A New Supervised Method for Blood Vessel Segmentation in Retinal Images by Using Gray-Level and Moment Invariants-Based Features

Diego Marín, Arturo Aquino*, Manuel Emilio Gegúndez-Arias, and José Manuel Bravo

***Abstract*—This paper presents a new supervised method for blood vessel detection in digital retinal images. This method uses a neural network (NN) scheme for pixel classification and computes a 7-D vector composed of gray-level and moment invariants-based features for pixel representation. The method was evaluated on the publicly available DRIVE and STARE databases, widely used for this purpose, since they contain retinal images where the vascular structure has been precisely marked by experts. Method performance on both sets of test images is better than other existing solutions in literature. The method proves especially accurate for vessel detection in STARE images. Its application to this database (even when the NN was trained on the DRIVE database) outperforms all analyzed segmentation approaches. Its effectiveness and robustness with different image conditions, together with its simplicity and fast implementation, make this blood vessel segmentation proposal suitable for retinal image computer analyses such as automated screening for early diabetic retinopathy detection.**

***Index Terms*—Diabetic retinopathy, moment invariants, retinal imaging, telemedicine, vessels segmentation.**

I. Introduction

DIABETIC retinopathy (DR) is the leading ophthalmic pathological cause of blindness among people of working age in developed countries [1]. It is provoked by diabetes-mellitus complications and, although diabetes affection does not necessarily involve vision impairment, about 2% of the patients affected by this disorder are blind and 10% undergo vision degradation after 15 years of diabetes [2], [3] as a consequence of DR complications. The estimated prevalence of diabetes for all age groups worldwide was 2.8% in 2000 and 4.4% in 2030, meaning that the total number of diabetes patients is forecasted to rise from 171 million in 2000 to 366 million in 2030 [4].

The main cause of DR is abnormal blood glucose level elevation, which damages vessel endothelium, thus increasing vessel permeability. The first manifestations of DR are tiny capillary dilations known as microaneurysms. DR progression also causes neovascularization, hemorrhages, macular edema and, in later stages, retinal detachment.

Although DR is not a curable disease, laser photocoagulation can prevent major vision loss if detected in early stages [1], [5]. However, DR patients perceive no symptoms until visual loss develops, usually in the later disease stages, when the treatment is less effective. So, to ensure the treatment is received in time, diabetic patients need annual eye-fundus examination [6]. However, this preventive action involves a huge challenge for Health Systems due to the huge number of patients needing ophthalmologic revision, thus preventing many patients from receiving adequate treatment. Therefore, DR also becomes a great economic problem for Administrations since, only in U.S., cost of ophthalmic chronic complications caused by diabetes exceeded 1 billion dollars in 2007 [7].

The employment of digital images for eye diseases diagnosis could be exploited for computerized early detection of DR. A system that could be used by nonexperts to filtrate cases of patients not affected by the disease, would reduce the specialists' workload, and increase the effectiveness of preventive protocols and early therapeutic treatments. Furthermore, it would also result in economic benefits for public Health Systems, since cost-effective treatments associated to early illness detection lead to remarkable cost savings [8].

Since vascular anomalies are one of DR manifestations, automatic assessment of eye-fundus blood vessels is necessary for automated detection of DR. As a previous step, vessel assessment demands vascular tree segmentation from the background for further processing. Knowledge on blood vessel location can be used to reduce the number of false positives in microaneurysm and hemorrhage detection [9]–[12]. Besides these applications motivated by automated early detection of DR, vascular tree segmentation proves useful for other clinical purposes: evaluation of the retinopathy of prematurity [15], arteriolar narrowing [16], [17], vessel tortuosity to characterize hypertensive retinopathy [18], vessel diameter measurement to diagnose hypertension and cardiovascular diseases [19]–[21], and computer-assisted laser surgery [22], [23], among others. On the other hand, the vascular tree can also be useful as valuable information to locate other fundus features such as the optic disc [24]–[26] and the fovea [27]. Moreover, it may serve as a mean for the registration of multimodal images [13], [14].

In this paper, a new methodology for blood vessel detection is presented. It is based on pixel classification using a 7-D

Manuscript received July 02, 2010; accepted July 28, 2010. *Asterisk indicates corresponding author*.

D. Marín and J. M. Bravo are with the Department of Electronic, Computer Science and Automatic Engineering, "La Rábida" Polytechnic School, University of Huelva, 21819 Palos de la Frontera, Spain (e-mail: diego.marin@diesia.uhu.es; caro@diesia.uhu.es).

*A. Aquino iswith the Department of Electronic, Computer Science and Automatic Engineering, "La Rábida" Polytechnic School, University of Huelva, 21819 Palos de la Frontera, Spain (e-mail: arturo.aquino@diesia.uhu.es).

M. E. Gegúndez-Arias is with the Department of Mathematics, "La Rábida" Polytechnic School, University of Huelva, 21819 Palos de la Frontera, Spain (e-mail: gegundez@uhu.es).

Digital Object Identifier 10.1109/TMI.2010.2064333

feature vector extracted from preprocessed retinal images and given as input to a neural network. Classification results (real values between 0 and 1) are thresholded to classify each pixel into two classes: vessel and nonvessel. Finally, a postprocessing fills pixel gaps in detected blood vessels and removes falsely-detected isolated vessel pixels.

Despite its simplicity, the high accuracy achieved by this method in blood vessel detection is comparable to that reported by the most accurate methods in literature. Moreover, it offers a better behavior against images of different conditions. This fact is especially relevant if we keep in mind that the main aim of implementing a vessel segmentation algorithm is its integration in systems for automated detection of eye diseases. This kind of systems should require no user interaction and, therefore, be robust enough to analyze different kinds of images. Within the field of retinal imaging, this involves a huge challenge, since large variability is observed in the image acquisition process and a natural variation is reported in the appearance of the retina.

The rest of the paper is organized as follows. Next section reviews other published vessel segmentation solutions. Section III describes the material used in this study. Section IV explains and illustrates the proposed method for retinal vessel segmentation, while Section V presents its results and compares them to those obtained with other existing methods. Finally, the authors' conclusions and discussion conclude this paper.

II. State of Art

Many methods for retinal vessel segmentation have been reported. These can be divided into two groups: rule-based methods and supervised methods. In the first group, we highlight methods using vessel tracking, mathematical morphology, matched filtering, model-based locally adaptive thresholding or deformable models. On the other hand, supervised methods are those based on pixel classification (implementing some kind of classifier).

Regarding rule-based methods, vessel tracking methods [28]–[33] attempt to obtain the vasculature structure by following vessel center lines. Starting from an initial set of points established automatically or by manual labeling, vessels are traced by deciding from local information the most appropriate candidate pixel from those close to that currently under evaluation. Other methods use mathematical morphology [15], [34]–[36] to benefit from *a priori*-known vasculature shape features, such as being piecewise linear and connected. Then, by applying morphological operators, the vasculature is filtered from the background for final segmentation. Matched filtering techniques [37]–[42] usually use a 2-D linear structural element with a Gaussian cross-profile section, extruded or rotated into three dimensions for blood vessel cross-profile identification (typically a Gaussian or Gaussian-derivative profile). The kernel is rotated into many different orientations (usually 8 or 12) to fit into vessels of different configuration. The image is then thresholded to extract the vessel silhouette from the background. Regarding model-based locally adaptive thresholding, a general framework based on a verification-based multithreshold probing scheme was presented by Jiang *et al.* in [43]. These authors enriched this generic methodology by incorporating relevant information related to retinal vessels into the verification process with the aim of enabling its application to retinal images. On the other hand, deformable or snake models have been also used in [44] and [45]. A snake is an active contour model that, once placed on the image near the contour of interest, can evolve to fit the shape of the desired structure by an iterative adaption. Other rule-based methods for retinal blood vessel segmentation were reported in [46] and [47]. Martinez *et al.* [46] proposed a method based upon multiscale feature extraction. The local maxima over scales of the gradient magnitude and the maximum principal curvature of the Hessian tensor were used in a multiple pass region growing procedure. Growth progressively segmented the blood vessels by using both feature and spatial information. In the method presented in [47], blood vessel-like objects were extracted by using the Laplacian operator and noisy objects were pruned according to centerlines, detected by means of the normalized gradient vector field.

On the other hand, supervised methods are based on pixel classification, which consists on classifying each pixel into two classes, vessel and non-vessel. Classifiers are trained by supervised learning with data from manually-labeled images. Gardner *et al.* [48] proposed a back propagation multilayer neural network (NN) for vascular tree segmentation. After histogram equalization, smoothing and edge detection, the image was divided into 20×20 pixel squares (400 input neurons). The NN was then fed with the values of these pixel windows for classifying each pixel into vessel or not. Sinthanayothin *et al.* [49] also used a multilayer perceptron NN. Each pixel in the image was classified by using the first principal component, and the edge strength values from a 10×10 pixel subimage centered on the pixel under evaluation, as input data. Niemeijer *et al.* [50] implemented a K-nearest neighbor (kNN) classifier. A 31-component pixel feature vector was constructed with the Gaussian and its derivatives up to order 2 at 5 different scales, augmented with the gray-level from the green channel of the original image. The assumption that vessels are elongated structures is the basis for the supervised ridge-based vessel detection method presented by Staal *et al.* [51]. Ridges were extracted from the image and used as primitives to form line elements. Each pixel was then assigned to its nearest line element, the image thus being partitioned into patches. For every pixel, 27 features were firstly computed and those obtaining the best class separability were finally selected. Feature vectors were classified by using a kNN-classifier and sequential forward feature selection. Soares *et al.* [52] used a Gaussian mixture model Bayesian classifier. Multiscale analysis was performed on the image by using the Gabor wavelet transform. The gray-level of the inverted green channel and the maximum Gabor transform response over angles at four different scales were considered as pixel features. Finally, Ricci and Pefetti [53] used a support vector machine (SVM) for pixel classification as vessel or nonvessel. They used two orthogonal line detectors along with the gray-level of the target pixel to construct the feature vector.

