Pré-Publicações do Departamento de Matemática
Universidade de Coimbra
Preprint Number 14–40

AUTOMATED LESION DETECTORS IN RETINAL FUNDUS IMAGES

ISABEL NARRA FIGUEIREDO, SUNIL KUMAR, CARLOS MANTA OLIVEIRA, JOÃO DIOGO RAMOS AND BJORN ENGQUIST

Abstract: Diabetic retinopathy (DR) is a sight-threatening condition occurring in persons with diabetes, which causes progressive damage to the retina. The early detection and diagnosis of DR is vital for saving the vision of diabetic persons. The early signs of DR which appear on the surface of the retina are the dark lesions such as microaneurysms (MAs) and hemorrhages (HEMs), and bright lesions (BLs) such as exudates. In this paper, we propose a novel automated system for the detection and diagnosis of these retinal lesions through the processing of retinal fundus images. We devise appropriate binary classifier detectors for these three different types of lesions. Some novel contextual/numerical features are derived, for each lesion type, depending on its inherent properties. This is performed by analysing several wavelet bands (resulting from the isotropic undecimated wavelet transform decomposition of the retinal image green channel) and by using an appropriate combination of Hessian multiscale analysis, variational segmentation and cartoon+texture decomposition. The proposed methodology has been validated on several medical data sets using standard performance measures such as sensitivity and specificity. The individually performance, per frame, of the MAs detector is 93% sensitivity and 89% specificity, of the HEMs detector is 86% sensitivity and 90% specificity, and of the BLs detector is 90% sensitivity and 97% specificity. Regarding the collective performance of these binary detectors, as an automated screening system for DR (meaning that a patient is considered to have DR if it is a positive patient for at least one of the detectors) it achieves in average 95%-100% of sensitivity and 70% of specificity at a per patient basis.

Keywords: Microaneurysm Detector, Hemorrhage Detector, Bright Lesion Detector, Retinal Fundus Image, Computer-aided Diagnosis, Wavelets, Multiscale Analysis, Variational Segmentation, Cartoon+Texture Decomposition.

1. Introduction

Diabetic retinopathy (DR) is a sight-threatening consequence of Diabetes Mellitus which affects the small blood vessels in the retina, the nerve layer that lines the back of your eye. Often there are no visual symptoms in the early stages of diabetic retinopathy. Early diagnosis and treatment are essential to prevent significant vision loss from DR. So, it is vital, for a diabetic

Received November 24, 2014.

patient to have regular eye examinations. One of the easiest way to diagnose DR is by analyzing fundus images. They are the visual records which document the current ophthalmoscopic appearance of a patient's retina, the retinal vasculature, and the optic nerve head (optic disc) from which the retinal vessels enter the eye. The acquisition of fundus images is inexpensive, non-invasive and easy to perform. Therefore they are adapted for the large scale screening purposes.

DR can be mainly classified into two classes: Non-proliferative diabetic retinopathy (NPDR) and Proliferative diabetic retinopathy (PDR). NPDR is the earliest stage of DR and can be described as mild, moderate or severe. With this condition, the walls of the blood vessels in the retina become weak. Tiny bulges protrude from the vessel walls, sometimes leaking or oozing fluid and blood into the retina. Also, sometimes, deposits of cholesterol or other fats from the blood may leak into the retina. NPDR can cause several changes in the eye, including microaneurysms (MAs), hemorrhages (HEMs) and bright lesions (BLs) (as for instance hard and soft exudates). PDR is the more advanced form of the disease, where fragile new blood vessels form on the surface of the retina over time. These abnormal vessels can bleed or develop scar tissue causing severe loss of vision. PDR may cause more severe vision loss than NPDR because it can affect both central and peripheral vision. Due to this reason it is very significant to diagnose and treat DR in the non-proliferative stage.

Computer-aided detection and diagnosis of DR with retinal fundus images significantly lessens the burden of the implementation of a large scale screening of the diabetic patients. Recent years have seen the development of the methods for the accurate detection of MAs, HEMs and BLs by considering them individually and in a collective way. Without being exhaustive we mention here some of these related works. A survey of various methods related to DR can be found in [28]. The segmentation of exudates using fuzzy c-means clustering algorithm is done in [22]. After extracting color, size, edge strength, and texture features from the regions and a multilayer neural network classifier is used to classify the images. In [3] a new hybrid classifier as an ensemble of Gaussian mixture model and support vector machine is proposed for exudate detection. In [29] a new candidate segmentation method, based on mathematical morphology, is proposed. Extracting some features, a random forest algorithm is used to detect the exudates among the candidates. An automatic MAs detector based on template matching in wavelet

space using a Gaussian template is proposed in [24]. The detection of MAs and HEMs based on recursive region growing segmentation algorithm and a moat operator is done in [25]. After extracting the potential candidates and formulating a set of features of MAs, HEMs and exudates, [2] used a hybrid classifier which is a weighted combination of multivariate m-Mediods and a Gaussian mixture model. An ensemble-based framework for the detection of MAs is proposed in [4]. The algorithm in [23] assessed the need for referral in DR detection with a decision based on the fusion of results by meta-classification. In [12] the detection of HEMs is done based on the hue, saturation and value (HSV) model and the Mahalanobis distance. A novel splat feature classification method is proposed in [27] to detect retinal hemorrhages.

The present paper describes an intelligent automated system for the binary classification of different types of DR lesions through the processing of eye fundus images.

We derive some novel contextual detectors for the lesions (MAs, HEMs and BLs) depending on their intrinsic properties. The numerical features, which define each detector, rely on the analysis of several wavelet bands (resulting from the isotropic undecimated wavelet transform [26] decomposition of the retinal image), and includes an appropriate combination of Hessian multiscale analysis (more precisely, using the Hessian eigenvalues), variational segmentation [6, 9] and cartoon+texture decomposition [7]. These features are then used to classify the retinal image into separate categories of DR. We perform a thorough testing of the proposed method on several rich data sets to ensure its good performance in realistic environment.

The rest of the paper is structured as follows. Section 2 presents in detail the methodology, by explaining the definition of the MAs, HEMs and BLs detectors. This includes the extraction of the candidates for the different lesions, description of all the features and the binary classifiers for the lesions. Section 3 presents the experimental results on several medical data sets prepared by the experts. Then, the paper ends with a discussion of the results and some conclusions in Section 4.

2. MAs, HEMs and BLs Detectors

In this section we present the detailed description of the method for the detection of the different lesions. The procedure relies on four major mathematical techniques (i- wavelets, ii- multiscale analysis of the image Hessian

matrix, iii- variational image segmentation, iv- cartoon and texture image decomposition) which are appropriately and sequentially applied and combined. Firstly we outline these 4 techniques in an abstract framework for a scalar image. Then we proceed with the description of the first step of the method: the pre-processing step and the optic disk detection in retinal fundus images. Finally we give a detailed explanation of the three lesion detectors: microaneurysms (MAs), hemorrhages (HEMs) and bright lesions (BLs). Once these feature detectors are extracted, the decision criteria, on whether a frame has or does not have one of these type of lesions, are based on simple thresholding approaches.

2.1. The Isotropic Undecimated Wavelet Transform. The Isotropic Undecimated Wavelet Transform (IUWT) [26] is characterized by a simple implementation of its direct and inverse transform. Denoting by $C_0 = I$ a scalar input image, at each subsequent iteration j, the scaling coefficients C_j are computed by lowpass filtering (where filtering is applied by convolution and separably along each dimension), and the wavelet coefficients W_j (hereafter called wavelet levels) are defined just by scaling coefficient subtraction, i.e, $W_{j+1} := C_{j+1} - C_j$. Starting with the filter $h_0 = [1/16, 1/4, 3/8, 1/4, 1/16]$, which is derived from the cubic B-spline, at each scale $j \geq 2$ the filter h_j is augmented by inserting $2^j - 1$ zeros between each pair of adjacent coefficients of h_0. The scaling coefficient C_{j+1} is defined from the previous coefficient C_j using separable convolution with h_j in each dimension. The reproduction of the original image I is achieved just by adding all the wavelet coefficients and the final computed scaling coefficient in the following way: $I := C_m + \sum_{j=1}^{m} W_j$. In addition we denote by I_n a partial reconstruction of I defined by

$$I_n = \sum_{j=2}^{n} W_j, \quad n = 2, 3, 4, 5, \tag{1}$$

with W_j the wavelet level at iteration j.

2.2. Multiscale Analysis of the Image Hessian Matrix. For looking for blob structures in a scalar image, we use herein a strategy based on the eigenvalues of the image Hessian matrix and multiscale image analysis. For a scalar image $I : \Omega \subseteq \mathbb{R}^2 \to \mathbb{R}$, where Ω is the pixel domain, we first define

E-mail address: jramos@critical-health.com
URL: http://www.retmarker.com/

Bjorn Engquist
Department of Mathematics and the Institute for Computational Engineering and Sciences, The University of Texas at Austin, Austin, USA.
E-mail address: engquist@ices.utexas.edu

