

AUTOMATIC RETINA EXUDATES SEGMENTATION WITHOUT A MANUALLY LABELLED TRAINING SET

Luca Giancardo, Fabrice Meriaudeau, T.P. Karnowski, Li Yi, Kenneth Tobin, Edward Chaum

▶ **To cite this version:**

Luca Giancardo, Fabrice Meriaudeau, T.P. Karnowski, Li Yi, Kenneth Tobin, et al.. AUTOMATIC RETINA EXUDATES SEGMENTATION WITHOUT A MANUALLY LABELLED TRAINING SET. International Symposium on Biomedical Imaging, Mar 2011, Russia. pp.1. hal-00585177

HAL Id: hal-00585177
https://hal.science/hal-00585177

Submitted on 15 Apr 2011

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

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

AUTOMATIC RETINA EXUDATES SEGMENTATION WITHOUT A MANUALLY LABELLED TRAINING SET

L. Giancardo[⋆†] *F. Meriaudeau*[⋆] *T.P. Karnowski*[†] *Y. Li*[‡] *K.W. Tobin Jr.*[†] *E. Chaum*[‡]

[⋆] University of Burgundy
[†] Oak Ridge National Laboratory
[‡] University of Tennessee Health Science Center

ABSTRACT

Diabetic macular edema (DME) is a common vision threatening complication of diabetic retinopathy. In a large scale screening environment DME can be assessed by detecting exudates (a type of bright lesions) in fundus images. In this work, two new methods for the detection of exudates are presented. The methods do not require a lesion training set so the need to ground-truth data is avoided with significant time savings and independence from human error. We evaluate our algorithm with a new publicly available dataset from various ethnic groups and levels of DME. Also, we compare our results with two recent exudate segmentation algorithms on the same dataset. In all of our tests, our algorithms are either outperforming or in line with existing methods. Additionally, the computational time is one order of magnitude less than similar methods.

Index Terms— One, two, three, four, five

1. INTRODUCTION

Diabetic retinopathy (DR) is a progressive eye disease that currently affects 250 million of people worldwide. Diabetic macular edema (DME) is a complication of DR and it is a common cause of vision loss and blindness. DME occurs from swelling of the retina in diabetic patients due to leaking of fluid from microaneurysms within the macula. Ophthalmologists can infer the presence of the fluid that causes the retina thickening in diabetic patients by the presence of accompanying lipid deposits called exudates. They appear as bright structures with well defined edges and variable shapes.

The approaches to exudate segmentation presented in the literature can be roughly divided into four different categories. *Thresholding methods* base the exudate identification on a global or adaptive grey level analysis. A first attempt was presented in [1] and recently a more sophisticated method based on image normalisation and distribution analysis was presented in [2]. *Region growing methods* segment the images using the spacial contiguity of grey levels; a standard region growing approach is used in [3], which is very computationally expensive by being employed to the whole image. In [4] the computational issues are limited by employing edge detection to limit the size of regions. *Morphology methods* employ greyscale morphological operators to identify all structures with predictable shapes (such as vessels). These structures are removed from the image so that exudates can be identified [5, 6]. *Classification methods* build a feature vector for each pixel or pixel cluster, which are then classified by employing a machine learning approach into exudates or not exudates [7, 8, 9] or additional types of bright lesions [10, 11] (such as drusen and cotton wool spots).

In this paper we present two variations of a new exudate segmentation method that falls into the category of *thresholding methods*. Hence, the methods do not require any manual segmented lesion map for training to achieve good performance. In order to reduce the problem of noise we have developed a new way to normalise the fundus images. We directly compare our method with an implementation of a morphology based technique [6] and another thresholding based technique [2]. We selected the techniques that were tested with the largest datasets in the respective class.

In Section 2 we introduce the new public dataset used for testing the algorithm; Section 3 presents the details of the two automatic exudate segmentation methods; Section 4 presents the results by comparing them to other two published techniques; finally, Section 5 concludes the discussion.

2. MATERIALS

For the past 5 years, our group has been designing a semi-automated, HIPAA-compliant, teleophthalmology network for the screening of diabetic retinopathy and related conditions by employing a single macula centred fundus image acquired with a Zeiss Visucam PRO fundus camera. We have randomly extracted a set of 169 images representative of various degree of DME and patient ethnicity (the appearance of the retinal fundus greatly varies between ethnic groups). Each image of the dataset was manually segmented by one of the co-authors (E. Chaum, a practising Retina specialist). The manual segmentation allowed us to test the algorithm and it was not employed in the training phase. We have decided to

release the dataset to the research community on the website: `http://risa.uthsc.edu/dmed`.

