Shedding Light on Early Central Nervous System Changes for Alzheimer's Disease through the Retina: An Animal Study

Rui Bernardes[1,2][a], Hugo Ferreira[1][b], Pedro Guimarães[1][c] and Pedro Serranho[1,3][d]

[1]*Coimbra Institute for Biomedical Image and Translational Research, Faculty of Medicine, University of Coimbra, Edifício do ICNAS, Polo 3 Azinhaga de Santa Comba 3000-548 Coimbra, Portugal*

[2]*University of Coimbra, CACC - Clinical Academic Center of Coimbra, Faculty of Medicine (FMUC), Coimbra, Portugal*

[3]*Department of Science and Technology, Universidade Aberta, Rua da Escola Politécnica, 147, 1269-001 Lisbon, Portugal*

Keywords: Optical Coherence Tomography, Retina, Biomarkers, Texture, Convolutional Neural Network, Alzheimer's Disease, Mouse Model, Diagnosis.

Abstract: The World Health Organization (WHO) 2015 projections estimated 75.6 million people living with dementia in 2030, an update from the 66 million estimated in 2013. These figures account for all types of dementia, but Alzheimer's disease stands out as the most common estimated type, representing 60% to 80% of the cases. An increasing number of research groups adopted the approach of using the retina as a window to the brain. Besides being the visible part of the central nervous system, the retina is readily available through non-invasive imaging techniques such as optical coherence tomography (OCT). Moreover, cumulative evidence indicates that neurodegenerative diseases can also affect the retina. In the work reported herein, we imaged the retina of wild-type and the triple-transgenic mouse model of Alzheimer's disease, at the ages of one-, two-, three-, four-, eight-, twelve- and sixteen-months-old, by OCT and segmented gathered data using a developed convolutional neural network into distinct layers. Group differences through texture analysis of computed fundus images for five layers of the retina, normative retinal thickness data throughout the observation period of the ageing mice, and findings related to the estimation of the ageing effect of the human genes present in the transgenic group, as well as the classification of individual fundus images through convolutional neural networks, will be presented and thoroughly discussed in the Special Session on "New Developments in Imaging for Ocular and Neurodegenerative Disorders".

1 INTRODUCTION

A biomarker of dementia from a non-invasive, widely available and inexpensive imaging technique is of utmost importance. Biomarkers can provide insights into the underlying pathophysiology and may be used as inclusion criteria and outcome measures for clinical trials (Ahmed et al., 2014). Also, the technique should not be based on the loss of tissue (since it occurs in the late stages of the disease) and have the property of not being specific for the type of dementia. Such a technique is of paramount importance to shed light on the changes unfolding in the retina.

Even though the number of people with dementia is increasing rapidly (Alzheimer's & Dementia, 2020)

[a] https://orcid.org/0000-0002-6677-2754
[b] https://orcid.org/0000-0003-1199-8489
[c] https://orcid.org/0000-0002-9465-4413
[d] https://orcid.org/0000-0003-2176-3923

and several research groups focused on this significant problem using the eye as a window to the brain, two major issues persist with the non-invasive followed approaches to date. First, research groups focused their attention on the measurement of the thickness of the retinal nerve fibre layer (RNFL) of the imaged retina and reported the thinning of the RNFL, even though with inconsistent results. Moreover, this thinning occurs in the late stages of the disease, making it useless as a screening technique. Second, the most common form of dementia –Alzheimer's disease (AD)– can remain undiagnosed for years (Habib et al., 2017; Krantic and Torriglia, 2014) and can only be definitely diagnosed post-mortem (Krantic and Torriglia, 2014), despite that "...the degree of atrophy on MRI correlates well with Braak staging at autopsy." (Jack et al., 2010) (MRI – magnetic resonance imaging).

Both these problems have the same drawback:

used biomarkers are valid only in the late stages of the disease. Thus, none can determine its onset when therapeutic intervention may have effective results on its progression (Hart et al., 2016). These drawbacks call for a distinct approach from the current and foreseen ones (Ahmed et al., 2014).

We took the initial steps towards a potential solution by conducting exploratory tests using a novel approach to tackle the above problems. Using a supervised machine learning algorithm – support vector machine (SVM) – we explored the hypothesis that the human retina's optical coherence tomography (OCT) data embeds information on neurodegeneration not perceivable by direct observation. The underlying rationale is the sensitive nature of OCT to subtle refractive index changes and the amount of data gathered from the ocular fundus. Even though minor differences cannot be individually detected, they may still influence the statistics of the global data. Preliminary studies showed the possibility of discriminating healthy controls by age group and discriminating between healthy controls and patients diagnosed with Alzheimer's, Multiple Sclerosis and Parkinson's disease (Nunes et al., 2019), thus confirming our working hypothesis and rationale.

Keeping in mind the second problem, that AD can only be confirmed post-mortem, the alternative comes from the triple-transgenic mouse model of Alzheimer's disease ($3\times$Tg-AD) (Oddo et al., 2003). It develops age-dependent amyloid-beta and tau pathologies, activated microglia, loss of synapses and neurodegeneration (Kitazawa et al., 2012).

Therefore, in the work herein reported, we rely on the $3\times$Tg-AD to identify the very early signs of disease and monitor the evolution of these signs. This approach allows us to determine which structures (layers) of the retina will be affected, their relative levels of change to the healthy condition and the sequence of events, that is, which structures show differences to the healthy state and in which chronological order.

Our preliminary data using healthy controls (WT – wild-type) and $3\times$Tg-AD mouse model imaged by an OCT allowed concluding that the results are consistent despite the human/mouse biological differences and the differences in instrumentation. It is possible to discriminate between healthy and unhealthy subjects in both cases. These findings pave the way to study further the changes over time from the animal model of disease and look for similar changes in available data from the human retina, considering different pathologies and healthy controls. Interestingly, even though studies on the central nervous system (CNS) in animals do not commonly readily translate to humans, our data suggest a parallelism between OCT data from animals and humans.

