



**HAL**  
open science

## Exudate detection in color retinal images for mass screening of diabetic retinopathy

Xiwei Zhang, Guillaume Thibault, Etienne Decencière, Beatriz Marcotegui, Bruno Laÿ, Ronan Danno, Guy Cazuguel, Gwénolé Quellec, Mathieu Lamard, Pascale Massin, et al.

### ► To cite this version:

Xiwei Zhang, Guillaume Thibault, Etienne Decencière, Beatriz Marcotegui, Bruno Laÿ, et al.. Exudate detection in color retinal images for mass screening of diabetic retinopathy. *Medical Image Analysis*, 2014, 18 (7), pp.1026 - 1043. 10.1016/j.media.2014.05.004 . hal-01082809

**HAL Id: hal-01082809**

<https://minesparis-psl.hal.science/hal-01082809v1>

Submitted on 14 Nov 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Exudate Detection in Color Retinal Images for Mass Screening of Diabetic Retinopathy

Xiwei Zhang<sup>a</sup>, Guillaume Thibault<sup>a</sup>, Etienne Decencière<sup>a</sup>, Beatriz Marcotegui<sup>a</sup>, Bruno Laj<sup>d</sup>, Ronan Danno<sup>d</sup>, Guy Cazuguel<sup>e,f</sup>, Gwénolé Quellec<sup>f</sup>, Mathieu Lamard<sup>e,f</sup>, Pascale Massin<sup>b</sup>, Agnès Chabouis<sup>c</sup>, Zeynep Victor<sup>b</sup>, Ali Erginay<sup>b</sup>

<sup>a</sup>Centre for mathematical morphology, Mathematics and Systems department, MINES ParisTech, 35 rue Saint-Honoré, Fontainebleau, France

<sup>b</sup>Service d'ophtalmologie, hôpital Lariboisière, APHP, 2, rue Ambroise-Paré, 75475 Paris cedex 10, France

<sup>c</sup>Direction de la politique médicale, parcours des patients et organisations médicales innovantes télémédecine, Assistance publique Hôpitaux de Paris, 3, avenue Victoria, 75184 Paris cedex 04, France

<sup>d</sup>ADCIS, 3 rue Martin Luther King, 14280 SAINT-CONTEST, France

<sup>e</sup>Télécom Bretagne, Institut Mines-Télécom, ITI Department, Brest, France

<sup>f</sup>Inserm UMR 1101 LaTIM U650, bâtiment I, CHRU Morvan, 29200 Brest, France

## Abstract

The automatic detection of exudates in colour eye fundus images is an important task in applications such as diabetic retinopathy screening. The presented work has been undertaken in the framework of the TeleOphta project, whose main objective is to automatically detect normal exams in a tele-ophthalmology network, thus reducing the burden on the readers. A new clinical database, e-ophta EX, containing precisely manually contoured exudates, is introduced. As opposed to previously available databases, e-ophta EX is very heterogeneous. It contains images gathered within the OPHDIAT telemedicine network for diabetic retinopathy screening. Image definition, quality, as well as patients condition or the retinograph used for the acquisition, for example, are subject to important changes between different examinations. The proposed exudate detection method has been designed for this complex situation. We propose new preprocessing methods, which perform not only normalization and denoising tasks, but also detect reflections and artifacts in the image. A new candidates segmentation method, based on mathematical morphology, is proposed. These candidates are characterized using classical features, but also novel contextual features. Finally, a random forest algorithm is used to detect the exudates among the candidates. The method has been validated on the e-ophta EX database, obtaining an AUC of 0.95. It has been also validated on other databases, obtaining an AUC between 0.93 and 0.95, outperforming state-of-the-art methods.

**Keywords:** Diabetic Retinopathy screening, Exudates Segmentation, Mathematical Morphology, e-ophta EX database.

## 1. Introduction

Diabetic Retinopathy (DR) is the main cause of blindness among the middle-aged population in developed countries. The first stage of the disease is silent, therefore a regular annual follow-up is recommended to diabetic patients (Massin and Erginay (2010)). Moreover, given the increasing prevalence of diabetes, on the one hand, and the limited – or even decreasing, in some countries – number of specialists, on the other hand, automatic methods to help reducing the burden on specialists are actively developed (Scotland et al. (2007), Philip et al. (2007), Abrmoff et al. (2008), Dupas et al. (2010), Agurto et al. (2011)). Telemedicine networks can contribute to a solution, by improving patient follow-up. The OPHDIAT telemedicine network for DR screening (Massin et al. (2008); Erginay et al. (2008)) was established in 2004 by Assistance Publique - Hôpitaux de Paris. Statistics show that 75% of the exams in this network are considered as normal by the readers (Erginay et al. (2008)). In this context, the TeleOphta project aims at developing an automatic system for detecting normal exams (Decencière et al. (2013)).

In this paper, we focus on the detection of exudates, one of the main clinical signs of the presence of DR. Exudates appear as white/yellow “soft” structures in color retinal images. Their size is variable: they can be as small as a microaneurysm (i.e. a few pixels on a digital image), or as large as the optic disc. Moreover, they also show a large variability in shape and contrast. Fig. 1 shows an example of exudates on a color retinal image and the manual annotation provided by an ophthalmologist.

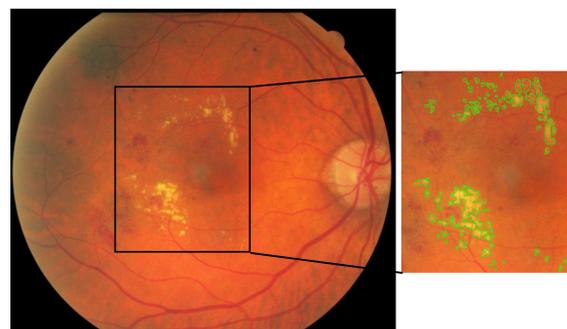


Figure 1: Color retinal image, zoom and manual annotation produced by an ophthalmologist.

Email addresses: xiwei.zhang@mines-paristech.fr (Xiwei Zhang), (Guillaume Thibault), etienne.decenciere@mines-paristech.fr (Etienne Decencière)

The automatic detection of exudates under complex clinical conditions remains an open question. In the literature, recent approaches usually start with a normalization of the image and the removal of the Optic Disc (OD). Then, a first detection method provides a set of candidates, i.e. structures which are similar to exudates. Finally, a selection or a classification procedure is applied, based on features computed on each candidate, in order to keep only exudates. Machine learning methods are typically used for this last step.

Region growing methods are frequently used to obtain the set of candidates (see for example Sinthanayothin et al. (2002)). Usher et al. (2004) proposed a method combining recursive region growing and an adaptive intensity thresholding to detect the candidates. Sánchez et al. (2009) use a mixture model to fit the histogram and estimate the mean and variance of a normal distribution. The estimated parameters are used to perform a dynamic thresholding. After computing a rough detection of exudates, Walter et al. (2002) use a morphological reconstruction to obtain a more precise localization of the exudate boundaries. The same method has been used by Sopharak et al. (2008) and Harangi et al. (2012). We have also adopted it to refine the result of our exudate candidates segmentation method.

The risk of each candidate being an exudate can be modelled as a probability map. Walter et al. (2002) use the local variance to estimate this probability. Giancardo et al. (2011) and Sánchez et al. (2009) use Kirsch operator to estimate it. They show that its value along the edges of each candidate is a good feature to predict the risk. Giancardo et al. (2011) also use stationary Haar wavelets.

More complex machine learning methods can be used for candidates selection. Different sorts of features have been proposed in the literature. Pixel-wise features have been the first to be used. Niemeijer et al. (2007), Sopharak et al. (2009) and Harangi et al. (2012) use pixelwise features like standard deviation, original intensity, hue channel, Sobel edge detector response, or differences of gaussians. In these cases, the authors have used fuzzy C-means and naive Bayes as classifiers. Gardner et al. (1996) proposed another approach: The image is divided into small squares, which are assigned to a class (vessel, exudates, background, etc.) by a neural network. Features based on connected components as candidates have also attracted some interest. Connected component features can be directly derived from pixel-wise features, but supplementary features can also be used, as area, length or perimeter. This approach was used by Fleming et al. (2007) and Sánchez et al. (2010). Sánchez et al. (2012) were the first to use contextual features in this domain. These features take into account the environment of a given candidate. Their results show that the performance is largely improved by using contextual information. Furthermore, Giancardo et al. (2012) proposed an image level classification method. The system classifies an image into one of two classes: “healthy” or “presence of diabetic macular edema”. Based on their previous work, Kirsch’s edge filter is used to get the candidates. Meanwhile, wavelet analysis and different color decompositions are applied to the original image. Then, within the binary mask of candidates, the global mean and standard deviation of the results of the filtered im-

ages are taken as features for the entire image. Each image has one feature vector, which is passed to a support vector machine classifier. The resulting method is computationally efficient. Furthermore it does not need a lesion-level ground truth, but only an image-level ground truth, which is enough for screening applications.

The work presented here builds mainly upon the results of Walter et al. (2002), Sánchez et al. (2012) and Giancardo et al. (2012). However, the application of these methods to our clinical database showed some drawbacks, specially concerning bright structures (such as vessels reflections and optical artifacts), which caused false detections, and small exudates. Concerning this last point, given that our database contained images with very small exudates (a few pixels in size), we decided to avoid image sub-sampling, a step most state-of-the-art methods resort to when dealing with database containing images of different sizes. This was made possible by the introduction of a new spatial calibration method. Moreover, the lack of a publicly available database containing accurately contoured exudates made precise comparisons between methods difficult.

After presenting e-optha EX, the databases used in this work, we describe in the next section the exudates detection method which has been specifically developed with this kind of application in mind. Therefore, it needs to be robust with respect to image variability, and artifacts. In order to achieve these goals, it automatically estimates the size parameters of the detection algorithm, and uses operators which are contrast invariant. Moreover, specific modules have been developed to detect artifacts, such as reflections. Finally, a random forest is used to extract exudates among all the detected candidates. It has been validated on several databases, including e-optha EX. Our work is mainly inspired by Walter et al. (2002), and uses contextual features similar to those proposed by Sánchez et al. (2012).

## 2. Materials: e-optha EX, a new exudates database

There are several databases publicly available for the evaluation of exudate detection methods. Messidor (Dupas et al. (2010)) is a large database containing 1200 images. It provides a DR diagnostic for each image, but does not contain exudate contours. The DIARETDB1\_v2 database, introduced by Kauppi et al. (2007), contains 89 images, and provides rough exudates annotations. However, these annotations are not precise enough to evaluate a segmentation algorithm at the lesion level. HEI-MED (Giancardo et al. (2012)) is a database dedicated to train and test image processing algorithms for the detection of exudates and diabetic macular edema. It contains 169 eye fundus images. Each image of the database was manually segmented by an expert. But it shows the same problem as DIARETDB1: the segmentation is not precise enough for a lesion level evaluation.

