OPTICAL COHERENCE TOMOGRAPHY (OCT) - GUIDED OPHTHALMIC THERAPY

By

Soohyun Lee

A dissertation submitted to Johns Hopkins University in conformity with
the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland

July 2021

Abstract

Optical coherence tomography (OCT), which provides cross-sectional images noninvasively with a micro-scale in real-time, has been widely applied to the diagnosis and treatment guidance for various ocular diseases. In this work, we demonstrate OCT-based guidance of two ophthalmic therapies, subretinal injection and selective retina therapy (SRT).

Firstly, the "SMART," a hand-held robotic surgical device actively guided by a common-path OCT (CP-OCT) distal sensor, improves in two aspects for being applied to subretinal injection: (i) A high-performance fiber probe based on high index epoxy lensed-fiber to enhance the CP-OCT retinal image quality in a wet environment; (ii) Automated retinal layer identification and tracking : retinal layer boundaries, as well as retinal surface, are tracked using convolutional neural network (CNN)-based segmentation for accurate subretinal injection guidance. It is shown that properly designed high index epoxy lensed-fiber probe improves the signal-to-noise ratio (SNR) and retinal image quality of the CP-OCT system. We propose, implement, and study real-time retinal boundary tracking of A-scan OCT images using CNNs for automatic depth targeting of a selected retinal boundary and accurate localization of a surgical tool (i.e. needle) tip. A simplified 1-D U-net is used for the retinal layer segmentation of A-scan OCT images which are obtained by the lensed-fiber probes. A Kalman filter, combining retinal boundary position measurement by CNN-based segmentation and velocity measurement by cross-correlation between consecutive A-scan images, is applied to optimally estimate the retinal boundary position. Unwanted axial motions of the surgical tools are compensated by a piezo-electric linear motor based on the retinal boundary tracking. A CNN-based CP-OCT distal sensor successfully tracks retinal boundaries, especially the PR/CH boundary for subretinal injection, and automatically guides the needle's axial

position in real-time. The micro-scale depth targeting accuracy of our system shows its promising possibility for clinical application.

We also propose and demonstrate SRT monitoring based on speckle variance OCT (svOCT) for dosimetry control. SvOCT quantifies speckle pattern variation caused by moving particles or structural changes in biological tissues. M-scans, time-resolved sequence of A-scans, of a phantom, *ex vivo* bovine iris, and *ex vivo* bovine retina are obtained by a swept-source OCT system during laser pulses irradiation. SvOCT images are calculated as interframe intensity variance of the sequence, and they show abrupt speckle variance change induced by laser pulse irradiation. The axially averaged svOCT signals show a sharp peak corresponding to each laser pulse, and the peak values are proportional to irradiated laser pulse energy. For the *ex vivo* retinal study, microscopic images of treated spots are obtained before and after removing the upper neural retinal layer to assess the damage in both RPE and neural layers. Spatial and temporal temperature distributions in the retina are numerically calculated in a 2D retinal model using COMSOL Multiphysics. We find that the svOCT peak values have a reliable correlation with the degree of retinal lesion formation. The temperature at the neural retina and RPE is estimated from the svOCT peak values using numerically calculated temperature, which is consistent with the observed lesion creation.

Candidate:

Soohyun Lee

Dissertation Committee:

Dr. Jin U. Kang (advisor)

Dr. Israel Gannot (Second reader)

Dr. Mark Foster

Acknowledgments

I would like to thank the following people for their help and support during my PhD study. I would first like to thank my advisor, Dr. Jin U. Kang, for his patient support and all of the opportunities I was given to further my research. I would also like to thank my committee members, Dr. Israel Gannot and Dr. Mark foster, for their valuable comments and support. I would like to thank all the photonics and optoelectronics group members for their support and for all the fun we have had in the last five years. In addition, I must express my very profound gratitude to my parents and to my husband for their constant love and support throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them. Thank you.

Contents

List of Tables

List of Figures

1

Introduction

1.1 Motivations

Subretinal injection is becoming increasingly popular in both scientific research and clinical communities as an efficient way of treating retinal diseases. The treatments deliver drugs or stem cells directly into subretinal space between the RPE and photoreceptor layer, thereby effectively affecting resident cells and tissues in the subretinal space. However, the procedure requires micro-scale precision due to the delicate anatomy of the retina. It makes the procedure challenging because of surgeons' physiological hand tremor [1,2] and limited depth perception and limited visual feedback from a traditional stereo-microscopic *en-face* view.

Optical coherence tomography (OCT)-guided robotic systems have been developed to reduce the unintended physiological motion and overcome the limited visual feedback during ocular microsurgery. OCT, which provides micro-scale resolution cross-sectional images in real-time [3], enables improved visualization and accurate guidance of robotic systems. Among them, fiber-optic common-path OCT (CP-OCT) distal sensor integrated hand-held surgical devices have been developed to implement simple, compact, and cost-effective microsurgical systems [4-7]. In those systems, a single fiber probe attached to a surgical tooltip (i.e., needle or micro-forceps) guided the hand-held surgical device by real-time A-scan-based surface tracking. However, surface tracking-based guidance could induce inaccurate depth targeting for subretinal injection because of retinal thickness variations and irregular morphological features caused by retinal diseases. The target or near target retinal boundary tracking, which is RPE and photoreceptor

boundary tracking for subretinal injection, allows precision guidance. Therefore, we apply convolutional neural network (CNN)-based retinal layer segmentation, which already have shown promising results for 2D or 3D OCT retinal image segmentation [8-11], for retinal boundary tracking and implement hand-held surgical device compensating undesired hand tremor in real-time using the retinal boundary tracking.

Selective retina therapy (SRT) is an effective laser treatment method for various retinal diseases associated with a degradation of the retinal pigment epithelium (RPE) [12-17]. The RPE, which contains a high concentration of melanosomes, can be selectively targeted by using a pulse laser having pulse duration shorter than a thermal relaxation time of the RPE (~10 µs) [18]. The SRT reduces negative side effects by avoiding thermal damages of the adjacent layers. However, the selection of proper laser energy—which is crucial for successful SRT without excessive burning and collateral damage—is challenging because lesions in the RPE are invisible ophthalmoscopically. In addition, melanin concentration variations among patients or regions even within an eye [19] make it impossible to set a static threshold value of pulse energy of a therapeutic irradiation window. Fundus fluorescence angiography (FFA) is an accurate method to detect the lesions, but it requires the use of fluorescent dye injection [20] and a long delay between treatment and detection. SvOCT is expected to work effectively for real-time non-invasive SRT monitoring by detecting speckle variation changes induced by morphological and structural changes of retinal tissue by laser irradiation. Therefore, we propose and demonstrate svOCT-based SRT monitoring method for dosimetry control.

1.2 Dissertation overview

This dissertation presents the development of CNN-based real-time OCT distal sensor and svOCT-based SRT monitoring system. Chapter 2 reviews the fundamental principles of OCT and CNN. Chapter 3 presents design and analysis of high-index lensed fiber probe for retinal imaging. Chapter 4 describes CNN-based CP-OCT distal sensor integrated with a subretinal injector for retinal boundary tracking and injection guidance. Chapter 5 presents SRT monitoring method using svOCT and temperature estimation. Chapter 6 summarizes the dissertation and proposes future directions of research.

2

Introduction to OCT and Neural Networks

2.1 Optical coherence tomography

2.1.1 Time domain and Fourier domain optical coherence tomography

A Michelson-type interferometer is typically used for an OCT system with a low-coherence light source as shown in Fig.1. The low coherent, broadband light is split into a reference and a sample arm by a beam splitter, and the returning light from each arm, which is reflected and backscattered from a reference mirror or sample, is recombined and produces an interference pattern. A

Reference reflector
r_R

$$E_R = \frac{E_i}{\sqrt{2}} r_R e^{i2kz_R}$$

z_R

$$E_S = \frac{E_i}{\sqrt{2}} \left(r_S(z_S) \otimes e^{i2kz_S} \right)$$

$$E_i = s(k, \omega) e^{i(kz - \omega t)}$$

Sample

Light source

z_{S1}

z_{S2}

z_{S3}

$$r_S(z_S) = \sum_{n=1}^{N} r_{Sn} \delta(z_S - z_{Sn})$$

$$I_D = \langle |E_R + E_S|^2 \rangle$$

Photodetector
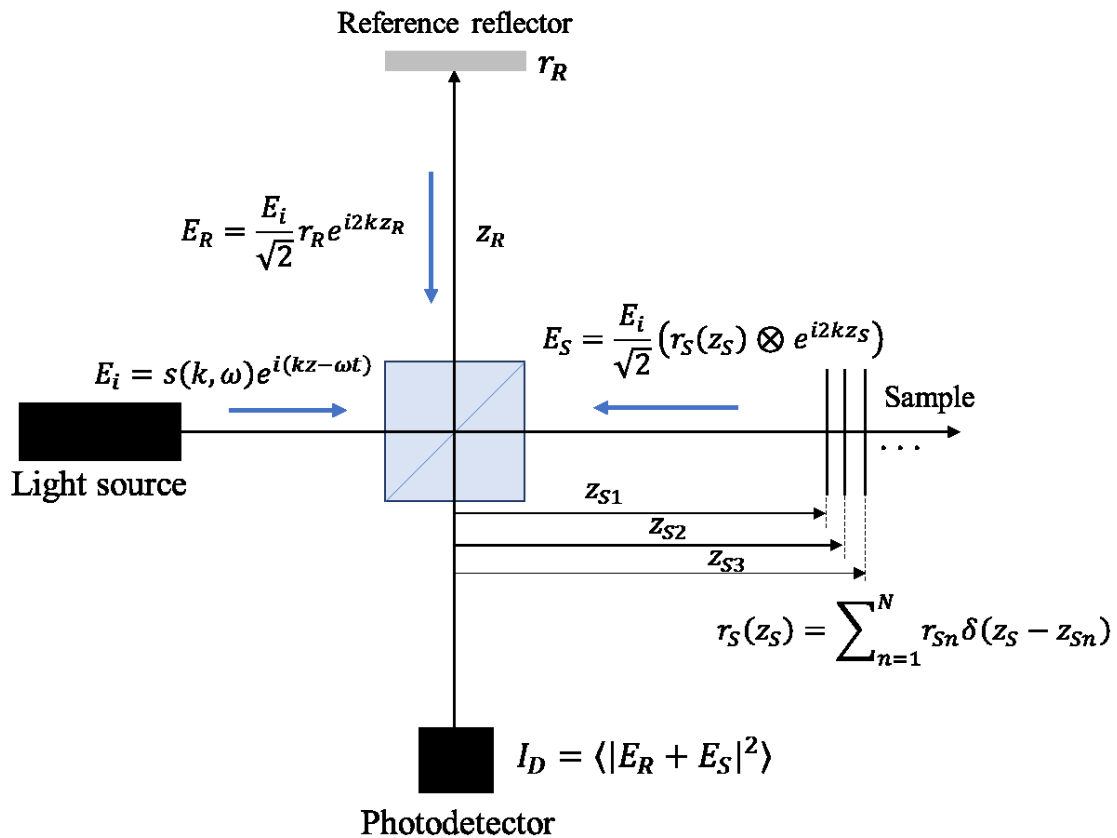
Figure 1 Schematic of a Michelson-type interferometer used in OCT

reflectivity profile of a sample can be reconstructed from the interferometric measurements, and its mathematical derivation will be described in this section.

The light source emits a polychromatic plane wave, whose electric field can be written as $E_i = s(k, \omega)e^{i(kz-\omega t)}$. Here, $s(k, \omega)$ is the electric field amplitude as a function of wave number k and angular frequency ω. The reference reflector is assumed to have electric field reflectivity r_R and power reflectivity $R_R = |r_R|^2$. In general, depth-dependent electric field reflectivity profile of sample, $r_S(z_S)$, is continuous, resulting from continuously changing refractive index of samples. However, for simplification, we assume a series of N discrete, real delta-function reflections of the form $r_S(z_S) = \sum_{n=1}^{N} r_{Sn}\delta(z_S - z_{Sn})$, with each reflection characterized by its electric field reflectivity, r_{Sn}, and path length from the beam splitter, z_{Sn}. When the surface of beam splitter is set to $z = 0$, returning light from the reference and sample arms are given by

$$E_R = \frac{E_i}{\sqrt{2}} r_R e^{i2kz_R} \tag{1}$$

$$E_S = \frac{E_i}{\sqrt{2}}\left(r_S(z_S) \otimes e^{i2kz_S}\right) = \frac{E_i}{\sqrt{2}}\sum_{n=1}^{N} r_{Sn}e^{i2kz_{Sn}} \tag{2}$$

where $\otimes$ represents convolution operation. The intensity of the interference pattern obtained by the returning light can be expressed as

$$I_D(k, \omega) = \frac{1}{2}\langle|E_R + E_S|^2\rangle = \frac{1}{2}\langle(E_R + E_S)(E_R + E_S)^*\rangle$$

$$= \frac{1}{2}\langle\left|\frac{s(k,\omega)}{\sqrt{2}} r_R e^{i(2kz_R-\omega t)} + \frac{s(k,\omega)}{\sqrt{2}}\sum_{n=1}^{N} r_{Sn}e^{i(2kz_{Sn}-\omega t)}\right|^2\rangle, \tag{3}$$

where the factor of two counts for the second pass of each field through the beam splitter and the angular brackets indicate integration over the response time of the detector. Because the response time of current detectors is much longer than light wave oscillations, the terms dependent on the temporal angular frequency, ω, can be eliminated by integrating the terms over the detector response time. Then, we can obtain temporally invariant interference intensity expressed as

$$I_D(k) = \frac{1}{4}\left[S(k)[R_R + \sum_{n=1}^{N} R_{Sn}]\right] + \frac{1}{4}\left[S(k)\left[\sum_{n=1}^{N} \sqrt{R_R R_{Sn}}\left(e^{i2k(z_R - z_{Sn})} + e^{-i2k(z_R - z_{Sn})}\right)\right]\right]$$

$$+ \frac{1}{4}\left[S(k)\left[\sum_{n \neq m=1}^{N} \sqrt{R_{Sn} R_{Sm}}\left(e^{i2k(z_{Sn} - z_{Sm})} + e^{-i2k(z_{Sn} - z_{Sm})}\right)\right]\right], \tag{4}$$

where $S(k) = \langle |s(k,\omega)|^2 \rangle$ is a wave number power spectrum of the light source, and $R_{Sn} = |r_{Sn}|^2$ is power reflectivity of the sample reflectors. Using Euler's rule, Eq. (4) is simplified to

$$I_D(k) = \frac{1}{4}\left[S(k)[R_R + \sum_{n=1}^{N} R_{Sn}]\right] + \frac{1}{2}\left[S(k)\left[\sum_{n=1}^{N} \sqrt{R_R R_{Sn}} \cos\left(2k(z_R - z_{Sn})\right)\right]\right]$$

$$+ \frac{1}{2}\left[S(k)\left[\sum_{n \neq m=1}^{N} \sqrt{R_{Sn} R_{Sm}} \cos\left(2k(z_{Sn} - z_{Sm})\right)\right]\right]. \tag{5}$$

The first part of Eq. (5) is DC terms independent of the path length z_R and z_{Sn}, and it is scaled by the light source power spectrum and the sum of power reflectivities of the reference mirror and sample reflectors. This is the largest component of the interference intensity because the reference reflectivity generally dominates the sample reflectivity which is typically very small on the order of 10^{-4} to 10^{-5}. The second part is cross-correlation terms representing the interference between the beams from the reference mirror and each sample reflector. These are the desired terms for OCT image reconstruction that visualize sample reflectivity profile. Since these terms are proportional to the multiplication of the square root of the reference reflectivity and sample reflectivity, they are typically smaller than the DC component. The last part is autocorrelation terms representing interference between the different sample reflectors, and it is considered as artifacts in a typical OCT system. These terms are linearly dependent on the power reflectivity of the sample and small compared to the DC and cross-correlation terms with dominant reference reflectivity.

In a time-domain OCT (TDOCT) system, a single photodetector, which cannot resolve the individual contributions of the wavenumber k, is used to measure the interference intensity $I_D(k)$ in Eq. (5), while the reference delay z_R is scanned to reconstruct an approximation of the internal

sample reflectivity profile $\sqrt{R_S(z_S)}$. The measurement corresponds to integration of $I_D(k)$ over all k as

$$I_D(z_R) = \frac{1}{4}\left[\int_0^\infty S(k)dk\left[R_R + \sum_{n=1}^N R_{Sn}\right]\right] + \frac{1}{2}\left[\sum_{n=1}^N \sqrt{R_R R_{Sn}}\int_0^\infty S(k)\cos\left(2k(z_R - z_{Sn})\right)dk\right]$$

$$= \frac{1}{4}\left[S_0\left[R_R + \sum_{n=1}^N R_{Sn}\right]\right] + \frac{1}{2}\left[S_0 \sum_{n=1}^N \sqrt{R_R R_{Sn}}\ \gamma(z_R - z_{Sn})\cos\left(2k_0(z_R - z_{Sn})\right)\right], \qquad (6)$$

where S_0 is the spectrally integrated power of the source, and the coherence function $\gamma(z_R - z_{Sn})$ is the inverse Fourier transform of the normalized light source spectrum $S(k)$. Autocorrelation terms are removed as being assumed to be negligible compared to the DC terms and cross-correlation terms. The sample reflectivity $\sqrt{R_{Sn}}$ is convolved with the coherence function $\gamma(z_R - z_{Sn})$ and modulated by a cosinusoidal carrier at a frequency proportional to the central wave number of the light source spectrum, k_0. The sample reflectivity profile could be obtained by envelop detection of the cross-correlation terms as a form of convolved reflectivity. In the followings, the interference intensity as a function of z_R, A-scans, of a discrete-reflectors sample is derived assuming a Gaussian light source, which approximates the shape of actual light sources and also has useful Fourier transform properties as below:

$$\gamma(z) = \exp(-z^2\Delta k^2) \overset{F}{\leftrightarrow} S(k) = \frac{1}{\Delta k\sqrt{\pi}}\exp\left(-\left(\frac{k-k_0}{\Delta k}\right)^2\right). \qquad (7)$$

Here, k_0 represents the central wave number of the light source spectrum, and Δk represents its spectral bandwidth. The interference intensity is expressed as

$$I_D(z_R) = \frac{1}{4}\left[S_0\left[R_R + \sum_{n=1}^N R_{Sn}\right]\right]$$

$$+ \frac{1}{2}\left[S_0 \sum_{n=1}^N \sqrt{R_R R_{Sn}}\exp(-(z_R - z_{Sn})^2\Delta k^2)\cos\left(2k_0(z_R - z_{Sn})\right)\right] \qquad (8)$$

and Fig. 2 shows the illustrative example of the measurement.
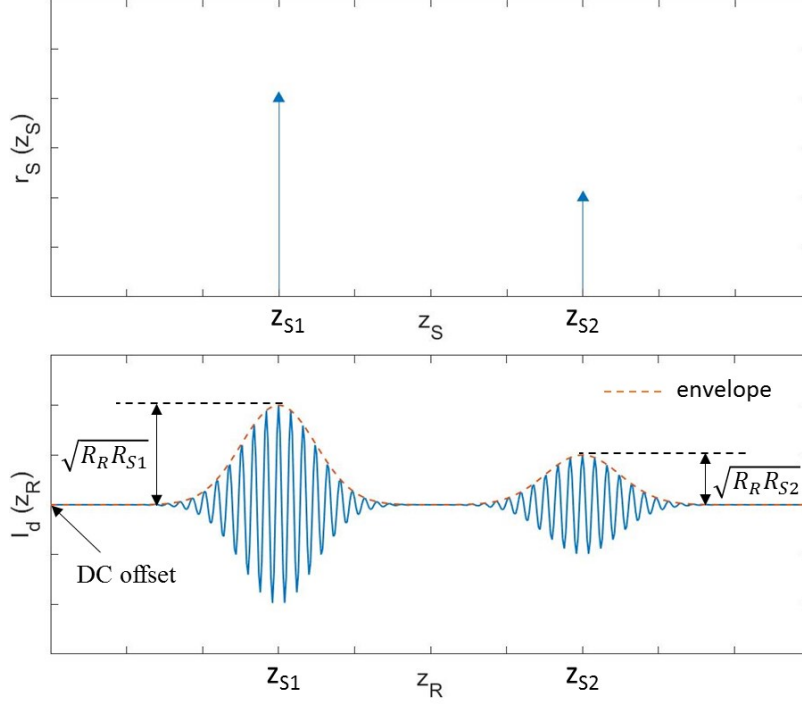
Figure 2 The reflectivity profile of a discrete-reflectors sample and the resultant A-scan of TDOCT

In a Fourier-domain OCT (FD-OCT) system, the wave number-dependent interference intensity $I_D(k)$ in Eq. (5) is captured and processed using Fourier analysis to reconstruct an approximation of the sample reflectivity profile $\sqrt{R_S(z_S)}$. In contrast to TD-OCT, the interferogram contains information for the entire depth profile of the sample simultaneously and does not need mechanical scanning of the optical path length. Using the Fourier transform pair $\frac{1}{2}[\delta(z - z_0) + \delta(z + z_0)] \overset{F}{\leftrightarrow} \cos(kz_0)$ and the convolution property of Fourier transform, the inverse Fourier transform of Eq. (5) can be expressed as

$$i_D(z) = \frac{1}{4}\left[\gamma(z)[R_R + \sum_{n=1}^{N} R_{Sn}]\right] + \frac{1}{4}\left[\gamma(z) \otimes \left[\sum_{n=1}^{N} \sqrt{R_R R_{Sn}}[\delta(z \pm 2(z_R - z_{Sn}))]\right]\right]$$

$$+ \frac{1}{4}\left[\gamma(z) \otimes \left[\sum_{n \neq m=1}^{N} \sqrt{R_{Sn} R_{Sm}}[\delta(z \pm 2(z_{Sm} - z_{Sn}))]\right]\right], \tag{9}$$

where the sample reflective profile $r_S(z_S) = \sum_{n=1}^{N} r_{Sn}\delta(z_S - z_{Sn})$ which we would like to obtain is embedded within cross-correlation terms of Eq. (9). Conducting convolutions by using shift

8

property of the delta function, A-scan, the reflectivity profile of sample in depth, is expressed as below:

$$i_D(z) = \frac{1}{4}\big[\gamma(z)[R_R + \sum_{n=1}^{N} R_{Sn}]\big] + \frac{1}{4}\Big[\big[\sum_{n=1}^{N} \sqrt{R_R R_{Sn}}[\gamma(2(z_R - z_{Sn})) + \gamma(-2(z_R - z_{Sn}))]\big]\Big]$$

$$+ \frac{1}{4}\Big[\big[\sum_{n \neq m=1}^{N} \sqrt{R_{Sn} R_{Sm}}[\gamma(2(z_{Sm} - z_{Sn})) + \gamma(-2(z_{Sm} - z_{Sn}))]\big]\Big]. \tag{10}$$

The sample reflectivity profile is reconstructed in the cross-correlation terms with some modifications. It appears as a function of the difference between the reference reflector distance z_R and the sample reflector distance z_S rather than just z_S. The displacement of each sample reflector from the reference position is doubled in A-scan because the interferometer measures the round-trip distance to each reflector. In addition, each sample reflector is broadened or blurred out to a width of about a coherence length by convolution with the coherence function $\gamma(z)$.

Additional image artifacts also appear in each term of Eq. (10). As seen in the cross-correlation terms, a mirror image of the blurred reflectors appears on the opposite side of zero path length, which is the reference reflector distance. It is called mirror image artifact or complex conjugate artifact in FDOCT, and it is simply understood from the fact that inverse Fourier transform of real function (interferometric spectrum) must be Hermitian symmetric. This artifact can be easily dealt with by displaying only the positive or negative distances if the sample is placed on one side of zero path length. However, if the sample strays over the zero-path length border, it starts to overlap its mirror image, and the mirror artifact cannot be removed by image processing alone. The DC terms produce a large artifactual signal centered at zero path length difference. The signal amplitude is so much larger than the desired cross-correlation terms due to the dominant reference reflectivity. A simple method to eliminate that component is to record the amplitude of the spectral signal with only the reference reflector and then to subtract this signal component from each subsequent spectral interferometric signal acquired. The autocorrelation terms in Eq. (10)

also give rise to artifactual signals. However, if reference reflectivity is sufficiently high, which is the most case in practice, its amplitude could be negligible compared to cross-correlation terms. The spectral interferogram and A-scan for the example of a discrete-reflectors sample and Gaussian-shaped source spectrum is illustrated in Fig. 3 including all the artifacts described previously.
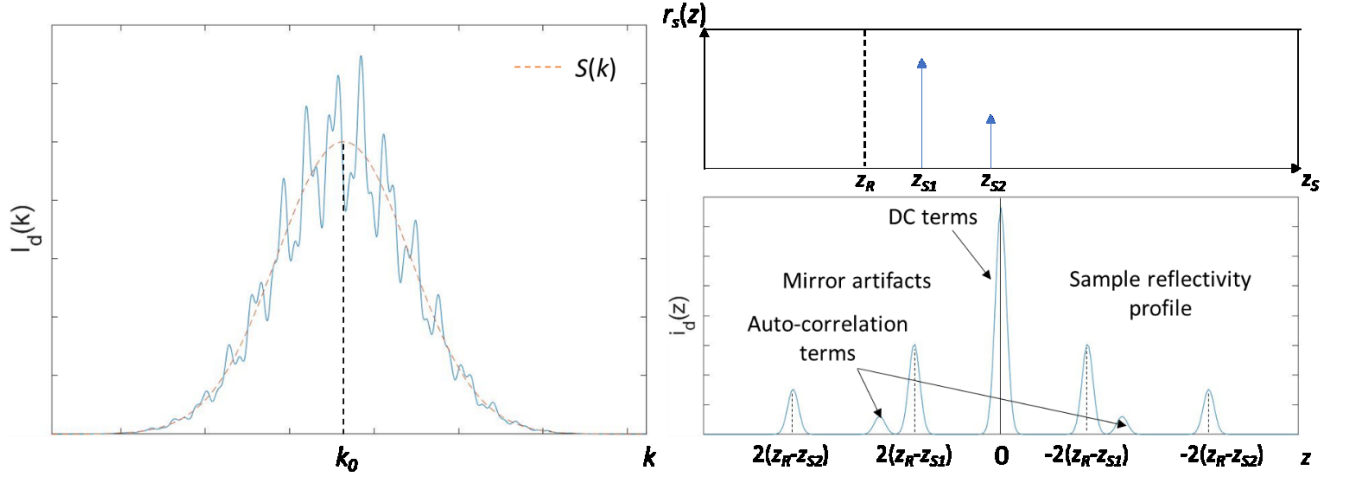
Figure 3 (Left) Spectral interferogram. (Right) The reflectivity profile of a discreate-reflectors sample and resultant A-scan obtained by FDOCT

The axial and transverse resolution of OCT is decoupled from each other. The axial resolution is determined by the coherence length of a light source, which is the full width at half the maximum (FWHM) of coherence function $\gamma(z)$. It is an explicit function of the light source bandwidth, stated in both wave number and wavelength terms as

$$\delta z = \frac{2\sqrt{ln2}}{\Delta k} = \frac{2\ln(2)}{\pi}\frac{\lambda_0^2}{\Delta\lambda} \tag{11}$$

where $\lambda_0 = \frac{2\pi}{k_0}$ is the center wavelength of the light source and $\Delta\lambda$ is its wavelength bandwidth defined as the FWHM of its wavelength bandwidth (so that $\Delta k = \frac{\pi}{\sqrt{\ln(2)}}\frac{\Delta\lambda}{\lambda_0^2}$). The transverse resolution is determined by the minimum spot size of the focused sample beam, a parameter which is inversely proportional to the numerical aperture (NA) of the focusing lens [21]:

$$\delta x = 0.37 \frac{\lambda_0}{NA} \tag{12}$$

It should be noted that there is a trade-off between transverse resolution and depth of field. A high NA features a great focusing power which allows high transverse resolution with a corresponding short depth of field. Meanwhile, a low NA would result in a greater diameter of the beam at the focal point but a large depth of field.

2.1.2 Spectral-domain and swept-source optical coherence tomography

FD-OCT techniques have allowed a dramatic improvement in signal-to-noise ratio (SNR) and imaging speed [22-24] compared to TD-OCT. FD-OCT could be implemented either through the use of a wavelength-swept light source and a standard photodiode receiver or a broadband light source and a spectrometer. The configuration using a spectrometer has been referred to as spectral domain OCT (SD-OCT); and the configuration using the wavelength-swept light source has been referred to as swept source OCT (SS-OCT).

