Automatic Detection of Epiretinal Membrane in OCT Images by Means of Local Luminosity Patterns

Sergio Baamonde[1], Joaquim de Moura[1], Jorge Novo[1], and Marcos Ortega[1]

VARPA Group*, Departament of Computer Science,
University of A Coruña, A Coruña, Spain
{sergio.baamonde,joaquim.demoura,jnovo,mortega}@udc.es

Abstract. This work presents a novel approach for automatic detection of the epiretinal membrane in Optical Coherence Tomography (OCT) images. A tool able to detect this pathology is very valued since it can prevent further ocular damage by doing an early detection. This approach is based in the location of the inner limiting membrane (ILM) layers of the retina. Then, the detected locations are classified using a local-feature based vector in order to determine presence of the membrane. Different tests are run and compared to establish the appropriateness of the approach as well as its practical validity.

Keywords: epiretinal membrane, retinal layers, medical imaging, Optical Coherence Tomography

1 Introduction

Epiretinal membrane (ERM), also called macular pucker, is a macular pathology that can cause minor damage to the retina, like central vision decrease and metamorphopsia [11]. This disease can be caused by changes in the vitreous humor [4] and, consequently, the response of immune system to protect the retina can sometimes provoke that a number of cells converge on the macular area. This situation produces a transparent layer (Fig. 1) that, like every scar tissue, contracts causing tension on the retina, specifically on the inner limiting membrane (ILM). This phenomenon contributes to the appearance of ERM.

Since this pathology is frequently asymptomatic, it is imperative to develop a reliable system of detection to avoid further complications caused by its increasing severity.

In order to detect the ERM, ophtalmologists can work with the patient clinical history, looking for diabetes and ocular diseases or surgeries. Also, specialists

* This work is supported by the Instituto de Salud Carlos III, Government of Spain and FEDER funds of the European Union through the PI14/02161 and the DTS15/00153 research projects and by the Ministerio de Economía y Competitividad, Government of Spain through the DPI2015-69948-R research project.

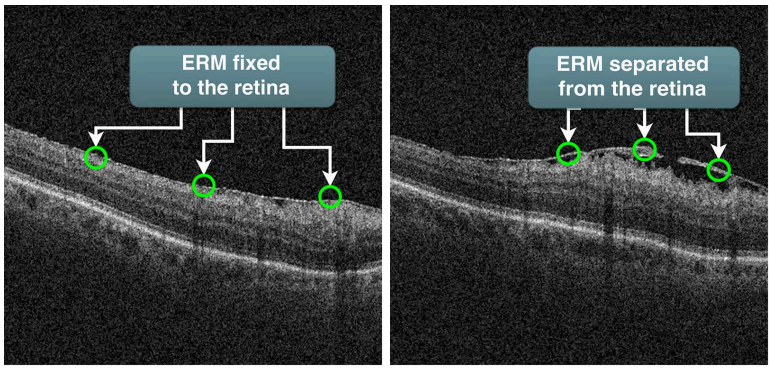


Fig. 1: Different appearances of ERM

can perform a complete ophthalmological evaluation to check for ERM, but at an additional cost and work hours.

The most precise way to evaluate the retinal morphology is doing an optical coherence tomography (OCT) scan [2], since the ERM appears as a bright layer on the retina [1]. Also, detecting irregularities on the retinal surface and/or retinal thickening, between others, can also mean that ERM is present on the patient.

Surgery may be needed when facing symptomatic ERM, e.g. vision loss, diplopia or debilitating metamorphopsia. When indicated, pars plana vitrectomy is performed [9]. However, ERM can recur and require further surgery. This recurrence rate can be reduced by undergoing ILM peeling [7].

The detection of the ERM is a manual process done by a specialist, but some tools have been developed to help with this task. Wilkins *et al.* [13] work with OCT pictures in real time, correcting patient's eye movement with image processing algorithms. Once the images are obtained, the specialist manually places computer cursors on the superficial and deep retinal boundaries. These boundaries are based on reflectivity and thickness differences between different areas of the retina.

Comparatively, other studies [8,3,6] work with spectral-domain OCT (SD-OCT). Its main advantage in comparison to time-domain OCT (TD-OCT) is the easier visualization of intraretinal layers (as the photoreceptor layer) through higher resolution pictures and the possibility of obtaining 3D images. This technique allows the specialists to obtain accurate surface maps and capture tension lines caused by the ERM on the ILM.

