Detection of New Vessels on the Optic Disc Using Retinal Photographs

Keith A. Goatman*, Alan D. Fleming, Sam Philip, Graeme J. Williams, John A. Olson, and Peter F. Sharp

Abstract—Proliferative diabetic retinopathy is a rare condition likely to lead to severe visual impairment. It is characterized by the development of abnormal new retinal vessels. We describe a method for automatically detecting new vessels on the optic disc using retinal photography. Vessel-like candidate segments are first detected using a method based on watershed lines and ridge strength measurement. Fifteen feature parameters, associated with shape, position, orientation, brightness, contrast and line density are calculated for each candidate segment. Based on these features, each segment is categorized as normal or abnormal using a support vector machine (SVM) classifier. The system was trained and tested by cross-validation using 38 images with new vessels and 71 normal images from two diabetic retinal screening centers and one hospital eye clinic. The discrimination performance of the fifteen features was tested against a clinical reference standard. Fourteen features were found to be effective and used in the final test. The area under the receiver operator characteristic curve was 0.911 for detecting images with new vessels on the disc. This accuracy may be sufficient for it to play a useful clinical role in an automated retinopathy analysis system.

I. INTRODUCTION

DIABETIC retinopathy is one of the leading causes of visual impairment in the developed world [1], [2]. In the United Kingdom everyone with diabetes aged twelve and above should be invited, at least annually, for retinal screening using digital retinal photography [3], [4]. The aim of the screening programmes is to detect potentially sight threatening disease sufficiently early to allow timely and effective treatment. Three million people are estimated to have diabetes in the United Kingdom and this figure is predicted to double in the next 15–30 years [5]. There is consequently considerable interest in the potential of automated retinal image analysis to mitigate this projected increase in the screening workload.

*K. Goatman is with the Division of Applied Medicine, School of Medicine and Dentistry, University of Aberdeen, Foresterhill, AB25 2ZD Aberdeen, U.K. (e-mail: k.a.goatman@abdn.ac.uk).

A. Fleming is with the Division of Applied Medicine, School of Medicine and Dentistry, University of Aberdeen, Foresterhill, AB25 2ZD Aberdeen, U.K.

S. Philip, G. Williams, and J. Olson are with the Grampian Diabetic Retinal Screening Programme, David Anderson Building, Foresterhill, AB25 2ZP Aberdeen, U.K.

P. Sharp is with the School of Medical Sciences, University of Aberdeen, Foresterhill, AB25 2ZDAberdeen, U.K.

Diabetic retinopathy causes several types of retinal lesion. The most common lesion, and usually the first sign of retinopathy, is the microaneurysm. These appear in the screening photograph as small red dots. As the disease progresses capillary vessels may begin to leak, forming exudates, lipid deposits that appear as bright yellow/white lesions in the photograph. Larger, dark red blot haemorrhages may also form at this stage. As the disease progresses to its proliferative stage retinal ischaemia can trigger abnormal vessel changes, such as venous beading, intra-retinal microvascular abnormalities (IRMA) and new vessel growth. New vessels are classified according to their position: if they occur on or within one optic disc diameter of the disc they are classified as new vessels at the disc, otherwise they are designated new vessels elsewhere.

There are two main mechanisms by which vision is lost due to diabetic retinopathy: macular oedema and proliferative retinopathy. Macular oedema is the accumulation of fluid in the macula. Although it cannot be seen directly on monoscopic retinal photographs, its presence may be inferred by surrogate indicators such as exudates. Proliferative retinopathy is the more serious condition as it involves the development of new vessels which are prone to bleed, leading to vitreous haemorrhage, fibrosis, and ultimately retinal detachment (Fig. 1). Although the prevalence of proliferative retinopathy is low, typically 0.4% of the screening population [6], the associated risk of the rapid onset of vision loss mean it must be detected reliably. New vessels have a characteristic appearance: they tend to have a narrower calibre and are more tortuous and convoluted than normal vessels. Nevertheless, the differences between normal and abnormal vessels are usually subtle, being towards the limits of the spatial and contrast resolution of the photograph. Vessels which grow out of the focal plane of the photograph can be particularly difficult to identify. Fig. 2 shows some examples of normal and abnormal optic disc vessels.

A number of studies have investigated automatic detection of microaneurysms or haemorrhages [7]–[10], and exudates [11]–[15]. In contrast, little work has been done to detect vessel abnormalities directly, with the exception of venous beading [16], [17]. Jelinek *et al.*, in a study of 27 fluorescein angiogram images, examined vessel characteristics to predict the presence of proliferative disease [18].

