Retinal Blood Vessel Segmentation Using Line Operators and Support Vector Classification

Elisa Ricci and Renzo Perfetti*

Abstract—In the framework of computer-aided diagnosis of eye diseases, retinal vessel segmentation based on line operators is proposed. A line detector, previously used in mammography, is applied to the green channel of the retinal image. It is based on the evaluation of the average grey level along lines of fixed length passing through the target pixel at different orientations. Two segmentation methods are considered. The first uses the basic line detector whose response is thresholded to obtain unsupervised pixel classification. As a further development, we employ two orthogonal line detectors along with the grey level of the target pixel to construct a feature vector for supervised classification using a support vector machine. The effectiveness of both methods is demonstrated through receiver operating characteristic analysis on two publicly available databases of color fundus images.

Index Terms—Classification, line detector, retinal imaging, support vector machines (SVM), vessel segmentation.

I. INTRODUCTION

DIGITAL fundus imaging in ophthalmology plays an important role in medical diagnosis of several pathologies like hypertension, diabetes, and cardiovascular disease [1]. Computer-aided image analysis of the eye fundus is highly desirable in many cases. For example, the diagnosis of diabetic retinopathy, the leading cause of blindness in the Western World, requires the screening of a large number of patients from specialized personnel and can be extremely facilitated with the adoption of automatic tools [2].

One task of the utmost importance is the segmentation of the vasculature in retinal fundus images. Several morphological features of retinal veins and arteries, like diameter, length, branching angle, and tortuosity, have diagnostic relevance and can be used to monitor the progression of diseases [3]. Moreover blood vessels, being invariant features, may be very useful for registration of retinal images of the same patient gathered from different sources [4]. Conversely, in some applications the vasculature must be excluded from the analysis to ease automatic detection of pathological elements like exudates or microaneurysms [5]. Finally, its tree-like geometry makes it also useful to determine the position of other features as fovea or optic disk [6]. Accurate vessel segmentation is a difficult task for several reasons: the presence of noise, the low contrast between vasculature and background, the variability of vessel width, brightness, and shape. Moreover, due to the presence of lesions, exudates, and other pathological effects, the image may have large abnormal regions.

Several methods have been proposed in the literature to address these problems including matched filtering [7], tracking methods [8]–[10], multithreshold probing [11], mathematical morphology [12], or a combination of nonlinear filtering and morphological operations [13].

Interesting results have been obtained by pixel classification based on supervised learning. The goal of classification approaches is to assign each pixel to one of two classes, i.e., *vessel* and *non-vessel*, based on some features extracted from the image in the neighborhood of the considered pixel. In [14], a neural network is trained by the backpropagation learning algorithm. The feature vectors are constructed using the first principal component values in a 10×10 window centered on the pixel being classified and the corresponding edge strength values in the same window (200 features). In [15], the outputs of multiscale Gaussian filters and the original grey level of the green channel image (31 features) are given as input to a k-nearest neighbor classifier. Better results have been obtained in [16] where ridge profiles are extracted from which 27 features are computed. Then, a feature selection scheme is applied to select those corresponding to the better class separability. In [17], a multiscale analysis is performed on the image using the Gabor wavelet transform. Only five features are considered: the grey level of the inverted (green plane) image and the maximum Gabor transform response over angles at four different scales. For supervised classification a Gaussian mixture model (GMM) Bayesian classifier is adopted.

Some of the above-mentioned methods are based on the detection of linear structures through angular scanning. In [7], the image is convolved with 12 rotated versions of a two-dimensional matched filter impulse response ($15°$ of angular resolution), retaining the maximum response with respect to the angle. A similar approach is adopted in [17] where the orientation spans from 0 to $170°$ at steps of $10°$, and in [13] where only four orientations are used. In [12], linear structures at 12 different orientations are detected by computing the supremum of openings using a linear structuring element. Suppression of undesirable patterns is then performed using the Laplacian and an original alternating filter. The final result is thresholded.

In this paper, we explore a computationally simpler but more effective realization of the same concept, inspired by a paper of Zwiggelaar *et al.* [18] about the detection of linear structures in mammographic images. Among the techniques compared in [18], we select the line operator introduced in [19], which has been modified to take into account the peculiarities of retinal vessel structure. The resulting image is thresholded

E. Ricci is with the Department of Electronic and Information Engineering, University of Perugia, I-06125 Perugia, Italy.

*R. Perfetti is with the Department of Electronic and Information Engineering, University of Perugia, I-06125 Perugia, Italy (e-mail: perfetti@diei.unipg.it).

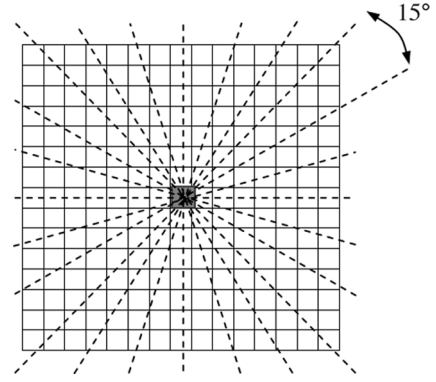Fig. 1. Inverted green channel of non-mydriatic retinal image from DRIVE database.

Fig. 2. Twelve orientations to evaluate line strength of shaded pixel.

Fig. 3. (a) Basic line detector. (b) Line detector with its orthogonal line.

to obtain unsupervised binary pixel classification. Then, to further improve the performance, a supervised method is proposed where two line detectors are used to compute the feature vectors, while a linear support vector machine (SVM) [20] is chosen as a classifier. With respect to previous supervised techniques, the proposed method exhibits the following desirable properties: it requires less features, feature extraction is computationally simpler, and fewer examples are needed for training.

The performance of both methods is evaluated on the publicly available STARE [21] and DRIVE [15] databases through receiver operating characteristic (ROC) analysis. From the area under the ROC curve and from the accuracy, we can see that our approach, despite its simplicity, matches or slightly outperforms the most recent methods.

