Morales, S.; Colomer, A.; Mossi García, JM.; Del Amor, R.; Woldbye, D.; Klemp, K.; Larsen, M.... (2021). Retinal layer segmentation in rodent OCT images: Local intensity profiles & fully convolutional neural networks. Computer Methods and Programs in Biomedicine. 198:1-14. https://doi.org/10.1016/j.cmpb.2020.105788

Retinal layer segmentation in rodent OCT images: Local intensity profiles & fully convolutional neural networks

Sandra Morales[a,*], Adrián Colomer[a], José M. Mossi[b], Rocío del Amor[a], David Woldbye[c], Kristian Klemp[d,e], Michael Larsen[d,e], Valery Naranjo[a]

[a] *Instituto de Investigación e Innovación en Bioingeniería, I3B, Universitat Politècnica de València, Camino de Vera s/n, 46022 Valencia, Spain.*
[b] *ITEAM Research Institute, Universitat Politècnica de València, Camino de Vera s/n, 46022 Valencia, Spain.*
[c] *Laboratory of Neural Plasticity, Department of Neuroscience, University of Copenhagen, Denmark.*
[d] *Dept. of Ophthalmology, Rigshospitalet-Glostrup, Glostrup, Copenhagen, Denmark.*
[e] *Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark.*

Abstract

Background and Objective: Optical coherence tomography (OCT) is a useful technique to monitor retinal layer state both in humans and animal models. Automated OCT analysis in rats is of great relevance to study possible toxic effect of drugs and other treatments before human trials. In this paper, two different approaches to detect the most significant retinal layers in a rat OCT image are presented. **Methods:** One approach is based on a combination of local horizontal intensity profiles along with a new proposed variant of watershed transformation and the other is built upon an encoder-decoder convolutional network architecture. **Results:** After a wide validation, an averaged absolute distance error of 3.77 ± 2.59 and 1.90 ± 0.91 μm is achieved by both approaches, respectively, on a batch of the rat OCT database. After a second test of the deep-learning-based method using an unseen batch of the database, an averaged absolute distance error of 2.67 ± 1.25 μm is obtained. The rat OCT database used in this paper is made publicly available to facilitate further comparisons. **Conclusions:** Based on the obtained results, it was demonstrated the compet-

*Corresponding author
Email address: sanmomar@i3b.upv.es (Sandra Morales)

October 16, 2020

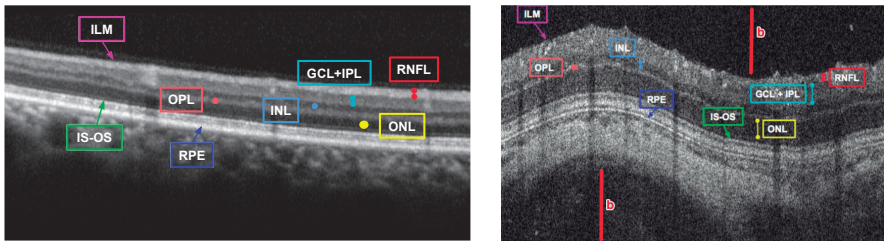
itiveness of the first approach since outperforms the commercial Insight image segmentation software (Phoenix Research Labs) as well as its utility to generate labelled images for validation purposes speeding significantly up the ground truth generation process. Regarding the second approach, the deep-learning-based method improves the results achieved by the more conventional method and also by other state-of-the-art techniques. In addition, it was verified that the results of the proposed network can be generalized to new rat OCT images.
Keywords: Optical coherence tomography, rodent OCT, rat OCT, layer segmentation, convolutional neural networks, intensity profile.

1. Introduction

Optical coherence tomography (OCT) is a powerful imaging technology able to visualize high-resolution cross sections of the retina that has become a prominent ophthalmic diagnostic technique and is routinely used to monitor retinal damage [31]. A way of monitoring the retinal damage is through the analysis of retinal layer thickness. To obtain proper layer thickness measurements, a previous accurate segmentation of the retinal layers from the OCT images is essential. Due to the fact that manual OCT segmentation is tedious, time-consuming and suffers from inter and intra-observer variability, the development of automatic layer segmentation algorithms is of importance. However, automatic layer segmentation can be very challenging because of factors such as speckle noise and low image contrast that are present in OCT images [21].

OCT can be adapted for rodent eye imaging in order to be used in research studies of ophthalmic diseases and also for treatment evaluation. It must be taken into account that human and rodent retina present significant differences. For example, rodents have no macula or fovea and their lenses are relatively larger. Regarding the thickness of the retinal layers, the ganglion cell layer (GCL) in rodents is not visually distinguishable (app. 2μm) while in humans is around 20-60μm. The third innermost layer, the inner plexiform layer (IPL), on the contrary, is relatively thick in the rat retina (app. 60 μm). In Figure 1,

the differences between human and rodent OCT images can be observed.



(a) (b)

Figure 1: Retinal layers in two OCT scans: (a) Human retina, Spectralis OCT (Heidelberg Eng.). (b) Rodent retina, Micron IV (Phoenix Research Labs). ILM: Internal Limiting Membrane, RNFL: Retinal Nerve Fiber Layer. GCL: Ganglion Cell Layer. IPL: Inner Plexiform Layer. INL: Inner Nuclear Layer. OPL: Outer Plexiform Layer. ONL: Outer Nuclear Layer. IS: Inner Segment. OS: Outer Segment. RPE: Retinal Pigment Epithelium. b: image background.

While numerous methods for retinal layer segmentation have been proposed in the literature for human OCT images [11, 8, 18, 9, 15, 24, 10, 16, 25, 19, 13, 12, 27, 17], few have addressed the segmentation of rodent eyes [21, 31, 29, 2, 3]. Yazdanpanah et al. proposed a semi-automated algorithm based on the Chan-Vese active contours without edges to address the segmentation of 6 retinal layer boundaries. It was a multi-phase, level-set Mumford-Shah model that incorporated a shape prior based on expert anatomical knowledge of the retinal layers [31]. They applied their algorithm to 80 retinal OCT images of seven rats and achieved an average Dice similarity coefficient of 0.84 over all segmented retinal layers. Mishra et al. used a two-step kernel-based optimization to segment retinal layers on a set of OCT images from healthy and diseased rodent retina [21]. The proposed algorithm segmented all intra-retinal layers (9 layers) in healthy retinas but in the case of diseased retinas only RNFL, IPL, ISOS and RPE were detected. Only two representative cases were shown but no quantitative results were provided. Srinivasan et al. presented an automatic approach that used sparsity based denoising, support vector machines, graph theory and dynamic programming to segment retinal layer boundaries, which could range from seven

to ten, in wild-type (WT) and rhodopsin knockout (Rho(-/-)) mice [29]. They calculated the absolute value of the average pixel difference between automatic and manual segmentations for every layer boundary in addition to compare to the commercial automated Diver segmentation software (Bioptigen). Their automatic results differed from the more experienced manual grader by $2.17\mu m$ for the WT group and $1.96\mu m$ for the Rho(-/-) group, respectively. Note that the aforementioned works are based on classical image processing techniques. In fact, machine learning methods have still not widely used to segment the different layers of the rodent's retina although some works can be found. The work proposed in [2] used a random forest classifier to design a cost function to be used as a part of a graph-theoretic method to segment ten retinal surfaces from SD-OCT mice scans. This method provided a overall unsigned border position error of $3.35 \pm 0.62\mu m$ over all segmented retinal layers. Antony et at. developed a flexible graph-based algorithm named ASIMOV for mouse OCT volume segmentation, which segmented ten retinal surfaces in healthy mice and six surfaces in light-damaged (LD) mice [3]. It was an extension of the work described by [18] that was used for human retinal layer segmentation. They learned a random forest classifier to identify the boundary pixels between layers, producing an accurate probability map for each boundary, which was then processed to finalise the boundaries. The overall mean error (along all segmented layers) with respect to three raters was $1.96 \pm 0.61\mu m$, $2.63 \pm 0.94\mu m$ and, $2.14 \pm 0.62\mu m$, respectively, in the normal scans. In the LD scans, the overall error was $2.41 \pm 0.56\mu m$, $2.71 \pm 0.73\mu m$ and $2.19 \pm 0.49\mu m$, respectively. These two last approaches are based on a hand-craft feature extraction stage and classification algorithms. However, very few works have focused on the segmentation of the rodent's retina layers by means of automatic feature extraction using deep learning techniques [1, 7]. So far, this kind of methodology has been mainly proposed to segment human retinal layers and the architectures proposed for humans do not have discriminatory capability enough to detect the thinnest layers in rodents because of the anatomical differences between both retinas. For example, the basic scheme of fully convolutional networks (FCNs)

was followed in [24] and [10] for human retinal layer segmentation using Gaussian Processes and the graph search technique as post-processing methods, respectively. 3D FCNs were proposed in [16] to extract the spatial and inter-frame informacion from human OCTs. FCN architectures composed by a contracting and expansive paths were proposed in [25] and [19] with the aim of encoding the relevant information of the human retinal layers for building a prediction map by means of the contracting path. Two FCNs in cascade were introduced in [13] to classify the human retinal layers with the first one and to remove topology errors with the second FCN. A combination of a convolutional neural network to extract the relevant features and long short term memory units to trace the layer boundaries is another DL-based approach to segment human retina layers using OCT [12]. The combination of a full convolutional network with a multiphase level set is also explored in [27] for automatic segmentation of nine boundaries in retinal OCT images belonging to patients with central serous chorioretinopathy. Fully-convolutional deep learning methods are proposed in [17] to determine retinal and choroidal segmentation in OCT images from children. Regarding the deep-learning methods applied for rodent layer segmentation, Amor et al. demonstrated that the FCN architecture proposed in [25] for the segmentation of human OCTs, even followed by a robust post-processing stage, only is able to obtain satisfactory results for the segmentation of three layer boundaries in rodents [1]. Chen et al. proposed three different strategies (training from scratch, transfer learning and continued training from a pre-trained model of a different animal cohort) to segment the contour of three layers leading similar performance in all of them. In this last work, the results achieved by each layer are presented through boxplots instead of a quantitative way which made the comparison difficult in addition to the fact that neither the code nor the images used are publicly available. However, to the best of the authors' knowledge, no previous studies have achieved to segment more than three retinal layers in rodent OCTs by means of automatic feature extraction using deep learning techniques.