III. Materials

To evaluate the vessel segmentation methodology described in the next section, two publicly available databases containing

retinal images, the DRIVE [54] and STARE [55] databases, were used. These databases have been widely used by other researchers to test their vessel segmentation methodologies since, apart from being public, they provide manual segmentations for performance evaluation.

The DRIVE database comprises 40 eye-fundus color images (seven of which present pathology) taken with a Canon CR5 nonmydriatic 3CCD camera with a 45° field-of-view (FOV). Each image was captured at 768×584 pixels, 8 bits per color plane and, in spite of being offered in LZW compressed TIFF format, they were originally saved in JPEG format. The database is divided into two sets: a test set and a training set, each of them containing 20 images. The test set provides the corresponding FOV masks for the images, which are circular (approximated diameter of 540 pixels) and two manual segmentations generated by two different specialists for each image. The selection of the first observer is accepted as ground truth and used for algorithm performance evaluation in literature. The training set also includes the FOV masks for the images and a set of manual segmentations made by the first observer.

The STARE database, originally collected by Hoover *et al.* [38], comprises 20 eye-fundus color images (ten of them contain pathology) captured with a TopCon TRV-50 fundus camera at 35° FOV. The images were digitalized to 700×605 pixels, 8 bits per color channel and are available in PPM format. The database contains two sets of manual segmentations made by two different observers. Performance is computed with the segmentations of the first observer as ground truth.

IV. Proposed Vessel Segmentation Method

This paper proposes a new supervised approach for blood vessel detection based on a NN for pixel classification. The necessary feature vector is computed from preprocessed retinal images in the neighborhood of the pixel under consideration. The following process stages may be identified: 1) original fundus image preprocessing for gray-level homogenization and blood vessel enhancement, 2) feature extraction for pixel numerical representation, 3) application of a classifier to label the pixel as vessel or nonvessel, and 4) postprocessing for filling pixel gaps in detected blood vessels and removing falsely-detected isolated vessel pixels.

Input images are monochrome and obtained by extracting the green band from original RGB retinal images. The green channel provides the best vessel-background contrast of the RGB-representation, while the red channel is the brightest color channel and has low contrast, and the blue one offers poor dynamic range. Thus, blood containing elements in the retinal layer (such as vessels) are best represented and reach higher contrast in the green channel [56].

All parameters described below were set by experiments carried out on DRIVE images with the aim of contributing the best segmentation performance on this database (performance was evaluated in terms of average accuracy—a detailed description is provided in Sections V-A and V-B). Therefore, they refer to retinas of approximately 540 pixels in diameter. The application of the methodology to retinas of different size (i.e., the diameter in pixels of STARE database retinas is approximately 650 pixels) demands either resizing input images to fulfil this condition or adapting proportionately the whole set of used parameters to this new retina size.

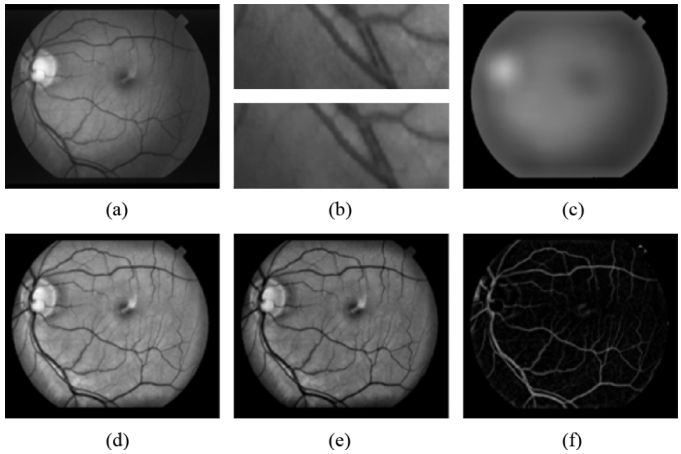


Fig. 1. Illustration of the preprocessing process: (a) Green channel of the original image. (b) The upper image is a fragment of the original image containing a vessel with central light reflex, while the bottom image shows the effect of reflex removal. (c) Background image. (d) Shade-corrected image. (e) Homogenized image. (f) Vessel-enhanced image.

A. Preprocessing

Color fundus images often show important lighting variations, poor contrast and noise. In order to reduce these imperfections and generate images more suitable for extracting the pixel features demanded in the classification step, a preprocessing comprising the following steps is applied: 1) vessel central light reflex removal, 2) background homogenization, and 3) vessel enhancement. Next, a description of the procedure, illustrated through its application to a STARE database fundus image (Fig. 1), is detailed.

1) Vessel Central Light Reflex Removal: Since retinal blood vessels have lower reflectance when compared to other retinal surfaces, they appear darker than the background. Although the typical vessel cross-sectional gray-level profile can be approximated by a Gaussian shaped curve (inner vessel pixels are darker than the outermost ones), some blood vessels include a light streak (known as a light reflex) which runs down the central length of the blood vessel.

To remove this brighter strip, the green plane of the image is filtered by applying a morphological opening using a three-pixel diameter disc, defined in a square grid by using eight-connexity, as structuring element. Disc diameter was fixed to the possible minimum value to reduce the risk of merging close vessels. I_γ denotes the resultant image for future references.

An example of vessel central light reflex and its removal from a fundus image by means of opening filtering operation is shown in Fig. 1(a) and (b).

2) Background Homogenization: Fundus images often contain background intensity variation due to nonuniform illumination. Consequently, background pixels may have different intensity for the same image and, although their gray-levels are usually higher than those of vessel pixels (in relation to green channel images), the intensity values of some background pixels is comparable to that of brighter vessel pixels. Since the feature vector used to represent a pixel in the classification stage is

formed by gray-scale values, this effect may worsen the performance of the vessel segmentation methodology. With the purpose of removing these background lightening variations, a shade-corrected image is accomplished from a background estimate. This image is the result of a filtering operation with a large arithmetic mean kernel, as described below.

Firstly, a 3×3 mean filter is applied to smooth occasional salt-and-pepper noise. Further noise smoothing is performed by convolving the resultant image with a Gaussian kernel of dimensions $m \times m = 9 \times 9$, mean $\mu = 0$ and variance $\sigma^2 = 1.8^2$, $G^m_{\mu,\sigma^2} = G^9_{0,1.8^2}$. Secondly, a background image I_B, is produced by applying a 69×69 mean filter [Fig. 1(c)]. When this filter is applied to the pixels in the FOV near the border, the results are strongly biased by the external dark region. To overcome this problem, out-of-the FOV gray-levels are replaced by average gray-levels in the remaining pixels in the square. Then, the difference D between I_γ and I_B is calculated for every pixel

$$D(x,y) = I_\gamma(x,y) - I_B(x,y). \tag{1}$$

To this respect, literature reports shade-correction methods based on the subtraction of the background image from the original image [10], [12], [57] or the division of the latter by the former [58], [59]. Both procedures rendered similar results upon testing. Moreover, none of them showed to contribute any appreciable advantage relative to the other. The subtractive approach in (1) was used in the present work.

Finally, a shade-corrected image I_{SC} is obtained by transforming linearly RD values into integers covering the whole range of possible gray-levels ([0–255], referred to 8-bit images). Fig. 1(d) shows the I_{SC} corresponding to a nonuniformly illuminated image. The proposed shade-correction algorithm is observed to reduce background intensity variations and enhance contrast in relation to the original green channel image.

Besides the background intensity variations in images, intensities can reveal significant variations between images due to different illumination conditions in the acquisition process. In order to reduce this influence, a homogenized image I_H [Fig. 1(a)] is produced as follows: the histogram of I_{SC} is displaced toward the middle of the gray-scale by modifying pixel intensities according to the following gray-level global transformation function:

$$g_{\mathrm{Output}} = \begin{cases} 0, & \text{if } g < 0 \\ 255, & \text{if } g > 255 \\ \mathrm{g}, & \text{otherwise} \end{cases} \tag{2}$$

where

$$g = g_{\mathrm{Input}} + 128 - g_{\mathrm{Input\_Max}} \tag{3}$$

and g_{Input} and g_{Output} are the gray-level variables of input and output images (I_{SC} and I_H, respectively). The variable denoted by $g_{\mathrm{Input\_Max}}$ defines the gray-level presenting the highest number of pixels in I_{SC}. By means of this operation, pixels with gray-level $g_{\mathrm{Input\_Max}}$, which are observed to correspond to the background of the retina, are set to 128 for 8-bit images. Thus, background pixels in images with different illumination conditions will standardize their intensity around this value.

