# Automatic Segmentation and Intuitive Visualisation of the Epiretinal Membrane in 3D OCT Images Using Deep Convolutional Approaches

**MATEO GENDE**[1,2], **JOAQUIM DE MOURA**[1,2], **JORGE NOVO**[1,2],
**PABLO CHARLÓN**[3,4], **AND MARCOS ORTEGA**[1,2]

[1]Centro de Investigación en Tecnologías de la Información y las Comunicaciones (CITIC), Universidade da Coruña, 15071 A Coruña, Spain
[2]Grupo VARPA, Instituto de Investigación Biomédica de A Coruña (INIBIC), Universidade da Coruña, 15006 A Coruña, Spain
[3]Instituto Oftalmológico Victoria de Rojas, 15009 A Coruña, Spain
[4]Hospital HM Rosaleda, 15701 Santiago de Compostela, Spain

Corresponding author: Joaquim de Moura (joaquim.demoura@udc.es)

**ABSTRACT** Epiretinal Membrane (ERM) is a disease caused by a thin layer of scar tissue that is formed on the surface of the retina. When this membrane appears over the macula, it can cause distorted or blurred vision. Although normally idiopathic, its presence can also be indicative of other pathologies such as diabetic macular edema or vitreous haemorrhage. ERM removal surgery can preserve more visual acuity the earlier it is performed. For this purpose, we present a fully automatic segmentation system that can help the clinicians to determine the ERM presence and location over the eye fundus using 3D Optical Coherence Tomography (OCT) volumes. The proposed system uses a convolutional neural network architecture to classify patches of the retina surface. All the 2D OCT slices of the 3D OCT volume of a patient are combined to produce an intuitive colour map over the 2D fundus reconstruction, providing a visual representation of the presence of ERM which therefore facilitates the diagnosis and treatment of this relevant eye disease. A total of 2.428 2D OCT slices obtained from 20 OCT 3D volumes was used in this work. To validate the designed methodology, several representative experiments were performed. We obtained satisfactory results with a Dice Coefficient of $0.826 \pm 0.112$ and a Jaccard Index of $0.714 \pm 0.155$, proving its applicability for diagnosis purposes. The proposed system also demonstrated its simplicity and competitive performance with respect to other state-of-the-art approaches.

**INDEX TERMS** Epiretinal membrane, machine learning, medical diagnostic imaging, optical coherence tomography.

## I. INTRODUCTION

Recent advances in computing power, together with the development of new algorithms based on artificial neural networks, are enabling enormous progress in the field of

The associate editor coordinating the review of this manuscript and approving it for publication was Jinhua Sheng.

the automatic analysis of medical imaging, as well as in the development of new and more efficient Computer-Aided Diagnosis (CAD) systems. As a result, deep learning architectures are now able to more efficiently extract relevant features from images and detect new computational biomarkers, among other capabilities. In particular, the ability of these architectures to isolate signal from noise in
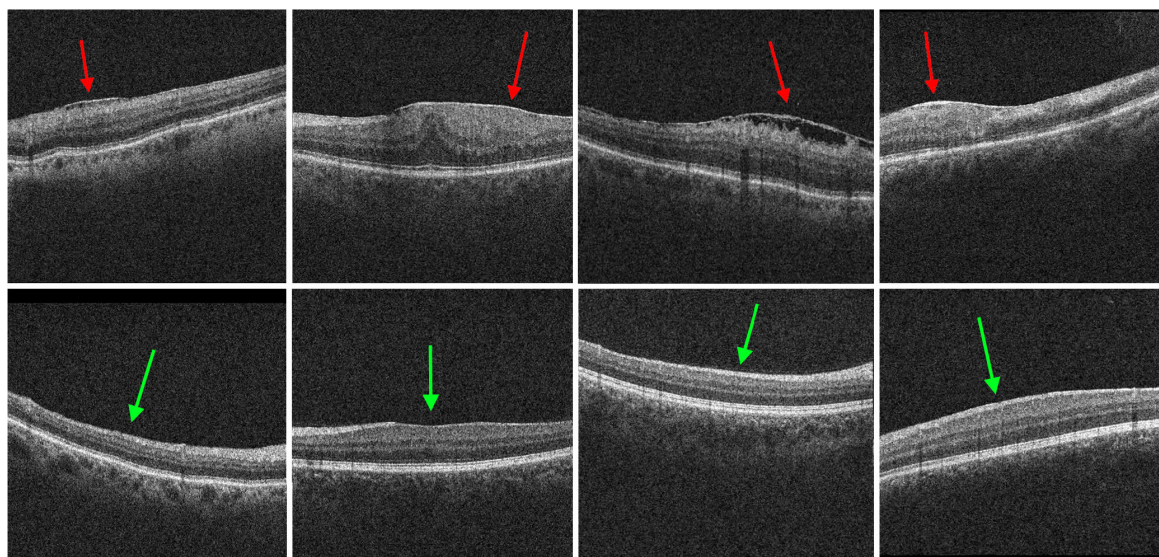
**FIGURE 1.** Examples of ERM presence and absence in OCT scans. 1$^{st}$ row, ERM appearing as a bright surface over the ILM. 2$^{nd}$ row, OCT scans without the presence of ERM.

the data makes them exceptionally valuable as a diagnostic tool [1], [2].

Traditionally, image-based CAD systems relied on hand-crafted object features, followed by a statistical classification designed specifically for the task [3]. This often required the development of multiple purpose-built classifiers, involving a long and arduous process of feature adjustment and evaluation. In contrast, the training process of convolutional neural networks performs this feature-extraction and selection process implicitly from annotated images. It works as an end-to-end system that can be trained directly from the data. In this way, the whole process can be simplified and generalised to other similar medical imaging domains.

In ophthalmology, Optical Coherence Tomography (OCT) is a non-invasive technique that can obtain high-resolution, volumetric digitalisations of eye tissues [4]. This makes OCT one of the most common medical imaging techniques when studying relevant retinal structures and diseases. In particular, OCT has successfully been used to study nervous tissue thickness [5] and angle-closure assessment [6] in glaucoma; to diagnose Age-Related Macular Degeneration (AMD or ARMD) [7], [8] and Diabetic Macular Edema (DME) [9], [10]; to assess the size of the optic nerve [11]; and to visualise the retinal layers [12], [13] and the retinal vasculature [14], [15], among others. Its ability to create *in vivo* visualisations of pathological structures that may be present in the retina makes it a valuable imaging technique for early diagnostic and treatment of a wide range of relevant eye diseases.

Similarly, OCT is playing an increasingly important role in the identification and monitoring of the Epiretinal Membrane (ERM), also known as macular pucker. ERM is a reflective, fibrocellular layer that develops over the Inner Limiting Membrane (ILM) in the eye macula. When an ERM forms, it may start to harden and contract, distorting

the macula and resulting in blurred vision or metamorphopsia [16], manifesting itself as a bowing when observing regular, parallel lines. Its most common treatment involves peeling away the ERM via pars plana vitrectomy [17]. This surgery preserves more vision for patients the earlier it is performed [18].

The ERM, although usually idiopathic, can also appear as a secondary factor to other related eye diseases, such as Posterior Vitreous Detachment (PVD) [19] or DME. This is usually caused by an immune system response to protect the retina. Other factors such as old age or ocular surgeries also play a part in its appearance [20]. In particular, DME has a high prevalence of related ERM appearance [21]–[23], with between 27% [24] and 34% [25] of DME patients presenting ERM. Its symptoms include an abnormal accumulation of fluid in the macular region as a result of leaking blood vessels. This results in a thickening of the macular region of the retina and a loss of vision. Consequently, DME is considered the main cause of visual impairment in patients with diabetes mellitus [26], representing a concerning public health problem.

In OCT imaging, the ERM appears as a thin, bright layer located in the boundary between the ILM layer and the vitreous. This high reflectivity allows it to be differentiated from healthy retinal tissues. Additionally, in advanced stages, its traction over the eye tissue tends to form additional *"wrinkles"* or *puckers* under the membrane that may also aid the detection. This makes OCT an effective modality for the diagnosis and detection of the ERM. Fig. 1 shows some examples of how the ERM appears in OCT scans. Red arrows indicate the presence of the ERM, a reflective layer on the top of the ILM, while green shows healthy tissues.

