variance along segment, (10) distance of segment centroid from disc centre, (11) vessel density, (12) number of segments, (13) mean ridge strength (κ), (14) mean estimated vessel width, (15) mean estimated vessel wall gradient.

A Support Vector Machine (SVM) was chosen as the classifier[1] for its rapid training phase and good classification. All the features were normalised to have zero mean and unit variance. The SVM was trained and tested by cross validation.

In order to be useful, features must discriminate normal and abnormal vessels. Furthermore no two features should be strongly correlated to prevent redundancy. Two non-parametric statistical tests were used to infer discrimination ability: the Wilcoxon rank sum test to test whether the normal and abnormal median feature values differ and the Ansari-Bradley test to determine whether the dispersions of the normal and abnormal values differ. If neither the median nor the dispersion differ significantly then the feature is unlikely to be useful for classification, and indeed could simply add noise and degrade performance. Correlation was tested using the non-parametric Kendall Tau test.

3 Results

From table 1 the top two features according to the Wilcoxon test are 12 and 15. The poorest two features are numbers 4 and 1, where in neither case is the median for normal segments significantly different from that of the abnormal segments. Referring to the Ansari-Bradley test results the dispersion of feature 4 is not significant either, so this feature is unlikely to add any value to the classification. In contrast, the dispersion for feature 1 is significant and so, despite there being no difference in the median values, this feature may still be useful for classification. This was confirmed by leaving out the features one at a time. Performance was degraded in all cases except when feature 4 was excluded, when classification performance improved.

Figure 2 shows the Kendall Tau correlation coefficients for all feature combinations. None of the correlations are particularly strong. Features 2 and 3, and features 5 and 6 have the strongest correlation (greater than 0.5) but leaving any of these features out degrades the classification performance. Figure 3 shows the ROC curve for detection of abnormal segments and abnormal images. Per image performance is better than the per segment performance as abnormal images contain many abnormal segments. Maximum accuracy of 84.4% is achieved at a sensitivity of 84.2% and specificity of 85.9%. An alternative operating point gives a sensitivity of 92.1% and a specificity of 73.2%.

The MATLAB code took 35 seconds on an Intel 5160 Xeon processor (3 GHz) to calculate the fifteen features for each image and classification took less than one second per image.

[1] Chih-Chung Chang and Chih-Jen Lin, LIBSVM: a library for support vector machines. Available from http://www.csie.ntu.edu.tw/~cjlin/libsvm

Wilcoxon rank test		
Feature	Score	p value
12	22.72	2.7×10^{-114}
15	−15.94	3.6×10^{-57}
8	15.37	2.5×10^{-53}
14	−14.60	2.7×10^{-48}
13	−12.23	2.2×10^{-34}
2	9.10	8.8×10^{-20}
3	7.72	1.2×10^{-14}
9	7.46	8.5×10^{-14}
11	6.57	5.1×10^{-11}
7	−4.79	1.7×10^{-06}
5	3.84	0.00012
10	−3.15	0.0016
6	3.11	0.0019
1	−1.76	0.078
4	−0.00	1

Ansari-Bradley test		
Feature	Score	p value
15	−17.66	8.9×10^{-70}
13	−14.48	1.6×10^{-47}
14	−6.81	9.6×10^{-12}
1	−5.95	2.7×10^{-09}
8	−5.60	2.1×10^{-08}
10	−4.62	3.9×10^{-06}
9	4.21	2.5×10^{-05}
2	3.68	0.00023
7	−3.37	0.00076
3	3.12	0.0018
5	2.74	0.0061
4	−1.56	0.12
12	−1.22	0.22
11	0.85	0.4
6	−0.25	0.8

Table 1: Performance of the fifteen feature parameters assessed using the Wilcoxon rank and Ansari-Bradley tests. The most significant scores at the top of the tables.

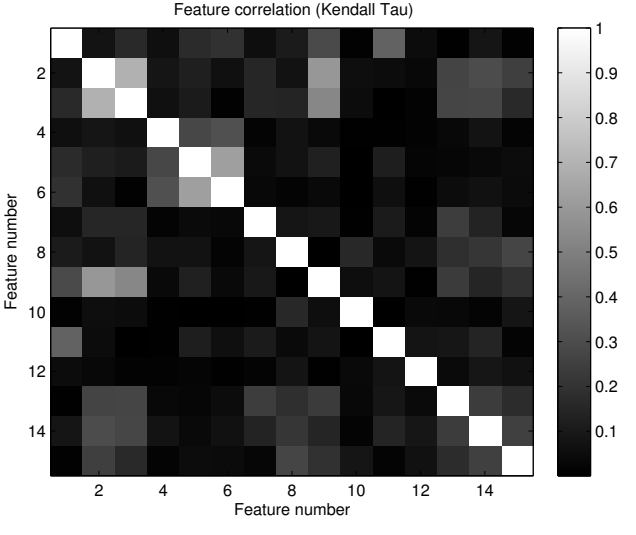

Figure 2: Feature-feature correlation tested using the Kendall Tau test. The brighter the grey level the stronger the correlation. The distribution is symmetrical about the leading diagonal.

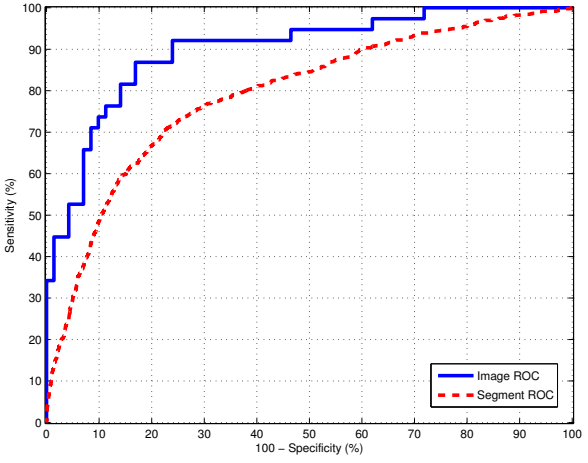

Figure 3: ROC curves using 14 features for abnormal image detection (solid line) and abnormal segment detection (dashed line).

4 Discussion

An automated system for the detection of abnormal vessels on the optic disc has been outlined based on fourteen image features. The features were shown to have good discrimination and low correlation to other features. The area under the ROC curve of 0.911 means that if two images are selected at random, one known to be normal and the other abnormal, the system will classify the abnormal image as the more abnormal 91.1% of the time.

In practice the system could be used in two ways. Firstly as part of a system to detect the most serious retinopathy requiring referral to an eye hospital. Alternatively, since new vessels on the optic disc have the worst prognosis of all the features of proliferative retinopathy, it could be used to automatically triage images, so that those classified as having new vessels could be seen sooner by a human grader.

References

[1] H F Jelinek, M J Cree, J J G Leandro, J V B Soares, R M Cesar Jr, and A Luckie. Automated segmentation of retinal blood vessels and identification of proliferative diabetic retinopathy. *Journal of the Optical Society of America A*, 24:1448–1456, 2007.

[2] T Lindeberg. Scale-space theory: A basic tool for analysing structures at different scales. *Journal of Applied Statistics*, 21:225–270, 1994.

[3] S Philip, A D Fleming, K A Goatman, S Fonseca, P Mcnamee, G S Scotland, G J Prescott, P F Sharp, and J A Olson. The efficacy of automated "disease/no disease" grading for diabetic retinopathy in a systematic screening programme. *British Journal of Ophthalmology*, 91:1512–1517, 2007.