400 kHz Spectral Domain Optical Coherence Tomography for Corneal Imaging

by

Lin Kun Chen

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Science
in
Physics

Waterloo, Ontario, Canada, 2021

AUTHOR'S DECLARATION

This thesis contains material all of which I authored or co-authored: see the Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contribution

The new system is heritage from the 34 kHz system which was developed by Dr. Bizheva and her group before I joined the group.

Le Han helped in explaining the pre-processing of the raw data and logics of the control software developed in the lab that is used with the old 34 kHz UHR-OCT system. He also helped me extensively for the characterization of the system, measuring Point Spread Function (PSF), imaging the resolution target etc, data collected were presented in Chapter 3. Le Han and Dr. Bizheva helped me with system tuning prior the imaging session as well as imaging the samples presented in Chapter 3.4 result section.

Dr. Bizheva provides extremely valuable opinions and feedbacks when proofreading the thesis.

Abstract

The cornea is the transparent, outermost layer of the human eye that contributes approximately 70% of the refractive power of the eye in air. It is composed of five major tissue layers: the epithelium, the Bowman's membrane, the stroma, the Descemet's membrane, and the endothelium. Corneal diseases such as Keratoconus and Fuchs' dystrophy can change the morphology of some or all of the corneal layers, which can lead to vision impairment and eventually blindness. For example, Keratoconus causes localizes thinning and thickening of the corneal epithelium, damage to the collagen structure of the corneal stroma (scarring) and alteration of the corneal curvature. All of these changes result in blurred and double vision, and in severe cases can lead to corneal blindness that would require corneal replacement surgery. Fuchs' dystrophy is a genetic disease that damages the endothelium cell. Since the endothelial cells are responsible for maintaining the fluid level in the stroma, impairment or death of the endothelial cells leads to dehydration or edema of the cornea that results in partial or full corneal blindness. Systemic diseases such as diabetes also affect the physiology and morphology of the cornea. Diabetes affects all the corneal cells and leads to abnormalities such as neuropathy, keratopathy, stromal edema, decrease in endothelial cell density, low tear secretion etc. Although there have been many clinical studies of these diseases, knowledge of the early-stage changes in the corneal morphology at the cellular level remains unclear. Understanding the early stage of disease development with the help of high speed and ultra-high resolution optical coherence tomography (UHR-OCT) corneal imaging can improve the early diagnostics of corneal diseases and well as monitoring the effectiveness of different therapies such as surgical intervention or administration of pharmaceutical drugs

The main objectives of my research project were: a) to upgrade the 34 kHz OCT system with a new camera that offered a 400 kHz data acquisition rate and 8192-pixel linear array sensor, b) test the performance of the 400 kHz OCT system for ex-vivo and in-vivo corneal imaging, and c) develop pre-processing for the interferogram and post-processing algorithms for the images. Implementing a camera with a faster acquisition rate will help to reduce the motion artifact caused by involuntary eye motions. Also, compared to 4500 pixels used in the 34 kHz camera, the new system utilizes all the pixels, resulting in a larger scanning range. Although new camera has smaller sensor size (30% smaller), vertical binning is applied to ensure the light signal is all captured. However, due to the faster acquisition rate (~11 times faster), about 10 dB of SNR will suffers from the reduced integration time. Doubling the sample arm power while keep all other conditions the same can boost the SNR by about 3 dB. Therefore, incident power at the sample arm will be raised carefully according to the maximum permissible exposure calculated using the

American National Standard for Ophthalmics – Light Hazard Protection for Ophthalmics instruments provided by ANSI.

The result from the technical tests shows that the 400 kHz SD-OCT system offers 1 µm axial resolution in biological tissue with an extended scanning range of 2.8 mm (compared to 1.2 mm of the 34 kHz system). It has a lateral resolution of 1.04 µm/pix and can resolve group 7 element 6 of the USAF target with a 20x objective. It can provide 83 dB SNR with 0.95 mW of incident power at a 400 kHz image acquisition rate which should be sufficient to image semi-transparent biological tissues such as the human retina and cornea. So far, the performance of the 400 kHz OCT system has been tested by imaging plant tissues (cucumber) and ex-vivo pig corneas, due to the cancellation of all in-vivo human and animal studies imposed by COVID-19.

Acknowledgements

Table of Contents

List of Figures

List of Tables

List of Abbreviations

Adam	Adaptive Moment estimation
AdaGrad	Adaptive Gradient Algorithm
AMD	Age-related Macular Degeneration
ANSI	American National Standard Institute
ASL	Acellular Stromal Layer
BM	Bowman's Membrane
CCD	Charge-Coupled Device
CL	Collimating Lens
CMOS	Complementary Metal-Oxide Semiconductors
CNV	Choroidal Neovascularization
dB	Decibels
DC	Dispersion Compensation
DFT	Discrete Fourier Transform
DM	Descemet's Membrane
DME	Diabetic Macular Edema
END	Endothelium
EPI	Epithelium
FC	Fiber Coupler
FCL	Fully Connected Layer
FEyeM	Fixational Eye Movement
FFT	Fast Fourier Transform
FL	Focusing Lens
FOV	Field of View
FWHM	Full Width Half Maximum
HT	Hilbert Transform
iff	if and only if
KC	Keratoconus
MIP	Maximum Intensity Projection
MO	Microscope Objective
MPE	Maximum Permissible Exposure
NA	Numerical Aperture

NDF	Neutral Density Filter
NI	National Instruments
NIR	Near Infrared
OCT	Optical Coherence Tomography
OPD	Optical Path Difference
PC	Polarization Controller
PCIe	Peripheral Component Interconnect Express
PNG	Portable Network Graphics
PSF	Point Spread Function
QE	Quantum Efficiency
RGB	Red Green Blue
RMSProp	Root Mean Square Propagation
SD	Standard Deviation
SD-OCT	Spectral Domain Optical Coherence Tomography
STR	Stroma
TD-OCT	Time Domain Optical Coherence Tomography
TIFF	Tag Image File Format
TS	Translation Stage
TTL	Transistor-Transistor Logic
UHR-OCT	Ultra-High Resolution Optical Coherence Tomography
UI	User Interface
USAF	United State Air Force
VGG	Visual Geometry Group
VPHG	Volumetric Phase Holographic Grating

Chapter 1
Introduction

Human perceives the environment with five senses, sight, hearing, smell, touch, and taste. Visual impairment reduces quality of life significantly since approximately 83.0 percent of the sensory information obtained from the outside world are coming from sight (Rosenblum, 2011). The cornea is the transparent outer most part of the human eye. Due to its unique shape and morphology, it contributes approximates two-thirds of the eye's total refractive power (Patel & Tutchenko, 2019). Corneal diseases like as keratoconus will gradually change the corneal shape from prolate ellipse to conical (Rabinowitz, 1998) by localized thickening and thinning in the corneal epithelium layer and damage the collagen structure at the corneal stroma. Therefore, shifting the focal plane away from the retina, causes blurry and double vision. Other diseases affect the corneal shape by dystrophy, damaging the endothelial cells. Low in endothelial cell density will also have significant effect on visual acuity (Wilson, Lin, Klyce, Reidy, & Insler, 1990) (Wacker, McLaren, & Patel, 2015). Diabetes will have effect on all the corneal cells and causing abnormalities. These common diseases are significant affecting patient's lifestyle yet the changes in the early-stage corneal cell morphology remains unclear. While there are plenty of corneal imaging modalities available, none can do cellular level, non-contact, in-vivo imaging of the corneal structure. This motivates the thesis project, to upgrade the existing 34 kHz UHR-OCT developed in 2016 by professor Kostadinka Bizheva with a 400 kHz camera for faster acquisition and achieve motion artefact-free, non-contact, in-vivo imaging for corneal cell morphology study.

Currently, multiple imaging modalities are used clinically for eye diagnosis and research. Slit lamp examination of the corneal is currently the most popular clinical method for initial screening and diagnosis of corneal diseases. The slit lamp is an optical microscope with a slit of bright light that shines onto the corneal surface for the ophthalmologist to visually check for corneal abnormalities. However, the slit lamp microscope has very poor spatial resolution both in axial and lateral direction and does not allow for imaging of the individual corneal layers or cells. Ultrasound biomicroscopes, which measure the echo of the acoustic reflections occurring due to acoustic impedance mismatch at different interfaces, has been a popular tool for eye imaging in the past. It requires physical contact of the probe and the patient's cornea and would need to apply the gel for impedance matching. A commercial device like ArcScan Insight® 100 can achieve 35 μm axial resolution and 65 μm lateral resolution, with a 2 Hz scanning rate. The low spatial resolution, prolonged imaging time and requirement of physical contact with the patient make it not suitable for cellular level in-vivo corneal imaging. Confocal

microscopy (CM) is another optical imaging modality for corneal diagnosis. Commercial CM such as the ConfoScan 4 from Nidek Technologies can provide lateral imaging resolution of 1.2 μm/pix for non-contact imaging with 20x objective and 0.6 μm/pix for immersive gel imaging with 40x objective. However, CM has limited axial resolution (in the order of several tens of micrometers), limited field-of-view (FOV), typically < 400 μm × 400 μm, and has fairly long 3D image acquisition time due to slower scanning in the axial direction.

Optical coherence tomography (OCT) is an optical imaging modality, based on interferometric principle, that can provide non-contact, non-invasive, in-vivo images of ocular tissues with very high spatial resolution (in the order of 1 μm, sufficient for visualization of individual corneal cells) and scanning range in the order of millimeters. However, the main remaining challenge for using OCT for in-vivo imaging of the corneal cellular structure is involuntary eye motion such as drift, tremor and microsaccades, that causes motion artefacts in the OCT images. In 2016, our research group designed a 34 kHz A-scan rate UHR-OCT system based on supercontinuum laser. The broad spectrum provided by that laser allowed the OCT system to achieve 0.95 μm axial resolution in the corneal tissue, 2 μm lateral resolution with 10x objective, and 95 dB sensitivity for 680 μW incident power. However, the camera in the 34 kHz OCT system has a relatively slow acquisition rate compared to the involuntary eye motions, which typically have frequencies from a few Hz (microsaccade) up to a hundred Hz (tremor). Therefore, phase noise and motion artifact will present and significantly affect the imaging quality.

1.1 Corneal Structure

Corneas is the of outer most part of the eyeball, together with sclera it forms the outer shell for protection of the liquid and structures inside. Human, and other primates have corneas consist of five layers: corneal epithelium, Bowman's membrane, corneal stroma, Descemet's membrane, and corneal endothelium. Normal human cornea can have a thickness in the range from 551 to 565 microns (Feizi, Jafarinasab, Karimian, Hasanpour, & Masudi, 2014).
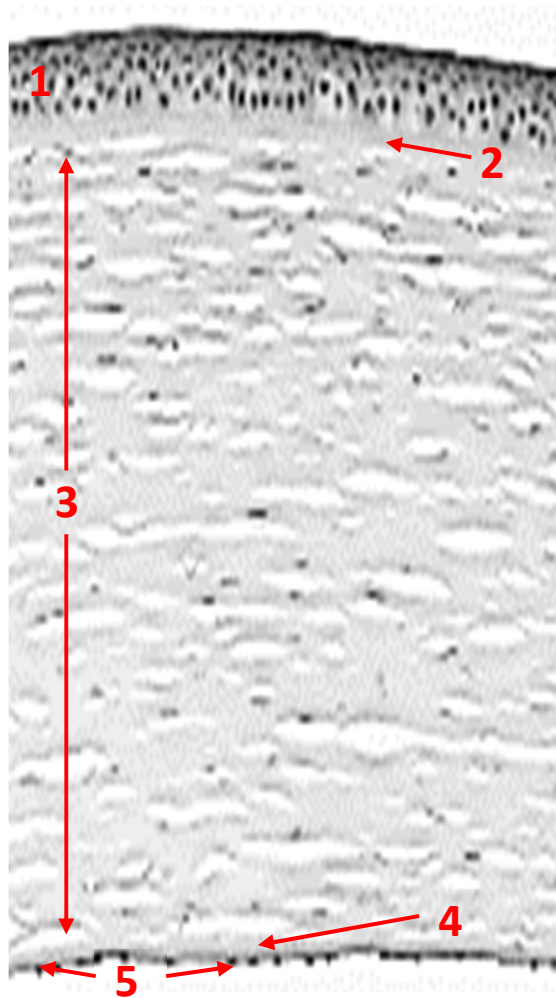
Figure 1.1 Histology of human cornea (Reproduced with permission from Prof. Denise Hileeto)
1 – Corneal Epithelium; 2 – Bowman's Membrane; 3 – Corneal Stroma; 4 – Descemet's
membrane; 5 – Corneal Endothelium

Corneal Epithelium

The corneal epithelium consists of about 6 layers populated with different types of cells (superficial cells, wing cells and basal cells). The corneal tissue is almost transparent for visible light and there are no blood vessels in the cornea to nourish the cells, as these can obstruct the visual field. One of the main functions of the epithelium cell is to absorb oxygens from air and nutrients from the tear to nourish the cells beneath. It also provides barrier to prevent the foreign dust, water and germs getting into the eye. (The Cornea and Corneal Disease, 2017)

Bowman's membrane

Located immediately below the basal cell layer of the corneal epithelium is the Bowman's membrane (#2 in Figure 1.1). It is not a true membrane but composed of strong layered collagens fibers arranged in a single layer to help cornea to maintains its shape (Sridhar, 2018) (The Cornea and Corneal Disease, 2017) The Bowman's membrane has about 17.7 μm thickness in a healthy cornea (Tao, et al., 2011), and will get thinner with aging. Bowman's membrane is not regenerative and will form scars when injured. (Sridhar, 2018)

Corneal Stroma

#3 in Figure 1.1 is the stroma region. Corneal stroma constitutes approximately 90% of the corneal thickness. It consists of about 78% of water, 16% of collagens. Keratocyte cells in the corneal stroma help to maintain the stromal homeostasis by synthesizing the collagen molecules and glycosaminoglycans (Zhang, et al., 2015) necessary for the regeneration and repair of the corneal stroma after injury or trauma. The strong collagen structure of the stroma provides mechanical strength of the cornea, while the ordered arrangement of the collagen fibrils allows for the cornea to appear almost transparent to visible light. The typical thickness of the healthy, normal corneal stroma is about 500 μm. (Reinstein, Archer, Gobbe, Silverman, & Coleman, 2009)

Descemet's membrane

The Descemet's membrane is under the stroma region (#4 in Figure 1.1) It is a strong sheet of collagen fibres and has thickness of 10 μm (Bizheva, et al., 2016). The collagen in Descemet's membrane is different from the collagens in the stroma. They are made by the anchored endothelial cells right beneath it therefore is regenerative and can be healed after injury. It acts as another protection layer against infection and injuries. It allows the nutrient traffic into the corneal stroma and is critical to the corneal hydration and transparency (Eghrari, Riazuddin, & Gottsch, 2015) (Lwigale, 2015) (Saikia, Medeiros, Thangavadivel, & Wilson, 2018).

Corneal Endothelium

The endothelium is the most posterior layer of the cornea (#5 in Figure 1.1). It consists of a single layer of endothelial cells with thickness of 4 microns. These endothelial cells hexagonally shaped and ~ 20 μm is diameter in the transverse direction. The endothelial cells are responsible for pumping out the excess fluid from the stroma to keep the stroma from swelling. The endothelial cells are lost forever

if destroyed by diseases or injuries. Insufficient amount of endothelial cells density will cause corneal edema and eventually leads to blindness. Currently, the corneal transplantation of donor tissue is the only available therapy to restore vision (Hertsenberg & Funderburgh, 2015).

1.2 OCT Theory

Most of the OCT technology (clinical or research-grade) is based on Michelson interferometry. Figure 1.2 shows a diagram of a typical SD-OCT system based on a fiber-optic Michelson interferometer. The SD-OCT system is powered by a low coherent light source, where the central wavelength and the spectral bandwidth of the light source determine the OCT axial resolution. Then the light propagates through a fibre coupler, which splits the input beam into a sample arm and reference arm beams. In the sample arm, the beam is projected on the imaged sample target though a series of optical elements. In a similar way, the optical beam in the reference arm is projected onto a reference mirror mounted on a translation stage to allow for adjustment of the reference pathlength. Light from the imaged sample and the reference mirror, propagates back to the beamsplitter and generates an interference pattern there. The interference pattern is then projected on to a linear-array sensor camera though a spectrometer that separates the broadband optical beam into a series of a narrow bandwidth beam. Each of those beams is focused onto individual pixels of the camera. The output signal from the camera is processed with a computer using Fast Fourier Transform (FFT) to convert the spectral signal into depth-depended spatial location in the imaged sample where the reflected signal originated. (Izatt & Choma, 2008) (Aumann, Donner, Fischer, & Müller, 2019)
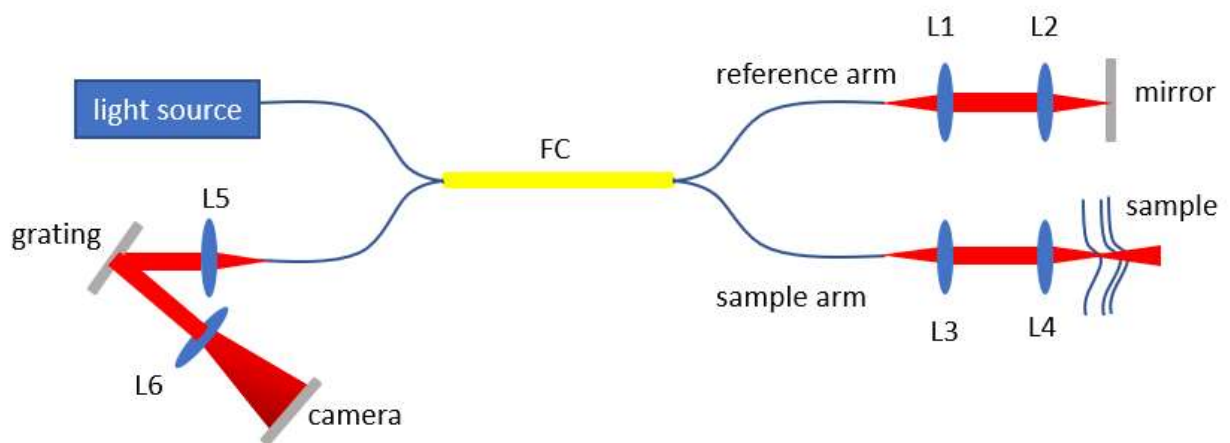
Figure 1.2 fiber-based SD-OCT typical setup

