Automatic Segmentation of Exudates in Ocular Images using Ensembles of Aperture Filters and Logistic Regression

2013 J. Phys.: Conf. Ser. 477 012021

(http://iopscience.iop.org/1742-6596/477/1/012021)

Automatic Segmentation of Exudates in Ocular Images using Ensembles of Aperture Filters and Logistic Regression

Marco Benalcázar[1,2,3], Marcel Brun[3], and Virginia Ballarin[3]

[1] Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina.
[2] Secretaría Nacional de Educación Superior, Ciencia, Tecnología e Innovación (SENESCYT), Ecuador.
[3] Grupo de Procesamiento Digital de Imágenes, Universidad Nacional de Mar del Plata, Mar del Plata, Argentina.

E-mail: marco_benalcazar@hotmail.com

Abstract. Hard and soft exudates are the main signs of diabetic macular edema (DME). The segmentation of both kinds of exudates generates valuable information not only for the diagnosis of DME, but also for treatment, which helps to avoid vision loss and blindness. In this paper, we propose a new algorithm for the automatic segmentation of exudates in ocular fundus images. The proposed algorithm is based on ensembles of aperture filters that detect exudate candidates and remove major blood vessels from the processed images. Then, logistic regression is used to classify each candidate as either exudate or non-exudate based on a vector of 31 features that characterize each potensial lesion. Finally, we tested the performance of the proposed algorithm using the images in the public HEI-MED database.

1. Introduction

Exudates are one the main signs of diabetic macular edema (DME), which occurs when the retina swells as a complication of diabetic retinopathy. There are two kinds of exudates, hard and soft. Hard exudates are composed of lipid and protein deposits that are leaked from the bloodstream into the retina. Soft exudates are superficial retinal microinfarcts. Both hard and soft exudates are direct consequences of the obstruction of the ocular blood vessels due to elevated blood sugar levels. The presence of exudates within the central macula, where the majority of photoreceptors are concentrated, can cause vision loss or even blindness in diabetic patients. In fundus ocular images of the retina, exudates appear as yellowish structures with well defined edges and variable shapes [1-4].

Diabetic retinopathy, and consequently DME, can be inhibited by early diagnosis and appropriate treatment. However, since diabetic retinopathy is asymptomatic in early stages, many people remain undiagnosed until they start losing their vision due to the increasing amount of hard and soft exudates and other pathologies that accompany diabetic retinopathy such as mycroaneurysms and new blood vessels. In the ideal case, a mass screening of all diabetic patients, even of those experiencing no vision issues, would help to diagnose DME early enough for optimal treatment [1]. At this time, such an undertaking would be too work and time intensive because each image must be analyzed by a specialist, making the early diagnosis of this pathology difficult. On the other hand, for those cases when diabetic retinopathy has been already diagnosed, the detection of exudates in fundus ocular

images generates valuable information for treatment planning. Namely, the position of exudates is important for laser treatment, which is the photocoagulation of the lesions to stop their progress and prevent blindness [3].

Several issues must be addressed when applying computer aided methods for the segmentation of exudates in fundus ocular images. First, hard and soft exudates have similar attributes to those of the optic disk in terms of color, brightness, contrast, and texture. Second, hard exudates have well defined structures; whereas, soft exudates have unclear boundaries. Third, in fundus ocular images, variations in illumination, contrast, and color in regions within the images (intra-image variability) as well as among the images (inter-image variability) occur. Intra-image variations are caused by differences in light diffusion inside the eye, lesions, and variations in reflectivity and thickness of the retina. Inter-image variations are caused by differences in the acquisition conditions of the images (e.g., illumination and acquisition angle), materials (e.g., cameras), and the ethnic group from which the screened patients come [1,3].

Because of the aforementioned issues, the automatic segmentation of exudates is still an open field of research, especially when we want to develop algorithms that perform well in a broad spectrum of application. In the present work, we combine tools from digital image processing, pattern recognition, and machine learning to propose a new algorithm for the automatic segmentation of both hard and soft exudates. It is important to mention that in this work, we make no distinction between hard and soft exudates. One reason is because differentiating between them is problematic even for specialized ophthalmologists. Additionally, both kinds of exudates are direct consequences of diabetic retinopathy, and their differentiation poses no significant clinical advantage for the diagnosis of DME [1-4]. Therefore, hard and soft exudates will be referred to simply as exudates from now on.

This paper is organized as follows. In Section 2, we describe both the source of the images as well as the methodology used in this work. In Section 3, we present the proposed algorithm for automatically segmenting exudates. Section 4 contains the results obtained and compares them with the results presented in recent works published in the scientific literature. Finally, in Section 5, we summarize the main contributions of this work and outline future possible developments.