The rest of the paper is organized as follows. In the next section we illustrate the proposed methods of vessel segmentation. In Section III, we review the main properties of SVMs. In Section IV, the experimental results are presented and compared to several existing methods. Some comments conclude the paper.

II. PROPOSED METHODS

In the analysis of red-green-blue (RGB) non-mydriatic images, usually the color components are considered separately because the green channel exhibits the best vessel/background contrast while the red and blue ones tend to be very noisy. Therefore, we work on the inverted green channel images, where vessels appear brighter than the background. An example is shown in Fig. 1. To preserve at most the vessel structure we avoid any preprocessing on the image. Indeed, low-pass filtering makes vessel boundaries less sharp and causes loss of thinnest vessels. The contrast enhancement suggested in [14] has been tried but it was discarded due to the increased amount of noise. Also, experiments with lighting and contrast equalization have been performed but without any improvement. Details on these experiments will be given in Section IV.

The basic line detector is illustrated in Fig. 2. The average grey level is evaluated along lines of fixed length l passing through the target pixel (i, j) at different orientations. We consider 12 angles (15° of angular resolution). The line with
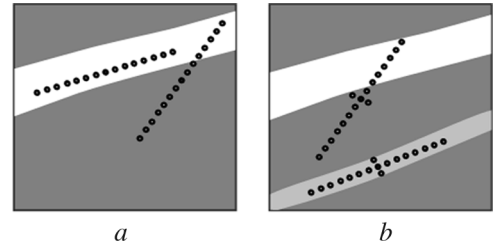
the largest average grey level is found; its value is denoted with $L(i, j)$. The difference $S(i, j) = L(i, j) - N(i, j)$ represents the *line strength* of the pixel [18], where $N(i, j)$ is the average grey level in the square window, centered on the pixel, with edge length equal to l. The line strength is large if the winning line is aligned within a vessel; otherwise, we have a partial overlap and the line strength is lower [Fig. 3(a)]. This difference allows us to discriminate vessel pixels from non-vessel pixels. Lines of different lengths have been considered; the best performance was obtained with $l = 15$ pixels (see Section IV for a comparison of different lengths). This result is in accordance with [12] where a line operator of length 15 pixels is used as structuring element to perform morphological operations.

Conversely to [18], the square window is not oriented like the winning line but it is kept fixed. Moreover, the average grey level of lines is not obtained by interpolation, but the pixels to be averaged are found by rounding the coordinates of the points on the ideal line. The approximated lines for 15°, 30°, 60°, and 75° are shown in Fig. 4. Lines with angles −15°, −30°, −60°, and −75° are turned right-left with respect to those shown in Fig. 4. Lines with angles 0°, 90°, and ±45° are exact. Simulation results have demonstrated that both simplifications do not affect the segmentation performance while computational cost is significantly reduced.

There is a certain resemblance between the proposed line detector and the "tramline" filter proposed in [22], so we point out the differences between the two schemes. The tramline filter consists of three parallel lines of nine pixels, where the outer lines are displaced five pixels on each side of the inner line. The filter response is the difference between the minimum pixel intensity in the outer lines and the maximum intensity level in the inner line (with vessels darker than the background). The filter is rotated through 12 angles to select the maximum response. It
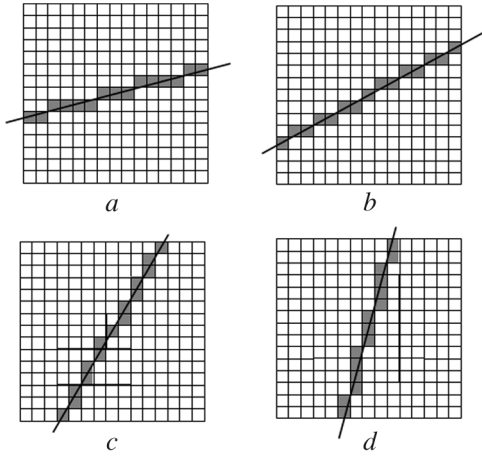
Fig. 4. Lines of 15 pixels with orientation: (a) 15°, (b) 30°, (c) 60°, and (d) 75°.
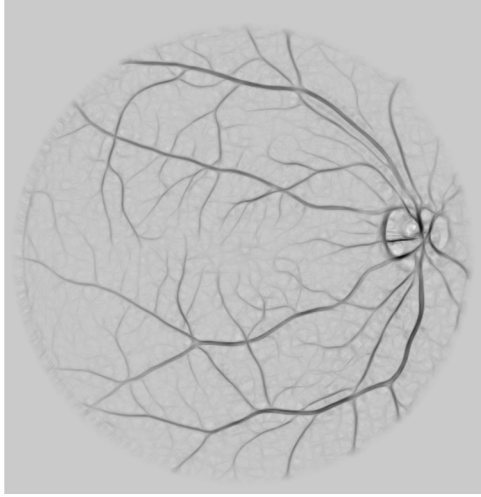

Fig. 5. Inverted line strength of image in Fig. 1.

gives a strong response when the inner line is aligned within a blood vessel and the outer lines are on either side of the vessel. Thus, it is suited to detect vessel centerlines, but not for vessel segmentation. When it is applied to large vessels, only the innermost pixels are emphasized. Moreover, the proposed line detector has a lower computational cost and it is less sensitive to noise due to averaging.

As can be seen in Fig. 1, the eye fundus images have a roughly circular field of view (FOV) surrounded by a bright frame. When the line detector is applied to the pixels inside the FOV near the border, the line strength is strongly biased by the external bright region. As a consequence, false vessel detection occurs along the borderline. To overcome this problem, applying the proposed method to border pixels, the out-of-the-FOV grey levels in the 15×15 square centered on the target pixel are replaced by the average grey level of the remaining pixels in the square. Identification of the FOV is not necessary for the database DRIVE where it is given. We identified the FOV only for STARE using routine morphological operations. In Fig. 5, the line strength image corresponding to Fig. 1 is shown. For better visualization the image has been inverted.