SD-OCT is depicted in Fig. 4 (left): it is similar to TD-OCT, but the point detector is replaced by a spectrometer and a charge-coupled device (CCD) or CMOS linear array. The spectrometer uses a diffractive element to spatially separate the different wavelength contributions into a line image which is recorded by a CCD array. A superluminescent diode (SLD) is commonly used as a broadband light source, because it has a broad bandwidth and a relatively high-power output. The depth range (z_{max}) is determined by the wavenumber spacing between pixels (δk) of the spectrometer, and is given by

$$z_{max} = \frac{\pi}{2\delta k} \tag{13}$$

$$\delta k = \frac{2\pi\Delta\lambda}{N\lambda_0^2} \tag{14}$$

where N is pixel number of the detector the spectrum is imaged on. The acquisition speed is limited by the line-scan rate of the CCD camera. In addition, non-linear k-sampling of SD-OCT, which is due to spectrometer angular dispersion, should be properly converted during the preprocessing of the wavelength to k-space mapping to avoid a depth-dependent broadening of the coherence function similar in appearance to dispersion in structural OCT images.

SS-OCT also has a similar optical setup with TD-OCT, but the broadband light source is replaced by an optical source which rapidly sweeps a narrow line-width over a broad range of wavelengths as shown in Fig. 4 (right). During one sweep, each wavelength component of the interferometric signal is detected sequentially by a high-speed photodetector. In SS-OCT, the acquisition rate is given by the sweep rate of the swept-source and subsequent analog-digital (AD) conversion, and commercially available sources can realize high sweep rates (>100 kHz), which require ultrafast detection and AD conversion in the GHz range. One wavelength sweep constitutes a spectral interferogram with fringe patterns, as in SD-OCT. The depth range (z_{max}) is also determined by the wavenumber resolution (δk), which is limited by the instantaneous line shape of the sweeping laser source and expressed by Eq. (14), where N is the number of readouts of the photodetector during one sweep of the light source. In addition, linear k-sampling can be achieved in a hardware scheme by using external sampling clocks, called k-clocks, coupled with high-end acquisition electronics capable of nonuniform sampling frequencies.
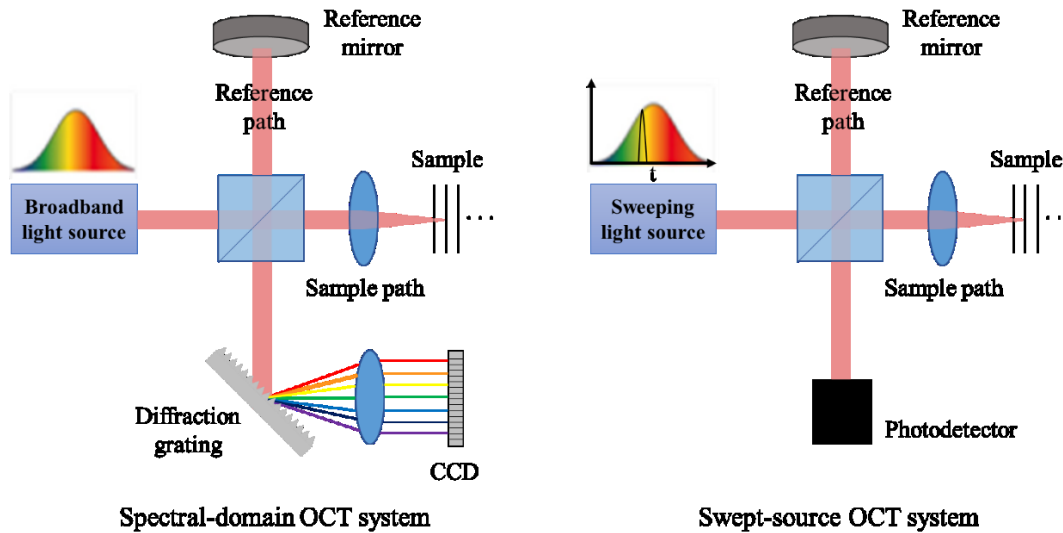
Figure 4 Schematic of (left) SD-OCT system and (right) SS-OCT system

2.1.3 Common-path optical coherence tomography

Common-path OCT (CP-OCT) uses a shared beam path for the reference and the sample arms. An optical surface (i.e. glass plate surface, cleaved fiber surface) placed near the sample serves as a reference reflector. Light reflected from the reference reflector and the sample interferes with each other and is directed back to the optical sensor. Figure 5 shows a schematic of SD-CPOCT in free space, in which the back surface of the glass plate serves as a reference reflector, and fiber-optic based SS-CPOCT, in which cleaved fiber surface serves as a reference reflector. The CP-OCT approach requires no alignment and has higher image stability [25]. The architecture could be simple, compact, and cost-effective. Especially, fiber-optic-based CP-OCT system allows a single optical fiber to work as an OCT probe and be easily integrated into existing medical instrumentation (i.e. a microsurgical tool) [26]. In addition, it is free from polarization and dispersion mismatch caused by optical elements in the interferometer and insensitive to vibration [27].
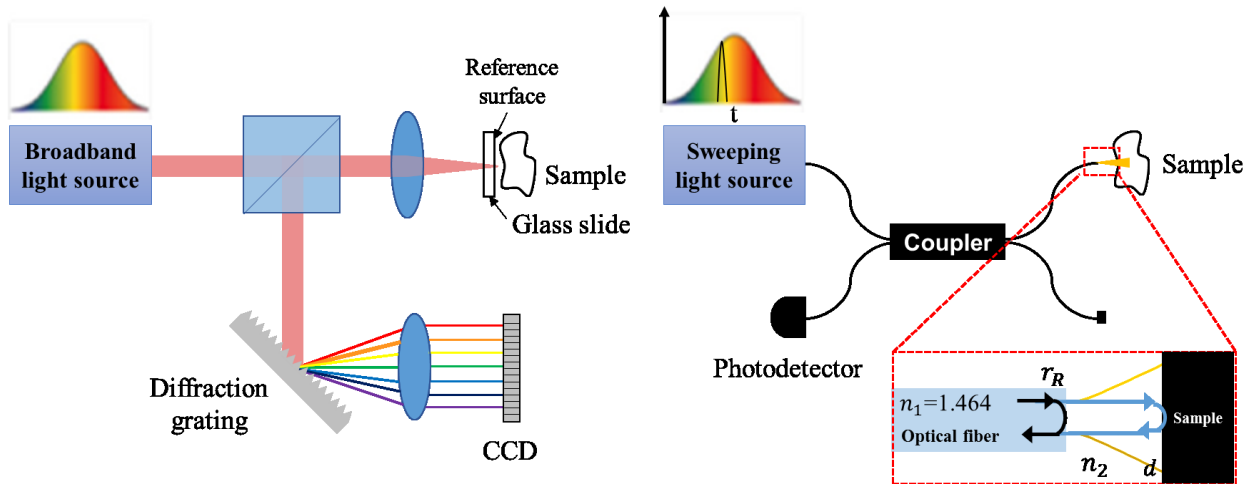
Figure 5 Schematic of (left) free space SD-CPOCT system and (right) fiber optic-based SS-CPOCT system

Nevertheless, one issue of using CP-OCT system is the unadjustable reference reflection which mostly comes from the Fresnel (partial) reflection at a reference surface. Fresnel reflection at a reference surface is expressed by

$$R = \left(\frac{n_1-n_2}{n_1+n_2}\right)^2 \tag{15}$$

for the case of normal incidence, where n_1 and n_2 is refractive index of incident and refracted medium. It is more challenging for the CP-OCT system to adjust reference reflection to optimal reflection than standard OCT system. Table 1 shows the reflectivity at the optical surface of various medium.

Table 1 Reflectivity at the optical surface of various medium.

Incident medium(n_1)	Glass (1.5)	Glass (1.5)	Fiber core (1.464)	Fiber core (1.464)
Refracted medium (n_2)	Air (1)	Water (1.33)	Air (1)	Water (1.33)
Reflection	4×10^{-2}	3.6×10^{-3}	3.56×10^{-2}	2.3×10^{-3}

2.1.4 Speckle-variance optical coherence tomography

Speckle is a random phenomenon generated by interference of mutually coherent waves with random phases. It is a fundamental property of signals and images acquired by all types of narrowband detection systems, which include radar, ultrasound, and radio astronomy. In addition to the optical properties and motion of the target object, the speckle is influenced by the size and temporal coherence of the light source, multiple scattering and phase aberrations of the propagating beam, and the aperture of the detector. All of these variables contribute to the observed characteristics of speckle in optical coherence tomography of living tissue [28]. Its role in imaging, however, is mainly that of noise, so considerable attention has been devoted to methods for speckle reduction. In medical OCT, speckle can mask diagnostically significant image features and reduce the accuracy of segmentation algorithms.

At the same time, there have been many attempts to utilize speckle as an information carrier. One of the attempts is the development of speckle variance OCT (svOCT), which is mostly applied for OCT angiography. If an OCT image is acquired in a stationary object, the speckle pattern is temporally stationary as well [29]. When moving particles exist in biological tissues (such as red blood cells), the speckle pattern varies with time and can be quantified by speckle variance calculations using either interframe or interline comparisons. SvOCT evaluates the speckle variance in the OCT structure intensity across the desired number of B-scan images, preferably acquired at the same location, using the following equation:

$$I_{SV(j,k)} = \frac{1}{N}\sum_{i=1}^{N}\left(I_{i,j,k} - \frac{1}{N}\sum_{i=1}^{N}I_{i,j,k}\right)^2 \qquad (16)$$

where N is the number of B-scans, j and k are the lateral and depth indices of the B-scan images, respectively, and i denotes the B-scan slice index within N.

SvOCT angiography allows detection of small blood vessel networks because of its angle independence and fast image acquisition capability. It has also been applied to monitor protein denaturation and coagulation [30] and estimate tissue temperature during laser therapy [31-32].

2.2 Convolutional neural network (CNN)

2.2.1 Building blocks of CNN architecture [33-35]

2.2.1.1 Convolutional layer

The convolutional layer is the core building block of CNNs that performs feature extraction. The layer's parameters consist of a set of learnable filters (or kernels) having a small receptive field, but extend through the full depth of the input volume. Each neuron is only connected to a small local region of the input volume by using the kernel smaller than the input. The local connectivity of the convolutional layer allows to exploit the spatial local correlation of the input (for a natural image, a pixel is more correlated to the nearby pixels than to the distant pixels). They detect small, meaningful features such as edges with kernels that occupy only tens or hundreds of pixels. The small kernels also allow to store fewer parameters, which both reduces memory requirements of the model and improve its statistical efficiency, and to reduce computational load. In addition, each member of the kernel is slid across the width and height of the input and used at every spatial position of the input. The parameter sharing reduces the number of parameters for efficiency of expression, efficiency of learning, and good generalization [36].

The output of the convolutional layer is calculated by dot products between the input, I, and kernel, K, at every spatial position. The output is calculated by

$$O(i,j) = K * I(i,j) = \sum_{x=-a}^{a} \sum_{y=-b}^{b} \sum_{z=1}^{n} K(x,y,z)I(i-x,j-y,z) \qquad (17)$$

where kernel size is 2*a*+1 by 2*b*+1 and input and kernel depth are *n*. Figure 6 depicts how the convolution operation works with an input image of size 5 by 5 and a kernel of size 3 by 3. Both stride and zero padding are set to 1. The resultant output spatial size is expressed by

$$\frac{W-F+2P}{S} + 1 \qquad\qquad (18)$$
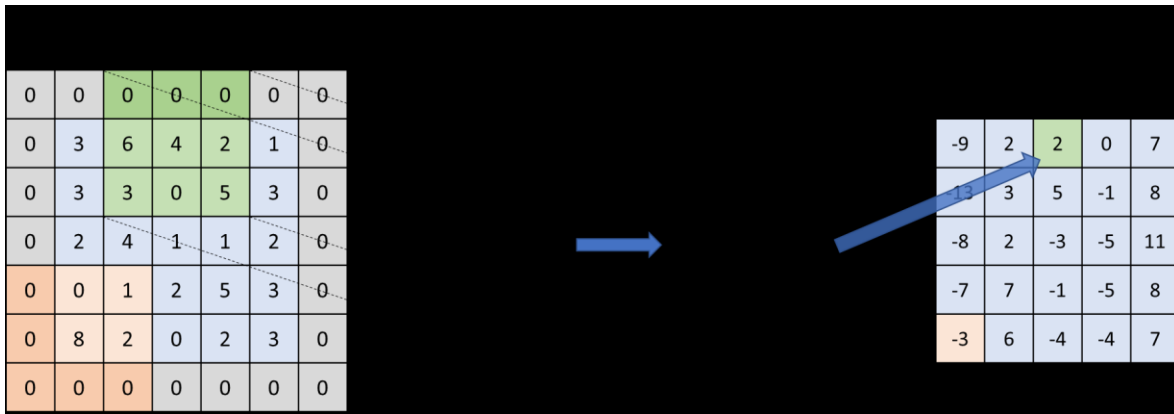
Figure 6 Image convolution with an input image of size 5 by 5 and a kernel of size 3 by 3. Both stride and zero padding are set to 1.

where *W* is the input size, *F* is the kernel size, *S* is the stride, and *P* is the amount of zero padding used on the border. Output depth is determined by the number of kernels.

2.2.1.2 Activation layer

An activation layer applies a function that decides whether a neuron should be activated or not. It helps neural networks learn complex patterns in the data by introducing nonlinearity into neural networks. The important characteristic of an activation function is that it should be differentiable to enable error backpropagation to train the model. There are many widely used activation functions including rectified linear unit (ReLU), Leaky ReLU, sigmoid function, and hyperbolic tangent function. The activation functions are depicted in Fig. 7.

17
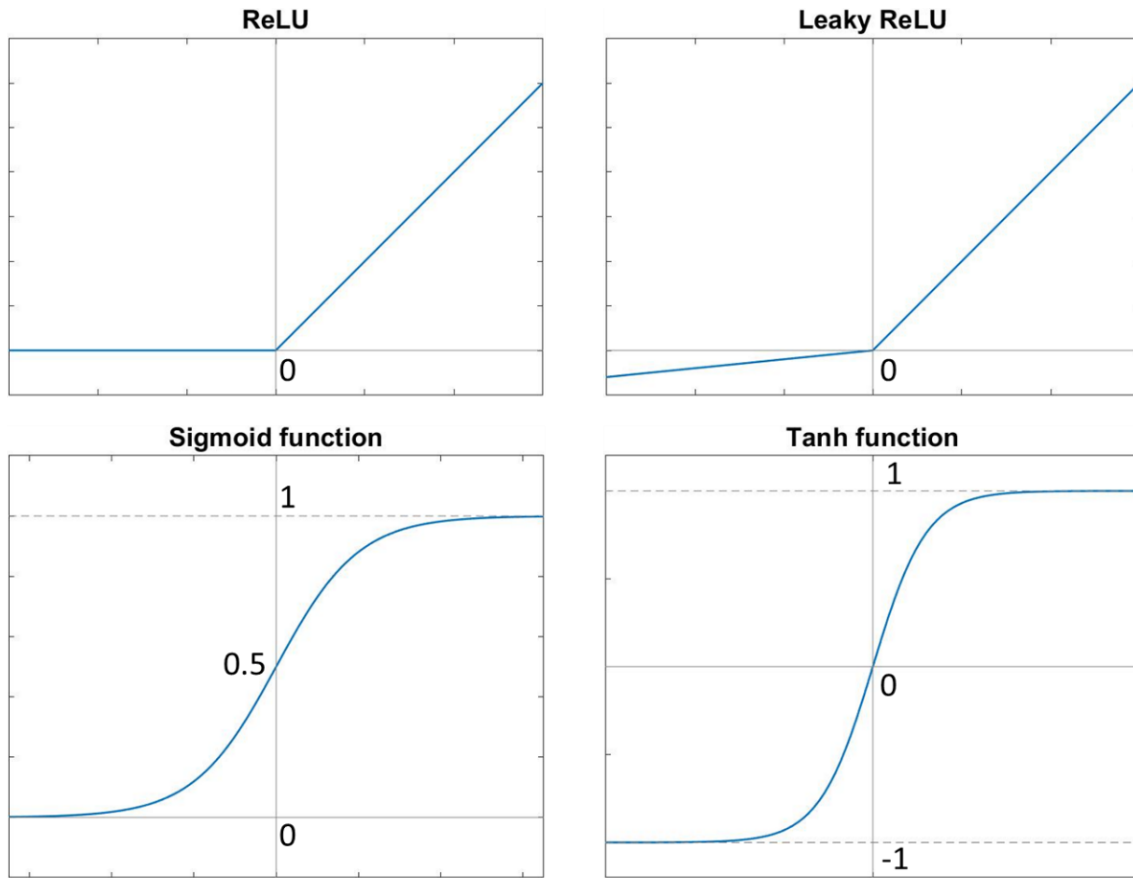
Figure 7 Widely used activation functions. Rectified linear unit (ReLU), leaky ReLU, sigmoid function, and hyperbolic tangent function

The ReLU is the most widely used activation function defined as

$$f(x) = \max(0, x). \tag{19}$$

It is a piecewise linear function that outputs the input directly if it is positive, and outputs zero otherwise. ReLU is computationally efficient and allows the network to converge very quickly. However, it has the dying ReLU problem – if too many activations get negative values then most of the neurons in the network with ReLU will simply output zero, in other words, die and thereby prohibiting learning. To solve the problem, leaky ReLU is applied, and it is defined as

$$f(x) = \begin{cases} x & if\ x > 0 \\ ax & if\ x < 0 \end{cases} \tag{20}$$

where *a* is a small positive constant. The leaky ReLU has an advantage of the nonzero gradient at all points (except 0 where the gradient is not defined).

The sigmoid function is a saturating function that transforms the input into a value between 0 and 1. It is especially used for models predicting the probability as an output because probability always exists only between the range of 0 and 1. It is defined as

$$f(x) = \frac{1}{1+e^{-x}} = \frac{e^x}{e^x+1}. \tag{21}$$

It has the useful property that its gradient is defined everywhere and smooth. However, it is computationally expensive and has vanishing gradient problem for very high or very low values of input.

The hyperbolic tangent function is another saturating function producing output a value between -1 and 1. Its shape is similar to that of the sigmoid function. It is defined as

$$f(x) = \frac{e^x-e^{-x}}{e^x+e^{-x}}. \tag{22}$$

The advantage of this function is that the negative inputs will be mapped to negative and the zero inputs will be mapped near zero in the tanh graph.

2.2.1.3 Pooling layer

A pooling layer is used to down-sample feature maps and reduce the number of parameters, memory footprint, and computation in the network, and thus also control overfitting. It is common to periodically insert a pooling layer between successive convolutional layers in a CNN architecture, and each one is typically followed by an activation function (i.e. ReLU layer). The pooling helps to make the representation approximately invariant to small translations of the input. Invariance to translation means that if we translate the input by a small amount, the values of most

of the pooled outputs do not change. Invariance to local translation can be a useful property if we care more about the feature's rough location relative to other features than its exact location.

There are two common pooling methods: max pooling and average pooling method. As the name suggests the max pooling calculates the maximum value for each rectangular subregion of the feature map. The average pooling involves calculating the average value for each subregion of the feature map. Figure 8 shows how the max and average pooling work on a 4 by 4 feature map with filter size of 2 by 2 and stride of 2.
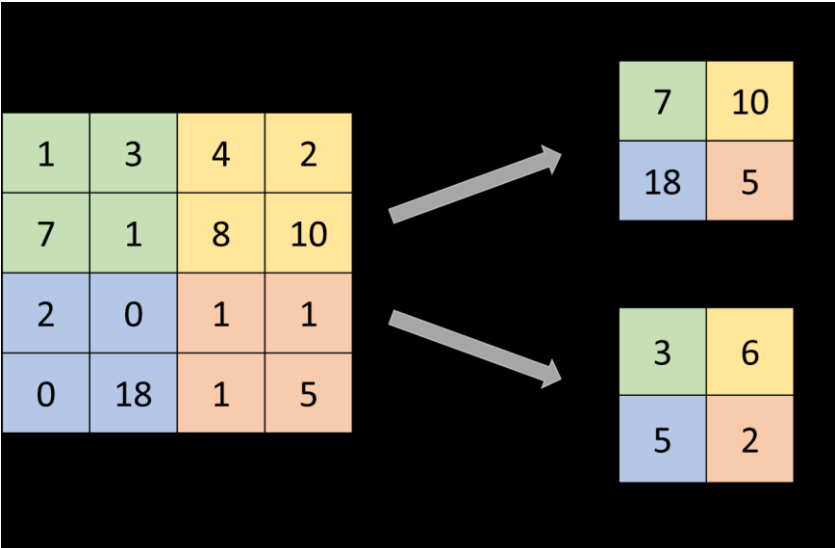
Figure 8 Max pooling and average pooling with a filter size of 2 by 2 and stride of 2

The pooling layer applied to input size of m by n yields output size of $(m - f) / s$ by $(n - f) / s$ where f is pool size and s is stride. In addition, it generally operates independently on every depth channel of the input, so the depth dimension remains unchanged.

2.2.1.4 Fully connected layer

Neurons in a fully connected layer have connections to all activations in the previous layer. When the previous layer is a final pooling or convolutional layer, not the fully connected layer, output from the previous layer is flattened and then fed into the fully connected layer (Fig. 9). If present,

they typically form the last few layers of the network to classify the data into various classes using features extracted from previous convolutional and max pooling layers.
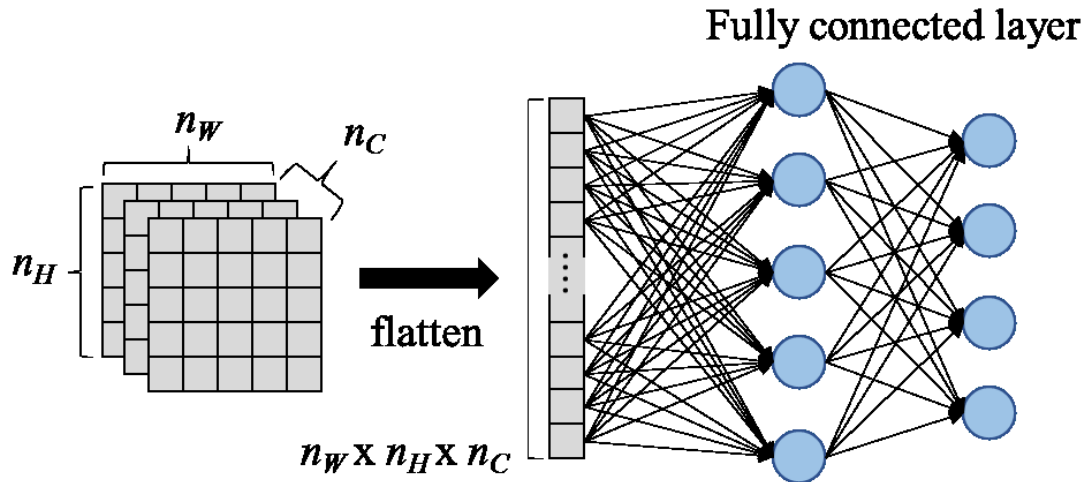
Figure 9 Illustration of fully connected layer

2.2.2 Training a neural network

Training a network is a process of finding kernels in convolution layers and weights in fully connected layers which minimize error (loss) between output predictions and given ground truth labels on a training dataset. The backpropagation algorithm is the method commonly used for training neural networks where loss function and gradient descent optimization algorithm play essential roles. A model performance under particular kernels and weights is calculated by a loss function through forward propagation on a training dataset, and learnable parameters, namely kernels and weights, are updated according to the loss value through an optimization algorithm called backpropagation and gradient descent.

2.2.2.1 Backpropagation

Backpropagation is an algorithm used for training feed-forward neural networks. Given an artificial neural network and a loss function (cost function), the method calculates the gradient of the loss function with respect to the neural network's weights. Computing an analytical expression for the gradient is straightforward, but numerically evaluating such an expression can be computationally expensive. The backpropagation algorithm does so by computing the gradient by the chain rule. Partial computations of the gradient from one layer are reused in the computation of the gradient for the previous layer. This backward flow of the cost information provides efficient computation of the gradient at each layer versus the naive approach of calculating the gradient of each layer separately. This efficiency allows using of gradient methods such as gradient descent or its variant stochastic gradient descent for training multilayer networks, updating weights to minimize cost.

The chain rule of calculus states that

$$\frac{dz}{dx} = \frac{dz}{dy}\frac{dy}{dx} \tag{23}$$

where f and g both are functions mapping from a real number to a real number such as $y=g(x)$ and $z=f(y)=f(g(x))$. It can be generalized to vector notation for $\boldsymbol{x} \in \mathbb{R}^m$, $\boldsymbol{y} \in \mathbb{R}^n$, g maps from $\mathbb{R}^m$ to $\mathbb{R}^n$, and f maps form $\mathbb{R}^n$ to $\mathbb{R}$ such as

$$\nabla_{\boldsymbol{x}}z = \left(\frac{d\boldsymbol{y}}{d\boldsymbol{x}}\right)^T \nabla_{\boldsymbol{y}}z \tag{24}$$

where $\frac{d\boldsymbol{y}}{d\boldsymbol{x}}$ is the n by m Jacobian matrix of function g. From this we see that the gradient of a variable $\boldsymbol{x}$ can be calculated by multiplying a Jacobian matrix $\frac{d\boldsymbol{y}}{d\boldsymbol{x}}$ and a gradient $\nabla_{\boldsymbol{y}}z$. The back-propagation algorithm consists of performing such a Jacobian-gradient product for each layer in the neural network.

Backpropagation for a simple two layers neural network is illustrated as an example (Fig. 10). The forward propagation is calculated by

$$z^{[i]} = W^{[i]} a^{[i-1]} \tag{25}$$

$$a^{[i]} = a(z^{[i]}) \tag{26}$$

where input is a vector $x = [x_1 \; x_2]^T = a^{[0]}$ and output is a scalar $\hat{y} = a^{[2]}$. $W^{[i]}$ is a matrix of which elements are weights of ith layer and $a(.)$ is an element-wise activation function. Then, the gradient of loss function L with respect to network weights $W^{[i]}$ is calculated by

$$dW^{[2]} = \frac{\partial L}{\partial W^{[2]}} = \frac{\partial L}{\partial a^{[2]}} \frac{\partial a^{[2]}}{\partial z^{[2]}} \frac{\partial z^{[2]}}{\partial W^{[2]}} \tag{27}$$

$$dW^{[1]} = \frac{\partial L}{\partial W^{[1]}} = \frac{\partial L}{\partial a^{[2]}} \frac{\partial a^{[2]}}{\partial z^{[2]}} \frac{\partial z^{[2]}}{\partial a^{[1]}} \frac{\partial a^{[1]}}{\partial z^{[1]}} \frac{\partial z^{[1]}}{\partial W^{[1]}} \tag{28}$$

and used to update the weights by optimization algorithms described in the next section.
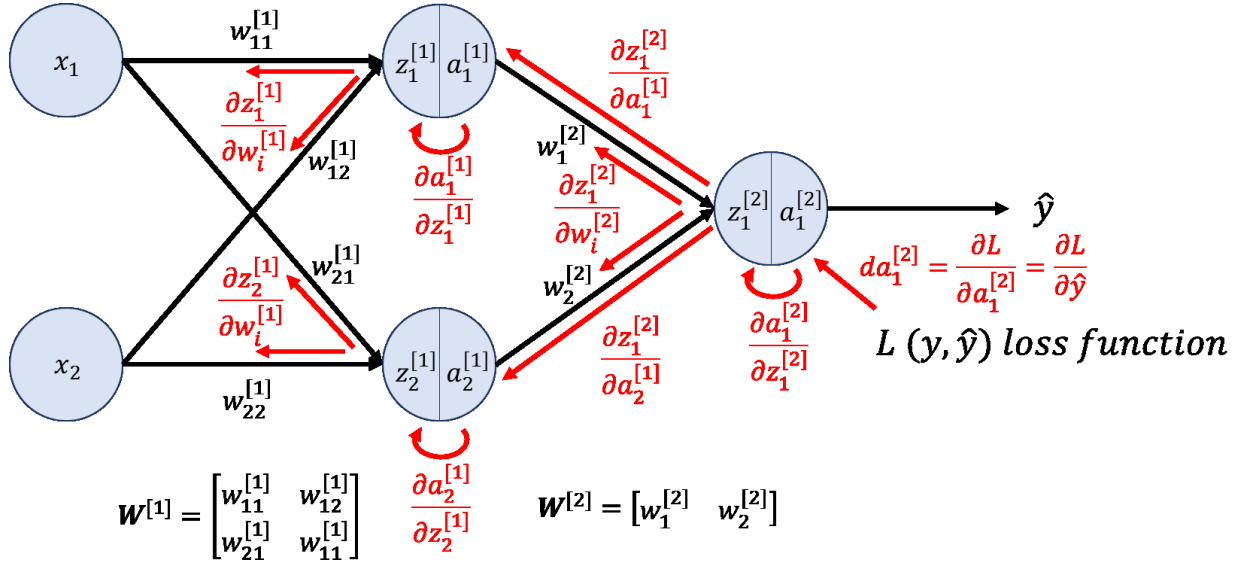
Figure 10 Illustration of backpropagation in a simple two-layer neural network

2.2.2.2 Neural network optimization algorithms

An optimization algorithm is needed to find weights minimizing the loss function. Optimization algorithms commonly used for CNN training is described in the following.

