UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

FORWARD-VIEW ENDOSCOPIC OPTICAL COHERENCE TOMOGRAPHY FOR NEEDLE-BASED INTERVENTIONS

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

CHEN WANG
Norman, Oklahoma

2024

Downloaded from SHAREOK. Persistent Handle URL: https://hdl.handle.net/11244/340291

FORWARD-VIEW ENDOSCOPIC OPTICAL COHERENCE TOMOGRAPHY FOR NEEDLE-BASED INTERVENTIONS

A DISSERTATION APPROVED FOR THE
STEPHENSON SCHOOL OF BIOMEDICAL ENGINEERING

BY THE COMMITTEE CONSISTING OF

Dr. Qinggong Tang

Dr. Wei Chen

Dr. Javier A. Jo Pun Kay

Dr. Song Fang

© Copyright by CHEN WANG 2024
All Rights Reserved.

ACKNOWLEDGEMENT

During my five-year Ph.D. academic career at the University of Oklahoma, I would first express my deepest gratitude to Dr. Qinggong Tang. I really appreciate your assistance with both my research studies and daily life. At the beginning of my Ph.D. journey, I was unfamiliar with many aspects of the program and faced a lot of challenges. Dr. Tang was consistently kind and patient, offering guidance whenever I had difficulties in this new environment. He has helped me in different aspects of my new life at OU and made me become accustomed to the new research projects. As a student who had no prior experience in the biomedical field, I was often confused in the engineering principles, research papers, and experimental work. Dr. Tang always patiently taught me how to approach research problems and execute specific tasks effectively. He has also provided me with lots of suggestions regarding my future development plans. His intelligence and character will always inspire me in my future career.

I would like to express my sincere gratitude to all my collaborators who have contributed significantly to our research projects. Special thanks to Dr. Wei Chen, an outstanding professor of biomedical engineering. He has given me a lot of valuable advice in different academic conferences and activities. During our collaboration I have gained a lot of experience about biological experiments and cancer treatment. I am also grateful to Dr. Chongle Pan from the School of Computer Science. He is very knowledgeable in computer science, machine learning and artificial intelligence. His work has been very helpful in the establishment of our medical imaging platform. Also, I extend my thanks to Dr. Kar-ming Fung for his assistance in pathological work and providing guidance on various aspects of cancer diagnosis and treatment. Moreover, I would like to acknowledge Dr. Javier A. Jo Pun

Kay and Dr. Song Fang for their substantial support throughout my Ph.D. program. Their help has been important to my academic career.

I would also like to thank all my lab members and friends from different countries. Without your help and support I would never have the achievements today. I will always memorialize our experiences on our collaborations, the conferences we attended together, and our trips to the national parks.

Finally, I would like to express my thanks to my family for their unwavering support throughout my whole life. My love for you is endless and will always endure!

ABSTRACT

Needle-based interventions (NBIs) are procedures that require a minimally invasive approach to gain access to tissue structures of interest. Each year, more than 30 million of these interventions are performed in the US. Due to the lack of proper visual feedback guiding navigation, up to 33% of NBIs are associated with complications, incurring significant human and economic costs. The objective of my dissertation is to establish a new research paradigm shift in the methodology of NBIs through integrating a unique forward-imaging multi-contrast optical coherent tomography (MC-OCT) needle technology with advanced machine-learning-based computer-aided diagnosis algorithms. Optical coherence tomography (OCT) is a novel biomedical imaging technique which provides micron-level spatial resolution. Compared to conventional imaging modalities, it provides the imaging results with a lot more details of localized tissue. OCT has been widely used in various applications including ophthalmology, biopsy procedures, different disease diagnosis, etc. Furthermore, OCT can perform real-time imaging based on the high-speed scanner and data acquisitor. Therefore, OCT has the potential in real-time guidance in different NBIs. To successfully acquire depth-resolved imaging during surgeries, we developed forward-view endoscopic OCT systems using gradient-index (GRIN) lenses. Additionally, we utilized polarization-sensitive OCT (PS-OCT) and Doppler OCT to build MC-OCT for enhanced tissue contrast and at-risk blood vessels detection. PS-OCT is a function extension of traditional OCT because it provides polarization-related biological features based on birefringence and diattenuation phenomena. Doppler OCT is effective in fluid flow imaging, so we utilized it for blood flow detection. By introducing PS-OCT and Doppler OCT, our endoscopic MC-OCT systems have been proved

to be feasible in recognizing different tissue type and blood vessels. In this dissertation, we demonstrated endoscopic MC-OCT systems in surgical navigations in different NBIs including percutaneous nephrostomy (PCN), epidural anesthesia, Veress needle for pneumoperitoneum establishment, and different cancer diagnosis. To improve the working efficiency and reduce learning curves, we further applied deep learning methods in our studies to automate the navigation procedures. Different architectures were tested for tissue classifications and blood vessel detections. This innovated imaging platform has established a paradigm shift, offering the potential for widespread applications across various NBIs.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION 1
1.1 Working principles of OCT 1
1.2 OCT signal measurement 3
1.3 Applications of OCT 7
1.3.1 Early ophthalmologic applications of OCT technology 7
1.3.2 Applications of OCT in surgical guidance 8
1.4 Endoscopic OCT 9
1.4.1 Development of endoscopic OCT 10
1.4.2 Clinical applications of endoscopic OCT 13
1.5 Polarization Sensitive OCT (PS-OCT) 15
1.5.1 Polarization state of light 16
1.5.2. Diattenuation and birefringence 18
1.5.3. Jones formalism 20
1.5.4 Mueller-Stokes Formalism 27
1.6 Doppler OCT 30
1.6.1 Working principles of Doppler OCT 30
1.6.2 Applications of Doppler OCT 31
1.7 Outline of Dissertation 32
CHAPTER 2: DEEP-LEARING AIDED FORWARD OPTICAL COHERENCE TOMOGRAPHY ENDOSCOPE FOR PERCUTANEOUS NEPHROSTOMY GUIDANCE 38
2.1 Introduction 38
2.2 Materials and Methods 42
2.2.1 Experimental setup 42
2.2.2 Data acquisition 44
2.2.3 Training of the convolutional neural networks 46
2.2.4 Nested cross-validation and testing 47
2.3 Results 49
2.3.1 Forward OCT imaging of different renal tissues 49

2.3.2 CNN development and benchmarking results 50
2.4 Discussion 55
CHAPTER 3: COMPUTER-AIDED VERESS NEEDLE GUIDANCE USING ENDOSCOPIC OPTICAL COHERENCE TOMOGRAPHY AND CONVOLUTIONAL NEURAL NETWORKS 64
3.1 Introduction 64
3.2 Methods 66
3.2.1 Experimental setup 66
3.2.2 Data acquisition 68
3.2.3 CNN method for tissue classification 69
3.2.4 CNN regression for distance measurement 71
3.3 Results 72
3.3.1 Imaging results of endoscopic OCT system 72
3.3.2 Multi-class classification 74
3.3.3 Distance estimation 77
3.4 Discussion 80
CHAPTER 4: LAPAROSCOPIC OPTICAL COHERENCE TOMOGRAPHY SYSTEM FOR 3D BLADDER TUMOR DETECTION 86
4.1 Introduction 86
4.2 System Setup 88
4.3 Results and discussion 89
4.4 Conclusion and future work 91
CHAPTER 5: AUTOMATIC RENAL CARCINOMA BIOPSY GUIDANCE USING FORWARD-VIEWING ENDOSCOPIC OPTICAL COHERENCE TOMOGRAPHY AND DEEP LEARNING 93
5.1 Introduction 93
5.2 Materials and Methods 96
5.2.1Experimental setup 96
5.2.2 Data acquisition 98
5.2.3 Attenuation coefficient method for tissue recognition 99
5.2.4 CNN methods for tissue identification 100
5.3 Results 102

5.3.1 Imaging results of renal tumor and normal tissues 102
5.3.2 Classification using attenuation coefficient method 103
5.3.3 Tissue recognition results using CNN.....108
5.4 Discussion 112
5.5 Conclusion 115
CHAPTER 6: EPIDURAL ANESTHESIA NEEDLE GUIDANCE BY FORWARD-VIEW ENDOSCOPIC OPTICAL COHERENCE TOMOGRAPHY AND DEEP LEARNING 119
6.1 Introduction.....119
6.2 Results.....123
6.2.1 OCT images of five epidural layer categories.....123
6.2.2 Multi-class classification of OCT images by tissue layers using sequential binary method.....124
6.2.3 Estimation of the distance between the needle tip and dura mater by regression..128
6.3 Discussion 130
6.4 Method 132
6.4.1 Experiment setup.....132
6.4.2 Data acquisition.....134
6.4.3 Convolutional neural networks 136
6.4.4 Epidural distance prediction using deep learning 137
CHAPTER 7: ENHANCING EPIDURAL NEEDLE GUIDANCE USING A POLARIZATION-SENSITIVE OPTICAL COHERENCE TOMOGRAPHY PROBE WITH CONVOLUTIONAL NEURAL NETWORK 144
7.1 Introduction.....144
7.2 Method 147
7.2.1 Experimental setup.....147
7.2.2 Deep learning procedure 150
7.3 Results.....152
7.3.1 PS-OCT imaging results 152
7.3.2 Prediction results by deep learning 154
7.4 Discussion 159
CHAPTER 8: CONCLUSION AND FUTURE WORK 165

LIST OF FIGURES

Figure 1.1. Basic schematic of OCT. 1

Figure 1.2. An example of transverse scanning of OCT. 2

Figure 1.3. Light signal transmission in OCT. 6

Figure 1.4. Schematics of (A) SD-OCT and (B) SS-OCT. 7

Figure 1.5 (A) Side-view and (B) Forward-view endoscopic OCT. 11

Figure 1.6. Schematic of capsule/balloon endoscopic OCT. 12

Figure 1.7. Schematic of an ultracompact endoscopic OCT. 12

Figure 1.8. Two orthogonal polarization states. 16

Figure 1.9. Changes of polarization states by optical element rotation. 23

Figure 1.10. Schematic of endoscopic PS-OCT. 26

Figure 1.11. Doppler OCT for blood flow calculations. 31

Figure 2.1. Schematic of endoscopic OCT system. MZI: Mach-Zehnder interferometer; FC: fiber coupler; PC: polarization controller; GSM: galvonometer scanning mirror; BD: balanced detector; DAQ: data acquisition board. 43

Figure 2.2. Illustration of data acquisition and processing steps. 45

Figure 2.3. Sample kidney used for endoscopic OCT imaging; (B-D) Representative 3D OCT images, cross-sectional OCT images and the histology results of renal cortex (B), medulla (C) and calyx (D). 50

Figure 2.4. ROC multi-class testing curves of (A) kidney 5, (B) kidney 10 and (C) the average of all kidneys. 53

Figure 2.5. Class activation heatmaps of (A) RI ResNet50, and (B) PT ResNet50 for three representative images in each tissue type. The class activation heatmaps used jet colormap and were superimposed on the input images. 55

Figure 3.1. Schematic of endoscopic OCT system for Veress needle guidance. MZI: Mach-Zehnder interferometer, BD: Balanced detector, DAQ: Data acquisition, GSM: Galvanometer scanning mirror, FC: Fiber coupler, PC: Polarization controller. 67

Figure 3.2. Process of the data acquisition. 68

Figure 3.3. Introduction of the nested cross-validation, cross-testing, and 8-fold cross-validation process. 71

Figure 3.4. Examples of OCT images of different tissue types. 73

Figure 3.5. Aggregated ROC across all 8 testing folds using Xception model. 76

Figure 3.6. Tissue classification activation heatmap obtained from sample 1. Scale bar: 250μm. 76

Figure 3.7. (A) Examples of abdominal space images with different distances between small intestine and needle tip. Yellow value: estimated distance; Blue value: Manually calculated distance. (B) Violin plots showing the absolute error and relative error on the testing set for subject 1. 79

Figure 4.1. Schematic of laparoscopic OCT system. FC: fiber coupler; BD: balanced detector; DAQ: data acquisition; PC: polarization controller; GSM: Galvonometer scanning mirror. 89

Figure 4.2. (A) Image of the laparoscopic OCT system; (B) Camera image of UPII-SV40T transgenic mouse bladder tissue (The cancerous tissues were labelled by the blue dashed circle); (C) Tissue attenuation coefficient image map to recognize cancerous tissues. (D)&(E) 3D OCT image of the same bladder tissue with the tissue attenuation coefficient map. Image size: X×Y×Z=4.5×4.5×2 (mm3). GSM: galvonometer scanning mirror, GRIN: Gradient Index lenses. 90

Figure 4.3. (A) Image of the new-generation handheld OCT imaging probe. (B) Image of the whole laparoscopic OCT imaging system including the OCT machine, hand-held probe, and computer. 91

Figure 5.1. (A) Schematic of the forward-viewing OCT probe. FC: fiber coupler; PC: polarization controller; MZI: Mach–Zehnder interferometer; BD: balanced detector; DAQ: Data acquisition. (B) Picture of the human kidney with renal carcinoma. (C) Data acquisition of renal tumor tissue and normal tissue. (D) Attenuation coefficient and two CNN architectures used for tissue classification. 97

Figure 5.2. OCT imaging results of normal renal tissues, carcinoma, and oncocytoma (Scale bar: 250μm). Corresponding histology results (Scale bar: 500μm). 102

Figure 5.3. (A) Attenuation coefficient distribution results between normal tissue and tumor tissue. (B) ROC curves of the classification with attenuation coefficient between tumor tissue and other tissues. (C) Attenuation coefficient distribution results of all five tissue types. (D)

ROC curves of the classification with attenuation coefficient among different normal tissues. 105

Figure 5.4. ROC curves and pixel importance heatmaps of (A) ResNet50 and (B) InceptionV3. 109

Figure 6.1. (A) Endoscopic OCT scanner setup and the representative OCT images of five epidural tissue layer categories. (B) Histology results of different tissue layers. 124

Figure 6.2. Class activation heatmaps for Subject 7 using ResNet50 in cross-testing (A) and video captures of the insertion process (B). 127

Figure 6.3. (A) Examples of epidural space images with different distances between needle tip and spinal cord surface. G: labeled ground truth value (μm); P: prediction value (μm); Scale bar: 250μm. (B) The distribution of the predicted absolute percentage errors and absolute error in testing fold 7 with 3000 testing images. 129

Figure 6.4. Schematic of forward-view OCT endoscope system. 133

Figure 6.5. Data acquisition process 135

Figure 7.1. Experimental setup of the PS-OCT probe. 148

Figure 7.2: Data acquisition of the epidural tissues with the PS-OCT probe. 151

Figure 7.3. PS-OCT imaging results of epidural tissues and corresponding histology. 153

Figure 7.4. Grad-CAM and heatmap examples of different imaging results in four modes. 156

Figure 7.5. ROC curves of different imaging modes. 158

LIST OF TABLES

Table 2.1: The averages and standard errors of the validation accuracies from the nested cross-validation and testing procedure. The architecture with the highest validation performance in each testing fold is indicated in bold. 51

Table 2.2. Results from the 10-fold cross-testing. Average and standard error of accuracy are shown for two thresholds. 53

Table 2.3. Confusion matrix for classification of the three tissue types. 54

Table 3.1: The average nested 7-fold cross-validation accuracies and standard errors. 74

Table 3.2: The 8-fold cross-testing accuracies on the testing set in each testing fold. 75

Table 3.3. Average nested 7-fold cross-validation MAPE with standard error 78

Table 3.4. MAPE with Standard Error and MAE with Standard Error during 8-fold cross-testing. 78

Table 5.1. (A) Confusion matrix of the averaged five nested cross-validation kidney samples. (B) Prediction results: Accuracy; Precision; Recall; and F1 Score of different tissue types. 106

Table 5.2. (A) Confusion matrix of the test results from the hold-out set with additional four kidney samples. (B) Prediction results: Accuracy; Precision; Recall; and F1 Score of different tissue types. 107

Table 5.3. Prediction results of oncocytoma tissues using attenuation coefficient. 108

Table 5.4. Confusion matrices and performance metrics of the cross-validation and cross-testing using five kidney samples by (A) ResNet50 and (B) InceptionV3. 110

Table 5.5. (A) Confusion matrix and (B) performance metrics of the final production model on the hold-out set containing four additional kidney samples with carcinoma. 111

Table 5.6. Predictions of oncocytoma by the final CNN model. 112

Table 6.1. Average accuracies and standard error based on the practical tissue layer sequence during puncture for cross-validation. 125

Table 6.2. Average and standard error for cross-testing for the four binary comparisons for ResNet50. 126

Table 6.3. The average loss for each model type in cross-validation for each testing fold. 128

Table 7.1: Accuracies for cross-validation and cross-testing results in different imaging modes. 156

Table 7.2: Precision, Recall, and F1 of different tissue types in each model (Red font for the highest value and green for the lowest value). 157

CHAPTER 1: INTRODUCTION

Optical coherence tomography (OCT) is a novel optical biomedical imaging technique which performs high resolution (several micrometers) cross-sectional imaging of internal biological tissues. Real-time high-quality imaging can be achieved, so OCT has a broad range of the applications in biomedical research. This chapter provides a description of OCT technique, in terms of its working principles and clinical applications.

1.1 Working principles of OCT

OCT was first demonstrated in 1991 by Huang et al [1]. The basic principles of OCT are based on low coherence interferometry. Interferometry is a technology that uses light interference patterns containing wave superpose information, which including both amplitudes and phase of light waves. Coherence demonstrates the correlation between different wave phases. Broad linewidth laser source is applied to produce low coherence phenomenon. Low coherence interferometry can be used in biomedical applications for morphology imaging particularly in OCT.

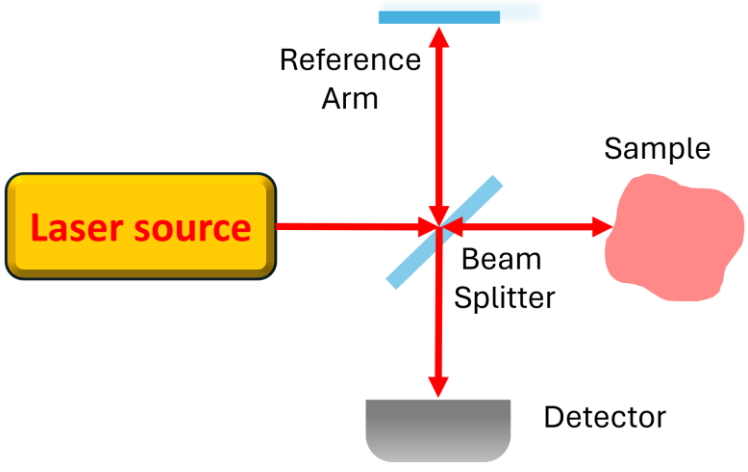


Figure 1.1. Basic schematic of OCT

OCT basic structure is based on a Michelson interferometer as shown in Figure 1.1. The input laser is directed to a beam splitter and separated to two light arms: reference arm and sample arm. The reference beam is reflected by a mirror, and the sample beam is backscattered by the sample tissue. The two beams combine again at the beam splitter and are received by a detector. The length of the reference arm is changed when the mirror in the reference arm is moved, leading to the interference phenomena. A wide range of interference of the relative length of the two arms can be detected when a high coherence (narrow linewidth) laser source is utilized. To achieve the imaging in specific distance, low coherence (broad linewidth) laser source is required in OCT system. Only in coherence length can the interference be detected while applying low coherence light. Through demodulating the interference signals, information of the backscattered light can be acquired. Cross-sectional imaging can be realized by continuous axial scanning in different transverse locations. In Figure 1.2, a two-dimensional (2D) OCT image of a kidney tissue under the observation of our endoscopic OCT system. Red arrows represent different in-depth axial scanning of the tissue.

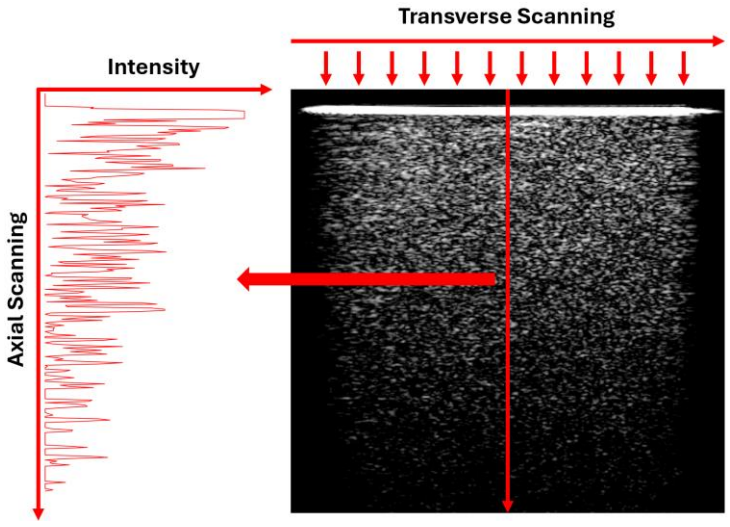


Figure 1.2 An example of transverse scanning of OCT

.

Compared to other imaging modalities, OCT provides very high axial imaging resolution. The axial resolution of OCT system is determined by the light coherence length, which is the spatial width of autocorrelation of the interference.

For conventional ultrasound system used in clinical applications, its working frequency is generally 10 MHz with the axial resolution of around 150 μm. For OCT scanning, axial resolution is determined by the coherence length of the laser source, and the transverse resolution is determined by the numerical aperture (NA) of the scanning lens. Currently, the OCT axial resolution reaches several micrometers. Some high-resolution OCT can provide 1 μm axial resolution for high quality tissue imaging.

1.2 OCT signal measurement

Since light is a type of wave with extremely high frequency, a high-speed measurement method is required to obtain the optical signal. In OCT, an approach for high-speed optical echo measurement is to apply second harmonic generation. The light source of OCT is pulse laser, and the backscattered light from the imaged samples in the sample arm can be obtained with the reflected pulse from the reference arm. We set the time delay between sample arm and reference arm as τ. Here we define I_s as the sample arm signal, and I_R as the reference arm signal, the received signal can be given as:

$$S(\tau) = \int_{-\infty}^{+\infty} I_s(t) I_R(t-\tau) dt \quad (1)$$

Time delay and the intensity of the signal can be acquired based on the above method. In OCT, low coherence interferometry was applied to measure the optical signal field. If we describe the electric field of a light (functional form) as:

$$E(t) = E\cos\left(2\pi\nu t - \frac{2\pi}{\lambda} z\right) \quad (2)$$

where $2\pi\nu$ is the angle frequency, and $2\pi/\lambda$ is the wave number. When the lights from sample arm and reference arm are combined (at the beam splitter), the output interferometry light signal can be given as:

$$E_{Output} = E_s(t) + E_R(t) \tag{3}$$

A light detector (balanced detector in OCT) is used to detect the intensity of the output interferometry signal, which can be described as follows:

$$I_{Output} = \frac{1}{4}|E_s(t)|^2 + \frac{1}{4}|E_R(t)|^2 + E_s E_R \cos\left(2 \cdot \frac{2\pi}{\lambda} \cdot \Delta l\right) \tag{4}$$

where Δl represents the length difference between reference arm and sample arm. When the length of the reference arm changes (by moving the reflector/mirror), the interference will happen under different positions in sample arm, namely different depth of the sample. Coherence length is inversely proportional to the linewidth of the light. A laser source with large linewidth is applied in the OCT system, so the coherence length of the light is relatively low, and the interference can only occur in a small range. Only when the optical lengths of the two arms are similar can the interference be observed. Based on this phenomenon, OCT can obtain the tissue structural information in a specific imaging depth of the sample.

The resolution of OCT imaging is related to the laser source. Here we discuss axial resolution and transverse resolution, respectively.

The axial resolution is based on the coherence length of the OCT laser source. Since the interferometer of the OCT produces autocorrelation field, its envelope can be considered as the power spectrum under Fourier transform. The bandwidth of the autocorrelation field represents the axial resolution, which is inversely proportional to the power spectrum bandwidth. If we consider the spectrum as an ideal Gaussian distribution, the OCT axial resolution can be given by:

$$\Delta z = \frac{2\ln 2}{\pi}\left(\frac{\lambda^2}{\Delta\lambda}\right) \tag{5}$$

where λ is the centered wavelength of the laser source, and Δ represents the full width at half maximum (FWHM). Thus, the axial resolution of OCT is inversely proportional to the laser linewidth.

The transverse resolution of OCT, similar to other microscopy, is determined by the spot size at the focal point of the light beam. This geometric feature is determined by the objective length of the OCT scanner, and the transverse resolution can be given by:

$$\Delta x = \frac{4\lambda}{\pi}\left(\frac{f}{d}\right) \tag{6}$$

where f represents the focal length, and d represents the spot size at the focal point produced by the objective lens. Therefore, transverse resolution is determined by the numerical aperture (NA). If we use the objective lens with larger NA a higher transverse resolution can be provided. However, larger NA leads to the limitation of imaging depth, so this trade-off has to be considered when designing the OCT system.

The basic optical structure of OCT is based on a Michaelson Interferometer as shown in Figure 1.1. Here we use E_s and E_R to represent the lights in sample arm and reference arm, respectively.

If we set the length of the reference arm as L_R, and the sample arm as L_S, the amplitude of E_R as A_R and E_s as A_s. E_R and $E_s$can be given as follows:

$$E_R = A_R exp\left(2\pi\nu t - \frac{2\pi}{\lambda}\cdot 2L_R\right) \tag{7}$$

$$E_s = A_s exp\left(2\pi\nu t - \frac{2\pi}{\lambda}\cdot 2L_S\right) \tag{8}$$

At the detector, the output photocurrent can be described by:

$$I = \langle\frac{\eta e}{h\nu}\frac{|E_R+E_S|^2}{2\eta_0}\rangle = \frac{\eta e}{h\nu}\cdot\frac{1}{\eta_0}\left[\frac{1}{2}|A_R|^2 + \frac{1}{2}|A_S|^2 + A_R A_S\cos\left(2\cdot\frac{2\pi}{\lambda}\cdot\Delta L\right)\right] \tag{9}$$

where η represents the quantum efficiency of the detector, e represents the electronic charge, $h\nu$ represents the energy of photon, η_0 represents the intrinsic impedance corresponding to free space, < > represents the average sign, and ΔL represents the difference of length between reference arm and sample arm. Signal transmission can be described as shown in Figure 1.3. Therefore, the received photocurrent signal is relative to the ΔL, imaging information of sample in different depth can be described. Currently, OCT includes two basic types: time domain OCT (TD-OCT) and Fourier domain OCT (FD-OCT).

TD-OCT is the traditional type of OCT. In TD-OCT, the interference signal from a sample in different depth is captured based on different reference arm length. L_R works as a function of time, and the carrier frequency is generated to make it convenient for lock-in detection, enhancing the signal sensitivity and improving the quality of the measurement.

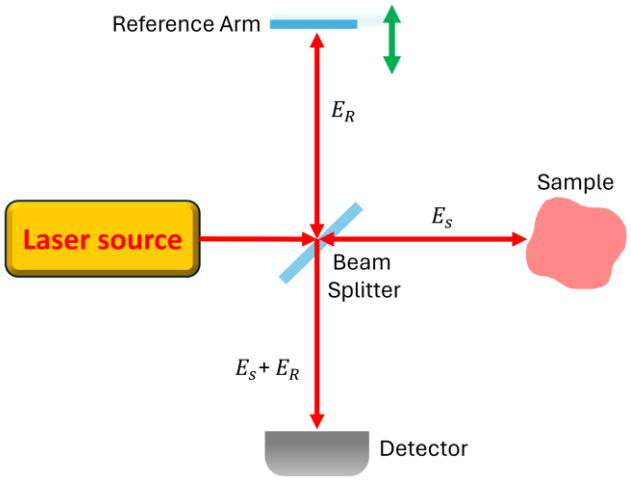


Figure 1.3. Light signal transmission in OCT

FD-OCT provides the signal using Fourier analysis method to reconstruct the tissue information in different imaging depths. It includes two basic types as shown in Figure 1.4: spectral-domain OCT (SD-OCT) and swept-source OCT (SS-OCT). For SD-OCT, it utilizes a laser source with broad linewidth, and a spectrometer to obtain the interference optical signals. It captured the entire spectrum of the backscattered light signal. For SS-OCT, it

utilizes a tunable laser as the light source, sweeping across a specific range of wavelengths. The signals upon different frequencies can be obtained based on a sequence.

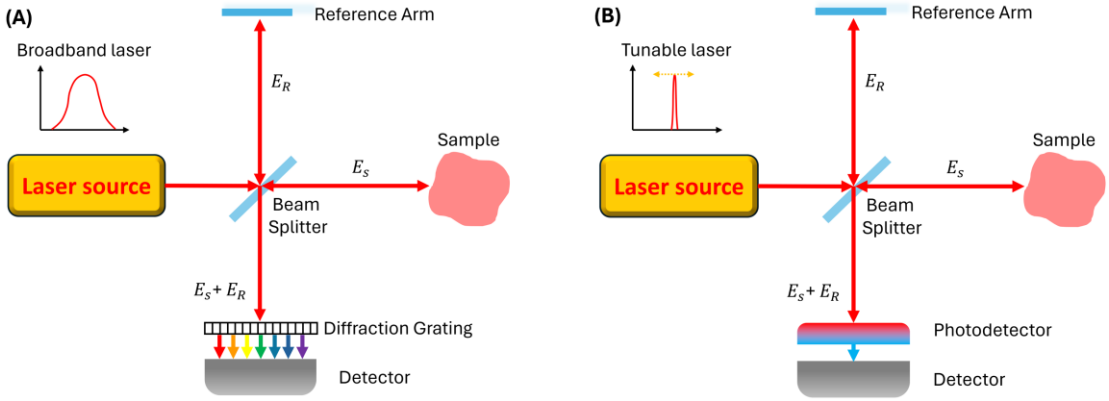


Figure 1.4. Schematics of (A) SD-OCT and (B) SS-OCT.

1.3 Applications of OCT

1.3.1 Early ophthalmologic applications of OCT technology

Because eye is accessible for light to enter, OCT technique was initially widely used in ophthalmology. OCT images of in vitro human retina, optic disk, and coronary artery were reported by Huang, et al. in the paper first demonstrating OCT technology [1]. The laser source they used was a low coherence super luminescent diode with the central wavelength of 830 nm. Imaging results were performed with 15 μm axial resolution. Anatomic structures of renal nerve fibers could be identified in the OCT images because of the cylindrical features of nerve fibers. Coronary artery was successfully examined by OCT, clearly showing the fatty-calcified plaque and plaque-normal arterial wall. In vivo human retina OCT imaging was realized by Hee et al [2]. The anatomic features of the nerve fiber could be clearly demonstrated in the high-quality OCT imaging results. The system utilized a laser source with 800 nm centered wavelength, performing 10 μm axial resolution.

OCT was investigated for ophthalmologic disease diagnoses in early studies in 1990s, such as macular diseases [3], cataract [4], glaucomatous eyes [5], optic disk pits and associated maculopathy [6], etc. Features of retinal nerve fiber, such as the thickness, could be interpreted by using OCT.

1.3.2 Applications of OCT in surgical guidance

OCT has been utilized in clinical applications other than ophthalmology. Light with longer wavelengths penetrates deeper through tissues because they have lower scattering coefficient [7], so OCT with long wavelength laser source has also been feasible in diagnoses during different surgeries. It has been reported that OCT could be employed to help the biopsy procedures for surgical diagnostics [8]. A super luminescent laser diode with a central wavelength of 1300 nm was used, performing an axial scan resolution of 16 μm. Different tissues including peripheral nerves fascicles, internal microvascular elastic membranes, and the granular layer of the cerebellum. This is potentially applicable in surgical interventions to help reduce the morbidity.

Except for 1300 nm laser source, other wavelengths have also been applied in OCT imaging of different clinical applications. OCTs with 1060 nm laser source have been used for the imaging of brain of Alzheimer's disease patients [9], human skin evaluations [10], intra-cochlear microstructural visualization [11], etc. OCTs using 1700 nm laser source have been applied for thyroid gland imaging [12], high-contrast deep tissue imaging of mouse brains [13], characterization of atherosclerosis [14], etc.

OCT has been utilized in different surgical procedures and diagnosis. It has been reported that OCT was applied in kidney related clinical applications. A 1300 nm high-speed OCT was used to observe the in situ living kidneys and accurately evaluate the rat kidney

features, such as the change of lining epithelium and the volumetric changes in tubular lumens under different conditions [15]. Also, the imaging condition during renal ischemia was conducted and analyzed. It has been proved that the morphology of uriniferous tubules and corpuscles could be observed by 3D OCT imaging. The volume of tubular lumens can be evaluated by 3D visualization based on OCT imaging results. Therefore, OCT can be considered as a non-invasive method for observing living kidneys. For kidney transplantation, OCT has been used to image the renal microstructures of living human kidneys and evaluate the acute tubular necrosis [16]. With the imaging depth of 2 mm, an entire human kidney can be non-invasively imaged by OCT within minutes, proving its potential in characterization of human kidney for transplantation. OCT could also help reveal different anatomic renal tissues such as glomeruli, convoluted tubules, renal collecting tubules, and loops of Henle [17]. Renal tumors can also be identified through the morphological differences shown by OCT imaging results. In addition, OCT has been applied for the percutaneous coronary intervention in chronic renal diseases [18], diabetic nephropathy observation [19], and renal tubular function and morphology evaluation [20].