The works of La Morgia et al. (Morgia et al., 2016) and Schön et al. (Schön et al., 2012) show, respectively, the accumulation of amyloid-beta deposits in and around melanopsin retinal ganglion cells and the presence of hyperphosphorylated tau in the innermost layers of the retina from AD patients. These findings support the hypothesis that the analysis of fundus projection images computed from OCT data (see (Guimarães et al., 2014)) may embed useful information on the content of the retina concerning CNS changes.

The texture analysis of the computed fundus images for distinct layers of the mouse retina revealed differences in texture between the WT and the $3\times$Tg-AD groups and the nature of these differences over time, suggesting a cumulative effect of the ageing and that of the disease's associated genes for the transgenic mice group.

Despite this exploratory approach, achieved results strongly suggest the presence of useful information on the status of the CNS within the computed fundus references, in agreement with (Morgia et al., 2016; Schön et al., 2012). The importance of such a finding is fivefold. First, texture-based metrics may provide additional information to strengthen the biomarker of CNS neurodegeneration. Second, it may lead to discrimination between different forms of neurodegeneration. Third, it may lead to a very early biomarker of dementia and, in consequence, shed light on the pathogenesis of the disease. Fourth, it may identify the window for treatment before a late-stage where treatment of CNS degeneration is ineffective and, fifth, for the efficient monitoring of therapy efficacy.

Despite the hurdles, we could confirm, over time and using different populations of controls and patients, that gathered data from the ocular fundus embeds valuable information on the human CNS. Furthermore, using WT and the triple-transgenic mouse model of AD and different instrumentation, we could replicate findings in humans and mice solely based on OCT data of the retina, bridging the gap between human and animal findings on the CNS status, based on the fact that the exploratory results are consistent in humans and mice. This bridge is of high interest by opening the possibility of developing biomarkers of AD using non-invasive and widely available instrumentation.

The independent analysis of the distinct retinal layers calls for the segmentation of OCT data, where the images are divided into regions corresponding to retinal layers or layer-complexes and the background

(vitreous and choroid). Over time, several approaches have been proposed to address this challenging problem since the introduction of the OCT in clinical practice, starting with time-domain systems, when the resolution was only a fraction of the current one.

Current methods can be divided into mathematical methods and machine learning approaches. Mathematical methods encode all retinal information known *a priori* into mathematical models towards the segmentation. These models are, however, difficult to obtain, and even more so to capture variations from different mice strains and various pathologies, primarily where significant differences in the structure appearance of the retina (Berger et al., 2014; Soukup et al., 2019) and significant differences in image contrast are found within the same B-scan, across different B-scans, and eyes.

The mathematical approaches aiming to detect retinal interfaces bounding the respective retinal layers include Markov modelling (Koozekanani et al., 2021), active contour modelling (Yazdanpanah et al., 2011), variational methods (Rathke et al., 2014), and graph-based methods (Chiu et al., 2010; Dufour et al., 2013; Srinivasan et al., 2014; Fang et al., 2017). In particular, graph-based approaches have shown promising results for the segmentation of several retinal interfaces. However, graph-based techniques typically use constraints to achieve better segmentation performances (Dufour et al., 2013). Unfortunately, while these constraints help produce good results for the datasets over which the methods were developed, these methods are tuned for a particular dataset. Furthermore, these methods typically make use of pre-processing (e.g. denoising (Srinivasan et al., 2014)) and post-processing steps, as well as heuristics, to achieve reported results (Chiu et al., 2010). More recently, a combination of graph-based methods and machine learning has been explored, where the machine learning output is used as the basis for the graph-search method (Fang et al., 2017).

Machine learning approaches address the problem distinctly. These typically employ pixel classification techniques, either by classifying each pixel as belonging to a layer (semantic segmentation) or a boundary (interface detection). Deep learning methods have gained momentum in general computer vision applications and eye research (Sarhan et al., 2020), especially convolutional neural networks (CNN). Over the last few years, deep learning approaches have been proposed to solve the retinal segmentation problem (Roy et al., 2017; Antony et al., 2017; Ngo et al., 2020; Venhuizen et al., 2017). However, despite achieved results, most published works in this field address the human retina, therefore not suitable for

mice due to the vast differences in the retinal structure between human and rodent retinas.

While there is a growing body of work on deep learning applied to rodents' OCT data, these models are generally trained on a narrow age range and use a single mice strain/pathology (Roy et al., 2017; Morales et al., 2021). Consequently, age-related changes, both in healthy ageing and potentially more so in neurodegenerative diseases, like Alzheimer's disease (Chiu et al., 2012; London et al., 2013), are not considered, again restricting the applicability of these methods to a limited strain, age range, and condition.

In this work, we make use of a deep learning process to address the identification of 10 image regions (7 retinal layers, 1 layer-complex, vitreous, and choroid) while imposing the following constraints into the process: a neural network as small as possible aiming the minimum number of parameters, no significant pre-processing (e.g. no denoising required), take the known organisation into account and cope with a wide range of ages, mice strains, healthy and pathological cases (diabetes and neurodegenerative cases). We leave out anatomical abnormalities like cystoids, retinal detachments, holes, and similar changes in the retina, as these are not within the scope of our research.

2 MATERIALS AND METHODS

2.1 Data

2.1.1 Ethics Statement

This study was approved by the Animal Welfare Committee of the Coimbra Institute for Clinical and Biomedical Research (iCBR), Faculty of Medicine, University of Coimbra, and by *Direção-Geral de Alimentação e Veterinária (DGAV)*. All procedures involving mice were conducted as per the Association for Research in Vision and Ophthalmology statement for animal use, and in agreement with the European Community Directive Guidelines for the care and use of nonhuman animals for scientific purposes (2010/63/EU), transposed into the Portuguese law in 2013 (DL113/2013).

2.1.2 Mouse Characterisation

In this study, 57 mice from each strain, C57BL6/129S and 3×Tg-AD, were used, at the ages of one to four months old and at eight, twelve and sixteen months old, and considering both left and right eyes. All

mice were housed and maintained at the vivarium of iCBR, Faculty of Medicine, University of Coimbra, and were on a 12-h light/dark cycle with free access to both food and water.

