# Feature definition and selection for epiretinal membrane characterization in Optical Coherence Tomography images

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

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

**Abstract.** Optical Coherence Tomography (OCT) is a common imaging technique for the detection and analysis of optical diseases, since it is a non invasive method that generates in vivo a cross-sectional visualization of the retinal tissues. These characteristics contributed to the use of OCT imaging in the analysis of pathologies as, for instance, vitreomacular traction, age-related macular degeneration or hypertension. Among its applications, OCT imaging can be used in the detection of any present epiretinal membrane section in the retina, a critical issue to prevent further complications caused by this pathology.

This work analyzed the main characteristics of the epiretinal membrane to define a complete and heterogeneous set of intensity and texture-based features. Those features were studied using representative selectors, as Correlation Feature Selection (CFS) and Relief-F, to identify the optimal subsets that offer the higher discriminative power. K-Nearest Neighbor (kNN), Naive Bayes and Random Forest were finally tested in a method for the automatic detection of the epiretinal membrane in OCT images. Previous works do not focus on automatic procedures and, on the contrary, depend on manual markers or supervised detections, while our method improves significantly this task by automating the search of the region of interest and the classification of the pixels belonging to that area.

The methodology was tested using a dataset of 129 OCT images. 120 samples were equally obtained from those scans, featuring both zones with and without epiretinal membrane. The best results were provided by the Random Forest classifier that, using a window size of 15 pixels, a quantity of 13 histogram bins and 28 features, achieved an accuracy of 93.89%.

**Keywords:** computer-aided diagnosis, retinal imaging, Optical Coherence Tomography, epiretinal membrane, feature selection, classification

* 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.

## 1 Introduction

Retinal image analysis is an important issue for the diagnosis of various optical diseases. To this end, it is necessary to identify precisely the pertinent structures of the eye fundus as, for instance, the optic disc [13] and the arterio-venular tree [12]. With this information, a characterization of cardiovascular complications [7] or pathologies such as diabetes [17] can be achieved.

Macular pucker, more commonly known as epiretinal membrane (ERM), is a fibrocellular tissue that can cause metamorphopsia, central vision decrease or blurred vision [1,10]. Moreover, epiretinal membranes are associated with different types of cysts (macular, paravascular, lamellar macular) [11], further contributing to the eyesight distortion or reduction.

Idiopathic ERMs are the most common, but retinal vascular diseases or changes in the vitreous humor [4] can induce a response from the immune system to protect the retina. This response causes, sometimes, that the retinal cells converge on the macular region, creating a transparent layer. This layer, that is scar tissue, causes tension on the retina by contraction, further increasing the chances of secondary ERMs to appear.

Optical Coherence Tomography (OCT) imaging [3] is frequently used to analyze the retinal morphology and detect the presence of ERM. ERM appears as a thin reflective layer on the retina [2], fact that can be used for its detection on OCT images. Irregularities on the retinal surface or retinal thickening can also indicate the presence of ERM on the patient.

The asymptomatic nature of this pathology makes necessary a reliable and accurate detection system. With an appropriate method, ERM can be early detected and treated before further complications appear. Those methods are usually based in the manual detection of the ERM by a specialist [14]. Similarly, the method of Wilkins *et al.* [16] uses real-time OCT images and, after an specialist establish manual markers on the image, ERM is detected by the use of information about the reflectivity and thickness of the retina on the selected points.

In this work, we aimed for the automation of the process by developing an algorithm that selects autonomously the region of interest (ROI) where the ERM can be present. We analyzed the main characteristics of the ERM and designed a complete and heterogeneous set of features that helped to characterize the regions where the ERM is present. Optimal subsets of those features were selected and used to train representative classifiers. We use those trained classifiers to identify automatically the points belonging to the region of interest and pinpoint the presence or absence of ERM in the selected area. This method aims to improve the general error tolerance of the process by avoiding the use of manual markers for ROI initialization and making them non-dependent of human interaction.

