



# Motion-invariant SRT treatment detection from direct M-scan OCT imaging

Tatiana Fountoukidou<sup>1</sup> · Philippe Raisin<sup>1</sup> · Daniel Kaufmann<sup>2</sup> · Jörn Justiz<sup>2</sup> · Raphael Sznitman<sup>1</sup> · Sebastian Wolf<sup>3</sup>

Received: 29 January 2018 / Accepted: 22 February 2018  
© CARS 2018

## Abstract

**Purpose** Selective retina therapy (SRT) is a laser treatment targeting specific posterior retinal layers. It is focused on inducing damage to the retinal pigment epithelium (RPE), while sparing other retinal tissue compared to traditional photocoagulation. However, the targeted RPE layer is invisible with most imaging modalities and induced SRT lesions cannot be monitored. In this work, imaging scans acquired from an experimental setup that couples the SRT laser beam with an optical coherence tomography (OCT) beam are analyzed in order to evaluate the treatment as they occur.

**Methods** We isolated a small part of the time-resolved scan corresponding to the end of the treatment, for which we have microscopic evidence of the SRT outcome. We then use a convolutional neural network to correspond each scan to the treatment result. We explore which aspects of the scan convey more valuable information for a robust therapy evaluation. By only using this adequately small part, we can achieve an online estimation, while being resilient to eye movement.

**Results** The available dataset consists of time-resolved OCT scans of 98 ex vivo porcine eyes, treated with different energy levels. The proposed method yields high performance in the task of predicting whether the applied energy was adequate for SRT treatment, by focusing on the immediate OCT signal acquired during treatment time.

**Conclusions** We propose a strategy toward online noninvasive SRT treatment assessment, able to provide a satisfying evaluation of a treatment status, that therefore could be used for the planning of the treatment continuation.

**Keywords** Computer assisted intervention · Time-resolved OCT · Online dosimetry · Selective retina Therapy

## Introduction

Selective retina therapy (SRT) is a novel laser treatment for eye diseases associated with the retinal pigment epithelium (RPE) layer of the retina [1,9]. The RPE layer is located in the posterior part of the retina, between the photoreceptors and the choroid. Among other things, it is responsible for the nourishment of the retinal visual cells and is thus critical for functioning sight [14].

The basic principle of SRT lies in the fact that when laser pulses with a short duration ( $< 10 \mu\text{s}$ ) and a low repetition rate are applied to the retina, their effect is localized

to dominant light-absorbing structures such as the intracellular melanosomes of the RPE, while the surrounding layers remain unaffected [10]. It is stipulated that SRT lesions are caused by thermomechanical disruption of the tissue, due to microbubble formation and vaporization around the melanosomes within the RPE cells. This is in contrast to conventional photocoagulation whose effect is thermal and spreads over all retinal layers. With SRT, the induced damage is inflicted to the RPE, which is then inherently healed by the reordering of the cells around the damaged area, leading to a desired tissue rejuvenation process.

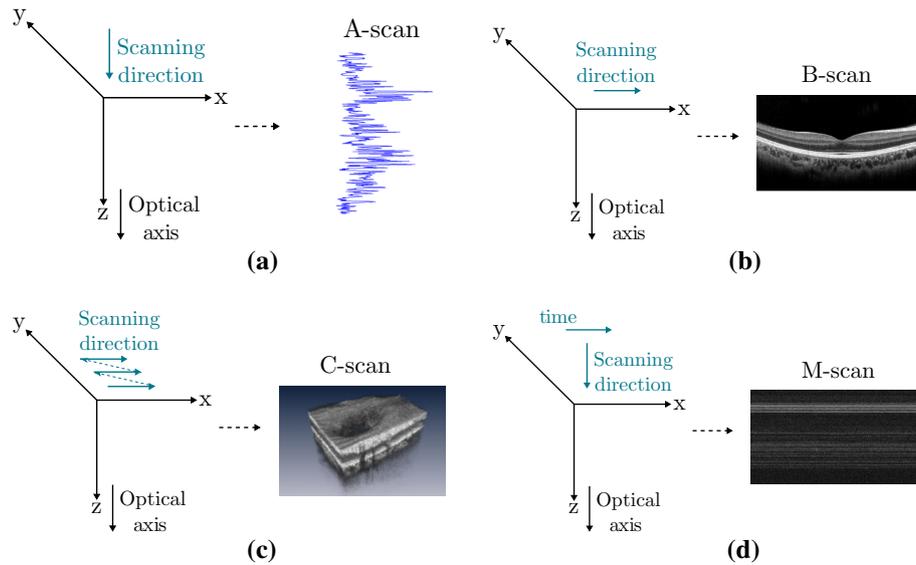
There are unfortunately two main limitations that hinder SRT from becoming more widespread. The first is that if the applied energy exceeds a certain limit, the microbubbles collapse. After a certain energy threshold is surpassed, the damage expands to the photoreceptors of the neural retina or the choroid, causing undesirable collateral harm to these structures. Establishing and modulating the energy limit before application is hence critical yet nontrivial since it

✉ Tatiana Fountoukidou  
tatiana.fountoukidou@artorg.unibe.ch

<sup>1</sup> ARTORG Center, University of Bern, Bern, Switzerland

<sup>2</sup> Engineering and Information Technology, Berner Fachhochschule, Biel/Bienne, Switzerland

<sup>3</sup> Inselspital, University Hospital of Bern, Bern, Switzerland



**Fig. 1** OCT scanning modes. A-scans (top left) provide depth information at the imaged spot. B-scans (top right) are repeated A-scans distributed over a line, forming a cross section of the image sample. C-scans (bottom right) are repeated B-scans in a perpendicular direction to the B-scan direction, forming a image volume of the sample.

