Impact of the Region of Analysis on the Performance of the Automatic Epiretinal Membrane Segmentation in OCT Images

Mateo Gende[1,2][0000−0003−1686−7189], Daniel Iglesias Morís[1,2][0000−0002−5239−0421], Joaquim de Moura[1,2*][0000−0002−2050−3786], Jorge Novo[1,2][0000−0002−0125−3064], and Marcos Ortega[1,2][0000−0002−2798−0788]
{m.gende, daniel.iglesias.moris, joaquim.demoura, jnovo, mortega}@udc.es

[1] Grupo VARPA, Instituto de Investigación Biomédica de A Coruña (INIBIC), Universidade da Coruña, A Coruña, (Spain)
[2] Centro de investigación CITIC, Universidade da Coruña, A Coruña, (Spain)

Abstract. The Epiretinal Membrane (ERM) is an ocular pathology that can cause permanent visual loss if left untreated for long. Despite its transparency, it is possible to visualise the ERM in Optical Coherence Tomography (OCT) images. In this work, we present a study on the impact of the analysis region on the performance of an automatic ERM segmentation methodology using OCT images. For this purpose, we tested 5 different sliding windows sizes ranging from 14×14 to 224×224 pixels to calibrate the impact of the field of view under analysis. Furthermore, 3 different approaches are proposed to enable the analysis of the regions close to the edges of the images. The proposed approaches provided satisfactory results, with each of them interacting differently with the variations in window size.

Keywords: Computer-aided Diagnosis · Optical Coherence Tomography · Epiretinal Membrane · Segmentation · Deep Learning

1 Introduction

The Epiretinal Membrane (ERM) is an ocular pathology consisting of a thin fibrocellular layer of tissue that may appear over the retina. Although this film is mainly transparent, its development over the photosensitive tissue of the eye can lead to complications. Once it reaches a certain stage, the ERM may start to contract over itself. Since it is adhered to the tissue of the retina, it may exert a traction. This traction deforms the underlying tissue, producing wrinkles or puckers. If the ERM is present over the macula, the part of the eye that is responsible for sharp vision, the deformations it might induce can lead to visual distortion, a loss of vision and, eventually, may cause a macular hole. If treated on time, the ERM may be removed while preserving patient vision. Conversely, a late diagnosis may lead to irreversible deformations of the tissue and the consequent vision loss [12].

Optical Coherence Tomography (OCT) is an ocular imaging technique that uses low coherence light to sweep the tissue and obtain three-dimensional visualisations of the underlying histological structure [8, 10]. Despite being mostly transparent, the ERM can be visualised in OCT imaging. Its appearance is that of a bright film located in the boundary between the vitreous body and the retinal tissue, an area that is known as the Inner Limiting Membrane, or ILM. Due to the high visibility and the advantages that this imaging technique supposes in terms of cross-sectional visualisation, OCT is a standard imaging modality used for the detection and assessment of the ERM [5].

Typically, the ERM is diagnosed by an ophthalmologist visually inspecting each OCT slice. This process can become tiresome and repetitive, due to the great volume of images to analyse and may lead to subjectivity in the detection. Because of this, various works have approached the automatisation of this screening process. Wilkins *et al.* [14] first proposed a semi-automatic approach, involving an initial manual annotation made by an expert and a progressive automatic refinement of the affected area. More recently, some works have introduced fully automatic methods for ERM presence classification in OCT images. Sonobe *et al.* [13] trained a Convolutional Neural Network (CNN) to determine the pathology that each image presented, and compared the results with those obtained with Support Vector Machines (SVMs). In this comparison, the deep learning methods achieved much better results than the SVMs. Parra-Mora *et al.* [11] used four convolutional neural network architectures for the classification of OCT slices. These models achieved high performance in discriminating ERM cases.

On the other hand, few studies have addressed the segmentation of the ERM in OCT slices. As reference, Baamonde *et al.* proposed different ways to characterise the ERM in OCT images using conventional hand-crafted features [4, 3, 1]. Additionally, the authors present a multi-stage methodology for the conversion of a segmentation problem into a classification one by extracting a series of vertical image patches from the OCT slices [2]. This classification is performed by extracting and selecting a subset of relevant features and using classical machine learning techniques. In the work of Gende *et al.* [6], the authors presented a complete methodology in which a series of square windows is extracted around the ILM. These windows are then classified by a CNN, automatising the feature extraction and selection process by allowing the deep learning models to be trained directly from the image patches. This methodology was able to outperform the previous methods while also considerably simplifying the inference process. However, the effects of different window sizes that calibrate the field of view under analysis on this methodology has not been assessed. Furthermore, this methodology is limited in the amount of tissue it can analyse close to the image edges, since windows cannot be extracted outside of the image bounds, as illustrated in Fig. 1.