2.1.3 Experimental Setup

Concerning OCT imaging preparation, mice were anaesthetised using a mixture of 80 mg/kg of ketamine and 5 mg/kg of xylazine. The pupils were dilated using a solution of 0.5% tropicamide and 2.5% phenylephrine. Additionally, oxibuprocaine, a local anaesthetic, was used. Eyes were regularly lubricated using eye drops (1% carmellose).

All retinas were imaged by a Micron IV OCT System (Phoenix Technology Group, Pleasanton, CA, USA). It creates a volume per acquisition composed of 512 B-scans, each with 512 A-scans of 1024 values in length; B-scans are saved as a non-compressed TIFF file image.

The system presents an imaging depth of 1.4 mm and an axial resolution of 3 μm, as determined by the superluminescent diode' bandwidth and central wavelength, 160 and 830 nm, respectively. All scans were taken by the same operator in the same retinal region using the optic disc as a landmark, centred horizontally with the optic disc and vertically above it.

2.2 Segmentation Neural Network

The neural network for the retina segmentation builds upon the U-Net (Ronneberger et al., 2015). This fully convolutional neural network combines local and global information to create a more accurate image segmentation process. This architecture consists of two symmetric paths, an encoder path, where the size of the feature maps is reduced to capture the global information, and a decoder path, where feature maps produced by the encoder are upscaled to match the input image size. In addition, skip-connections were also used to transfer features maps from the encoding path to the respective level of the decoding path (Figure 1), which allows achieving better data localisation. Finally, the classification block allows for semantic segmentation from the feature maps, where classes represent retinal layers/layer-complex, vitreous, and choroid, towards identifying classes' boundaries (interfaces).

Furthermore, the proposed architecture uses residual learning (He et al., 2016; Zhang et al., 2018; Jha et al., 2019; Diakogiannis et al., 2020). Residual learning has improved neural network learning by helping with derivative propagation during the training process (He et al., 2016).

All convolutions made use of *stride* 1 and *kernel* 3×3 in size. Also, *zero-padding* was used to preserve the dimensions of feature maps in each convolutional layer. Batch Normalisation (BN) and ReLu (Rectified Linear Unit) were used to prevent over-fitting of the neural network (Ioffe and Szegedy, 2015) and introduce non-linearity in the training process. For details on the neural network, see Table 1.
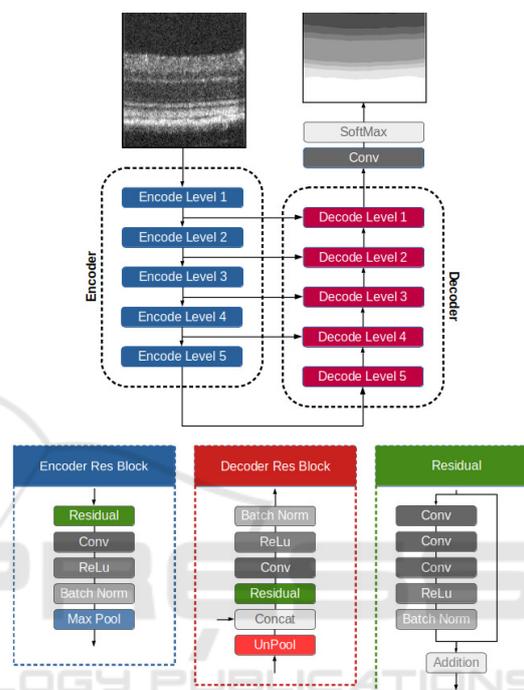
Figure 1: Semantic segmentation neural network scheme employed to segment OCT's B-scans.

2.2.1 Encoding / Decoding Blocks

Each block of the encoding path starts with a residual learning block, followed by a convolution, BN and ReLu activation layers (Figure 1). During this process, the number of channels is doubled. Finally, the encoding block ends with a max-pooling layer to halve feature maps' size, using a 2×2 *kernel* and *stride* 2.

The decoding blocks start with the upscaling of the respective feature maps, combined with the feature maps from the corresponding encoding level. The remaining steps are similar to the encoding blocks. First, a residual block is used, followed by a convolution, BN and ReLu activation layer (Figure 1). In each decoding/encoding block, the number of channels is halved/doubled.

The classification is achieved using a convolutional layer with a 1×1 *kernel* and reduces the number of channels to match the number of classes (10),

Table 1: Network structure.

	Unit Level	Layer / Block	Filter (*Kernel* / channels)	*Stride*	Output Size
Input					768 x 512 x 1
Encoder	Level 1	Residual	3x3 / 16	1	768 x 512 x 16
		Conv	3x3 / 16	1	768 x 512 x 16
	Level 2	Residual	3x3 / 32	1	384 x 256 x 32
		Conv	3x3 / 32	1	384 x 256 x 32
	Level 3	Residual	3x3 / 64	1	192 x 128 x 64
		Conv	3x3 / 64	1	192 x 128 x 64
	Level 4	Residual	3x3 / 128	1	96 x 64 x 128
		Conv	3x3 / 128	1	96 x 64 x 128
	Level 5	Residual	3x3 / 256	1	48 x 32 x 256
		Conv	3x3 / 256	1	48 x 32 x 256
Decoder	Level 5	Residual	3x3 / 256	1	48 x 32 x 256
		Conv	3x3 / 256	1	48 x 32 x 256
	Level 4	Residual	3x3 / 128	1	96 x 64 x 128
		Conv	3x3 / 128	1	96 x 64 x 128
	Level 3	Residual	3x3 / 64	1	192 x 128 x 64
		Conv	3x3 / 64	1	192 x 128 x 64
	Level 2	Residual	3x3 / 32	1	384 x 256 x 32
		Conv	3x3 / 32	1	384 x 256 x 32
	Level 1	Residual	3x3 / 16	1	768 x 512 x 16
		Conv	3x3 / 16	1	768 x 512 x 16
Output		Conv	1x1 / 10	1	768 x 512 x 10
		Output			768 x 512 x 1

followed by a *softmax* activation.

2.2.2 Loss Function

The cross-entropy loss function is used (eq. 1) where ω is the weigh function (with custom weight per pixel), g_l is the ground truth for class l, and p_l is the neural network prediction for class l. This weighting scheme aims two-fold: to produce accurate classifications near interface locations; and reduce the likelihood that misclassifications occur near the interfaces.