2. Images and Methodology

2.1. Images

We used the public Hamilton Eye Institute Macular Edema database (HEI-MED) [2]. This database is composed of 169 fundus ocular images (color images) with their associated ground truth images (binary images). These images, color and binary, are formed by 2196x1958 pixels. The binary images correspond to manual segmenetations of the regions containing exudates. Small lesions were clustered together for those cases in which the individual ones were not clearly discernible. The database contains images from a mixture of ethnicities and stages of diabetic retinopathy and includes images without any pathology. For this work, we randomly divided the 169 images into 3 subsets: TrS_1, TrS_2 and TtS. TrS_1 and TrS_2 are composed of 30 images each and are used for the training stage of the proposed algorithm. TtS contains the reamining 109 images, which were used for testing.

2.2. Methodology

In this subsection, we explain some theoretical aspects of the methods that compose the different steps of the proposed algorithm.

2.2.1. Color to Grayscale Transformation

Let the function $G:\mathbf{E}\to L$ be a *grayscale imag*e. $\mathbf{E}$ is a nonempty subset of $\mathbf{Z^2}$, and $L=[0,l]$, with $l\in\mathbf{Z}^+$, represents a gray-level interval. For usual applications we consider $l=255$. The set of all possible images from $\mathbf{E}$ to L is denoted by $L^{\mathbf{E}}$. A *color image* is modeled as a function $\overline{H}:\mathbf{E}\to[L_R\times L_G\times L_B]$, usually $L_R=L_G=L_B=[0,l]$, such that $\overline{H}=(H_R,H_G,H_B)$, with $H_R,H_G,H_B\in L^{\mathbf{E}}$, where H_R, H_B and H_G are the channels of the image $\overline{H}$. In the particular case of RGB images, H_R, H_B and H_G are the red, green and blue channels, respectively. The set of possible color images from $\mathbf{E}$ to $[L_R\times L_G\times L_B]$ is denoted by $[L_R\times L_G\times L_B]^{\mathbf{E}}$. A *pixel* t of the image $\overline{H}$ is a coordinate in $\mathbf{E}$ of the form $t=(x,y)$ that takes the triplet $\overline{H}(t)=(H_R(t),H_G(t),H_B(t))$, where $H_R(t),H_G(t),H_B(t)\in[0,l]$.

Given an RGB image $\overline{H}$ and an arbitrary point $t\in\overline{H}$, the function $\Omega:[L_R\times L_G\times L_B]^{\mathbf{E}}\to L^{\mathbf{E}}$ is a transformation of RGB into grayscale images, such that

$$G(t)=\Omega(\overline{H}(t))=round\left(\alpha_1H_R(t)+\alpha_2H_G(t)+\alpha_3H_B(t)\right), \tag{1}$$

where $\alpha_1+\alpha_2+\alpha_3=1$, and $round(\bullet)$ rounds to nearest integer.

2.2.2. Aperture Filters

Formally, let a *spatial window* $W=\{w_1,w_2,...,w_n\}$, with $w_i\in\mathbf{Z^2}$ and $i=1,2,...,n$, be a small subset of $\mathbf{Z^2}$, centered at the origin of $\mathbf{E}$: $(0,0)\in W$. The number of points of W is denoted by $n=|W|$. Moreover, let consider a *range window* $K=[-k,k]$, with $k\in\mathbf{Z}^+$ and $k<<l$, be an interval that represents the number of input gray-levels for filtering. Finally, let an *aperture* $W\times K$ be the Cartesian product of a spatial window $W=\{w_1,w_2,...,w_n\}$ and a range window $K=[-k,k]$. Therefore, given the grayscale image O and the aperture $W\times K$, an *aperture configuration* (also called *observation*) is a function $\mathbf{v}:W\to K$, so that $\mathbf{v}=(v_1,...,v_n)$ with $n=|W|$. We obtain an aperture configuration $\mathbf{v}$ from the image O by applying the following two steps: 1. The *spatial translation* of O by t: $O_{-t}=O(t'-t)$, being $t,t'\in\mathbf{E}$, and then the observation of the pixel values of O_{-t} within the spatial window W. This step returns the vector $\mathbf{u}_{O(t)}=(O(w_1-t),...,O(w_n-t))$. For simplicity of notation, we will refer to vector $\mathbf{u}_{O(t)}$ with $\mathbf{u}$ from now on. 2. The *range translation* of $\mathbf{u}$ by a scalar $z\in\mathbf{Z}^+$: $\mathbf{u}_{-z}=(O(w_1-t)-z,...,O(w_n-t)-z)$, followed by the projection of points of $\mathbf{u}_{-z}$ outside $K=[-k,k]$ to either $+k$ or $-k$. Formally, and denoting the projection operation by $(\bullet)^*$, each point $v_i\in\mathbf{v}$, with $i=1,2,...,n$, is obtained from a given observation $\mathbf{u}=(u_1,...,u_n)$ as follows:

$$v_i=(u_i)^*=\min(\max(-k,u_i-z),k). \tag{2}$$

Given the definitions above, let an *aperture filter* $\mathbf{\Psi}:L^{\mathbf{E}}\to\{0,1\}^{\mathbf{E}}$ be a function that maps grayscale images into binary images. That is, given the grayscale image O, then $I=\mathbf{\Psi}(O)$ is a binary image so that $I:\mathbf{E}\to\{0,1\}$. Aperture filters are characterized by functions of the form $\psi:K^W\to\{0,1\}$, called *characteristic functions*, so that for each $t\in\mathbf{E}$:

$$I(t)=\mathbf{\Psi}(O)(t)=\psi((O(w_1-t)-z)^*,...,(O(w_n-t)-z)^*)=\psi(\mathbf{v}_{O(t)}), \tag{3}$$

where z is a function $z: L^W \to L$ that allows us to place the aperture $W \times K$ vertically in the pixel $t \in O$. In this work, we chose z to be $z = median(u_1,...,u_n)$.

Moreover, in the language of pattern recognition, the characteristic function ψ can be seen as a *classifier* that maps observations $\mathbf{v} \in \{\mathbf{v}_1,...,\mathbf{v}_m\}$ to labels from the set $\{0,1\}$. Based on this assumption, the segmentation of a grayscale image O is reduced to a classification of its pixels within $W \times K$. Finally, aperture filters fulfill three important properties: 1. *spatial translation invariance*: $\mathbf{\Psi}(O_t) = \mathbf{\Psi}(O)_t$ for each pixel $t \in O$; 2. *range translation invariance*: $\mathbf{\Psi}(O(t)+z) = \mathbf{\Psi}(O(t)) + z$ for any $z \in \mathbf{Z}^+$; and 3. *local definition* within $W \times K$: $\mathbf{\Psi}(O)(t) = \mathbf{\Psi}(O \cap (W \times K))(t)$ [5,6].

2.2.3. Statiscal Design of Aperture Filters

In the section above, we reduced the segmentation of an image to a classification of its pixels. Based on this fact, the statistical design of aperture filters is reduced to the well known problem of designing classifiers in the field of pattern recognition. In this context, we model the images to be processed as random realizations of the two processes ***O*** and ***I***, which are stochastic, discrete, stationary, and joint**.** Process ***O*** produces observed images, and process ***I*** produces ideal images [6,7]. *Observed images* represent the problem to be solved; whereas, *ideal images* represent the desired output of the processing.

Thus, given a pair of images (*O*,*I*) drawn from the processes (***O***,***I***), the goal of the statistical design is to find an aperture filter $\mathbf{\Psi}$ so that $\mathbf{\Psi}(O)$ is as close as possible to I. If the measure of closeness between $\mathbf{\Psi}(O)$ and *I* is the mean square error (MSE), then the cost function to be minimized is

$$\varepsilon(\mathbf{\Psi}) = MSE = E[(1/|\mathbf{E}|)\sum_{t \in \mathbf{E}} (I(t) - \mathbf{\Psi}(O)(t))^2], \tag{4}$$

where *E* is the expectation operator and t is an arbitrary point of $\mathbf{E}$, which is possible because (***O***,***I***) are stationary processes. When $\mathbf{\Psi}$ is characterized by ψ; the processes are stationary; and the observation $\mathbf{v} = (v_1,...,v_n)$ is an *n*-upla of random variables, v_i, then equation 4 yields

$$\varepsilon(\mathbf{\Psi}) = \varepsilon(\psi) = \sum_{\{\mathbf{v}:\, \mathbf{v} \in O \cap W(t) \times K\}} \Pr(I(t) \neq \psi(\mathbf{v}) | \mathbf{v}) \Pr(\mathbf{v}), \tag{5}$$

where $\Pr(I(t) \neq \psi(\mathbf{v})|\mathbf{v})$ represents the probability of misclassifying an observation $\mathbf{v}$ seen by $W \times K$ placed at $t \in \mathbf{E}$, and $\Pr(\mathbf{v})$ represents the probability of the occurrence of $\mathbf{v}$. Therefore, the statistical design of an aperture filter $\mathbf{\Psi}$ reduces to find its optimal characteristic function ψ_{opt} (classifier), which minimizes equation 5 through the following *decision rule*:

$$\psi_{opt}(\mathbf{v}) = \begin{cases} 1 & \text{if} \quad \Pr(I(t) = 1 | \mathbf{v}) \geq \tau \\ 0 & \text{otherwise} \end{cases}. \tag{6}$$

In the equation above, if $\tau = 0.5$, then ψ_{opt} is the Bayes classifier. However, for practical applications, we select the value for the threshold τ based on a tradeoff between the false positive and false negative rates of ψ_{opt}.