In this paper a method for detecting standard screening photographs which show new vessels on the optic disc is described and evaluated. Of all the features of proliferative retinopathy, new vessels at the disc carry the worst prognosis and detection of these is most likely to add value to an automated grading system.
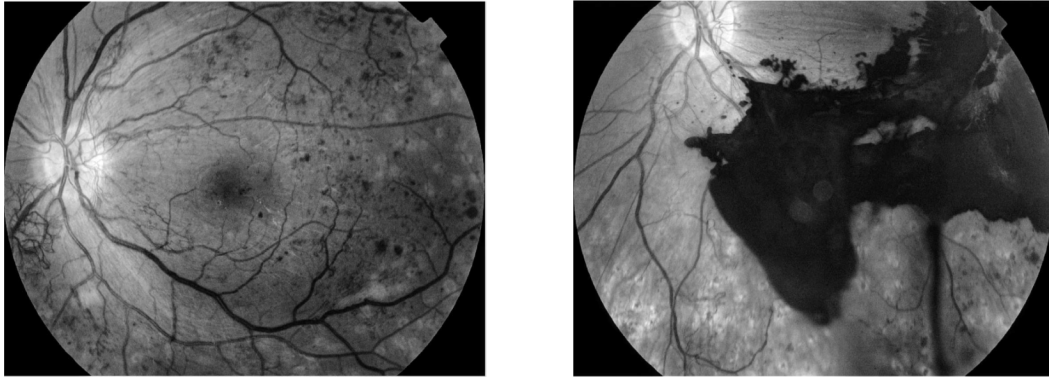
Fig. 1.   Left: Proliferative retinopathy. Right: A slightly different view of the same eye at a later date showing a large, dark haemorrhage.
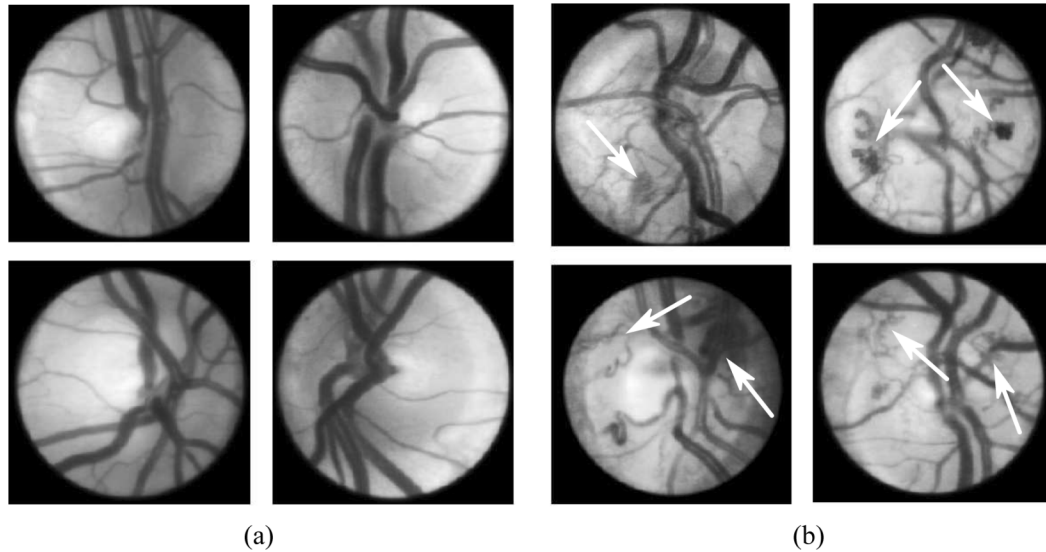
(a)   (b)

Fig. 2.   (a) Examples of (a) normal discs and (b) discs containing abnormal vessels. The white arrows indicate some of the abnormal vessels.

II. METHOD

A. Image Acquisition

Due to the low prevalence of new vessels in the screening population, images were collected from three sources: two retinal screening programmes and a hospital eye clinic, covering a population of approximately 30 000 people with diabetes. A total of 109 images were included in the data set: 38 with confirmed new vessels and a further 71 images without new vessels. Images with inactive new vessels or evidence of previous retinal laser treatment were excluded from the study. All image data were anonymized. The image data from each source were as follows.

1) Grampian Retinal Screening Programme: 28 normal images and seven images with new vessels at the disc acquired with a Canon CR6-NM fundus camera with a Canon EOS-D30 digital SLR camera. 45° nonmydriatic photographs were taken using the highest quality camera JPEG setting.

2) Aberdeen Royal Infirmary Eye Clinic: 32 images showing various pathology but normal discs and 25 images with new vessels at the disc were collected. 50° mydriatic images were acquired from 1) a Topcon TRC-50IX with a Nikon D1H digital SLR having 2000 × 1312 pixels, and 2) a Topcon TRC-50IX with a Nikon D1x digital SLR having 3008 × 1960 pixels. No lossy compression was applied to the images.

3) Tayside Retinal Screening Programme: 11 normal images and 6 with new vessels at the disc were collected. 45° nonmydriatic images were acquired using 1) a Canon CR6-NM fundus camera with an EOS-D60 digital SLR set at 1536 × 1024 pixels and 2) a Canon DGi fundus camera with a Canon EOS-20D digital SLR set to 1728 × 1152 pixels. Images were captured at the highest quality camera JPEG settings.

B. Image Annotation

Two experienced graders (G. J. Williams and S. Philip) annotated the disc vessels which they considered abnormal by drawing freehand on the images using a graphics tablet. Where the new vessels were densely packed or blurred (due to being out of the focal plane of the retina) a region, rather than individual vessels, was annotated. Vessel segments were only considered abnormal where both graders marked them as abnormal.

C. Preprocessing

The green color plane was used in the analysis since it shows the best contrast between the vessels and the background retina.

The grey levels were normalized by stretching the image contrast to cover the full pixel dynamic range, excluding the surrounding dark border pixels and any image labels.

The diameter of a typical optic disc is approximately $1800\,\mu m$. All the images were resized to the same scale prior to analysis, so that the diameter of the optic disc was approximately 300 pixels, making each pixel approximately 6 μm square. The position of the optic disc was located using a method we have described previously [19].

D. Detection of Candidate New Vessels

Several authors have described methods for segmenting normal retinal vasculature, e.g., [19]–[22]. Compared with the normal vasculature abnormal disc vessels are smaller and more tortuous. However, detection is aided by the bright background of the lamina cribrosa, which gives greater vessel/background contrast on the disc, and the deep cup structure of the disc, which means any distractors tend to be out of the vessel focal plane.

The dark ridges formed by the vessel center lines may be detected using the ridge strength (contour curvature), κ, given by [23]

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

where L is the Gaussian filtered image (the standard deviation of the Gaussian function determines the scale of the ridges detected). 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 is positive for ridges such as the vessel center lines, and negative in the valleys between vessels. κ is undefined in regions where the gradient is zero in both the x and y directions.

Applying an empirically derived threshold, κ_{thresh}, to κ allows segmentation of the vessels. However, as Fig. 3(b) shows, using $\kappa_{\text{thresh}} = 0.1$ generates a large number of disjoint candidate segments due to image noise and partially obscured vessels. A hysteresis technique, similar to that used in the Canny edge detector [24], was tried in an attempt to overcome this problem, as shown in Fig. 3(c). However, the method below, based on the morphological watershed transform and shown in Fig. 3(d), was found to be more effective.