The second goal is of particular interest as it is not intended to boost efficiency. Indeed, its contribution in this regard may be negligible. However, it helps the overall method by reducing the number of classification errors close to interfaces that would be difficult, if not impossible, to solve with further processing. On the other hand, errors away from interfaces can be solved based on *a priori* information on the retina, namely that each layer is present (mandatory) for the entire B-scan and that layers are strictly stratified.

$$L_{cross} = -\sum_l \sum_{\mathbf{x}} \omega(\mathbf{x}, l) g_l(\mathbf{x}) \log(p_l(\mathbf{x})). \quad (1)$$

The weight function (ω) is composed of two components: one captures the differences between classes, and another differentiates the relevance of pixels within the same class. The first component is the class weights determined by the inverse of the relative frequency of each class $\left(\Gamma(l) = \frac{1}{f(l)}\right)$.

The second component uses the distance to the nearest interface to calculate a weight per pixel. As determined above, this gives higher weights for pixels closer to interfaces. This weight is computed by:

$$\alpha(\mathbf{x}, l) = e^{-\frac{d(\mathbf{x}, l)}{\sigma^2}}, \quad (2)$$

where $d(\mathbf{x}, l)$ is the distance to the nearest interface, and $\sigma = 15$ as determined based on the thickness of the thinnest layer. The weight function was built such that the pixel weight is added as a factor for the class weight:

$$\omega(\mathbf{x}, l) = \Gamma(l)(1 + \alpha(\mathbf{x}, l)). \quad (3)$$

2.3 Datasets

Mice were separated into two sets, one for training and validation (N=96) and one for testing (N=18), each with a balanced distribution of WT and $3 \times$Tg-AD mice. This mice-based split ensures the complete separation between data used for training and testing.

The training set was built by randomly considering 25 volumes from each group and the first four time-points. From each volume, 15 B-scans were randomly selected, with a minimum separation of 5 B-scans to reduce the similarities between selected B-

scans, yielding a training set of 3000 B-scans (25 volumes $\times$ 2 groups $\times$ 4 time points $\times$ 15 B-scans).

Horizontally mirroring and sinusoidal vertical displacements were used towards data augmentation to reduce over-fitting and introduce further variability in the used dataset. The latter uses a sine function $(t(x) = K \sin(2\pi f x + \phi))$ to determine the vertical translation of each A-scan across the B-scan. In this work, we used two sets of range for two of the parameters, K and f, one focusing on low amplitude and high frequency distortions and one focusing on high amplitude and low frequency. These parameters were randomly set for each B-scan in the training dataset. The phase, ϕ, was not constrained in any way. The resulting images and the original B-scan example are shown in Figure 2.
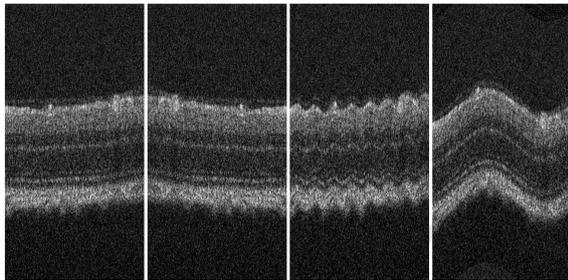
Figure 2: Example of data augmentation, from left to right: original B-scan, mirrored B-scan, sinusoidal displacements, respectively low amplitude/high frequency and high amplitude/low frequency.

The ground truth used to train a neural network is often provided by experts' manual segmentation of images composing the training set. However, this is not a viable option for the current number of images because of the burden, the time required, and all associated costs. Hence, a semi-automated process was used where segmentations were performed based on image and signal processing techniques (Ferreira et al., 2020), first and then validated and corrected, wherever required, by experts.

From the split into training and validation sets (see above), 180 volumes (92 160 B-scans — 18 mice) composed the test set.

These volumes were segmented by the developed neural network (Figure 1) to produce five mean value fundus images (MVF) (Guimarães et al., 2014), each corresponding to a distinct anatomical layer/layer-complex of the retina, specifically the retinal nerve fibre and ganglion cell layer complex (RNFL-GCL), the inner plexiform layer (IPL), the inner nuclear layer (INL), the outer plexiform layer (OPL), and the outer nuclear layer (ONL).

2.4 Training

The Adam method (Kingma and Ba, 2015) was used during training using a batch size of 8. The initial learning rate was set to 10^{-3} and subsequently reduced by a factor of 0.75 when error plateaus. Early stopping was used to prevent over-fitting, and both learning rate and early stopping used the validation accuracy as the performance metric.

All processing was done in Python 3.7.9, with the neural network models using the Keras 2.7 (Chollet, 2015) framework with TensorFlow 2.7 as backend (Abadi et al., 2016). The training was performed in a Ubuntu desktop, equipped with an AMD Rysen 9 3900x CPU @3.8 GHz with 12 cores, 64 GB RAM, and an Nvidia RTX 3060 with 12 GB of memory, using version 11.5 of CUDA.

2.5 Interface Detection

The CNN proposed herein produces the semantic segmentation of a single B-scan towards the ultimate goal of determining the interfaces between retinal layers/layer-complexes, as shown in Figure 3. A validation step was implemented at the level of individual A-scans, and an A-scan is considered correctly segmented if fulfilling the following criteria:

1. all layers are present;
2. any layer can only appear once;
3. the order of the layers is fixed;
4. each layer thickness is consistent across all A-scans of this B-scan.

These criteria are possible because the optic disc is of no interest for the present study and only retinas close to the healthy status are of interest. Therefore, thickness' variability throughout the B-scan is expected to be low to all layers.