With this work we aim to create an automatic tool to detect epiretinal membrane presence on OCT pictures. The methodology consists on the processing of the OCT picture to locate the ILM layer of the retina, continuing with the extraction of relevant features of this layer. Finally, we will classify these data using classifiers trained beforehand to identify presence of ERM in the vicinity or adherent to the retina.

2 Methodology

Our methodology is based on the classification of the ILM located points to determine the presence of the membrane. To reach this goal, several stages are proposed, as shown on Fig. 2.

Fig. 2: Methodology used for the developing process

The first step is a preprocessing stage in order to remove undesired structures in the OCT input image as well as enhancing relevant ones.

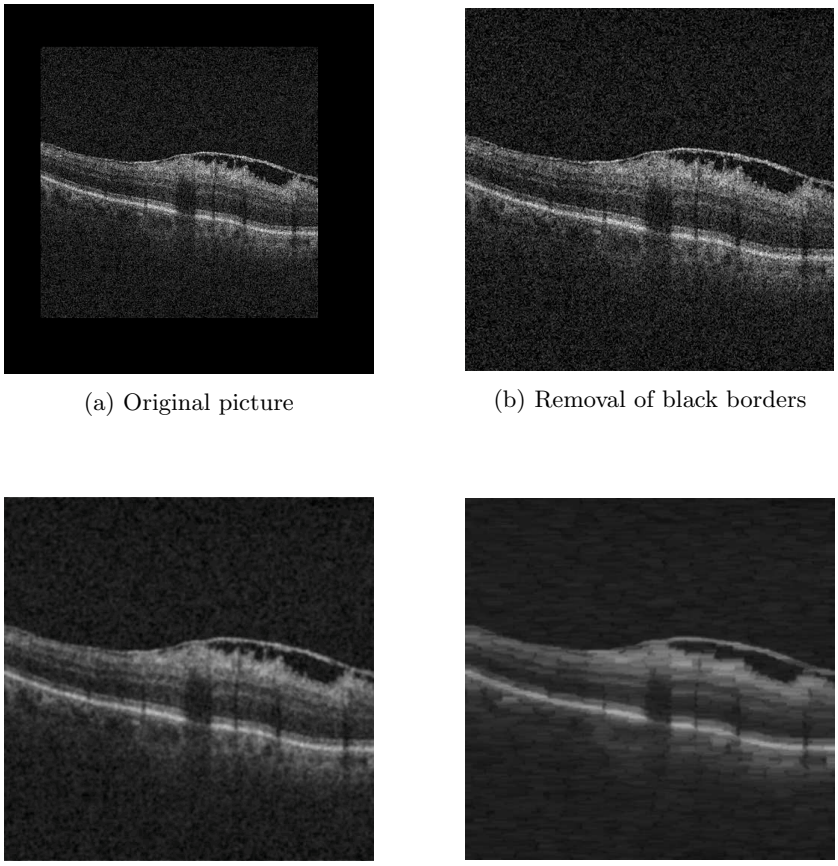
Afterwards, the goal is to locate the ILM as it represents our region of interest (ROI) given the fact that it is the location where the membrane appears. To this end, an active contour model (Snake) [5] is used to get the location of the topmost layer of the retina in the picture. This model will try to adapt its shape to the shape of the inner limiting membrane. Consequently, a fair amount of information is available about the split between background and eye zones.

In the next step, once the ILM is located, a feature extraction procedure takes place in each ILM location point in order to establish the presence of the membrane by using a trained classifier on these features, which would be the last step of our method. The feature vector for a ILM point would be defined in a small local window of the image surrounding that particular point. The idea behind this is to check for the ERM also in the zones where it is not adhered to the retina.

In order to be able to classify the existence of membrane, the vector will be based on local histograms of intensity, as this is the main characteristic of the membrane to be recognized. Following sections explain each step in more detail.

2.1 Region of Interest Segmentation

Preprocessing. In order to be able to correctly fit the shape of the Snake on top of the ILM, some preprocessing operations are performed on the image to avoid unnecessary elements. Figure 3 shows an example of the different steps at this stage. We remove every black border surrounding the OCT image and then we apply a Gaussian filter with $\sigma = 1.5$. This value was found to be good at preserving relevant features while filtering significantly. Finally, we apply a morphological operator (opening) to finish the cleaning of the picture and ease up the execution of the geometrical model.