The Watershed transform is a morphological region-based segmentation operation. It divides an image into regions based on a topographic map of the image grey level. The dividing lines between hypothetical topographical catchment areas are known as the watershed lines. In a retinal image the dark vessels form topographical valleys and hence the grey level is inverted so that the vessel center lines form the watershed ridges. The inverted image was filtered with a 2-D Gaussian function (with a standard deviation, σ_{ws}, equal to 2 pixels) to prevent over-segmentation. The watershed regions are calculated using Meyer's algorithm [25], implemented in the MATLAB image processing toolbox (Mathworks Inc.). The watershed transform was also used by Walter et al. to detect vessels across the retina [26]. The binary image of the watershed lines is thinned, such that only pixels at vessel bifurcations have more than two neighbors.
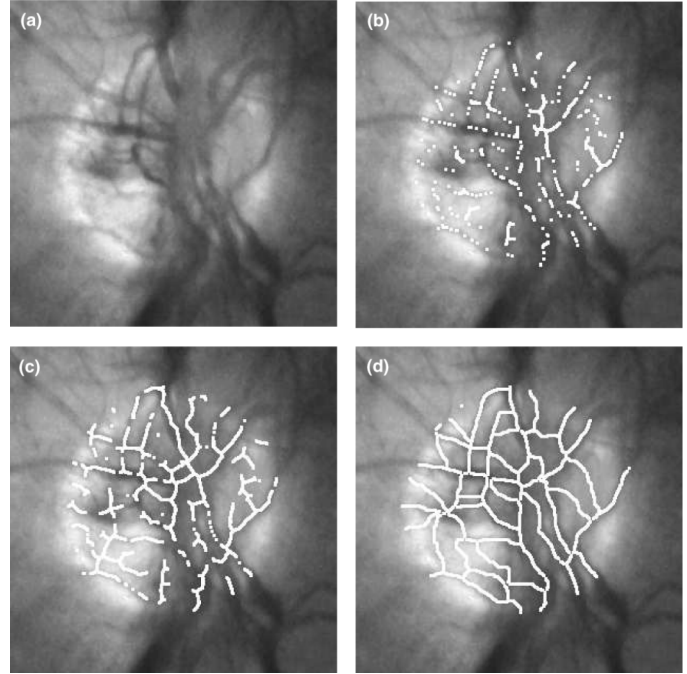

Fig. 3. Example new disc vessel segmentation (shown in white): (a) original image; (b) threshold κ ridge strength; (c) κ threshold with hysteresis; (d) Watershed transform combined with κ threshold.

Candidate segments are separated by removing pixels at bifurcations, with more than two eight-way neighbors. The Watershed transform generates closed regions connected by the watershed lines, not all of which represent vessels. To remove the nonvessel segments the mean value of κ along each candidate segment is calculated and candidates with mean values less than κ_{thresh} are discarded. Small segments, consisting of fewer than seven pixels (i.e., up to 36 μm, roughly the size of a microaneurysm), were also discarded as being too small to classify.

E. Segment Feature Measurements

Fifteen segment features were proposed, based partly on observation of the characteristics human observers use to recognize abnormal vessels, such as their shape, position, orientation, brightness, contrast and line density.

Several feature measures require the angle of the tangent at each point in the segment. Linear regression was used to fit a line through the point itself and its two closest neighbors on either side. The angle was then calculated from the gradient of the fitted line and constrained to lie in the right-hand side of the plane, i.e., $\theta_{\text{tangent}} \in [-\pi/2, \pi/2]$. The angles were later adjusted to avoid discontinuities about $\pm\pi/2$ when tracking segment paths.

As the origin of the major vessels from the disc does not always coincide with the geometric center of the disc, the vessel origin was estimated as follows. Firstly, a 19×19 pixel median filter was applied to remove smaller vessels. Next a threshold was applied to select the darkest 20% of pixels, which were assumed to belong to the major blood vessels. The centroid of the result was taken as the approximate origin of the major vessels.

The following features were calculated for each segment.

1) **Segment length:** The length of the segment in pixels.
2) **Gradient:** The mean gradient magnitude along the segment using the Sobel gradient operator.
3) **Gradient variation:** The standard deviation of the Sobel gradient along the segment. This feature is based on the observation that abnormal vessels are less well defined, being less homogeneous with more contrast variation than normal vessels.
4) **Direction:** The angle between a tangent to the segment center point and a line from its center point to the vessel origin. The feature is based on the observation that normal vessels tend to radiate from the vessel origin towards the edge of the disc, whereas the direction of new vessels is more random. It was calculated using the inverse cosine of the normalized dot product and constrained to $\theta_{\text{direction}} \in [0, \pi]$.
5) **Tortuosity measure 1:** Three measures were assessed to describe the observation that new vessels tend to take more tortuous paths than normal vessels.
   The first measure, T_1, was the sum of the absolute changes in the tangential direction along segment path

$$T_1 = \frac{1}{n-1} \sum_{i=1}^{n-1} |\theta_{i+1} - \theta_i| \qquad (2)$$

where θ_i is the tangential angle at the ith element and n is the number of segment elements.
6) **Tortuosity measure 2:** The second tortuosity measure, T_2, was the difference in the angular extrema of the segment tangents

$$T_2 = \max_{i=1...n} \{\theta_i\} - \min_{i=1...n} \{\theta_i\}. \qquad (3)$$

This measure is insensitive to small changes (jitter) along the segment.
7) **Tortuosity measure 3:** The third tortuosity measure was the mean change in direction per pixel along the segment. This feature is similar to T_1, but is more sensitive to jitter