[19] P. V. C. Hough. Methods and means for recognizing complex patterns. U.S. Patent 3 069 654, Dec. 1962.
[20] C. Kimme, D. Ballard, and J. Sklansky. Finding circles by an array of accumulators. *Commun. ACM*, 18:120–122, 1975.
[21] A. Mamonov, I. N. Figueiredo, P. N. Figueiredo, and R. Tsai. Automated polyp detection in colon capsule endoscopy. *IEEE Transactions on Medical Imaging*, 33(7):1488–1502, 2014.
[22] A Osareh, B. Shadgar, and R. Markham. A computational-intelligence-based approach for detection of exudates in diabetic retinopathy images. *IEEE Transactions on Information Technology in Biomedicine*, 13:535–545, 2009.
[23] R. Pires, H. F. Jelinek, J. Wainer, S. Goldenstein, E. Valle, and A. Rocha. Assessing the need for referral in automatic diabetic retinopathy detection. *IEEE Transactions on Biomedical Engineering*, 60:3391–3398, 2013.
[24] G. Quellec, M. Lamard, P.M. Josselin, G. Cazuguel, B. Cochener, and C. Roux. Optimal wavelet transform for the detection of microaneurysms in retina photographs. *IEEE Transactions on Medical Imaging*, 27:1230–1241, 2008.
[25] C. Sinthanayothin, J. F. Boyce, T. H. Williamson, H. L. Cook, E. Mensah, S. Lal, and D. Usher. Automated detection of diabetic retinopathy on digital fundus images. *Diabetic Medicine*, 19:105–112, 2002.
[26] J.-L. Starck, J. Fadili, and F. Murtagh. The undecimated wavelet decomposition and its reconstruction. *IEEE Transactions on Image Processing*, 16:297–309, 2007.
[27] L. Tang, M. Niemeijer, J. M. Reinhardt, M. K. Garvin, and M.D. Abràmoff. Splat feature classification with application to retinal hemorrhage detection in fundus images. *IEEE Transactions on Medical Imaging*, 32:364–375, 2013.
[28] R. J. Winder, P. J. Morrow, I. N. McRitchie, J. R. Bailie, and P.M. Hart. Algorithms for digital image processing in diabetic retinopathy. *Computerized Medical Imaging and Graphics*, 33:608 – 622, 2009.
[29] X. Zhang, G. Thibault, E. Decencière, B. Marcotegui, B. Laÿ, R. Danno, G. Cazuguel, G. Quellec, M. Lamard, Pascale M., A. Chabouis, Z. Victor, and A. Erginay. Exudate detection in color retinal images for mass screening of diabetic retinopathy. *Medical Image Analysis*, 18:1026 – 1043, 2014.

Isabel Narra Figueiredo
CMUC, Department of Mathematics, University of Coimbra, Portugal.
E-mail address: isabelf@matuc.pt
URL: http://www.mat.uc.pt/∼isabelf

Sunil Kumar
CMUC, Department of Mathematics, University of Coimbra, Portugal.
E-mail address: skumar@mat.uc.pt

Carlos Manta Oliveira
Universidade Nova de Lisboa, Portugal and Retmarker
E-mail address: carlos.a.oliveira@critical-health.com
URL: http://www.retmarker.com/

João Diogo Ramos
Retmarker

the Hessian matrix of one point $(x, y) \in \Omega$, and at a scale s, by

$$H_s(x,y) := \begin{pmatrix} I^s_{xx}(x,y) & I^s_{xy}(x,y) \\ I^s_{xy}(x,y) & I^s_{yy}(x,y) \end{pmatrix}.$$

Here I^s_{xx}, I^s_{xy} and I^s_{yy} are the second-order partial derivatives of I and the scale s is involved in the calculation of these derivatives (for instance using Gaussian convolution). Suppose $\lambda_{s,1}$ and $\lambda_{s,2}$ are the two eigenvalues of the Hessian matrix H_s. Note that in a blob region, these two eigenvalues have the same sign and similar magnitudes (see [17]). Without loss of generality we assume that $|\lambda_{s,1}| \leq |\lambda_{s,2}|$. Then, setting $F_s := \lambda^2_{s,1} + \lambda^2_{s,2}$

$$f_1 := \exp\left(-\beta F_s^2\right) \text{ and } f_2 := 1 - \exp\left(-\alpha \left(\frac{\lambda_{s,1}}{\lambda_{s,2}}\right)^2\right),$$

and motivated by [17, 14], we define the blob (B_s) detector (at each point of the domain) by

$$B_s := \begin{cases} 0, & \text{if } \lambda_{s,1}\lambda_{s,2} < 0 \text{ or } |\lambda_{s,2} - \lambda_{s,1}| > \delta \\ (1 - f_1) f_2, & \text{otherwise.} \end{cases}$$

Here α and β are parameters which control the sensitivity of the function B_s and δ a user chosen threshold. The values of α, β and δ are considered to be 1, 1 and 10^{-4}, respectively. In order to automatically detect blobs of different sizes, a multiscale approach is necessary. The response of the detector function B_s will be maximum at a scale that approximately matches the size of the blob to be detected. Hence, at each point of the domain, we define the function

$$B := \max_{s_{min} \leq s \leq s_{max}} (B_s.I), \tag{2}$$

where s_{min} and s_{max} are the minimum and maximum scales at which the blobs are expected to be found. Finally, in the subsequent Sections we consider the maximum of the function B as an input to our binary (blob) classifier

$$B^{max} := \max_{i,j} B_{i,j},\ 1 \leq i \leq N_x,\ 1 \leq j \leq N_y, \tag{3}$$

where N_x and N_y are the width and height of the image in pixels.

2.3. Variational Image Segmentation. The segmentation scheme, we have applied, relies on a reformulation of the Chan and Vese variational model [6, 9], and is described as follows: We first compute the constants c_1 and c_2 (representing the averages of a scalar image $I : \Omega \subseteq \mathbb{R}^2 \to \mathbb{R}$,

in a two-region image partition). We then solve the following minimization problem

$$\min_{u,v}\left\{TV_g(u)+\frac{1}{2\theta}\|u-v\|^2_{L^2(\Omega)}+\int_\Omega\Big(\lambda\, r(I,c_1,c_2)\, v+\alpha\,\nu(v)\Big)\,dx\,dy\right\} \quad (4)$$

where $TV_g(u) := \int_\Omega g(x,y)|\nabla u|\,dx\,dy$ is the total variation norm of the function u, weighted by a positive function g; $r(I,c_1,c_2)(x,y) := \big(c_1 - I(x,y)\big)^2 - \big(c_2 - I(x,y)\big)^2$ is the fitting term, $\theta > 0$ is a fixed small parameter, $\lambda > 0$ is a constant parameter weighting the fitting term, and $\alpha\,\nu(v)$ is a term resulting from a reformulation of the model as a convex unconstrained minimization problem (see [6, Theorem 3]). Here, u represents the two-phase segmentation and v is an auxiliary unknown. The segmentation curve, which divides the image into two disjoint parts, is a level set of u, $\{(x,y)\in\Omega :\ u(x,y)=\mu\}$, where in general $\mu = 0.5$ (but μ can be any number between 0 and 1, without changing the segmentation result, because u is very close to a binary function).

The above minimization problem is solved by minimizing u and v separately, and iterated until convergence. In short we consider the following two steps:

1. v being fixed, we look for u that solves

$$\min_u\left\{TV_g(u)+\frac{1}{2\theta}\|u-v\|^2_{L^2(\Omega)}\right\}. \quad (5)$$

2. u being fixed, we look for v that solves

$$\min_v\left\{\frac{1}{2\theta}\|u-v\|^2_{L^2(\Omega)}+\int_\Omega\Big(\lambda\, r(I,c_1,c_2)\, v+\alpha\,\nu(v)\Big)\,dx\,dy\right\}. \quad (6)$$

We refer to [6, Propositions 3, 4] for the solution of (5)-(6). For our experiments g is considered to be the edge indicator function $g(\nabla u) := \frac{1}{1+\beta\|\nabla u\|^2}$ with $\beta = 10^{-3}$ ($\|.\|$ denotes the Euclidean norm). The values of θ and λ are considered to be 1 and 0.01, respectively.

2.4. Cartoon and Texture Image Decomposition. Given a scalar image $I : \Omega \subseteq \mathbb{R}^2 \to \mathbb{R}$, we decompose it into the cartoon (geometric) part C and the texture (oscillatory) part T using an algorithm of Buades et al. [7]. This algorithm is briefly described as follows. We define the local total variation

References

[1] M. D. Abràmoff, J. M. Reinhardt, S. R. Russell, J. C. Folk, V. B. Mahajan, M. Niemeijer, and G. Quellec. Automated early detection of diabetic retinopathy. *Ophthalmology*, 117:1147 – 1154, 2010.

[2] M. U. Akram, S. Khalid, A. Tariq, S. A. Khan, and F. Azam. Detection and classification of retinal lesions for grading of diabetic retinopathy. *Computers in Biology and Medicine*, 45:161 – 171, 2014.

[3] M. U. Akram, A. Tariq, S. A. Khan, and M. Y. Javed. Automated detection of exudates and macula for grading of diabetic macular edema. *Computer Methods and Programs in Biomedicine*, 114:141 – 152, 2014.

[4] B. Antal and A. Hajdu. An ensemble-based system for microaneurysm detection and diabetic retinopathy grading. *IEEE Transactions on Biomedical Engineering*, 59:1720–1726, 2012.

[5] P. Bankhead, C. N. Scholfield, J.G. McGeown, and T. M. Curtis. Fast retinal vessel detection and measurement using wavelets and edge location refinement. *PloS one*, 7(3):e32435, 2012.

[6] X. Bresson, S. Esedoglu, P. Vandergheynst, J.-P. Thiran, and S. Osher. Fast global minimization of the active contour/snake model. *J Math. Imaging Vis.*, 28:151–167, 2007.

[7] A. Buades, T. M. Le, J. M. Morel, and L. A Vese. Fast cartoon + texture image filters. *IEEE Transactions on Image Processing*, 19:1978–1986, 2010.

[8] J. Canny. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8:679–698, 1986.

[9] T. F. Chan and L. A. Vese. Active contours without edges. *IEEE Transactions on Image Processing*, 10:266–277, 2001.

[10] J. Ferreira, R. Bernardes, P. Baptista, and J. Cunha-Vaz. Earmarking retinal changes in a sequence of digital color fundus photographs. In *IFMBE Proc*, volume 11, pages 1727 – 1983, 2005.

[11] I. N. Figueiredo and S. Kumar. Automatic optic disc detection in retinal fundus images based on geometric features. In *Image Analysis and Recognition, International Conference ICIAR 2014*. Aurelio Campilho and Mohamed Kamel, 2014.