2.6 Texture Analysis

A MVF was computed for each of the segmented layers/layer-complexes (Figure 4), with texture analysis being carried out based on the grey-level co-occurrence matrices (GLCM) (Haralick et al., 1973) after reducing the number of grayscale levels from 65 536 (16 bits) to 16 (4 bits). Twenty features were determined per block as in (Soh and Tsatsoulis, 1999) and (Haralick et al., 1973), and direction. Furthermore, four directions (0°, 45°, 90° and 135°) with symmetry ON and scale (pixel distance) of one pixel led to four GLCM per block (24 $\times$ 24 blocks of 20 $\times$ 20 pixels each). Each feature was then considered the
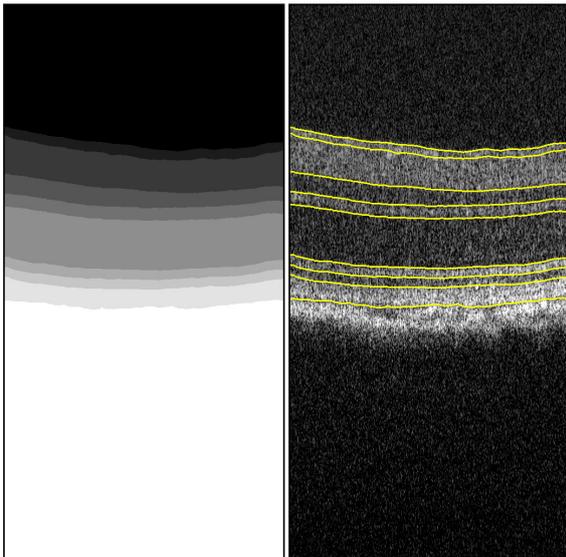
Figure 3: Semantic segmentation (left) and corresponding interface detection (right) for an example B-scan.

maximum across the four directions to decouple features' values from the orientation, resulting in 20 features per block.
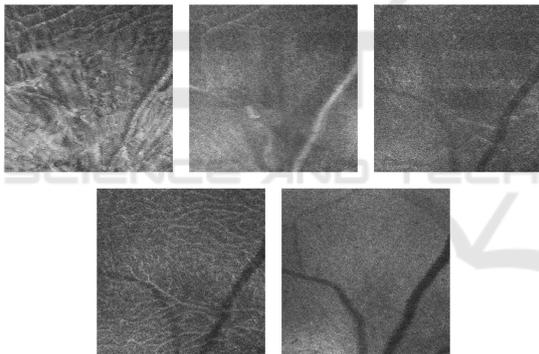
Figure 4: Mean value fundus images of a triple-transgenic mouse model of Alzheimer's disease. From left to right and top to bottom: retinal nerve fibre layer-ganglion layer complex; inner plexiform layer; inner nuclear layer; outer plexiform layer, and; outer nuclear layer.

3 RESULTS

3.1 Segmentation Capability

In addition to the performance of the segmentation, concerning the ground truth segmentation, it is essential to assess the feasibility of the segmentation process to deliver results for a significant number of cases. This work deals with retinas with significantly different appearances due to the study duration. Indeed, the retina of a mouse at one-month-old does

look significantly different from that at twelve or sixteen months old.

The histogram of the fraction of A-scans correctly segmented, following the criteria outlined in section 2.5, is presented in Figure 5. Here, 163 of the 180 volumes of the test set present 90% to 100% of A-scans with valid segmentations at all layers.
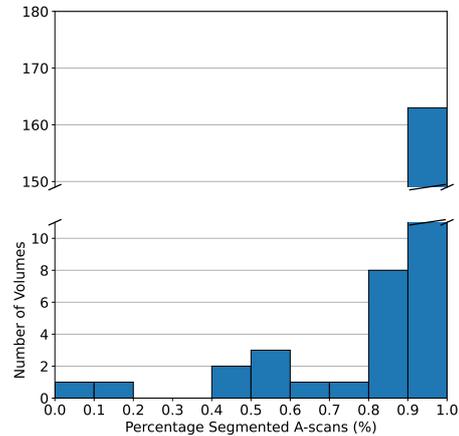
Figure 5: Histogram of the fraction of A-scans with valid segmentations at all layers.

A threshold was established from the histogram above, and only data from volumes with at least 90% of A-scans correctly segmented were considered. Table 2 presents the fraction of these volumes from the test set by group, eye and age. The number of volumes on the test set varies because of the exclusion of data due to image quality and the death of mice during the studied period.

Overall, the segmentation works well in all cases. However, two particular instances stand out at three and sixteen months old, both for the control group. These volumes belong to the same three mice in both cases. This fact suggests that these consistent errors are mice specific rather than a segmentation problem.

The texture analysis found extensive and statistically significant differences between groups.

3.2 Thickness

3.2.1 Normative Data

A normative database was generated based on the 512×512 thickness value maps by computing the average and standard deviation for 3×3 blocks. Indeed, a normative 3×3 thickness map was computed per group and time point. An example of a normative map is provided in Figure 6, along with two application examples. A detailed and thorough analysis of the normative database will be provided in a subsequent publication.

Table 2: Fraction of volumes of the test set correctly segmented by group, eye and age.