M-scans (bottom right) are repeated A-scans at a given location, allowing time-lapse image of single spot. **a** An A-scan is a depth profile of a single retinal spot. **b** A B-scan is a 2-D cross-sectional image of the retina. **c** A C-scan is a 3-d representation of the retina. **d** An M-scan is the depth profile of a single retinal spot, over time

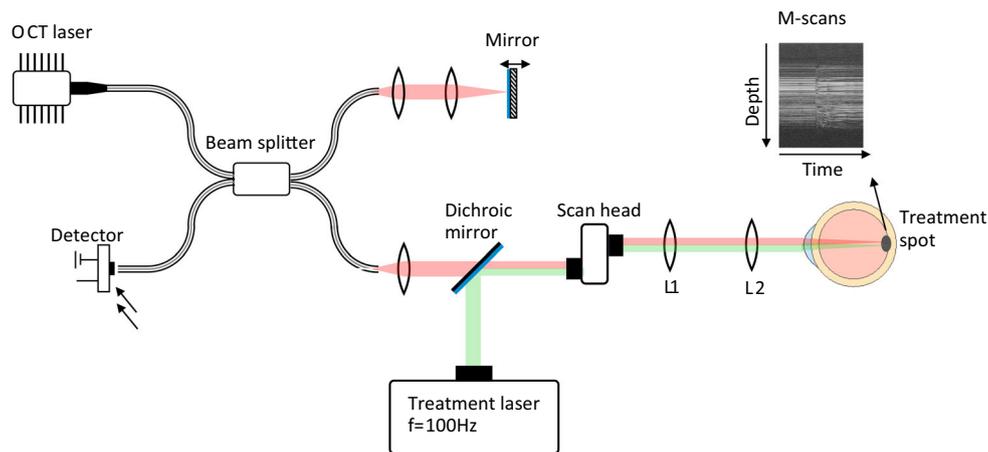
is above all patient specific. Second, imaging the RPE directly is challenging as it is embedded in the posterior part of the retina. That is, most common imaging modalities used by the ophthalmologists, such as fundus imaging, are unable to provide a direct view of the layer during the laser application. In effect, waiting to observe positive treatment outcomes from such traditional imaging can take days and is thus ill-suited to monitor the treatment process. One alternative way to visualize the state of the RPE, however, is using fundus fluorescein angiography (FFA) [2,10] which leverages injected fluorescein dye to reveal the circulation patterns of the retina and exposes leakages in the tissue structure. In particular, leakage through the RPE indicates damage in this layer. However, FFA is at best a follow-up method to evaluate the treatment efficacy and is often not compatible with day-to-day clinical workflow, since there is currently no system that combines it with SRT, and potential interaction between the SRT laser and the fluorescein dye is unknown.

To overcome these shortcomings, a number of hardware-based solutions have been proposed, to give more immediate indications to appropriate treatment. In [12], a piezoelectric transducer is embedded in a contact lens. The transducer measures the optoacoustic transient that is provoked by the SRT laser pulses, and the transient signal is observed through time. There is evidence that the amount of RPE damaged cells is proportional to pulse-to-pulse fluctuation observed, according to FFA. While this approach can provide a strategy for SRT energy level monitoring, it does not provide visual feedback of the treatment and requires using the con-

tact lens which is a discomfort for patients. Alternatively, in [8], SRT lesions were induced to Chinchilla Bastard rabbits by applying a series of 30 pulses in a linearly increasing ramping mode. The SRT energy was controlled through a reflectometer, which divides the SRT beam and measures the correlation between the entrance and the backscattered light. It is thought that the larger the deviation between the two signals, the more likely it is that microbubble formation has taken place. When a predefined threshold is surpassed, laser emission is halted within 30 pulses. The results show that the majority of induced lesions are within the therapeutic window of the SRT and that reflectometry could be used to control the energy dosage of the treatment.

More recently, there has been growing evidence that Optical Coherence Tomography (OCT) [4,15] can be used to image the retina layers as they are being treated and provide more accurate indications of successful SRT laser applications [6,13,16]. At its heart, OCT uses infrared light and interferometry to image internal structures within the retina. While high-frequency B-scan and C-scan OCT image data (see Fig. 1) remain challenging to acquire fast enough, A-scans taken over time at a single location, also known as *M-scans*, have shown promise in this direction.

In this context, automatic assessment of OCT M-scans to determine whether or not SRT laser application has been correctly and effectively applied is promising as it can also provide direct visual confirmation of treatment and the ability to dynamically monitor treatment as it is occurs. This could lead to reduction in energy applications and thus diminish the



**Fig. 2** Schematic of the experimental setup. The SRT and the OCT beams are coupled and guided to the eye. The resulting M-scan consists a single depth profile of the treatment spot through time

risk of over-treatment. Along this line, [16] demonstrated an automated strategy that used complete M-scans to determine whether treatment was achieved appropriately. This method used handcrafted image features and a random forest in order to classify M-scans. While effective, the major limitation of this method was its reliance on a complete and continuous M-scan to perform an assessment at test-time. As we will show in our experimental section, this leads this method to “count” laser pulses observed in the M-scan and is potentially dangerous, as treatment at a location is approximate as the patient eye is not immobilized and moves during the laser application. As such, counted pulses could be accumulated over several adjacent positions and not necessarily at a specific location. Involuntary eye movement frequency varies between 20 and 200 Hz [3], therefore making it possible for the eye to move from 6 to 60 times during an SRT 30-pulse train application (SRT frequency is set to 100 Hz).

To overcome this limitation, we propose here a novel approach that allows small sequences of A-scans that are unlikely to have been sampled from more than one location to be assessed for treatment effect. That is, our method evaluates treatment outcome based on extremely small OCT M-scan segments, ensuring that the treatment assessment is not based on pulse counting. Our approach is based on a CNN that evaluates pairs of non-treated and potentially treated M-scan sequences at a given location. This allows for both location specific energy dosimetry and accurate classification of treatments. In the remainder of the paper, we outline our approach and the experimental setup used to validate our method.