[12] I. N. Figueiredo and S. Kumar. Wavelet-based computer-aided detection of bright lesions in retinal fundus images. In *CompIMAGE 2014, LNCS 8641*, pages 234–240, 2014.

[13] I. N. Figueiredo, S. Kumar, and P. N. Figueiredo. An intelligent system for polyp detection in wireless capsule endoscopy images. In *Computational Vision and Medical Image Processing, VIPIMAGE 2013*, pages 229–235. João Tavares & Natal Jorge (eds), 2014 Taylor & Francis Group, London, ISBN 978-1-138-00081-0 (ECCOMAS Thematic Conference on Computational Vision and Medical Image Processing).

[14] I. N. Figueiredo, S. Kumar, C. Leal, and Pedro N. Figueiredo. Computer-assisted bleeding detection in wireless capsule endoscopy images. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 1:198–210, 2013.

[15] P. N. Figueiredo, I. N. Figueiredo, S. Prasath, and R. Tsai. Automatic polyp detection in pillcam colon 2 capsule images and videos: preliminary feasibility report. *Diagnostic and Therapeutic Endoscopy*, pages 1–7, 2011.

[16] M. Foracchia, E. Grisan, and A. Ruggeri. Luminosity and contrast normalization in retinal images. *Medical Image Analysis*, 9(3):179 –190, 2005.

[17] A. F. Frangi, W. J. Niessen, K. L. Vincken, and M. A. Viergever. Multiscale vessel enhancement filtering. In *Medical Image Computing and Computer-Assisted Intervention*, pages 130–137, Springer, Berlin, 1998.

[18] R. M. Haralock and L. G. Shapiro. *Computer and Robot Vision*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1991.

structures, such as the optic disk and blood vessels, so that they cannot interfere in the detection of the lesions. The extraction of the candidates for different types of lesions (MAs, HEM and BLs) is done by analysing different subband wavelet images (resulting from the isotropic undecimated wavelet transform decomposition of the green channel of the retinal image), and applying Hessian multiscale analysis, variational segmentation, cartoon+texture decomposition. These techniques are combined in a sequential and suitable way and this combination depends on the visual information transmitted by the different lesions (that by nature are very diverse in shape, size, texture and color). After that some novel contextual features are derived and quantified, for each lesion depending on their properties. Then the correspondent numerical feature values are used to obtain a binary classification method. Firstly, the testing of the proposed methods for the MAs, HEMs and BLs detection is done individually. Secondly, we consider the proposed methods as a system for DR detection and performed a thorough testing of the system on rich data sets to ensure its good performance in realistic situations. Standard performance measures such as sensitivity and specificity are used for its evaluation.

For general use of an automated system, high sensitivity is a safety issue and is therefore more important than high specificity, which is an efficacy issue [1]. The main function of any automated system is to decrease the amount of frames that require manual inspection. The proposed system achieves a very high sensitivity (between 90% and 100%), and equally importantly, a very good specificity (approximately 70%, in datasets with an average good quality images).

Acknowledgment

This work was partially supported by the project PTDC/MATNAN/0593/2012, and also by CMUC (Center for Mathematics, University of Coimbra) and FCT (Portugal), through European program COMPETE/ FEDER and project PEst-C/MAT/UI0324/2013. The authors would also like to thank Dr. Gonçalo Quadros for having suggested the study of the interesting and challenging retina topic.

(LTV) for each pixel x in Ω, by the convolution (denoted by '$*$')

$$LTV_\sigma(I)(x) := G_\sigma * |\nabla I|(x)$$

where G_σ is a Gaussian kernel of variance σ. The main characteristic of a textural region is its high total variation (due to its oscillations) and the property of decreasing very fast under low-pass filtering (as opposed to cartoon, or equivalently, non-oscillatory regions). We then compute the relative reduction rate of LTV by

$$\lambda(x) := \frac{LTV_\sigma(I)(x) - LTV_\sigma(\mathcal{R}_\sigma * I)(x)}{LTV_\sigma(I)(x)},$$

where $\mathcal{R}_\sigma$ is a low pass filter. λ is a local indicator of the oscillations in I: if $\lambda(x) \approx 0$ there is little reduction of LTV with the low-pass filter, but if $\lambda(x) \approx 1$, this means the reduction is strong, and consequently the pixel x belongs to a textured region.

Depending on the relative reduction of LTV, taking the weighted average of I and $\mathcal{R}_\sigma * I$, we get the cartoon part C, and hence the texture part T,

$$\begin{aligned} C(x) &:= w\big(\lambda(x)\big)\big(\mathcal{R}_\sigma * I\big)(x) + \Big(1 - w\big(\lambda(x)\big)\Big)I(x), \\ T(x) &:= I(x) - C(x). \end{aligned} \tag{7}$$

The weight $w : [0,1] \to [0,1]$ is a nondecreasing piecewise affine function that is constant and equal to zero, near zero, and constant and equal to 1, near 1. If $\lambda(x)$ is very small, then I is non oscillatory around x and therefore $w \approx 0$, so $C \approx I$, that is the image is cartoon. If instead $\lambda(x)$ is big, then the image I is locally oscillatory in the neighborhood of x.

2.5. Pre-processing and Optic Disk Detection. Retinal fundus images normally contain a main region at the center of the image, which is called the field of view (FOV), surrounded by dark background pixels. For automated detection of lesions we only need to process the main region, so we separate it from the background, in a preprocessing step. The separation is performed by eroding the green channel of the image, using a square structuring element.

It is also needed to extract the optic disc (OD) from each image to eliminate any spurious and false regions caused by their similarities with bright lesions. The segmentation of the OD is done in two steps. We first localize the OD using its geometric features and then use the circular Hough transform (CHT) in the localized region to detect the OD boundary. The herein proposed

approach for the OD detection was presented by the authors in a conference paper in [11], and it is summarized here for convenience of the reader.

It is well known that the OD appears most contrasted in the green channel compared to the red and blue channels of the RGB image. More importantly, it appears as a concave region (protrusion) in the green channel. Motivated from [13, 15], we define

$$\mathcal{OD}_{loc} := -\mathcal{K}\min(\mathcal{H}, 0),$$

where $\mathcal{K}$ and $\mathcal{H}$ are, respectively, the Gaussian and mean curvatures of the green channel image. This geometrical function was introduced in [13, 15] as a protrusion measure for detecting colonic polyps (which are round-shaped protusion objects) in wireless capsule endoscopy images. Mathematically, the value of the $\mathcal{OD}_{loc}$ function is closely related to the size of the protrusions in the images. Therefore, the optic disc location in the frame can be inferred by identifying the locations where $\mathcal{OD}_{loc}$ is higher. Hence, applying a threshold we can separate the candidate region(s) for the location of the optic disc. We extract the highest 15% of pixels of $\mathcal{OD}_{loc}$. The entire thresholded image is scanned to count the number of connected components. The connected components are found using an algorithm by Haralick and Shapiro [18]. For each connected component, in the thresholded image, we compute the area and the $\mathcal{OD}_{loc}$ mean pixel intensity. The component corresponding to the maximum of the product between the area and the $\mathcal{OD}_{loc}$ mean pixel intensity is considered to be the location of the OD. We then compute the centroid of the same component, which is identified with the estimated OD center and consequently the OD location.

Afterwards, the detection of the OD boundary is done based on the CHT [20, 19], which relies on equations for circles. The parametric representation of a circle is $x := a + r * cos(\theta)$, $y := b + r * sin(\theta)$, where a and b represent the coordinates for the center, $\theta \in [0, 2\pi]$ and r is the circle radius. We consider r between r_{min} and r_{max} which are, respectively, one-tenth and one-fifth of the image (as these measurements are estimated to be the average OD diameter) divided by two. The minimum value makes sure that the OD cup (an internal circular object in the OD) is not considered, while the maximum value is such that too wide areas are eliminated. The CHT transforms a set of feature points in the image space into a set of accumulated votes in a parameter space. Then, for each feature point, votes are accumulated in an accumulator array for all parameter combinations. The array elements that

artifacts (in fact these images are nonmydriatic photographs without pupil dilation, which originates lower quality images). This leads to an increase of false positives and some false negatives (some examples are shown in Figure 9).

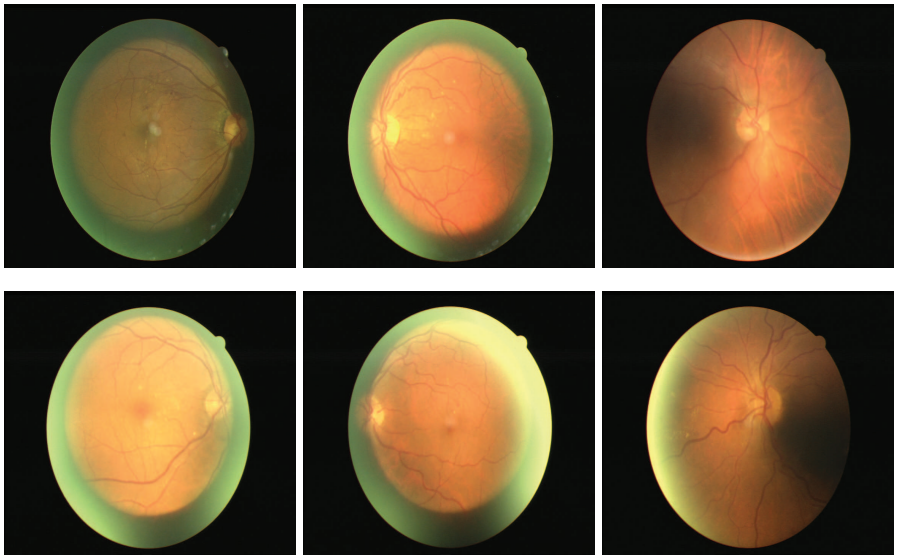


FIGURE 9. Some examples of bad quality DR retinal fundus images, not detected by the proposed system, from the data sets DR3 and DR4.