Group	Eye	One Month	Two Months	Three Months	Four Months	Eight Months	Twelve Months	Sixteen Months
WT	OD	9 / 9 (100%)	6 / 6 (100%)	6 / 6 (100%)	6 / 6 (100%)	5 / 6 (83.3%)	6 / 6 (100%)	5 / 6 (83.3%)
	OS	8 / 9 (88.9%)	6 / 6 (100%)	3 / 6 (50%)	6 / 6 (100%)	6 / 6 (100%)	5 / 6 (83.3%)	3 / 6 (50%)
	Total	17 / 18 (94.4%)	12 / 12 (100%)	9 / 12 (75%)	12 / 12 (100%)	11 / 12 (91.7%)	11 / 12 (91.7%)	8 / 12 (66.7%)
AD	OD	8 / 8 (100%)	7 / 9 (77.8%)	8 / 8 (100%)	5 / 6 (83.3%)	6 / 6 (100%)	5 / 5 (100%)	3 / 3 (100%)
	OS	8 / 8 (100%)	7 / 8 (87.5%)	8 / 8 (100%)	5 / 6 (83.3%)	6 / 7 (85.7%)	4 / 5 (80%)	3 / 3 (100%)
	Total	16 / 16 (100%)	14 / 17 (82.4%)	16 / 16 (100%)	10 / 12 (83.3%)	12 / 13 (92.3%)	9 / 10 (90%)	6 / 6 (100%)
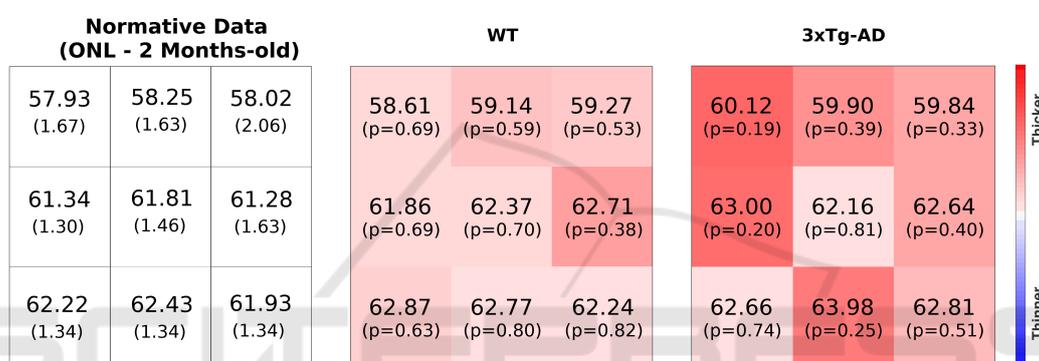
Figure 6: Left: normative thickness values (μm) for the outer nuclear layer (ONL) of WT mice (two-months-old). Thickness maps for a WT and a 3×Tg-AD mouse (both two-months-old). WT – wild-type; 3×Tg-AD – triple-transgenic mouse model of Alzheimer's disease.

3.2.2 Group Differences

The comparison between groups shows consistent retinal layer differences with the 3×Tg-AD group presenting a decreased layer thickness except for the RNFL-GCL complex, where the 3×Tg-AD group presents an increased thickness, and the ONL, where thickness values match one another. This finding (at the RNFL-GCL complex) is against common results reporting decreased RNFL thickness. Nevertheless, in the present study, two layers aggregated are considered because, in these animal groups, it is not possible to distinguish the RNFL from the GCL.

3.2.3 Age Differences

We demonstrated different behaviours for the retina's layers with ageing based on collected data. While some get thicker over time, some get thinner, and some keep the same for the reported period, establishing a thickness normative database (Ferreira et al., 2021) for each layer, location and time point that can be used for other comparative studies. A thorough analysis will be the subject of a subsequent publication.

3.3 Texture Analysis

From MVF images computed per retinal layer/layer-complex, as shown in Figure 4, texture analysis metrics were computed as disclosed in section 2.6. Four retinal fundus regions (quadrants) were considered by averaging 12×12 blocks to find widespread statistically significant differences between groups over the imaged area and across all retinal layers/layer-complex. Details for the ONL are presented in Table 3 as an example. A thorough analysis will be the subject of a subsequent publication.

Table 3: Features with statistically significant differences between wild-type and transgenic mice groups for the outer nuclear layer (ONL). The number of symbols identifies the number of quadrants where the significance was observed for each feature shown. The green-coloured circles (🟢) represent a p-value ≤ 0.05, the orange-coloured squares (🟧) show p-values ≤ 0.01 and the red-coloured asterisks (✱) represent a p-value ≤ 0.001.

Layer	Feature	Age (months)			
		1	2	3	4
ONL	Autocorrelation	✱✱✱✱	✱✱✱✱	✱✱✱	🟢✱
	Contrast	✱✱✱✱	✱✱✱✱	✱✱✱	✱
	Correlation	✱✱🟧	🟢	🟢	🟢✱
	Cluster Prominence	✱✱✱✱	✱✱✱✱	✱✱✱	✱
	Cluster Shade	✱✱✱✱	✱✱✱✱	✱✱✱✱	🟢🟢
	Dissimilarity	✱✱✱✱	✱✱✱✱	✱✱✱✱	✱✱🟢
	Energy	✱✱✱✱	✱✱✱✱	✱✱✱✱	✱🟧🟢
	Entropy	✱✱✱✱	✱✱🟧✱	✱✱	
	Homogeneity	✱✱✱✱	✱✱✱✱	✱✱✱	✱
	Maximum Probability	✱✱✱✱	✱✱✱✱	✱✱✱	✱
	Sum of Squares	✱✱✱✱	✱✱✱✱	🟧✱✱🟧	
	Sum Average	✱✱✱✱	✱✱✱🟧	✱✱	🟢✱
	Sum Variance	✱✱✱✱	✱✱✱✱	✱✱✱✱	✱✱🟢
	Sum Entropy	✱✱✱✱	✱✱✱✱	✱✱✱	✱
	Difference Variance	✱✱✱✱	✱✱✱✱	✱✱✱	✱
	Difference Entropy	🟢		🟧🟧	🟧
	IMC1	✱✱✱✱	✱✱✱✱	✱✱✱✱	🟧🟢🟧
	IMC2	🟢		🟧🟧	🟧🟢
	INN	✱✱✱✱	✱✱✱🟧	✱✱✱	🟢✱
	IDN	✱✱✱✱	✱✱✱🟧	🟢✱✱	🟢✱

4 DISCUSSION AND CONCLUSIONS

Along the work developed over three years, we continuously monitor two mice groups, a WT control group and the $3\times$Tg-AD, by frequent imaging their retina by OCT to gather knowledge on the healthy ageing, the ageing of transgenic mice and the difference between groups in particular how these compare over time.

This is the first time mice are consistently imaged since their young age (one-month-old), demonstrating an immature retina and shedding light on changes at the early stages of life.

We have developed a neural network model for retina segmentation based on the U-Net architecture and using a new weighting scheme to adapt it to our objectives. This weighting scheme is pixel-based and calculated based on the pixels' distance to the nearest interface, allowing the prioritisation of the areas close to the interfaces to produce fewer errors at interface locations.

To train this model, we used 114 mice from the two groups. A semi-automated strategy was employed to determine the ground truth of the selected B-scans, as the amount of data available makes it difficult, if not impossible, to produce manual segmentation for all B-scans. A previously detailed algorithm (Ferreira et al., 2020) was used to create the initial segmentation, with manual review and corrections, wherever needed, from expert graders.