1.4 Endoscopic OCT

Based on the high performance in tissue imaging, OCT has potential to be applied inside the body for disease diagnosis and surgical navigation. Currently, endoscopy with visualization and biopsy has been the gold standard of disease diagnosis in different clinical applications, especially cancerous tissue recognition [21]. The cure rate of any disease can be significantly improved if the disease, such as cancer, is treated at an early stage before metastasizing. Therefore, early detection of lesions is of great importance in clinical treatments. By using endoscopy with biopsy, lesions can be first recognized by imaging

results, and then recognized by histopathology analysis. The development of OCT has offered the feasibility of combining endoscopy techniques. Therefore, endoscopic OCT was proposed and successfully implemented across a variety of imaging applications for internal organs and tissues. Since OCT can provide the spatial resolution of several micrometers, it is applicable to identify the lesion tissue type and distinguish different normal tissues during the endoscopic imaging procedures. In this section, we will endoscopic OCT and its applications in different clinical areas.

1.4.1 Development of endoscopic OCT

OCT can provide high-quality imaging results with a lot of details of the tissue. However, the imaging depth of OCT can only reach several millimeters due to the strong light attenuation. To achieve subsurface imaging, endoscopic OCT has been developed and has become a critical component for the OCT imaging of internal organs.

The basic working principle of endoscopic OCT system is to transmit the light from the laser source to the sample tissue through the endoscope, collect the backscattered light signal with the tissue information, and transmit back to the OCT system through the same endoscope again. In 1996, G. J. Tearney, et al. first reported a single-mode OCT catheter [22]. They demonstrated an OCT catheter-endoscope using a gradient-index (GRIN) lens to realize the light transmission between the sample and OCT scanner. It was tested on *in vitro* human venous morphology. From then on, endoscopic OCT systems for different applications have been developed. According to the scanning directions, endoscopic OCT systems include two basic types: side-view endoscopic OCT and forward-view endoscopic OCT. Schematics of these two types are shown as follows in Figure 1.5.

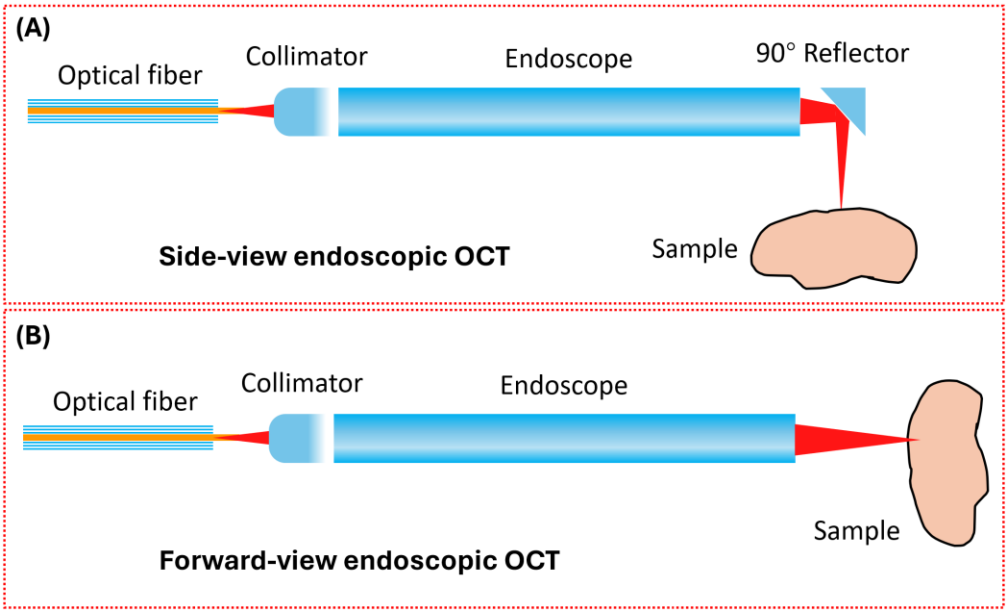


Figure 1.5 (A) Side-view and (B) Forward-view endoscopic OCT.

A side-view endoscopic OCT is usually used for large area scanning, especially for lumen inner imaging and inside organ scanning. By integrating specifically made motors and probe rotation components, the side-view endoscopic OCT is able to scan a whole round of the inner structure of the sample. By contrast, a forward-view endoscopic OCT is generally used for scanning a small and localized area, where the specific tissue type should be identified, such as surgical guidance and tissue biopsy. A cover (plastic or stainless-steel sheath) of the endoscope is commonly used for easy disinfection and protecting the endoscope from being damaged. The endoscope can keep the spatial resolution when the beam is transferred from one end to the other, so the tissue information will not be lost during the imaging procedure. For specific applications, endoscopic OCT has been designed to be different structures.

Capsule/Balloon endoscopic OCT has been designed to realize the imaging inside lumen areas such as esophagus or blood vessels [23-25]. Under those conditions the imaging probe requires a long working distance to arrive at the scanning destination. For that purpose,

a small capsule/balloon structure which includes the side-view endoscope was proposed as shown in Figure 1.6. The capsule/balloon is placed in the imaging area and taken out after the imaging procedure. For instance, a capsule based endoscopic OCT has been designed for esophagus internal imaging [26]. The capsule is attached by a tether and can be swallowed by the patients without any anesthesia requirement. When the capsule is moving by the peristaltic force of the esophagus or pulling the tether, OCT imaging of different areas in the esophagus can be conducted. By rotating the optical fiber rotary joint, the scanning direction of the capsule can be changed. Three-dimensional imaging can be achieved by the movement and rotation of the capsule. Based on the ultrafast micromotor and scanning components, a high-speed OCT internal imaging has been successfully realized.

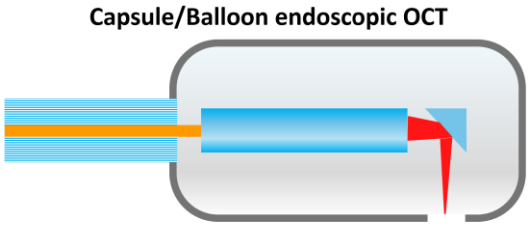


Figure 1.6. Schematic of capsule/balloon endoscopic OCT.

In order to image small internal tissue or lumens, ultracompact endoscopic OCT has been developed (Figure 1.7). Similar to capsule/balloon endoscopic OCT, the needle head of ultracompact endoscopic OCT is protected by a sheath which is polished to be sharp and compact. This structure makes it convenient for the access to the inside tissue or small lumens, and the post biopsy procedure.

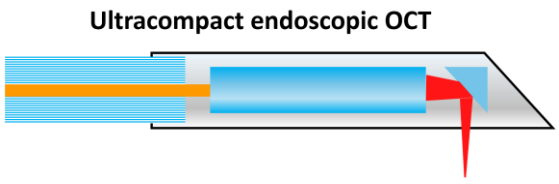


Figure 1.7. Schematic of an ultracompact endoscopic OCT.

1.4.2 Clinical applications of endoscopic OCT

Gastrointestinal (GI) tract imaging: One of the most common utilizations of endoscopic OCT is the diagnosis of Barrett's Esophagus (BE) through the GI tract imaging. BE is a complication that the swallow tube becomes damaged by gastroesophageal reflux disease (GERD) which affects 25%-35% of the U.S. population. Statistically, the population of BE is approximately 3.3 million in the U.S [27]. The normal squamous mucosa tissue will transform to glandular mucosa, which leads to over 30 times greater risk of developing into esophageal adenocarcinoma (EAC). However, EAC is often diagnosed at a late stage when the symptoms have been severe. The five-year survival rate of EAC is approximately 13%.

Therefore, early monitoring of BE is of great importance. Currently the BE is usually diagnosed based on the endoscopy biopsy. Nevertheless, the biopsy procedure requires a lot of time, and it is expensive. Based on the high spatial resolution, OCT has been an available imaging technology for BE lesion tissue recognition. Endoscopic OCT was first proposed for human GI imaging in 1999 [28]. A linear-scanning optical fiber-based catheter OCT was developed for *in vivo* human esophagus imaging. A cross-sectional OCT image could be obtained within 0.25s. Different layers of esophagus, including squamous epithelium, lamia propria, muscularis mucosa, submucosa, muscularis propria, with surrounding glands and vessels, could be clearly distinguished from the OCT imaging results. With the developments of capsule/balloon-based OCT probe and the high-speed scanner/data processor, novel endoscopic OCT was proposed to further identify the abnormal tissue at different stages. By applying an ultrahigh resolution OCT (UHR OCT) whose axial resolution reached 5μm, normal esophagus, BE, high-grade dysplasia and EAC could be differentiated from the exhibited OCT imaging features [29]. Further studies have been conducted to improve the

endoscopic imaging in different aspects, such as the device design, operational procedures, capsule contact, and patient toleration [30].

Vascular system imaging: Internal imaging of intravascular disease is also a common application of endoscopic OCT. Various OCT needle systems have been developed to realize the inside imaging of blood vessels. Coronary artery wall imaging is an important area of using endoscopic OCT. By applying the side-view endoscopic OCT, the artery wall can be 3D reconstructed and evaluated. Based on its high spatial resolution, endoscopic OCT can provide the detailed information of vascular features, such as the thrombus, lipid, cholesterol crystals, calcium, etc. [31-33], to assist the evaluation of the vascular plaque types. Different levels of those biological features could be shown by the OCT imaging characteristics. For instance, fibrosis plaque demonstrated a bright and thick band, but calcium plaque usually showed patterns that were not heterogeneous in OCT imaging results. Therefore, endoscopic OCT can assist in stent placement and peripheral artery atherectomy procedures.

In recent years, endoscopic OCT has been further developed to achieve more functions. A flexible micro endoscopic OCT has been proposed for 3D arterial microstructure imaging [34]. It was designed based on a micro-OCT (μOCT) which enables the visualization of tissues at sub-cellular level. A binary phase spatial filter was utilized to further increase the imaging depth. For the purpose of reducing the trauma to tissue, miniaturization of the endoscopic OCT probe is important for minimally invasive surgical procedures. An ultrathin probe, which was produced using 3D printing techniques, has been proposed to overcome the limitations of the current probe size [35]. A probe with a diameter of less than 0.5 mm was developed and tested feasible on human arteries.

Urinary tract imaging: Bladder cancer originates from the cells in the urothelial tissue. It requires frequent follow-up tests after treatment because of the high relapse rate. The tumor cells usually exist under the surface of the bladder, so conventional endoscopy technique is difficult to detect them. Since OCT can provide subsurface imaging of tissue, it is applicable in the detection of bladder cancerous tissue underneath the bladder. The greatest advantage of using the OCT probe in bladder cancer diagnosis is that a noninvasive examination can be achieved. Relative cystoscopy OCT probes have been proposed to identify muscle-invasive bladder tumor tissue. It has been reported that through using OCT with a conventional cystoscopy, bladder tumor recognition accuracy could reach over 90% [36]. Moreover, endoscopic OCT was integrated with optical coherence angiography technology to realize the visualization of microvasculature, assisting the pathology of bladder normal and abnormal tissues [37].

Except for bladder, endoscopic OCT has been applied in ureter related disease diagnosis. Different layers of the ureteral wall could be visualized and distinguished [38]. It has been reported that endoscopic OCT could provide intraoperative information of ureter tumor grade and stage, which could help urologist make the decision for treatment. Moreover, endoscopic OCT is able to visualize different tissue types of prostates, including neurovascular bundles, fat tissue, prostate capsule, and pelvic fascia [39]. As a minimally invasive imaging modality, endoscopic OCT works as an effective tool for intraoperative diagnosis of prostate cancer [40].

1.5 Polarization Sensitive OCT (PS-OCT)

Polarization-sensitive OCT (PS-OCT) utilizes the polarization features of light to provide contrast to the biological sample imaging results. By analyzing the polarization state

of light in the recorded interference fringe intensity, PS-OCT can provide more detailed spatial information based on the polarization of light backscattered by tissue [41]. Owing to the diversity in the impact of each scattering event on the incident polarization state, the cumulative effect of these events leads to changes in different polarization directions. This phenomenon can happen on specific tissues, such as fibrosis tissues, muscle, tendon, cartilage, etc.

1.5.1 Polarization state of light

To articulate the polarization states of light signals within PS-OCT, the electromagnetic waves can be represented in the form of complex electric field vectors, as follows:

$$\vec{E}(z,t) = Ee^{i(kz-\omega t)} \tag{10}$$

where the temporal angular frequency $\omega = 2\pi c/\lambda$, the spatial angular frequency (wave number) $k = 2\pi/\lambda$. Additionally, z denotes the optical path that the light transmits through, and t corresponds to the time associated with this transmission.

Here we decompose the light into two components with orthogonal vibration directions: x as the horizontal component, and y as the vertical component, to describe orthogonal polarization states as shown in Figure 1.8.

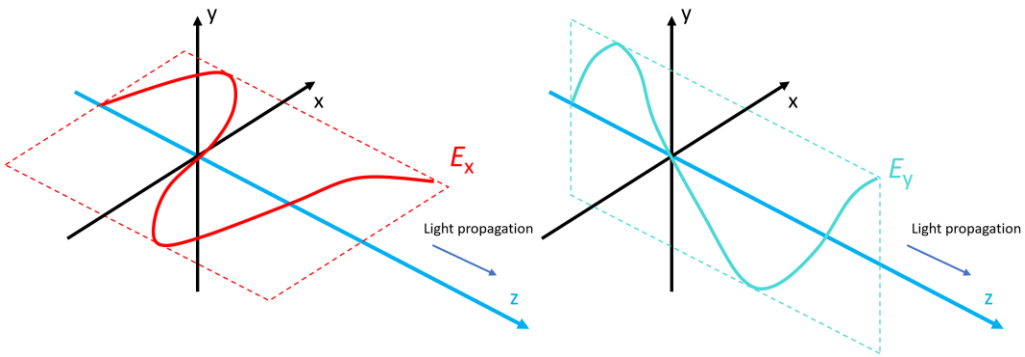


Figure 1.8. Two orthogonal polarization states.

Here we describe the electric field as the following format:

$$\vec{E}(z,t) = \overrightarrow{E_x} + \overrightarrow{E_y} = A_x e^{-i(kz-\omega t-\delta_x)}\overrightarrow{e_x} + A_y e^{-i(kz-\omega t-\delta_y)}\overrightarrow{e_y}$$
$$= E_x e^{-i(kz-\omega t)}\overrightarrow{e_x} + E_y e^{-i(kz-\omega t)}\overrightarrow{e_y} \quad (11)$$

where $E_x = A_x e^{i\delta_x}$,$E_y = A_y e^{i\delta_y}$. A_x , A_y represent the amplitudes, $\overrightarrow{e_x}$ and $\overrightarrow{e_y}$ represent orthonormal vectors on horizontal (x) and vertical (y) axes, and δ_x , δ_y are the phases. To demonstrate the relationship between time and amplitude for the electric field E, the real part of it can be shown as follows, when the optical path is zero:

$$E_{Re} = Re\left[\vec{E}(0,t)\right] = A_x \cos(\omega t + \delta_x)\overrightarrow{e_x} + A_y \cos(\omega t + \delta_y)\overrightarrow{e_y} \quad (12)$$

Here, E_{Re} is referred to as the vibration ellipse, which serves as a tool for identifying different light polarization states. Here the phase difference is defined as $\Delta\delta = \delta_x - \delta_y$. When $A_x = 0$, the amplitude in horizontal vibration axis is 0. When the vibration ellipse collapses to a straight line oriented along the vertical axis, this configuration signifies that the light is linearly polarized in the vertical direction. Similarly, when $A_y = 0$, the linearly polarized light vibrates in horizontal axis.

Under the condition where the amplitudes $|A_x| = |A_y|$, the phase difference $\Delta\delta$ emerges as a critical parameter in describing the polarization state of light. Specifically:

When $\Delta\delta = n\pi$, where n belongs to the set of integers $\mathbb{Z}$, the light is in linear polarization state.

When $\Delta\delta = n\pi + \frac{\pi}{2}$, where $n \in \mathbb{Z}$, the light is in circular polarization state.

In conditions other than those specified, the light is in an elliptical polarization state, which represents a general type of polarization where the electric field vector forms an ellipse in the plane perpendicular to the propagation direction.

Here we use the vectors $\overrightarrow{e_x}$ and $\overrightarrow{e_y}$ as orthonormal basis to demonstrate the special polarization states as follows:

$\overrightarrow{e_{+\frac{\pi}{4}}} = \frac{\sqrt{2}}{2}\left(\overrightarrow{e_x} + \overrightarrow{e_y}\right)$ and $\overrightarrow{e_{-\frac{\pi}{4}}} = \frac{\sqrt{2}}{2}\left(\overrightarrow{e_x} - \overrightarrow{e_y}\right)$ represent the linear polarization states at +45° and -45°.

$\overrightarrow{e_R} = \frac{\sqrt{2}}{2}\left(\overrightarrow{e_x} - i\overrightarrow{e_y}\right)$ and $\overrightarrow{e_L} = \frac{\sqrt{2}}{2}\left(\overrightarrow{e_x} + i\overrightarrow{e_y}\right)$ represent the right-hand and left-hand circular polarization states.

1.5.2. Diattenuation and birefringence

The polarization state of light can be characterized by the amplitude ratio and phase difference between the x and y components, which are able to represent any orthogonal vibration directions. These parameters change as light propagates through a specific medium, illustrating the dynamic nature of polarization state of light in response to the biological tissue properties.

Diattenuation refers to the property of a material or a tissue in which the transmittance varies based on the polarization state of the incident light. Tissues with diattenuation features exhibit preferential absorption along one specific axis, resulting in differential transmission for light polarized along different axes. This property can significantly influence the polarization characteristics of light passing through such tissues, leading to changes in its polarization state dependent on the polarization orientation relative to the optical axes of tissue. The amplitude ratio between orthogonal components of light can be altered by diattenuation phenomena. Diattenuation occurs when light passes through materials equipped with a mechanism, such as dichroic material, leading to differing transmittances for light polarized in two orthogonal directions. For instance, previous studies have reported that brain

tissue provided diattenuation effects on 3D-polarized light imaging [42, 43]. The degree of diattenuation is determined by the ratio of the difference to the sum of transmittances along the orthogonal vibration axes. As to the dichroic material, if we define TD_x and TD_y as the transmittances parallel and perpendicular to the optical axis, respectively, the diattenuation effect observed when linearly polarized light is aligned with the optical axis of the material can be quantified as follows:

$$diattenuation = \frac{TD_x - TD_y}{TD_x + TD_y} \tag{13}$$

Birefringence is precisely defined as the phenomenon of double refraction of light within a transparent, molecularly ordered samples. This effect is characterized by the presence of polarization orientation-dependent differences in the refractive index. Light traveling through such a sample tissue can be split into two polarized directions at different transmitting speeds. each direction corresponding to one of the two distinct refractive indices. This divergence leads to a phase difference between them, altering the polarization state based on the initial polarization orientation and the optical axis. Because of the difference of refractive index between the directions along and perpendicular to the optic axis, a phase retardation is provided to the orthogonal electric field components.

$$\eta = 2\pi\Delta n x/\lambda \tag{14}$$

where Δn is the refraction index difference between the polarization direction parallel and vertical to the birefringent sample tissue optic axis, x is the distance that light transmits through, and λ is the light wavelength. It is worthy to note that a shift of $\pi/2$ is required in the PS-OCT system. An optical element that provides a phase shift of $\pi/2$ between polarization components is known as a quarter-wave plate (QWP), which is utilized in the reference and sample arm of PS-OCT system to introduce phase difference between the two split lights.

1.5.3. Jones formalism

The propagation of polarized light waves and their changes in polarization state can be effectively described using the Jones formalism. Within this framework, light waves are represented by Jones vectors, which encode the amplitude and phase of the electric field components of the wave in a specific polarization state. Optical elements, such as waveplates, mirrors, or lenses, which can change the polarization state of the light, are characterized by Jones matrices. The Jones matrices function on the Jones vectors to describe how different optical elements modify the polarization state of the light transversing them. Through this formalism, changes in the light amplitude and phase on orthogonal polarization orientations of the electric field vector can be precisely calculated and quantified.

Jones vector describes the amplitude ratio and phase retardation of the two orthogonal polarization axes. It can be described in a form of 1×2 complex vector:

$$\vec{E} = E_x\overrightarrow{e_x} + E_y\overrightarrow{e_y} = \begin{bmatrix} E_x \\ E_y \end{bmatrix} = \begin{bmatrix} A_x e^{-i\delta_x} \\ A_y e^{-i\delta_y} \end{bmatrix} \quad (15)$$

Here for the six common examples of polarization states, their Jones vectors are shown below:

Linear polarized in horizontal (x) axis: $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$

Linear polarized in vertical (y) axis: $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$

Linear polarized at +45°: $\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$

Linear polarized at +45°: $\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$

Right hand circular polarized state: $\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -i \end{bmatrix}$

Left hand circular polarized state: $\frac{1}{\sqrt{2}}\begin{bmatrix}1\\+i\end{bmatrix}$

The Jones vector adheres to the superposition principle, indicating that for monochromatic lights sharing the same propagation direction, a single Jones vector can be derived from the sum of all their individual Jones vectors. This principle demonstrates the linear nature of Jones formalism. By accumulating the Jones vectors of each independent light wave, one Jones vector can accurately describe the resultant polarization state of the combined light waves. It offers a framework for analyzing the interaction and superposition of multiple polarized monochromatic light waves within the same propagation direction.

The Jones matrix is used to describe the transformation in the polarization state of light as it transmits through different optical elements. It mathematically represents the effects of optical components, such as lenses, mirrors, waveplates, and polarizers, on the polarization characteristics of light. By applying a Jones matrix to a Jones vector, the final polarization state of the light can be quantitatively described. Jones matrix is a 2×2 complex matrix that transform the input polarization state $\vec{E}$ to the output polarization state $\vec{E}'$ as shown below:

$$\vec{E}' = J\vec{E} = \begin{bmatrix}J_{xx} & J_{xy}\\ J_{yx} & J_{yy}\end{bmatrix}\begin{bmatrix}E_x\\ E_y\end{bmatrix} = \begin{bmatrix}E_x'\\ E_y'\end{bmatrix} \tag{16}$$

where $E_x' = J_{xx}E_x + J_{xy}E_y$, $E_y' = J_{yx}E_x + J_{yy}E_y$.

For example, the Jones matrices of a linear polarizer with the optic axis of horizontal direction (x) and vertical direction (y) are shown as:

$$J_H = \begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix} \text{ and } J_V = \begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix} \tag{17}$$

The right circular polarizer and left circular polarizer can be described as:

$$J_R = \frac{1}{2}\begin{bmatrix}1 & i\\ -i & 1\end{bmatrix} \text{ and } J_L = \frac{1}{2}\begin{bmatrix}1 & -i\\ i & 1\end{bmatrix} \tag{18}$$

It should be noted that within Jones formalism, the direction of rotation is described from the perspective of the wave receiver. This convention is crucial for determining the polarization state and understanding the effects of optical elements.

For a phase retarder, where the horizontal (x) and vertical (y) components of the beam are aligned to the fast and slow axes of the wave plate, respectively, the corresponding Jones matrix can be shown as follows [44]:

$$J_{WP} = \begin{bmatrix} e^{i\eta/2} & 0 \\ 0 & e^{-i\eta/2} \end{bmatrix} \tag{19}$$

where η represents the phase retardation, which is the phase difference between the lights along the fast and slow axes, introduced by the phase retarder.

When a light beam transmits through multiple optical components, the Jones vector of the output polarization state can be calculated by the product of all the Jones matrices from each individual optical device as:

$$\vec{E}^{(n)} = J_n J_{n-1} \dots \dots J_2 J_1 \vec{E}^{(1)} \tag{20}$$

Nevertheless, polarization states are not always aligned to the optic axis of the elements ideally. When a polarizer or a wave plate is rotated by an azimuth of θ, rotation matrices are necessary to be considered in the calculation of Jones formalism to accurately account for the adjustments. Here if we describe the input rotation and output rotation in Figure 1.9 as follows:

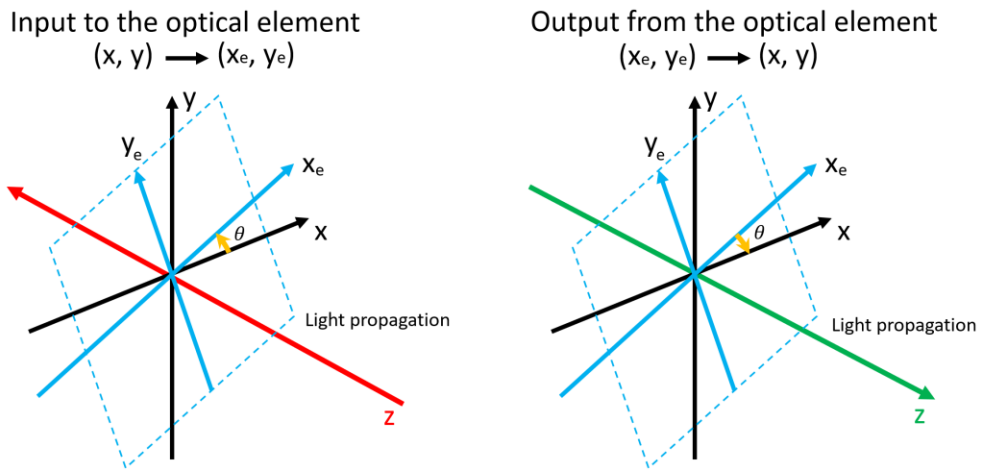


Figure 1.9. Changes of polarization states by optical element rotation.

From the above figure, the input rotation and the output rotation of the Jones matrices can be expressed as:

$$R_{in} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \text{ and } R_{out} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \tag{21}$$

Here we define R_{out} as the rotation matrix $R(\theta)$. The Jones matrix for any wave plate with an azimuth of θ can be represented as follows:

$$J'_{WP} = R_{out} J_{WP} R_{in} = R(\theta) J_{WP} R(-\theta) \tag{22}$$

Under more general case, circularity should be included [45].

$$R_{\phi} = \begin{bmatrix} e^{-i\phi/2} & 0 \\ 0 & e^{i\phi/2} \end{bmatrix} \tag{23}$$

$$J_{WP}'' = R_{\phi} J_{WP}' R_{-\phi} \tag{24}$$

For tissue samples with diattenuation features, here we define the transmission coefficients for the horizontal (x) and vertical (y) axes as T_x and T_y. Since the Jones vector of the input beam is represented as $[E_x \quad E_y]^T$, and the output beam is represented as $[T_x E_x \quad T_y E_y]^T$, the Jones matrix of any tissue sample qqwith diattenuation features can be

given by: $\boldsymbol{J}_{diattenuation} = [T_x, \quad 0; \quad 0, \quad T_y]$. Therefore, the Jones matrix of a diattenuation optical system with an azimuth of θ can be given by:

$$\boldsymbol{J_d} = \begin{bmatrix} e^{-i\phi/2} & 0 \\ 0 & e^{i\phi/2} \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} T_x & 0 \\ 0 & T_y \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} e^{i\phi/2} & 0 \\ 0 & e^{-i\phi/2} \end{bmatrix}$$

$$= \begin{bmatrix} T_x \cos^2\theta + T_y \sin^2\theta & (T_x - T_y)\cos\theta \sin\theta e^{-i\phi} \\ (T_x - T_y)\cos\theta \sin\theta e^{i\phi} & T_x \sin^2\theta + T_y \cos^2\theta \end{bmatrix} \tag{25}$$

where ϕ is the circularity rate. For linear optical elements, $\phi = 0$.

Similarly, for any tissue sample with birefringence features, as the previous definition of the phase retardation as η, its Jones matrix can be given by $\boldsymbol{J}_{birefringence} = [e^{i\eta/2}, \quad 0; \quad 0, \quad e^{-i\eta/2}]$. Therefore, the Jones matrix of a birefringence optical system with an azimuth of θ can be given by:

$$\boldsymbol{J_b} = \begin{bmatrix} e^{-i\phi/2} & 0 \\ 0 & e^{i\phi/2} \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} e^{i\eta/2} & 0 \\ 0 & e^{-i\eta/2} \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} e^{i\phi/2} & 0 \\ 0 & e^{-i\phi/2} \end{bmatrix}$$

$$\begin{bmatrix} e^{i\eta/2}\cos^2\theta + e^{-i\eta/2}\sin^2\theta & \left(e^{i\eta/2} - e^{-i\eta/2}\right)\cos\theta \sin\theta e^{-i\phi} \\ \left(e^{i\eta/2} - e^{-i\eta/2}\right)\cos\theta \sin\theta e^{i\phi} & e^{i\eta/2}\sin^2\theta + e^{-i\eta/2}\cos^2\theta \end{bmatrix} \tag{26}$$

Practically, a tissue sample has both diattenuation and birefringence properties. T_x and T_y are the eigenvalues of J_d. $e^{i\eta/2}$ and $e^{-i\eta/2}$ are the eigenvalues of J_b. Based on the formula above, it can be derived that:

$$\boldsymbol{J_d} \begin{bmatrix} \cos\theta \\ e^{i\phi}\sin\theta \end{bmatrix} = T_x \begin{bmatrix} \cos\theta \\ e^{i\phi}\sin\theta \end{bmatrix},$$

$$\boldsymbol{J_d} \begin{bmatrix} -\sin\theta \\ e^{i\phi}\cos\theta \end{bmatrix} = T_y \begin{bmatrix} -\sin\theta \\ e^{i\phi}\cos\theta \end{bmatrix},$$

$$\boldsymbol{J_b} \begin{bmatrix} \cos\theta \\ e^{i\phi}\sin\theta \end{bmatrix} = e^{i\eta/2} \begin{bmatrix} \cos\theta \\ e^{i\phi}\sin\theta \end{bmatrix},$$

$$\boldsymbol{J_b} \begin{bmatrix} -\sin\theta \\ e^{i\phi}\cos\theta \end{bmatrix} = e^{-i\eta/2} \begin{bmatrix} -\sin\theta \\ e^{i\phi}\cos\theta \end{bmatrix}. \tag{27}$$

When we consider diattenuation and birefringence phenomena characterized by the same azimuth θ and circularity rate ϕ, the Jones matrices share a common set of eigenvectors because when two matrices have the same set of eigenvectors they satisfy the commutative law, allowing them to be multiplied in any order without changing the multiplication result [46]. In PS-OCT system, therefore, the sequence of the optical elements which light passes through does not affect the output polarization state when their optic axes are parallel or orthogonal to each other. Since diattenuation and birefringence effects both exist in different tissues, we must consider both of them in PS-OCT imaging procedures.