4. Discussion and Conclusions

We developed a system for automated detection and diagnosis of DR, through the processing of retinal fundus images, by devising appropriate MAs, HEMs and BLs binary classifier detectors. The preprocessing phase involves the extraction of the background pixels that enables further computations on the foreground pixels only, as well as the removal of main image

TABLE 3. Performance evaluation of each detector (per frame)

Detector	Sensitivity	Specificity
MAs	93.45	88.92
HEMs	86.34	90.19
BLs	89.81	97.47

Next we evaluate the performance of the methods in a collective way as a system for the detection and diagnosis of DR. Since here we have the classification that is on a per patient basis, the assessment of the performance of our system was done on a per patient basis. For this purpose we use a modified definition of sensitivity and specificity, in (20) and (21), where there TP, FN, FP and TN are considered on a per patient basis instead of a per frame basis. A patient is labelled to be a positive patient (i.e. belonging to the class of DR) if at least one frame of the patient is classified as a positive frame by our methodology, *i.e.* if it is a positive for at least one of the three detectors (MAs, HEMs or BLs). Accordingly we compute the number of TP, FN, FP and TN, and hence the sensitivity and specificity.

TABLE 4. Performance evaluation of the proposed system for DR detection (per patient)

Data sets	Sensitivity	Specificity
DR1	100	70.7
DR2	100	73
DR3	89.3	58.9
DR4	92	57.6

The evaluation of the proposed system in terms of sensitivity and specificity is given in Table 4. This Table shows that the sensitivity and specificity values in data sets DR3 and DR4 are worse than in DR1 and DR2. This is because the two latter datasets contain images with very bad quality and

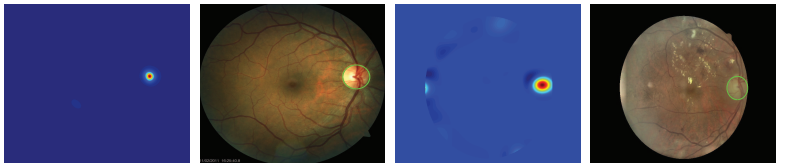


FIGURE 1. Two pairs of $\mathcal{OD}_{loc}$ function in the green channel and the corresponding optic disk boundary determined by CHT.

contains the highest number of votes indicate the presence of the desired shape, i.e. the OD in our case. Based on our previous optic disc localization (the above mentioned computed centroid), we define a region of interest (ROI) of size $m \times n$, where m and n are one-ninth of the respective dimensions of the image multiplied by two. Finally to segment the OD boundary from the retinal image we apply the CHT to the edges of the ROI, which are computed using Canny edge detector [8].

The OD detection for a disease and a non-disease image (exhibiting BLs) is shown in Figure 1.

2.6. Microaneurysm Detection. A microaneurysm is a round dark-red structure (a very small spot whose size is approximately less than 125 microns in its longest dimension) caused by thickening of capillary basement membranes within the retina. We have observed that MAs are visible and enhanced at wavelet level W_2, or at the sum of wavelets levels W_2 and W_3, of the green channel of the input fundus image. These facts have motivated us to devise a MAs detector, based on circular blob detection in these two wavelet images, as described in the sequel of the current Section.

For each image we take $I := I_G$ (in Section 2.1), with I_G the green channel of the input fundus image, and we compute (1), *i.e.* $I_n = \sum_{j=2}^{n} W_j$, $n = 2, 3$. Then, we extract the 20% darkest pixels within the FOV of I_n to identify the MA candidates. We then take the product of this thresholded image with I_n, and remove the pixels corresponding to the vessels and OD. The resulting image is denoted by T_n. The segmentation of the vessels is done by using an approach similar to [5] [briefly it includes the following procedure described in points i) to iii) : i) We compute the sum of the wavelet levels W_2 and W_3 of the green channel of the image and extract the darkest 10% of the pixels within the FOV of the same. ii) The vasculature can be seen

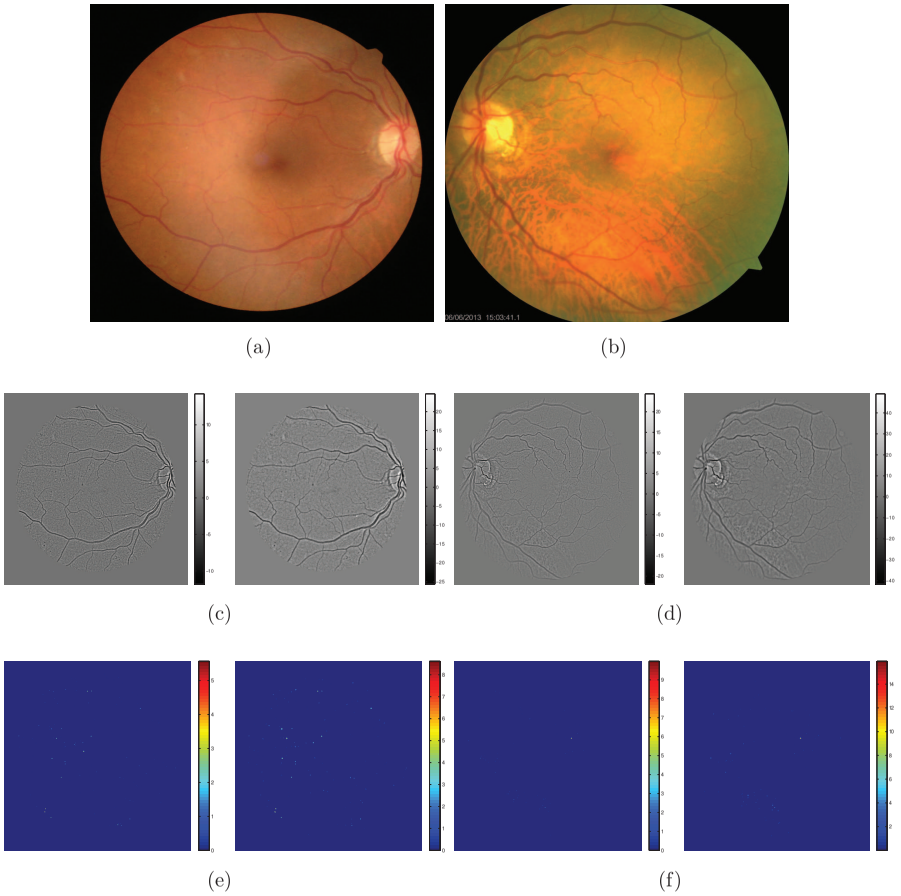


FIGURE 2. Visualization of the main steps of MAs Detector. (a) and (b) Retinal fundus images having MAs. (c) and (d) Pair of wavelet level W_2 and sum of wavelet levels W_2 and W_3, for (a) and (b), respectively. (e) and (f) Pair of blob functions MA_2 and MA_3, for (a) and (b), respectively.

as a large connected structure in the binary image, along with some isolated small objects, which are removed from the binary image. In a similar way we

TABLE 2. Values of the parameters used in the classifier

Detector	Parameter	Value	Equation	Section
MAs	MA_2^L	3.5	9	2.6
	MA_3^L	5.5	9	2.6
HEMs	HEM^{L_1}	12	10	2.7
	HEM^{L_2}	6	11	2.7
	H_{mean}^L	9	11	2.7
	H_{ecc}^U	0.7	11	2.7
	HEM^{L_3}	6	14	2.7
	$\mathcal{S}^L$	0.7	14	2.7
	HEM^{L_4}	7	15	2.7
	$\mathcal{S}^L$	0.7	15	2.7
	HEM^{L_5}	8	16	2.7
	$\mathcal{S}^L$	0.7	16	2.7
BLs	BL^{L_1}	11	17	2.8
	BL^{L_2}	6	18	2.8
	$\mathcal{N}_t^L$	1.3	18	2.8
	$\mathcal{E}^U$	0	18	2.8
	B^L	18	19	2.8
	B^U	40	19	2.8
	$\mathcal{N}_c^L$	0.01	19	2.8
	$\mathcal{N}_c^U$	0.2	19	2.8
	C^L	35	19	2.8
	C^U	60	19	2.8

finally for BLs in (17)-(18)-(19). The values were selected from the common sense considerations, not from the goal of improving the performance for any particular data set that we used. In fact we utilized several big datasets and images acquired with different devices, and it can be seen from the tables displayed in this section that the detectors have a consistent performance. Moreover, we would like to stress that the generality of these thresholds were set by analysing the ROC (receiver operating characteristic) curves for each individual criterion, in large data sets.

TABLE 1. Expert Classification of DR Data sets

Data sets	DR Patients (Images)	Normal Patients (Images)	Total Patients (Images)
DR1	11 (44)	345 (1351)	356 (1395)
DR2	18 (72)	341 (1364)	359 (1436)
DR3	466 (1854)	4961 (19651)	5427 (21505)
DR4	225 (900)	5144 (20534)	5369 (21434)

3.2. Results. We use sensitivity and specificity to assess the performance of our methodology, which are widely used measures in the medical community. Their definitions are illustrated in the following formulas

$$\text{sensitivity} = \frac{\text{number of TP}}{\text{number of TP} + \text{number of FN}}, \tag{20}$$

$$\text{specificity} = \frac{\text{number of TN}}{\text{number of TN} + \text{number of FP}}, \tag{21}$$

where TP, FN, FP and TN represent the number of true positives, false negatives, false positives and true negatives frames, respectively. If the frame belongs to the class of disease (DR) frames and it is classified by the algorithm as negative, it is counted as a false negative; if it is classified as positive, it is counted as a true positive. If the frame belongs to the class of non-disease (*i.e.* normal) frames and it is classified as positive, it is counted as a false positive; if it is classified as negative, it is counted as a true negative.