The training and test sets were separated by mouse, guaranteeing the complete separation between the two groups. Furthermore, only data from the first four time-points were considered during training, resulting in 3000 B-scans being used for training. These were augmented to increase the training set and introduce variations in the layer location and shape to reduce the possibility of over-fitting.

The results from the test set showed that the model successfully segmented over 90% of A-scans for the large majority of the volumes in the test set. We present the number per group, eye and age by taking this value as a threshold to consider a volume correctly segmented.

While we do not see any bias towards any particular group or time point, only 50% of the left eyes for mice three and sixteen months old were correctly segmented. A thorough analysis revealed that these eyes came only from three mice, implying that these errors might be specific to these mice and not a general limitation of the segmentation process.

The developed normative data for retinal thickness, both layer-specific and total retina thickness, up to the age of sixteen months old, is a considerable asset by demonstrating the thickness evolution of the same animals and providing a comparative reference to other studies.

Also, the consistent imaging and computing of fundus images from the retina of WT and 3×Tg-AD and the texture analysis of those allowed to detect differences between groups since one-month-old and raise novel scientific questions on their meaning.

In conclusion, these preliminary results from the ongoing study demonstrate the massive differences between the retinas of WT and 3×Tg-AD mice since one-month-old, both from the thickness and texture analysis viewpoints.

ACKNOWLEDGEMENTS

REFERENCES

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., Kudlur, M., Levenberg, J., Monga, R., Moore, S., Murray, D. G., Steiner, B., Tucker, P., Vasudevan, V., Warden, P., Wicke, M., Yu, Y., and Zheng, X. (2016). TensorFlow: A system for large-scale machine learning. In *Proceedings of the 12th USENIX Conference on Operating Systems Design and Implementation*, OSDI'16, page 265–283, USA. USENIX Association.

Ahmed, R. M., Paterson, R. W., Warren, J. D., Zetterberg, H., O'Brien, J. T., Fox, N. C., Halliday, G. M., and Schott, J. M. (2014). Biomarkers in dementia: clinical utility and new directions. *Journal of Neurology, Neurosurgery & Psychiatry*, 85(12):1426.

Alzheimer's & Dementia (2020). 2020 Alzheimer's disease facts and figures. *Alzheimer's & Dementia*, 16(3):391–460.

Antony, B. J., Kim, B.-J., Lang, A., Carass, A., Prince, J. L., and Zack, D. J. (2017). Automated segmentation of mouse OCT volumes (ASiMOV): Validation & clinical study of a light damage model. *PLOS ONE*, 12(8):e0181059.

Berger, A., Cavallero, S., Dominguez, E., Barbe, P., Simonutti, M., Sahel, J.-A., Sennlaub, F., Raoul, W., Paques, M., and Bemelmans, A.-P. (2014). Spectral-domain optical coherence tomography of the rodent eye: Highlighting layers of the outer retina using signal averaging and comparison with histology. *PLoS ONE*, 9(5):e96494.

Chiu, K., Chan, T.-F., Wu, A., Leung, I. Y.-P., So, K.-F., and Chang, R. C.-C. (2012). Neurodegeneration of the retina in mouse models of Alzheimer's disease: what can we learn from the retina? *Age*, 34(3):633–649.

Chiu, S. J., Li, X. T., Nicholas, P., Toth, C. A., Izatt, J. A., and Farsiu, S. (2010). Automatic segmentation of seven retinal layers in SDOCT images congruent with expert manual segmentation. *Optics Express*, 18(18):19413–19428.

Chollet, F. (2015). Keras.

Diakogiannis, F. I., Waldner, F., Caccetta, P., and Wu, C. (2020). ResUNet-a: A deep learning framework for semantic segmentation of remotely sensed data. *IS-PRS Journal of Photogrammetry and Remote Sensing*, 162:94–114.

Dufour, P. A., Ceklic, L., Abdillahi, H., Schröder, S., Dzanet, S. D., Wolf-Schnurrbusch, U., and Kowal, J. (2013). Graph-based multi-surface segmentation of OCT data using trained hard and soft constraints. *IEEE Transactions on Medical Imaging*, 32(3):531–543.

Fang, L., Cunefare, D., Wang, C., Guymer, R. H., Li, S., and Farsiu, S. (2017). Automatic segmentation of nine retinal layer boundaries in OCT images of non-exudative AMD patients using deep learning and graph search. *Biomedical Optics Express*, 8(5):2732.

Ferreira, H., Martins, J., Nunes, A., Moreira, P. I., Castelo-Branco, M., Ambrósio, A. F., Serranho, P., and Bernardes, R. (2020). Characterization of the retinal changes of the 3×Tg-AD mouse model of Alzheimer's disease. *Health and Technology*, 10(4):875–883.

Guimarães, P., Rodrigues, P., Lobo, C., Leal, S., Figueira, J., Serranho, P., and Bernardes, R. (2014). Ocular fundus reference images from optical coherence tomography. *Computerized Medical Imaging and Graphics*, 38(5):381–389.

Habib, M., Mak, E., Gabel, S., Su, L., Williams, G., Waldman, A., Wells, K., Ritchie, K., Ritchie, C., and O'Brien, J. T. (2017). Functional neuroimaging findings in healthy middle-aged adults at risk of Alzheimer's disease. *Ageing Research Reviews*, 36:88–104.

Haralick, R. M., Shanmugam, K., and Dinstein, I. (1973). Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-3(6):610–621.

Hart, N. J., Koronyo, Y., Black, K. L., and Koronyo-Hamaoui, M. (2016). Ocular indicators of Alzheimer's: exploring disease in the retina. *Acta Neuropathologica*, 132(6):767–787.

He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. *2016 IEEE Con-

ference on Computer Vision and Pattern Recognition (CVPR), pages 770–778. HugoFerreira.

Ioffe, S. and Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. In Bach, F. and Blei, D., editors, *Proceedings of the 32nd International Conference on Machine Learning*, volume 37 of *Proceedings of Machine Learning Research*, pages 448–456, Lille, France. PMLR.