5

A simplified example would be looking at a Michelson interferometer as illustrated in Figure 1.3
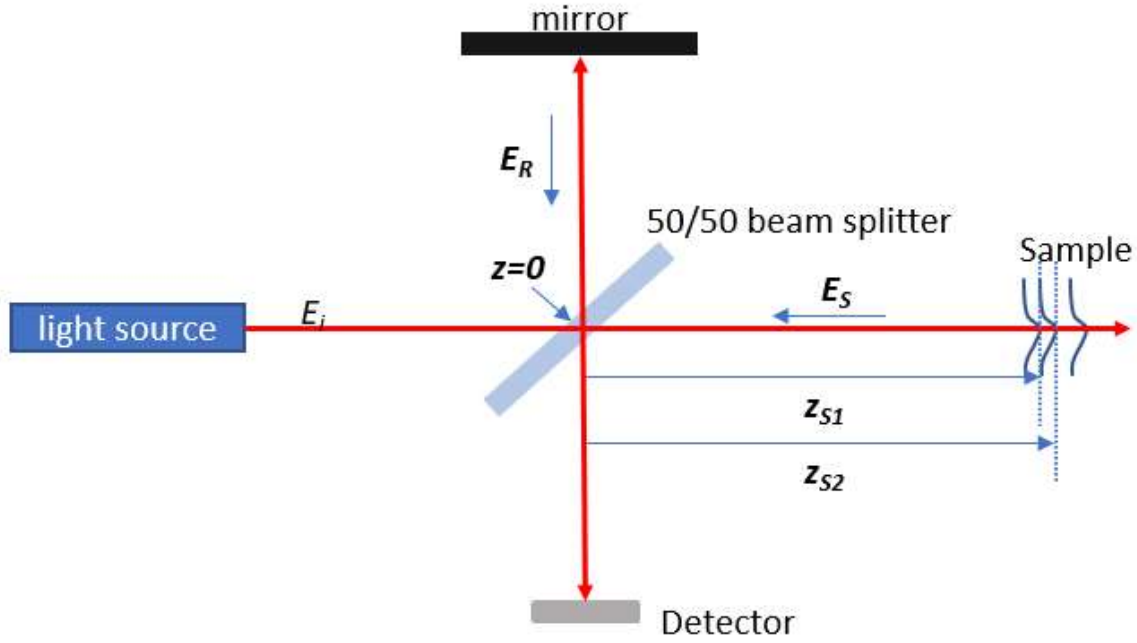
Figure 1.3 Michelson Interferometer setup in OCT

The light source has electric field of $E_i = s(k,\omega)e^{i(kz-\omega t)}$, with $s(k,\omega)$ representing the wave number and angular frequency dependent amplitude. The field travelling the sample arm is then,

$$E_s = \frac{E_i}{\sqrt{2}}\left[r_s(z_s) * e^{i2kz_s}\right] \qquad (1.1)$$

Where $r_s(z_s)$ is electric field reflectivity from sample arm at given depth z_s. the reference arm will have electric field E_R,

$$E_R = \frac{E_i}{\sqrt{2}}\left[r_R * e^{i2kz_R}\right] \qquad (1.2)$$

The intensity detected at camera end is the time averaged dot product of the electric fields from reference arm and sample arm.

$$I_d = \rho\langle(\vec{E}_s + \vec{E}_R)\cdot(\vec{E}_s + \vec{E}_R)\rangle = \rho\langle|E_R + E_S|^2\rangle \qquad (1.3)$$

Where ρ is the responsivity of the detector. Consider a 50/50 beam spliter, the intensity will reduce to half because of the double pass. The full expression of intensity at detector becomes,

$$I_d(k,\omega) = \frac{\rho}{2}\left\langle\left|\frac{s(k,\omega)}{\sqrt{2}}r_R e^{i(2kz_R)} + \frac{s(k,\omega)}{\sqrt{2}}\sum_{n=1}^{N} r_{Sn}e^{i(2kz_s-\omega t)}\right|^2\right\rangle \qquad (1.4)$$

Where $r_{Sn}(z_s) = \sum_{n=1}^{N} r_{Sn}\delta(z_s - z_{Sn})$, assuming the sample have N discrete different layers with electric field reflectivity r_{sn}. Multiply by the direct delta function with offset z_{Sn} away from the beam splitter. Since frequency of the light v oscillates much faster than the detector response, the frequency dependence $\omega t = 2\pi v t$ can be eliminated. The power reflectivity is the square of the electric field reflectivity, $R_s = r_s^2$ and $R_R = r_r^2$, the equation (1.4) will reduce to,

$$\begin{aligned}
I_d(k) = \frac{\rho}{4}S(k)\Bigg[&\left(R_R + \sum_{n=1}^{N} R_{Sn}\right) \\
&+ 2\sum_{n=1}^{N}\sqrt{R_R R_{Sn}}(\cos[2k(z_i - z_R)]) \\
&+ 2\sum_{n \neq m=1}^{N}\sqrt{R_{Sm}R_{Sn}}\,(\cos[2k(z_{Sn} - z_{Sm})])\Bigg]
\end{aligned} \qquad (1.5)$$

There are three terms in the equation:

1. The first term is the path length independent "DC" term which mainly came from direct reflection from the reference mirror as we assume the biological sample has exceedingly small reflectivity compares to the sample arm mirror. This term has the most contribution to the detector current.
2. The second term in the equation is "cross-correlation" term, representing the interferences between multiple reflection from the different depth of biological sample with the reference spectrum. This is the desired signal component in OCT imaging.
3. The third term is called "auto-correlation", it is the interference generated between different layers in the biological. This term is generally small compares to the DC term and cross-correlation term.

Perform Fourier-transform to signal $I_d(k)$ will obtain the $I_d(z)$, which is the depth dependent reflectivity profile,

$$I_d(z) = \frac{\rho}{8}\gamma(z)\left[\left(R_R + \sum_{n=1}^{N} R_{Sn}\right)\right.$$

$$+ 2\sum_{n=1}^{N}\sqrt{R_R R_{Sn}}(\gamma[2(z_R - z_{Sn})] + \gamma[-2(z_R - z_{Sn})]) \tag{1.6}$$

$$\left. + 2\sum_{n \neq m=1}^{N}\sqrt{R_{Sm}R_{Sn}}(\gamma[2(z_{Sn} - z_{Sm})] + \gamma[-2(z_{Sn} - z_{Sm})])\right]$$

Where $\gamma(z)$ is the Fourier transform pair of $S(k)$, the cross-correlation term can be used to extract the depth dependent reflectivity of the sample arm $R_s(z_s)$. Because the DC terms and the auto-correlation terms were located close to zero-delay, the cross-correlation signal can be separated by adjusting the optical path difference (OPD) of the sample with respect to the beam splitter, shifting the signal away from the zero-delay.

Low coherence laser light source is commonly used in OCT technology. Narrow bandwidth highly coherent laser light can maintain the interference pattern over a long distance. Therefore, when the signal received by the detector, it has no depth information because it cannot determine where reflection happen. Low coherence source will create a noticeable fringe pattern when the optical path of the two beams became less than the coherence length of the light source.

1.2.1 Axial Resolution

In digital signal processing, the auto-correlation function is found by taking Fourier-transform of a spectrum signal. Similarly, coherence function $\gamma(z)$ of light source can be found by finding the Fourier-transform of the power spectrum density function $S(k)$. With normalized Gaussian function $S(k)$, its Fourier transform pair is found to be,

$$\gamma(z) = e^{-z^2\Delta k^2} \xleftarrow{\text{F.T. pair}} S(k) = \frac{1}{\Delta k\sqrt{\pi}}e^{-\frac{(k-k_0)^2}{\Delta k^2}} \tag{1.7}$$

Where k_0 is the central wavenumber of the spectrum, and Δk is the half width of the spectrum at $1/e$ of the intensity. The coherence length l_c is defined Full-With-Half-Maximum (FWHM) of the coherence function,

$$l_c = 2z_{1/2} = \frac{2\sqrt{\ln 2}}{\Delta k} = \frac{2\ln 2}{\pi}\cdot\frac{\lambda_0^2}{\Delta\lambda} \tag{1.8}$$

Here λ_0 is the central wavelength defined by $2\pi/k_0$, and $\Delta\lambda$ is the FWHM in the wavelength spectrum (hence, $\Delta k = \frac{\pi}{\sqrt{\ln 2}}\frac{\Delta\lambda}{\lambda_0^2}$). The coherence function is the PSF of the signal, and the coherence length is then the axial resolution of the OCT system.

1.2.2 Digital Scanning Range

An A-scan is one interferogram and the result is the depth profile the axial resolution is determined by the light source property. However, the scanning range is limited by the maximum fringe frequencies the detector is capable to capture because the maximum frequency of the fringe pattern in equation (1.4) contains the reflectivity of the maximum depth. The image depth is defined as,

$$Z_{max} = \frac{1}{4}\frac{\lambda_0{}^2}{\delta\lambda} \tag{1.9}$$

Here the $\delta\lambda$ is the spectral resolution of the detector. Maximum scanning range divide by half number of pixels in the detector will determine the digital axial resolution, with the units of micrometer per pixel.

1.2.3 Lateral Resolution of OCT

The lateral resolution of the OCT system is determined by the FWHM of the spot size of the probing beam. Taking $w_0 = \frac{2\lambda_0}{\pi}\frac{f}{d}$ as the beam waist (defines where the intensity of the beam drops to $1/e^2$), the theoretical limit for lateral resolution is,

$$\Delta x = \sqrt{2\ln 2}\, w_0 = \sqrt{2\ln 2}\,\frac{\lambda_0}{\pi NA} = 0.37\frac{\lambda_0}{NA} \tag{1.10}$$

where λ_0 is the center wavelength, and $NA = n\sin\theta = n\sin\left(\tan^{-1}\left(\frac{D}{2f}\right)\right) \approx n\frac{d}{2f}$. Here d is the size of the incident beam on the objective lens, and f is the focal length.

The depth of focus (DOF) b, also called focal parameters is square proportional to the lateral resolution of the system. It is defined as twice the Rayleigh range:

$$b = 2\cdot Z_R = \frac{2\pi n w_0{}^2}{\lambda} = \frac{n\lambda_0}{2\pi\, NA^2} \tag{1.11}$$

While having higher NA will grant the system better lateral resolution, lower NA setup will provide larger depth of focus with poorer but relatively uniform lateral resolution.
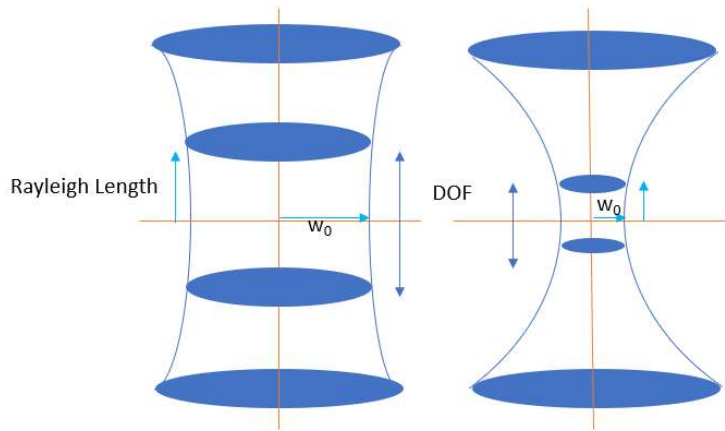
Figure 1.4 Examples Illustrating effect of having lower NA (left) and Larger NA (right)

The lateral field of view depends purely on the objective and the scanning beam incident angle as illustrated in the Figure 1.5, where $FOV_{Lateral} = 2f\theta_{Max}$
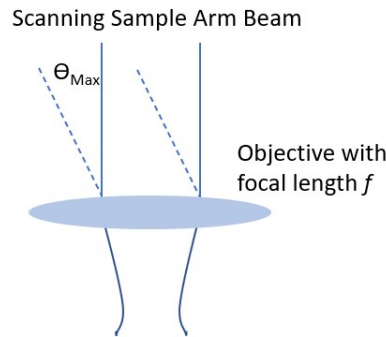
Figure 1.5 Lateral Field of View

1.2.4 Signal Noise Ratio and SNR roll off

The signal noise ratio (SNR) of the system is measured in decibels (dB) and is defined as the ratio of the OCT power value to standard deviation (SD) of the background power. OCT system with 100dB corresponds to the ability to detect the sample reflectivity (R) as low as R=10^{-10}. SD-OCT typically has 20dB and more sensitivity compares to TD-OCT because it utilizes N pixel detector where TD-OCT normally uses single pixel detector. This theoretically grants SNR of the SD-OCT to be N/2 times larger than SNR of the TD-OCT.

In a shot noise limited SD-OCT system, the photon distribution is said to obey Poisson distribution (Choma, Sarunic, Yang, & Izatt, 2003) and the theoretical sensitivity can be approximated to

$$Sensitivity = \frac{\rho PT}{2e} \tag{1.12}$$

Where ρ is the photodetector responsivity of the detector, P is the optical power illuminated at the detector assumed to be perfectly reflected from sample arm mirror, T is the integration time for the detector and e is the electron charge. SNR at zero delay can be then determined by

$$SNR = 10\log_{10}(Sensitivity) \tag{1.13}$$

The SNR will decrease as function of depth. In SD-OCT, reflectivity profile is caught by the interferogram. A perfect measurement of the interferogram requires that each pixel measures a single and exact wavelength. The inability of measuring higher frequencies due to finite pixel size (Leitgeb, Hitzenberger, & Fercher, 2003) and finite spectral resolution of the spectrometer (Yun, Tearney, Bouma, Park, & de Boer, 2003), causes sensitivity roll-off. Crosstalk between neighbour pixels can also contributes to the roll-off effect.

1.3 Brief History of OCT and Current Status

Optical coherence tomography (OCT) is a non-invasive imaging modality that measures the interference of back scattered light to analyze the depth encoded reflectivity profile of the sample. It was first introduced as an imaging modality with potential for in-vivo non-invasive ophthalmic imaging back in 1991 (Huang, et al., 1991). This publication showed the first in-vivo Time Domain OCT (TD-OCT) cross-sectional image (B-scan) of the human retina. As TD-OCT was based on mechanical scanning of the reference mirror, this resulted in low scanning rates in the order of 1 kHz or less. The first spectral domain OCT (SD-OCT) retinal images were presented by Wojtkoski et al. in 2002 (Wojtkowski, et al., 2002). The SD-OCT utilizes a spectrometer to disperse the broadband spectrum on to a photon detection device, typically an array of charge-coupled device (CCD) or a complementary metal-oxide semiconductors (CMOS) sensor. This type of setup increases the imaging process significantly because lack of moving part at reference arm for scanning and the depth reflectivity profile is decoded by the spectrometer and the sensors. Since then, a lot of effort was made in improving the SD-OCT system in image acquisition speed and axial resolution. Major development of OCT focuses on upgrading the light source, faster camera, as well as designing new optics and mechanics to be adaptive with respect to the new light source and detection. Because that the OCT axial resolution is decoupled with the lateral resolution, by choosing a source light with wider spectrum, micron level and even submicron axial resolution can be achieved (Yadav, et al., 2011) (Werkmeister, et al., 2013) (Bizheva, et al., 2016). Since eye motion will cause fringe averaging, OCT imaging requires high

11

acquisition speed to reduce motion blur caused by the eye motion without extra eye motion tracking hardware. Cellular level resolution was achieved for in-vivo imaging of both animals (Yao, et al., 2019) (Chen, et al., 2017) and humans (Tan, et al., 2018) corneal endothelium.

1.4 Eye motion

The SD-OCT technologies were only suitable for B-scan images if the acquisition speed cannot overcome the disruption occurred on the volumetric data and enface image that is caused by the eye motion. The slow speed of the commercial OCT system will result not only smaller FOV, but also preserves the eye motion in the image data. Commercial system acquisition speed of 20 kHz up to 100 kHz often combine the scanning laser ophthalmoscope (SLO) imaging modality to help correct the motion between the OCT image slices. Even with the fixation stage for corneal or retinal imaging, fixational eye movement (FEyeM) can still occur. FEyeM tracking and objective measurement started during late 19th century where EB Heuy first reported the eye movement of the during reading with steady fixation (Huey, 1900). later research confirmed the FEyeM and categorized it into three different types of movements: drift, tremor and micro saccades. Drift is a tiny random walk like eye movement. It has low velocity, and the fixated object can move tens of photoreceptors in the retina during the drift. Tremor is irregular, wave-like eye motion occurred at the same time as drift. The average frequency of 84 Hz from 105 normal participants were measured in a recent article (Bolger, Bojanic, Sheahan, Coakley, & Malone, 1999). Micro saccades are jerk-like, fast eye movement occurred in between drifts. It has duration about 25 milliseconds and can shift the retinal images across hundreds of photoreceptor width (Martinez-Conde, Macknik, & Hubel, 2004). The micro saccades can distort the OCT volumetric and enface data significantly since typical lateral field of view can be small as few hundreds of photoreceptors when imaging the retina.

There are another two main source of eye motion during the OCT imaging. Since the OCT has micron level resolution, involuntary head movement, even heartbeat and the respiration during imaging will also introduce shifts of the retina position, causing axial motion artifact (de Kinkelder, et al., 2011). There are hardware and software approach for reducing the artefact caused by the eye motion. One is applying eye motion tracking hardware with OCT system, so the eye motion is collected simultaneously while scanning. Software based approach approximates the eye motion by referencing a base image and perform image registration on the dataset. Hardware solution usually have superior result since it has approximately exact information about eye motion for correction. Software based methods are still capable of producing moderate result with no extra cost, but their performance is highly dependent on

the OCT data qualities. If the data were significantly distorted by the fast eye motion like micro saccades, the algorithms cannot help with reconstruct the data. A camera with faster acquisition rate will also help to reduce eye motion related artifacts in the OCT images, as well as allow for expanding the FOV while sampling it with higher transverse resolution.

Chapter 2
400 kHz UHR-OCT

To achieve corneal imaging at cellular level, the SD-OCT system requires to have fast temporal resolution as well as micron level spatial resolution. Integrating the fast camera will grant a high temporal resolution thus reduce the artifacts caused by eye motion. Because the axial and the lateral resolutions are decoupled, a broadband light source will provide a better axial resolution. Although the light source with longer wavelength in the near infrared (NIR) can have better penetration depth, shorter wavelength will provide a higher axial OCT resolution because of the squared relation between center wavelength of the source spectrum and the axial resolution described in equation (1.8). As the axial OCT resolution is related to the OCT scanning range, increasing the axial OCT resolution while keeping the number of camera pixels the same will result in decreased OCT scanning range. Since the water has low light absorption at 800nm, local minima in the absorption spectrum at 1060nm and 1300nm and the infra-red-light source is less visible or invisible to human eye. Historically, OCT systems operating in the 800 nm, 1060nm and 1300 nm central wavelength have been used for in-vivo imaging of the human cornea, each of those spectral regions providing a combination of spectral resolution and scanning range. (Leitgeb & Baumann, 2018). For an OCT system with high axial resolution to be able to image individual cells in the corneal tissue, a similarly high transverse resolution, in the order of ~ 1um, is required. Since the OCT transverse resolution is determined by the optics of the OCT imaging probe, using high NA objective will ensure the required transverse resolution, though at the expense of shorter depth-of-focus and limited FOV. Furthermore, when deciding the spot size of the beam illuminating on the imaged target, maximum permissible exposure (MPE) also needs to be carefully considered and calculated for the system before safely imaging the live sample since concentrating massive energy to a fine spot will potentially damage the tissue.