We introduce what we think is the best database for exudates segmentation evaluation: e-optha EX <sup>1</sup>. The images of this

---

<sup>1</sup>Details and download of e-optha EX database refer to <http://www.adcis.net/en/Download-Third-Party/E-Ophtha.html>

database have been provided by Assistance Publique - Hôpitaux de Paris. They were captured within the OPHDIAT telemedical network for DR screening (Massin et al. (2008), Erginay et al. (2008)). More than 30 screening centers are members of this network, each working with its own equipment. The trained professionals in charge of acquisitions follow a common protocol, however differences can be seen between images, not the least image size and quality. Moreover, given that the images are sent through the network to the reading center, they are compressed using the JPEG compression standard. All images are acquired with the same field of view angle, i.e.  $45^\circ$ .

In the framework of the TeleOphta project (Decencière et al. (2013)), all images acquired by OPHDIAT between January 2008 and December 2009 were extracted from the database, and made anonymous. These images, with contextual data, constitute the e-ophta database. An ophthalmologist randomly chose 47 images containing exudates from this database, and carefully contoured them using software developed by ADCIS. A second ophthalmologist reviewed the annotations and, when necessary, consensus was reached between the two specialists. This resulted in 2278 annotated exudates. This number might seem very high; it is due to the fact that each exudate component has been individually and carefully contoured by the expert. For example, in Fig. 2, more than 30 exudates are annotated. In addition, 35 normal images, containing no exudates, were chosen from e-ophta by image processing experts. These images often contain structures which tend to mislead exudate segmentation methods, as optical artifacts and vessels reflections, typical on young patients (see for example Fig. 3(d)), and artifacts, as shown in Figs.3(b) and 3(c)). Ophthalmologists checked that these images indeed did not contain exudates. The resulting database, called e-ophta EX, contains therefore 82 images, with four different image sizes, ranging from  $1440 \times 960$  pixels to  $2544 \times 1696$  pixels.

We introduce in the following section a robust method to detect exudates in these difficult conditions.

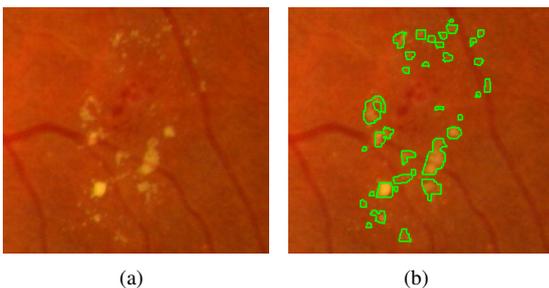


Figure 2: Example of manual exudates annotation. Here, 33 exudates have been contoured.

### 3. Spatial calibration and detection of anatomical structures

Before starting the main processing, we have to introduce several operations which are not specific to exudate detection,

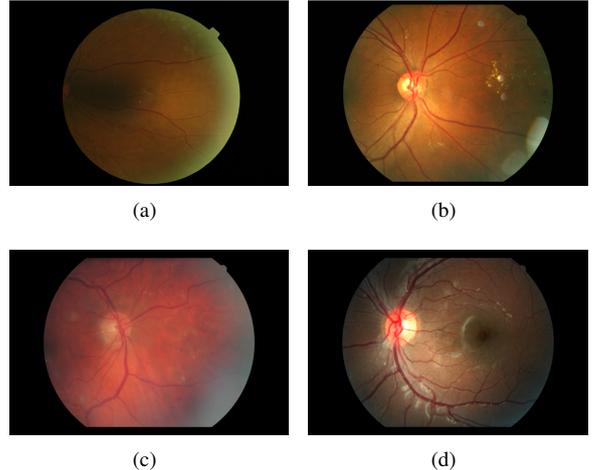


Figure 3: Examples of fundus images in e-ophta EX.

but necessary to provide input information to the exudate segmentation method.

#### 3.1. Spatial calibration

First of all, in order to take into account the different image resolutions, we use a recently introduced spatial calibration method for eye fundus images (Zhang et al. (2012b), Zhang et al. (2012a)). The main idea is to use the diameter  $D$  of the field of view (i.e. the circular region containing the eye fundus information, as opposed to the dark background) as a size invariant. This hypothesis is reasonable as all images from the e-ophta database are acquired with the same field of view angle, equal to  $45^\circ$ .

The field of view (FOV) can be efficiently segmented by a threshold on the sum of the three RGB channels. The largest connected component of the result is supposed to be the FOV. The width of this region gives in practice a good approximation of the FOV diameter  $D$  (see Fig. 4). Once this invariant is computed, it is used to parametrize our exudates detection algorithm, or other detection methods, such as microaneurysms detection (Zhang et al. (2011)), as illustrated in Fig. 4. Images are not re-sized. Our algorithm uses three size parameters:

- $d_1$  is the average diameter of the optic disc;
- $d_2$  corresponds to the maximal width of vessels;
- $d_3$  is the size of the smallest lesions.

In the case of eye fundus images obtained with a FOV angle of  $45^\circ$ , we have experimentally set these parameters to  $d_1 = D/5$ ,  $d_2 = D/74$  and  $d_3 = D/99$ . This is the case for instance of the e-ophta database.

### 4. Anatomical structures detection

We firstly need a rough segmentation of the vascular network of the retina. Any state-of-the-art method would probably be sufficient, as here we are only interested by the main vessels. In

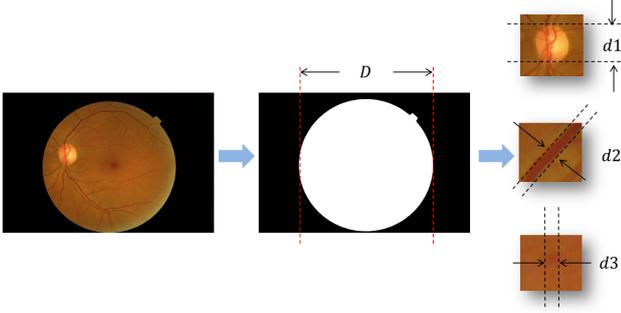


Figure 4: Spatial calibration and parameter estimation.  $D$  is the width of the field of view. Parameters  $d_1$ ,  $d_2$  and  $d_3$  are respectively the optic disc diameter, the maximal vessel width and the size of the smallest lesions (here, a microaneurysm). They can all be derived from  $D$ .

practice, our method is largely inspired by methods based on the residue of a morphological closing (see for instance Walter and Klein (2001)), except that here we use an alternate sequential filter (Sternberg (1986) and Serra (1988)). The proposed vessel segmentation method is detailed in the annex.

The OD is usually the largest and brightest structure in eye fundus images. Given that exudate detection methods can be misled by the OD's white and contrasted aspect, it is a common practice to segment it in a preprocessing phase. The method we use here, as described in the annex, combines intensity information (the OD is supposed to be bright) with vessel structure information (the vessel network is rooted at the OD). In addition, the proposed method has been designed to return a void region when no OD disk is present in the field of view.

## 5. A robust method for exudate segmentation

The exudates segmentation method is composed of three steps: 1) preprocessing; 2) exudate candidates detection; 3) classification and individual risk evaluation.

### 5.1. Pre-processing

Given the complex problems introduced by the clinical database we are dealing with, preprocessing will not only deal with noise removal, but also with more complex problems, such as the detection of reflection zones. We proceed in two steps in order to remove spurious structures. First, we remove all dark structures, including vessels and dark lesions. Second, we get rid of bright artifacts. A clean image will be obtained after this process, simplifying exudate candidates extraction.

Two original ideas constitute the main novelties of this section:

- Bright structures, including reflections, are removed by an adaptive template;
- Bright regions along the borders of the field of view are segmented using the blue channel of the image.

#### 5.1.1. Dark structures removal

Dark structures, including vessels and dark lesions, induce local intensity variations, which can mislead exudate candidates detection methods based on local contrast. We propose a morphological inpainting to remove dark elements. Let  $I_{orig}$  be the original image, and  $\gamma_B^n$  and  $\varphi_B^n$  respectively the morphological opening and closing of size  $n$ , with structuring element  $B$ . In this work, a hexagonal structuring element is used, as it is more isotropic than a square structuring element, while remaining computationally efficient. In order to compute the **inpainted image**  $I_{inp}$ , we proceed in two steps:

$$I^* = \gamma_B^{n+1}(\varphi_B^n(I_{orig})), \quad (1)$$

$$I_{inp} = I_{orig} \vee I^*, \quad (2)$$

where  $\vee$  denotes the supremum operator. Fig. 5 shows an example of inpainted image. Vessels, as well as dark and small structures, are removed. Moreover, given that we want to get rid of dark structures up to the size of the largest vessels, we choose  $n = d_2$ .

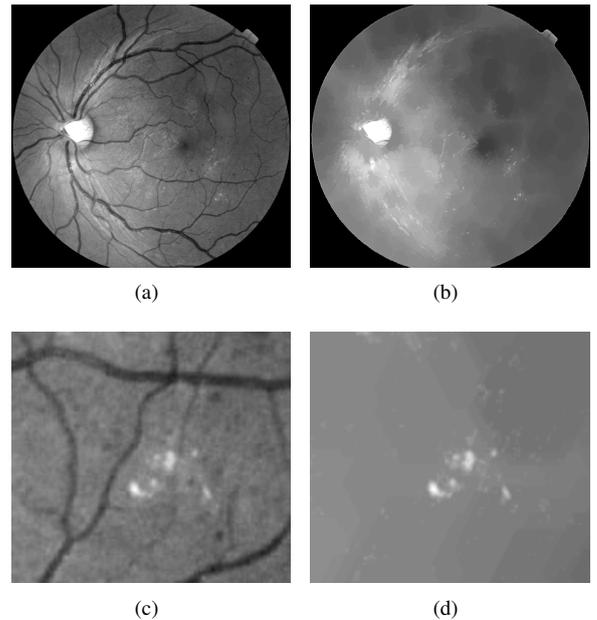


Figure 5: Dark structures removal based on morphological inpainting. a) Original image, b) Inpainted image, c) and d) Some details.

