FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

Detection of red lesions associated with diabetic retinopathy

Tânia Filipa Fernandes de Melo

DISSERTATION

Mestrado Integrado em Bioengenharia

Supervisor: Ana Maria Mendonça

Co-supervisor: Aurélio Campilho

October, 2016

Detection of red lesions associated with diabetic retinopathy

Tânia Filipa Fernandes de Melo

Mestrado Integrado em Bioengenharia

October, 2016

Abstract

Diabetic retinopathy is a disease that affects the vision of a high number of diabetic people. Since microaneurysms are one of the first signs of this disease, a correct detection of these lesions can facilitate diabetic retinopathy diagnosis.

In this way, the main goal of this master thesis is to develop new methodologies for the detection of microaneurysms in color retinal photographs. This includes the development and validation of new methodologies for the enhancement, detection and classification of these lesions.

Regarding to the methodologies that were already developed for microaneurysm detection, they can be divided into two main categories: the approaches based on mathematical morphology and those based on shape analysis that do not use morphological operations.

The methodology developed and herein reported can be included in the second group due to the use of two filters that enhance round objects: the Laplacian of Gaussian and the Sliding Band filter. Although both filters present strong responses in the microaneurysms' location, they also enhance undesirable structures, such as blood vessels and noisy regions. So, the combination of the responses of the two filters as well as the removal of candidates lying on the vasculature lead to the decrease of false positives. A region growing procedure was also applied to obtain the final candidates. Since the seed points which are part of the vessels or the background give rise to large objects, many grown objects overlap. In order to avoid the loss of microaneurysms due to overlapping, for each set of overlapped objects only the one with the smallest area remains.

For reducing the number of false detections, a classification step was also included. In this stage, 36 different features based on the candidates' shape, intensity, color and filters' response were used. For obtaining the final results, 5 different classifiers were compared: a classification tree, a linear and a quadratic discriminant analysis classifier, a 9-NN classifier and a support vector machine classifier. The best result was achieved when the support vector machine classifier was applied (sensitivity = 64.8% and specificity = 79.7%.).

In the future, it is important to study the effect of feature extraction and selection on the classifiers' performance as well as validate the results in other datasets.

Acknowledgments

This work provided me a period of personal and professional growth, accompanied by many good and some difficult moments. Fortunately, these moments were shared with people that in their professional capabilities contributed to the development of this work. Therefore, I would like to acknowledge them.

I dedicate a very special acknowledgment to Prof. Ana Maria by the scientific guidance of this work as well as her permanent availability. Her help was very important during all phases of this project.

My acknowledgments also go to Prof. Campilho that accompanied the project from te beginning and created the necessary material conditions for the realization of this work.

I also want to thank all my friends who accompanied me throughout my academic pathway, especially in the most difficult moments. For all the experiences, sharing of knowledge, fears and victories, my heartfelt thanks.

Last but not least, I want to thank all my family, especially my mammy and my boyfriend, who always gave me the support I needed. For all the love and patience, thank you very much.

I dedicate this thesis to my daddy... Although he is no longer here with me, he is still alive in my heart!

"This work is financed by the ERDF – European Regional Development Fund through the Operational Programme for Competitiveness and Internationalisation - COMPETE 2020 Programme, and by National Funds through the FCT – Fundação para a Ciência e a Tecnologia within project CMUP-ERI/TIC/0028/2014".

"The rung of a ladder was never meant to rest upon, but only to hold a man's foot long enough to enable him to put the other somewhat higher."

Thomas Huxley

Contents

List of Figures

List of Tables

Abbreviations

AUC	Area under the curve
CAD	Computer-aided diagnosis
DM	Diabetes mellitus
DR	Diabetic retinopathy
EA	Explore agents
FP	False positives
FOV	Field of view
FROC	Free-response receiver operating characteristic
k-NN	K-nearest neighbors
LoG	Laplacian of Gaussian
LDA	Linear discriminant analysis
R_{max}	Maximum radius
MA(s)	Microaneurysm(s)
R_{min}	Minimum radius
QDA	Quadratic discriminant analysis
ROC	Receiver operating characteristic
RA	Region agents
SBF	Sliding band filter
SVM	Support vector machine

Chapter 1

Introduction

Diabetic retinopathy (DR) is a progressive disease that affects the vision of a high number of diabetic people. It is characterized by the presence of several retinal lesions, being a leading cause of blindness. In order to prevent the development of this disease, regular eye examinations need to be done in the patients. However, in a screening program, there is a large number of retinal images to be examined manually by ophthalmologists, which is labor intensive, time consuming and expensive. Therefore, there is a great interest in the development of computer-aided diagnosis (CAD) systems for DR diagnosis. Using these tools, only the images where lesions were detected by the system would require the examination by an ophthalmologist.

In sections 1.1 and 1.2, more detailed descriptions of diabetes mellitus (DM) and DR are presented.

1.1 Diabetes Mellitus

1.1.1 Definition and classification

DM is a group of metabolic disorders that are characterized by high levels of blood glucose (Kasper et al., 2015). Depending on the processes that lead to the hyperglycaemia, DM can be classified in three main types:

1. Type 1 DM - Occurs when pancreas' beta cells are destroyed and insulin - the hormone responsible for the uptake of glucose by body's cells - is no longer produced. In this type, patients need daily insulin administration to survive.

2. Type 2 DM - Although there is insulin production, the amount is not sufficient, the body can not process it (insulin resistance) and/or the glucose production is very high. In this type (the most common), daily administration of insulin is not usually required.

3. Gestational diabetes - During pregnancy, women develop glucose intolerance due to metabolic changes. This situation usually reverts post-partum (IDF Diabetes Atlas, 2013).

1.1.2 Prevalence

DM is a public health problem that affected, approximately, 415 million people worldwide in 2015 (46.5% of them are undiagnosed). The majority lives in low and middle income countries and 77% of them are aged between 20 and 64. It is also estimated that the number of diabetic people will increase to more than 50% in the next 25 years (IDF Diabetes Atlas, 2015).

1.1.3 Complications

The high blood glucose levels affect several organs and systems, such as the eye, brain, kidney, cardiovascular and peripheral nervous system. In this way, DM is associated with several complications. Morbidity and mortality are usually related to them.

Those complications can be divided into the following groups and subgroups:

1. Vascular
   (a) Microvascular - Involve small vessels, such as capillaries, and they are diabetes-specific (e.g. retinopathy, neuropathy and nephropathy)
   (b) Macrovascular - Involve large vessels, such as arteries and veins, and they are similar to those in nondiabetics, but occur more frequently in diabetics (e.g. coronary heart disease, peripheral arterial disease and cerebrovascular disease)
2. Non-vascular - e.g. dermatological changes, uropathy, among others (Kasper et al., 2015)

In this work, only retinopathy will be described in detail.

1.2 Diabetic Retinopathy

1.2.1 Definition and classification

One of the organs that is affected by hyperglycaemia is the eye (Kasper et al., 2015). Approximately 33% of diabetic people suffer from retinopathy (National Eye Institute, 2015). It is a microvascular complication that affects the portion of the eye that is sensitive to light, the retina, and it is a leading cause of blindness.

Two distinct stages of diabetic retinopathy can be distinguished:

1. Non-proliferative DR - This stage is characterized by the increase of retinal vascular permeability, alterations in retinal blood flow and abnormal microvasculature, which can lead to retinal ischemia.
2. Proliferative DR - At this advanced stage, there is an abnormal neovascularization as a result of retinal hypoxemia.

Non-proliferative DR does not always progress to proliferative DR. However, the more severe the first stage is, the biggest the probability of evolution to the second one (Kasper et al., 2015).

1.2.2 Signs

The different stages of DR are characterized by the presence of several lesions.

As illustrated in figure 1.1, the main signs of this disease are:

1. Microaneurysms (MAs) - First clinical sign of DR; MAs result from capillary wall outpounching; they are circular, small (their diameter varies from 10 to 100 microns) and appear as red dots (Jelinex and Cree, 2009);

2. Haemorrhages - They are caused by MA rupture and are sometimes indistinguishable from them; the haemorrhages' size can be related to the severity of DR;

3. Hard exudates - They are caused by the breakdown of blood-retina barrier and vary in size;

4. Cotton-wool spots or soft exudates - They appear in more severe stages of DR and are caused by the swelling of local nerve fiber axons that results from the occlusion of precapillary arterioles (Bhavsar et al., 2015).

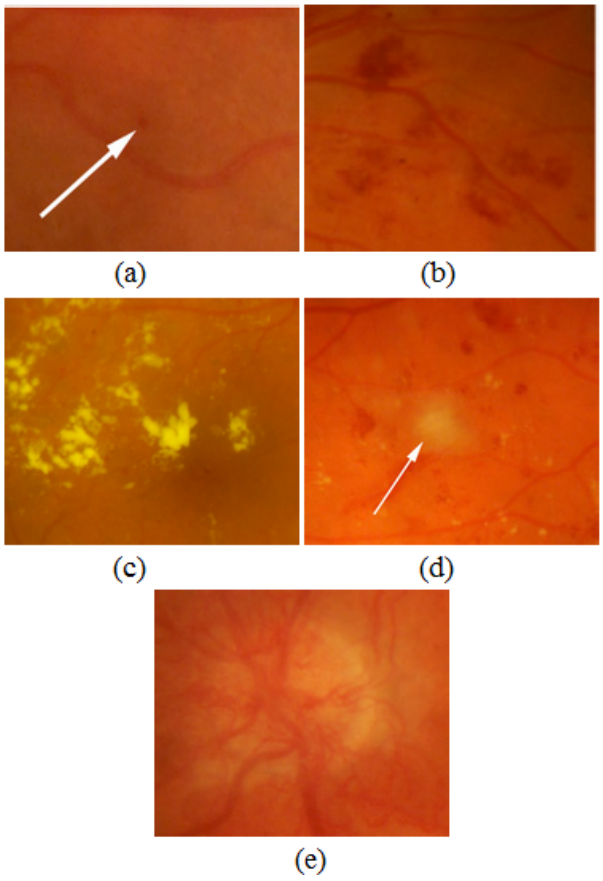

Figure 1.1: Signs of DR: (a) Microaneurysms (marked with an arrow); (b) Haemorrhages; (c) Hard exudates; (d) Cotton-wool spots (marked with an arrow); (e) Neovascularization (Kauppi et al.).

1.3 Objectives

Since it is intended to diagnose diabetic retinopathy at an early stage (through regular examinations) and, at the same time, reduce the number of images to be examined by the ophthalmologists, the main objective of this master thesis is to develop dedicated algorithms based on automated image analysis for detecting microaneurysms in retinal photographs. This includes the development and validation of new methodologies for the enhancement, detection and classification of these lesions. At the end, a comparison between the methodology implemented and the methodologies already developed should also be done.

1.4 Contributions

The main contributions of the work herein reported are:

1. Implementation of new methodologies for MA enhancement based on the sliding band filter (SBF) and the Laplacian of Gaussian (LoG). Although both methodologies have been implemented in other areas (for example, to detect lung nodules or segment cell nuclei and cytoplasm), they have not been used for this purpose.

2. Study of different features and classifiers for the implementation of MA detection.

1.5 Outline of the dissertation

This dissertation has 6 chapters. This chapter gave a brief overview of diabetes mellitus (including its complications) and diabetic retinopathy (definition, its signs and prevention). Furthermore, it focuses on the importance of the development of automated systems for DR screening. Chapter 2 – "State-of-the-Art: Methodologies for MA detection" – presents some of the approaches described in the literature for microaneurysm detection. Chapter 3 – "Methodology for MA detection" – describes the methods implemented and includes images that illustrate the most relevant intermediate steps. In chapter 4, the results obtained are presented and discussed. Chapter 5 – “Conclusions” – presents the achievements of this work and chapter 6 some perspectives to improve the results in the future.

Chapter 2

State-of-the-Art: Methodologies for MA detection

In this chapter, a detailed description of the methodologies developed for detecting MAs in retinal photographs is presented.

Those methodologies can be divided into two main groups: mathematical morphology-based approaches and shape analysis-based approaches which do not use morphological operations (Antal and Hajdu, 2013). For each category, a description of relevant methods is included.

2.1 Methodologies based on mathematical morphology

2.1.1 Early approaches

Läy (1983) is considered the author of the first algorithm for MA detection, which was applied to fluorescein angiography negatives and is based on mathematical morphology. As can be seen in figure 2.1, in this type of images, MAs appear as dark dots. So, local minima were obtained using a bottom-hat transformation. Knowing that some minima corresponded to vessels, the vasculature was extracted with the skeletonization of a further bottom-hat transformation which uses a linear structuring element at different directions. The result of this operation was then subtracted from the image with local minima for obtaining only the red lesions.

Later, several authors presented enhanced algorithms based on the method proposed by Läy (1983). For instance, Baudoin et al. (1984) applied an additional morphological step in order to reduce false detections. Spencer et al. (1992) also included, in the preprocessing step, a radiometric correction for the illumination of the negatives and a shade correction to remove choroidal fluorescence. A matched filtering was also applied, after the vessel removal, to enhance MAs.

2.1.2 The standard approach

Aware of the limitations of the methodologies cited before, Spencer et al. (1996) proposed a new method, which will be here described. Since this approach was the basis for a large number of

Figure 2.1: Example of a fluorescein angiography negative. Microaneurysms appear as dark dots (Spencer et al., 1996).

other detectors, it is named «standard approach».

As in the scheme of figure 2.2, initially a shade correction is performed by estimating the background and subtracting it from the original image. The image background can be estimated using a median or mean filter larger than the blood vessels and MAs' width. In order to eliminate the candidates detected on the vessels, the vasculature is extracted applying a top-hat transformation to the shade corrected image. Afterwards, a matched filter with a 2D Gaussian kernel is applied to the image without vessels in order to enhance the lesions. An example of the images obtained in each processing stage is presented in figure 2.3.

The locations of the first candidates are obtained by thresholding the result of matched filtering. These positions are then used as seed points in a region growing algorithm. In order to reduce the

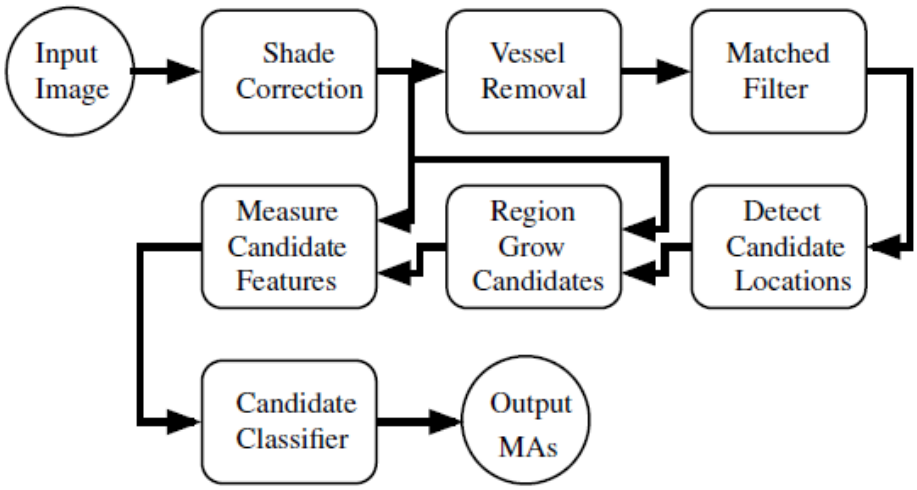

Figure 2.2: Standard approach proposed by Spencer et al. (1996) for microaneurysm detection (Jelinex and Cree, 2009).

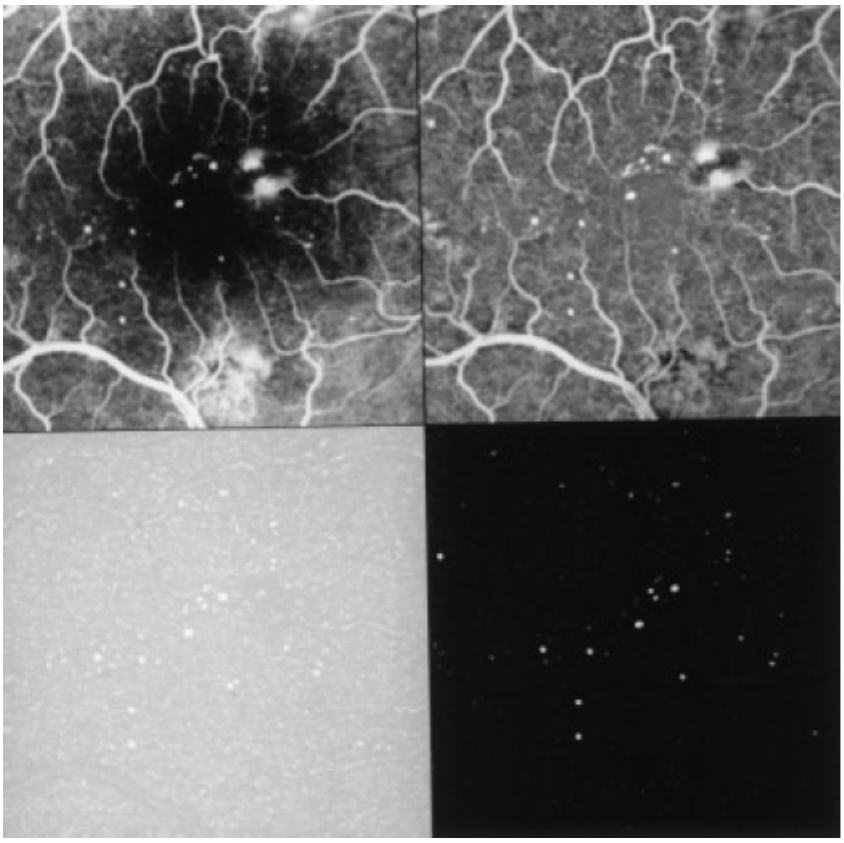

Figure 2.3: An example of the images obtained when the processing stages proposed by Spencer et al. (1996) are applied. (Top-left) Original image (positive version). (Top-right) Shade corrected image. (Bottom-left) Result of the top-hat transformation using a linear structuring element with 11 pixels. (Bottom-right) Result of matched filtering with a Gaussian kernel of size 11 x 11 pixels (σ = 1 pixel).

incorrect detections (false positives - FP), a refinement step is also included. Some features that allow to distinguish between MA and non-MA are extracted from each candidate and introduced in a classifier to obtain the final result.

Cree et al. (1996) improved the method described above by making the region growing algorithm computationally efficient and redesigning the classifier. Extra features based on the intensities of the normalized images and the matched filter response were also added.

2.1.3 Limitations and extensions of the standard approach

Although the previous approach was the first to present a performance similar to the clinicians' analysis detection, it has undergone several extensions and modifications over time. Those changes were done at several levels: in the type of images that were used for diagnosing diabetic retinopathy, the preprocessing methods and MA extractors applied, as well as the type of features and classifiers used in the refinement stage. A description of those extensions will be done in this section.

- **Type of images used for diagnosing DR**

  All the methods described in the previous sections were applied in fluorescein angiographies. Although the contrast between MAs and background is higher in this type of images when compared with color photographs, it is an invasive method and may have some risk for the patient. Furthermore, the conditions necessary to do the exam (such as the injection of fluorescein and the waiting time until images can be obtained) made its application for large scale screening impossible. In this way, Hipwell et al. (2000) extended the method proposed by Cree et al. (1996) to red-free photographs. Later, some improvements to this method were applied by Fleming et al. (2006).

  Methods that use color photographs commonly use the green channel in the preprocessing and detection stages, since this is the component that presents a better contrast between MAs and the background. However, the full color information is used in the final classifier (Niemeijer et al., 2005).