Here the optic axes of diattenuation and birefringence components in the calculation are aligned, the polarization properties can be effectively described within one Jones matrix as follows [47]:

$$\boldsymbol{J_s}(\eta,\theta,\phi)$$

$$= \begin{bmatrix} e^{-\frac{i\phi}{2}} & 0 \\ 0 & e^{\frac{i\phi}{2}} \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} T_x & 0 \\ 0 & T_y \end{bmatrix} \begin{bmatrix} e^{i\eta/2} & 0 \\ 0 & e^{-i\eta/2} \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} e^{i\phi/2} & 0 \\ 0 & e^{-i\phi/2} \end{bmatrix}$$

$$= \begin{bmatrix} T_x e^{i\eta/2}\cos^2\theta + T_y e^{-i\eta/2}\sin^2\theta & \left(T_x e^{i\eta/2} - T_y e^{-i\eta/2}\right)\cos\theta\sin\theta e^{-i\phi} \\ \left(T_x e^{i\eta/2} - T_y e^{-i\eta/2}\right)\cos\theta\sin\theta e^{i\phi} & T_x e^{i\eta/2}\sin^2\theta + T_y e^{-i\eta/2}\cos^2\theta \end{bmatrix} \quad (28)$$

For instance, if we put a QWP in the light path, the phase retardation $\eta = \pi/2$, $\phi = 0$. The corresponding QWP can be described as:

$$\boldsymbol{J_{QWP}}\left(\frac{\pi}{2},\theta,0\right) = \begin{bmatrix} \frac{\sqrt{2}}{2}(1 + i\cos 2\theta) & i\frac{\sqrt{2}}{2}\sin 2\theta \\ i\frac{\sqrt{2}}{2}\sin 2\theta & \frac{\sqrt{2}}{2}(1 - i\cos 2\theta) \end{bmatrix} \quad (29)$$

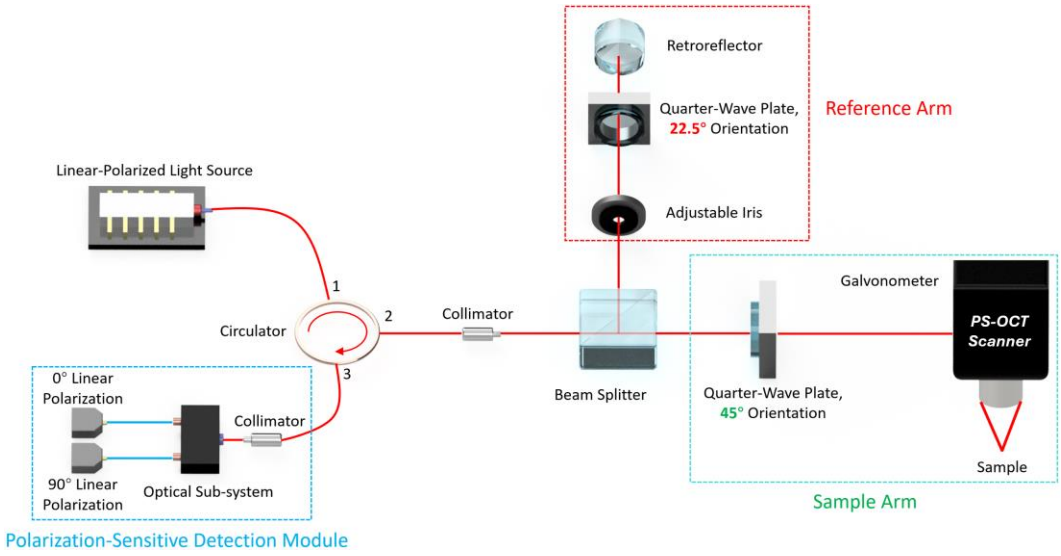


Figure 1.10. Schematic of endoscopic PS-OCT.

A basic schematic of PS-OCT is demonstrated in Figure 1.10. For the post processing of the interference signals, a QWP with a 22.5° orientation is placed in reference arm and another one with 45° orientation is placed in sample arm. Therefore, the corresponding Jones matrices of them can be given as:

$$\boldsymbol{J_{QWP,Reference}}\left(\frac{\pi}{2},\frac{\pi}{8},0\right)=\begin{bmatrix}\frac{\sqrt{2}}{2}(1+i) & \frac{1}{2}i\\ \frac{1}{2}i & \frac{\sqrt{2}}{2}(1-i)\end{bmatrix},$$

$$\boldsymbol{J_{QWP,Sample}}\left(\frac{\pi}{2},\frac{\pi}{4},0\right)=\begin{bmatrix}\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2}i\\ \frac{\sqrt{2}}{2}i & \frac{\sqrt{2}}{2}\end{bmatrix} \tag{30}$$

To describe the light transmission in reference arm, Jones formalism can be utilized. It is worth noting that when the light transmits through the opposite direction,

$$\theta_{-}=-\theta \tag{31}$$

The Jones matrix of a mirror can be given as:

$$\boldsymbol{M}=\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix} \tag{32}$$

The change of polarization state in reference arm can be described as:

$$\boldsymbol{E_R}' = \boldsymbol{J_{QWP-\frac{\pi}{8}} M J_{QWP+\frac{\pi}{8}} E_R} \tag{33}$$

If we set the output power of the linear polarized laser source as:

$$\boldsymbol{E} = E(z=0)\begin{bmatrix}1\\0\end{bmatrix} \tag{34}$$

Therefore,

$$\boldsymbol{E_R} = \boldsymbol{E_S} = \frac{\sqrt{2}}{2}E(z=0)\begin{bmatrix}1\\0\end{bmatrix} \tag{35}$$

In reference arm:

$$\boldsymbol{E_R'} = \frac{\sqrt{2}}{2}E(2z_R)\begin{bmatrix}\frac{\sqrt{2}}{2}+\frac{i}{2} & -\frac{i}{2}\\ -\frac{i}{2} & \frac{\sqrt{2}}{2}-\frac{i}{2}\end{bmatrix}\begin{bmatrix}1 & 0\\0 & -1\end{bmatrix}\begin{bmatrix}\frac{\sqrt{2}}{2}+\frac{i}{2} & \frac{i}{2}\\ \frac{i}{2} & \frac{\sqrt{2}}{2}-\frac{i}{2}\end{bmatrix}\begin{bmatrix}1\\0\end{bmatrix}$$

$$= \frac{\sqrt{2}}{2}E(2z_R)\times\frac{1}{4}\begin{bmatrix}2\sqrt{2}i & 2\sqrt{2}i\\ 2\sqrt{2}i & 2\sqrt{2}i\end{bmatrix}\begin{bmatrix}1\\0\end{bmatrix}$$

$$= \frac{i}{2}E(2z_R)\begin{bmatrix}1\\1\end{bmatrix} \tag{36}$$

In sample arm, similar calculations can be conducted based on the tissue characteristics.

The Jones matrices can be decomposed based on the commutative law. Thus, the diattenuation parameters T_x, T_y and phase retardation parameter η can be obtained to analyze the characteristics of the scanned tissues.

1.5.4 Mueller-Stokes Formalism

Jones matrix has been proved to be available in the description of polarization state in PS-OCT system. Nevertheless, Jones formalism cannot describe partially polarized light and

depolarization process. Therefore, Mueller-Stokes formalism is introduced to completely demonstrate the changes of polarization.

Stokes vectors, including the measurements of polarization characteristics, are used to accurately describe the polarization states as follows:

$$I = E_x E_x^* + E_y E_y^*,$$

$$Q = E_x E_x^* - E_y E_y^*,$$

$$U = E_{+45^o} E_{+45^o}^* - E_{-45^o} E_{-45^o}^* = E_x E_y^* + E_y E_x^*,$$

$$V = E_R E_R^* - E_L E_L^* = i\left(E_x E_y^* - E_y E_x^*\right). \tag{37}$$

where I represents the overall irradiance; Q represents the irradiance difference between the orthogonal (x and y) components; U represents the phase difference between the orthogonal components (to characterize the linear polarized light completely); and V represents the circularity of the light. Therefore, we can summarize the relation among the Stokes parameters as follows:

$$I^2 = Q^2 + U^2 + V^2 + 4\left(\langle a_x^2\rangle\langle a_y^2\rangle - \langle a_x a_y e^{i\delta}\rangle\langle a_x a_y e^{-i\delta}\rangle\right) \geq Q^2 + U^2 + V^2 \tag{38}$$

where the phase difference $\delta = \delta_x - \delta_y$. The degree of polarization can be defined as:

$$Degree\ of\ Polarization = \sqrt{\frac{Q^2+U^2+V^2}{I^2}} \tag{39}$$

Similar to Jones formalism, Mueller matrix is utilized to describe the change of light polarization from the optical elements:

$$S' = MS = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \\ m_{41} & m_{42} & m_{43} & m_{44} \end{bmatrix} \begin{bmatrix} I \\ Q \\ U \\ V \end{bmatrix} \tag{40}$$

Based on the measurements of Stokes parameters, the relationship between Mueller-Stokes formalism and Jones formalism can be given as:

$$\boldsymbol{S} = \boldsymbol{A}\left(\vec{E} \otimes \vec{E}^{*}\right),$$

$$\boldsymbol{M} = \boldsymbol{A}(\boldsymbol{J} \otimes \boldsymbol{J}^{*})\boldsymbol{A}^{-1}. \tag{41}$$

where $\otimes$ represents the Kronecker tensor product, and $\boldsymbol{A}$ can be described as:

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 0 & -1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & i & -i & 0 \end{bmatrix} \tag{42}$$

In practical system establishments, Mueller-Stokes formalism is commonly used to analyze the PS-OCT signal processing. The system we utilized in the studies introduced in the following chapters applies Mueller-Stokes formalism to demonstrate the imaging results. From the schematic shown in Figure 1.10, the PS-OCT signal detection module has two receivers for orthogonal polarized signals. Here we define their overall irradiances as I_1 and I_2, where

$$I_1 = E_1 \times E_1^{*},\ I_2 = E_2 \times E_2^{*} \tag{43}$$

Therefore, the other Stokes parameters can be given as:

$$Q = E_1E_1^{*} - E_2E_2^{*},$$

$$U = E_{+45^o}E_{+45^o}^{*} - E_{-45^o}E_{-45^o}^{*} = E_1E_2^{*} + E_2E_1^{*},$$

$$V = E_RE_R^{*} - E_LE_L^{*} = i(E_1E_2^{*} - E_2E_1^{*}). \tag{44}$$

To evaluate the polarization features of the tissue imaging results, three imaging modes including: 1) Phase retardation; 2) Optic axis; and 3) Degree of polarization uniformity, were proposed as follows to demonstrate the polarization features:

$$Phase\ retardation = tan^{-1}\left(\sqrt{I_1/I_2}\right) \in [0,\ \pi/2],$$

$$Optic\ Axis = \frac{1}{2}atan2(U, Q) \in [-\pi/2,\ \pi/2],$$

$$DOPU = \sqrt{\langle Q\rangle^2 + \langle U\rangle^2 + \langle V\rangle^2} \in [0,\ 1]. \tag{45}$$

1.6 Doppler OCT

Among all the wave phenomena, there is a change in frequency based on the relative motion of the wave source and observer. The frequency becomes higher when the source and observer approach each other, and lower when they move away from each other. Therefore, Doppler effect can be an effective tool for speed detection. In biomedical areas, Doppler effect has been applied to small particle detections. Doppler ultrasound has been widely used in blood flow detection by imaging the red blood cell motions. Blood vessel distribution can be identified in cardiovascular evaluation [48], peripheral vascular disease diagnosis [49], fetal assessment [50], etc. However, due to the limitation of ultrasound spatial resolution, it is still difficult for Doppler ultrasound to image the relatively small blood vessels.

Doppler OCT technique, also known as optical Doppler tomography (ODT), integrates the Doppler effect principles within conventional OCT system. Doppler OCT has the high spatial resolution of OCT, thus it enables the blood flow assessment of small blood vessels with more precise imaging results. Doppler OCT system was proposed to realize the imaging of fluid flow velocity [51]. With its development in recent years, Doppler OCT has become an efficient biomedical imaging tool in clinical applications.

1.6.1 Working principles of Doppler OCT

Based on Doppler effect, the light source in Doppler OCT system illuminates the target moving particle and the backscattered signal is collected. The Doppler frequency shift of the coherence light signal can be given by:

$$f_D = \frac{1}{2\pi}\left(\boldsymbol{k_{input}} - \boldsymbol{k_{output}}\right) \cdot \boldsymbol{v} \tag{46}$$

where k_{output} and k_{input} are the wave vectors of the output/input light, and v represents the velocity vector of the moving particles. If the Doppler angle is set as θ as shown in Figure 1.11, the Doppler frequency shift can be calculated as:

$$f_D = \frac{2v\cos\theta}{\lambda_0} \tag{47}$$

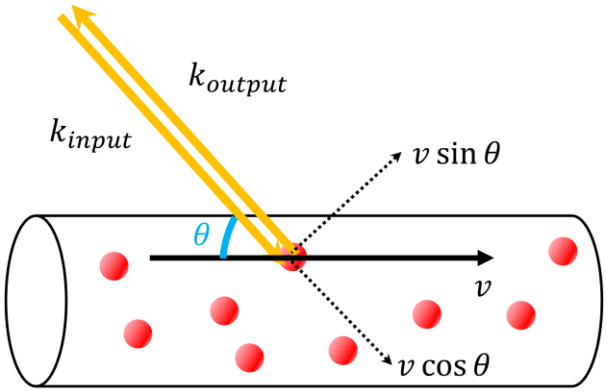


Figure 1.11. Doppler OCT for blood flow calculations.

where the λ_0 represents the center wavelength of the OCT laser source. By measuring the frequency change, the particle motion can be detected based on post signal processing.

1.6.2 Applications of Doppler OCT

Since Doppler OCT was proposed, its main function is based on the visualization and quantification of blood flow for the purpose of vascular evaluation [52]. With the development of laser, data acquisition devices and other optical components, the working efficiency of Doppler OCT has become better in resolution, sensitivity, and image processing speed. Compared to other blood vessel imaging techniques such as fluorescence, a prime benefit of Doppler OCT is its non-invasiveness without requiring any dye which has potential risk to human health. Therefore, it has potential in diagnosis of vascular diseases in different human organs.

Retina vascular visualization: Ophthalmology is a common use of OCT technique because the eye is transparent enough for light to transmit through. It has been reported that the blood vessels near the retinal optic nerve head could be visualized by Doppler OCT [53]. Retinal blood flow could be quantitatively evaluated, so the retinal diseases could be more effectively diagnosed [54]. Furthermore, the evaluation of retinal blood vascular system can help in sepsis, acute coronary syndrome, and systemic inflammatory bowel disease [55].

Other applications: Except for retinal blood vessel, Doppler OCT has also been applied in other clinical situations such as otology and osteonecrosis diseases. In otology, Doppler OCT can assist the measurement of the motion information of tympanic membrane non-invasively [56]. The oscillation amplitudes and the thickness of tympanic membrane can be effectively evaluated to help the diagnosis of otologic diseases. In osteonecrosis, the blood flow can be assessed by Doppler OCT to enable the differentiation between normal bone and necrosis areas. Blood flow rates can be quantitatively evaluated for the diagnosis of osteonecrosis [57].

1.7 Outline of Dissertation

In Chapter 1, the fundamentals of OCT technique are introduced. OCT working principles, different types of OCT, and their different clinical applications are demonstrated. Endoscopic OCT is demonstrated, along with its diagnostic capabilities in various clinical situations. Moreover, the developments of PS-OCT and Doppler OCT and their applications have been introduced. This chapter provides an overview of OCT and its significance in the medical imaging area.

In Chapter 2, the use of a forward-view endoscopic system for percutaneous nephrostomy guidance is introduced. OCT imaging was conducted on ten pig kidney samples.

In addition, we applied deep learning methods in automatic recognition of different kidney tissues. The results validated the feasibility of using endoscopic OCT in percutaneous nephrostomy needle guidance.

In Chapter 3, the research work of using endoscopic OCT for Veress needle guidance is introduced. Veress needle is used in pneumoperitoneum establishment, and its placement is of great importance. Based on the experimental results, our endoscopic OCT can successfully distinguish different tissues during Veress needle insertion.

In Chapter 4, a laparoscopic OCT is introduced for bladder tumor recognition. Experiments were conducted on mice bladder tumor. Our laparoscopic OCT could identify the boundary of the tumor on bladder, demonstrating the feasibility bladder tumor diagnosis

In Chapter 5, our endoscopic OCT was proved to be applicable for renal carcinoma biopsy guidance. Human carcinoma samples were used in the experiment, and our system could clearly identify the cancerous tissue from normal kidney tissues. Moreover, benign tumor could be distinguished from malignant tumor on kidney.

In Chapter 6, we introduced the use of our endoscopic OCT in epidural anesthesia needle guidance. Experiments on pig backbone samples were conducted. By using a binary CNN recognition model, tissue recognition through the epidural needle insertion sequence could be achieved.

In Chapter 7, we further developed an endoscopic PS-OCT to improve the epidural anesthesia guidance. Based on the biological features provided by PS-OCT signals, the tissue recognition efficiency could be significantly improved.

In Chapter 8, a conclusion of my previous work on endoscopic OCT is introduced. Moreover, future research plans for improving the system and its applications are proposed.

Reference:

1. Huang, D., et al., *Optical coherence tomography.* Science, 1991. **254**(5035): p. 1178-81.
2. Hee, M.R., et al., *Optical coherence tomography of the human retina.* Arch Ophthalmol, 1995. **113**(3): p. 325-32.
3. Puliafito, C.A., et al., *Imaging of macular diseases with optical coherence tomography.* Ophthalmology, 1995. **102**(2): p. 217-29.
4. Kamensky, V., et al., *In-situ monitoring of the mid-IR laser ablation of cataract-suffered human lens by optical coherent tomography*. BiOS Europe '96. Vol. 2930. 1996: SPIE.
5. Mistlberger, A., et al., *Heidelberg retina tomography and optical coherence tomography in normal, ocular-hypertensive, and glaucomatous eyes.* Ophthalmology, 1999. **106**(10): p. 2027-32.
6. Krivoy, D., et al., *Imaging congenital optic disc pits and associated maculopathy using optical coherence tomography.* Arch Ophthalmol, 1996. **114**(2): p. 165-70.
7. Brezinski, M.E., et al., *Optical coherence tomography for optical biopsy. Properties and demonstration of vascular pathology.* Circulation, 1996. **93**(6): p. 1206-13.
8. Brezinski, M.E., et al., *Optical biopsy with optical coherence tomography: feasibility for surgical diagnostics.* J Surg Res, 1997. **71**(1): p. 32-40.
9. Lichtenegger, A., et al., *Ex-vivo Alzheimer's disease brain tissue investigation: a multiscale approach using 1060-nm swept source optical coherence tomography for a direct correlation to histology.* Neurophotonics, 2020. **7**(3): p. 035004.
10. Alex, A., et al., *Multispectral in vivo three-dimensional optical coherence tomography of human skin.* J Biomed Opt, 2010. **15**(2): p. 026025.
11. Kim, S., et al., *Multiple-Wavelength Optical Coherence Tomography Assessments for Enhanced Ex Vivo Intra-Cochlear Microstructural Visualization.* Electronics, 2018. **7**(8).
12. Yamanaka, M., N. Hayakawa, and N. Nishizawa, *High-spatial-resolution deep tissue imaging with spectral-domain optical coherence microscopy in the 1700-nm spectral band.* J Biomed Opt, 2019. **24**(7): p. 1-4.
13. Yamanaka, M., N. Hayakawa, and N. Nishizawa, *Signal-to-background ratio and lateral resolution in deep tissue imaging by optical coherence microscopy in the 1700 nm spectral band.* Sci Rep, 2019. **9**(1): p. 16041.
14. Li, Y., et al., *Intravascular Optical Coherence Tomography for Characterization of Atherosclerosis with a 1.7 Micron Swept-Source Laser.* Sci Rep, 2017. **7**(1): p. 14525.
15. Andrews, P.M., et al., *High-resolution optical coherence tomography imaging of the living kidney.* Lab Invest, 2008. **88**(4): p. 441-9.
16. Andrews, P.M., et al., *Optical coherence tomography of the living human kidney.* Journal of Innovative Optical Health Sciences, 2014. **7**(2).
17. Lee, H.C., et al., *Integrated optical coherence tomography and optical coherence microscopy imaging of ex vivo human renal tissues.* J Urol, 2012. **187**(2): p. 691-9.

18. Galougahi, K.K., et al., *Optical coherence tomography-guided percutaneous coronary intervention in pre-terminal chronic kidney disease with no radio-contrast administration.* European Heart Journal, 2016. **37**(13): p. 1059-1059.
19. Ale-Chilet, A., et al., *Optical Coherence Tomography Angiography in Type 1 Diabetes Mellitus-Report 2: Diabetic Kidney Disease.* J Clin Med, 2021. **11**(1).
20. Mukherjee, P., et al., *Renal tubular function and morphology revealed in kidney without labeling using three-dimensional dynamic optical coherence tomography.* Scientific Reports, 2023. **13**(1).
21. Gora, M.J., et al., *Endoscopic optical coherence tomography: technologies and clinical applications [Invited].* Biomed Opt Express, 2017. **8**(5): p. 2405-2444.
22. Tearney, G.J., et al., *Scanning single-mode fiber optic catheter-endoscope for optical coherence tomography: erratum.* Opt Lett, 1996. **21**(12): p. 912.
23. Xi, J., et al., *High-resolution OCT balloon imaging catheter with astigmatism correction.* Opt Lett, 2009. **34**(13): p. 1943-5.
24. Fu, H.L., et al., *Flexible miniature compound lens design for high-resolution optical coherence tomography balloon imaging catheter.* J Biomed Opt, 2008. **13**(6): p. 060502.
25. Vakoc, B.J., et al., *Comprehensive esophageal microscopy by using optical frequency-domain imaging (with video).* Gastrointest Endosc, 2007. **65**(6): p. 898-905.
26. Gora, M.J., et al., *Tethered capsule endomicroscopy enables less invasive imaging of gastrointestinal tract microstructure.* Nat Med, 2013. **19**(2): p. 238-40.
27. Health, C. *Facts about Barrett's Esophagus*. 2024; Available from: https://www.covenanthealth.com/thompson-cancer-survival-center/facts-about-barretts-esophagus/#:~:text=Genetic%20variation%20in%20two%20chromosomes,esophageal%20adenocarcinoma%20is%20approximately%2013%25.
28. Bouma, B.E. and G.J. Tearney, *Power-efficient nonreciprocal interferometer and linear-scanning fiber-optic catheter for optical coherence tomography.* Opt Lett, 1999. **24**(8): p. 531-3.
29. Chen, Y., et al., *Ultrahigh resolution optical coherence tomography of Barrett's esophagus: preliminary descriptive clinical study correlating images with histology.* Endoscopy, 2007. **39**(7): p. 599-605.
30. Liang, K., et al., *Tethered capsule en face optical coherence tomography for imaging Barrett's oesophagus in unsedated patients.* BMJ Open Gastroenterol, 2020. **7**(1).
31. Kume, T., et al., *Assessment of coronary arterial thrombus by optical coherence tomography.* American Journal of Cardiology, 2006. **97**(12): p. 1713-1717.
32. Nishimura, S., et al., *Cholesterol crystal as a new feature of coronary vulnerable plaques: An optical coherence tomography study.* Journal of Cardiology, 2017. **69**(1-2): p. 253-259.
33. Yabushita, H., et al., *Characterization of human atherosclerosis by optical coherence tomography.* Circulation, 2002. **106**(13): p. 1640-5.

34. Kim, J., et al., *Flexible endoscopic micro-optical coherence tomography for three-dimensional imaging of the arterial microstructure.* Sci Rep, 2020. **10**(1): p. 9248.
35. Li, J., et al., *Ultrathin monolithic 3D printed optical coherence tomography endoscopy for preclinical and clinical use.* Light Sci Appl, 2020. **9**: p. 124.
36. Goh, A.C., et al., *Optical Coherence Tomography as an Adjunct to White Light Cystoscopy for Intravesical Real-Time Imaging and Staging of Bladder Cancer Urology (vol 72, pg 133, 2008).* Urology, 2009. **73**(1): p. 217-217.
37. González-Cerdas, G., et al., *Preclinical-grade microendoscope for optical coherence tomography and angiography inside the bladder.* Journal of Optical Microsystems, 2023. **3**(1).
38. Wang, H., et al., *Three-dimensional imaging of ureter with endoscopic optical coherence tomography.* Urology, 2011. **77**(5): p. 1254-8.
39. Aron, M., et al., *Second prize: preliminary experience with the Niris optical coherence tomography system during laparoscopic and robotic prostatectomy.* J Endourol, 2007. **21**(8): p. 814-8.
40. Bus, M.T., et al., *Volumetric in vivo visualization of upper urinary tract tumors using optical coherence tomography: a pilot study.* J Urol, 2013. **190**(6): p. 2236-42.
41. de Boer, J.F., et al., *Two-dimensional birefringence imaging in biological tissue by polarization-sensitive optical coherence tomography.* Opt Lett, 1997. **22**(12): p. 934-6.
42. Menzel, M., et al., *Diattenuation of brain tissue and its impact on 3D polarized light imaging.* Biomed Opt Express, 2017. **8**(7): p. 3163-3197.
43. Menzel, M., et al., *Diattenuation Imaging reveals different brain tissue properties.* Sci Rep, 2019. **9**(1): p. 1939.
44. de Boer, J.F. and T.E. Milner, *Review of polarization sensitive optical coherence tomography and Stokes vector determination.* J Biomed Opt, 2002. **7**(3): p. 359-71.
45. Gil, J.J. and E. Bernabeu, *Obtainment of the polarizing and retardation parameters of a non-depolarizing optical system from the polar decomposition of its Mueller matrix.* Optik (Stuttgart), 1986. **76**(2): p. 67-71.
46. Apostol, T.M., *Calculus*. 2d ed. A Blaisdell book in pure and applied mathematics. 1967, Waltham, Mass.,: Blaisdell Pub. Co.
47. Jiao, S., et al., *Optical-fiber-based Mueller optical coherence tomography.* Opt Lett, 2003. **28**(14): p. 1206-8.
48. Evans, D.H., *On the measurement of the mean velocity of blood flow over the cardiac cycle using Doppler ultrasound.* Ultrasound Med Biol, 1985. **11**(5): p. 735-41.
49. Sigel, B., *A brief history of Doppler ultrasound in the diagnosis of peripheral vascular disease.* Ultrasound Med Biol, 1998. **24**(2): p. 169-76.
50. Neilson, J.P. and Z. Alfirevic, *Doppler ultrasound for fetal assessment in high risk pregnancies.* Cochrane Database Syst Rev, 2000(2): p. CD000073.

51. Chen, Z., et al., *Optical Doppler tomographic imaging of fluid flow velocity in highly scattering media*. Opt Lett, 1997. **22**(1): p. 64-6.
52. Leitgeb, R.A., et al., *Doppler Optical Coherence Tomography.* Progress in Retinal and Eye Research, 2014. **41**: p. 26-43.
53. Yazdanfar, S., A.M. Rollins, and J.A. Izatt, *In vivo imaging of human retinal flow dynamics by color Doppler optical coherence tomography.* Arch Ophthalmol, 2003. **121**(2): p. 235-9.
54. Tan, O., et al., *Doppler optical coherence tomography of retinal circulation*. J Vis Exp, 2012(67): p. e3524.
55. Courtie, E., et al., *Retinal blood flow in critical illness and systemic disease: a review.* Ann Intensive Care, 2020. **10**(1): p. 152.
56. Seong, D., et al., *Doppler Optical Coherence Tomography for Otology Applications: From Phantom Simulation to In Vivo Experiment*. Applied Sciences-Basel, 2021. **11**(12).
57. Wang, D., et al., *Application of OCT for osteonecrosis using an endoscopic probe based on an electrothermal MEMS scanning mirror.* International Journal of Optomechatronics, 2021. **15**(1): p. 87-96.

CHAPTER 2: DEEP-LEARING AIDED FORWARD OPTICAL COHERENCE TOMOGRAPHY ENDOSCOPE FOR PERCUTANEOUS NEPHROSTOMY GUIDANCE

This chapter is reproduced from "Chen Wang, Paul Calle, Nu Bao Tran Ton, Zuyuan Zhang, Feng Yan, Anthony M. Donaldson, Nathan A. Bradley, Zhongxin Yu, Kar-ming Fung, Chongle Pan, and Qinggong Tang, *Deep-learning-aided forward optical coherence tomography endoscope for percutaneous nephrostomy guidance*, **Biomed. Opt. Express**, 2021."

2.1 Introduction

Percutaneous nephrostomy (PCN) was first described in 1955 as a minimally invasive, x-ray guided procedure in patients with hydronephrosis [1]. PCN needle placement has since become a valuable medical resource for minimally invasive access to the renal collecting system for drainage, urine diversion, the first step for percutaneous nephrolithotomy (PCNL) surgery and other therapeutic intervention, especially when the transurethral access of surgical tools into the urological system is difficult or impossible [1-7]. Despite being a common urological procedure, it remains technically challenging to insert the PCN needle correctly in the right place. During PCN, a needle penetrates the cortex and medulla of the kidney to reach the renal pelvis. Conventional imaging modalities have been used in PCN puncture. Ultrasound technique, as a commonly used medical diagnostic imaging method, has been utilized in PCN surgery for decades [8-11]. Additionally, fluoroscopy and computed tomography (CT) are also employed in PCN guidance and sometimes they are used with ultrasonography simultaneously [10-14]. However, due to the limited spatial resolution, these standard imaging modalities have proven to be inadequate for accurately locating the needle tip position [15, 16]. The failure rate of PCN needle placement is up to 18%, especially in

nondilated systems or for complex stone diseases [17, 18]. Failure of inserting the needle into the targeted location in the kidney through a suitable route might result in severe complications [19-21]. Moreover, fluoroscopy has no soft tissue contrast and, therefore, cannot differentiate critical tissues, such as blood vessels, which are important to avoid during the needle insertion. Rupture of renal blood vessels by needle penetrations can cause bleeding. Temporary bleeding after PCN placement occurs in ~95% of cases [2, 17]. Retroperitoneal hematomas have been found in 13% [17]. When PCNL is followed, hemorrhage requiring transfusion has increased to 12-14% of the patients [22]. Additionally, needle punctures during PCN can lead to infectious complications such as fever or sepsis, thoracic complications like pneumothorax and hydrothorax, and other complications like urine leak or rupture of pelvicalyceal system [23, 24].

Therefore, the selection of position and route of the puncture is important in PCN needle placement. It is recommended to insert the needle into the renal calyx through calyx papilla because fewer blood vessels are distributed on this route, leading to lower possibility of vascular injury [23, 24]. Nevertheless, it is always difficult to precisely identify this preferred inserting route in complicated clinical setting even for experienced urologists [25]. If PCN puncture was executed for multiple times, the likelihood of renal injury will increase and the operational time will lengthen, resulting in higher risks of complications [26].

To better guide PCN needle placement, substantive research work has been done to improve the current guidance practice [9, 17, 27-30]. Ultrasound with technical improvements in many aspects has been utilized. For instance, contrast-enhanced ultrasound has been proved to be a potential modality in the guidance of PCN puncture [30, 31]. Tracked ultrasonography snapshot is also a promising method to improve the needle guidance [32]. In order to resolve

the bleeding during needle puncture, combined B-mode and color Doppler ultrasonography has been applied in PCN surgeries and it provides promising efficiency in decreasing the major hemorrhage incidence [33]. Moreover, developments in other techniques such as cone-beam CT [34], retrograde ureteroscopy [35] and magnetic field-based navigation device [36] have been utilized to improve the guidance of PCN needle access. On the other hand, endoscope is an effective device that can be assembled within PCN needle [37, 38]. It can effectively improve the precision of PCN needle punctures, resulting in lower risks of complications and fewer times of insertions [38]. However, most of the conventional endoscopic techniques involving charge coupled device (CCD) cameras can only provide two-dimensional information and cannot detect subsurface tissue (in other words, visualize the tissue before the needle tip damage it.) [39]. Thus, there is a critical need to develop new guiding techniques which have depth-resolved capability for PCN. Optical coherence tomography (OCT) is a well-established non-invasive biomedical imaging modality which can image subsurface tissue with the penetration depth of several millimeters [40-44]. By obtaining and processing the coherent infrared lights backscattered from reference arm and sample arm, OCT can provide two-dimensional (2D) cross-sectional images with high axial resolution (~10 μm), which is 10–100 times higher than conventional medical imaging modalities (*e.g.,* CT and MRI) [45-47]. Owing to the high speed of laser scanning and data processing, three-dimensional (3D) images of the detected sample formed by numerous cross-sectional images can be obtained in real time [46, 48]. Because of the differences in tissue structures among renal cortex, medulla and calyx, OCT has the potential to distinguish different renal tissue types. Due to the penetration limitation in biological tissues (1-2 mm), studies in kidney using OCT have mainly focused on renal cortex [44, 45, 49-51]. OCT can be

integrated with fiber-optic catheters and endoscopes for internal imaging applications [52-54]. For example, endoscopic OCT imaging has been demonstrated in the human GI tract [55-58] to detect Barrett's esophagus (BE) [59, 60], dysplasia [61] and colon cancer [62-64]. In the previous study, our lab has developed a portable hand-held forward-imaging endoscopic OCT needle device for real-time epidural anesthesia surgery guidance [65]. This endoscopic OCT setup holds the promise in PCN guidance.

Given enormous accumulation of images and inter- and intra-observer variation from subjective interpretation, computer-aided automatic methods have been utilized to accurately and efficiently classify these data [66, 67]. In automated OCT image analysis, convolutional neural networks (CNN) [66, 68, 69] has been demonstrated to be promising in various applications, such as hemorrhage detection of retina versus cerebrum and tumor tissue segmentation [67, 68, 70-73].

Herein we demonstrated a forward OCT endoscope system to image the kidney tissues lying ahead of PCN needle during PCN surgery. Images of renal cortex, medulla and calyx were obtained from ten porcine kidneys using our system. A tissue type classifier was developed using the ResNet34, ResNet50, and MobileNetv2 CNN architectures. Nested cross-validation and testing [74-76] was used for model selection and performance benchmarking to account for the large biological variability among kidney through uncertainty quantification. The predictions from the CNN models were interpreted to identify the important regions in the representative OCT images used by CNN for the classification.