Sensitivity represents the ability of the method to correctly classify the frame containing a lesion (in our case MAs, HEMs and BLs) as lesion frame, while specificity represents the ability of the method to correctly classify a non lesion frame as normal frame. The above definitions are on a per frame basis, and are used to assess the performance of each individual detector (MAs or HEMs or BLs). The performance of the MAs detector, HEMs detector and BLs detector in terms of sensitivity and specificity is shown in Table 3 (with respect to the previous mentioned data sets D_{MA}, D_{HEM} and D_{BL}, respectively).

Our binary classifier, for each detector, depends on certain parameters which are listed in Table 2 along with their values. These parameters are thresholds for the binary classifiers. Their values were chosen based on the intrinsic definitions of the associated criteria: see the qualitative justification of these values for MAs in (9), for HEMs in (10)-(11) and after in (16), and

fill small holes present within the thresholded regions. iii) We remove objects of size less than 150 pixels and fill holes of size greater than 20 pixels]. After removing the vessels from T_n we also apply to T_n the adaptive Wiener filter, using neighborhoods of size $[15, 15]$, for smoothing and enhancing. We have observed that there are still some elongated structures left in T_n. To remove them we apply a technique similar to the vessel segmentation. We take the sum of the wavelet levels W_1, W_2 and W_3 of T_n and extract the darkest 10% of the pixels within the FOV of the same. From the same binary image we remove objects of size less than 75 pixels and fill holes of size greater than 20 pixels. The pixels corresponding to the resulting binary image are then removed from T_n and the corresponding image is denoted by F_n. Finally, we perform the inversion of the image in the following way

$$M_n := \left(\max_{i,j} F_{n;i,j}\right) - F_{n;i,j},\ 1 \le i \le N_x,\ 1 \le j \le N_y, \tag{8}$$

where N_x and N_y are the width and height of the image F_n in pixels. We then look for blob structures in each image M_n $(n = 2, 3)$, using the eigenvalues of the Hessian matrix of M_n and multiscale image analysis (as described in Section 2.2, with $I := M_n$). Hence, at each point of the domain we define the functions, see (2),

$$\text{MA}_n := \max_{s_{min} \le s \le s_{max}} \left(B_s . M_n\right),$$

where s_{min} and s_{max} are the minimum and maximum scales at which the MAs are expected to be found. For our experiments we consider scale s to vary between 2 and 8.

Finally, we consider the maximum of each function MA_n as an input to our MA binary classifier (or equivalently the Mas detector), see (3)

$$\text{MA}_n^{max} := \max_{i,j} \text{MA}_{n;i,j},\ 1 \le i \le N_x,\ 1 \le j \le N_y.$$

Consequently, if for a frame

$$(\text{MA}_2^{max} \ge \text{MA}_2^L) \wedge (\text{MA}_3^{max} \ge \text{MA}_3^L) \tag{9}$$

then it is classified as a frame containing a MA, otherwise the frame is considered to be a normal frame. Here the pair of lower bounds $(\text{MA}_2^L, \text{MA}_3^L)$ filters out frames that are unlikely to have MAs (in the experiments we consider the values (3.5, 5.5), see Table 2 in Section 3).

The previous functions, I_2, I_3, MA_2, MA_3, involved in the definition of the MA detector are illustrated in Figure 2 for two retinal fundus images exhibiting MAs.

2.7. Hemorrhage Detection. Due to the difficulty of identifying HEMs, resulting of their diversity and variability in appearance (shape, size, texture) we have devised 3 criteria, where the third is based directly on 3 different HEMs sizes. Therefore there is a total of five criteria for HEMs. Then a frame is considered as containing an HEM if it satisfies at least one of these five criteria. The motivation for the definition of each criterion (that relies essentially on the hemorrhage visible information) is justified for each case. The input for each HEM criterion is always the green channel of the retinal fundus image, because it exhibits the best contrast for visual HEM detection in wavelets bands.

2.7.1. *HEM Criterion 1.* This first criterion is quite similar to the detection of MAs. The main difference is that because hemorrhages have a bigger size than MAs, and as we know that larger features are visible with improved contrast on higher wavelet levels, we use $I_4 = \sum_{j=2}^{4} W_j$ (see (1), where now the input scalar image is the green channel of the fundus image), instead of I_2 or I_3 as for MAs. To identify the hemorrhage regions we extract the 20% darkest pixels, which we call R_h. Hemorrhages are know be medium size dark red spots, so we create a binary mask by extracting objects of size between 40 and 500 pixels from R_h and call this mask M_h. Finally, we consider the product of I_4, R_h, M_h, and remove the pixels corresponding to the OD. To the resulting image we apply the adaptive Wiener filter using neighborhoods of size $[15, 15]$ and do the inversion similar to (8) to get the image H_h. As a result of these thresholding and smoothing operations, the HEM lesions appear as isolated bright patches or blobs outstanding to their surroundings in H_h. We then look for blob structures in H_h, using the technique described previously in Section 2.2 (with $I := H_h$), and where now s_{min} and s_{max} are the minimum and maximum scales at which the HEMs are expected to be found. From that we get the measure HEM^{max} for our binary classifier (see (3)). If the frame passes the criterion

$$\mathrm{HEM}^{max} \geq \mathrm{HEM}^{L_1} \tag{10}$$

then the frame is labelled as a hemorrhage frame. The threshold HEM^{L_1} is chosen very high (12 in the experiments, see Table 2 in Section 3) so that

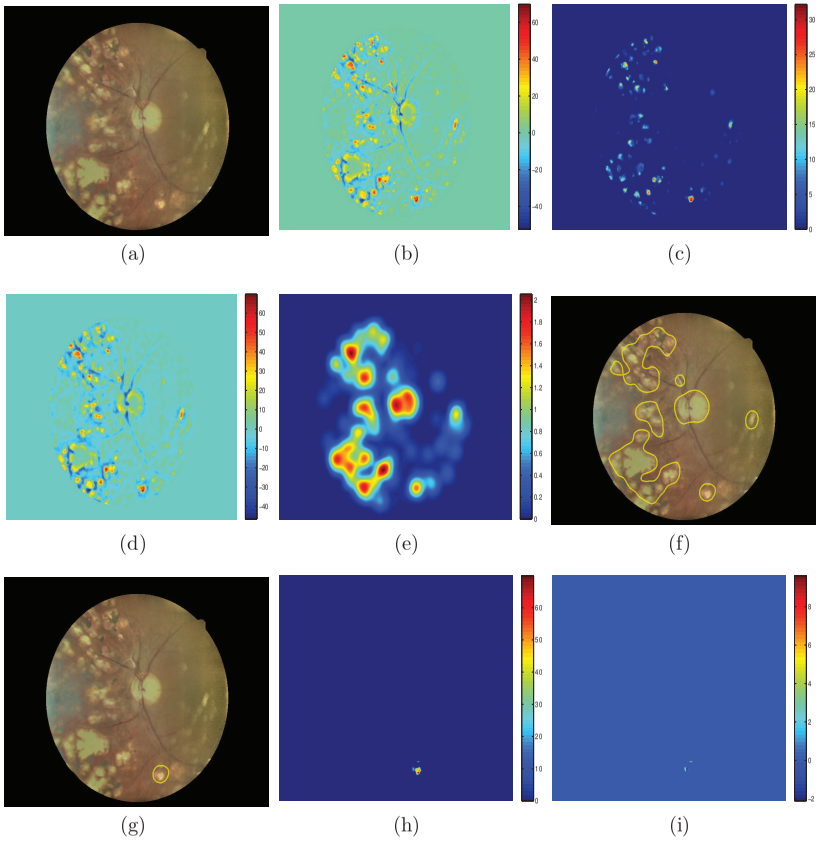


FIGURE 8. Pictures of the principal stages of *BL Criterion 3*. (a) Original image with medium/large BLs. (b) Sum I_b of wavelet levels W_2, W_3 and W_4. (c) Computed blob function in the sum I_b. (d) Cartoon C_b. (e) Convolution type transform χ_c. (f) Segmentation of χ_c superimposed on the original image. (g) Final segmentation region $\mathcal{M}_c$ superimposed on the original image. (h) Cartoon C_b in restricted to the selected segmented candidate $\mathcal{M}_c$. (i) Texture T_b in the final region (for computing $\mathcal{N}_c$).

Figure 8 displays the computations of this *BL Criterion 3*, for one frame which has medium/large BL lesions.

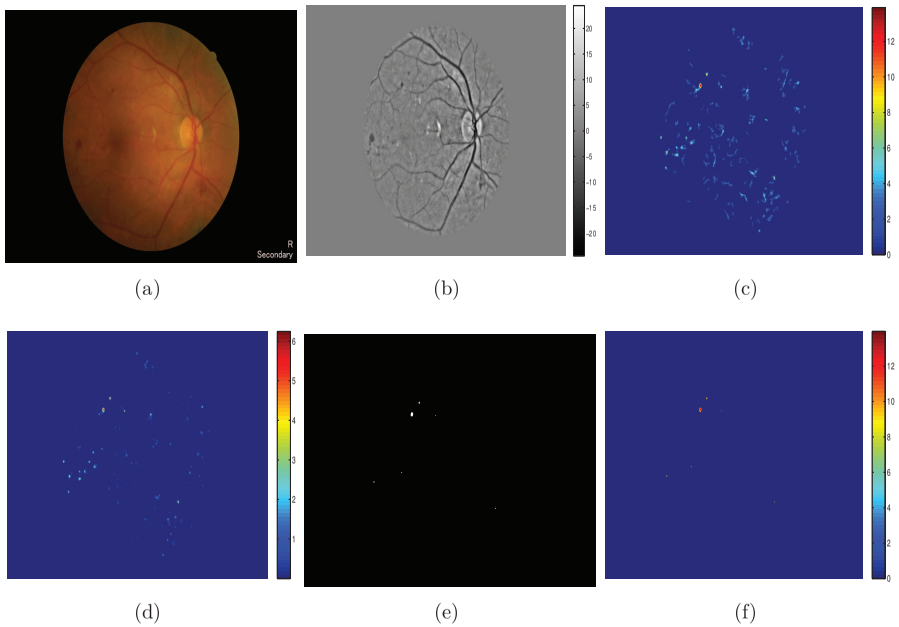
3. Experiments

We now assess the performance of the proposed methodology for automated detection and diagnosis of DR. We have done two different evaluations: first by applying each detector individually and then by considering the detectors in a collective way (that is, a patient is considered a DR patient if he is a positive patient for at least one detector).