Corneal diseases such as keratoconus, Fuchs' dystrophy etc, can cause corneal tissue damage and eventually leads to permanent corneal blindness. Right now, the only treatment right now is to perform corneal transplant. The ability to perform in-vivo imaging on the corneal structure and see the tissue damage at cellular level over a large FOV, can help to study and diagnosis the early-stage corneal diseases and advance the development of the more effective treatment. Due to involuntary eye motion, acquisition of volumetric images of the human cornea at cellular level resolution requires OCT technology that combines high axial resolution and high image acquisition rates.

Previously Bizheva's research group introduced 34 kHz UHR-OCT system that has the cellular resolution of the corneal tissue show in Figure 2.1. The system has sufficient axial resolution to identify the corneal layers and can even resolve corneal cells. However, due to the slow acquisition speed, it is not possible to acquire the motion artifact free in-vivo 3D volumetric data of human cornea. This thesis introduces the upgraded UHT-OCT system with 400 kHz camera that can eliminate the disadvantage previous generation.
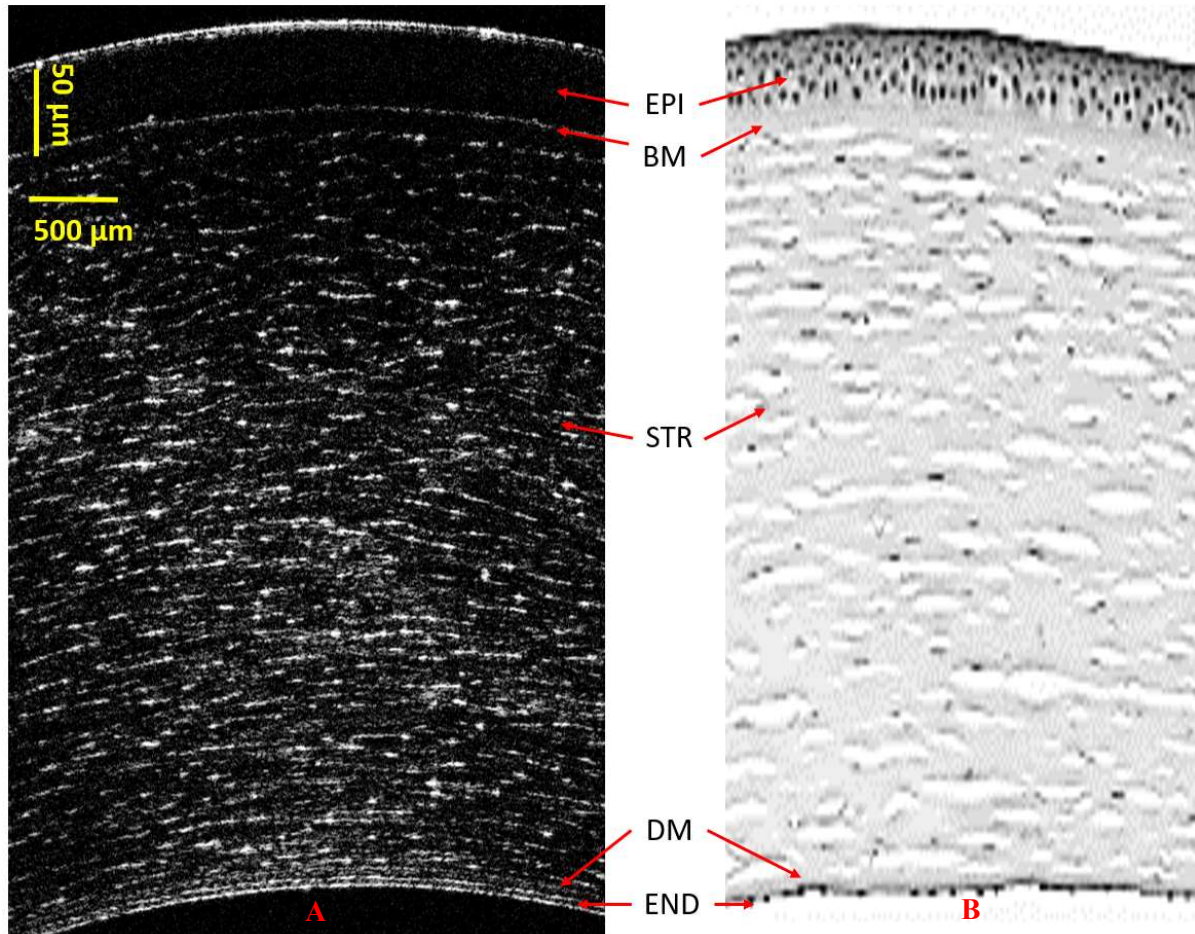

Figure 2.1 In-vivo imaging of human corneal tissue with the 34kHz system (A), a healthy human cornea histology image (B), EPI – epithelium; BM – Bowman's membrane; STR – stroma; ASL – acellular stromal layer; DM – Descemet's membrane; END – endothelium.

2.1 Maximum permissible exposure calculation for corneal imaging

American National Standard Institute (ANSI) provides the American National Standard for Ophthalmics – Light Hazard Protection for Ophthalmics instruments (ANSI Z80.36-2016). It classifies the ophthalmic instruments into two groups and defines the exposure limits.

The imaging probe of the 400 kHz UHR-OCT system focuses the imaging beam to spot of 10 µm diameter on the cornea with the 5x objective. The spot is continuously scanned over the cornea; therefore, it stays at one location on the corneal surface only for a duration of 10/400 kHz = 2.5×10^{-5} seconds, taking a maximum 90% overlapping illuminating area into consideration. The spectrum of the sample arm is presented in Figure 3.1. According to ANSI Z80.36-2016, for a light source with spectral range between 380 nm and 1200 nm, the limit value of the unweighted anterior segment visible and infrared radiation exposure, ($H_{\text{VIR-AS}}$) is defined as: $25\ t^{1/4}$ J/cm^2, where t is the exposure time.

The MPE is defined as $H_{\text{VIR-AS}}$ divide by time and then multiply the area of exposure,

$$MPE_{\text{VIR-AS}} = (25t^{-3/4}\ W/cm^2\ ) \times Exposure\ Area \qquad \textbf{(2.1)}$$

Normally human eye will blink at least every 10 seconds and eyelid will naturally block out the light. for safety reasons, t is set to 10 seconds in this calculation. According to the ANSI standards, it is averaging the radiant power within a circular area having a diameter of 1.0 mm, the MPE is then determined to be

$$MPE_{\text{VIR-AS}} = \left(25 \times 10^{-3/4} W/cm^2\right) \times \pi \left(\frac{0.10\ cm}{2}\right)^2 = 0.0349\ W \approx 35\ mW$$

For our corneal imaging studies, optical power of the imaging beam at the corneal surface will be adjusted to 2.75 mW which is lower than the limit value calculated under the ANSI guideline for anterior segment imaging.

During corneal imaging, since the optical beam is focused on the cornea, it will produce a large diffuse illumination of the retina, which will pose no safety concerns in term of retinal damage, since the optical imaging power is distributed over a large area.

2.2 Methods

2.2.1 Hardware

The 34kHz SD-OCT system is based on the fiber-optics Michelson Interferometer. A supercontinuum laser (SuperK, NKT Photonics, Denmark) is used with a custom filter unit (Figure 2.2, CF). The custom filter unit contains multiple bandpass filters (BPF) and neutral density filters (NDF). Since the laser source have extremely wide spectrum, these filters are chosen to filter out unwanted band and only leave a spectrum centered around 800 nm with a spectral bandwidth about 390 nm. The NDFs are used to reduce the optical power down to around 20mW before coupling into the 50/50 broadband fiber coupler (Gould Fiberoptics, USA).

The reference arm of the system comprised of a collimator, a neutral density filter, a beam expander, a focusing lens and a silver mirror. A two BK7 prism mounted on miniature translation stages are used form the custom hardware dispersion compensator (Figure 2.2, DC). Manual control of the overlap between the prism can vary the thickness the glass present in the optical path. Other than the collimator, all the other components are mounted on a 12 inches manual translation stage, as indicated in Figure 2.2 with the red dashed rectangle.

The sample arm consists of a collimator, a x-y galvanometric scanner, beam expander and a microscope objective (5x, 10x or 20x depending on the required trade-off between lateral resolution and FOV). The new OCT system utilizes a customized commercially available spectrometer (Cobra, Wasatch Photonics, USA) which is equipped with 8192-pixels CMOS camera which can perform acquisition at 400kHz rate. Together it provides a spectral resolution of about 0.0476nm. the polarization controller is used to perform manual spectrum reshaping so the shape and magnitude of system's point spread function (PSF can be optimized.

For in-vivo imaging of the human cornea, the optical power of the imaging beam incident on the corneal surface will be set to <2.75mW, which is significantly lower than the maximum permissible exposure power as specified by the ANSI (American National Standards Institute) standard.
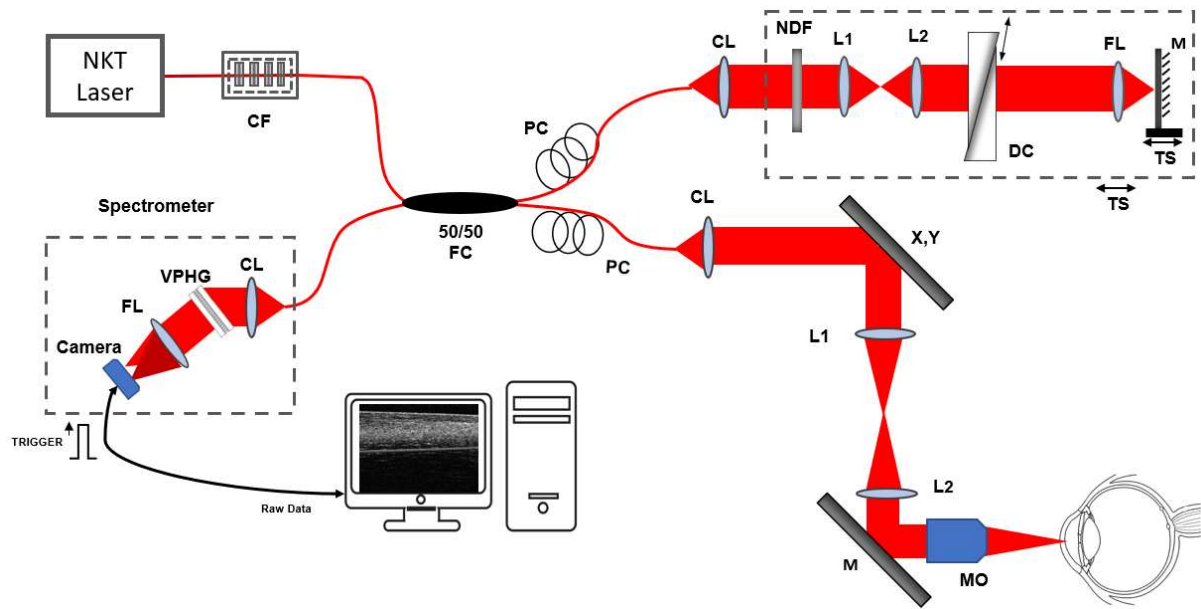
Figure 2.2 Schematic of the OCT system. CCD-camera; CF-custom filter; CL-collimator; DC - dispersion compensation unit; FC-fiber coupler; FL-focusing lens; L1 and L2-lenses; M-silver mirror; MO-microscope objective; NDF-neutral density filter; PC-polarization controller; TS-translation stage; VPHG-volumetric phase holographic grating; XY-galvanometric scanners.

2.2.2 Acquisition System Setup

CamExpert from Teledyne DALSA company was used for high-speed image acquisition, and LabVIEW was used for scanner control as well as camera trigger signal generation with NI PCIe-6321 Multifunctional I/O device. Figure 2.3 shows the system layout. The trigger signals use BNC connectors. The camera connection to the Xtium2 frame grabber uses camera link cable to allow the high data throughput required for the raw data transfer.
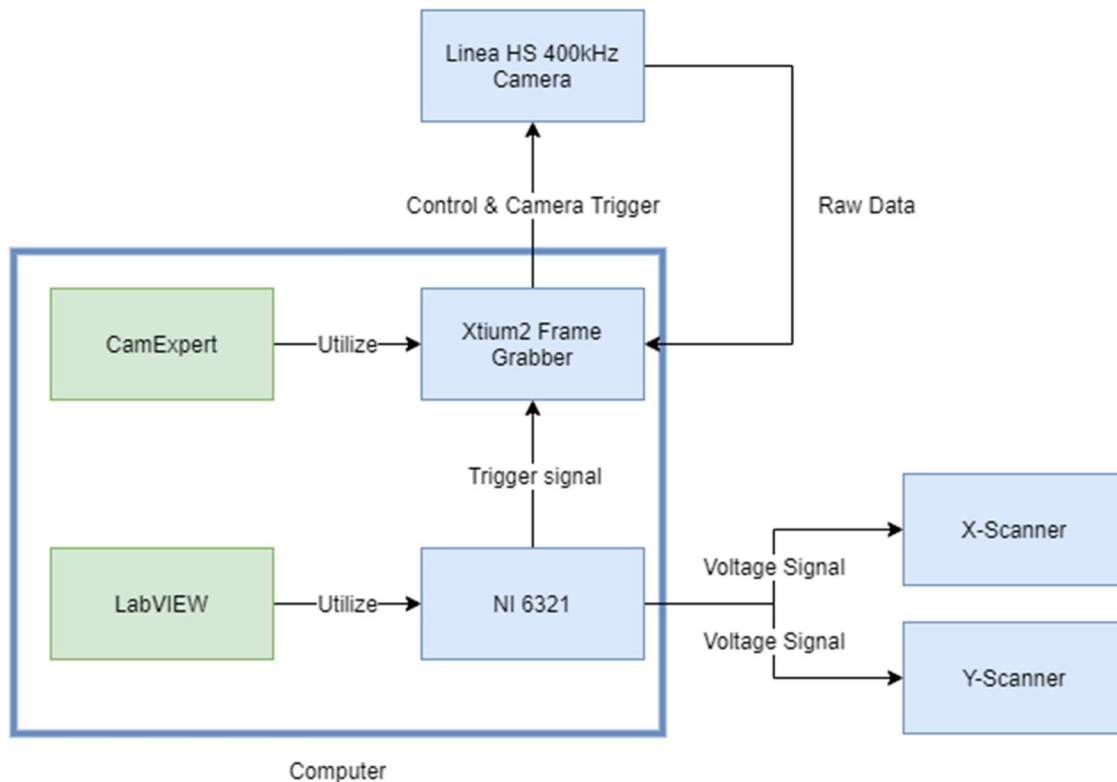
Figure 2.3 System Layout, blue indicates hardware system, green is the software. Xtium2 Frame grabber and NI PCIe-6321is installed on the computer; one computer runs both CamExpert and LabVIEW software

Scanning Pattern

Detailed schematic for ex-vivo imaging setup is show in Figure 2.3. The x-y scanners were controlled by LabVIEW and NI PCIe-6321 devices, and raster scan patter were implemented. 80% duty cycle were used for one b-scan, the rest 20% is saved for resetting of the x-scan mirror to starting position. For example, by setting up a squared imaging area, 500 by 500 points in x and y direction will result 400 in x direction (fast scan) and 500 in y direction (slow scan). Although this introduces an uneven distribution (0.8 to 1 for fast and slow scans) of the scanning points. It can be adjusted to have a 1 to 1 by either having close to 100% duty cycle or use 625 by 500 raster pattern. However, maximizing the duty cycle will potentially harm the x-scanner since it allows less time for it to scan through the length of the FOV. Increasing the raster scan points will consume 20% more computer memory.
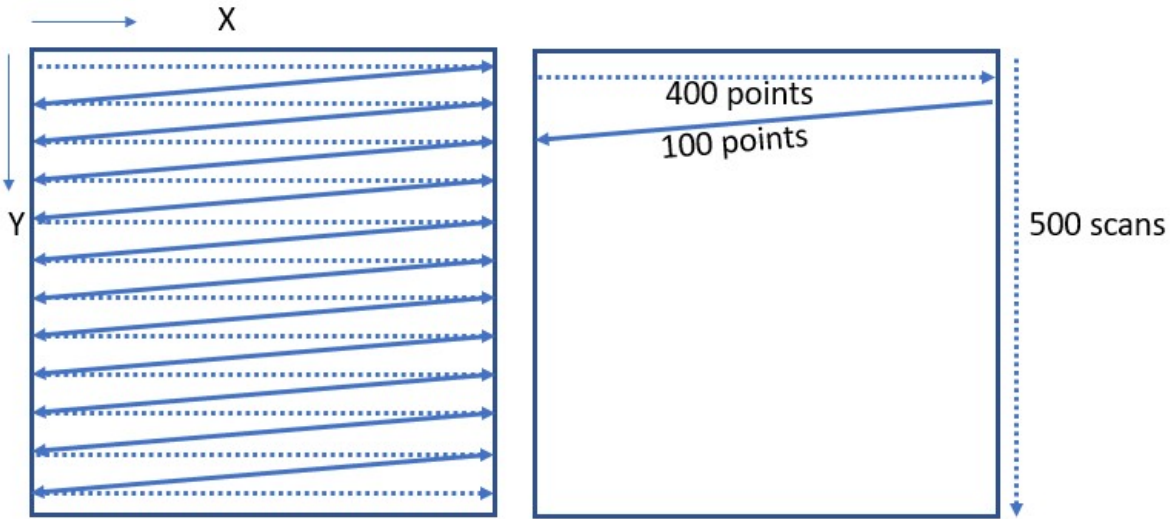
Figure 2.4 Illustration of raster scan used, dotted line representing the one b-scan, solid line representing fly-back of the x-scan mirror (left); example of one b scan and point distribution with 80% duty cycle (right).