## Methods

We begin by outlining our physical system and the process by which we acquire OCT M-scans, as well as ground truth data of SRT treatment outcomes. We then describe our classification approach.

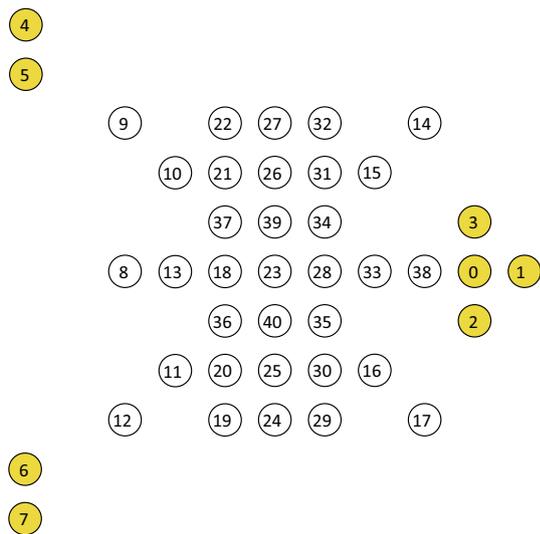
## System description

Similar to [6], we used a system with two lasers in order to collect OCT M-scans and apply SRT treatment. The first is the SRT system, which is the treatment laser. The SRT pulses are executed using a frequency-doubled Nd:YLF laser with pulse width of  $1.7 \mu\text{s}$  and a pulse repetition rate of 100 Hz at a wavelength of 527 nm for 30 pulse trains. The second one is the OCT system, which is the imaging laser. The OCT system has a line scan frequency of 50 kHz and a spectral bandwidth of 120 nm centered at 800 nm, with an axial resolution of  $2.25 \mu\text{m}$  in air. The two beams are coupled and guided to the retina. The resulting OCT signal is effectively a 2-D image M-scan consisting of a sequence of 1-D depth profiles (i.e., A-scans) at the treatment spot over time. The setup is detailed more thoroughly in [6], and a schematic is shown in Fig. 2.

## Dataset acquisition and description

Our dataset consists of a total of 1045 M-scans from ex vivo porcine enucleated eyes.<sup>1</sup> In total, 98 different eyes were used. The eyes were SRT laser-treated at different locations with different energy levels, and their corresponding M-scans were acquired simultaneously. In order to keep track of the retinal location corresponding to each M-scan, a treatment protocol was used. At first, 8 high energy ( $300 \mu\text{J}$  per SRT pulse), microscopically visible lesions were induced to serve as microscopic landmarks. Then, a location pattern with a specified sequence defined the rest of the treatment spots, where different energy levels were used. The treatment pattern is shown in Fig. 3, where the landmark locations are highlighted in yellow.

<sup>1</sup> The eyes were collected from a local slaughterhouse, from pigs that were already killed for commercial purposes.

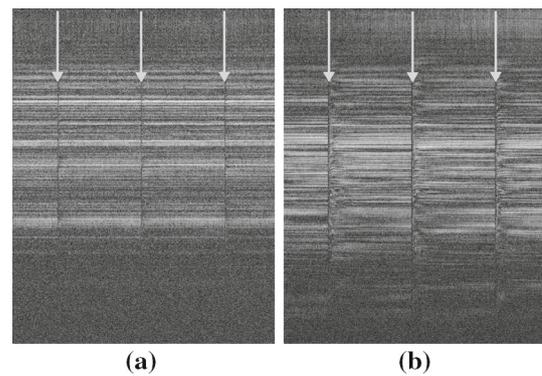


**Fig. 3** Treatment pattern. Locations 0–7 (in yellow) are treated with high energy, causing microscopically visible lesions. They therefore serve as landmarks for the localization of all the other treatment spots as well. The numbering indicates the order of the applied treatment pulse trains

The treatment energy used varied between 20 and 300  $\mu\text{J}$ , and the treatment mode was either classic (all 30 pulses are of the same energy) or ramp mode (every pulse had a higher energy than the previous one, the increase being linear through the pulse train). Different energies aim at investigating the differences in SRT treatment level, since the extend of tissue damage is not linearly related to the absorbed energy. The OCT acquisition frequency also varied between 15 and 50 kHz, producing M-scans of different sizes (same depth resolution, but different time resolution, therefore different widths). The acquisition frequencies were chosen on site during the experiment, to examine the ability to capture an SRT-related effect of different M-scan resolutions. A lower acquisition rate would result in a smaller M-scan image, therefore easier to process in real time, at the cost of reduced time resolution. Also, by choosing different OCT A-scan rates we aimed at building a system able to be independent of the commercial OCT device's acquisition parameters. For each eye, a microscopic image of the RPE layer was taken after the treatment, so that each M-scan has a corresponding microscopic view indicating whether there is a visible RPE lesion or not. This view can be treated equivalent to an FFA examination, as it demonstrates the laser effect at the RPE layer.<sup>2</sup>

For each OCT M-scan, a manual treatment assignment was performed. Treatment visibility is defined as present if at least one SRT pulse in the sequence created a visible effect in the OCT signal. In practice, this involves observing a signal

<sup>2</sup> Doing so for a patient is unfeasible as it involves physically removing the top layers of the retina.