In this paper, two different approaches to detect the six most significant

retinal layer boundaries in a rat OCT image are presented, one based on classical image processing techniques and the other based on deep learning. The first one combines the analysis of local horizontal intensity profiles, unlike other state-of-the-art works based on global vertical projections, with a new variant of the marker-controlled watershed transformation, the enclosed watershed, which is not affected by layers lacking continuity or very close to each other, common facts in this type of images. The second one proposes an encoder-decoder fully convolutional network (FCN) architecture, built upon ReLayNet network [25], to segment the rat retinal layers. The main novelty of the proposed architecture is the ability for capturing the significant information contained in the thinnest rodent layers by means of addition of five convolutional blocks and modifications in loss function. Both methods are widely validated. The different experiments carried out in this study let confirm that the best results are achieved with the deep-learning-based approach which also outperforms other works of the literature. However, it must be taken into account that the first approach has the advantage of being a more versatile method, able to properly work on rodent, other animal model and human OCTs without needing a specific training. In addition, it does not require a high number of labelled samples for that training, one of the main drawbacks of the deep learning algorithms. Thus, the classical approach presented in the paper can be used to generate the labelled images needed for training convolutional networks or to validate models pre-trained with different images, as it was carried out in this paper. In that way, although some images will require little manual modifications, the labelling process is highly speeded up and it is much more efficient than manually labelling from scratch.

Another contribution of this work is the presentation of a new rat OCT database where the segmentation of the retinal layers was manually expert-reviewed so that it can be used for further comparisons with other works. It is possible to find in the literature human public databases of these characteristics but not for rats which makes that most state-of-the-art works focused on rat OCT analysis were validated on private databases. The goal of this database

is to make possible a reliable comparison between segmentation methods for rodent OCTs. The presented database is composed of two batches which allows using the first batch for performance evaluation and the second one to analyse the generalization ability of the method when it is applied in an image set with different conditions to that used for training, as was done in this work.

2. Materials and Methods

2.1. Rodent model

For this study, we used two batches of a database of OCT images which were acquired at different conditions and belong to different rodents. In particular, a set of 22 Sprague-Dawley rats were used. For both batches, animal experiment permission was granted by the Danish Animal Experimentation Council (license number: 2017-15-0201-01213). Rat OCT images were taken with the Micron IV equipment (Phoenix Research Labs, Pleasanton, USA) in different time points and by different specialists. All rats were anesthetized previously to image acquisition. Several OCT images were acquired before and after intravitreal injection of endothelin-1 (ET-1). ET-1 causes vasoconstriction of retinal vessels and subsequent ischemia, which contributes to the degeneration of the retinal layers. These effects make rat model of intravitreal injection of ET-1 suitable for research studies about retinal layer thickness monitoring.

Specifically, for the first batch of the database, a follow-up 3, 7 and 14 days after injection was performed on 10 rats making a total of 129 OCT images of 1024x1024 pixels with 0.9775 μm/pixel. For the second batch, a follow-up 3 and 7 days after injection was performed obtaining a total of 115 OCT images belonging to 12 rats with the same size and resolution that the first batch images. On average, 13 images per rat for the first batch and 10 images for the second were recorded from circular scans located on different positions at different times during the study. Tables 1 and 2 detail the content of the two batches of the database.

Table 1: Content of the first batch of the rat OCT database. Before injection: Day 0. After ET-1 injection: Day 3, Day 7 and Day 14.

Rat ID	#images	Day 0	Day 3	Day 7	Day 14
C23 GP1A A1	16	4	4	4	4
C23 GP1A A2	17	4	5	4	4
C23 GP1A A3	16	4	4	4	4
C23 GP1A A4	16	4	4	4	4
C24 GP2B A1	16	4	4	4	4
C24 GP2B A2	16	4	4	4	4
C24 GP2B A3	16	4	4	4	4
C24 GP2B A4	4	4	0	0	0
C25 GP1A A1	8	4	4	0	0
C25 GP2B A3	4	4	0	0	0

Table 2: Content of the second batch of the rat OCT database. Before injection: Day 0. After ET-1 injection: Day 3 and Day 7

Rat ID	#images	Day 0	Day 3	Day 7
Rat01	7	5	2	0
Rat02	11	5	3	3
Rat03	10	5	3	2
Rat04	7	5	0	2
Rat05	11	5	2	4
Rat06	10	5	2	3
Rat07	11	4	4	3
Rat08	9	4	4	4
Rat09	14	5	4	5
Rat10	7	4	2	1
Rat11	10	3	4	3
Rat12	8	4	2	2

For the first batch of the database, the most significant retinal layer boundaries that were visibly distinguishable on these images were manually segmented by an expert for training and validation purposes. However, the second batch of the database, which was used to perform an alternative test, it was not manually segmented from scratch but it was based on an automatic segmentation manually reviewed and modified by experts. How the second batch is used will be widely explained in Section 3.2. The goal of the two different approaches presented in this paper will be to segment six layer boundaries (ILM, IPL-INL, INL-OPL, OPL-ONL, IS-OS, RPE) as it is shown in Figure 2. The complete database, including original images and the ground truth of the two batches, is publicly available in [22].

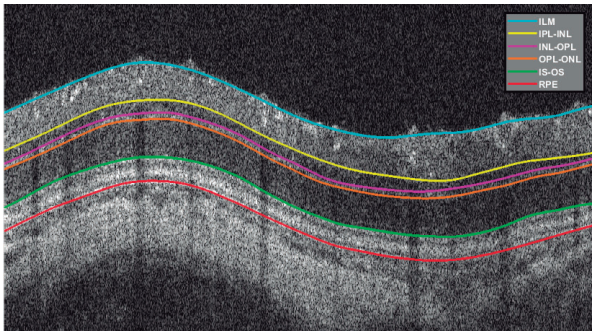


Figure 2: Ground truth of retinal layer boundaries.

2.2. Method 1: Local horizontal intensity profiles (M1)

Due to rat OCT image characteristics, outer nuclear layer (ONL) is the thickest hyporeflective layer of the retina as shown in Figure 1 shows. This prior knowledge makes the ONL one of the most stable layers taking into account the existing noise and the variability among different OCT scans. Thus, the retinal layer segmentation method presented in this subsection is focused on a first rough ONL detection as basis for subsequent OCT image processing steps. Other works in the literature are based on a first estimation of the Retinal Pigment Epithelium (RPE) asserting that the RPE is the most hyperreflective layer [8]. However, the choroid or the photoreceptors in some OCT images may have higher reflectivity levels and cause subsequent errors. In the flowchart of the Figure 3, the process to detect the six retinal layer boundaries is depicted. Although not all layer boundaries are obtained in the same way due to their own particularities, the layer-dependent parameters were equally set up for all the images of the database. Each stage will be detailed throughout this section. Note that M1 is directly designed to detect the six layer boundaries and M2 first detects the layers and, secondly, their boundaries are extracted.

2.2.1. Outer Nuclear Layer (ONL) Estimation

ONL estimation is based on horizontal projection of the image intensity, that is a way to avoid noise without losing resolution in the vertical direction. To

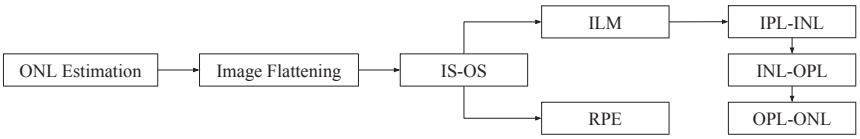


Figure 3: Flowchart of the proposed method M1 for retinal layer segmentation.

obtain a proper ONL estimation, the first step is to filter the image with a 2-D Gaussian smoothing kernel and then to limit the image region to be analyzed. As ONL is a hyporeflective layer, it is necessary to identify the retina region and mask the image background to avoid misidentification. As background we mean the dark region above the ILM towards the upper limit of the image, corresponding to the vitreous humor and the dark region below the lowest bright region of the image and the lower limit of the image. Both regions are noted in Fig. 1b as "b". The identification of the region of interest, i.e. the area that corresponds with the retina, is performed by means of Otsu's threshold after image filtering. Afterwards, the filtered and masked image is divided into 32 vertical sections and they are processed separately. The horizontal projection of each section is computed and the most prominent minimum between the two most prominent maxima of the horizontal profile are found. The prominence of a peak measures the intrinsic height of that peak with respect to other peaks. The set of prominent minima of all vertical sections are fitted by a spline of grade 6 so that the profile followed by ONL is estimated such as it can be observed in Figure 4.