In this work, we present a study on the impact that the size of the extracted windows has on the performance of the classification models for the segmentation of the ERM. In this study, we compare the results obtained using five window

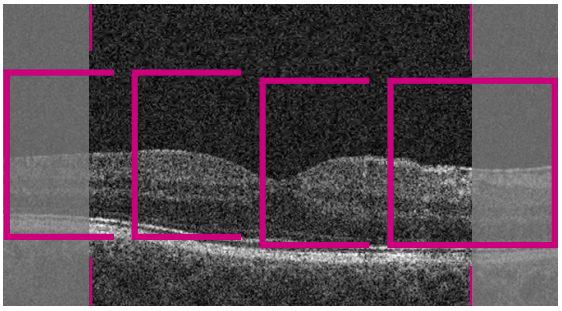

Fig. 1. Illustration of the limitations of the existing methodology concerning image boundaries. Those spots lying on the highlighted areas cannot be directly analysed.

sizes ranging from 14×14 to 224×224. Furthermore, we analyse 3 different approaches to overcome the limitations posed by the existing methodologies in terms of analysing the areas that are close to the image bounds: A basic, Zero Padding approach in which the part of the windows lying outside of the image bounds is filled with zeros; a Border Extension approach in which the final column is extended outwards and a Border Reflection approach in which the image is reflected along the vertical axis around the image border.

2 Methodology

2.1 ERM segmentation methodology

In order to segment the ERM in each OCT slice, we followed the three-step methodology proposed by Gende *et al.* [6]. The first step is to obtain the location of the ILM, the area that is susceptible to ERM proliferation. Next, a series of windows is extracted from the OCT image. Each of these windows is centred on an ILM pixel, and contains the visual information surrounding the point of interest. Finally, each window is classified by using a DenseNet-121 [9] classifier, returning a label for each ILM pixel indicating if the spot is healthy or diseased. This architecture has been used in similar medical image segmentation tasks with encouraging results [6, 7]. By accumulating all of these labels and assigning them to the original ILM pixels, a segmentation of the ERM can be produced, as illustrated in Fig. 2.

2.2 Impact of the region of analysis

The size of the extracted windows has an impact on the amount of visual information that is incorporated into the classification process, that is, the field of view that is particularly under analysis to determine each potentially pathological case. This in turn translates into an impact on the ability of the system to

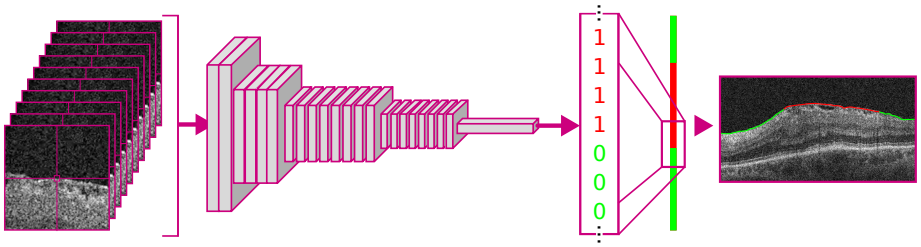

Fig. 2. Summary of the ERM segmentation methodology. Windows are extracted from the original OCT slice. These windows are classified by a CNN, and the resulting class labels are used to reconstruct a segmentation of the ERM.

discriminate between healthy and diseased spots. Additionally, in the window extraction process, pixels close to the edges of the image cannot be directly analysed, since the area surrounding each of these pixels would lie out of bounds. This constraint becomes more restrictive the more the window width is increased.

In order to study the impact that the field of view has on the performance of the ERM segmentation methodology, different configurations were tested and analysed. On the one hand, a series of DenseNet-121 models were trained using five different window sizes: 14×14, 28×28, 56×56, 112×112, and 224×224 pixels. Furthermore, 3 different approaches aimed at overcoming the limitations related to image boundaries were tested. A schematic representation of the proposed approaches can be seen in Fig. 3.