Data Throughput

The camera was set to acquire 8192 pixels of data per scan with 8 bits pixel depth although only about 7570 pixels were illuminated on. Setting the camera to operate at full speed, the system will generate extensive amount of data,

$$data\ size = 8192\ pixels/line \times 1\ Byte/pixel \times 400000\ lines/sec \approx 3.05\ GB/sec$$

With such huge amount of data needs to be transferred and stored, the old frame grabber from National Instrument (NI) used for 34 kHz UHR-OCT system is no longer an option. NI does not have anything up to the speed and commercially available yet. Xtium2-CLHS frame grabber from Teledyne DALSA is the only available frame grabber that satisfies the compatibility with the 400 kHz camera and the high data throughput requirement.

Trigger Signals

The Xtium2-CLHS frame grabber is set to have internal trigger of 400kHz while the NI PCIe-6321 is programmed to generate a TTL trigger signal long enough for the scanning process. Figure 2.5 shows an example of generating camera trigger. The Xtium2-CLHS frame grabber takes the trigger signal from the NI PCIe-6321, "AND" with the internal 400 kHz clock (outputs high iff both signals are high), it reshapes the input and generates camera trigger signal.

Figure 2.5 Example of trigger timing diagram

When taking 500 by 500 area scan, a total of 250,000 scans were required. This will require a trigger signal from NI PCIe-6321 to last 250,000 / 400,000 = 0.625 seconds.

Image acquisition settings

In the CamExpert image settings, the image was set to acquire a 2D image, with size of (number of A-scans times number of B-scans) by 8192 pixels. When acquiring a square area of 400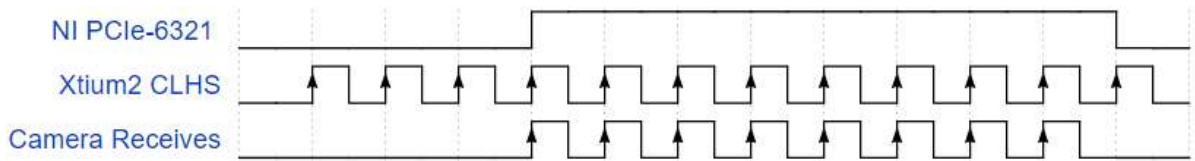 by 500 raster scans, with duty cycle of 80%, a total of 400 ÷ 0.8 × 500 =250,000 lines scans were required. The camera will generate a 2D image of 250,000 by 8192 8-bit TIFF grey image. A total of 2048000000 pixels were contained in the image with 1 byte per pixel, occupying approximately 1.91 GB in the computer hard drive. TIFF is non-compressed data format and therefore preserves all the sensor reading details with no compression loss.

2.2.3 MATLAB GUI for Pre-processing of the raw Data

The stored TIFF file has the 8-bit gray scale info about the interferogram of the scans. MATLAB is used for preprocessing of the raw data to generate the B-scan images and enface images. A MATLAB application with graphical user interface (GUI) were created for the task as shown in Figure 2.6. It allows defines parameters such as duty cycle, number of A-scans for the tasks. It will process the raw interferogram data and display the Xth image for preview on the whitespace located at the right panel. Cropping of the preview is allowed by specifying the "Top", "Left", "Height" and "Width" of the previewing image. It can be selectively to generate "enface" images, "raw dat" file "xzIntensity" files and the phase file, allowing re-accessing the intermediate data file for different purpose. i.e., accessing the phase file for functional Doppler application.

The GUI is set to have adjustable DC constants up to 7th order. The B-scan images on the right panel will update the preview image as any of the constants got changed. The program will pick one A-scan specified by user for image quality calculation.
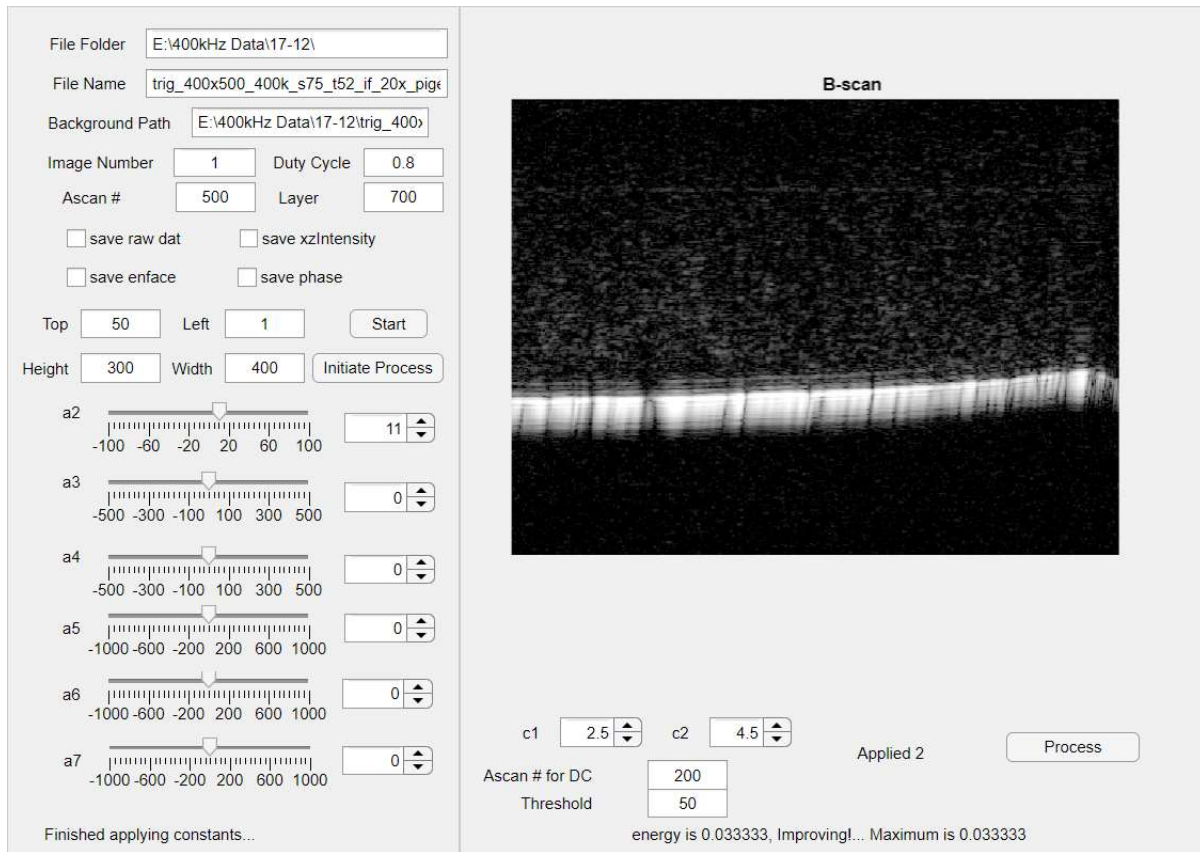
Figure 2.6 MATLAB application GUI for pre-processing of the raw data. illustrated is B-scan image of an inverted pig cornea with high reflection at the cornea-tear film interface

A sharpness metric of finding maximum energy concentration were used for quantifying the image quality after dispersion compensation (Wojtkowski, et al., 2004). By finding the number of points in the axial reflectivity profile that exceeds predetermined threshold, the sharpness metric is defined by one over the total number of points. With this metric, manual dispersion compensation can be applied and the result can be quantified. Contrast constants c_1 and c_2 are upper limit and lower limit for normalizing the intensity value, it is used for better contrast adjustment, and it will affect the sharpness calculation as well.

2.2.4 Pre-processing Steps

The preprocessing of the raw data to a B-scan image is shown in Figure 2.8. the raw tiff image was loaded with MATLAB program. Since the trigger signal suffers delay from the computer to the scanners and to the camera as illustrated in Figure 2.7. MATLAB circular shift function can shift the data array

in one direction as illustrated in the same diagram. This allows a software correction to the data for synchronization, forms correct b-scan and enface images. The number of rows to shift only need to be calibrated when the hardware systems setup was complete. Recalibration would be required when the electrical wiring got changed.
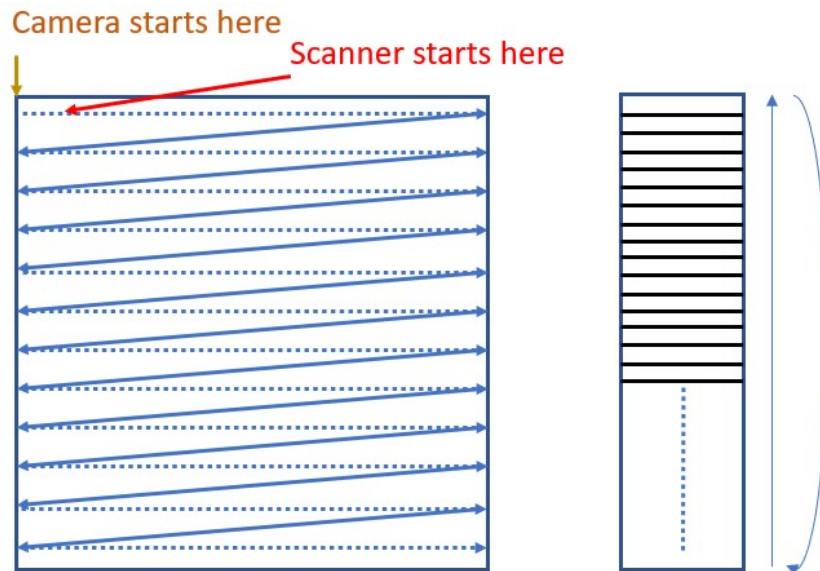
Figure 2.7 Example of asynchronization between Camera and scanner (left) and graphical illustration of MATLAB circular shift function on the tiff data arrays, arrays of interferogram been shifted upward and the squeezed-out arrays circulate to the bottom (right)

The first B-scan image is defined by the A-scan numbers, and duty cycle represent the valid A-scan data. with A-scan number equal to 500, and duty cycle of 0.8, the first 400 lines of interferograms would be valid to generate the b-scan image. The rest 100 A-scans are the data taken during the flyback.

The 400 arrays of interferogram are sliced out for further processing. The interferograms were collected over wavelength space, but fast Fourier Transform function in MATLAB performs discrete Fourier transform and it requires the signal to be in linear wavenumber (k) space instead of linear λ-space. The most typical solution for this is to perform numerical resampling using interpolation (Dorrer, Belabas, Likforman, & Joffre, 2000). MATLAB spline interpolation is used for such tasks, and Hilbert Transform (HT) is used prior the FFT to acquire complex phase info of the signal. 7TH order of DC is used in the pre-processing of the image, but it requires manual tuning of the constants. DC is achieved by multiplying the signal in Hilbert space with the DC factor:

$$dQ = e^{\sum_{i=2}^{7} a_i (\omega - \omega_0)^i}$$

The DC signal is equal to the Hilbert space signal multiplies the DC factor.
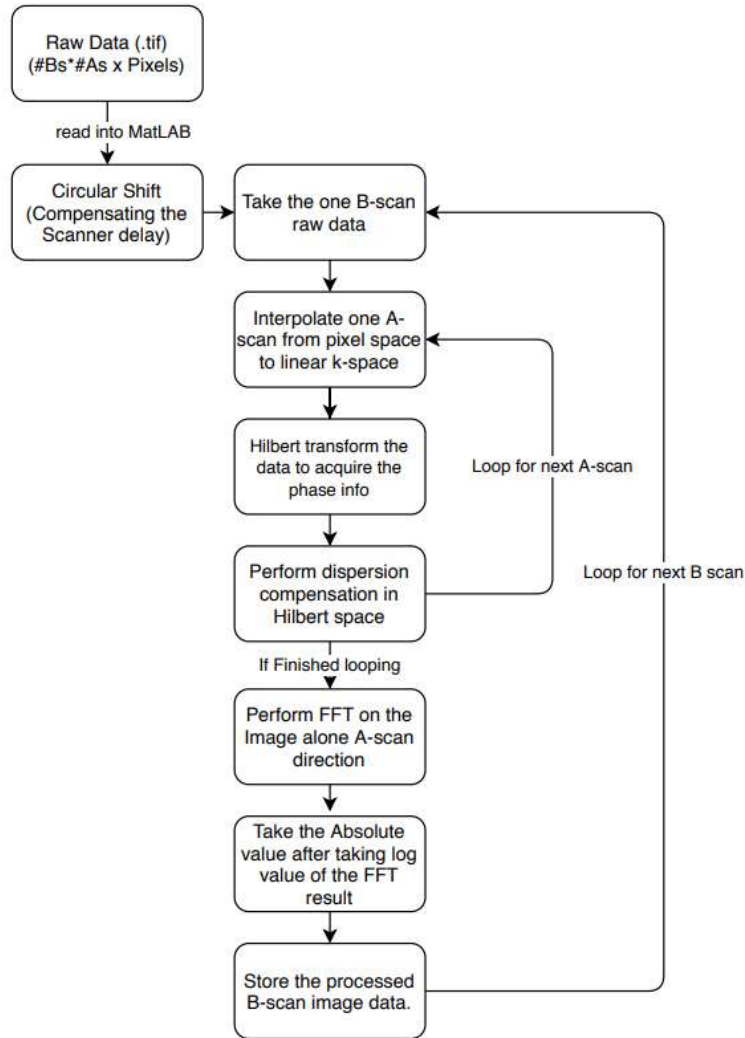
Figure 2.8 Pre-processing algorithms flow chart

Performing FFT on the DC signals will get the axial reflectivity profile of the A-scan. The absolute value of the log value of the reflectivity signal is used to generate b-scan images in Portable Network Graphics (PNG) format for storage. The software will loop through the data arrays until it finishes all the B-scans.

The enface images will be generated as illustrated in Figure 2.9, the volumetric OCT data has x,y,z dimensions and the b scan is on the xz plane. The enface image is the xy plane of the data.
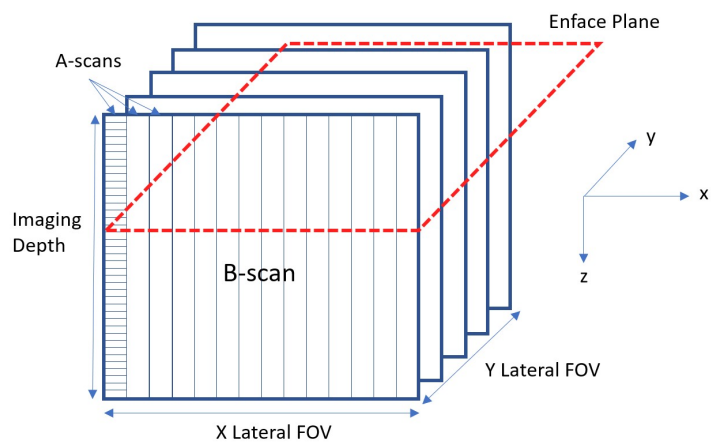
Figure 2.9 Enface image generation from OCT volumetric data

Chapter 3

System Characterization and Result

This chapter contains result the tests for characterizing the 400 kHz UHR-OCT system. The system utilizes spectrum from 600 nm to 980 nm. With center frequency of 802 nm and spectrum band with of 256 nm, it can provide an axial resolution of 1.37 μm in air. With a 20x objectives, it is capable of providing 1.04 μm/pixel of lateral resolution on the slow axis and 1.31 μm/pixel on the fast axis. Element 7 group 6 of the USAF-1951 target can be resolved. Providing 0.95 mW incident power at the sample arm, it can achieve 83 dB SNR with about 10 dB roll-off over the 1mm scanning range. Detailed result and calculations are displayed in each subsection with figures and plots.

3.1 Spectrums

The sample arm and reference arm spectral were measured with at the detection end of the 400kHz OCT system by using a silver mirror in place of the imaged object. Figure 3.1 shows the results from the test measurements.
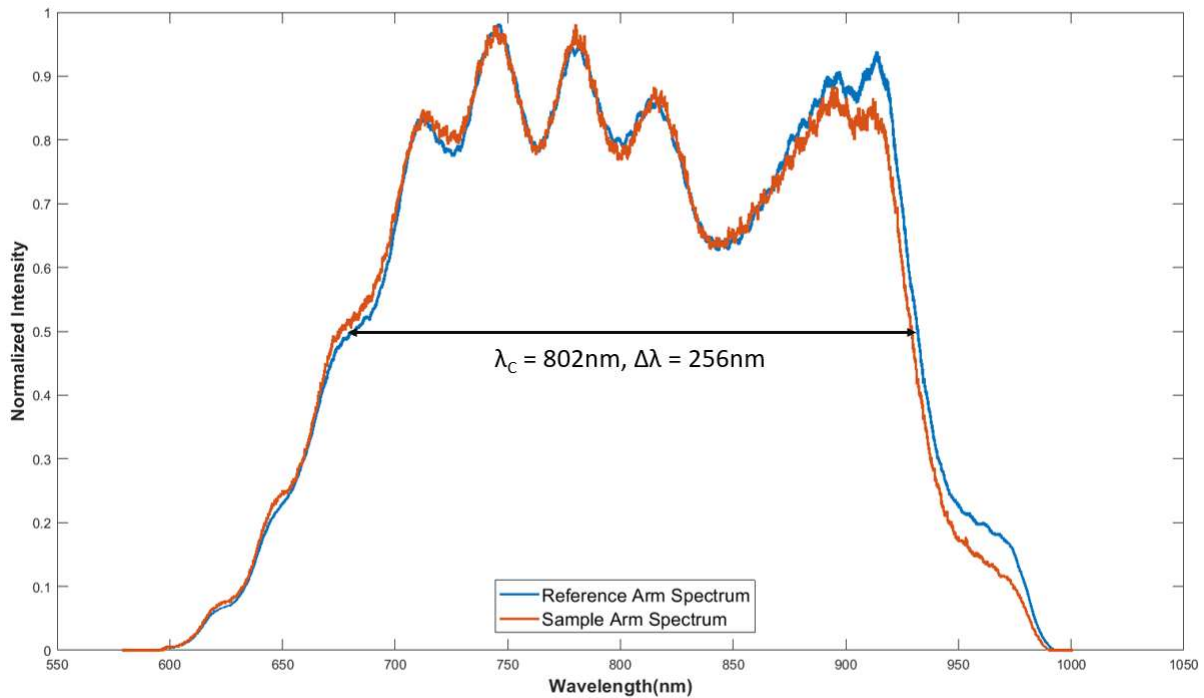
Figure 3.1 Sample and reference arm spectra measured at the detection end of the system