- **Shade correction and contrast enhancement**

  In order to enhance the contrast and remove the noise of the green channel of the color retinal photographs, Walter and Klein (2002) applied, in the preprocessing stage, a new local contrast enhancement operator followed by a Gaussian filtering step. That operator can be defined through the following equations:

$$u = \Gamma(t) = \begin{cases} \frac{\frac{1}{2}(u_{max} - u_{min})}{(\mu_f - t_{min})^r} \cdot (t - t_{min})^r + u_{min} & : t \leq \mu_f \\ \frac{-\frac{1}{2}(u_{max} - u_{min})}{(\mu_f - t_{max})^r} \cdot (t - t_{max})^r + u_{max} & : t > \mu_f \end{cases} \tag{2.1}$$

  where t is the gray level of the original image in the position x, u the corresponding gray level in the enhanced image, $\mu_f(x)$ the mean value of the original image in a window centered in x and r a transition parameter. In figure 2.4, the graph of the corresponding gray level transformation is shown. For $r = 1$ it corresponds to a linear contrast stretching. Figure 2.5 also illustrates the result of the operator application in the green channel of a color retinal photography.

  On the other hand, since in the standard approach the shade corrected image is obtained by subtracting the image background from the original image, Niemeijer et al. (2005) set to zero the pixels of that image with a value higher than zero in order to prevent the appearance of bright pathologies as false positives.

  Recently, Rasta et al. (2015) published a comparative study of preprocessing techniques that can be used for illumination correction and contrast enhancement of retinal fundus photographs. In this work, authors concluded that three techniques, namely the dividing method (which consists in the division of the original image by an estimation of the illumination pattern obtained using the median filter), the quotient-based approach and the homomorphic filtering, may be used for

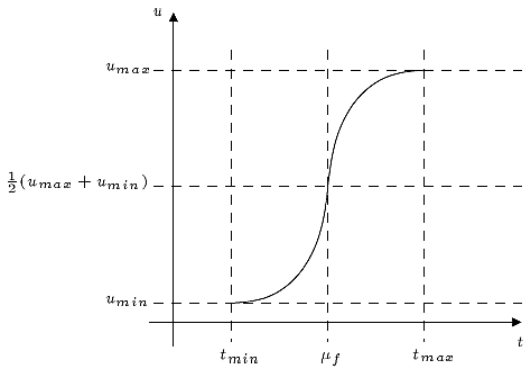

Figure 2.4: The graph of the gray level transformation for a given μ_f (Walter and Klein, 2002).

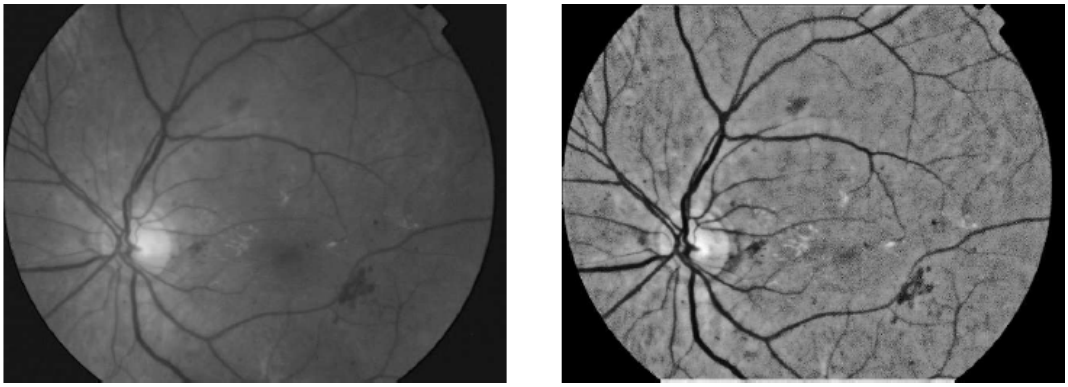

Figure 2.5: The result of the contrast enhancement operator proposed by Walter and Klein (2002). (Left) Original grayscale image. (Right) Enhanced image (r = 5).

correcting the illumination. Furthermore, the contrast-limited adaptive histogram equalization technique showed good potential for enhancing the contrast of this type of images.

- **MA detection**

  Although the vast majority of the morphological-based approaches use the top-hat transformation to obtain the lesions, this method has several limitations. For instance, the length of the linear structuring element must be higher than the microaneurysms' diameter, but not too larger that does not fit the curved vessels. This aspect makes the choice of the length value a difficult task. The authors that used this methodology tend to choose the value of this parameter empirically (based on their training set), which leads to a decrease of the algorithm robustness. This is the main reason that led Niemeijer et al. (2005) to implement an hybrid method for selecting the MA candidates through a combination of candidates detected using mathematical morphology with candidates obtained with supervised pixel classification (the classifier is trained with foreground pixels - from vasculature and MAs - and background pixels).

  Since some false positives also correspond to fragments of tortuous vessels, Walter et al. (2007) proposed an alternative morphological operation for candidates' detection, namely a diameter

closing. This morphological operation allows to remove all dark structures with a diameter smaller than λ pixels. Mathematically, it corresponds to the minimum of all closings using structuring elements with diameter greater than or equal to λ pixels. The preprocessed image is then subtracted from the result of diameter closing, and a threshold is applied in order to obtain the candidates. The results of these operations are shown in figure 2.6.

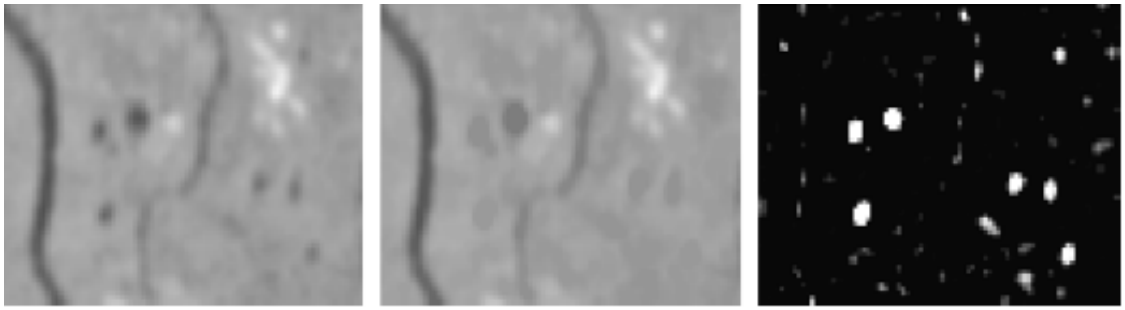

Figure 2.6: Detection of microaneurysm candidates. (Left) Preprocessed image. (Middle) Result of the diameter closing. (Right) Result of the corresponding top-hat transformation, which consists in the subtraction of the preprocessed image (left) to the result of diameter closing (middle) (Walter et al., 2007).

Sánchez et al. (2009), despite having used preprocessing techniques similar to those of the standard approach, selected an unsupervised mixture-model based clustering for the extraction of MA candidates. This technique determines a suitable threshold that separates the foreground regions from the background pixels according to its statistical intrinsic structure. Vasculature was also removed for reducing false detections. Several features based on texture, color and shape were then extracted and introduced in a trained logistic regression classifier.

- **Features and classification**

  The features, as well as the classifiers, used in the refinement stage are different from one author to another, and are also dependent on the type of images being processed. While the algorithms applied to fluorescein angiographies use essentially shape and intensity features, the algorithms applied to color photographs also include color features. With regard to the classifiers, there is also some diversity. For instance, Niemeijer et al. (2005) use a k-Nearest Neighbors (k-NN) classifier and Cree (2008) a Naïve-Bayes classifier.

2.2 Methodologies based on shape analysis with non-morphological operations

This category focuses on the shape of MAs and does not use morphological operations to distinguish MAs from other structures of the retinal images. Furthermore, it presents a larger diversity in comparison with the morphological-based approaches. While there are methods that use, for

instance, template-matching to detect MAs (Quellec et al., 2008; Zhang et al., 2010), other approaches apply operators that allow the extraction of circular objects, such as the double-ring filter (Mizutani et al., 2009). As in the mathematical-morphology based methodologies, in this type of approaches, the MA detection stage is also followed by a classification step for reducing the number of false detections.

In this section, the shape analysis-based methodologies were grouped according to the type of operator used for detecting MAs.

2.2.1 Gaussian matching

Since MAs can be modeled with 2D Gaussian functions, Quellec et al. (2008) implemented a method based on template-matching in the wavelet domain. This domain is particularly interesting, since, by removing sub-bands corresponding to high and low frequencies, noise and intensity variations can be also removed. Models with different standard deviations are used in order to match MAs with different sizes. After the subtraction of the coefficients of the wavelet transform of the image to the coefficients of the wavelet transform of the model, a threshold is applied in order to obtain MA locations. These authors also proposed a vessel segmentation algorithm based on wavelet analysis to remove candidates that lie on the vasculature.

Knowing that the correlation coefficient can give information about the similarity between two functions, Zhang et al. (2010) proposed a methodology based on multi-scale correlation filtering. It corresponds to the application of a sliding filter with Gaussian kernels (with different standard deviations) to the fundus image and then compute the correlation coefficient between the Gaussian function and the intensities distribution. Results at different scales are combined and a threshold is applied to obtain the candidates. The vasculature is removed using adaptive thresholding. Figure 2.7 shows the results of the several steps just described. In order to remove false detections, a set of 31 features are also extracted and a discrimination table is constructed with the minimum and the maximum values of each feature in true lesions. So, candidates with a value for a given feature that is out of the range are eliminated. More recently, the authors of this approach improved the classification stage with the incorporation of a dictionary learning with sparse representation classifier (Zhang et al., 2012).

2.2.2 Other circularity-based operators

Taking into account the round shape of MAs, Mizutani et al. (2009) proposed an approach that detects MAs using a double-ring filter. In order to remove the candidates corresponding to vessels, a different double-ring filter is also applied to segment the vasculature. After obtaining the MA locations, a region growing is performed and 12 features based on shape, color, contrast and intensity are extracted from each candidate. A final classification is obtained using an artificial neural network.

In a different way, but also considering the circularity of MAs, Giancardo et al. (2011) proposed a new technique for MA segmentation based on the Radon transform. The Radon based

features allow the detection of MAs in images with minimal processing (for instance, vessel segmentation is not necessary). The maximum dimension of each feature vector is 37. All feature vectors are then projected to a hyperplane of dimension equal to 10 through principal component analysis. The final results are obtained using a support vector machine classifier.

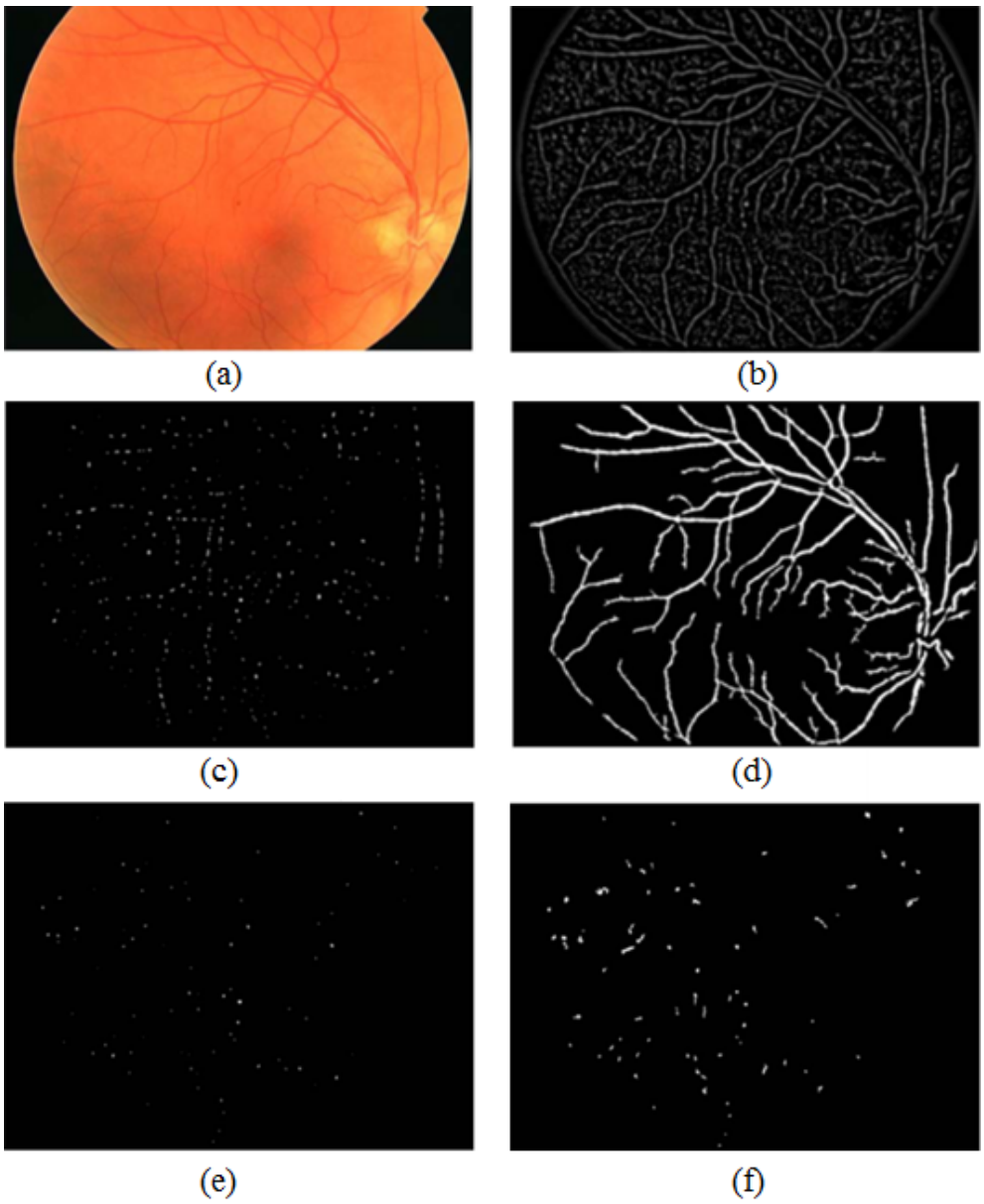

Figure 2.7: Results of the methods proposed by Zhang et al. (2010) for MA detection. (a) Original image. (b) Result of the multi-scale correlation filtering (σ = 1.1, 1.2, 1.3, 1.4 and 1.5). (c) Result of the thresholding applied to the image presented in b (threshold = 0.4). (d) Vascular map obtained using adaptive thresholding. (e) Detected candidates after blood vessel removal. (f) MA candidates after the application of a region growing algorithm.

2.3 Other methodologies

There are some methodologies which are novel and very different from the approaches presented before. These methods are described in this section.

Lazar and Hajdu (2013) proposed an approach that detects MAs through the analysis of directional cross-section profiles centered on the local maxima of image after preprocessing. For each profile, the peak is detected and a set of values that describe size, shape and height of the peak are computed. As illustrated in figure 2.8, different regions of the original image have different cross-section profiles. The features are then introduced in a naïve-Bayes classifier. At the end, a score is calculated for each candidate and a threshold is applied to obtain the final result.

On the other hand, Antal and Hajdu (2012) proposed an ensemble-based system for MA detection and DR diagnosis, as illustrated in figure 2.9.

In that system, the ensembles are sets of <preprocessing method, MA extractor> pairs. The selection of the optimal ensemble requires the evaluation of all pairs using the training images. So,

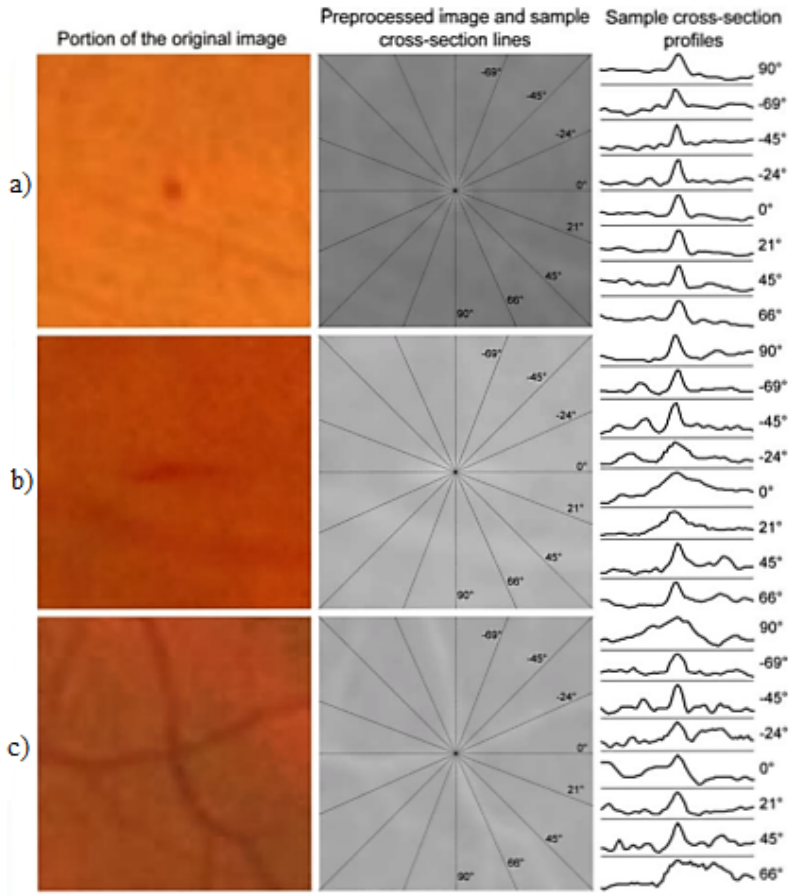

Figure 2.8: Sample cross-section profiles of three different portions of the image at several orientations: (a) Region with a MA; (b) Region with an elongated non-MA object, such as a blood vessel; (c) Region with a vessel crossing (Lazar and Hajdu, 2013).

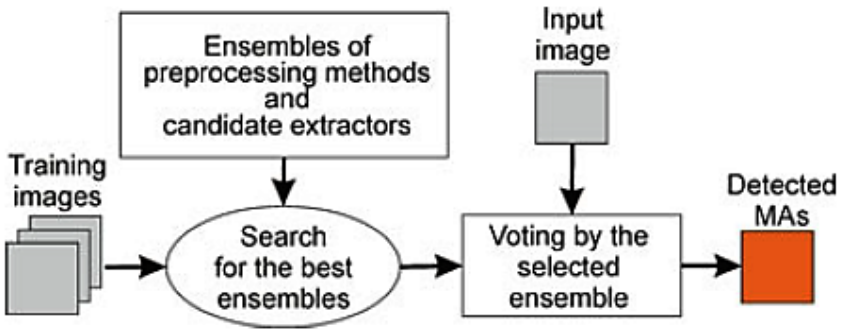

Figure 2.9: Flowchart of the ensemble-based system proposed by Antal and Hajdu (2012).

the use of a search algorithm in these situations can be useful. After obtaining the best ensemble, the corresponding pair <preprocessing method, MA extractor> is applied to the test image and the confidence values associated to each candidate are computed. In order to obtain the final candidates, a thresholding is applied.

More recently, the same authors proposed two improvements to the previous system: a context-aware selection and an adaptive weighting. After the categorization of the MAs according to their visibility and spatial location, each preprocessing method is applied to the training images and the results obtained (using the same MA extractor) are compared. Based on these results, a selection of the best performing preprocessing method for each category is done (context-aware selection). In order to obtain the MA candidates in the test images, the results obtained using the preprocessing methods previously selected are merged. The adaptive weighting is a technique that combines the output of different pairs by weighting. The weights associated with each candidate depend on the <preprocessing method, MA extractor> pair that detected the candidate, its visibility and spacial location. These weights are then used to compute the confidence value.