(a) Original picture

(b) Removal of black borders

(c) Gaussian filter

(d) Opening operator

Fig. 3: Preprocessing applied to an OCT image sample

Layer segmentation. After preparing the image, since we want to approximate the shape of the inner limiting membrane, we use the active contour model mentioned beforehand. This model is initialized above all the layers of the retina, near the top border of the picture. When executed, it will try to converge on the topmost layer and adapt to its shape.

In this particular case we use a different approach of the Snake. We only allow downwards movement and, if the energy does not decrease, for a particular node this is stopped. This way, we ensure every point remains on their respective start columns. Also, with this approach, the Snake does not converge around an object, but instead lands on top of the upper layer, behaving as intended.

The energy of a Snake is defined in (1).

$$E = \int (\alpha(s)E_{\text{cont}} + \beta(s)E_{\text{curv}} + \gamma(s)E_{\text{img}})ds \tag{1}$$

Snake is defined as N points $p_1, p_2, ..., p_N$, so the formulations for each energy term are explained on (2), (3) and (4).

$$E_{\text{cont}} = \|p_i - p_{i-1}\|^2 \tag{2}$$

$$E_{\text{curv}} = \|p_{i-1} - 2p_i + p_{i+1}\|^2 \tag{3}$$

$$E_{\text{img}} = -\|\nabla I\| \tag{4}$$

where ∇I is the gradient of the intensity computed at each Snake point.

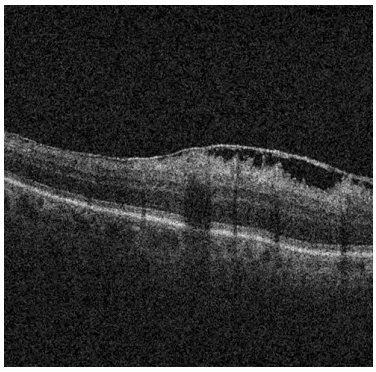
In order to get the Snake to adapt to the region of interest (ROI), an external energy is built based on the principle of distance to gradient. The main idea is to give the Snake an indication of the distance to the ROI (ERM and ILM) being the first relevant gradients in each image column on the ILM which is very strong also in the image. To achieve this energy, first edges are calculated via Sobel [12]. In general, we will aim to detect the limit between the background and the inner limiting membrane as a border, since we want the Snake to position above this sector.

Once we have this region segmented, we apply the Euclidean distance transform for the edge image. The resulting picture will be passed to the Snake as the external energy parameter. This way, the Snake will try to stick to the zones of less energy, that is, the zones where a border exist (the closest one being the border on the inner limiting membrane). An example of this procedure can be found on Fig. 4. Figure 4b shows the borders detected by the edge extraction algorithm, symbolized as white pixels. Figure 4c represent the external energy of the Snake, where dark areas are the zones of minimal energy. Lastly, Fig. 4d shows the final result of the Snake after finishing its iterations. Green crosses mark the topmost border (ideally ERM or ILM).

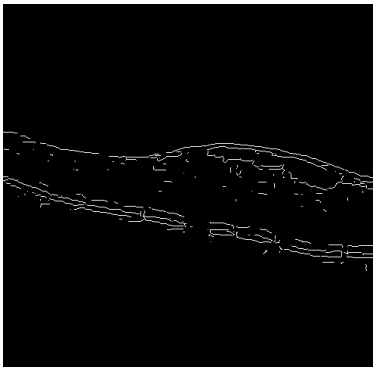
2.2 Feature extraction

Once the ROI is located in the image it is needed to establish the presence or absence of epiretinal membrane along the retina surface of a particular image. The hypothesis to achieve this is that luminosity of membrane differs sufficiently from ILM, the retina and image background. Thus, local features based on intensity can be defined on a vicinity of each Snake node to determine the existence of membrane in it by analyzing luminosity patterns. This is, a location with darker values above and under the central point should be a floating ERM, while if it only has dark values above and bright points under, can be a ERM next to the ILM or ERM nonexistence, depending on the intensity of the central window (brighter values are associated with ERM presence).