$$T_3 = \frac{1}{n} \sum_{i=1}^{n} k_i \qquad (4)$$

where

$$k_i = \begin{cases} 1, & \text{if } \begin{cases} \left| \frac{d^2x}{di^2}(i) \right| = 1 \text{ and } \frac{d^2y}{di^2}(i) = 0 \\ \frac{d^2x}{di^2}(i) = 0 \text{ and } \left| \frac{d^2y}{di^2}(i) \right| = 1 \end{cases} \\ \sqrt{2}, & \text{if } \left| \frac{d^2x}{di^2}(i) \right| = 1 \text{ and } \left| \frac{d^2y}{di^2}(i) \right| = 1 \\ 0, & \text{otherwise} \end{cases} \qquad (5)$$

and x and y are the Cartesian coordinates of the ith segment element.
8) **Grey Level:** The normalized mean segment grey level

$$g_{\text{norm}} = \frac{1}{G_{\max} - G_{\min}} \left[\left(\frac{1}{n} \sum_{i=1}^{n} g_i \right) - G_{\min} \right] \qquad (6)$$

where g_i is the grey level of the ith segment pixel and $G_{\max}$ and $G_{\min}$ are the maximum and minimum grey level values in the original image, respectively.
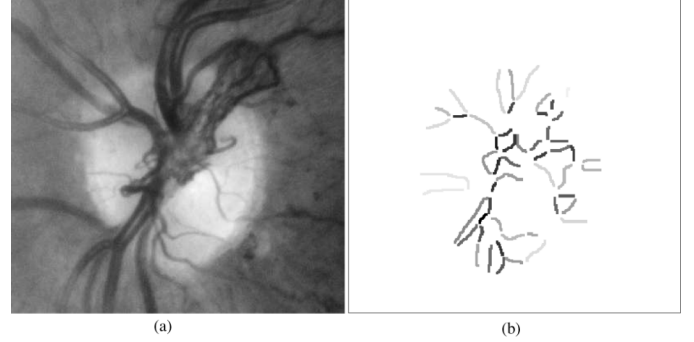

Fig. 4. Example vessel density map: (a) original image; (b) vessel density image, where darker grey levels indicate higher vessel density.

9) **Grey Level coefficient of variation:** This measure was based on the observation that new vessels appear less homogeneous than normal vessels. It is calculated as the ratio of the mean and standard deviation of the segment grey level values.
10) **Distance from vessel origin:** The distance (in pixels) from the center of the segment to the disc vessel origin. This feature was included to test for any positional dependency of the segment within the disc, e.g., whether abnormal vessels are more likely to occur towards the edge of the disc than at the center.
11) **Vessel density:** This feature was based on the observation that new vessels often occur in denser groups than normal vessels. The segments surrounding segment a were determined using

$$O(a) = \{b \mid (s(a) \oplus D) \cap s(b) \neq \emptyset\}$$

where

$$b \in \mathcal{S} \text{ and } b \neq a \qquad (7)$$

where $s(a) \oplus D$ represents the path of segment a dilated by structuring element D, and $\mathcal{S}$ is the set of all segments. A circular disc with a radius of 10 pixels was used as the structuring element in order to include all segments within 10 pixels of any point in the segment. The density was taken as the number of elements in $O(a)$ divided by the length of segment a. An example density map is shown in Fig. 4.
12) **Number of segments:** The total number of segments following candidate segmentation. Images with new vessels tend to have a higher number of segments.
13) **Mean ridge strength:** Mean value of κ, defined in (1), along the segment.
14) **Mean vessel width:** An edge map was generated using the Canny edge detector. The distance from each segment point to the closest edge point is assumed to be the vessel half-width at that point. If the distance is too large it is assumed that the true edge has not been identified, probably due to a very fine vessel, and a width of zero is used.
15) **Mean vessel wall gradient:** As for the mean vessel width above, the closest point on the edge map is assumed to correspond with the vessel wall. The mean value of the

Sobel gradient magnitude for all the vessel wall points is a useful measure of how well defined the vessel wall is.

F. Classification

A support vector machine (SVM) [27], [28] was chosen as the classifier for its rapid training phase and good classification performance. The original SVM algorithm is a linear classifier which finds the best hyperplane separating two classes. However, a kernel function can be used to transform the features to a higher dimensional space. Although the SVM finds a linear hyperplane in the transformed space, the chosen hyperplane is likely to be nonlinear in the original feature space. The kernel function used here was a radial basis function, given by

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2) \quad \gamma > 0 \tag{8}$$

where $\mathbf{x}_i$ and $\mathbf{x}_j$ are the feature vectors for the two classes and γ is a configurable parameter. In addition to γ, a cost or penalty function weight, C, is also configurable.

All the features were normalized before classification, using

$$\hat{f} = \frac{f - m}{s} \tag{9}$$

where f is the feature value to normalize and $\hat{f}$ the normalized value. m and s, respectively, are the mean and standard deviation for this feature. They are calculated from the features used to train the classifier so that all the training features have zero mean and unit variance.

The SVM estimates a probability of abnormality for each vessel segment [29]. For the detection of abnormal images the single segment with the highest abnormality probability was selected and compared with a threshold. Different operating points may be chosen by varying the abnormality score threshold. In this way receiver operator characteristic (ROC) curves may be generated for the new vessel detector. The LIBSVM SVM library was used in this study[1].

G. Detector Training and Testing

The optimal SVM parameters, γ and C, were determined by a cross-validation logarithmic grid search. During training candidate segments were assumed to be part of abnormal vessels if one third or more of their length coincided with the combined manual annotation.

As there were too few images with new vessels for separate training and test sets, the SVM was trained and tested simultaneously by leave-one-out cross-validation. The SVM was trained using all the images in the test set except the single test image, and this process was repeated for each image. The feature value normalization was also recalculated each time, leaving out the test image.

The area under the ROC curve was used as the main measure of detector performance. It was calculated using the Wilcoxon rank sum test [30].