Jack, C. R., Knopman, D. S., Jagust, W. J., Shaw, L. M., Aisen, P. S., Weiner, M. W., Petersen, R. C., and Trojanowski, J. Q. (2010). Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade. *The Lancet. Neurology*, 9(1):119–28.

Jha, D., Smedsrud, P. H., Riegler, M. A., Johansen, D., Lange, T. d., Halvorsen, P., and Johansen, H. D. (2019). ResUNet++: An advanced architecture for medical image segmentation. In *Proceedings - 2019 IEEE International Symposium on Multimedia, ISM 2019*, volume 00, pages 225–230. Institute of Electrical and Electronics Engineers Inc.

Kingma, D. P. and Ba, J. (2015). Adam: A method for stochastic optimization. In Bengio, Y. and LeCun, Y., editors, *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*.

Kitazawa, M., Medeiros, R., and LaFerla, F. M. (2012). Transgenic mouse models of Alzheimer disease: Developing a better model as a tool for therapeutic interventions. *Current Pharmaceutical Design*, 18(8):1131–1147.

Koozekanani, D., Boyer, K., and Roberts, C. (2021). Retinal thickness measurements from optical coherence tomography using a Markov boundary model. *IEEE Transactions on Medical Imaging*, 20(9):900–916.

Krantic, S. and Torriglia, A. (2014). Retina: source of the earliest biomarkers for Alzheimer's disease? *Journal of Alzheimer's disease : JAD*, 40(2):237–43.

London, A., Benhar, I., and Schwartz, M. (2013). The retina as a window to the brain–from eye research to CNS disorders. *Nature Reviews Neurology*, 9(1):44–53.

Morales, S., Colomer, A., Mossi, J. M., Amor, R. d., Woldbye, D., Klemp, K., Larsen, M., and Naranjo, V. (2021). Retinal layer segmentation in rodent OCT images: Local intensity profiles & fully convolutional neural networks. *Computer Methods and Programs in Biomedicine*, 198:105788.

Morgia, C. L., Ross-Cisneros, F. N., Koronyo, Y., Hannibal, J., Gallassi, R., Cantalupo, G., Sambati, L., Pan, B. X., Tozer, K. R., Barboni, P., Provini, F., Avanzini, P., Carbonelli, M., Pelosi, A., Chui, H., Liguori, R., Baruzzi, A., Koronyo-Hamaoui, M., Sadun, A. A., and Carelli, V. (2016). Melanopsin retinal ganglion cell loss in Alzheimer disease. *Annals of Neurology*, 79(1):90–109.

Ngo, L., Cha, J., and Han, J.-H. (2020). Deep neural network regression for automated retinal layer segmentation in optical coherence tomography images. *IEEE Transactions on Image Processing*, 29:303–312.

Nunes, A., Silva, G., Duque, C., Januário, C., Santana, I., Ambrósio, A. F., Castelo-Branco, M., and Bernardes, R. (2019). Retinal texture biomarkers may help to discriminate between Alzheimer's, Parkinson's, and healthy controls. *PLOS ONE*, 14(6):e0218826.

Oddo, S., Caccamo, A., Shepherd, J. D., Murphy, M., Golde, T. E., Kayed, R., Metherate, R., Mattson, M. P., Akbari, Y., and LaFerla, F. M. (2003). Triple-transgenic model of Alzheimer's disease with plaques and tangles intracellular aβ and synaptic dysfunction. *Neuron*, 39(3):409–421.

Rathke, F., Schmidt, S., and Schnörr, C. (2014). Probabilistic intra-retinal layer segmentation in 3-D OCT images using global shape regularization. *Medical Image Analysis*, 18(5):781–794.

Ronneberger, O., Fischer, P., and Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 9351, pages 234–241. Springer Verlag.

Roy, A. G., Conjeti, S., Karri, S. P. K., Sheet, D., Katouzian, A., Wachinger, C., and Navab, N. (2017). ReLayNet: retinal layer and fluid segmentation of macular optical coherence tomography using fully convolutional networks. *Biomedical Optics Express*, 8(8):3627.

Sarhan, M. H., Nasseri, M. A., Zapp, D., Maier, M., Lohmann, C. P., Navab, N., and Eslami, A. (2020). Machine learning techniques for ophthalmic data processing: A review. *IEEE Journal of Biomedical and Health Informatics*, 24(12):3338–3350.

Schön, C., Hoffmann, N. A., Ochs, S. M., Burgold, S., Filser, S., Steinbach, S., Seeliger, M. W., Arzberger, T., Goedert, M., Kretzschmar, H. A., Schmidt, B., and Herms, J. (2012). Long-term in vivo imaging of fibrillar tau in the retina of P301S transgenic mice. *PLoS ONE*, 7(12):e53547.

Soh, L.-K. and Tsatsoulis, C. (1999). Texture analysis of SAR sea ice imagery using gray level co-occurrence matrices. *IEEE Transactions on Geoscience and Remote Sensing*, 37(2):780–795.

Soukup, P., Maloca, P., Altmann, B., Festag, M., Atzpodien, E.-A., and Pot, S. (2019). Interspecies variation of outer retina and choriocapillaris imaged with optical coherence tomography. *Investigative Opthalmology & Visual Science*, 60(10):3332.

Srinivasan, P. P., Heflin, S. J., Izatt, J. A., Arshavsky, V. Y., and 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. *Biomedical Optics Express*, 5(2):348.

Venhuizen, F. G., Ginneken, B. v., Liefers, B., Grinsven, M. J. J. P. v., Fauser, S., Hoyng, C., Theelen, T., and Sánchez, C. I. (2017). Robust total retina thickness segmentation in optical coherence tomography images using convolutional neural networks. *Biomedical Optics Express*, 8(7):3292.

Yazdanpanah, A., Hamarneh, G., Smith, B. R., and 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(2):484–496.

Zhang, Z., Liu, Q., and Wang, Y. (2018). Road extraction by deep residual U-Net. *IEEE Geoscience and Remote Sensing Letters*, 15(5):749–753.