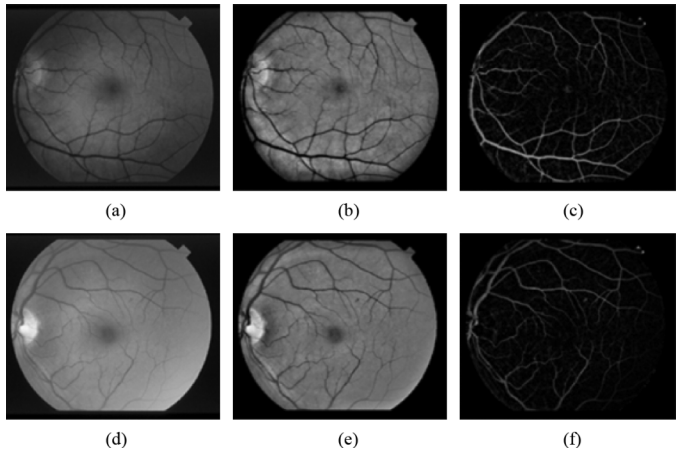


Fig. 2. Two examples of application of the preprocessing on two images with different illumination conditions. (a), (d) Green channel of the original images. (b), (e) Homogenized images. (c), (f) Vessel-enhanced images.

Fig. 2 (a), (b) and (d), (e), shows this effect for two fundus images in the STARE database.

3) Vessel Enhancement: The final preprocessing step consists on generating a new vessel-enhanced image (I_{VE}), which proves more suitable for further extraction of moment invariants-based features (see Section IV-B).

Vessel enhancement is performed by estimating the complementary image of the homogenized image I_H, I^c_H, and subsequently applying the morphological *Top-Hat transformation* [Fig. 1(f)]

$$I_{\mathrm{VE}} = I^c_H - \gamma(I^c_H) \tag{4}$$

where γ is a morphological opening operation using a disc of eight pixels in radius. Thus, while bright retinal structures are removed (i.e., optic disc, possible presence of exudates or reflection artifacts), the darker structures remaining after the opening operation become enhanced (i.e., blood vessels, fovea, possible presence of microaneurysms or hemorrhages). Samples of vessel enhancement operation results are shown in Fig. 2(c) and (f) for two fundus images with variable illumination conditions.

B. Feature Extraction

The aim of the feature extraction stage is pixel characterization by means of a feature vector, a pixel representation in terms of some quantifiable measurements which may be easily used in the classification stage to decide whether pixels belong to a real blood vessel or not. In this paper, the following sets of features were selected.

- *Gray-level-based features*: features based on the differences between the gray-level in the candidate pixel and a statistical value representative of its surroundings.
- *Moment invariants-based features*: features based on moment invariants for describing small image regions formed by the gray-scale values of a window centered on the represented pixels.

1) Gray-Level-Based Features: Since blood vessels are always darker than their surroundings, features based on describing gray-level variation in the surroundings of candidate

pixels seem a good choice. A set of gray-level-based descriptors taking this information into account were derived from homogenized images I_H considering only a small pixel region centered on the described pixel (x,y). $S^w_{x,y}$ stands for the set of coordinates in a $w \times w$ sized square window centered on point (x,y). Then, these descriptors can be expressed as

$$f_1(x,y) = I_H(x,y) - \min_{(s,t)\in S^9_{x,y}} \{I_H(s,t)\} \quad (5)$$

$$f_2(x,y) = \max_{(s,t)\in S^9_{x,y}} \{I_H(s,t)\} - I_H(x,y) \quad (6)$$

$$f_3(x,y) = I_H(x,y) - \operatorname*{mean}_{(s,t)\in S^9_{x,y}} \{I_H(s,t)\} \quad (7)$$

$$f_4(x,y) = \operatorname*{std}_{(s,t)\in S^9_{x,y}} \{I_H(s,t)\} \quad (8)$$

$$f_5(x,y) = I_H(x,y). \quad (9)$$

2) Moment Invariants-Based Features: The vasculature in retinal images is known to be piecewise linear and can be approximated by many connected line segments. For detecting these quasi-linear shapes, which are not all equally wide and may be oriented at any angle, shape descriptors invariant to translation, rotation and scale change may play an important role. Within this context, moment invariants proposed by Hu [60] provide an attractive solution and are included in the feature vector. In this paper, they are computed as follows.

Given a pixel (x,y) of the vessel-enhanced image I_{VE}, a subimage is generated by taking the region defined by $S^{17}_{x,y}$. The size of this region was fixed to 17×17 so that, considering that the region is centered on the middle of a "wide" vessel (8-9-pixel wide and referred to retinas of approximately 540 pixels in diameter), the subimage includes an approximately equal number of vessel and nonvessel pixels. For this subimage, denoted by $I^{S^{17}_{x,y}}_{\mathrm{VE}}$, the 2-D moment of order $(p+q)$ is defined as

$$m_{pq} = \sum_i \sum_j i^p j^q I^{S^{17}_{x,y}}_{\mathrm{VE}}(i,j) \quad p,\ q = 0,1,2,\ldots \quad (10)$$

where summations are over the values of the spatial coordinates i and j spanning the subimage, and $I^{S^{17}_{x,y}}_{\mathrm{VE}}(i,j)$ is the gray-level at point (i,j).

The corresponding central moment is defined as

$$\mu_{pq} = \sum_i \sum_j (i-\bar{i})^p (j-\bar{j})^q I^{S^{17}_{x,y}}_{\mathrm{VE}}(i,j) \quad (11)$$

where

$$\bar{i} = \frac{m_{10}}{m_{00}}, \quad \bar{j} = \frac{m_{01}}{m_{00}} \quad (12)$$

are the coordinates of the center of gravity of the subimage.

The normalized central moment of order $(p+q)$ is defined as

$$\eta_{pq} = \frac{\mu_{pq}}{(\mu_{00})^{\gamma}} \quad p,\ q = 0,1,2,\ldots \quad (13)$$

where

$$\gamma = \frac{p+q}{2} + 1; \quad (p+q) = 2,3,\ldots. \quad (14)$$

A set of seven moment invariants under size, translation, and rotation, known as Hu moment invariants, can be derived from combinations of regular moments. Among them, our tests have revealed that only those defined by

$$\phi_1 = \eta_{20} + \eta_{02} \quad (15)$$

$$\phi_2 = (\eta_{20} + \eta_{02})^2 + 4\eta_{11}^2 \quad (16)$$

constitute the combination providing optimal performance in terms of average accuracy (see Section V-B). The inclusion of the remainder moments result in decreasing classification performance and increasing computation needed for classification. Moreover, the module of the logarithm was used instead of its values themselves. Using the logarithm reduces the dynamic range and the module prevents from having to deal with the complex numbers resulting from computing the logarithm of negative moment invariants.

Fig. 3 shows several samples of pixels, marked with solid white dots on an I_{VE} image [Fig. 3(a)], and the subimages $I^{S^{17}_{x,y}}_{\mathrm{VE}}$ generated around them [Fig. 3(b)–(i)]. Pairs of pixels were selected for different vessels: one inside and the other outside the vessel but near enough so that both subimages contain the vessel. Table I shows the moment values corresponding to each subimage. It can be checked that numbers are close, thus indicating a high degree of invariance to size, rotation and translation. Moments computed as mentioned above characterize numerically a vessel independently of its width, orientation and location in the subimage. However, they are not useful to describe the central pixel of the subimage in terms of vessel or nonvessel, as their values do not distinguish between these two situations.

To overcome this problem, moments are computed on new subimages I_{Hu} produced by multiplying the original ones, $I^{S^{17}_{x,y}}_{\mathrm{VE}}$, by an equal-dimension matrix (17×17) of Gaussian values, whose mean is 0 and variance is 1.7^2, $G^{17}_{0,1.7^2}$. That is, for every point of coordinates (i,j)

$$I_{Hu}(i,j) = I^{S^{17}_{x,y}}_{\mathrm{VE}}(i,j) \times G^{17}_{0,1.7^2}(i,j). \quad (17)$$

With this choice of parameters, the $9x9$ central values in $G^{17}_{0,1.7^2}$ contain the 97% of the area of the represented Gaussian distribution, the remainder values being close to 0 (supposing that the central pixel of $I^{S^{17}_{x,y}}_{\mathrm{VE}}$ is located on the middle of a "wide" vessel, these 9×9 central values in $G^{17}_{0,1.7^2}$ correspond to vessel pixels in $I^{S^{17}_{x,y}}_{\mathrm{VE}}$). This operation is illustrated in Fig. 3(b)–(i) and (j)–(q). The effect of this multiplication is clearly observed in these new subimages and their associated moment values (Table II). These values become sensitive for describing vessel and nonvessel central pixels, as they now reflect significant differences between them. Both ϕ_1 and ϕ_2 values, in comparison with their original ones, increase if they describe vessel pixels and decrease otherwise.