2.2.2. Image Flattening

The estimation of the ONL is used for image flattening. Image is flattened by vertically (up or down) shifting the points of the ONL profile so that they are completely aligned. Once the image is flattened, all the layers will be practically parallel. This knowledge will make layer segmentation more precise because layers will be more visible in the horizontal projections of each vertical section if its main direction is, indeed, horizontal. Figure 5(a) shows a non-flattened

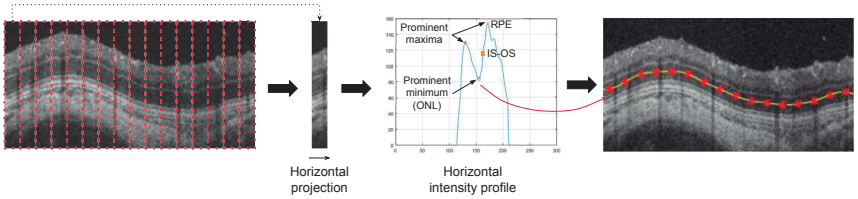


Figure 4: Graphic representation of ONL estimation. First, the image is divided into 32 vertical sections (for better understanding only some of them are shown in the figure). Secondly, the horizontal profile of each section is computed. Finally, the most prominent minimum between the two most prominent maxima of each horizontal profile is detected as the ONL estimation.

original image, and Figure 5(b) the resulting image after flattening process.

2.2.3. Inner Segment - Outer Segment (IS-OS) Detection

The boundary between IS and OS is the most enhanced because it is found between hypo and hyperreflective layers. Making use of the peaks (prominent maxima) and valleys (prominent minima) calculated in ONL estimation for each vertical section, the region of interest to find the IS-OS boundary is determined as the area located between the valley, that corresponds to the estimated ONL, and the peak on the right side, that should correspond to the RPE (see Figure 4). The IS-OS boundary is detected by calculating the point of maximum slope within that area. After IS-OS boundary detection, the original image is flattened once again but using the IS-OS boundary as reference because it is more accurate. The detection of the remaining boundaries is performed on the resulting flattened image. A correct IS-OS boundary detection is key for the segmentation of the other layers because this boundary can be used to limit the image area where other layers will be looked for.

2.2.4. ILM Detection

Because of the presence of noise in the vitreo area, the search of peaks on the local intensity profile may misidentifies the ILM so, for that reason, marker-controlled watershed is used for ILM detection. The watershed transformation

is a segmentation technique for gray-scale images where the image is considered as a topographic surface so that light pixels represent high elevations and dark pixels represent low elevations [6]. The basic idea consists of flooding the image from its minima to split the image into the catchment basins and the watershed lines. Due to this fact, the input image of this method is usually a gradient image $\varrho(f)$ and it can be noted by $WS(\varrho(f))$. The marker-controlled variant, $WS_{f_{mrk}}$, consists in flooding the topographic surface from a previously defined set of markers f_{mrk} to avoid over-segmentation caused by the existence of numerous local minima in the image. In the case of ILM detection, the upper limit of the retina (f_{UR}), previously identified, and the IS-OS boundary (f_{IS-OS}) are used as markers.

$$f_{mrk}(\mathbf{x}) = \begin{cases} 1 & \text{if } \mathbf{x} \in f_{UR} \cup f_{IS-OS} \\ 0 & Otherwise. \end{cases} \tag{1}$$

Retinal OCT images with prominent vessels may be a challenge for layer segmentation because they result in vertical hyporeflective regions that disturb the continuity of the different layers. Watershed transformation, as other techniques, can be affected by this issue. To address this problem, vessel regions are considered as missing pixels and they are restored using surrounding information through a diffusion-based inpainting technique [5]. This technique involves filling the selected parts of the image by propagating external information into the region to be inpainted. Vessels are inpainted from the original images previous to gradient computation. For vessel detection, the vertical pixel intensity profile of the flattened image is analysed. Due to the fact that the layers are practically parallel, the vertical profile should not contain high-contrasted peaks. Once the vertical profile is computed, the peaks whose intensity is greater than the Otsu's threshold are defined as vessel regions [20]. See Figure 5(b) and 5(c).

Let $\Omega(\mathbf{x})$ stand for the vessel regions to be inpainted and $\Upsilon(f, \Omega)(\mathbf{x})$ be the vessel-inpainted image of the original image $f(\mathbf{x})$. The ILM boundary is computed as the marker-controlled watershed of the vessel-inpainted image imposed by f_{mrk}, $WS(\varrho(\Upsilon))_{f_{mrk}}$. Afterwards, the resulting boundary of the marker-

controlled watershed is smoothed to remove irrelevant fluctuations. In Figure 5 the difference to apply the watershed transformation for ILM detection with and without vessel inpainting can be observed.

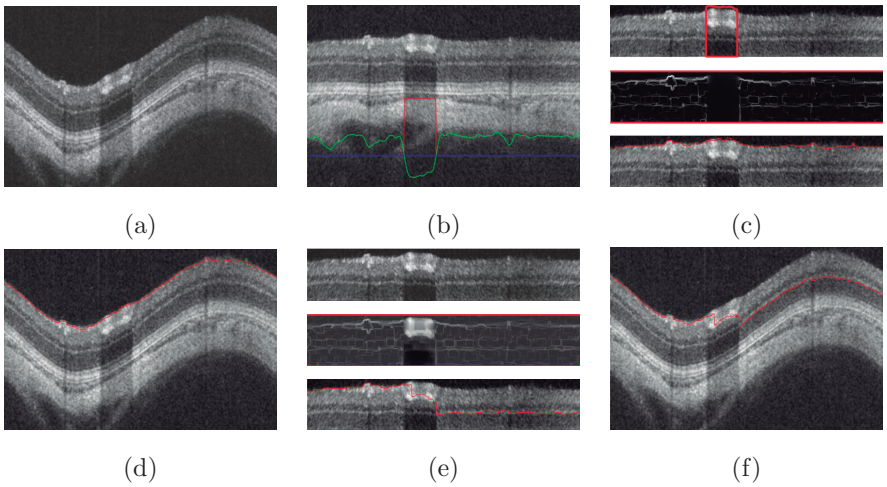


Figure 5: ILM detection: (a) Original OCT image. (b) Prominent vessel detection on the flattened image (vessels marked in red, Otsu's threshold in blue, intensity vertical profile in green). (c) Watershed transformation after inpainting (upper row: region of interest with the area to be inpainted in red, middle row: gradient after inpainting, watershed result). (d) ILM detection after inpainting on the original image. (e) Watershed transformation without inpainting (upper row: region of interest, middle row: gradient without inpainting, watershed result). (f) ILM detection without inpainting on the original image.

2.2.5. Inner Plexiform Layer - Outer Nuclear Layer (IPL-ONL) Detection

For the detection of IPL-INL, INL-OPL and OPL-ONL boundaries, the same approach followed in IS-OS detection is performed but restricting the search for prominent maxima and minima of the horizontal intensity profile to the area limited by the ILM and the ONL boundary previously estimated as it is depicted in Figure 6. The horizontal intensity profile of each vertical section is filtered with a Gaussian smoothing kernel ($\sigma = 4$) to eliminate small fluctuations before the search for the most significant peaks.

It is considered the IPL corresponds to the first maximum of each horizontal intensity profile in this restricted area, the INL to the first minimum and the

OPL to the next maximum. After peak identification, IPL-INL, INL-OPL and OPL-ONL boundaries are detected as the point of maximum slope among the corresponding peaks. Finally, the points detected in the 32 vertical sections are fitted by a spline of grade 6.

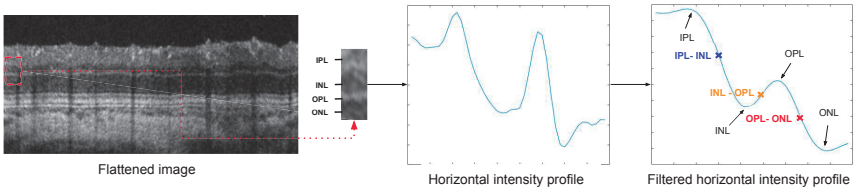


Figure 6: Graphic representation of IPL-ONL detection. IPL corresponds to the first maximum of the horizontal profile, INL to the first minimum and OPL to the next maximum. IPL-INL, INL-OPL and OPL-ONL boundaries are detected as the point of maximum slope among these points.

2.2.6. Retinal Pigment Epithelium (RPE) Detection

Inherent choroid characteristics, the lack of continuity in RPE and its proximity with other layers make either the method proposed for ILM detection or that based on the peaks of the horizontal intensity profile inappropriate for RPE identification. Therefore, a new variant of the marker-controlled watershed that combines both approaches is proposed. We named this variant enclosed watershed transformation and it is based on adapting the markers as their reliability grade to avoid leaks of the watershed lines caused by discontinuity of the RPE which would provoke subsequent errors in the segmentation.

First, the area for RPE search is restricted to the area between the IS-OS and a fixed threshold above the lower retina limit (50 pixels). See Figure 7-1. Due to the fact that the RPE is a very thin layer, instead of detecting its upper and lower edge, the goal of the enclosed watershed is to detect the middle line of the RPE. For that reason, the watershed transformation is not applied on a gradient image as usual but on the result of a morphological top hat performed with a vertical structuring element (Figure 7-3). Top-hat transformation is the residue between the original image and its morphological opening and its aim is to sup-

press slow trends and enhance the contrast of some features in images, according to size and shape criteria. In this case, it is used to extract the light contrasted components of specific shape and size. The size of the structuring element must be a bit lager than the RPE width. Prior knowledge about RPE characteristics is used to determine these parameters (horizontal light lines of less than 5 pixel width). To remove small structures and get a cleaner background, a supremum of 3 openings calculated using a 7-pixel line of angle 0°, +25°and -25°, is applied (Figure 7-4). Afterwards, the horizontal intensity profiles of the 32 vertical sections are computed (Figure 7-5) and a small intensity image of 32 pixel width is composed by combining the resulting horizontal profiles (Figure 7-6). Marker-controlled watershed will be then applied on this image, obtaining the markers to be used as follows. Note that each horizontal profile is individually analysed to determine those markers. Let $P_i(x)$ be the horizontal profile of the vertical section i and $M_i = P_i(x_{max})$ the most prominent maximum of the profile. The marker f_{mrk} of a section i is computed as:

$$f_{mrk_i} = \begin{cases} m_i(x) & \text{if } \bigcap_{i-2 \leq k \leq i+2} \{M_k >= t\} \\ n_i(x) & \text{if } Otherwise \end{cases}, \tag{2}$$

t being a fixed threshold used to determine the minimum value of M_i to be a representative maximum. However, a representative maximum M_i only is considered reliable if the most prominent maximum of its k neighbouring sections also are representative. In that case, the marker of a section i with a representative and reliable maximum is defined as:

$$m_i(x) = \begin{cases} 0 & \text{if } x \in (x_{max} - h, x_{max} + h) \\ 1 & Otherwise. \end{cases}, \tag{3}$$

and if the maximum is not reliable, the marker would be equal to:

$$n_i(x) = \begin{cases} 1 & \text{if } x = 1 \cup x = end \\ 0 & Otherwise. \end{cases}, \tag{4}$$

where h is a 10% of the height of the region of interest where RPE is being searched. The final marker image will be formed by the concatenation of the markers of the 32 sections, $f_{mrk} = \left\{ f_{mrk_i} \right\}_{i=1,\ldots,32}$. In that way, the markers will closely enclose the RPE when the maxima are reliable, forcing the path of the watershed lines, and allow a major degree of freedom when their reliability is unknown. The more reliable maxima are, the path of the watershed will be more restricted and the possible leaks because of the discontinuity of the RPE will be hindered.