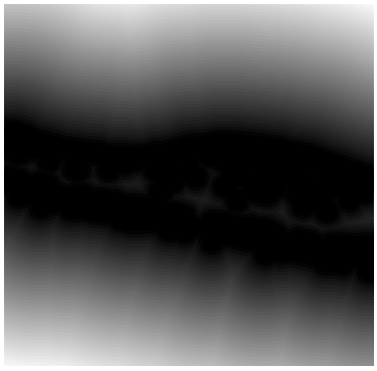
Following this hypothesis, after the Snake finishes its execution in the previous stage, local intensity features are computed for each node. These features need to contain information of the surroundings of the obtained points, more



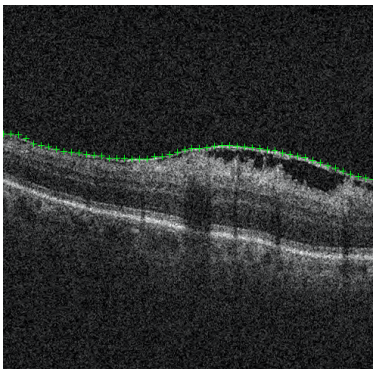
(a) Original picture



(b) Edge extraction algorithm



(c) Snake external energy



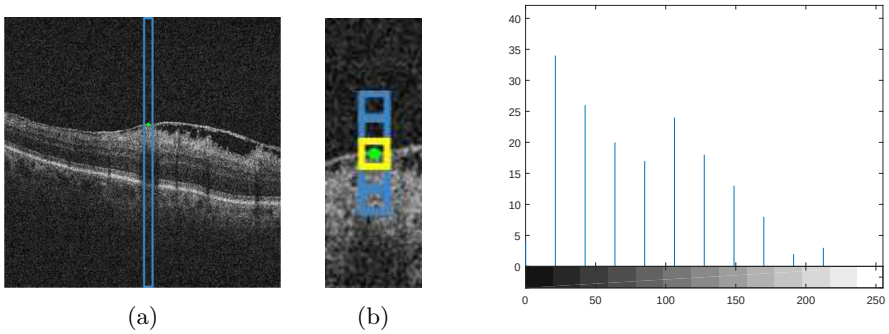
(d) Snake execution after 111 iterations

Fig. 4: Region of Interest segmentation procedure steps

precisely from the vertical area around the point. Having this information allows us to differentiate between points situated on the background from the ones situated on the ILM.

To this end, we will be using a series of vertical areas centered on the points of the Snake (Fig. 5). This area is divided in a series of W squared windows. For each of these windows, we will calculate afterwards the intensity histogram with N bins for the area. By appending all the bin values the feature vector is built. Lastly, all W feature vectors are combined in one full feature vector containing $N \times W$ elements. This vector represents the intensity values of the entire vertical area of the point.

For this work five regions are considered centered around the node. As Fig. 5 shows, data located above or under the limits on the defined windows do not contribute with any meaningful data for ERM location as it only adds redundant information. In the result section, several studies are conducted to establish a



(a) (b) (c)

Fig. 5: Vertical window around a Snake point. (a): Area around a Snake point. (b): Feature window of the associated point with $W_{\text{size}} = 13$. (c): Histogram of the central squared window of (b).

suitable value for window size. Nevertheless, size of the regions is matter of study in our experiment section.

We calculate the histogram with N bins for each squared window. This process will give us N discrete values for each window. Afterwards, all of these values will be converted to a full feature vector containing $N \times W$ values total. This vector contains all relevant information about the point and the surrounding area, more precisely its luminosity.

2.3 Layer classification

The final stage of the methodology is to perform a classification based on the intensity feature vectors. This way, each node is labeled as including ERM or not.

We will classify the points extracted from the image using a series of classifiers trained previously by using a 10-fold Cross-validation method with a set of samples manually labeled by a clinician. Each fold will use 90% of the samples as training samples and 10% as test samples. The models being used on this section are a Naive Bayes classifier, a Multilayer Perceptron and a Random Forest. We will generate different classifiers for each class with different parameters: number of bins and size of the squared windows. An example of the ERM recognition is shown on Fig. 6.

3 Experimental Results

OCT scans were obtained with a tomograph CIRRUS$^{\text{TM}}$ HD-OCT Zeiss, with Spectral Domain Technology. The resolution of the images is 490×500 pixels.