3.2 Spatial resolution

3.2.1 Axial Resolution

The theoretical axial resolution can be calculated from equation (1.8), $\Delta z = 1.11 \ \mu m$. However, this equation assumes a Gaussian shape of the spectrum, while as shown in Fig. 3.1. the measured refence and sample spectra are not exactly Gaussian. The true axial resolution was measured experimentally by using a silver mirror as the imaged object. Because the reference and sample arms of the OCT system contain different optical components made of exotic types of glass that generates higher orders of dispersion in the propagating optical beam, numerical dispersion compensation up to the 5th order was applied to the measured axial PSF of the OCT system. The symmetric axial PSF is shown in Figure 3.2 and has a FWHM of 1.37 μm in free space. The difference between the theoretically calculated and the experimentally measured axial OCT resolution can be attributed to the fact that the shape of the reference and sample spectra is very different than a Gaussian one. An additional result from this difference is the appearance of small sidelobes adjacent to the main lobe of the experimentally measured PSF. The measured PSF function has FWHM of 1.37 μm in free space, which corresponds to about 1μm of axial resolution in corneal tissue assuming an average corneal refractive index of 1.375 (Meek, Dennis, & Khan, 2003).
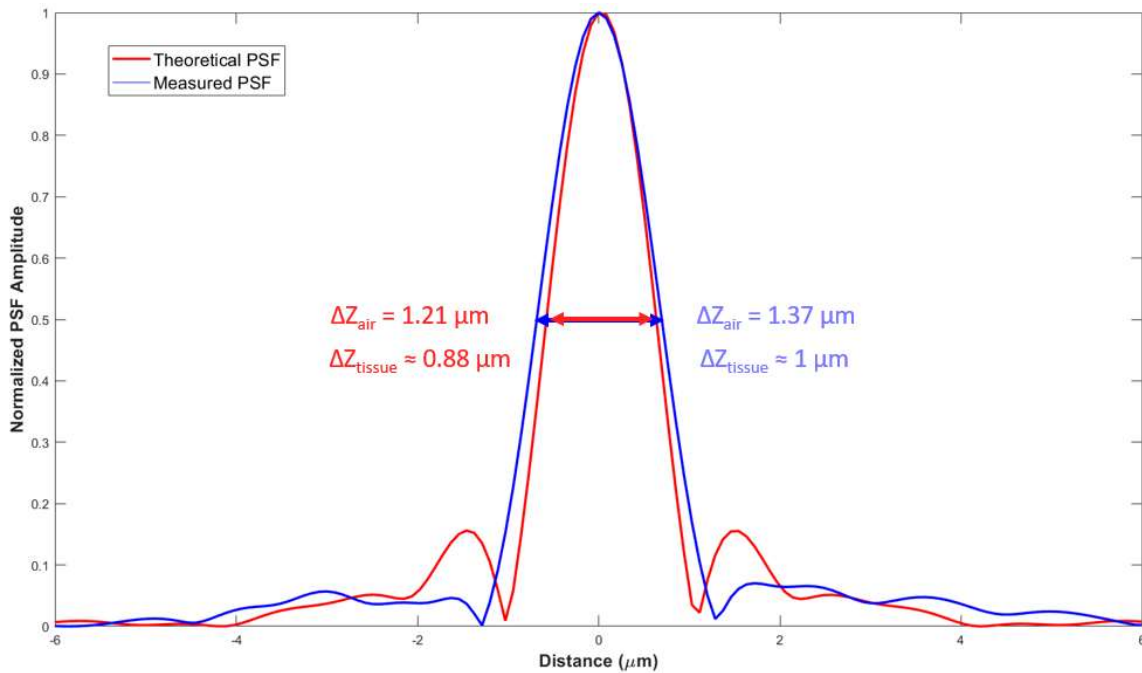
Figure 3.2 Theoretical PSF and measured PSF

27

The theoretical PSF is calculated by multiplying the sample spectrum with reference spectrum. Then adding a cosine term to it so the FFT result, the theoretical PSF function is shifted away from zero delay line for ease of comparison. The theoretical PSF has lower FWHM is because the measured PSF's hardware dispersion compensation is not perfect enough and the polarization state of the fibres could be changed in between the measurements.

3.2.2 Lateral OCT resolution evaluation

An USAF-1951 resolution target was used to evaluate the lateral resolution of the OCT imaging system. The resolution target image shown in Figure 3.3 was acquired with a 20x microscope objective (Mitutoyo, Infinity, IR corrected). The image shows Elements 6 and 7, which are the highest resolution groupings for this version of the USAF-1951 resolution target.
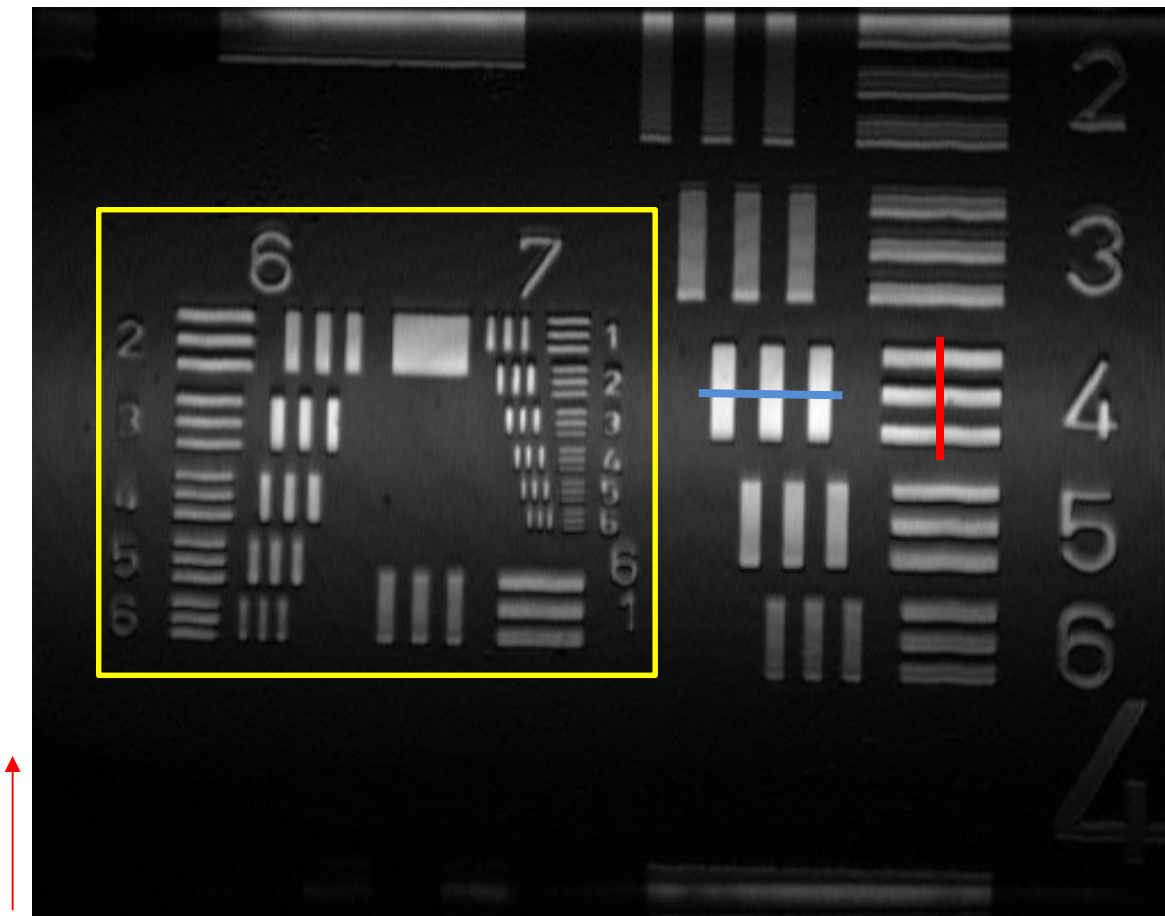
Figure 3.3 USAF-1951 with 20x objectives and 30 layers maximum intensity projection the arrow indicates the x-scanning direction

Two lines draw on the picture were used to determine the pixel resolution. They are element 4 of group 5, and the line pair resolution is $2^{\left(\text{Group}+\frac{\text{element}-1}{6}\right)}$, which is 45.3 lp/mm.
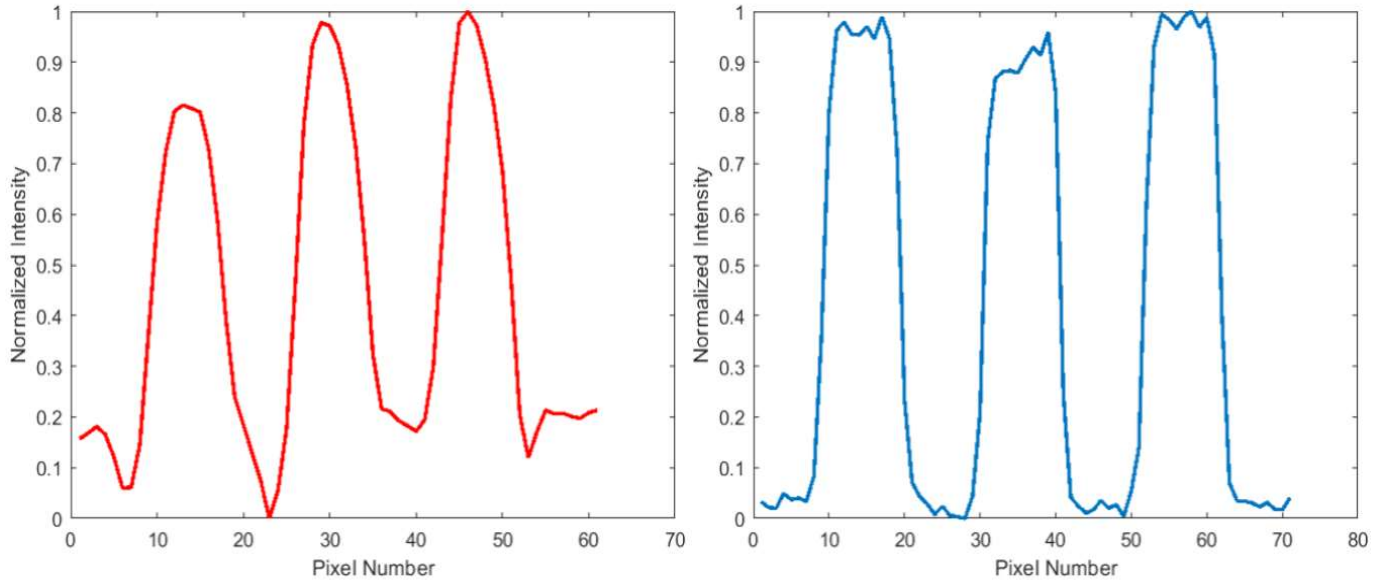
Figure 3.4 Line profiles of the imaging result for group 5 element 4

Taking the normalized intensity at its half maximum, the pixel run across 2.5 line-pairs are 42 pixels for the x and 53 pixels for y. the pixel wise resolution is calculated to be

$$x_{res} = \frac{2.5 \text{ lp}}{45.3 \text{ lp/mm}} \times \frac{1}{42 \text{ pixels}} = 1.31 \text{ μm/pixel}$$

$$y_{res} = \frac{2.5 \text{ lp}}{45.3 \text{ lp/mm}} \times \frac{1}{53 \text{ pixels}} = 1.04 \text{ μm/pixel}$$

The result is consistent with the pixel distribution since a square area is scanned over 400 points in x direction and 500 points in y direction.
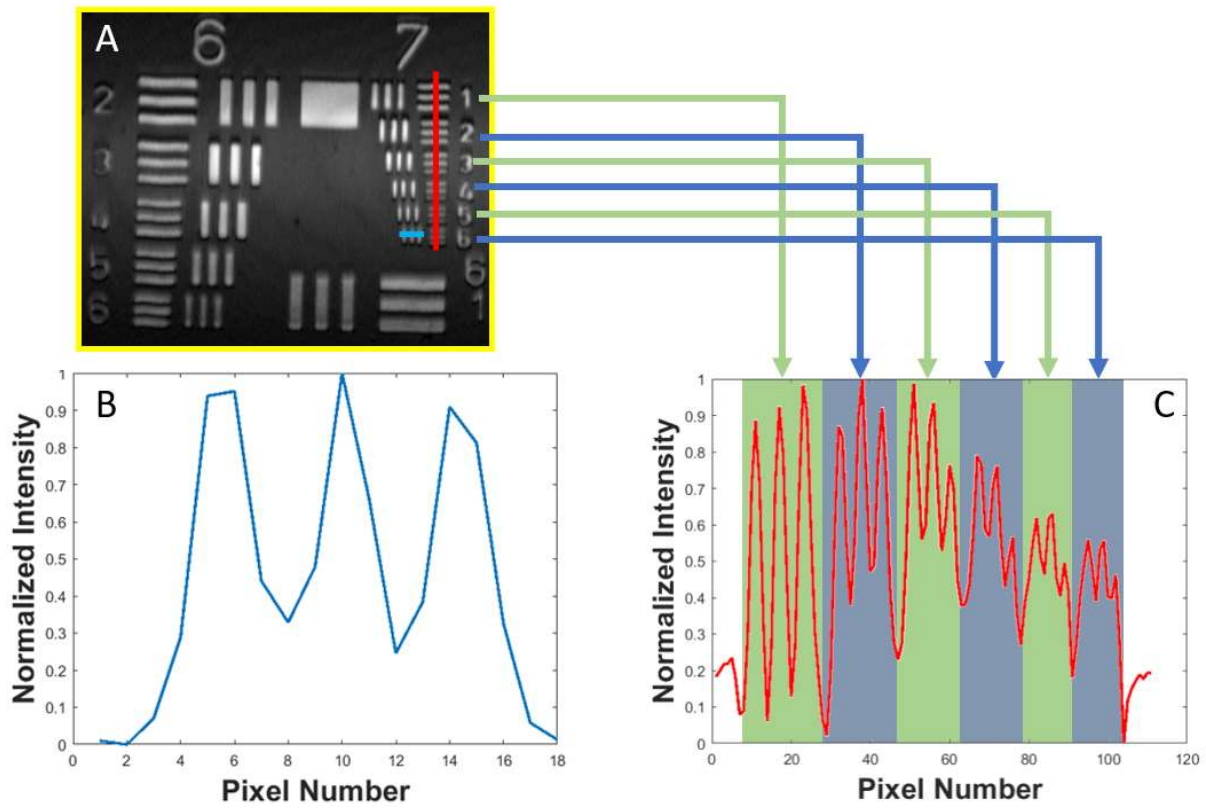
Figure 3.5 Group 6 and 7 regions of the USAF imaged with 400kHz UHR-OCT system (A), normalized intensity profile of the group 7 element 6 (B), normalized intensity profile of the group 7 elements 1 to 6 (C)

In the y direction of the group 6-7 region, the smallest scale can be resolved easily with the 400kHz UHR-OCT system, representing lower contrast. This is expected in the presence of off axis optical aberrations, including field curvature and off axis astigmatism, if the image plane is not on the disk of least confusion. The x-y directions do not have same contrast is because the scanning is not symmetrical, (i.e., 400 points by 500 points of raster scan.).

3.3 SNR and SNR Roll-off

The quantum efficiency (QE) of the Linea HS 400 kHz CLHS TDI cameras is show in Figure 3.6
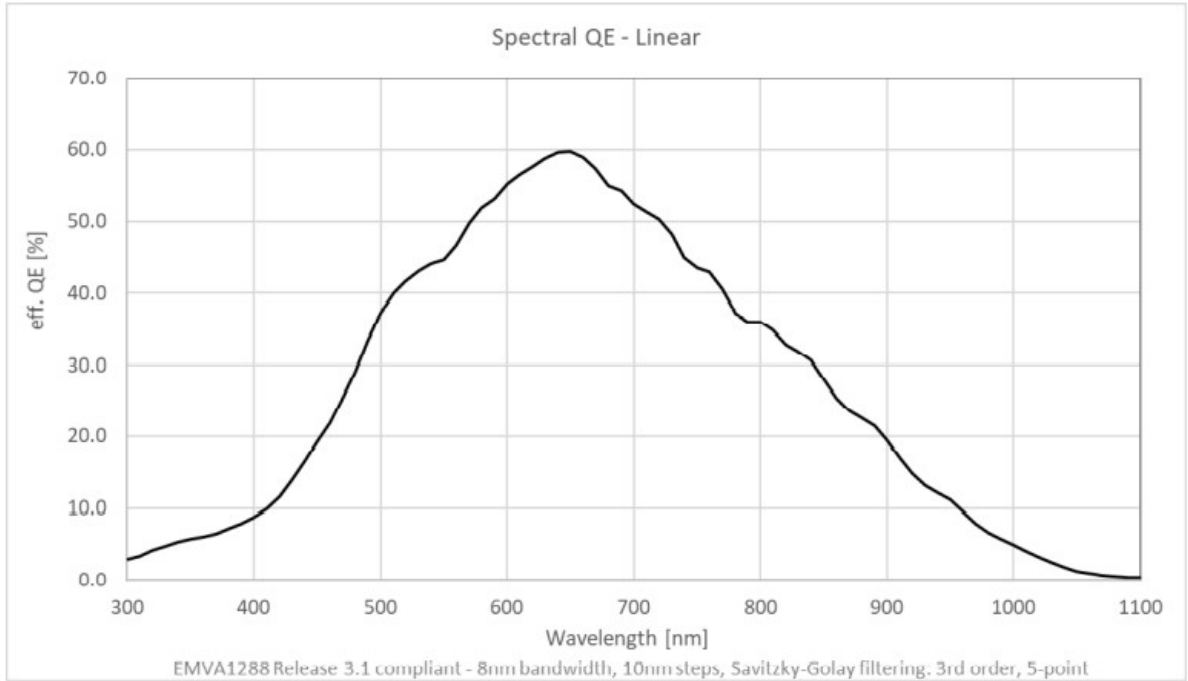
Figure 3.6 Spectral QE of the 400 kHz camera (source: Teledyne DALSA)

For simplified theoretical SNR calculation, taking the QE to be 0.35 since our spectrum is around 670 nm to 926 nm at FWHM centered at 802 nm. The calculated photodetector responsivity at 802nm is approximately 0.23 by converting the quantum efficiency with the equation:

$$\rho = \eta \times \frac{q}{hf} \qquad\qquad (3.1)$$

Where η is the quantum efficiency, q is the electron charge, h is the Plank's constant and f is the frequency of the light. When the sample arm has incident power of 0.95 mW, at full speed of 400 kHz, the camera's theoretical SNR is calculated to be approximately 92 dB. In real practice, a lower SNR is generally measured since factors like the power attenuation in the fibre and intensity fluctuation associated with the light source will also contribute approximately 5 to 10 dB to the SNR loss.