Pereira et al. (2014) also proposed an approach that uses a multi-agent system for detecting MAs in color photographs. In this methodology, different autonomous entities, the agents, are situated in the environment (preprocessed image) and interact with each other. The segmentation of the image is commanded by their behaviour. The agents can be classified as explore agents (EA) or region agents (RA). The first agents explore the image and launch a region agent when a region of interest is found, the other agents segment and analyse their region. An overview of this methodology is shown in figure 2.10.

2.4 Concluding remarks

The majority of the methodologies developed for MA detection can be divided into two main groups: mathematical morphology-based approaches and shape analysis-based approaches which do not use morphological operations. Recently, ensemble-based systems were also introduced in this context.

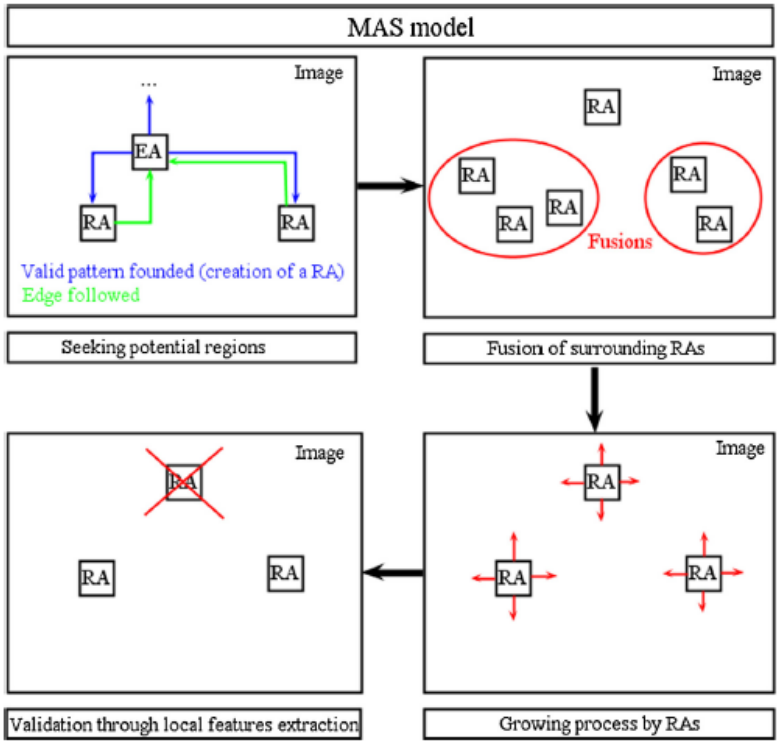

Figure 2.10: Overview of the multi-agent system proposed by Pereira et al. (2014).

Chapter 3

Methodology for MA detection

In this chapter, the methodology implemented for detecting microaneurysms in retinal photographs will be described in detail. Such methodology has as input the original retinal photography (RGB image) and as output a binary image where the regions classified as MAs are presented in white (value equal to 1) and the other regions in black (zero value). A scheme with the main steps of the methodology is presented in figure 3.1. The most relevant stages are:

1. Preprocessing, which includes the mask generation and the homogenization of the image background (shade correction).

2. Candidate detection, which comprises the enhancement of the potential MA regions and the removal of candidates lying on the vasculature. The output is a first indication of the candidates' location and has usually a high number of false positives (FP).

3. Candidate classification, which includes the measurement of local features at the detected locations and the classification of candidates as MAs or non-MAs.

3.1 Mask generation

In the color retinal photographs, the field of view (FOV) corresponds to the portion of retina that is visible through the camera. So, the pixels outside the FOV are black and have no information about retina (figure 3.2a). However, many of them have non-zero intensities. Since this can cause some interferences during the processing phase, it is important to detect those pixels and set their intensity to zero.

In order to obtain the mask of the FOV, a threshold is applied to the red component of the RGB image, since it presents the highest difference between the intensities of the pixels inside and outside the FOV (figure 3.2b). Although many authors use fixed thresholds, in this work the threshold is obtained through the analysis of the histogram of the red component. Similarly to the method described in Chong and Suniel (2015), the main steps performed for obtaining the optimal threshold are:

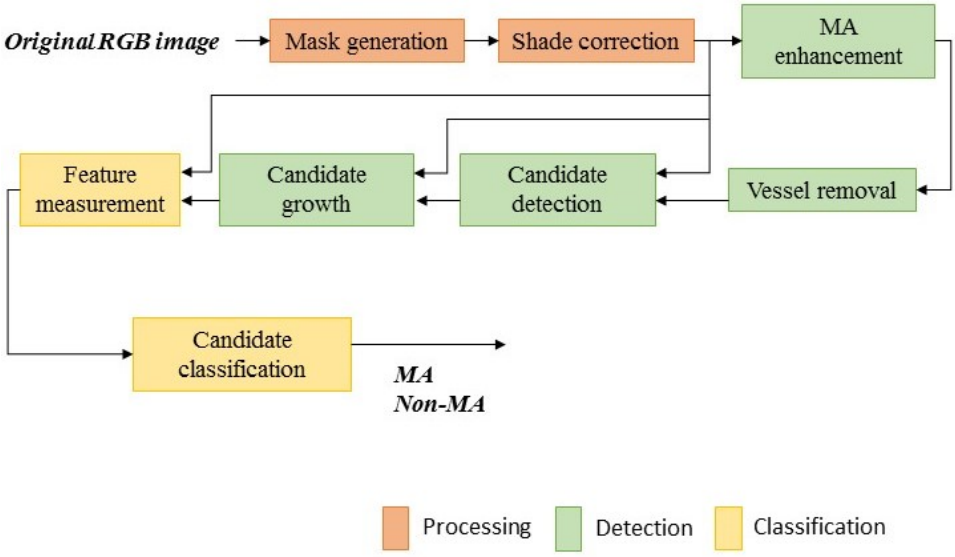

Figure 3.1: Scheme of the methodology implemented in this work.

1. Extract the red component of the original RGB image and calculate the histogram using just 32 bins;

2. Find the bin immediately before the first regional maximum of the histogram and convert it to the range [0 , 255];

3. Compute the sum of the values of the histogram (in this case, with 256 bins), considering a window with a size equal to 7 bins and centered on the bin obtained in the previous step plus 4 (in order to the first bin of the window be equal to the bin obtained in the previous step plus 1);

4. Shift the window one pixel to the right and repeat step 3. Stop when the window is centered on the bin 128, since only the peaks present in the first half of the histogram (lower intensities) are searched.

5. Find the window where the minimum sum is achieved. If there are several windows where this occurs, select the window centered on the highest bin. The threshold corresponds to the bin which is in the middle of that window;

6. Apply the threshold.

The binary image obtained by thresholding can have some holes inside the FOV and a rough boundary. In order to improve it, the highest object is selected and its holes are filled.

As some images also present a white contour around the FOV (figure 3.2a), which affects the result of the preprocessing stage, an erosion of the mask with a diamond structuring element with radius dependent on the contour's width is also performed.

The final result is an image in which the pixels outside the FOV and the pixels of white contour are assigned a zero value and the other pixels are assigned a value equal to 1 (figure 3.2d).

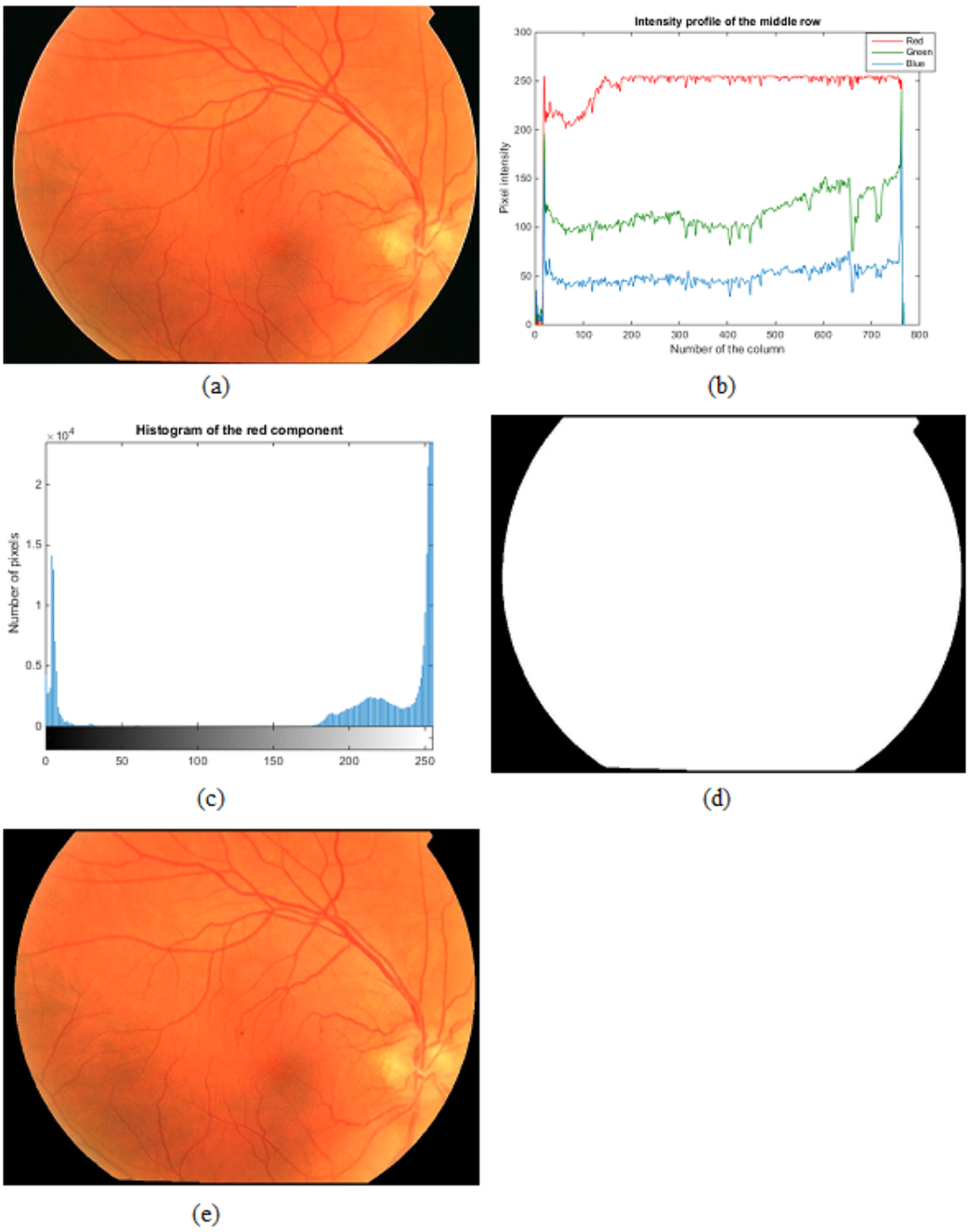

Figure 3.2: Mask generation. (a) Original RGB image. (b) Intensity profile of the middle row for the different RGB components. (c) Histogram of the red channel. (d) Final mask. (e) Result of the multiplication of the original RGB image by the mask.

3.2 Shade correction

Through the observation of the different bands of the original color image, it is possible to verify that the green channel is the one where the contrast between MAs and background is higher (figure 3.3d). Therefore, image preprocessing is performed on the green component of the original RGB image.

By observing the figure 3.3d, it is also possible to verify that there are some intensity variations on the background of the image. To overcome this problem, the image background is estimated and subtracted from the green component. For estimating the background, a median filter is applied to

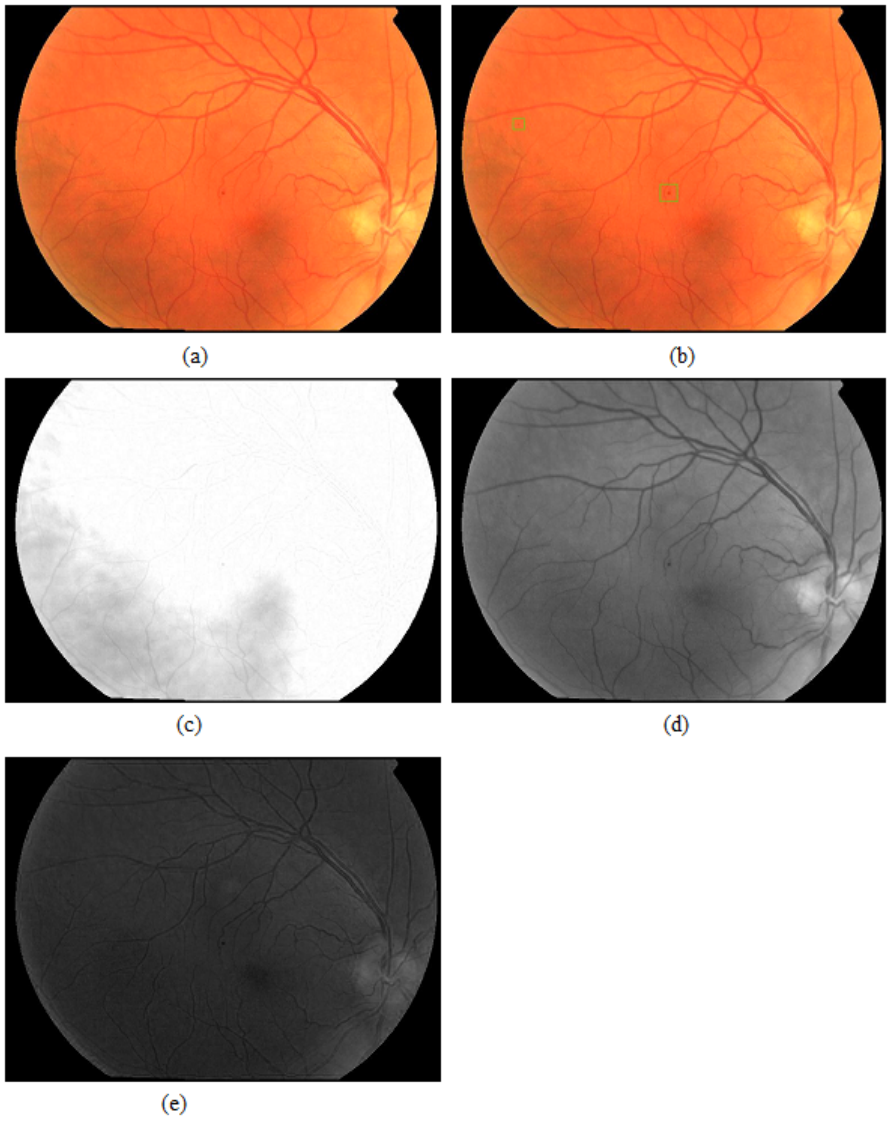

Figure 3.3: (a) Original RGB image. (b) Ground truth. (c-e) Red, green and blue channel of (a), respectively.

the green component, considering a n x n neighborhood. The value of n can be derived as follows:

$$n = n_s \times \frac{\text{diameter of the FOV}}{d_m} \tag{3.1}$$

where n_s is the ideal size of the neighborhood to obtain the image background of the smaller images and d_m the mean diameter of the FOV of those images.

The quotient between the diameter of the FOV of a given image and d_m is applied as scale factor. So, for images with different sizes, the number of pixels of the neighborhood considered to compute the median is also different.

Since the microaneurysms are always darker than the background, after the subtraction of the green component to the image background, all pixels with a non positive value are set to 0. Then, the intensities are scaled to the full range [0,1], giving rise to the shade corrected image (figure 3.4).

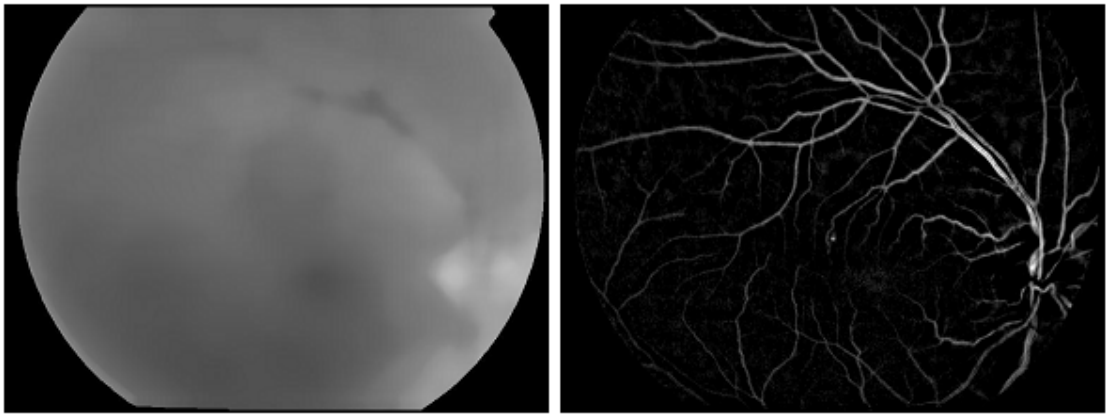

Figure 3.4: Shade correction. (Left) Image background corresponding to the green channel (figure 3.3d). (Right) Shade corrected image.

3.3 Microaneurysm enhancement

As microaneurysms (MAs) usually present a round shape, in order to enhance the regions that are potential candidates two different filters are used: a scale-normalized Laplacian of Gaussian (LoG) filter and a sliding band filter (SBF). In the following sections, these filters are described in detail.

3.3.1 Scale-normalized Laplacian of Gaussian filter

As mentioned in Esteves et al. (2012), the LoG filter is widely used as blob detector. Since the MAs appear as dark blobs in the green component of the original color image, the application of this filter in this image allows the enhancement of these lesions.

Mathematically, the filter can be defined as follows (Lindeberg, 1994):

$$LoG(x,y) = \nabla^2_{g\ normalized} = \frac{x^2+y^2-2\sigma^2}{2\pi\sigma^4}\ e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{3.2}$$

where g is a Gaussian kernel at scale σ^2.

Thus, for a given input image, the filter's response only depends on the scale of the Gaussian kernel (the single parameter of the filter). The LoG filter has strong positive responses for dark blobs with a radius equal to $\sqrt{2}\sigma$ and strong negative responses for bright blobs of the same size.

In order to enhance MAs with different radii, the green channel is filtered at several scales. The range of scales used only depends on the minimum (R_{min}) and maximum radius (R_{max}) of the lesions to be detected.

After applying the LoG filter at all scales in the range $[R_{min}/\sqrt{2}\ ,\ R_{max}/\sqrt{2}]$, all responses are combined, such that for each position (x,y) the maximum of all responses is considered (figure 3.5).

Then, in order to obtain the first candidates, only the regional maxima are considered (figure 3.6).

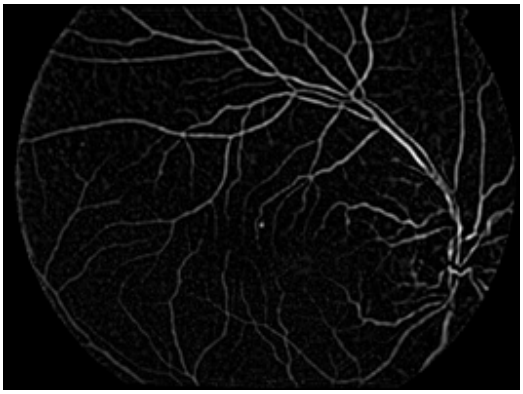

Figure 3.5: Result of the application of the LoG filter in the green component of the RGB image (figure 3.3d) at several scales in the range $[1/\sqrt{2}\ ,\ 5/\sqrt{2}]$.