3.1. Data sets. For evaluating the performance of the MAs, HEMs and BLs detectors, individually, three data sets, named D_{MA}, D_{HEM} and D_{BL} were prepared by the experts: D_{MA} contains 484 frames out of which 316 are normal and 168 frames contain MAs, D_{HEM} contains 492 frames out of which 309 are normal and 183 frames contain HEMs and D_{BL} contains 607 frames out of which 396 are normal and 211 frames contain the BLs.

Secondly for using the proposed detectors as a system for automated diagnosis of diabetic retinopathy (DR), and test the system performance, we consider four data sets prepared by the experts, named DR1, DR2, DR3 and DR4 (see Table 1).

All these images are provided by the company Retmarker (http://www.retmarker.com/), and were obtained from patients screened according to the Diabetic Retinopathy Screening Program of Portugal. The screening includes the acquisition of 2 images per eye and per patient (except in DR3 and DR4, where a few of the normal patients have less than 4 images) without pupil dilation. The images corresponding to the data sets DR1 and DR2 are taken by the Topcon TRC NW-100 non-mydrictic retinal camera, and for the data sets DR3 and DR4 by a nonmydriatic 45-degree fixed Canon CR6-45NM fundus camera attached to a Sony Power HD 3CCD digital color camera. The images were stored in jpeg files (no compression) with resolutions of 1024×682, 768×584, and 768×768 pixels. The true classification of the disease patients, with DR, and normal patients is shown in Table 1. The classification is done on a per patient basis. This means that if at least one image of a patient contains any type of lesion related to the definition of DR, it is considered as a disease (DR) patient. This ground truth classification is also provided by the company Retmarker and is the human grading at the Diabetic Retinopathy Screening Program.



(a) (b) (c)

(d) (e) (f)

FIGURE 3. Illustration of *HEM Criterion 2* (and also *HEM Criterion 1* in subfigure (d)). (a) Original image with HEMs. (b) Sum of wavelet levels W_2, W_3 and W_4. (c) Image H_h. (d) Computed blob function in image H_h. (e) Binary segmentation image S (via thresholding). (f) Image H_h in the regions of S.

we can detect HEM with marginal false positives. In Figure 3, the subfigure (c) depicts the blob function, and there we can see that HEM^{max} attains a higher value than 12.

2.7.2. *HEM Criterion 2.* One of the most frequent features of some hemorrhages is that it appears as a red spot, with uneven density and surrounding a smaller punctuate lesion considered to be a MA. Thus this second criterion is a procedure that intends to quantify this property.

In this case we compute essentially other measures from H_h and also use HEM^{max} (both defined in *HEM Criterion 1*). We use a binary segmentation

via thresholding, such that for a pixel x

$$S(x) := \mathcal{H}(H_h(x) - \theta)$$

where $\mathcal{H}$ is the Heaviside function taken pixel-wise

$$\mathcal{H}(x) := \begin{cases} 0, & \text{if } x < 0, \\ 1, & \text{if } x \geq 0, \end{cases}$$

and the scalar threshold θ is defined by

$$\theta := \frac{1}{2} \max_{i,j} H_{h;i,j} \; 1 \leq i \leq N_x, \; 1 \leq j \leq N_y.$$

We find the connected components of S using an algorithm by Haralick and Shapiro [18]. For each connected component we compute the mean pixel intensity in H_h. We consider the maximum of the mean intensities for our binary classification and denote it by H_{mean}. This maximum value is considered since we expect the mean of the pixels to be high corresponding to the selected HEM region. We also compute the eccentricity of the component corresponding to the maximum mean and denote it by H_{ecc}. This information is useful since we expect the selected HEM region to be round in shape after thresholding and smoothing operations. Thus criterion 2 is

$$(\text{HEM}^{max} \geq \text{HEM}^{L_2}) \wedge (H_{mean} \geq H^L_{mean}) \wedge (H_{ecc} \leq H^U_{ecc}). \tag{11}$$

The frame is checked for this criterion. If it passes, then it is classified as a HEM frame, otherwise we continue to the next criterion of our HEMs classification. The lower bounds HEM^{L_2}, H^L_{mean} and the upper bound H^U_{ecc} discard frames without HEMs (the chosen values in the experiments are 6, 9 and 0.7, respectively, see Table 2 in Section 3).

The main steps of this *HEM Criterion 2* are illustrated in Figure 3.

2.7.3. *HEM Criterion 3.* This final criterion includes 3 subcriteria denoted by 3-a, 3-b and 3-c. It is based on different sizes of HEMs. Essentially we look for three sizes of hemorrhages regardless of their shapes : between 100 and 500 pixels, or between 500 and 1100 pixels, or big objects but less than 1500 pixels.

To this end, and for each size class, we essentially perform, firstly, a suitable segmentation of the HEM regions in the frame, then select (with an appropriate procedure that involves the cartoon part of the image) the best candidate hemorrhage region and finally quantify same features in this selected candidate for the binary HEM classification. We explain with all the

Finally the criterion for the detection of clusters of small dotted BLs is

$$(\text{BL}^{max} \geq \text{BL}^{L_2}) \wedge (\mathcal{N}_t \geq \mathcal{N}_t^L) \wedge (\mathcal{E} \leq \mathcal{E}^U). \tag{18}$$

If the frame satisfies this criterion, then it is a BL frame, otherwise we continue to our final criterion of the BLs classification. In the experiments (see Table 2 in Section 3) the lower bounds BL^{L_2} and $\mathcal{N}_t^L$ are 6 and 1.3 and the upper bound $\mathcal{E}^U$ is 0.

Figure 7 illustrates the main computations, described in this *BL Criterion 2*, for one frame, with very bad quality, but clearly exhibiting a cluster of dotted BLs.

2.8.3. *BL Criterion 3.* In our final criterion we look for the large/medium BLs that are more homogeneous. Such regions are missed by the previous BL criteria. Firstly, we consider $I_b.I_b^t$ and remove the pixels corresponding to the OD. We also apply the adaptive Wiener filter using neighborhoods of size $[15, 15]$. Similarly to the procedure described in *BL Criterion 1*, for the wavelet level W_2, we apply the blob detection technique, starting from $I_b.I_b^t$ (and not W_2), and obtain the feature $\text{BL}^{max}_{2:4}$ (as in Section 2.2, (3), and the subscript $_{2:4}$ means that in I_b the sum of wavelet levels from 2 up to 4 is used). Secondly, we extract the candidate regions corresponding to the large/medium BLs. For that (and similarly to the clusters of small dotted BL), we compute the nonlinear convolution-type transform of $I_b^t.C_b$ Afterwards we perform a two-phase segmentation on χ_c (as described in Section 2.3) to get the segmentation mask containing the candidates. We then remove the pixels corresponding to the OD and search for the segmented region that matches the location of the $\text{BL}^{max}_{2:4}$, determined above, to get the selected candidate region $\mathcal{M}_c$. Then we compute the maximum value in the image region $\mathcal{M}_c \cap \left(C_b.I_b^t\right)$ and denote it by C^{max}. We also compute the feature $\mathcal{N}_c = \frac{1}{|\mathcal{M}_c|}||\nabla(I_b^t.T_b)||_{L^2(\mathcal{M}_c)}$, where $|\mathcal{M}_c|$ is the sum of the pixels in $\mathcal{M}_c$. As opposed to $\mathcal{N}_t$ in *BL Criterion 2*, a low value of $\mathcal{N}_c$ indicates that $\mathcal{M}_c$ is a cartoon region (and not a textured region). Our final criterion is

$$(B^L \leq \text{BL}^{max}_{2:4} \leq B^U) \; \wedge \; (\mathcal{N}_c^L \leq \mathcal{N}_c \leq \mathcal{N}_c^U) \wedge \; (C^L \leq C^{max} \leq C^U). \tag{19}$$

If the frame satisfies this criterion, then it considered to be a BL frame, otherwise it is considered to be a normal frame. The values for the lower and upper bounds are indicated in Table 2 in Section 3. The upper bounds B^U, C^U are imposed for deleting fames with artifacts and $[\mathcal{N}_c^L, \mathcal{N}_c^U]$ is an estimated low valued interval for the feature $\mathcal{N}_c$.