Currently, the diagnosis of ERM is determined by visual inspection of OCT scans. This visual inspection is not only a tedious and time-consuming routine, but it is also prone to subjectivity on the part of clinical specialists. Therefore, this

necessity for early diagnosis, combined with the requirement to accurately localise the pathological regions in order that the surgeon can remove them, motivates the search for automated methodologies that can make more accurate and reproducible identifications, thus helping clinical specialists to make more precise diagnoses and appropriate treatments.

Due to the relevance of this ocular pathology, different authors have approached its analysis in recent years, proposing some methodologies that use OCT images as a source of information. In this line, previous semi-automatic methods for ERM detection were usually based in manual identification of the ERM by a clinician visually inspecting the OCT images [4], [27], [28]. In particular, Wilkins *et al.* [29] proposed a semi-automatic method based around an initial manual annotation of the ERM, using an estimation of its reflectivity to locate the rest of the ERM regions. This, however, relies on the expert to diagnose and mark an approximate preliminary location of the ocular region affected by the presence of ERM.

Other relevant recent studies address the problem of ERM detection by the automatic analysis of OCT scans. In [30], the authors proposed a methodology to identify the ILM layer and then analyse the neighbouring region and determine the presence or absence of the ERM disease using a classical machine learning strategy. In the work of Lo *et al.* [31], the authors used a ResNet architecture for the screening of ERM in cross-sectional OCT images. Kuwayama *et al.* [32] proposed a deep learning strategy with image augmentation for automated detection of different macular diseases, including ERM. In [33], the authors compared Support Vector Machines (SVMs) and deep learning methods for detecting ERM cases in OCT macular surface volumes, with deep learning clearly outperforming SVMs.

As we can observe from the literature, the presented methods only considered the problem of automatic identification, a problem that is easier to address than precise segmentation. In this regard, only Baamonde *et al.* [34] presented a fully automatic method for ERM segmentation by performing a sliding window classification over the ILM layer, determining the presence or not of the ERM disease in 3D OCT volumes. In particular, the authors added a post-processing stage over the classification results to integrate information from the surrounding layers into the segmentation maps. In addition, the classification stage is done by means of a series of feature sets such as domain-related window features, histograms of oriented gradients, Laws features and local binary patterns. A subset of these features is selected by using the Spatial Uniform ReliefF algorithm. The classification stage is performed by using SVM although Random Forest and k-Nearest Neighbours were also tested and compared, for overall similar results. This process involves the extraction of 452 features for each point of interest over the ILM, a statistical analysis of the relevance of each feature and a comparison of the classical classification methods and their combinations with different feature selector strategies.

In this work, we propose a fully automatic methodology for the segmentation and characterisation of ERM regions using 3D OCT volumes. For this purpose, we use a Convolutional Neural Network (CNN) architecture, given its simplicity, computational performance and adequate results for many relevant image segmentation issues. By training this neural network directly from annotated images, the feature extraction and selection process is automatically performed, contributing to simplify the development of the system. In a similar manner to [34], a sliding window is passed over the barrier between the retina and the vitreous. This way, a CNN architecture classifies each point based on the centre pixel of the window, providing a precise segmentation of each OCT scan. Finally, all the scans are combined to produce an intuitive representation of the pathological regions, providing a map of the precise location of the ERM on the retina that facilitates the posterior analysis and diagnosis of the patients by the clinical experts.

This paper is organised in the following structure: section II will describe the materials used to develop the system, with subsection II-D dedicated to detailing the main characteristics of the proposed methodology. Section III introduces the discussion of the experimental results. Finally, section IV presents the conclusions that were reached and describes possible future lines of work.

## II. MATERIALS AND METHODS

In this section, we will proceed to describe in detail the required materials and resources for the implementation and full reproduction of this work. In particular, the reader can see the information of the used dataset (subsection II-A), as well as the different software and hardware resources (subsection II-B). In subsection II-D an explanation of the main characteristics of the methodology is presented.

### A. DATASET

In this work, we used a total of 20 OCT 3D volumes obtained from 20 different patients. The images ranged in size from $479 \times 501$ to $693 \times 676$ pixels for a total of 2.428 2D OCT scans, 892 being cases of ERM presence and 1.536 cases of ERM absence. These images were acquired using a CIRRUS$^{TM}$ HD-OCT Carl Zeiss Meditec confocal scanning laser ophthalmoscope. In addition, all the OCT scans were manually labelled by an expert, identifying the presence or absence of ERM cases and serving as ground truth for the entire segmentation process. This study was performed in accordance with the Declaration of Helsinki, approved by the local Ethics Committee of Investigation from A Coruña/Ferrol (2014/437) the $24^{th}$ of November, 2014.

### B. SOFTWARE AND HARDWARE RESOURCES

The programming language chosen to develop this system was Python (version 3.7.7) due to its flexibility and access to the PyTorch [35] machine learning library (version 1.7.1), which was used as the framework with which the neural networks were trained and validated. Additionally,
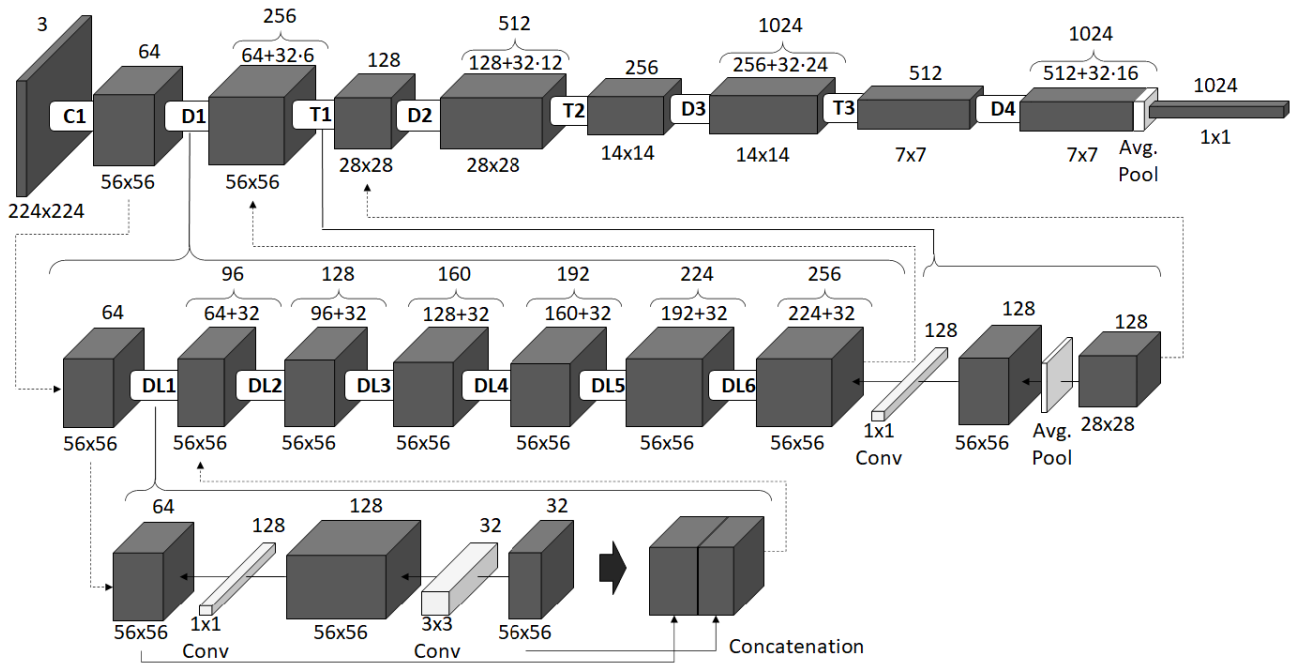
**FIGURE 2.** DenseNet-121 architecture summary. 1$^{st}$ row shows the whole model. 2$^{nd}$ row illustrates the structure of every dense block and transition layer. 3$^{rd}$ row details how every dense layer works by performing two convolutions and concatenating the input and the results.

OpenCV (version 3.4.8) and NumPy (version 1.15.0) libraries were used for all the image processing requirements, while Scikit-Learn (version 0.22.1) [36] was used to calculate the evaluation metrics. The training process was carried out on a computer that consisted of an Intel Xeon E5-2640 CPU, an NVIDIA GeForce GTX TITAN X GPU and 64 GB of RAM.

### C. NETWORK ARCHITECTURE

In this work, we used a Densely Connected Convolutional Network (DenseNet) [37] architecture to perform the classification stage. DenseNet has been used extensively in medical imaging for classification, segmentation and detection tasks, surpassing other current CNN architectures [38]–[40] in many issues. A schematised summary of the components that make up the neural network can be found in Fig. 2.