The basic line detector described above can be improved for retinal vessel segmentation, since the vessels can have different width and intensity. In proximity of a large and bright vessel, the line strength of a pixel can be comparable with that of the pixels inside a darker and thinner vessel. A possible way to overcome this problem is shown in Fig. 3(b). A line of three pixels is considered, centered on the midpoint of the main line and orthogonal to it. Its average grey level is denoted with $L_o(i,j)$. Its strength is obtained again subtracting the average intensity in the 15×15 square, i.e., $S_o(i,j) = L_o(i,j) - N(i,j)$. The angle of the orthogonal line is approximated with the closest principal angle ($0°$, $45°$, $90°$, $135°$). For example the "orthogonal" direction for the orientations $30°$, $45°$, and $60°$ is $135°$.

The strength of the shorter line will be negligible for out-of-the-vessel pixels, while it is comparable to the main line strength for inside pixels. This additional information helps to discriminate between inside and outside pixels, decreasing the number of false positives. Moreover, we consider the grey level $I(i,j)$ of the pixel as a third feature; we note that it helps to reduce false detection due to pathology or to the proximity of the optic disk. Then, we construct the feature vector $\mathbf{x} = [S(i,j)\ S_o(i,j)\ I(i,j)]$, used to train a supervised classifier. A normal transformation is applied to the feature vector [23]

$$\overline{x}_i = \frac{x_i - \mu_i}{\sigma_i} \tag{1}$$

where x_i is the ith feature ($i = 1, 2, 3$), while μ_i and σ_i are the corresponding mean value and standard deviation, respectively. Normalization is applied separately to each image in order to compensate for illumination variation between the images. For supervised binary classification we adopt an SVM because of its well-known good generalization [20]. We carried out experiments with both linear and nonlinear SVMs. However, the results do not demonstrate any improvement of nonlinear SVM with respect to the linear SVM. Since the latter is much less demanding in terms of computation, it seems more suited to this application. Hence, in the paper we present only the linear case. A review on linear SVMs is given in the following section.

III. SVMs

Let $(\mathbf{x}_i, y_i)$, $i = 1, \ldots, N$ represent the training examples of the classification problem, where $\mathbf{x}_i \in \mathrm{R}^d$ and $y_i \in \{-1, +1\}$. In the case of linearly separable classes, there exists a separating hyperplane given by

$$\mathbf{w}^T \mathbf{x} + b = 0 \tag{2}$$

where $\mathbf{x}$ is the feature vector, $\mathbf{w}$ is the weight vector, and b is the bias. We may thus write

$$y_i(\mathbf{w}^T \mathbf{x}_i + b) > 0 \quad i = 1, \ldots, N. \tag{3}$$

The distance between the separating hyperplane and the closest points is called *margin*. The support vector machine identifies the *optimal separating hyperplane* (OSH) for which the margin of separation is maximized. We can rescale the

weights $\mathbf{w}$ and the bias b so that the constraints (3) can be rewritten as

$$y_i(\mathbf{w}^T\mathbf{x}_i + b) \geq 1 \quad i = 1,\ldots,N. \tag{4}$$

As a consequence, the margin of separation is $1/\|\mathbf{w}\|$ and the maximization of the margin is equivalent to the minimization of the Euclidean norm of the weight vector $\mathbf{w}$. The corresponding weights and bias represent the OSH.

Introducing the Lagrange multipliers $\alpha_1 \ldots \alpha_N$, the minimization of $\|\mathbf{w}\|^2$ under constraints (4) can be recast in the following dual form [20].

Minimize

$$J(\alpha) = \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j y_i y_j \mathbf{x}_i^T \mathbf{x}_j - \sum_{i=1}^{N} \alpha_i \tag{5}$$

subject to the linear constraints

$$\sum_{i=1}^{N} \alpha_i y_i = 0 \tag{6a}$$

$$\alpha_i \geq 0 \quad i = 1,\ldots,N. \tag{6b}$$

There is a Lagrange multiplier α_i for every training point. Those points $\mathbf{x}_i$ corresponding to $\alpha_i > 0$ are called *support vectors* and lie closest to the OSH.

Solving the quadratic programming (QP) problem (5)–(6) we can compute the optimum weight vector $\mathbf{w}$ as follows:

$$\mathbf{w} = \sum_{i=1}^{N} \alpha_i y_i \mathbf{x}_i. \tag{7}$$

For a support vector $\mathbf{x}_i$, it is $\mathbf{w}^T\mathbf{x}_i + b = y_i$, from which the optimum bias b can be computed by substitution of (7)

$$b = y_i - \sum_{k=1}^{N} \alpha_i y_i \mathbf{x}_i^T \mathbf{x}_k. \tag{8}$$

In practice, it is better to average the values obtained by considering the set of all support vectors. The class of a new point $\mathbf{x}$ is given by the sign of the following decision function:

$$f(\mathbf{x}) = \mathbf{w}^T\mathbf{x} + b. \tag{9}$$

SVM formulation can be extended to the case of linearly non-separable data, introducing the slack variables ξ_i; optimal class separation can be obtained by minimization of

$$\frac{1}{2}\mathbf{w}^T\mathbf{w} + C \sum_{i=1}^{N} \xi_i \tag{10}$$

under constraints

$$y_i(\mathbf{w}^T\mathbf{x}_i + b) \geq 1 - \xi_i, \quad i = 1,\ldots,N \tag{11a}$$

$$\xi_i \geq 0 \quad i = 1,\ldots,N. \tag{11b}$$

In this case, the margin of separation is said to be *soft*. The constant $C > 0$ is user-defined and controls the tradeoff between the maximization of the margin and the minimization of classification errors on the training set. The dual formulation of the QP is the same as (5)–(6) with the only difference in the bound constraints

$$0 \leq \alpha_i \leq C \quad i = 1,\ldots,N. \tag{12}$$

In this case, for the Lagrange multipliers we have: $0 \leq \alpha_i < C$ for points correctly classified with $\xi_i = 0$ and $\alpha_i = C$ for points with $\xi_i > 0$ (misclassified if $\xi_i > 1$). In the soft margin case, (7) and (8) still hold.