3. METHOD

3.1. Preprocessing

3.1.1. Background Estimation

In our approach we employ noth the green channel I_g of the RGB colour space and the I_i channel of the HSI colour space. We start the analysis by estimating the background by the means of a large median filter, whose size is $\frac{1}{30}$ the height of the fundus image on I_i. This approach has been used by other authors (such as [12, 13]) and it has great computation performance advantages over the method from Foracchia et al., mainly because it does not use bilinear interpolation or multiple passes. In other median filtering normalisation techniques, the background is subtracted from the original image in order to obtain a normalised version. In our approach, we enhance the normalisation with the addition of a morphological reconstruction step [14]. This seems to improve the removal of nerve fibre layer and other structures at the edges of the ON, at little or no expense for the exudates. The following pseudocode illustrates this step in more details.

Algorithm 1 Background Estimation

```
function MorphBgEst(I_g)
    bgEst ← medianFilter(I_g)
    initialise(bgMask)                          ▷ set all the pixels to 0
    for y = 0 to height(I_g)−1 do
        for x = 0 to width(I_g)−1 do
            if I_g(x, y) < bgEst(x, y) then
                bgMask(x, y) ← bgEst(x, y)
            else
                bgMask(x, y) ← I_g(x, y)
            end if
        end for
    end for
    bgEst_2 ← morphReconstr(bgEst, bgMask)
return bgEst_2
end function
```

3.1.2. Image Normalisation

Once the background $bgEst_2$ is estimated, it is subtracted from the original image. The subtraction operation is performed with 16 bit signed precision in order to maintain the negative pixels. The image obtained shows a peculiar greylevel distribution: the highest peak of the histogram is always centred on zero regardless of the ethnicity of the patient or disease stratification. The histogram provides a clear distinction of two classes of structures: *dark structures* and *bright structures*. *Dark structures* are located in the left side of the histogram and represent the vasculature, the macula, dark lesions (such as microaneurysms and haemorrhages) and other structures due to the retinal pigment epithelium layer. Their distribution seem to vary depending on the ethnicity; the reason is related to the pigmentation of the patient, the less the pigmentation the more the choroidal vasculature is visible, hence it is visible as a bump in the negative side of the histogram. On the other hand *bright structures* are found in the positive side of the histogram which contains the optic nerve, bright lesions (such exudates and cotton wool spots), nerve fibre layer reflectance residuals (from the background removal), artefacts due to dirty lenses and other structures related to very bright retinal pigment epithelium layer. In this case, the distribution is fairly constant across different ethnicities.

Because of the alignment of the histogram after normalisation we can select all the exudate candidates I_{cand} with a hard threshold th_{cand}. This has obvious computational advantages in comparison with any model fitting operations, that not only do not seem to provide any quantifiable advantage for the selection of a threshold, but are also more sensitive to suboptimal background estimation. We have selected $th_{cand} = 3$ given that the original image I_i has a grey level range between 0 and 255. This parameter has not been selected by trial and error on the whole dataset, but by empirically choosing a value slightly above 0 in order to accommodate small background estimation errors.

3.2. Exudate detection

The exudate detection is performed by assigning a score for each exudate candidate. The exudate candidates are selected by running a 8-neighbour connected component analysis on I_{cand}. We have implemented two ways to assign this score, one based on *Kirsch's Edges* [15] and the other based on *Stationary Wavelets*. Both try to capture the higher edge values of the exudates in comparison the one of other candidates.

3.2.1. Kirsch's Edges

Kirsch's edges try to capture the external edges of the lesion candidate. This edge detector is based on the kernel k evaluated at 8 different directions on I_g. The kernel outputs are combined together by selecting the maximum value found on each pixel output. The result is stored in the final I_{kirsch} image.

$$k = \begin{bmatrix} \frac{5}{15} & -\frac{3}{15} & -\frac{3}{15} \\ \frac{5}{15} & 0 & -\frac{3}{15} \\ \frac{5}{15} & -\frac{3}{15} & -\frac{3}{15} \end{bmatrix} \quad (1)$$

As proposed by Sanchez et al., the average edge outputs of I_{kirsch} under each lesion cluster are calculated and assigned to the lesion in its entirety. The thresholds used to evaluate the final output are $th_{fin} \in \{0 : 0.5 : 30\}$.