Finally, the size $|\mathrm{H}|$ of the space of hypothesis, H, for the statistical design of aperture filters grows exponentially as a function of both the size $|W|$ of the spatial window W and the size $|K|$ of the range window K, according to $|\mathrm{H}| = 2^{|K|^{|W|}}$.

2.2.4. Automatic Design of Aperture Filters using Logistic Regression

From equation 6, we can conclude that the statistical design of an aperture filter $\boldsymbol{\Psi}$ consists of estimating the value of $\Pr(I(t)=1|\mathbf{v})$ for each observation $\mathbf{v}$ that can be seen by $W \times K$. In this work we model the probability $\Pr(I(t)=1|\mathbf{v})$, which we will refer to simply as $\Pr(I=1|\mathbf{v})$ from now on, with the following *logistic function*:

$$\Pr(I=1|\mathbf{v},\boldsymbol{\beta}) = 1/(1+\exp(-f(\mathbf{v},\boldsymbol{\beta}))), \tag{7}$$

where $f: K^W \rightarrow \Re$ is definded by $f(\mathbf{v},\boldsymbol{\beta}) = \beta_0 + \beta_1 v_1 + ... + \beta_n v_n$. Furthermore, we used $\Pr(I=1|\mathbf{v},\boldsymbol{\beta})$ to denote the dependence of $\Pr(I=1|\mathbf{v})$ on $\boldsymbol{\beta}$. In the language of machine learning, equation 7 is called *logistic regression* [9,10]. We chose logistic regression because it is a low complex model that can be easily trained, avoids *overfitting*, and has good capabilities of prediction [10].

Using logistic regression to estimate the value of $\Pr(I=1|\mathbf{v})$ for each observation $\mathbf{v}$ that can be seen by $W \times K$ poses a constraint in the space of search H of the optimal aperture filter $\boldsymbol{\Psi}_{opt}$. In fact, the new space of search H', being $\mathrm{H}' \subset \mathrm{H}$, is composed by only linear classifiers. Although the use of constraints yields a suboptimal filter $\boldsymbol{\Psi}_{subopt}$ with increased error compared to the unconstrained optimal operator, $\varepsilon(\boldsymbol{\Psi}_{subopt}) > \varepsilon(\boldsymbol{\Psi}_{opt})$, using constraints produces lower estimation error. In general, a constraint is beneficial if the reduction in the error of estimation is larger than the error induced by the constraint [5-7].

Finally, by using logistic regression, the design of aperture filters is reduced to the task of estimating the vector $\boldsymbol{\beta} = (\beta_0,...,\beta_n)$ by maximizing the likelyhood $\Pr(\mathcal{D}|\boldsymbol{\beta})$ given by equation 8. This task is performed using a set of examples $\mathcal{D} = \{(\mathbf{v}_i, freq(I=0,\mathbf{v}_i), freq(I=1,\mathbf{v}_i))\}$, where $i = 1,2,...,m$, and $freq(I=0,\mathbf{v}_i)$ is the frequency with which the label 0 is seen given the observation $\mathbf{v}_i$ (a similar rule applies for $freq(I=1,\mathbf{v}_i)$). Observation $\mathbf{v}_i$ and frequencies $freq(I=0,\mathbf{v}_i)$ and $freq(I=1,\mathbf{v}_i)$ are obtained by translating $W \times K$ through the images in the training set $\mathbf{D} = \{(O_i, I_i)\}$, being $i = 1,2,...,p$.

$$\Pr(\mathcal{D}|\boldsymbol{\beta}) = \prod_{i=1}^{m} \Pr(I=0|\mathbf{v}_i,\boldsymbol{\beta})^{freq(I=0,\mathbf{v}_i)} \Pr(I=1|\mathbf{v}_i,\boldsymbol{\beta})^{freq(I=1,\mathbf{v}_i)}. \tag{8}$$

The maximization of the equation above is equivalent to minimizing $-ln(\Pr(\mathcal{D}|\boldsymbol{\beta}))$ [10].

2.2.5. Aperture Filter Design with Imbalanced Datasets

Imbalanced datasets are characterized by a high number of examples from one class (majority class), and only few, or even no examples, from the other class (rare class). Assuming that the majority class is labeled with 1, an imbalanced dataset is characterized by $\left(\sum_{i=1}^{m} freq(I=1,\mathbf{v}_i)\right) >> \left(\sum_{i=1}^{m} freq(I=0,\mathbf{v}_i)\right)$. If we train a classifier using imbalanced datasets, then the classifier will usually underestimate the

conditional probabilities of the rare class, because they tend to be biased towards the majority class. Several approaches have been proposed to address this issue, including undersampling examples from the majority class, or oversampling the examples from the rare class [11]. One important drawback of the former approach is the elimination of some examples from the majority class. In this work, we increased artificially (oversample) the frequency $freq(I=1,\mathbf{v}_i)$ of each example $\mathbf{v}_i \in \mathcal{D}$ from the minority class using the following expression:

$$freq_{balanced}(I=1,\mathbf{v}_i) = round(\, freq(I=1,\mathbf{v}_i)(\textstyle\sum_{i=1}^{m} freq(I=0,\mathbf{v}_i) / \sum_{i=1}^{m} freq(I=1,\mathbf{v}_i))), \quad (9)$$

where $freq_{balanced}(I=1,\mathbf{v}_i)$ is the new frequency with which the example $\mathbf{v}_i$ influences the maximization of equation 8.