Often, in many practical applications, it is important to distinguish between the two types of errors that can arise in a binary classification task, i.e., false positives and false negatives. This is particularly true for medical applications where performance is generally evaluated in terms of sensitivity (ratio of the number of true positives and the number of positive instances) and specificity (ratio of the number of true negatives and the number of negative instances).

One possible way to face this problem with SVMs is based on the idea of assigning different weights to the errors of the two classes. This leads to a slightly different formulation of the primal problem which becomes [24] the following.

Minimize

$$\frac{1}{2}\mathbf{w}^T\mathbf{w} + C^+ \sum_{i|y_i=+1} \xi_i + C^- \sum_{i|y_i=-1} \xi_i \tag{13}$$

under constraints

$$y_i(\mathbf{w}^T\mathbf{x}_i + b) \geq 1 - \xi_i, \quad i = 1,\ldots,N \tag{14a}$$

$$\xi_i \geq 0 \quad i = 1,\ldots,N. \tag{14b}$$

Therefore, also the dual formulation of the QP is the same of (5)–(6) but with the Lagrange multipliers α_i constrained as follows:

$$0 \leq \alpha_i \leq C^+ \quad \text{if } y_i = +1 \tag{15a}$$

$$0 \leq \alpha_i \leq C^- \quad \text{if } y_i = -1. \tag{15b}$$

To solve the QP problem in its dual formulation, in our experiments we used SVM*light* [25][1] which allows us to deal with unsymmetric cost factors. Here, varying the ratio $r = C^+/C^-$, different possible tradeoffs between the cost of false positives and false negatives can be obtained. In practical applications [26], often the parameters C^+ and C^- are chosen such that the cost of the false positives equals the cost of the false negatives, that is $r = n^-/n^+$, where n^+ and n^- are, respectively, the number of positive and negative examples.

IV. EXPERIMENTAL RESULTS

A. Databases

We use two publicly available databases of non-mydriatic color images. The first one, known as STARE, was collected by

[1]Software, available at http://svmlight.joachims.org/.

Hoover *et al.* [21].[2] It consists of 20 images (ten with pathologies) captured by a TopCon TRV-50 fundus camera at 35° FOV. The slides were digitized to 700 × 605 pixels with 8 bits per color channel. The FOV in the images is approximately 650 × 550 pixels. All the images were manually segmented by two observers. The first observer marked 10.4% of pixels as vessel, the second one 14.9%. Performance of the proposed method is evaluated with the segmentation of the first observer as ground truth.

The second database was collected by Niemeijer *et al.* [15] with the aim of comparing the performance of different segmentation methods. This database is known as DRIVE and it is publicly available.[3] It consists of 40 images (seven with pathologies) captured by a Canon CR5 3CCD camera with a 45° FOV. The images are of size 768 × 584 pixels with 8 bits per color channel. The FOV is circular with approximately 540 pixels in diameter. For each image, a mask image is provided that delineates the FOV. Hence, detection of the FOV border is not needed in this case. The images are compressed in JPEG format, which is a common practice in screening programs, and are divided into a training set and a test set, each containing 20 images. The test set has four images with pathology. The training set is useful to design supervised segmentation methods. All the images were manually segmented. Those of the test set were segmented twice, resulting in a set A and a set B. In set A, 12.7% of pixels were marked as vessel, against 12.3% for set B. Performance is evaluated on the test set using the segmentations of set A as ground truth.

B. Experimental Evaluation

It is common practice to evaluate the performance of retinal vessel segmentation algorithms using ROC curves [15]. An ROC curve plots the fraction of vessel pixels correctly classified as vessel, namely the true positive rate (TPR), versus the fraction of nonvessel pixels wrongly classified as vessel, namely the false positive rate (FPR). The closer the curve approaches the top left corner, the better is the performance of the system. The most frequently used performance measure extracted from the ROC curve is the value of the area under the curve (AUC) which is 1 for an optimal system. For both databases, TPR and FPR are computed considering only pixels inside the FOV. The ROC curve of the basic line detector can be obtained thresholding the line strength at different levels.

As concerns the proposed supervised method, we first select a training set to build the SVM classifier and then evaluate the performance on a test set. For the database STARE, training is performed using 20 000 manually segmented pixels randomly extracted from the 20 images (1000 pixels per image). Due to the small size of the training set (<0.5% of the entire database), classification performance is evaluated on the whole set of 20 images. For the database DRIVE 20 000 manually segmented pixels are randomly extracted from the available training set of 20 images (1000 pixels per image) and are used to train the classifier. Classification performance is then evaluated on the 20 images of the test set.
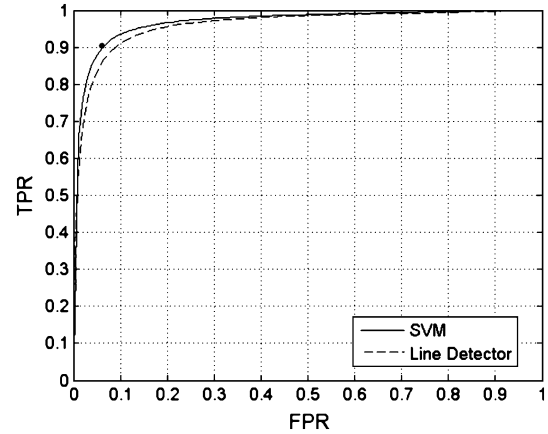
Fig. 6. ROC curves for STARE database: line detector (dashed line) and linear SVM (solid line). Manual segmentation of second observer compared to first (●).
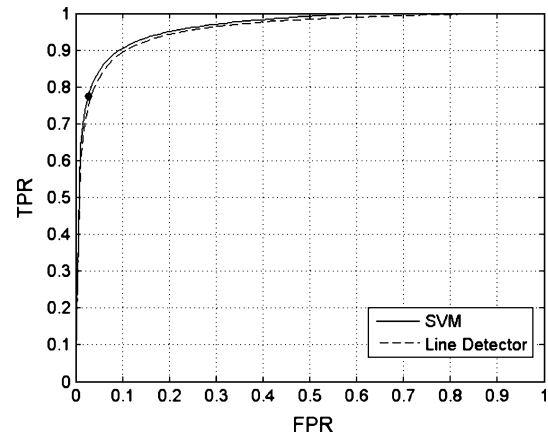

Fig. 7. ROC curves for DRIVE database: line detector (dashed line) and linear SVM (solid line). Manual segmentation of set B compared to set A (●).