Gradient descent is a first-order optimization gradient algorithm which is dependent on the first-order derivative of a cost function. The idea of gradient descent is to take repeated steps in the opposite direction of the gradient of the function at the current point, because this is the direction of the steepest descent. The weights are updated by

$$w_{t+1} = w_t - \eta \frac{\partial C}{\partial w_t} \tag{29}$$

$$\frac{\partial C}{\partial w_t} = \nabla_w C(w_t) \tag{30}$$

where η is the learning rate controlling the rate at which the algorithm updates the parameter estimates. $C(.)$ is the cost function and w_t are the weights at step t. Batch gradient descent updates the network weights after scanning the whole training dataset. Convergence takes much time because the whole dataset needs to be reevaluated at every step (i.e., epoch). It converges smoothly to the global minimum for convex loss functions, but it could converge to a local minimum for non-convex functions. Stochastic gradient descent was introduced in order to overcome the shortcomings of batch gradient descent. It updates the network weights for each training data and, thus, is generally noisier due to the high variance between different data. However, it is computationally much less expensive, even though it requires a higher number of iterations to reach the minima than batch gradient descent, and in most cases, especially with large data set, stochastic gradient descent is preferred for optimizing a learning algorithm. Mini-batch gradient descent was introduced to overcome the shortcomings of the previous two algorithms, because it allows for the weights to be updated per batch, and not per data. It makes a compromise between the fast convergence and the noise associated with gradient update, so it could be a more flexible and robust algorithm.

The mini-batch gradient descent still has an issue of oscillations during the updating of the weights. Momentum, also known as moving average gradients, helps to avoid oscillations in the

wrong directions. Momentum gradient descent adds a fraction of the previous update, which gives the optimizer the momentum needed to continue moving in the right direction. The weights are updated by

$$v_t = \lambda v_{t-1} + \eta \frac{\partial C}{\partial w_t} \tag{31}$$

$$w_{t+1} = w_t - v_t \tag{32}$$

where v is the velocity and initialized to 0. λ is used to select the amount of information needed from the previous update. η is the learning rate, w_t are the weights at step t, and $C(.)$ is the cost function.

There are also several algorithms using adaptive learning rates. AdaGrad establishes different updates for different weights. The learning rate is tuned automatically, by dividing the learning rate by the sum of squares of all previous gradients [37]. It gives a high learning rate for the least frequent gradients and a low learning rate for the more frequent gradients. Each weight is updated by

$$w_t^i = w_{t-1}^i - \frac{\eta}{\sqrt{\sum_{\tau=1}^t \left(\nabla_w C(w_\tau^i)\right)^2 + \varepsilon}} \nabla_w C(w_t^i) \tag{33}$$

where η is the learning rate, w_t is the weights at step t, $C(.)$ is the cost function, and $\nabla_w C(w_t)$ is the gradient of weight parameters wt. It is well-suited for dealing with sparse data, but it has a problem that the learning rate decreases monotonically.

RMSProp was introduced to address the problem of the monotonically decreasing learning rate. It divides the learning rate by an average of squared gradients, and each weight is updated by

$$G_{i,t} = \nabla_w C(w_t^i) \tag{34}$$

$$E[G_i^2]_t = \lambda E[G_i^2]_{t-1} + (1 - \lambda)G_{i,t}^2 \tag{35}$$

$$w_t^i = w_{t-1}^i - \frac{\eta}{\sqrt{E[G_i^2]_t + \varepsilon}} G_{i,t} \tag{36}$$

where η is the learning rate, w_t are the weights at step t, $C(.)$ is the cost function, and $\nabla_w C(w_t)$ is the gradient of weight parameters w_t. λ is used to select the amount of information needed from the previous update. $E[G^2]$ is the running average of the squared gradients.

Adaptive Moment Estimation (Adam) [38] is another method that computes adaptive learning rates for each parameter. In addition to storing an average of past squared gradients v_t like RMSprop, Adam also keeps an average of past gradients m_t, similar to momentum. The weights are updated by

$$w_t^i = w_{t-1}^i - \frac{\eta}{\sqrt{\widehat{v_t}}+\varepsilon} \widehat{m_t} \tag{37}$$

$$\widehat{m_t} = \frac{m_t}{1-\beta_1^t} \tag{38}$$

$$\widehat{v_t} = \frac{v_t}{1-\beta_2^t} \tag{39}$$

$$m_t = \beta_1 m_{t-1} + (1-\beta_1)G \tag{40}$$

$$v_t = \beta_2 v_{t-1} + (1-\beta_2)G^2 \tag{41}$$

where η is the learning rate hyperparameter, w_t are the weights at step t, $C(.)$ is the cost function, and $\nabla_w C(w_t)$ is the gradient of weight parameters w_t. β_i is used to select the amount of information needed from the previous update, where $\beta_i \in [0, 1]$, m_t is the running average of the gradients, also known as the first moment, v_t is the running average of the squared gradients, and known as the second moment.

2.2.2.3 Loss function

A loss function is used to evaluate how well the model performs over the training dataset and train neural network. There are various tasks neural networks can perform such as regression, classification and image segmentation, and each task requires a different type of loss function since

the output format is different. A type of loss function is one of the hyperparameters that needs to be determined according to the given tasks.

Mean squared error (MSE), also known as L2 loss, is the most commonly used regression loss function. MSE is the average of squared distances between target variable, y, and predicted values, $\hat{y}$, and it is expressed by

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2. \tag{42}$$

Mean Absolute Error (MAE) is another regression loss function. MAE is the average of absolute differences between target and predicted values, and it is expressed by

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|. \tag{43}$$

MAE is more robust to outliers, but its constant gradient, which means it has a relatively large gradient even for a small loss, make it harder to converge.

For classification, cross-entropy loss, or log loss, is commonly used. Cross-entropy builds upon the idea of entropy from information theory and is designed to quantify the difference between two probability distributions. The multi-class cross-entropy loss function is expressed by

$$Cross\ entropy\ loss = -\sum_{c=1}^{C} y_c \log \hat{y}_c \tag{44}$$

where C is the number of classes, y_c is the binary indicator (0 or 1) if class c is the correct classification, and $\hat{y}_c$ is the predicted probability value between 0 and 1.

The popular loss function for image segmentation is a pixel-wise cross entropy loss and dice loss function. A pixel-wise cross loss examines each pixel individually, comparing the class predictions (depth-wise pixel vector) to its one-hot encoded target vector using Eq. (44). Dice loss uses the Dice coefficient which is a measures of overlap between two segmented images. The Dice

coefficient ranges from 0 to 1 where the score of 1 denotes perfect and complete overlap. Dice loss is expressed by

$$Dice\ loss = 1 - 2\frac{\sum_{c=1}^{C}\sum_{m=1}^{M} y_{cm}\hat{y}_{cm}+\varepsilon}{\sum_{c=1}^{C}\sum_{m=1}^{M} y_{cm}+\sum_{c=1}^{C}\sum_{m=1}^{M}\hat{y}_{cm}+\epsilon} \qquad (45)$$

where C is the number of classes, M is the number of pixels and ϵ is used to avoid dividing by 0.

3

High index epoxy lensed fiber OCT probe for retinal imaging

3.1 Introduction

CPOCT-based compact all-fiber probes allow their insertion into small and tight areas of the body (i.e. eye ball), enabling cathedral-based endoscopic imaging. They can be integrated into small-scale surgical tools such as a needle and a forceps and work for guidance of ophthalmic surgery and microsurgery [4, 5, 39-43]. An example includes hand-held automatic CPOCT-distal sensor-guided ophthalmic surgical devices that can compensate physiological tremors during surgery [4-5, 40]. One of important advantage of CPOCT-based probes is that they can be easily disposed after having been inserted into body due to no need for optical path matching. CP-OCT system is also tolerable to system vibrations and free from dispersion and polarization mismatches.

However, bare fiber probes, commonly used as a sensing probe for CP-OCT, have two main drawbacks: small depth of focus caused by strong divergence of the sample beam; and dependence of OCT sensitivity on the refractive index of the working environment. To overcome the strong divergence of bare fiber probes, various lensed probe systems have been proposed and demonstrated. These include fused coreless fiber [44,45], ball lens [46-49], GRIN lens [50], and conical-frustum tip [51]. Whereas most work well for standard OCT configurations, they do not work well for common-path setups, especially in vitreous media (i.e., water). This is mainly caused by the inability to readily fabricate the reference surface with an optimized reflectivity. Those designed for CP-OCTs [45,47,50-51] used lens/air interfaces as the reference, and their

performance could degrade significantly for retinal imaging through vitreous media because the lens effectiveness and/or the reference power level are significantly reduced.

In order to resolve the problem, we designed and fabricated elliptical high-refractive-index epoxy-lensed fiber probes that exhibit long working distances and large depths of focus (DOF) in vitreous media (i.e., water) compared to standard bare or half-ball lensed-fiber probes. The refractive index difference between the high refractive index epoxy and fiber core keeps the reference beam power independent of the working environment. Additionally, the availability of a wide range of epoxies with different indices allows fine-tuning of the reference reflectivity by mixing them in proportion. To validate our approach, we fabricated multiple probes with a range of beam expansion lengths and lens curvatures. We then measured the SNR of the lensed OCT probes as a function of the target depth and compared them with theoretically calculated SNR. The endoscopic OCT probes were then tested for B-mode OCT imaging using a phantom target comprising several layers of tape to confirm the improved imaging performance of the proposed lensed-fiber probes.

3.2 Design of lensed fiber probe

A key performance metrics of OCT-distal sensors is beam size and effective sensing range, which is defined as the region of DOF outside the probe. The working distance, DOF, and beam waist depends on the beam expansion length and the curvature of lenses, and they can be calculated by the ABCD matrix method [52]. For the lensed-fiber, ABCD can be calculated as

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \frac{n_e - n_h}{n_e R} & \frac{n_h}{n_e} \end{bmatrix} \begin{bmatrix} 1 & L \\ 0 & 1 \end{bmatrix} \tag{46}$$

where L is the length of expansion region, R is surface radius of the lens, n_e and n_h are the refractive indices of the sample environment (i.e., in air or in water) and high-index epoxy, respectively. x is

the propagation distance from the lens surface. For the case of bare fiber probes, the second and third matrices are identity matrices, and x is the distance from the fiber end. Working distance can be calculated by finding x, satisfying $AC+a^2BD=0$, where a is $\lambda/n_g\pi\omega_0^2$ for lensed-fiber and $\lambda/\pi\omega_0^2$ for bare fiber. Beam waist diameter ($2\omega_f$) at the focal length is expressed by [53]

$$2\omega_f = 2\omega_0 \left[\left(\frac{n_g}{n_e}\right)\frac{A^2+a^2B^2}{AD-BC}\right]^{\frac{1}{2}}, \tag{47}$$

where $2\omega_0$ is the mode-field diameter of the fiber, and λ is the wavelength in vacuum. The DOF can be expressed as

$$DOF = \frac{2\pi n_e \omega_f^2}{\lambda}. \tag{48}$$

Figure 11 shows the calculated working distance, beam waist diameter, DOF, and effective sensing range as a function of beam expansion length and lens curvature. The working distance is a distance from lens surface to the beam waist where a plus/minus sign indicates which side of the lens surface the beam waist places. It is positive when the beam waist places outside the probe, and, conversely, it is negative when the beam waist is virtual and occurs inside the fiber probe. To obtain effective sensing range longer than 500 μm, 1,000 μm, 1,500 μm, or 2,000 μm, beam expansion length must be longer than 157 μm, 228 μm, 282 μm, and 327 μm, respectively. Lens curvature must be around 1/5 of the expansion length. Because the fiber has a 125-μm cladding, it is easier to make a lens of 60-μm curvature. Therefore, we set the lens curvature to 60 μm and targeted expansion region length around 300 μm, because it provides the longest effective sensing range (1.6 mm) for the curvature. The beam waist diameter of the targeted lensed-fiber was expected to be 31 μm in water. The red stars indicate the geometry of fabricated lenses.

31
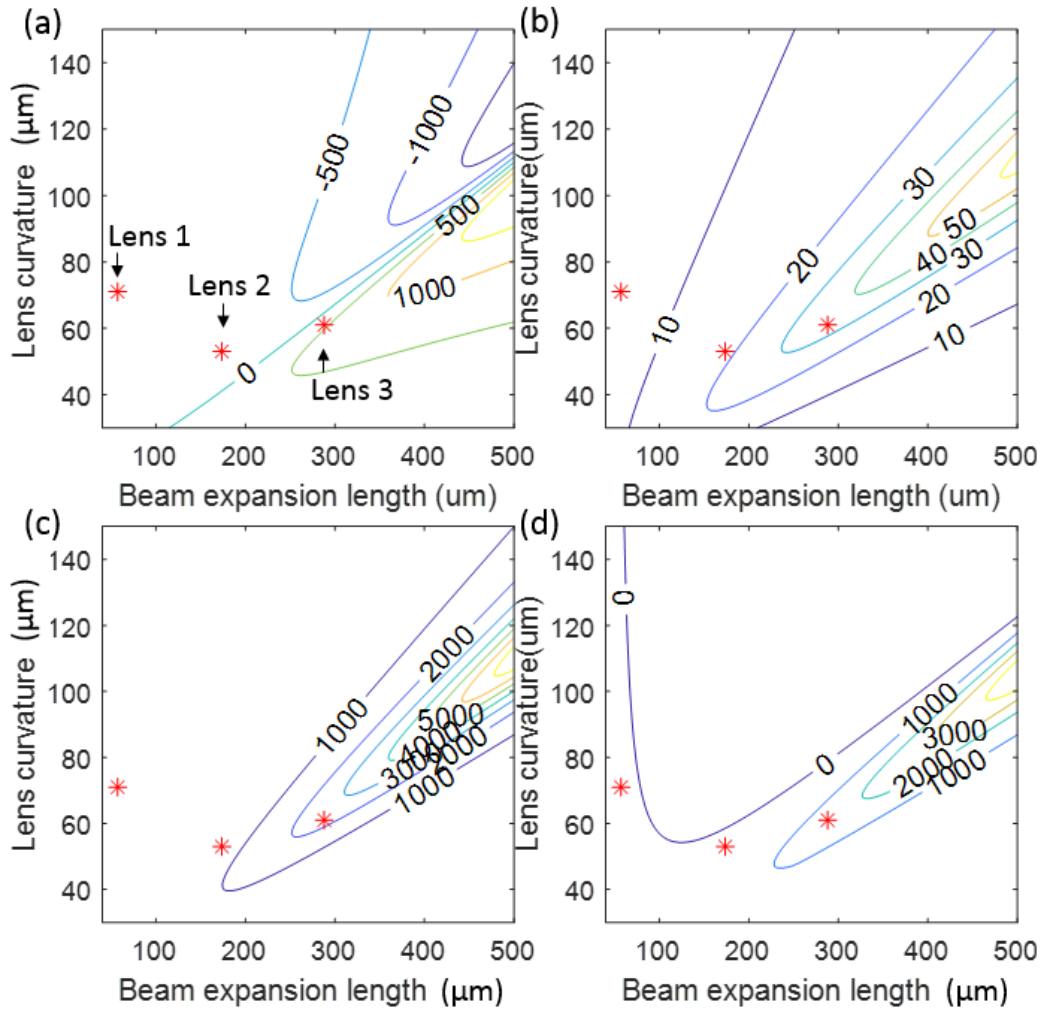
Figure 11 Calculated (a) working distance, (b) beam waist, (c) DOF and (d) effective sensing range as a function of beam expansion length and lens curvature. Red stars indicate geometry of lenses we made.

3.3 Fabrication of lensed fiber and lens geometry

3.3.1 Fabrication of lensed-fiber

The lensed-fiber OCT sensor was fabricated using UV curable epoxy (Norland Optical Adhesive, n=1.7). First, a semi-spherical lensed probe was fabricated by applying the epoxy to the fiber end, which naturally forms a semi-spherical shaped lens, owing to surface tension. Two fibers having semi-spherical epoxy ends were placed inline and brought together so that the epoxy lenses were in contact. We controlled axial and transverse positions of the fibers using *x-y-z* linear stages while

monitoring them using digital microscopes and our CP-SSOCT system to ensure they were perfectly aligned. A beam expansion rod was fabricated by curing the combined epoxy using ultra-violet light and breaking connection with one of the fibers. The length of the expansion rod was controlled by changing the distance between fiber ends. The fabricated expansion rod length was limited to around 120 μm because of the limited amount of epoxy that could be applied to bare fibers. A beam expansion rod of approximately 240 μm was fabricated by applying more epoxy on the 120-μm expansion rod via repeating the process. The end of the expansion rod was capped by applying a semi-spherical epoxy lens. This process can be highly reproducible when the direction and position of the fiber is controlled precisely. Another factor in the reproducibility is the amount of the epoxy picked up by the fiber. This is also related to the precise control of the fiber. An epoxy lens can be quickly distorted when even a small amount of epoxy touches the sidewall of the fiber or the expansion rod. However, once epoxy forms a spherical shape at fiber end and expansion rod end, the shape of the lens is highly repeatable and durable. Figure 12 schematically shows the fabrication steps from bare fiber to the designed elliptical-lensed-fiber. For comparison, we also constructed lensed fibers having different expansion region lengths and lens curvatures. Lens 1, Lens 2, and Lens 3 were fabricated by applying and curing spherical epoxy on bare fiber, 120 μm beam expansion rod, and 240 μm beam expansion rod, respectively.
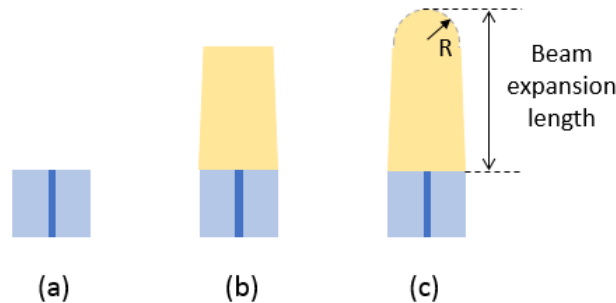
Figure 12 Illustration of fabrication process: (a) preparing bare fiber, (b) constructing expansion rod and (c) applying spherical surface on the expansion rod.

3.3.2 Lens geometry

Figure 13 shows *en-face* OCT images of the fabricated lensed-fiber probe with Lens 1, Lens 2, Lens 3, and a 240-μm expansion rod. The curvature of the spherical surface was measured using OCT images with customized Matlab code, which uses least-squares to fit a sphere on a manually selected edge. The length of beam expansion region was measured by our OCT system. Beam expansion lengths and curvatures of fabricated lenses are listed in Table 2. Lens 3 is the one closest to our design. Only Lens 3 has positive working distance because it has a sufficient beam expansion area that allows the focusing of the beam. Lens 1 and Lens 2 simply reduce beam divergence.
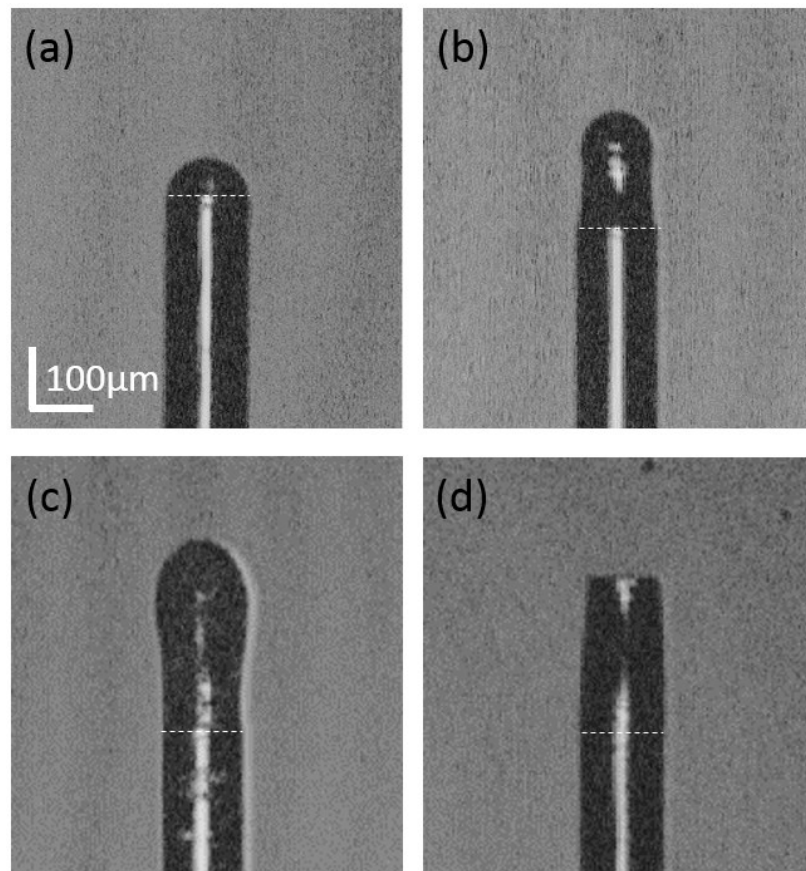
Figure 13 *En-face* **OCT images of fiber probe with (a) Lens 1, (b) Lens 2, (c) Lens 3 and (d) 240 μm expansion rod. The dashed line shows interface between bare fiber and epoxy.**

Table 2 Fabricated lens geometry and theoretically calculated working distance, beam waist size and effective sensing range

	Lens length (μm)	Surface radius (μm)	Working distance (μm)	Beam waist size (μm)	Effective sensing range (μm)
Lens 1	57.2	71	-43	8	14
Lens 2	173.6	53	-240	18	80
Lens 3	288	61	476	33	1600

We measured and verified the output-beam characteristics of the fabricated probes by comparing the beam spot diameter measured from knife-edge method [54] to simulated beam diameter via the ABCD method. Figure 14(a) and (b) show the erf function result using the knife-edge measurement and resultant beam shapes of bare fiber and each lensed-fiber at 1-mm distance. We performed knife-edge measurements of each lensed-fiber at various distances. Figure 14(c) shows the measured and calculated beam diameter in air. The mode-field diameter was 3.2 μm, and the lens geometry listed in Table 1 was used for calculation. Calculated and measured beam diameters match reasonably well throughout the distance ranges, confirming the measured geometry of our lens and calculation. Thus, whereas it is difficult to measure beam diameter in water, we can use the calculated values shown in Fig. 14(d).
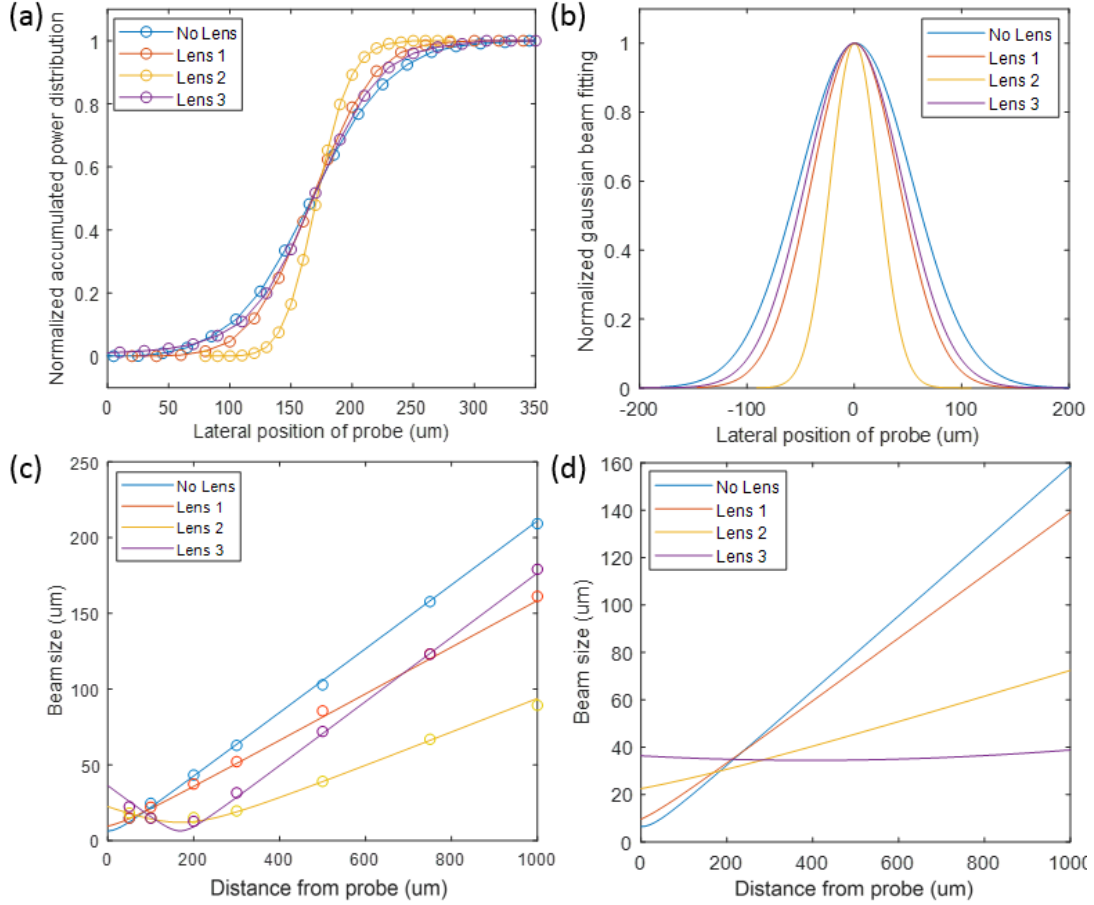
Figure 14 (a) The erf function result using knife-edge measurement at 1mm away from probe. (b) Beam shape from fitting the erf functions in (a). Measured (circle) and calculated (solid line) beam width of each lensed fiber in the (c) air and (d) water depending on distance from probe.

3.4 Performance of the lensed fiber probe

3.4.1 Signal-to-noise ratio

The SNR versus depth was measured to show improved performance of the lensed-fiber probes.

Theoretical SNR was also calculated using the following equation to confirm the measured SNR.

$$\frac{S}{N} = \frac{2\frac{1}{L_{Proc}}\left(\frac{\eta q}{h\upsilon}\right)^2 P_r P_s}{\left[\left(\frac{\eta q}{h\upsilon}\right)^2 NEP^2 + 2\frac{\eta q^2}{h\upsilon}(P_r + P_s) + \left(\frac{\eta q}{h\upsilon}\right)^2 RIN(P_r + P_s)^2\right]\frac{2B}{N}}. \qquad (49)$$

For the calculation, the detector quantum efficiency (η) was 0.9, noise equivalent power was 5pW/√Hz at 1-mm depth, relative intensity noise was -130 dB/Hz, the bandwidth of the optical

receiver (B) was 68.8 MHz, and the number of samples (N) was 1,376. The digital processing loss

was set to 1.8 dB for a Hann window [55]. Sample power at each depth was estimated by returned

sample beam diameter at the fiber end because the peak intensity of a Gaussian beam is inversely

proportional to the square of the beam diameter. Coupling loss was considered when P_r and P_s

were calculated. For reflected sample beam diameters, ABCD can be calculated by

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} 1 & L \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -\frac{1}{n_h-n_e} & \frac{n_e}{n_h} \\ \frac{n_hR}{n_hR} \end{bmatrix} \begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{n_e-n_h} & 0 \\ \frac{n_eR}{n_eR} & \frac{n_h}{n_e} \end{bmatrix} \begin{bmatrix} 1 & L \\ 0 & 1 \end{bmatrix}. \tag{50}$$

A mirror was used as a sample, and the power at the fiber end was set to 0.15 mW to avoid

saturation. Noise was limited by detector noise because of the low power. Detector noise has

frequency dependence [55]. Thus, we estimated NEP from the measurement of detector noise (i.e.,

A-scan image without an input signal) assuming NEP at 1 mm is 5 pW/√Hz.