### D. METHODOLOGY

The proposed method receives, as input, a 3D OCT volume and returns an intuitive representation of ERM presence in the macular region. The designed pipeline is comprised of a set of 3 main steps, as illustrated in Fig. 3. The first step is to determine the ILM layer, the boundary between the retina and the vitreous. The next step is a sliding window classification of each pixel of the ILM layer into pathological or healthy. Finally, we generate 2D complete reconstructions of the retinal surface that integrate the ERM presence information from 3D OCT volumes. Each of these steps is explained in more detail in the following subsections.

### 1) ILM SEGMENTATION

The first step is to determine the region of interest (ROI) where the classification will take place. As stated before, ERM manifests as a scar tissue on the surface of the retina around the macular area. Thus, it appears just between the retina and the vitreous, over the ILM layer. To establish the location of the ILM layer in the OCT image, we used a strategy based on the active contour (Snake) model [41]. The main idea is that the active contour model deforms and adapts its shape to the points that minimise its energy. In particular, in the context of OCT imaging, depth can be represented as a vertical axis, and the position of the ILM can be determined as a height value for each image column. Therefore, by initialising the Snake at the top of the OCT image and allowing it to contract downwards, it will eventually converge at the ILM layer, provided there is a boundary for it to stop. This boundary is obtained by using the Sobel operator [42] over the image. This way, points that maximise gradient, minimise the internal energy of the Snake as it settles over them. By using this technique, we were able to accurately acquire the position of each ILM pixel for every OCT image. With this position, it is possible to obtain a sample of the surrounding tissue and perform a classification. This process is summarised in Fig. 4.

### 2) CLASSIFICATION STAGE

As mentioned before, a DenseNet-121 architecture was used to perform the ERM segmentation. In particular, this is achieved by classifying each ILM pixel segmented in the previous step into either pathological or healthy by using their surrounding area. For this purpose, a window size of
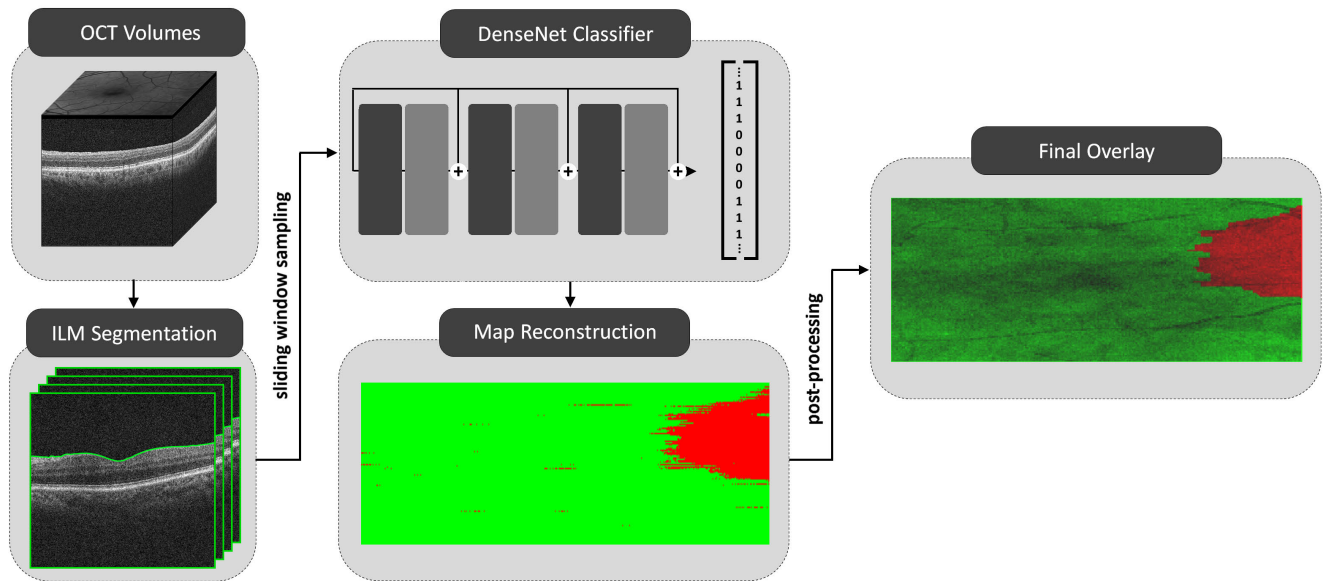
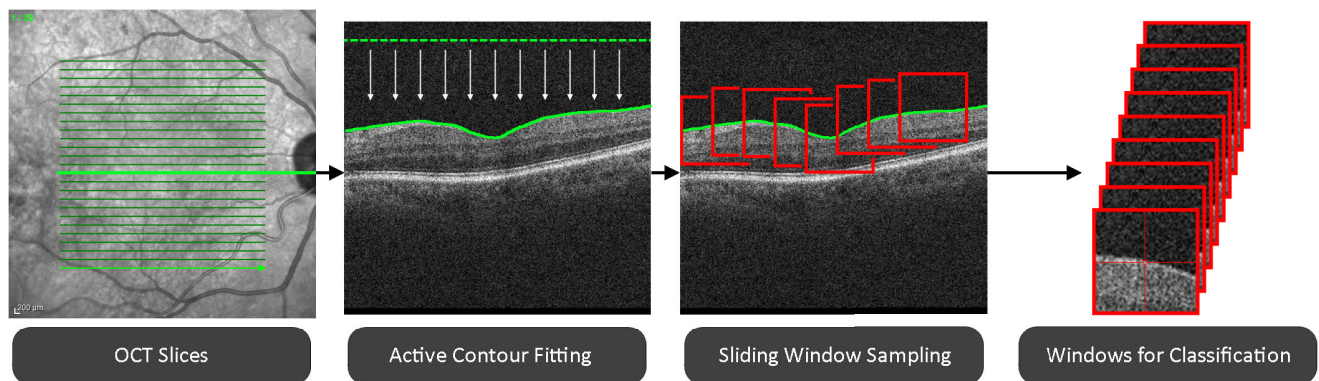**FIGURE 3.** A general overview of the proposed methodology.



**FIGURE 4.** First step of the ERM segmentation. An active contour is fitted to every slice in the OCT volume, segmenting the ILM. Next, a window is extracted, centered around every ILM pixel. These windows are classified in the next step.

$112 \times 112$ pixels was used as a compromise to incorporate surrounding visual information in each pixel classification while being able to accurately analyse regions close to the edges of the images. Each one of these windows is up-scaled to $224 \times 224$ pixels and then classified. Its corresponding class is then assigned to the original pixel, as shown in Fig. 5. This way, context information is integrated into the classification of each membrane point. This context information is crucial to distinguish isolated bright spots due to the presence of artefacts in data acquisition from real pathological regions.

In terms of training details, a pre-trained version of DenseNet-121 using ImageNet [43] was used in this work. Its final layer was replaced with a two-class fully-connected layer and it was then trained with available OCT data. In addition, as mentioned above, the data consisted of 8 diseased eyes and 12 healthy ones, because of this, a 4-fold cross-validation method was followed for the training. In each data partition, half the eyes were used as training examples,

a quarter for validation and another quarter for testing. This way, each healthy and pathological eye is represented in the testing stage.