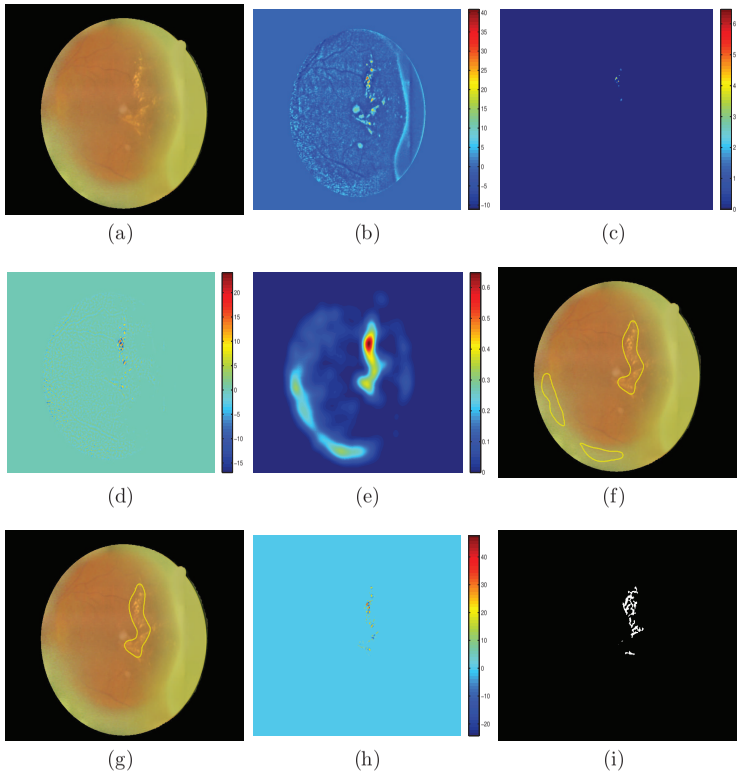


FIGURE 7. Depiction of the several main procedures for *BL Criterion 2* (and also for *BL Criterion 1* in subfigure (c)). (a) Original image with a cluster of dotted BLs. (b) Sum of wavelet levels W_2, W_3 and W_4. (c) Computed blob function in wavelet level W_2. (d) Texture T_b (e) Convolution type transform χ_t. (f) Segmentation of χ_t superimposed on the original image. (g) Selected segmented candidate mask $\mathcal{M}_t$ superimposed on the original image. (h) Sharped image $\mathcal{U}_b$ in the final region (for computing feature $\mathcal{N}_t$). (i) Intersection of the mask $\mathcal{M}_t$ with the vessels (for computing feature $\mathcal{E}$).

details criterion 3-a, and shorten the descriptions of criteria 3-b and 3-c, which are similar.

HEM Criterion 3-a. To achieve a better segmentation, we first enhance and smooth the hemorrhages, and consequently they will have a more roundish shape. To this end we compute the Laplacian of H_h (defined in *HEM Criterion 1*, but using the adaptive Wiener filter with neighborhoods of size $[25, 25]$). We set

$$L_s := -s^2 \Delta G_s * H_h, \tag{12}$$

where G_s is a Gaussian kernel of variance s, Δ is the Laplacian operator and $*$ represents the convolution. Taking the maximum with respect to the scales we get

$$L := \max_{s_{min} \le s \le s_{max}} L_s. \tag{13}$$

We then normalize the pixel intensity values of L and H_h in the range 0 to 255 and to emphasize the hemorrhage regions we consider $D_h := L + H_h$. Secondly, we apply the blob detection algorithm described previously in Section 2.2 to D_h and denote the resulting image by P_h (that is in (2) $I := D_h$, B is denoted by P_h and the scale vector $s = [10\ 12\ 14\ 16]$). We now perform the two-phase variational segmentation, described in Section 2.3, on the scalar image P_h, aiming at separating the hemorrhage regions, from the other parts: the segmentation mask is denoted by u, and contains the HEM candidates.

Decomposing $I_4 = C_h + T_h$, into the cartoon (geometric) part C_h and the texture (oscillatory) part T_h, using the algorithm of Section 2.4, we set then $S_h = C_h.R_h.u$, for discarding some superfluous segments from the segmentation mask u. In fact, because HEM regions appear as medium/large size geometric objects in I_4 they go to the cartoon part C_h (rather than to T_h), which is relevant for the procedure described in the next sentence. Afterwards, for each connected component in S_h we compute the area and the mean intensity of the pixels. We consider at most five components corresponding to the higher values of the product between the area and the mean intensity for the location of the HEMs. The resulting group of candidates is denoted by $\mathcal{C}$.

Then, within the FOV of the sum of the wavelet levels W_2 and W_3, we compute the darkest 10% of the pixels. From the resulting image a binary image is created and the objects of size less than 150 pixels are removed and the holes of size greater than 20 pixels are filled. We denote the resulting binary

image by Q_h. We shall compute some measures for hemorrhage classification from Q_h.

In this first criterion we quantify some properties related to HEMs of size between 100 and 500 pixels from a binary image M_h (created in a similar way as the image also called M_h in *HEM Criterion 1*, for objects between 40 and 500 pixels). Then to the product $I_4.R_h.M_h$ we proceed exactly as in *HEM Criterion 1* to generate an image H_h for which we compute HEM_a^{max} (following the technique of Section 2.2, with $I := H_h$, and the subscript letter a symbolizes *criterion 3-a*).

If the location of the previous value HEM_a^{max} matches with the location of a candidate region in $\mathcal{C}$, then based on the location of the centroid of the latter we define a region of interest (ROI) of size 60×60 pixels in Q_h. If there is no match found then we continue to the next step, *HEM criterion 3-b*. In the ROI we look for the object which is at the minimum Euclidian distance from the centroid of the matched component. For the same object we compute the solidity (a scalar value specifying the proportion of the pixels in the convex hull that are also in the object) and denote it by $\mathcal{S}$. If for the frame the following condition is satisfied

$$(\text{HEM}_a^{max} \geq \text{HEM}^{L_3}) \wedge (\mathcal{S} \geq \mathcal{S}^L), \tag{14}$$

(where HEM^{L_3} and $\mathcal{S}^L$ are pre-defined values) then the frame is a HEM frame, else we continue to look for the objects of other sizes.

HEM Criterion 3-b. This criterion is similar to the previous *HEM criterion 3-a* (including formulas (12), (13) and the group of candidates $\mathcal{C}$, obtained via I_4). The only difference now is that we extract the 20% darkest pixels from the sum of wavelet levels up to level 5, *i.e.* $I_5 = \sum_{j=2}^{5} W_j$, and then from there we select the objects of size between 500 and 1100 pixels, to create the binary image M_h. Again we get an image H_h (as in the previous criterion) and applying the blob detection technique of Section 2.2 to H_h we obtain the scalar value HEM_b^{max} (the subscript letter b symbolizes *criterion 3-b*). If the location of HEM_b^{max} matches with the location of a candidate region in $\mathcal{C}$ (as defined exactly in *HEM Criterion 3-a*) then based on the location of the centroid of the latter we apply the previously described procedure for checking the matching in Q_h (as defined in *HEM Criterion 3-a*) and obtain the solidity measure $\mathcal{S}$. If there is no match found then we continue to the next step, *HEM Criterion 3-c*. If for the frame

$$(\text{HEM}_b^{max} \geq \text{HEM}^{L_4}) \wedge (\mathcal{S} \geq \mathcal{S}^L), \tag{15}$$

are more homogeneous. To describe them we consider $I_b = \sum_{j=2}^{4} W_j$, where W_j is the wavelet level at iteration j of the green channel of the enhanced image, see (1). Let I_b^t denote the image corresponding to the 20% highest pixels in I_b. We decompose I_b into $I_b = C_b + T_b$, where C_b and T_b represent the cartoon and texture decomposition of I_b according to the algorithm of Section 2.4.

2.8.2. *BL Criterion 2.* To extract the candidate regions corresponding to the clusters of dotted BLs spread regions, we perform the nonlinear convolution-type transform of $I_b^t.T_b$ in the following way

$$\chi_t = \mathcal{L}_s(|I_b^t.T_b|^\tau),$$

where $\mathcal{L}_s$ is a linear operator convolving the image with a Gaussian kernel with variance s, and the positive scalar $\tau < 1$ de-emphasizes very small regions with strong texture (see [21], where we have used a similar argument). Here, the convolution is performed separably along each dimension. After computing χ_t we apply a two-phase segmentation on χ_t as described in Section 2.3. As a result of this we obtain the segmentation mask containing the candidate regions. After removing the pixels corresponding to the OD, we look for the region in the segmented mask that matches the location of the value BL^{max} in (17). The selected candidate is denoted by $\mathcal{M}_t$. To sharpen the lesions we apply the unsharp masking. The sharpening process works by utilizing a slightly blurred version of the image. We compute

$$\mathcal{U}_b = I_b^t.T_b + (I_b^t.T_b - \chi_t).$$

From $\mathcal{U}_b$ we get the feature $\mathcal{N}_t = \frac{1}{|\mathcal{M}_t|}||\nabla \mathcal{U}_b||_{L^2(\mathcal{M}_t)}$, where $|\mathcal{M}_t|$ is the sum of the pixels in $\mathcal{M}_t$ (and $L^2(\mathcal{M}_t)$ is the space of square-integrable functions in $\mathcal{M}_t$). A high value for $\mathcal{N}_t$ indicates that $\mathcal{M}_t$ is a high textured (oscillatory) region.

Our next feature, in this criterion, is based on the observation that clusters of small dotted BLs, that are more spread, create holes in the vessels segmentation. We obtain the segmentation of the vessels from the sum of the wavelet levels W_2 and W_3 of the green channel image using the method described in Section 2.6. Then in the region resulting from the intersection of $\mathcal{M}_t$ with the detected vessels we compute the feature $\mathcal{E}$, which is the difference between the number of components to the number of holes in the region.

2.8. Bright Lesion Detection. For the BLs detection a preprocessing step is essential to algorithm success, aimed at reducing the intra- and inter-image variability. We employ the technique proposed in [10, 16] to normalize luminosity and contrast in fundus images based on a statistical model of the image. In this section we always work with the corresponding enhanced images.