Figure 15(a) and (b) respectively show the SNR of each lensed-fiber versus its depth in

both air and water. Bare fiber showed relatively high SNR in air, despite its large divergence

caused by higher reference power. However, it degraded significantly in water. Lens 3 showed the

highest SNR in water. The 10-dB range of the bare fiber was 200 μm, and it increased to 1,600

μm with Lens 3. The measured and calculated SNR fits reasonably well throughout the distance

range, except for Lens 3 in water. The difference is probably caused by the large mode

mismatching caused by the tight beam focus.
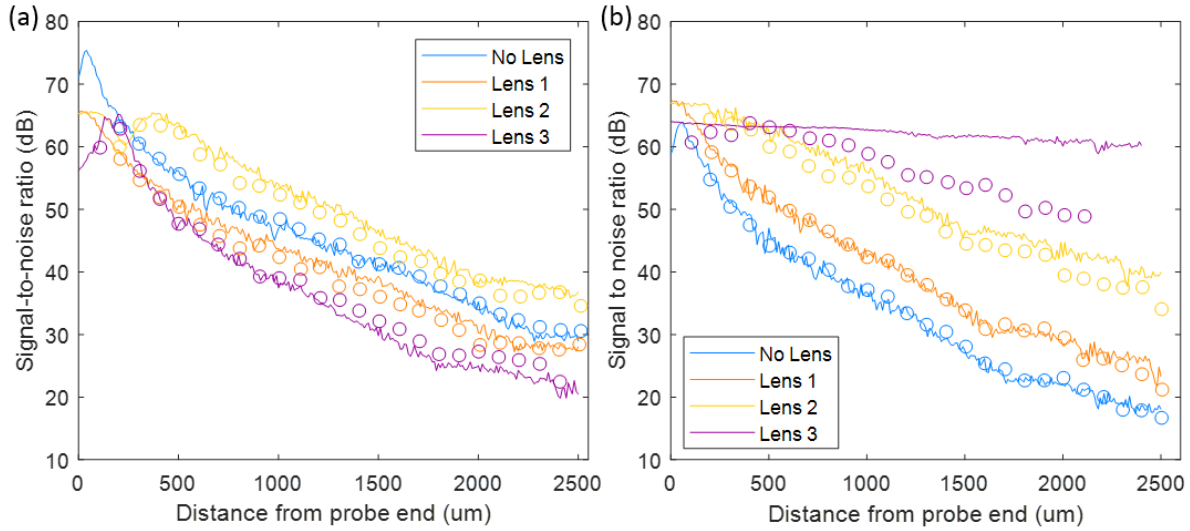
Figure 15 Theoretically calculated (solid line) and measured (circle) signal-to-noise ratio of each lensed fiber in the (a) air and (b) water

3.4.2 OCT imaging performance

To confirm the improved SNR and resolution, B-scan images of a phantom target comprising several layers of tape were obtained at various distances using the lensed fibers. The fiber probe was attached to a motorized translation stage (Thorlabs MT1-Z8), and the probe was scanned horizontally across the phantom. Figure 16 and 17 show the OCT images working in water and air, respectively. The OCT images showed similar trends as with the previous beam diameters and SNR measurements. In water, the designed lensed-fiber showed the best image quality at all distances. All probes showed similar image quality at 300-μm distance except bare fiber, which has a lower SNR because of low reflection at the fiber end. However, the designed lensed-fiber showed better quality than other lensed fibers when the probe was further away from the sample because of its long working distance and DOF. The quality improvement from the designed lensed-fiber can be clearly seen at the 1500 μm distance. In the air, fiber probes without a lens and with Lens 2 showed better quality, as expected. A fiber probe with Lens 3, optimized for working in

38

wet environments, showed the worst image quality because of strong divergence and smaller reference than bare fiber.
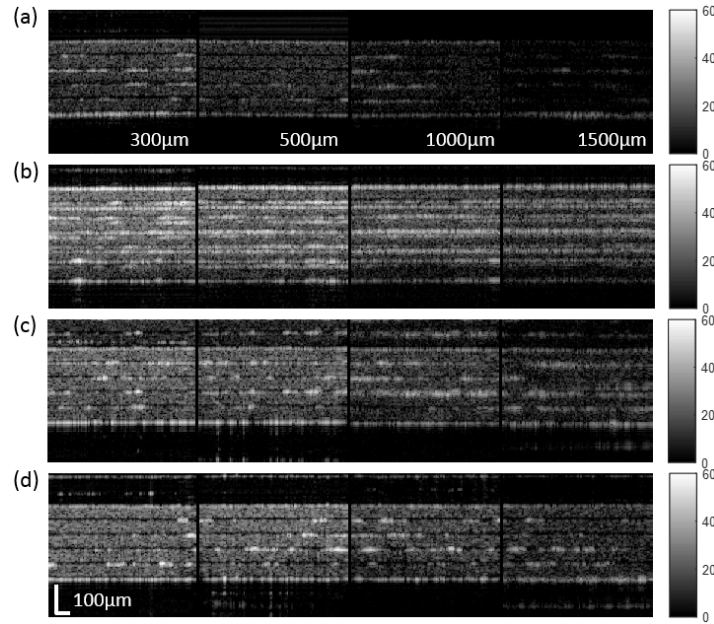
Figure 16 B-mode OCT images of a phantom target made up of several layers of tapes obtained by fiber probe (a) without lens (bare fiber), (b) with Lens 1, (c) with Lens 2, and (d) with Lens 3 in water
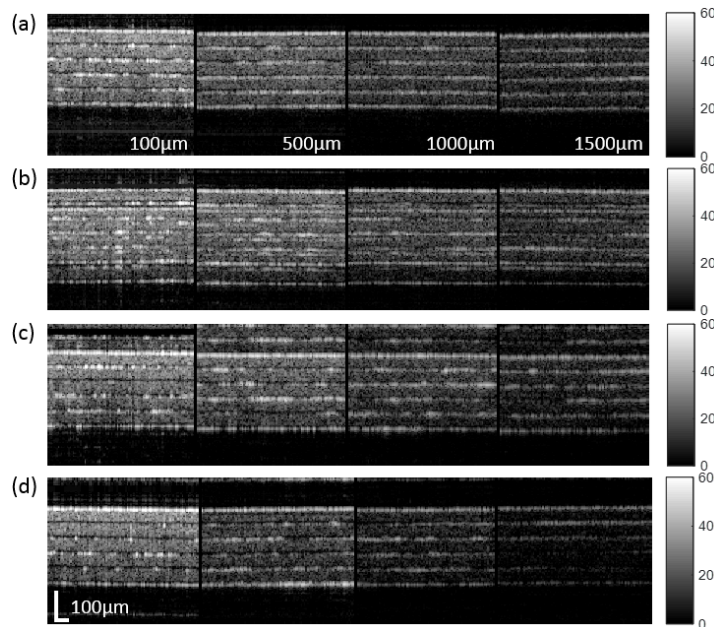
Figure 17 B-mode OCT images of a phantom target made up of several layers of tapes obtained by fiber probe (a) without lens (bare fiber), (b) with Lens 1, (c) with Lens 2, and (d) with Lens 3 in air

3.5 Conclusion

We have demonstrated that CP-SSOCT imaging and sensing in water can be enhanced by using a high refractive index elliptical epoxy-lensed fiber probe. We designed lensed fibers having long working distances and DOFs in water to obtain sufficiently long effective distal sensing ranges. Enhanced performance was confirmed by SNR and B-scan OCT images. The SNR of the proposed fiber probe, compared to a bare fiber at a distance of 1.5 mm, was increased by 25 dB, and the resolution improved from 476 μm to 47 μm. We can conclude that the improved performance in water indicates that such high refractive index epoxy-lensed fibers can work effectively in ophthalmic and vascular applications.

4

A CNN-based CP-OCT sensor integrated with a subretinal injector for retinal boundary tracking and injection guidance

4.1 Introduction

Subretinal injection is becoming increasingly prevalent in both scientific research and clinical communities as an efficient way of treating retinal diseases. It has been used for gene and cell transplant therapies to treat many degenerative vitreoretinal diseases, such as retinitis pigmentosa, age-related macular degeneration, and Leber's congenital amaurosis [56]. The treatments involve the delivery of drugs or stem cells into subretinal space between the RPE and photoreceptor layer, thereby directly affecting resident cells and tissues in the subretinal space. However, the procedure requires surgeons' high dexterity and micro-scale precision due to the delicate anatomy of the retina. The procedure is further complicated by the existence of physiological motions by patients, surgeons' hand tremor [1,2] and limited depth perception, and limited visual feedback from a traditional stereo-microscopic *en-face* view.

OCT-guided robotic systems have been developed to reduce the unintended physiological motion and overcome the limited visual feedback during ocular microsurgery. OCT, which provides micro-scale resolution cross-sectional images in real-time [3], enables improved visualization and accurate guidance of robotic systems. Microscope-integrated OCT systems were applied for surgical tool localization and robotic system guidance by intraoperatively providing volumetric images of tissues and surgical tools [57-61]. Fiber-optic CP-OCT distal sensor

integrated hand-held surgical devices have also been developed to implement simple, compact, and cost-effective microsurgical systems [4-7]. In those systems, a single fiber distal sensor attached to a surgical tooltip (i.e., needle or micro-forceps) guided the hand-held surgical device by real-time A-scan-based surface tracking. However, surface tracking-based guidance could induce inaccurate depth targeting for subretinal injection because of retinal thickness variations and irregular morphological features caused by retinal diseases. The target or near target retinal boundary tracking, which is RPE and photoreceptor boundary tracking for subretinal injection, allows precision guidance, but previous researches on retinal layer segmentation of OCT images using active contours [62,63], graph search [64-66], and shortest path methods [67,68] are not adequate for A-scan images due to the absence of lateral information. In recent years, convolutional neural network (CNN)-based retinal layer segmentation have been proposed and showed promising results [8-11]. Although the proposed CNN-based methods were developed for B-scan or C-scan OCT image segmentation, they could also be applied to A-scan images and operate in real-time by simplifying networks and using GPU parallel computing.

In order to resolve the problem, we present real-time retinal boundary tracking based on CNN segmentation of A-scan OCT images for accurate depth targeting of a selected retinal boundary. The U-net [69], which is widely used in medical image segmentation, was simplified and applied for segmentation on A-scan images. A Kalman filter, combining retinal boundary position measurement by CNN and velocity measurement by cross-correlation between subsequent A-scan images, is applied to estimate the retinal boundary position optimally. Undesired axial motions of the surgical tool are compensated by a piezo-electric linear motor using the tracked boundary position. An ex vivo bovine eye model is used to evaluate the retinal boundary tracking and depth targeting performance of the hand-held microsurgical device.

4.2 Experiment and Method

4.2.1 Network architecture and training for retinal layer segmentation

We applied a simplified 1-D U-net for A-scan retinal OCT image segmentation. The U-net is a fully convolutional neural network consisting of a contracting path to capture context followed by a symmetric expanding path that enables precise localization. In our design, double convolutional layers of the original U-net were reduced to a single convolutional layer, and the identical number of feature channels was used for all convolutional layers.

Figure 18(a) shows the 1-D U-net architecture we designed. The contracting path composed of four contracting blocks containing a convolutional layer, batch normalization layer, ReLU activation layer, and max-pooling layer in sequence. Similarly, the expanding path composed of four expanding blocks containing a transposed convolution layer, concatenation layer, convolutional layer, batch normalization layer, and ReLU activation layer in sequence. The convolutional kernel size of 15 by 1 was used to ensure the receptive field to be larger than the image size. The receptive field is expressed as [70]

$$r = s^b\big(1 + 2(k - 1)\big) - k \tag{51}$$

where s is the sampling size, which equals the kernel size of max-pooling layer and the transposed convolutional layer, b is the number of contracting blocks, and k is the convolutional kernel size. The kernel size of the max-pooling layer and the transposed convolutional layer was set to 2 by 1, and, in this case, the receptive field is calculated as 450 by 1. Since improving inference speed is important for our application, the 1-D U-net illustrated in Fig. 18(a) was simplified stepwise, and the performance of four architectures was compared. The number of contracting and expanding blocks was reduced to three while keeping other conditions the same, and also, max-pooing and transposed convolutional kernels were sized up to 4 by 1 for compensating reduced receptive field.

We then removed skip concatenation layers to see the effect of the skip connections, and the simplest 1-D U-net is illustrated in Fig. 18(b).

The 1-D U-net models were implemented using Pytorch on a computer with Intel i9-10900X CPU, NVIDIA Quadro RTX 4000 GPU, and 32GB RAM for training. A generalized dice loss function was used, and the network parameters were updated via back-propagation and Adam optimization process. Max epoch was 20, and the mini-batch size was 128. The learning rate was initialized as 10^{-3}, which then decreases by 10 times after 10 epochs.

The trained CNN model was implemented on CUDA by customized CUDA function, and the inference time of the CNN models on GPU was measured using the NVIDIA Nsight tool in Visual Studio on the workstation described in the section 4.2.4.
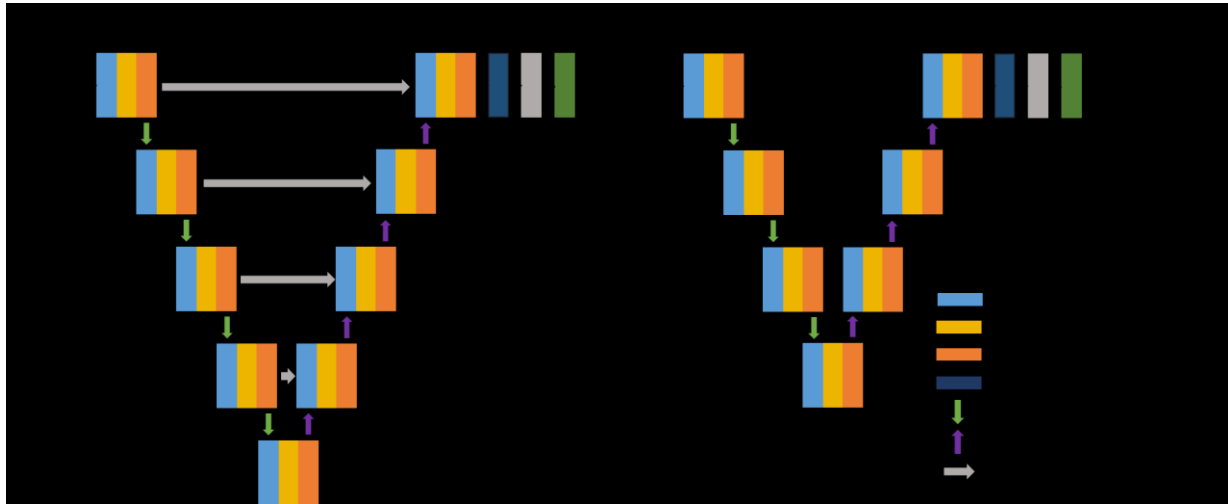
Figure 18 Network architectures of (a) our 1-D U-net and (b) the most simplified 1-D U-net we applied.
N: **kernel number,** *S*: **kernel size.**

4.2.2 Retinal boundary tracking

The axial distance between a fiber (needle) end and a target boundary can be measured from the target boundary position at A-scan images since the fiber end, working as a reference reflector, locates at the top edge of the images. A target boundary position was measured from a segmented image by averaging the bottommost pixel position of an adjacent upper layer and the topmost pixel position of an adjacent lower layer. Then, the Kalman filter [71] was applied to optimally estimate the boundary position using the dynamic and measurement model described as

$$\mathbf{x}_k = \begin{bmatrix} x_k \\ v_k \end{bmatrix} = F\mathbf{x}_{k-1} + Bu_{k-1} + \mathbf{w}_k = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix} \mathbf{x}_{k-1} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u_{k-1} + \begin{bmatrix} \frac{1}{2}\Delta t^2 \\ \Delta t \end{bmatrix} a_k \qquad (52)$$

$$z_k = H\mathbf{x}_k + \mathbf{n}_k = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \mathbf{x}_k + \mathbf{n}_k \qquad (53)$$

where x_k, v_k and a_k are the axial position, velocity, and acceleration of the target boundary. The control of the linear motor, u_k, is a distance that the linear motor moves forward or backward. The velocity, v_k, was measured by the ratio of movement distance of the sample (i.e. target boundary) to a known constant time duration. The movement distance was calculated by displacement of the sample in two consecutive A-scan images, which is the shift value maximizing cross-correlation between two consecutive A-scan images, subtracted by the previous control u_{k-1}. The u_k was defined as c $(x_{target} - x_k)$ using proportional control, where $(x_{target} - x_k)$ is an error and c is a proportional gain. The bias for control was set to zero because the linear motor is supposed to be stationary when the boundary position is at the target position. The proportional gain, c, was determined experimentally. The $\mathbf{w}_k$ and $\mathbf{n}_k$ are the process noise and observation noise, respectively, and they were assumed to be zero-mean Gaussian white noise. The algorithm works in a two distinctive process and is given by,

Prediction :

$$\hat{\mathbf{x}}_{k|k-1} = F\mathbf{x}_{k-1|k-1} + Bu_{k-1} \qquad (54)$$

$$P_{k|k-1} = FP_{k-1|k-1}F^T + Q \qquad (55)$$

Correction :

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + K_k(z_k - H_k\hat{\mathbf{x}}_{k|k-1}) \qquad (56)$$

$$K_k = P_{k|k-1} + P_{k|k-1}H^T(HP_{k|k-1}H^T + R_k)^{-1} \qquad (57)$$

$$P_{k|k} = (I - K_kH)P_{k|k-1} \qquad (58)$$

where *P, Q*, and *R* are covariance of error, process noise, and observation noise, and *K* is Kalman gain.

The quantitative evaluation of the retinal layer tracking performance was based on three metrics: mean signed error (MSE), mean unsigned error (MUE), and absolute maximum error (AME) of each layer boundary position.

4.2.3 Data set

A-scan OCT images of the retina were obtained from 11 ex vivo bovine eyes using endoscopic CP-OCT lensed fiber probes [72]. The cornea and lens of the eyes were removed, and the lensed fiber probes were inserted into the vitreous humor and horizontally scanned by a motorized linear translation stage (Z812B, Thorlabs, USA). More details about the CP-OCT system are described in section 4.2.4. Eight A-scan images were averaged to improve the signal-to-noise ratio. The resultant A-scan images were combined to present a quasi B-scan image for easy visualization as shown in Fig. 19(a). The quasi B-scan images were then manually segmented into the vitreous humor (VH), the six retinal layers, labeled as ganglion cell layer (GCL), inner plexiform layer (IPL), inner nuclear layer (INL) - outer plexiform layer (OPL), outer nuclear layer

(ONL) - external limiting membrane (ELM), photoreceptor layers (PR), and choroid (CH), and region below the retina by a single observer using ImageJ software. Figure 19(b) shows the manually segmented image. 8,400 A-scan OCT retinal images from 9 eyes were used for training, and 1,000 A-scan OCT retinal images from 2 eyes were used for testing.

A-scan images of 1 by 1024 pixels were cropped into 1 by 320 pixels along the axial direction, keeping only the region around retinal tissues, to reduce computation time. The retinal tissue area was found by using cross-correlation between the averaged A-scan image over all data set and each A-scan image. All A-scan images in the data set were averaged, after being shifted such that the retinal layer surface lays on zero position, and then thresholded to remove background noise. The upper graph of Fig. 19(c) shows the averaged A-scan image and a sampled A-scan image from Fig. 19(a), and the lower graph shows cross-correlation between the two A-scans as a function of displacement. Since the retinal surface position of the averaged A-scan is set to zero, the displacement maximizing the cross-correlation indicates approximately the retinal surface location of each A-scan image. Figure 19(d) is the cropped image obtained from Fig. 19(a).

The cropped images for the train data set were augmented by random vertical translation. For each A-scan image, five additional training samples were created with random translation values between -15 to 15. The final train and test sets consist of 46,530 A-scan images and 1,000 A-scan images, respectively. The image pixel size along the axial direction is 2.7 um.
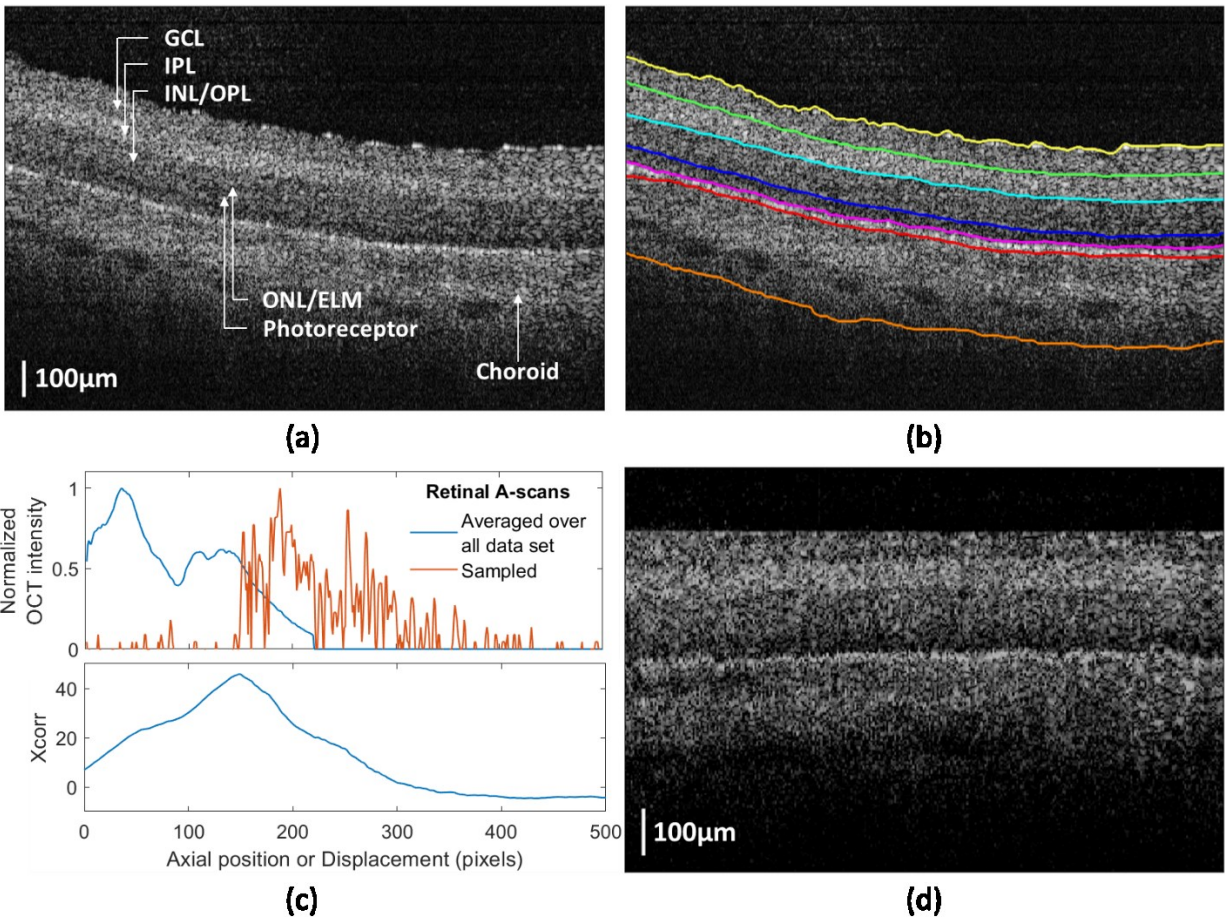
Figure 19 (a) A quasi B-scan OCT image of an *ex vivo* bovine eye obtained using an endoscopic CP-OCT lensed fiber probe. (b) A manually segmented OCT image. (c) The averaged retinal A-scan over all data set and a sampled retinal A-scan (upper graph) and cross-correlation between the two A-scans (lower graph). (d) A cropped quasi B-scan OCT image consisting of the cropped A-scan images in the train set.

4.2.4 CP-SSOCT distal sensor guided hand-held microsurgical tool system

Figure 20 shows the schematic of the CP-SSOCT distal sensor-guided hand-held microsurgical tool system and a signal processing flow chart. The CP-SSOCT system uses a commercial swept-source engine (Axsun Technologies Inc., Billerica, USA) operating at a 100 kHz sweep rate. The center wavelength and sweeping bandwidth of the system are 1060 nm and 100 nm, respectively. A lensed fiber probe of the CP-SSOCT system is encased in a 25-gauge blunt needle and fixed

along the needle using UV curable glue. The fiber probe guides the needle to maintain a specified distance from a target boundary using a piezo-electric linear motor (LEGS LT20, PiezoMotor, Uppsala, Sweden). More details about the microsurgical tool system are described in [5]. A workstation (Dell Precision T5810) with an NVIDIA Quadro K4200 GPU processes the sampled spectral data to measure a distance between a target boundary and a needle and controls the linear motor. Most parts of the signal processing including CNN inference are performed on GPU by CUDA to reduce processing time. Specifically, 128 spectra were transmitted from a frame grabber and processed at the same time. A-scan images were obtained by performing the fast Fourier transform on the spectral data. After background noise subtraction, eight sequential A-scan images were averaged to increase the signal-to-noise ratio and cropped into 16 by 320 pixels. CNN-based segmentation is performed on the 16 cropped images of 1 by 320 pixels, and a target boundary distance is measured as described in section 4.2.2. The Kalman filter is applied using the measured position and velocity, and the optimally estimated position was used for motor control.
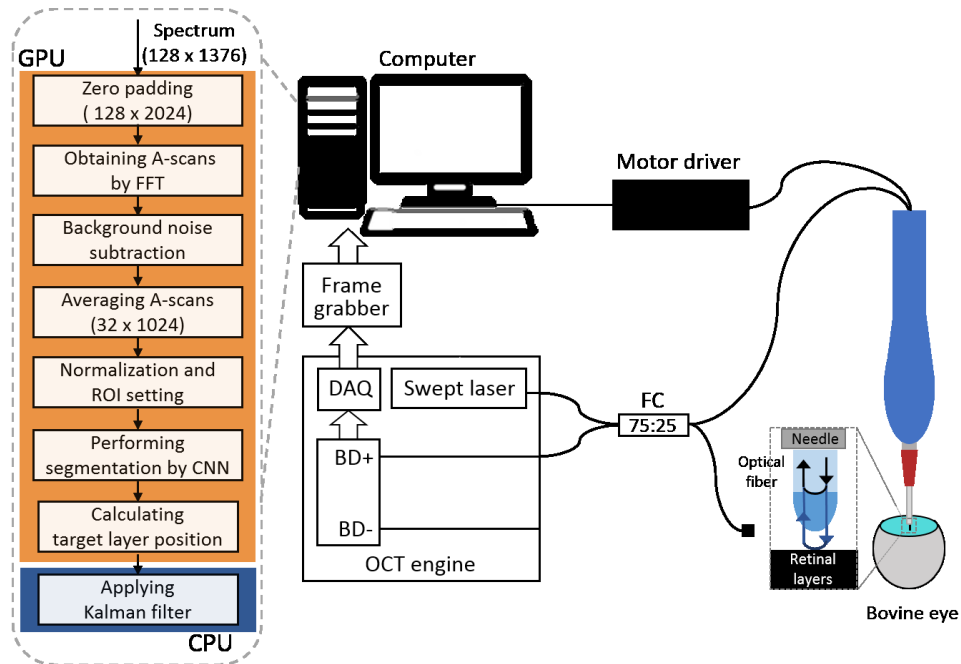
Figure 20 Schematic of CP-SSOCT distal sensor guided hand-held microsurgical tool system and a signal processing flow chart.

4.3 Experimental results

The CNN-based retinal layer segmentation performance was evaluated by mean intersection over union (IoU). The mean IoU is calculated by averaging the IoU score of each class as follows:

$$Mean\ IoU = \sum_{c=1}^{C} \frac{n_{c,TP}}{n_{c,TP}+n_{c,FP}+n_{c,FN}} \tag{59}$$

where $n_{c,TP}$, $n_{c,FP}$, and $n_{c,FN}$ are the number of true-positive pixels, false-positive pixels, and false-negative pixels of the class c, respectively, and C is the total number of classes.

Figure 21(a) shows the mean IoU on the train and test data set as a function of the number of feature channels calculated by networks described in section 4.2.1. Each CNN architecture was trained five times, and the plots indicate average values. As expected, mean IoU on the train set increases with learnable parameters, which increase with the number of contracting and expanding blocks, the number of feature channels, and sampling size, and mean IoU on the test set decreases or increases and then decreases with learnable parameters due to overfitting. Also, the removal of the skip concatenation connections does not degrade performance distinctively. This could be because our network is not very deep and high-resolution features passed from the contracting path to the expanding path do not advantageously affect the task due to the speckle noise of the images. We achieve the best mean IoU of 79.1 % on the test set with three contracting and expanding blocks and a sampling size of 4. The inference time of the trained networks on GPU was measured considering real-time axial tremor compensation. The most time-consuming layer is a convolutional layer, so inference time is significantly affected by the number of channels, sampling size, and skip concatenation connection as shown in Fig. 21(b). Inference time for 16 images of 1 by 320 pixels is at most 1.6 ms with an optimal number of features for each architecture. Physiological hand tremor has a frequency of 7 to 13 Hz, and its amplitude in the axial direction

is around 50 μm². The speed of physiological hand tremor is approximately calculated as 1 μm/ms assuming a frequency of 10Hz and linear movement. Therefore, inference time of 1.6 ms is considered reasonably fast for physiological tremor cancellation since other computation and communication delay of our system is around 1.5 ms and image pixel size along the axial direction, the smallest distance we can detect, is 2.7 μm.