Regarding the number of samples, we performed 3 preliminary tests. These consisted in sampling 10 equispaced windows from each diseased eye and 2 from healthy ones, 20 from diseased and 4 from healthy eyes and, lastly, 40 and 8 respectively. This was done to preserve a balance between the two classes and avoid over-representing healthy samples, as ERM-diseased eyes also display healthy tissue. In addition, to increase the effective number of samples that are available in the OCT dataset and to further improve the training, data augmentation techniques have been applied. In particular, window samples were further augmented using vertical and horizontal shear transformations, rotations, horizontal flips and brightness changes. Borders are replicated in a mirror-like manner to preserve continuity and prevent black edges from distorting the original appearance.
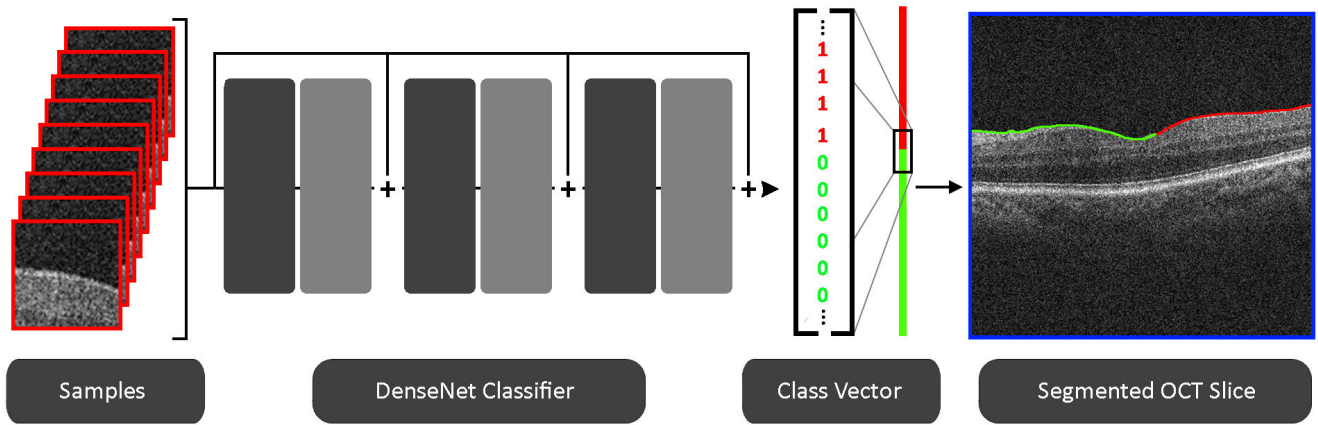
**FIGURE 5.** Second step of the ERM segmentation. DenseNet-121 is used to classify each of the 112 × 112 windows in pathological or healthy. Each of these classes are assigned to the centre pixel of every window.
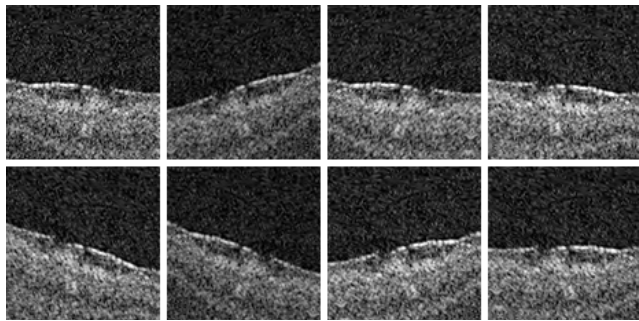


**FIGURE 6.** Example of augmented variations on a single diseased spot sample. Each transformation is a combination of shears, rotations, flips and brightness changes.

A representative example of the data augmentation applied in this work can be seen in Fig. 6.

The DenseNet-121 architecture was trained using Adaptive moment estimation (Adam) [44], configured with a learning rate of $1 \times 10^{-5}$, $\beta_1 = 0.9$ and $\beta_2 = 0.999$. In addition, training employed 4 samples per batch and was allowed to run for a maximum of 100 epochs for each network. Table 1 shows a summary of the chosen values for each hyper-parameter.

**TABLE 1.** Summary of sampling conditions and hyper parameters used in training.

| Window Conditions | | Training Conditions | |
| --- | --- | --- | --- |
| | | Batch size | 4 |
| Window Size | 112 × 112 | Epochs | 100 |
| Shear angle | [-15°,15°] | Optimiser | Adam |
| Rotation angle | [-15°,15°] | Learning rate | $1 \times 10^{-5}$ |
| | | $\beta_1, \beta_2$ | 0.9, 0.999 |

### 3) 2D COLOUR MAP RECONSTRUCTION AND POST-PROCESSING MAP REFINEMENT

Once each pixel of the retinal surface has been classified as healthy or pathological, we can generate an intuitive 2D colour map representing the regions affected by this relevant eye pathology. For this purpose, a 2D representation of the eye fundus must first be generated from the information provided by the 3D OCT volume. To do this, the average intensity of each column is calculated for every OCT slice. Subsequently, the information from all the slices is integrated to generate a single aerial view of the macular region, as illustrated in Fig. 7.

Once the 2D fundus image has been generated, we can create a 2D colour map to represent the analysed regions. To do so, each 1D array of classified pixels is taken from each 2D OCT slice and combined with all the slices from the same 3D OCT volume, generating a preliminary 2D colour map reconstruction (see Fig. 8).

On the other hand, the ERM is mostly an uniform and continuous pathological region on the surface of the retina, while the classification process can, in some cases, label pixels without taking into account information from their nearest neighbours, generating very small separate regions. To address this, a post-processing map refinement step is applied to clean the preliminary 2D colour map from isolated pixels that do not belong to the ERM (see Fig. 8). It consists of a contour filtering that rejects any separated blobs under 25 pixels and a morphological opening using a 10 × 4 pixels rectangular structuring element, both values empirically estimated. This makes the final 2D colour map more intuitive and easier to interpret clinically, as some isolated and misclassified pixels are removed while smoothing the ERM boundary. Finally, the refined 2D colour map can be overlaid over the 2D representation of the eye fundus to obtain an intuitive representation of the ERM in colour over the retina surface. Fig. 8 shows a representative example of the preliminary 2D colour map reconstruction stage, the post-processing map refinement process and the final overlay.

### E. EVALUATION METRICS

The designed methodology was validated using different statistical metrics commonly reported in the literature for similar medical imaging tasks. In this sense, Sensitivity (1),
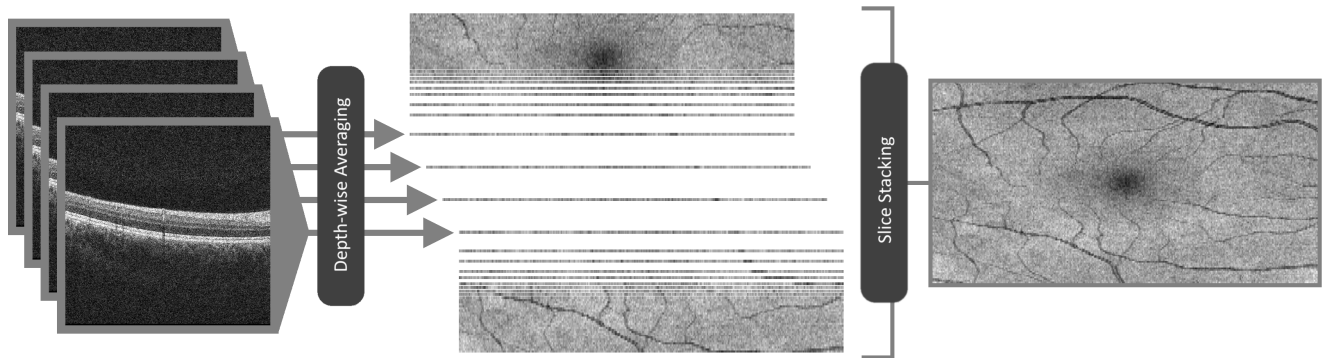
**FIGURE 7.** Example of how the 2D fundus image is reconstructed by vertically averaging the OCT slices and stacking them to form an aerial view of the macular region.
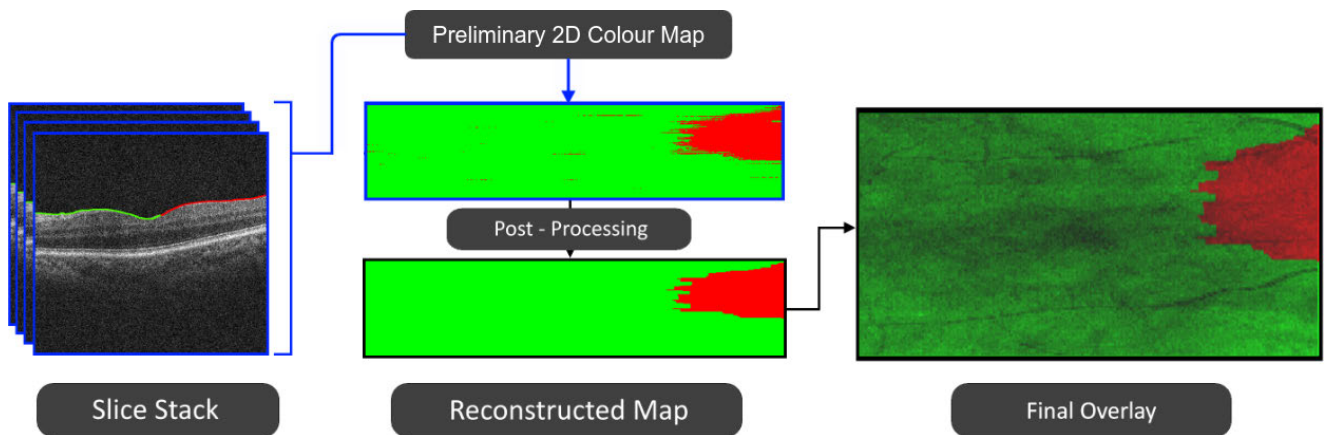


**FIGURE 8.** The third step begins by stacking the segmented OCT slices to form a segmentation map. A post-processing stage helps to clean spurious pixels and smooth the ERM boundary. With this refined map, an intuitive visual representation can be generated.