In conclusion, the following descriptors were considered to be part of the feature vector of a pixel located at (x,y)

$$f_6(x,y) = |\log(\phi_1)| \quad (18)$$

$$f_7(x,y) = |\log(\phi_2)| \quad (19)$$

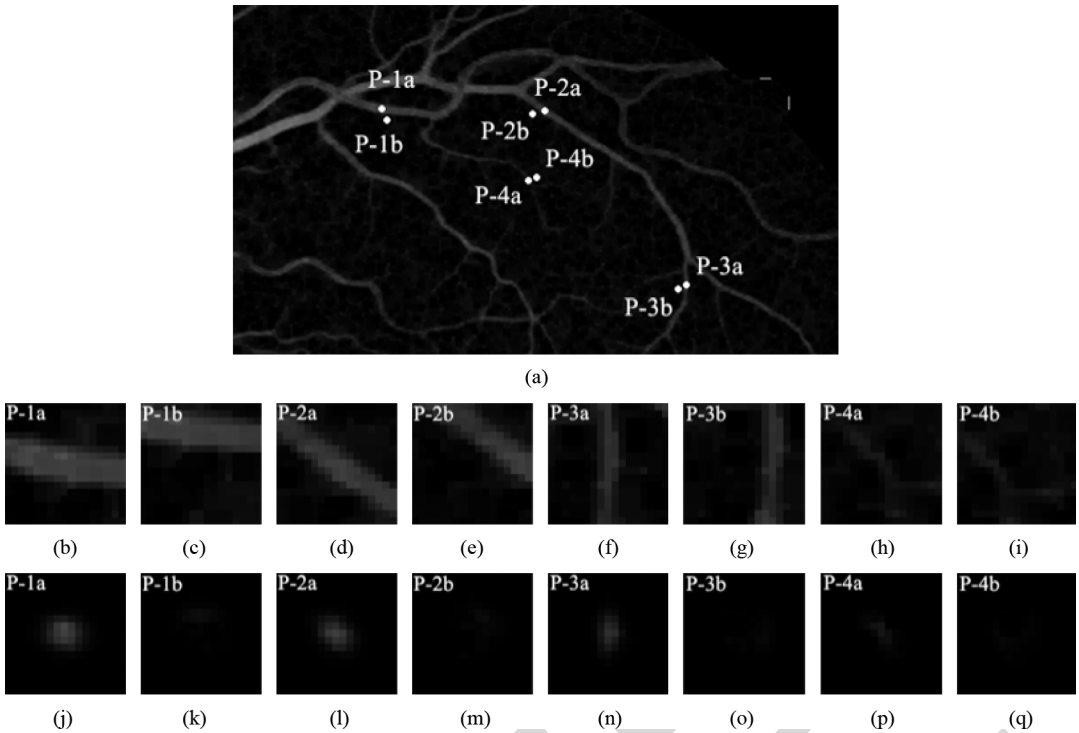


Fig. 3. Examples of obtaining pixel environments for moment invariants calculation. (a) Vessel enhanced subimage. Four pairs of pixels are marked with white dots: $\mathrm{P}-k\mathrm{a}$ and $\mathrm{P}-k\mathrm{b}$ with $k=1,2,3,4$; $\mathrm{P}-k\mathrm{a}$ are vessel pixels and $\mathrm{P}-k\mathrm{b}$ are background pixels close to their corresponding pair. (b)–(i) From left to right, extracted subimages $I_{\mathrm{VE}}^{S_{\mathrm{P}-k\delta}^{17}}$ with $k=1,2,3,4$ and $\delta=\mathrm{a},\mathrm{b}$. (j)–(q) Subimages I_{Hu} result of multiplying the original ones in (b)–(i) by the Gaussian matrix.

TABLE I
MODULE OF THE ϕ_1 AND ϕ_2 MOMENTS LOGARITHM CALCULATED FROM THE SUBIMAGES $I_{\mathrm{VE}}^{S_{x,y}^{17}}$ SHOWN IN FIG. 3 IMAGES (B)–(I)

	P-1a	P-1b	P-2a	P-2b	P-3a	P-3b	P-4a	P-4b
$\lvert\log(\phi_1)\rvert$	5.26	4.73	4.87	5.02	4.36	4.23	3.96	3.92
$\lvert\log(\phi_2)\rvert$	11.70	11.29	10.71	11.81	10.92	10.90	10.59	12.11

TABLE II
MODULE OF THE ϕ_1 AND ϕ_2 MOMENTS LOGARITHM CALCULATED FROM THE SUBIMAGES I_{Hu} SHOWN IN FIG. 3 IMAGES (J)–(Q)

	P-1a	P-1b	P-2a	P-2b	P-3a	P-3b	P-4a	P-4b
$\lvert\log(\phi_1)\rvert$	5.34	2.89	5.16	3.13	4.79	2.34	4.12	2.21
$\lvert\log(\phi_2)\rvert$	13.57	9.16	12.85	10.11	11.19	8.31	10.82	7.79

where ϕ_1 and ϕ_2 are the moment invariants given by (15) and (16) computed on the subimages I_{Hu}, generated according to (17).

C. Classification

In the feature extraction stage, each pixel from a fundus image is characterized by a vector in a 7-D feature space

$$F(x,y)=\left(f_1(x,y),\ldots,f_7(x,y)\right). \tag{20}$$

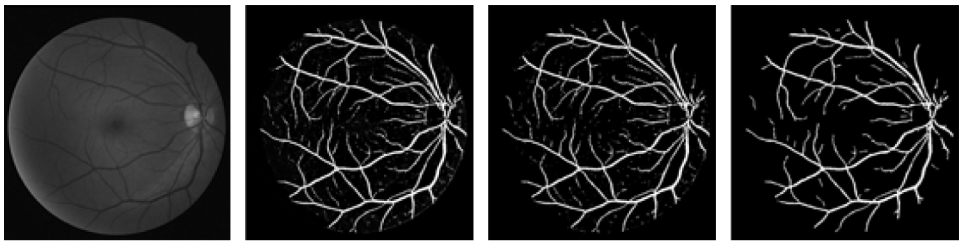
Now, a classification procedure assigns one of the classes $C1$ (vessel) or $C2$ (nonvessel) to each candidate pixel when its representation is known. In order to select a suitable classifier, the distribution of the training set data (described below) in the feature space was analyzed. The results of this analysis showed that the class linear separability grade was not high enough for the accuracy level required for vasculature segmentation in retinal images. Therefore, the use of a non linear classifier was necessary. The following nonlinear classifiers can be found in the existing literature on this topic: the kNN method [51] and [49], support vector machines [52], Bayesian classifier [50], or neural networks [43], [48]. A multilayer feedforward NN was selected in this paper.

Two classification stages can be distinguished: a design stage, in which the NN configuration is decided and the NN is trained, and an application stage, in which the trained NN is used to classify each pixel as vessel or nonvessel to obtain a vessel binary image.

1) Neural Network Design: A multilayer feedforward network, consisting of an input layer, three hidden layers and an output layer, is adopted in this paper. The input layer is composed by a number of neurons equal to the dimension of the feature vector (seven neurons). Regarding the hidden layers, several topologies with different numbers of neurons were tested. A number of three hidden layers, each containing 15 neurons, provided optimal NN configuration. The output layer contains a single neuron and is attached, as the remainder units, to a nonlinear logistic sigmoid activation function, so its output ranges between 0 and 1. This choice was grounded on the fact of interpreting NN output as posterior probabilities.

The training set, S_T, is composed of a set of N candidates for which the feature vector [F—(20)], and the classification result (C_1 or C_2: vessel or nonvessel) are known

$$S_T=\left\{\left(F^{(n)},C_k^{(n)}\right)\,|n=1,\ldots,N;\ k\in\{1,2\}\right\}. \tag{21}$$



(a) (b) (c) (d)

Fig. 4. (a) Green channel of the original image. (b) Obtained probability map represented as an image. (c) Thresholded image. (d) Postprocessed image.

The samples forming S_T were collected from manually labeled nonvessel and vessel pixels in the DRIVE training images. Specifically, around 30 000 pixel samples, fairly divided into vessel and non-vessel pixels, were used (as a reference, this number represents 0.65% of the total number of the DRIVE test image pixels that will be classified later on). Unlike other authors [52], [53], who selected their training set by random pixel-sample extraction from available manual segmentations of DRIVE and STARE images, we produced our own training set by hand.[1] As discussed in literature, gold-standard images may contain errors (see Bioux *et al.* [61] for a comprehensive discussion on this issue) due to the considerable difficulty involved by the creation of these handmade images. To reduce the risk of introducing errors in S_T and, therefore, of introducing noise in the NN, we opted for carefully selecting specific training samples covering all possible vessel, background, and noise patterns. Moreover, it should be pointed out that the network trained with the just defined S_T, in spite of taking information from DRIVE images only, was applied to compute method performance with both DRIVE and STARE databases.

Since the features f_i of F have very different ranges and values, each of these features is normalized to zero mean and unit variance independently by applying

$$\bar{f}_i = \frac{f_i - \mu_i}{\sigma_i} \tag{22}$$

where μ_i and σ_i stand for the average and standard deviation of the ith feature calculated over S_T.

Once S_T is established, NN is trained by adjusting the weights of the connections through error interpretation. The back-propagation training algorithm [62] was used with this purpose.

2) Neural Network Application: At this stage, the trained NN is applied to an "unseen" fundus image to generate a binary image in which blood vessels are identified from retinal background: pixels' mathematical descriptions are individually passed through the NN. In our case, the NN input units receive the set of features provided by (5)–(9) and (18) and (19), normalized according to (22). Since a logistic sigmoidal activation function was selected for the single neuron of the output layer, the NN decision determines a classification value between 0 and 1. Thus, a vessel probability map indicating the probability for the pixel to be part of a vessel is produced. Illustratively, the resultant probability map corresponding to a DRIVE database fundus image [Fig. 4(a)] is shown as an image in Fig. 4(b). The bright pixels in this image indicate higher probability of being vessel pixel. In order to obtain a vessel binary segmentation, a thresholding scheme on the probability map is used to decide whether a particular pixel is part of a vessel or not. Therefore, the classification procedure assigns one of the classes C_1 (vessel) or C_2 (nonvessel) to each candidate pixel, depending on if its associated probability is greater than a threshold Th. Thus, a classification output image I_{CO} [Fig. 4(c)], is obtained by associating classes C_1 and C_2 to the gray level values 255 and 0, respectively. Mathematically

$$I_{\mathrm{CO}}(x,y) = \begin{cases} 255\ (\equiv C1), & \text{if } p(C_1|F(x,y)) \geq Th \\ 0\ (\equiv C2), & \text{otherwise} \end{cases} \tag{23}$$

where $p(C_1|F(x,y))$ denotes the probability of a pixel (x,y) described by feature vector $F(x,y)$ to belong to class C_1. The optimal Th value is discussed in Section V-B.