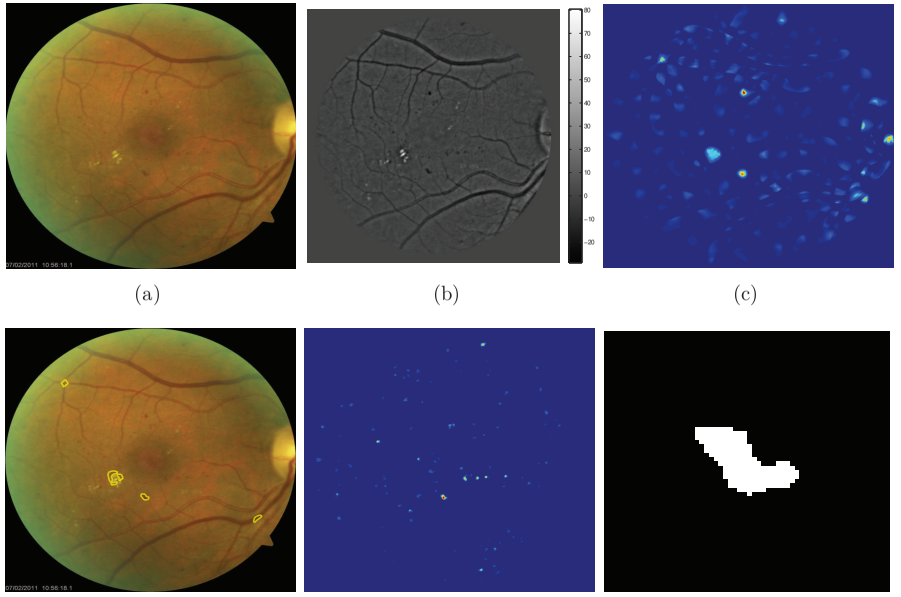
The herein proposed technique for the BLs detection is an improved and extended version of the approach presented by the authors in a conference paper in [12]. To detect BLs our strategy is based on identifying separately small and medium/large BL lesion regions. We propose three criteria. The first is similar to the MAs detection, and aims at identifying small dotted BLs. After that, for the other two criteria we utilize the high wavelet levels and also the texture and cartoon decomposition of the image. In fact, doing the texture+cartoon decomposition of the wavelet levels of the green channel of the enhanced image, we have observed that clusters of small dotted BLs, that are spread across the image, are more prominent in the texture part. On the other hand, large/medium BL regions, that are more homogeneous, go to the cartoon part and, in general, create void regions in the texture part. These facts are also in good agreement with the comments reported in Section 2.4

2.8.1. *BL Criterion 1.* We start by computing W_2, the wavelet level at iteration 2 of the green channel of the enhanced image. Extracting the 20% highest pixels in W_2 we consider the product of the thresholded image with W_2 and remove the pixels corresponding to the OD. The resulting image denoted by F_b contains the small dotted BLs. To smooth and enhance these lesions, we apply the adaptive Wiener filter to F_b using neighborhoods of size $[15, 15]$. Consequently, the small dotted BLs appear as isolated bright patches or blobs outstanding to their surroundings. We then apply the previously described blob detection technique (of Section 2.2) to the scalar image F_b and obtain the measure BL^{max} (see (2)) for the binary classification. If

$$\mathrm{BL}^{max} \geq \mathrm{BL}^{L_1} \tag{17}$$

then the frame is classified as containing a BL (the lower bound BL^{L_1} discards frames without BL and its value is 11, see Table 2 in Section 3).

The approach for BLs detection in [12] utilized only a criterion similar to (17). Our next two BL criteria are different from [12]. Now we look for small dotted BLs that are more spread in the images and large/medium BLs that



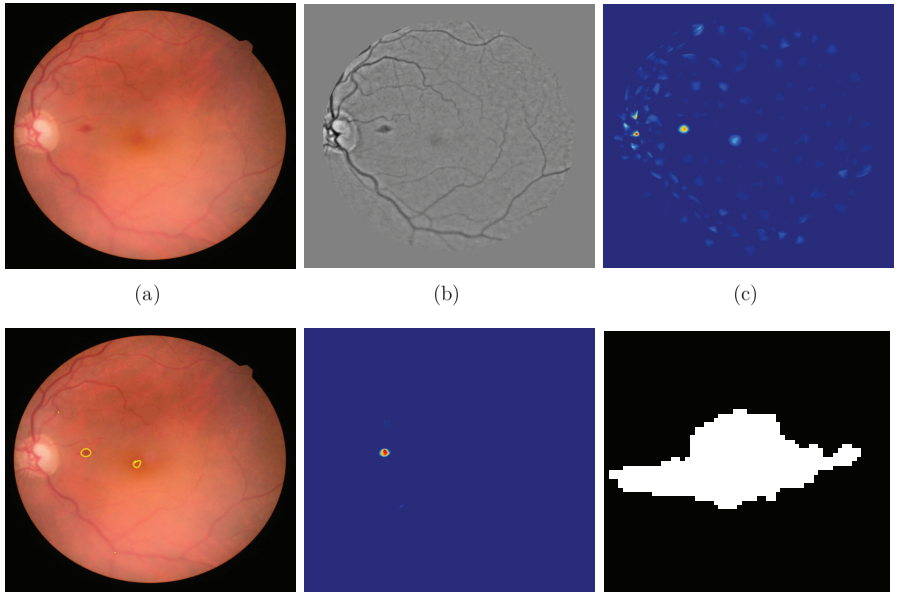
(a) (b) (c)

(d) (e) (f)

FIGURE 4. Visualization of the main steps of *HEM Criterion 3-a*. (a) Original image with HEMs. (b) sum of wavelet levels W_2, W_3 and W_4. (c) Image P_h for segmentation. (d) Final segmentation superimposed on the original image. (e) Blob function related to objects of size between 100 and 500 pixels. (f) Image Q_h in the ROI for computing the solidity.

(where HEM^{L_4} and $\mathcal{S}^L$ are pre-defined values) then the frame is a HEM frame, else we move to our final HEM step.

HEM Criterion 3-c. This criterion is again analogous to the previous HEM criteria (including formulas (12), (13) and the group of candidates $\mathcal{C}$, obtained via I_4). The main alteration is that the binary mask M_h is now created from the extraction of the 20% darkest pixels of I_5 and by removing objects of size less than 1500 pixels (in this process we also fill the holes of size greater than 20 pixels). Again we consider the product of I_5, M_h and the 20% darkest pixels of I_5, remove from it the OD, and the resulting image is smoothed



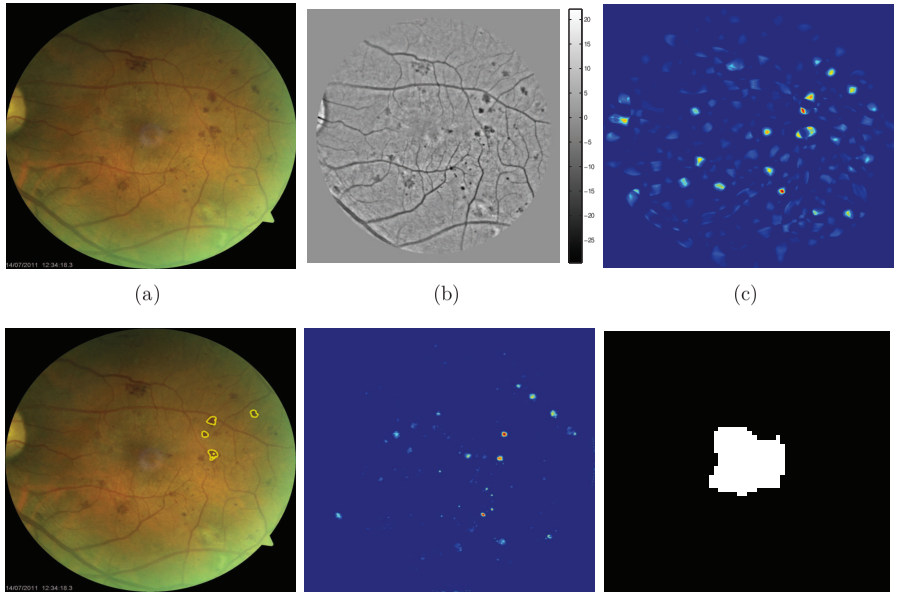
(a) (b) (c)

(d) (e) (f)

FIGURE 5. Illustration of the main steps for *HEM Criterion 3-b*. (a) Original image with HEMs. (b) Sum of wavelet levels W_2, W_3 and W_4. (c) Image P_h for segmentation. (d) Final segmentation $\mathcal{C}$ superimposed on the original image. (e) Blob function related to objects of size between 500 and 1100 pixels. (f) Image Q_h in the ROI for computing the solidity.

using the adaptive Wiener filter of neighborhoods of size $[15, 15]$. After that the inversion is performed similar to (8) to obtain H_h. We again apply the blob detection technique of Section 2.2 to H_h, to get a scalar value HEM_c^{max} (the subscript letter c symbolizes *criterion 3-c*), whose location is used to obtain the solidity measure $\mathcal{S}$ (in a similar way to the previous two criteria). If no match is found in the procedure then the frame is considered to be a normal frame. Our final criterion is

$$(\text{HEM}_c^{max} \geq \text{HEM}^{L_5}) \wedge (\mathcal{S} \geq \mathcal{S}^L). \tag{16}$$



(a) (b) (c)

(d) (e) (f)

FIGURE 6. Visual representation of the relevant parts of *HEM Criterion 3-c*. (a) Original image with HEMs. (b) Sum of wavelet levels W_2, W_3 and W_4. (c) Image P_h for segmentation. (d) Final segmentation $\mathcal{C}$ superimposed on the original image. (e) Blob function related to objects of size up to 1500 pixels. (f) Image Q_h in the ROI for computing the solidity.

If the frame satisfies this criterion then it is considered to be a HEM frame, otherwise a normal frame.

Again in these criteria 3-a, 3-b, 3-c, the pairs of lower bounds $(\text{HEM}^{L_3}, \mathcal{S}^L)$, $(\text{HEM}^{L_4}, \mathcal{S}^L)$, $(\text{HEM}^{L_5}, \mathcal{S}^L)$ filter out frames that are unlikely to have HEMs. In the experiments we consider the values $(6, 0.7)$, $(7, 0.7)$ and $(8, 0.7)$, respectively (see Table 2 in Section 3).

The segmentation of the HEM candidates as well the main steps of *HEM Criteria 3-a, 3-b and 3-c*, for retinal images with HEMs, are displayed in Figures 4, 5 and 6, respectively.