- **Zero Padding**: This baseline approach consists of filling the area of the window lying out of bounds with zeros. This way, no additional information is incorporated into the classification process. This approach creates a synthetic border at the point where the original image ends.
- **Border Extension**: The second approach consists of replicating the final row or column of original image pixels outwards until the edge of the window. This avoids generating a synthetic border at the edge of the image, but in turn may incorporate more noise in the window.
- **Border Reflection**: The final approach reflects the original image along the axis located at its edge. This way, the original look of the image is preserved, providing a more seamless transition where the original image border would be located and replicating the information already contained in the window.

3 Results and Discussion

The five window sizes that determine the fields of view under analysis were combined with the three approaches for a total of 15 different configurations. For each of these configurations, a series of CNNs were trained and evaluated. To train and validate these models, a dataset consisting of 2,427 OCT slices

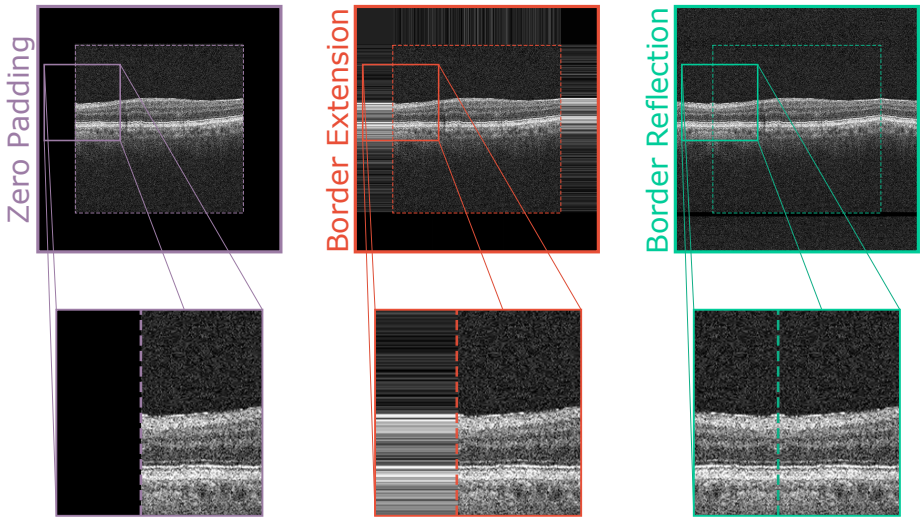

Fig. 3. Detailed viewts of each of the three proposed approaches that allow the analysis of pixels close to the image boundary.

belonging to 20 patients was used. Out of these 20 patients, 8 presented ERM signs while 12 were healthy. In each OCT image, ERM presence or absence was marked as a height value for each column. In total, 251,994 columns were annotated as pathological, out of a total of 1,308,160 image columns. The models were trained using a 4-fold cross-validation at the eye level, ensuring each eye appears in a test set once. Online augmentation was used in the form of random shear transformations, rotations, intensity variations and horizontal flipping. The loss employed was Cross-entropy, and Adam was used for optimisation. Training was allowed to run for a maximum of 75 epochs for each configuration, saving the model that performed best in validation for testing.

The results that were obtained by testing each of the proposed configurations can be found on Table 1. As for the proposed approaches that enable analysis close to the image edges, the Zero Padding approach is the most affected by window size, achieving Dice scores of 0.573 and 0.517 with the more extreme window sizes, comparatively worse than the 0.675 obtained with windows of the size 112×112. While the Border Extension approach decays in performance with larger windows, as is expected from the approach that has the biggest impact on the look of the patches, it outperforms the other two for the smaller window sizes of 14×14 and 28×28. The Border Reflection approach gives the overall best result out of the three proposed when combined with windows of size 56×56, with a Dice Coefficient of up to 0.699.

Complementarily, Fig. 4 shows an example of the segmentation results produced by models trained with the Border Reflection approach for different window sizes. As we can see, the obtained results show that intermediate window

Table 1. Test results for each of the proposed configurations. Bold indicates best results for each of the metrics.