**Fig. 4** Examples of OCT treatment visibility. The gray arrows indicate the timepoint where the signal loss is observed. The presence of a signal washout is connected to the refracting layers' moving speed with respect to the A-scan integration time of the sensor. **a** OCT without washout. **b** OCT with washout

loss extending to a few pixels that may or may not be followed by a signal washout (see Fig. 4). This effect is thought to be connected with shifting of the refracting surface along the optical axis during the A-scan integration time, caused by the tissue disruption and microbubble formation. If the surface is shifted at least  $\lambda/4$  per A-scan integration time in the direction of the optical axis, no washout is observed after the signal loss. If the refracting layers are moving, but the moving is slower than  $\lambda/4$  per A-scan integration time, a signal washout is depicted in the OCT [6]. Both types of OCT signal classify as a visible OCT signal effect. Since the SRT treatment is applied in the form of a pulse train of standard frequency, a periodic signal loss is expected (one for each SRT pulse). In some cases, not all SRT pulses result in a signal loss. If at least one SRT pulse of the pulse train shows in the OCT M-scan, the whole M-scan is characterized as having a visible OCT effect.

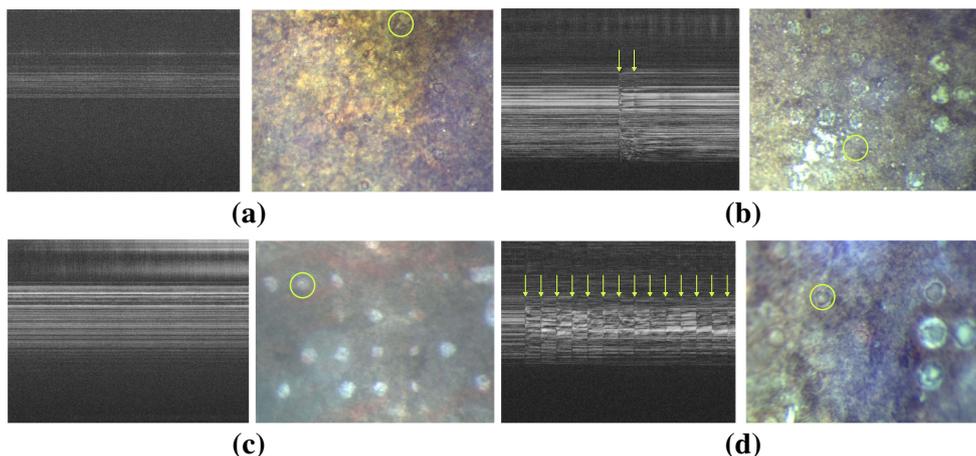
Formally, we define our data in the following way. We let  $\{X_m | m = 1, \dots, M\}$  be the set of  $M$  M-scans of size  $D \times T$ , with corresponding labels

$$y^{\text{oct}} = \begin{cases} 0, & \text{no visible OCT signal effect} \\ 1, & \text{at least one visible OCT signal effect} \end{cases}$$

and

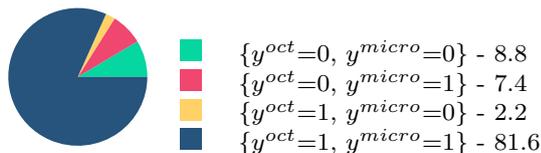
$$y^{\text{micr}} = \begin{cases} 0, & \text{no RPE lesion visible} \\ 1, & \text{visible RPE microscopic lesion} \end{cases}$$

An example for each one of the label combinations is shown in Fig. 5. An overview of the distribution of labels for the dataset is depicted in Fig. 6, where we see that for the vast majority of samples, the OCT and the microscopy labels coincide. However, since the exact correspondence of the OCT signal to the tissue damage is not fully understood, we are



**Fig. 5** Dataset examples. The left part of each subfigure is the OCT M-scan corresponding to the retinal spot that is indicated by the green circle on the right microscopic image. The right part of each subfigure is the microscopy image of the broader retinal region, when the layers above the RPE are removed. The green arrows indicate the points where OCT visibility is observed. The green circles indicate the region where

microscopic lesion should be observed, if induced. Note that the OCT label corresponds to the whole M-scan (since the M-scan is acquired from a single spatial point), while the microscopic label corresponds to the circled region (also indicating the same point on the retina). **a**  $y^{oct} = 0, y^{micro} = 0$ . **b**  $y^{oct} = 1, y^{micro} = 0$ . **c**  $y^{oct} = 0, y^{micro} = 1$ . **d**  $y^{oct} = 1, y^{micro} = 1$



**Fig. 6** Label distribution percentages

interested in inferring the microscopic labels, since they are a more reliable ground truth. The OCT labeling in this work serves mainly in identifying the SRT pulse locations.

**Pulse annotation**

The samples with OCT visibility ( $y^{oct} = 1$ ) were further analyzed to specify the exact pulse positions in the M-scan. In a first stage, an automatic annotator was applied to annotate samples with strong signal washout. The remaining inconclusive samples were manually annotated as a second stage. For the automatic annotation, we calculated the root-mean-square (RMS) for every A-scan as

$$RMS(t) = \sqrt{\frac{1}{n} \sum_{i=1}^n X_m^t(i)^2} \tag{2.1}$$

where  $X_m^t$  is the A-scan corresponding to time  $t$  of the M-scan  $X_m$  and  $X_m^t(i)$  is the  $i$ th pixel of the A-scan of time  $t$ , corresponding to a single depth point. This is motivated by the fact that the application of the SRT pulse leads to a sharp decrease in the intensity contrast in the A-scan, hence a lower RMS. This resulting curve was subsequently filtered with