Marker-controlled watershed is applied on the small intensity image using the enclosed markers previously computed (Figure 7-6). Due to the size of the input image, enclosed watershed transformation is very efficient computationally. Each pixel of the resulting enclosed watershed boundary will correspond with the central pixel of the 32 vertical sections which will be finally fitted by a spline of grade 6 (Figure 7-7). As seen throughout this subsection, Figure 7 summarises the process for RPE detection.

2.3. Method 2: Fully convolutional neural networks (M2)

With the aim of comparing a methodology based on classical image-processing techniques (see Section 2.2) with one of the most trending technique in the computer vision field (i.e. deep learning), an encoder-decoder fully convolutional network built upon ReLayNet [25] is presented. To the best of the author's knowledge, this is the first FCN architecture designed with the ability of automatically capture the relevant information of the thinnest retinal layers of rodent eyes.

2.3.1. Encoder-decoder architecture

The proposed network architecture able to segment five different retinal layers from rat OCTs can be observed in Figure 8. In particular, it is composed by a contracting path in charge of encoding the relevant information into feature maps and an expansive path responsible to transform the stacked features resulting from the last encoder block into probability maps with the same dimensions

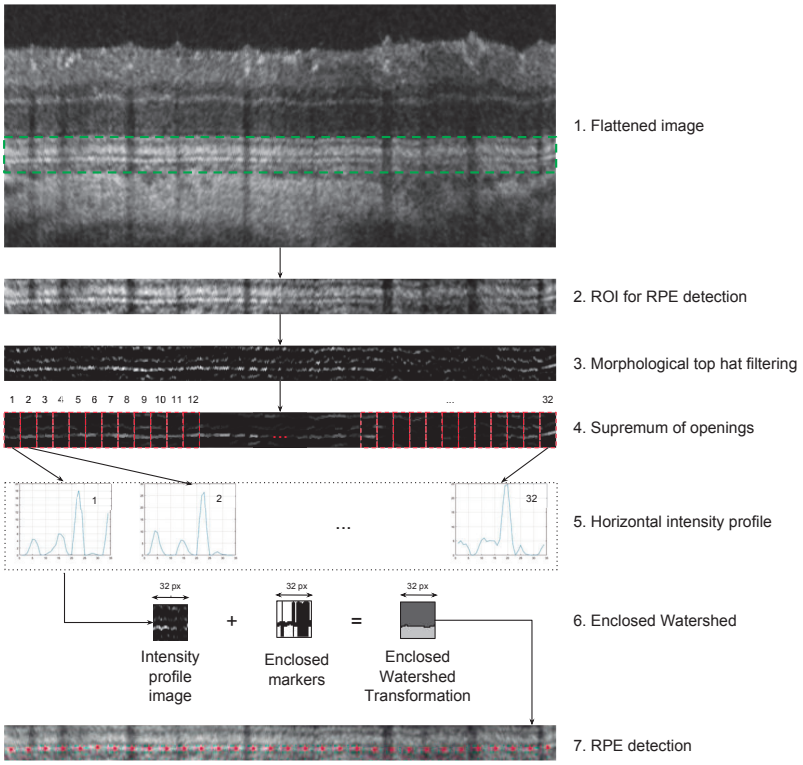


Figure 7: Graphic representation of RPE detection. The horizontal intensity profiles of the 32 sections are computed after image filtering. The combination of these 32 profiles generates the intensity profile image where the enclosed watershed transformation will be applied. The enclosed markers are determined according to the representativity and reliability of the maxima of each profile.

as the input image, recovering the spatial information.

Each encoder block is composed by four layers, in sequence: convolution layer, batch normalization layer, ReLU activation layer (represented together by the blue layer in Figure 8) and max-pooling layer. 64 rectangular kernels of 7×3 are defined in the convolutional layers of all encoder blocks. This kernel size is established with the aim of capturing the transitions between the different retinal layers. Note that zero-padding is used in each convolutional layer to guarantee the preservation of the feature map dimension after the convolution. In

addition, batch normalization technique is applied to avoid overfitting [14] and ReLU activation function is used to introduce non-linearities [23]. Max-pooling operation is introduced in the last layer of each encoder block. In particular, this layer simplifies the feature information reducing the spatial dimensions by half (i.e. 2×2 kernels with a stride equal to two). The pooling indexes of this operation are achieved and transferred to the corresponding unpooling layer in the decoder block to preserve spatial consistency.

Decoder blocks consist of five layers, in sequence: unpooling layer, concatenation layer, convolution layer, batch normalization and ReLU activation function. The unpooling layer upsamples the feature maps from the previous decoder block to a double resolution by using the achieved pooling indexes corresponding to the matched encoder block. After this step, a concatenation of the upsampled feature maps with the corresponding output feature maps of the matched encoder block is performed to enrich the information and avoiding vanishing gradient problems [28]. Finally, convolutional layer, batch normalization and ReLu are applied to the concatenated feature map. Note that the kernel size remains constant for all the decoder blocks and it is the same as we used in the encoder path (i.e. 7×3).

The final decoder block is responsible to associate each pixel to one of the seven possible classes (i.e. upper bound, RNFL+GCL+IPL, INL, OPL, ONL, IS+OS , bottom bound). For this purpose, it is composed by a convolutional layer with 1×1 kernel and the softmax activation function.

2.3.2. Training process

Attending to Section 2.1, the available database is composed of two different batches. The first batch of the database was used for training and validation purposes and the second one for testing. Therefore, in this section, only the first batch of the database was used.

Data conditioning. Training the proposed network with whole OCT images would require several GPUs working in parallel to provide the required RAM for this task. In order to avoid memory problems in the training stage,

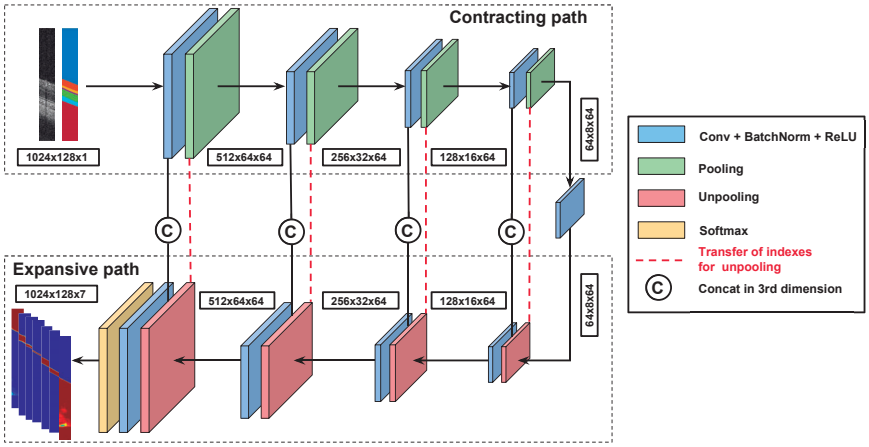


Figure 8: Proposed encoder-decoder architecture. In the contracting path, blue layers represent a convolution with a kernel 7×3 (stride equal to 2 and 64 filters) followed by the batch normalization operation and the ReLU activation function. The maxpooling layer defined by a kernel size of 2×2 and a stride of 2 is represented in green. Regarding to the expansive path, red layers correspond to the unpooling operation responsible of upsampling the feature maps from the previous decoder block to a double resolution (i.e. kernel size 2×2). For this purpose, the achieved pooling indexes corresponding to the matched encoder block are needed (this transference is illustrated by the dotted red lines.). After this step a concatenation of the upsampled feature map with the feature map of the corresponding encoder block is carried out. The final yellow layer is a convolutional layer with 1×1 kernel and the softmax activation function.

a patch-wise learning methodology is used in this work. This fact implicates a notable increasing of the database instances. In particular, using 1024×128 patches, a total of 1032 instances are generated. Further, we augment the sliced data by introducing random geometric transformations such as croppings, horizontal flips, rotations and translations. Memory requirements in the testing stage are more permissive allowing to predict a test sample in two slices of 1024×512 with a Titan V GPU.