2.2.6. Ensemble of Aperture Filters

Even though we are using a low complex model for classification (i.e., logistic regression), the estimation of $\boldsymbol{\beta}$ using all the triplets $(\mathbf{v}, freq(I=0,\mathbf{v}), freq(I=1,\mathbf{v})) \in \mathcal{D}$ is highly time and memory consuming. To circumvent this issue, we used an ensemble of aperture filters [8]. Under the ensemble approach, instead of designing a single aperture filter using all the m triplets from the set $\mathcal{D}$, we designed as many aperture filters $\boldsymbol{\Psi}_1,...,\boldsymbol{\Psi}_p$ (called base aperture filters) as number of training pairs of images we had in set $\mathbf{D}$, where $\boldsymbol{\Psi}_i$ was designed using the pair $(O_i, I_i) \in \mathbf{D}$, being $i \in \{1,2,...,p\}$. Then, to process a new image O', we weighted and then combined the predictions of all the p base aperture filters using the Shannon entropy. Formally, the prediction of an ensemble of aperture filters, $\boldsymbol{\Psi}_E$, is given by a weighted sum of the predictions of the base aperture filters $\boldsymbol{\Psi}_1,...,\boldsymbol{\Psi}_p$:

$$\Pr(I=1|\mathbf{v})_E = \textstyle\sum_{i=1}^{p} Pe_i \Pr(I=1|\mathbf{v})_i \,. \quad (10)$$

In the equation above, $\Pr(I=1|\mathbf{v})_E$ is the probability predicted by the ensemble $\boldsymbol{\Psi}_E$ that the observation $\mathbf{v}$ will be assigned label 1. $\Pr(I=1|\mathbf{v})_i$ is the probability predicted by the base aperture filter $\boldsymbol{\Psi}_i$ that the observation $\mathbf{v}$ will be assigned label 1. The weight Pe_i controls the influence of the prediction of $\boldsymbol{\Psi}_i$ in the final prediction of the ensemble and is computed by

$$Pe_i(\mathbf{v}) = (1 - H(\psi(\mathbf{v})_i)) / (p - \textstyle\sum_{j=1}^{p} H(\psi(\mathbf{v})_j)) \,. \quad (11)$$

In the equation above, $H(\psi(\mathbf{v})_i) = -\sum_{j=\{0,1\}} \Pr(I=j|\mathbf{v})_i \, Log_2(\Pr(I=j|\mathbf{v})_i)$ is the Shannon entropy of the prediction of $\boldsymbol{\Psi}_i$ regarding the observation $\mathbf{v}$.

3. Proposed Algorithm

In this section, we describe the algorithm we propose to segment automatically exudates in fundus ocular images.

a) Color to Grayscale Transformation: In this step, we transformed the color ocular image $\overline{H} \in [L_R \times L_G \times L_B]^{\mathbf{E}}$ (Figure 1a) to a grayscale image $H_1 \in L^{\mathbf{E}}$ (Figure 1b), with 256 gray-levels, using equation 1 with the weights $\alpha_1 = 0.07$, $\alpha_2 = 0.9$ and $\alpha_3 = 0.03$. We chose these weights because it was observed that the green channel H_G of the image $\overline{H}$ has the highest contrast between exudates

and the other components of the eye fundus. The red, H_R, and blue, H_B, channels have low contrast and some level of noise [1,8].