Experimentally, the signal strength is measured up to 2.8 mm with steps size about 100 microns. Results from the system's sensitivity test as a function of the scanning range are presented in Figure 3.7
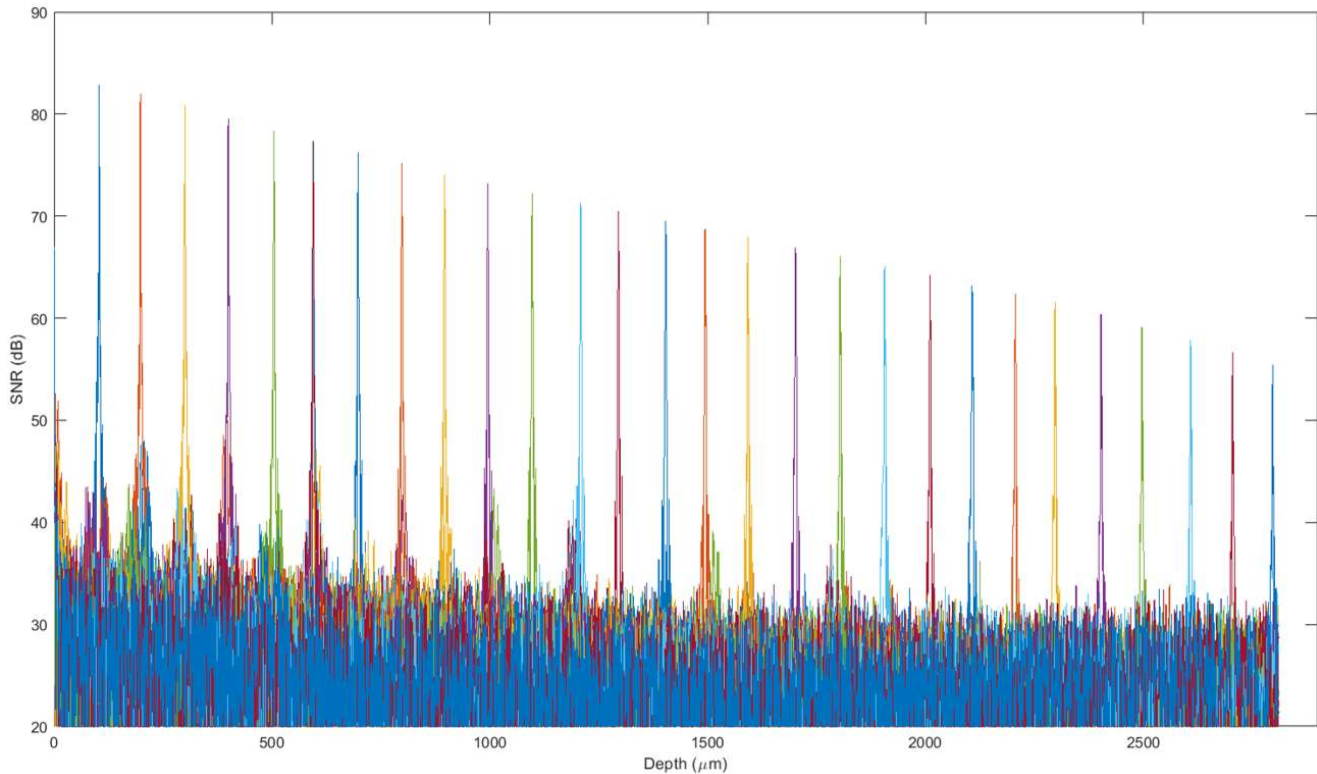
Figure 3.7 Sensitivity roll-off with scanning range

By projecting the 390 nm spectral bandwidth over the 8192 pixels, the system provides ~2.8 mm scanning range. The maximum SNR was approximately 83 dB measured at ~100 microns away from the zero-delay line for 0.95mW optical power incident on the silver mirror. The SNR roll-off was about 10 dB roll-off over the 1mm scanning range.

3.4 Images of biological tissue acquired with the 400 kHz OCT system

3.4.1 Cucumber

Due to COVID-19 related restrictions on research involving human subjects, we could not evaluate the 400 kHz system for in-vivo imaging of the human cornea. Instead, we caried ex-vivo imaging sessions on other biological tissues. Since the flesh of cucumber has similar water content and transparency of human corneal tissue in the visible spectral, as a first test, we imaged slices of cucumber near where the seeds are located. Figure 3.8 shows the enface result of the cucumber imaging with 20x objectives and 2.75mW illuminating power at the sample. A cell nucleus was found on the plane and

circled in red in the image. The images were post-processed with 8 layers of MIP in Amira software. The images on Figure 3.9 show the cellular structure of the cucumber tissue near the surface of a seed. Many small cells (~ 10 um in diameter) can be clearly identified in the B-scan and the enface image. The fact that the 400 kHz OCT system is able to visualize cells < 10um in size in cucumbers, means that the system should be able to visualize human corneal cells which are typically larger (~15 to 20 um in diameter).
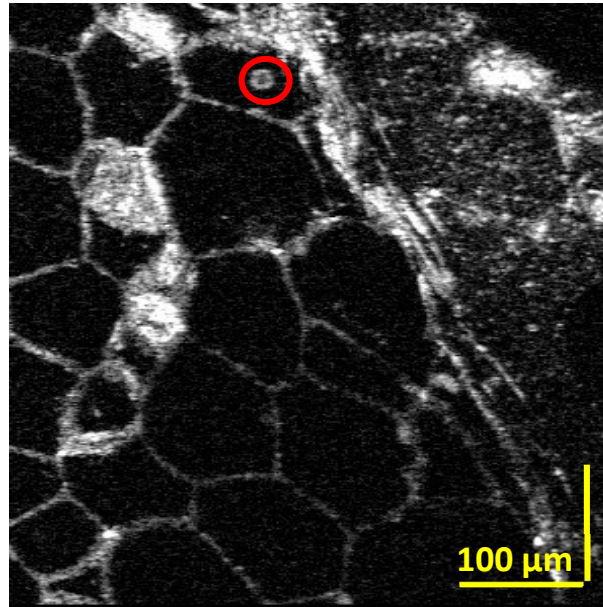
Figure 3.8 Enface image of a big cucumber seed, yellow bar shows the scale of 100 microns
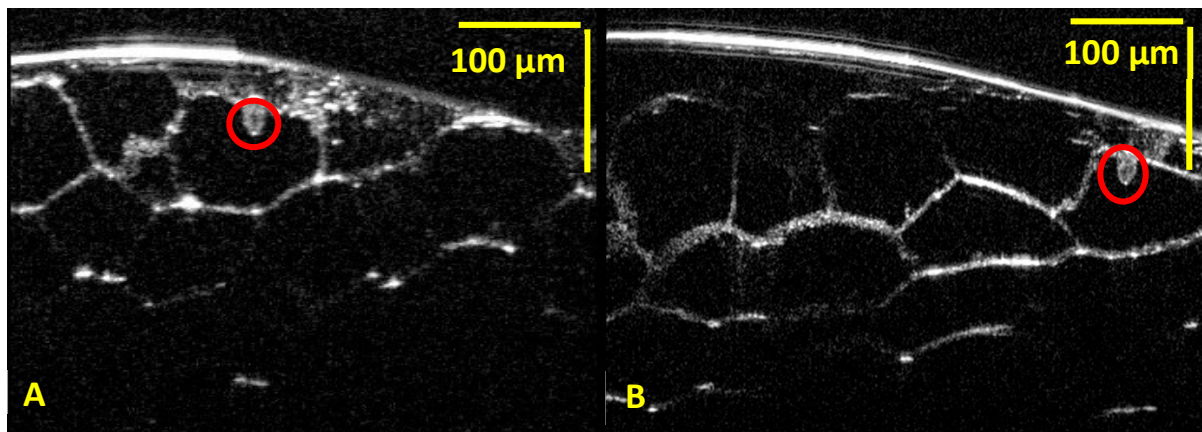
Figure 3.9 B-scan (xz plane) of the large cucumber seed (A), yz-plane of the large cucumber seed (B)

The B-scan and yz-plane of the same seed are also shown in Figure 3.9. The red circles indicate where the cell nucleus is located.
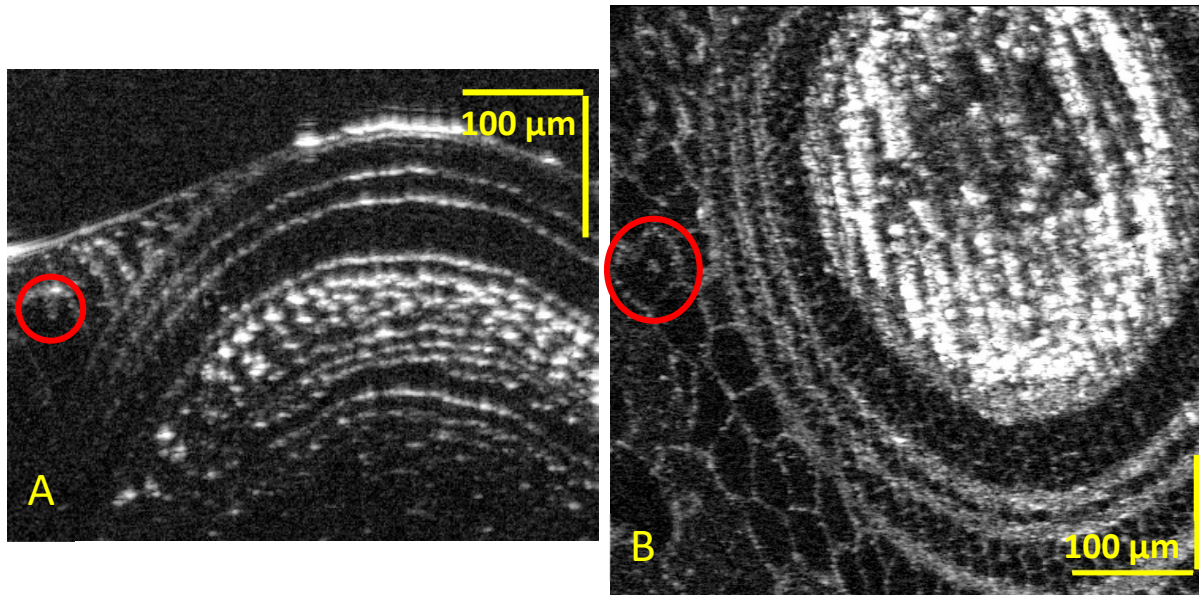
Figure 3.10 B-scan image of a smaller cucumber seed (A) Enface image of the same cucumber seed (B)

3.4.2 Images of porcine corneal epithelium

Freshly enucleated porcine eyes were obtained from a local meat processing plant and placed in buffer solution during transportation to the research lab. For the imaging session, the eyes were removed from the buffer solution and placed in a custom holder to allow alignment of the cornea relative to the imaging beam. A contact lens was placed on the surface of the cornea for 2 reasons: to keep the corneal epithelium hydrated for the duration of the imaging procedure and to reduce the refractive index mismatch between the epithelial tissue and air. Volumetric images (800 A-scans x 800 B-scans) of the corneal epithelium were acquired from a region directly adjacent to the corneal apex in order to avoid the specular reflection at the apex.

Figure 3.11 shows an enface image of the corneal epithelial layer at a depth collation corresponding to the basal cell layer. The image clearly shows individual epithelial cells and their nuclei.
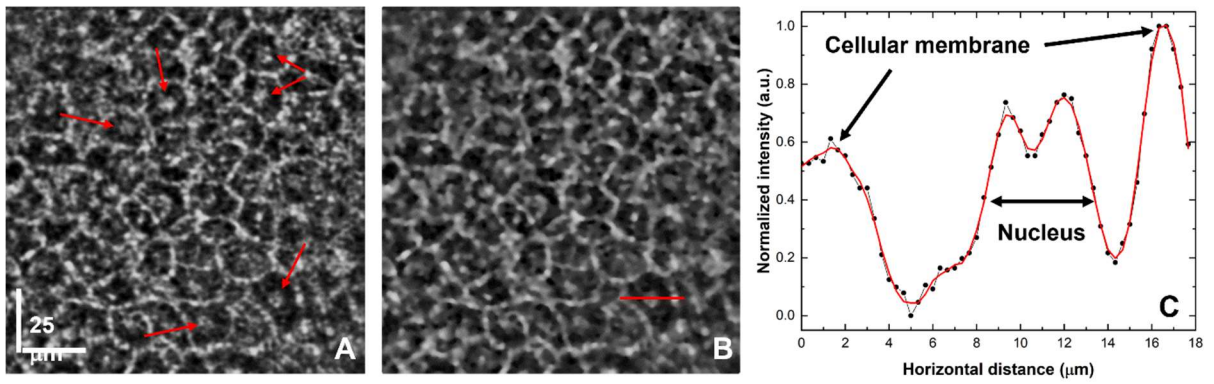
Figure 3.11 Human corneal epithelium. Enface OCT image acquired ex-vivo with 20x objective (A). Despeckled image (B). Intensity profile of a limbal cell (C) marked with the red line in (B).

Chapter 4
Challenges and Future Work

In conclusion, this thesis work provides an UHR-OCT system that combines 1 μm axial resolution in biological tissue with an extended scanning range of 2.8 mm, 400 kHz image acquisition rate and SNR of 83 dB with 0.95 mW incident power at the sample arm. The design should be sufficient to image semi-transparent biological tissues such as the human retina and cornea.

Images of Cucumber cells were presented in this thesis work, the cellular structure of the cucumber seeds can be resolved and even the nuclei of the larger seed cells were visible. An ex-vivo image of the pig cornea was imaged, and the epithelial layer is revealed.

4.1 Challenges

4.1.1 Software

Due to the huge amount of data throughput, the old LabVIEW software system used to control both the frame grabber and the scanner is no longer working properly. During the high-speed imaging process, the software acquires the B-scans intermittently due to the lack of the ability to acquire and process a huge amount of data with an incompatible frame grabber. The idea of using Teledyne DALSA CamExpert software to synchronize with the LabVIEW controlled scanner and test the system took a long time. However, it is not the perfect solution to the system because it is only for ex-vivo imaging. For in-vivo imaging, the scanner needs to be constantly scanning to avoid tissue damage of prolonged laser exposure.

4.1.2 COVID-19 Pandemic

COVID-19 lockdown delayed the development process. Even after the lockdown, communication between the camera engineer, software developers were delayed. The imaging policy needs to be updated and human subject imaging is restricted.

4.2 Future Work

The proposed solution for the software would be using visual studio (VS) to create a new imaging GUI for in-vivo imaging since Teledyne DALSA provides SDKs in VS for control of their Xtium2-CLHS frame grabber. This will solve the incompatibility of the frame grabber and LabVIEW software

issue. The logic of how to handle and process the raw data is confirmed and the in-vivo imaging protocol is defined. The new GUI should be able to perform similar tasks as the old LabVIEW software but without frame loss during the acquisition and saving data phase. The new method of acquisition should be sequential (grabbing all the data to the memory buffer) instead of grabbing one frame and process the frame at the same time.

Also, testing of the performance of the OCT system for in-vivo imaging of the animal and human ocular tissue is required once the COVID-19 restrictions to clinical studies have been lifted. Appendix A shows the fine-tuning of the modified VGG-16 with transferred learning on the OCT retinal images for image classification. The result suggests that with enough data for diseased and healthy corneal images are collected, a neural network can be trained to perform image classification.

Letter of Permission

Hi Dida,

I agree for the healthy cornea histology image that I have prepared and provided to be used as needed for the research work and MSc thesis of Lin Kun Chen.

Denise

Bibliography

Alonso-Caneiro, D., Read, S. A., Hamwood, J., Vincent, S. J., & Collins, M. J. (2018). Automatic segmentation of retinal and choroidal thickness in OCT images using convolutional neural networks.

Aumann, S., Donner, S., Fischer, J., & Müller, F. (2019). Optical coherence tomography (OCT): principle and technical realization. In *High Resolution Imaging in Microscopy and Ophthalmology* (pp. 59–85). Springer.

Bizheva, K., Haines, L., Mason, E., MacLellan, B., Tan, B., Hileeto, D., & Sorbara, L. (2016). In vivo imaging and morphometry of the human pre-Descemet's layer and endothelium with ultrahigh-resolution optical coherence tomography. *Investigative ophthalmology & visual science, 57*, 2782–2787.

Bolger, C., Bojanic, S., Sheahan, N. F., Coakley, D., & Malone, J. F. (1999). Dominant frequency content of ocular microtremor from normal subjects. *Vision research, 39*, 1911–1915.

Chen, S., Liu, X., Wang, N., Wang, X., Xiong, Q., Bo, E., . . . Liu, L. (2017). Visualizing micro-anatomical structures of the posterior cornea with micro-optical coherence tomography. *Scientific reports, 7*, 1–10.

Choma, M. A., Sarunic, M. V., Yang, C., & Izatt, J. A. (2003). Sensitivity advantage of swept source and Fourier domain optical coherence tomography. *Optics express, 11*, 2183–2189.

de Kinkelder, R., Kalkman, J., Faber, D. J., Schraa, O., Kok, P. H., Verbraak, F. D., & van Leeuwen, T. G. (2011). Heartbeat-induced axial motion artifacts in optical coherence tomography measurements of the retina. *Investigative ophthalmology & visual science, 52*, 3908–3913.

de Oliveira, R. C., & Wilson, S. E. (2020). Descemet's membrane development, structure, function and regeneration. *Experimental Eye Research*, 108090.

Dorrer, C., Belabas, N., Likforman, J.-P., & Joffre, M. (2000). Spectral resolution and sampling issues in Fourier-transform spectral interferometry. *JOSA B, 17*, 1795–1802.

Dozat, T. (2016). Incorporating nesterov momentum into adam.

Duchi, J., Hazan, E., & Singer, Y. (2011). Adaptive subgradient methods for online learning and stochastic optimization. *Journal of Machine Learning Research, 12*, 2121–2159.

Eghrari, A. O., Riazuddin, S. A., & Gottsch, J. D. (2015). Overview of the cornea: structure, function, and development. In *Progress in molecular biology and translational science* (Vol. 134, pp. 7–23). Elsevier.

Feizi, S., Jafarinasab, M. R., Karimian, F., Hasanpour, H., & Masudi, A. (2014). Central and peripheral corneal thickness measurement in normal and keratoconic eyes using three corneal pachymeters. *Journal of ophthalmic & vision research, 9*, 296.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE conference on computer vision and pattern recognition*, (pp. 770–778).

He, Y., Carass, A., Jedynak, B. M., Solomon, S. D., Saidha, S., Calabresi, P. A., & Prince, J. L. (2018). Topology guaranteed segmentation of the human retina from OCT using convolutional neural networks. *arXiv preprint arXiv:1803.05120*.

Hertsenberg, A. J., & Funderburgh, J. L. (2015). Stem cells in the cornea. In *Progress in molecular biology and translational science* (Vol. 134, pp. 25–41). Elsevier.

Huang, D., Swanson, E. A., Lin, C. P., Schuman, J. S., Stinson, W. G., Chang, W., . . . others. (1991). Optical coherence tomography. *science, 254*, 1178–1181.