For linear SVMs the ROC curve can be obtained by varying the threshold b as a free parameter [27]. To this end, the value of b obtained with (8) is considered as a starting point; from there if b diminishes, the TPR and FPR decrease, while increasing it the TPR and FPR approach to unity (vessel pixels correspond to class +1). It is worth noting that the ratio $r = C^+/C^-$ is unimportant in ROC analysis since varying r is equivalent to moving on the ROC curve. However, also in this case, we set r to 7 which is roughly equal to the ratio n^-/n^+. The parameter C^- is determined by cross-validation on the training set using the AUC as performance measure. For both databases the value $C^- = 1$ corresponds to the best generalization performance. However, experiments have shown that the classification error is almost insensitive to the value of C^- in the range 0.1, 10.

C. Results

The ROC curves for the two databases are depicted in Figs. 6 and 7: the solid line represents the results of segmentation with the linear SVM, while the dashed line corresponds to the line detector with line length $l = 15$. The dot shows the performance of manual segmentation: for the STARE database, the second observer gives FPR = 0.061 and TPR = 0.903, while for the DRIVE database results of set B are FPR = 0.0275 and

TABLE I
COMPARISON OF LINE LENGTH—STARE

l	AUC	ACCURACY
17	0.9582	0.9583
15	0.9602	0.9584
13	0.9523	0.9582
11	0.9448	0.9574

TABLE II
COMPARISON OF LINE LENGTH—DRIVE

l	AUC	ACCURACY
17	0.9555	0.9562
15	0.9558	0.9563
13	0.9454	0.9562
11	0.9492	0.9561
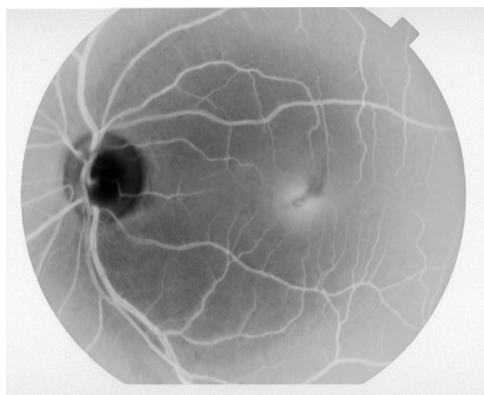
Fig. 8. Inverted green channel of retinal image from STARE database.

Fig. 9. Segmentation of image in Fig. 8 obtained using line detector.

a

b

Fig. 10. Segmentation in presence of a central reflex: (a) patch from STARE database with a central reflex and (b) segmentation obtained using line detector (threshold is set to 15).

TPR $= 0.775$. In Tables I and II, we show a comparison of performance for the basic line detector with varying line length l, in terms of AUC and accuracy. The accuracy is the maximum average accuracy at a certain "optimum" threshold. For both databases $l = 15$ seems the best choice; this value has been used in all the experiments presented in the following. However, we can see that the value of l is not critical since the performance slowly varies with it. Moreover, the line length can be easily tuned to images with different resolution.

In Fig. 8, the inverted green channel is shown for an image with nonuniform illumination from the STARE database. The segmentation obtained with the basic line detector is shown in Fig. 9. This result corresponds to the best accuracy (0.9630). Even if the brightness of the image increases radially outwards, the performance is quite satisfactory both in the darker central region and in the brighter outermost one. This is due to the local differential computation of the line strength which makes the line detector somewhat robust with respect to nonuniform illumination and contrast.

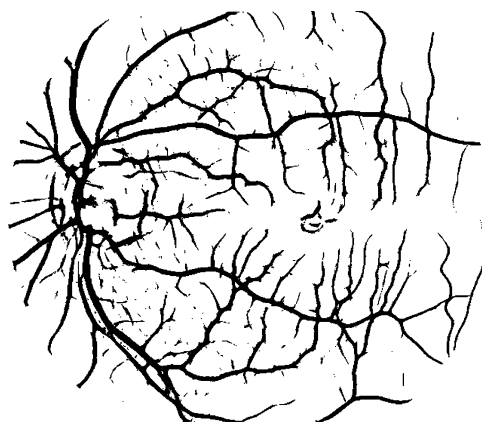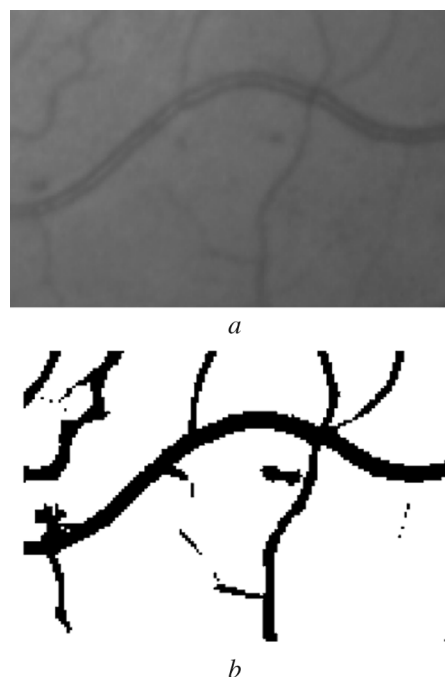We also tested two equalization techniques to ascertain their effect on the line detector behavior: the lighting compensation proposed by Hoover *et al.* [6] and the method of Foracchia *et al.* [28]. Both methods rely on a window size, so different values were tried to get the best performance. However, when the line detector is combined with equalization the accuracy of the detection is generally worse. For the sake of comparison, we consider again the image in Fig. 8. Using the lighting compensation in [6] the maximum accuracy was 0.9419 (49×49 window), while using the method of [28] the maximum accuracy was 0.9221 (80×80 window). The explanation for this reduced performance comes from the fact that these equalization techniques not only enhance vessels but also other structures in low contrast regions, increasing the effect of noise and the false detections of non-vessel pixels.