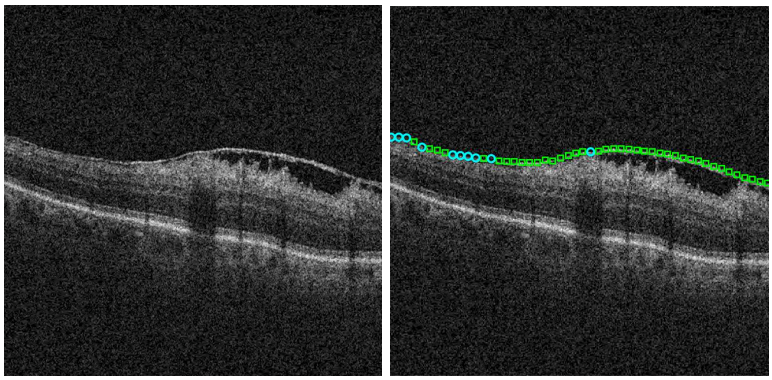


Fig. 6: Result of classification with Random Forest, $W_{\text{size}} = 13$ and $N_{\text{bins}} = 15$. Circles symbolize absence of ERM and squares presence of ERM

Our working set is comprised of 129 images showing different sections of the eye. ERM presence can be found in some of the pictures. Training samples have been randomly selected from all the pictures. In different experiments, separation between training and test is done accordingly as explained in following sections.

The energy terms used by the Snake in our experiment are shown on Table 1. These values have been selected because with the picture set we are using they give the Snake enough traction to provide a good approximation of the shape of the ILM. The high γ value allows the Snake to adapt to the ILM or ERM shape (zones of high energy), while α and β are less relevant because we only allow downwards movement so keeping the points clustered is not a relevant problematic.

Table 1: Energy terms used for the Snake

Energy type	Parameter	Value
E_{cont}	α	0.8
E_{curv}	β	0.4
E_{img}	γ	2.0

Our study of the methodology is done by performing two different experiments. First, we aim to separate the samples between 2 classes (*membrane* and *no membrane*) to get a first approximation about the presence or absence of ERM. Lastly, those samples will be divided instead on 4 classes, subdividing *membrane* class on *membrane* and *floating membrane* (ERM separated from the retina). Similarly, *no membrane* class is split on *no membrane* (points of the ILM with ERM absence) and *background*.

Our goal is to check what is the most accurate approximation (2 or 4 classes) while improving the behavior of the classifiers used by refining the parameters passed as input.

3.1 2-class classification

We will test first the behavior of the classifiers when using 2 different classes to split the data, as seen on Fig. 7:

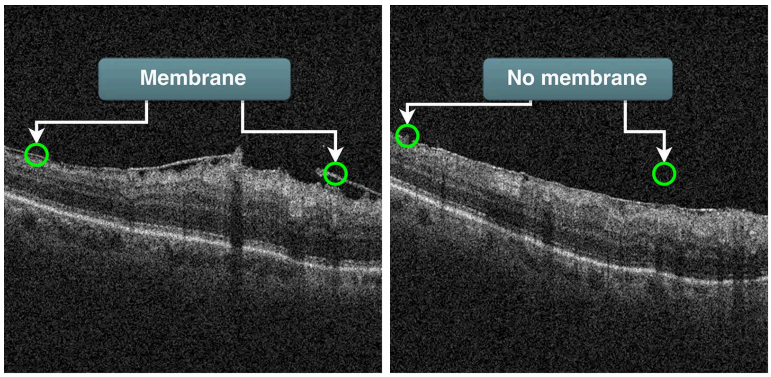


Fig. 7: Structure types used on classification

- 60 samples from class *membrane*. These points are the ones belonging to any point where ERM exists.
- 60 samples from class *no membrane*. This class contains any point not possessing ERM, either background or ILM without membrane.

These samples were used to train different classifiers. Each training iteration was repeated 10 times to obtain more accurate metrics. The results appearing here are the average of every iteration. Accuracy is defined in (5).

$$Acc = (TP + TN)/(P + N) \tag{5}$$

where TP and TN are True Positive and Negative values, while P and N are Positive and Negative values.

To evaluate the results, we use a k-fold cross-validation with $k = 10$ [10].