2.2 Materials and Methods

2.2.1 Experimental setup

In our project, the OCT endoscope was built based on the swept-source OCT (SS-OCT) system. Figure 2.1 shows the schematic of our forward endoscopic OCT system. The light source of the SS-OCT system has the center wavelength of 1300 nm and bandwidth of 100 nm [39]. The wavelength-swept frequency (A-scan) rate is 200 kHz with ~25 mW output power. As demonstrated in Figure 2.1, laser provided by the light source was first split by a fiber coupler (FC) into two paths with 3% and 97% of the whole laser power, respectively. The 3% of the laser power was delivered into the Mach-Zehnder interferometer (MZI), which generated a frequency-clock signal for triggering the OCT sampling procedure and provided it to data acquisition (DAQ) board. The remaining 97% laser transmitted through a circulator in which the light runs only in one direction. Therefore, the light entering port 1 only emitted out from port 2, and then it was evenly split to the reference arm and sample arm of a fiber-based Michelson interferometer. Backscattered light from both arms formed interference fringes at the FC and transmitted to the balanced detector (BD). The interference fringes from different depths received by the BD were encoded with different frequencies. The output signal from BD was further transmitted to the DAQ board and computer for processing. Cross-sectional information can be obtained through Fourier transform of interference fringes [65].

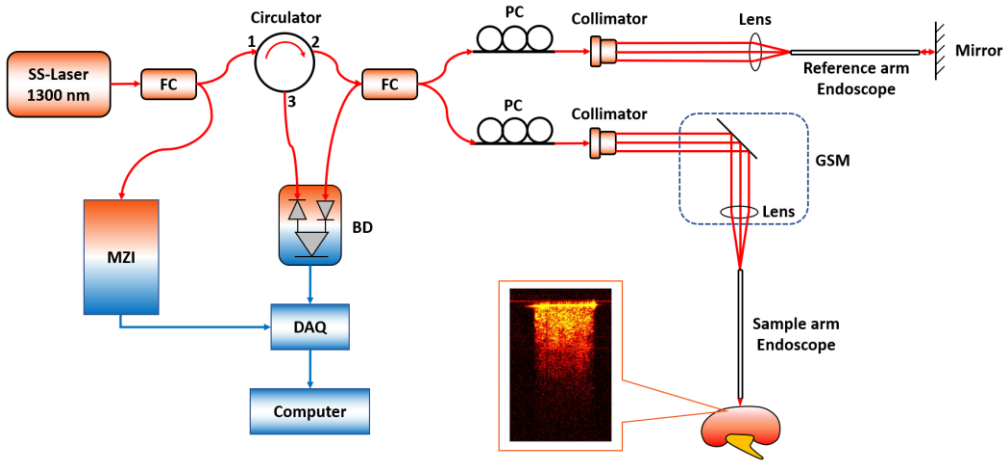


Figure 2.1. Schematic of endoscopic OCT system. MZI: Mach-Zehnder interferometer; FC: fiber coupler; PC: polarization controller; GSM: galvonometer scanning mirror; BD: balanced detector; DAQ: data acquisition board.

A gradient-index (GRIN) rod lens with a diameter of 1.3 mm was stabilized in front of the galvanometer scanning mirror (GSM). The endoscope we used has the diameter of 1.3 mm, the length of 138.0 mm and the view angle of 11.0°. The total outer diameter (O.D.) including the GRIN rod lens and the protective steel tubing is around 1.65 mm. The proximal GRIN lens entrance of the endoscope was placed close to the focal plane of the objective lens. The GRIN lens can preserve the spatial relationship between the entrance and the output (distal end) and further to the sample. Therefore, one or two directional scanning can be readily performed on the proximal GRIN lenses surface to create 2D or 3D images. In addition, the same GRIN rod lens was put in the light path of reference arm for the purpose of compensating light dispersion and expanding the length of reference arm. Polarization controllers (PCs) were put in both arms to decrease the background noise. The system had the axial resolution of ~11 μm and lateral resolution of ~20 μm in tissue. The lateral imaging

field-of-view (FOV) was around 1.25 mm. The sensitivity of system was optimized to 92 dB, calculated using a silver mirror with a calibrated attenuator.

2.2.2 Data acquisition

Ten fresh porcine kidneys were obtained from a local slaughterhouse. The cortex, medulla and calyx of the porcine kidneys were exposed and imaged in the experiment. Renal tissue types can be identified from the anatomic appearance. The OCT endoscope was placed against different renal tissues for image acquisition. To mimic the clinical situation, we applied some force during imaging the *ex-vivo* kidney tissues to generate tissue compression. The 3D images of 320×320×480 pixels on X, Y and Z axes (Z presents the depth direction) were obtained with the pixel size of 6.25 μm on all three axes. Therefore, the size of the original 3D images is 2.00 mm×2.00 mm×3.00 mm. For every kidney sample, we obtained at least 30 original 3D OCT images for each tissue type and each 3D tissue scanning took no more than 2 seconds. Afterwards, the original 3D images were separated to 2D cross-sectional images as shown in Figure 2.2. Since the GRIN lens is cylindrical, the 3D OCT images obtained were also in the cylindrical shape. Therefore, not all the 2D cross-sectional images contained the same structural signal of the kidney. Only the 2D images with sufficient tissue structural information (cross-sectional images that close to the center of the 3D cylindrical structures) were subsequently selected and utilized for the image preprocessing. At the end of imaging, tissues of cortex, medulla and calyx of the porcine kidneys were excised and processed for histology to compare with corresponding OCT results. The tissues were fixed with 10% formalin, embedded in paraffin, sectioned (4 μm thick) and stained with hematoxylin and eosin (H & E) for histological analysis. Images were taken by Keyence Microscope BZ-X800. Sectioning and H & E staining was carried out by the Tissue Pathology

Shared Resource, Stephenson Cancer Center (SCC), University of Oklahoma Health Sciences Center. The Hematoxylin (cat#3801571) and Eosin (cat# 3801616) were purchased from Leica biosystems and the staining was performed utilizing Leica ST5020 Automated Multistainer following the HE staining protocol at the SCC Tissue Pathology core.

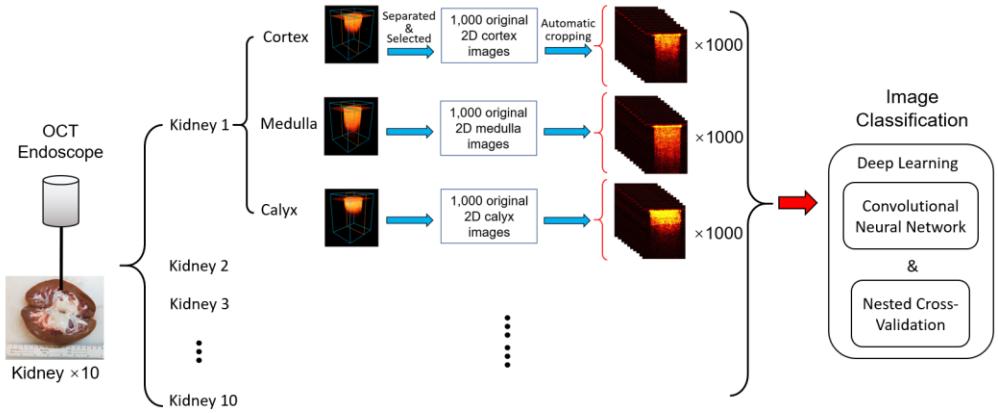


Figure 2.2. Illustration of data acquisition and processing steps.

Although the three tissue types showed different imaging features for visual recognition, it will take time and expertise for doctors to differentiate them during surgeries. In order to improve the efficiency, we developed deep learning methods for automatic tissue classification based on the imaging data. Figure 2.2 shows the overall process of the data acquisition and processing. In total, ten porcine kidneys were imaged in this study. For each kidney, 1,000 2D cross-sectional images were obtained for cortex, medulla, and calyx, respectively. For the purpose of convenient analysis and increasing the speed of deep-learning process of the OCT images, a custom MATLAB algorithm was designed to recognize the surface of the kidney tissue on the 2D cross-sectional images. It automatically cropped the images from the size of 320×480 to 235×301. Therefore, all the 2D cross-sectional images have the same dimensions and cover the same FOV before deep-learning processing.

2.2.3 Training of the convolutional neural networks

CNN was used to classify the images of renal cortex, medulla, and calyx. Three CNN architectures were tested, including Residual Network 34 (ResNet34) [77], Residual Network 50 (ResNet50) [77] and MobileNetv2 [78], using Tensorflow 2.3 [79] in open-ce version 0.1.

Pre-trained ResNet50 and MobileNetv2 models on the ImageNet dataset [80] were imported from the Keras library [81]. The output layer of the models was changed to one containing 3 softmax output neurons for cortex, medulla, and calyx. The input images were preprocessed by resizing to the 224x224 resolution, replicating the input channel to 3 channels, and scaling the pixel intensities to [-1, 1]. Model fine-tuning was conducted in two stages as described in [82]. First, the output layer was trained with all the other layers frozen. The optimizer, stochastic gradient descendent (SGD), was used with a learning rate of 0.2, a momentum of 0.3, and a decay of 0.01. Then, the entire model was unfrozen and trained. The SGD with Nesterov momentum optimizer was used with a learning rate of 0.01, a momentum of 0.9, and a decay of 0.001. Early stopping with a patience of 10 and a maximum number of epochs 50 was used for the Pre-trained ResNet50. Early stopping with a patience of 20 and a maximum number of epochs 100 was used for MobileNetv2.

The ResNet34 and ResNet50 architectures were also trained using randomly initialized weights. ResNet34 [77] was obtained from [82]. The mean pixel in the training dataset was used to center the training, validation, and test datasets. The input layer was modified to accept only one input channel in the OCT images and the output layer was changed for the classification of the three tissue types. For ResNet50, the optimizer SGD with Nesterov momentum with learning rate 0.01, momentum 0.9 and decay 0.01 was used. ResNet50 was trained with a maximum of 50 epochs, early stopping with a patience of 10,

and a batch size of 32. For ResNet34, the Adam optimizer was used with learning rate 0.001, beta1 0.9, beta2 0.9999 and epsilon 1E-7. ResNet34 was trained with a maximum of 200 epochs, early stopping with a patience of 10, and a batch size of 512.

2.2.4 Nested cross-validation and testing

A nested cross-validation and testing procedure [74, 76, 83] was used to estimate the validation performance and the test performance of the models across the 10 kidneys with uncertainty quantification. The pseudo-code of the nested cross-validation and testing is shown below.

```
# 10-fold cross-testing loop
for kidney i in the 10 kidneys do
    Hold out kidney i in the test set
    # model optimization loop
    for each model configuration do
        # 9-fold cross-validation loop
    for kidney j in the remaining 9 kidneys do
        Use kidney j as the validation set
        Train a model using the remaining 8 kidneys as the training set
        Benchmark the validation performance using kidney j
        end for
        Estimate the mean validation accuracy and its std. error
    end for
    Select the best model configuration based on the validation performance
    Train a model with the selected configuration using the 9 kidneys
    Benchmark the test performance of this model using kidney i
end for
Summarize the test performance of this procedure
```

In the 10-fold cross-testing, one kidney was selected in turn as the test set. In the 9-fold cross-validation, the remaining nine kidneys were partitioned 8:1 between the training set and the validation set. Each kidney contained a total of 3000 images, including 1000 images for each tissue type. The validation performance of a model was tracked based on its

classification accuracy on the validation kidney. The classification accuracy is the percentage of correctly labeled images out of all 3000 images of a kidney.

The 9-fold cross-validation loop was used to compare the performance of ResNet34, ResNet50, and MobileNetv2, and optimize the key hyperparameters of these models, such as pre-trained versus randomly initialized weights, learning rates, and number of epochs. The model configuration with the highest average validation accuracy was selected for the cross-testing loop. The cross-testing loop enabled iterative benchmarking of the selected model across all 10 kidneys, giving a better estimation of generalization error with uncertainty quantification.

Gradient-weighted Class Activation Mapping (Grad-CAM) [84] was used to explain the predictions of a selected CNN model by highlighting the important regions in the image for the prediction outcome. The interpretation implementation on ResNet50 was based on [85].

All the model development was performed on the Schooner supercomputer at the University of Oklahoma and the Summit supercomputer at Oak Ridge National Laboratory. The computation on Schooner used five computational nodes, each of which had 40 CPU cores (Intel Xeon Cascade Lake) and 200 GB of RAM. The computation on Summited used up to 15 nodes, each of which had 2 IBM POWER9 processors and 6 NVIDIA Volta Graphic cards. The source code of the model training is available at https://github.com/thepanlab/FOCT_kidney.

2.3 Results

2.3.1 Forward OCT imaging of different renal tissues

The imaging setup our OCT endoscope was demonstrated in Figure 2.3(A). An adapter was used to stabilize the endoscope in front of the OCT scan lens kit. From the kidney sample shown in Figure 2.3(A), we can visually distinguish different tissue types. Renal cortex was the brown tissue on the edge of the whole kidney; Medulla can be recognized by its red renal pyramid structures distributed on the inner side of the cortex; Calyx was featured by its obvious white structure in the central kidney. Three tissue types were imaged respectively following the procedure described in Section 2.2.

Figure 2.3(B-D) show representative 3D OCT images, 2D cross-sectional images and the histology results of three renal tissues. They were featured with different imaging depth and brightness. The renal calyx had the shallowest imaging depth, but the tissue close to the surface showed the highest brightness and density. Cortex and medulla both presented relatively homogeneous tissue structures in OCT images, and the imaging depth of medulla was larger than cortex. Furthermore, compared to cortex and medulla, calyx was featured with horizonal stripes and layered structure. The transitional epithelium and fibrous tissue in the calyx may explain the strip-like structures and significantly higher brightness in comparison to the other two renal tissues. This is the significant part for PCN insertion since the goal of PCN is to reach the calyx precisely. These imaging results demonstrated the feasibility of distinguishing renal cortex, medulla, and calyx with the endoscopic OCT system.

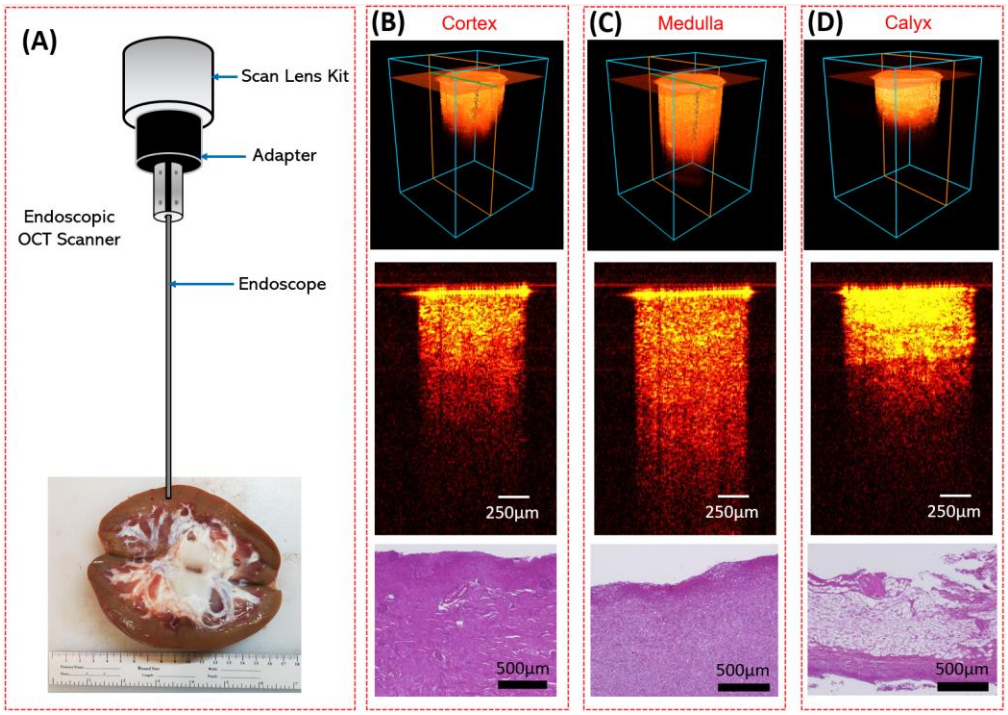


Figure 2.3. Sample kidney used for endoscopic OCT imaging; (B-D) Representative 3D OCT images, cross-sectional OCT images and the histology results of renal cortex (B), medulla (C) and calyx (D).

2.3.2 CNN development and benchmarking results

Table 2.1 shows the average validation accuracies and their standard errors for the pre-trained (PT) or randomly initialized (RI) model architectures after hyperparameter optimization. RI MobileNetv2 frequently failed to learn, so only the PT MobileNetv2 model was used here. The PT ResNet50 models outperformed the RI ResNet50 models in 6 of the 10 testing folds, which indicated only a small boost by the pre-training on ImageNet. For all the 10 testing folds, the validation accuracies of the ResNet50 models were significantly higher than those of the MobileNetv2 and ResNet34 models. Thus, the characteristic patterns of the three kidney tissues may require a deep CNN architecture to recognize.

Table 2.1: The averages and standard errors of the validation accuracies from the nested cross-validation and testing procedure. The architecture with the highest validation performance in each testing fold is indicated in bold.

Testing folds	PT MobileNetv2	RI ResNet34	RI ResNet50	PT ResNet50
K1	84.1%±3.5%	83.8%±3.7%	87.1%±2.9%	**89.2%±2.4%**
K2	82.1%±3.9%	84.6%±2.8%	87.0%±3.0%	**89.8%±2.4%**
K3	83.9%±4.2%	84.5%±3.5%	**88.5%±2.4%**	86.6%±2.4%
K4	78.0%±6.3%	83.5%±2.6%	85.0%±2.4%	**87.8%±2.3%**
K5	83.3%±2.7%	82.8%±4.2%	**87.8%±1.9%**	87.7%±2.3%
K6	86.2%±2.7%	84.5%±3.6%	**88.4%±2.1%**	88.1%±2.4%
K7	86.0%±3.8%	82.8%±3.0%	**88.6%±1.7%**	86.6%±2.4%
K8	76.1%±6.0%	82.1%±3.9%	87.7%±2.6%	**88.0%±2.5%**
K9	77.7%±4.8%	79.4%±4.3%	82.1%±3.6%	**85.1%±2.3%**
K10	81.8%±4.6%	86.4%±2.7%	87.5%±3.2%	**87.7%±2.1%**

Table 2.2 shows the test accuracy of the best-performing model in each of the 10 testing folds. The output layer of the CNN models estimated three softmax scores that summed up to 1.0 for the three tissue types. When the category with the highest softmax score was selected for an image (*i.e.*, a softmax score threshold of 0.333 to make a prediction), the CNN model made a prediction for every image (100% coverage) at a mean test accuracy of 82.6%. This was substantially lower than the mean validation accuracy of 87.3%, which suggested the overfitting to the validation set by the hyperparameter tuning and early stopping. The classification accuracy can be increased at the expense of lower coverage by increasing softmax score threshold, which allowed the CNN model to make only confident classifications. When the softmax score threshold was raised to 0.5, 89.9% of the images on average were classified to a tissue type and the mean classification accuracy increased to 85.6%±3.0%. For the uncovered images, the doctors can make a prediction with the help of other imaging modality and their clinical experience.

There was substantial variability in the test accuracy among different kidneys. While three kidneys had test accuracies higher than 92% (softmax score threshold of 0.333), the kidney in the sixth fold had the lowest test accuracy of 67.7%. Therefore, the current challenge in the image classification mainly comes from the anatomic differences among the samples. For instance, Figure 2.4(A) and (B) shows the receiver operating characteristic (ROC) curves of the prediction results from kidney No. 5 and No. 10. It is clear that the prediction of kidney 5 is much more accurate than that of kidney 10. Our nested cross-validation and testing procedure was designed to simulate the real clinical setting in which the CNN models trained on one set of kidneys need to perform well on a new kidney unseen by the CNN models until the operation. When a CNN model was trained on a subset of images from all kidneys and validated on a separate subset of images from all kidneys in cross-validation as opposed to partitioning by kidneys, it achieved accuracies over 99%. This suggested that the challenge of image classification mainly stemmed from the biological differences between different kidneys. The generalization of the CNN models across kidneys can be improved by expand our dataset with kidneys at different ages or physical conditions to represent different structural and morphological features.

Table 2.2. Results from the 10-fold cross-testing. Average and standard error of accuracy are shown for two thresholds.

Testing folds	Threshold = 0.333		Threshold = 0.5	
	Accuracy	Coverage	Accuracy	Coverage
K1	75.4%	100%	79.0%	89.3%
K2	87.9%	100%	90.0%	93.3%
K3	71.4%	100%	76.8%	79.3%
K4	92.1%	100%	95.9%	88.4%
K5	93.2%	100%	95.6%	93.7%
K6	67.7%	100%	68.9%	94.2%
K7	83.5%	100%	86.4%	91.8%
K8	88.2%	100%	92.3%	85.0%
K9	92.2%	100%	94.8%	94.2%
K10	74.1%	100%	76.6%	90.2%
Average	82.6%±3.0%	100%	85.6%±3.0%	89.9%±1.5%

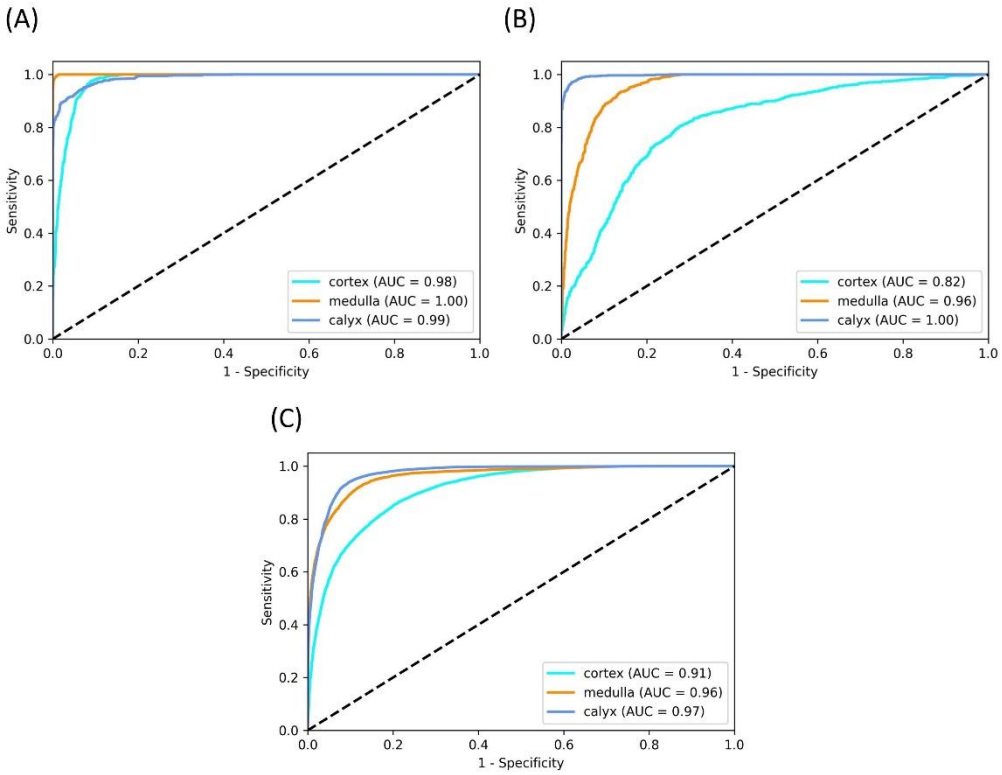


Figure 2.4. ROC multi-class testing curves of (A) kidney 5, (B) kidney 10 and (C) the average of all kidneys.

Figure 2.4(C) shows the average ROC curves for the three tissue types. The Area Under the ROC Curve (AUC) were 0.91 for cortex, 0.96 for medulla, and 0.97 for calyx. Table 2.3 shows the average confusion matrix for the 10 kidneys in the 10-fold cross-testing

with a score threshold of 0.333 and the average recall and precision for each type of tissue. Cortex was most challenging tissue type to be classified correctly and often mixed up with medulla. From the original images we found that the penetration depths of medulla were much larger than cortex in seven of the ten imaged kidneys. While in other three samples, these differences were insignificant. This may explain the challenging classification between cortex and medulla.

Table 2.3. Confusion matrix for classification of the three tissue types.

		Predicted			Recall
		Cortex	Medulla	Calyx	
True	Cortex	677 ± 109	204 ± 93	118 ± 93	**68% ± 11%**
	Medulla	75 ± 33	908 ± 41	17 ± 15	**91% ± 4%**
	Calyx	95 ± 30	14 ± 9	892 ± 32	**89% ± 3%**
	Precision	**79% ± 4%**	**85% ± 6%**	**91% ± 5%**	

To better understand how a CNN model classified different renal types, the class activation maps were generated to visualize heatmaps of class activation over input images. Figure 2.5 shows the class activation heatmaps for 3 representative images of each tissue type from the RI ResNet50 model and the PT ResNet50 model. The models and the representative images were selected from the fifth testing fold. The RI ResNet50 model performed the classification by paying more intention to the lower part of the images of cortex, to both the lower part and near the upper part of the medulla images, and to the area of the calyx images with high intensity near the needle tip. The PT ResNet50 model focused on both the upper part and lower part of the cortex images, on the middle part and/or lower part of the medulla images, and on the region close to the needle tip of the calyx images. Compared to the RI ResNet50 model, the PT ResNet50 model shifted its attention closer to the signal regions. The

class classification heatmaps provided an intuitive explanation of the classification basis for the two CNN models.

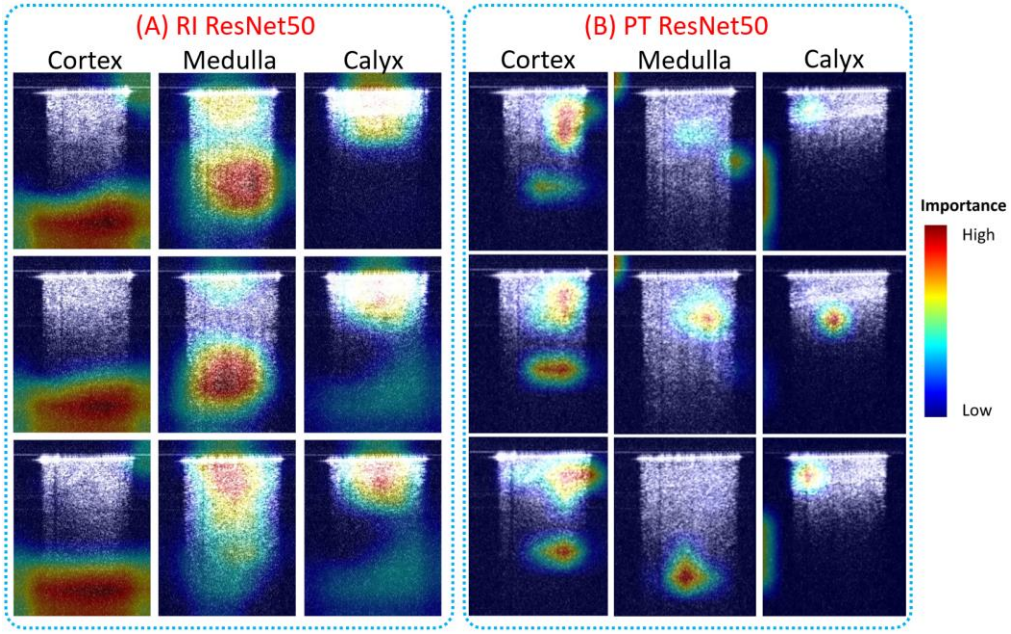


Figure 2.5. Class activation heatmaps of (A) RI ResNet50, and (B) PT ResNet50 for three representative images in each tissue type. The class activation heatmaps used jet colormap and were superimposed on the input images.

2.4 Discussion

We investigated the feasibility of OCT endoscopic system for PCN surgery guidance. Three porcine kidney tissue types: cortex, medulla and calyx were imaged. These three kidney tissues show different structural features which can be further used for tissue type recognition. To increase the image recognition efficiency and reduce the learning burden of the clinical doctors, we developed and evaluated CNN methods for image classification and recognition. ResNet50 had the best performance compared to ResNet34 and PT MobileNetv2 and achieved an average classification accuracy of 82.6%±3.0%.

In the current study, the porcine kidneys samples were obtained from a local slaughterhouse without controlling the sample preservation and time after death. Biological changes may have occurred in the *ex-vivo* kidneys, including collapse of some structures of nephrons such as the renal tubules. This may make the tissue recognition more difficult, especially the classification between cortex and medulla. Characteristic renal structures in the cortex can be clearly imaged by OCT in both well-preserved *ex-vivo* human kidneys and living kidneys as previously reported [44, 50, 86] and verified in an ongoing study in our lab using well-preserved human kidneys. Additionally, nephron structures distributed in renal cortex and medulla are different [87]. These additional features in renal cortex and medulla will improve the recognition of these two tissue types and increase the classification accuracy of our future CNN models when imaging *in-vivo* samples or well-preserved *ex-vivo* samples. The current study established the feasibility of automatic tissue recognition using CNN and provided information for the model selection and hyper-parameter optimization in future CNN model development using *in-vivo* pig kidneys and well-preserved *ex-vivo* human kidneys.

For translating the proposed OCT probe into clinics, we will assemble the endoscope with appropriate diameter and length into the clinical used PCN needle. In current PCN puncture, a trocar needle is inserted into the kidney. Since the trocar has a hollow structure, we can fix the endoscope within the trocar needle. Then our OCT endoscope can be inserted into the kidney together with the trocar needle. After the trocar needle tip arrives at the destination (such as the kidney pelvis), we will withdraw the OCT endoscope from the trocar needle and other surgical processes can be continued. During the whole puncture, no extra invasiveness will be caused. Since the needle will keep moving during the puncture, there will

be a tight contact between the needle tip and the tissue. Therefore, the blood (if any) will not accumulate in front of the needle tip. From our previous experience in the *in-vivo* pig experiment guiding the epidural anesthesia using our OCT endoscope, the presence of blood is not a big issue [65]. The diameter of the GRIN rod lens we used in the study is 1.3 mm. In the future study, we will further improve the current setup with smaller GRIN rod lens that can be fit inside the 18-gauge PCN needle which is clinically used in PCN puncture [88]. Furthermore, we will miniaturize the GSM device based on microelectromechanical system (MEMS) technology, which will enable ease of operation and is important for translating the OCT endoscope to clinical applications. The current employed OCT system has a scanning speed up to 200 kHz, the 2D tissue images in front of the PCN needle can be provided to surgeons in real time. Using ultra-high speed of laser scanning and data processing system, 3D images of the detected sample can be obtained in real time [46, 48]. In the next step, we will acquire 3D images that may further improve our classification accuracy, because of the added information content in 3D images. For example, Kruthika *et al* detected Alzheimer's disease from MRI images and showed improved performance of 3D Capsule Network (CapsNet) and 3D CNN over previous 2D approaches [89].

The CNN model training in this study used significant computational power. Each fold of the cross-validation took ~25 minutes for RI ResNet50 and ~45 minutes for PT ResNet50 using one NVIDIA Volta GPU. The 90 folds of the nested cross-validation for each model configuration were performed in parallel across multiple compute nodes on the Summit supercomputer. The inference was computationally efficient. It took ~50 seconds using one NVIDIA Volta GPU to perform inference on 1000 images (i.e., ~0.05 seconds per image on average), including model loading, image preprocessing, and the ResNet50 classification. The

ResNet50 classification used ~16 seconds for 1000 images or ~0.016 seconds per image. In future, the inference can be further accelerated through algorithm optimization and parallelization to make it more practical for surgical applications.