3.3.2 Sliding band filter

In order to compare the result of different filters for MA enhancement, the SBF is also applied to the green component of the original color image. The choice of this filter to also enhance MAs is based on the following aspects:

1. As a local convergence filter, it is based on gradient convergence and therefore can handle noise and low contrast-related problems (characteristic of this type of images);

2. Some of its parameters are intuitive and can be set based only on visual inspection of lesions' shape (Esteves et al., 2012).

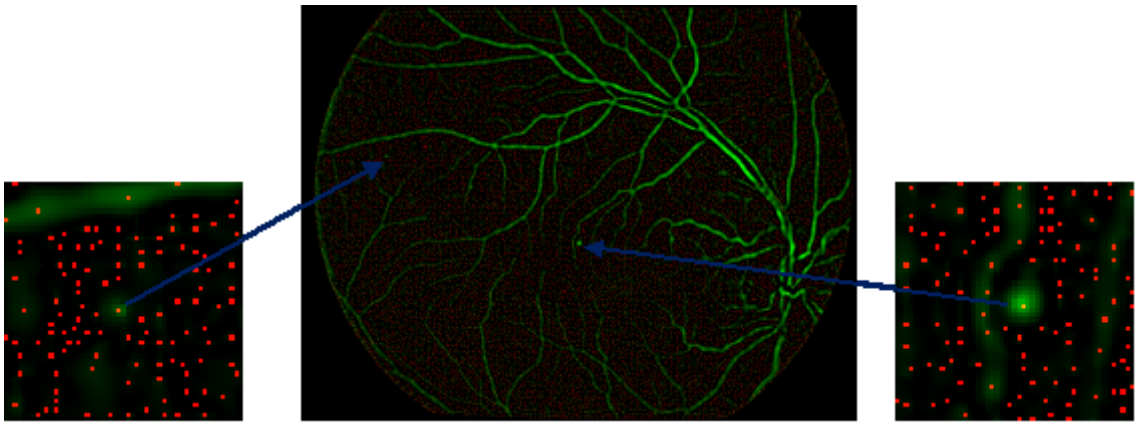

Figure 3.6: First candidate detection. The combination of the responses of all LoG filters applied is represented in the green channel and the regional maxima in the red channel. The images presented in the left and right side correspond to the regions where a microaneurysm is present (blue arrow).

The SBF combines the concept of two different convergence index filters: the iris filter, which adapts the radius of the support region in order to maximize the convergence in each direction independently, and the adaptive ring filter, that uses as support region a ring with a fixed width, which depends on the radius of maximum convergence (figure 3.7) (Esteves et al., 2012).

Thus, the SBF searches, for each radial direction i, the band of width d that corresponds to the maximum degree of convergence (c_i). This can be formulated as follows:

$$SBF(x,y) = \frac{1}{N}\sum_{i=0}^{N-1} c_{i_{max}} \tag{3.3}$$

$$c_{i_{max}} = \max_{R_{min}\leq n\leq R_{max}} \left(\frac{1}{d}\sum_{m=n}^{n+d} \cos\theta_{i,m} \right) \tag{3.4}$$

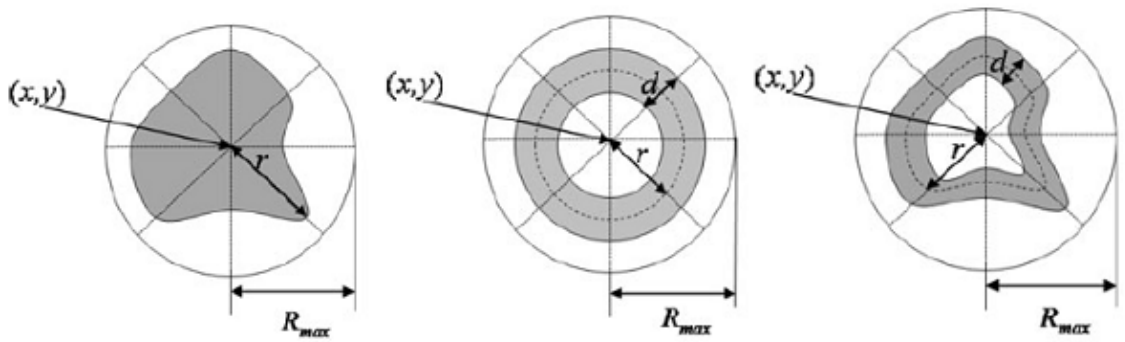

Figure 3.7: Convergence index filters with support regions represented in grey. (Left) Iris filter. (Middle) Adaptive ring filter. (Right) Sliding band filter (Esteves et al., 2012).

Based on its formulation, four parameters must be chosen for the application of the SBF:

1. minimum radius of the objects to be enhanced (R_{min});
2. maximum radius of the objects to be enhanced (R_{max});
3. number of radial directions for which convergence is evaluated (N);
4. width of the band in which the gradient is analyzed (d).

In addition to these parameters, the image where the filter is applied also affect the filter's response.

Similarly to the LoG filter, for obtaining the first candidates, only the regional maxima of the SBF response should be considered (figures 3.8 and 3.9).

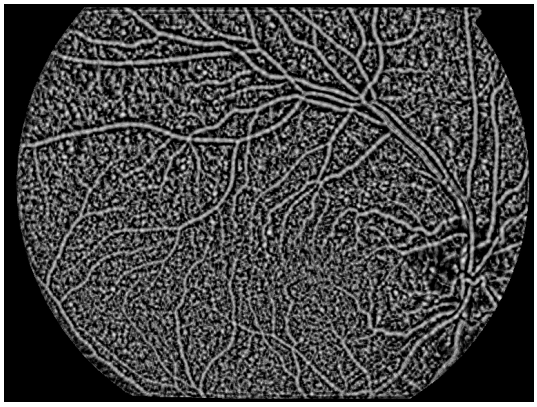

Figure 3.8: Result of the application of the SBF in the green component of the RGB image (figure 3.3d) ($R_{min} = 1$, $R_{max} = 5$, $N = 8$ and $d = 3$).

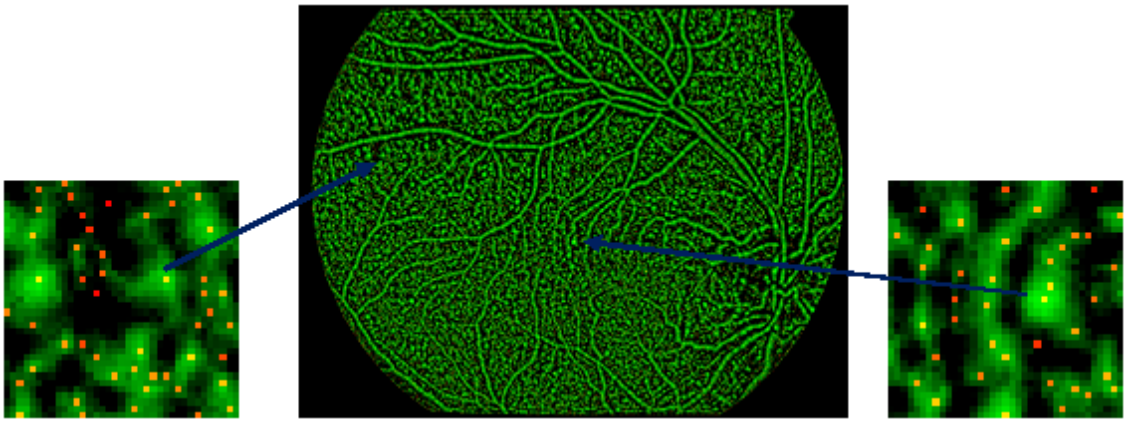

Figure 3.9: First candidate detection. The response of the SBF (figure 3.8) is represented in the green channel and the regional maxima in the red channel. The images presented in the left and right side correspond to the regions where a microaneurysm is present (blue arrow).

3.4 Vessel removal

A high number of candidates corresponds to pixels that belong to the vasculature. In order to remove those false positives, retinal vessels are segmented and the candidates lying on the vasculature are rejected.

For vessel segmentation the method described in Mendonça and Campilho (2006) was selected. An example of its application for obtaining the vasculature map is shown in figure 3.10 (left).

Since this method can lead to the segmentation of MAs which are near the vessels, the vasculature map should be post-processed. This post-processing phase includes the following steps:

1. Division of the vasculature map into segments, which includes the detection of the crossing points and the removal of those points after dilating them with a circular kernel of size equal to the highest diameter of the vessels;

2. Exclusion of the candidates with a size smaller than the mean size of the segments, since when a MA is segmented it usually gives rise to small segments.

The result of this step is illustrated in figure 3.10 (right).

After vessel segmentation, in order to obtain only the candidates that do not belong to the vasculature, the regional maxima of the LoG and SBF responses are multiplied separately by the negative of the vasculature map.

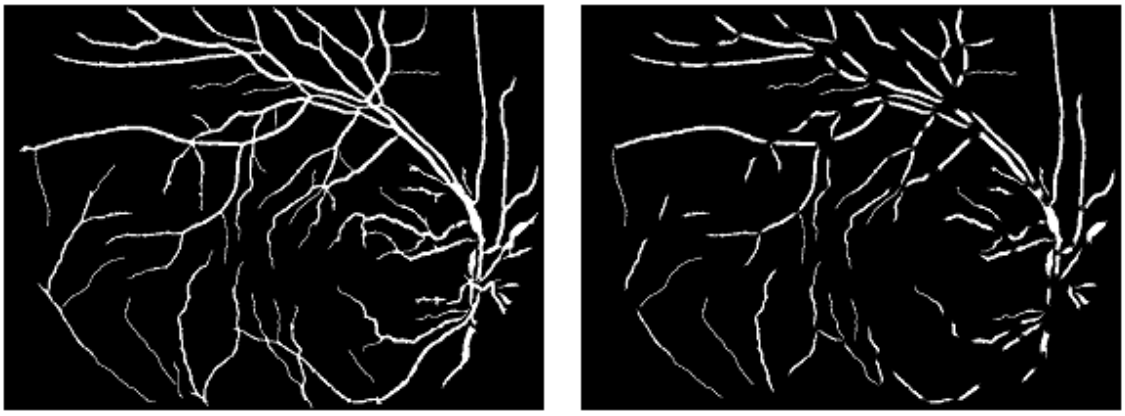

Figure 3.10: Vessel segmentation. (Left) Vasculature map obtained using the method described in Mendonça and Campilho (2006). (Right) Vasculature map after the post-processing step.

3.5 Candidate detection

Even after removing the candidates belonging to the vasculature, a high number of candidates remains in the result of both filters.

Since until now all regional maxima are being considered independently of the intensity of the filter response, a way of reducing the number of false detections is to maintain only the regional maxima where the filter response is higher than a threshold value. This threshold is calculated by averaging the filter response in the regional maxima locations.

Although there is a reduction of a significant number of candidates by applying the threshold in the filters' responses, a high number of false detections remains for both filters.

Taking into account that both filters have strong responses in the microaneurysms' location and that in the other regions (for example noisy regions) their responses may differ, the false positives can be reduced through the combination of the candidates that are detected by the two filters.

A direct intersection of the regional maxima of the LoG and SBF responses can not be performed, because in several lesions the filters give maximum responses in different pixels of the lesion, although normally very close. So, for maintaining the maximum number of lesions possible, the candidates of each filter are dilated with a square structuring element with 3 pixels of width and then intersected. An illustrative example is presented in figure 3.11.

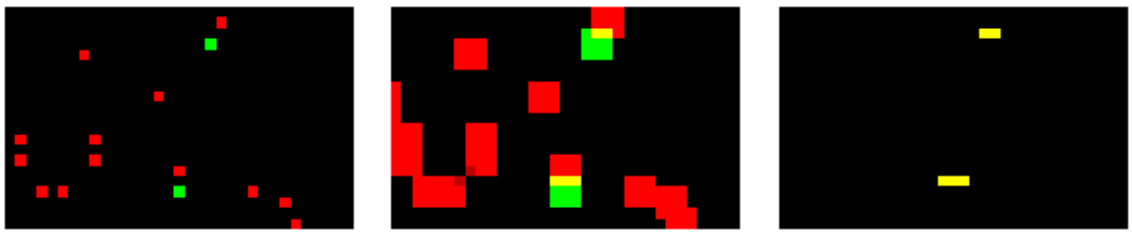

Figure 3.11: Illustrative example of the combination of the candidates obtained using 2 different filters: LoG and SBF. (Left) Insertion of the candidates of both filters in the same image. The candidates obtained using the LoG filter are represented in green and the candidates obtained through SBF application in red. (Middle) Result after dilating all candidates. (Right) Result of the intersection of the candidates of both filters after dilation. The locations where two candidates of different filters overlap are represented in yellow.

Since the LoG filter and the SBF only take into account the objects' shape but not their intensities, the strategy used for obtaining the final candidates is to apply a threshold based on the candidates' intensities.

An example of the final candidates obtained is presented in figure 3.12 (right).

3.6 Candidate growth

Since the candidates obtained are smaller than the true lesions, they need to be grown. For that, the region growing procedure described in Spencer et al. (1996) is applied. This algorithm comprises the following steps:

1. Reduction of each binary object (candidate) to a single point, which corresponds to the seed point. This point corresponds to the pixel of the object which presents the highest intensity in the inverted green channel.

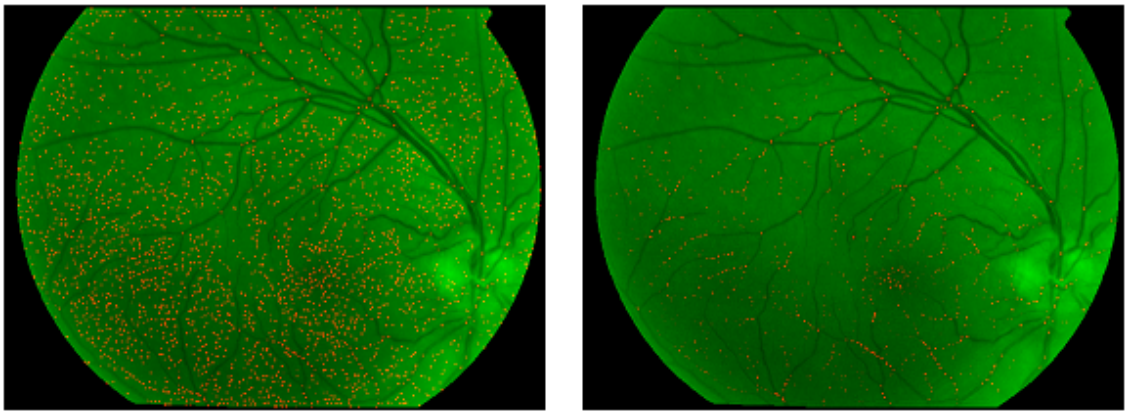

Figure 3.12: Final candidates. (Left) Candidates obtained after the combination of the candidates that come from the two different filters. (Right) Final candidates obtained after applying a threshold equal to 1.5 of the mean of all candidates in the image that results of the multiplication of the SBF response by the shade corrected image. The green channel of the image is represented in green and the candidates in red.

2. Aggregation of the neighbouring pixels with an intensity in the inverted green channel equal or higher than a threshold, which is defined as follows:

$$threshold = i_{peak} - x \cdot (i_{peak} - i_{bgnd}) \tag{3.5}$$

   where i_{peak} corresponds to the intensity of the inverted green channel in the seed point, i_{bgnd} is the intensity of the inverted image background in the position of the pixel to be attached to the grown object and x is a fraction between 0 and 1 which controls the maximum acceptable difference between the intensity of the seed point and the intensities of the neighbouring pixels (figure 3.13).

3. Repetition of step 2 until the intensity of the neighbouring pixels do not meet the inclusion criteria or when the area of the object reaches a predefined size (which will be referred as size limit).

This limitation of the size can reduce significantly the processing time and the overlapping of grown objects. Taking into account that the lesions' radius depends on the image resolution, the size limit applied should also be different for images with different resolutions. For that, the scale factor mentioned in the previous sections is also used here.

Since the seed points which are part of the background or the vessels give rise to large objects that overlap other objects with smaller dimensions (figure 3.14 (left)), which can correspond to real microaneurysms, for each set of overlapped objects only the object with the smallest number of pixels is maintained.

In this way, the final result of this step only comprises the isolated objects and the smallest object of each set of overlapped objects. An example is illustrated in figure 3.14 (right).

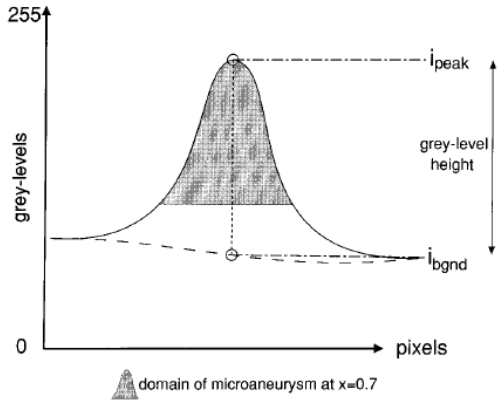

Figure 3.13: Typical intensity profile of a microaneurysm (Spencer et al., 1996). The grey area, which corresponds to the acceptable intensities, depends on the x parameter (equation 3.5).

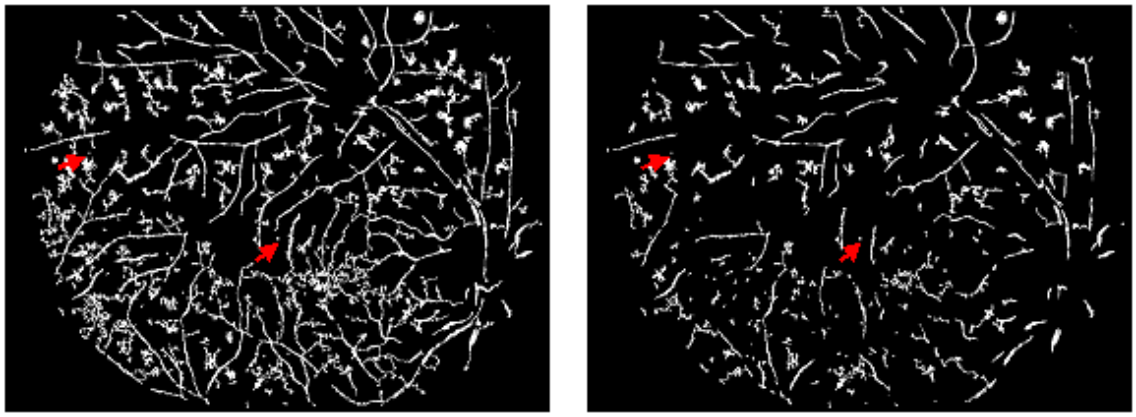

Figure 3.14: Candidate growth. (Left) All candidates after the region growing procedure. (Right) Candidates obtained after removing all overlaps. The microaneurysms are identified by the red arrows.

3.7 Feature measurement

Taking into account that the result of the detection phase presents a high number of false detections, it is necessary to extract from each candidate some relevant features which will be used for distinguishing the MAs from the non-MAs.

Since the microaneurysms present round shapes and a lot of false detections correspond to elongated structures (specially when their growth was guided through the vessels), several shape features must be computed (1-5 and 20-36). Furthermore, features based on the intensities (6-15), contrast and color (16-19) should also be computed, since the noisy regions that are also detected as microaneurysms present a lower contrast when compared with most MAs, as well as different intensities and colors.