## 2 Methodology

The proposed system tries to identify automatically the Inner Limiting Membrane (ILM), which is the boundary between the retina and the vitreous body, area where the ERM can appear if the pathology is present on the patient. Then, using these identifications, we analyze all the points belonging to the ILM, generating a rectangular-shaped window for each point and calculating the relevant features of the constructed window. Finally, every feature vector is used to classify its associated point and obtain information about the presence of ERM in the ILM retinal layer.

### 2.1 Identification of the region of interest

In this work, we employed a new method based on the use of an active contour model (Snake) [8], which adjusts its shape to the ILM contour. A predefined number of points are initialized on the uppermost part of the OCT scan. These points adapt its contour to the shape of the ILM by using information about the intensity of this layer in contrast to the rest of the layers. We designed an adapted version of the Snake, since we restrict its movement to the vertical axis, allowing only downwards movement. All the points of the Snake are moved progressively, approaching to the ILM layer. Finally, if a point does not modify its energy after an iteration, that point is fixed and is not processed again. With this method, the Snake behaves like a cascade of points instead of a contracting closed shape.

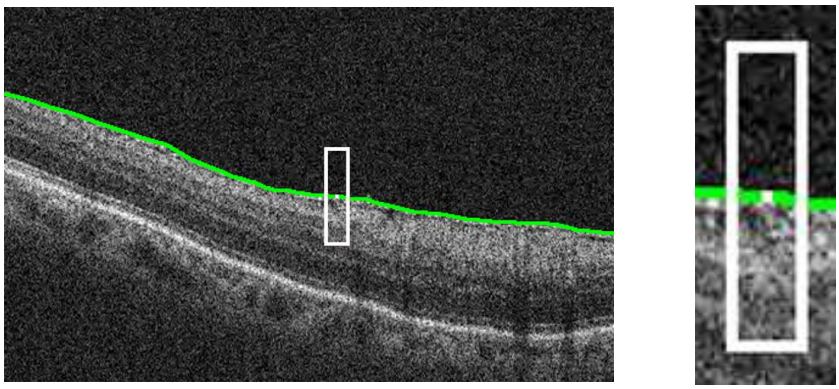
The Snake finally reaches the ROI (defined by the ILM position) to identify the ERM presence. In order to obtain relevant information from the ROI, a large set of heterogeneous features is obtained from each point of the Snake. These features are measured in the surrounding area of each point of interest. This area is defined as a rectangular window where $W_{\text{size}}$ is the width in pixels of the window and the height is $5 \times W_{\text{size}}$ (Fig. 1), offering enough information of the layer tissue with respect to its surrounding area.

### 2.2 Feature definition

Using the properties of the ILM with and without ERM presence, we selected a complete set of intensity and texture-based features of the windows obtained around the points of interest to be able to separate precisely the points with ERM from the normal ILM tissue.

The number of features varies between 223 and 263. This variability is caused by the use of the input parameter $N_{\text{bins}}$ of the window features, depicted below. The used features can be classified in the following groups:

**Window features** Each window obtained from the Snake is divided in five different square-shaped windows. Then, we calculate the histogram for each sub-window. The number of bins was empirically selected, so the resulting number of features obtained oscillates between 35 and 85, depending on the



(a) (b)

Fig. 1: Example of Region of Interest definition. (a): Snake situated on top of the ILM and window around a point of the ROI. (b): Zoom on the feature window

value of $N_{\text{bins}}$. This parameter is used to determine the number of bins for the histograms associated to each window. We tested configurations from 7 to 17 bins (in increments of 2). For each point we have 5 different sub-windows with an associated histogram, so the total number of features range from $5 \times 7 = 35$ to $5 \times 17 = 85$.

**Intensity features** 13 features are obtained with all the intensity information of the window: *maximum, minimum, mean, median, standard deviation, variance, first quartile, third quartile, skewness* and *maximum likelihood estimate* (for a normal distribution).