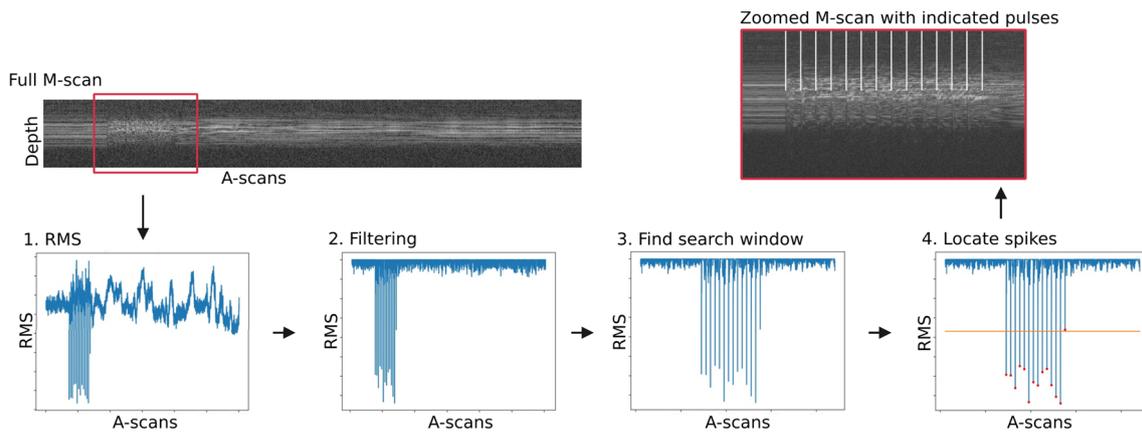
a high-pass Savitzky–Golay filter (order: 2, window length: 301) [11]. This process is visualized in Fig. 7.

To locate the coarse position of the pulse train in the M-scan, a short-time Fourier transform (STFT) with a window size given by the length of the expected pulse train was calculated. The window showing the highest amplitude at the treatment laser frequency ( $f = 100$  Hz) was chosen. In the resulting window, the lowest RMS value was assumed to result from signal washout. From this first pulse, another minima was searched both to the left and to the right of it. If the minima were (1) isolated, (2) below a threshold of 50% of the first observed minima, and (3) reasonably close to the expected pulse-to-pulse distance, then it was marked as a signal washout from the SRT application. After termination, the number of found signal washouts was compared to the number of applied pulses. If these numbers agreed (with a tolerance of  $\pm 1$  pulses, due to system latency), the locations were stored. This procedure was successful for a significant number of cases: 876 of the initial 1045 samples had at least one visible OCT pulse.

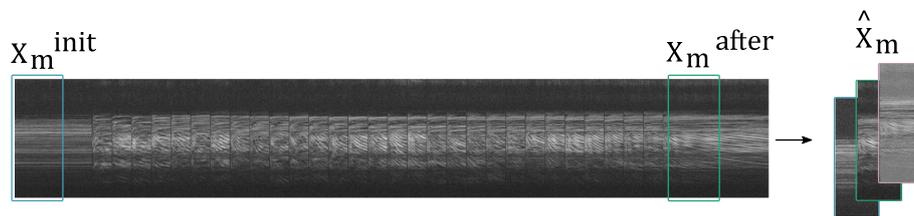
**Treatment evaluation**

To classify M-scans based on their treatment, we want to infer the microscopic visibility solely from the OCT data, since in a realistic patient scenario the microscopic information is not available. We therefore want to explore the correlation between the OCT signal and the resulting effect on the RPE layer.

More specifically, we define our data as a matrix  $\mathbf{X} \in \mathbb{R}^{M \times D \times T}$ , where  $M$  is the total number of M-scans,  $D$  is the number of pixels of the retinal depth and  $T$  is the number



**Fig. 7** Pulse annotation pipeline. The top right shows a zoom of the red rectangular of the original M-scan, with the automatically detected signal washouts indicated in white



**Fig. 8** Schematic of dataset. From each entire M-scan, we only isolate two small parts, corresponding to the before and after of the treatment process. The final image  $\hat{X}_m$  is composed of 3 channels, the first one being  $X_m^{\text{init}}$ , the second one  $X_m^{\text{after}}$  and the last one  $X_m^{\text{after}} - X_m^{\text{init}}$

of acquired time points. Each M-scan has microscopic label  $\mathbf{y} \in \{0, 1\}$  and has an associated vector  $\mathbf{p}$  of variable size, containing the sequence of SRT pulse time points determined by the pulse annotator.

Instead of attempting to evaluate an entire M-scan, as this is inappropriate to do to observed eye motion in patients, we will evaluate a subset of M-scans that are unlikely to be imaging different locations. To do this, we define  $X_m^{\text{init}}$  to be the initial block of  $\delta$  A-scans of the M-scan. That is, an initial block where no pulses have yet been applied. Similarly, we define  $X_m^{\text{after}}$  to be the  $\delta$ -size block right after the last treatment pulse observed. In the case of no OCT visibility, we define  $X_m^{\text{after}}$  as a block in the last 3rd of the M-scan. From these, we construct a new 3-channel image, of the form

$$\hat{X}_m = [X_m^{\text{init}}, X_m^{\text{after}}, X_m^{\text{after}} - X_m^{\text{init}}], \quad \forall m = 1, \dots, M. \quad (2.2)$$

The above process is depicted in Fig. 8. From such  $\hat{X}_m$ , we are interested in automatically determining the associated microscopic label.