Reference

1. Goodwin, W.E., W.C. Casey, and W. Woolf, *Percutaneous trocar (needle) nephrostomy in hydronephrosis*. J Am Med Assoc, 1955. **157**(11): p. 891-4.
2. Young, M. and S.W. Leslie, *Percutaneous Nephrostomy*, in *StatPearls*. 2019: Treasure Island (FL).
3. Bigongiari, L.R., et al., *Percutaneous nephrostomy. A non-operative temporary or long-term urinary diversion procedure.* J Kans Med Soc, 1978. **79**(3): p. 104-6,132.
4. Hellsten, S., et al., *Percutaneous nephrostomy. Aspects on applications and technique.* Eur Urol, 1978. **4**(4): p. 282-7.
5. Perinetti, E., et al., *Percutaneous nephrostomy: indications, complications and clinical usefulness.* J Urol, 1978. **120**(2): p. 156-8.
6. Stables, D.P., N.J. Ginsberg, and M.L. Johnson, *Percutaneous nephrostomy: a series and review of the literature.* AJR Am J Roentgenol, 1978. **130**(1): p. 75-82.
7. Lee, W.J., et al., *Emergency percutaneous nephrostomy: results and complications.* J Vasc Interv Radiol, 1994. **5**(1): p. 135-9.
8. Pedersen, J.F., *Percutaneous nephrostomy guided by ultrasound.* J Urol, 1974. **112**(2): p. 157-9.
9. Efesoy, O., et al., *The results of ultrasound-guided percutaneous nephrostomy tube placement for obstructive uropathy: A single-centre 10-year experience.* Turk J Urol, 2018. **44**(4): p. 329-334.
10. LeMaitre, L., et al., *Percutaneous nephrostomy: placement under laser guidance and real-time CT fluoroscopy.* Eur Radiol, 2000. **10**(6): p. 892-5.
11. Zegel, H.G., et al., *Percutaneous nephrostomy: comparison of sonographic and fluoroscopic guidance.* AJR Am J Roentgenol, 1981. **137**(5): p. 925-7.
12. Miller, N.L., B.R. Matlaga, and J.E. Lingeman, *Techniques for fluoroscopic percutaneous renal access.* J Urol, 2007. **178**(1): p. 15-23.
13. Barbaric, Z.L., et al., *Percutaneous nephrostomy: placement under CT and fluoroscopy guidance.* AJR Am J Roentgenol, 1997. **169**(1): p. 151-5.
14. Hacker, A., et al., *A biological model to teach percutaneous nephrolithotomy technique with ultrasound- and fluoroscopy-guided access.* J Endourol, 2007. **21**(5): p. 545-50.
15. de Sousa Morais, N., et al., *Percutaneous nephrostomy vs ureteral stent for hydronephrosis secondary to ureteric calculi: impact on spontaneous stone passage and health-related quality of life-a prospective study.* Urolithiasis, 2019. **47**(6): p. 567-573.
16. Gupta, S., M. Gulati, and S. Suri, *Ultrasound-guided percutaneous nephrostomy in non-dilated pelvicaliceal system.* J Clin Ultrasound, 1998. **26**(3): p. 177-9.

17. Ramchandani, P., et al., *Quality improvement guidelines for percutaneous nephrostomy.* J Vasc Interv Radiol, 2003. **14**(9 Pt 2): p. S277-81.
18. Montvilas, P., J. Solvig, and T.E.B. Johansen, *Single-centre review of radiologically guided percutaneous nephrostomy using "mixed" technique: Success and complication rates.* European Journal of Radiology, 2011. **80**(2): p. 553-558.
19. Ozbek, O., et al., *Rapid percutaneous nephrostomy catheter placement in neonates with the trocar technique.* Diagn Interv Imaging, 2017. **98**(4): p. 315-319.
20. Vignali, C., et al., *Vascular injuries after percutaneous renal procedures: treatment by transcatheter embolization.* Eur Radiol, 2004. **14**(4): p. 723-9.
21. Dagli, M. and P. Ramchandani, *Percutaneous nephrostomy: technical aspects and indications.* Semin Intervent Radiol, 2011. **28**(4): p. 424-37.
22. Lee, W.J., et al., *Complications of percutaneous nephrolithotomy.* American Journal of Roentgenology, 1987. **148**(1): p. 177-180.
23. Hausegger, K.A. and H.R. Portugaller, *Percutaneous nephrostomy and antegrade ureteral stenting: technique-indications-complications.* Eur Radiol, 2006. **16**(9): p. 2016-30.
24. Kyriazis, I., et al., *Complications in percutaneous nephrolithotomy.* World J Urol, 2015. **33**(8): p. 1069-77.
25. Radecka, E. and A. Magnusson, *Complications associated with percutaneous nephrostomies. A retrospective study.* Acta Radiol, 2004. **45**(2): p. 184-8.
26. Egilmez, H., et al., *The place of computed tomography as a guidance modality in percutaneous nephrostomy: analysis of a 10-year single-center experience.* Acta Radiol, 2007. **48**(7): p. 806-13.
27. Sonawane, P., et al., *A modified Malecot catheter design to prevent complications during difficult percutaneous nephrostomy.* Arab Journal of Urology, 2019. **17**(4): p. 330-334.
28. Alken, P., et al., *Percutaneous Stone Manipulation.* Journal of Urology, 2017. **197**(2): p. S154-S157.
29. Liu, Q., et al., *Fluoroscopy versus ultrasound for image guidance during percutaneous nephrolithotomy: a systematic review and meta-analysis.* Urolithiasis, 2017. **45**(5): p. 481-487.
30. Liu, B.X., et al., *Contrast-enhanced US-assisted Percutaneous Nephrostomy: A Technique to Increase Success Rate for Patients with Nondilated Renal Collecting System.* Radiology, 2017. **285**(1): p. 293-301.
31. Cui, X.W., et al., *Feasibility and Usefulness of Intra-Cavitary Contrast-Enhanced Ultrasound in Percutaneous Nephrostomy.* Ultrasound Med Biol, 2016. **42**(9): p. 2180-8.
32. Ungi, T., et al., *Tracked ultrasonography snapshots enhance needle guidance for percutaneous renal access: a pilot study.* J Endourol, 2014. **28**(9): p. 1040-5.
33. Lu, M.H., et al., *A comparative study of clinical value of single B-mode ultrasound guidance and B-mode combined with color doppler ultrasound guidance in mini-*

invasive percutaneous nephrolithotomy to decrease hemorrhagic complications. Urology, 2010. **76**(4): p. 815-20.
34. Hawkins, C.M., et al., *Use of cone-beam CT and live 3-D needle guidance to facilitate percutaneous nephrostomy and nephrolithotripsy access in children and adolescents.* Pediatric Radiology, 2016. **46**(4): p. 570-574.
35. Uribe, C.A., et al., *A New Technique for Percutaneous Nephrolithotomy Using Retrograde Ureteroscopy and Laser Fiber to Achieve Percutaneous Nephrostomy Access: The Initial Case Report.* J Endourol Case Rep, 2019. **5**(3): p. 131-136.
36. Krombach, G.A., et al., *US-guided nephrostomy with the aid of a magnetic field-based navigation device in the porcine pelvicaliceal system.* Journal of Vascular and Interventional Radiology, 2001. **12**(5): p. 623-628.
37. Miller, R.A. and J.E. Wickham, *Percutaneous nephrolithotomy: advances in equipment and endoscopic techniques.* Urology, 1984. **23**(5 Spec No): p. 2-6.
38. Isac, W., et al., *Endoscopic-guided versus fluoroscopic-guided renal access for percutaneous nephrolithotomy: a comparative analysis.* Urology, 2013. **81**(2): p. 251-6.
39. Tang, Q.G., et al., *Depth-resolved imaging of colon tumor using optical coherence tomography and fluorescence laminar optical tomography.* Biomedical Optics Express, 2016. **7**(12): p. 5218-5232.
40. Huang, D., et al., *Optical Coherence Tomography.* Science, 1991. **254**(5035): p. 1178-1181.
41. Fercher, A.F., et al., *Optical coherence tomography - principles and applications.* Reports on Progress in Physics, 2003. **66**(2): p. 239-303.
42. Andrews, P.M., et al., *High-resolution optical coherence tomography imaging of the living kidney.* Laboratory Investigation, 2008. **88**(4): p. 441-449.
43. Fujimoto, J.G., et al., *Optical Biopsy and Imaging Using Optical Coherence Tomography.* Nature Medicine, 1995. **1**(9): p. 970-972.
44. Andrews, P.M., et al., *Optical coherence tomography of the living human kidney.* Journal of Innovative Optical Health Sciences, 2014. **7**(2).
45. Li, Q., et al., *Automated quantification of microstructural dimensions of the human kidney using optical coherence tomography (OCT).* Optics Express, 2009. **17**(18): p. 16000-16016.
46. Israelsen, N.M., et al., *Real-time high-resolution mid-infrared optical coherence tomography.* Light-Science & Applications, 2019. **8**.
47. Fujimoto, J. and E. Swanson, *The Development, Commercialization, and Impact of Optical Coherence Tomography.* Investigative Ophthalmology & Visual Science, 2016. **57**(9): p. Oct1-Oct13.
48. Wieser, W., et al., *High definition live 3D-OCT in vivo: design and evaluation of a 4D OCT engine with 1 GVoxel/s.* Biomedical optics express, 2014. **5**(9): p. 2963-2977.
49. Wang, H.W. and Y. Chen, *Clinical applications of optical coherence tomography in urology.* Intravital, 2014. **3**(1): p. e28770.

50. Konkel, B., et al., *Fully automated analysis of OCT imaging of human kidneys for prediction of post-transplant function*. Biomedical Optics Express, 2019. **10**(4): p. 1794-1821.
51. Ding, Z., et al., *Multi-modality Optical Imaging of Rat Kidney Dysfunction: In Vivo Response to Various Ischemia Times*, in *Oxygen Transport to Tissue XXXVIII*. 2016, Springer. p. 345-350.
52. Tearney, G.J., et al., *In vivo endoscopic optical biopsy with optical coherence tomography.* Science, 1997. **276**(5321): p. 2037-2039.
53. Xi, J., et al., *Diffractive catheter for ultrahigh-resolution spectral-domain volumetric OCT imaging.* Opt Lett, 2014. **39**(7): p. 2016-9.
54. Sharma, G.K., et al., *Long-Range Optical Coherence Tomography of the Neonatal Upper Airway for Early Diagnosis of Intubation-related Subglottic Injury.* American Journal Of Respiratory And Critical Care Medicine, 2015. **192**(12): p. 1504-1513.
55. Chen, Y., et al., *Ultrahigh resolution optical coherence tomography of Barrett's esophagus: preliminary descriptive clinical study correlating images with histology.* Endoscopy, 2007. **39**(7): p. 599-605.
56. Sergeev, A.M., et al., *In vivo endoscopic OCT imaging of precancer and cancer states of human mucosa*. Optics Express, 1997. **1**(13): p. 432-440.
57. Bouma, B.E., et al., *High-resolution imaging of the human esophagus and stomach in vivo using optical coherence tomography.* Gastrointestinal Endoscopy, 2000. **51**(4): p. 467-474.
58. Gora, M.J., et al., *Tethered capsule endomicroscopy enables less invasive imaging of gastrointestinal tract microstructure.* Nat Med, 2013. **19**(2): p. 238-40.
59. Poneros, J.M., et al., *Diagnosis of specialized intestinal metaplasia by optical coherence tomography.* Gastroenterology, 2001. **120**(1): p. 7-12.
60. Liang, K., et al., *Ultrahigh speed en face OCT capsule for endoscopic imaging*. Biomed Opt Express, 2015. **6**(4): p. 1146-63.
61. Isenberg, G., et al., *Accuracy of endoscopic optical coherence tomography in the detection of dysplasia in Barrett's esophagus: a prospective, double-blinded study.* Gastrointestinal Endoscopy, 2005. **62**(6): p. 825-831.
62. Winkler, A.M., et al., *Quantitative tool for rapid disease mapping using optical coherence tomography images of azoxymethane-treated mouse colon.* Journal of biomedical optics, 2010. **15**(4): p. 041512-041512-10.
63. Jäckle, S., et al., *In vivo endoscopic optical coherence tomography of the human gastrointestinal tract-toward optical biopsy.* Endoscopy, 2000. **32**(10): p. 743-749.
64. Pfau, P.R., et al., *Criteria for the diagnosis of dysplasia by endoscopic optical coherence tomography.* Gastrointestinal Endoscopy, 2003. **58**(2): p. 196-202.
65. Tang, Q.G., et al., *Real-time epidural anesthesia guidance using optical coherence tomography needle probe.* Quantitative Imaging in Medicine and Surgery, 2015. **5**(1): p. 118-124.
66. Wang, C., et al., *Adversarial convolutional network for esophageal tissue segmentation on OCT images.* Biomedical Optics Express, 2020. **11**(6): p. 3095-3110.

67. Schmidt-Erfurth, U., et al., *Artificial intelligence in retina*. Prog Retin Eye Res, 2018. **67**: p. 1-29.
68. Fang, L.Y., et al., *Attention to Lesion: Lesion-Aware Convolutional Neural Network for Retinal Optical Coherence Tomography Image Classification*. Ieee Transactions on Medical Imaging, 2019. **38**(8): p. 1959-1970.
69. Fang, L.Y., et al., *Extinction Profiles Fusion for Hyperspectral Images Classification*. Ieee Transactions on Geoscience and Remote Sensing, 2018. **56**(3): p. 1803-1815.
70. Date, R.C., S.J. Jesudasen, and C.Y. Weng, *Applications of Deep Learning and Artificial Intelligence in Retina*. Int Ophthalmol Clin, 2019. **59**(1): p. 39-57.
71. Dou, Q., et al., *Automatic Detection of Cerebral Microbleeds From MR Images via 3D Convolutional Neural Networks*. Ieee Transactions on Medical Imaging, 2016. **35**(5): p. 1182-1195.
72. Pereira, S., et al., *Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images*. Ieee Transactions on Medical Imaging, 2016. **35**(5): p. 1240-1251.
73. van Grinsven, M.J.J.P., et al., *Fast Convolutional Neural Network Training Using Selective Data Sampling: Application to Hemorrhage Detection in Color Fundus Images*. Ieee Transactions on Medical Imaging, 2016. **35**(5): p. 1273-1284.
74. Raschka, S., *Model evaluation, model selection, and algorithm selection in machine learning*. arXiv preprint arXiv:1811.12808, 2018.
75. Varma, S. and R. Simon, *Bias in error estimation when using cross-validation for model selection*. BMC Bioinformatics, 2006. **7**(1): p. 91.
76. Iizuka, N., et al., *Oligonucleotide microarray for prediction of early intrahepatic recurrence of hepatocellular carcinoma after curative resection*. The Lancet, 2003. **361**(9361): p. 923-929.
77. He, K., et al. *Deep residual learning for image recognition*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.
78. Sandler, M., et al. *Mobilenetv2: Inverted residuals and linear bottlenecks*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018.
79. Abadi, M., et al. *Tensorflow: A system for large-scale machine learning*. in *12th {USENIX} symposium on operating systems design and implementation ({OSDI} 16)*. 2016.
80. Deng, J., et al. *ImageNet: A large-scale hierarchical image database*. in *2009 IEEE Conference on Computer Vision and Pattern Recognition*. 2009.
81. Chollet, F., *Keras*. 2015.
82. Géron, A., *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems*. 2019: O'Reilly Media.
83. Krstajic, D., et al., *Cross-validation pitfalls when selecting and assessing regression and classification models*. Journal of Cheminformatics, 2014. **6**(1): p. 10.

84. Selvaraju, R.R., et al. *Grad-cam: Visual explanations from deep networks via gradient-based localization*. in *Proceedings of the IEEE international conference on computer vision*. 2017.
85. Chollet, F., *Deep learning with Python*. Vol. 361. 2018: Manning New York.
86. Chen, Y., et al., *High-resolution three-dimensional optical coherence tomography imaging of kidney microanatomy ex vivo.* J Biomed Opt, 2007. **12**(3): p. 034008.
87. Lote, C.J., *Principles of renal physiology*. 4th ed. 2000, Dordrecht ; Boston: Kluwer Academic Publishers. x, 203 p.
88. Dyer, R.B., et al., *Percutaneous nephrostomy with extensions of the technique: Step by step.* Radiographics, 2002. **22**(3): p. 503-525.
89. Kruthika, K.R., Rajeswari, and H.D. Maheshappa, *CBIR system using Capsule Networks and 3D CNN for Alzheimer's disease diagnosis.* Informatics in Medicine Unlocked, 2019. **14**: p. 59-68.

CHAPTER 3: COMPUTER-AIDED VERESS NEEDLE GUIDANCE USING ENDOSCOPIC OPTICAL COHERENCE TOMOGRAPHY AND CONVOLUTIONAL NEURAL NETWORKS

This chapter is reproduced from "Chen Wang, Justin C. Reynolds, Paul Calle, Avery D. Ladymon, Feng Yan, Yuyang Yan, Sam Ton, Kar-Ming Fung, Sanjay G. Patel, Zhongxin Yu, Chongle Pan, Qinggong Tang, *Computer-aided Veress needle guidance using endoscopic optical coherence tomography and convolutional neural networks*, **J. Biophotonics**, 2022."

3.1 Introduction

Laparoscopy is a modern and minimally invasive surgical technique used in the diagnosis and therapeutic purposes [1-3]. With the development of video camera and other medical auxiliary instruments, laparoscopy has become a procedure widely used in various surgeries such as cholecystectomy [4], appendectomy [5], herniotomy [6], gastric banding [7] or colon resection [8]. In the first step of the laparoscopic procedure, a trocar/Veress needle is inserted into the patient's abdominal cavity through a small skin incision [9]. The Veress needle penetrates subcutaneous fat and muscle before reaching the abdominal cavity [10]. Once entry to the peritoneal cavity has been achieved, gas insufflation is used to establish pneumoperitoneum for the next surgical process [11, 12]. The pneumoperitoneum establishment step does not take a long time; however, more than 50% of all laparoscopic procedure complications occur during this step [13, 14]. In most of the current practices, the Veress needle is blindly inserted, and appropriate needle positioning is largely dependent on prior experience from the surgeon. Complications such as subcutaneous emphysema and gas embolism, or injury to internal organs during abdominal entry could happen when the Veress needle is not appropriately inserted [15]. While the average incidence rate of severe needle

injury is below 0.05%, there are more than 13 million laparoscopic procedures performed annually worldwide. Thousands of patients suffer from needle insertion injuries each year [16-18]. The most common injuries are lesions of abdominal organs, especially small intestine injuries[19].

Imaging methods have been proposed in guiding the Veress needle insertion. For instance, ultrasound has been used to visualize the different layers of abdominal wall in guiding the Veress needle insertion [20, 21]. Magnetic resonance (MRI) imaging has been utilized to assist accurately measure the Veress needle insertion depth [22]. In addition, virtual reality technique has been proved to be a useful tool for Veress needle insertion [23]. Nevertheless, these techniques cannot accurately locate the needle tip because of the limited resolution and tissue deformation during needle insertion [24, 25]. Therefore, new techniques that can better guide the Veress needle is critically needed.

Optical coherence tomography (OCT) is an established biomedical imaging technique that can visualize subsurface tissue [26]. OCT provides high axial resolution at ~10 μm and several millimeters' imaging depth [27], thus OCT has the potential for providing better imaging quality in Veress needle guidance. However, benchtop OCT cannot be directly used for Veress needle guidance due to the limited penetration depth. Endoscopic OCT systems has been applied in many surgical guidance procedures such as the investigation of colon cancer [28], vitreoretinal surgery [29] and nasal tissue detection [30]. In our previous work, we developed OCT endoscope based on gradient-index (GRIN) rod lens and demonstrated its feasibility in real-time percutaneous nephrostomy (PCN) guidance [31] and epidural anesthesia guidance [32].

In this study, we adapted the OCT endoscope for Veress needle guidance. To simulate the laparoscopy procedure, the endoscopic OCT system was used to image different tissue layers including subcutaneous fat, muscle, abdominal space, and small intestine from swine abdominal tissue. These tissues can be recognized based on their distinct OCT imaging features. To assist doctors, convolutional neural networks (CNNs) [33, 34] were developed for recognizing the different types of tissues and estimating the exact distance between needle tip and the small intestine from OCT images. OCT images were taken from four tissue layers (Subcutaneous fat, muscle, abdominal space, and small intestine) along the path of the Veress needle. These images were then used to train and test a classification model for tissue layer recognition and a regression model for estimation of the distance from the tip of the needle to the small intestine. The CNN architectures used for the classification and regression tasks included ResNet50 [35], InceptionV3 [36], and Xception [37]. Results from these three architectures were analyzed and benchmarked [38-40]. To the best of our knowledge, this is the first report to combine endoscopic OCT system with CNN as a novel imaging platform for guiding the Veress needle procedure and these preliminary results demonstrated the feasibility of this novel imaging strategy.

3.2 Methods

3.2.1 Experimental setup

The endoscopic OCT system was established on a swept-source OCT (SS-OCT) system which applied the laser source with 1300nm-center wavelength and 100nm-bandwidth [31]. Its schematic was shown in Figure 1. The laser source provided output power of around 25 mW. The system had the axial scanning (A-scan) rate of up to 200 kHz. The light was initially split by a coupler into two different parts, one of which accounted for 97% of the

total power and the other one took the rest 3%. A Mach-Zehnder interferometer (MZI) received the rest 3% power and generated a frequency-clock signal to trigger the imaging sampling process. The 97%-light was further transmitted to an optical circulator. When the light exited from port 2, it was split evenly into the sample arm and reference arm of the OCT system. Polarization controllers were assembled in each coherence arm to control the noise levels. The interference signal of the backscattered light from sample arm and the reflected beam from reference arm was sent to the balanced detector (BD) for further noise reduction, and then to the data acquisition (DAQ) board and computer for post processing. Cross-sectional OCT images from different imaging depths of the sample can be provided through Fourier transform.

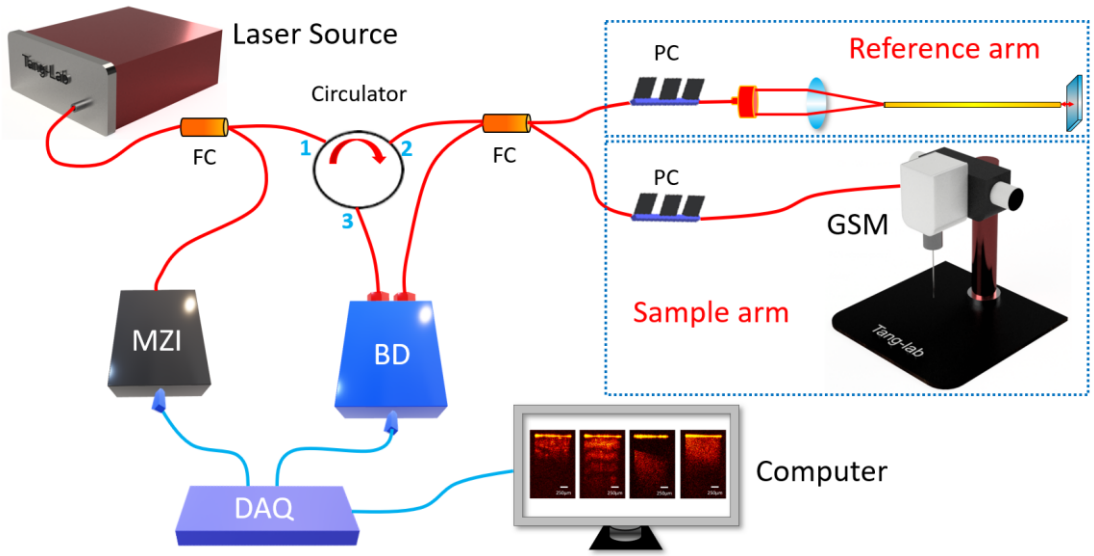


Figure 3.1. Schematic of endoscopic OCT system for Veress needle guidance. MZI: Mach-Zehnder interferometer, BD: Balanced detector, DAQ: Data acquisition, GSM: Galvanometer scanning mirror, FC: Fiber coupler, PC: Polarization controller.

To build the endoscopic system, we used gradient-index (GRIN) rod lens as endoscopes. One GRIN lens was stabilized in front of the galvanometer scanner as demonstrated in Figure 3.1. The GRIN lens transmitted the imaging information from the distal end to the proximal end with the spatial resolution remained constant. We placed the proximal surface of the GRIN lens at the focal point of the OCT scanner to acquire the tissue images in front of the GRIN lens. To compensate for light dispersion, we placed another identical GRIN lens into the light path of the reference arm. The GRIN lenses in our experiment had the length of 138 mm and diameter ~1.30 mm. Stainless steel tubes were assembled to protect them. Lateral field-of-view (FOV) of our system reached ~1.25 mm, and the sensitivity was calibrated to be ~92 dB. The endoscopic OCT achieved ~11 μm-axial resolution and 20 μm-transverse resolution.

3.2.2 Data acquisition

The OCT images of subcutaneous fat, muscle, abdominal space, and small intestine were taken from eight pigs. For each sample, 1,000 2D OCT cross-sections were selected. A total 32,000 images were utilized for CNN tissue classification model development. For the distance estimation task, there were a total of 8,000 OCT images of abdominal space used, and distance between the GRIN lens end and the small intestine varied in these images. The flow diagram of data acquisition and process was demonstrated in Figure 3.2.

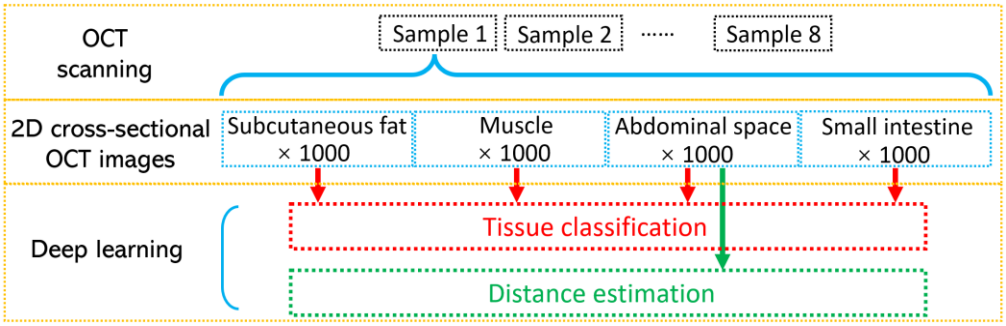


Figure 3.2. Process of the data acquisition.

The original size of each image was 320 (transverse/X axis) × 480 (longitudinal axis/depth/Z direction). The pixel size on both axes were 6.25μm. To decrease the computation burden, the size of the 2D images was reduced by cropping the unnecessary part of the images' edges. They were cropped to 216×316 (X by Z) for tissue classification and 180×401(X by Z) for distance estimation.

3.2.3 CNN method for tissue classification

CNNs were used to accomplish the task of identifying the layer of tissues in which an image was taken. The four layers analyzed for classification include fat, muscle, abdominal space, and intestine. For each subject, there were 1,000 images taken from each tissue layer, and the layer in which the image was taken was manually annotated and represented the ground truth label for the classification task. The CNN model architectures used for model development included ResNet50 [35], InceptionV3 [36] and Xception [37], which contained 25.6 million, 23.9 million and 22.9 million parameters, respectively. Training took place over 20 epochs with a batch size of 32. The cross-entropy optimizer was used with Nesterov momentum, a learning rate of 0.9, and a decay rate of 0.01. The loss function was sparse categorical cross-entropy and accuracy was used as the primary evaluation metric for the tissue classification task. ResNet50, Xception, and InceptionV3 were used because of their demonstrated performance on similar image prediction tasks and relatively comparable network depths [35-37]. Initially, a wide range of architectures were selected and tested, including EfficientNet (B3, B4, and B5), InceptionV3, NasNetLarge, NasNetMobile, ResNet50, ResNet101, and Xception. On average, the ResNet50, Xception, and InceptionV3 architectures supported the best performing models in regard to accuracy and efficiency.

The accuracy was calculated as:

$$Accuracy = (TP + TN)/(TP + TN + FP + FN) \quad (1)$$

where TP: true positive; TN: true negative; FP: false positive; FN: false negative.

Nested cross-validation and cross-testing were used to select an optimal model and evaluate the performance [31]. For nested cross-validation and testing, images were separated into eight folds based on the subject in which the images were taken. With nested cross-validation, the performance of the three model architectures were compared, and the architecture with the highest accuracy was selected for the corresponding testing phase. Once the optimal model architecture was determined for a given nested cross-validation fold, the cross-testing phase required training a new model with all images except those from the corresponding testing fold, and the new model's accuracy on the unseen testing images was recorded. This process was illustrated in Figure 3.3. Training and testing took place on ten compute nodes containing GPUs on the Summit supercomputer at Oak Ridge National Laboratory.

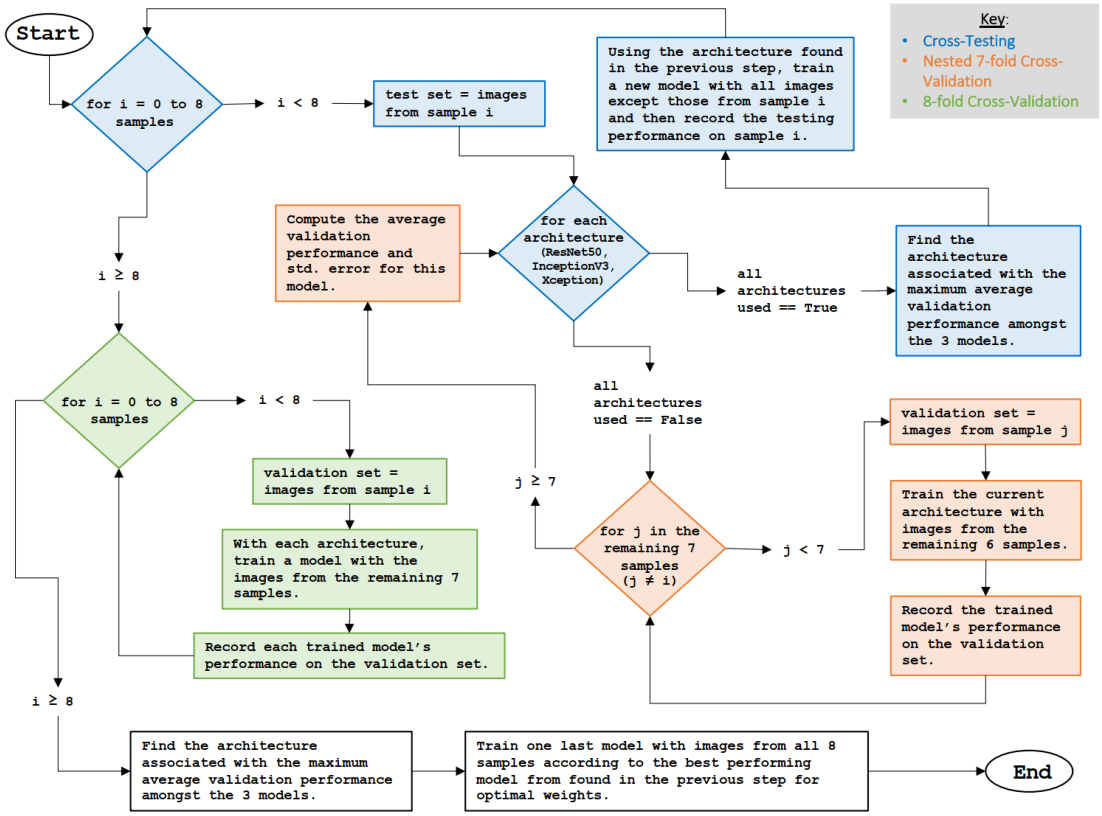


Figure 3.3. Introduction of the nested cross-validation, cross-testing, and 8-fold cross-validation process.

3.2.4 CNN regression for distance measurement

Regression CNNs were used to estimate the distance from the Veress needle lens to the intestine. The distance from the needle tip to the intestine represented the ground truth label and was manually annotated.

CNN regression models were constructed with the same three architectures used for classification, which includes ResNet50, InceptionV3, and Xception, and the same nested cross-validation and cross-testing approach was used for training and performance evaluation. For the regression model architecture, the final output layer was changed to a single neuron with an identity activation function. Training involved using the SGD optimization algorithm,

a learning rate of 0.01, decay rate of 0.09, and Nesterov momentum. Training took place over 20 epochs with a batch size of 32. The loss function was the mean absolute percentage error (MAPE). Regression model development was accomplished on a private workstation containing two NVIDIA RTX 3090 GPUs. Here, MAPE and mean absolute error (MAE) were utilized to evaluate the estimation accuracy of the distance. They can be described as:

$$MAPE(\%) = (100\%/n)\sum_{i=1}^{n}|Y_i - X_i|/Y_i \quad (2)$$

$$MAE = (1/n)\sum_{i=1}^{n}|Y_i - X_i| \quad (3)$$

Where X_i was the estimated value, Y_i was the ground truth value (manually labeled), and n was the number of the images.

3.3 Results

3.3.1 Imaging results of endoscopic OCT system

We used the subcutaneous fat, muscle, abdominal space, and small intestine tissues from eight different swine samples to mimic the practical tissue layers that the Veress needle traverses. Fat and muscle were both taken from the abdominal areas. In the experiment, the GRIN lens in the sample arm was inserted into the three different tissues for imaging. Moreover, to replicate the condition when the needle tip was in the abdominal cavity, we kept the needle tip at different distances in front of the small intestine and took the OCT images as the abdominal space layer.