The different types of classifiers we used are a Multilayer Perceptron, a Naive Bayes classifier and a Random Forest classifier. These three approximations provide us a vast array of behaviors, allowing us to conclude what is the best approach for this problem.

The following tables (Table 2a, Table 2b and Table 2c) show the results of obtaining the accuracy of each classifier for different values of window size

Table 2: Accuracy of different classifiers

(a) Multilayer Perceptron accuracy

	N_{bins}				
W_{size}	5	10	15	20	25
5	75.75%	73.50%	72.67%	72.00%	69.83%
9	83.17%	80.92%	77.75%	76.75%	76.17%
13	**86.92%**	83.33%	84.92%	79.25%	77.75%
17	84.08%	79.50%	81.92%	78.08%	77.50%
21	79.75%	79.25%	81.83%	79.83%	84.17%
25	82.25%	82.33%	79.25%	79.50%	82.25%

(b) Naive Bayes classifier accuracy

	N_{bins}				
W_{size}	5	10	15	20	25
5	68.92%	70.17%	64.58%	68.17%	63.33%
9	70.92%	76.08%	73.58%	73.92%	74.42%
13	72.33%	**80.83%**	79.08%	73.33%	74.42%
17	80.50%	78.75%	75.08%	70.33%	70.83%
21	71.83%	76.83%	75.33%	76.75%	70.08%
25	79.92%	73.92%	73.53%	75.83%	69.75%

(c) Random Forest classifier accuracy

	N_{bins}				
W_{size}	5	10	15	20	25
5	79.67%	78.00%	78.00%	75.17%	75.25%
9	84.83%	85.83%	85.58%	83.92%	82.08%
13	87.33%	91.08%	**91.25%**	88.08%	86.75%
17	86.33%	85.92%	85.92%	85.67%	85.83%
21	83.83%	84.42%	83.58%	85.75%	83.75%
25	84.67%	81.25%	84.75%	84.58%	81.83%

(W_{size}) and number of bins (N_{bins}). The most accurate classifier in each series is bolded for clarity.

A more in-depth comparison between the best approximation of each class is done by comparing side to side the ROC curves of each classifier (Fig. 8). We can conclude that with the results we have obtained (Fig. 9), the best classifier is the one based in Random Forest method, accuracy-wise and with better ROC values.

3.2 4-class classification

In this experiment the goal was to assess the performance of a classifier able to distinguish four scenarios:

- Class *membrane*. These points are the ones belonging to zones where the ERM is fixed to ILM layer.
- Class *floating membrane*. Here we group ERM points situated on the background.
- Class *no membrane*. This class contains points from the retina not containing ERM.
- Class *background*. Background points not belonging to any of the other classes are classified here.

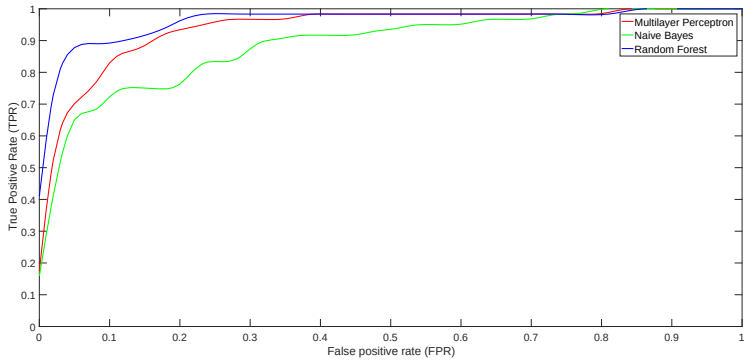


Fig. 8: ROC Curves for each best classifier. Random Forest scores above the other 2 classifiers