In Fig. 10, we show an example of segmentation concerning a vessel with "central reflex" (a brighter strip along the center line present sometimes in wider vessels). In (a), a patch is shown
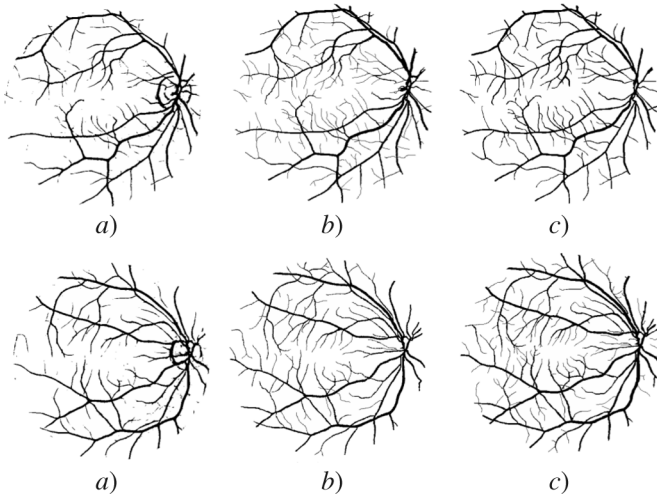
Fig. 11. Example of segmentation results for two images of DRIVE database. (a) Vasculature reconstructed with linear SVM. (b) Manual segmentation of observer A. (c) Manual segmentation of observer B. Second image corresponds to Fig. 1.

TABLE III
COMPARISON OF SEGMENTATION METHODS—STARE

	AUC	ACCURACY
Line detector	0.9602	0.9584
Orthogonal lines - LMSE	0.9615	0.9595
Orthogonal lines - SVM	**0.9680**	**0.9646**
Soares at al. [17]	0.9671	0.9480
Mendonça et al. [13]	-	0.9479
Staal et al. [16]	0.9614	0.9516
Jiang et al. [11]	0.9298	0.9009
Hoover [21]	0.7590	0.9275
Second observer	-	0.9351

TABLE IV
COMPARISON OF SEGMENTATION METHODS—DRIVE

	AUC	ACCURACY
Line detector	0.9558	0.9563
Orthogonal lines - LMSE	0.9582	0.9577
Orthogonal lines – SVM	**0.9633**	**0.9595**
Soares at al. [17]	0.9614	0.9466
Mendonça et al. [13]	-	0.9463
Staal et al. [16]	0.9520	0.9441
Chaudhuri et al. [7]	0.7878	0.8773
Matched filter k-NN [15]	0.9294	0.9416
Jiang et al. [11]	0.9327	0.8911
Second observer	-	0.9473

from the STARE database where a central reflex is evident. In (b), the segmentation obtained by the line detector is shown (the accuracy on the patch is 0.9572). Generally, the line detector behavior in presence of a central reflex is quite satisfactory. The reason is that the winner line rotates with respect to the vessel center line, so the line strength is not much affected by the central strip intensity. When the central reflex is very large it may happen that the line detector splits the vessel into parallel lines, but this occurrence is very rare.

Fig. 11 shows two segmented images of the DRIVE database obtained using the linear SVM. In both cases the result corresponding to the best accuracy is considered: 0.9604 for the first image and 0.9703 for the second. The associated values of (TPR, FPR) are (0.7283, 0.0168) and (0.8346, 0.0175), respectively. The manual segmentations of the two observers A and B are also depicted.

Tables III and IV compare our approach with the most recent methods in terms of AUC and accuracy. The performance for the methods of Jiang *et al.* [11] and Hoover *et al.* [21] are taken from [16]. The results for the matched filter method of Chaudhuri *et al.* [7] and the matched filter based pixel classification are taken from [15]. As concerns the methods of Mendonça *et al.* [13] and Soares *et al.* [17], several variants have been proposed by the authors, so we report only the best results (for the method in [13] the AUC is not available).

To summarize the results, we can say that the performance of the basic line detector, despite its simplicity, is comparable to or slightly better than some of the existing segmentation methods. The addition of the other two features and the use of the SVM classifier make our approach superior considering both AUC and accuracy. Even if the improvement is in the order of hundredths, this difference can be meaningful due to the huge number of pixels being classified. For both databases, the results obtained with a linear SVM have been also compared to those of a linear minimum squared error (LMSE) classifier. Even in this case the ROC curve is obtained varying the threshold b as a free parameter. From Tables III and IV, it is easy to see that

linear SVM clearly outperforms the LMSE, trained on the same training sets (20 000 data points), considering both AUC and accuracy. From our experiments we found that LMSE needs larger training sets to reach the same performance of SVM, as expected, due to the good generalization performance of the latter.

To ascertain the sensitivity of the SVM performance with respect to the training set choice, 20 different training sets of the same size (20 000 points) have been considered for the DRIVE database. The standard deviations for AUC and accuracy are 0.916×10^{-3} and 0.828×10^{-3}, respectively.