Specificity (2), Sørensen-Dice Coefficient (3), Jaccard Index (4) and Accuracy (5) were calculated for the quantitative validation of the obtained results.

$$Sensitivity = \frac{TP}{TP + FN} \quad (1)$$

$$Specificity = \frac{TN}{TN + FP} \quad (2)$$

$$Dice = \frac{2 \times TP}{2 \times TP + FP + FN} \quad (3)$$

$$Jaccard = \frac{TP}{TP + FP + FN} \quad (4)$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (5)$$

## III. RESULTS AND DISCUSSION

This section covers all the results that were acquired during the training, validation and testing of the models in terms of classification and segmentation metrics. It begins by covering the preliminary sampling strategy test where we determined the best number of windows to extract from each healthy and pathological image. Next, an evaluation of the classification stage is presented. Finally, we present the final results, including the 2D colour map reconstruction and post-processing

map refinement stage, as well as a comparison between the proposed method and the previous work.

### A. EVALUATION OF THE SAMPLING STRATEGIES

As mentioned in section II, a representative set of preliminary experiments were conducted to determine an appropriate sampling strategy prior to the training stage. In this line, 3 different configurations were analysed: 10 equispaced windows from each diseased eye and 2 from healthy ones, 20 from diseased and 4 from healthy eyes and, lastly, 40 and 8 respectively. The balanced (macro-average) accuracy of the 3 considered sampling strategies is represented in Fig. 9. As we can see, the best results were obtained by extracting 40 equispaced windows from images belonging to diseased eyes and 8 from healthy ones. This strategy allows a compromise between acquiring a representative sample of pathological tissue without over-representing the healthy class. As stated before, the OCT images range between 479 and 693 pixels in width while extracted samples are 112 pixels wide. By sampling 40 windows from each image, there is a consistent overlapping between them.
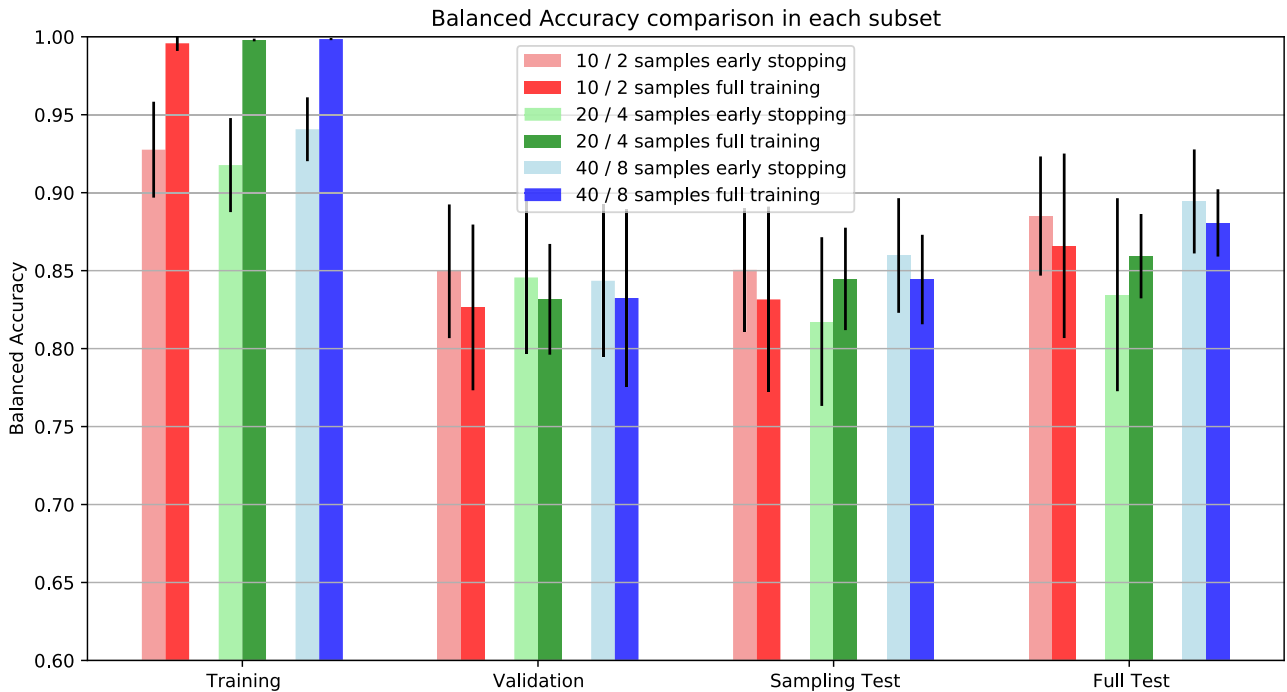
**FIGURE 9.** Comparison of balanced accuracy of the 3 sampling strategies, where the light columns show early stopping of the model by validation. Sampling test represents the results obtained from the classification of test samples obtained using the same sampling strategy as for training and testing over the test set. Full test represents the results obtained from the classification of all pixels using the test dataset.

## B. EVALUATION OF THE CLASSIFICATION STAGE

Based on the results obtained in the previous experiments, the best sampling strategy was used in the training and validation of the DenseNet-121 models. In addition, to quantitatively assess whether the trained model has learned about the domain-specific context, a comprehensive analysis of its classification ability was performed. In this line, 21.667 and 26.293 samples of ERM and Non-ERM categories were used respectively. Table 2 shows the results that were obtained by classifying each pixel of the ILM layer using the test dataset, averaged for the 4 folds. As we can see, these results show that the proposed system is able to accurately detect the presence of ERM, correctly identifying more than 82% of the ERM-related pixels, while keeping specificity over 94%.



**FIGURE 10.** Representative example of OCT scans of healthy eyes with the presence of hyper-reflective surfaces in the ILM layer that are due to acquisition conditions and can be misclassified as ERM regions.

**TABLE 2.** Accuracy, Sensitivity and Specificity of test results for 4 folds in terms of mean and standard deviation.

| | Classification Stage | | |
|---|---|---|---|
| | Accuracy | Sensitivity | Specificity |
| **Mean** | 0.918 | 0.828 | 0.944 |
| **SD** | ±0.023 | ±0.074 | ±0.028 |

Despite the good results that were obtained, the method presents some misclassifications that are mainly due to the presence of small artefacts and/or reflections, both inherent characteristics of the OCT capture devices. In this sense, Fig. 10 shows a representative example of OCT images that show healthy regions that are very similar to ERM regions
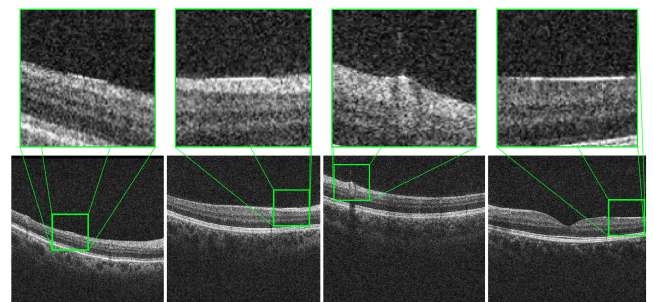
and, therefore, lead to some of the classification errors presented in this proposal.