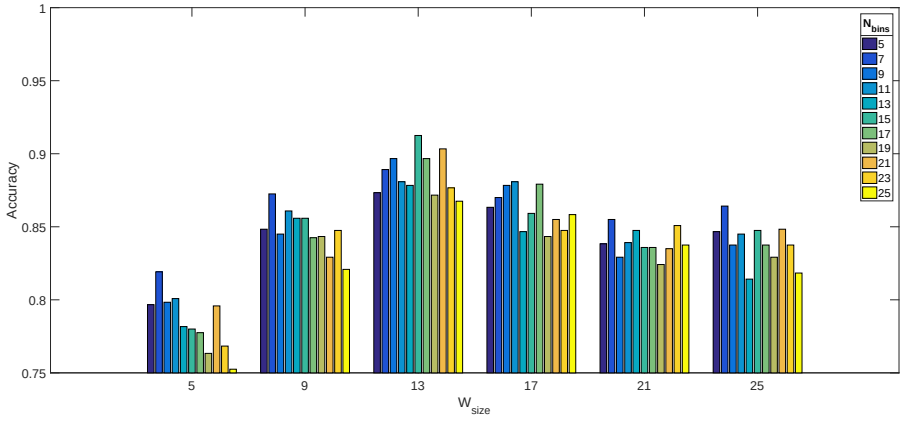


Fig. 9: Accuracy for Random Forest classifiers with 2 classes. Best results are found with $W_{\text{size}} = 13$ and $N_{\text{bins}} = 15$

For each class we have obtained 30 samples that will be used for training, test and validation of the classifiers. In Fig. 7 a sample of each different class is shown.

Based on the previous results, we will work with a Random Forest classifier with default parameters and same input data as the last section. Results can be seen on Fig. 10.

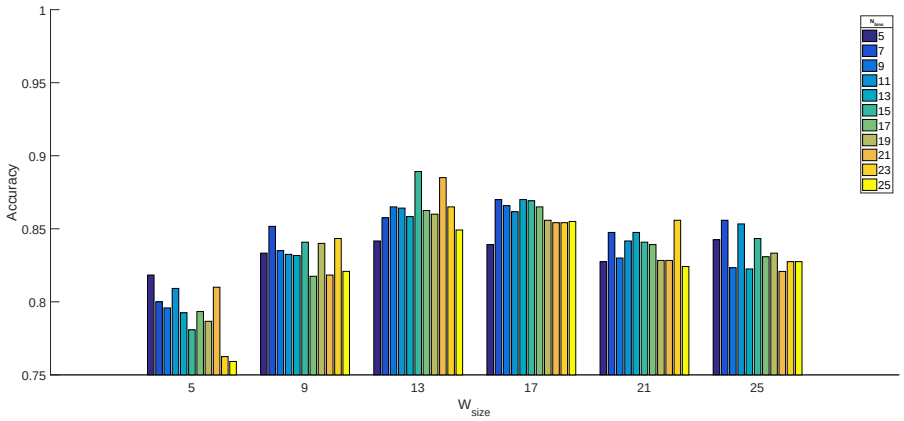


Fig. 10: Accuracy for Random Forest classifiers with 4 classes. Best results are found with $W_{\text{size}} = 13$

As before, we choose the most accurate classifier. With the information of Table 3 we can conclude that the best model is, as before, the one with $N_{\text{bins}} = 15$. In this case, to make a deeper analysis, we extract its confusion matrix (Table 4). With this data, we can deduct that differentiating between *membrane* and *no membrane* class is the process that contributes the most to the inaccuracy of the classifier. This is coherent with our last approach (splitting between those both classes only) being the main focus of this work.

Table 3: Accuracy for Random Forest with 4 classes and $W_{\text{size}} = 13$

Number of bins (N_{bins})										
5	7	9	11	13	15	17	19	21	23	25
84.17%	85.75%	86.50%	86.42%	85.83%	**88.34%**	86.25%	86.00%	88.05%	86.50%	84.92%

We can see that the behavior of each approximation is similar: a low value of W_{size} or too high causes inaccuracy in the classifiers since we are introducing noise instead of useful information. Also, we see better accuracy with N_{bins}

Table 4: Confusion matrix for Random Forest with $W_{\text{size}} = 13$ and $N_{\text{bins}} = 15$

Ground Truth \ Obtained results	Membrane	Floating membrane	No membrane	Background	
Membrane	24	1	5	0	80.0% 20.0%
Floating membrane	0	30	0	0	100.0% 0.0%
No membrane	7	0	23	0	76.7% 23.3%
Background	0	1	0	29	96.7% 3.3%
	77.4% 22.6%	93.8% 6.2%	79.3% 20.7%	100.0% 0.0%	88.3% 11.7%

values in the range of 13 to 17. As with window size, values too low or too high provoke inconsistencies in the classification. Better values for 2 classes and 4 classes are situated on $W_{\text{size}} = 13$, giving us a good approximation about what is a good value to get the most information but avoiding unnecessary noise.