In this work, the following features were used:

1. Area (a), which corresponds to the number of pixels of the candidate. Since the size limit defined in section 3.6 is higher than the maximum area of the lesions, the candidates with a number of pixels very close to the size limit correspond usually to candidates for which the seed point was set in the background or within a vessel and grew until the object reaches the size limit.

2. Perimeter (p), which was obtained using the Freeman chain codes (Freeman, 1961) of the object: $p = n_{even} + n_{odd} \times \sqrt{2}$, where n_{even} and n_{odd} are the number of even and odd chain codes, respectively.

3. Inverse of the aspect ratio (1/r), which corresponds to the ratio between the major axis length (λ_1) and the minor axis length (λ_2) of the ellipse that has the same normalized second central moments as the region: $\frac{1}{r} = \frac{\lambda_2}{\lambda_1}$. For round shapes, λ_1 is approximately equal to λ_2 and this feature presents a value very close to 1; otherwise, for elongated structures, λ_1 is much higher than λ_2 and the value of this feature tends to zero. The inverse was used in order to obtain values in the range]0 , 1].

4. Inverse of the circularity (1/c), which is dependent on the area and the perimeter of the object: $\frac{1}{c} = \frac{p^2}{4 \times \pi \times a}$. Since the squared perimeter of the round objects are approximately equal to $4 \times \pi \times area$, the value of this feature is approximately 1 in these situations. For elongated structures, this feature tends to have a value lesser than 1. The inverse was used in order to obtain values in the range]0 , 1].

5. Compactness (v), which is calculated as follows: $v = \sqrt{\frac{\sum_{j=1}^{n} |d_j - \bar{d}|}{n}}$, where d_j is the distance between the j^{th} boundary pixel and the centroid of the object, $\bar{d}$ the mean of all the distances from the centroid of the object to all edge points and n the number of edge points. For round shapes, the compactness is equal to zero, since the distance from the centroid to each edge pixel is the same and equal to the mean. For non-round shapes, the compactness has a value higher than zero.

6. Sum of the intensities under the object in the green component of the original RGB image (i_{green}), since this is the unprocessed image that presents an higher contrast between microaneurysms and background.

7. As item 6 but using the shade corrected image, which is the processed image that results from background subtraction.

8. Mean intensity under the object in the green component of the original RGB image (m_{green}): $m_{green} = \frac{i_{green}}{a}$.

9. As item 8 but using the shade corrected image.

10. The normalized intensity in the green component of the original RGB image: $NI_{green} = \frac{1}{\sigma}(i_{green} - \bar{x})$, where σ is the standard deviation and $\bar{x}$ the mean of the intensity of the image background within the object.

11. As item 10 but using the shade corrected image.

12. The normalized mean intensity in the green component of the original RGB image: $NM_{green} = \frac{1}{\sigma}(m_{green} - \bar{x})$, where σ is the standard deviation and $\bar{x}$ the mean of the intensity of the image background within the object.

13. As item 12 but using the shade corrected image.

14-15. The intensity of the region growing seed in the LoG and SBF outputs, respectively.

16. Difference between the mean pixel value inside the object and the mean value in a circular region centered on the object in the red channel of the original RGB image. The outer diameter of the circular region corresponds to twice the maximum distance between the centroid of the object and the edge pixels of the corresponding bounding box plus 3.

17. As item 16 but using the green channel of the original RGB image.

18. As item 16 but using the blue channel of the original RGB image.

19. As item 16 but using the hue image taken from the HSI color space. The hue image can be obtained using the following equations (Gonzalez and Woods, 2007):

$$f(n) = \begin{cases} 2\pi - \theta & \text{if } B > G \\ \theta & \text{if } B \leq G \end{cases} \tag{3.6}$$

$$\theta = \arccos\left(\frac{{}^{1}/_{2}((R-G)+(R-B))}{\sqrt{(R-G)^2+(R-B)(G-B)}}\right) \tag{3.7}$$

where R, G and B are the intensities in the red, green and blue components normalized to the range [0 , 1].

20-27. Mean of the LoG filter outputs under the object for values of σ in the range $[1/\sqrt{2}\,,\, 8/\sqrt{2}]$, since it is expected that the radii of the lesions to be detected vary between 1 and 8 pixels.

28-35. Standard deviation of the LoG filter outputs under the object for values of σ in the range $[1/\sqrt{2}\,,\, 8/\sqrt{2}]$.

36. The average output of the SBF under the object.

Most of the features here presented are derived from those that were applied in the work developed by Niemeijer et al. (2005).

3.8 Candidate classification

In order to classify the candidates as MAs or non-MAs, different supervised learning algorithms were applied: a classification tree, two discriminant analysis classifiers (linear and quadratic), a k-nearest neighbour classifier (k-NN) and a support vector machine (SVM) classifier.

While the classification trees and the discriminant analysis classifiers are commonly used regarding the ease of interpreting how they classify the data and the prediction speed, their accuracy tends to be lower than the accuracy of the k-NN and the SVM classifiers. In the following paragraphs, each classifier will be better described in order to understand the advantages and disadvantages of choosing one or another classifier.

The classification trees predict the response associated with an observation by following the decisions in the tree from the root node down to a leaf node, which contains the final response. Each step in a prediction involves checking the value of one feature. An example of a classification tree is shown in figure 3.15.

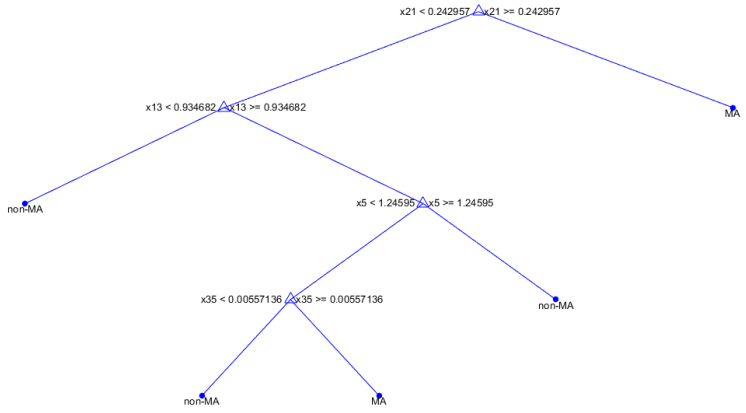

Figure 3.15: Example of a classification tree. The triangles represent the decision nodes.

The creation of this type of classifiers includes the following steps (Breiman et al., 1984):

1. Inspection of all possible binary splits on every predictor of the training set;

2. Selection of the split with the best optimization criterion: the minimum Gini's diversity index, which is a measure of the node impurity (a pure node is a node with just one class and a Gini's diversity index equal to zero);

3. Repeat steps 1 and 2 for the two child nodes until the node is pure or any split imposed on the node leads to a children with too few observations. The minimum number of leaf node observations should be the one that leads to a lower cross-validation error.

In a different way, the discriminant analysis classifiers assume that each class generates data based on different Gaussian distributions. While for linear discriminant analysis (LDA) classifiers the model has the same covariance matrix for each class, varying only the means; for quadratic discriminant analysis (QDA) classifiers, both the means and the covariances of each class vary (van der Heijden et al., 2004). In order to predict the class of a new observation, the trained classifier finds the class with the smallest misclassification cost. The space of the features is divided into regions considering the features and the labels of the training set. Those regions

are separated by straight lines for linear discriminant analysis, and by conic sections (ellipses, hyperbolas, or parabolas) for quadratic discriminant analysis, as can be seen in figure 3.16.

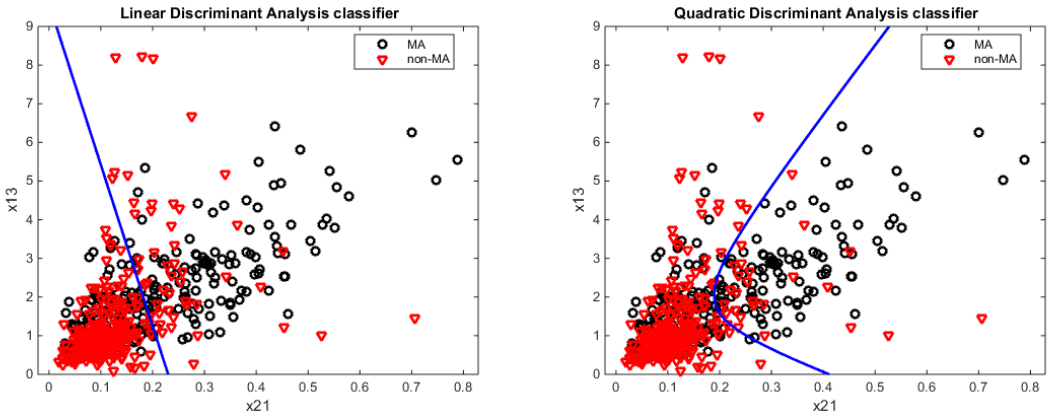

Figure 3.16: Example of a linear (left) and a quadratic (right) discriminant analysis classifier.

The k-NN classifiers are widely used as benchmark learning rules. They categorize query points based on their distance to points in the training set. For example, given a training set and a distance function (such as the Euclidean distance), the k-nearest neighbour classifier finds the k closest points in the training set to a query point and assigns to that point the class more frequent in those neighbours (Duda et al., 2004). An illustrative example is shown in figure 3.17. The k parameter of the k-NN classifier should correspond to the one that leads to a minimum cross-validation error.

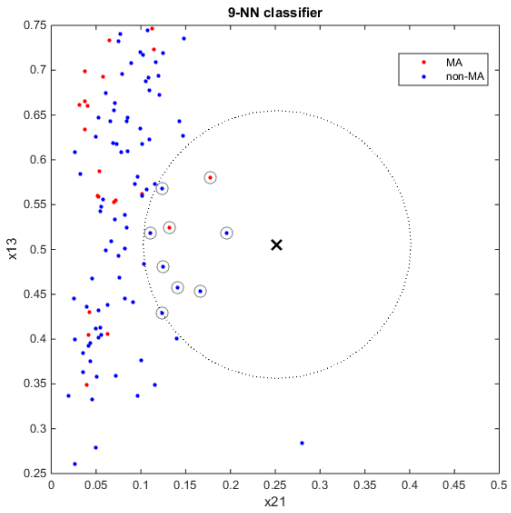

Figure 3.17: Example of a 9-NN classifier. The query point is represented by the X and the 9 points of the training set closest to the query point are marked with a circle. Using a rule based on the majority vote of the 9 nearest neighbours, the query point is classified as non-MA.

In comparison with the previous classifiers, the SVM classifier has a more complex formulation and it can be difficult to understand how it classifies the data. Furthermore, this type of classifiers can only be applied to problems with two classes (which is the case of this work). However, their prediction accuracy is usually higher than the other classifiers as mentioned above.

The SVM classifier classifies the data by finding the best hyperplane that separates all data points of one class from those of the other class, as illustrated in figure 3.18. The best hyperplane corresponds to the one with the largest margin between the two classes and the support vectors correspond to the data points that are closest to that hyperplane (van der Heijden et al., 2004).

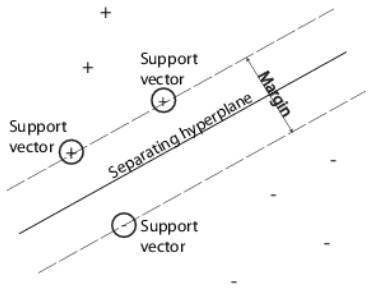

Figure 3.18: Illustation of how the SVM classifier classifies the data. + indicates the data points of type 1 and – the data points of type –1.

In the cases where the data can not be separable by an hyperplane, the SVM classifier can use a soft margin, that is, an hyperplane that separates most data points but not all.

In order to measure the performance of the different classifiers, several metrics can be used.

In this case, since, in the beginning of the classification step, the number of candidates that correspond to true microaneurysms is much lower than the number of candidates which are not microaneurysms, the receiver operating characteristic (ROC) curve is used.

The advantage of using this type of curves in these situations is that they summarize the space of possible trade-offs between sensitivity and specificity (equations 3.8 and 3.9, respectively).

$$Sensitivity = \frac{\text{Number of MAs which are correctly detected}}{\text{Total number of MAs}} \tag{3.8}$$

$$Specificity = \frac{\text{Number of candidates which are correctly identified as non-MAs}}{\text{Total number of candidates which are not MAs}} \tag{3.9}$$

For constructing these curves, several thresholds must be applied to the scores returned by the classifiers. For each threshold, the sensitivity and specificity are measured.

The area under the ROC curve (AUC) is commonly used as a measure of the detection system performance. The higher the AUC, the greater the average probability of the system making a correct decision.

3.9 Concluding remarks

The methodology developed and presented in this chapter is divided into blocks very similar to those of the standard approach developed by Spencer et al. (1996) and presented in figure 2.2.

However, the methods implemented in the candidate detection phase as well as the features and classifiers used in the classification step are different from the methods already described in the literature.

The uniqueness of this methodology is mainly related to the filters that were used for MA enhancement and whose responses were also used as features in the classification step. Although they were already used for detecting other round objects, they were never applied to this context.

Chapter 4

Results and discussion

This chapter begins with the description of the dataset used in this work. Then, the results obtained in each phase are presented and discussed, with a reference to the main problems found and the solutions proposed to overcome them.

4.1 Dataset

Although there are some datasets publicly available with retinal photographs, as well as the annotations of their respective retinal lesions, the Retinopathy Online Challenge dataset is the one that is most commonly used by the teams that developed methodologies for MA detection. In order to compare the performance of those methodologies with the method herein presented, this dataset was also used in this work.

As can be seen in table 4.1, the dataset is relatively heterogeneous, since the images were obtained with different cameras. Figure 4.1 shows an example of an image of each type.

Table 4.1: Different types of images in the Retinopathy Online Challenge dataset.

Type	Image dimensions (height x width *in pixels*)	Coverage of the retina	Number of images in the training set	Number of images in the test set
I	576 x 768	45º	22	22
II	1061 x 1059	45º	3	6
III	1388 x 1389	45º	25	22

The dataset has 100 images that are equally divided into two sets (training and test set). All images of the dataset were examined by 4 ophthalmologists, in order to obtain the locations of the lesions in each image (Niemeijer et al., 2010). However, only the annotations of the training images are publicly available on the website of the challenge (University and the ROC organizers, 2015). These annotations are in a XML file. Using an annotator reader, the annotations that are contained in the file *annotations-consensus-ma-only.xml* can also be saved in text files. Each annotation indicates the coordinates of the center of the lesion and its radius.

Figure 4.2 shows an example of an original image and the corresponding ground truth.

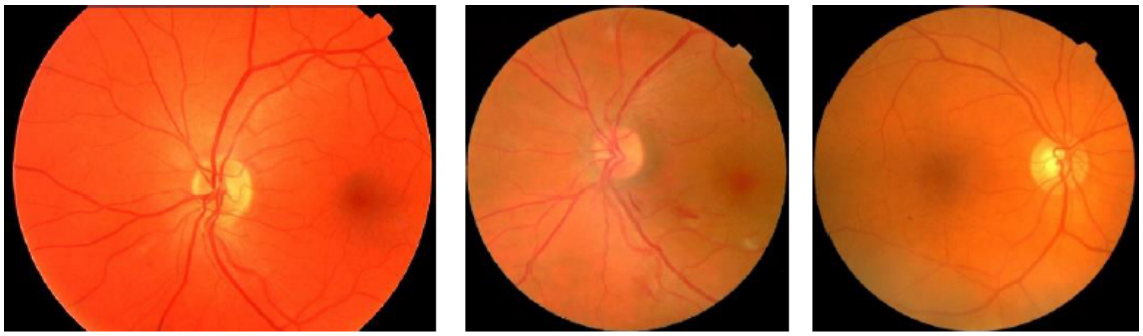

Figure 4.1: An example of an image of type I, II and III, respectively. Images were scaled for display.

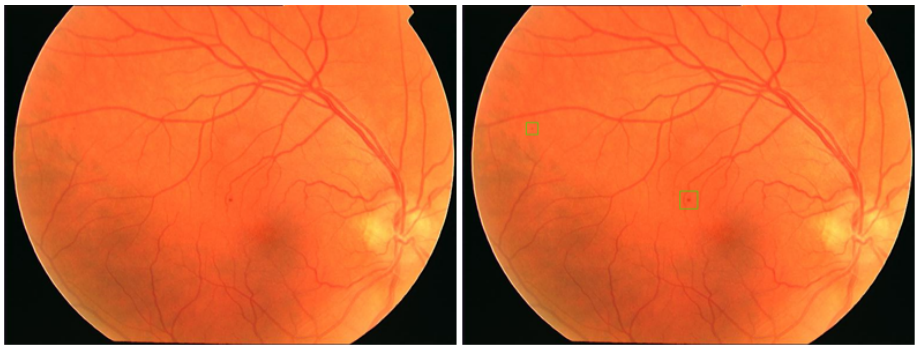

Figure 4.2: Example of an original image (left) and the corresponding ground truth (right). The lesions are marked with green squares.

4.2 Retinopathy Online Challenge evaluation algorithm

In order to make the comparison of the results obtained by different approaches possible, the Retinopathy Online Challenge authors also proposed an evaluation algorithm, described in detail in Niemeijer et al. (2010), that has a set of pairs (sensitivity, average number of false positives per image) as output. Those pairs are then used to obtain the free-response receiver operating characteristic (FROC) curve. This curve is typically plotted with the sensitivity on the vertical axis and the average number of false positives per image on the horizontal axis. A final score can be obtained by averaging the sensitivities at specific values of the average number of false positives per image: (1/8), (1/4), (1/2), 1, 2 4 and 8 (Niemeijer et al., 2010).

4.3 Results

In this section, the results of all intermediate steps are presented and discussed. In order to facilitate the interpretation of those results, this section is divided as chapter 3.

4.3.1 Mask generation

Although a high number of authors uses fixed thresholds for obtaining the masks of the field of view (FOV), in this work a threshold dependent on the histogram of the red component of the original image was used.

Table 4.2 shows the minimum and maximum values, as well as the mean and standard deviation of the thresholds applied to the red component of the 50 training images. By analysing those values, it is possible to see that the difference between the minimum and the maximum thresholds applied is high and that the mean is approximately in the middle of that range. This shows the importance of using adaptive thresholds instead of fixed thresholds empirically chosen.

Table 4.2: Minimum value, maximum value, mean and standard deviation of two variables involved in the generation of the masks of the FOV of the 50 training images: the threshold applied to the red component and the radius of the diamond kernel used for removing the white contour around the FOV.

	Minimum value	Maximum value	Mean	Standard deviation
Threshold applied to the red component for obtaining the mask of the FOV	0.0980	0.5020	0.3309	0.1046
Radius of the diamond kernel used for removing the white contour around the FOV (*in pixels*)	2	4	2.5	0.5

Since there are images where the "width" of the white contour is larger than others and some small lesions are located in the peripheries of the FOV, the radius of the diamond kernel used for removing that contour was also adapted for each image (table 4.2). This was achieved through the identification of the positions where the variation of the derivative is maximum and minimum in the blue component of the original image, which is the one where the contour is more highlighted. The contour's width corresponds to the distance between these two positions. This approach avoids the mask to be eroded in excess and some lesions to be lost *a posteriori*.

Although the methodology here described for mask generation is attractive due to the use of adaptive parameters, its performance can not be evaluated quantitatively because, for the dataset used in this work, the masks corresponding to the retinal photographs are not given. However, by visual inspection of the results (figure 4.3), it is possible to observe that the generated masks are very similar to the FOV.