**Gray-level Intensity Histogram (GLIH)** The histogram of the full window is calculated. From it, we obtain the following metrics: *obliquity, kurtosis, energy* and *entropy*.

**Gray-level Co-ocurrence Matrix (GLCM)** These features provide information about the spatial relationship of the pixels [15]. We use a distance of 2 pixels and 4 directions as proposed by Haralick *et al.* [6], for a total of 16 features.

**Histogram of Oriented Gradients (HOG)** Gradient orientation can be an useful feature, since it can contribute to the detection of the different patterns of gradients when the epiretinal membrane is found in contrast to its absence. Besides, HOG features are suitable to recognize gradient patterns in the ROI since they are invariant to scale, rotation or translation modifications. We used 9 HOG windows with 9 bins, obtaining a total of 81 features.

**Local Binary Patterns (LBP)** Local Binary Patterns also help to detect patterns of intensity changes in the selected window. Another advantage is their low sensitivity to intensity changes, since variation in illumination is usual in OCT images. We use a total of 64 features that will give a extended range of information.

### 2.3 Feature selection and classification

Once the feature set was specified, we proceeded with the analysis and selection of those most relevant features that contain meaningful data and provide the highest discriminative power. This way, we can optimize the classification process and obtain better results by avoiding the introduction of unneeded and redundant information to the classifiers.

Feature selection was performed using representative strategies. Correlation Feature Selection (CFS) algorithm [5], which works by selecting features correlated or predictive of the class that, otherwise, are irrelevant. Relief-F algorithm [9] is also used, consisting on the repeatedly sampling of a random instance and checking the distance to the nearest instances of the same or different class.

Finally, representative classifiers with proven utility in medical imaging applications were trained and tested using the selected feature subsets: K-Nearest Neighbor (kNN), Naive Bayes and Random Forest classifiers. For each classifier, we test three window sizes and six different number of bins for the window features mentioned beforehand.

## 3 Results and Discussion

The validation of the method was done by the use of a set of 129 OCT images. These scans were obtained with a tomograph CIRRUS$^{TM}$ HD-OCT Zeiss, with Spectral Domain Technology. The resolution of the scans was 490×500 pixels without any preprocessing stage.

The scans were labeled by an expert clinician, identifying the areas where ERM is present and absent, respectively. With this groundtruth, we selected a set of 120 samples, divided in 60 samples with ERM presence and 60 with ERM absence. Furthermore, each group of samples can also be split in the following classes (Fig. 2):

1. *Membrane* class. Points with ERM presence on top of the ILM.
2. *Floating membrane* class. Points with ERM presence on the background.
3. *Non-membrane* class. Points situated on the first layer of the retina but without ERM presence.
4. *Background* class. Points not situated on the ILM layer, but on the background.

This way, we have labeled datasets for two and four class approximations. We performed experiments facing both approximations to test the capabilities of the designed method. To evaluate the results, we use the accuracy of the classifiers as our control metric. Table 1a and Table 1b present the results obtained using the two-class approximation. Results are very similar across all configurations, only obtaining a slight improvement with $W_{\text{size}} = 17$ for the Random Forest classifier.

Table 2a and Table 2b show the results for the 4 class approximation. In this case, the improvement in the performance is more accentuated when using