[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

TABLE I
PERFORMANCE OF THE 15 FEATURE PARAMETERS ASSESSED USING THE WILCOXON RANK TEST. FEATURES ARE LISTED IN ORDER OF SCORE, WITH THE HIGHEST ABSOLUTE SCORES AT THE TOP OF THE TABLE

Wilcoxon rank test		
Feature	Score	p value
12	22.72	$< 10^{-6}$
15	-15.94	$< 10^{-6}$
8	15.37	$< 10^{-6}$
14	-14.60	$< 10^{-6}$
13	-12.23	$< 10^{-6}$
2	9.10	$< 10^{-6}$
3	7.72	$< 10^{-6}$
9	7.46	$< 10^{-6}$
11	6.57	$< 10^{-6}$
7	-4.79	1.7×10^{-6}
5	3.84	0.00012
10	-3.15	0.0016
6	3.11	0.0019
1	-1.76	0.078
4	-0.00	1

H. Feature Selection

Not all features are equally useful; some even degrade classification [31]. It is therefore important to determine which features are helping the classification process and which, if any, are hindering it. Given a small feature set, or limitless computing power, a full cross-validation can be performed on every possible combination of the features. It was estimated that a leave-one-out cross-validation of every combination of five or more of the 15 features would take more than three years to complete. A k-fold cross-validation would be faster, but is more likely to suffer bias [32].

A much more rapid assessment of feature relevance was made using a series of heuristic tests. Two nonparametric statistical tests were used to assess the usefulness of each of the 15 features. The first was the Wilcoxon rank sum test, which tests whether the normal and abnormal class median values differ significantly. However, even if the median values do not differ significantly the feature may still be useful if the class distributions differ in shape. This was assessed using the Ansari–Bradley test, which determines if the dispersions of the normal and abnormal class feature values differ significantly following subtraction of the median values. If neither the median nor the dispersion differ between normal and abnormal classes then the feature was *unlikely* to be useful for classification.

III. RESULTS

A. Feature Selection

From Table I the top five features, according to the magnitude of the Wilcoxon test score, are 12, 15, 8, 14, and 13 (which are the total number of segments, mean vessel wall gradient, normalized grey level, mean vessel width and mean ridge strength, respectively).

The poorest two features from the table are numbers 4 and 1, which are the segment direction and length, respectively. In both cases there was no significant difference between the median values in the normal and abnormal classes. Referring to the Ansari–Bradley test results in Table II, for feature 4 there is likewise no difference in the dispersion between the normal and

TABLE II
PERFORMANCE OF THE 15 FEATURE PARAMETERS ASSESSED USING THE
ANSARI–BRADLEY TEST. FEATURES ARE LISTED IN ORDER OF SCORE,
WITH THE HIGHEST ABSOLUTE SCORES AT THE TOP OF THE TABLE

Ansari-Bradley test		
Feature	Score	p value
15	-17.66	$< 10^{-6}$
13	-14.48	$< 10^{-6}$
14	-6.81	$< 10^{-6}$
1	-5.95	$< 10^{-6}$
8	-5.60	$< 10^{-6}$
10	-4.62	3.9×10^{-6}
9	4.21	2.5×10^{-5}
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

Fig. 6. Grid search result to optimise SVM parameters C and γ. The grey level value represents classifier accuracy.
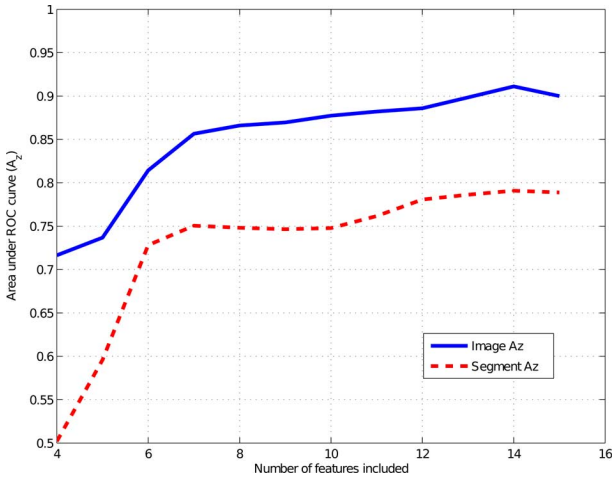

Fig. 5. Performance of the new vessel detector, measured by area under the ROC curve (A_z), on a per segment (dashed line) and per image (solid line) basis versus different numbers of features.

TABLE III
AREA UNDER THE ROC CURVE FOR DIFFERENT SEGMENTATION
PARAMETERS σ_{ws}, THE STANDARD DEVIATION OF THE GAUSSIAN LOW-PASS
FILTER APPLIED BEFORE THE WATERSHED TRANSFORM, AND κ_{thresh},
THE RIDGE STRENGTH THRESHOLD

	σ_{ws} (pixels)					
	0.5	1.0	1.5	2.0	2.5	3.0
0.000	0.623	0.749	0.847	0.888	0.874	0.834
0.100	0.610	0.739	0.848	0.897	0.866	0.830
0.200	0.598	0.759	0.873	0.909	0.874	0.857
0.400	0.548	0.723	0.768	0.748	0.624	0.675

(leftmost column label: κ_{thresh})

C. Segmentation Parameters

Table III shows the variation in detector performance, expressed as the area under the ROC curve, for different segmentation parameters. The peak area of 0.909 is achieved for Gaussian low-pass filtering with a standard deviation, σ_{ws}, of 2 pixels and a ridge strength threshold, κ_{thresh}, of 0.2. Increasing either σ_{ws} or κ_{thresh} results in fewer candidate vessels being detected. Since a vessel cannot be classified as abnormal unless it is first detected in the initial candidate segmentation larger σ_{ws} or κ_{thresh} result in poorer performance due to candidate vessels not being detected. Similarly, decreasing σ_{ws} or κ_{thresh} results in more candidate vessels being detected. One approach would be to accept a low vessel candidate specificity in order to guarantee a very high sensitivity—relying on the classifier to exclude the false positive candidates. However, Table III also shows a fall in performance for smaller σ_{ws} and κ_{thresh} due to over-segmentation, the false positive candidates introducing noise into the classifier.