Figs. 6 and 7 show a small difference between ROC curves obtained with the line detector and with linear SVM, especially for the database DRIVE. A careful analysis reveals that the supervised method is useful in case of images containing artefacts due to pathologies. In the database STARE, in fact, there are ten images with pathologies, while for the test set of DRIVE we have only four. Moreover, in STARE the abnormal regions are wider. An example is given concerning an image from the STARE database: Fig. 12 shows the associated inverted green channel. The ROC curves for both methods are depicted in Fig. 13. The AUC is 0.9491 for the linear SVM classifier while
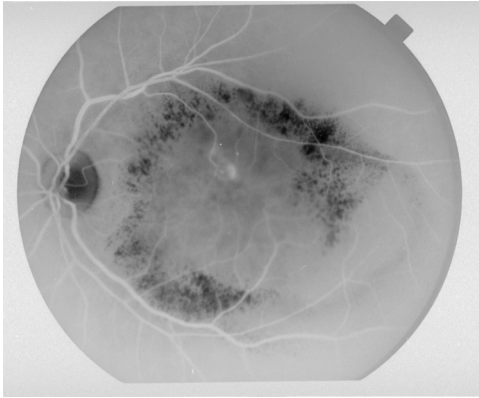
Fig. 12.   Inverted green channel of non-mydriatic retinal image with pathology from STARE database.
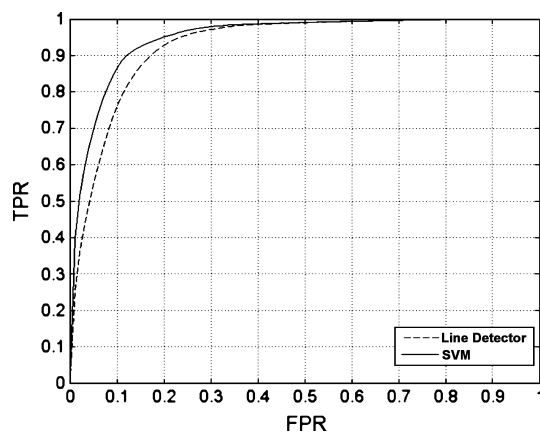

Fig. 13.   ROC curves for image in Fig. 12: line detector (dashed line) and linear SVM (solid line).
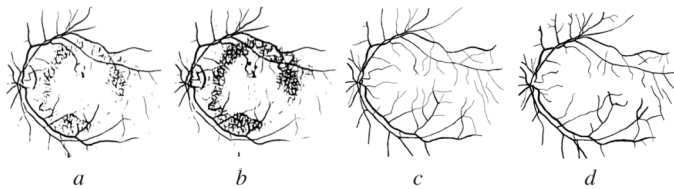

Fig. 14.   Vasculature segmentation of image in Fig. 12: (a) linear SVM, (b) line detector, (c) manual segmentation of the first observer, and (d) manual segmentation of second observer.

it is only 0.9281 for the line detector. The difference is clearly visible from the segmented images shown in Fig. 14. The first image is obtained with the linear SVM and corresponds to the value of b giving the best accuracy (0.9395). The second image is obtained using the line detector: the associated accuracy is 0.9147. With the same TPR (0.63), we have far less false positives for the first method (FPR $= 0.038$) with respect to the second (FPR $= 0.065$), making the linear SVM approach more robust to artefacts due to pathologies. The manual segmentation of the two observers is also depicted.

Finally, to investigate the dependence of the proposed pixel classification method on the dataset, we trained the classifier on each of the DRIVE and STARE databases and then we tested it on the other. The maximum accuracy is 0.9452 when training

on DRIVE and 0.9266 in the second case. The performance is worse, as expected, due to the different image types (a similar behavior is observed in [17]). Hence retraining of the classifier is needed when using a new dataset. However, this is not a problem with the proposed classification approach since the training set size is small, training is quite easy (no parametric model must be assumed, no local nonoptimal solutions can exist) and automatic (C^- should be tuned by the user but in practice it can be set to one with negligible effect on the performance).

V. DISCUSSION AND CONCLUSION

In this paper, we have presented an effective retinal vessel segmentation technique based on simple line operators. In developing our method the usual pattern recognition paradigm has been followed: first feature extraction and then classification. The basic line detector is computationally simple and gives very good results with respect to existing unsupervised methods. The proposed operator retains some advantages of the matched filter [7], avoiding some of its drawbacks. The averaging implicit in the line strength computation reduces noise but it is simpler than a convolution with a two-dimensional kernel. Since no hypothesis is assumed on the cross-section profile, we obtain robust detection in presence of vessels with different size. Moreover, the one-pixel width of the rotating line gives good edge localization and increased selectivity between inside and outside pixels.

As concerns the proposed supervised approach, it requires fewer features with respect to existing methods in the literature [16], [17]. SVMs guarantee good generalization performance even with a small training set. Hence, the manual segmentation effort of the ophthalmologist is extremely reduced. Since a linear SVM has been chosen, the evaluation of the decision function in the test phase is very fast. Performance estimated both in terms of AUC and accuracy shows that our supervised method is slightly superior than the existing approaches. With respect to the line detector alone, in presence of abnormal regions, like hemorrhages, the supervised approach gives far less false positives producing cleaner segmented images.

From a visual inspection of the segmented images, vessels seem to have a few gaps. Therefore, our method is a good starting point to develop algorithms for accurate detection of the skeleton and the branching points [3]. Concerning noise and other artefacts, our approach is quite robust but there is still room for improvement. In particular, some false positives are found around the border of the optic disk and in proximity of the pathological regions. Some problems occur also for the segmentation of the thinnest vessels or when the local contrast is quite low; in these cases, vessels can be captured only at the expense of an high FPR. In this condition, however, the segmented image tends to be quite noisy. Our ongoing work is devoted to facing these problems.

As a final remark, it should be pointed out that evaluating the classification performance pixelwise is not void of disadvantages. The importance of wide, well contrasted vessels is overestimated while false positive pixels are underestimated, due to the huge number of non-vessel pixels. So, a different widely accepted criterion which avoids the above-mentioned drawbacks would be desirable.