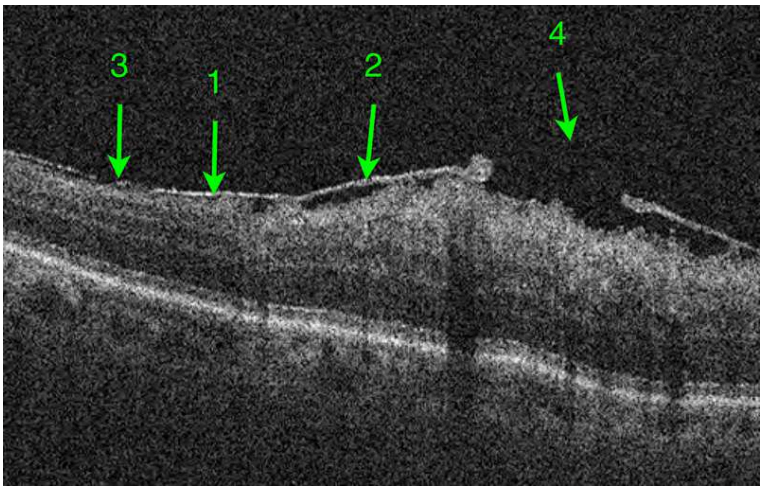



Fig. 2: Example of the different types of classes to be identified

Table 1: 2-class classification accuracy results

(a) CFS algorithm

| | **kNN** | | | **Naive Bayes** | | | **Random Forest** | | |
|---|---|---|---|---|---|---|---|---|---|
| | $W_{\text{size}}$ | | | $W_{\text{size}}$ | | | $W_{\text{size}}$ | | |
| $N_{\text{bins}}$ | 13 | 15 | 17 | 13 | 15 | 17 | 13 | 15 | 17 |
| 7 | 81.80% | 88.68% | 87.44% | 81.58% | 89.71% | 86.75% | 83.79% | 89.46% | 90.45% |
| 9 | 82.80% | 87.46% | 87.24% | 80.85% | 88.74% | 87.26% | 83.58% | 89.22% | 89.99% |
| 11 | 84.22% | 85.98% | 88.95% | 81.09% | 88.73% | 85.78% | 83.81% | 89.93% | 89.73% |
| 13 | 81.56% | 86.19% | 87.46% | 80.61% | 89.23% | 85.29% | 83.57% | 88.72% | **90.48%** |
| 15 | 81.29% | 87.41% | 85.48% | 81.09% | 89.70% | 86.05% | 82.10% | 88.47% | 90.23% |
| 17 | 84.51% | 88.19% | 86.67% | 80.61% | 88.74% | 85.54% | 82.10% | 90.21% | 90.22% |

(b) Relief algorithm

| | **kNN** | | | **Naive Bayes** | | | **Random Forest** | | |
|---|---|---|---|---|---|---|---|---|---|
| | $W_{\text{size}}$ | | | $W_{\text{size}}$ | | | $W_{\text{size}}$ | | |
| $N_{\text{bins}}$ | 13 | 15 | 17 | 13 | 15 | 17 | 13 | 15 | 17 |
| 7 | 81.14% | 82.30% | 85.73% | 83.10% | 82.34% | 85.49% | 81.88% | 84.74% | 86.00% |
| 9 | 81.60% | 82.04% | **89.22%** | 84.30% | 82.08% | 86.29% | 82.82% | 82.54% | 86.02% |
| 11 | 81.65% | 81.05% | 87.48% | 84.55% | 82.08% | 86.24% | 82.64% | 84.00% | 86.53% |
| 13 | 81.42% | 85.76% | 88.23% | 83.84% | 84.81% | 87.01% | 84.57% | 85.51% | 86.77% |
| 15 | 83.32% | 84.79% | 84.29% | 84.33% | 82.85% | 84.54% | 84.80% | 85.77% | 83.83% |
| 17 | 85.82% | 85.78% | 88.48% | 85.53% | 84.34% | 87.04% | 84.84% | 85.54% | 85.07% |

CFS algorithm, both with kNN and Random Forest classifiers for $W_{\text{size}} = 15$. Generally, the best accuracy is found for $W_{\text{size}} = 15$, while $N_{\text{bins}}$ is a parameter with more discrepancy. Normally, extreme values on this parameter decrease the

overall accuracy of the classifier. The best results are obtained for the 4 class approach, both with kNN and Random Forest classifiers, reaching an accuracy of 93.89% when using CFS algorithm for feature selection and Random Forest. Random Forest algorithm obtains a slightly higher accuracy than the kNN method. 28 features were selected for this configuration.