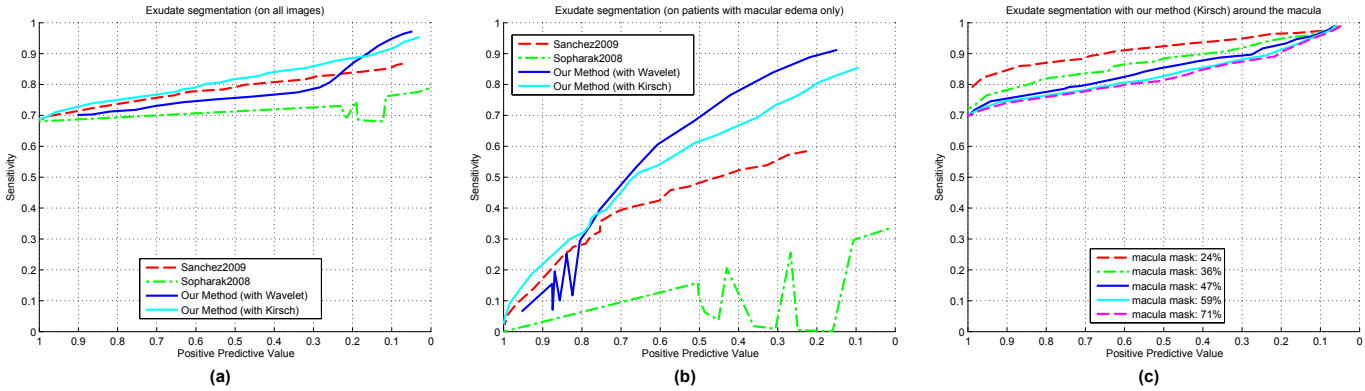


Fig. 1. FROC curves. (a) the exudate segmentation performances on all images; (b) the performance on the images diagnosed with DME only; (c) the exudate segmentation performances of our method using Kirsch's edges in the macula area; each line corresponds to a growing circular area centred on the fovea.

3.2.2. Stationary Wavelets

Quellec et al. [16] proposed a method for the detection of retina microaneurysms in the wavelet space. We try to capture the strong peak at the centre of exudates by developing a method that employs a similar wavelet as Quellec et al., but that evaluates the results in image space. A stationary Haar wavelet analysis is performed up to the second level on I_g. The process is inverted maintaining the last vertical, diagonal and horizontal details only, as these are the wavelet coefficients that seem to contain most of the foreground structures. By transforming back to the image space we obtain a background-less image I_{wav} with a strong response at the centre of exudates. It is interesting to notice that the distribution of I_{wav} has similar properties as the image obtained during the image normalisation phase (i.e. centred at 0 and with the exudates located on the positive side).

Similarly to the previous approach, we estimate the chance of being an exudate by evaluating the I_{wav} area which corresponds to each lesion cluster of I_{cand}. However, instead of calculating the average coefficient, we attempt to detect spikes in greyvalues as follows.

$$wavScore = \frac{max(px_{wav}) - min(px_{wav})}{max(px_{wav})} \qquad (2)$$

where px_{wav} are the pixels of I_{wav} corresponding to a lesion candidate cluster. The thresholds used to evaluate the final output are $th_{fin} \in \{0 : 0.05 : 1\}$.

4. RESULTS

In our results, we compare our two techniques described in this paper with our implementation of [6] and [2]. In all instances, we evaluate the performance of the lesion segmentation algorithms on a lesion by lesion basis for each image.

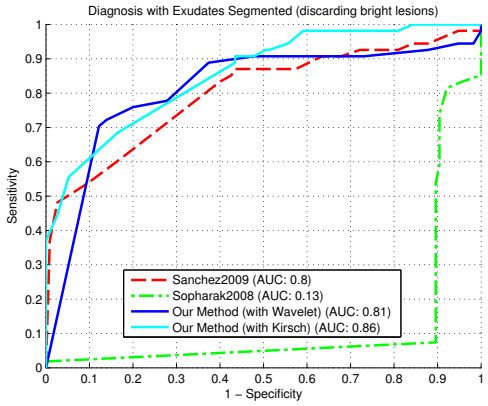


Fig. 2. ROC curves of the diagnosis of DME by employing only the lesions segmented, i.e. an image is considered positive if it shows at least a lesion. All the segmented lesions overlapping with the areas marked as "Other Bright Lesion" in the ground-truth are not taken in consideration.