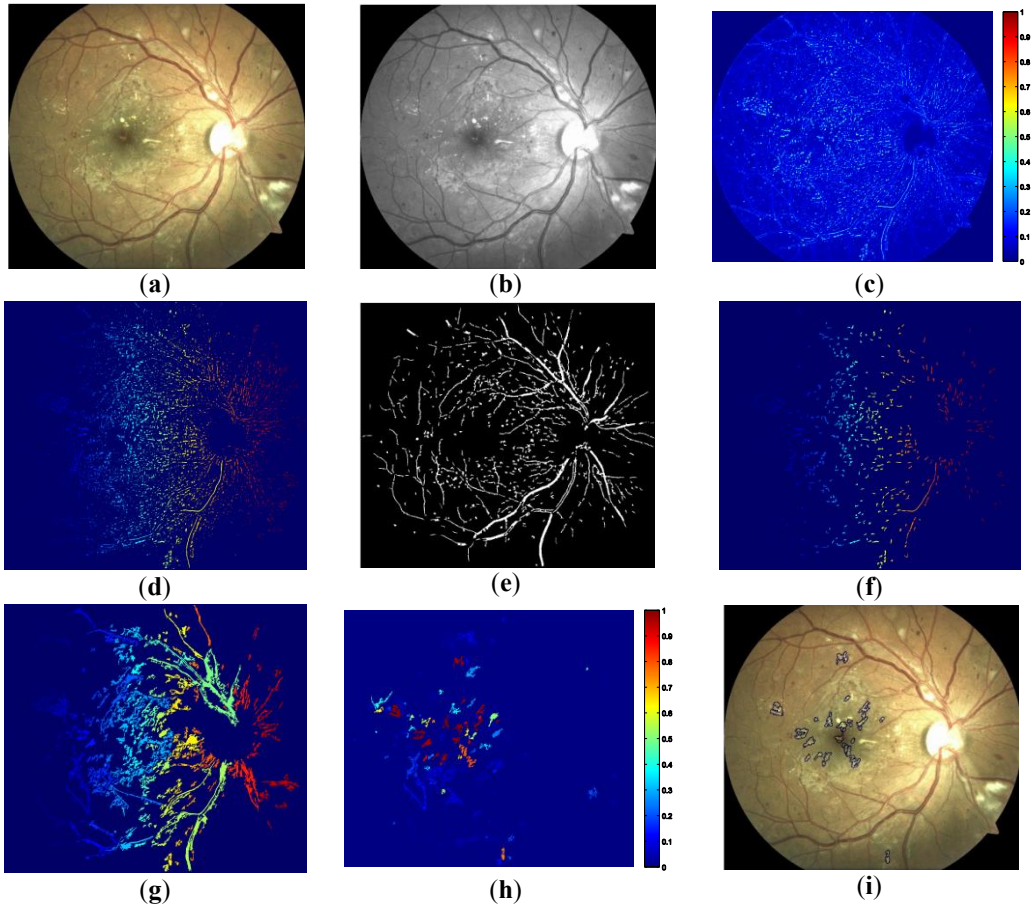

Figure 1: Example of the automatic segmentation of exudates: (**a**) Original image (**b**) Grayscale image (**c**) Lesion probability map (**d**) Exudate candidate markers (Each marker is shown by a different color) (**e**) Image with blood vessels automatically segmented (**f**) Image without blood vessels and noise (**g**) Exudate candidates (Each candidate is shown by a different color) (**h**) Exudate candidate probability map (**i**) Edges of the segmented exudates superimposed on the original image

b) Exudate Candidate Marker Detection: This step starts with the generation of a lesion probability map, in which we assigned a probaility of being part of an exudate region to each pixel $t \in H_1$ (Figure 1c). Each pixel from the image H_1 was processed using an ensemble of aperture filters (Section 2.2.6). In this case, we used the aperture $W \times K$, where W is a 15x15-squared spatial window and $K = [10,10]$ is a range window. The base aperture filters were trained using the training set TrS_1 (Section 2.1) and the procedure described in Section 2.2.4. In the ocular images the number of pixels that belong to exudates is, in general, much less than the number of background pixels. For this reason, before designing the base aperture filters, we balanced the training triplets using the procedure described in Section 2.2.5. Then, we applied a threshold at $\tau = 0.2$ and grouped pixels from the lesion probability map using a connectivity of 8, obtaining a binary image $H_2 \in \{0,1\}^{\mathbf{E}}$ composed of clusters of pixels that belong to exudate candidate markers (Figure 1d).

c) Blood Vessels and Noise Removal: In this step, we started by removing the clusters from the binary image H_2 that are composed of pixels that belong to major blood vessels using an algorithm developed in our previous work [8]. The threshold we used to classify each pixel from image H_1 as either blood vessel or background was $\tau = 0.5$ (Figure 1e). Then, we defined small clusters, those with fewer than 100 pixels, as noise and removed them to obtain the binary image H_3 (Figure 1f).

d) Exudate Candidate Detection: This step consists of obtaining a binary image H_4 composed by clusters of pixels that are potential exudates (Figure 1g). Image H_4 is obtained by morphological reconstruction [12] of the binary image H_3 (marker) from the binary image obtained by thresholding the lesion probability map from step b at $\tau = 0.1$ (mask).

e) Feature Extraction and Exudate Candidate Classification: Image H_4 contains spurious clusters of pixels as consequence of false markers detected in step b. In this step, we classified each cluster from image H_4 as either exudate or non-exudate by using logistic regression. The parameters of the logistic regression were estimated maximizing equation 8 (Section 2.2.4) using the training pairs $(\mathbf{x}_i, y_i)$, being that $i = 1,2,...,q$. q denotes the total number of potential exudates detected in images from the training set TrS_2. Vector $\mathbf{x}_i$ is composed by 31 features, and y_i takes label 1 if the region covered by the exudate candidate in the ideal image belongs to an exudate or otherwise label 0.