REFERENCES

[1] J. J. Kanski, *Clinical Ophthalmology: A Systematic Approach*. London, U.K.: Butterworth-Heinemann, 1989.

[2] D. C. Klonoff and D. M. Schwartz, "An economic analysis of interventions for diabetes," *Diabetes Care*, vol. 23, no. 3, pp. 390–404, 2000.

[3] M. E. Martinez-Perez, A. D. Hughes, A. V. Stanton, S. A. Thom, N. Chapman, A. B. Bharath, and K. H. Parker, "Retinal vascular tree morphology: A semi-automatic quantification," *IEEE Trans. Biomed. Eng.*, vol. 49, no. 8, pp. 912–917, Aug. 2002.

[4] F. Zana and J. C. Klein, "A multimodal registration algorithm of eye fundus images using vessels detection and Hough transform," *IEEE Trans. Med. Imag.*, vol. 18, no. 5, pp. 419–428, May 1999.

[5] M. Niemeijer, B. van Ginneken, J. 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.

[6] A. Hoover and M. Goldbaum, "Locating the optic nerve in a retinal image using the fuzzy convergence of the blood vessels," *IEEE Trans. Med. Imag.*, vol. 22, no. 8, pp. 951–958, Aug. 2003.

[7] S. Chauduri, S. Chatterjee, N. Katz, M. Nelson, and M. Goldbaum, "Detection of blood vessels in retinal images using two-dimensional matched filters," *IEEE Trans. Med. Imag.*, vol. 8, no. 3, pp. 263–269, Mar. 1989.

[8] I. Liu and Y. Sun, "Recursive tracking of vascular networks in angiograms based on the detection-deletion scheme," *IEEE Trans. Med. Imag.*, vol. 12, no. 2, pp. 334–341, Feb. 1993.

[9] Y. A. Tolias and S. M. Panas, "A fuzzy vessel tracking algorithm for retinal images based on fuzzy clustering," *IEEE Trans. Med. Imag.*, vol. 17, no. 4, pp. 263–273, Apr. 1998.

[10] A. Can, H. Shen, J. N. Turner, H. L. Tanenbaum, and B. Roysam, "Rapid automated tracing and feature extraction from retinal fundus images using direct exploratory algorithms," *IEEE Trans. Inform. Technol. Biomed.*, vol. 3, no. 2, pp. 125–138, Jun. 1999.

[11] X. Jiang and D. Mojon, "Adaptive local thresholding by verification based multithreshold probing with application to vessel detection in retinal images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 1, pp. 131–137, Jan. 2003.

[12] F. Zana and J. 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.

[13] A. M. Mendonça and A. Campilho, "Segmentation of retinal blood vessels by combining the detection of centerlines and morphological reconstruction," *IEEE Trans. Med. Imag.*, vol. 25, no. 9, pp. 1200–1213, Sep. 2006.

[14] C. Sinthanayothin, J. F. Boyce, H. L. Cook, and T. H. Williamson, "Automated localisation of the optic disc, fovea and retinal blood vessels from digital colour fundus images," *Br. J. Ophtalmol.*, vol. 83, pp. 902–910, 1999.

[15] M. Niemeijer, J. Staal, B. van Ginneken, M. Loog, and M. D. Abramoff, "Comparative study of retinal vessel segmentation methods on a new publicly available database," in *SPIE Med. Imag.*, J. M. Fitzpatrick and M. Sonka, Eds., 2004, vol. 5370, pp. 648–656.

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

[17] J. V. B. Soares, J. J. G. Leandro, R. M. Cesar, Jr., H. F. Jelinek, and M. J. Cree, "Retinal vessel segmentation using the 2D Gabor wavelet and supervised classification," *IEEE Trans. Med. Imag.*, vol. 25, no. 9, pp. 1214–1222, Sep. 2006.

[18] R. Zwiggelaar, S. M. Astley, C. R. M. Boggis, and C. J. Taylor, "Linear structures in mammographic images: Detection and classification," *IEEE Trans. Med. Imag.*, vol. 23, no. 9, pp. 1077–1086, Sep. 2004.

[19] R. Dixon and C. Taylor, "Automated asbestos fibre counting," *Inst. Phys. Conf. Ser.*, vol. 44, pp. 178–185, 1979.

[20] N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines*. Cambridge, U.K.: Cambridge Univ. Press, 2000.

[21] 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.

[22] J. Lowell, A. Hunter, D. Steel, A. Basu, R. Ryder, and R. L. Kennedy, "Measurement of retinal vessel widths from fundus images based on 2-D modeling," *IEEE Trans. Med. Imag.*, vol. 23, no. 10, pp. 1196–1204, Oct. 2004.

[23] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, 2nd ed. New York: Wiley, 2001.

[24] K. Veropoulos, C. Campbell, and N. Cristianini, "Controlling the sensitivity of support vector machines," in *Proc. Int. Joint Conf. Artificial Intelligence*, T. Dean, Ed., Stockholm, Sweden, 1999, pp. 55–60.

[25] T. Joachims, "Making large-scale SVM learning practical," in *Advances in Kernel Methods, Support Vector Learning*, B. Scholkopf, C. Burges, and A. Smola, Eds. Cambridge, MA: MIT Press, 1999.

[26] K. Morik, P. Brockhausen, and T. Joachims, "Combining statistical learning with a knowledge-based approach—A case study in intensive care monitoring," in *Proc. 16th Int. Conf. Machine Learning (ICML-99)*, 1999.

[27] D. D. Lewis, Y. Yang, T. G. Rose, and F. Li, "RCV1: A new benchmark collection for text categorization research," *J. Machine Learning Res. 5*, pp. 361–397, 2004.

[28] M. Foracchia, E. Grisan, and A. Ruggeri, "Luminosity and contrast normalization in retinal images," *Med. Image Anal.*, vol. 9, no. 3, pp. 179–190, 2005.