D. Postprocessing

Classifier performance is enhanced by the inclusion of a two-step postprocessing stage: the first step is aimed at filling pixel gaps in detected blood vessels, while the second step is aimed at removing falsely detected isolated vessel pixels.

From visual inspection of the NN output, vessels may have a few gaps (i.e., pixels completely surrounded by vessel points, but not labeled as vessel pixels). To overcome this problem, an iterative filling operation is performed by considering that pixels with at least six neighbors classified as vessel points must also be vessel pixels. Besides, small isolated regions misclassified as blood vessel pixels are also observed. In order to remove these artifacts, the pixel area in each connected region is measured. In artifact removal, each region connected to an area below 25 is reclassified as nonvessel. An example of the final vessel segmented image after this further processing stage is shown in Fig. 4(d).

V. Experimental Results

A. Performance Measures

In order to quantify the algorithmic performance of the proposed method on a fundus image, the resulting segmentation is compared to its corresponding gold-standard image. This image is obtained by manual creation of a vessel mask in which all vessel pixels are set to one and all nonvessel pixels are set to zero. Thus, automated vessel segmentation performance can be assessed.

[1]The used training set is available online at http://www.uhu.es/retinopathy/eng/bd.php

TABLE III
CONTINGENCY VESSEL CLASSIFICATION

	Vessel present	Vessel absent
Vessel detected	True Positive (TP)	False Positive (FP)
Vessel not detected	False Negative (FN)	True Negative (TN)

In this paper, our algorithm was evaluated in terms of sensitivity (Se), specificity (Sp), positive predictive value (Ppv), negative predictive value (Npv), and accuracy (Acc). Taking Table III into account, these metrics are defined as

$$Se = \frac{TP}{TP+FN} \tag{24}$$

$$Sp = \frac{TN}{TN+FP} \tag{25}$$

$$Ppv = \frac{TP}{TP+FP} \tag{26}$$

$$Npv = \frac{TN}{TN+FN} \tag{27}$$

$$\mathrm{Acc} = \frac{TP+TN}{TP+FN+TN+FP}. \tag{28}$$

Se and Sp metrics are the ratio of well-classified vessel and nonvessel pixels, respectively. Ppv is the ratio of pixels classified as vessel pixel that are correctly classified. Npv is the ratio of pixels classified as background pixel that are correctly classified. Finally, Acc is a global measure providing the ratio of total well-classified pixels.

In addition, algorithm performance was also measured with receiver operating characteristic (ROC) curves. A ROC curve is a plot of true positive fractions (Se) versus false positive fractions ($1 - Sp$) by varying the threshold on the probability map. The closer a curve approaches the top left corner, the better the performance of the system. The area under the curve (AUC), which is 1 for a perfect system, is a single measure to quantify this behavior.

B. Proposed Method Evaluation

This method was evaluated on DRIVE and STARE database images with available gold-standard images. Since the images' dark background outside the FOV is easily detected, Se, Sp, Ppv, Npv and Acc values were computed for each image considering FOV pixels only. Since FOV masks are not provided for STARE images, they were generated with an approximate diameter of 650×550. The results are listed in Tables IV and V[2]. The last row of the tables shows average Se, Sp, Ppv, Npv, and Acc values (denoted as $\overline{\mathrm{Acc}}$ for future references), for the 20 images in each database.

TABLE IV
PERFORMANCE RESULTS ON DRIVE DATABASE IMAGES

Image	*Se*	*Sp*	*Ppv*	*Npv*	*Acc*
1	0.7778	0.9734	0.8149	0.9667	0.9477
2	0.7665	0.9793	0.8671	0.9597	0.9474
3	0.7225	0.9711	0.8099	0.9535	0.9349
4	0.7032	0.9855	0.8818	0.9557	0.9479
5	0.6815	0.9872	0.8931	0.9518	0.9457
6	0.6296	0.9893	0.9060	0.9420	0.9385
7	0.6780	0.9861	0.8817	0.9525	0.9453
8	**0.5704**	**0.9917**	**0.9083**	**0.9414**	0.9388
9	0.6530	0.9882	0.8809	0.9554	0.9489
10	0.6967	0.9846	0.8599	0.9599	0.9502
11	0.7164	0.9745	0.8068	0.9585	0.9410
12	0.7399	0.9785	0.8312	0.9634	0.9486
13	0.6522	0.9866	0.8896	0.9450	0.9392
14	0.7773	0.9702	0.7769	0.9702	0.9474
15	**0.8139**	**0.9444**	**0.6290**	**0.9777**	**0.9308**
16	0.7327	0.9816	0.8573	0.9606	0.9491
17	0.6629	0.9851	0.8623	0.9541	0.9454
18	0.7106	0.9781	0.8078	0.9630	0.9473
19	0.8058	0.9813	0.8552	0.9736	**0.9602**
20	0.6436	0.9861	0.8463	0.9586	0.9495
Average	0.7067	0.9801	0.8433	0.9582	0.9452

The performance results shown in Tables IV and V were obtained considering the same Th threshold value for all the images in the same database (0.63 and 0.91 for DRIVE and STARE images, respectively). These values were set to provide maximum average accuracy ($\overline{\mathrm{Acc}}_{\mathrm{MAX}}$) in each database in the following way. For a given Th value, one Acc value is obtained for each of the 20 images selected for testing on a given database. These 20 Acc values are then averaged to obtain a single performance measure, $\overline{\mathrm{Acc}}$, linked to the selected Th value. Several $\overline{\mathrm{Acc}}$ values are obtained at certain thresholds applying these operations for different Th values. The final Th threshold value selected for a given database is that providing the maximum $\overline{\mathrm{Acc}}$ value, $\overline{\mathrm{Acc}}_{\mathrm{MAX}}$. Fig. 5 shows the $\overline{\mathrm{Acc}}$ values calculated for Th values from 0 to 1 (step of 0.02). The results for both DRIVE and STARE databases are shown. The $\overline{\mathrm{Acc}}_{\mathrm{MAX}}$ values and their corresponding Th values are marked for every database in this figure. It is worth mentioning that $\overline{\mathrm{Acc}}$ variation shows no significant dependence on Th. As it can be observed in Fig. 5, although different optimum Th values are reached depending on the database on which performance is computed, a wide range of Th values provides $\overline{\mathrm{Acc}}$ values very close to $\overline{\mathrm{Acc}}_{\mathrm{MAX}}$. Therefore, Th values can be concluded not to be a critical method to assess performance in terms of Acc, since it slowly varies according to it. This influence of Th on system performance is also visible in the ROC curves for the two databases shown in Fig. 6. These curves were produced by calculating the true and the false positive fraction on all test images through Th-threshold variations. The AUC measured for both curves was 0.9588 and 0.9769 for the DRIVE and STARE databases, respectively.

On the other hand, the spacial distribution of the classification errors produced by the segmentation algorithm, FN and FP, was studied. The following four situations were considered: FN produced in thin and non-thin vessel pixels, and FP produced

[2]The final vessel segmented images are available online at http://www.uhu.es/retinopathy/eng/bd.php

TABLE V
PERFORMANCE RESULTS ON STARE DATABASE IMAGES

Image	*Se*	*Sp*	*Ppv*	*Npv*	*Acc*
1	0.5997	0.9844	0.8245	0.9527	0.9425
2	0.5074	0.9931	0.8798	0.9527	0.9489
3	0.6534	0.9892	0.8422	0.9699	0.9619
4	**0.4159**	0.9953	0.9094	**0.9379**	**0.9365**
5	0.5884	0.9857	0.8508	0.9452	0.9372
6	0.7958	0.9712	0.7250	0.9803	0.9559
7	0.8183	0.9771	0.8142	0.9772	0.9593
8	0.8682	**0.9673**	0.7512	0.9848	0.9572
9	0.7729	0.9818	0.8355	0.9731	0.9595
10	0.6670	0.9809	0.8110	0.9601	0.9467
11	0.8109	0.9724	0.7601	0.9795	0.9567
12	**0.8781**	0.9675	0.7603	**0.9854**	0.9581
13	0.7796	0.9770	0.8246	0.9697	0.9530
14	0.7765	0.9787	0.8374	0.9687	0.9537
15	0.6910	0.9850	0.8598	0.9599	0.9504
16	0.6802	0.9825	0.8639	0.9497	0.9402
17	0.7039	0.9882	0.8926	0.9599	0.9534
18	0.5840	**0.9961**	**0.9172**	0.9698	0.9675
19	0.6776	0.9872	0.7694	0.9799	**0.9689**
20	0.6225	0.9763	**0.7245**	0.9628	0.9441
Average	0.6944	0.9819	0.8227	0.9659	0.9526

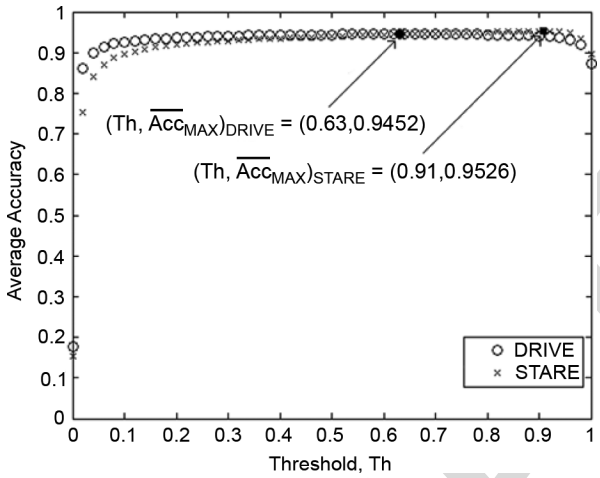


Fig. 5. Acc of the segmentation algorithm as a function of the threshold parameter Th for the DRIVE and STARE databases.