Figure 21 Mean IoU of trained networks on the train and test data sets. (b) Inference time on GPU for segmentation of 16 A-scan OCT images of 320 by 1 pixel. NB: the number of contracting and expanding blocks, SS: sampling size, NSC: no skip concatenation connection.

Tables 2, 3, and 4 show the MSE, MUE, and AME of retinal boundary position calculated with an optimal number of feature channels before and after applying the Kalman filter. The MSE, MUE, and AME are defined as follows:

$$MSE = \frac{1}{N} \sum_{n=1}^{N} \widehat{p_\iota} - p_i \tag{60}$$

$$MUE = \frac{1}{N} \sum_{n=1}^{N} |\widehat{p_\iota} - p_i| \tag{61}$$

$$AME = \max_{n=1..N} |\widehat{p_\iota} - p_i| \tag{62}$$

where $\widehat{p_\iota}$ and p_i are the estimated and true retinal boundary position of the i-th A-scan images, and N is the total number of A-scan images in the test set. The Kalman filtering does not affect MSE distinctively, but it reduces MUE and AME by removing unexpected high-frequency motion of

tracked position. Overall, the errors are comparable for the networks presented in the tables, and we selected the last network architecture for our tremor cancellation system because it tracked boundaries more stably with lower MUEs and AMEs. Using the selected parameters for CNN, MSE, MUE, and AME of PR/CH boundary after Kalman filtering were -0.45 pixels (-1.2 µm), 2.48 pixels (6.7 µm), and 16 pixels (43.2 µm), respectively. Relatively larger errors than that of other CNN-based OCT retinal segmentation could be caused by the absence of lateral information and limited image quality obtained from a fiber probe.

Table 3 Mean signed error of retinal boundary position (pixels), NB : the number of contracting and expanding blocks, NC: the number of feature channels, SS: sampling size, NSC: no skip concatenation connections.

Retinal Boundary	Convolution Filter Size and Number, and Sampling Size							
	NB 4, NC 4, SS 2		NB 3, NC 8, SS 2		NB 3, NC 4, SS 4		NB 3, NC 12 SS 4, NSC	
	CNN	KF	CNN	KF	CNN	KF	CNN	KF
VH/GCL	-1.46	-1.54	**0.22**	**0.14**	-0.72	-0.79	0.67	0.60
GCL/IPL	1.34	1.27	**0.35**	**0.28**	0.96	0.88	1.54	1.48
IPL/INL-OPL	0.46	0.39	-5.37	-5.51	**-0.017**	**-0.09**	-0.068	-0.13
INL-OPL/ ONL-ELM	-2.49	-2.57	-1.36	-1.43	-1.45	-1.53	**-1.06**	**-1.12**
ONL-ELM/PR	-1.08	-1.16	**-0.92**	**-1.00**	**-0.92**	**-1.00**	-1.41	-1.49
PR/CH	**-0.23**	**-0.31**	-0.83	-0.90	-0.59	-0.69	-0.38	-0.45

Table 4 Mean unsigned error of retinal boundary position (pixels)

Retinal Boundary	Convolution Filter Size and Number, and Sampling Size							
	NB 4, NC 4, SS 2		NB 3, NC 8, SS 2		NB 3, NC 4, SS 4		NB 3, NC 12 SS 4, NSC	
	CNN	KF	CNN	KF	CNN	KF	CNN	KF
VH/GCL	3.81	3.01	2.86	**2.04**	3.11	2.50	**2.68**	2.15
GCL/IPL	**4.10**	**3.71**	5.09	4.61	4.24	3.75	4.21	3.79
IPL/INL-OPL	3.82	3.37	8.72	7.94	3.93	3.34	**3.61**	**3.14**
INL-OPL/ ONL-ELM	4.06	3.88	3.71	3.45	3.55	3.36	**3.29**	**3.01**
ONL-ELM/PR	2.79	2.43	2.97	2.51	**2.65**	**2.37**	2.93	2.56
PR/CH	2.99	2.56	3.44	2.84	3.20	2.77	**2.88**	**2.48**

Table 5 Absolute maximum error of retinal boundary position (pixels)

Retinal Boundary	Convolution Filter Size and Number, and Sampling Size							
	NB 4, NC 4, SS 2		NB 3, NC 8, SS 2		NB 3, NC 4, SS 4		NB 3, NC 12 SS 4, NSC	
	CNN	KF	CNN	KF	CNN	KF	CNN	KF
VH/GCL	27.5	19.9	61	21.2	39.5	22.1	**15**	**14.6**
GCL/IPL	21	**15.3**	45	31.3	21	17	**16**	17.2
IPL/INL-OPL	32.5	18.6	61.5	53	48	18.6	**22**	**15**
INL-OPL/ ONL-ELM	35.5	17.9	53	24.4	**18**	**15.1**	31.5	16.2
ONL-ELM/PR	31	17.1	41	16.8	25.5	15.4	**22**	**14.3**
PR/CH	27	18	48.5	22.9	88	40.6	**24.5**	**16**

4.3.1 Train and test results of CNN-based segmentation and boundary tracking

The CNN-based retinal layer segmentation performance was evaluated by mean intersection over union (IoU). The mean IoU is calculated by averaging the IoU score of each class as follows:

$$Mean\ IoU = \sum_{c=1}^{C} \frac{n_{c,TP}}{n_{c,TP}+n_{c,FP}+n_{c,FN}} \tag{59}$$

Figure 22(a) shows the mean IoU on the train and test data set as a function of the number of feature channels calculated by networks described in section 4.2.1. Each CNN architecture was trained five times, and the plots indicate average values. As expected, mean IoU on the train set increases with learnable parameters, which increase with the number of contracting and expanding blocks, the number of feature channels, and sampling size, and mean IoU on the test set decreases or increases and then decreases with learnable parameters due to overfitting. Also, the removal of the skip concatenation connections does not degrade performance distinctively. This could be because our network is not very deep and high-resolution features passed from the contracting path to the expanding path do not advantageously affect the task due to the speckle noise of the images. We achieve the best mean IoU of 79.1 % on the test set with three contracting and expanding

blocks and a sampling size of 4. The inference time of the trained networks on GPU was measured considering real-time axial tremor compensation. The most time-consuming layer is a convolutional layer, so inference time is significantly affected by the number of channels, sampling size, and skip concatenation connection as shown in Fig. 22(b). Inference time for 16 images of 1 by 320 pixels is at most 1.6 ms with an optimal number of features for each architecture. Physiological hand tremor has a frequency of 7 to 13 Hz, and its amplitude in the axial direction is around 50 μm^2. The speed of physiological hand tremor is approximately calculated as 1 $\mu m/ms$ assuming a frequency of 10Hz and linear movement. Therefore, inference time of 1.6 ms is considered reasonably fast for physiological tremor cancellation since other computation and communication delay of our system is around 1.5 ms and image pixel size along the axial direction, the smallest distance we can detect, is 2.7 μm.
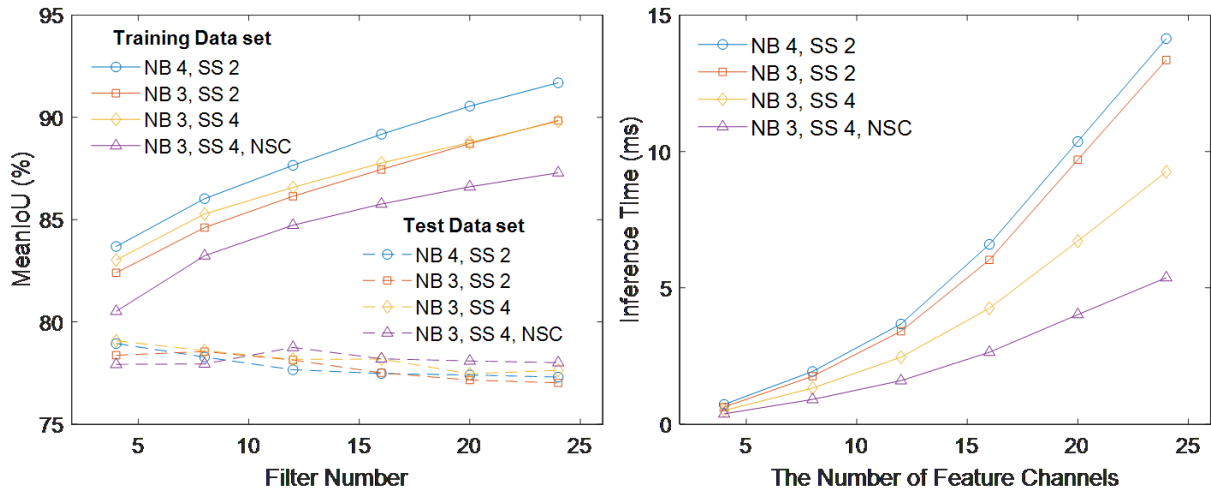
Figure 22 Mean IoU of trained networks on the train and test data sets. (b) Inference time on GPU for segmentation of 16 A-scan OCT images of 320 by 1 pixel. NB: the number of contracting and expanding blocks, SS: sampling size, NSC: no skip concatenation connection.

Tables 2, 3, and 4 show the MSE, MUE, and AME of retinal boundary position calculated with an optimal number of feature channels before and after applying the Kalman filter. The MSE, MUE, and AME are defined as follows:

$$MSE = \frac{1}{N}\sum_{n=1}^{N}\widehat{p_i} - p_i \tag{60}$$

$$MUE = \frac{1}{N}\sum_{n=1}^{N}|\widehat{p_i} - p_i| \tag{61}$$

$$AME = \max_{n=1..N}|\widehat{p_i} - p_i| \tag{62}$$

where $\widehat{p_i}$ and p_i are the estimated and true retinal boundary position of the i-th A-scan images, and N is the total number of A-scan images in the test set. The Kalman filtering does not affect MSE distinctively, but it reduces MUE and AME by removing unexpected high-frequency motion of tracked position. Overall, the errors are comparable for the networks presented in the tables, and we selected the last network architecture for our tremor cancellation system because it tracked boundaries more stably with lower MUEs and AMEs. Using the selected parameters for CNN, MSE, MUE, and AME of PR/CH boundary after Kalman filtering were -0.45 pixels (-1.2 μm), 2.48 pixels (6.7 μm), and 16 pixels (43.2 μm), respectively. Relatively larger errors than that of other CNN-based OCT retinal segmentation could be caused by the absence of lateral information and limited image quality obtained from a fiber probe.

4.3.2 Real-time ex vivo bovine retinal boundary tracking and tremor cancellation

We evaluated the retinal boundary tracking and depth targeting performance of the hand-held microsurgical instrument guided by CNN using an ex vivo bovine retina model.

At first, we produced an estimate of noise for retinal boundary tracking by measuring standard deviations (SDs) of VH/GCL and PR/CH boundary position using a stationary OCT distal

sensor. Fig. 23(a) shows the M-scan OCT images for 1s and tracked boundary position obtained using the stationary OCT distal sensor. Overall speckle pattern doesn't change as expected, but local intensity variations, which could be caused by OCT noise and micro-oscillations inside a sample, induce small fluctuations of tracked retinal boundary positions. Therefore, although the SDs of boundary positions are supposed to be zero because the distance between the retina and the OCT distal sensor doesn't vary, the SDs acquired from 13 trials of 5 eyes are 2.83 ± 0.69 um (1.04 ± 0.26 pixel) for VH/GCL boundary and 3.09 ± 0.92 um (1.14 ± 0.34 pixel) for PR/CH boundary.

Depth targeting system noise was then evaluated using a piezo-electric motor fixed to a stationary stage. The motor was integrated with an OCT sensor attached needle and activated for depth targeting of the needle. Ideally, the motor should be stabilized when the needle reaches a target depth since both the motor and the sample are stationary. However, due to retinal boundary tracking noise and control error, the motor kept working actively as shown in Fig. 23(b) and (c). Fig. 23(b) and (c) show M-scan OCT images for 1s when VH/CGL boundary and PR/CH boundary are targeted, respectively. The SDs of VH/GCL and PR/CH boundary position during depth targeting were measured with 13 trials from 5 eyes and shown in Fig. 23(d). The SDs of VH/GCL and PR/CH boundary position are 2.75 ± 0.35 um (1.02 ± 0.13 pixel) and 4.8 ±1.46 um (1.78 ±0.54 pixel), respectively, when the VH/CGL boundary is targeted. When the PR/CH boundary is targeted, the SDs of VH/GCL and PR/CH boundary positions are 4.41 ± 0.31 um (1.63 ± 0.12 pixel) and 4.28 ±1.02 μm (1.58 ±0.38 pixel), respectively. Theoretically, the speckle pattern doesn't change with axial motion only, so the overall speckle pattern doesn't change significantly except shifts in the axial direction. However, local intensity variations of the speckle pattern increase with axial motion because the OCT sensing beam is not perfectly perpendicular to the retina surface and axial motion could induce slight transverse motion. Moreover, since the sensing

beam is focused on the retina, the axial motion changes the integration volume inside the retina, which also could increase local intensity variations. Therefore, PR/CH boundary tracking, which has a larger tracking error, is degraded more by the intensity variations and shows larger SDs than that of VH/GCL boundary tracking.
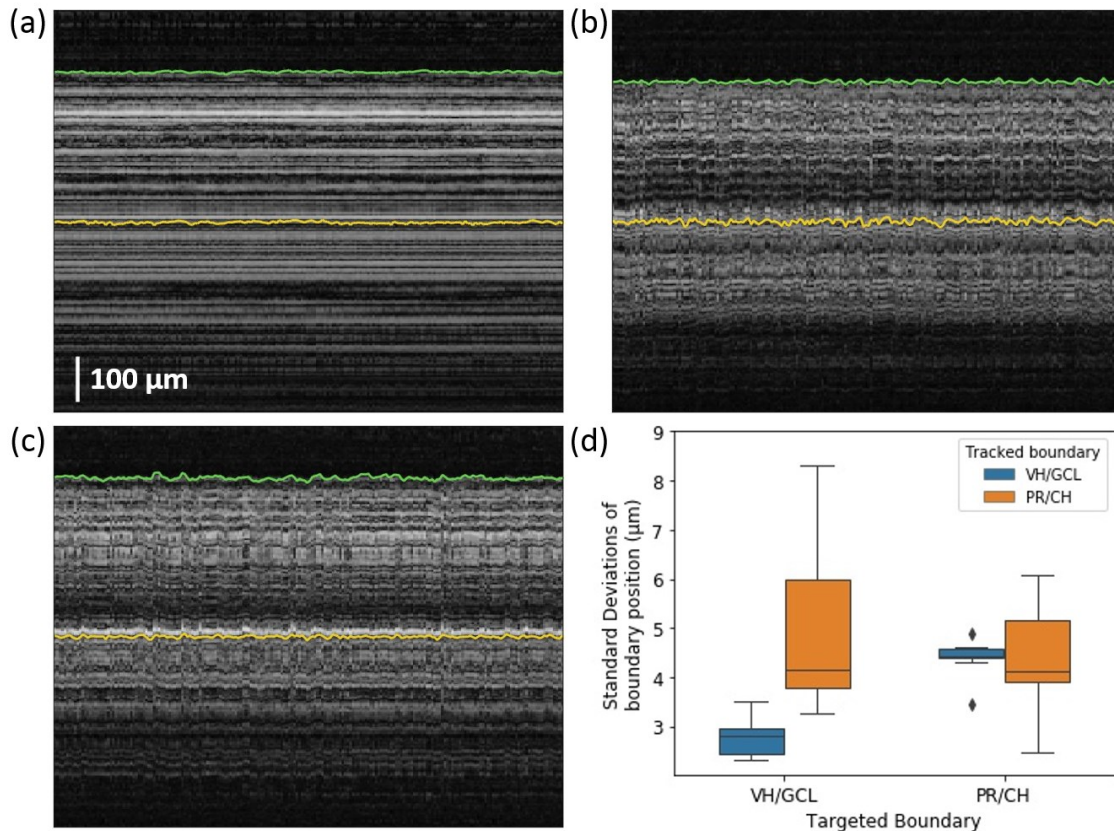
Figure 23 M-scan OCT images of ex vivo bovine eyes acquired using (a) a stationary OCT distal sensor and an OCT distal sensor attached to fixed motor activated for (b) VH/GCL boundary targeting and (c) PR/CH boundary targeting. The green and yellow solid lines represent tracked VH/GCL and PR/CH boundaries, respectively. (d) SDs of tracked boundary positions during depth targeting by an OCT distal sensor attached to fixed motor

Tremor compensation and depth targeting performance were evaluated for a hand-held microsurgical instrument. The microsurgical instrument was held by a free-hand and proceeded toward the retina until automatic depth targeting was activated. We used a tremor compensation algorithm we developed earlier and more details can be found in our previous work[13]. A VH/GCL boundary, as well as a PR/CH layer boundary, were tracked, and one of them was used for depth

targeting. We performed 12 trials of depth targeting each for VH/CGL and PR/CH boundaries using 5 eyes. Figure 24 shows the M-scan OCT images of the bovine retina obtained with and without tremor compensation for ~ 13 seconds. In Fig. 24(a), the VH/GCL boundary (yellow line) was used for depth targeting, and its target depth represented by dashed lines was set to 700 um away from fiber probe ends. Similarly, in Fig. 24(b), the PR/CH boundary (yellow line) was targeted, and its target depth was set to 1000 um. The green solid lines are untargeted boundaries (VH/GCL or PR/CH), and the white vertical lines indicate the moment when motion compensation has been activated. The left side of the vertical line with a highly irregular boundary profile represents duration without the tremor compensation, however, once the tremor compensation has been activated (right side of the vertical line), the targeted boundary becomes flat and fixed around the target depth indicating that the motion compensation is working effectively. As expected, when VH/GCL or PR/CH boundary is targeted, the axial variation of another boundary position increases, and it is quantitatively verified by comparing the MSEs and SDs of the tracked boundary positions for each trial. Here, the MSE is defined as follows:

$$MSE = \frac{1}{N}\sum_{n=1}^{N}\widehat{p_i} - p_{target} \tag{63}$$

where $\widehat{p_i}$ and p_{target} are the estimated and targeted retinal boundary position, respectively, and N is the total number of A-scan images of each trial. In Fig. 25(a), the MSEs of targeted boundaries are -0.15 ± 1.02 um for VH/CGL boundary and -0.11± 0.96 um for the PR/CH boundary, and the MSEs of untargeted boundaries are -319.52 ± 10.13 um for the VH/GCL boundary and 325.72 ± 11.35 um for PR/CH boundary. Untargeted boundaries have almost ten times larger variations of MSEs than that of targeted boundaries because of retinal thickness variations between different eyes and different areas, and this result supports the necessity of PR/CH boundary tracking rather than just surface tracking for accurate subretinal injection guidance. The SDs of targeted

boundaries are 9.42 ± 0.80 um for VH/GCL boundary and 10.8 ± 0.90 um for the PR/CH boundary. The axial motion mostly from the hand tremor, which includes low-frequency draft in the order of hundreds of micrometers and physiological tremor in the order of tens of micrometers, are reduced significantly. The residual variations are caused by a boundary tracking error and the time delay between the signal processing and motor control. The slightly better performance of the VH/GCL boundary targeting could be explained by the more accurate tracking of the VH/GCL boundary as shown in section 4.3.1. The SDs of untargeted boundaries are 13.03 ± 1.96 um for VH/GCL boundary and 13.67 ± 1.79 um for the PR/CH boundary. Retinal thickness variations within an eye increased SDs of the untargeted boundaries due to lateral motion of hand tremor.
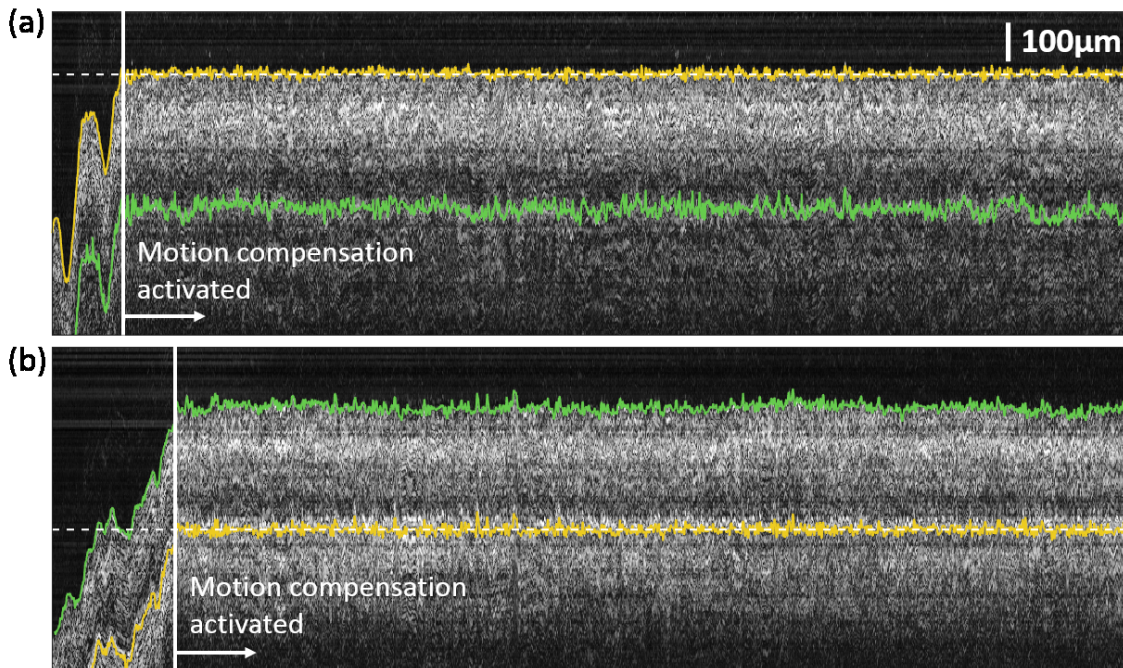
Figure 24 M-scan OCT images of ex vivo bovine eyes with and without tremor cancellation when (a) a boundary between VH and GCL is targeted and when (b) a boundary between PR and CH is targeted. The yellow and green solid lines are targeted boundary and untargeted another boundary, respectively. The dashed line represents target depth, and white vertical lines indicate the moment when motion compensation has been activated.
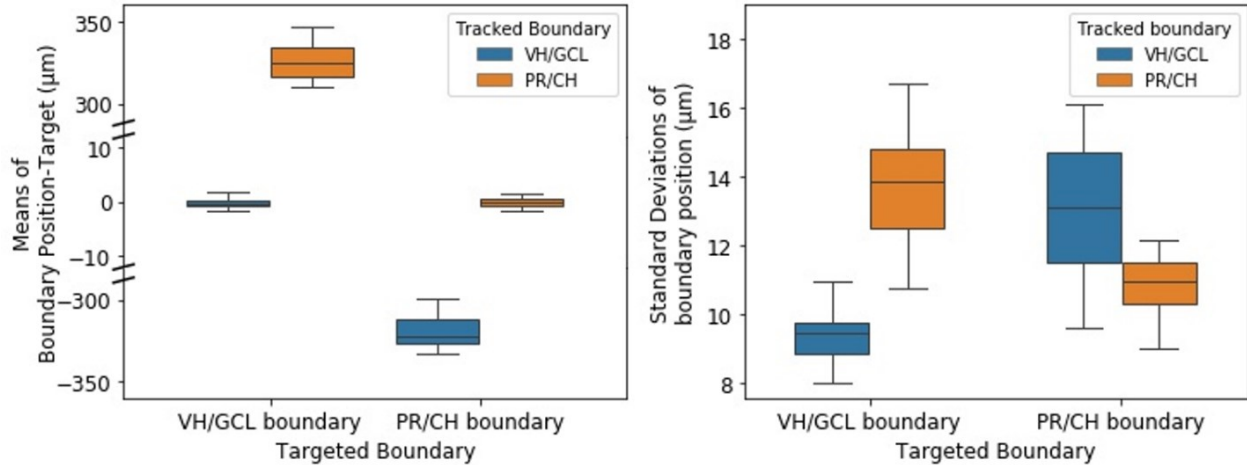
Figure 25 Box plots of (a) MSEs and (a) SDs of the VH/GCL and PR/CH boundary positions during VH/GCL boundary targeting and PR/CH boundary targeting.

It is difficult to obtain a precise ground-truth segmentation label from our M-scan OCT images (Fig. 24) because of high-frequency longitudinal fluctuations and speckle noise and, thus, to evaluate the accuracy of the tracked boundary positions quantitatively. Nevertheless, we can assess it visually by checking how flat and smooth the targeted retinal boundary is when each A-scan image is aligned to the tracked boundary position. The more accurate boundary tracking brings the flatter and smoother target boundaries in the aligned M-scan images. Figure 26 (a) and (b) show the aligned M-scan images to the targeted boundaries, the VH/GCL boundary, and the PR/CH boundary, represented by yellow dashed lines. High-frequency fluctuations shown in Fig. 24 were significantly reduced in the regions around the targeted boundaries, and we could infer that retinal boundary tracking works effectively.
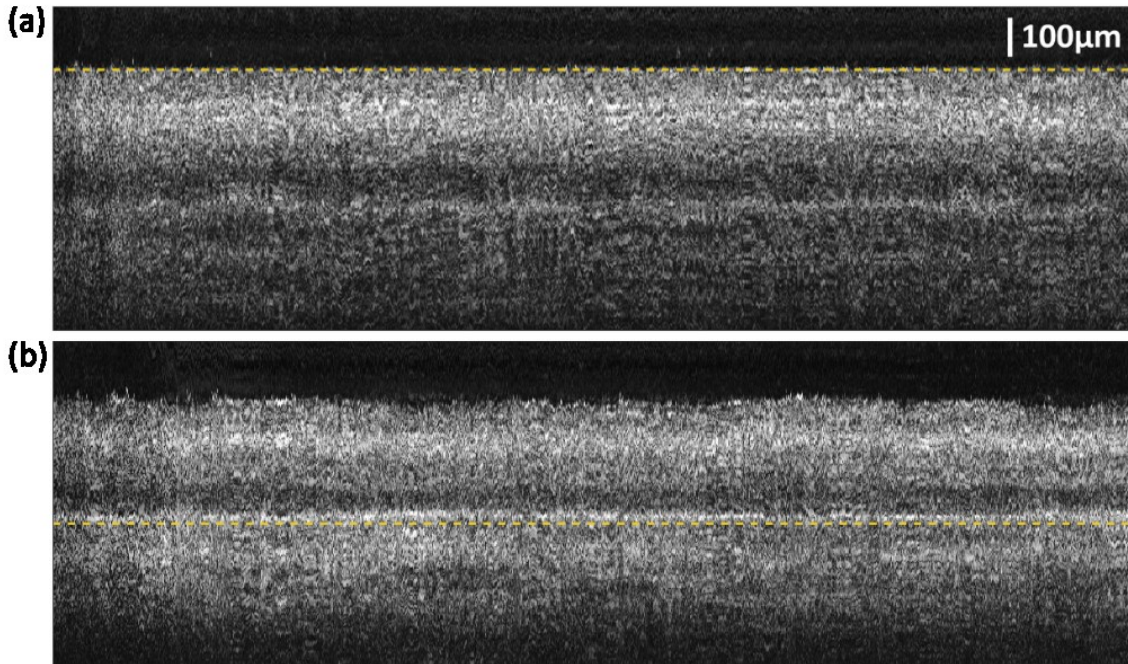
Figure 26 M-scan OCT images of ex vivo bovine eyes when each A-scan image is aligned to the targeted boundaries, (a) the VH/GCL boundary and (b) the PR/CH boundary.

4.4 Conclusion

In this work, we presented real-time A-scan-based CNN segmentation and automatic retinal boundary targeting for hand-held subretinal needle guidance. A-scan retinal OCT images are segmented using a simplified 1D U-net, and the Kalman filter reduces retinal boundary tracking error by combining boundary position measurement and velocity measurement. We achieve the MUE of around 3 pixels (8.1 μm) using an *ex vivo* bovine retina model. GPU parallel computing allows real-time inference (~1.6 ms) and, thus, real-time retinal boundary tracking. The MSE between target depth and target boundary position of the depth targeting experiment is -0.15 μm and 0.11 μm for the VH/GCL and the PR/CH boundary, respectively. Involuntary tremors, which include low-frequency draft in the order of hundreds of micrometers and physiological tremor in the order of tens of micrometers, are reduced significantly, and the SDs of target boundary

positions are 9.42 μm for the VH/GCL boundary and 10.8 μm for the PR/CH boundary. Our networks currently work only for normal bovine retina, but, in the future, we will expand its utility to diseased retina having irregular morphology by including diseased retinal images into our train data set. We also plan to perform *ex vivo* and *in vivo* studies of subretinal injection using our system to validate its clinical applicability.