#### 5.1.2. Removal of bright structures

Besides the OD, there are other bright structures in fundus images which can mislead exudate detection methods. Fig. 6 shows three kinds of common bright structures. Arrow  $a$  shows optic nerve fibers, which come out of the OD and are mainly visible along the main vessels. Arrow  $b$  points at reflections in the middle of the vessels. Arrow  $c$  shows reflections which are especially common on young patients, given the extreme transparency of the eye lens and vitreous humour. This kind of reflections is usually close to the vessels. They will be taken into

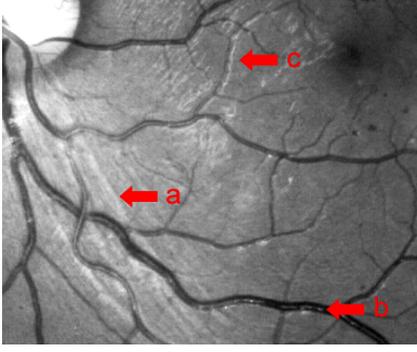


Figure 6: Three kinds of bright structures: a) Optic nerve fibers, b) and c) Reflections common on young patients.

account in the classification section thanks to a specific feature: the distance of a candidate exudate to the nearest vessel.

We propose to use an adaptive template to remove most bright structures. See for example the simple case pictured in Fig. 7, where we see a profile ( $I$ ) containing an exudate  $c$  and some bright structures  $a$  and  $b$ . If we can generate a template ( $T$ ) like the red profile in Fig. 7, which has higher values in the reflection regions, plus a “forbidden” region, where all maxima are ignored, the reflections can be easily removed by combining a subtraction with morphological reconstructions:

$$I_r = Rec_I(I - T) \quad (3)$$

where  $Rec_X(Y)$  means reconstruction under  $X$  by taking  $Y$  as markers. The result  $I_r$  is the third profile in Fig. 7.

In order to recover the original shape of exudates, another reconstruction is done by:

$$I_f = Rec_I(Maxima(I_r)) \quad (4)$$

where  $Maxima()$  extracts the local maxima, and sets them to maximum intensity (255). The green line ( $I_f$ ) is the expected result.

For different kinds of bright structures, the adaptive templates are generated with the methods described hereafter.

Firstly, the normalized local mean value is used to remove non-uniform background illumination effects. The global mean value (within the FOV) is calculated from the original image, denoted  $\mu$ . Then, a mean filter with a window size of  $D/10$  is applied on the same image. The result is named  $I_{mean}$ . The first template is given by:

$$T_{mean} = \frac{\mu I_{mean}}{\max(I_{mean})}, \quad (5)$$

where  $\max(I_{mean})$  is the maximum grey level value of image  $I_{mean}$ .

Secondly, template  $T_{vessel}$ , corresponding to reflections within the vessels, is simply obtained by using the segmented vessel mask. The value of the mask is set to  $\mu$ .

Most reflections or optic fibers are found along the main retina vessels, as shown in Fig. 8(a). They lie inside a parabolic region passing through the OD. Therefore, inspired by work done by Tobin et al. (2007b), we fit a symmetric double

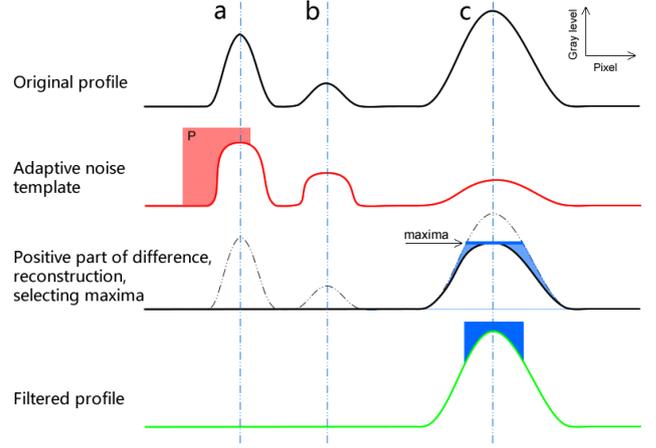


Figure 7: Adaptive template for the removal of bright structures - illustration on an image profile.  $a$  and  $b$  are bright structures, and  $c$  is an exudate. The bright structures can be removed by the following steps: Subtract the template from the original profile and keep the positive part. Perform a reconstruction to restore the profiles of remaining structures. Maxima touching the restricted region  $P$  are removed. Another reconstruction by the maxima is used to restore the exudate.

parabolic region passing through the OD center, to the largest retina vessels (for details refer to the annex). The mask is denoted  $T_{parabola}$ . Its value is set to  $\mu$ . Within this region, we define a “forbidden” sub-region  $P$  corresponding to the parabolic region which bends around the macula, up to a horizontal distance to the OD equal to  $D/4$ , where the value is set to the maximal authorized value of the image (here 255). This region is depicted in white in Fig. 8(b).

The final template is given by the supremum of the first three templates:

$$T_{final} = T_{mean} \vee T_{vessel} \vee T_{parabola}. \quad (6)$$

Fig. 8(b) illustrates the result.

This template image is used as reference to analyse the maxima of the inpainted image: all regional maxima that are lower than the template image  $T_{final}$ , or that touch region  $P$ , are erased. The final result will be referred to as the **preprocessed image**. This treatment is illustrated in Fig. 8. Note that all bright structures which lie in regions where reflections and optical fibers are currently found are suppressed when their value is lower than the template value at the same position. Note that if the OD is absent or not detected, this part will be skipped.

### 5.1.3. Bright border regions detection

Some images contain excessively bright regions along the FOV border (see for example Fig. 9(a), red arrow). These bright regions can cause false detections. To deal with them, we introduce a simple idea: their segmentation will be based on the analysis of the blue channel of the image. Indeed, this channel conveys little information on the retinal structures, whereas these bright regions are clearly visible in it, as shown in Fig. 9(b).

The contour of the field of view (Fig. 9(c)) is taken as marker. A mean filter, with a window size equal to  $D/10$ , is applied to

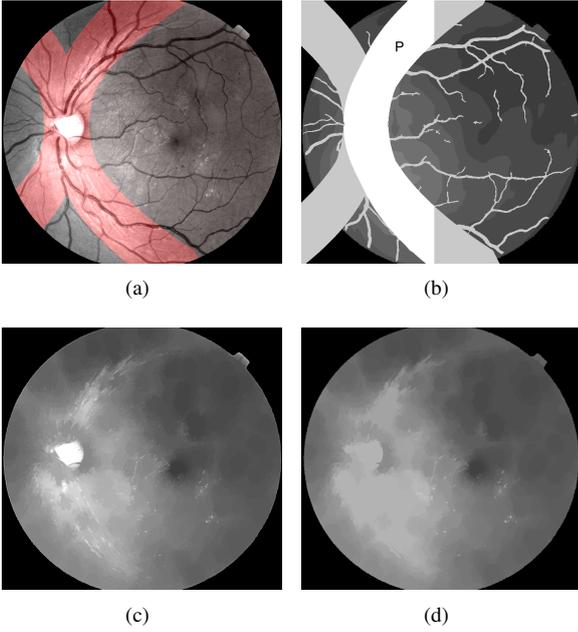


Figure 8: Reflection level estimation and computation of preprocessed image. a) Original image, with overlaid parabolic region, b) Adaptive template, c) In-painted image, d) Preprocessed image, obtained after removing maxima which touch region  $P$ , or that are lower than the template image.

the blue channel, which gives an estimation of the background. The background is removed from the original image, and only positive values are kept. The resulting image is taken as the mask image. A morphological reconstruction is applied to this image, using the previously defined marker, in order to keep only those bright structures which are adjacent to the FOV border. The result of this procedure is illustrated in Fig. 9(d)). Finally, after a threshold (empirically set to 3) and a length opening (which removes all structures with a geodesic length smaller than  $1.3d_1$ ), we get the binary mask of bright border regions. Fig. 10(b) illustrates the result, together with the FOV contour. It will be used in the following step for exudate candidates extraction.

## 5.2. Candidates extraction

In this section, we present a novel two-scale exudate candidates segmentation method. Large exudate candidates are obtained from the preprocessed image, thanks to a mean filter followed by a reconstruction. Small exudate candidates are directly computed on the green channel of the original image, by means of a morphological top-hat.

In the case of large exudates, firstly we compute a mask, which is the union of the FOV contour, the OD mask (modelled as a disc), and bright border regions (if any). This mask is taken as marker to perform a reconstruction under the preprocessed image (Fig. 10(a)). A comparison between the reconstructed image and the preprocessed image is done, and pixels where the values are equal in both images are set to zero. The result is illustrated in Fig. 10(c). This step keeps only relevant regions. The candidates will be only extracted from these regions. Secondly, a mean filter with window size equal to  $D/10$  is applied

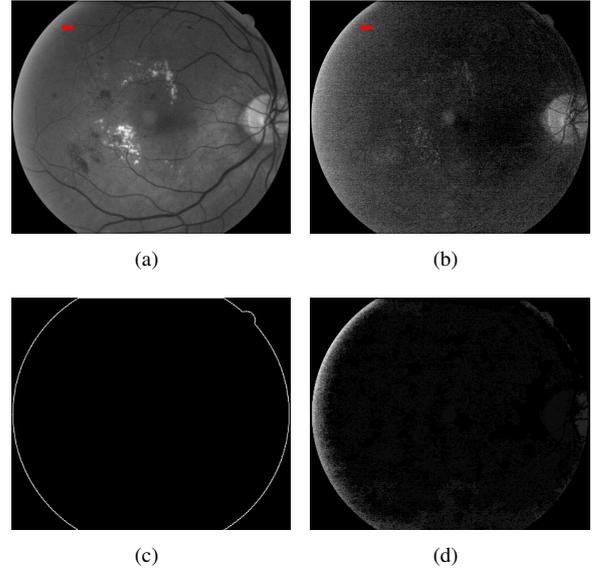


Figure 9: Bright border regions segmentation. a) Green channel, b) Blue channel, c) FOV contour, d) Result of morphological reconstruction.

on the preprocessed image, and the result is subtracted from the relevant regions image; only the positive part of the result is kept. Finally, a threshold of 10 gives the rough candidates mask (see Fig. 10(d)).

In order to get a more precise shape of the candidates - an important feature for the subsequent candidates classification - an efficient technique based on a morphological reconstruction, initially proposed by Walter et al. (2002), is used. We begin by slightly dilating the obtained rough candidates mask, using a structuring element of size  $d_2/2$ . All pixels belonging to the resulting mask are set to zero in the original image, as shown in Fig. 10(e). Then, a morphological reconstruction is applied by taking the previous image as marker and the original image as mask (see Fig. 10(f)). Finally, the difference between the reconstructed image and the original one gives the **large exudate candidates image** (see Fig. 10(g) for an illustration).