The 31 features that compose the vector $\mathbf{x}_i$ represent the main characteristics of each exudate candidate as follows:

- *shape* (extracted from image H_4): area, perimeter, eccentricity, major and minor axis length;
- *texture* (extracted from image H_1): edge strength (mean pixel value within the candidate border detected by a morphological gradient using a disk-shaped structuring element of radius 1), mean and standard deviation of the pixels within the candidate, and local pixel contrast (mean/standard deviation pixel value within the candidate subtracted from the mean/standard deviation pixel value from a 6-pixels-border-wide around the potential lesion); and
- *color* (extracted from image $\overline{H}$): median, mean and standard deviation from the LUV (U and V channels), RGB (all the three channels), and HSV (H and S channels) color spaces.

Finally, we applied the trained logistic regression classifier to each lesion of the image H_4, assigning to each a probability of being an exudate (Figure 1h). Candidates having a probability equal to or greater than the threshold $\tau = 0.7$ were classified as exudates (Figure 1i).

4. Results and Discussion

In this section, we present the results obtained by applying the proposed algorithm to each of the 109 images that comprise the testing set TtS (Section 2.1). The ultimate goal of the automatic segmentation of exudates is the diagnosis of the DME condition. Therefore, we evaluated the performance of our algorithm based on its ability to classify patients as either having or not DME. We defined a patient with DME as one in which at least one exudate was detected. In Figure 2, we show the receiver operating characteristic (ROC) curve [10] obtained, where $Sensitivity = TP/(TP + FN)$ and $Specificity = TN/(TN + FP)$. TP is the number of patients correctly classified as having DME. FN is the number of patients that our algorithm incorrectly classified as being healthy. TN is the number of patients correctly classified as healthy. FP is the number of patients wrongly classified as having DME. The area under the ROC curve (AUC) is 0.77. This value and the shape of the ROC curve indicate that our algorithm tends to detect exudates in some images that belong to healthy patients.

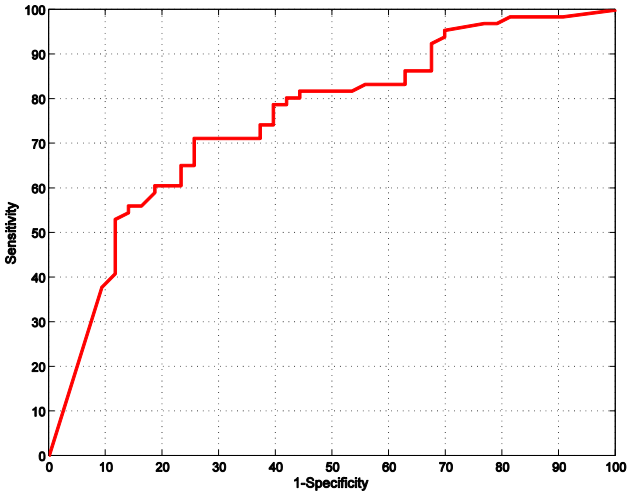

Figure 2: ROC curve of the proposed algorithm for the diagnosis of diabetic macular edema with AUC = 0.77.