Data partitioning. As Section 2.1 details, the first batch of the dataset is composed by 129 images coming from ten different rodents. With the aim of avoiding biased results due to a specific partition of the database into training and test subsets and obtaining the segmentation of the 129 images, an external

K-fold cross-validation technique was carried out. Specifically, $K = 8$ partitions were created according to the number of images acquired from each rat. The images belonging to rats 1-7 composed seven individual partitions while the images from the last three rats were grouped forming the eighth fold (see Table 3). Consequently, in the training process, $K - 1$ different folds in each external iteration were used while the remaining partition was utilized to test the model performance. In addition, an internal leave-one-out cross-validation was carried out, using the images from one different training fold in each internal iteration as validation set. As a result of this process, $(K - 1) \times K$ models were learned. This technique guarantees reliable results and robust models.

Table 3: Distribution of the images belonging to the first batch of the database throughout the eight folds.

Rat ID	#images	Fold
C23 GP1A A1	16	1
C23 GP1A A2	17	2
C23 GP1A A3	16	3
C23 GP1A A4	16	4
C24 GP2B A1	16	5
C24 GP2B A2	16	6
C24 GP2B A3	16	7
C24 GP2B A4	4	8
C25 GP1A A1	8	8
C25 GP2B A3	4	8

Loss function. The proposed network is trained by optimizing a weighted multi-class logistic function. This logistic loss provides a probabilistic measure of similarity between the prediction and the ground truth. Let $p_c(x)$ be the probability for the pixel x of belonging to the class c and $g_c(x)$ the ground truth probability, the basic logistic loss function can be defined as:

$$L(x) = -\sum_{x} \omega(x) g_c(x) log(p_c(x)) \tag{5}$$

The weights $\omega(x)$ are introduced in loss function with two objectives: (i) compensating the effects produced by the imbalance character of the classes. The different thickness of each retinal layer and the huge quantity of background pixels propitiates imbalanced pixel-wise samples. (ii) boosting the accurate segmentation of the retinal layer boundaries. The pixels near to the tissue-

transition regions are the most challenging cases to be correctly segmented due to diffused boundaries, speckle noise and the limited OCT resolution. Given the cumulative frequency f_c of class c in the training data (i.e. the prior class probability), the logic operator I, a ground truth image M and the 2D gradient operator Δ, the $\omega(x)$ term is formulated as:

$$\omega(x) = \sum_{c} I(M(x) = c) \cdot \frac{\tilde{\mathbf{f}}}{f_c} + \omega_0 \cdot I(|\Delta M(x)| > 0) \tag{6}$$

where $\mathbf{f} = [f1, f2, ..., f_c]$ is a vector containing all frequencies. The first term models median frequency balancing and compensates for the classes imbalance problem by enhancing classes with low probability [4]. The second term puts higher weight on anatomical boundary regions to emphasize on the correct segmentation of contours [26]. ω_0 balances both terms.

Hyper-parameter configuration. The proposed network is learned through the stochastic gradient descent (SGD) optimizer using mini batches of eight samples. A momentum value of 0.9 is set to compensate for this small batch size. The learning rate is initially established to 0.001 and reduced by one order after every 20 epochs. The training stage is composed by sixty epochs. As it was mentioned above, the Titan V GPU was used to carry out this process. An overview of the whole training/testing process can be observed in Figure 9.

3. Results

3.1. First batch images: Performance evaluation

To evaluate the performance of the proposed approaches, the results of our segmentation methods (M1 and M2) obtained on the first batch of the database are directly compared to other state-of-the-art approaches making use of the expert manual segmentation (ground truth). The two other compared methods are the proposed by Chiu et al. [8] and the commercial Insight software from Phoenix Research Labs which offers a tool for automatic segmentation. Note that Chiu's method was focused on human OCT segmentation. This method was chosen with the aim of checking if a reproducible and open source method

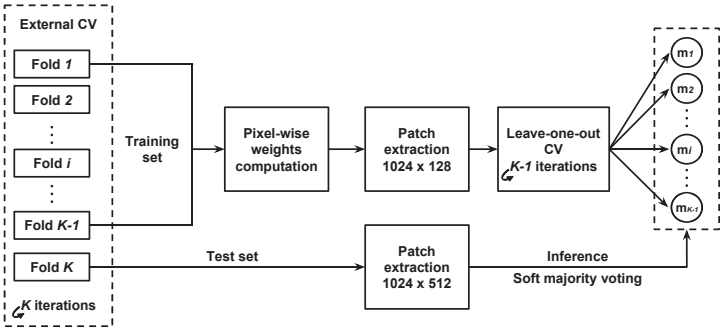


Figure 9: Pipeline of the training/testing process. The data is divided in $K = 8$ folds to remove the bias of a specific data partitioning. Pixel-wise weights for the training images are computed using the ground-truth to mitigate the class imbalance problem and to emphasize on the correct segmentation of contours. Slices of 1024×128 pixels are extracted from the images to learn the models. An internal leave-one-out cross-validation results on a committee of $K - 1$ trained networks.

that is widely referred by the research community for human OCT segmentation was able to adapt itself and achieve good results on rodent OCTs.

First, for a qualitative evaluation, Figure 10 shows three examples of representative segmentation results. Secondly, quantitative measures were computed: absolute distance error between layer boundaries, layer thickness error and Dice similarity coefficient. See Tables 4-6. Due to the fact that the absolute distance error of the Insight software was the greatest among the compared methods, an overall error of 10.97 ± 6.51 μm, and not all the layers are segmented, the value of layer thickness and Dice coefficient of this method was not included in Tables 5 and 6.

Next, the results of the presented approaches were also compared with those obtained by [31], [2] and [3], three methods specifically designed for rodent OCT segmentation (Table 7-8). Same measures performed in those works were calculated. These methods were tested on a different image database thus the comparison is not fully reliable. However, it is useful to see if the performance of our algorithms is similar to other state-of-the-art methods. Note that, to best of author's knowledge, there is no a publicly available rat OCT segmentation

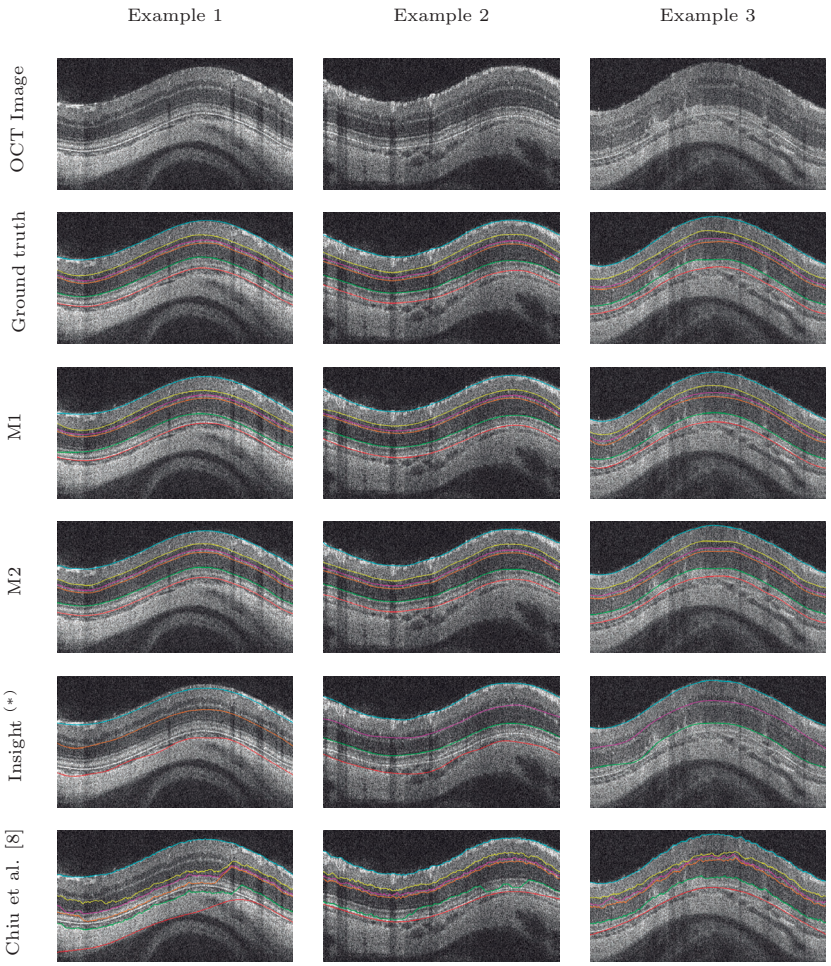


Figure 10: Automated segmentation results on three representative examples of the first batch of the rodent OCT images using different algorithms. Color code: ILM (cyan), IPL-INL (yellow), INL-OPL (magenta), OPL-ONL (orange), IS-OS (green), RPE (red). (*) Note that Insight software does not segment a fixed number of layers and either identifies the layer that is being segmented, thus it is considered that the segmented layer corresponds to the closest ground truth layer.

method that can be tested on our own images to perform a direct comparison because most methods designed for this goal are not open access or require a specific image format. For example, the code of the ASiMOV tool presented in [3] was made available but it only works for Bioptigen images. In addition,

Table 4: Absolute distance errors (mean and standard deviation in μm) calculated comparing the results of the proposed methods (M1 and M2), InSight software and Chiu et al. [8] with the ground truth on the first batch of database. Color code used in Figure 10: ILM (cyan), IPL-INL (yellow),INL-OPL (magenta), OPL-ONL (orange), IS-OS (green), RPE (red). $^{(*)}$ Note that Insight software does not segment all layers under analysis. Only the segmented layers are taken into account to calculate the absolute distance error (out of 129 OCT images, ILM is only detected in 113, IPL-INL in 26, ONL-OPL in 48, OPL-ONL in 19, IS-OS in 70 and RPE in 75).