4.3.2 Shade correction

Since the retinal photographs often present variations in the background of the FOV which can deteriorate the performance of the candidate detection algorithms, a shade correction was performed by estimating the image background through a median filter.

Figure 4.4 illustrates the effect of using the same median filter to obtain the image background for images with different resolutions. As can be seen in that figure, the 45×45 median filter is sufficient to obtain the background corresponding to the images of smaller dimensions but it is not

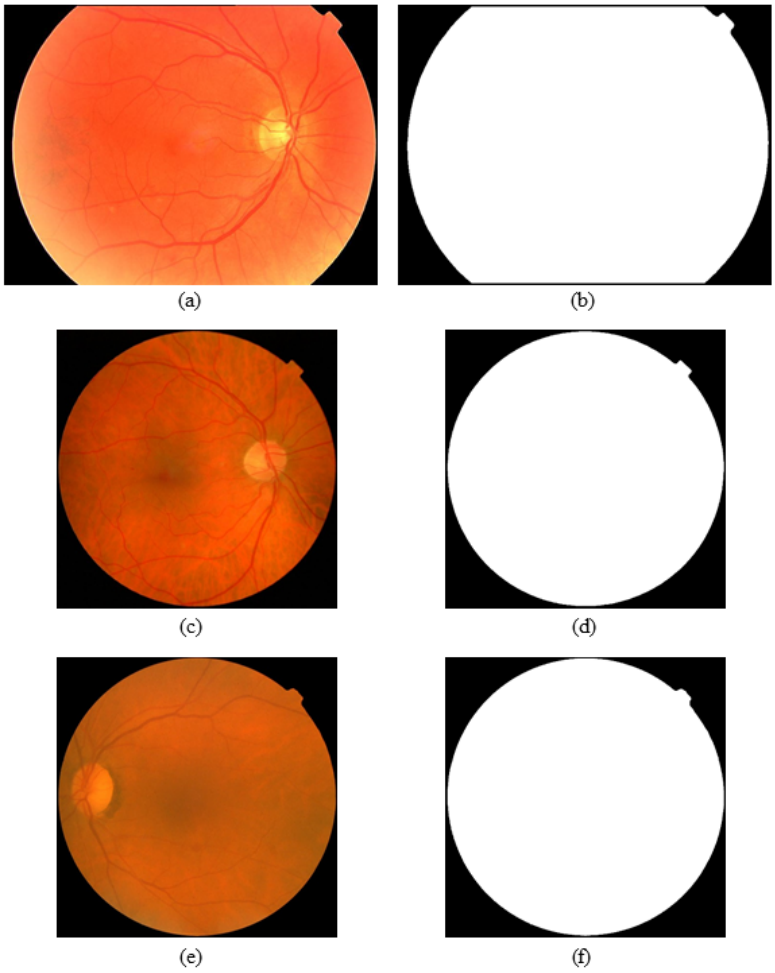

Figure 4.3: Masks of the FOV of three images with different resolutions. b is the mask corresponding to image a (type I image), d the mask corresponding to c (type II image) and f the mask corresponding to e (type III image).

sufficient to obtain the background for the larger images (since the vessels are not discriminated from the background).

Table 4.3 shows the scale factors that were used for each type of images and the size of neighborhood to be considered by the median filter after applying those scale factors.

By comparing the figures 4.4 (d) and 4.5 (bottom-right), which correspond to the same original

Table 4.3: Mean diameter of the FOV, scale factor and the size of the neighbourhood where the median filter is applied for each type of images.

Image type	**Mean diameter of the FOV (*in pixels*)**	**Scale factor**	**Size of the neighbourhood where the median filter is applied (*in pixels*)**
I	742	1.00	45 x 45
II	1044	1.41	63 x 63
III	1365	1.84	83 x 83

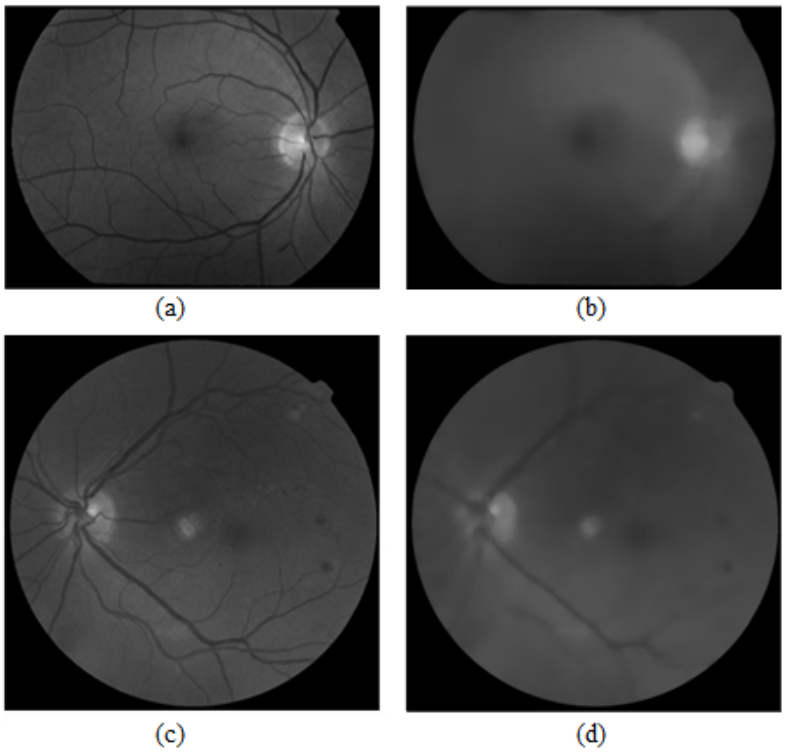

Figure 4.4: Effect of using the same size of neighborhood (45×45 pixels) to obtain the image background of different types of images. b is the image background of a (type I image) and d is the image background of c (type III image).

image, it can be concluded that when different neighbourhoods are used to obtain the background of images with different resolutions, better estimations of the background are obtained.

Thus, all filters must be scaled before their application, based on the resolution of the image where they will be applied.

4.3.3 Microaneurysm enhancement

In order to enhance the regions that were potential candidates, two filters were applied to the green component of the original image: a LoG filter and a SBF.

Regarding to the LoG filter, it was applied at several scales for enhancing MAs with different radii. The range of scales to be used was chosen based on the minimum (R_{min}) and maximum radius (R_{max}) of the lesions to be detected.

Since in the ground truth the lesions presented a radius higher than their real radius, a LoG operator was also applied at several scales to obtain their real radius.

For that, the green channel was filtered with several LoG filters, whose standard deviations varied between $1/\sqrt{2}$ and $10/\sqrt{2}$ (10 corresponds to the maximum radius of the ground truth). Then, for each lesion, it was found the standard deviation for which the filter response in the

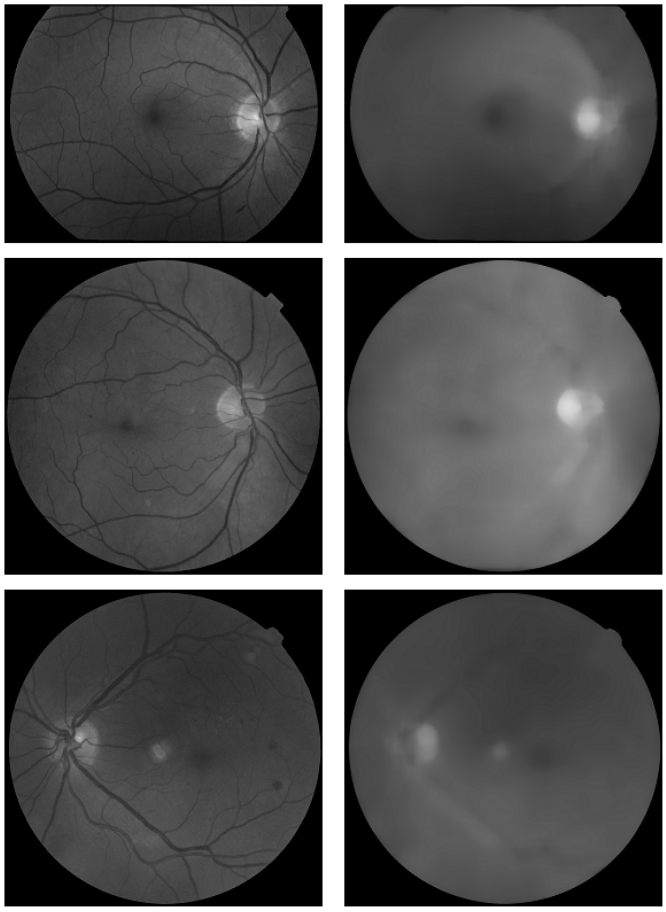

Figure 4.5: Estimates of the background obtained using different neighborhoods for images with different resolutions. For each image in the column of the left side, the corresponding background is shown in the column of the right side. The first row corresponds to a type I, the second to a type II and the last one to a type III image.

centroid of that lesion was maximum. Knowing this value, it was possible to obtain the real radius of that lesion (figure 4.6).

Taking into account that the lesions' radius depends on the image dimensions and considering the values presented in the table 4.4, the value of R_{min} was set to 1 for all types of images, while the value of R_{max} was adapted to the resolution of the image that was being processed. For that, the same scale factors that were used for background estimation (table 4.3) were also applied here. In this case, the value of reference was 5 for the smaller images.

Regarding to the SBF, the values of R_{min} and R_{max} used in this filter corresponded to those used in the LoG filter. Besides these two parameters, the number of radial directions for which the convergence is evaluated (N) and the width of the band in which the gradient is analyzed (d) had to be chosen.

After applying to the green component several filters resulting from different combinations of these two parameters, it was verified that the best result was achieved when $N = 8$ and $d = 3$.

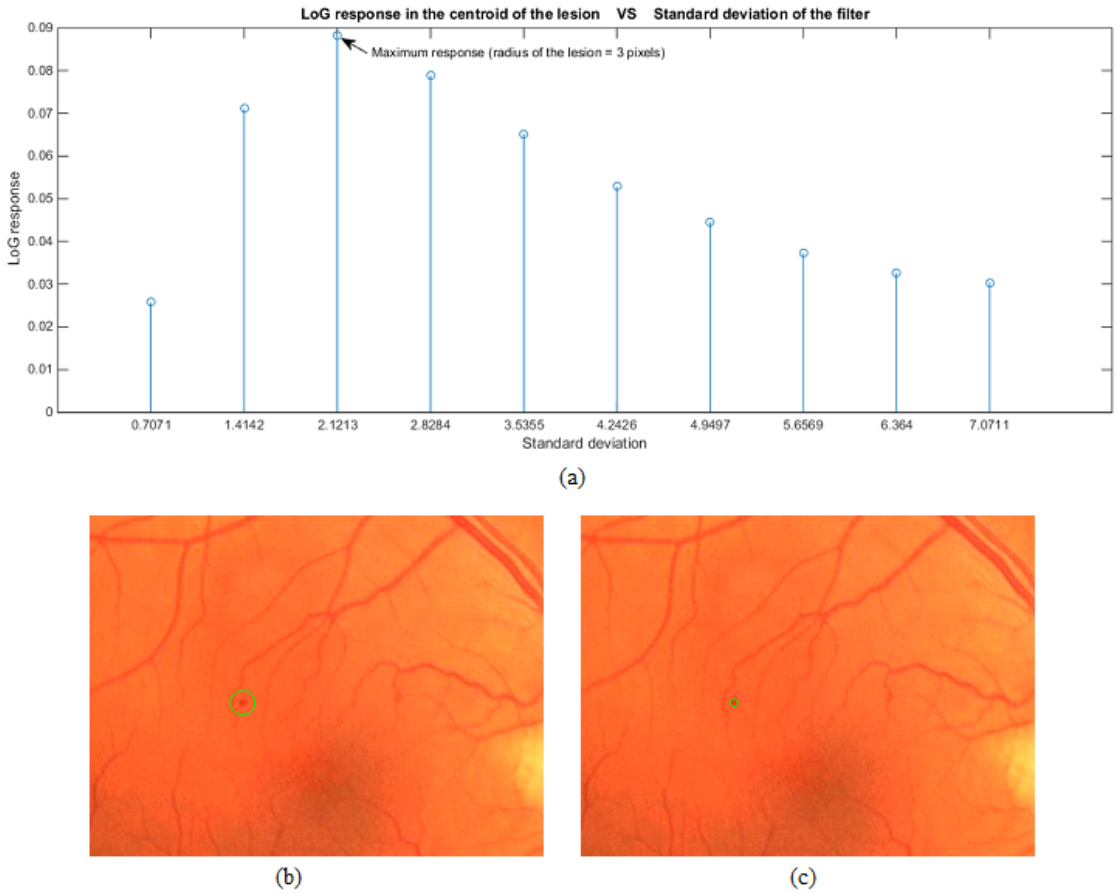

Figure 4.6: Determination of the true radius of the lesions. (a) LoG responses in the centroid of the lesion as a function of the standard deviation of the filter. (b) Ground truth, where the lesion is represented by a green circle with radius higher than the real radius of the lesion. (c) Ground truth, where the lesion is represented by a green circle with a radius equal to the real radius of the lesion (in this case, 3 pixels).

Table 4.4: Minimum, maximum and mean radius of the lesions presented in the ground truth.

Image dimensions (height x width *in pixels*)	**Minimum radius (in *pixels*)**	**Maximum radius (in *pixels*)**	**Mean radius (in *pixels*)**
576 x 768	1	5	2.91
1061 x 1059	1	5	3.51
1388 x 1389	1	8	3.75

Figure 4.7 shows the sensitivity and the average number of false positives per image when all regional maxima of the LoG and SBF responses were considered.

As shown in figure 4.7, all MAs were detected when all regional maxima of the filters' response were considered (sensitivity = 100%). Comparing the two filters, for the same sensitivity, the SBF gave rise to, approximately, more 9% false detections per image than the LoG filter (most of them correspond to blood vessels). Through these results, it is possible to say that these filters

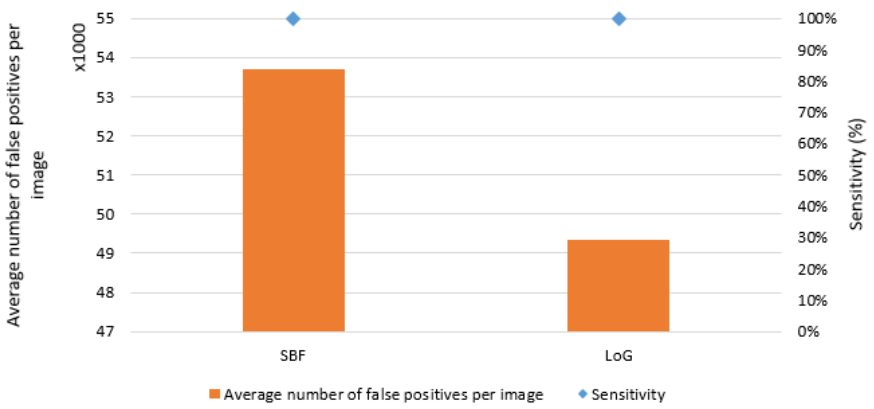

Figure 4.7: Sensitivity and average number of false positives per image obtained when all regional maxima of the LoG and SBF responses are considered.

enhance well the microaneurysms but also another structures, such as the blood vessels and noisy regions.

4.3.4 Vessel removal

In order to eliminate the candidates corresponding to blood vessels, a removal of the vasculature was also performed.

However, in some cases, the segmentation of the vessels failed and some regions corresponding to microaneurysms were also segmented, as shown in figure 4.8.

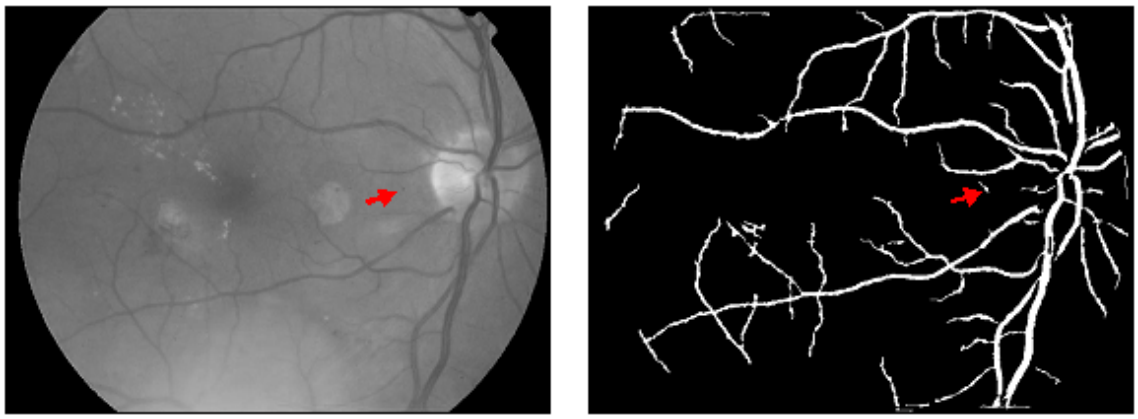

Figure 4.8: Example of a microaneurysm which is included in the vasculature map (right). The red arrow indicates the location of that microaneurysm.

In order to not lose lesions due to errors in the segmentation of the vasculature, a threshold based on size was applied to all segments of the vasculature map, as explained in the section 3.4.

The results obtained after vessel removal are presented in figure 4.9. Through the analysis of that figure, it is possible to see that:

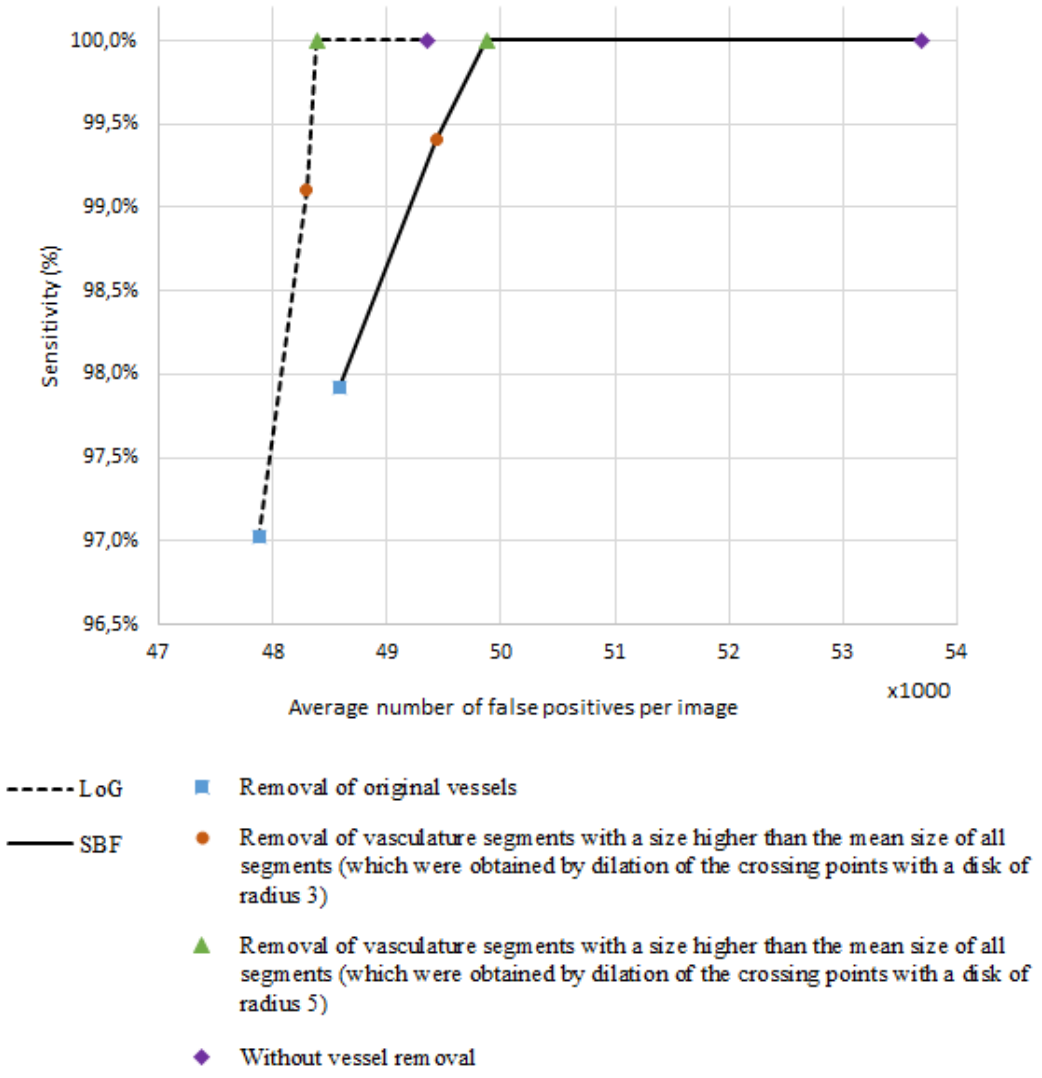

Figure 4.9: FROC curves obtained when the regional maxima which lie on the vessels are removed using different vasculature maps.