## C. EVALUATION OF THE 2D COLOUR MAP RECONSTRUCTION AND POST-PROCESSING MAP REFINEMENT

By stacking each classified slice of the same 3D OCT volume, a preliminary 2D colour map can be reconstructed to facilitate the visual inspection of the segmented ERM regions. While this preliminary map shows a complete spatial representation of the classification results and segmented regions, only the isolated classifications of each OCT slice are taken into account. Therefore, as explained and motivated in the methodological section, this may result in the presence of
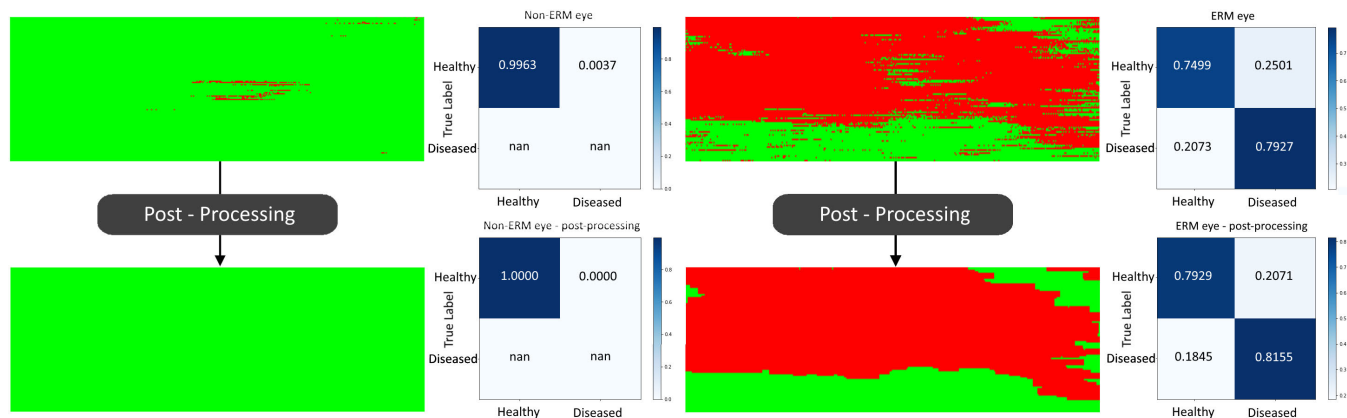
**FIGURE 11.** Representative example of the effects of the post-processing map refinement on a healthy and a diseased 3D OCT volumes respectively, with accompanying confusion matrices. $1^{st}$ row shows the output of the classification stage. $2^{nd}$ row is the result of post-processing.

**TABLE 3.** Dice and Jaccard index results before and after applying the post-processing map refinement in the test set.

|  |  | Segmentation | Post-processing |
|---|---|---|---|
| **Dice** | **Mean** | 0.800 | 0.826 |
| | **SD** | ±0.100 | ±0.112 |
| **Jaccard** | **Mean** | 0.676 | 0.714 |
| | **SD** | ±0.141 | ±0.155 |

**TABLE 4.** Comparison of Dice and Jaccard indexes of the previous work and the deep learning method proposed in this work for classification and post-processing stages.

|  |  | Baamonde *et al.* [45] | | Our proposal | |
|---|---|---|---|---|---|
|  |  | Segmentation | Post-processing | Segmentation | Post-processing |
| **Dice** | **Mean** | 0.670 | 0.780 | 0.810 | 0.833 |
| | **SD** | ±0.110 | ±0.092 | ±0.099 | ±0.091 |
| **Jaccard** | **Mean** | 0.515 | 0.649 | 0.689 | 0.725 |
| | **SD** | ±0.140 | ±0.128 | ±0.139 | ±0.129 |

very small (even at pixel level) and isolated ERM regions, which do not represent the uniform and regional characteristics of this relevant eye pathology. In some cases, this may also make the pathological boundary regions more irregular, as some OCT slices belonging to the same 3D OCT volume may be more reflective than others. To mitigate this possible effect, a post-processing map refinement is applied for the integration of cross-sectional information in the 2D colour map representation process. As we can see in Table 3, satisfactory results were achieved for the segmentation, reaching Dice coefficient scores of $0.800 \pm 0.100$ and $0.826 \pm 0.112$, before and after the application of the post-processing map refinement, respectively.

Fig. 11 shows a couple of representative examples of the post-processing stage, where the FP and FN rates are reduced, improving the overall results of the proposed method and, consequently, the final representation of the 2D colour map.

Regarding the methodological novelty and the potential relevance of the proposed method with respect to other similar approaches, we would like to point out that in the literature, to date, there is only one single methodology specifically designed for ERM segmentation in 3D OCT volumes. The state-of-the-art proposal evaluates the segmentation stage using only ERM-positive eyes [45]. As we can see in Table 4, our proposal obtains better results than the state of the art even before the application of the post-processing map refinement. In this context, our work achieved Dice coefficient scores of $0.810 \pm 0.099$ and $0.833 \pm 0.091$, with and without the application of the post-processing map refinement, respectively.
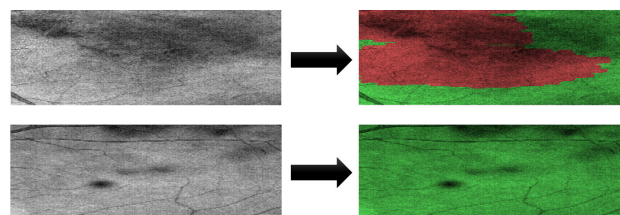


**FIGURE 12.** Representative examples of overlaying a 2D fundus image with the acquired segmentation for an eye presenting ERM (top) and a healthy one (bottom).

Finally, by averaging the OCT 3D volume values in-depth, we obtain a 2D fundus image representing the analysed region. The resulting 2D colour map from the previous steps can be combined with this 2D fundus image to generate a more intuitive visualisation of the ERM location over the retina. In addition, this final visualisation module, as shown in Fig. 12, can also be used to determine the bounds where the ERM must be peeled by vitrectomy in a simple way.

## IV. CONCLUSION

In this work, we have presented a complete methodology for the ERM segmentation and characterisation, capable of generating a complete reconstruction of an intuitive 2D colour map, using a 3D OCT volume as source of information. This fully automatic system provides crucial information that facilitates the diagnostic process and monitoring of the patients with ERM.

For this purpose, we first restrict the search space using, as a reference, the location of the ILM layer, region where the ERM potentially appears. A sliding window is then passed over the ILM layer to extract a set of samples. Next, a deep learning architecture is used to classify each sample between ERM and non-ERM classes. Finally, all OCT scans from the same 3D OCT volume are combined to produce a 2D colour map over the 2D fundus reconstruction, providing an intuitive representation of the presence of ERM and facilitating the diagnosis and treatment of this relevant eye disease.

As final result, the proposed system provided a Dice score of $0.826 \pm 0.112$ and a Jaccard coefficient of $0.676 \pm 0.714$ for the 3D OCT volumes, demonstrating its applicability for diagnosis purposes. Furthermore, from the obtained experimental results, we can conclude that the proposed system obtains better results than the state of the art even before the application of the post-processing map refinement. Consequently, this indicates that our deep learning-based approach can better incorporate the surrounding information than the classical machine learning-based approach. Additionally, by employing a convolutional neural network for the classification, the relevant graphical features are automatically extracted by the model. This greatly simplifies the training procedure, as it bypasses the manual feature selection process performed by other works in the field.

Plans for future work include refining this methodology and incorporating surrounding slices into the classification process by using a fully 3-dimensional CNN architecture that can extract volumetric patterns from the ILM and potentially improve the obtained results. Furthermore, with a dataset appropriately labelled by experts, we could extend this proposal by incorporating different grades of ERM disease. For example, the differentiation between ERM-attached and ERM-detached could provide a better characterisation of this disease and offer more complete visualisation of the ERM presence for clinicians, highlighting the type of damage that appears over the retinal tissues. In this way, more information is provided to help in the diagnosis and treatment of this and other related eye pathologies.

## REFERENCES

[1] D. Shen, G. Wu, and H.-I. Suk, "Deep learning in medical image analysis," *Annu. Rev. Biomed. Eng.*, vol. 19, no. 1, pp. 221–248, Jun. 2017, doi: 10.1146/annurev-bioeng-071516-044442.

[2] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. W. M. van der Laak, B. van Ginneken, and C. I. Sánchez, "A survey on deep learning in medical image analysis," *Med. Image Anal.*, vol. 42, pp. 60–88, Dec. 2017, doi: 10.1016/j.media.2017.07.005.

[3] M. Goldbaum, S. Moezzi, A. Taylor, S. Chatterjee, J. Boyd, E. Hunter, and R. Jain, "Automated diagnosis and image understanding with object extraction, object classification, and inferencing in retinal images," in *Proc. 3rd IEEE Int. Conf. Image Process.*, vol. 3, Sep. 1996, pp. 695–698, doi: 10.1109/ICIP.1996.560760.

[4] C. A. Puliafito, M. R. Hee, C. P. Lin, E. Reichel, J. S. Schuman, J. S. Duker, J. A. Izatt, E. A. Swanson, and J. G. Fujimoto, "Imaging of macular diseases with optical coherence tomography," *Ophthalmology*, vol. 102, no. 2, pp. 217–229, Feb. 1995, doi: 10.1016/s0161-6420(95)31032-9.