Layer Boundaries	M1	M2	InSight $^{(*)}$	Chiu et al.[8]
ILM	3.52 ± 1.11	1.79 ± 0.55	6.97 ± 4.59	2.57 ± 0.58
IPL-INL	3.65 ± 1.30	2.61 ± 0.72	10.49 ± 4.06	11.50 ± 14.08
INL-OPL	2.73 ± 1.91	1.71 ± 1.00	7.90 ± 3.70	9.04 ± 17.20
OPL-ONL	3.19 ± 2.12	1.76 ± 0.92	7.99 ± 5.53	7.87 ± 16.90
IS-OS	6.18 ± 5.93	1.89 ± 1.12	8.67 ± 5.90	14.75 ± 26.84
RPE	3.33 ± 3.15	1.62 ± 1.17	23.81 ± 15.29	11.62 ± 20.00
Overall	3.77 ± 2.59	$\mathbf{1.90 \pm 0.91}$	10.97 ± 6.51	8.19 ± 15.93

Table 5: Thickness errors (mean and standard deviation) calculated comparing the results of the proposed methods (M1 and M2) and Chiu et al. [8] with the ground truth (GT) on the first batch of database. Color code used in Figure 10: RNFL+GCL+IPL (cyan-yellow), INL (yellow-magenta),OPL (magenta-orange), ONL (orange-green), IS+OS (green-red)

Layers	Average thickness (μm)				Absolute thickness errors (μm)			Relative thickness errors		
	GT	M1	M2	Chiu et al.	M1	M2	Chiu et al.	M1	M2	Chiu et al.
RNFL+GCL+IPL	66.81	65.17	66.42	71.66	1.92 ± 1.49	1.51 ± 1.10	7.96 ± 7.77	0.028 ± 0.022	0.022 ± 0.016	0.116 ± 0.114
INL	24.81	22.41	24.85	19.77	2.70 ± 2.02	1.53 ± 1.18	5.79 ± 4.62	0.106 ± 0.080	0.060 ± 0.046	0.228 ± 0.182
OPL	12.12	11.70	12.67	12.59	1.39 ± 1.09	0.91 ± 0.85	2.21 ± 2.51	0.112 ± 0.088	0.073 ± 0.069	0.178 ± 0.202
ONL	60.67	67.50	60.13	72.62	6.98 ± 6.18	1.30 ± 1.17	12.42 ± 11.45	0.112 ± 0.100	0.021 ± 0.019	0.200 ± 0.184
IS+OS	38.79	35.54	38.63	37.58	4.99 ± 5.45	1.35 ± 1.24	10.66 ± 15.94	0.126 ± 0.137	0.034 ± 0.031	0.269 ± 0.402
Overall	203.20	202.31	202.69	214.22	3.59 ± 3.24	$\mathbf{1.32 \pm 1.11}$	7.81 ± 8.46	0.097 ± 0.085	$\mathbf{0.042 \pm 0.036}$	0.198 ± 0.217

neither the code nor the images used in [31] and [2] are publicly available.

Finally, a time analysis was performed on an Intel i7 @ 3.10 GHz of 16 GB of RAM with a Titan V GPU. Note that both methods were executed in MATLAB 2018a and MatConvNet framework was used [30] for M2. The averaged prediction time was 1.1668 and 5.1687 seconds for M1 and M2, respectively.

3.2. Second batch images: Generalization ability of the encoder-decoder FCN

As it was observed in Section 3.1, M2 is the method that obtained the best results. To analyse the generalization ability of this method, a new test was carried out. The segmentation of images belonging to the second batch of the database was predicted using the models trained with the first batch. To obtain

Table 6: Dice similarity coefficient (mean and standard deviation) calculated comparing the results of the proposed methods (M1 and M2) and Chiu et al. [8] with the ground truth on the first batch of database.

Layers	M1	M2	Chiu et al.
RNFL+GCL+IPL	0.945 ± 0.015	0.967 ± 0.008	0.899 ± 0.098
INL	0.866 ± 0.054	0.913 ± 0.029	0.703 ± 0.164
OPL	0.770 ± 0.102	0.865 ± 0.057	0.674 ± 0.168
ONL	0.927 ± 0.050	0.969 ± 0.019	0.856 ± 0.154
IS+OS	0.865 ± 0.108	0.954 ± 0.028	0.707 ± 0.226
Overall	0.875 ± 0.066	$\mathbf{0.934 \pm 0.028}$	0.768 ± 0.162

Table 7: Comparison between the results of the proposed methods (M1 and M2) and ACWOE-SW [31] based on thickness errors. Analysis carried out on different datasets.

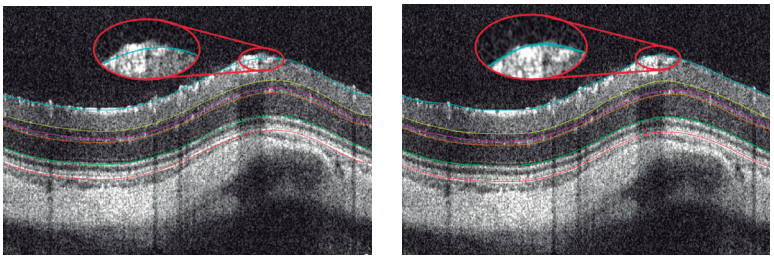
Layers	Absolute thickness error (μm)			Relative thickness error		
	M1	M2	ACWOE-SW	M1	M2	ACWOE-SW
RNFL+GCL+IPL	1.92 ± 1.49	1.51 ± 1.10	3.76 ± 1.47	0.03 ± 0.02	0.02 ± 0.02	0.06 ± 0.02
INL	2.70 ± 2.02	1.53 ± 1.18	2.94 ± 1.11	0.11 ± 0.08	0.06 ± 0.05	0.16 ± 0.06
OPL	1.39 ± 1.09	0.91 ± 0.85	3.06 ± 1.38	0.11 ± 0.09	0.07 ± 0.07	0.14 ± 0.06
ONL	6.98 ± 6.18	1.30 ± 1.17	2.54 ± 1.16	0.11 ± 0.10	0.02 ± 0.02	0.04 ± 0.02
IS+OS	4.99 ± 5.45	1.35 ± 1.24	3.43 ± 1.77	0.13 ± 0.14	0.03 ± 0.03	0.09 ± 0.05
Overall	3.59 ± 3.24	$\mathbf{1.32 \pm 1.11}$	3.15 ± 1.37	0.10 ± 0.09	$\mathbf{0.04 \pm 0.04}$	0.10 ± 0.04

the ground truth of images of the second batch and validate the results obtained on these images, M1 was used. As it was explained in Section 2.2, M1 was an unsupervised method based on classic image processing techniques so it does not require specific training stage based on labelled images. However, as it can be observed in Figure 11(a), the M1 method made some minor faults in layer segmentation so that small manual modifications should be performed to obtain a quality ground truth. The final ground truth is shown in Figure 11(b).

First, for a qualitative evaluation, Figure 12 shows three examples of representative segmentation results on the second batch of the database. There

Table 8: Comparison between the results of the proposed methods (M1 and M2), the work presented by Antony et al. [2] and ASiMOV [3] based on absolute distance error. Analysis carried out on different datasets.

Layer boundaries	M1	M2	Antony et al.	ASiMOV
ILM	3.52 ± 1.11	1.79 ± 0.55	2.48 ± 0.72	1.40 ± 0.38
IPL-INL	3.65 ± 1.30	2.61 ± 0.72	3.24 ± 0.83	2.90 ± 1.06
INL-OPL	2.73 ± 1.91	1.71 ± 0.67	2.59 ± 0.71	2.02 ± 0.67
OPL-ONL	3.19 ± 2.12	1.76 ± 0.92	3.98 ± 1.33	2.34 ± 0.77
IS-OS	6.18 ± 5.93	1.89 ± 1.12	2.89 ± 0.80	2.28 ± 1.23
RPE	3.33 ± 3.15	1.62 ± 1.17	3.72 ± 1.15	1.79 ± 1.48
Overall	3.77 ± 2.59	1.90 ± 0.91	3.15 ± 0.92	2.12 ± 0.93



(a) (b)

Figure 11: Workflow for ground truth generation of the second batch of database: (a) Segmentation obtained by M1 (b) Ground truth after of the manual modification.

exists more variability between images belonging to this batch than to the first one. In addition, these images present more artefacts (large black spots) that compromise the correct definition of the different retinal layers in OCT images.

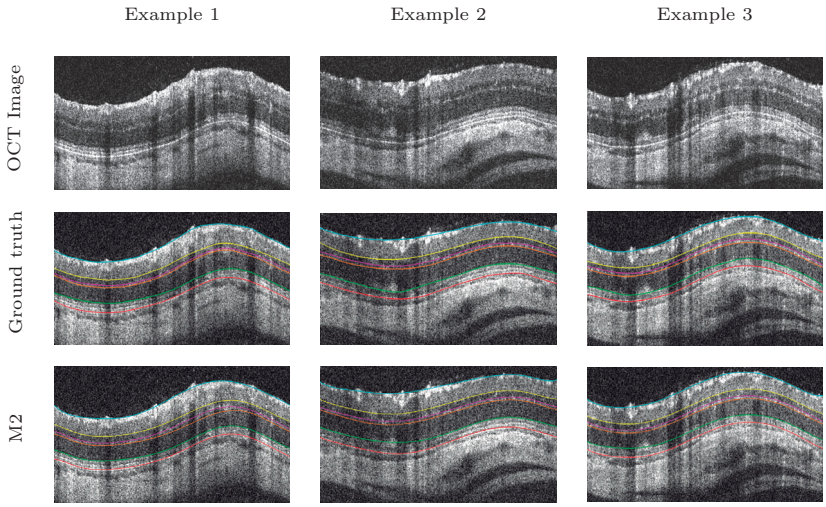


Figure 12: Segmentation results on three representative examples of the second batch of the rodent OCT images. Color code: ILM (cyan), IPL-INL (yellow), INL-OPL (magenta), OPL-ONL (orange), IS-OS (green), RPE (red).