Table 2: 4-class classification accuracy results

(a) CFS algorithm

| | **kNN** | | | **Naive Bayes** | | | **Random Forest** | | |
|---|---|---|---|---|---|---|---|---|---|
| | $W_{\text{size}}$ | | | $W_{\text{size}}$ | | | $W_{\text{size}}$ | | |
| $N_{\text{bins}}$ | 13 | 15 | 17 | 13 | 15 | 17 | 13 | 15 | 17 |
| 7 | 83.85% | 89.77% | 88.08% | 86.60% | 86.08% | 88.77% | 88.53% | 92.47% | 88.55% |
| 9 | 85.34% | 91.20% | 88.32% | 86.84% | 89.23% | 90.75% | 85.86% | 93.67% | 87.57% |
| 11 | 85.33% | 92.17% | 88.79% | 88.11% | 88.04% | 90.23% | 86.85% | 91.96% | 89.48% |
| 13 | 85.37% | 89.24% | 88.05% | 85.93% | 88.05% | 89.53% | 88.09% | **93.89%** | 89.25% |
| 15 | 88.31% | 91.45% | 88.08% | 85.64% | 87.51% | 88.02% | 87.34% | 91.71% | 89.00% |
| 17 | 86.14% | 91.34% | 88.28% | 86.61% | 87.55% | 89.98% | 86.63% | 91.94% | 89.24% |

(b) Relief algorithm

| | **kNN** | | | **Naive Bayes** | | | **Random Forest** | | |
|---|---|---|---|---|---|---|---|---|---|
| | $W_{\text{size}}$ | | | $W_{\text{size}}$ | | | $W_{\text{size}}$ | | |
| $N_{\text{bins}}$ | 13 | 15 | 17 | 13 | 15 | 17 | 13 | 15 | 17 |
| 7 | 88.29% | 86.54% | 87.04% | 86.58% | 85.11% | 83.88% | **92.19%** | 90.43% | 89.22% |
| 9 | 87.55% | 86.55% | 86.09% | 85.61% | 83.38% | 82.68% | 90.24% | 88.49% | 88.74% |
| 11 | 87.56% | 87.81% | 87.06% | 87.04% | 83.66% | 83.17% | 88.78% | 87.29% | 88.73% |
| 13 | 90.98% | 87.54% | 88.32% | 85.87% | 84.65% | 82.45% | 89.97% | 88.02% | 89.48% |
| 15 | 89.02% | 87.03% | 87.28% | 87.80% | 84.88% | 83.16% | 90.21% | 88.97% | 87.29% |
| 17 | 88.79% | 85.83% | 85.86% | 86.57% | 84.15% | 84.13% | 89.48% | 88.73% | 89.97% |

Regarding feature selection, we present the results using the 4-class approximation. The analysis and conclusions are analogous to the other 2-class approximation. About the selected features, Fig. 3 shows the ones provided by each selector in their best configuration, for a total of 28, grouped by type. The most relevant features that were provided by both strategies belong to the window features. More precisely, information about the first bins on the third sub-window (this is, the central sub-window, which is the one where the ERM should be located) were included in the first positions. As we can observe, these features are selected because the core of the class differentiation is done by the use of information of the center of the window (luminosity values indicate presence or absence of ERM). In a lower degree, it is also relevant the information about the fourth sub-window (the second from the bottom) and the fifth sub-window (the bottom one). This is congruent with the theory that the information under the central sub-window contributes to the differentiation between *floating membrane*

and *membrane* classes, the first having lower intensity values on those windows than the *membrane* class. The rest of the selected features are mostly HOG values and a few features from the LBP group with the CFS selection, as they provide higher information about intensity and patterns on the ROI and contribute to improve the discrimination between the ERM presence or absence. On the contrary, this information is better represented by mainly GLCM features using the Relief-F selected subset.