	14x14	28x28	56x56	112x112	224x224
Zero Padding Approach					
Accuracy	0.821±0.039	0.863±0.019	0.869±0.016	**0.879±0.015**	0.859±0.050
Precision	0.526±0.131	0.633±0.135	0.621±0.122	**0.657±0.126**	0.647±0.156
Recall	0.660±0.243	0.692±0.201	**0.726±0.165**	0.709±0.075	0.565±0.324
Specificity	0.860±0.037	0.906±0.029	0.899±0.019	**0.920±0.011**	0.905±0.081
Jaccard	0.414±0.152	0.486±0.141	0.506±0.145	**0.513±0.088**	0.373±0.197
Dice	0.573±0.152	0.645±0.122	0.663±0.125	**0.675±0.075**	0.517±0.245
Border Extension Approach					
Accuracy	0.832±0.042	0.859±0.020	0.858±0.042	**0.870±0.022**	0.861±0.032
Precision	0.559±0.135	0.625±0.111	0.604±0.141	**0.682±0.098**	0.649±0.140
Recall	0.678±0.203	0.701±0.150	**0.703±0.186**	0.652±0.219	0.604±0.212
Specificity	0.871±0.033	0.898±0.024	0.891±0.023	**0.920±0.045**	0.915±0.051
Jaccard	0.440±0.148	0.490±0.111	**0.493±0.168**	0.481±0.130	0.434±0.145
Dice	0.601±0.137	**0.652±0.096**	0.647±0.153	0.641±0.122	0.594±0.142
Border Reflection Approach					
Accuracy	0.827±0.036	0.859±0.031	**0.882±0.017**	0.871±0.025	0.840±0.039
Precision	0.549±0.123	0.635±0.171	**0.697±0.142**	0.656±0.142	0.618±0.111
Recall	0.665±0.179	0.687±0.177	**0.744±0.172**	0.684±0.157	0.649±0.188
Specificity	0.869±0.019	0.906±0.032	**0.917±0.041**	0.917±0.008	0.888±0.067
Jaccard	0.429±0.131	0.485±0.148	**0.544±0.118**	0.505±0.144	0.440±0.079
Dice	0.592±0.123	0.644±0.128	**0.699±0.094**	0.662±0.128	0.608±0.077

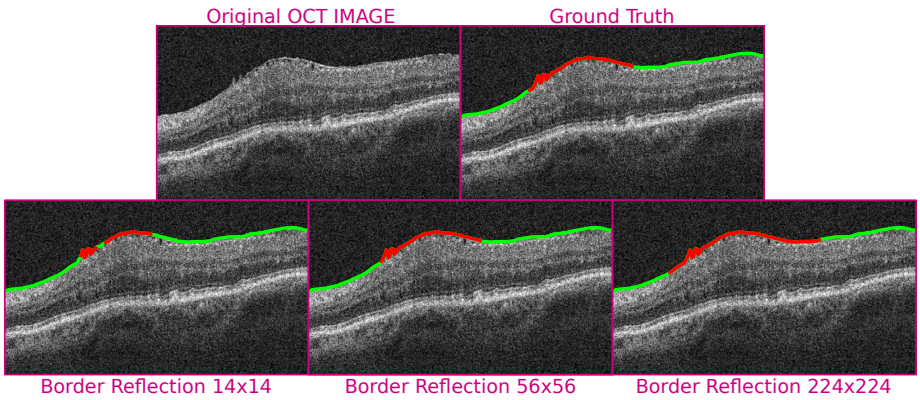

Fig. 4. Examples of the results obtained for different window sizes combined with the Border Reflection approach.

sizes tend to give better results than either of the extreme sizes. In addition, a smaller field of view tends to leave out relevant information for the classification, while a wider one may include visual features of regions unrelated to the patch on which the window is centred.

4 Conclusions

The early detection of the ERM in OCT images is paramount in order to preserve patient vision. In this work, we propose a series of improvements for the automatic segmentation of this relevant pathology in OCT slices. These consist in three different approaches that enable the analysis of the slices close to the image borders. Furthermore, we study the impact that five different window sizes have on the performance of the detection of the ERM and its combination with the three proposed approaches. These combinations were trained and validated on a representative dataset consisting of 2,427 OCT scans. The satisfactory results that were obtained show that intermediate window sizes provide the best performance, with the Zero Padding approach being most affected by either very small or very large windows. For the two smaller window sizes, the Border Extension approach performed better than the other two. Out of all the tested configurations, the Border Reflection approach combined with windows of size 56×56 provides the best results, achieving a Dice Coefficient of 0.699 ± 0.094, this way providing a robust and accurate segmentation of the ERM, contributing to the early diagnosis of this pathology.