### Network details

To determine the associated label of an  $\hat{X}_m$ , we train a convolutional neural network (CNN) to learn whether a before

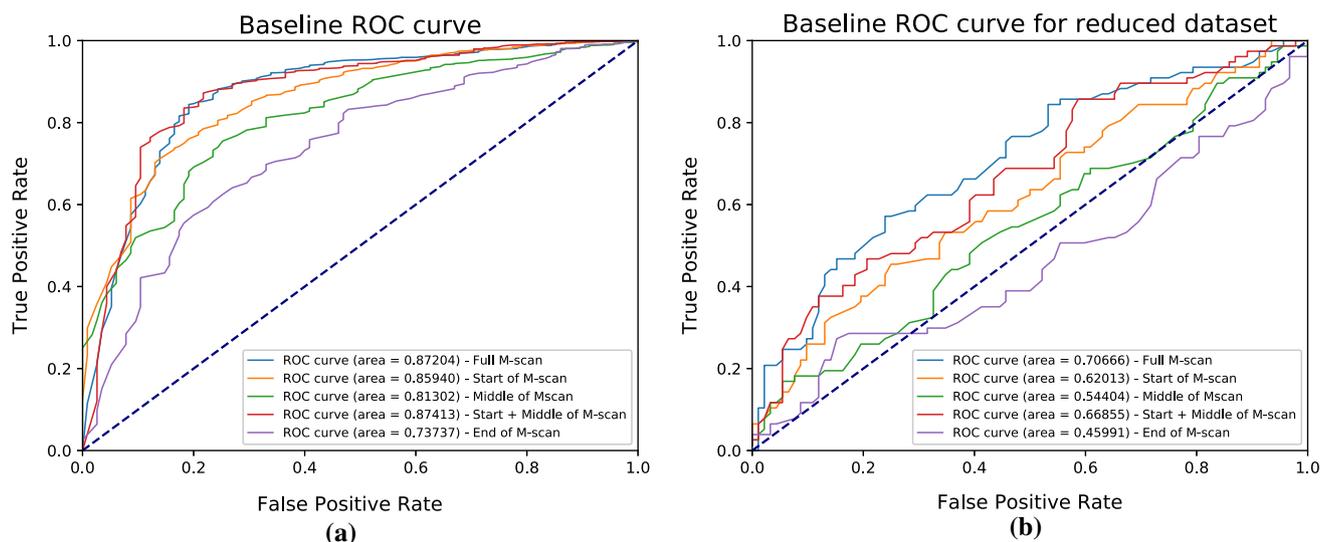
and after pair of M-scan blocks corresponds to visible RPE damage, which would indicate a successful SRT treatment.

We modified a ResNet architecture to perform this task such that it can accept inputs of smaller sizes [5]. In addition, we also substituted the first  $7 \times 7$  filter layer with one of size  $3 \times 3$ . This is because the  $7 \times 7$  filter size results in an overly large receptive field in the first layers, which is not appropriate for the fine signal variations that we want to capture. A smaller filter size has proven to be more appropriate for the problem we aim to tackle. The fully connected and classification layers were removed, and new ones were trained for the desired task. For the layers that were kept common with the original architecture, the pretrained weights learned on the imagenet [7] dataset were used as a starting training point, for faster convergence. To tackle the large class imbalance, we weighted the cross-entropy loss during backpropagation to penalize the misclassification of the rarer class.

## Results

### Baseline performance

As the closest relevant baseline method, we computed the features described in [16] for this dataset. The results for a fivefold cross-validation scheme are shown in Fig. 9a.



**Fig. 9** Baseline performance. **a** ROC curve of baseline for the complete dataset. **b** ROC curve of baseline for the subset of the dataset with  $y^{\text{oct}} = 0$

We observed, however, that the performance of this baseline scheme highly depends on the presence of visible pulse effects in the OCT signal. To test this, we isolated a subset of the complete dataset, consisting only of the samples without OCT visibility. Of them, 45.6% corresponded to visible microscopic lesions, and the remaining 54.4% were not microscopically observed. The results on this reduced dataset are shown in Fig. 9b and assert our intuition that this method focuses primarily on “counting” or evaluating the presence of visible pulses. Given that in real human patient, eye motions are constant (i.e., the eye is not immobilized), such a strategy would be ill-suited.

Note that since the testset has a large class imbalance, the use of ROC curves is not appropriate. We show it here purely for demonstration purposes and to be in line with what the authors of [16] report. We will show results using a more appropriate metric in what follows and will use only the baseline results for the complete M-scan for comparison.

## CNN performance

A series of experiments was performed to explore our method’s robustness. We followed a fivefold cross-validation scheme, with the same split for both the CNN version and the baseline, and recorded the weighted accuracy, where each class was weighted inversely proportional to its size. The cross-validation split was done on an eye level, to avoid classification bias.

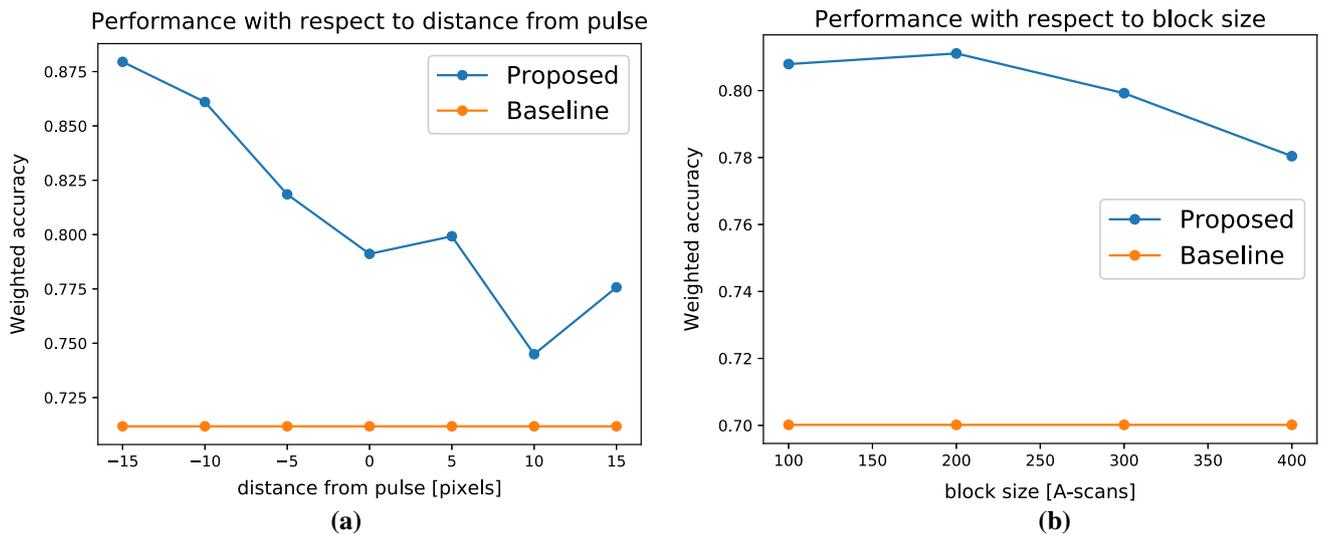
The first experiment investigates the performance’s dependence on the distance from the SRT pulse. The outcome of this experiment is shown in Fig. 10a. We can see that as the distance from the SRT pulse grows, the performance tends to drop. This confirms that the OCT observable effect of the