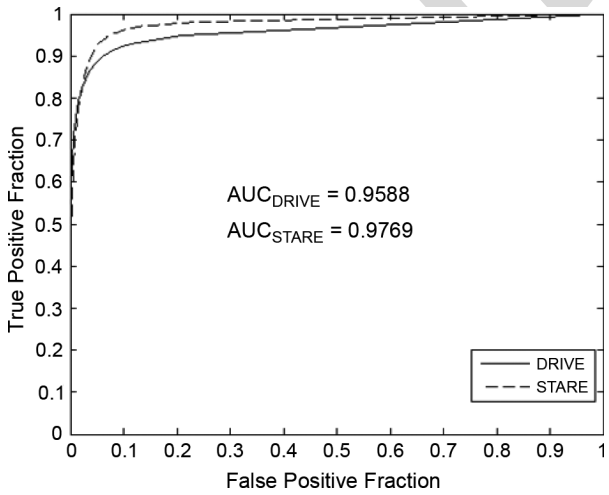


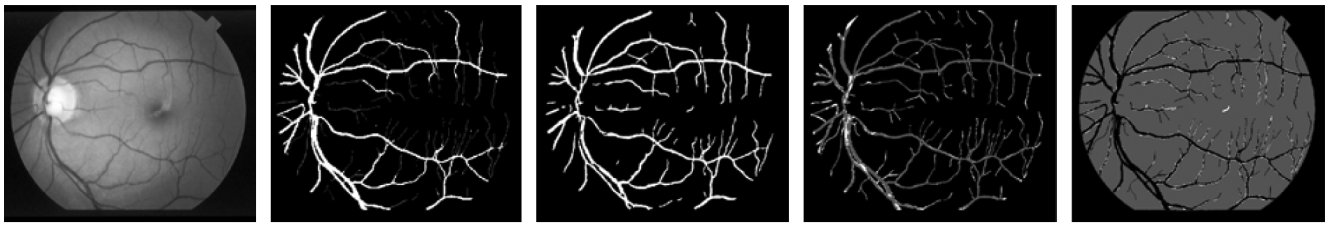
Fig. 6. ROC curves for the DRIVE and STARE databases. Measured AUC values are given.

in pixels near to and far from vessel borders. For that, thin and non-thin vessels were separated in each gold standard image [as an example, see Fig. 7(b)]. A vessel was considered thin if its width is lower than 50% of the width of the widest optic disc vessel. Otherwise the vessel is considered non-thin. On the other hand, a FP is considered to be far from a vessel border if the distance from its nearest vessel border pixel in the gold-standard is over two pixels. Otherwise, the FP is considered to be near. Table VI summarizes the results of this study. This table shows the average ratio of FN and FP provided by the segmentation algorithm for the 20 test images in the DRIVE and STARE databases. The average percent of FN and FP corresponding to the different spacial locations considered are also shown. For both databases, the percent of FN produced in non-thin vessel pixels was higher than that in thin vessel pixels. However, taking into account that thin vessels are composed by a considerably lower number of pixels than non-thin ones, the value obtained for thin vessels bears a more negative impact. This can be observed in Fig. 4(d), and Fig. 7(d). While FN in non-thin vessels involve no degradation in the segmented structure, FN in thin vessels produce disconnections in some of them. Regarding FP distribution, FPs tend to be near vessel borders. As it can be checked in Fig. 7(e), this means that most FPs produced by the segmentation algorithm tend to slightly enlarge the vessels and not to introduce meaningless isolated noise.

C. Comparison to Other Methods

In order to compare our approach to other retinal vessel segmentation algorithms, $\overline{\text{Acc}}$ and AUC were used as measures of method performance. Since these measurements were performed by other authors, this choice facilities comparing our results to theirs. Tables VII and VIII show performance comparison results in terms of $\overline{\text{Acc}}$ and AUC, respectively, with the following published methods: Chaudhuri *et al.* [37], Hoover *et al.* [38], Jiang and Mojon [43], Niemeijer *et al.* [50], Staal *et al.* [51], Mendonça *et al.* [36], Soares *et al.* [52], Martinez-Perez *et al.* [46], Ricci and Perfetti, [53], and Cinsdikici and Aydin, [42]. All these supervised or rule-based methods have been briefly commented in Section II. The values shown in both tables are presented for each database as reported by their authors. If they are not available for a specific database or were not calculated for the 20 images selected for testing, they were not included in the tables, thus appearing as gaps. The values in the last column of each table indicate the overall $\overline{\text{Acc}}$ and AUC when both databases are taken into account.

An overview of the segmentation results on DRIVE images shows our proposed method reaches better performance than most of the other methods, being comparable to or approximating the performance of other detection techniques. The $\overline{\text{Acc}}$ value achieved with our algorithm is outperformed only by Soares *et al.* [52], Mendonça *et al.* [36], and Ricci and Perfetti [53]. Regarding the approaches by Soares *et al.* [52] and Mendonça *et al.* [36], it is important to point out that our method clearly outperforms the $\overline{\text{Acc}}$ these authors reported on STARE images. Therefore, our approach renders better overall $\overline{\text{Acc}}$ for both databases than theirs. The same conclusions are drawn when these methods are compared in terms of AUC. On DRIVE database images, the AUC value provided by our proposal is only lower than those reported by Soares *et al.* [52] and Ricci and Perfetti [53] (Mendonça *et al.* [36] did not report



(a) (b) (c) (d) (e)

Fig. 7. Illustration of the spacial location of classification errors on a segmentation of a STARE image. (a) Green channel of the original image. (b) Thin and non-thin blood vessels extracted from the manual segmentation in white and dark-gray colors, respectively. (c) Segmentation of (a) generated by the presented algorithm. (d) FN and TP obtained by the proposed algorithm represented in white and dark-gray colors, respectively. (e) FP and TN obtained by the proposed algorithm represented in white and dark-gray colors, respectively.

TABLE VI
STUDY OF FN AND FP SPACIAL LOCATION ON THE DRIVE AND STARE DATABASES

	FN		FP	
Database	**FN average rate**		**FP average rate**	
DRIVE	0.2933		0.3056	
STARE	0.0199		0.0181	
	Thin Vessels	***Non-thin Vessels***	***Near to Vessels***	***Far from Vessels***
DRIVE	46.08%	53.92%	86.25%	13.75%
STARE	34.60%	65.40%	83.55%	16.45%

TABLE VII
PERFORMANCE RESULTS COMPARED TO OTHER METHODS ON THE STARE AND DRIVE DATABASES IN TERMS OF AVERAGE ACCURACY

Method Type	Method	DRIVE	STARE	DRIVE + STARE
Supervised	Staal *et al.* [51]	0.9441	-	-
	Niemeijer *et al.* [50]	0.9417	-	-
	Soares *et al.* [52]	0.9466	0.9480	0.9473
	Ricci and Perfetti [53]	**0.9595**	**0.9646**	**0.9621**
	Marin *et al.* (this work)	0.9452	0.9526	0.9489
Ruled-Based	Chaudhuri *et al.* [37]	0.8773	-	-
	Hoover *et al.* [38]	-	0.9275	-
	Jiang and Mojon [43]	0.8911	0.9009	0.8960
	Mendonça *et al.* [36]	0.9463	0.9479	0.9471
	Martinez-Perez *et al.* [46]	0.9344	0.9410	0.9377
	Cinsdikici and Aydın [42]	0.9293	-	-

TABLE VIII
PERFORMANCE RESULTS COMPARED TO OTHER METHODS ON THE STARE AND DRIVE DATABASES IN TERMS OF AREA UNDER ROC CURVE

Method Type	Method	DRIVE	STARE	DRIVE + STARE
Supervised	Staal *et al.* [51]	0.9520	-	-
	Niemeijer *et al.* [50]	0.9294	-	-
	Soares *et al.* [52]	0.9614	0.9671	0.9642
	Ricci and Perfetti [53]	**0.9633**	0.9680	0.9656
	Marin *et al.* (this work)	0.9588	**0.9769**	**0.9678**
Ruled-Based	Chaudhuri *et al.* [37]	0.7878	-	-
	Hoover *et al.* [38]	-	0.7590	-
	Jiang and Mojon [43]	0.9327	0.9298	0.9312
	Mendonça *et al.* [36]	-	-	-
	Martinez-Perez *et al.* [46]	-	-	-
	Cinsdikici and Aydın [42]	0.9407	-	-

AUC values). However, due to the excellent AUC result on the STARE database, our approach reaches the highest average AUC when both databases are considered.