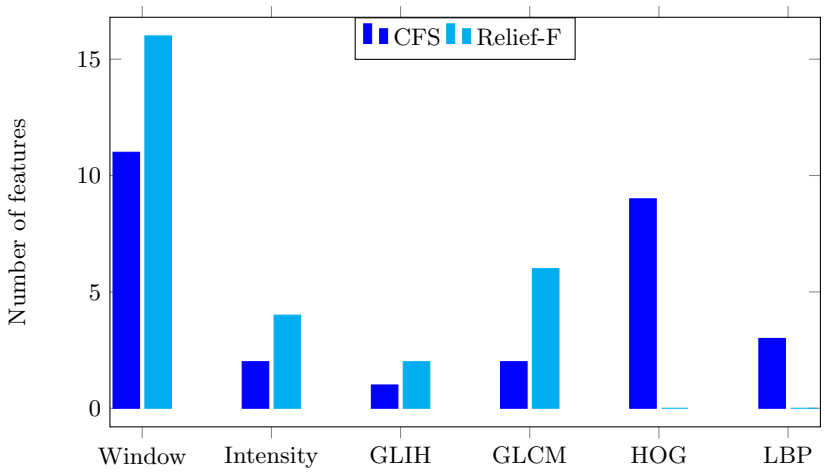



Fig. 3: Number of selected features, of a total of 28, for each feature selector method

Figure 4 shows the accuracy of the classification stage with the most accurate configuration (Random Forest with $W_{\text{size}} = 15$ and $W_{\text{size}} = 13$) when using progressive larger subsets of features with both selector strategies. In almost all situations, CFS selector performs better than the Relief-F selector.

Figure 5 presents a representative classification result of an OCT image using the most accurate configuration. As we can see, most of the points are classified correctly and the ERM presence is detected in almost all the ILM surface. Differentiation between the ERM with or without separation from the ILM is also done correctly.

In contrast, Fig. 6 show a common incorrect classification on the right points of the image. In this case, the Snake algorithm cannot be locally adjusted to the lower zone of the retina, being penalized the final adjustment and the points get detected at background positions.

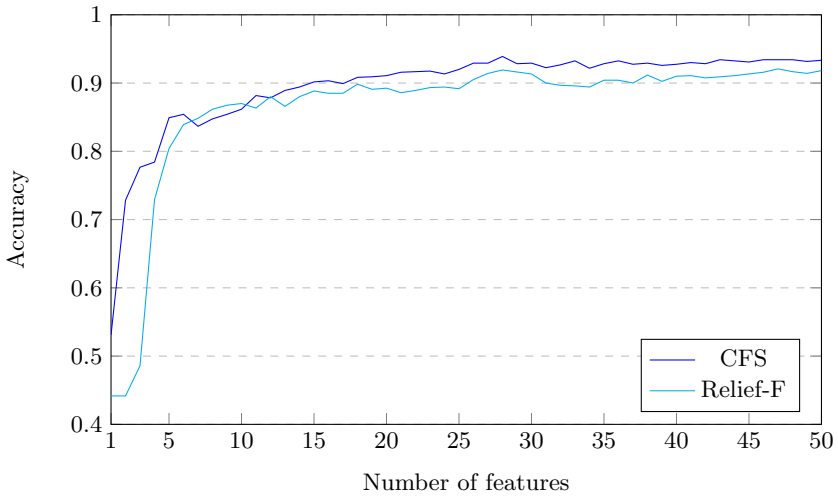



Fig. 4: Evolution of accuracy using Random Forest classifier and progressive larger feature subsets with the analyzed feature selectors