SRT is short in duration and that tissue relaxation occurs shortly after laser application [13]. It is therefore important to be able to evaluate the OCT signal in real time and in considerably short time intervals, to be sure not to miss crucial information. Also, when the distance is negative, meaning the pulse effect is in the evaluated block, the performance is better, something that is expected since the signal loss and its immediate aftermath contain most of the valuable signal information.

The second experiment probes the importance of the M-scan block size,  $\delta$ , and its results are shown in Fig. 10b. We observe that smaller M-scan sizes are not only adequate to evaluate the treatment, but even more appropriate than larger ones. This confirms the assumption that the immediate SRT effect is short in time and renders the fast evaluation necessary. To evaluate the applicability of the method for an eye movement scenario, we performed a test where we trained the network as described above, but in the testing set we randomly switched the initial blocks of each sample and substituted them with an initial block from another sample, of the same eye. We observed no significant change in the performance (0.78 weighted accuracy for the ‘moved’ samples and 0.79 for the still ones). We therefore trust that the presented results apply in the eye movement case.

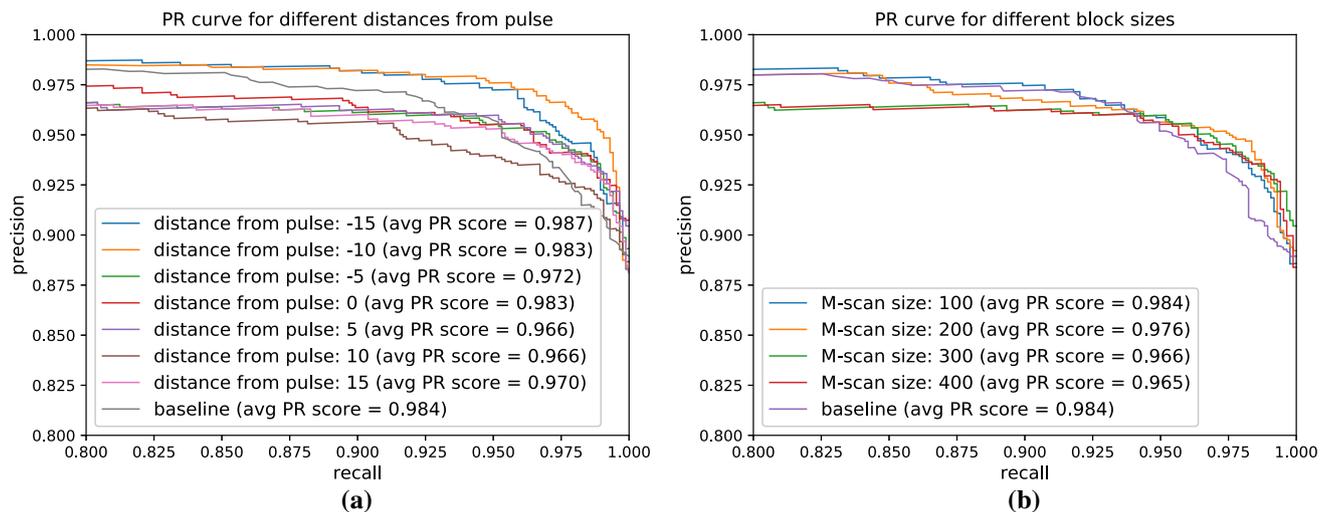
We also show the precision–recall curve for the aforementioned two experiments in Fig. 11.

In general, the precision–recall curves’ conclusions agree with the weighted accuracy ones. We can observe in the precision–recall curves that in some cases the baseline slightly outperforms the proposed method. This is to some extent expected given the nature of the task and the dataset content. Even in those cases, however, where the strong OCT visibility makes the baseline perform better, the difference



**Fig. 10** Evaluation of method robustness (weighted accuracy). **a** Weighted accuracy with respect to how far from the SRT pulse the evaluation was made (the block size was kept constant to a value of

$\delta = 300$ ). **b** Weighted accuracy with respect to the block size (the distance from the pulse was kept constant to a value of 5)



**Fig. 11** Evaluation of method robustness (PR curve). **a** PR curve for different distances from the SRT pulse (the block size was kept constant to a value of  $\delta = 300$ ). **b** PR curve for different block sizes (the distance from the pulse was kept constant to a value of 5)

is not significant and the proposed method compensates by managing to infer the treatment with just a short look at the signal, sometimes not even looking at the OCT effect. This guarantees the desired movement invariance which may not be necessary in ex vivo data, but is critical for an in vivo clinical setup.