Acknowledgements This research was funded by Instituto de Salud Carlos III, Government of Spain, [DTS18/00136]; Ministerio de Ciencia e Innovación y Universidades, Government of Spain, [RTI2018-095894-B-I00]; Ministerio de Ciencia e Innovación, Government of Spain through the research project [PID2019-108435RB-I00]; Consellería de Cultura, Educación e Universidade, Xunta de Galicia, Grupos de Referencia Competitiva, [ED431C 2020/24] and predoctoral grants [ED481A 2021/161] and [ED481A 2021/196]; Axencia Galega de Innovación (GAIN), Xunta de Galicia, grant ref. [IN845D 2020/38]; CITIC, Centro de Investigación de Galicia ref. [ED431G 2019/01] receives financial support from Consellería de Educación, Universidade e Formación Profesional, Xunta de Galicia through the ERDF (80%) and Secretaría Xeral de Universidades (20%).

References

1. Baamonde, S., de Moura, J., Novo, J., Charlón, P., Ortega, M.: Automatic identification and characterization of the epiretinal membrane in OCT images. Biomedical Optics Express **10**(8), 4018 (Jul 2019). https://doi.org/10.1364/boe.10.004018
2. Baamonde, S., de Moura, J., Novo, J., Charlón, P., Ortega, M.: Automatic identification and intuitive map representation of the epiretinal membrane presence in 3d OCT volumes. Sensors **19**(23), 5269 (Nov 2019). https://doi.org/10.3390/s19235269

3. Baamonde, S., de Moura, J., Novo, J., Ortega, M.: Automatic detection of epiretinal membrane in oct images by means of local luminosity patterns. In: Rojas, I., Joya, G., Catala, A. (eds.) Advances in Computational Intelligence. pp. 222–235. Springer International Publishing, Cham (2017)
4. Baamonde, S., de Moura, J., Novo, J., Rouco, J., Ortega, M.: Feature definition and selection for epiretinal membrane characterization in optical coherence tomography images. In: Image Analysis and Processing - ICIAP 2017, pp. 456–466. Springer International Publishing (2017). https://doi.org/10.1007/978-3-319-68548-9_42
5. Chua, P.Y., Sandinha, M.T., Steel, D.H.: Idiopathic epiretinal membrane: progression and timing of surgery. EYE **36**(3), 495–503 (Mar 2022). https://doi.org/https://doi.org/10.1038/s41433-021-01681-0
6. Gende, M., De Moura, J., Novo, J., Charlón, P., Ortega, M.: Automatic segmentation and intuitive visualisation of the epiretinal membrane in 3d oct images using deep convolutional approaches. IEEE Access **9**, 75993–76004 (2021). https://doi.org/10.1109/ACCESS.2021.3082638
7. Gende, M., de Moura, J., Novo, J., Ortega, M.: End-to-end multi-task learning approaches for the joint epiretinal membrane segmentation and screening in oct images. Computerized Medical Imaging and Graphics **98**, 102068 (Jun 2022). https://doi.org/10.1016/j.compmedimag.2022.102068
8. Huang, D., Swanson, E.A., Lin, C.P., Schuman, J.S., Stinson, W.G., Chang, W., Hee, M.R., Flotte, T., Gregory, K., Puliafito, C.A., Fujimoto, J.G.: Optical coherence tomography. Science **254**(5035), 1178–1181 (1991). https://doi.org/10.1126/science.1957169, https://www.science.org/doi/abs/10.1126/science.1957169
9. Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q.: Densely connected convolutional networks. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 2261–2269 (2017). https://doi.org/10.1109/CVPR.2017.243
10. de Moura, J., Novo, J., Charlón, P., Barreira, N., Ortega, M.: Enhanced visualization of the retinal vasculature using depth information in OCT. Medical & Biological Engineering & Computing **55**(12), 2209–2225 (2017)
11. Parra-Mora, E., Cazañas-Gordon, A., Proença, R., da Silva Cruz, L.A.: Epiretinal membrane detection in optical coherence tomography retinal images using deep learning. IEEE Access **9**, 99201–99219 (2021). https://doi.org/10.1109/ACCESS.2021.3095655
12. Rahman, R., Stephenson, J.: Early surgery for epiretinal membrane preserves more vision for patients. Eye **28**(4), 410–414 (Jan 2014). https://doi.org/10.1038/eye.2013.305
13. Sonobe, T., Tabuchi, H., Ohsugi, H., Masumoto, H., Ishitobi, N., Morita, S., Enno, H., Nagasato, D.: Comparison between support vector machine and deep learning, machine-learning technologies for detecting epiretinal membrane using 3d-OCT. International Ophthalmology **39**(8), 1871–1877 (Sep 2018). https://doi.org/10.1007/s10792-018-1016-x
14. 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**(12), 2142–2151 (Dec 1996). https://doi.org/10.1016/s0161-6420(96)30377-1