1. The SBF presents more regional maxima inside the vessels than the LoG filter.

2. The best result was achieved when a post-processed vasculature map was used for removing the regional maxima which lie on the vasculature. That vasculature map comprises the segments (which result from the dilation of the crossing points with a disk of radius 5) that have a size higher than the mean size of all segments.

3. Although this step led to a significant reduction of the average number of false positives per image, this value is still very high.

4.3.5 Candidate detection

After removing the candidates which lie on the vessels, a threshold based on the filters' response was applied to the regional maxima of each filter, such that only the candidates with a filter response higher than the mean of all candidates remained. The results obtained with this procedure are shown in figure 4.10.

As can be seen in figure 4.10, this step led to the loss of a few number of lesions. Most of the lesions which were lost in this phase are very difficult to observe in the original image (figure 4.11b). Since those lesions can not be well distinguished from the background, the LoG and SBF responses also present low values in these regions.

When the candidates of both filters were joined, it was verified that there was a lot of regions where both filters detected candidates, usually very close, but there were also some regions where

Figure 4.10: Sensitivity and average number of false positives per image obtained when only the regional maxima with a filter response higher than the mean of all candidates are considered.

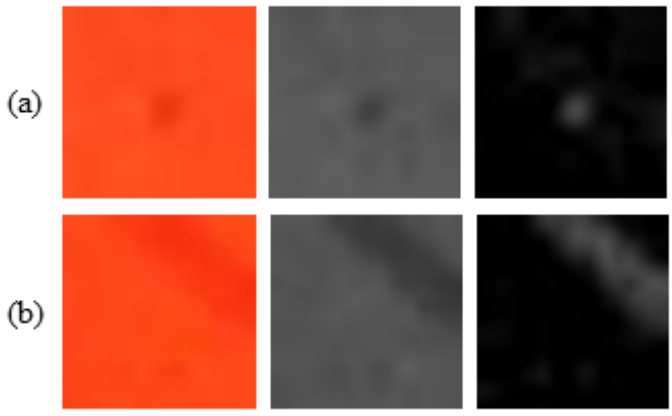

Figure 4.11: Example of a region where there is a visible lesion (a) and a subtle lesion (b). In each column, that region is shown for different images: RGB image, green channel and the shade corrected image, respectively.

the candidates were detected by just one filter. This may be associated with the sensitivity of the filter to the image noise.

As shown in figure 4.12, a simple intersection of both candidates could not be performed, because this would lead to a high reduction in the number of MAs which would be correctly detected. This happened since for most lesions the SBF and LoG filter gave a maximum response in different pixels of the lesion, although normally very close.

In this way, the results obtained were much better when a dilation of the candidates was performed before the intersection. Although there was a small decrease in the value of the sensitivity (that is related to the MAs which were only detected by one of the filters), the average number of false positives per image was reduced to approximately 50%.

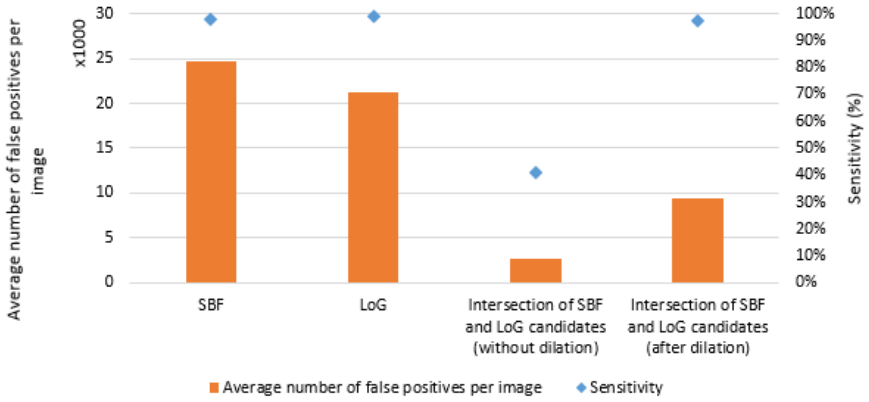

Figure 4.12: Sensitivity and average number of false positives per image obtained when the candidates that come from each filter are combined.

In order to study the effect of applying a threshold based on the candidates' intensities (to the candidates obtained after intersection) in the value of the sensitivity and the average number of false positives per image, 8 FROC curves were obtained through the application of several thresholds in 8 different images (figure 4.13):

1. Shade corrected image;
2. Green component of the RGB image;
3. SBF response multiplied by the shade corrected image;
4. SBF response multiplied by the green component;
5. LoG response multiplied by the shade corrected image;
6. LoG response multiplied by the green component;
7. SBF response multiplied by the LoG response and the shade corrected image;
8. SBF response multiplied by the LoG response and the green component.

By analysing figure 4.13, it is possible to see that the best result was obtained when a threshold equal to $1.5 \times$ *mean of all candidates* was applied to the image containing the result of the multiplication of the SBF response by the shade corrected image. This threshold corresponds to the point where a better trade-off between the sensitivity and the average number of false positives was achieved.

Although the number of false detections is still high, approximately 4000 false positives per image were removed with this thresholding operation.

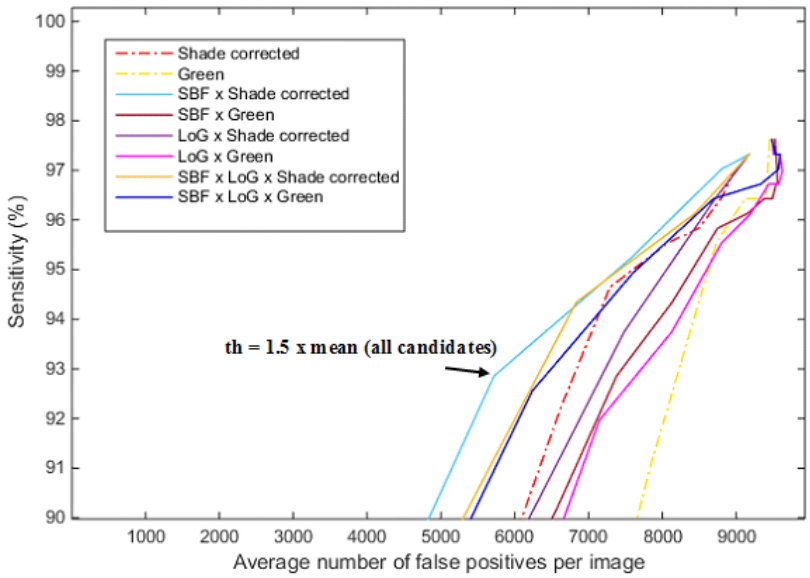

Figure 4.13: FROC curves obtained when several thresholds are applied to 8 different images in order to obtain the final candidates. A zoom in the range of sensitivities between 90% and 100% was done for display.

4.3.6 Candidate growth

Since the candidates obtained in the candidate detection step had a size smaller than the real microaneurysms, they were grown using the procedure described in section 3.6.

Although in the works developed by Spencer et al. (1996) and Niemeijer et al. (2005) the parameter x of the region growing algorithm (see section 3.6) was set to 0.5, this value led to an over-segmentation of the MAs which were located near the vessels (figure 4.14 (middle)), giving rise to large objects. Thus, in order to obtain a better segmentation of the MAs, the value of that parameter was set to 0.25 in this work (figure 4.14 (right)).

In order to reduce the number of overlaps, as well as the processing time, a limit to the size of the grown objects was applied. Since the lesions' radius varies with the image resolution, different limits were applied to different images. For that, the scale factor used in the background estimation step was also used here. As reference, for the smaller images the maximum acceptable area was

200 pixels, considering that the area of a square with a width equal to 11 (diameter of the largest lesions in this type of images) is 121 pixels.

The time required for the growth of all candidates of an image was on average 168 seconds, using the MATLAB R2014b and a computer with an Intel Core i7-4710HQ and 12 GB RAM.

After growing the candidates, it was noticed that the result presented several overlapping candidates. This often occurs when a seed point is placed in a normal region of the retina, due to the noise for example, leading the corresponding object to grow into neighbouring vessels and lesions. The task of removing all overlaps is very time consuming, since all intersections between objects are inspected. The results obtained after the removal of all overlaps are presented in figure 4.15.

By analysing this figure, it is possible to conclude that approximately a third of all candidates were removed in this step due to the fact that they gave rise to large objects that overlapped other smaller candidates.

This procedure also led to the decrease of the sensitivity. This effect is related to the fact that, for the MAs which present a low contrast with the background, the region growing grows until the

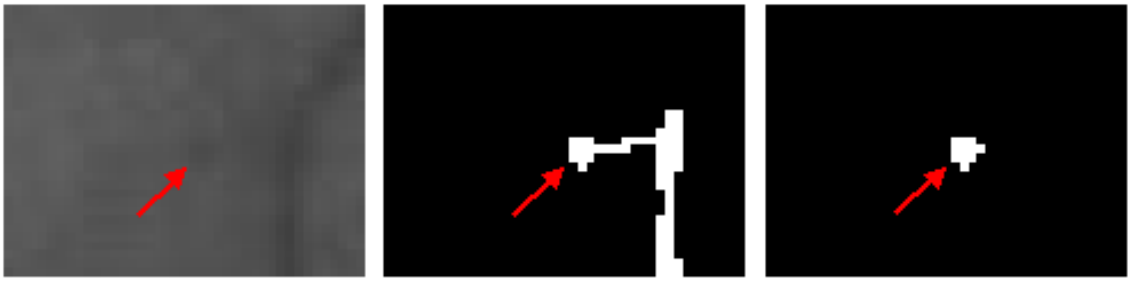

Figure 4.14: Influence of the x parameter of the region growing in the growth of the candidates which are located near the vessels. (Left) Region where a microaneurysm is present (green channel image). (Middle) Result of the region growing procedure for $x = 0.5$. (Right) Result of the region growing procedure for $x = 0.25$. The microaneurysm is marked with the red arrow.

Figure 4.15: Sensitivity and average number of false positives per image obtained before and after the candidate growth.

size limit is reached, giving rise to large objects.

However, more than 80% of all microaneurysms were preserved, which is good, taking into account the low quality of the images and the fact that there are microaneurysms which are marked in the ground truth but are very subtle.

4.3.7 Feature measurement and candidate classification

In order to distinguish the candidates that correspond to true microaneurysms, 36 features were extracted from each final candidate.

Since each image has on average 2000 false positives, the feature measurement task is also time consuming.

In order to study the performance of each classifier in the classification of the candidates as MAs or non-MAs, rotative training and test sets were used. This was necessary because only the training images presented ground truth. So, they had to be used as training and test images (not at the same time).

Thus, the following steps had to be performed:

1. Partition of the 50 training images into 3 folds;
2. Consider the images of one of the folds for test and the others for training;
3. Construct the training set, which includes all candidates of the training images that are real microaneurysms and some candidates (also from the training images) that are not microaneurysms (the number of objects of the two classes was, approximately, equal);
4. Construct the test set, which comprises all candidates of the test images;
5. Train the classifier with the training set and test it with the test set;
6. Repeat steps 3 and 4, rotating the test and training sets.
7. Repeat steps 1 to 5 two more times.

This procedure allowed to train the 5 classifiers with 9 different training sets (3 folds $\times$ 3 repetitions) and test each one in an independent test set.

Regarding to the classification tree, figure 4.16 shows how the minimum number of leaf node observations (which is a parameter of the tree) affected the cross-validation error. As can be seen, a minimum cross-validation error was obtained when the minimum number of leaf node observations was 80.

Figure 4.17 presents the 9 ROC curves obtained when a classification tree was trained with 9 different datasets. In this case, all classification trees were constructed considering a minimum number of leaf node observations equal to 80.

By analysing figure 4.17, it can also be concluded that the performance of the classifier depends on the dataset used to train it, as expected. In this case, the best results were achieved when the classifier was trained with the dataset 9 (associated to the highest area under the ROC curve).

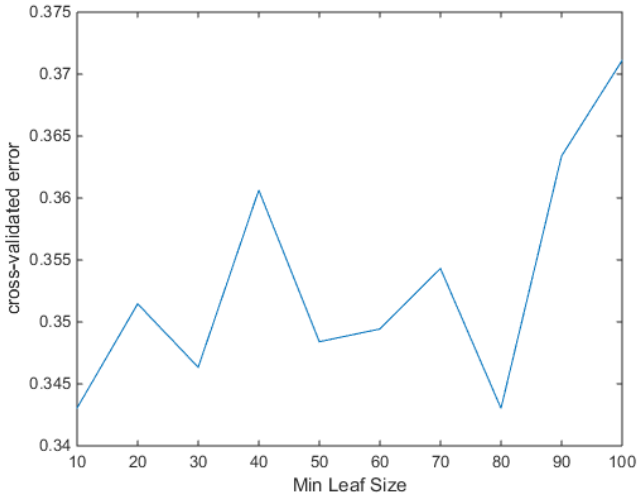

Figure 4.16: Cross-validation error as a function of the minimum number of leaf node observations of the classification tree.

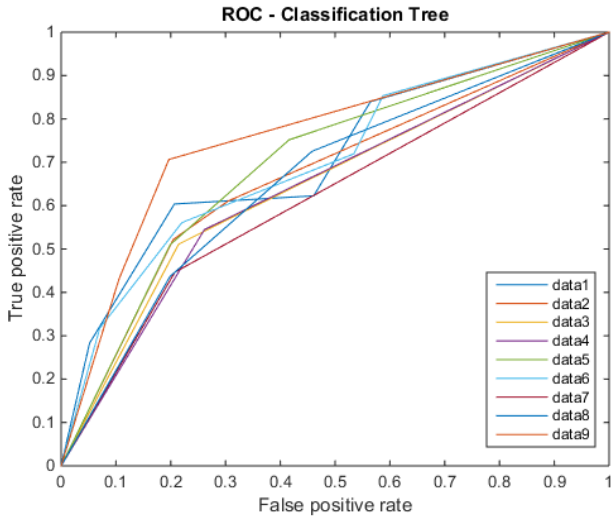

Figure 4.17: ROC curves of 9 different classification trees. Each classifier was trained with a different dataset.

The ROC curves obtained when a LDA classifier was trained with 9 different datasets are presented in figure 4.18. Contrarily to what happens for the classification trees, in this case, there is not a classifier that stands out from the others. This is related to the fact that the best classifiers present areas under the ROC curves very similar.

Figure 4.19 presents the ROC curves that correspond to the QDA classifiers trained with different datasets. By analysing this figure, it is possible to see that for false positive rates below 0.5, the best results were also achieved for the classifier trained with the dataset 9.

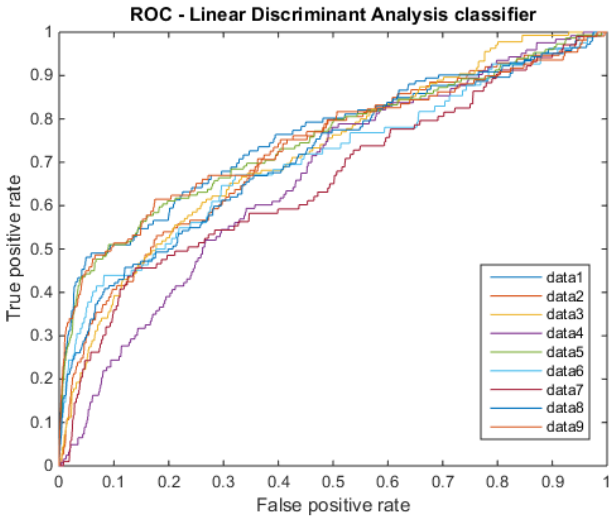

Figure 4.18: ROC curves of 9 different LDA classifiers. Each classifier was trained with a different dataset.

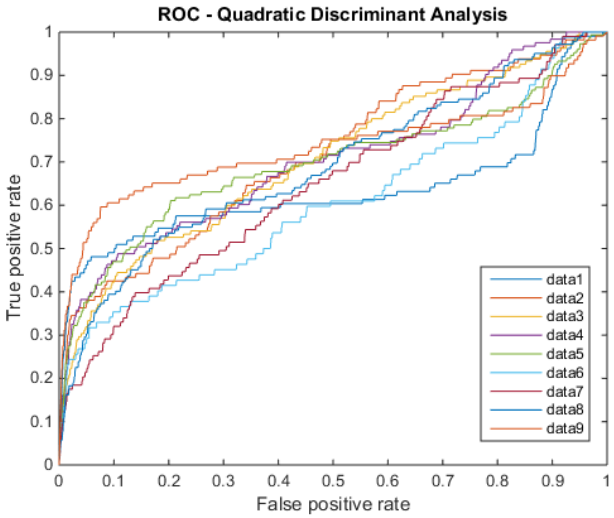

Figure 4.19: ROC curves of 9 different QDA classifiers. Each classifier was trained with a different dataset.

Regarding to the k-NN classifier, figure 4.20 shows how the number of neighbours considered for the classification affected the cross-validation error. As can be seen, a minimum cross-validation error was obtained when the number of neighbours used was 9.

Figure 4.21 presents the 9 ROC curves obtained when the 9-NN classifier is trained with 9 different datasets. By observing this figure, it is possible to say that the worst results were obtained when the 9-NN classifier was trained with the dataset 1 and the best results were obtained by training the classifier with the dataset 2.

The ROC curves obtained when the SVM classifier was trained with 9 different datasets are

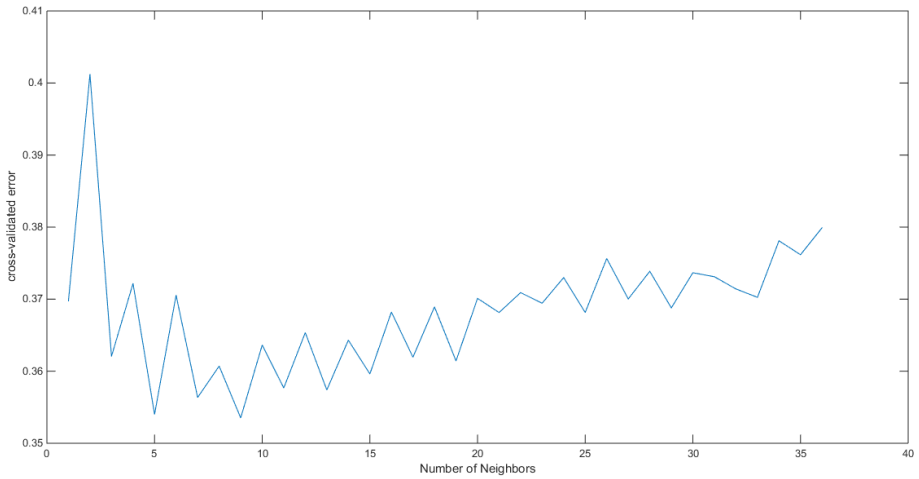

Figure 4.20: Cross-validation error as a function of the number of neighbours used by the k-NN classifier for classifying data.