Secondly, the same measures computed in Section 3.1 were presented. See Tables 9-10.

Table 9: Absolute distance errors (mean and standard deviation in μm) calculated on the second batch of the database.

Layer Boundaries	M2
ILM	2.64 ± 0.87
IPL-INL	2.85 ± 0.94
INL-OPL	2.74 ± 1.51
OPL-ONL	2.34 ± 1.12
IS-OS	2.33 ± 1.12
RPE	2.82 ± 2.17
Overall	2.67 ± 1.25

Table 10: Average thickness, Absolute thickness error and Relative thickness error calculated on the second batch of database.

Layers	Average thickness (μm)		Absolute distance error (μm)	Relative thickness error (μm)	DICE
	GT	M2	M2	M2	M2
RNFL+GCL+IPL	68.58	66.80	2.72 ± 1.79	0.04 ± 0.03	0.96 ± 0.01
INL	23.07	25.81	23.24 ± 1.51	0.14 ± 0.06	0.89 ± 0.03
OPL	12.02	11.98	1.14 ± 0.99	0.09 ± 0.08	0.77 ± 0.08
ONL	63.71	61.63	2.36 ± 2.02	0.04 ± 0.03	0.96 ± 0.02
IS+OS	39.76	38.18	2.97 ± 2.39	0.07 ± 0.06	0.94 ± 0.03
Overall	207.16	204.41	2.49 ± 1.74	0.08 ± 0.05	0.90 ± 0.03

4. Discussion

A wide validation of the proposed methods was performed in the previous section. Methods M1 and M2 achieve more similar results to the ground truth than the other analysed solutions, as shown in Figure 10 and Tables 4-6. It must be emphasized that the commercial software Insight does not segment a fixed number of layers and its segmentation is not sufficiently accurate. Thus, most times, requires manual segmentation or modification through some specific tool as that provided in the software. The results achieved by Chiu et al. are also less accurate compared to those obtained by the proposed methods although it should be taken into account that this method was not initially designed for rodent OCT segmentation but for humans. Therefore, it makes sense that the best results were achieved in the ILM identification because is where the difference between rodents and humans is less appreciable. Regarding absolute distance errors (Table 4), M1 achieves the best results in the INL-OPL (magenta) and OPL-ONL (orange) identification and M2 in INL-OPL (magenta) and RPE (red). Concerning thickness errors (Table 5), the first layer

formed by RNFL, GCL and IPL (from the cyan to the yellow boundary) is that whose error is the lowest with M1 and the same layer along with ONL (from the orange to the green boundary) achieve the best results with M2. The same is appreciated in Table 6 as for the Dice similarity coefficient. Based on the data of those tables, M2, the deep-learning-based approach, outperforms M1.

As it was mentioned in Section 2.3, M2 is built upon ReLayNet architecture [25] although it had to be modified to adapt itself to rodent OCTs by providing the network with a deeper level through the addition of an extra convolutional block. This new architecture is key to achieve the required precision to detect the thinnest retinal layers. Other essential contributions that improve the segmentation results on rodent layers were to modify the loss function used in ReLayNet and optimize the encoder-decoder CNN performance following a grid-search methodology to find the optimal learning rate value, optimizer, batch size and number of epochs. All these modifications make that the performance of the proposed method was quite different to that obtained by the original ReLayNet on rodent images as Figure 13 shows. That figure compares the results of the original ReLayNet architecture with those obtained by our proposed architecture on the same representative examples of the Figure 10. As can be observed, in spite of the promising results of ReLayNet in human OCTs, if the additional convolutional block is not included in the rodent's case, the network has not discriminator capability to detect the OPL, that is the thinnest layer of those analyzed (shown in purple color in the results of M2) and the remaining layers are misidentified, so the architecture proposed in M2 can be considered as an RelayNet improvement. In fact, the authors already demonstrated in a previous publication that the original Relaynet architecture, even followed by a robust post-processing stage, only is able to obtain satisfactory results for the segmentation of the three most significant retinal layer boundaries instead of the six boundaries detected in this work [1]. Thus, it is verified that the direct transfer of the method from human retinal OCT images to rodents images does not properly work when the thinnest retinal layers are involved.

Due to the lack of public rodent OCT databases, M1 and M2 could not be

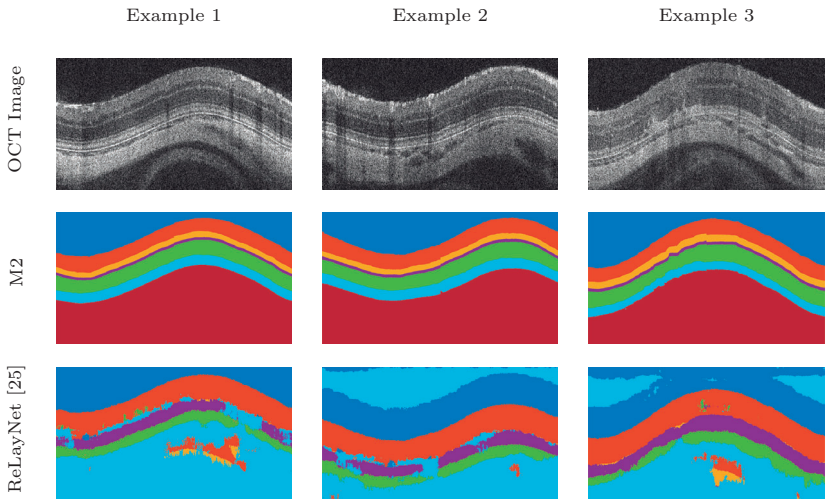


Figure 13: Comparison of M2 and ReLayNet [25] performance on the same representative examples shown in Figure 10.

tested on the same images used in other state-of-the-art methods originally proposed for rodent retinal layer segmentation. Thus, Table 7 and 8 are not a fully reliable evaluation but they are useful as an indirect comparison. From these tables, it is possible to see that M1 has a similar performance to ACWOE-SW [31] and M2 reduces significantly its average relative thickness error. Moreover, the averaged absolute distance error of M2 is even lower than ASiMOV [3] and that proposed by [2]. To avoid this fact in the future and facilitate direct comparisons, we have made our own database publicly accessible.

Regarding prediction time, it must be emphisized that, although M1 is based on classical image processing techniques, it is faster than M2 when $m = K - 1$ prediction models are used (1.1668 vs 5.1687 s). If only one model was used for M2, the time would have to be divided by the number of models ($m = 7$ in this work), resulting in 0.7383 s per image. In that case, the prediction with both models is almost immediate. However, M1 does not require a training stage with a high number of manually segmented images which makes possible to use M1 to facilitate ground truth generation.

Concerning the alternative test with the second batch of the database, its

results are very rewarding in spite of the variety and artefacts existing in this batch, as is shown in Figure 12. This demonstrates that M2 can be generalized and does not need to be retrained to segment images belonging to different databases to those previously used to train the model if the time, hardware or software required is not available. In addition, the use of M1 to obtain a first approximation of the ground truth in these images considerably reduces the time required to perform a manual segmentation because only requires small modifications.

5. Conclusion

Automated OCT analysis in rats is of great relevance because it can facilitate fast evaluation of data from a large study population of animals. This fact is especially important when screening for a possible toxic effect of drugs and other treatments before human trials. In that context, this paper proposes two approaches to segment six retinal layer boundaries in rat OCT images, one is a more conventional approach based on classical image processing techniques and the other makes use of deep learning algorithms. The conventional approach is based on a combination of local horizontal intensity profiles along with two variants of the watershed transformation. One of these variants, the enclosed watershed, is presented for the first time in this work and could be satisfactorily used in other cases where the segmentation of boundaries that lack continuity is required. The deep-learning-based method is an adaptation of the FCN architecture presented in [25] which allows increasing the network precision to detect the rodent retinal layers. Although the proposed network in this paper is built upon ReLayNet architecture, it also incorporates own contributions such as the addition of an extra convolutional block to encode the information of the very thin layers and the modification of the loss function, which are essential to obtain satisfactory segmentation results in rodent OCTs. After a wide validation, it was demonstrated that both methods significantly outperform the commercial Insight image segmentation software (Phoenix Research Labs), ob-

taining an averaged absolute distance error of 3.77 ± 2.59 and 1.90 ± 0.91 μm, respectively, on the first batch of the database. Moreover, an averaged absolute distance error of 2.67 ± 1.25 μm on the second batch of the database confirms that the deep-learning-based method can be used to generate the segmentation of images of other rodent OCT databases previously unseen and with different conditions to that used for training.

Because of the first approach proposed in this paper is an unsupervised method can be used as an easy way to generate labelled images. Note that, in that case, the labelled images should to be manually reviewed and modified if required but, taking into account that labelling process is very time-consuming and the results of M1 are very competitive, it significantly speeds up the process.

It is a fact that the limited existence of rodent OCT databases in the literature conditions the evaluation of retinal layer segmentation methods applied to rat OCT images. With the aim of facilitating further comparisons and avoid having to use different databases to evaluate the goodness of two segmentation methods, the dataset used in this work along with their ground truth (expert-reviewed segmentations) were made publicly available in [22].

From a technical perspective, the future lines of work will focus on adapting the network architecture to segment a greater number of layers through the inclusion of extra convolutional blocks and optimizing the different hyperparameters involved (i.e. number of filters, kernel size, etc.). In the case of having OCT volumes, a 3D fully CNN could be proposed or the segmentation could be performed by capturing information between adjacent scans making use of a CNN + LSTM architecture. After the validation performed in the paper, a clinical study to quantitatively evaluate the retinal layer degeneration caused by the ET-1 injection over the time could be also carried out.