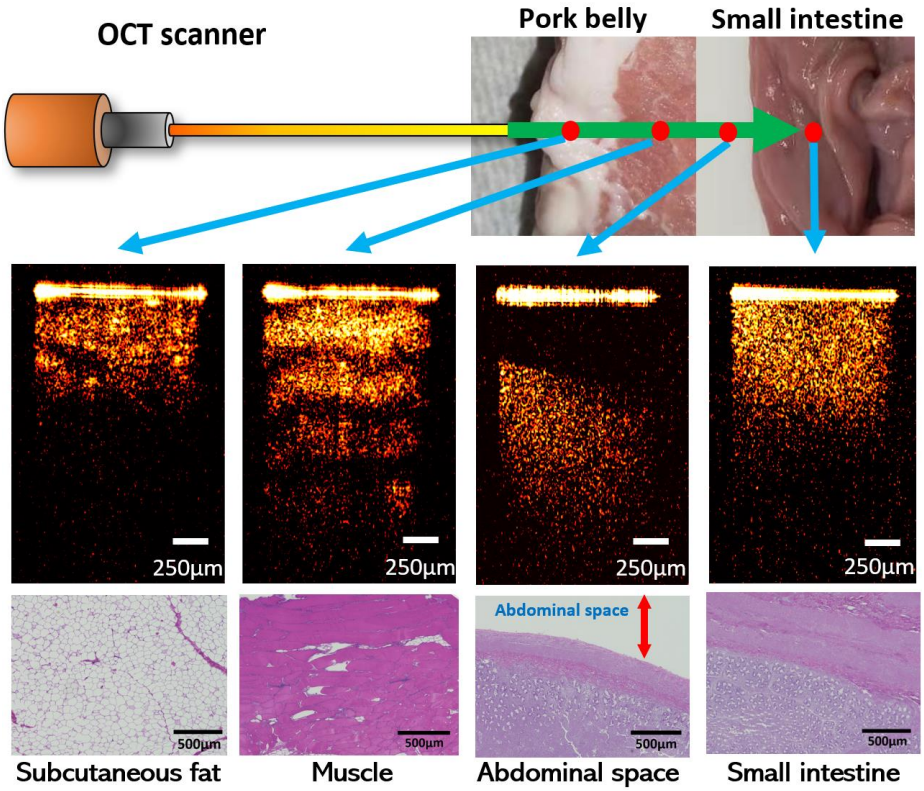


Figure 3.4. Examples of OCT images of different tissue types.

Figure 3.4 showed the examples of 2D OCT results of the three different tissues and abdominal space. The 2D results demonstrated clear differences between each other: Abdominal space could be easily recognized from the gap between the tip of the GRIN lens and the small intestine tissue on the bottom. Among the three tissues, muscle had clear transverse fiber structures which was shown as different light and dark layers on the OCT image and had the largest penetration depth. The imaging result of small intestine showed homogeneous tissue density and brightness. Regarding the subcutaneous fat images, some granular structures occurred because of the existence of adipocytes. The corresponding histology results were also included. Different tissues presented distinct cellular structures and distributions and correlated well with their OCT images. These results proved the potential of using OCT to distinguish different tissue layers.

3.3.2 Multi-class classification

There were 32,000 total images with the size of 216×316 (X by Z) pixels taken from eight subjects used for training and testing the tissue classification model. Three CNN models: ResNet50 [35], InceptionV3 [41] and Xception [42] were applied. The average nested 7-fold cross-validation accuracies for tissue-layer classification were shown in Table 3.1. All models achieved 90% accuracy or higher on the validation set. For InceptionV3 and Xception models, the average accuracies were both higher than 97%.

Table 3.1: The average nested 7-fold cross-validation accuracies and standard errors.

Fold	ResNet50		InceptionV3		Xception	
	Accuracy	SE*	Accuracy	SE*	Accuracy	SE*
S1	92.69%	2.11%	98.31%	0.58%	98.93%	0.40%
S2	97.41%	0.94%	98.45%	0.47%	98.86%	0.38%
S3	94.63%	1.67%	98.15%	0.58%	98.44%	0.73%
S4	96.93%	0.54%	98.54%	0.38%	98.49%	0.40%
S5	91.94%	1.72%	97.84%	0.60%	99.02%	0.33%
S6	95.88%	0.97%	97.07%	1.15%	98.56%	0.51%
S7	90.75%	2.60%	98.01%	0.55%	98.60%	0.46%
S8	96.90%	1.14%	97.95%	0.56%	98.55%	0.38%

*** SE is the standard error of the average accuracy.**

Cross-testing was further performed to provide an unbiased evaluation of the classification performance on the set of test images (i.e., images that were not used during nested cross-validation). The architecture associated with the highest nested cross-validation accuracy in each cross-validation fold was used to train a new model for the corresponding cross-testing fold. During cross-testing, a new model was trained with images from seven subjects (7,000 images) in both the training and validation folds and tested with images from one subject (1,000 images) in the test fold. The testing results along with inferencing times were shown in Table 3.2. The best testing accuracy on the testing images was 99.825% in the

S6 testing fold. There was a tie for the lowest testing accuracy (97.200%) in the S2 and S3 testing fold. The classification inferencing time is the amount of time it took for the classification model to predict the tissue layer on a single image. The inferencing time for Xception was slightly greater than that of InceptionV3.

Table 3.2: The 8-fold cross-testing accuracies on the testing set in each testing fold.

Fold	Architecture	Inferencing time (ms)	Testing accuracy
S1	Xception	1.765	97.95%
S2	Xception	1.743	97.20%
S3	Xception	1.738	97.20%
S4	InceptionV3	1.262	97.98%
S5	Xception	1.758	98.80%
S6	Xception	1.765	99.83%
S7	Xception	1.740	99.80%
S8	Xception	1.745	99.50%
Average			98.53 ± 0.39%

From the classification cross-testing benchmarking results, the Xception architecture was used in 7 out of the 8 cross-testing folds, and the InceptionV3 architecture was used once (selected via cross-validation). The aggregated ROC curve across all eight testing folds (i.e., 32,000 images) was shown in Figure 3.5. The ROC curve showed that the models were able to classify the images pertaining to each tissue-layer with high accuracy. The average ROC AUC score was 0.998967 and the r^2 value was 0.9839.

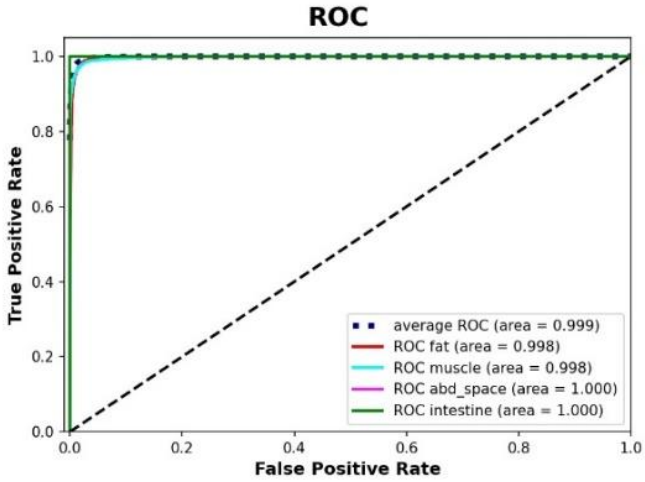


Figure 3.5. Aggregated ROC across all 8 testing folds using Xception model.

To visually explain the model predictions of the different tissue layers, a tissue classification activation heatmap from sample 1 based on Xception was illustrated in Figure 3.6. The activation distributions differed in position and size among the four tissue layer types. In subcutaneous fat and small intestine, the activation was mainly concentrated in the bottom of the images. However, the size of the activated area on the image of subcutaneous fat was larger than the activated area on the image of the small intestine, and there was also increased attention given to the area right below the needle tip. For muscle and abdominal space, the activated areas on the image were near the top, but there was greater focus given to the area just under the needle tip in muscle images. In contrast to the muscle images, the activated areas of the abdominal space images were more concentrated over the needle tip.

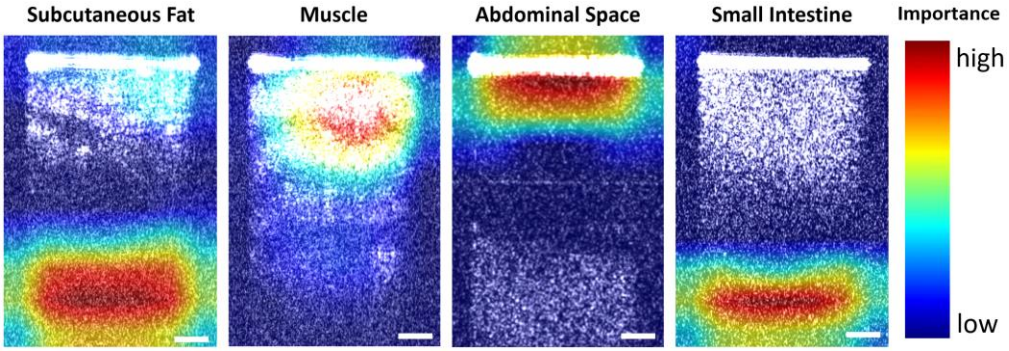


Figure 3.6. Tissue classification activation heatmap obtained from sample 1. Scale bar: 250μm.

After the performance of our model development procedure was benchmarked by the cross-testing, a final model was generated using this procedure in two steps. First, an architecture was selected using 8-fold cross-validation. As expected, the Xception architecture provided higher accuracy on average (98.37±0.59%) than ResNet50 (95.92±1.40%) and InceptionV3 (97.46±0.75%). Then, a final model was trained using all 8 folds with the Xception architecture.

3.3.3 Distance estimation

There were 8000 images in total used for the regression task (1000 images per subject). The images of the abdominal space were taken at a range of distances between approximately 0.2 mm and 1.5 mm from the needle tip to the intestine. To estimate the distance values between the needle tip and the surface of small intestine, the same three architectures (ResNet50, InceptionV3, and Xception) were also utilized for the regression task. The MAPE was used to evaluate the distance estimation error. Nested cross-validation and cross-testing were performed in the same fashion as done the tissue-layer classification task. Table 3.3 showed the average nested cross-validation MAPE with standard error in each nested cross-validation fold. The distance estimation nested cross-validation results indicated that the InceptionV3 achieved the lowest average error in six out of eight folds, and Xception achieved the lowest average error in the other two folds.

Table 3.3. Average nested 7-fold cross-validation MAPE with standard error.

Fold	ResNet50		InceptionV3		Xception	
	MAPE	SE*	MAPE	SE*	MAPE	SE*
S1	5.25%	0.78%	**4.47%**	0.79%	5.06%	0.89%
S2	5.01%	0.40%	**3.79%**	0.39%	4.24%	0.44%
S3	5.43%	0.74%	**4.74%**	0.65%	5.06%	0.76%
S4	5.42%	0.53%	5.16%	0.76%	**5.05%**	0.54%
S5	5.47%	0.82%	**4.43%**	0.44%	4.92%	0.47%
S6	5.62%	0.72%	**5.03%**	0.60%	5.42%	0.63%
S7	5.89%	0.64%	**4.74%**	0.60%	5.39%	0.75%
S8	4.91%	0.53%	4.90%	0.64%	**4.35%**	0.40%

*** SE is the standard error of the mean MAPE.**

Similar to tissue classification, cross-testing was used to get an unbiased evaluation of the performance on the distance estimation task, and results were shown in Table 3.4. InceptionV3 provided the highest MAPE of 6.66% with MAE of 56.1 µm in fold S2, and the lowest MAPE of 2.07% with MAE of 16.7 µm in fold S7. Overall, the MAPEs were all under 7% and the MAEs never exceeded 60 µm. Examples of comparisons between the manually labeled results (ground truth value) and predicted results were shown in Figure 3.7(a). Violin plots from sample one was shown in Figure 3.7(b) to show the error distribution based on distance.

Table 3.4. MAPE with Standard Error and MAE with Standard Error during 8-fold cross-testing.

Testing fold	Architecture	MAPE	MAE (µm)	Inference time (ms)
S1	**InceptionV3**	6.45%	52.9	1.395
S2	**InceptionV3**	6.66%	56.1	1.392
S3	**InceptionV3**	5.09%	45.6	1.380
S4	**Xception**	3.38%	31.7	1.669
S5	**InceptionV3**	3.78%	26.2	1.395
S6	**InceptionV3**	3.43%	31.6	1.415
S7	**InceptionV3**	2.07%	16.7	1.394
S8	**Xception**	4.46%	27.9	2.044
Average		4.42% ± 0.56%	36.09 ± 4.92	

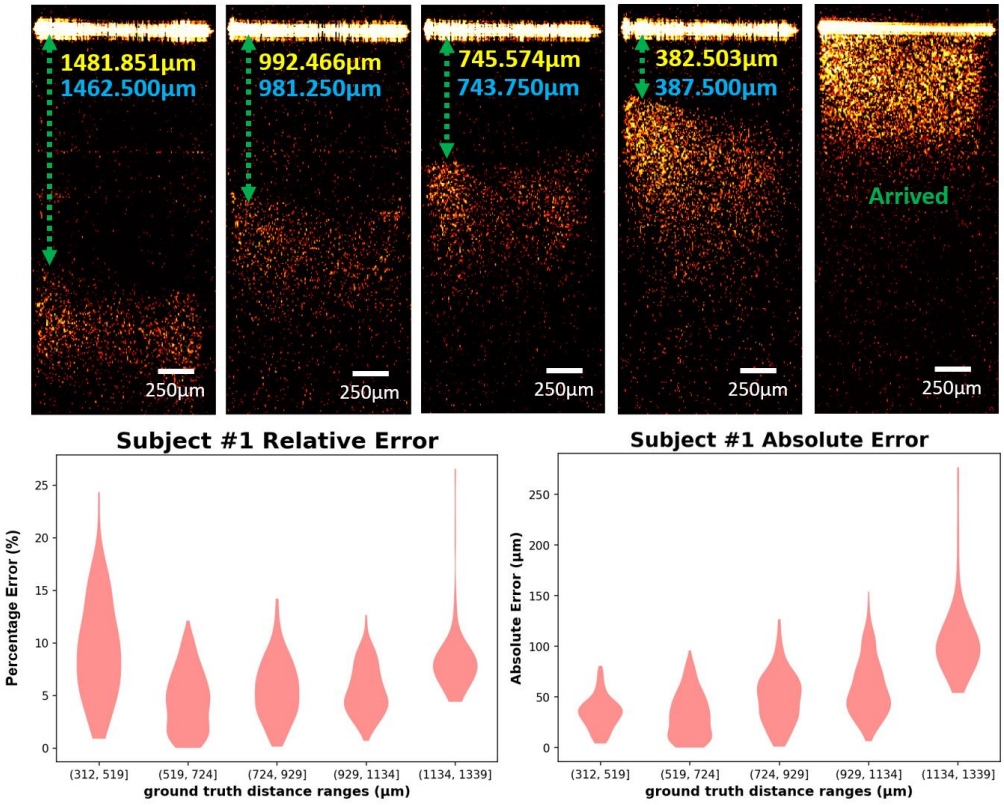


Figure 3.7. (A) Examples of abdominal space images with different distances between small intestine and needle tip. Yellow value: estimated distance; Blue value: Manually calculated distance. (B) Violin plots showing the absolute error and relative error on the testing set for subject 1.

After the regression model development procedure was benchmarked by cross-testing, this procedure was repeated to produce a final model. Because no data needed to be held back for testing, all 8 folds were used in the cross validation for final architecture selection. The models trained with the InceptionV3 architecture provided lower error on average (4.44±0.43%) than ResNet50 (5.29±0.56%) and Xception (4.77±0.58%). After the architecture was selected, because no data needed to be held back for validation or testing, all 8 folds were used to train the final model.

3.4 Discussion

In this study, we demonstrated the feasibility of our forward-view endoscopic OCT system for Veress needle guidance. Compared to conventional imaging methods, OCT can provide more structural details of the subsurface tissues to help recognize the tissue type in front of the needle tip. Four tissue layers following the sequence of the tissues that Veress needle passing through during the surgery were imaged by our endoscopic OCT system, including subcutaneous fat, muscle, abdominal space, and small intestine. The OCT images of these four layers could be distinguished by their unique imaging features. By fitting the rigid OCT endoscope inside the hollow bore of the Veress needle, no additional invasiveness will be introduced from the OCT endoscope. The OCT endoscope will provide the images of the tissues in front of the Veress needle during insertion, thus indicating the needle tip location in real time and facilitating the precise placement of the Veress needle.

Deep learning was used to automate the OCT imaging data processing. Three CNN architectures, including ResNet50, InceptionV3 and Xception, were cross-validated for both tasks. These three architectures were used because of their demonstrated performance on similar image prediction tasks and relatively comparable network depths [35-37]. Initially, a wide range of architectures were selected and tested, including EfficientNet (B3, B4, and B5), InceptionV3, NasNetLarge, NasNetMobile, ResNet50, ResNet101, and Xception. On average, the ResNet50, Xception, and InceptionV3 architectures supported the best performing models in regard to accuracy and efficiency. Among these three architectures, the best architecture was found to be Xception for tissue layer classification and InceptionV3 for estimating the distance from needle tip to small intestine surface. However, all three architectures provided very high prediction performance and had only insignificant performance differentials among

them. We used a nested cross-validation and cross-testing to provide an unbiased performance benchmarking of our model development procedure from architecture selection to model training. The average testing accuracy of our procedure was 98.53 ± 0.39% for tissue layer classification. The average MAPE of our procedure was 4.42% ± 0.56% for the distance estimation.

For the classification task, the training time per fold over 28,000 images was ~98 minutes on average for the Xception architecture and ~32 minutes for the InceptionV3 architecture during cross testing. The average inferencing time (i.e., the time it took for a trained model to make a prediction on a single image) for the Xception models during cross-testing was 1.75 milliseconds, while the InceptionV3 model had an interpretation time of 1.26 milliseconds. The classifiers were trained and tested using NVIDIA Volta GPUs. For the regression task, the average training time for the InceptionV3 model during cross-testing was ~367 minutes, which took place over 7,000 images. The average training time for the Xception model was ~1,244 minutes. The average interpretation time for the InceptionV3 models was 1.30 milliseconds. And the average interpretation time for the Xception model was ~1.86 milliseconds. Regression model training and testing took place on NVIDIA RTX 3090 GPUs.

Our current study has shown the feasibility of using endoscopic OCT and deep learning methods in Veress needle guidance. Next, we will apply our system in the *in-vivo* swine experiments. It is worth mentioning that blood flows exist during *in-vivo* experiments. Except for the lesions of abdominal organs, injury to the blood vessels is also a major complication in the Veress needle insertion [43]. Major vascular injuries, especially to aorta, vena cava or iliac vessels, are risky to patients' lives [44]. The mortality rate can reach up to

17% when injury to large vessels happen [45]. Doppler OCT is an extension of our current OCT endoscope that can help detect the flows [46, 47]. Doppler OCT endoscope has been used for detecting the at-risk blood vessels within sheep brain in real time [48], in colorectal cancer (CRC) diagnosis [49], management of pulmonary nodules [50], and human GI tract imaging and treatment [51]. Therefore, the proposed OCT endoscope has potential to solve the problem of blood vessel injury during Veress needle insertion. As to hardware, we will redesign the OCT scanner to make it easier for surgeons to operate. We will continue to accelerate the models through knowledge distillation and weight pruning. Furthermore, Since the proposed OCT endoscope system can distinguish different tissue types in front of needle tip, it also has potential for guiding other needle-based interventions such as percutaneous nephrostomy (PCN) needle guidance in kidney surgery [52], epidural anesthesia imaging guidance in painless delivery [32], the tumor tissue detection in cancer diagnosis [53], and a variety of needle biopsy procedures.

Reference

1. Hasson, H.M., et al., *Open laparoscopy: 29-year experience.* Obstet Gynecol, 2000. **96**(5 Pt 1): p. 763-6.
2. McKernan, J.B. and J.K. Champion, *Access techniques: Veress needle--initial blind trocar insertion versus open laparoscopy with the Hasson trocar.* Endosc Surg Allied Technol, 1995. **3**(1): p. 35-8.
3. Vilos, G.A., et al., *Laparoscopic entry: a review of techniques, technologies, and complications.* J Obstet Gynaecol Can, 2007. **29**(5): p. 433-447.
4. Reynolds, W., Jr., *The first laparoscopic cholecystectomy.* JSLS, 2001. **5**(1): p. 89-94.
5. Frazee, R.C., et al., *A prospective randomized trial comparing open versus laparoscopic appendectomy.* Ann Surg, 1994. **219**(6): p. 725-8; discussion 728-31.
6. Giseke, S., et al., *A true laparoscopic herniotomy in children: evaluation of long-term outcome.* J Laparoendosc Adv Surg Tech A, 2010. **20**(2): p. 191-4.
7. Belachew, M., et al., *Laparoscopic adjustable gastric banding.* World J Surg, 1998. **22**(9): p. 955-63.

8. Jacobs, M., J.C. Verdeja, and H.S. Goldstein, *Minimally invasive colon resection (laparoscopic colectomy)*. Surg Laparosc Endosc, 1991. **1**(3): p. 144-50.
9. Srivastava, A. and A. Niranjan, *Secrets of safe laparoscopic surgery: Anaesthetic and surgical considerations*. J Minim Access Surg, 2010. **6**(4): p. 91-4.
10. Netter, F.H., *Atlas of human anatomy*, in *Netter Basic Science*. 2019, Saunders/Elsevier,: Philadelphia, PA. p. 1 online resource.
11. Volz, J., et al., *Pathophysiologic features of a pneumoperitoneum at laparoscopy: a swine model*. Am J Obstet Gynecol, 1996. **174**(1 Pt 1): p. 132-40.
12. Tanaka, C., et al., *Optical trocar access for initial trocar placement in laparoscopic gastrointestinal surgery: A propensity score-matching analysis*. Asian J Endosc Surg, 2019. **12**(1): p. 37-42.
13. Vilos, G.A., *Litigation of laparoscopic major vessel injuries in Canada*. J Am Assoc Gynecol Laparosc, 2000. **7**(4): p. 503-9.
14. Vilos, G.A., *Laparoscopic bowel injuries: forty litigated gynaecological cases in Canada*. J Obstet Gynaecol Can, 2002. **24**(3): p. 224-30.
15. Azevedo, J.L., et al., *Injuries caused by Veress needle insertion for creation of pneumoperitoneum: a systematic literature review*. Surg Endosc, 2009. **23**(7): p. 1428-32.
16. Jansen, F.W., et al., *Complications of laparoscopy: an inquiry about closed-versus open-entry technique*. Am J Obstet Gynecol, 2004. **190**(3): p. 634-8.
17. Magrina, J.F., *Complications of laparoscopic surgery*. Clin Obstet Gynecol, 2002. **45**(2): p. 469-80.
18. iDataResearch. *Over 13 Million Laparoscopic Procedures are Performed Globally Every Year* https://idataresearch.com/over-13-million-laparoscopic-procedures-are-performed-globally-every-year/. 2020.
19. Schafer, M., M. Lauper, and L. Krahenbuhl, *Trocar and Veress needle injuries during laparoscopy*. Surg Endosc, 2001. **15**(3): p. 275-80.
20. Ra, Y.S., et al., *The analgesic effect of the ultrasound-guided transverse abdominis plane block after laparoscopic cholecystectomy*. Korean J Anesthesiol, 2010. **58**(4): p. 362-8.
21. Pratap, A., D. Oleynikov, and V. Kothari, *Real time ultrasound guided insertion of Veress needle in obese patients*. Ann R Coll Surg Engl, 2018. **100**(2): p. 158-159.
22. Afifi, Y., et al., *New nomogram for safe laparoscopic entry to reduce vascular injury*. J Obstet Gynaecol, 2011. **31**(1): p. 69-72.
23. Okrainec, A., et al. *Development of a virtual reality haptic Veress needle insertion simulator for surgical skills training*. in *MMVR*. 2009.
24. Beregi, J.P. and J. Greffier, *Low and ultra-low dose radiation in CT: Opportunities and limitations*. Diagn Interv Imaging, 2019. **100**(2): p. 63-64.
25. Carvalho, J.C., *Ultrasound-facilitated epidurals and spinals in obstetrics*. Anesthesiol Clin, 2008. **26**(1): p. 145-58, vii-viii.
26. Huang, D., et al., *Optical coherence tomography*. Science, 1991. **254**(5035): p. 1178-81.

27. Fujimoto, J. and E. Swanson, *The Development, Commercialization, and Impact of Optical Coherence Tomography.* Invest Ophthalmol Vis Sci, 2016. **57**(9): p. OCT1-OCT13.
28. Tumlinson, A.R., et al., *Miniature endoscope for simultaneous optical coherence tomography and laser-induced fluorescence measurement.* Appl Opt, 2004. **43**(1): p. 113-21.
29. Han, S., et al., *Handheld forward-imaging needle endoscope for ophthalmic optical coherence tomography inspection.* J Biomed Opt, 2008. **13**(2): p. 020505.
30. Schulz-Hildebrandt, H., et al., *Novel endoscope with increased depth of field for imaging human nasal tissue by microscopic optical coherence tomography.* Biomed Opt Express, 2018. **9**(2): p. 636-647.
31. Wang, C., et al., *Deep-learning-aided forward optical coherence tomography endoscope for percutaneous nephrostomy guidance.* Biomed Opt Express, 2021. **12**(4): p. 2404-2418.
32. Tang, Q.G., et al., *Real-time epidural anesthesia guidance using optical coherence tomography needle probe.* Quantitative Imaging in Medicine and Surgery, 2015. **5**(1): p. 118-124.
33. Litjens, G., et al., *A survey on deep learning in medical image analysis.* Medical Image Analysis, 2017. **42**: p. 60-88.
34. Shen, D., G. Wu, and H.-I. Suk, *Deep Learning in Medical Image Analysis.* Annual Review of Biomedical Engineering, 2017. **19**(1): p. 221-248.
35. He, K., et al. *Deep residual learning for image recognition*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.
36. Szegedy, C., et al. *Rethinking the inception architecture for computer vision*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.
37. Chollet, F. *Xception: Deep learning with depthwise separable convolutions*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.
38. Raschka, S., *Model evaluation, model selection, and algorithm selection in machine learning.* arXiv preprint arXiv:1811.12808, 2018.
39. Varma, S. and R. Simon, *Bias in error estimation when using cross-validation for model selection.* BMC Bioinformatics, 2006. **7**(1): p. 91.
40. Iizuka, N., et al., *Oligonucleotide microarray for prediction of early intrahepatic recurrence of hepatocellular carcinoma after curative resection.* The Lancet, 2003. **361**(9361): p. 923-929.
41. Xia, X., C. Xu, and B. Nan. *Inception-v3 for flower classification*. in *2017 2nd International Conference on Image, Vision and Computing (ICIVC)*. 2017. IEEE.
42. Kassani, S.H., et al. *Diabetic retinopathy classification using a modified xception architecture*. in *2019 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*. 2019. IEEE.

43. Schafer, M., M. Lauper, and L. Krahenbuhl, *Trocar and Veress needle injuries during laparoscopy*. Surgical Endoscopy-Ultrasound and Interventional Techniques, 2001. **15**(3): p. 275-280.
44. Azevedo, J.L.M.C., et al., *Injuries caused by Veress needle insertion for creation of pneumoperitoneum: a systematic literature review*. Surgical Endoscopy and Other Interventional Techniques, 2009. **23**(7): p. 1428-1432.
45. Olsen, D.O., *Laparoscopic Cholecystectomy*. American Journal of Surgery, 1991. **161**(3): p. 339-344.
46. Izatt, J.A., et al., *In vivo bidirectional color Doppler flow imaging of picoliter blood volumes using optical coherence tomograghy*. Optics Letters, 1997. **22**(18): p. 1439-1441.
47. Chen, Z., et al., *Optical Doppler tomographic imaging of fluid flow velocity in highly scattering media*. Optics Letters, 1997. **22**(1): p. 64-66.
48. Liang, C.P., et al., *A forward-imaging needle-type OCT probe for image guided stereotactic procedures*. Opt Express, 2011. **19**(27): p. 26283-94.
49. Welge, W.A. and J.K. Barton, *In vivo endoscopic Doppler optical coherence tomography imaging of the colon*. Lasers in Surgery and Medicine, 2017. **49**(3): p. 249-257.
50. Pahlevaninezhad, H., et al., *Endoscopic Doppler optical coherence tomography and autofluorescence imaging of peripheral pulmonary nodules and vasculature*. Biomedical Optics Express, 2015. **6**(10): p. 4191-4199.
51. Yang, V.X.D., et al., *Endoscopic Doppler optical coherence tomography in the human GI tract: initial experience*. Gastrointestinal Endoscopy, 2005. **61**(7): p. 879-890.
52. Wang, C., et al., *Deep-learning-aided forward optical coherence tomography endoscope for percutaneous nephrostomy guidance*. Biomedical Optics Express, 2021. **12**(4): p. 2404-2418.
53. Wang, C., et al., *Laparoscopic optical coherence tomography system for 3D bladder tumor detection*. SPIE BiOS. Vol. 11619. 2021: SPIE.

CHAPTER 4: LAPAROSCOPIC OPTICAL COHERENCE TOMOGRAPHY SYSTEM FOR 3D BLADDER TUMOR DETECTION

This chapter is reproduced from "Chen Wang, Venkateshwar Madka, Feng Yan, Sanjay G. Patel, Chinthalapally V. Rao, Qinggong Tang, *Laparoscopic Optical coherence tomography system for 3D bladder tumor detection*, Proc. SPIE 11619, **Advanced Photonics in Urology**, 116190J, 2021."

4.1 Introduction

The estimated number of new cases for bladder cancer in the United States in 2021 is 83,730 (about 64,280 in men and 19,450 in women) [1]. Bladder cancer originates in the urothelium (innermost lining) of the bladder and is curable if diagnosed and treated early but has a high mortality rate in advanced stages (after spreading to the bladder's muscle wall and beyond) [2, 3]. However, accurate detection of bladder tumor in early stage is still a critical challenge in bladder cancer diagnosis and treatment.

Endoscopy (blue light or white light cystoscopy) guided excisional biopsy and histopathology is currently the gold standard for bladder cancer diagnosis. The procedure to biopsy an abnormal area is a transurethral resection (TUR). For this operation, a type of rigid cystoscope called resectoscope is placed into the bladder through the urethra. The resectoscope has a wire loop at its end to remove any abnormal tissues or tumors. The removed tissue is sent to a lab to be looked at by a pathologist to determine if someone has bladder cancer and, if so, whether the cancer has invaded the muscle layer of the bladder wall (depth of tumor penetration), which is critical for deciding the following treatment plan. Nevertheless, the current guided biopsy can only image the bladder surface. Cystoscopy cannot achieve subsurface visualization, which results that the conventional guided biopsy has relatively high false negative rates [2-4]. In addition to obtaining histopathologic information,

TUR is also the most effective treatment for early-stage or superficial (non-muscle invasive) bladder cancers. Unfortunately, the current cystoscopy can only provide tissue structure on the urothelium surface, it is difficult to identify 3D tumor margin and residual cancer cells during TUR, which results in tumor recurrence or excessive tissue resection. Because of this, up to 70% of patients develop recurrence and up to 30% can develop metastatic progression of disease which needs lifelong follow-up and repeated treatments, making bladder cancer one of the most expensive cancers to manage [1]. Therefore, a method that can help surgeons to check if there is subsurface residual tumor tissue is of great importance.

Optical coherence tomography (OCT) is a novel biomedical imaging technology for subsurface imaging of tissue with high image resolution (~10 μm) and a penetration depth of several millimeters [5]. OCT system measures the intensity of backscattered light, which provides the information of axial depth position of each scanning point [5, 6]. Three-dimensional (3D) tissue structural reconstruction can be formed based on scanning a specific area of the tissue. Additionally, OCT can be integrated with fiber-optic catheters and endoscopes for internal imaging applications [6-9]. OCT has been proved to be potential in bladder imaging and bladder tumor detection and diagnosis [10, 11]. In our study, in order to image the internal bladder cancer, we developed a laparoscopic OCT imaging system based on a rigid Gradient Index (GRIN) lenses (diameter ~ 4.5 mm). The feasibility of the laparoscopic OCT system was demonstrated by imaging the bladder tumors from UPII-SV40T mice model and algorithm was developed to quantify sub-surface tumor extension. This laparoscopic OCT system is expected to have a major impact on detection, diagnosis and characterization of bladder cancer and other epithelial cancers.

4.2 System Setup

Figure 4.1 shows the schematic diagram of our laparoscopic OCT system [12]. The swept-source OCT (SS-OCT) system we use has the light source with the wavelength 1300 nm. Its axial-scanning rate is 200 kHz with the imaging depth of around 8 mm in the air and 6 mm in water. The laser is split into two parts by a fiber coupler. About 97% of the laser power is input to the circulator, half of which transmits to the reference arm and the other half to the sample arm. The rest 3% enters the Mach-Zehnder interferometer (MZI) to generate the frequency-clock signal to trigger the sampling process. The balanced detector receives the interference fringes of the signals from the two arms and transmit them to data acquisition (DAQ) and computer for the following processing. A GRIN lenses is integrated into the OCT system. The scanning mechanism is placed at the proximal GRIN lenses entrance. The GRIN lenses can preserve the spatial relationship between the entrance and the output (distal end) and further to the sample. A same GRIN lenses is employed in the reference arm of OCT to minimize the dispersion mismatch. This design is extremely compact, solid, and reliable and could be modified for required sizes to perform 3D laparoscopic imaging in a range of clinical settings. The GRIN lens has a diameter ~ 4.5 mm, thus the catheter-based imaging device will be easily integrated with the imaging channel of the standard 27 Fr resectoscope and guide the surgical procedures.