5

Selective retina therapy monitoring by speckle variance OCT

5.1 Introduction

Selective retina therapy (SRT) is an effective laser treatment method for various retinal diseases associated with a degradation of the retinal pigment epithelium (RPE), such as diabetic macular edema, central serous chorioretinopathy, and age-related macular degeneration [12-17]. The RPE, which contains a high concentration of melanosomes, absorbs 50~60% of incident green light. However, in order to selectively target the RPE layer, the laser pulse duration needs to be shorter than a thermal relaxation time of the RPE (~10 μs) [18]. The SRT reduces negative side effects and facilitates healing of the induced retinal lesions by avoiding thermal damages of the adjacent photoreceptors, the neural retina, and the choroid. However, the selection of proper laser energy—which is crucial for successful SRT without excessive burning and collateral damage—is challenging because lesions in the RPE are invisible ophthalmoscopically. In addition, different melanin concentrations among patients or regions even within an eye [19] make it impossible to set a static threshold value of pulse energy of a therapeutic irradiation window.

Fundus fluorescence angiography (FFA) is an accurate method to detect the lesions, but it requires the use of fluorescent dye injection [20] and a long delay between treatment and detection. For real-time non-invasive SRT monitoring, several approaches have been proposed. These include the detection of microbubble formation and collapse, which induce mechanical disruption and damage to RPE cells [73]. This approach measures the acoustic transient [74] or light reflection

changes [75]. Although these methods have already been used in several clinical studies [74,76,77], they do not provide visual feedback during the treatment. Optical coherence tomography (OCT), which can provide depth-resolved imaging, was also applied for the SRT monitoring [78-82]. The treatments are considered successful when OCT signal variations, i.e., intensity decrease, are detected, and the results show good agreement with the evaluation of lesions by FFA.

Speckle variance OCT (svOCT) quantifies the speckle pattern variation caused by moving particles or structural changes in biological tissues. It calculates the interframe intensity variance of a sequence of structural OCT images. The svOCT has been developed extensively in recent years for OCT angiography, which is used to visualize retinal micro-vasculatures [83]; it has also been applied to monitor protein denaturation and coagulation [30] and to estimate tissue temperature during laser therapy [31].

Thus, it is expected that svOCT could be an effective way to detect speckle variation changes induced by morphological and structural changes of retinal tissue during the thermal-induced micro-bubble formation and collapse by laser irradiation. In this work, we studied and demonstrated SRT monitoring based on the svOCT. At first, a SS-OCT imaging system integrated with a micro-second pulsed laser system was tested for phantom (floppy disk film) and *ex vivo* bovine iris to examine its availability for monitoring of each laser-pulse irradiation in real-time. The svOCT values of phantom and iris tissue were averaged along the axial direction, and peak values of the averaged svOCT at each pulse laser irradiation were analyzed. The microscopic images of the treated spots were taken after laser-pulse irradiation. Then, the SS-OCT imaging system was used for *ex vivo* bovine retina study. SvOCT images corresponding to various laser pulse energies and the various number of frames were obtained during laser-pulse irradiation. Similarly, the svOCT values of RPE and photoreceptor layers were averaged along the axial

direction, and peak values of the svOCT at each pulse laser irradiation were analyzed. The microscopic images of the treated spots were taken before and after removing the upper neural retinal layers to assess the degree of retina and RPE damage. Spatial and temporal thermal effects in the retina induced by pulse laser irradiation were simulated and correlated to the peak values of svOCT.

5.2 Phantom and *ex vivo* bovine iris experiment

5.2.1 Experimental method

An in-house built swept-source OCT imaging system was integrated into a R:GEN system as shown in Fig. 27. The OCT system uses a commercial swept source engine (Axsun Technologies, Inc.), operating at 100 kHz sweep rate, the center wavelength of 1060 nm and the sweeping bandwidth of 100 nm. The wavelength of ophthalmic pulse laser is 527 nm, and the laser pulse train consists of 15 pulses with 100Hz repetition rate and 1.7 µs duration. Iris was extracted from *ex vivo* bovine eyes and cut into small pieces. A phantom and bovine iris used for experiment are shown in Fig. 28.

M-scan OCT images of a phantom (floppy disk film) and bovine iris were acquired during laser-pulse trains irradiation. For phantom, we calculated variance of each A-scan from the averaged A-scan before laser pulse radiation by

$$SV_{ij} = [I_i(j) - I_{mean}(j)]^2 \qquad (64)$$

where i and j are indices of frame and axial position of the M-scan, and I_{mean} is average of 1024 A-scans before laser pulse radiation. 80 pixels around the sample surface were set as ROI and the svOCT images in the ROI was averaged in axial direction. For *ex vivo* bovine iris, ten A-scans were averaged to reduce high frequency noise after background subtraction and numerical

dispersion compensation, and the averaged M-scans were used for obtaining svOCT images. The svOCT signal was calculated by Eq. (16).

We tested two different radiation modes of the laser system called classic and ramping mode, where classic mode radiated constant energy for all the pulses while the ramping mode radiated laser pulse energy linearly increased for a laser pulse train.
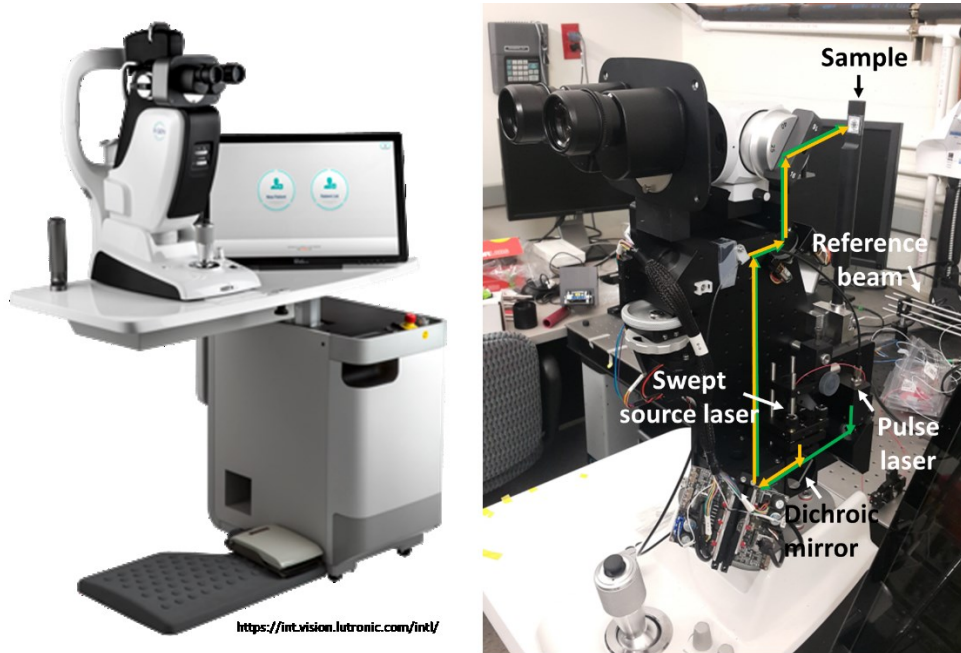
Figure 27 (Left) The R:GEN system and (right) set up of the R:GEN system combined with swept-source OCT imaging system.

Figure 28 Samples for experiment (Left) Floppy disk film and (right) *ex vivo* bovine iris

5.2.2 Results

M-scan OCT image of a phantom is shown in Fig. 29(a). It is consisting of 20,000 A-scans, and it shows visible temporal signal variations by a laser pulse train marked with white arrows. The intensity of A-scans was reduced abruptly after laser pulse radiation and recovered slowly during the time between each laser pulse. Overall shapes of A-scans almost don't change. The phantom has a simple structure consisting of multi-plastic layer on a substrate, and its dominant reflection is Fresnel reflection rather than backscattering. It was thought that a sufficient laser pulse energy would melt the phantom surface, and the reflection from the surface would decrease without any substructural change. In order to see the intensity changes more clearly, which is different from signal variations of biological sample, we calculated variance of each A-scan from the averaged A-scan before laser pulse radiation by Eq. (64). Figure 29(b) and(c) are the svOCT image of Fig. 5.3(a) averaged svOCT signal in the ROI, respectively. Each laser pulse irradiation shows abrupt increase and exponential decay in the averaged svOCT signal. These results seem consistent with the sharp increase in the temperature following laser energy absorption and exponential cooling of the sample without the laser pulse.

Figure 30 shows averaged svOCT signals for laser-pulse train with different energy level in classic mode. The pulse energy was set to 9uJ, 18uJ, 45uJ, 90uJ, 135uJ and 180uJ. The averaged svOCT signals have 15 distinctive peaks corresponding to 15 pulses, and the peak intensities seem to linearly depend on the laser energy. We calculated average and standard deviation value of the svOCT peak signals as a function of the pulse energy, and the average peak value linearly increased with the energy until 135uJ where it saturated as shown in Fig. 31(a). The picture of the laser-pulse irradiated floppy disk surface in Fig 31(b) also shows similar result where the burn mark expands with increasing energy level until 135uJ and it stays relatively constant for the energy

higher than 135uJ. In addition, the svOCT signal shows distinctive peaks with 9uJ radiation, although the picture doesn't show any burn mark at the energy level.
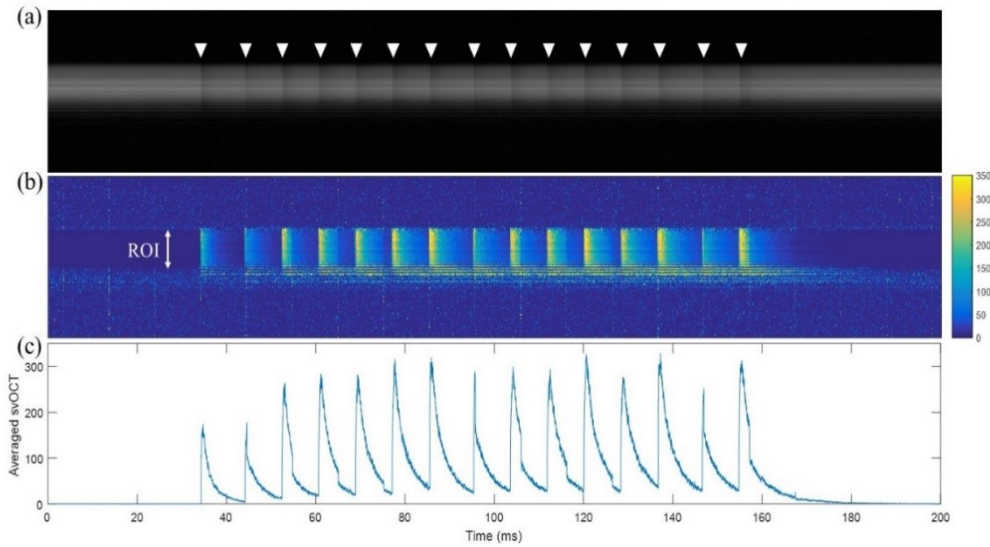
Figure 29 (a) M-scan structural OCT image, (c) M-scan svOCT image, and (d) Averaged svOCT signal in ROI during laser pulse train irradiation.
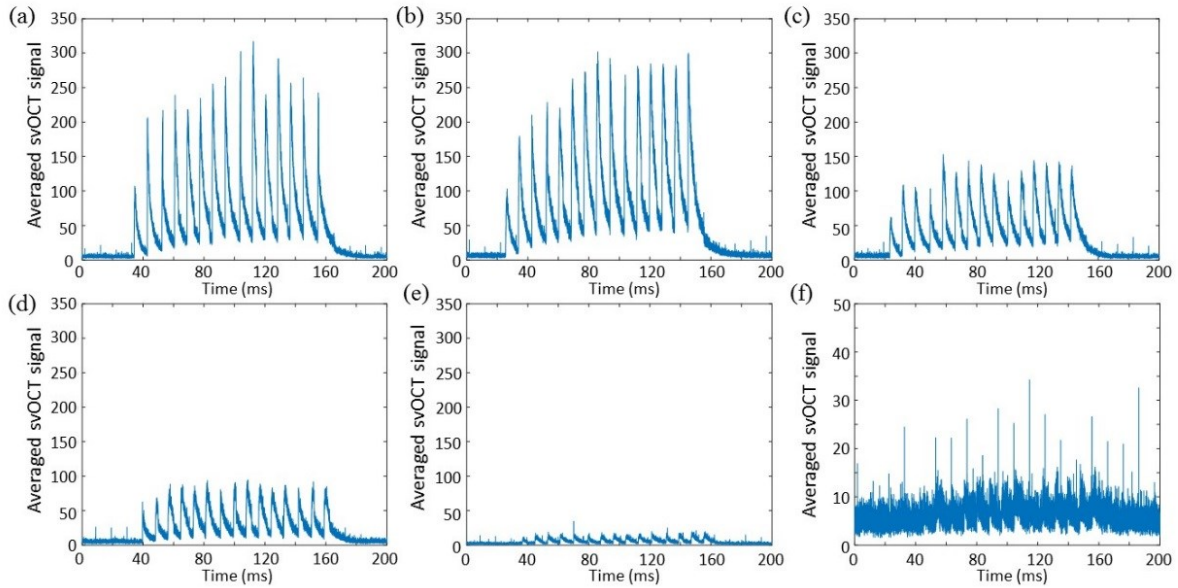
Figure 30 Averaged svOCT signal during a laser-pulse train irradiation with irradiation energy of (a) 180uJ, (b) 135uJ, (c) 90uJ, (d) 45uJ, (e) 18uJ and (f) 9uJ in classic mode
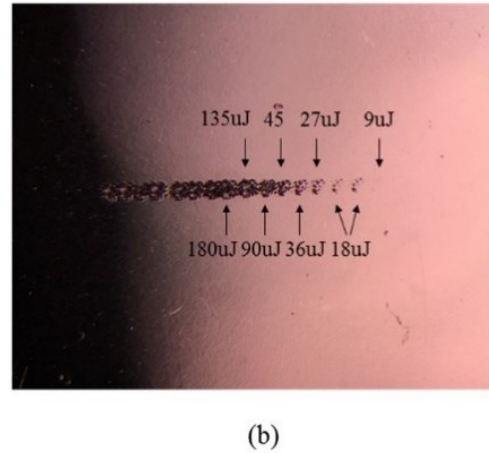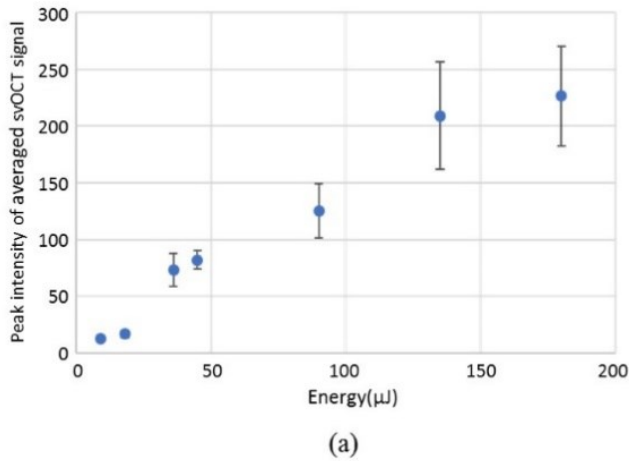
Figure 31 (a) Average (blue point) and standard deviation (bar) of svOCT signal peak value dependent on radiation energy in classic mode. (b) Microscopic image of a laser-pulse irradiated phantom surface.

In ramping mode, the laser pulse energy level ramped up during the pulse train and the corresponding svOCT signal peaks increased proportional to the increasing laser pulse energy. Figure 32(a)-(d) are the averaged svOCT signals with the peak energy level of pulse trains of 180uJ, 90uJ, 45uJ and 18uJ, respectively. The ratio of each peak to the 15th peak was calculated and box-plotted in Fig. 32(e), and it showed linear dependence over the15 laser-pulses during a single pulse train sequence. The 15th peak values at which laser-pulse is irradiated with peak energy were also linearly dependent on the target energy until 180uJ as shown in Fig. 32(f). The 15th peak value in ramping mode were much smaller (less than 60%) than peaks of classic mode except at 180uJ energy level. It was thought to be caused by smaller accumulated thermal energy by ramping mode, and it is expected that it would saturate with energy higher than 180uJ.

Figure 33 shows M-scan OCT images and its averaged svOCT signals of *ex vivo* bovine iris during laser-pulse train irradiation. The temporal signal variations by each laser pulse irradiation are identifiable in the M-scan OCT images, and they become more significant as energy level increases. Averaged svOCT signals show 15 distinctive peaks for the laser-pulse train irradiation, and peak intensities increase with the energy level until it is saturated at 135uJ. The

peak intensities as a function of energy level are boxplotted in Fig. 34(a). Relatively high standard deviation of the peaks at each energy level is possibly caused by the different melanin concentration and the wetness of each sample. At the energy level of 27 uJ, the svOCT signal shows distinctive peaks, but microscopic image doesn't show any burn mark in Fig. 34(b). From this, it is expected that peak detection of the svOCT signal could be more sensitive and effective than analysis of the microscopic image of lesion for laser therapy monitoring.
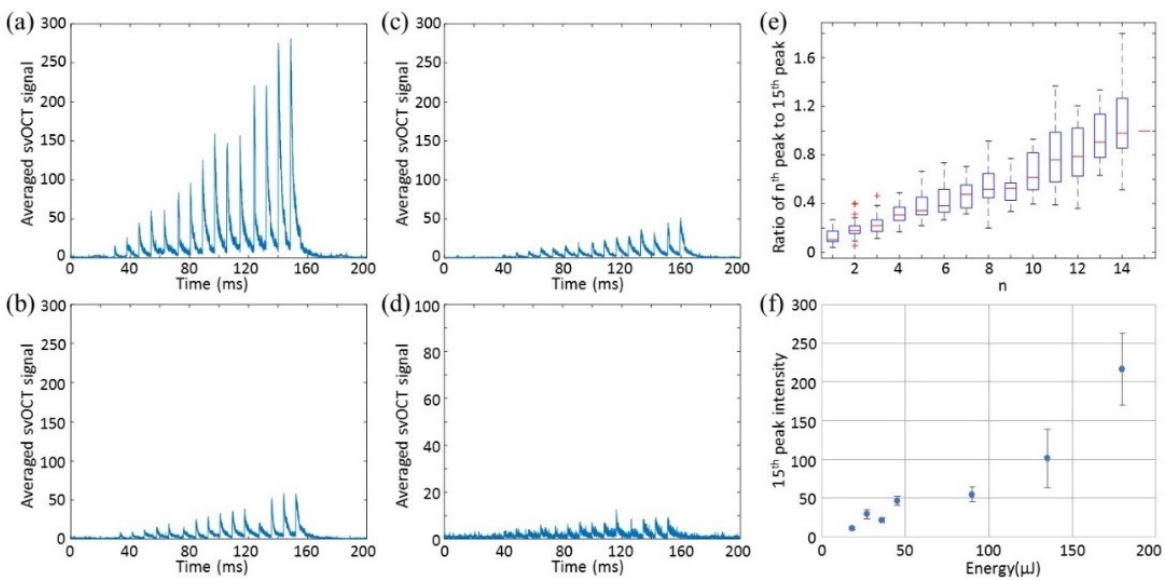
Figure 32 Averaged svOCT signal of a laser-pulse train with radiation energy of (a) 180uJ, (b) 90uJ, (c) 45uJ and (d) 18uJ. (e) Ratio of each peak to 15th peak of averaged svOCT signal. (f) Average (blue point) and standard deviation (bar) of 15th peak values depending on radiation energy in ramping mode.
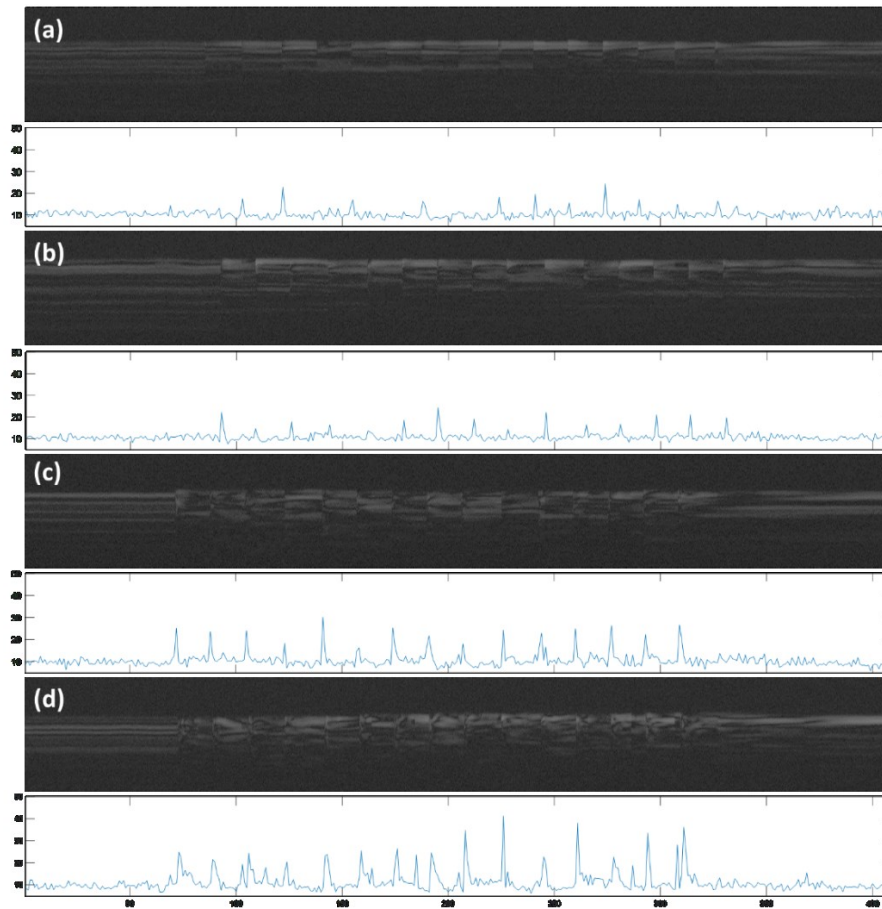
Figure 33 M-scan structural OCT image and its corresponding averaged svOCT signal in ROI during laser pulse train irradiation for the energy level of (a) 27 μJ, (b) 45 μJ, (c) 135 μJ, and (d) 225 μJ.
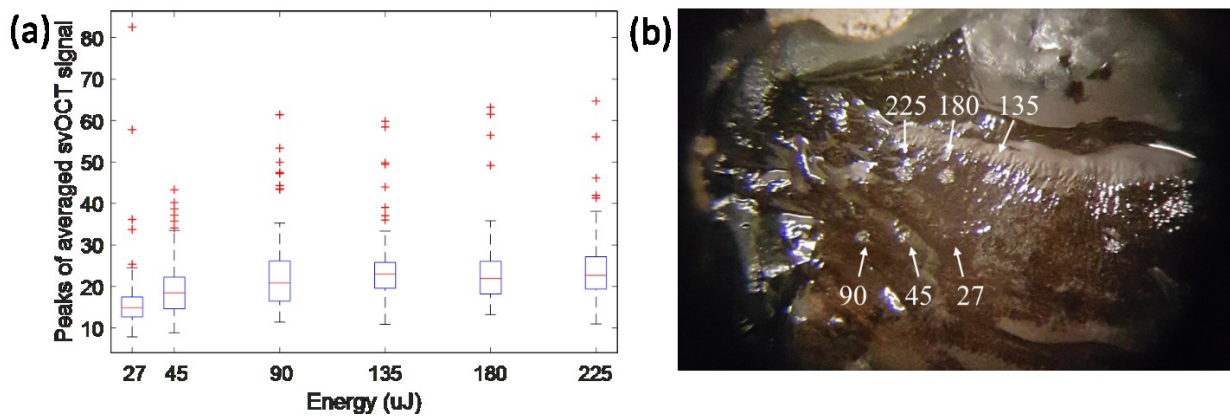
Figure 34 (a) Box plot of peak intensities of averaged svOCT signal as a function of energy level and (b) microscopic image of a laser-pulse irradiated *ex vivo* bovine iris.

5.3 *Ex vivo* bovine retinal experiment and temperature estimation

5.3.1 Experimental method

An in-house built swept-source OCT imaging system was integrated with a frequency-doubled Nd:YLF laser based SRT system (Lutronic, Goyang, Korea). The schematic of the system is shown in Fig. 35. The OCT system used a commercial swept-source engine (Axsun Technologies Inc., Billerica, USA) operating at 100 kHz sweep rate. The center wavelength and sweeping bandwidth of the system was 1060 nm and 100 nm, respectively. The OCT laser was combined with the pulse laser using a dichroic mirror; a Galvano mirror was used to direct the OCT laser to a treated spot on the retina. The wavelength of the pulse laser was 527 nm, and the pulse laser operated at 100 Hz repetition rate and 1.7 µs duration. The pulse laser energy was adjusted from 22 µJ to 190 µJ using neutral-density filters. In this study, fresh ex vivo bovine eyes were acquired from a local butcher and immersed in a cooled saline solution. Bovine eyes have tapetum fibrosum, which has the retinal epithelial layer completely unpigmented, over the central and mid-region of the retina [84]. Because the RPE layer exists in the periphery of the retina, it was difficult to focus the beam on the RPE layer using the crystalline lens of an eye itself. Therefore, the bovine eye's cornea and lens were removed, and the beam was focused on a tilted eye using an objective lens. A total of 39 treated spots were tested on two eyes. M-scan OCT images of the bovine retina were acquired during the laser-pulse irradiation.

SvOCT images were calculated by Eq. (16). The photoreceptor and RPE layers, which are highly scattering and absorptive, were set as a region of interest (ROI), and the svOCT values in the ROI were averaged along the axial direction. Microscopic images of the treated spots were obtained using a CCD camera (DCC1645C, Thorlabs, USA) and a 10X magnification zoom lens.
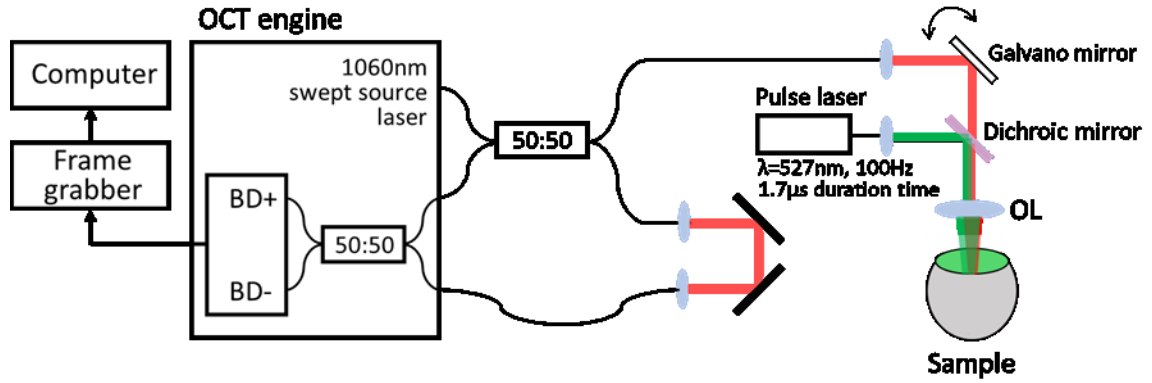
Figure 35 Schematic of a swept-source OCT system integrated into a pulse laser system. BD, balanced detector; OL, objective lens.