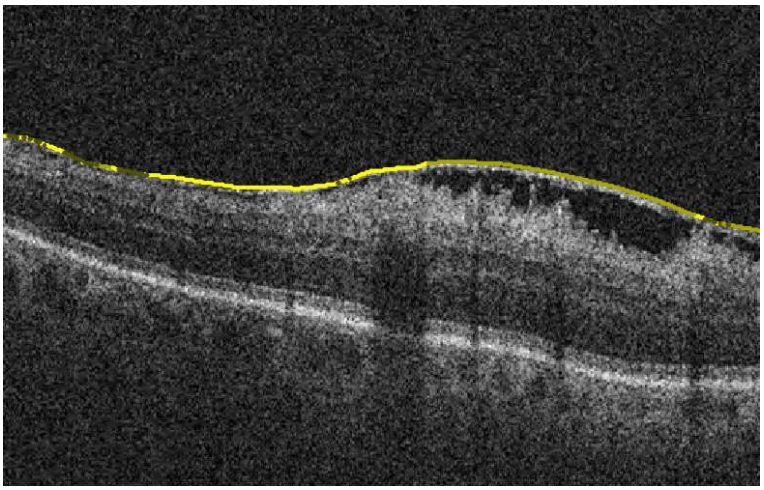


Fig. 5: Example of 4-class classification of an OCT image. Bright points represent the ERM on the ILM layer. Medium-intensity points symbolize ERM separated from the retina. Dark points show absence of epiretinal membrane

## 4 Conclusions

The accurate identification of the presence of the ERM is an important issue in the retinal analysis as its early detection improves the chances of success of ERM removal surgery, avoiding the complications that its presence derive.

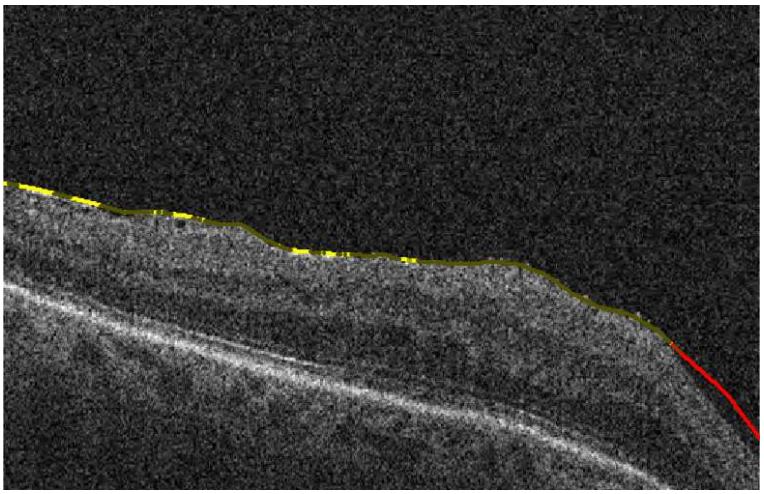


Fig. 6: Example of a region incorrectly identified and classified. Medium-intensity points on the right side symbolize background points, not contributing to the task at hand

In this work, we proposed an automatic method to detect the ERM in OCT images. The method is fully automatic, instead of the few previous approaches that are based on manual detections by the specialist. Furthermore, we have achieved a higher level of tolerance to errors by using a deformable model to detect the ROI compared to the detection of this region based on manual markers.

The method firstly uses a deformable model (Snake) to initially identify the ILM retinal layer, ROI where the ERM is originated. Then, a complete and heterogeneous set of features were measured, based on the properties of the ERM. Representative feature selectors as CFS and Relief-F were used with the entire feature set to identify those that provide the highest discriminative power.

We defined a set of 223-263 features that were then filtered by a process of feature selection, obtaining 28 features with the method that provided the most accuracy afterwards (CFS). Different suitable classifiers were tested, for a total of 216 different configurations. For testing, we used a set of 120 samples, distributed equally between the different classes in both two and four class approximations.