D. Overall Classification Performance

Fig. 7 shows the result for one of the abnormal discs (a). The candidate vessel segment detection in (c) includes many of the vessels annotated as abnormal in (b). Several of these candidate segments are correctly classified as abnormal in (d). From the ROC plot in Fig. 8 it is clear that the per image performance is better than the per segment performance, since abnormal images like Fig. 7 typically contain many abnormal segments. Maximum accuracy is achieved at an operating point giving a sensitivity of 84.2% and specificity of 85.9%. An alternative op-

abnormal segments. It is therefore unlikely that this feature adds any value to the classification process. However, in contrast, the dispersion of feature 1 does differ significantly between normal and abnormal segments. Hence despite there being no difference in the normal and abnormal segment median values for feature 1, it is still likely be useful for classifying the candidate segments. Fig. 5 shows the new vessel detector performance as area under the ROC curve, A_z, versus number of features included in the classification. The features were added in the order given by the Wilcoxon test scores in Table I. The performance is seen to increase up to 14 features, but fall slightly on adding the final feature (number 4). The top 14 features were therefore used for all subsequent tests and feature 4 was discarded.

B. SVM Parameters

A contour plot for the SVM parameters C and γ is given in Fig. 6, which shows that the classification performance is stable over a range of C and γ values (accuracy range 84%–86.5% for $C \in [2^{-6}, 2^{6}]$ and $\gamma \in [2^{-6}, 2^{0}]$).
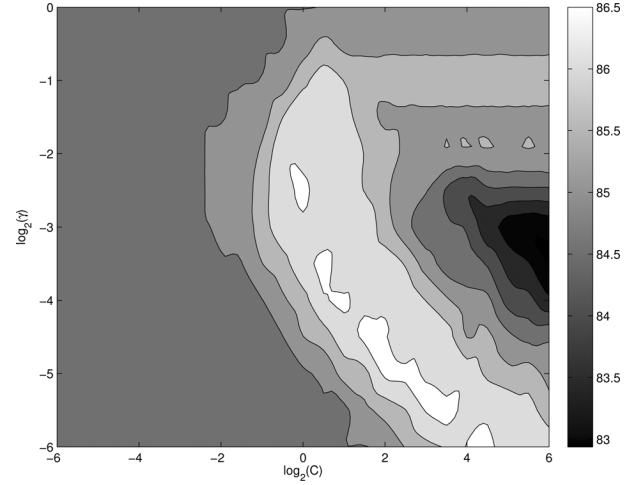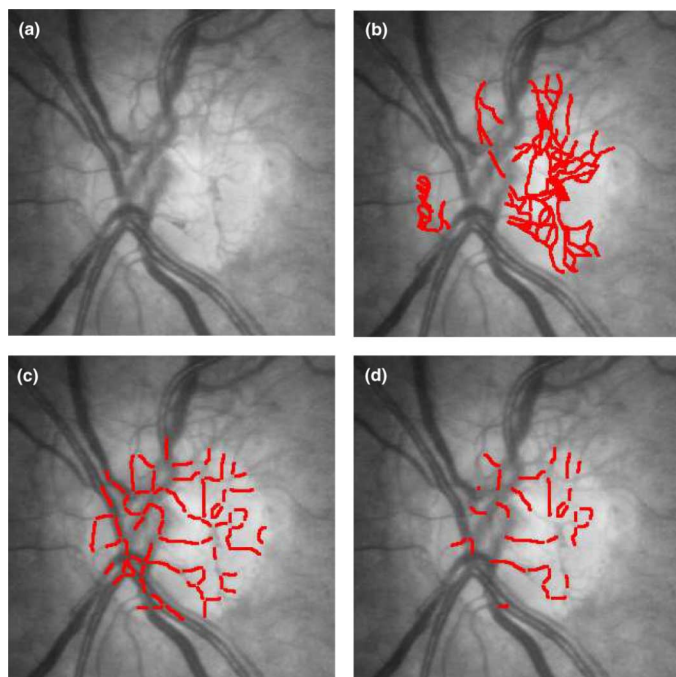
Fig. 7. Example result showing: (a) the original disc image; (b) manual annotation; (c) candidate vessel segmentation; (d) candidates classified as abnormal.
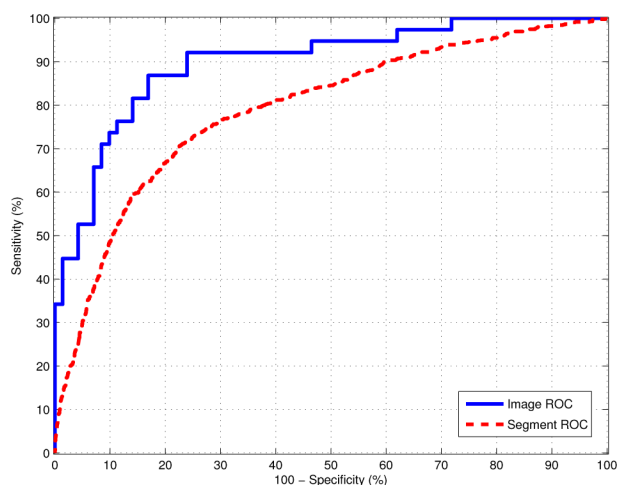

Fig. 8. ROC curves using 14 features for abnormal image detection (solid line) and abnormal segment detection (dashed line).

erating point gives a sensitivity of 92.1% and a specificity of 73.2%.

The MATLAB code took 35 s on an Intel 5160 Xeon processor (3 GHz) to calculate the fifteen features for each image. Classification took less than one second per image. The classifier training phase took 2 min, but this only has to be performed once prior to using the system.