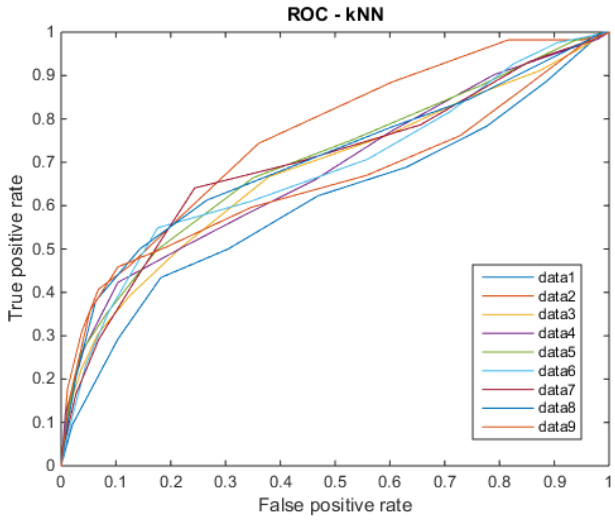

Figure 4.21: ROC curves of 9 different 9-NN classifiers. Each classifier was trained with a different dataset.

presented in figure 4.22. Observing the figure, it is possible to see that the best results were achieved when the SVM classifier was trained with the dataset 8.

Since the comparison of the performance of different classifiers can be difficult only by inspection of the curves, the average areas under the curves (AUC) were computed for all classifiers individually and are presented in table 4.5.

Based on these values, the SVM classifier is the one that gave better results, while the classification tree presented the worst.

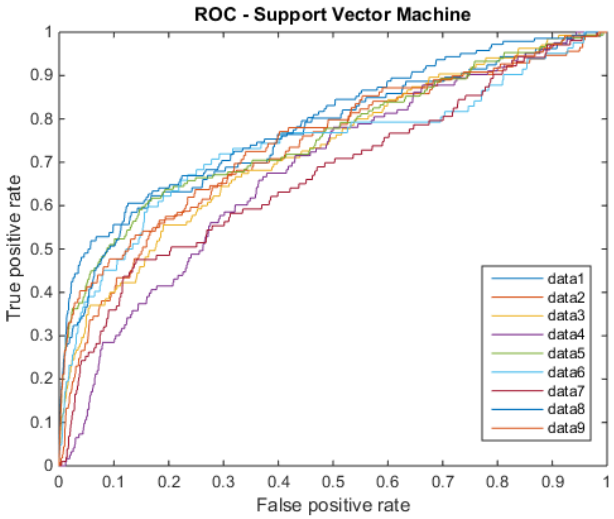

Figure 4.22: ROC curves of 9 different SVM classifiers. Each classifier was trained with a different dataset.

Table 4.5: Average area under the ROC curves for the 5 classifiers used in this work.

Classifier	Average area under the ROC curves
Classification tree	0.6790
QDA classifier	0.6807
k-NN classifier	0.6854
LDA classifier	0.7164
SVM classifier	0.7354

The SVM classifier which gave rise to the best results presented a sensitivity of 64.8% and a specificity equal to 79.7%.

Due to the lack of time, it was not studied the effect of feature extraction and selection on the performance of each classifier. However, these approaches can lead to better results.

The classification results here presented can not be compared directly with the other methodologies that also used the Retinopathy Online Challenge dataset, since the ground truth of the test images is not available publicly. While the other methodologies were trained with the complete training set (50 images) and tested in an independent test set (50 images), in this work the 50 training images were used for training and test.

However, the results obtained in the detection phase for the training dataset using the methodology herein presented can be compared with the results of the other candidate extractors (table 4.6).

By analysing table 4.6, it can be concluded that our method outperforms the methods proposed by Walter et al. (2007), Zhang et al. (2010) and Abdelazeem (2002) (which are associated to a

Table 4.6: Comparison of different candidate extractors on the Retinopathy Online Challenge training set.

Candidate extractor	Average number of false positives per image	Sensitivity	Sensitivity of our method	Sensitivity obtained by Dai et al. (2016)
Spencer et al. (1996)	20.30	0.12	0.09	**0.29**
Lazar et al. (2011)	73.94	**0.48**	0.24	0.47
Walter et al. (2007)	154.42	0.36	0.38	**0.55**
Zhang et al. (2010)	328.30	0.33	0.48	**0.62**
Abdelazeem (2002)	505.85	0.28	0.53	**0.68**

high number of false positives per image). Nevertheless, the method recently published by Dai et al. (2016) presented the highest sensitivity for all situations, except for 73.94 false positives per image. In this case, the method proposed by Lazar and Hajdu (2011) was the one that gave rise to the highest sensitivity.

4.4 Concluding remarks

The dataset used to test the viability of the methodology herein reported is very heterogeneous, including images with different resolutions. The low quality of these images as well as the low contrast between the microaneurysms and background make the detection of these lesions a very difficult task.

Although the LoG and the SBF enhance well the microaneurysms, these filters also present strong responses in the blood vessels and noisy regions, giving rise to a high number of false detections.

The thresholding operations can reduce significantly the number of false positives, but they also lead to the decrease of the sensitivity. So, a trade-off between these two variables has to be done.

For reducing the number of false detections, the incorporation of a classification stage is crucial. From the main types of supervised classifiers, the SVM classifier is the one that presents best performances in the classification of the candidates as MAs or non-MAs.

The best result presents a sensitivity of 64.8% and a specificity of 79.7%.

Chapter 5

Conclusions

Despite the main objective of this master thesis was to detect red lesions in retinal photographs - microaneurysms and haemorrhages -, the methodology herein proposed only detects microaneurysms and dot haemorrhages (which are almost indistinguishable). The large haemorrhages can not be detected by this method, due to the diversity of their shapes.

By analysing retinal photographs, it was possible to verify that the contrast between the red lesions and the background is low in these type of images. From all color channels, the green component is the one where the red lesions are better distinguished from the background.

As in a screening program, the dataset used presents images with different resolutions. For applying the same method to all types of images, a scale factor must be applied to the filters in order to achieve better results.

Although the LoG and SBF, used in this work for MA enhancement, presented strong responses in the MA locations, they also enhanced undesirable structures, such as the blood vessels and noisy regions. For the maximum sensitivity, the SBF is the one that presented a high number of false detections. However, through the combination of both responses and the removal of the candidates lying on the vasculature, the number of false detections was reduced significantly.

In this type of CAD systems, the incorporation of a classification phase is crucial to obtain the final lesions. Since most of the false positives correspond to elongated structures, features related to the shape of the candidates should be used. Besides these features, the color as well as the filters' response can also be used to distinguish the regions corresponding to microaneurysms from the others. Considering all supervised classifiers used in this work, the SVM is the one that better classifies the candidates as MAs or non-MAs, presenting a sensitivity of 64.8% and a specificity of 79.7%.

The methodology developed in this work for extracting the candidates from the retinal photographs outperforms the candidate extractors proposed by Walter et al. (2007), Zhang et al. (2010) and Abdelazeem (2002). However, the best results are achieved using the methodology recently published by Dai et al. (2016).

Regarding to the results of the classification phase, they could not be compared directly with the other methodologies, since the ground truth of the test images is not available publicly.

Chapter 6

Future work

Although many efforts were done to minimize the number of false detections and maximize the sensitivity of the CAD system, this proved to be a very difficult task, since a trade-off between the number of lesions correctly detected and the number of false positives has to be done. So, in the future, it is important to find alternative techniques for reducing the false positives without compromising the number of lesions detected.

Furthermore, the method used for removing the overlaps after the candidate growth should be optimized, in order to minimize the computational time associated with this task.

Regarding to the classification step, it should be studied the effect of feature extraction and selection in the performance of the classifiers.

Since the methodology herein reported only detects microaneurysms and dot haemorrhages, a methodology for detecting large haemorrhages should be developed in the future.

In order to compare directly the results of the classification phase obtained using this methodology with the others already developed, the results obtained for the test images have to be submitted in the Retinopathy Online Challenge website.

This methodology should also be validated for other datasets, such as the DiaretDB1 2.1 database (Kauppi et al., 2015).

References

S. Abdelazeem. Micro-aneurysm detection using vessels removal and circular hough transform. In *Radio Science Conference, 2002. (NRSC 2002). Proceedings of the Nineteenth National*, pages 421–426, 2002. doi: 10.1109/NRSC.2002.1022650.

B. Antal and A. Hajdu. An ensemble-based system for microaneurysm detection and diabetic retinopathy grading. *IEEE Transactions on Biomedical Engineering*, 59(6):1720–1726, June 2012. ISSN 0018-9294. doi: 10.1109/TBME.2012.2193126.

B. Antal and A. Hajdu. Improving microaneurysm detection in color fundus images by using context-aware approaches. *Computerized Medical Imaging and Graphics*, 37(5–6):403–408, 2013. ISSN 0895-6111. doi: http://dx.doi.org/10.1016/j.compmedimag.2013.05.001.

C. E. Baudoin, B. J. Läy, and J. C. Klein. Automatic detection of microaneurysms in diabetic fluorescein angiographies. *Revue d Épidémiologie et de Santé Publique*, pages 254–261, February 1984.

A. R. Bhavsar, J. H. Drouilhet, and N. H. Atebara. Diabetic retinopathy, April 2015. `http://emedicine.medscape.com/article/1225122-overview`.

L. Breiman, J. Friedman, C. J. Stone, and R. A. Olshen. *Classification and Regression Trees*. CRC Press, 1984. ISBN 978-04-120-4841-8.

R. M. Chong and J. C. Suniel. *Histogram-based masking technique for retinal fundus images*, pages 561–567. Springer Netherlands, Dordrecht, 2015. ISBN 978-94-017-9618-7. doi: 10.1007/978-94-017-9618-7_57.

M. Cree. The waikato microaneurysm detector. Technical report, University of Waikato, Hamilton, New Zealand, July 2008.

M. J. Cree, J. A. Olson, K. C. McHardy, J. V. Forrester, and P. F. Sharp. Automated microaneurysm detection. In *Image Processing, 1996. Proceedings., International Conference on*, volume 3, pages 699–702 vol.3, Sep 1996. doi: 10.1109/ICIP.1996.560763.

B. Dai, X. Wu, and W. Bu. Retinal microaneurysms detection using gradient vector analysis and class imbalance classification. *PLoS ONE*, 11(8), 2016. doi: 10.1371/journal.pone.0161556.

R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification, 2nd Edition*. John Wiley & Sons, Ltd, 2004. ISBN 978-04-700-9013-8.

T. Esteves, P. Quelhas, A. M. Mendonça, and A. Campilho. Gradient convergence filters and a phase congruency approach for in vivo cell nuclei detection. *Machine Vision and Applications*, 23(4):623–638, 2012. ISSN 1432-1769. doi: 10.1007/s00138-012-0407-7.

International Diabetes Federation. *IDF Diabetes atlas*. Sixth edition, 2013. ISBN 2-930229-85-3.

International Diabetes Federation. *IDF Diabetes atlas*. Seventh edition, 2015. ISBN 2-930229-81-2.

A. D. Fleming, S. Philip, K. A. Goatman, J. A. Olson, and P. F. Sharp. Automated microaneurysm detection using local contrast normalization and local vessel detection. *IEEE Transactions on Medical Imaging*, 25(9):1223–1232, Sept 2006. ISSN 0278-0062. doi: 10.1109/TMI.2006.879953.

H. Freeman. On encoding arbitrary geometric configurations. *IRE Trans. Electron. Comput.*, 10: 235–245, 1961.

L. Giancardo, F. Meriaudeau, T. P. Karnowski, Y. Li, K. W. Tobin, and E. Chaum. Microaneurysm detection with radon transform-based classification on retina images. In *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 5939–5942, Aug 2011.

R. Gonzalez and R. Woods. *Digital Image Processing*. Pearson Education, 2007. ISBN 9780131687288.

J. H. Hipwell, F. Strachan, J. A. Olson, K. C. McHardy, J. V. Forrester, and P. F. Sharp. Automated detection of microaneurysms in digital red-free photographs: a diabetic retinopathy screening tool. *Diabetic Medicine*, pages 588–594, August 2000.

National Eye Institute. Facts about diabetic eye disease, September 2015. `https://nei.nih.gov/health/diabetic/retinopathy`.

H. Jelinex and M. J. Cree. *Automated image detection of retinal pathology*. CRC Press, First edition, 2009. ISBN 978-0-8493-7556-9.

D. Kasper, A. Fauci, S. Hauser, D. Longo, J. Jameson, and L. Loscalzo. *Harrison's Principles of Internal Medicine 19/E (Vol.1 & Vol.2)*. McGraw-Hill Education, 2015. ISBN 9780071802161.

T. Kauppi, V. Kalesnykiene, J. Kamarainen, L. Lensu, I. Sorri, H. Uusitalo, H. Kälviäinen, and J. Pietilä. Diaretdb0: Evaluation database and methodology for diabetic retinopathy algorithms. Technical report.

T. Kauppi, V. Kalesnykiene, J. Kamarainen, L. Lensu, I. Sorri, A. Raninen, R. Voutilainen, J. Pietilä, H. Kälviäinen, and H. Uusitalo. Diaretdb1 - standard diabetic retinopathy database, December 2015. `http://www.it.lut.fi/project/imageret/diaretdb1/`.

I. Lazar and A. Hajdu. Microaneurysm detection in retinal images using a rotating cross-section based model. *Proceedings IEEE International Symposium on Biomedical Imaging*, pages 1405–1409, 2011. doi: 10.1109/ISBI.2011.5872663.

I. Lazar and A. Hajdu. Retinal microaneurysm detection through local rotating cross-section profile analysis. *IEEE Transactions on Medical Imaging*, 32(2):400–407, Feb 2013. ISSN 0278-0062. doi: 10.1109/TMI.2012.2228665.

T. Lindeberg. Scale-space theory: a basic tool for analysing structures at different scales. *Journal of Applied Statistics*, 21(2):225–270, Sept 1994. doi: 10.1080/757582976.

B. Läy. Analyse automatique des images angiofluorographiques au cors de la rétinopathie diabétique. Technical report, School of Mines, Paris, June 1983.

A. M. Mendonça and A. Campilho. Segmentation of retinal blood vessels by combining the detection of centerlines and morphological reconstruction. *IEEE Transactions on Medical Imaging*, 25(9):1200–1213, Sept 2006. ISSN 0278-0062. doi: 10.1109/TMI.2006.879955.

A. Mizutani, C. Muramatsu, Y. Hatanaka, S. Suemori, T. Hara, and H. Fujita. Automated microaneurysm detection method based on double-ring filter in retinal fundus images. *SPIE Medical Imaging 2009: Computer-aided diagnosis*, February 2009.

M. Niemeijer, B. van Ginneken, J. Staal, M. S. A. Suttorp-Schulten, and M. D. Abramoff. Automatic detection of red lesions in digital color fundus photographs. *IEEE Transactions on Medical Imaging*, 24(5):584–592, May 2005. ISSN 0278-0062. doi: 10.1109/TMI.2005.843738.

M. Niemeijer, B. van Ginneken, M. J. Cree, A. Mizutani, G. Quellec, C. I. Sanchez, B. Zhang, R. Hornero, M. Lamard, C. Muramatsu, X. Wu, G. Cazuguel, J. You, A. Mayo, Q. Li, Y. Hatanaka, B. Cochener, C. Roux, F. Karray, M. Garcia, H. Fujita, and M. D. Abramoff. Retinopathy online challenge: Automatic detection of microaneurysms in digital color fundus photographs. *IEEE Transactions on Medical Imaging*, 29(1):185–195, Jan 2010. ISSN 0278-0062. doi: 10.1109/TMI.2009.2033909.

C. Pereira, D. Veiga, J. Mahdjoub, Z. Guessoum, L. Gonçalves, M. Ferreira, and J. Monteiro. Using a multi-agent system approach for microaneurysm detection in fundus images. *Artificial Intelligence in Medicine*, 60(3):179 – 188, 2014. ISSN 0933-3657. doi: http://dx.doi.org/10.1016/j.artmed.2013.12.005.

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 Trans. Med. Imaging*, 27(9):1230–1241, 2008.

S. H. Rasta, M. E. Partovi, H. Seyedarabi, and A. Javadzadeh. A comparative study on preprocessing techniques in diabetic retinopathy retinal images: illumination correction and contrast enhancement. *Journal of Medical Signals and Sensors*, pages 40–48, January 2015.

C. I. Sánchez, R. Hornero, A. Mayo, and M. García. Mixture model-based clustering and logistic regression for automatic detection of microaneurysms in retinal images. *SPIE Medical Imaging 2009: computer-aided diagnosis*, February 2009.

T. Spencer, R. P. Phillips, P. F. Sharp, and J. V. Forrester. Automated detection and quantification of microaneurysms in fluorescein angiograms. *Graefe's Archive for Clinical and Experimental Ophthalmology*, 230(1):36–41, 1992. ISSN 1435-702X. doi: 10.1007/BF00166760.

T. Spencer, J. A. Olson, K. C. McHardy, P. F. Sharp, and J. V. Forrester. An image-processing strategy for the segmentation and quantification of microaneurysms in fluorescein angiograms of the ocular fundus. *Comput. Biomed. Res.*, 29(4):284–302, August 1996. ISSN 0010-4809. doi: 10.1006/cbmr.1996.0021.

Iowa University and the ROC organizers. Retinopathy online challenge, November 2015. `http://webeye.ophth.uiowa.edu/ROC/`.

F. van der Heijden, R. P. W. Duin, D. de Ridder, and D. M. J. Tax. *Classification, Parameter Estimation and State Estimation: An Engineering Approach Using MATLAB*. John Wiley & Sons, Ltd, 2004. ISBN 9780470090138.

T. Walter and J.-C. Klein. *Automatic detection of microaneurysms in color fundus images of the human retina by means of the bounding box closing*, pages 210–220. Springer Berlin Heidelberg, Berlin, Heidelberg, 2002. ISBN 978-3-540-36104-6. doi: 10.1007/3-540-36104-9_23.

T. Walter, P. Massin, A. Erginay, R. Ordonez, C. Jeulin, and J. C. Klein. Automatic detection of microaneurysms in color fundus images. *Medical Image Analysis*, pages 555–566, December 2007.

B. Zhang, X. Wu, J. You, Q. Li, and F. Karray. Detection of microaneurysms using multi-scale correlation coefficients. *Pattern Recogn.*, 43(6):2237–2248, June 2010. ISSN 0031-3203. doi: 10.1016/j.patcog.2009.12.017.

B. Zhang, F. Karray, Q. Li, and L. Zhang. Sparse representation classifier for microaneurysm detection and retinal blood vessel extraction. *Information Sciences*, 200:78 – 90, 2012. ISSN 0020-0255. doi: http://dx.doi.org/10.1016/j.ins.2012.03.003.