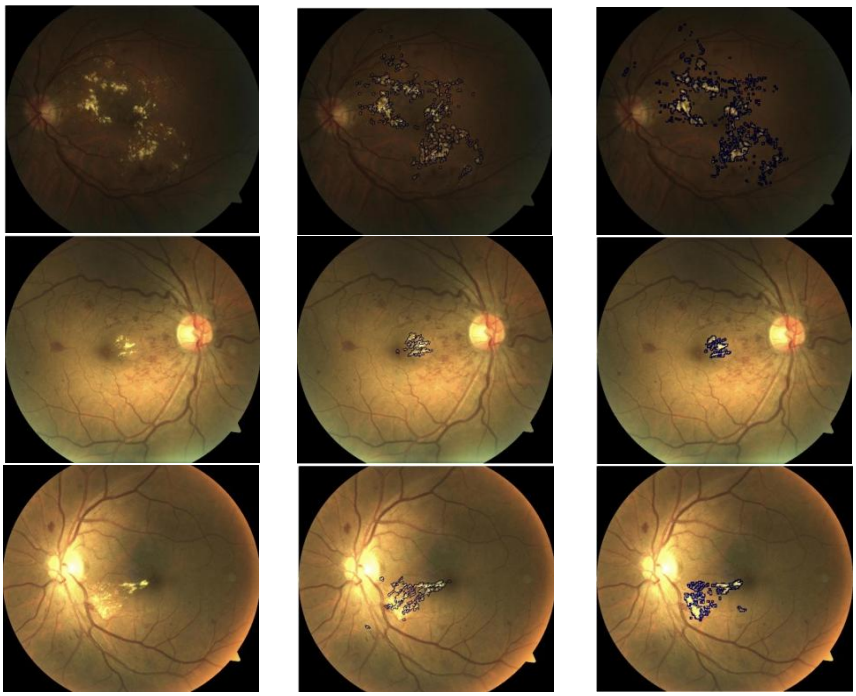

Figure 3: Examples of the segmentation of exudates applying the proposed approach to images from the HEI-MED public database: Each row contains a different test image. In the first column are the original images. The second column shows the results obtained by applying the proposed algorithm. The third column contains the ground truth images of the exudates.

In the scientific literature several algorithms have been proposed to address automatic segmentation of exudates. However, in most of the cases the performance evaluation is done using either private databases with only a few images or public databases that have no ground truth images showing the borders of the exudates. Therefore, to perform a benchmarking between the different proposed algorithms is very difficult. To our knowledge, we only found that [2] and [3] evaluated their algorithms for the diagnosis of DME using the HEI-MED database. They obtained values for the AUC of 0.81 and 0.82, respectively. Although these algorithms perform better than the proposed algorithm in diagnosing DME, they do not perform exudate segmentation. Therefore, one potential issue

regarding these works can be the diagnosis of DME based on false positive exudates. Conversely, our algorithm provides information on the borders and positions of the exudates (Figure 3).

5. Concluding Remarks

In this work, we proposed a new approach for the automatic segmentation of exudates. The information generated by our approach is useful not only for the diagnosis of DME, but also for treatment and blindness prevention. The proposed method is based on ensembles of aperture filters that were used for both detecting exudate candidates and removing parts of the major blood vessels from the processed images. Finally, a logistic regression classifier was used to assign a probability of being an exudate to each candidate detected.

We also treated the aperture filter design as a pattern recognition problem. Using this approach, we reduced the aperture filter design to a classifier design by basing it on logistic regression. Additionally, the practical complexity issue that occurs when designing aperture filters with large training sets of images was circumvented by using an ensemble method. The ensemble method allowed us to combine the predictions of several base aperture filters. Each base aperture filter was designed using only one training pair of images.

Lastly, the results obtained in this paper and other publications demonstrate that automatic methods can be used to assist in the diagnosis of DME. However, we still need to improve the performance of such methods. In the case of our particular approach, future work includes the development of new methods for designing aperture filters to process color images directly without a color to grayscale transformation.

6. References

[1] Winder R, Morrow P, McRitchie I, Bailie J, Hart P, 2009, *Algorithms for digital image processing in diabetic retinopathy*, Computerized medical imaging **33** 608–622.
[2] Giancardo L, Meriaudeau F, Karnowski T, Li Y, Tobin Jr, Chaum E, 2011, *Automatic retina exudates segmentation without a manually labelled training set*, International symposium on biomedical imaging, 1396–1400.
[3] Giancardo L, Meriaudeau F, Karnowski T, Li Y, Garg S, Tobin Jr K, Chaum E, 2012, *Exudate-based diabetic macular edema detection in fundus images using publicly available datasets*, Medical image analysis **16** 216–226.
[4] Youssef D and Solouma H, 2012, *Accurate detection of blood vessels improves the detection of exudates in color fundus images*, Computer methods and programs in biomedicine **108** 1052–1061.
[5] Hirata R, Brun M, Barrera J, Dougherty E, 2006, *Advances in nonlinear signal and image processing*, Chapter 2, Hindawi.
[6] Hirata R, Dougherty E, Barrera J, 2000, *Aperture filters*, Signal processing **80** 697–721.
[7] Benalcázar M, Brun M, Ballarin V, 2013, *Artificial neural networks applied to statistical design of window operators*, Pattern recognition letters **34** 970–979.
[8] Benalcázar M, Brun M, Ballarin V, 2012, *Segmentación de vasos sanguíneos en angiografías retinales usando ensambles de filtros aperture*, III Simposio internacional de innovación y tecnología (SIIT), 125–129.
[9] Duda R, Hart P, Stork D, 2001, *Pattern classification*, Wiley.
[10] Murphy k, 2012, *Machine learning: a probabilistic perspective (adaptive computation and machine learning series*, Masachusets Institute of Technology.
[11] Weiss G, 2004, *Mining with rarity: a unifying framework*, ACM SIGKDD Explorations Newsletter **6** 7–19.
[12] Vincent L, 1993, *Morphological grayscale reconstruction in image analysis: applications and efficient algorithms*, IEEE Transactions on image processing **2** 176–201.