Small exudate candidates are missed by this method. In order to detect them, we apply a morphological top-hat with a structuring element of size  $d_3$  to the green channel of the original image. The resulting image contains the **small exudate candidates**. The supremum between this image and the large exudate candidates image gives the **candidates contrast image**. This grey level image will be extremely useful when computing candidate descriptors, as it contains precious information on their shape and contrast. Finally, by thresholding the candidates contrast image at level 5, and removing all connected components which contain less than 5 pixels (the size is fixed for all types of images because noise is almost resolution independent) we obtain the **exudate candidates mask**. Each connected component of this binary image is considered in the following section as an exudate candidate.

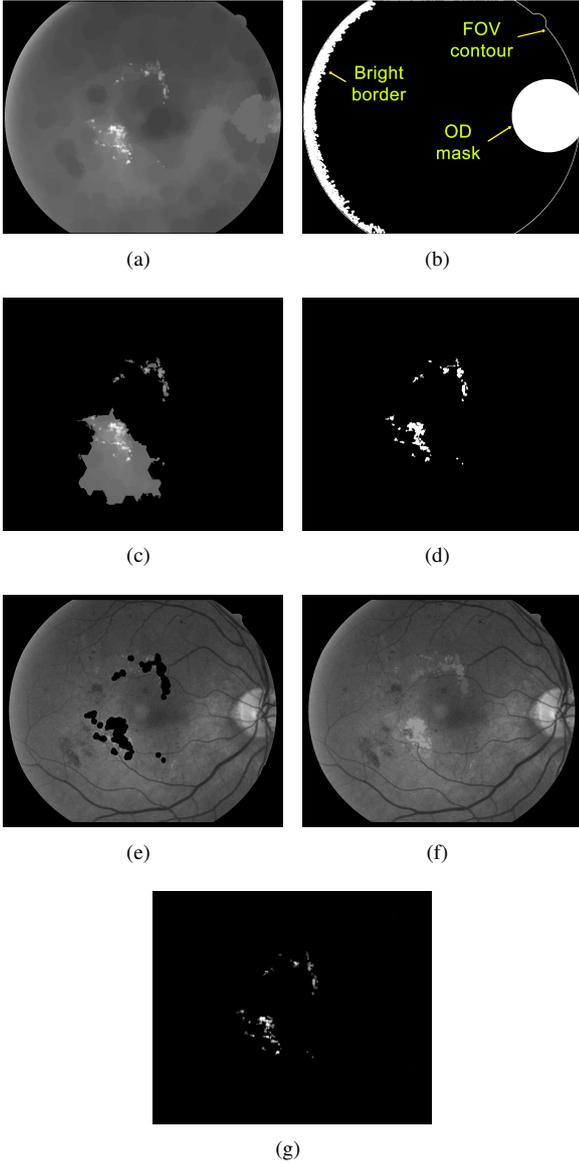


Figure 10: Large exudate candidates extraction illustration. a) Preprocessed image, b) FOV contour, OD and bright border regions mask, c) Relevant regions, d) Rough mask of large candidates, e) Green channel with candidates set to zero, f) Reconstruction, g) large exudate candidates.

### 5.3. Classification and risk evaluation

The candidates set contains exudates, but also other bright structures. In order to classify the candidates, we adopt a classic strategy: we compute features on them, and then use a machine learning technique - here a random forest method (Breiman (2001)) - to classify them.

Two novelties are introduced in this section:

- some features are not only computed on the whole candidate, but also on a smaller support, which corresponds to a different level of the candidates contrast image;
- new contextual and textural features are introduced.

#### 5.3.1. Features

The chosen features are based on the following images: the candidates mask, the candidates contrast image (*contrast*), the inpainted image (*inp*), the preprocessed image (*pre*), and the green and saturation channels of the original image. A typical feature name results from appending the code corresponding to the image taken into consideration, with the code corresponding to the measure. For example, *meanContrast* corresponds to the mean value of the candidates contrast image. The final value is obtained by integrating the measure on the considered image over the exudate candidate. In some cases, the image is not specified in the feature name, either because it is implicit (like for geometric features, e.g. *perimeter*), or for simplicity.

As hinted above, we have discovered that in some cases considering different levels for the same exudate candidate brings useful information into the classification. Therefore we introduce a second level, computed by applying a morphological reconstruction to the candidates contrast image divided by two, under the initial candidates contrast image. The final regions where the integration will take place correspond to the regional maximums of the result. This process is illustrated in Fig. 11. In this case, instead of a single feature, such as *varInp*, we have two derived features *varInp<sub>1</sub>*, computed on the basic layer, and *varInp<sub>2</sub>*, computed on the second layer.

*Intensity features.* The maximum, minimum, mean and median values of each candidate (coded respectively as *max*, *min*, *mean* and *median*) in the candidates contrast and *green* channel images, give the first basic intensity based descriptors. For example, *maxContrast* corresponds to the maximum value of the given candidate in the candidates contrast image. Note that *minContrast* is not taken into account, as its value is always equal to 1 in practice. Descriptor *diffGreen* is the difference between *maxGreen* and *minGreen*.



Figure 11: Candidate layers. Left: exudate candidate in candidates contrast image; middle: same structure divided by 2; right: second layer, obtained with a morphological reconstruction.

The saturation channel is used to distinguish another sort of optical artifacts, which can be introduced by the camera optics. They can appear anywhere on the image, as shown in Fig. 12(a). They are difficult to distinguish from normal exudates in any of the RGB channels of the image (Fig. 12(b)), but we have found that they can be easily detected in the saturation channel (Fig. 12(c)).

The saturation values vary between different image sources. But for the same type of image, the reflections are darker than other bright structures in the saturation channel. Thus, we normalize the saturation channel by dividing it by its global mean value. The final descriptors *meanSat<sub>1</sub>* and *meanSat<sub>2</sub>* are ob-

tained by computing on layers 1 and 2 of each exudate candidate the mean value of the normalized saturation channel.

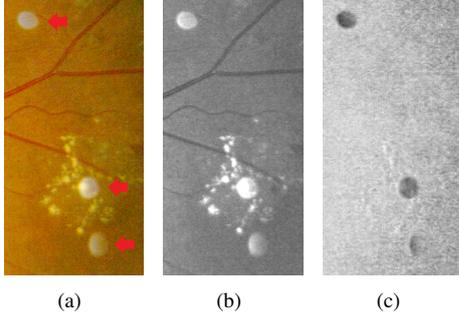


Figure 12: Optical artifacts: a) Color image with three artifacts and small exudates, b) green channel, c) saturation channel.

*Geometric features.* Features *area* and *perimeter* are obtained by counting the number of relevant pixels. Feature *integral-Contrast* is the accumulation of the intensity in the candidates contrast image. These three size dependent features are normalized by dividing them by  $d_2^2$ . The circularity of an exudate candidate is given by:

$$circ = \frac{4A}{\pi L^2}, \quad (7)$$

where  $A$  is its area, and  $L$  its geodesic diameter (see Lantuéjoul and Maisonneuve (1984); Morard et al. (2011)).

*Textural features.* Local variance is used to estimate the contrast change. It is calculated in a sliding square window of width  $d_2$ . The result is shown in Fig. 14(b). We compute it on the inpainted image  $I_{inp}$ , in order to avoid the variance introduced by dark structures. It is moreover computed at two levels, leading to two features  $varInp_1$  and  $varInp_2$ .

The gradient is classically used for edge characterization. Here, we use it to characterize the roughness inside a given region. After computing the morphological gradient of the inpainted image with a hexagon of size 2 inside a given exudate candidate, we compute the value  $n$  of the flooding of the region which removes all minima. Fig. 13 illustrates the method. We call the resulting descriptor *swamping*, as the name of this procedure in mathematical morphology.

*A hybrid feature: the ultimate opening.* The ultimate opening is a multi-scale morphological operator introduced by Beucher (2005). It extracts the most contrasted structures and the corresponding size information from an image. More precisely, for each pixel the ultimate opening keeps the largest difference, denoted  $R$ , between consecutive openings, as well as the size of the opening corresponding to this largest difference (not used here).

In our case, we apply the ultimate opening to the inpainted image. For each candidate, the mean value of  $R(I_{inp})$  within the exudate candidate is computed. Two derived features are obtained:  $uoInp_1$  and  $uoInp_2$ .

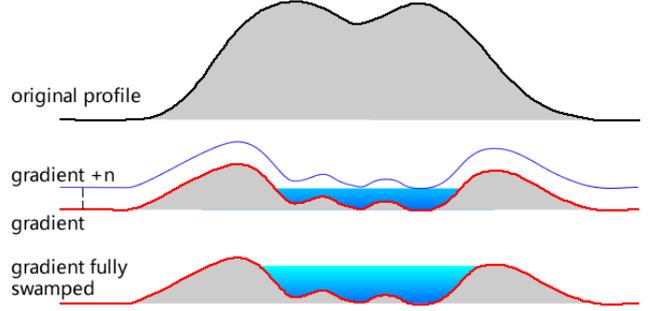


Figure 13: Computing the swamping feature. Top: original profile of an exudate; Middle: its gradient, in red, and the result of adding a constant  $n$  to the gradient. Bottom: when  $n$  is large enough, after reconstruction all local minima are removed.

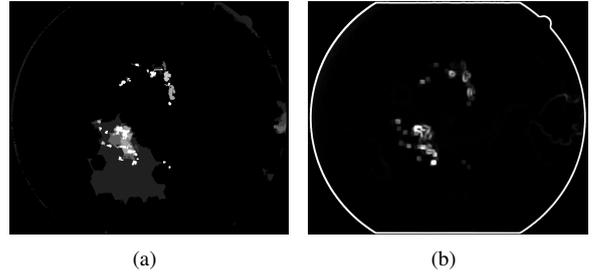


Figure 14: a) ultimate opening, b) local variance computed on the preprocessed image.

*Contextual features.* Reflections and optic fibers often lie beside the vessels, as shown in Fig. 15. This observation was first used by Sánchez et al. (2012) to introduce two contextual features, the distance  $distCenter$  between the barycenter of the candidate to the nearest vessel, and the minimum distance  $distMin$  of the candidate to the nearest vessel. They are normalized by dividing by  $d_2$ .

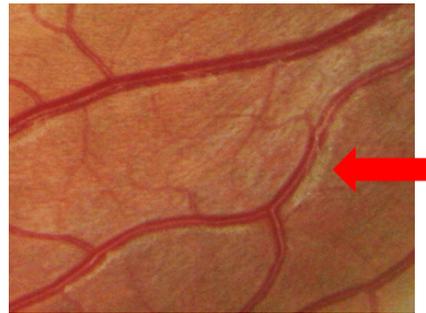


Figure 15: Reflections beside the vessels.