[5] J. S. Schuman, "Quantification of nerve fiber layer thickness in normal and glaucomatous eyes using optical coherence tomography," *Arch. Ophthalmol.*, vol. 113, no. 5, p. 586, May 1995, doi: 10.1001/archopht.1995.01100050054031.

[6] H. Hao, Y. Zhao, Q. Yan, R. Higashita, J. Zhang, Y. Zhao, Y. Xu, F. Li, X. Zhang, and J. Liu, "Angle-closure assessment in anterior segment OCT images via deep learning," *Med. Image Anal.*, vol. 69, Apr. 2021, Art. no. 101956, doi: 10.1016/j.media.2021.101956.

[7] L. S. Lim, P. Mitchell, J. M. Seddon, F. G. Holz, and T. Y. Wong, "Ophthalmology 1: Age-related macular degeneration," *Lancet*, vol. 379, no. 9827, pp. 1728–1738, May 2012. [Online]. Available: https://www-proquest-com.accedys.udc.es/scholarly-journals/ophthalmology-1-age-related-macular-degeneration/docview/1013605664/se-2?accountid=17197

[8] C.-H. Lin, W.-M. Liao, J.-W. Liang, P.-H. Chen, C.-E. Ko, C.-H. Yang, and C.-K. Lu, "Denoising performance evaluation of automated age-related macular degeneration detection on optical coherence tomography images," *IEEE Sensors J.*, vol. 21, no. 1, pp. 790–801, Jan. 2021, doi: 10.1109/JSEN.2020.3014254.

[9] J. D. Moura, J. Novo, and M. Ortega, "Deep feature analysis in a transfer learning-based approach for the automatic identification of diabetic macular edema," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2019, pp. 1–8, doi: 10.1109/IJCNN.2019.8852196.

[10] S. Vujosevic, C. Toma, E. Villani, A. Muraca, E. Torti, G. Florimbi, F. Leporati, M. Brambilla, P. Nucci, and S. De Cilla, "Diabetic macular edema with neuroretinal detachment: OCT and OCT-angiography biomarkers of treatment response to anti-VEGF and steroids," *Acta Diabetol.*, vol. 57, no. 3, pp. 287–296, Mar. 2020, doi: 10.1007/s00592-019-01424-4.

[11] G. Savini, "Correlation between retinal nerve fibre layer thickness and optic nerve head size: An optical coherence tomography study," *Brit. J. Ophthalmol.*, vol. 89, no. 4, pp. 489–492, Apr. 2005, doi: 10.1136/bjo.2004.052498.

[12] Y. He, A. Carass, Y. Liu, B. M. Jedynak, S. D. Solomon, S. Saidha, P. A. Calabresi, and J. L. Prince, "Structured layer surface segmentation for retina OCT using fully convolutional regression networks," *Med. Image Anal.*, vol. 68, Feb. 2021, Art. no. 101856, doi: 10.1016/j.media.2020.101856.

[13] A. González-López, J. de Moura, J. Novo, M. Ortega, and M. G. Penedo, "Robust segmentation of retinal layers in optical coherence tomography images based on a multistage active contour model," *Heliyon*, vol. 5, no. 2, Feb. 2019, Art. no. e01271, doi: 10.1016/j.heliyon.2019.e01271.

[14] S. S. Ong, M. O. Linz, X. Li, T. Y. A. Liu, I. C. Han, and A. W. Scott, "Retinal thickness and microvascular changes in children with sickle cell disease evaluated by optical coherence tomography (OCT) and OCT angiography," *Amer. J. Ophthalmol.*, vol. 209, pp. 88–98, Jan. 2020, doi: 10.1016/j.ajo.2019.08.019.

[15] J. de Moura, J. Novo, J. Rouco, M. G. Penedo, and M. Ortega, "Automatic detection of blood vessels in retinal OCT images," in *Proc. Int. Work-Conf. Interplay Between Natural Artif. Comput.* Cham, Switzerland: Springer, 2017, pp. 3–10, doi: 10.1007/978-3-319-59773-7_1.

[16] A. Watanabe, S. Arimoto, and O. Nishi, "Correlation between metamorphopsia and epiretinal membrane optical coherence tomography findings," *Ophthalmology*, vol. 116, no. 9, pp. 1788–1793, Sep. 2009, doi: 10.1016/j.ophtha.2009.04.046.

[17] M. H. Suh, J. M. Seo, K. H. Park, and H. G. Yu, "Associations between macular findings by optical coherence tomography and visual outcomes after epiretinal membrane removal," *Amer. J. Ophthalmol.*, vol. 147, no. 3, pp. 473.e3–480.e3, Mar. 2009, doi: 10.1016/j.ajo.2008.09.020.

[18] R. Rahman and J. Stephenson, "Early surgery for epiretinal membrane preserves more vision for patients," *Eye*, vol. 28, no. 4, pp. 410–414, Jan. 2014, doi: 10.1038/eye.2013.305.

[19] M. W. Johnson, "Posterior vitreous detachment: Evolution and complications of its early stages," *Amer. J. Ophthalmol.*, vol. 149, no. 3, pp. 371.e1–382.e1, Mar. 2010, doi: 10.1016/j.ajo.2009.11.022.

[20] S. Fraser-Bell, M. Guzowski, E. Rochtchina, J. J. Wang, and P. Mitchell, "Five-year cumulative incidence and progression of epiretinal membranes: The blue mountains eye study," *Ophthalmology*, vol. 110, no. 1, pp. 34–40, Jan. 2003, doi: 10.1016/s0161-6420(02)01443-4.

[21] J. Haller, H. Qin, R. S. Apte, R. R. Beck, N. M. Bressler, D. J. Browning, R. P. Danis, A. R. Glassman, J. M. Googe, C. Kollman, and A. K. Lauer, "Vitrectomy outcomes in eyes with diabetic macular edema and vitreomacular traction," *Ophthalmology*, vol. 117, no. 6, pp. 1087.e3–1093.e3, Jun. 2010, doi: 10.1016/j.ophtha.2009.10.040.

[22] C. J. Flaxel, A. R. Edwards, L. P. Aiello, P. G. Arrigg, R. W. Beck, N. M. Bressler, S. B. Bressler, F. L. Ferris, S. K. Gupta, J. A. Haller, H. S. Lazarus, and H. Qin, "Factors associated with visual acuity outcomes after vitrectomy for diabetic macular edema," *Retina*, vol. 30, no. 9, pp. 1488–1495, Oct. 2010, doi: 10.1097/iae.0b013e3181e7974f.

[23] A. Ophir, M. R. Martinez, P. Mosqueda, and A. Trevino, "Vitreous traction and epiretinal membranes in diabetic macular oedema using spectral-domain optical coherence tomography," *Eye*, vol. 24, no. 10, pp. 1545–1553, Jun. 2010, doi: 10.1038/eye.2010.80.

[24] T. Yamamoto, N. Akabane, and S. Takeuchi, "Vitrectomy for diabetic macular edema: The role of posterior vitreous detachment and epimacular membrane," *Amer. J. Ophthalmol.*, vol. 132, no. 3, pp. 369–377, 2001, doi: 10.1016/s0002-9394(01)01050-9.

[25] N. G. Ghazi, J. B. Ciralsky, S. M. Shah, P. A. Campochiaro, and J. A. Haller, "Optical coherence tomography findings in persistent diabetic macular edema: The vitreomacular interface," *Amer. J. Ophthalmol.*, vol. 144, no. 5, pp. 747–754, 2007, doi: 10.1016/j.ajo.2007.07.012.

[26] D. R. Matthews, I. M. Stratton, S. J. Aldington, R. R. Holman, and E. M. Kohner, "Risks of progression of retinopathy and vision loss related to tight blood pressure control in type 2 diabetes mellitus," *Arch. Ophthalmol.*, vol. 122, no. 11, p. 1631, Nov. 2004, doi: 10.1001/archopht.122.11.1631.

[27] H. Koizumi, R. F. Spaide, Y. L. Fisher, K. B. Freund, J. M. Klancnik, and L. A. Yannuzzi, "Three-dimensional evaluation of vitreomacular traction and epiretinal membrane using spectral-domain optical coherence tomography," *Amer. J. Ophthalmol.*, vol. 145, no. 3, pp. 509.e1–517.e1, Mar. 2008, doi: 10.1016/j.ajo.2007.10.014.