The proposed method proves especially useful for vessel detection in STARE images. Its application to this database resulted in the second highest accuracy score among all experiments (only behind Ricci and Perfetti's approach [53]) and the first when AUC is the reference measurement. This result gains more importance by the fact that our classifier was trained only on DRIVE images, unlike the other supervised approaches presented in Tables VII and VIII. For instance, since there are no available labeled training images for STARE images, Soares *et al.* [52] performed leave-one-out tests on this database (i.e., every image is classified by using samples from the other 19 images), while Ricci and Perfetti [53] built its classifier by using a training set comprising samples randomly extracted from test images. In our case, with the purpose of using one and the same trained classifier for testing the method on the 20 STARE images and including no sample belonging to the test set in the training, we opted for forming the training set by collecting pixels from DRIVE training images. Thus, the method's suitability for being applied on any fundus image can be checked in a more realistic way. We should also mention that these good results with respect to other existing approaches were obtained on images containing pathological artifacts. The STARE database contains ten images with pathologies, while the test of DRIVE only contains four. Moreover, abnormal regions are wider in STARE.

Regarding performance comparison in terms of $\overline{\text{Acc}}$ when results are jointly analyzed for DRIVE and STARE images (Table VII, last column), our algorithm renders greater accuracy than others authors' algorithms, being outperformed only by Ricci and Perfetti's proposal [53]. However, this method proved very dependent on the training set. Ricci and Pefetti [53], to research the dependence of their classification method on the dataset, carried out an experiment based on, firstly, training the classifier on each of the DRIVE and STARE databases, and then, testing it on the other. Their maximum accuracy values are shown in Table IX. It can be observed that performance is worse now, since $\overline{\text{Acc}}$ strongly decreases from 0.9595 to 0.9266 on DRIVE and 0.9646 to 0.9452 on STARE database images. Therefore, as assumed by these authors, classifier retraining is necessary before applying their methodology on a new database. To verify our method dependence on the training set, the same experiment was completed. Thus, performance was computed on the DRIVE database, training the classifier with STARE images (as previously mentioned, our accuracy on STARE was already obtained by training on DRIVE images). The resulting $\overline{\text{Acc}}$ values are shown in Table IX to facilitate comparisons between both methods under identical conditions. In this case, it is clearly observed that our estimated performance in terms of method accuracy is higher, thus proving higher training set robustness.

TABLE IX
PERFORMANCE RESULTS COMPARED TO RICCI AND PERFETTI'S METHOD WITH CROSS TRAINNING IN TERMS OF AVERAGE ACCURACY

Method	DRIVE (training on STARE)	STARE (training on DRIVE)
Ricci and Perfetti [53]	0.9266	0.9452
Marin *et al.* (this work)	**0.9448**	**0.9526**

TABLE X
AVERAGE ACCURACY ON THE DRIVE AND STARE DATABASES USING ONE TYPE OF FEATURE: GRAY-LEVEL (G-L) OR MOMENT INVARIANTS-BASED FEATURES (M-I)

Feature Vector Composition	DRIVE	STARE
G-L.	0.9407	0.9514
M-I.	0.9398	0.9410

VI. DISCUSSION AND CONCLUSION

Previous methods for blood vessel detection in retinal images can be classified into rule-based and supervised methods. This study proposes a method within the latter category. This method is based on a NN scheme for pixel classification, being the feature vector representing each pixel composed of gray-level and moment invariants-based features. To the best of our knowledge, although moment invariants have been widely used over the years as features for pattern recognition in many areas of image analysis (typical examples include optical character recognition and shape identification), they have never been applied within this framework.

The experiments aimed at evaluating the efficiency of the applied descriptors prove this method is capable of rendering accurate results, even when these types of features are used independently. Table X shows its accuracy performance when a feature vector composed of the five gray-level-based features, on one hand, and the two moment invariants-based features, on the other hand, is considered. According to these results, both sets of descriptors are suitable for pixel classification as vessel or nonvessel. Even when the feature vector is built as mentioned above, our algorithm outperforms other segmentation proposals (e.g., Jiang and Mojon [43] or Martinez-Perez *et al.* [46]). However, these pixel representations are not optimal by themselves. Discriminative power increases when both sets of different feature types are jointly considered. Thus, accuracy improves up to 0.9452 and 0.9526 for the 20 test images in the DRIVE and STARE databases, respectively. Therefore, the method finally adopts a 7-D feature vector composed by the five gray-level and the two moment invariants-based features.

The proposed method uses a NN for pixel classification as vessel or non-vessel. This classifier was selected after method accuracy assessment by means of a kNN and a SVM (used as in [51] and [53], respectively), instead of a NN. Table XI shows this performance comparison in terms of $\overline{\text{Acc}}$ measured on DRIVE and STARE test images. Results before and after the postprocessing stage application are presented. NN showed better accuracy than kNN and SVM for all cases.

Tables VII and VIII show a performance overview on previously-published vessel segmentation methods in terms of $\overline{\text{Acc}}$ and AUC, respectively, for test DRIVE and STARE images. Among them, those providing best performance results include

TABLE XI
AVERAGE ACCURACY ON THE DRIVE AND STARE DATABASES OF THE PROPOSED METHOD WHEN NN, KNN, AND SVM CLASSIFIERS ARE USED

	Before postprocessing		After postprocessing	
Classifier	**DRIVE**	**STARE**	**DRIVE**	**STARE**
NN	0.9439	0.9517	0.9452	0.9526
kNN	0.9383	0.9500	0.9415	0.9524
SVM	0.9393	0.9498	0.9427	0.9519

the rule-based proposal by Mendonça *et al.* [36] (AUC is not available for this method), and the supervised methods by Soares *et al.* [52] and Ricci and Perfetti [53].

When the results on both databases are jointly analysed, the approach proposed in this paper provides the highest average AUC. Regarding $\overline{\text{Acc}}$, the value reached in this work is slightly higher than that obtained by Mendonça *et al.* [36] and Soares *et al.* [52]. Although Ricci and Perfetti's approach [53] seems to offer more accurate behavior, the results in Table IX indicate that their method shows strong performance loss when confronted to an "unseen" database. These authors obtain excellent results when both training and testing are performed on the same database but accuracy significantly worsens when the method is trained and tested on a different database. A possible explanation to this fact is that authors avoid image preprocessing to preserve vessel structure at most. However, since the different images are not homogenized, the method proves very sensitive to the training set. This is an important disadvantage for practical application, since a blood vessel detection tool must work on retinal images from multiple origins and be used by different operators working with different equipment. On the contrary, our proposal proves robust regarding the dataset used: even when trained on the DRIVE database, its application to the STARE database results in the second highest accuracy score among all experiments, and the first if the Ricci and Pefetti's [53] results are recorded under the same conditions. Therefore, the training set robustness shown by our method allows its automated application to images taken under different conditions. To this respect, it should be pointed out that the proposed method, in its different stages (preprocessing, feature extraction, classification, and postprocessing), uses a set of parameters fixed to provide the best accuracy on DRIVE test images. That is, parameter values were optimized to retinas of 540 pixels in diameter. It is important to point out that, since the features used for pixel representation are invariant to scale change, the method is also applicable for processing images with different resolution without retraining. To make the method applicable for processing images of any resolution, its implementation should include either resampling the image to fulfil the mentioned size condition, or adapting all parameter to the new retina size.

In addition, method simplicity should also be highlighted. Its pixel classification procedure is based on computing only seven features for each pixel, thus needing shorter computational time. The total time required to process a single image is less than approximately one minute and thirty seconds, running on a PC with an Intel Core2Duo CPU at 2.13 GHz and 2 GB of RAM. Since our implementation is experimental, this performance might still be improved.

The demonstrated effectiveness and robustness, together with its simplicity and fast implementation, make this proposed automated blood vessel segmentation method a suitable tool for being integrated into a complete prescreening system for early DR detection.

ACKNOWLEDGMENT

The authors would like to thank J. J. Staal and his colleagues and A. Hoover for making their databases publicly available. This work was carried out as part of the Expert System for Early Automated Detection of DR by Analysis of Digital Retinal Images Project, supported and funded by the Health Ministry of the Andalusian Regional Government (Spain).

REFERENCES

[1] H. R. Taylor and J. E. Keeffe, "World blindness: A 21st century perspective," *Br. J. Ophthalmol.*, vol. 85, pp. 261–266, 2001.

[2] R. Klein, S. M. Meuer, S. E. Moss, and B. E. Klein, "Retinal microaneurysm counts and 10-year progression of diabetic retinopathy," *Arch. Ophthalmol.*, vol. 113, pp. 1386–1391, 1995.

[3] P. Massin, A. Erginay, and A. Gaudric, *Rétinopathie Diabétique*. New York: Elsevier, 2000.

[4] 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.

[5] S. J. Lee, C. A. McCarty, H. R. Taylor, and J. E. Keeffe, "Costs of mobile screening for diabetic retinopathy: A practical framework for rural populations," *Aust. J. Rural Health*, vol. 8, pp. 186–192, 2001.

[6] D. S. Fong, L. Aiello, T. W. Gardner, G. L. King, G. Blankenship, J. D. Cavallerano, F. L. Ferris, and R. Klein, "Diabetic retinopathy," *Diabetes Care*, vol. 26, pp. 226–229, 2003.

[7] "Economic costs of diabetes in the U.S. in 2007," in *Diabetes Care*. : American Diabetes Association, 2008, vol. 31, pp. 596–615.

[8] American Academy of Ophthalmology Retina Panel, Preferred Practice Pattern Guidelines. Diabetic Retinopathy. San Francisco, CA, Am. Acad. Ophthalmo., 2008 [Online]. Available: http://www.aao.org/ppp

[9] T. Spencer, J. A. Olson, K. C. McHardy, P. F. Sharp, and J. V. Forrester, "An image-processing strategy for the segmentation and quantification of microaneurysms in fluorescein angiograms of the ocular fundus," *Comput. Biomed. Res.*, vol. 29, no. 4, pp. 284–302, 1996.