IV. DISCUSSION

In a typical U.K. screening programme approximately two thirds of patients have no visible signs of retinopathy [6] and are simply recalled for screening after twelve months [3]. If all these normal images could be identified automatically it would reduce the manual grading workload by two thirds. As practically all

images with retinopathy contain at least one microaneurysm, detecting the normal images is equivalent to detecting images without microaneurysms. In practice issues such as how to deal automatically with inadequate images must also be addressed. We have shown previously that an automated system, based on image quality assessment and detection of microaneurysms, would reduce the manual grading workload in Scotland and significantly reduce the cost of running the screening programme [33]. Adding a detector for exudates has been shown to improve the detection of maculopathy [34]. Detection of new vessels on the optic disc may similarly improve the detection of proliferative disease. However, the high sensitivity achieved using microaneurysm detection alone, coupled with the low prevalence of new vessels, implies that a very large study would be required to reveal any potential improvement.

Of the one third of patients with visible retinopathy the majority have only microaneurysms. Their subsequent clinical pathway is identical to the patient with no microaneurysms, i.e., rescreening in twelve months. Hence the potential exists to reduce the manual workload still further by automatically screening out images which contain microaneurysms alone; only images with more serious retinopathy would be passed to the manual graders. The difficulty with this approach is that it requires every possible referable feature to be detected reliably for the system to be safe; many of the referable features, such as new vessels and intra-retinal microvascular abnormalities, are difficult to detect, even by eye. Furthermore, as more retinal features are added it becomes increasingly difficult to maintain a low false positive rate. Nevertheless, if good detectors are devised for the other referable features the system described here could have an important role detecting new vessels at the disc.

The most effective strategy at the present time may be to use the additional features to optimize the manual grading queue. This could be performed where there is a delay in grading images, enabling images with suspected proliferative retinopathy to be prioritized and moved up the grading queue. The new vessel detector described here could usefully triage patients: for example, 92.1% of the images with new vessels at the disc would be promoted to the head of the queue, together with 26.8% of the images without new vessels on the disc.

The clinical definition of new vessels at the disc includes new vessels outside the disc but within one optic disc diameter of the disc as well as the new vessels on the disc itself. The system described here detects the new vessels on the disc itself. The new vessels outside the disc are more similar to new vessels elsewhere on the retina than the new vessels on the disc itself, and would be detected by a different detector.

There are a number of areas which merit further investigation. For example, for the new vessels on the optic disc to be detected it is necessary first to correctly locate the disc itself. The method used here is reported to detect the disc successfully in 98.4% of cases [19]. Other methods have reported similarly high detection rates (e.g., [35]). Nevertheless, it would be helpful if the detector calculated a confidence measure that allowed more doubtful disc locations to be passed to a human observer for confirmation. Also, for a clinical system, images would first have to be tested for adequate image clarity [9], [36]. Table III

showed how performance is tied to candidate segmentation performance. This suggests that improvements to the candidate selection have the potential to increase overall performance. Although the features and classifier described are fast and have good accuracy, the potential advantages of other features or classifiers could also be investigated.

This paper has demonstrated an automated system which is able to distinguish normal and abnormal vasculature on the optic disc. It could form part of a system to reduce manual grading workload or a tool to prioritize patient grading queues.

REFERENCES

[1] I. Kocur and S. Resnikoff, "Visual impairment and blindness in europe and their prevention," *Br. J.Ophthalmol.*, vol. 86, pp. 716–722, 2002.

[2] M. A. Bamashmus, B. Matlhaga, and G. N. Dutton, "Causes of blindness and visual impairment in the west of scotland," *Eye*, vol. 18, pp. 257–261, 2004.

[3] K. Facey, E. Cummins, K. Macpherson, A. Morris, L. Reay, and J. Slattery, Organisation of services for diabetic retinopathy screening health technology assessment report 1 (Technical Report Health Technology Board for Scotland Glasgow) 2002.

[4] U.K. Nat. Screening Committee, Essential Elements in Developing a Diabetic Retinal Screening Programme. Workbook Version 4.3 2009.

[5] S. Wild, G. Roglic, A. Green, R. Sicree, and H. King, "Global prevalence of diabetes: estimates for the year 2000 and projections for 2030," *Diabetes Care*, vol. 27, pp. 1047–1053, 2004.

[6] 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," *Br. J. Ophthalmol.*, vol. 91, pp. 1512–1517, 2007.

[7] J. H. Hipwell, F. Strachan, J. A. Olson, K. C. McHardy, P. F. Sharp, and J. V. Forrester, "Automated detection of microaneurysms in digital red-free photographs: A diabetic retinopathy screening tool," *Diabetic Med.*, vol. 17, pp. 588–594, 2000.

[8] M. Niemeijer, B. V. Ginneken, J. Staal, M. S. A. Suttorp-Schulten, and M. D. Abramoff, "Automatic detection of red lesions in digital color fundus photographs," *IEEE Trans. Med. Imag.*, vol. 24, no. 5, pp. 584–592, May 2005.

[9] A. D. Fleming, S. Philip, K. A. Goatman, J. A. Olson, and P. F. Sharp, "Automated assessment of diabetic retinal image quality based on clarity and field definition," *Investigative Ophthalmol. Vis. Sci.*, vol. 47, pp. 1120–1125, 2006.

[10] T. Walter, P. Massin, A. Erginay, R. Ordonez, C. Jeulin, and J. C. Klein, "Automatic detection of microaneurysms in color fundus images," *Med. Image Anal.*, vol. 11, pp. 555–566, 2007.

[11] T. Walter, J. C. Klein, P. Massin, and A. Erginay, "A contribution of image processing to the diagnosis of diabetic retinopathy-detection of exudates in color fundus images of the human retina," *IEEE Trans. Med. Imag.*, vol. 21, no. 10, pp. 1236–1243, Oct. 2002.