Spatial and temporal temperature distribution in the retina were numerically calculated by COMSOL Multiphysics software. Spatial distribution was calculated in 2D, and the geometry of the bovine retina model used for the simulation is shown in Fig. 36(a). The RPE was modeled as a 7-μm layer containing melanosomes that were assumed as spheres of radius 0.3 μm by a discrete absorber model [85]. Melanosomes were diagonally distributed, and the distance between adjacent melanosomes was set to 1.2 μm. The model applied a heat equation shown below,

$$\rho C_p \frac{\partial T}{\partial t} = \nabla \cdot (k\nabla T) + Q, \tag{65}$$

where T is temperature as a function of time (t) and spatial coordinate x and y, ρ is the density, k is the thermal conductivity, and C_p is the heat capacity of the material. The Q refers to the heat source from laser irradiation. The coefficient values used are shown in Table 5, which include the thickness, absorption coefficient and thermal physical constant values [87] of each retinal layer. The absorption coefficient of retinal melanosomes at 532 nm wavelength was estimated from 2370 cm^{-1} to 13000cm^{-1} [88-89], and 6500 cm^{-1}, which makes the absorption in RPE around 50% of the total, was used for our simulation.
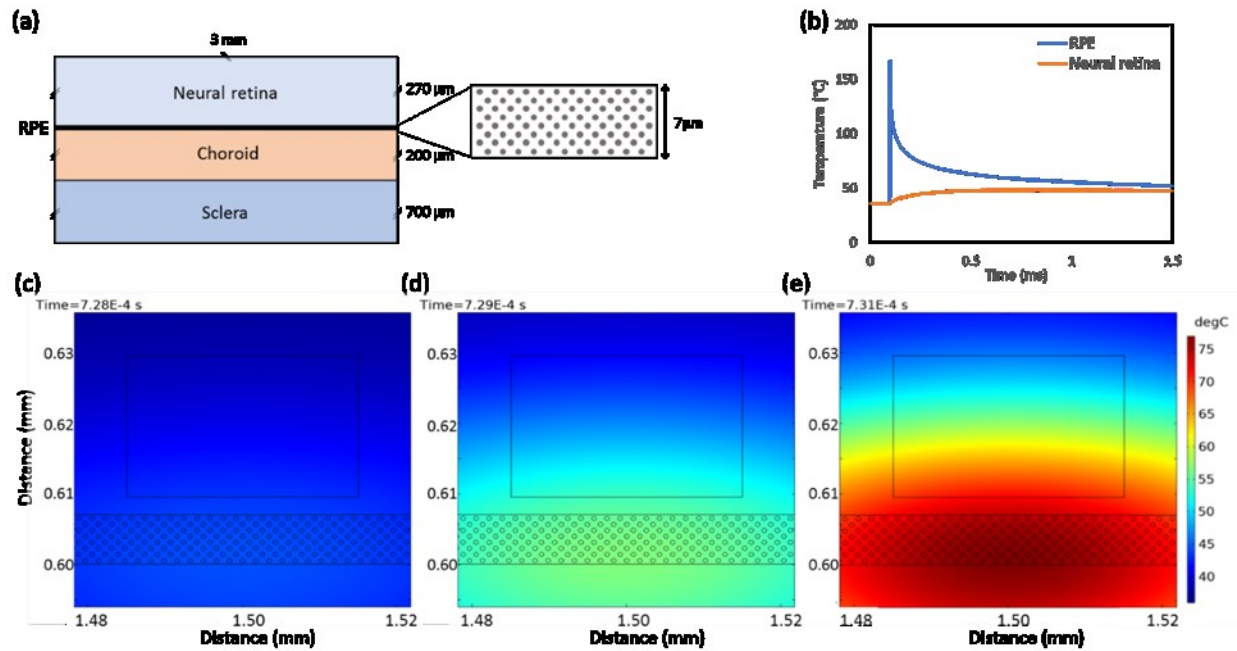
Figure 36 (a) The geometry of the bovine retina model. The retina was assumed to consist of two layers, neural retina and RPE, and have immediate contact with a choroid. The RPE was modeled as a 7-μm layer containing melanosomes that were assumed as diagonally distributed spheres of radius 0.3 μm. (b) Temperature time dependence in the neural retina and at the melanosome surface in RPE when laser pulse of energy 50 μJ irradiated. Spatial distribution of temperature around RPE, when the temperature of the neural retina reached a maximum after a (c) 20 μJ, (d) 50 μJ and (e) 100 μJ pulse irradiation.

Table 6 Thickness, absorption coefficient, and thermal properties of each retinal layer

	Neural retina	RPE layer outside of melanosome	Melanosome	Choroid	Sclera
Thickness[μm]	270	7	0.3 (radius)	200	700
Absorption coefficient[cm^{-1}]	10.4[86]	0	6500	245[86]	4.9
Heat capacity [J/Kg·K]	3680	3680	3680	3680	4178
Density [Kg/m³]	1000	1000	1000	1000	1000
Thermal conductivity [W/m·K]	0.565	0.565	0.565	0.530	0.58

The duration and frequency of the pulse laser were set to 2 μs and 100 Hz, respectively; the peak temperature was calculated over 1.5ms. Gaussian laser-beam profile with a diameter of 150 μm was used for the simulation.

The temperature of the neural retina was calculated by averaging the temperature in a rectangular region (30 μm by 20 μm) 2.5 μm away from RPE, and the temperature of the melanosome surface in RPE was calculated by averaging the temperature on seven melanosome surfaces at the first two melanosome layers located in the center of the Gaussian beam. Fig. 36(b) shows the time-dependent temperature variation in the neural retina (in red) and at the melanosome surface in RPE (in blue). Spatial distributions of temperature around RPE, when the temperature of neural retina reached a maximum after a 20 μJ, 50 μJ and 100 μJ pulse irradiation, are shown in Fig. 36(c)-(e). The calculated peak temperatures in each region were correlated to the peak values of svOCT and tissue damage.

5.3.2 Results

Figure 37(a) and (b) show an M-scan OCT image and the corresponding svOCT image of the bovine retina when 108 μJ energy per pulse irradiated. The M-scan OCT image shows the visible temporal signal variations induced by the laser pulse at the moment marked with white triangles. The signal variation increases the svOCT value. Figure 37(c) shows svOCT values averaged in ROI, and it shows a distinctive peak for each pulse irradiation.

To find appropriate N for calculating the svOCT image, we first tested different values of N and compared the average peak values of the svOCT spikes, the standard deviation of those peaks, and background noise level. Since the speckle variation induced by each laser pulse lasts for only around 50 μs, only 5 frames of the speckle variances show high signal while the rests show very low signal. Intuitively, if we choose a larger window size N, the lower portion of sv signal in the window will be significant, and it will decrease the overall value of the sv signal inside the window, i.e., the peak value of svOCT spikes. Furthermore, the variance of those peak
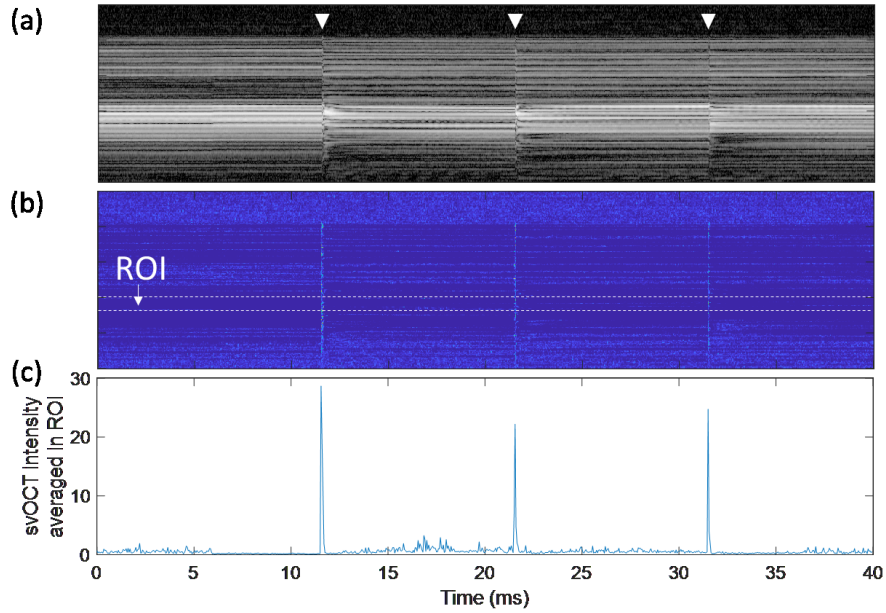
Figure 37 (a) M-scan OCT image of the bovine retina and corresponding (b) svOCT image. Photoreceptor and RPE layers, which are highly scattering and absorptive, were set as an ROI. (c) Axially averaged svOCT values in the ROI during pulse laser irradiation. White triangles mark the moment when each laser pulse (108 µJ) irradiated.

values will also decrease. Fig. 38(a)-(d) show how the average peak values (shapes) and standard deviation (error bar) change depending on laser energy level when N is 2, 5, 10 and 20, respectively. We can see that, as expected, both the average peak value and the standard deviation decrease when N increases. Relative standard deviation, defined as the ratio of the standard deviation to the mean, was 0.53, 0.46, 0.43 and 0.46 for N of 2, 5, 10 and 20, respectively. Therefore, a mid-range of N values between 5-10 are suitable for calculating svOCT in terms of the precision and repeatability. In addition, note that the background noise level increases with increasing N. For N of 2, 5, 10 and 20, the average upper bound levels of background noise were 0.54 ± 0.24, 0.74 ± 0.31, 1.03 ± 0.41 and 1.64 ± 0.66, respectively. The background noise levels were bounded by $\mu + 3\sigma$ upper limit, where μ was svOCT values averaged in ROI before or after laser irradiation, and σ was the standard deviation. The increase in background noise was caused by the bulk motion of the sample and other environmental changes.
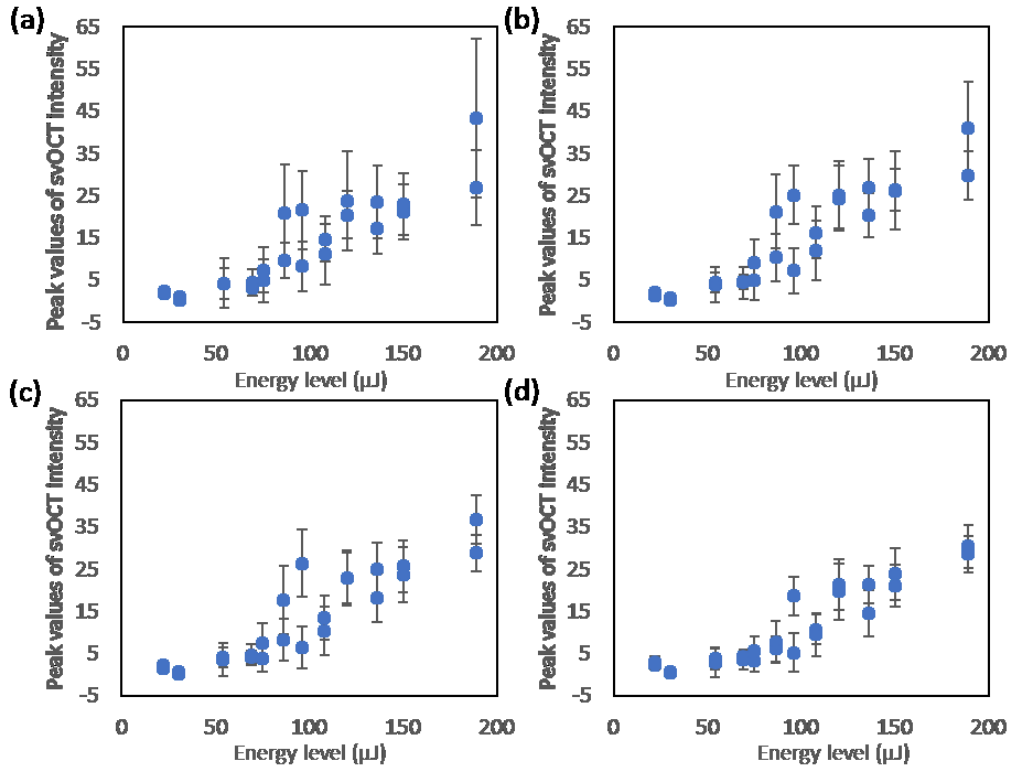
Figure 38 Mean (shapes) and standard deviation (error bar) of peak values of svOCT values averaged in ROI depending on laser pulse energy when window size N is (a) 2, (b) 5, (c) 10 and (d) 20.
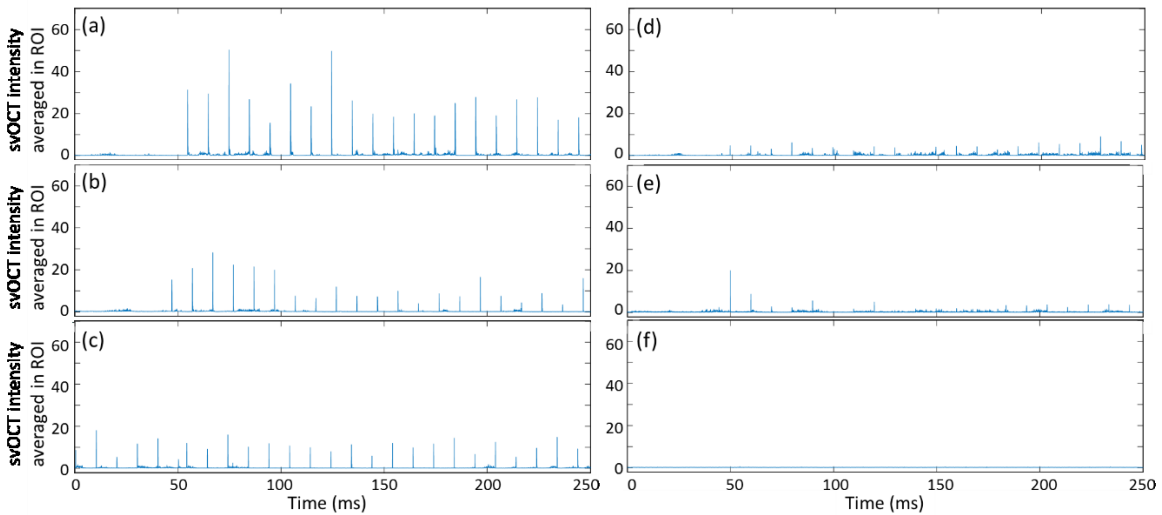
Figure 39 SvOCT values averaged in ROI when pulse laser energy is (a) 150 μJ, (b) 108 μJ, (c) 86 μJ, (d) 69 μJ, (e) 54 μJ and (f) 30 μJ.

Shorter integration time will decrease these effects and so will the smaller value of N. Considering both the effects of relative standard deviation and background noise levels, we choose N as 5.

Figure 39 shows svOCT values averaged in ROI depending on pulse laser energy when N is 5. Peak values increased with increasing pulse laser energy, and the distinctive peaks were observed when the laser pulse energy was as low as 54 μJ. Figure 40(a) and (b) show microscopic images of the treated spots before and after peeling upper neural retinal layers off, respectively. The energy level of each spot is shown in Fig. 40(c). The denaturation of neural retina can be observed as whitish spots (pointed by white triangles) in Fig. 40(a), but the lesions confined only to the RPE layer are invisible. The leftmost and rightmost columns are high energy lesions marking pattern of the spots. In Fig. 40(b), lesions in the RPE layer (pointed by arrows) can be detected by peeling off the upper neural retinal layers.

Average peak values of svOCT depend on the laser pulse energy; the result is summarized in Fig. 41(a). The data with blue circles indicate when the laser pulses induce a lesion in the upper neural layers in addition to inducing a lesion in the RPE layer (corresponding to the indicated spots in Fig. 40(a)). The data represented by orange squares represent the cases when the lesions were induced only on the RPE layer, indicated by white triangle arrows on Fig. 5.14(b). The laser-induced lesion confined only to the RPE layer can be considered as a successful treatment. Fifteen peaks for each spot are averaged, and the standard deviation is shown by the error bar. As expected, it was difficult to define the threshold energy level that induces lesion only in the RPE layer. If the null hypothesis is defined as a successful treatment and the treatment is decided to be successful when the energy level is in the range from 41.9 μJ to 92.1 μJ, which is determined by logistic regression, the type I error and the type II error were 26.7 % and 25 %, respectively. Compared to the energy level, the average peak values of svOCT showed a better correlation with the lesion

creation. Average sv peak values ranged from 12.4 to 38.7 when the induced lesions were observed in both neural and RPE layers; sv values ranged from 0.3 to 18.3 when the lesion was confined only to the RPE layer. In the range of sv peak values from 0.7 to 1.9, no induced lesion was observed at all. When the treatment is decided to be successful with the average svOCT peak value
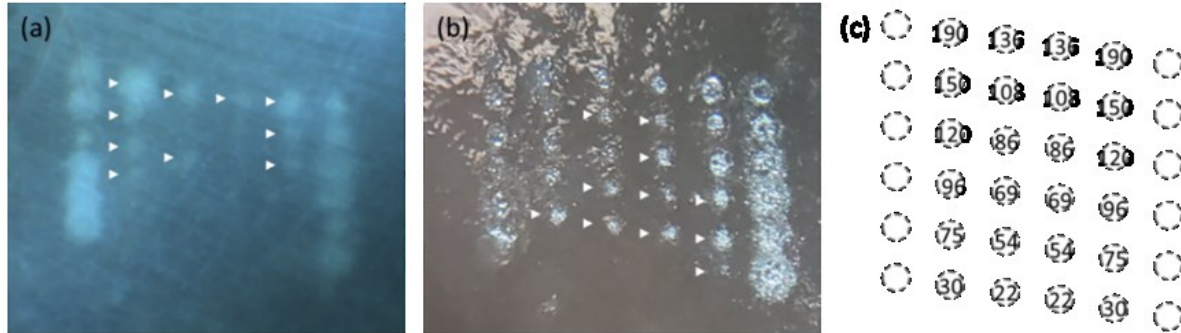
Figure 40 Microscopic image of the retina (a) before and (b) after peeling neural retinal layers off. (c) The energy level of treated spots.

between 1.88 and 15.3 based on logistic regression, the type I error and the type II error were 20% and 0.083%, which was better than the case when the threshold was set by the energy level. For dosimetry control, the method is designed to be used with a power ramped pulsed mode in which laser energy increases linearly from pulse to pulse within one pulse train and automatically stop the next pulse irradiation when the svOCT peak values reached the predetermined threshold value. In addition, the average peak value of svOCT intensities was correlated to the simulated temperature of the neural retina and melanosome surface in RPE. The linear regression of average peak values of svOCT on laser energy was calculated for photocoagulated lesions and selectively damaged lesions. Then, each of them was correlated to the simulated temperature of neural retina and RPE because the tissue damage process was different from each other. Since the simulated temperature of melanosome surface in the RPE was also linear to pulse laser energy, it was
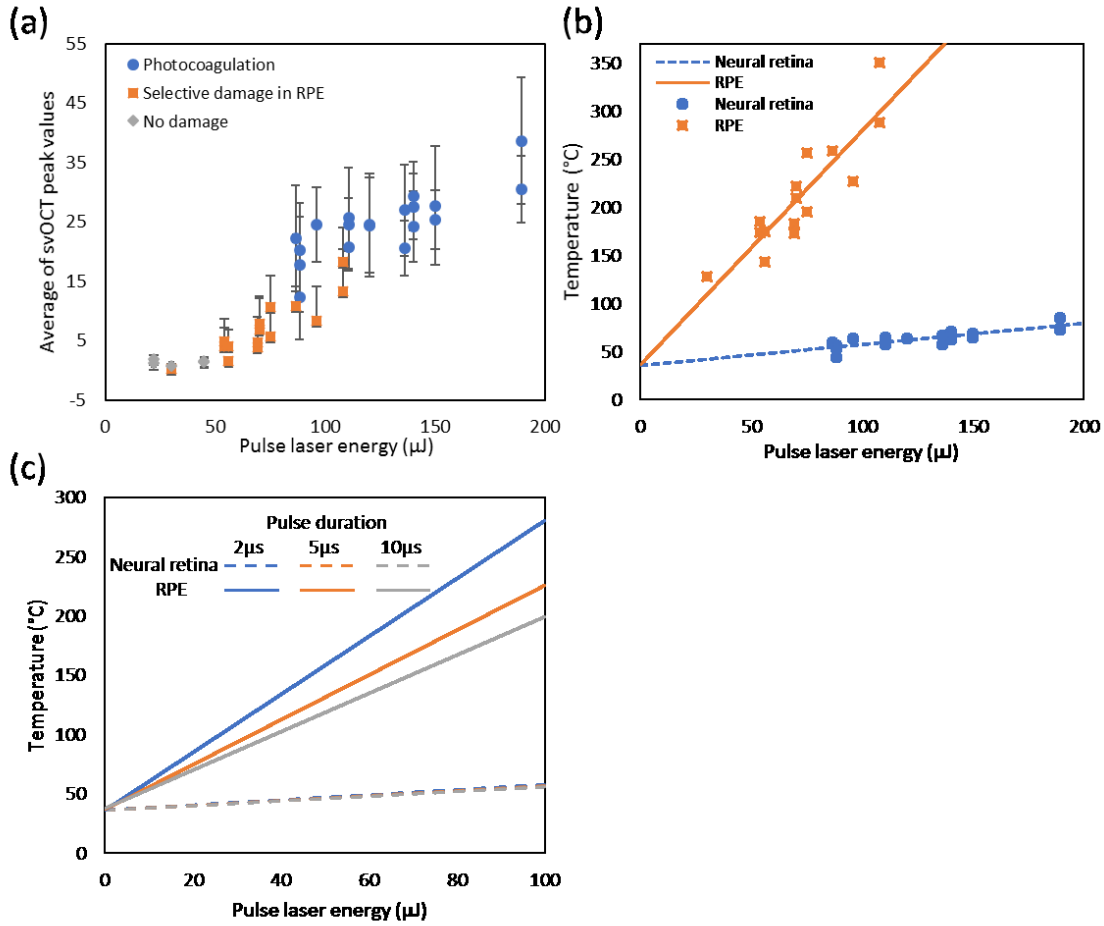
Figure 41 (a) Averaged peak values depending on pulse laser energy and damage range. (b) Simulated (lines) and estimated temperature from the svOCT intensity (shapes) at neural retina and RPE. (c) Simulated temperature at the neural retina and the RPE as a function of laser energy level for three pulse durations, 2 μs, 5 μs, and 10 μs.

correlated to the average peak values of selectively damaged lesions as,

$$T_M = 124.5 + 12.4P, \tag{66}$$

temperature of melanosome surface in the RPE was also linear to pulse laser energy, it was

correlated to the average peak values of selectively damaged where T_M is the temperature of

melanosome surface in RPE, and P is the average peak values of svOCT. Fig. 41(b) shows the

simulated temperature (solid line) and estimated temperature (square shapes) from the average

peak values of svOCT. Most of the estimated temperatures at RPE were higher than 150 °C, while

the lowest temperature was 127.7 °C. This is reasonable since the micro-vaporization is known to occur at around 150 °C. Similarly, the average peak values of the photocoagulated lesions were correlated to the temperature of the neural retina and empirically fitted to a line as,

$$T_N = 25.5 + 1.5P. \tag{67}$$

Fig. 41(b) shows that the temperature of the neural retina was mostly estimated to increase higher than 50 °C when the neural retina was photocoagulated.

We also simulated temperatures of the neural retina and the melanosome surface in RPE as a function of laser energy level for three pulse durations, 2 μs, 5 μs, and 10 μs as shown in Fig. 41(c). The temperature of the neural retina does not change significantly, but the temperature of the melanosome surface in RPE decreases as pulse duration increases. The decrease in temperature can be explained by less heat confinement with longer pulse duration due to heat diffusion during irradiation.

5.4 Conclusion

In conclusion, it was shown that the SRT could be successfully monitored by the svOCT imaging system when integrated with the SRT system. We tested our SS-OCT imaging system integrated with a micro-second pulsed laser system using a phantom and bovine iris models in order to examine its availability for monitoring of each laser-pulse irradiation in real-time. Then, our system performance was tested using *ex vivo* bovine eyes; the svOCT showed distinctive signal variation corresponding to each laser pulse irradiation. The signal variations were proportional to pulse energy levels, and it had a reliable correlation with the creation of lesion within the retina. The temperature at the neural retina and RPE was estimated by svOCT peak values using temperature simulation results, which was consistent with the observed lesion creation. However, we could have missed some minor tissue damages when assessing the photocoagulation and RPE

cell damage from the microscopic images. Therefore, more studies that incorporate further analysis supported by histology or fluorescence microscopy would be needed to obtain a more accurate correlation between the svOCT signal and retinal damage range. In addition, we plan to perform *in vivo* studies using a live animal model to fully validate the utility of this method as an automatic dosimetry control in clinical SRT systems.

6

Conclusions

In this Chapter, the technical achievements and contribution of this dissertation are summarized along with some discussion of future research directions.

6.1 Summary of contributions

We presented automatic axial motion guidance of microsurgical tools (i.e., a subretinal injector) using a fiber-optic CP-SSOCT distal sensor. A high-index epoxy lensed fiber was designed and fabricated to have sufficiently long effective distal sensing ranges in water (or vitreous humor) and, thus, to obtain improved retinal image quality. Enhanced performance was confirmed by SNR measurement and B-scan OCT imaging. The SNR of the proposed fiber probe, compared to a bare fiber, was increased by 25 dB at a distance of 1.5 mm, and the resolution improved from 476 μm to 47 μm. Real-time A-scan-based CNN segmentation was applied for retinal boundary tracking and automatic retinal boundary targeting. A-scan retinal OCT images are segmented using a simplified 1D U-net, and the Kalman filter reduces retinal boundary tracking error by combining boundary position measurement and velocity measurement. We achieve the MUE of around 3 pixels (8.1 μm) using an *ex vivo* bovine retina model. GPU parallel computing allows real-time inference (~1.6 ms) and, thus, real-time retinal boundary tracking. The MSE between target depth and target boundary position of the depth targeting experiment is -0.15 μm and 0.11 μm for the VH/GCL and the PR/CH boundary, respectively. Involuntary tremors, which include low-frequency draft in the order of hundreds of micrometers and physiological tremor in the order of

tens of micrometers, are reduced significantly, and the SDs of target boundary positions are 9.42 µm for the VH/GCL boundary and 10.8 µm for the PR/CH boundary.

It was shown that the SRT could be successfully monitored by the svOCT imaging system when integrated with the SRT system. We found a reliable correlation between the svOCT peak values and the degree of retinal lesion formation, which can be used for selecting proper laser energy during SRT. SvOCT images show abrupt speckle variance changes when samples are irradiated by laser pulses. The averaged svOCT intensities along the axial direction show sharp peaks corresponding to each laser pulse, and the peak values were proportional to the laser pulse energy. The peak values also had a reliable correlation with the creation of lesion within the retina. The temperature at the neural retina and RPE was estimated by svOCT peak values using temperature simulation results, which was consistent with the observed lesion creation.

6.2 Future work

So far, progress has been made in applying OCT imaging system to ophthalmic therapy guidance. The method to stably integrate the CP-OCT lensed-fiber probe to injector needle needs to be considered, and we plan to perform *ex vivo* and *in vivo* studies of subretinal injection using our system to validate its clinical applicability. In addition, *in vivo* studies using a live animal model will be performed to fully validate the utility of svOCT-based monitoring method as an automatic dosimetry control in clinical SRT systems.

Bibliography

[1] L. F. Hotraphinyo and C. N. Riviere, "Three-dimensional accuracy assessment of eye surgeons," in 2001 ConferenceProceedings of the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, vol. 4, pp. 3458–3461 (2001).

[2] S. P. N. Singh and C. N. Riviere, "Physiological tremor amplitude during retinal microsurgery," in Proceedings of the IEEE 28th Annual Northeast Bioengineering Conference (IEEE Cat. No.02CH37342), pp. 171–172 (2002).

[3] A. F. Fercher, W. Drexler, C. K. Hitzenberger et al., "Optical coherence tomography – principles and applications," Reports on Progress in Physics 66, 239–303 (2003).

[4] C. Song, D. Y. Park, P. L. Gehlbach, S. J. Park, and J. U. Kang, "Fiber-optic OCT sensor guided "smart" micro-forceps for microsurgery," Biomed. Opt. Express 4, 1045–1050 (2013).

[5] G. W. Cheon, Y. Huang, J. Cha, P. L. Gehlbach, and J. U. Kang, "Accurate real-time depth control for CP-SSOCT distal sensor based handheld microsurgery tools," Biomed. Opt. Express 6, 1942–1953 (2015).

[6] G. W. Cheon, B. Gonenc, R. H. Taylor, P. L. Gehlbach, and J. U. Kang, "Motorized microforceps with active motion guidance based on common-path SSOCT for epiretinal membranectomy," IEEE/ASME Transactions on Mechatronics 22, 2440–2448 (2017).

[7] J. U. Kang and G. W. Cheon, "Demonstration of subretinal injection using common-path swept source OCT guided microinjector," Appl. Sci. 8, 1287 (2018).

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

[9] A. Shah, L. Zhou, M. D. Abrámoff, and X. Wu, "Multiple surface segmentation using

convolution neural nets: application to retinal layer segmentation in OCT images," Biomed. Opt. Express 9, 4509–4526 (2018).

[10] S. K. Devalla, P. K. Renukanand, B. K. Sreedhar, G. Subramanian, L. Zhang, S. Perera, J.-M. Mari, K. S. Chin, T. A.Tun, N. G. Strouthidis, T. Aung, A. H. Thiéry, and M. J. A. Girard, "DRUNET: a dilated-residual U-Net deep learning network to segment optic nerve head tissues in optical coherence tomography images," Biomed. Opt. Express9,3244–3265 (2018).

[11] S. Borkovkina, A. Camino, W. Janpongsri, M. V. Sarunic, and Y. Jian, "Real-time retinal layer segmentation of OCT volumes with GPU accelerated inferencing using a compressed, low-latency neural network," Biomed. Opt. Express11, 3968–3984 (2020).

[12] J. Roider, S. H. M. Liew, C. Klatt, et al., "Selective retina therapy (SRT) for clinically significant diabetic macular edema," Graefes Arch. Clin. Exp. Ophthalmol. 248(9), 1263 (2010).