[10] A. J. Frame, P. E. Undrill, M. J. Cree, J. A. Olson, K. C. McHardy, P. F. Sharp, and J. V. Forrester, "A comparison of computer based classification methods applied to the detection of microaneurysms in ophthalmic fluorescein angiograms," *Comput. Biol. Med.*, vol. 28, no. 3, pp. 225–238, 1998.

[11] M. Larsen, J. Godt, N. Larsen, H. Lund-Andersen, A. K. Sjølie, E. Agardh, H. Kalm, M. Grunkin, and D. R. Owens, "Automated detection of fundus photographic red lesions in diabetic retinopathy," *Investigat. Opht. Vis. Sci.*, vol. 44, no. 2, pp. 761–766, 2003.

[12] 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.

[13] 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.

[14] G. K. Matsopoulos, P. A. Asvestas, N. A. Mouravliansky, and K. K. Delibasis, "Multimodal registration of retinal images using self organizing maps," *IEEE Trans. Med. Imag.*, vol. 23, no. 12, pp. 1557–1563, Dec. 2004.

[15] C. Heneghan, J. Flynn, M. O'Keefe, and M. Cahill, "Characterization of changes in blood vessel width and tortuosity in retinopathy of prematurity using image analysis," *Med. Image Anal.*, vol. 6, pp. 407–429, 2002.

[16] E. Grisan and A. Ruggeri, "A divide and impera strategy for the automatic classification of retinal vessels into arteries and veins," in *Proc. 25th Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2003, pp. 890–893.

[17] Y. Hatanaka, H. Fujita, M. Aoyama, H. Uchida, and T. Yamamoto, "Automated analysis of the distribuitions and geometries of blood vessels on retinal fundus images," *Proc. SPIE Med. Imag. 2004: Image Process.*, vol. 5370, pp. 1621–1628, 2004.

[18] M. Foracchia, E. Grisan, and A. Ruggeri, "Extraction and quantitative description of vessel features in hypertensive retinopathy fundus images," in *Book Abstracts 2nd Int. Workshop Comput. Asst. Fundus Image Anal.*, 2001, p. 6.

[19] X. Goa, A. Bharath, A. Stanton, A. Hughes, N. Chapman, and S. Thom, "A method of vessel tracking for vessel diameter measurement on retinal images," *Proc. ICIP*, pp. 881–884, 2001.

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

[21] 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.

[22] D. E. Becker, A. Can, J. N. Turner, H. L. Tanenbaum, and B. Roysam, "Image processing algorithms for retinal montage, synthesis, mapping and real-time location determination," *IEEE Trans. Biomed. Eng.*, vol. 45, no. 1, pp. 115–118, Jan. 1998.

[23] H. Shen, B. Roysam, C. V. Stewart, J. N. Turner, and H. L. Tanenbaum, "Optimal scheduling of tracing computations for real-time vascular landmark extraction from retinal fundus images," *IEEE Trans. Inf. Technol. Biomed.*, vol. 5, pp. 77–91, Mar. 2001.

[24] 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.

[25] M. Foracchia, E. Grisan, and A. Ruggeri, "Detection of optic disc in retinal images by means of a geometrical model of vessel structure," *IEEE Trans. Med. Imag.*, vol. 23, no. 10, pp. 1189–1195, Oct. 2004.

[26] A. A. H. A. R. Youssif, A. Z. Ghalwash, and A. R. Ghoneim, "Optic disc detection from normalized digital fundus images by means of a vessels' direction matched filter," *IEEE Trans. Med. Imag.*, vol. 27, no. 1, pp. 11–18, Jan. 2008.

[27] H. Li and O. Chutatape, "Automated feature extraction in color retinal images by a model based approach," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 2, pp. 246–254, Feb. 2004.

[28] O. Chutatape, L. Zheng, and S. Krishman, "Retinal blood vessel detection and tracking by matched Gaussian and Kalman filters," in *Proc. IEEE Int. Conf. Eng. Biol. Soc.*, 1998, vol. 20, pp. 3144–3149.

[29] 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. 2, pp. 263–273, Apr. 1998.

[30] 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.

[31] L. Gagnon, M. Lalonde, M. Beaulieu, and M.-C. Boucher, "Procedure to detect anatomical structures in optical fundus images," *Proc. SPIE Med. Imag.: Image Process.*, vol. 4322, pp. 1218–1225, 2001.

[32] 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, Jun. 1993.

[33] L. Zhou, M. S. Rzeszotarski, L. J. Singerman, and J. M. Chokreff, "The detection and quantification of retinopathy using digital angiograms," *IEEE Trans. Med. Imag.*, vol. 13, no. 4, pp. 619–626, Dec. 1994.

[34] T. Walter and J. C. Klein, "Segmentation of color fundus images of the human retina: Detection of the optic disc and the vascular tree using morphological techniques," in *Medical Data Analysis, ser. Lecture Notes in Computer Science*, J. Crespo, V. Maojo, and F. Martin, Eds. Berlin, Germany: Springer-Verlag, 2001, pp. 282–287.

[35] 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.

[36] 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.

[37] S. Chaudhuri, 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, Sep. 1989.

[38] 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.

[39] L. Gang, O. Chutatape, and S. M. Krishnan, "Detection and measurement of retinal vessels in fundus images using amplitude modified second-order Gaussian filter," *IEEE Trans. Biomed. Eng.*, vol. 49, pp. 168–172, Feb. 2002.

[40] M. Al-Rawi and H. Karajeh, "Genetic algorithm matched filter optimization for automated detection of blood vessels from digital retinal images," *Comput. Methods Programs Biomed.*, vol. 87, pp. 248–253, 2007.

[41] M. Al-Rawi, M. Qutaishat, and M. Arrar, "An improved matched filter for blood vessel detection of digital retinal images," *Comput. Biol. Med.*, vol. 37, pp. 262–267, 2007.

[42] M. G. Cinsdikici and D. Aydin, "Detection of blood vessels in ophthalmoscope images using MF/ant (matched filter/ant colony) algorithm," *Comput. Methods Programs Biomed.*, vol. 96, pp. 85–95, 2009.

[43] 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.

[44] T. McInerney and D. Terzopoulos, "T-snakes: Topology adaptive snakes," *Med. Imag. Anal.*, vol. 4, pp. 73–91, 2000.

[45] L. Espona, M. J. Carreira, M. Ortega, and M. G. Penedo, "A snake for retinal vessel segmentation," *Pattern Recognition and Image Analysis*, vol. 4478, Lecture Notes Comput. Sci., pp. 178–185, 2007.

[46] M. E. Martinez-Perez, A. D. Hughes, S. A. Thom, A. A. Bharath, and K. H. Parker, "Segmentation of blood vessels from red-free and fluorescein retinal images," *Med. Imag. Anal.*, vol. 11, pp. 47–61, 2007.

[47] B. S. Y. Lam and H. Yan, "A novel vessel segmentation algorithm for pathological retina images based on the divergence of vector fields," *IEEE Trans. Med. Imag.*, vol. 27, no. 2, pp. 237–246, Feb. 2008.

[48] G. G. Gardner, D. Keating, T. H. Williamson, and A. T. Elliott, "Automatic detection of diabetic retinopathy using an artificial neural network: A screening tool," *Br. J. Ophthalmol.*, vol. 80, pp. 940–944, 1996.

[49] 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.

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

[51] J. Staal, M. D. Abràmoff, 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.

[52] 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.

[53] E. Ricci and R. Perfetti, "Retinal blood vessel segmentation using line operators and support vector classification," *IEEE Trans. Med. Imag.*, vol. 26, no. 10, pp. 1357–1365, Oct. 2007.

[54] Research Section, Digital Retinal Image for Vessel Extraction (DRIVE) Database. Utrecht, The Netherlands, Univ. Med. Center Utrecht, Image Sci. Inst. [Online]. Available: http://www.isi.uu.nl/Re-search/Databases/DRIVE

[55] STARE ProjectWebsite. Clemson, SC, Clemson Univ. [Online]. Available: http://www.ces.clemson.edu/

[56] 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.

[57] B. Zhang, X. Wu, J. You, Q. Li, and F. Karray, "Detection of microaneurysms using multi-scale correlation coefficients," *Pattern Recognit.*, vol. 43, pp. 2237–2248, 2010.

[58] L. Streeter and M. J. Cree, "Microaneurysm detection in colour fundus images," in *Image Vision Comput.*, Nov. 2003, pp. 280–284.

[59] R. Jagoe, J. Arnold, C. Blauth, P. L. C. Smith, K. M. Taylor, and R. Wootton, "Measurement of capillary dropout in retinal angiograms by computerised image analysis," *Pattern Recognit. Lett*, vol. 13, pp. 143–151, 1992.

[60] M. K. Hu, "Visual pattern recognition by moment invariants," *IRE Trans. Inform. Theory*, vol. IT-8, pp. 179–187, 1962.

[61] S. Bioux, M. Martin-Fernandez, L. Ungar, M. Nakamura, M. -. Koo, R. W. McCarley, and M. E. Shenton, "On evaluating brain tissue classifiers without a ground truth," *NeuroImage*, vol. 36, no. 4, pp. 1207–1224, 2007.

[62] D. E. Rumelhart and J. L. McClellandthe, *Parallel Distributed Processing, Explanations in the Micro Structure of Cognition, 1: Foundations. A Bradford Book*. Cambridge, MA: MIT Press, 1988.