4 Conclusions

Identifying the appearance of epiretinal membrane is an important process in the opthalmologic field, since it can improve the results of ERM extraction surgery.

In this paper, we have developed an automatic process to detect the ERM on OCT pictures with deformable models. First, we situated a number of points on the suitable area where the ERM can appear. Then, we extracted information from a series of windows situated around those points. With this information, we generated a feature vector from the values of the histograms of those windows. Lastly, we used different classifiers to classify those feature vectors and obtained the classes associated to each point.

The methodology is very recent and open for improvement. Nevertheless, results have been so far very promising, justifying further development within this field. These results may be improved by increasing the number of samples used in training, to split better the classes *membrane* and *no membrane* and increase the precision of the classifiers. Also, the use of more samples will provide the classifiers with better data about each class, improving the overall robustness of the system.

A first proof about the ideal number of classes was also developed, allowing us to conclude that the approximation using only 2 classes is more accurate than introducing another 2 classes, giving us 4 in total. In future works, a tool to separate the points of ERM fixed on the retina from the ones on the background will need to be developed to provide that information to specialists.

References

1. Brancato, R.: Optical coherence tomography (OCT) in macular edema. Documenta Ophthalmologica 97, 337-339 (1999).
2. Do, D.V., Cho, M., Nguyen, Q.D., Shah, S.M., Handa, J.T., Campochiaro, P.A., Zimmer–Galler, I., Sung, J.U., Haller, J.A.: Impact of Optical Coherence Tomography on Surgical Decision Making for Epiretinal Membranes and Vitreomacular Traction. Retina 27, 552-556 (2007).
3. Falkner–Radler, C.I., Glittenberg, C., Hagen, S., Benesch, T., Binder, S.: Spectral-Domain Optical Coherence Tomography for Monitoring Epiretinal Membrane Surgery. Ophthalmology 117, 798-805 (2010).
4. Foos, R.Y.: Vitreoretinal juncture; epiretinal membranes and vitreous. Invest. Ophthalmol. Vis. Sci. 16, 416-422 (1977).
5. Kass, M., Witkin, A., Terzopoulos, D.: Snakes: Active contour models. Int J Comput Vision 1, 321-331 (1988).
6. Koizumi, H., Spaide, R.F., Fisher, Y.L., Freund, K.B., Klancnik Jr, J.M., Yannuzzi, L.A.: Three-Dimensional Evaluation of Vitreomacular Traction and Epiretinal Membrane Using Spectral-Domain Optical Coherence Tomography. American Journal of Ophthalmology 145, 509-517.e1 (2008).
7. Kwok, A. K., Lai, T. Y. and Yuen, K. S.: Epiretinal membrane surgery with or without internal limiting membrane peeling. Clinical & Experimental Ophthalmology, 33: 379-385. (2005).
8. Legarreta, J.E., Gregori, G., Knighton, R.W., Punjabi, O.S., Lalwani, G.A., Puliafito, C.A.: Three-Dimensional Spectral-Domain Optical Coherence Tomography Images of the Retina in the Presence of Epiretinal Membranes. American Journal of Ophthalmology 145, 1023-1030.e1 (2008).
9. Machemer R.: A New Concept for Vitreous Surgery. 7. Two Instrument Techniques in Pars Plana Vitrectomy. Arch Ophthalmol. 92(5), 407–412. (1974).
10. McLachlan, Geoffrey J., Do, Kim–Anh, Ambroise, Christophe.: Analyzing microarray gene expression data. Wiley (2004).
11. Medina, C.A., Townsend, J.H., Singh, A.D. (Eds.): Manual of Retinal Diseases. Springer International Publishing (2016).
12. Sobel, I., Feldman, G.: A 3x3 Isotropic Gradient Operator for Image Processing. Talk at the Stanford Artificial Intelligence Project (SAIL), 271-272 (1968).
13. Wilkins, J.R., Puliafito, C.A., Hee, M.R., Duker, J.S., Reichel, E., Coker, J.G., Schuman, J.S., Swanson, E.A., Fujimoto, J.G.: Characterization of Epiretinal Membranes Using Optical Coherence Tomography. Ophthalmology 103, 2142-2151 (1996).