Exudates often appear in a bunch within a small region. Thus, if more than one high risk candidate appears, the risk associated to the corresponding region should be raised. Sánchez et al. (2012) proceeded in two steps to apply this idea: they first classified the candidates into a class (lesion, background, etc.) using local features; then, they used as contextual features the number of candidates withing a neighbour region, and the distance to the nearest candidate, to obtain the final classification.

In order to simplify this process, we propose using the area

Intensity:	<i>maxContrast, meanContrast, medianContrast, minGreen, maxGreen, meanGreen, medianGreen, diffGreen, meanSat<sub>1</sub>, meanSat<sub>2</sub></i>
Geometric:	<i>area<sub>1</sub>, area<sub>2</sub>, volume<sub>1</sub>, volume<sub>2</sub>, perimeter, circularity</i>
Textural:	<i>varInp<sub>1</sub>, varInp<sub>2</sub>, swamping</i>
Hybrid features:	<i>uoInp<sub>1</sub>, uoInp<sub>2</sub></i>
Contextual:	<i>distCenter, distMin, nbMaxH, nbMaxL, areaH, areaL</i>

Table 1: Features list

(code name *area*) and the number of local maxima (code name *nbMax*) close to a given candidate as contextual features. We take as neighbour region a rectangle with a margin of  $d_2$  around the candidate. Within this region, the original green channel is thresholded with two values: a high value  $H$  equal to *meanGreen*, and a low value  $L$  equal to *minGreen*. This leads to four contextual features, *areaH*, *areaL*, *nbMaxH* and *nbMaxL*. This procedure is illustrated in Fig. 16.

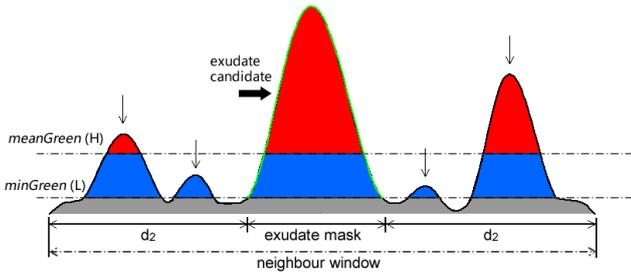


Figure 16: Neighbor structures. An exudate is represented in the middle of the image, with four surrounding structures. Two thresholds are used to select the relevant neighbouring structures. The number of the maxima and their total area are used as descriptors.

*Final set of features.* Finally, twenty eight features are computed on each candidate. They are grouped in Table1.

### 5.3.2. Classification

A Random Forest (Breiman (2001)) is used to perform the classification of the exudate candidates. This method has been chosen as it gives good results in general applications, and is easy to parametrize. The number of trees is set to 500. After a certain threshold, increasing the number of the trees will not bring any improvement, but can reduce the variance of the result. In exchange, computing time increases. We have found that 500 trees provide a good and stable classification in our case. Moreover, according to Breiman (2001), random forests do not overfit. Another important parameter, the number of features to consider for a best split, is set to the square root of the total features number, as recommended by Breiman.

We have used the feature importance definition proposed by Breiman (2001) to evaluate their interest. Given the probabilistic nature of random forests, each run gives a slightly different result. Fig. 17 shows the result which is obtained by taking the

mean value of 5 runs for each feature. As expected, the integral of all the values is equal to 1. We can observe that:

- Contextual features play an important role in the classification. Indeed, among the first five features, three are contextual. This result confirms those of Sánchez et al. (2012). Note also that our new contextual feature *nbMaxH*, appears in the top 5.
- The new hybrid features *uoInp1* and *uoInp2*, based on the ultimate opening, appear also in the top 5.
- The new features based on the image saturation obtain good scores. This good result is probably linked to their discriminant power with respect to reflections.
- Features based on the green channel (such as *meanGreen* and *maxGreen*) do not perform as well as one could expect, probably because the grey level of the green channel is not discriminant enough.
- The geometric features (*area*, *perimeter* etc.) are not interesting for exudates. This result is not unexpected, as exudate size and perimeter show great variability. But *circularity* is an exception, because most reflections have an elongated shape.

The Random Forest algorithm gives a probability of being an exudate to each exudate candidate. Fig. 18, shows such a probability map, which can be compared with the manual annotation. Observe that in this example a low probability is given to each optical artifact, while the probability associated to exudates is high.

## 6. Results

In this section, we evaluate the proposed method on the new database, e-optha EX, as well as on three other publicly available databases: Messidor, DiaRetDB1.v2 and HEI-MED.

The evaluation of the results can be done at two different levels. When an accurate delineation of the exudates is available, the evaluation can be done at the individual exudate level. This approach is especially interesting when comparing different exudate segmentation methods. From a clinical point of view, and especially for screening applications, it is more interesting to evaluate the results at the image level (i.e. presence of absence of lesions).

### 6.1. Exudate level validation

To the limit of our knowledge, this is the first time that an exudate segmentation method can be evaluated at the pixel level on a publicly available database, containing precise lesions annotations. This is possible thanks to the new e-optha EX database.

The evaluation can be classically done by counting the number of pixels which are correctly classified. However, as other authors, we considered this approach inappropriate for exudate segmentation evaluation. Indeed, imagine the situation shown

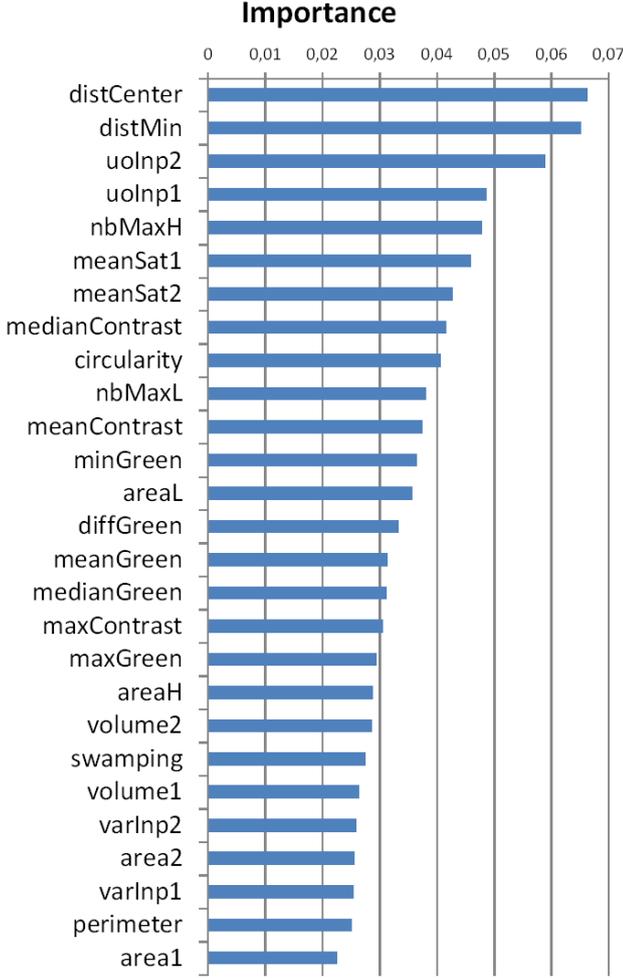


Figure 17: Features importance.

in Fig. 19. In the middle, there is a detected exudate (in blue) and the corresponding ground truth (in red). In practice, most people would consider that this exudate was correctly detected, even if the contours do not match perfectly. If we only count the pixels belonging to the intersection as true positives, we would get half blue pixels as false positives and half red pixels as false negatives. Moreover, this kind of measure would tend to under-estimate errors on small connected components. This is the reason why authors such as Giancardo et al. (2011) and Sánchez et al. (2012) have resorted to connected component level validation: a connected component candidate is considered as True Positive (TP) if, and only if, it touches the ground truth. Thus the connected components of the segmentation and of the ground truth can be classified as True Positives (TP), False Positives (FP) and False Negatives (FN). However, attributing the same weight to a large exudate and a small exudate seems inappropriate. Moreover, with this approach a single very large detection mask would produce excellent results as long as it covers the whole ground truth set. Therefore, a hybrid validation method is proposed below, where a minimal overlap ratio between ground truth and candidates is required.

The main problem is how to evaluate the overlaps between

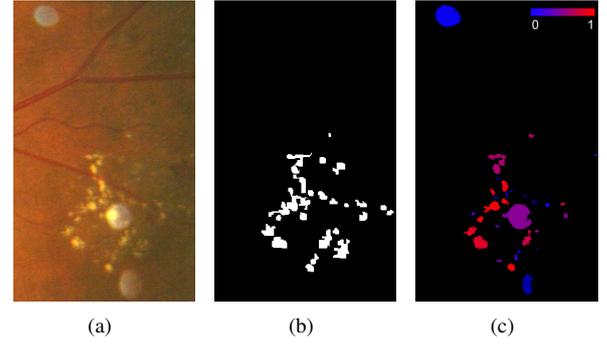


Figure 18: Risk function. a) Original image, containing exudates and three optical artifacts; b) Manual annotations, given by the e-optha database c) Probability map.

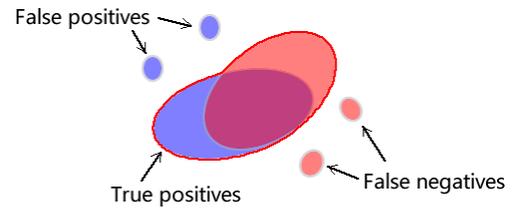


Figure 19: Illustration of the definition of True Positive and False Positive pixels. In blue: detected candidates; in red: ground truth.

detected candidates and ground truth. To proceed with this evaluation, we take as starting point evaluation methods used for document analysis introduced by Wolf and Jolion (2006). The set of candidates is  $\{D_1, D_2, \dots, D_N\}$ , where  $D_i$  corresponds to a connected component. Similarly, the set of ground-truth exudates is  $\{G_1, G_2, \dots, G_M\}$ . The masks of candidates and exudates are therefore respectively

$$D = \bigcup_{1 \leq i \leq N} D_i \text{ and} \quad (8)$$

$$G = \bigcup_{1 \leq j \leq M} G_j. \quad (9)$$

A pixel is considered as a True Positive if, and only if, it belongs to any of the following sets:

- $D \cap G$ ;
- $D_i$  such that  $\frac{|D_i \cap G|}{|D_i|} > \sigma$ ;
- $G_j$  such that  $\frac{|G_j \cap D|}{|G_j|} > \sigma$ ;

where  $|\cdot|$  is the cardinal of a set, and  $\sigma$  is a parameter belonging to  $[0, 1]$ .