Since our aim is toward a system that is able to evaluate M-scans during treatment, we would like to be independent of the presence and number of visible pulse effects in the OCT. Also, we want to be able to infer the treatment status from as short an M-scan block as possible. This request is linked to the real-time aspect of the task. The shorter a block

is, the faster it is evaluated. In addition, in an in vivo patient scenario the unavoidable eye movement causes the resulting full M-scan to be comprised of A-scans of different retinal spots. The shorter a block is, the more likely it describes a more concentrated retinal area. This way, the evaluation of such a block can provide information about the accumulated treatment effect on that specific area. The number of visible SRT pulse effects on the other hand does not take into account the acquisition location, and the analysis of a full M-scan does not reveal the true SRT effect at different locations due to movement, thus being more prone to overestimating the treatment effect.

## Conclusions

In this work, we introduced a method that uses OCT M-scans to automatically classify SRT lesions. The descriptive capability of a convolutional neural network allowed us to evaluate the accumulated SRT energy deposition effect by only looking at a short signal portion. The results on ex vivo porcine data show a promising ability to assess the progress of a treatment by just evaluating a small M-scan block, therefore leaving room to early judgment before more energy is deposited and thus potentially avoiding over-treatment. More importantly, the adequate small M-scan size needed for classification allows us to safely assume eye motion invariance of our setup system. Involuntary eye movement occurs approximately every 5–50 ms [3]. The lower limit of 5 ms corresponds to an M-scan block of 75 and 250 A-scans, for an OCT acquisition frequency of 15 and 50 kHz, respectively. The results of the current work indicate that the evaluation of such a block can lead to good performance.

We evaluated our approach on a dataset that included a variety of acquisition frequencies, energy levels and treatment modes (ramping mode and classical mode). We also observed high inter-data variability, due to different eyes and differences in postmortem time interval from the measurement time. We showed that the proposed pipeline is robust to all these parameters and exhibits good generalization ability, without the need to provide the network with specific information regarding them.

Hence, OCT feedback can fill a gap in the area of sub-threshold therapies, both by providing a direct view of the target tissue and by serving as an automatic stopping mechanism, when the classification algorithm indicates an adequate RPE lesion. This system has the strong potential to assist in a wider use of SRT therapy as a minimally invasive and minimally harmful option when confronting a number of retinal diseases. Future work will include a more thorough examination of the performance quality needed for clinical application. Given, however, that current clinical practice relies on fundus images and quite commonly applies more energy than needed, there is a strong possibility that the proposed method would not introduce additional risks.

**Acknowledgements** The Titan X NVIDIA GPU used for this research was donated by the NVIDIA Corporation. We also thank Lutronic Corporation for their support in this research and for providing the R:GEN laser used in this study. We too thank Christian Burri for the data collection, and Stefanos Apostolopoulos for the valuable scientific support on the OCT data manipulation.

**Funding** This work was partly funded by the Swiss National Science Foundation Grant No. 325230-141189.

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

**Ethical approval** No ethical approval was necessary in this work.

## References

1. Brinkmann R, Roeder J, Birngruber R (2006) Selective retina therapy (SRT): a review on. *Bull Soc Belge Ophthalmol* 302:51–69
2. Chopdar A (2014) Fundus fluorescein angiography. In: *Multimodal retinal imaging*, chap 1. JP Medical Ltd, pp 1–10
3. Findlay J (1971) Frequency analysis of human involuntary eye movement. *Biol Cybern* 8(6):207–214
4. Fujimoto JG, Pitris C, Boppart SA, Brezinski ME (2000) Optical coherence tomography: an emerging technology for biomedical imaging and optical biopsy. *Neoplasia* 2(1–2):9–25
5. He K, Zhang X, Ren S, Sun J (2015) Deep residual learning for image recognition. *arXiv preprint arXiv:1512.03385*
6. Kaufmann D, Burri C, Arnold P, Koch VM, Meier C, Považay B, Justiz J (2017) Dosimetry control and monitoring of selective retina therapy using optical coherence tomography. In: *European conference on biomedical optics*, p 1041604
7. Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. In: *Advances in neural information processing systems*, pp 1097–1105
8. Park YG, Seifert E, Roh YJ, Theisen-Kunde D, Kang S, Brinkmann R (2014) Tissue response of selective retina therapy by means of a feedback-controlled energy ramping mode. *Clin Exp Ophthalmol* 42(9):846–855
9. Roeder J, Hillenkamp F, Flotte T, Birngruber R (1993) Microphotocoagulation: selective effects of repetitive short laser pulses. *Proc Natl Acad Sci* 90(18):8643–8647
10. Roeder J, Liew SHM, Klatt C, Elsner H, Poerksen E, Hillenkamp J, Brinkmann R, Birngruber R (2010) Selective retina therapy (SRT) for clinically significant diabetic macular edema. *Graefes Arch Clin Exp Ophthalmol* 248(9):1263–1272
11. Savitzky A, Golay MJ (1964) Smoothing and differentiation of data by simplified least squares procedures. *Anal Chem* 36(8):1627–1639
12. Schuele G, Elsner H, Framme C, Roeder J, Birngruber R, Brinkmann R (2005) Optoacoustic real-time dosimetry for selective retina treatment. *J Biomed Opt* 10(December):064,022–064,022
13. Steiner P, Ebnetter A, Berger LE, Zinkernagel M, Považay B, Meier C, Kowal JH, Framme C, Brinkmann R, Wolf S, Sznitman R (2015) Time-resolved ultra-high resolution optical coherence tomography for real-time monitoring of selective retina therapy. *Invest Ophthalmol Vis Sci* 56(11):6654–6662
14. Strauss O (2005) The retinal pigment epithelium in visual function. *Physiol Rev* 85(3):845–881
15. Yaqoob Z, Wu J, Yang C (2005) Spectral domain optical coherence tomography: a better OCT imaging strategy. *Biotechniques* 39:6–13
16. Zbinden S, Kucur ŞS, Steiner P, Wolf S, Sznitman R (2016) Automatic assessment of time-resolved OCT images for selective retina therapy. *Int J Comput Assist Radiol Surg* 11(6):863–871