The analysis starts from the labelled image in the dataset and each lesion is compared to the automatic segmentation one by one. A lesion is considered a true positive (TP) if it overlaps at least in part with the ground-truth; a false negative (FN) if no corresponding lesion is found in the automatic segmentation; a false positive (FP) if an exudate is found in the automatic segmentation but no corresponding lesion has been manually segmented. In the evaluation of the segmentation, we did not employ any true negatives (TN). As such, we avoid any evaluation of specificity, which is inherently high in images where the lesions represent a very small percentage of the total image area. In order to compare the methods fairly, all the images in the dataset have been resized to the resolution used in the original papers (but maintaining the original

width/height ratio). It should be noticed that in the test results of this section all the FP corresponding to bright lesions other than exudates have not been taken into account.

Fig. 1(a) shows the free-response receiver operating characteristic (FROC) analysis [17] of the implemented algorithms on the entire dataset. The global Sensitivity is evaluated as $\frac{TP}{TP+FN}$ on each image and then averaged together. The global Positive Predictive Value (PPV) is computed in a similar way but employing $\frac{TP}{TP+FP}$. It can be noticed how our two methods and the technique developed by Sanchez et al. performed comparably, with our technique (with Kirsch's edges) performing somewhat better.

Similar results are shown in Fig. 1(b). In this case only the 54 image with DME are used in order to have a better evaluation of the segmentation performance. In this case both of our methods seem to have a higher sensitivity than the other two methods implemented.

When exudation is close to the fovea it is more likely to cause vision loss. Therefore, we tested the ability of identifying lesions close to the macula of our method with Kirsch's edges, which seems to be the best performing algorithm of the previous tests. We have evaluated the lesion segmentation on the most dangerous area which might be affected by exudationi.e. A growing area around the fovea. All the lesions outside these areas have been discarded. Fig. 1(c) shows how the lesion segmentation improves when the lesion is closer to the fovea. This means that our algorithm is more effective on more sight threatening lesions. This is because the macular pigment in the area around fovea tends to be darker than the rest of the image, therefore exudates have more contrast. Also, when the area close to the vessels is discarded there are less chances of false positives due to the reflections of nerve fibre layer or to the muscular tissues of large vessels.

When developing an exudate segmentation method it must be taken into account that, in itself, the localisation of the lesions is not the final aim of an automatic diagnosis tool, but rather only a step towards the final diagnosis. As such, the detection of every single exudate in an image with many of them is not as important as finding a single lesion on a image with a single one. Therefore, we have evaluated the algorithms based on their ability to discern patients with or without DME by employing the hard threshold of a single lesion. If one or more exudates are found, the image is diagnosed with DME, otherwise the patient is classified as being negative. Fig. 2 shows the results of this experiment by the means of a standard ROC analysis on our dataset. Again, our methods seem to perform better or comparably to the other algorithms. Although the area under the ROC curve (AUC) is highest for our method with Kirsch's edges, the technique with wavelet shows a higher sensitivity at a higher specificity, a useful aspect for the development of a automatic DME screening tool.

The computational performance are evaluated on a Dual Core 2.6 GHz machine with 4 GB of RAM. All the algorithms are implemented in unoptimised Matlab code. Table 1 shows the average time to segment an image.

Table 1. Computational Performance

Method	Seconds per image
Sanchez et al.	$\sim$39
Sopharak et al.	$\sim$36
Our method (with Wavelet)	$\sim$2.4
Our method (with Kirsch)	$\sim$1.9

The reason for such performance discrepancy is likely to be due to the bilinear interpolation and expectation maximization for Sanchez et al. The calculation of the local standard deviation for each pixel is the cause for the computational performance degradation with the method by Sopharak et al.

5. DISCUSSION AND CONCLUSION

We have evaluated the global segmentation performance with the real distribution of patients in a screening setting and on only the patients showing signs of exudation. The results are particularly encouraging especially because of the comparison with the other techniques by Sopharak et al. and Sanchez et al. The method by Sanchez et al. is somewhat close to ours tests, however, the image normalisation procedure gives a substantial advantage to our method. The median filter with morphological reconstruction approach maintains a good contrast of the foreground structures by limiting the effects of the noise due to nerve fibre layer reflections and other small artefacts.

Also, we have evaluated our algorithms ability in identifying patients with DME based on the segmentation of one or more lesion in a fundus image. This is a naive method for DME diagnosis because neither it takes into account the many other features available in an image nor it employs any state of the art classification techniques. However, this test provides a baseline of the possible screening performances that can be achieved employing the output of the segmentation as a classification feature.