[12] A. Osareh, M. Mirmehdi, B. Thomas, and R. Markham, "Automated identification of diabetic retinal exudates in digital colour images," *Br. J. Ophthalmol.*, vol. 87, pp. 1220–1223, 2003.

[13] M. Niemeijer, B. V. Ginneken, S. R. Russel, M. S. A. Suttorp-Schulten, and M. D. Abramoff, "Automated detection and differentiation of drusen, exudates and cotton-wool spots in digital color fundus photographs for diabetic retinopathy diagnosis," *Investigate Ophthalmol. Vis.Sci.*, vol. 48, pp. 2260–2267, 2007.

[14] A. D. Fleming, S. Philip, K. A. Goatman, J. A. Olson, and P. F. Sharp, "Automated detection of exudates for diabetic retinopathy screening," *Phys. Med. Biol.*, vol. 52, pp. 7385–7396, 2007.

[15] C. I. Sánchez, M. García, A. Mayo, M. I. López, and R. Hornero, "Retinal image analysis based on mixture models to detect hard exudates," *Med. Image Anal.*, vol. 13, pp. 650–658, 2009.

[16] P. H. Gregson, Z. Shen, R. C. Scott, and V. Kozousek, "Automated grading of venous beading," *Comput. Biomed. Res.*, vol. 28, pp. 291–304, 1995.

[17] C. W. Yang, D. J. Ma, S. C. Chao, C. M. Wang, C. H. Wen, C. S. Lo, P. C. Chung, and C. I. Chang, "Computer-aided diagnostic detection system of venous beading in retinal images," *Opt. Eng.*, vol. 39, pp. 1293–1303, 1995.

[18] H. F. Jelinek, M. J. Cree, J. J. G. Leandro, J. V. B. Soares, R. M. C. Jr, and A. Luckie, "Automated segmentation of retinal blood vessels and identification of proliferative diabetic retinopathy," *J. Opt. Soc. Am. A*, vol. 24, pp. 1448–1456, 2007.

[19] A. D. Fleming, S. Philip, K. A. Goatman, J. A. Olson, and P. F. Sharp, "Automatic detection of retinal anatomy to assist diabetic retinopathy screening," *Phys. Med. Biol.*, vol. 52, pp. 331–345, 2007.

[20] A. Hoover, V. Kouznetsova, and M. Goldbaum, "Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response," *IEEE Trans. Med. Imag.*, vol. 19, no. 3, pp. 203–210, Mar. 2000.

[21] F. Zana and J. C. Klein, "Segmentation of vessel-like patterns using mathematical morphology and curvature evaluation," *IEEE Trans. Image Process.*, vol. 10, no. 7, pp. 1010–1019, Jul. 2001.

[22] J. J. Staal, M. D. Abr'amoff, M. Niemeijer, M. A. Viergever, and B. V. Ginneken, "Ridge based vessel segmentation in color images of the retina," *IEEE Trans. Med. Imag.*, vol. 23, no. 4, pp. 501–509, Apr. 2004.

[23] T. Lindeberg, "Scale-space theory: A basic tool for analysing structures at different scales," *J. Appl. Stat.*, vol. 21, pp. 225–270, 1994.

[24] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 8, no. 6, pp. 679–698, Nov. 1986.

[25] F. Meyer, "Topographic distance and watershed lines," *Signal Process.*, vol. 38, pp. 113–125, 1994.

[26] T. Walter and J. C. Klein, "Ch. Automatic Analysis of Color Fundus Photographs and its Application to the Diagnosis of Diabetic Retinopathy," in *Handbook of Biomedical Image Analysis*. New York: Kluwer Academic/Plenum, 2005, vol. II, Segmentation Models, pt. B, pp. 315–368.

[27] B. E. Boser, I. Guyon, and V. Vapnik, "A training algorithm for optimal margin classifiers," in *Proc. 5th Annu. Workshop Computat. Learn. Theory*, 1992, pp. 144–152.

[28] V. N. Vapnik, *Statistical Learning Theory*. New York: Wiley, 1998.

[29] T. F. Wu, C. J. Lin, and R. C. Weng, "Probability estimates for multi-class classification by pairwise coupling," *J. Mach. Learn. Res.*, vol. 5, pp. 975–1005, 2004.

[30] J. A. Hanley and B. J. McNeil, "The meaning and use of the area under a receiver operating characteristic (ROC) curve," *Radiology*, vol. 143, pp. 29–36, 1982.

[31] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *J. Mach. Learn. Res.*, vol. 3, pp. 1157–1182, 2003.

[32] T. Hastie, *The Elements of Statistical Learning*, 2nd ed. New York: Springer, 2009.

[33] G. S. Scotland, P. McNamee, S. Philip, A. D. Fleming, K. A. Goatman, G. J. Prescott, S. Fonseca, P. F. Sharp, and J. A. Olson, "Cost-effectiveness of implementing automated grading within the national screening programme for diabetic retinopathy in scotland," *Br. J. Ophthalmol.*, vol. 91, pp. 1518–1523, 2007.

[34] A. D. Fleming, K. A. Goatman, S. Philip, G. J. Williams, G. J. Prescott, G. S. Scotland, P. McNamee, G. P. Leese, W. Wykes, P. F. Sharp, and J. A. Olson, "The role of haemorrhage and exudate detection in automated grading of diabetic retinopathy," *Br. J. Ophthalmol.*, vol. 94, no. 6, pp. 706–711, 2010.

[35] M. Niemeijer, M. D. Abràmoff, and B. van Ginneken, "Fast detection of the optic disc and fovea in color fundus photographs," *Med. Image Anal.*, vol. 13, pp. 869–870, 2009.

[36] M. Niemeijer, M. D. Abràmoff, and B. van Ginneken, "Image structure clustering for image quality verification of color retina images in diabetic retinopathy screening," *Med. Image Anal.*, vol. 10, pp. 888–898, 2006.