Huey, E. B. (1900). On the psychology and physiology of reading. I. *The American Journal of Psychology, 11*, 283–302.

Izatt, J. A., & Choma, M. A. (2008). Theory of optical coherence tomography. In *Optical coherence tomography* (pp. 47–72). Springer.

Kermany, D. S., Goldbaum, M., Cai, W., Valentim, C. C., Liang, H., Baxter, S. L., . . . others. (2018). Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell, 172*, 1122–1131.

Kinga, D., & Adam, J. B. (2015). A method for stochastic optimization. *International Conference on Learning Representations (ICLR)*, *5*.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, (pp. 1097–1105).

Lee, C. S., Baughman, D. M., & Lee, A. Y. (2017). Deep learning is effective for classifying normal versus age-related macular degeneration OCT images. *Ophthalmology Retina, 1*, 322–327.

Leitgeb, R. A., & Baumann, B. (2018). Multimodal optical medical imaging concepts based on optical coherence tomography. *Frontiers in Physics, 6*, 114.

Leitgeb, R., Hitzenberger, C. K., & Fercher, A. F. (2003). Performance of fourier domain vs. time domain optical coherence tomography. *Optics express, 11*, 889–894.

Lwigale, P. Y. (2015). Corneal development: different cells from a common progenitor. In *Progress in molecular biology and translational science* (Vol. 134, pp. 43–59). Elsevier.

Martinez-Conde, S., Macknik, S. L., & Hubel, D. H. (2004). The role of fixational eye movements in visual perception. *Nature reviews neuroscience, 5*, 229–240.

Meek, K. M., Dennis, S., & Khan, S. (2003). Changes in the refractive index of the stroma and its extrafibrillar matrix when the cornea swells. *Biophysical journal, 85*, 2205–2212.

Nugroho, K. A. (2018). A Comparison of Handcrafted and Deep Neural Network Feature Extraction for Classifying Optical Coherence Tomography (OCT) Images. *arXiv preprint arXiv:1809.03306*.

Patel, S., & Tutchenko, L. (2019). The refractive index of the human cornea: A review. *Contact Lens and Anterior Eye, 42*, 575–580.

Rabinowitz, Y. S. (1998). Keratoconus. *Survey of ophthalmology, 42*, 297–319.

ReinsteinZ.Dan, ArcherJ.Timothy, GobbeMarine, SilvermanH.Ronald, & ColemanJacksonD. (2009). Stromal thickness in the normal cornea: three-dimensional display with Artemis very high-frequency digital ultrasound. Journal of Refractive Surgery, 25, 776‑786.

RosenblumD.Lawrence. (2011). See what I'm saying: The extraordinary powers of our five senses. WW Norton & Company.

Ruder, S. (2016). An overview of gradient descent optimization algorithms. *arXiv preprint arXiv:1609.04747*.

Saikia, P., Medeiros, C. S., Thangavadivel, S., & Wilson, S. E. (2018). Basement membranes in the cornea and other organs that commonly develop fibrosis. *Cell and tissue research, 374*, 439–453.

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

Spaide, R. F., Klancnik, J. M., & Cooney, M. J. (2015). Retinal vascular layers imaged by fluorescein angiography and optical coherence tomography angiography. *JAMA ophthalmology, 133*, 45–50.

SridharS.Mittanamalli. (2018). Anatomy of cornea and ocular surface. Indian journal of ophthalmology, 66, 190.

Srinivasan, P. P., Kim, L. A., Mettu, P. S., Cousins, S. W., Comer, G. M., Izatt, J. A., & Farsiu, S. (2014). Fully automated detection of diabetic macular edema and dry age-related macular degeneration from optical coherence tomography images. *Biomedical optics express, 5*, 3568–3577.

Swanson, E. A., & Fujimoto, J. G. (2017). The ecosystem that powered the translation of OCT from fundamental research to clinical and commercial impact. *Biomedical optics express, 8*, 1638–1664.

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., . . . Rabinovich, A. (2015). Going deeper with convolutions. *Proceedings of the IEEE conference on computer vision and pattern recognition*, (pp. 1–9).

Tan, B., Hosseinaee, Z., Han, L., Kralj, O., Sorbara, L., & Bizheva, K. (2018). 250 kHz, 1.5 μm resolution SD-OCT for in-vivo cellular imaging of the human cornea. *Biomedical Optics Express, 9*, 6569–6583.

Tao, A., Wang, J., Chen, Q., Shen, M., Lu, F., Dubovy, S. R., & Abou Shousha, M. (2011). Topographic thickness of Bowman's layer determined by ultra-high resolution spectral domain–optical coherence tomography. *Investigative ophthalmology & visual science, 52*, 3901–3907.

The Cornea and Corneal Disease. (2017, 03 17). Retrieved from The Eye Center: https://theeyecenter.com/eye-health-education/cornea-corneal-disease/

Wacker, K., McLaren, J. W., & Patel, S. V. (2015). Directional posterior corneal profile changes in Fuchs' endothelial corneal dystrophy. *Investigative ophthalmology & visual science, 56*, 5904–5911.

Werkmeister, R. M., Alex, A., Kaya, S., Unterhuber, A., Hofer, B., Riedl, J., . . . others. (2013). Measurement of tear film thickness using ultrahigh-resolution optical coherence tomography. *Investigative ophthalmology & visual science, 54*, 5578–5583.

Wilson, S. E., Lin, D. T., Klyce, S. D., Reidy, J. J., & Insler, M. S. (1990). Topographic changes in contact lens-induced corneal warpage. *Ophthalmology, 97*, 734–744.

Wojtkowski, M., Leitgeb, R., Kowalczyk, A., Bajraszewski, T., Fercher, A. F., & others. (2002). In vivo human retinal imaging by Fourier domain optical coherence tomography. *Journal of biomedical optics, 7*, 457–463.

Wojtkowski, M., Srinivasan, V. J., Ko, T. H., Fujimoto, J. G., Kowalczyk, A., & Duker, J. S. (2004). Ultrahigh-resolution, high-speed, Fourier domain optical coherence tomography and methods for dispersion compensation. *Optics express, 12*, 2404–2422.

Yadav, R., Lee, K.-S., Rolland, J. P., Zavislan, J. M., Aquavella, J. V., & Yoon, G. (2011). Micrometer axial resolution OCT for corneal imaging. *Biomedical optics express, 2*, 3037–3046.

Yao, X., Devarajan, K., Werkmeister, R. M., Dos Santos, V. A., Ang, M., Kuo, A., . . . others. (2019). In vivo corneal endothelium imaging using ultrahigh resolution OCT. *Biomedical Optics Express, 10*, 5675–5686.

Yun, S. H., Tearney, G. J., Bouma, B. E., Park, B. H., & de Boer, J. F. (2003). High-speed spectral-domain optical coherence tomography at 1.3 μm wavelength. *Optics express, 11*, 3598–3604.

Zeiler, M. D. (2012). ADADELTA: an adaptive learning rate method. *arXiv preprint arXiv:1212.5701*.

Zhang, Y., Yeh, L.-K., Zhang, S., Call, M., Yuan, Y., Yasunaga, M., . . . Liu, C.-Y. (2015). Wnt/β-catenin signaling modulates corneal epithelium stratification via inhibition of Bmp4 during mouse development. *Development, 142*, 3383–3393.

Appendix A

OCT Image Classification

This an unpublished project work done for SYDE 677 class back in the Fall 2018, the test accuracy done were at the that moment the highest to the best of my knowledge. The project was done on the OCT retinal images. But with sufficient OCT corneal images, the neural network can be re-tuned with transferred learning method to perform image classification on the diseased and healthy corneal images.

Introduction

Optical Coherence Tomography is a fast-growing technology that uses coherence property of light and gets cross-sectional information of the target. It helped ophthalmologist to obtained cross-sectional information of the patient's eye retina. A dataset of 84484 OCT retinal images is available and consisting of four classes, Choroidal neovascularization (CNV), Diabetic Macular Edema (DME), Drusen (DRUSEN) and normal (NORMAL). This proposed study aims to fine-tuning of VGG-16 deep neural network with customized top layers to perform image classification of the dataset and achieves test accuracy of 98.14%.

Recent development in Spectral Domain Optical Coherence Tomography provides ophthalmologist with a perfect tool to perform non-invasive in-vivo image of the human eyes. This breakthrough technology detects the interference signal coming back from the sample arm and reference arm and transform the signal into the axial depth information of the target which can achieve the axial resolution in microns (Wojtkowski, et al., 2004). a report has shown that compare to traditional angiography, OCT can achieve better imaging result and it is non-invasive (Spaide, Klancnik, & Cooney, 2015). Approximately 30 million OCT images are produced every year, without a smart analysis tool, interpretation and analysis of these images would be exceedingly difficult and time-consuming (Swanson & Fujimoto, 2017).
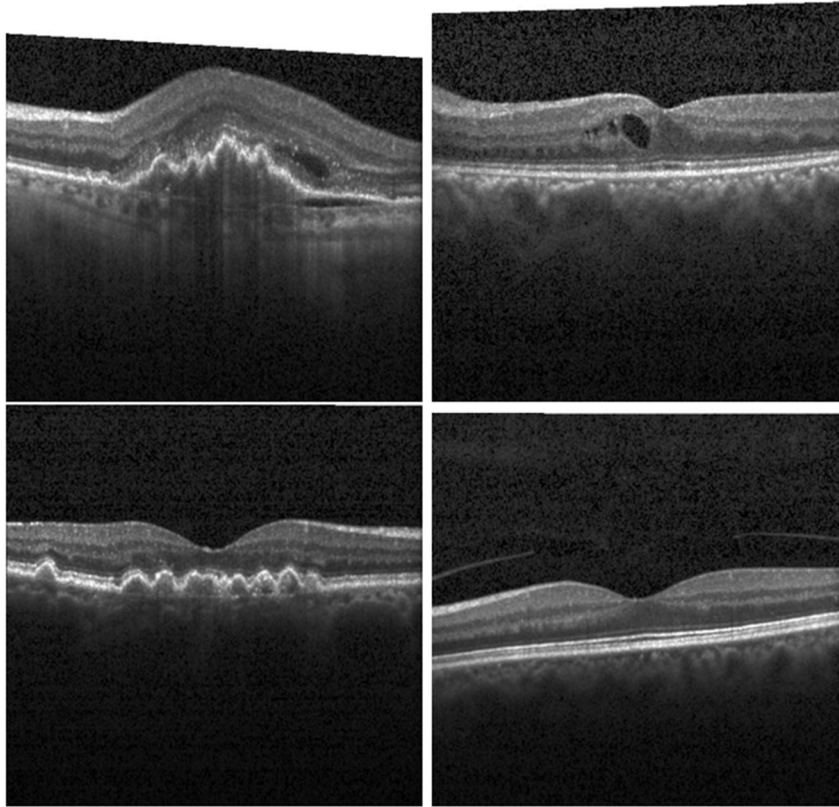
Figure A.1 OCT Images of Choroidal neovascularization (CNV)(Top-left), Diabetic Macular Edema (DME)(Top-right), Drusen (Bottom-left), and Normal Retina (Bottom-Right)

Figure A.1 shows some examples of the retinal diseases with the comparison to the healthy retina. CNV is the creation of new blood in the choroid layer of the eye and makes the patient visually impaired in a few weeks. DME occurs when fluid and proteins accumulate on or under the macular region of the eye. Macular is the region where most of the cones sit, accumulation of unexpected material may cause the visual distortion and patient will lose the ability to see details and colours. Drusen made up of lipid, although it does not cause any damage, having drusen will increase the risk of development of age-related macular degeneration (AMD).

Those OCT images can be very confusing to people because some of them look remarkably similar, a computer aid tool that helps to exam the image and provides suggestion to ophthalmologist would be extremely useful. These processing and analysis tasks in OCT can include segmentation and classification of the diseases. Researchers have tried both handcrafted method and neural network approach for extracting features. Lots of deep learning architectures are available for image classification. Alex et al (Krizhevsky, Sutskever, & Hinton, 2012) introduce AlexNet which uses

convolutional neural networks (CNN) technique in his deep learning model. This architecture significantly reduced the error by almost 10% from the previous year and won the ILSVRC-2012 competition. Since then, more architectures like VGG (Simonyan & Zisserman, 2014), GoogLeNet (Szegedy, et al., 2015) and ResNet (He, Zhang, Ren, & Sun, 2016) came up with different convolutional layer combinations and won the ILSVRC from 2014 to 2015.

Pratul et. al. studied OCT retina images with a histogram of oriented gradient (HOG) for feature extraction and support vector machine for classification. For 45 subjects, 15 normal, 15 patients with dry age-related macular degeneration (AMD), and 15 patients with DME, his algorithm was able to correctly identify 100%, 100% and 86.77% for AMD, DME and normal cases respectively (Srinivasan, et al., 2014).

Kuntoro's study compared the handcrafted and deep neural network feature extraction for classifying the OCT images into CNV, DME, DRUSEN and NORMAL. histogram of oriented gradient approach used in the study showed test accuracy of 50.1%. Local binary pattern (LBP) only got test accuracy of 42.3%. Whereas DenseNet-169 gave 88.0% and 89.2% (Nugroho, 2018). Another reported study in OCT images using VGG16 model with Xavier initialization successfully identify AMD with test accuracy of 87.63% (Lee, Baughman, & Lee, 2017).

He et al. and Alonson-Caneiro et al demonstrated deep learning feature extraction on their study (He, et al., 2018) (Alonso-Caneiro, Read, Hamwood, Vincent, & Collins, 2018). U-Net followed by S-Net and R-Net were proposed to obtain topology guaranteed segmentation (He, et al., 2018). In comparison, CNN was also trained to compute the probability map for each boundary position that was traced with a graph-search technique (Alonso-Caneiro, Read, Hamwood, Vincent, & Collins, 2018). Both studies shown the superiority of deep learning feature extraction compare to handcrafted approaches.

However, those studies mostly were done with a small dataset, typically less than 1000 images. This could be fine for handcrafted approaches but could be an issue for deep learning model. Kermany et al. provides an OCT dataset with over eighty thousand of classified images in 2017. This study will focus on fine-tuning modified VGG16 model and train the data to step up the test accuracy compare to 88.0% and 89.2% from Kuntoro's study, who uses the same dataset (Nugroho, 2018).

Methods and Experimental Design

1. Data Re-sampling

Dataset is provided by Kermany et al. with his study (Kermany, et al., 2018). The dataset is separated into four classes, they are Choroidal neovascularization (CNV), Diabetic Macular Edema (DME), Drusen (DRUSEN) and normal (NORMAL). However, the provided dataset contains limited validation set, only 32 images compare to 83484 training images. The original dataset was trained first, details will be discussed in the result section. The dataset was re-sampled for the training and fine-tuning of the model. Detailed count comparison of the datasets is shown in Table A.3 and Table A.4 below:

Table A.1 Original Dataset

Parameters	Train	Val	Test	Total
CNV	37205	8	242	37455
DME	11384	8	242	11598
DRUSEN	8616	8	242	8866
NORMAL	26315	8	242	26565
Total	83484	32	968	84484

Table A.2 Resampled Dataset

Parameters	Train	Val	Test	Total
CNV	36705	508	242	37455
DME	10884	508	242	11598
DRUSEN	8116	508	242	8866
NORMAL	25815	508	242	26565
Total	83484	2032	968	84484

500 images from each class were randomly selected and moved to the valid data. The test data remained the same.

2. Deep Learning Model and Platform

There are several platforms available for applying deep learning model, this study would be using Keras Platform. VGG16 will be used as the base model of the project, the model was loaded initially with no top layers and the structure is explained in Figure A.2.
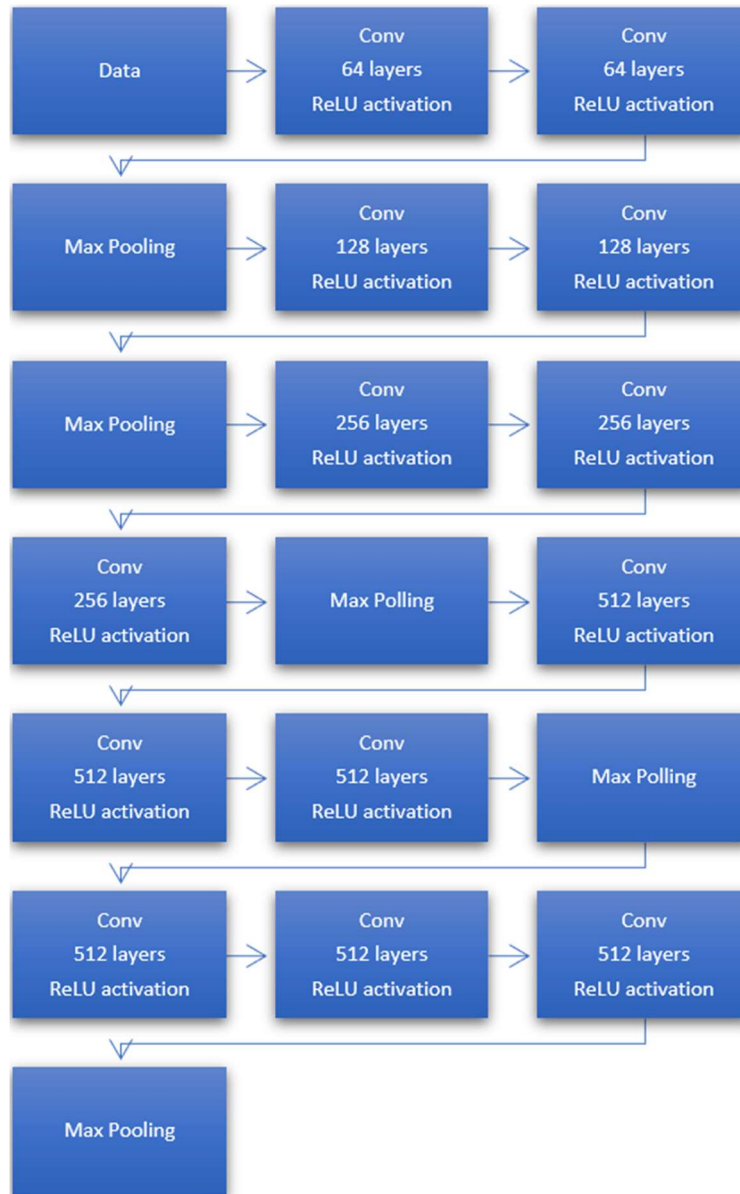