[28] C. I. Falkner-Radler, C. Glittenberg, S. Hagen, T. Benesch, and S. Binder, "Spectral-domain optical coherence tomography for monitoring epiretinal membrane surgery," *Ophthalmology*, vol. 117, no. 4, pp. 798–805, Apr. 2010, doi: 10.1016/j.ophtha.2009.08.034.

[29] J. R. Wilkins, C. A. Puliafito, M. R. Hee, J. S. Duker, E. Reichel, J. G. Coker, J. S. Schuman, E. A. Swanson, and J. G. Fujimoto, "Characterization of epiretinal membranes using optical coherence tomography," *Ophthalmology*, vol. 103, no. 12, pp. 2142–2151, Dec. 1996, doi: 10.1016/s0161-6420(96)30377-1.

[30] S. Baamonde, J. de Moura, J. Novo, J. Rouco, and M. Ortega, "Feature definition and selection for epiretinal membrane characterization in optical coherence tomography images," in *Image Analysis and Processing*. Cham, Switzerland: Springer, 2017, pp. 456–466, doi: 10.1007/978-3-319-68548-9_42.

[31] Y.-C. Lo, K.-H. Lin, H. Bair, W. H.-H. Sheu, C.-S. Chang, Y.-C. Shen, and C.-L. Hung, "Epiretinal membrane detection at the ophthalmologist level using deep learning of optical coherence tomography," *Sci. Rep.*, vol. 10, no. 1, May 2020, Art. no. 8424, doi: 10.1038/s41598-020-65405-2.

[32] S. Kuwayama, Y. Ayatsuka, D. Yanagisono, T. Uta, H. Usui, A. Kato, N. Takase, Y. Ogura, and T. Yasukawa, "Automated detection of macular diseases by optical coherence tomography and artificial intelligence machine learning of optical coherence tomography images," *J. Ophthalmol.*, vol. 2019, pp. 1–7, Apr. 2019, doi: 10.1155/2019/6319581.

[33] T. Sonobe, H. Tabuchi, H. Ohsugi, H. Masumoto, N. Ishitobi, S. Morita, H. Enno, and D. Nagasato, "Comparison between support vector machine and deep learning, machine-learning technologies for detecting epiretinal membrane using 3D-OCT," *Int. Ophthalmol.*, vol. 39, no. 8, pp. 1871–1877, Sep. 2018, doi: 10.1007/s10792-018-1016-x.

[34] S. Baamonde, J. de Moura, J. Novo, P. Charlón, and M. Ortega, "Automatic identification and characterization of the epiretinal membrane in OCT images," *Biomed. Opt. Exp.*, vol. 10, no. 8, p. 4018, Jul. 2019, doi: 10.1364/boe.10.004018.

[35] A. Paszke *et al.*, "Pytorch: An imperative style, high-performance deep learning library," in *Advances in Neural Information Processing Systems*, vol. 32, H. Wallach, H. Larochelle, A. Beygelzimer, F. d' Alché-Buc, E. Fox, and R. Garnett, Eds. Red Hook, NY, USA: Curran Associates, 2019, pp. 8026–8037. [Online]. Available: https://proceedings.neurips.cc/paper/2019/file/bdbca288fee7f92f2bfa9f7012727740-Paper.pdf

[36] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and É. Duchesnay, "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, no. 85, pp. 2825–2830, 2011. [Online]. Available: http://jmlr.org/papers/v12/pedregosa11a.html

[37] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 2261–2269, doi: 10.1109/cvpr.2017.243.

[38] B. Al-Bander, B. Williams, W. Al-Nuaimy, M. Al-Taee, H. Pratt, and Y. Zheng, "Dense fully convolutional segmentation of the optic disc and cup in colour fundus for glaucoma diagnosis," *Symmetry*, vol. 10, no. 4, p. 87, Mar. 2018. [Online]. Available: https://www.mdpi.com/2073-8994/10/4/87, doi: 10.3390/sym10040087.

[39] S. Wang, C. Tang, J. Sun, and Y. Zhang, "Cerebral micro-bleeding detection based on densely connected neural network," *Frontiers Neurosci.*, vol. 13, p. 422, May 2019, doi: 10.3389/fnins.2019.00422.

[40] O. Yildirim, M. Talo, B. Ay, U. B. Baloglu, G. Aydin, and U. R. Acharya, "Automated detection of diabetic subject using pre-trained 2D-CNN models with frequency spectrum images extracted from heart rate signals," *Comput. Biol. Med.*, vol. 113, Oct. 2019, Art. no. 103387. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0010482519302641, doi: 10.1016/j.compbiomed.2019.103387.

[41] K. Gawlik, F. Hausser, F. Paul, A. U. Brandt, and E. M. Kadas, "Active contour method for ILM segmentation in ONH volume scans in retinal OCT," *Biomed. Opt. Exp.*, vol. 9, no. 12, pp. 6497–6518, Dec. 2018. [Online]. Available: http://www.osapublishing.org/boe/abstract.cfm?URI=boe-9-12-6497, doi: 10.1364/BOE.9.006497.

[42] N. Kanopoulos, N. Vasanthavada, and R. L. Baker, "Design of an image edge detection filter using the sobel operator," *IEEE J. Solid-State Circuits*, vol. SSC-23, no. 2, pp. 358–367, Apr. 1988, doi: 10.1109/4.996.

[43] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 248–255, doi: 10.1109/CVPR.2009.5206848.

[44] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. 3rd Int. Conf. Learn. Represent. (ICLR)*, Y. Bengio and Y. LeCun, Eds., San Diego, CA, USA, May 2015, pp. 1–15. [Online]. Available: http://arxiv.org/abs/1412.6980

[45] S. Baamonde, J. de Moura, J. Novo, P. Charlón, and M. Ortega, "Automatic identification and intuitive map representation of the epiretinal membrane presence in 3D OCT volumes," *Sensors*, vol. 19, no. 23, p. 5269, Nov. 2019, doi: 10.3390/s19235269.

**MATEO GENDE** was born in Spain, in 1995. He received the degree in computer engineering from the University of A Coruña, Spain, in 2018, where he is currently pursuing the Ph.D. degree in computer science and artificial intelligence. Since 2018, he has been working as a Research Assistant with the Centre for Research in Information and Communication Technologies (CITIC), Spain. His research interests include computer vision, medical image analysis, pattern recognition, and machine learning.

**JOAQUIM DE MOURA** received the degree in computer engineering, the M.Sc. degree in computer engineering, and the Ph.D. degree *(cum laude)* in computer science and artificial intelligence from the University of A Coruña, Spain, in 2014, 2016, and 2019, respectively. He worked as a Visiting Researcher with INESC-TEC, Portugal, in the development of different methodologies for DME diagnosis with OCT images. He is currently a Postdoctoral Researcher with the Centre for Research in Information and Communication Technologies (CITIC), University of A Coruña. His research interests include computer vision, biomedical image processing, and video processing.

**JORGE NOVO** was born in Spain, in 1983. He received the M.Sc. and Ph.D. degrees *(cum laude)* in computer science from the University of A Coruña, in 2007 and 2012, respectively. He worked as a Visiting Researcher with CMR images in the detection of landmark points at Imperial College London, and a Postdoctoral Research Fellow with INEB and INESC-TEC research institutes in the development of CAD systems for lung cancer diagnosis with chest CT images. He is currently an Assistant Professor with the Department of Computer Science and Information Technologies, University of A Coruña. His research interests include computer vision, pattern recognition, and biomedical image processing.

**MARCOS ORTEGA** received the degree in computer science, in 2004, and the doctor's degree in computer science, in 2009. He currently serves as an Associate Professor for the University of A Coruña, teaching mainly in the Faculty of Computer Science, also serving as his Secretary. He is also a Researcher with the Research Centre in Information and Communication Technologies (CITIC), and a member of its scientific committee and its representative in the ECHAlliance. He is also a member of the Institute of Biomedical Research of A Coruña (INIBIC) with the rank of Principal Investigator. His research interests include computer vision, medical image processing, and medical informatics.

● ● ●

**PABLO CHARLÓN** received the M.Sc. degree from the University of Santiago de Compostela, Spain, in 2011, with a specialising in the field of vision science research. He is currently working as a Clinical Optometrist with the Victoria de Rojas Ophthalmological Institute, A Coruña, Spain, and the HM Rosaleda Hospital. He is also the President of the Galician Society of Clinical Optometry. His research interests include vision sciences and the eye, where he has published a number of research articles on OCT imaging.