A pixel will be considered as a False Positive if, and only if, it belongs to any of the following sets:

- $D_i$  such that  $D_i \cap G = \emptyset$ ;
- $D_i \cap \bar{G}$  such that  $\frac{|D_i \cap G|}{|D_i|} \leq \sigma$ .

A pixel will be considered as a False Negative if, and only if, it belongs to any of the following sets:

- $G_j$  such that  $G_j \cap D = \emptyset$ ;
- $G_j \cap \bar{D}$  such that  $\frac{|G_j \cap D|}{|G_j|} \leq \sigma$ .

All other pixels are considered as True Negatives (TN).

Given that these four classes are, in our case, clearly unbalanced, as TP, FN and FP are in practice negligible with respect to TN, computing the specificity, i.e.  $TN/(FP+TN)$ , and therefore a ROC (Receiver operating characteristic) curve, does not seem appropriate.

We will finally compute the sensitivity  $S$  of the detection on the one hand:

$$S = \frac{TP}{TP + FN}, \quad (10)$$

and the positive prediction value (PPV) on the other hand:

$$PPV = \frac{TP}{TP + FP}. \quad (11)$$

Parameter  $\sigma$  has been set to 0.2. We chose a small value in order to correctly take into account situations such as the one depicted in Fig. 19. We did not want to use  $\sigma = 0$ , which would correspond to the approach proposed by Giancardo et al. (2011) and Sánchez et al. (2012), with the drawbacks described above. Note however that the exact value of  $\sigma$  is not critical, as our main objective is to compare different methods – or to optimize a given method. Fig. 21(b) shows an evaluation result on an image. Note the FNs in the image bottom. Only a few points are detected as candidates; not enough to satisfy the criterion associated to  $\sigma = 0.2$ . Thus, the other pixels in the connected component of the ground truth are regarded as FNs.

The test is done on the e-ophtha EX database with leave-one-out cross-validation. Candidates are grouped by image. The candidates from the test image are left out, while the rest are used to train a model. The test is then performed on the test image, using the ground truth. The procedure is repeated for each image in the database. As we increase the threshold on the probability given by the classification to obtain the detection mask, the sensitivity decreases, and the PPV tends to increase (note however that PPV is not an increasing function of the threshold). Fig. 20 sums up the result.

Table. 2 gives some values, corresponding to different probability thresholds. We can see that a number of pixels are not detected even with the lowest threshold. These are mainly small low contrasted exudates. Fig. 21(d) shows such a case, containing 2400 TP pixels and 1037 FN pixels. Some FPs are due to the presence of other bright lesions, like cotton wool spots and drusens, which are often considered as exudates. This is not a major problem for a screening application.

Before proceeding to the next section, we would like to insist upon the fact that the exudate segmentation evaluation method presented in this section aims at evaluating and comparing exudate detection algorithms. For a clinical evaluation, we believe that image level evaluation, introduced below, is more appropriate.

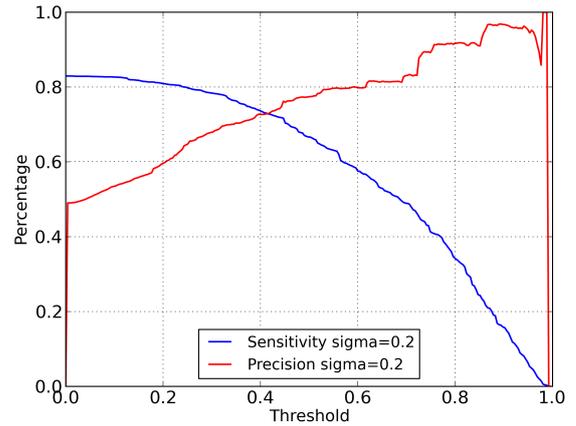


Figure 20: Sensitivity and precision (positive predictive value) for pixel level evaluation.

## 6.2. Image level validation

The result of our exudate detection method is a set of connected components, each of them accompanied by a probability. In order to compute an image level probability, i.e. a probability that the image contains at least one exudate, we take the maximum of all individual probabilities.

In order to evaluate our algorithm on the e-ophtha EX database the same leave-one-out cross-validation is adopted. The resulting ROC curve is shown in Fig. 22(a). The corresponding AUC (Area Under the Curve) is 0.95. For example, with a specificity of 89%, we obtain a sensitivity of 96%. This is a very satisfactory result from a diabetic retinopathy screening point of view.

We have applied to the same database the method proposed by Giancardo et al. (2011), one of the best state-of-the-art methods for exudates segmentation, which was developed for the HEI-MED database. The resulting ROC curve is given on the same diagram. The AUC is 0.87. Our method gives better results on e-ophtha. For example, to achieve the same 96% sensitivity with the method by Giancardo et al. (2011), specificity would drop to 63%.

In order to evaluate the robustness of our method, we have applied it to the other three publicly available databases, DiaRetDB1\_v2, HEI-MED and Messidor, after learning on the whole e-ophtha EX database. Again, note that the image processing (or the candidates extraction) part is automated, for example, the parameterizations of the filters and the thresholdings. We do not change anything while performing the test on the other databases, except for DiaRetDB1\_v2, because their images are acquired with a field of view angle  $50^\circ$ . We adjust the coefficients of spatial calibration. The resulting ROC curves are given on Fig. 22(b). The corresponding AUC values show that the performance remains similar. This result shows that our method is robust with respect to changes in the acquisition conditions.

In Giancardo et al. (2012), the authors proposed a method to predict per image exudate presence probability. The method was tested on the three previously cited databases. Table 3 sums

Threshold	TP pixels	FN pixels	FP pixels	Sensitivity	PPV
0	293,466	60,331	305,279	83%	49%
0.4	261,771	92,026	99,569	74%	72%
0.8	103,311	250,486	9,396	30%	92%

Table 2: Details of pixel level validation.

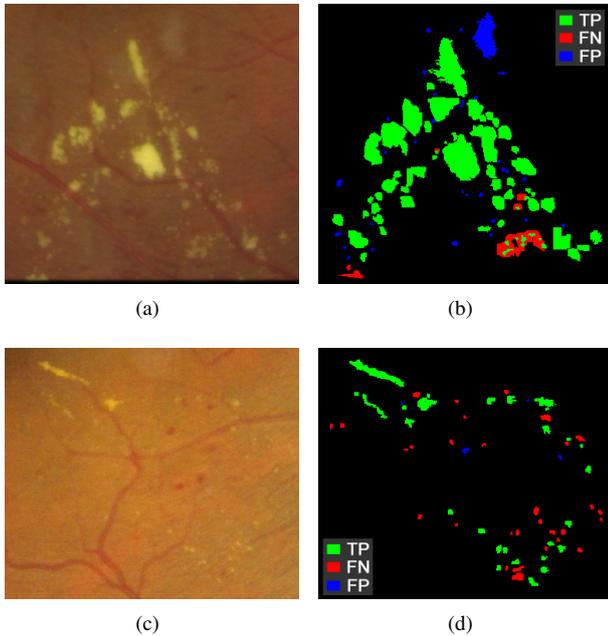


Figure 21: Example of pixel level validation. a) and c) Original images, b) and d) Result of pixel level validation with  $\sigma = 0.2$ , and minimum threshold of probability.

	Proposed method	Giancardo et al. (2012)
DiaRetDB1_v2	0.95	0.93
Messidor	0.93	0.90
HEI-MED	0.94	0.94

Table 3: Comparison of AUC on three public database. The proposed method, optimized for e-optha EX, has been directly applied to the other databases. The method by Giancardo et al. (2012) has been individually optimized for each database.

up the results, where learning and testing have been done on the same database, using leave-one-out cross-validation. In spite of the fact that our method has been optimized for the e-optha EX database, we can see that it obtains better or similar results than the method by Giancardo et al. (2012). Moreover, the performance of this method decreases when learning is done in one database, and testing on another database. For example, according to Giancardo et al. (2012), a model trained on Messidor and tested on HEI-MED gives an AUC 0.82. This shows that the proposed method, developed for a heterogeneous database, tends to be more robust than existing methods when directly applied to other databases.

### 6.3. Processing time

The application of our method to the image of the e-optha EX database of size  $1440 \times 960$  approximately takes 9 seconds on a conventional computer. It takes 35 seconds for an image of size  $2544 \times 1696$ . If needed, this time could be reduced by optimizing the code and computing only the best features used for the classification. In the second case, the computation of the local variance and of the swamping features take respectively 9 seconds and 12 seconds, which occupies 60% computing time. However, it should be noted that this computing time includes the segmentation of the retinal structures and preprocessing steps, which is reused for other lesions detection. Moreover, it is already compatible with off-line diabetic retinopathy screening applications, such as the one developed in the framework of the TeleOphta project.

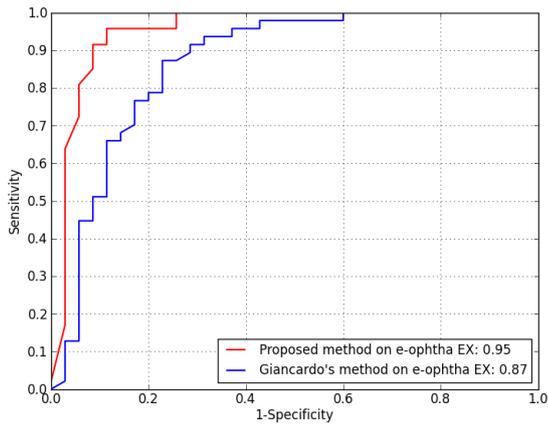
## 7. Conclusion

In this paper, a new method for exudates segmentation on colour fundus images has been introduced. The method is able to deal with images showing large variability in terms of quality, definition and presence of artifacts. It is, to the limit of the authors knowledge, the first method to be able to successfully process images containing reflections, which are frequently found on young patients. In order to reach this objective, the proposed method combines a precise candidates segmentation step, with a classification step, where new features are introduced.

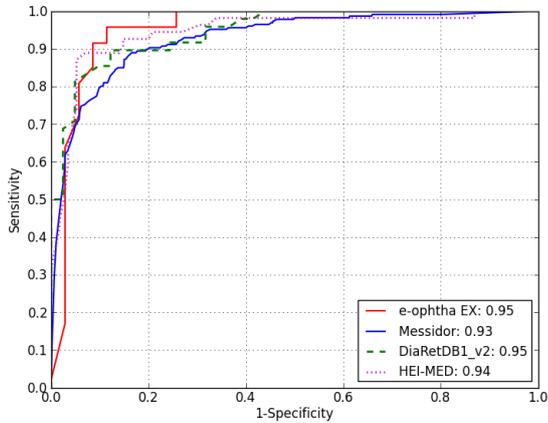
This paper also introduces a new data base, e-optha EX, which contains on the one hand images where the exudates have been accurately contoured, and on the other hand healthy images, without exudates. The proposed method has been validated on this database, and has been compared with state-of-the-art methods on other publicly available databases. The results show that the new method not only performs better than the previous ones, but that it is very robust, as the results remain good on all tested databases, without changing any parameters.