[13] Y.G. Park, J. R. Kim, S. Kang, et al., "Safety and efficacy of selective retina therapy (SRT) for the treatment of diabetic macular edema in Korean patients," Graefes Arch. Clin. Exp. Ophthalmol. 254, 1703 (2016).

[14] H. Elsner, E. Pörksen, C. Klatt, et al. "Selective retina therapy in patients with central serous chorioretinopathy," Graefe's Arch. Clin. Exp. Ophthalmol. 244(12), 1638 (2006).

[15] C. Klatt, M. Saeger, T. Oppermann, et al., "Selective retina therapy for acute central serous chorioretinopathy," Br. J. Ophthalmol. 95, 83-88 (2011).

[16] C. Framme, A. Walter, L. Berger, P. Prahs, C. Alt, D. Theisen-Kunde, J. Kowal, and R. Brinkmann, "Selective retina therapy in acute and chronic-recurrent central serous chorioretinopathy," Ophthalmologica 234, 177-188 (2015).

[17] C. Framme, R. Brinkmann, R. Birngruber, and J. Roider, "Autofluorescence imaging after

selective RPE laser treatment in macular diseases and clinical outcome: a pilot study," Br. J. Ophthalmol. 86, 1099-1106 (2002).

[18] R. Brinkmann, J. Roider, R. Birngruber, "Selective retina therapy (SRT): a review on methods, techniques, preclinical and first clinical results," Bull. Soc. Belge. Ophtalmol. 302, 51 (2010).

[19] J. J. Weiter, F. C. Delori, G. L. Wing, K. A. Fitch, "Retinal pigment epithelial lipofuscin and melanin and choroidal melanin in human eyes," Invest. Ophthalmol. Vis. Sci. 27(2), 145-152 (1986).

[20] A. Chopdar, T. Aung, "Ch 1 Fundus fluorescein angiography," in Multimodal retinal imaging, JP Medical Ltd, Victoria (2014).

[21] Izatt J.A., Choma M.A., "Theory of Optical Coherence Tomography," In: Drexler W., Fujimoto J.G. (eds) Optical Coherence Tomography. Biological and Medical Physics, Biomedical Engineering. Springer, Berlin, Heidelberg (2008).

[22] Choma M, et al., "Sensitivity advantage of swept source and Fourier domain optical coherence tomography," Opt. Express. 11(18):2183 (2003).

[23] Leitgeb R, Hitzenberger C, Fercher A. Performance of fourier domain vs time domain optical coherence tomography. Opt Express. 11(8):889 (2003).

[24] Johannes F. de Boer, Barry Cense, B. Hyle Park, Mark C. Pierce, Guillermo J. Tearney, and Brett E. Bouma, "Improved signal-to-noise ratio in spectral-domain compared with time-domain optical coherence tomography," Opt. Lett. 28, 2067-2069 (2003)

[25] Andrei B. Vakhtin, Daniel J. Kane, William R. Wood, and Kristen A. Peterson, "Common-path interferometer for frequency-domain optical coherence tomography," Appl. Opt. 42, 6953-6958 (2003)

[26] Kang J.U., Han J.H., Liu X, Zhang K. Common-Path Optical Coherence Tomography for

Biomedical Imaging and Sensing. J Opt Soc Korea. 14(1):1-13 (2010).

[27] Zhang K, Wang W, Han J, Kang JU. "Surface Topology and Motion Compensation System for Microsurgery Guidance and Intervention based on Common-Path Optical Coherence Tomography," IEEE Trans. Biomed. Eng. 56:2318–2321 (2009).

[28] Joseph M. Schmitt, S. H. Xiang, and Kin Man Yung "Speckle in optical coherence tomography," J. Biomed. Opt. 4(1) (1999).

[29] X. Liuet al., "Spectroscopic-speckle variance OCT for microvasculature detection and analysis," Biomed. Opt. Express 2 (11), 2995 –3009 (2011).

[30] Changho Lee, Gyeongwoo Cheon, Do-Hyun Kim, Jin U. Kang, "Feasibility study: protein denaturation and coagulation monitoring with speckle variance optical coherence tomography," J. Biomed. Opt. 21(12) 125004 (2016).

[31] S. Guo, S. Wei, S. Lee, M. Sheu, S. Kang and J. U. Kang, "Intraoperative Speckle Variance Optical Coherence Tomography for Tissue Temperature Monitoring During Cutaneous Laser Therapy," in *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 7, pp. 1-8, Art no. 1800608 (2019).

[32] Soohyun Lee, Shuwen Wei, Shoujing Guo, Jongmin Kim, Bongkyun Kim, Gihoon Kim, Jin U. Kang, "Selective retina therapy monitoring by speckle variance optical coherence tomography for dosimetry control," J. Biomed. Opt. 25(2) 026001 (2020).

[33] S. Albawi, T. A. Mohammed and S. Al-Zawi, "Understanding of a convolutional neural network," *2017 International Conference on Engineering and Technology (ICET)*, 2017, pp. 1-6, doi: 10.1109/ICEngTechnol.2017.8308186.

[34] Ghosh A., Sufian A., Sultana F., Chakrabarti A., De D. (2020) Fundamental Concepts of Convolutional Neural Network. In: Balas V., Kumar R., Srivastava R. (eds) Recent Trends

and Advances in Artificial Intelligence and Internet of Things. Intelligent Systems Reference Library, vol 172. Springer, Cham.

[35] Yamashita, R., Nishio, M., Do, R.K.G. et al. Convolutional neural networks: an overview and application in radiology. Insights Imaging 9, 611–629 (2018).

[36] I. Goodfellow, Y. Bengio, and A. Courville, "Chapter 9 Convolutional Networks," Deep learning, MIT Press (2016). http://www.deeplearningbook.org.

[37] Duchi, J.C.; Hazan, E.; Singer, Y. "Adaptive Subgradient Methods for Online Learning and Stochastic Optimization," J. Mach. Learn. Res. 12, 2121–2159 (2011).

[38] Kingma, D.; Ba, J. Adam, "A Method for Stochastic Optimization," Available online: https://arxiv.org/abs/1412. 6980 (accessed on 22 August 2020).

[39] Shuo Han, Marinko V. Sarunic, Jigang Wu, Mark S. Humayun M.D., Changhuei Yang, "Handheld forward-imaging needle endoscope for ophthalmic optical coherence tomography inspection," J. Biomed. Opt. 13(2) 020505(2008)

[40] J. U. Kang, J. -H. Han, X. Liu, K. Zhang, C. G. Song and P. Gehlbach, "Endoscopic Functional Fourier Domain Common-Path Optical Coherence Tomography for Microsurgery," in IEEE Journal of Selected Topics in Quantum Electronics, vol. 16, no. 4, pp. 781-792 (2010).

[41] Karen M. Joos and Jin-Hui Shen, "Miniature real-time intraoperative forward-imaging optical coherence tomography probe," Biomed. Opt. Express 4, 1342-1350 (2013).

[42] Tetsu Asami, Hiroko Terasaki, Yasuki Ito, Tadasu Sugita, Hiroki Kaneko, Junpei Nishiyama, Hajime Namiki, Masahiko Kobayashi, Norihiko Nishizawa; Development of a Fiber-Optic Optical Coherence Tomography Probe for Intraocular Use. *Invest. Ophthalmol. Vis. Sci.* 57(9): OCT568-OCT574 (2016).

[43] Mura, M., Iannetta, D., Nasini, F., Barca, F., Peiretti, E., Engelbrecht, L., de Smet, M.D. and

Verbraak, F., Use of a new intra-ocular spectral domain optical coherence tomography in vitreoretinal surgery. Acta Ophthalmol, 94: 246-252 (2016).

[44] Seon Young Ryu, Hae Young Choi, Jihoon Na, Woo June Choi, and Byeong Ha Lee, "Lensed fiber probes designed as an alternative to bulk probes in optical coherence tomography," Appl. Opt. 47, 1510-1516 (2008).

[45] Y. Qiu, Y. Wang, K. D. Belfield, and X. Liu, "Ultrathin lensed fiber-optic probe for optical coherence tomography," Biomed. Opt. Express 7, 2154-2162 (2016).

[46] Mingtao Zhao, Yong Huang, and Jin U. Kang, "Sapphire ball lens-based fiber probe for common-path optical coherence tomography and its applications in corneal and retinal imaging," Opt. Lett. 37, 4835-4837 (2012).

[47] Kanwarpal Singh, Daisuke Yamada, Guillermo Tearney, "Common Path Side Viewing Monolithic Ball Lens Probe for Optical Coherence Tomography," Sovremennye tehnologii v medicine 7(1): 29–33 (2015).

[48] Sy-Bor Wen, Vijay M. Sundaram, Daniel McBride, and Yu Yang, "Low-cost, high-precision micro-lensed optical fiber providing deep-micrometer to deep-nanometer-level light focusing," Opt. Lett. 41, 1793-1796 (2016).

[49] Marica Marrese, Hidde Offerhaus, Erik Paardekam, and Davide Iannuzzi, "70 μm diameter optical probe for common-path optical coherence tomography in air and liquids," Opt. Lett. 43, 5929-5932 (2018).

[50] Donglin Wang, Can Duan, Xiaoyang Zhang, Zhao Yun, Antonio Pozzi, and Huikai Xie, "Common-path optical coherence tomography using a microelectromechanical-system-based endoscopic probe," Appl. Opt. 55, 6930-6935 (2016).

[51] J. Kim, J. Han and J. Jeong, "Common-Path Optical Coherence Tomography Using a Conical-

Frustum-Tip Fiber Probe," in *IEEE* Journal of Selected Topics in Quantum Electronics, vol. 20, no. 2, pp. 8-14, March-April 2014, Art no. 6800407

[52] Herwig Kogelnik, "On the Propagation of Gaussian Beams of Light Through Lens like Media Including those with a Loss or Gain Variation," Appl. Opt. 4, 1562-1569 (1965).

[53] W. Emkey and C. Jack, "Analysis and evaluation of graded-index fiber lenses," in *Journal of Lightwave Technology*, vol. 5, no. 9, pp. 1156-1164, September 1987, doi: 10.1109/JLT.1987.1075651.

[54] J. A. Arnaud, W. M. Hubbard, G. D. Mandeville, B. de la Clavière, E. A. Franke, and J. M. Franke, "Technique for Fast Measurement of Gaussian Laser Beam Parameters," Appl. Opt. 10, 2775-2776 (1971).

[55] B. Johnson et al., "Swept light sources," Optical Coherence Tomography, 639 –658 Springer, Cham (2015).

[56] Y. Peng, L. Tang, and Y. Zhou, "Subretinal injection: a review on the novel route of therapeutic delivery for vitreoretinal diseases," Ophthalmic Res.58, 217–226 (2017).

[57] Jin U. Kang, Yong Huang, Kang Zhang, Zuhaib Ibrahim, Jaepyeong Cha, W.P. Andrew Lee, Gerald Brandacher, and Peter L. Gehlbach, "Real-Time 3-D Fourier-Domain Optical Coherence Tomography Video Image Guided Microsurgeries," J. Biomed. Opt. 17, 081403 (2012).

[58] Kang Zhang and Jin U. Kang, "Real-time intraoperative 4D full-range FD-OCT based on the dual graphics processing units architecture for microsurgery guidance," Biomedical Optics Express, Vol. 2, Issue 4, pp. 764-770 (2011).

[59] M. Draelos, G. Tang, B. Keller, A. Kuo, K. Hauser, and J. A. Izatt, "Optical coherence tomography guided robotic needle insertion for deep anterior lamellar keratoplasty," IEEE

Transactions on Biomed. Eng.67, 2073–2083 (2020).

[60] M. Zhou, Q. Yu, K. Huang, S. Mahov, A. Eslami, M. Maier, C. P. Lohmann, N. Navab, D. Zapp, A. Knoll, and M. A.Nasseri, "Towards robotic-assisted subretinal injection: A hybrid parallel-serial robot system design and preliminary evaluation," IEEE Transactions on Ind. Electron.67, 6617–6628 (2020).

[61] M. Sommersperger, J. Weiss, M. A. Nasseri et al., "Real-time tool to layer distance estimation for robotic subretinal injection using intraoperative 4D OCT," Biomed. Opt. Express 12, 1085–1104 (2021).

[62] A. Yazdanpanah, G. Hamarneh, B. Smith, and M. Sarunic, "Intra-retinal layer segmentation in optical coherence tomography using an active contour approach," in Medical Image Computing and Computer-Assisted Intervention– MICCAI 2009,G.-Z. Yang, D. Hawkes, D. Rueckert, A. Noble, and C. Taylor, eds. (Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 649–656) (2009).

[63] A. González-López, J. de Moura, J. Novo, M. Ortega, and M. Penedo, "Robust segmentation of retinal layers inoptical coherence tomography images based on a multistage active contour model," Heliyon5, e01271 (2019).

[64] Kang Li, Xiaodong Wu, D. Z. Chen, and M. Sonka, "Optimal surface segmentation in volumetric images-a graph-theoretic approach," IEEE Transactions on Pattern Analysis Mach. Intell.28, 119–134 (2006).

[65] M. K. Garvin, M. D. Abramoff, X. Wu, S. R. Russell, T. L. Burns, and M. Sonka, "Automated 3-d intraretinal layer segmentation of macular spectral-domain optical coherence tomography images," IEEE Transactions on Med.Imaging28, 1436–1447 (2009).

[66] Z. Hu, X. Wu, A. Hariri, and S. R. Sadda, "Multiple layer segmentation and analysis in three-

dimensional spectral-domain optical coherence tomography volume scans," J. Biomed. Opt.18, 1 − 9 (2013).

[67] S. J. Chiu, X. T. Li, P. Nicholas, C. A. Toth, J. A. Izatt, and S. Farsiu, "Automatic segmentation of seven retinal layers in SDOCT images congruent with expert manual segmentation," Opt. Express18, 19413–19428 (2010).

[68] J. Tian, B. Varga, G. M. Somfai, W.-H. Lee, W. E. Smiddy, and D. C. DeBuc, "Real-time automatic segmentation of optical coherence tomography volume data of the macular region," PLoS ONE10, e0133908.

[69] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedicalimage segmentation," inMedical Image Computing and Computer-Assisted Intervention −MICCAI 2015, N. Navab, J. Hornegger, W. M. Wells et al., Eds., 234–241, Springer Interna-tional Publishing, (Cham) (2015).

[70] F. G. Venhuizen, B. van Ginneken, B. Liefers et al., "Robust total retina thickness segmentation in optical coherence tomography images using convolutional neural networks," Biomed. Opt. Express 8, 3292–3316 (2017).

[71] G. Welch and G. Bishop, "An Introduction to the Kalman Filter," Siggraph Course8, 1−16 (2006).

[72] S. Lee, C. Lee, R. Verkade, G. W. Cheon, and J. U. Kang, "Common-path all-fiber optical coherence tomography probe based on high-index elliptical epoxy-lensed fiber," Opt. Eng.58, 1 − 5 (2019).

[73] R. Brinkmann, G. Huttmann, J. Rogener, J. Roider, R. Birngruber, C. P. Lin, "Origin of retinal pigment epithelium cell damage by pulsed laser irradiance in the nanosecond to microsecond time regimen," Lasers in Surgery and Medicine 27, 451–464 (2000).

[74] G. Schuele, H. Elsner, C. Framme, J. Roider, R. Birngruber, and R. Brinkmann, "Optoacoustic real-time dosimetry for selective retina treatment," J. Biomed. Opt. 10(6), 064022 (2005).

[75] E. Seifert, Y.-J. Roh, A. Fritz, Y. G. Park, S. Kang, D. Theisen-Kunde, and R. Brinkmann, "Automatic irradiation control by an optical feedback technique for selective retina treatment (SRT) in a rabbit model," in Medical Laser Applications and Laser-Tissue Interactions VI, L. Lilge and R. Sroka, Ed., Proc. SPIE 8803, 880303 (2013).

[76] S. Kang, Y. G. Park, J. R. Kim, E. Seifert, T.-K. Dirk, B. Ralf, Y. J. Roh, "Selective retina therapy in patients with chronic central serous chorioretinopathy: a pilot study," Medicine (Baltim.), 95(3), e2524 (2016).

[77] Y. J. Kim, Y. G. Lee, D. W. Lee, and J. H. Kim, "Selective retina therapy with real-time feedback-controlled dosimetry for treating acute idiopathic central serous chorioretinopathy in Korean Patients," Journal of Ophthalmology, vol. 2018, Article ID 6027871 (2018).

[78] P. Steiner, A. Ebneter, L. E. Berger, M. Zinkernagel, B. Považay, C. Meier, J. H. Kowal, C. Framme, R. Brinkmann, S. Wolf, and R. Sznitman, "Time-resolved ultra–high resolution optical coherence tomography for real-time monitoring of selective retina therapy," Invest. Ophthalmol. Vis. Sci. 56(11), 6654-6662 (2015).

[79] S. Zbinden, Ş.S. Kucur, P. Steiner et al., "Automatic assessment of time-resolved OCT images for selective retina therapy," Int. J. CARS. 11, 863 (2016).

[80] D. Kaufmann, C. Burri, P. Arnold, V. M. Koch, C. Meier, B. Považay, and J. Justiz, "Dosimetry control and monitoring of selective retina therapy using optical coherence tomography," Proc. SPIE 10416, Optical Coherence Imaging Techniques and Imaging in Scattering Media II, 1041604 (2017).

[81] T. Fountoukidou, P. Raisin, D. Kaufmann, et al., "Motion-invariant SRT treatment detection

from direct M-scan OCT imaging" Int J CARS 13: 683 (2018).

[82] D. Kaufmann, C. Burri, P. Arnold, et al., "Selective retina therapy enhanced with optical coherence tomography for dosimetry control and monitoring: a proof of concept study," Biomed. Opt. Express. 9(7), 3320-3334 (2018).

[83] M. S. Mahmud, D. W. Cadotte, B. Vuong, C. Sun, T. W. H. Luk, A. Mariampillai, and V. X. D. Yang, "Review of speckle and phase variance optical coherence tomography to visualize microvascular networks," J. Biomed. Opt. 18(5) 050901 (2013).

[84] F. J. Ollivier, D. A. Samuelson, D. E. Brooks, P. A. Lewis, M. E. Kallberg, and A. M. Komáromy, "Comparative morphology of the tapetum lucidum (among selected species)," Veterinary Ophthalmology 7, 11-22 (2004).

[85] J. Roider, F. Hillenkamp, T. Flotte, and R. Birngruber, "Microphotocoagulation: selective effects of repetitive short laser pulses," Proc. Natl. Acad. Sci. USA. 90(18):8643-7 (1993).

[86] D. K. Sardar, F. S. Salinas, J. J. Perez, and A. T. Tsin, "Optical characterization of bovine retinal tissues," J. Biomed. Opt. 9(3) (2004).

[87] S. A. Mirnezami, M. Rajaei Jafarabadi, and M. Abrishami, "Temperature distribution simulation of the human eye exposed to laser radiation," Journal of lasers in medical sciences, 4(4), 175-181 (2013).

[88] S. L. Jacques, R. D. Glickman, and J. A. Schwartz, "Internal absorption coefficient and threshold for pulsed laser disruption of melanosomes isolated from retinal pigment epithelium," Proc. SPIE 2681, Laser-Tissue Interaction VII, (1996).

[89] R. Brinkmann, G. Hüttmann, J. Rögener, J. Roider, R. Birngruber, and C. P. Lin, "Origin of retinal pigment epithelium cell damage by pulsed laser irradiance in the nanosecond to microsecond time regimen," Lasers in Surgery and Medicine 27:451–464 (2000).

Curriculum Vitae

Soohyun Lee
+1-443-875-4953

slee452@jhu.edu, soohyunlee87@gmail.com

EDUCATION

Johns Hopkins University   Baltimore, MD
Ph. D., Department of Electrical Engineering   Aug. 2021 (Expected date)
Advisor: Prof. Jin U. Kang

Korea Advanced Institute of Science and Technology (KAIST)   Daejeon, Korea
M.S., Department of Electrical Engineering   Feb. 2012
Advisor: Prof. Seunghyup Yoo

Korea Advanced Institute of Science and Technology (KAIST)   Daejeon, Korea
B.S., Department of Electrical Engineering   Jan. 2010

SKILLS

Software: C, C++, Python, Matlab, DiffactMOD
Language: English, Korean (native)

WORK & RESEARCH EXPERIENCE

Johns Hopkins University   Baltimore, MD
Research Assistant, Photonics and optoelectronics Lab.   Sep. 2016 – Present
- Developed retinal boundary tracking method using 1D CNN-based segmentation of A-scan optical coherence tomography (OCT) retinal images
- Realized real-time retinal boundary tracking by parallel programming on CUDA and integrated it into OCT-guided microsurgical tool for subretinal injection
- Optimized and fabricated lensed-fiber OCT probe for retinal imaging
- Proposed and developed real-time laser therapy monitoring method using speckle variance OCT

Electronics and Telecommunications Research Institute (ETRI)   Daejeon, Korea
Research Engineer, Digital Holography Research Group   Sep. 2012 – May. 2016
- Designed a computer-generated-hologram algorithm for viewing-window-based digital holographic display system
- Researched noise reduction of holographic 3D images based on optical image processing
- Simulated observation of holographic display using numerical reconstruction of holographic 3D images
- Proposed integral imaging-based objective quality assessment of holographic 3D images

Korea Advanced Institute of Science and Technology (KAIST)   Daejeon, Korea
Research Assistant, Integrated Organic Electronics Lab.   Jul. 2009 – Aug. 2012
- Analyzed and optimized metallic multilayer electrode using thin-film optics for semitransparent, parallel tandem, and inverted organic photovoltaic cells
- Simulated efficiency of diffractive grating-based non-tracking planar solar concentrator combining geometrical and diffractive optics

PUBLICATIONS

Journal Paper

1. **Soohyun Lee**, Jin U. Kang, "CNN-based CP-OCT sensor integrated with a subretinal injector for retinal boundary tracking and injection guidance," J. Biomed. Opt. 26(6) 068001 (2021).

2. **Soohyun Lee**, Shuwen Wei, Shoujing Guo, Jongmin Kim, Bongkyun Kim, Gihoon Kim, Jin U. Kang, "Selective retina therapy monitoring by speckle variance optical coherence tomography for dosimetry control," J. Biomed. Opt. 25(2) 026001 (2020).

3. **Soohyun Lee**, Changho Lee, Ruben Verkade, Gyeong Woo Cheon, Jin U. Kang, "Common-path all-fiber optical coherence tomography probe based on high-index elliptical epoxy-lensed fiber," Opt. Eng. 58(2) 026116 (2019).

4. Shoujing Guo, Shuwen Wei, **Soohyun Lee**, Mary Sheu, Sewon Kang, and Jin U. Kang, "Intraoperative Speckle Variance Optical Coherence Tomography for Tissue Temperature Monitoring During Cutaneous Laser Therapy," IEEE Journal of Translational Engineering in Health and Medicine, vol. 7, pp. 1-8, Art no. 1800608 (2019).

5. Yongjun Lim, Keehoon Hong, Hwi Kim, Hyun-Eui Kim, Eun-Young Chang, **Soohyun Lee**, Taeone Kim, Jeho Nam, Hyon-Gon Choo, Jinwoong Kim, and Joonku Hahn, "360-degree tabletop electronic holographic display," Opt. Express 24, 24999-25009 (2016).

6. **Soohyun Lee**, Tae Eui Kang, Donggeon Han, Hoyeon Kim, Bumjoon J Kim, Jongjin Lee, Seunghyup Yoo, "Polymer/small-molecule parallel tandem organic solar cells based on MoOx–Ag–MoOx intermediate electrodes," Sol. Ener. Mater. and Sol. Cells 137, 34-43 (2015).

7. Hoyeon Kim, **Soohyun Lee**, Donggeon Han, Seunghyup Yoo "High-density organic photovoltaic modules: Mask-free fabrication using nozzle jet printing and oblique deposition," Sol. Ener. Mater. and Sol. Cells, 120, pp.561-565 (2014)

8. Donggeon Han, **Soohyun Lee**, H. Kim, S. Jeong, Seunghyup Yoo, "Cathodic multilayer transparent electrodes for ITO-free inverted organic solar cells," Org. Electron, 14(6), pp.1477-1482 (2013)

9. Jungmin Choi, Tae-Wook Koh, **Soohyun Lee**, and Seunghyup Yoo, "Enhanced light extraction in organic light-emitting devices: using conductive low-index layers and micropatterned indium tin oxide electrodes with optimal taper angle," App. Phys. Lett. 100, 233303 (2012)

10. Maengsun Eo, **Soohyun Lee**, Myung Hwan Park, Min Hyung Lee, Seunghyup Yoo, Youngkyu Do, "Vinyl-Type Polynorbornenes with Pendant PCBM: A Novel Acceptor for Organic Solar Cells," Macromol. Rap. Comm. 33(13), 1119-1125 (2012)

11. Donggeon Han, Hoyeon Kim, **Soohyun Lee**, Myungsoo Seo, Seunghyup Yoo, "Realization of efficient semitransparent organic photovoltaic cells with metallic top electrodes: utilizing the tunable absorption asymmetry," Opt. Express, Vol. 18, No.S4, A513-A521 (2010)

Conference

1. Mandeep S Singh, Shoujing Guo, Shuwen Wei, **Soohyun Lee**, Jin Kang, "Optical coherence tomography distal-sensor guided manual injection device for transscleral subretinal access," Invest. Ophthalmol. Vis. Sci. 2019;60(9):5800, ARVO Annual Meeting (2019, Vancouver, Canada)

2. Shoujing Guo, Shuwen Wei, Nicolas Sarfaraz, **Soohyun Lee**, William G. Gensheimer, Axel Krieger, Jin Kang, "Optical coherence tomography distal-sensor guided manual trephine/dissection system for DALK," *Invest. Ophthalmol. Vis. Sci.* 2019;60(9):1870, ARVO Annual Meeting (2019, Vancouver, Canada)

3. **Soohyun Lee**, Changho Lee, Gyeongwoo Cheon, Jongmin Kim, Dongki Jo, Jihoon Lee, Jin U. Kang, "Ophthalmic laser system integrated with speckle variance optical coherence tomography for real-time temperature monitoring," Proc. SPIE 10483, Optical Coherence Tomography and Coherence Domain Optical Methods in Biomedicine XXII, 104832C (2018, San Francisco).

4. Kristina Irsch, **Soohyun Lee**, Sanjukta N. Bose, Jin U. Kang, "Motion-compensated optical coherence tomography using envelope-based surface detection and Kalman-based prediction," Proc. SPIE 10484, Advanced Biomedical and Clinical Diagnostic and Surgical Guidance Systems XVI, 104840Q (2018, San Francisco, USA).

5. **Soohyun Lee**, Changho Lee, Soohyun Lee, J Jeremy Chae, Gyeongwoo Cheon, Berk Gonenc, Peter L Gehlbach, Jin U Kang, "Evaluation of optical coherence tomography distal sensor with high-index elliptical cone epoxy lens," 2017 Conference on Lasers and Electro-Optics (CLEO), JTu5A. 3 (2017, San jose, USA).

6. **Soohyun Lee**, Eun-Young Chang, Hyon-Gon Cho and Jinwoong Kim, "Computer-Generated-Hologram for Viewing Window Based Holographic Display and Removal of Conjugate Images," 3DSA 2015 (2015, Taipei, Tiwan).

7. Joongki Park, **Soohyun Lee**, Min Sung Yoon, Jaehan Kim, Hyon-Gon Choo, Jinwoong Kim and Taegeun Kim, "Reconstruction of Optically-scanned Hologram Using Amplitude-only or Phase-only Spatial Light Modulator," 3DTV-CON 2015 (2015, Lisbon, Portugal).

8. Eun-Young Chang, **Soohyun Lee**, Hyon-Gon Choo, Jinwoong Kim, "Computer-generated Stereo Hologram Using Inter-view Redundancy," 3DSA 2015 (2015, Taipei, Taiwan).

9. **Soohyun Lee**, Jeho Nam, Eun-Young Jang, Sung-Keun Lee, Kyungae Moon, Jinwoong Kim, "Measurement of depth representation using integral imaging for quality evaluation of computer-generated hologram," Three-Dimensional Imaging, Visualization, and Display 2014, 911706 (2014, Baltimore, USA).

10. **Soohyun Lee**, Hyon-Gon Choo, Kyungae Moon, and Jinwoong Kim, "Floating image display system combined with eye-tracking for simulating observation of holographic display," 3DSA 2014 (2014, Seoul, Korea)

11. Seunghyup Yoo, Donggeon Han, Hoyeon Kim, Sooyeon Lim, **Soohyun Lee,** "Metal-based transparent electrodes for versatile organic photovoltaics," 2011 Spring European Material Research Society Meeting (2010, Nice, France)