Figure A.2 VGG16 model without top layers

3. Top Layers

Customized top layers consist of flattening layer, Fully Connected layers (FCL) and dropout layers. SoftMax activation was used for the last FCL. Detailed info can be found in Figure A.3. Dropout layers introduced here helps the neurons to learn more robust features and helps to reduce the over-fitting (Krizhevsky, Sutskever, & Hinton, 2012)
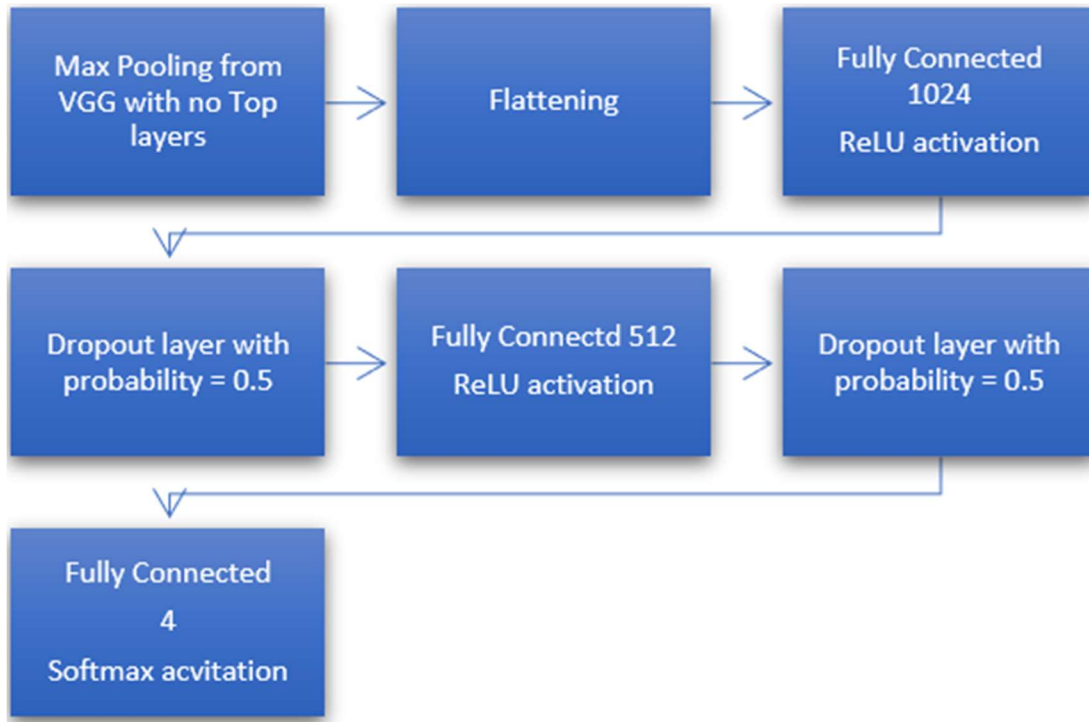

Figure A.3 VGG16 model without customized top layers

4. Optimizer

There are a lot of options for optimizer available. Some of them are stochastic gradient descent (SGD), root mean square propagation (RMSProp), AdaGrad, Adadelta, Adam, Adamax and Nadam. RMSProp is an optimizer that divides the gradient by a running average of its recent magnitude. AdaGrad has adaptive learning rate, it takes larger steps for infrequent and smaller steps for frequent parameters (Duchi, Hazan, & Singer, 2011). Adadelta is an extension from AdaGrad, but it adapts the learning rate based on a moving window whereas AdaGrad accumulates all the past gradients (Zeiler, 2012). Adam is another optimizer which combines both features of Adadelta and RMSProp, it keeps the exponentially decayed average of past gradients. Adamax is just another variant of Adam which is based on the infinity norm (Kinga & Adam, 2015). Nadam is Adam RMSProp with Nesterov

momentum (Dozat, 2016). All of them are similar algorithms and has similar good performance in similar cases (Ruder, 2016). However, due to the bias-corrected feature, Adam optimizer slightly outperform RMSProp towards the end of the optimization (Kinga & Adam, 2015), and this is the optimizer chosen to produce the best result.

5. Inputs and Parameters

To simplify and speed up the training process, transfer learning method is used in this study, pre-trained weights available from Kaggle website are loaded into the model and frozen for non-Top layers. Only customized top layers weights were trained in this study. Although OCT images are theoretically grayscale images of the reflected signal, Images from the dataset have 512 x 496 pixels with jpeg format. However, the VGG model requires RGB channels with 224 pixels of 2D-squared input data so the images were fed into the model with RGB format and resized to 224 x 224 x 3. The resizing of the images uses nearest neighbour interpolation.

Hyperparameters used for fine tuning are number of batch sizes, learning rate, number of epochs and dropout rates. Adam optimizer with categorical cross-entropy loss function (for one hot encoded labels) were used in this model. The final model contains 40.9 million parameters, 14.7 million of them are frozen and 26.2 million of them are trainable. For training of the model, the train images and validation images were shuffled while feeding into the model.

Fine Tuning and Result

As mentioned previously, the original dataset only contains 32 images for validation. For the first trial, the validation accuracy shoots up to 100% and fluctuate between 90% to 100% easily. Therefore 2000 images from the training set were chosen randomly and moved into the validation set.

Table A.3 Parameters Tuned for Trial 1 to 4

Parameters	Trial 1	Trial 2	Trial 3	Trial 4
Dropout Layer 1	0.7	0.7	0.5	0.5
Dropout Layer 2	0.5	0.5	0.5	0.5
Learning Rate	1e-4	1e-4	1e-5	1e-5
Batch Size	16	16	32	32
Epochs	10	10	10	9/50

When training the model second time, it got stuck in roughly 70% train accuracy and 50% validation accuracy. This may be caused by the large learning rate and dropout probability used. Therefore, the learning rate would be reduced to one order of magnitude less for the next trial and dropout rate changed to 50%.

After 10 epochs of trial 3, it got training accuracy of 89.33% with validation accuracy of 82.29%. weights were saved and loaded for trial 4. Epochs for trial 4 were increased to 50 but an early stopping function was applied with patience equal to 5. this time it stopped at epochs number 9. This model produces a test accuracy of 95.8%.

Table A.4 Parameters Tuned for Trial 5 and 6

Parameters	Trial 5	Trial 6
Dropout Layer 1	0.5	0.5
Dropout Layer 2	0.5	0.5
Learning Rate	1e-5	1e-5
Batch Size	32	64
Epochs	50	50

In trial 5, the epochs number were increased to 50 epochs and checkpoint were set to monitor the validation accuracy. Early stopping function was disabled. although the validation accuracy maxed out in epochs number 33. It still produces the best result so far. Trial 6 was attempted to see if increasing batch size could push the result better. However, doubling the batch size did not make it converge faster or slower, the rates were about the same, but the training time for each epoch increased about 1/6. And the final test accuracy is the same.

Table A.5 Confusion matrix of best model

	CNV	DME	DRUSEN	NORMAL
CNV	242	0	0	0
DME	4	243	0	5
DRUSEN	8	0	234	0
NORMAL	0	0	1	241

Overall, the test accuracy of the best model is

$$\frac{242+243+234+241}{242+233+234+241+4+8+5+1}=\frac{950}{968}=0.9814$$

with the precision and recall score shown in Table A.6.

Table A.6 Best model precision and recall result

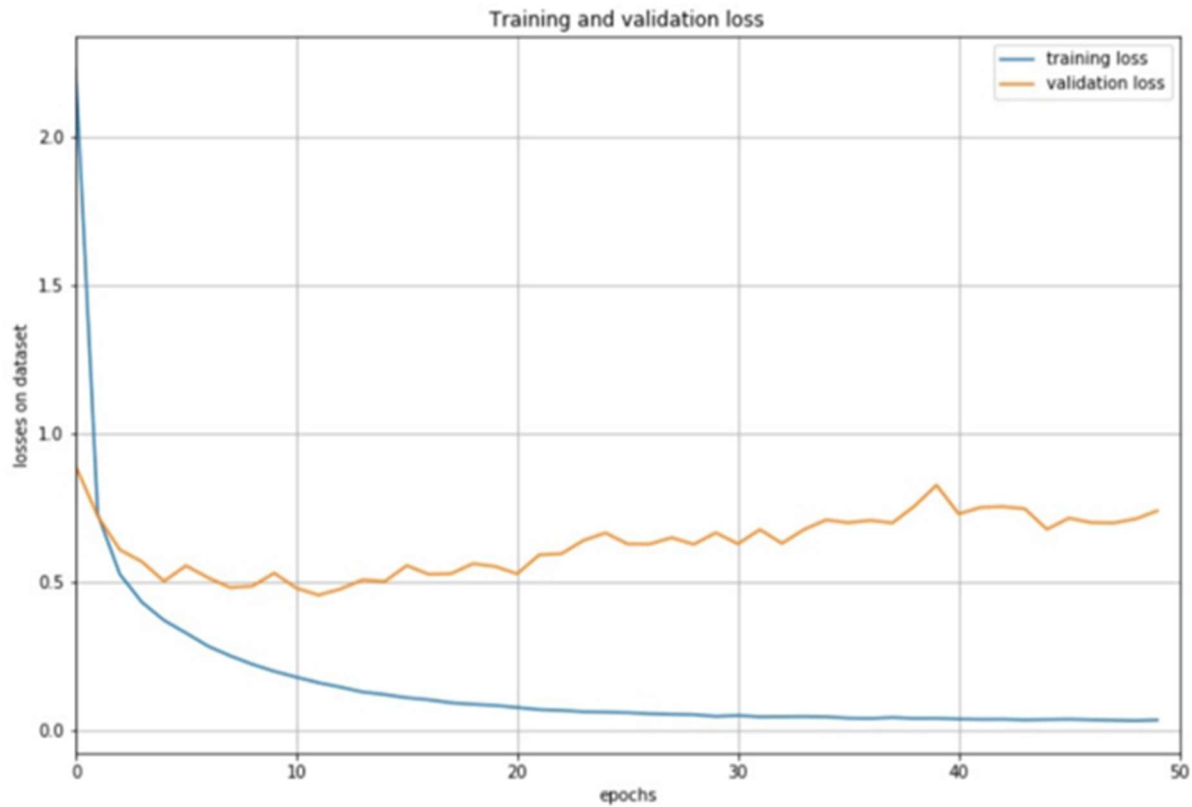|	Precision	Recall
CNV	0.95	1.00
DME	1.00	0.96
DRUSEN	1.00	0.97
NORMAL	0.98	1.00

Figure A.4 Training loss vs validation loss

The training accuracy vs validation accuracy in the Figure A.4 shows that the model reaches the 90% training accuracy and 85% validation accuracy in first six epochs, rest of the epochs contributed little accuracy and after 11 epochs, the validation loss starts to increase as shown in Figure A.5. This might indicate that the model is over-fitting. The weight at 11 epochs was loaded and the test accuracy was

97.62%. Later epochs just make the model better at fitting the training data while increasing the validation loss.
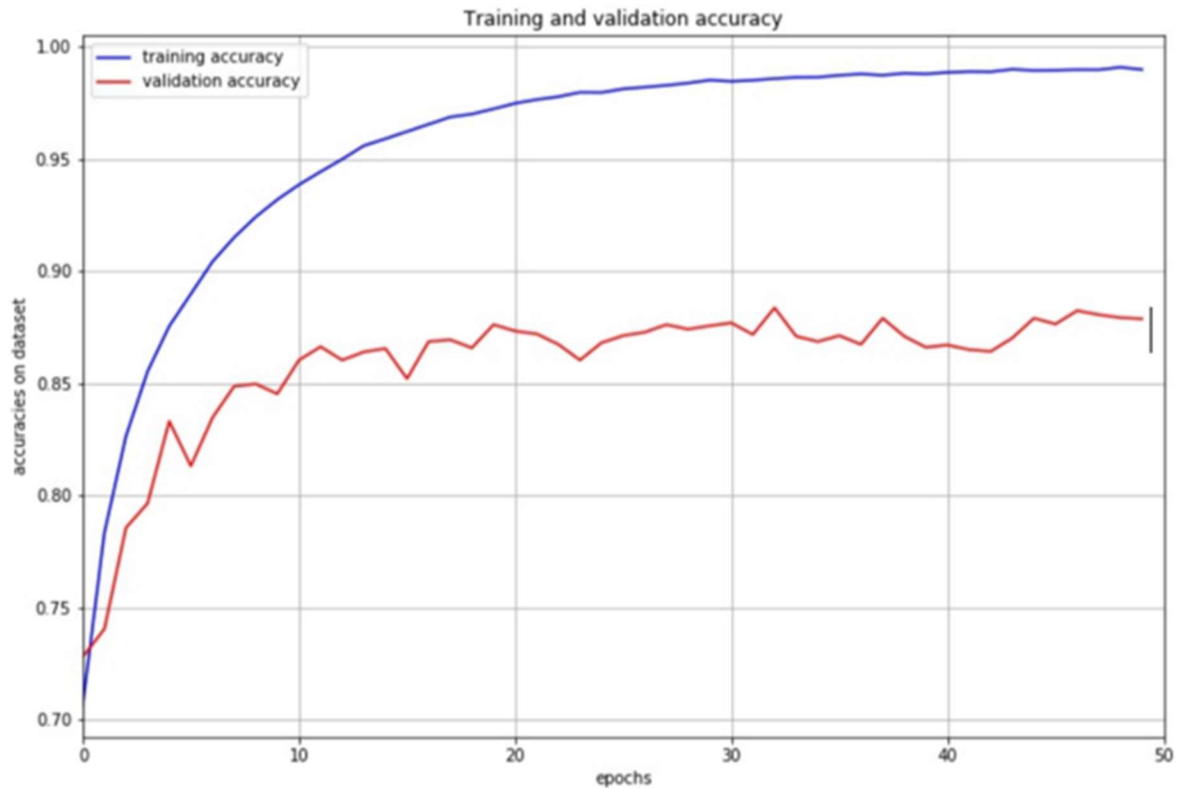

Figure A.5 Training accuracy vs validation accuracy

Conclusion

Customized VGG16 model with customized top layers was used to train OCT dataset of 81484 images and valid on 2032 images. This model of VGG16 uses 3 FCL with 2 dropout layers in between to classify OCT images into CNV, DME, DRUSEN and NORMAL classes. With the learning rate of 1e-5, batch sizes of 32 images, 50% dropout rates for both dropout layers for 50 epochs, best test accuracy were achieved 98.15% with 968 test images.

Appendix B

Python Codes for OCT Image Classification with VGG-16

```python
import os
import keras
import warnings
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt

from keras.models import Sequential, Model
from keras.layers import Conv2D,MaxPool12D, Dense, Dropout, Softmax, Input, Flatten
from keras.optimizers import Adam,RMSprop,SGD
from keras.layers.merge import add
from keras.layers import Dense, Activation, Flatten
from keras.layers import Conv2D, MaxPooling2D, BatchNormalization
from keras.layers import BatchNormalization
from keras.callbacks import ModelCheckpoint, Callback, EarlyStopping
from keras.applications.vggi6 import VGG16
from keras.applications.inception_v3 import Inceptionv3
from keras.metrics import categorical_accuracy
from sklearn.metrics import roc_auc_score,roc_curve,accuracy_score,recall_score
from sklearn.metrics import classification_report, confusion_matrix
from keras.preprocessing.image import ImageDataGenerator
from tensorflow import set_random_seed

warnings .filterwarnings ("ignore")

os.environ[ "PYTHONHASHSEED"] = "0"

np.random.seed(1)
set_random_seed(2)

imageSize = 224

weight_pathl = '/home/shared/Practice/OCT/keras-pretrained-models/vgg16_weights_tf_dim_ordering tf_kernels_notop.h5"
#weight_path2 = '/home/shared/Practice/OCT/keras -pretrained-models/inception_v3_weights_tf_dim_ordering_tf_kernels_notop.h5'
pretrained_VGG = VGG16(weights = weight_path1, include_top=False, input_shape=(imageSize, imageSize, 3))
#pretrained_IncepV3 = InceptionV3(weights = weight_path2, include_top=False, input_shape=(imageSize, imageSize, 3))

base_model = pretrained_VGG
#base_model = pretrained_IncepvV3

X = base_model.output

for layer in base_model.layers:
layer.trainable = False
x = Flatten(name = 'flatten')(x)
x = Dense(1024, activation = 'relu', name = 'fc1')(x)
x = Dropout(0.5, name = 'dropout1') (x)
x = Dense(512, activation = 'relu', name = 'fc2")(x)
x = Dropout(0.5, name = 'dropout2')(x)
predictions = Dense(4, activation="softmax' )(x)
model = Model(inputs=base_model.input, outputs=predictions)

model. summary()
batch_size = 32
nu_of_epoches = 59
learning_rate = 1e-5

#es = EarlyStopping(patience = 18)
# Setup a callback to save the best model

chkpt = ModelCheckpoint (filepath='/home/shared/Practice/OCT/OCT_2017/model.{epoch:02d}-{val_acc: .2f}.hdf5',
                monitor='val_acc', verbose=1,mode = 'max', save_best_only = True, period = 1)

model. compile(Adam(1r=learning_rate),loss="categorical_crossentropy", metrics=["accuracy"])}

train_dir = "/home/shared/Practice/OCT/OCT_2017/0CT_2017/train/"
val_dir = "/home/shared/Practice/OCT/OCT_2017/0CT_2017/val/"
test_dir = "/home/shared/Practice/OCT/OCT_2017/0CT_2017/test/"
# ['DME'', 'CNV', 'NORMAL', '.DS_Store', 'DRUSEN']
```

```python
datagen = ImageDataGenerator()
train_set = datagen.flow_from_directory(train_dir,target_size=(imageSize, imageSize),
                                        color_mode="rgb", shuffle=True, seed=1, batch_size=batch_size)
valid_set = datagen.flow_from_directory(val_dir, target_size=(imageSize, imageSize),
                                        color_mode="rgb", shuffle=True, seed=1, batch_size=batch_size)
test_set = datagen.flow_from_directory(test_dir, target_size=(imageSize, imageSize),
                                       color_mode="rgb", shuffle=False, batch_size=8)

#modeL.lLoad_weights("/home/shared/Practice/OCT/OCT_2017/best_model_todate")
history = model.fit_generator(train_set,steps_per_epoch=81484//batch_size,
                              epochs=nu_of_epoches,
                              validation_data=valid_set,
                              validation_steps=968//8,
                              callbacks = [chkpt])

# Evaluate model
#model.save_weights('/home/shared/Practice/OCT/OCT_2017/50epLR2e5_weights_cont')
y_pred = model.predict_generator(test_set, verbose=True, steps = 968//8)
y_pred = np.argmax(y_pred,axis = -1)

print('Confusion Matrix')
cm = confusion_matrix(test_set.classes, y_ pred)
print(cm)
print('Classification Report')
target_names = ['CNV','DME', 'DRUSEN", 'NORMAL' ]
print(classification_report(test_set.classes, y_pred, target_names=target_names))
```