The results were highly successful, obtaining an accuracy of 93.89% for the CFS algorithm, with a Random Forest classifier with $W_{\text{size}}$ of 15 and a $N_{\text{bins}}$ of 13.

For further works, an increase in the number of samples for training is planned in order to improve even further the accuracy of the classifiers. Furthermore, wrapper-based feature selection methods will be tested as well as a larger variability of classifiers.

## References


1. Agarwal, A.: Gass Atlas of Macular Diseases. Elsevier Health Sciences. (2011).
2. Brancato, R.: Optical coherence tomography (OCT) in macular edema. Documenta Ophthalmologica 97, 337–339 (1999).
3. Do, D.V., Cho, M., Nguyen, Q.D., Shah, S.M., Handa, J.T., Campochiaro, P.A. et al.: Impact of Optical Coherence Tomography on Surgical Decision Making for Epiretinal Membranes and Vitreomacular Traction. Retina 27, 552–556 (2007).
4. Foos, R.Y.: Vitreoretinal juncture; epiretinal membranes and vitreous. Invest. Ophthalmol. Vis. Sci. 16, 416–422 (1977).
5. Hall, M.A.: Correlation-based feature selection for machine learning. The University of Waikato (1999).
6. Haralick, R.M., Shanmugam, K., others: Textural features for image classification. IEEE Transactions on systems, man, and cybernetics 3, 610–621 (1973).
7. Ikram, M.K., de Jong, F.J., Bos, M.J., Vingerling, J.R., Hofman, A., Koudstaal, P.J. et al.: Retinal vessel diameters and risk of stroke: the Rotterdam Study. Neurology 66, 1339–1343 (2006).
8. Kass, M., Witkin, A., Terzopoulos, D.: Snakes: Active contour models. Int J Comput Vision 1, 321–331 (1988).
9. Kira, K., Rendell, L.A.: A practical approach to feature selection. Proceedings of the ninth international workshop on Machine learning 249–256 (1992).
10. Medina, C.A., Townsend, J.H., Singh, A.D. (Eds.): Manual of Retinal Diseases. Springer International Publishing (2016).
11. Meuer, S.M., Myers, C.E., Klein, B.E.K., Swift, M.K., Huang, Y., Gangaputra, S. et al.: The Epidemiology of Vitreoretinal Interface Abnormalities as Detected by SD-OCT: the Beaver Dam Eye Study. Ophthalmology 122, 787–795 (2015).
12. Moura, J. de, Novo, J., Ortega, M., Charln, P.: 3D Retinal Vessel Tree Segmentation and Reconstruction with OCT Images. International Conference Image Analysis and Recognition, Springer, Cham, 716–726 (2016).
13. Novo, J., Penedo, M.G., Santos, J.: Optic Disc Segmentation by Means of GA-Optimized Topological Active Nets. International Conference Image Analysis and Recognition, Springer, Berlin, Heidelberg, 807–816 (2008).
14. Puliafito, C.A., Hee, M.R., Lin, C.P., Reichel, E., Schuman, J.S., Duker, J.S. et al.: Imaging of Macular Diseases with Optical Coherence Tomography. Ophthalmology 102, 217–229 (1995).
15. Ramamurthy, B., Chandran, K.R.: Content based medical image retrieval with texture content using gray level co-occurrence matrix and k-means clustering algorithms. Journal of Computer Science 8, 1070 (2012).
16. Wilkins, J.R., Puliafito, C.A., Hee, M.R., Duker, J.S., Reichel, E., Coker, J.G. et al.: Characterization of Epiretinal Membranes Using Optical Coherence Tomography. Ophthalmology 103, 2142-2151 (1996).
17. Wong, T.Y., Klein, R., Sharrett, A.R., Schmidt, M.I., Pankow, J.S. et al.: Retinal arteriolar narrowing and risk of diabetes mellitus in middle-aged persons. JAMA 287, 2528–2533 (2002).