Many other exudate segmentation methods have been published throughout the years, however all of them seem to employ different datasets and different evaluation methods. This makes a direct comparison almost impossible, as is underlined by many of the authors themselves. We contribute to a solution to this problem by making available our heterogeneous, ground-truthed dataset to the research community. The need of having a common dataset to compare different techniques has been confirmed in this paper, where the two algorithms implemented did not perform as well as in the respective datasets employed by the original authors.

6. ACKNOWLEDGEMENTS

These studies were supported in part by grants from Oak Ridge National Laboratory, the National Eye Institute, (EY017065), by an unrestricted UTHSC Departmental grant from Research to Prevent Blindness (RPB), New York, NY, Fight for Sight, New York, NY, by The Plough Foundation, Memphis, TN and by the Regional Burgundy Council, France. Dr. Chaum is an RPB Senior Scientist.

7. REFERENCES

[1] R Phillips, J Forrester, and P Sharp, "Automated detection and quantification of retinal exudates," *Graefes Arch Clin Exp Ophthalmol*, vol. 231, no. 2, pp. 90–94, Feb 1993.

[2] C I Sanchez, M Garcia, A Mayo, M I Lopez, and R Hornero, "Retinal image analysis based on mixture models to detect hard exudates.," *Medical Image Analysis*, vol. 13, no. 4, pp. 650–658, Aug 2009.

[3] C Sinthanayothin, J F Boyce, T H Williamson, H L Cook, E Mensah, S Lal, and D Usher, "Automated detection of diabetic retinopathy on digital fundus images," *Diabetic Medicine*, vol. 19, no. 2, pp. 105–112, Feb 2002.

[4] H Li and O Chutatape, "Automated feature extraction in color retinal images by a model based approach," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 2, pp. 246–254, Feb 2004.

[5] T Walter, J K Klein, P Massin, and A Erginay, "A contribution of image processing to the diagnosis of diabetic retinopathy–detection of exudates in color fundus images of the human retina.," *IEEE Transactions on Medical Imaging*, vol. 21, no. 10, pp. 1236–1243, Oct 2002.

[6] Akara Sopharak, Bunyarit Uyyanonvara, Sarah Barman, and Thomas H Williamson, "Automatic detection of diabetic retinopathy exudates from non-dilated retinal images using mathematical morphology methods.," *Computerized Medical Imaging and Graphics*, vol. 32, no. 8, pp. 720–727, Dec 2008.

[7] G G Gardner, D Keating, T H Williamson, and A T Elliott, "Automatic detection of diabetic retinopathy using an artificial neural network: a screening tool," *Br J Ophthalmol*, vol. 80, no. 11, pp. 940–944, Nov 1996.

[8] A Osareh, M Mirmehdi, B Thomas, and R Markham, "Automated identification of diabetic retinal exudates in digital colour images," *British Journal of Ophthalmology*, vol. 87, no. 10, pp. 1220–1223, Oct 2003.

[9] M Garcia, C I Sanchez, M I Lopez, D Abasolo, and R Hornero, "Neural network based detection of hard exudates in retinal images.," *Comput Methods Programs Biomed*, vol. 93, no. 1, pp. 9–19, Jan 2009.

[10] M Niemeijer, B Van Ginneken, S R Russell, M S A Suttorp-Schulten, and M D Abrmoff, "Automated detection and differentiation of drusen, exudates, and cotton-wool spots in digital color fundus photographs for diabetic retinopathy diagnosis," *Investive Ophthalmology and Visual Science*, vol. 48, no. 5, pp. 2260–2267, May 2007.

[11] A D Fleming, K A Goatman, and S Philip, "The role of haemorrhage and exudate detection in automated grading of diabetic retinopathy," *British Journal of Ophthalmology*, August 2009.

[12] 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 Trans Med Imag*, vol. 24, no. 5, pp. 584–592, 2005.

[13] 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*, vol. 25, no. 9, pp. 1223–1232, Sept. 2006.

[14] L Vincent, "Morphological grayscale reconstruction in image analysis: applications and efficient algorithms," *IEEE Journal of Image Processing*, vol. 2, no. 2, pp. 176–201, 1993.

[15] R A Kirsch, "Computer determination of the constituent structure of biological images.," *Computers and Biomedical Research*, vol. 4, no. 3, pp. 315–328, Jun 1971.

[16] 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 on Medical Imaging*, vol. 27, no. 9, pp. 1230–1241, 2008.

[17] P Bunch, G Hamilton, J Sanderson, and A Simmons, "A free-response approach to the measurement and characterization of radiographic-observer performance," *Journal of Applied Photographic Engineering*, vol. 4, pp. 166–171, 1978.