We have invested a great effort into reducing the number of parameters as much as possible. The fact that the results are competitive not only in the e-optha EX database, which was the primary aim of this study, but also on other publicly available databases, without any specific learning or parameter tuning, shows that this goal has been reached. However, there are still some parameters left, mostly corresponding to thresholds. One of our current goals is to use a grey level calibration of the images, based on the same philosophy as our spatial calibration, in order to automatically adapt these parameters to the image content.

The results of this exudates segmentation algorithm are used by the TeleOphta diabetic retinopathy screening system, among



(a)



(b)

Figure 22: ROC curves and AUC values. a) Proposed method (with cross validation) and method from Giancardo et al. (2011) on e-ophtha EX, b) Proposed method (with cross validation) on e-ophtha EX and validation on three public database with the model trained on e-ophtha EX.

other descriptors, to estimate if an examination is normal or not. However, the resulting segmentations can also be used in the scope of other applications, such as therapy follow-up, to quantify the evolution of a pathology.

## Acknowledgements

This work has been supported by the French National Research Agency (ANR) through the TECSAN programme (ANR-09-TECS-015).

X. Zhang and E. Decenci ere would like to thank J.P. Vert and T. Walter for their recommendations concerning the choice of the machine learning method and software (scikit-learn, Pedregosa et al. (2011)).

## Annex A: Main vessels segmentation

The objective of the presented method is to efficiently segment the main vessels of the retina. It is inspired by methods based on the residue of a morphological closing (see for instance Walter and Klein (2001)), except that here we use an alternate sequential filter (see Sternberg (1986) and Serra (1988)).

We have in practice used a threshold on the negative part of the residue of an alternate sequential filter of the green channel. The size of the filter goes from 1 to  $d_2$  (Fig. 23(b)). In order to automatically choose the threshold value, we choose it in such a way that the resulting mask represents around 13% of the FOV area. However, we consider a minimum threshold value of 5, in order to take into account images with few or no vessels. Finally, an area opening (with a size criterion of  $d_2^2$ ) is used to remove small fragments. The method is illustrated in Fig. 23. The main vessels help detecting the Optic Disc (OD) and vascular reflections.

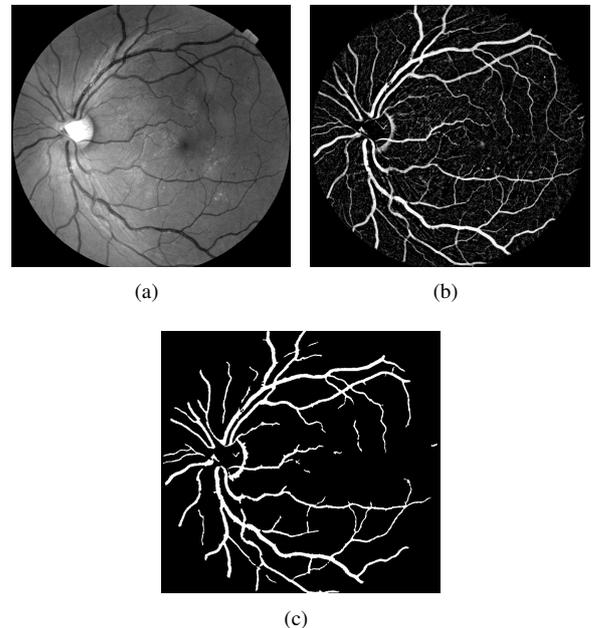


Figure 23: Main vessels segmentation. a) Original image; b) Negative residues of alternate sequential filter; c) Main vessels mask.

## Annex B: Optic Disc detection

There are several sound OD detection methods in the literature. For example, Gagnon et al. (2001), Lalonde et al. (2001) and Eswaran et al. (2008) proposed methods based on OD local properties, such as intensity, size and shape. Other methods are based on the fact that retina vessels come from the OD (Foracchia et al. (2004), Tobin et al. (2007a) and Abdel-Razik Youssif et al. (2008)). These methods tend to be more robust than the first ones. Our OD detection method combines ideas presented above.

Firstly, we use the intensity and size information to get OD candidates. The average of the three original color channels

is taken as the input image ( $I_{input}$ ). A closing is applied to remove dark structures like vessels. Then, a large mean filter with kernel size  $D/7$  is used to estimate the background, and is subtracted from the previous image. Another mean filter with size  $d_1$  is applied to smooth the resulting image. Then a thresholding starts from the maximum intensity value of this image and is lowered until the area of the thresholded image reaches  $0.6 \times d_1^2$ . A binary mask of the OD candidates is obtained, denoted  $I_{ODcandi}$ . Then, vessel information is used to select and verify the candidates.

Secondly, from the vessel mask, width and orientation information is extracted based on a local skeleton analysis (Fig. 24(b)). Vessels usually extend vertically after coming out of the OD. Based on this fact, we first compute the vertical projection of vertical vessels whose width is included between  $0.4d_2$  and  $d_2$ . The profile of such a projection is shown in Fig. 24(c). The position of its maximal response ( $x_{proj}$ ) gives an estimation of the OD horizontal position. Based on the fact that vessel density is large in the OD region, a rectangular window (Fig. 24(a)) is used to compute a vessel density map, as proposed by Tobin et al. (2007a). The location which has the largest density gives a second estimation of the OD position ( $x_{dens}, y_{dens}$ ). The ratio between the number of vertical vessel pixels and other directions ( $R_V$ ) is also computed.

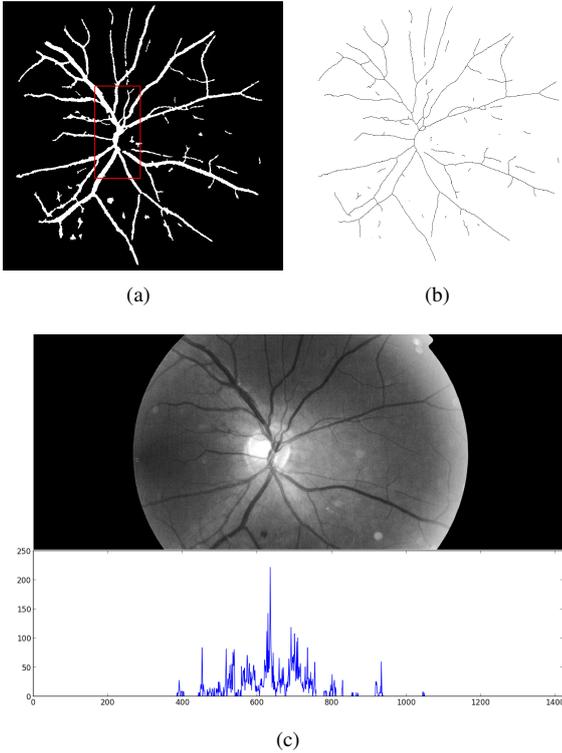


Figure 24: Vessel analysis for OD localization. a) Rectangular moving window for vessel density computation; b) Skeleton of the binary vessel mask; c) Vessels vertical projection.

The final selection is performed based on  $I_{ODcandi}$ ,  $x_{proj}$ ,  $x_{dens}$  and  $R_V$ . First of all, we cope with images from which the OD is absent, such as photographs taken at the peripheral region of

the retina. We found that for this kind of image, there are few vertical vessels, while horizontal vessels tend to cross the entire FOV. Thus, if the  $R_V \leq 0.2$ , we conclude that there is no OD in the image and a void region is returned. Otherwise, we consider two situations:

- If  $|x_{proj} - x_{dens}| < d_1$ , we take  $x_{OD} = (x_{proj} + x_{dens})/2$ . Then, we search for the vertical position of the OD with a moving window ( $d_1 \times d_1$  pixels), along the y-axis on  $I_{input}$ , horizontally limited to  $[x_{OD} - 0.5d_1, x_{OD} + 0.5d_1]$ . The y-coordinate of the position with the maximum mean value within the window is noted  $y_{OD}$ . This process is illustrated in Fig. 25. If there is a candidate OD within  $0.5d_1$  pixels of  $(x_{OD}, y_{OD})$ , it is considered as the detected OD. Otherwise, a void region is returned.

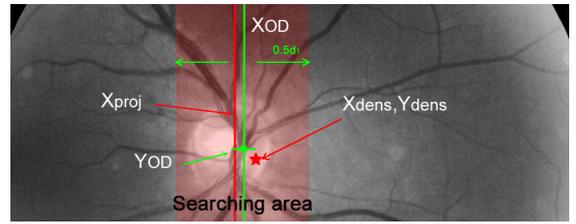


Figure 25: Estimation of OD position using vessel information.

- If  $|x_{proj} - x_{dens}| \geq d_1$ , then we look for a convenient OD candidate around position  $x_{proj}$ , using the method described above. If no candidate is found, then the procedure is repeated, starting this time with  $x_{dens}$ . Finally, if no convenient candidate is found, a void region is returned.

### Annex C: Parabola fitting

A simplified version of the method proposed by Tobin et al. (2007a) is described below. The parabola fitting is based on choosing the best parameter from a predefined set of parameters. First, we remove from the segmented binary vessel skeleton the points corresponding to vessel branches which are less than  $(2/3)d_2$  pixels wide. Then a parabola is formulated as:

$$x - x_{OD} = a(y - y_{OD})^2 \quad (12)$$

where  $x_{OD}$  and  $y_{OD}$  are the coordinates of the OD center. Parameter  $a$  is selected from  $\{0.02, 0.013, 0.01, 0.005, 0.003, 0.0025, 0.002\}$ . The generated parabola is dilated in the horizontal direction by a linear structuring element of width  $(1/15)D$ , to obtain a parabolic region. Thus, according to the parameters list, 7 parabolic regions are generated. The one containing the most filtered vessels is selected.

### References

- Abdel-Razik Youssif, A., Ghalwash, A., Abdel-Rahman Ghoneim, A., 2008. Optic disc detection from normalized digital fundus images by means of a vessels' direction matched filter. Medical Imaging, IEEE Transactions on 27 (1), 11–18.