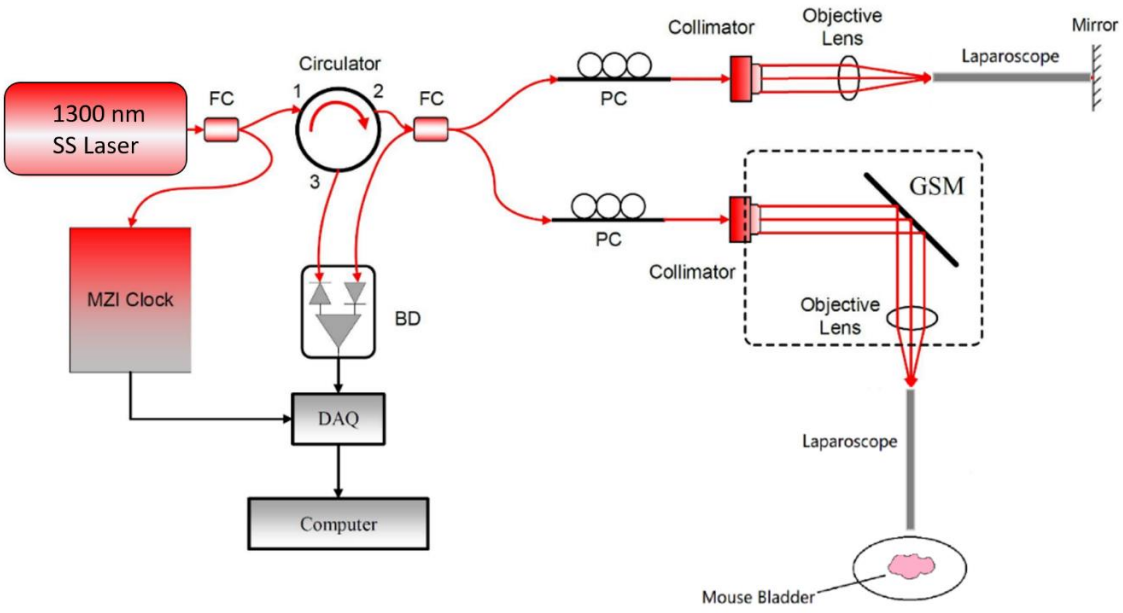


Figure 4.1. Schematic of laparoscopic OCT system. FC: fiber coupler; BD: balanced detector; DAQ: data acquisition; PC: polarization controller; GSM: Galvonometer scanning mirror.

4.3 Results and discussion

The laparoscopic OCT system is shown in Figure 4.2(A) and the image of the bladder tissue is shown in Figure 4.2(B). The cancerous tissues are indicated by the dashed blue line. To quantitatively characterize the cancerous and normal tissues of mouse bladder, we calculated the attenuation coefficients of the 3D image in axial direction as biomarker to indicate local tissue scattering changes associated with neoplasia. The basic principle is based on Beer-Lambert Law: $I(z) = I_0 \times \exp(-2 \times \mu_t \times z)$, where I_0 is the initial intensity, μ_t is the extinction coefficient which consists of absorption coefficient and scattering coefficient, and z is the axial depth. The intensity of the image pixels decreases according to Beer-Lambert law. Compared to normal tissue, cancerous tissue is relatively denser, so its attenuation coefficient is higher than normal tissue.

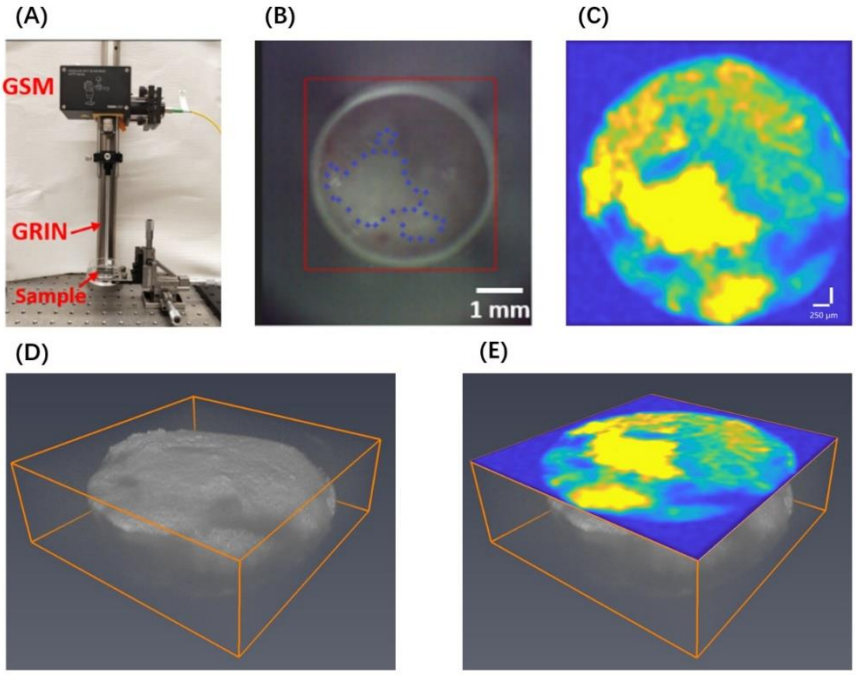


Figure 4.2. (A) Image of the laparoscopic OCT system; (B) Camera image of UPII-SV40T transgenic mouse bladder tissue (The cancerous tissues were labelled by the blue dashed circle); (C) Tissue attenuation coefficient image map to recognize cancerous tissues. (D)&(E) 3D OCT image of the same bladder tissue with the tissue attenuation coefficient map. Image size: X×Y×Z=4.5×4.5×2 (mm^3). GSM: galvonometer scanning mirror, GRIN: Gradient Index lenses.

Therefore, this characteristic can be utilized for tissue differentiation. We calculated the attenuation coefficients on each axial direction of the whole 3D OCT image and then formed the 2D attenuation coefficient image map. Figure 4.2(C) demonstrates the axial attenuation coefficient map of the mouse bladder we imaged. The yellow parts show higher coefficients, which correspond to tumor regions. Figure 4.2(D) shows the 3D OCT image of the bladder tissue and Figure 4.2(E) shows the fused 3D OCT image with the tissue attenuation coefficient map. Therefore, our laparoscopic OCT method can help differentiate

tumors and normal tissues of bladder. By using this system, subsurface bladder tumor can be located during the surgery so doctors can be guided to resect the cancerous region of bladder, which will decrease the risk of the recurrence of bladder tumor, leading to lower patient mortality.

4.4 Conclusion and future work

We have developed a laparoscopic OCT guidance system and tested its feasibility in mice bladder tumor imaging. 3D mice bladder images were successfully obtained. Through calculating the attenuation coefficients on each axial line of the 3D OCT images, the normal bladder tissue and cancerous tissue were distinguished clearly. Therefore, our method is potential in the detection and diagnosis of bladder cancers in the future.

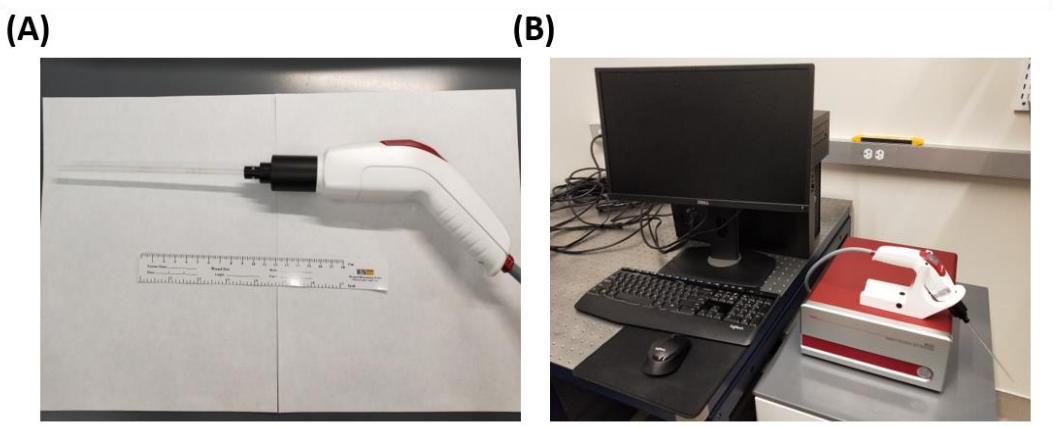


Figure 4.3. (A) Image of the new-generation handheld OCT imaging probe. (B) Image of the whole laparoscopic OCT imaging system including the OCT machine, hand-held probe, and computer.

For future clinical applications, the size of our device should be miniaturized. In collaboration with Thorlabs, a compact handheld laparoscopic OCT probe was developed. This portable device is shown in Figure 4.3. It makes the inserting process of the needle easier

in operations. By applying a swept-source laser with axial scanning rate up to 200 kHz, our machine can provide the imaging guidance in nearly real time. Our current work has demonstrated the visualizing capability of our laparoscopic OCT in bladder tumor detection and its potential in practical surgeries.

Reference

1. Society, A.C. *Key Statistics for Bladder Cancer 2021*. 2021; Available from: https://www.cancer.org/cancer/bladder-cancer/about/key-statistics.html.
2. Bostrom, P.J., et al., *Staging and Staging Errors in Bladder Cancer.* European Urology Supplements, 2010. **9**(1): p. 2-9.
3. Wang, H.W. and Y. Chen, *Clinical applications of optical coherence tomography in urology.* Intravital, 2014. **3**(1): p. e28770.
4. Herr, H.W. and S.M. Donat, *Quality control in transurethral resection of bladder tumours.* BJU Int, 2008. **102**(9 Pt B): p. 1242-6.
5. Huang, D., et al., *Optical coherence tomography.* Science, 1991. **254**(5035): p. 1178-81.
6. Zagaynova, E., et al., *Endoscopic OCT with forward-looking probe: clinical studies in urology and gastroenterology.* J Biophotonics, 2008. **1**(2): p. 114-28.
7. Park, H.C., et al., *Forward imaging OCT endoscopic catheter based on MEMS lens scanning.* Opt Lett, 2012. **37**(13): p. 2673-5.
8. Kang, W., et al., *Motion artifacts associated with in vivo endoscopic OCT images of the esophagus.* Opt Express, 2011. **19**(21): p. 20722-35.
9. Gora, M.J., et al., *Endoscopic optical coherence tomography: technologies and clinical applications [Invited].* Biomed Opt Express, 2017. **8**(5): p. 2405-2444.
10. Xie, T., M. Zeidel, and Y. Pan, *Detection of tumorigenesis in urinary bladder with optical coherence tomography: optical characterization of morphological changes.* Opt Express, 2002. **10**(24): p. 1431-43.
11. Xie, T., et al., *Endoscopic optical coherence tomography with a modified microelectromechanical systems mirror for detection of bladder cancers.* Appl Opt, 2003. **42**(31): p. 6422-6.
12. Tang, Q., et al., *Depth-resolved imaging of colon tumor using optical coherence tomography and fluorescence laminar optical tomography.* Biomed Opt Express, 2016. **7**(12): p. 5218-5232.

CHAPTER 5: AUTOMATIC RENAL CARCINOMA BIOPSY GUIDANCE USING FORWARD-VIEWING ENDOSCOPIC OPTICAL COHERENCE TOMOGRAPHY AND DEEP LEARNING

This chapter is reproduced from "Chen Wang, Haoyang Cui, Qinghao Zhang, Paul Calle, Yuyang Yan, Feng Yan, Kar-Ming Fung, Sanjay G. Patel, Zhongxin Yu, Sean Duguay, William Vanlandingham, Chongle Pan, Qinggong Tang, *Automatic renal carcinoma biopsy guidance using forward-viewing endoscopic optical coherence tomography and deep learning*, **Communication Engineering**, 2024"

5.1 Introduction

Renal cancer is a disease characterized by the uncontrolled growth of kidney cells, resulting in the formation of tumor tissue. The lifetime risk of renal cancer was estimated to be ~2.02% in males and ~1.03% in females [1]. American Cancer Society projected approximately 81,800 new cases of renal cancer in the United States in 2023, including 52,360 males and 29,440 females, and 14,890 (9,920 males and 4,970 females) would die from it [1]. Renal cell carcinoma (RCC) is the most prevalent type of renal cancer, constituting approximately 95% of all renal cancer cases, and it has shown a consistent rise over the past few decades [2, 3]. Accurate identification and diagnosis of renal tumor malignancy is critical for devising effective treatment plans. At present, percutaneous renal biopsy (PRB) is the most commonly used procedure to extract renal tissue for pathological analysis [4]. Since being proposed in 1951, PRB has been widely applied in the diagnosis and prognosis of kidney diseases [5]. It has been considered as an effective method for the assessment of renal tumors from patients with RCC. During the PRB procedure, the patient is placed in a prone position. An exploratory needle is then inserted into the kidney to establish a path for the subsequent insertion of the biopsy needle [6]. The biopsy needle is employed to

extract a tissue sample from the tumor. The obtained tissue sample is processed, analyzed, and classified as benign, malignant, or nondiagnostic for patients with RCC [7].

Despite being widely used for kidney disease diagnosis and renal tumor tissue evaluation, PRB still faces challenges in the tissue localization. It has been reported that even with the assistance of a radiologist, approximately 92% of PRBs required two or more attempts in order to accurately extract the targeted tissue [8]. More insertion attempts raise the likelihood of complications, including hematoma and infection [9, 10]. Therefore, accurate guidance of needle insertion is of great importance for effective and safe renal carcinoma biopsy. Conventional imaging modalities have been employed for PRB needle guidance. Ultrasound has been harnessed to assist in needle guidance, leading to a significant reduction in the incidence of severe PRB complications over the past few decades [11, 12]. Computed tomography (CT) has also been utilized in the diagnosis of metastatic lesion during percutaneous biopsy [13]. Three-dimensional cone-beam CT guidance has been documented as beneficial for the secure and effective determination of needle trajectory, as well as for performing biopsies on small renal masses [14]. Moreover, the guidance of PRB through magnetic resonance imaging (MRI) exhibits promising outcomes in precisely targeting the renal tumor and accurately locating the PRB needle tip, particularly in obese patients [15]. Although these macroscopic imaging methods, including ultrasound, CT and MRI, are effective in needle trajectory determination, they cannot accurately identify tissue types ahead of the needle tip due to their limited spatial resolutions [16]. As a result, it was reported that ~10% - 20% of the biopsy results were non-diagnostic due to the limitation of imaging guidance of tumor location [16]. Hence, there is an unmet need for an imaging method capable of accurately identifying tissue types ahead of the PRB needle.

Optical coherence tomography (OCT) is a well-established imaging technique with micron-level spatial resolution [17]. Therefore, OCT has the potential for providing enhanced fine-scale visualization during PRB needle guidance. Moreover, OCT has already demonstrated feasibility in diagnosing renal masses, with the ability to differentiate renal tumor tissue from normal renal tissue based on distinct attenuation coefficients in OCT images [18-20]. OCT can also provide real-time imaging [21]. Therefore, we hypothesized that OCT can be used for real-time guidance of the PRB needle by detecting and identifying the renal tumor tissue. To enable tissue imaging ahead of the PRB needle, we have developed a forward-view OCT probe that incorporates gradient-index (GRIN) lenses functioning as a forward-viewing endoscope. Our previous work successfully demonstrated the application of endoscopic OCT in percutaneous nephrostomy (PCN) guidance, accurately identifying various kidney tissues within pig models [22, 23]. Hence, the OCT probe system holds the promise in PRB needle guidance and renal tumor recognition.

Because visually identifying tissue from OCT images incurs a significant learning curve for radiologists and PRB operators, we have developed automatic imaging processing techniques, including tissue classification using attenuation coefficient and deep learning methods. Mapping optical attenuation coefficients has already been employed in distinguishing between cancerous and non-cancerous tissues [24]. Given that cancerous tissue tends to exhibit higher density in the imaging results, its attenuation coefficient typically surpasses that of normal tissues. Therefore, attenuation coefficient can be a promising feature used for tumor tissue recognition during PRB needle guidance. Furthermore, deep learning has been demonstrated to be an effective approach in tissue classification and recognition for surgical guidance [23, 25, 26]. In this study, we have applied convolutional neural networks

(CNNs) for automatic identification of carcinoma from normal tissue and oncocytoma. The performance of these methods was benchmarked for renal tissues classification and tumor diagnosis.

5.2 Materials and Methods

5.2.1 Experimental setup

We constructed our forward-viewing optical coherence tomography (OCT) probe based on a swept-source OCT (SS-OCT) system [23, 26, 27]. The light source utilized was a 1,300 nm swept-source laser with a bandwidth of 100 nm. The scanning rate of our system reached 200 kHz. As illustrated in Figure 5.1, the light was initially divided using a fiber coupler (FC) into two distinct paths: one directed with 97% of power towards a circulator for the OCT imaging, and the other with 3% power directed to a Mach-Zehnder interferometer (MZI) to initiate the sampling process. Upon traversing the circulator, the laser underwent another division through another FC into the reference arm and sample arm. Subsequently, the light reflected by the mirror in the reference arm and the light backscattered by the tissue in the sample arm converged, generating an interference signal. This signal was then denoised using a balanced detector (BD), acquired by the data acquisition (DAQ) board, and finally processed and visualized on a computer. Polarization controllers (PC) were employed in both arms to reduce background noise.

In PRB procedures, 14 gauge (G) or 16 G biopsy needles are commonly used [28]. To assess feasibility, we employed GRIN lenses with a diameter of 1.3 mm and length of 138.0 mm, allowing them to be accommodated within the 14 G or 16 G biopsy needles. These GRIN lenses possess a refractive index distribution that varies perpendicular to the optical axis, facilitating effective image transmission from one end to the other [29]. The GRIN lens

in our system was featured with a pitch length of four integers, thus enabled the direct relay of laser profiles or images from one end to the other [22, 23].

In the sample arm, the proximal end of the GRIN lens was precisely positioned at the focal point of the scanner lens, while the distal end was in contact with the sample. Consequently, the laser profile was directly transmitted from the proximal surface to the sample. Simultaneously, spatial information from the distal surface (tissue sample) of the GRIN lens, transmitted to the proximal surface, was further captured by the OCT scanner. Additionally, another same GRIN lens was situated in the reference arm to counteract light dispersion. Our system features an axial resolution of approximately 11 μm and a lateral resolution of around 20 μm. It provides a sensitivity of 92 dB and covers a field of view (FOV) of approximately 1.25 mm in diameter, rendering it well-suited for our intended application.

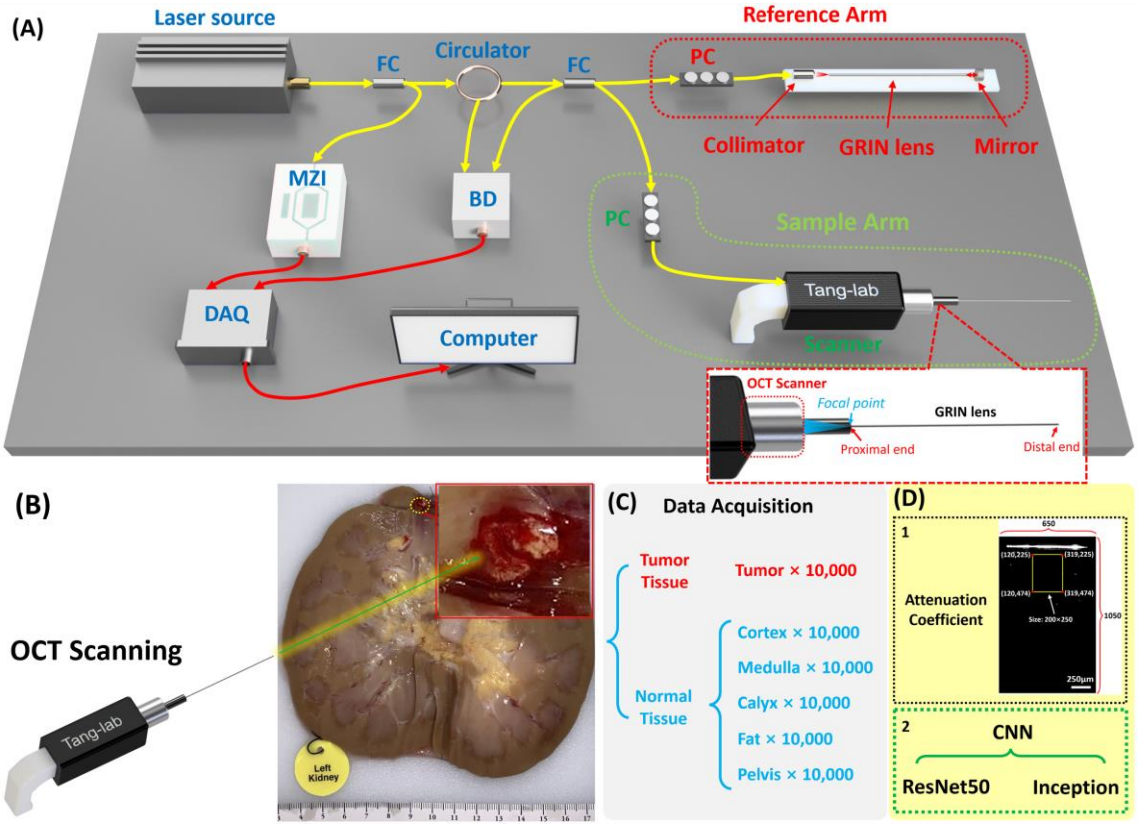


Figure 5.1. (A) Schematic of the forward-viewing OCT probe. FC: fiber coupler; PC: polarization controller; MZI: Mach–Zehnder interferometer; BD: balanced detector; DAQ: Data acquisition. (B) Picture of the human kidney with renal carcinoma. (C) Data acquisition of renal tumor tissue and normal tissue. (D) Attenuation coefficient and two CNN architectures used for tissue classification.

5.2.2 Data acquisition

The study was approved by the University of Oklahoma Institutional Review Board (Study number: 12462). Written informed consent was obtained by Lifeshare of Oklahoma. Patient demographics were collected upon obtaining consent. For our dataset, we utilized nine human kidney samples with renal carcinoma. OCT images of carcinoma tissue and different normal tissues were imaged by our OCT probe. Furthermore, to compare benign renal tumor against malignant renal carcinoma in this study, we conducted OCT imaging on one kidney sample with oncocytoma, which is a well-defined benign solid tumor arising from proximal renal tubule [30]. To keep the organ as fresh as possible, all the kidney samples were preserved by hypothermic machine perfusion (HMP), and we started the experiment as soon as the sample was acquired.

Figure 5.1(B) illustrates one of the kidney samples with carcinoma, clearly showcasing the tumor tissue alongside other normal renal tissues including cortex, medulla, calyx, fat, and pelvis. To emulate a practical biopsy procedure during the experiment, we carefully inserted the OCT probe into different renal tissue types, applying controlled force to compress the tissue. OCT image acquisition was successfully performed in the final.

For quantitative identification of tumor tissue, we utilized both conventional methods, involving the calculation of attenuation coefficients, and deep learning techniques for

automated recognition of renal tissue types. To train the deep learning model, we obtained a dataset of 10,000 OCT images from the tumor and five different normal renal tissues in each kidney sample, respectively (Figure 5.1(C)). The size of all output images was set as lateral X depth: 650×1050 (X×Z: 1.30 mm×2.10 mm), so each result covered the same FOV for attenuation coefficient method and CNN process (Figure 5.1(D)). For each method (attenuation coefficient and CNN), data from five kidney samples with carcinoma were used to train and benchmark the model by nested cross-validation, and the other four kidney samples with carcinoma were used in a hold-out set to independently benchmark the prediction performance. 10,000 oncocytoma OCT images were further tested to compare the biological features of oncocytoma with carcinoma.

5.2.3 Attenuation coefficient method for tissue recognition

Our OCT findings provided in-depth structural insights into renal tissues. The attenuation coefficient serves as a crucial parameter, signifying the decrease in signal intensity observed in OCT images as depth increases [31]. In structural OCT images, the imaging intensity reflects the backscattering level of the tissue sample, so different attenuation coefficients demonstrate different tissue features and can be used for tissue classification. The use of attenuation coefficient for cancer detection has been extensively documented [32-34]. Compared to normal tissues, cancerous tissue is typically denser thus its imaging signal value attenuated faster in vertical direction and represented by higher attenuation coefficient. The attenuation coefficient of each A-scan can be extracted and used to generate a spatial attenuation coefficient distribution, which helps visually distinguish tumor and normal tissues [24].

Here we used Beer-Lambert law to describe the structural OCT intensity signals in depth as follows:

$$I_{(z)} = I_0 e^{-2\mu z} \quad (1)$$

where I_0 represents the initial incident light intensity, μ is the attenuation coefficient value which represents the decay rate, and z is the depth [35]. Thus, μ can be described as:

$$\mu = -\frac{1}{2}.\frac{d\log I_{(Z/0)}}{d_z} \quad (2)$$

$I_{(Z/0)}$ is the ratio of recorded OCT intensity in depth to the initial incident light intensity. The intrinsic optical attenuation coefficient was calculated according to the slope of OCT intensity in a 250-pixel depth window and then averaged in a 200-pixel width region as shown in Figure 5.1(D).

5.2.4 CNN methods for tissue identification

Two CNN architectures, ResNet50 and InceptionV3, were evaluated. ResNet50 has ~ 25.6 million parameters and InceptionV3 has ~ 22 million parameters. Five kidney samples with carcinoma were used to develop the CNN architectures. We used a nested 4-fold cross-validation and 5-fold cross-testing procedure [23], in which each iteration rotated three kidneys to the training dataset, one kidney to the validation dataset, and one kidney to the test dataset. This division enabled us to obtain a more accurate estimation of the generalization error, as every kidney was included in the test dataset once. Early stopping was implemented with a patience of 20. The optimizer used for the training was stochastic gradient descent with Nesterov accelerated gradient, utilizing a momentum value of 0.9 and a learning rate of 0.01. Additionally, the learning decay rate was set to be 0.1. The average epoch count from the cross-validation folds was subsequently employed to train the model intended for use on the

test dataset. The test performances of the obtained models were benchmarked using the 5-fold cross-testing.

A final model was trained using all five kidneys included in the nested cross-validation with the identified optimum hyperparameters. The performance of the final model was benchmarked using a separate hold-out set which contains four additional kidney samples with carcinoma. Furthermore, one oncocytoma sample was also used to test the capability of the final model for distinguishing oncocytoma from carcinoma. The classification performance was measured using four metrics, including accuracy, precision, recall, and F_1 score as defined below:

$$Accuracy = (TP + TN)/(TP + FP + TN + FN) \quad (3)$$

$$Precision = TP/(TP + FP) \quad (4)$$

$$Recall = TP/(TP + FN) \quad (5)$$

$$F_1 \text{ score} = 2 \cdot \frac{\text{Precision} \cdot Recall}{Precision + Recall} = \frac{TP}{TP + \frac{1}{2}(FP + FN)} \quad (6)$$

where TP is true positive, FP is false positive, TN is true negative, and FN is false negative.

To interpret the prediction of deep learning models [36, 37], GradCAM [38] was used to generate pixel saliency heatmaps. One of the five kidneys' images were used to generate the heatmaps for both CNN models: InceptionV3 and ResNet50. For example, all images of the cortex from the chosen kidney were utilized. We calculated the sum of the corresponding pixel points of the images, then the sum was divided by the number of total images used, resulting in an average heatmap image of cortex. Same practice applied to the rest of the tissue types.

5.3 Results

5.3.1 Imaging results of renal tumor and normal tissues

Figure 5.2 illustrates the OCT imaging results obtained through our OCT probe, demonstrating malignant renal carcinoma, benign renal oncocytoma, and five distinct normal renal tissue types. Two dimensional (2D) cross-sectional OCT images of these tissues demonstrate different structural features. The corresponding histology results are also shown in Figure 5.2. The renal pelvis is the cavity for collecting urine, hence there is no histology result for it.

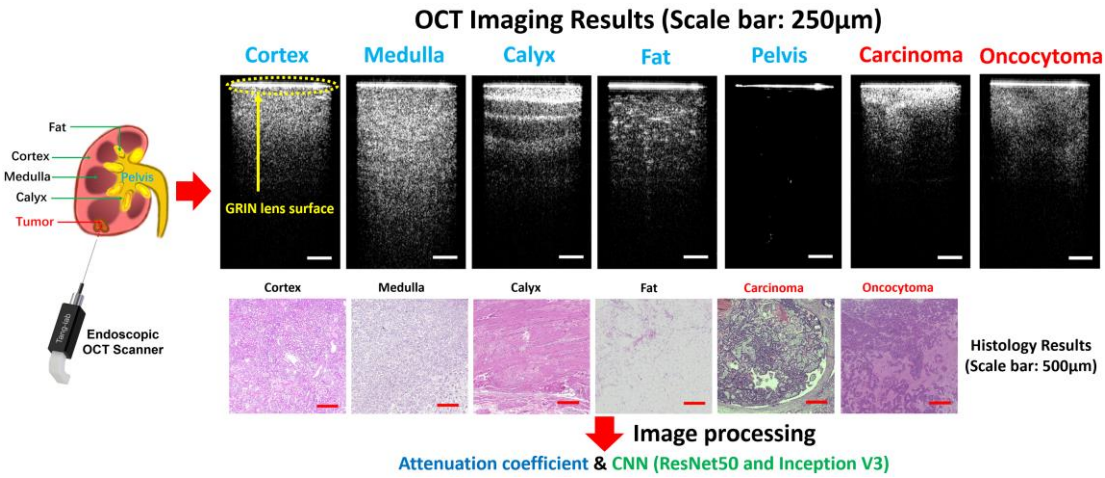


Figure 5.2. OCT imaging results of normal renal tissues, carcinoma, and oncocytoma (Scale bar: 250μm). Corresponding histology results (Scale bar: 500μm).

In the OCT results, there is a bright line on the top (pointed by a yellow arrow) which corresponds to the surface of the GRIN lens. The intensity signal distributions of both cortex and medulla are relatively homogeneous, although the imaging depth of medulla surpasses that of cortex. Calyx exhibits alternately shaded stripes, a result of its transitional epithelium and fibrous tissue composition. Fat presents the darkest pattern, accompanied by bright dots

characteristic of adipocytes. We injected synthetic urine into the pelvis and tied the ureter during the experiment and inserted the probe into the pelvis for imaging. As an empty space for collecting urine, the pelvis has no tissue signal under the GRIN lens surface. The renal carcinoma, being the pivotal tissue type for identification during PRB, showcases the shallowest imaging depth while presenting the highest brightness among all OCT results. Its structural features exhibit irregularities. Compared to carcinoma, the overall imaging density of oncocytoma is relatively lower, while there are still clear irregular structural features. From the corresponding histology results, these tissues showed different tissue structures and distributions and correlated well with their OCT results. Different normal renal tissues and the two types of tumor tissues can be categorized based on their distinctive OCT imaging features. As a result, our endoscopic OCT probe has the potential to assist in the recognition of tissue types ahead of the needle during PRB procedures.

5.3.2 Classification using attenuation coefficient method

We first assessed the performance of attenuation coefficient method in distinguishing renal tumor and normal tissues. The attenuation coefficient did not work for the pelvis, because the values of most pixels were zero and the logarithm could not be calculated. Therefore, only tumor tissue and other four normal tissue types were tested using the attenuation coefficient method.

This study involved the processing of a total of 50,000 OCT images from each of the five nested cross-validation samples (after removing pelvis category) with each tissue type contributing 10,000 images. The performance of the results was benchmarked first using the cross-testing procedure and then using a hold-out set containing five additional kidney samples (4 renal carcinoma and 1 benign renal oncocytoma). Among the five nested cross-

validation samples, four of them were randomly selected as the training dataset, while the remaining kidney served as the test dataset. To calculate the attenuation coefficient, the region of interest (ROI) window was selected to be 200×250 pixels for each OCT image. The ROIs were chosen from the central sections of the images and the coordinate of the area is selected in the top middle position which includes most tissue information (Figure 5.1(D)). The mean value of all the attenuation coefficient results from every vertical direction was used as the attenuation coefficient level for each OCT image.