Acknowledgements

Animal experiment permission was granted by the Danish Animal Experimentation Council (license number: 2017-15-0201-01213). We gratefully ac-

knowledge the support of NVIDIA Corporation with the donation of the Titan V GPU used for this research. This work has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation, under grant agreement No. 732613 (GALAHAD Project), the Spanish Ministry of Economy and Competitiveness through project DPI2016-77869 and GVA through project PROMETEO/2019/109.

References

[1] del Amor, R., Morales, S., Colomer, A., Mossi, J.M., Woldbye, D., Klemp, K., Larsen, M., Naranjo, V., 2019. Towards automatic glaucoma assessment: An encoder-decoder cnn for retinal layer segmentation in rodent oct images, in: Proceedings of the 2019 27th European Signal Processing Conference (EUSIPCO).

[2] Antony, B.J., Abràmoff, M.D., Harper, M.M., Jeong, W., Sohn, E.H., Kwon, Y.H., Kardon, R., Garvin, M.K., 2013. A combined machine-learning and graph-based framework for the segmentation of retinal surfaces in sd-oct volumes. Biomed. Opt. Express 4, 2712–2728. doi:10.1364/BOE.4.002712.

[3] Antony, B.J., Kim, B.J., Lang, A., Carass, A., Prince, J.L., Zack, D.J., 2017. Automated segmentation of mouse oct volumes (asimov): Validation & clinical study of a light damage model. PLOS ONE 12, 1–17. doi:10.1371/journal.pone.0181059.

[4] Badrinarayanan, V., Kendall, A., Cipolla, R., 2017. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence 39, 2481–2495. doi:10.1109/TPAMI.2016.2644615.

[5] Bertalmio, M., Bertozzi, A.L., Sapiro, G., 2001. Navier-stokes, fluid dynamics, and image and video inpainting, in: Proceedings of the 2001 IEEE

Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001, pp. I–355–I–362 vol.1. doi:10.1109/CVPR.2001.990497.

[6] Beucher, S., Meyer, F., 1993. The morphological approach to segmentation: the watershed transformation. Mathematical morphology in image processing. Optical Engineering 34, 433–481.

[7] Chen, A., Ong, C.Z.L., Luo, W., Lee, C.F., Chia, S.M., Galvao, J., Metzger, D., Gifford, E., Chin, C.L., Ali, A.A.B., 2020. Comparison of training strategies for the segmentation of retina layers in optical coherence tomography images of rodent eyes using convolutional neural networks, in: Išgum, I., Landman, B.A. (Eds.), Medical Imaging 2020: Image Processing, International Society for Optics and Photonics. SPIE. pp. 821 – 825. doi:10.1117/12.2549442.

[8] Chiu, S.J., Li, X.T., Nicholas, P., Toth, C.A., Izatt, J.A., Farsiu, S., 2010. Automatic segmentation of seven retinal layers in sdoct images congruent with expert manual segmentation. Opt. Express 18, 19413–19428. doi:10.1364/OE.18.019413.

[9] Dufour, P.A., Ceklic, L., Abdillahi, H., Schroder, S., Dzanet, S.D., Wolf-Schnurrbusch, U., Kowal, J., 2013. Graph-based multi-surface segmentation of oct data using trained hard and soft constraints. IEEE Transactions on Medical Imaging 32, 531–543. doi:10.1109/TMI.2012.2225152.

[10] Fang, L., Cunefare, D., Wang, C., Guymer, R.H., Li, S., Farsiu, S., 2017. Automatic segmentation of nine retinal layer boundaries in oct images of non-exudative amd patients using deep learning and graph search. Biomed. Opt. Express 8, 2732–2744.

[11] Garvin, M.K., Abramoff, M.D., Wu, X., Russell, S.R., Burns, T.L., Sonka, M., 2009. Automated 3-d intraretinal layer segmentation of macular spectral-domain optical coherence tomography images. IEEE Transactions on Medical Imaging 28, 1436–1447. doi:10.1109/TMI.2009.2016958.

[12] Gopinath, K., Rangrej, S.B., Sivaswamy, J., 2017. A deep learning framework for segmentation of retinal layers from oct images. CoRR abs/1806.08859.

[13] He, Y., Carass, A., Yun, Y., Zhao, C., Jedynak, B.M., Solomon, S.D., Saidha, S., Calabresi, P.A., Prince, J.L., 2017. Towards topological correct segmentation of macular oct from cascaded fcns, in: Fetal, Infant and Ophthalmic Medical Image Analysis, Springer International Publishing, Cham. pp. 202–209.

[14] Ioffe, S., Szegedy, C., 2015. Batch normalization: Accelerating deep network training by reducing internal covariate shift, in: Proceedings of the 32Nd International Conference on International Conference on Machine Learning - Volume 37, JMLR.org. pp. 448–456.

[15] Jing, T., Boglarka, V., Erika, T., Palya, F., Mark, S.G., E., S.W., Cabrera, D.D., 2016. Performance evaluation of automated segmentation software on optical coherence tomography volume data. Journal of Biophotonics 9, 478–489. doi:10.1002/jbio.201500239.

[16] Kiaee, F., Fahimi, H., Rabbani, H., 2018. Intra-retinal layer segmentation of optical coherence tomography using 3d fully convolutional networks, in: 2018 25th IEEE International Conference on Image Processing (ICIP), pp. 2795–2799.

[17] Kugelman, J., Alonso-Caneiro, D., Read, S.A., Hamwood, J., Vincent, S.J., Chen, F.K., Collins, M.J., 2019. Automatic choroidal segmentation in oct images using supervised deep learning methods. Scientific Reports 9. doi:10.1038/s41598-019-49816-4.

[18] Lang, A., Carass, A., Hauser, M., Sotirchos, E.S., Calabresi, P.A., Ying, H.S., Prince, J.L., 2013. Retinal layer segmentation of macular oct images using boundary classification. Biomed. Opt. Express 4, 1133–1152. doi:10.1364/BOE.4.001133.

[19] Liu, X., Fu, T., Pan, Z., Liu, D., Hu, W., Liu, J., Zhang, K., 2018. Automated layer segmentation of retinal optical coherence tomography images using a deep feature enhanced structured random forests classifier. IEEE Journal of Biomedical and Health Informatics , 1–1doi:10.1109/JBHI.2018.2856276.

[20] Lu, S., l. Cheung, C.Y., Liu, J., Lim, J.H., s. Leung, C.K., Wong, T.Y., 2010. Automated layer segmentation of optical coherence tomography images. IEEE Transactions on Biomedical Engineering 57, 2605–2608. doi:10.1109/TBME.2010.2055057.

[21] Mishra, A., Wong, A., Bizheva, K., Clausi, D.A., 2009. Intra-retinal layer segmentation in optical coherence tomography images. Opt. Express 17, 23719–23728. doi:10.1364/OE.17.023719.

[22] Morales, S., Colomer, A., Mossi, J.M., del Amor, R., Woldbye, D., Klemp, K., Larsen, M., Naranjo, V., 2020. Data for: Retinal layer segmentation in rodent oct images: Local intensity profiles & fully convolutional neural networks. Mendeley Data, v1. doi:10.17632/g42vvvzpwh.1.

[23] Nair, V., Hinton, G.E., 2010. Rectified linear units improve restricted boltzmann machines, in: Proceedings of the 27th International Conference on International Conference on Machine Learning, Omnipress, USA. pp. 807–814. URL: http://dl.acm.org/citation.cfm?id=3104322.3104425.

[24] Pekala, M., Joshi, N., Freund, D.E., Bressler, N.M., DeBuc, D.C., Burlina, P.M., 2018. Deep learning based retinal OCT segmentation. CoRR abs/1801.09749.

[25] Roy, A.G., Conjeti, S., Karri, S.P.K., Sheet, D., Katouzian, A., Wachinger, C., Navab, N., 2017a. Relaynet: retinal layer and fluid segmentation of macular optical coherence tomography using fully convolutional networks. Biomed. Opt. Express 8, 3627–3642.

[26] Roy, A.G., Conjeti, S., Sheet, D., Katouzian, A., Navab, N., Wachinger, C., 2017b. Error corrective boosting for learning fully convolutional networks with limited data, in: Medical Image Computing and Computer-Assisted Intervention MICCAI 2017, Springer International Publishing, Cham. pp. 231–239.

[27] Ruan, Y., Xue, J., Li, T., Liu, D., Lu, H., Chen, M., Liu, T., Niu, S., Li, D., 2019. Multi-phase level set algorithm based on fully convolutional networks (fcn-mls) for retinal layer segmentation in sd-oct images with central serous chorioretinopathy (csc). Biomed. Opt. Express 10, 3987–4002. doi:10.1364/BOE.10.003987.

[28] Shelhamer, E., Long, J., Darrell, T., 2017. Fully convolutional networks for semantic segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence 39, 640–651.

[29] Srinivasan, P.P., Heflin, S.J., Izatt, J.A., Arshavsky, V.Y., Farsiu, S., 2014. Automatic segmentation of up to ten layer boundaries in sd-oct images of the mouse retina with and without missing layers due to pathology. Biomed. Opt. Express 5, 348–365. doi:10.1364/BOE.5.000348.

[30] Vedaldi, A., Lenc, K., 2015. Matconvnet – convolutional neural networks for matlab, in: Proceeding of the ACM Int. Conf. on Multimedia.

[31] Yazdanpanah, A., Hamarneh, G., Smith, B.R., Sarunic, M.V., 2011. Segmentation of intra-retinal layers from optical coherence tomography images using an active contour approach. IEEE Transactions on Medical Imaging 30, 484–496. doi:10.1109/TMI.2010.2087390.