- Abrahamoff, M. D., Niemeijer, M., Suttorp-Schulten, M. S. A., Viergever, M. A., Russell, S. R., Ginneken, B. v., Jan. 2008. Evaluation of a system for automatic detection of diabetic retinopathy from color fundus photographs in a large population of patients with diabetes. *Diabetes Care* 31 (2), 193–198. URL <http://care.diabetesjournals.org/content/31/2/193>
- Agurto, C., Barriga, E. S., Murray, V., Nemeth, S., Crammer, R., Bauman, W., Zamora, G., Pattichis, M. S., Soliz, P., Jan. 2011. Automatic detection of diabetic retinopathy and age-related macular degeneration in digital fundus images. *Investigative Ophthalmology & Visual Science* 52 (8), 5862–5871. URL <http://www.iovs.org/content/52/8/5862>
- Beucher, S., 2005. Numerical residues. *Mathematical Morphology: 40 Years On*, 23–32.
- Breiman, L., 2001. Random forests. *Machine learning* 45 (1), 5–32.
- Decencière, E., Cazuguel, G., Zhang, X., Thibault, G., Klein, J.-C., Meyer, F., Marcotegui, B., Quellec, G., Lamard, M., Danno, R., et al., 2013. *TeleOphta: Machine learning and image processing methods for teleophthalmology*. IRBM.
- Dupas, B., Walter, T., Erginay, A., Ordonez, R., Deb-Joardar, N., Gain, P., Klein, J.-C., Massin, P., 2010. Evaluation of automated fundus photograph analysis algorithms for detecting microaneurysms, haemorrhages and exudates, and of a computer-assisted diagnostic system for grading diabetic retinopathy. *Diabetes & metabolism* 36 (3), 213–220.
- Erginay, A., Chabouis, A., Viens-Bitker, C., Robert, N., Lecleire-Collet, A., Massin, P., 2008. OPHDIAT: Quality-assurance programme plan and performance of the network. *Diabetes & metabolism* 34 (3), 235–242.
- Eswaran, C., Reza, A., Hati, S., 2008. Extraction of the contours of optic disc and exudates based on marker-controlled watershed segmentation. In: *Computer Science and Information Technology, 2008. ICCSIT'08. International Conference on. IEEE*, pp. 719–723.
- Fleming, A., Philip, S., Goatman, K., Williams, G., Olson, J., Sharp, P., 2007. Automated detection of exudates for diabetic retinopathy screening. *Physics in medicine and biology* 52, 7385.
- Foracchia, M., Grisan, E., Ruggeri, A., 2004. Detection of optic disc in retinal images by means of a geometrical model of vessel structure. *Medical Imaging, IEEE Transactions on* 23 (10), 1189–1195.
- Gagnon, L., Lalonde, M., Beaulieu, M., Boucher, M., 2001. Procedure to detect anatomical structures in optical fundus images. In: *Medical Imaging 2001. International Society for Optics and Photonics*, pp. 1218–1225.
- Gardner, G., Keating, D., Williamson, T., Elliott, A., 1996. Automatic detection of diabetic retinopathy using an artificial neural network: a screening tool. *British Journal of Ophthalmology* 80 (11), 940–944.
- Giancardo, L., Meriaudeau, F., Karnowski, T., Li, Y., Garg, S., Tobin, K., Chaum, E., 2012. Exudate-based diabetic macular edema detection in fundus images using publicly available datasets. *Medical Image Analysis* 16 (1), 216–226.
- Giancardo, L., Meriaudeau, F., Karnowski, T., Li, Y., Tobin, K., Chaum, E., 2011. Automatic retina exudates segmentation without a manually labelled training set. In: *Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on. IEEE*, pp. 1396–1400.
- Harangi, B., Antal, B., Hajdu, A., 2012. Automatic exudate detection with improved naïve-bayes classifier. In: *Computer-Based Medical Systems (CBMS), 2012 25th International Symposium on. IEEE*, pp. 1–4.
- Kauppi, T., Kalesnykiene, V., Kamarainen, J., Lensu, L., Sorri, I., Raninen, A., Voutilainen, R., Uusitalo, H., Kälviäinen, H., Pietilä, J., 2007. Diaretdb1 diabetic retinopathy database and evaluation protocol. *Proc. Medical Image Understanding and Analysis (MIUA)*, 61–65.
- Lalonde, M., Beaulieu, M., Gagnon, L., 2001. Fast and robust optic disc detection using pyramidal decomposition and Hausdorff-based template matching. *Medical Imaging, IEEE Transactions on* 20 (11), 1193–1200.
- Lantuéjoul, C., Maisonneuve, F., 1984. Geodesic methods in quantitative image analysis. *Pattern Recognition* 17 (2), 177187.
- Massin, P., Chabouis, A., Erginay, A., Viens-Bitker, C., Lecleire-Collet, A., Meas, T., Guillausseau, P.-J., Choupot, G., André, B., Denormandie, P., 2008. OPHDIAT: A telemedical network screening system for diabetic retinopathy in the ile-de-france. *Diabetes & metabolism* 34 (3), 227–234.
- Massin, P., Erginay, A., 2010. *Rétinopathie diabétique*. Elsevier Masson.
- Morard, V., Decencière, E., Dokladal, P., 2011. Geodesic attributes thinnings and thickenings. *Mathematical Morphology and Its Applications to Image and Signal Processing*, 200–211.
- Niemeijer, M., van Ginneken, B., Russell, S., Suttorp-Schulten, M., Abramoff, M., 2007. Automated detection and differentiation of drusen, exudates, and cotton-wool spots in digital color fundus photographs for diabetic retinopathy diagnosis. *Investigative ophthalmology & visual science* 48 (5), 2260–2267.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., et al., 2011. Scikit-learn: Machine learning in python. *The Journal of Machine Learning Research* 12, 28252830.
- Philip, S., Fleming, A. D., Goatman, K. A., Fonseca, S., McNamee, P., Scotland, G. S., Prescott, G. J., Sharp, P. F., Olson, J. A., Jan. 2007. The efficacy of automated disease/no disease grading for diabetic retinopathy in a systematic screening programme. *British Journal of Ophthalmology* 91 (11), 1512–1517. URL <http://bj.o.bmj.com/content/91/11/1512>
- Sánchez, C., García, M., Mayo, A., López, M., Hornero, R., 2009. Retinal image analysis based on mixture models to detect hard exudates. *Medical Image Analysis* 13 (4), 650–658.
- Sánchez, C., Niemeijer, M., Išgum, I., Dumitrescu, A., Suttorp-Schulten, M., Abramoff, M., van Ginneken, B., 2012. Contextual computer-aided detection: Improving bright lesion detection in retinal images and coronary calcification identification in ct scans. *Medical Image Analysis* 16 (1), 50–62.
- Sánchez, C., Niemeijer, M., Suttorp Schulten, M., Abramoff, M., van Ginneken, B., 2010. Improving hard exudate detection in retinal images through a combination of local and contextual information. In: *Biomedical Imaging: From Nano to Macro, 2010 IEEE International Symposium on. IEEE*, pp. 5–8.
- Scotland, G. S., McNamee, P., Philip, S., Fleming, A. D., Goatman, K. A., Prescott, G. J., Fonseca, S., Sharp, P. F., Olson, J. A., Jan. 2007. Cost-effectiveness of implementing automated grading within the national screening programme for diabetic retinopathy in scotland. *British Journal of Ophthalmology* 91 (11), 1518–1523. URL <http://bj.o.bmj.com/content/91/11/1518>
- Serra, J. (Ed.), 1988. *Image Analysis and Mathematical Morphology - Volume II: Theoretical Advances*. Academic Press, London.
- Sinthanayothin, C., Boyce, J., Williamson, T., Cook, H., Mensah, E., Lal, S., Usher, D., 2002. Automated detection of diabetic retinopathy on digital fundus images. *Diabetic Medicine* 19 (2), 105–112.
- Sopharak, A., Uyyanonvara, B., Barman, S., 2009. Automatic exudate detection from non-dilated diabetic retinopathy retinal images using fuzzy c-means clustering. *Sensors* 9 (3), 2148–2161.
- Sopharak, A., Uyyanonvara, B., Barman, S., Williamson, T., 2008. Automatic detection of diabetic retinopathy exudates from non-dilated retinal images using mathematical morphology methods. *Computerized Medical Imaging and Graphics* 32 (8), 720–727.
- Sternberg, S. R., 1986. Grayscale morphology. *Computer Vision, Graphics, and Image Processing* 35, 333355.
- Tobin, K., Chaum, E., Govindasamy, V., Karnowski, T., 2007a. Detection of anatomic structures in human retinal imagery. *Medical Imaging, IEEE Transactions on* 26 (12), 1729–1739.
- Tobin, K. W., Chaum, E., Govindasamy, V., Karnowski, T., Dec. 2007b. Detection of anatomic structures in human retinal imagery. *IEEE Transactions on Medical Imaging* 26 (12), 1729–1739.
- Usher, D., Dumskyj, M., Himaga, M., Williamson, T., Nussey, S., Boyce, J., 2004. Automated detection of diabetic retinopathy in digital retinal images: a tool for diabetic retinopathy screening. *Diabetic Medicine* 21 (1), 84–90.
- Walter, T., Klein, J., Massin, P., Erginay, A., 2002. A contribution of image processing to the diagnosis of diabetic retinopathy-detection of exudates in color fundus images of the human retina. *Medical Imaging, IEEE Transactions on* 21 (10), 1236–1243.
- Walter, T., Klein, J.-C., 2001. Segmentation of color fundus images of the human retina: Detection of the optic disc and the vascular tree using morphological techniques. In: *Proceedings of the Second International Symposium on Medical Data Analysis. ISMDA '01. Springer-Verlag, London, UK, UK*, p. 282287. URL <http://dl.acm.org/citation.cfm?id=646351.691036>
- Wolf, C., Jolion, J.-M., Sep. 2006. Object count/area graphs for the evaluation of object detection and segmentation algorithms. *International Journal of Document Analysis and Recognition (IJ DAR)* 8 (4), 280–296. URL <http://link.springer.com/article/10.1007/s10032-006-0014-0>
- Zhang, X., Thibault, G., Decencière, E., 2011. application of the morphological ultimate opening to the detection of microaneurysms on eye fundus images from clinical databases. In: *International Congress for Stereology*. Beijing,

China.

Zhang, X., Thibault, G., Decencière, E., 2012a. Procédé de normalisation d'échelle d'images ophtalmologiques (patent, filing number: 12 53929). "Patent".

Zhang, X., Thibault, G., Decencière, E., Quellec, G., Gazuguel, G., Erginay, A., Massin, P., Chabouis, A., 2012b. Spatial normalization of eye fundus images. In: International Symposium on Biomedical Imaging - ISBI. IEEE.