Figure 5.3 displays the attenuation coefficient distributions for the five tissue types (four normal renal tissues and renal carcinoma). Firstly, considering cortex, medulla, calyx, and fat are all normal tissues, these four types were initially consolidated into a single classification. In this testing, all the tumor tissue images from 5 samples (50,000 images in total) were utilized. To make the number of normal tissue images equal to the number of tumor images, we randomly selected 2,500 images from each normal tissue in every sample (2,500 images × 4 tissue types × 5 samples = 50,000 images). In general, the tumor tissue and the normal tissues displayed distinctive distribution characteristics, thereby enabling a clear differentiation between tumor and normal tissues as shown in Figure 5.3(A). Normal distributions were fitted to the attenuation coefficient distributions as shown in Figure 5.3(A). To quantify the separation of the tumor from normal tissues, the cross point (blue dot) of two curves was chosen as threshold and used to evaluate the accuracy. The receiver operating characteristic (ROC) curves were shown between tumor and normal renal tissues in Figure 5.3(B). The area under the ROC curves (AUC) (Tumor_Normal) was 99.95%, which validated the tumor recognition accuracy based on attenuation coefficients.

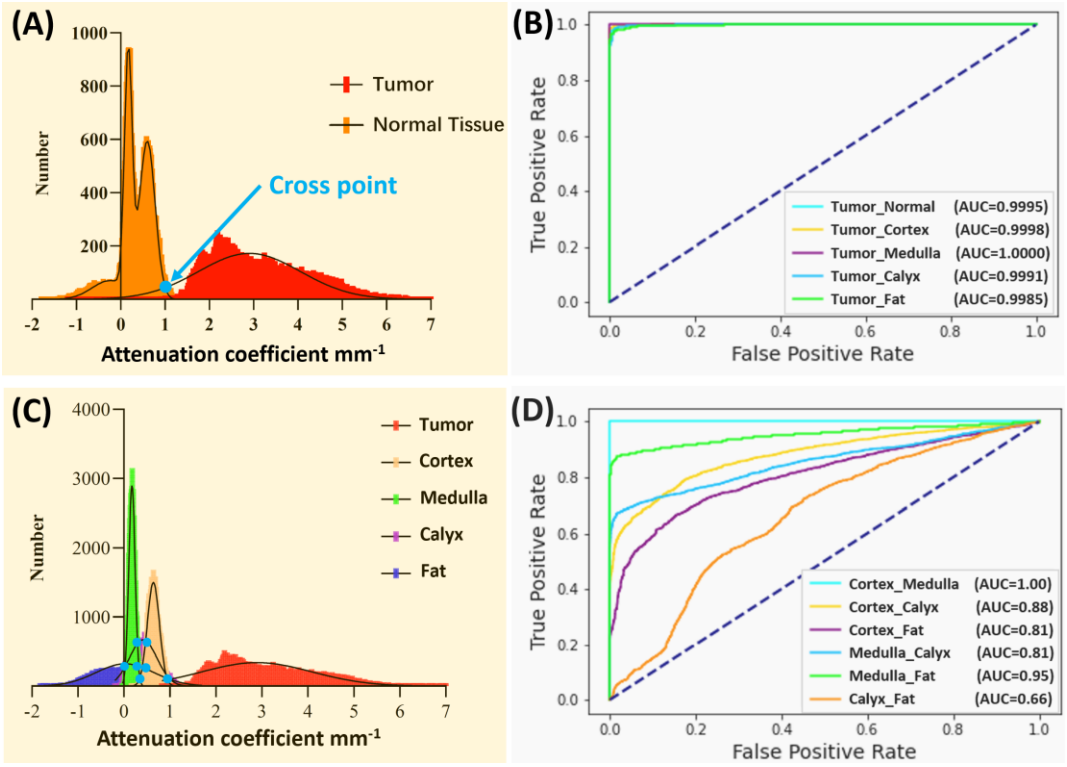


Figure 5.3. (A) Attenuation coefficient distribution results between normal tissue and tumor tissue. (B) ROC curves of the classification with attenuation coefficient between tumor tissue and other tissues. (C) Attenuation coefficient distribution results of all five tissue types. (D) ROC curves of the classification with attenuation coefficient among different normal tissues.

Table 5.1. (A) Confusion matrix of the averaged five nested cross-validation kidney samples. (B) Prediction results: Accuracy; Precision; Recall; and F_1 Score of different tissue types.

(A)

Tissue		Prediction of five samples using attenuation coefficient				
		Cortex	Medulla	Calyx	Fat	**Carcinoma**
True	Cortex	9145	14	723	0	118
	Medulla	3	9154	756	87	0
	Calyx	4326	2820	1947	347	560
	Fat	1895	1872	1083	4942	208
	Carcinoma	**17**	**0**	**0**	**0**	**9983**

(B)

Tissue	**Accuracy**	**Precision**	**Recall**	**F_1 Score**
Cortex	85.81%	59.44%	91.45%	72.05%
Medulla	88.90%	66.05%	91.54%	76.73%
Calyx	78.77%	43.18%	19.47%	26.84%
Fat	89.02%	91.93%	49.42%	64.28%
Carcinoma	**98.19%**	**91.85%**	**99.83%**	**95.67%**

We then evaluated classification of the renal tissue types by using 10,000 OCT images from 5 samples (50,000 images in total). The attenuation coefficient distributions were shown in Figure 5.3(C). To quantify the separation of the tumor from each normal tissue (i.e., tumor vs cortex, tumor vs medulla, tumor vs calyx, and tumor vs fat), the cross point (blue dot) of two curves was again chosen as threshold and used to evaluate the accuracy. The AUCs for differentiating the tumor tissue from each normal renal tissue were all over 99.85% (Figure 5.3(B)), which further proved the tumor recognition capacity of using attenuation coefficients.

However, different normal renal tissues were difficult to classify using attenuation coefficient method. The AUCs were 100% for cortex vs. medulla, 88% for cortex vs. calyx, 81% for cortex vs. fat, 86% for medulla vs. calyx, 95% for medulla vs. fat, and only 66% for calyx vs. fat (Figure 5.3(D)). The confusion matrix for multi-class classification was further shown in Table 5.1(A), and the corresponding prediction results were shown in Table 5.1(B). Tumor demonstrated high recognition performance, with 98.19% accuracy, 91.85% precision, 99.83% recall, and 95.67% F_1 score. However, the normal tissue recognition was challenging, especially for calyx (78.77% accuracy, 43.18% precision, 19.47% recall, and 26.84% F_1

score). The needle's precise placement can enhance surgical efficiency and overall effectiveness. The present results primarily demonstrated the promise of tumor recognition, but the performance in identifying other normal kidney tissues still needed improvement. In particular, the attenuation coefficient methods misclassified over 500 images of calyx as tumor tissue, causing potential sampling errors.

Table 5.2. (A) Confusion matrix of the test results from the hold-out set with additional four kidney samples. (B) Prediction results: Accuracy; Precision; Recall; and F_1 Score of different tissue types.

(A)

Tissue		Testing of the hold-out set using attenuation coefficient				
		Cortex	Medulla	Calyx	Fat	**Carcinoma**
True	Cortex	7418	306	2170	0	106
	Medulla	6	9225	756	13	0
	Calyx	2452	2612	1228	604	3104
	Fat	1007	1418	775	6679	121
	Carcinoma	**7**	**4**	**2**	**13**	**9974**

(B)

Tissue	**Accuracy**	**Precision**	**Recall**	**F_1 Score**
Cortex	87.89%	68.12%	74.18%	71.02%
Medulla	89.77%	68.01%	92.25%	78.29%
Calyx	75.05%	24.90%	12.28%	16.45%
Fat	92.10%	91.38%	66.79%	77.17%
Carcinoma	**93.29%**	**74.96%**	**99.74%**	**85.60%**

Further benchmarking with the hold-out set containing four kidney samples with carcinoma was conducted and the corresponding results were demonstrated in Table 5.2. Similar to previous cross-testing results, carcinoma provided high accuracy and recall of 93.29% and 99.74% respectively. However, the precision and F_1 score just reached 74.96% and 85.60%. This decrease can be attributed to the fact that over 3,000 calyx images were incorrectly predicted as carcinoma. Moreover, the prediction of normal tissue remained

unsatisfactory. Therefore, attenuation coefficient, as a conventional tool, is not an effective method for renal carcinoma recognition during PRB.

Table 5.3. Prediction results of oncocytoma tissues using attenuation coefficient.

		Testing of oncocytoma samples predictions by attenuation coefficient				
		Cortex	Medulla	Calyx	Fat	**Carcinoma**
True	Oncocytoma	3087	185	484	197	6047

In addition, we tested 10,000 oncocytoma OCT images using attenuation coefficient, with the results shown in Table 5.3. It demonstrated that over 60% of the oncocytoma imaging results were predicted as carcinoma and over 30% was misclassified as cortex. This further indicated that attenuation coefficient was not effective for distinguishing between oncocytoma and carcinoma.

5.3.3 Tissue recognition results using CNN

The recognition of kidney tissues using the ResNet50 and InceptionV3 CNNs was evaluated using cross-testing on the 5 nested cross-validation kidney samples with carcinoma. The ROC curves and pixel saliency heatmaps were shown in Figure 5.4. The confusion matrixes and performance metrics were shown in Table 5.4.

The ResNet50 CNN yielded a tumor recognition accuracy of 99.51%. For normal tissue recognition, the prediction accuracies were 98.96% for cortex, 99.78% for medulla, 99.09% for calyx, 99.27% for fat, and 99.95% for pelvis. Most of the tumor sample misclassifications (123 of 134) were erroneously categorized as cortex. None of cortex, medulla and fat images were predicted to be tumor, and 1.48% of the calyx images was falsely predicted as tumor. The InceptionV3 CNN reached 99.48% tumor prediction accuracy. The accuracies for normal tissue classification were 99.58% for cortex, 99.88% for medulla, 99.27% for calyx, 99.65% for fat, and 99.95% for pelvis. Calyx recognition has the lowest

accuracy, and 0.02% of tumor samples were erroneously categorized as calyx, which is different from the ResNet50 model. Both ResNet50 and InceptionV3 achieved >99.9% ROC AUC for all tissues.

InceptionV3 outperformed ResNet50 in overall precision for the classification of the normal renal tissues, with a slightly lower precision in the tumor classification. All normal tissues and tumor results using InceptionV3 have higher recalls than ResNet50 except for calyx. Additionally, calyx provide the lowest F1 scores in both models, but the overall performances for carcinoma are still outstanding (both > 98%).

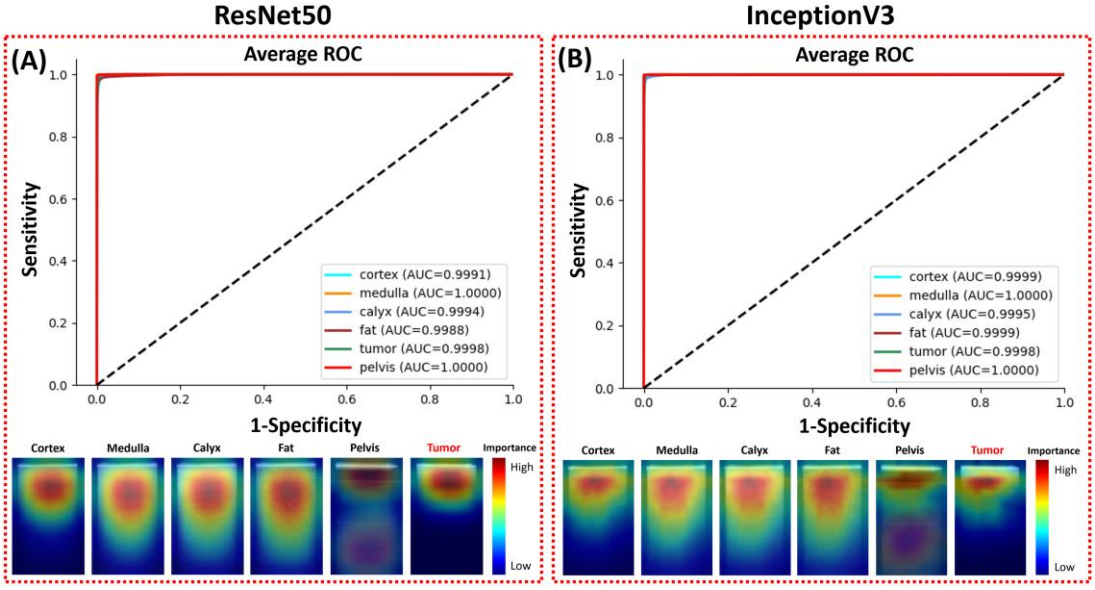


Figure 5.4. ROC curves and pixel importance heatmaps of (A) ResNet50 and (B) InceptionV3.

Table 5.4. Confusion matrices and performance metrics of the cross-validation and cross-testing using five kidney samples by (A) ResNet50 and (B) InceptionV3.

(A)

Tissue		Performance Benchmarking of ResNet50 by Nested Cross-Validation					
		Cortex	Medulla	Calyx	Fat	Pelvis	**Carcinoma**
True	Cortex	9875	75	23	27	0	0
	Medulla	43	9956	1	0	0	0
	Calyx	136	15	9553	148	0	148
	Fat	199	0	56	9745	0	0
	Pelvis	0	0	21	0	9969	10
	Carcinoma	**123**	**0**	**0**	**9**	**2**	**9866**

Tissue	Accuracy	Precision	Recall	F_1 Score
Cortex	98.96%	95.16%	98.75%	96.92%
Medulla	99.78%	99.10%	99.56%	99.33%
Calyx	99.09%	98.95%	95.53%	97.21%
Fat	99.27%	98.15%	97.45%	97.80%
Pelvis	99.95%	99.98%	99.69%	99.83%
Carcinoma	**99.51%**	**98.43%**	**98.66%**	**98.54%**

(B)

Tissue		Performance Benchmarking of InceptionV3 by Nested Cross-Validation					
		Cortex	Medulla	Calyx	Fat	Pelvis	**Carcinoma**
True	Cortex	9964	1	1	34	0	0
	Medulla	15	9978	7	0	0	0
	Calyx	69	38	9639	0	0	254
	Fat	121	0	30	9822	0	27
	Pelvis	0	11	15	1	9973	0
	Carcinoma	**10**	**0**	**22**	**0**	**0**	**9968**

Tissue	Accuracy	Precision	Recall	F_1 Score
Cortex	99.58%	97.89%	99.64%	98.76%
Medulla	99.88%	99.50%	99.78%	99.64%
Calyx	99.27%	99.22%	94.39%	96.74%
Fat	99.65%	99.64%	98.22%	98.92%
Pelvis	99.95%	100.00%	99.73%	99.86%
Carcinoma	**99.48%**	**97.26%**	**99.68%**	**98.46%**

The average pixel importance heatmaps of all the images from the two models were shown in Figure 5.4. These heatmaps highlighted the regions important for the models to distinguish different tissue types. The narrow area below the GRIN lens surface line was important for the tumor recognition. The important area for the pelvis recognition included a narrow part on the top and a wide area on the bottom. Heatmaps of the other four tissues all highlighted the top half of the images with strong signals. The important areas used by ResNet50 were sparser and wider than those used by InceptionV3. This could be attributed to the fact that ResNet50's feature map was more focused and localized within each residual block than InceptionV3, so it could capture fine-grained details in the images in pixel level.

Table 5.5. (A) Confusion matrix and (B) performance metrics of the final production model on the hold-out set containing four additional kidney samples with carcinoma.

(A)

Tissue		Testing of the hold-out set using the final model					
		Cortex	Medulla	Calyx	Fat	Pelvis	**Carcinoma**
True	Cortex	9997	0	0	3	0	0
	Medulla	0	10000	0	0	0	0
	Calyx	161	1	9697	90	0	51
	Fat	0	0	32	9968	0	0
	Pelvis	0	0	0	0	10000	0
	Carcinoma	**0**	**0**	**0**	**35**	**196**	**9796**

(B)

Tissue	Accuracy	Precision	Recall	F_1 Score
Cortex	99.73%	98.42%	99.97%	99.19%
Medulla	100.00%	99.99%	100.00%	100.00%
Calyx	99.44%	99.67%	96.97%	98.30%
Fat	99.73%	98.73%	99.68%	99.20%
Pelvis	99.67%	98.33%	100.00%	99.16%
Carcinoma	**99.53%**	**99.48%**	**97.96%**	**98.71%**

For the performance of nested cross-validation between two models, the average accuracies across five kidney samples with InceptionV3 and ResNet50 for predicting all tissue types were 98.26% and 97.32%, respectively. The overall prediction accuracies of two models were close, with InceptionV3 being higher marginally. In addition, the inference time was 18 ms per image for ResNet50, and 11 ms per image for InceptionV3. Thus, to make the prediction more efficient, we utilized InceptionV3 as the final model for the hold-out testing. The experiment was performed using a hold-out set with four additional samples with carcinoma, with the results displayed in Table 5.5. The model made erroneous carcinoma predictions at a rate of 0.35% to fat and 1.96% to pelvis, respectively. For F_1 scores of all tissue types in the prediction with the hold-out set, they outperformed the nested cross-validation results demonstrated in Figure 5.4(B).

To test the feasibility of distinguishing oncocytoma from carcinoma, we further tested the 10,000 oncocytoma OCT images using the final model. The predicted results were shown in Table 5.6. On average, the model presented a misclassification rate of 2.24%. Most of the

oncocytoma images were misclassified as cortex and fat. The model could effectively distinguish oncocytoma from carcinoma.

Table 5.6. Predictions of oncocytoma by the final CNN model.

		Testing of oncocytoma samples predictions by the final model					
		Cortex	Medulla	Calyx	Fat	Pelvis	**Carcinoma**
True	Oncocytoma	9215	0	0	553	8	224

5.4 Discussion

Although PRB is the most commonly performed procedure for evaluating renal tissue, it remains a challenge even for experienced urologists to precisely extract carcinoma tumor tissue with minimal sampling error. In this study, we tested the feasibility of using a forward-viewing OCT probe for RCC detection on *ex-vivo* human kidneys for the first time. In total, ten human kidney samples, nine of which have malignant renal carcinoma and one has benign renal oncocytoma, were utilized in our experiments to validate our approach. The two types of renal tumor tissues and different normal renal tissues, including cortex, medulla, calyx, renal fat, and pelvis were all included to evaluate the recognition accuracy of our system. We applied the probe on these tissues and obtained their OCT images. The OCT images of different renal tissues could be classified based on their distinct imaging characteristics. Moreover, the OCT probe's dimensions were designed to fit inside the PRB needle's hollow bore, enabling seamless integration without introducing any additional invasiveness. Under clinical settings, our OCT probe can image the tissue in front of the PRB needle in real time, effectively identifying the tissue type at the needle's tip.

To automate the process of tissue identification, we utilized and compared two classification methods in our study: attenuation coefficient and deep learning. The attenuation coefficient method has been widely used in OCT-based cancer diagnosis, including breast

cancer [33], prostate cancer [39], bladder cancer [34], colon cancer [40], etc. Renal mass diagnosis using OCT was also reported, and it exhibited highly promising outcomes in predicting RCC [41]. For the classification task, we utilized the data from five carcinoma samples to train and benchmark the models using nested cross-validation and employed a hold-out set for further testing. We first evaluated the attenuation coefficient for identification of tumor and normal tissues. Discrimination of tumor from normal tissues was promising with the accuracy of 98.19% in cross-testing, and 93.29% in the hold-out testing. Although the results were generally favorable, misrecognition still happened. Additionally, relying solely on the attenuation coefficient proved challenging when it came to distinguishing between different types of normal renal tissues. The classification accuracy results were not satisfactory, particularly for the calyx prediction accuracy which fell below 80% in both testing procedures. In addition, we tested the model performance for distinguishing between benign and malignant renal tumors. Over 60% of the oncocytoma data were erroneously predicted as carcinoma, further proving that attenuation coefficient was not an efficient tool in renal tissue recognition.

Accurate identification of normal renal tissues prior to biopsy needle insertion played a pivotal role in precisely tracking the needle's location and avoiding sampling error. To address this critical need, we further evaluated the deep learning methods in our study. Specifically, we employed two well-established CNN architectures, ResNet50 and InceptionV3, both of which demonstrated robust performance in tumor recognition and the classification of various normal renal tissues.

In cross-testing using the five carcinoma samples, InceptionV3 achieved tumor recognition rates of 99.48%, which surpassed the performance of the attenuation coefficient

method. Equally impressive was the accuracy in normal tissue recognition, with rates consistently exceeding 99% across all tissue types. ROC curves for both models showed AUC values consistently exceeding 99.9%. This observation underscores the effectiveness of the CNN-based approach for the classification of different normal renal tissues compared to the attenuation coefficient method. The outstanding performance of these models can be attributed, in part, to their substantial parameter sets. ResNet50 features 25.6 million and InceptionV3 features 22 million parameters, respectively, in contrast to the single parameter employed by the attenuation coefficient method. All computations were conducted on a server equipped with two NVIDIA RTX A6000 graphics cards. The cross-validation required an average of 1,435 minutes per fold. The average inference time over 320 images was 18 milliseconds per image for ResNet50 and 11 milliseconds per image for InceptionV3. To facilitate clinical translation, future experiments could investigate the use of smaller architectures with faster processing times, all while maintaining performance integrity. Theoretically, CNN architectures with fewer trainable parameters or less depth will reduce the processing time and improve the efficiency. For instance, we will consider using pretrained EfficientNet or VGG16 architectures, to reduce the training and inference time. The prediction results from the hold-out set containing four additional carcinoma samples revealed significant generalization capability of the final CNN model as it showed higher F_1 scores for the majority of the tissue types compared to the results in cross-testing. Therefore, the final CNN model trained with the architecture of InceptionV3 could be considered as a suitable one for carcinoma recognition during PRB using endoscopic OCT in our study. Furthermore, the final CNN model was further tested using a renal oncocytoma sample. Although no renal oncocytoma images were involved in the training procedure, it exhibited strong performance

in differentiating between benign and malignant renal tumors, demonstrating the capability for accurately recognizing the renal carcinoma.

5.5 Conclusion

We tested the feasibility of employing an endoscopic OCT probe for the diagnosis of renal cancer. Renal tumor and normal kidney tissues can be distinguished based on their OCT images. To automate this process, we utilized both attenuation coefficient and deep learning methods for renal tissue recognition. The attenuation coefficient enabled the differentiation between tumor and normal tissues, while the CNN method further classified tumor tissue as well as different normal renal tissues accurately. It is worthy to note that CNN also provided better performance for distinguishing benign and malignant renal tumors. By combining the deep learning approach with the forward-viewing OCT probe, we envision the creation of a precise imaging guidance tool for the PRB procedure. Moreover, our system can complement the established clinical PRB guiding methods, including ultrasound, fluoroscopy, and CT by providing high-resolution images in front of the PRB needle in real time.

Reference

1. *American Cancer Society, Key Statistics About Kidney Cancer* *https://www.cancer.org/cancer/kidney-cancer/about/key-statistics.html*. 2022.
2. Levi, F., et al., *The changing pattern of kidney cancer incidence and mortality in Europe*. BJU Int, 2008. **101**(8): p. 949-58.
3. Ferlay, J., et al., *Cancer incidence and mortality patterns in Europe: Estimates for 40 countries and 25 major cancers in 2018*. Eur J Cancer, 2018. **103**: p. 356-387.
4. Whittier, W.L., *Complications of the Percutaneous Kidney Biopsy*. Advances in Chronic Kidney Disease, 2012. **19**(3): p. 179-187.
5. Iversen, P. and C. Brun, *Aspiration Biopsy of the Kidney*. American Journal of Medicine, 1951. **11**(3): p. 324-330.
6. Kark, R.M. and R.C. Muehrcke, *Biopsy of Kidney in Prone Position*. Lancet, 1954. **1**(May22): p. 1047-1049.
7. Tsivian, M., et al., *Small renal mass biopsy--how, what and when: report from an international consensus panel*. BJU Int, 2014. **113**(6): p. 854-63.

8. Yesudas, S.S., et al., *Percutaneous real-time ultrasound-guided renal biopsy performed solely by nephrologists: A case series.* Indian J Nephrol, 2010. **20**(3): p. 137-41.
9. Munoz, A.T., et al., *Percutaneous renal biopsy of native kidneys: efficiency, safety and risk factors associated with major complications.* Archives of Medical Science, 2011. **7**(5): p. 823-831.
10. Whittier, W.L. and S.M. Korbet, *Timing of complications in percutaneous renal biopsy.* Journal of the American Society of Nephrology, 2004. **15**(1): p. 142-147.
11. Madaio, M.P., *Renal biopsy.* Kidney Int, 1990. **38**(3): p. 529-43.
12. Korbet, S.M., K.C. Volpini, and W.L. Whittier, *Percutaneous Renal Biopsy of Native Kidneys: A Single-Center Experience of 1,055 Biopsies.* American Journal of Nephrology, 2014. **39**(2): p. 153-162.
13. Schiavon, L.H.O., et al., *Computed tomography-guided percutaneous biopsy of abdominal lesions: indications, techniques, results, and complications.* Radiol Bras, 2018. **51**(3): p. 141-146.
14. Braak, S.J., et al., *3D cone-beam CT guidance, a novel technique in renal biopsy--results in 41 patients with suspected renal masses.* Eur Radiol, 2012. **22**(11): p. 2547-52.
15. Garnon, J., et al., *Evaluation of percutaneous biopsies of renal masses under MRI-guidance: a retrospective study about 26 cases.* Eur Radiol, 2015. **25**(3): p. 617-23.
16. Wang, Z.J., A.C. Westphalen, and R.J. Zagoria, *CT and MRI of small renal masses.* Br J Radiol, 2018. **91**(1087): p. 20180131.
17. Huang, D., et al., *Optical Coherence Tomography.* Science, 1991. **254**(5035): p. 1178-1181.
18. Barwari, K., et al., *Advanced Diagnostics in Renal Mass Using Optical Coherence Tomography: A Preliminary Report.* Journal of Endourology, 2011. **25**(2): p. 311-315.
19. Yan, F., et al., *Optical coherence tomography for multicellular tumor spheroid category recognition and drug screening classification via multi-spatial-superficial-parameter and machine learning.* Biomedical Optics Express, 2024. **15**(4): p. 2014-2047.
20. Yan, F., et al., *Polarization-sensitive optical coherence tomography for renal tumor detection in ex vivo human kidneys.* Opt Lasers Eng, 2024. **173**.
21. Israelsen, N.M., et al., *Real-time high-resolution mid-infrared optical coherence tomography.* Light Sci Appl, 2019. **8**: p. 11.
22. Wang, C., et al., *Forward optical coherence tomography endoscope for percutaneous nephrostomy guidance*. SPIE BiOS. Vol. 11631. 2021: SPIE.
23. Wang, C., et al., *Deep-learning-aided forward optical coherence tomography endoscope for percutaneous nephrostomy guidance.* Biomed Opt Express, 2021. **12**(4): p. 2404-2418.

24. Yuan, W., et al., *Robust and fast characterization of OCT-based optical attenuation using a novel frequency-domain algorithm for brain cancer detection*. Scientific Reports, 2017. **7**.
25. Wang, C., et al. *Laparoscopic optical coherence tomography system for 3D bladder tumor detection*. in *Advanced Photonics in Urology*. 2021. International Society for Optics and Photonics.
26. Wang, C., et al., *Computer-aided Veress needle guidance using endoscopic optical coherence tomography and convolutional neural networks.* Journal of Biophotonics, 2022. **15**(5).
27. Wang, C., et al., *Epidural anesthesia needle guidance by forward-view endoscopic optical coherence tomography and deep learning.* Sci Rep, 2022. **12**(1): p. 9057.
28. Cui, S., et al., *Needle Size and the Risk of Kidney Biopsy Bleeding Complications*. Kidney Int Rep, 2016. **1**(4): p. 324-326.
29. Marchand, E.W., *Gradient index optics*. 1978, New York: Academic Press. x, 166 p.
30. Wobker, S.E. and S.R. Williamson, *Modern Pathologic Diagnosis of Renal Oncocytoma*. J Kidney Cancer VHL, 2017. **4**(4): p. 1-12.
31. Faber, D.J., et al., *Quantitative measurement of attenuation coefficients of weakly scattering media using optical coherence tomography.* Optics Express, 2004. **12**(19): p. 4353-4365.
32. Gong, P.J., et al., *Parametric imaging of attenuation by optical coherence tomography: review of models, methods, and clinical translation*. Journal of Biomedical Optics, 2020. **25**(4).
33. Scolaro, L., et al., *Parametric imaging of the local attenuation coefficient in human axillary lymph nodes assessed using optical coherence tomography*. Biomedical Optics Express, 2012. **3**(2): p. 366-379.
34. Cauberg, E.C.C., et al., *Quantitative measurement of attenuation coefficients of bladder biopsies using optical coherence tomography for grading urothelial carcinoma of the bladder.* Journal of Biomedical Optics, 2010. **15**(6).
35. Rodriguez, C.L.R., et al., *Decreased light attenuation in cerebral cortex during cerebral edema detected using optical coherence tomography.* Neurophotonics, 2014. **1**(2).
36. Badre, A. and C.L. Pan, *Explainable multi-task learning improves the parallel estimation of polygenic risk scores for many diseases through shared genetic basis*. Plos Computational Biology, 2023. **19**(7).
37. Badre, A. and C.L. Pan, *LINA: A Linearizing Neural Network Architecture for Accurate First-Order and Second-Order Interpretations*. Ieee Access, 2022. **10**: p. 36166-36176.
38. Selvaraju, R.R., et al., *Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization*. International Journal of Computer Vision, 2020. **128**(2): p. 336-359.

39. Muller, B.G., et al., *Prostate cancer diagnosis: the feasibility of needle-based optical coherence tomography.* J Med Imaging (Bellingham), 2015. **2**(3): p. 037501.
40. Zhao, Q.L., et al., *Ex vivo determination of glucose permeability and optical attenuation coefficient in normal and adenomatous human colon tissues using spectral domain optical coherence tomography.* Journal of Biomedical Optics, 2012. **17**(10).
41. Wagstaff, P.G.K., et al., *Percutaneous Needle Based Optical Coherence Tomography for the Differentiation of Renal Masses: a Pilot Cohort.* Journal of Urology, 2016. **195**(5): p. 1578-1584.

CHAPTER 6: EPIDURAL ANESTHESIA NEEDLE GUIDANCE BY FORWARD-VIEW ENDOSCOPIC OPTICAL COHERENCE TOMOGRAPHY AND DEEP LEARNING

This chapter is reproduced from "Chen Wang, Paul Calle, Justin C. Reynolds, Nu Bao Tran Ton, Feng Yan, Anthony M. Donaldson, Avery D. Ladymon, Pamela R. Roberts, Alberto J. De Armendi, Kar-Ming Fung, Chongle Pan, Qinggong Tang, *Epidural anesthesia needle guidance by forward endoscopic optical coherence tomography and ensemble deep learning*, **Sci. Rep**., 2022."

6.1 Introduction

Epidural anesthesia has become a well-established anesthetic method widely used in painless delivery [1], thoracic surgeries [2], orthopedic surgeries [3], organ transplantation surgeries [4], abdominal surgeries [5], and chronic pain relief [6]. Epidural anesthesia uses a needle to inject the anesthetic medications into the epidural space, which averages 1-6 millimeters in width and several centimeters in depth behind the skin layer [7]. During the placement of the epidural needle, the epidural needle penetrates subcutaneous fat, supraspinous ligament, interspinous ligament and ligamentum flavum before reaching the epidural space (between flavum and dura mater) to inject the medications [8]. Therefore, accurate positioning of the needle in the epidural space is critical for safe and effective epidural anesthesia.

Inadvertent penetration and damage to neurovascular structures leads to several complications, such as headache, transient paresthesia, and severe epidural hematomas [1]. Puncturing the dura will cause excessive loss of cerebrospinal fluid (CSF) and can damage nerves in the spinal cord [9]. It has been reported that more than 6% of patients have abnormal feelings during the placement of needle, and this has been proved to be a risk factor

of persistent paresthesia [10]. Post dural puncture headache (PDPH) is one of the most common complications in epidural anesthesia [11]. It occurs in over 50% of the accidental dural puncture cases [12]. Some researchers reported the PDPH incidence rate for females was two to three times greater than men, and pregnancy could further increase the possibility of PDPH [13]. Besides PDPH, more serious consequences such as spinal cord damage, paralysis or epidural hematoma and the development of an abscess might occur due to inaccurate puncture [14, 15]. Moreover, the neurologic injury caused by inadvertent puncture can lead to other symptoms like fever or photophobia [16, 17].

In current clinical practice, accurate placement of the needle relies on the experience of the anesthesiologist [18]. The most common method of detecting arrival of the needle in the epidural space is based on the loss of resistance (LOR) [19]. To test the LOR, the anesthesiologist keeps pressing on the plunger of a syringe filled with saline or air during the inserting the epidural needle [20]. When the needle tip passes through the ligamentum flavum and arrives at the epidural space, there is a sudden decrease of the resistance that can be sensed by anesthesiologists [21]. Nevertheless, this method has been proved to be inaccurate in predicting needle location and actual needle insertion could be further inside the body than the expectation [22]. Up to 10% of patients undergoing epidural anesthesia are not provided with adequate analgesia by using LOR [23, 24]. And the LOR technique can fail in up to 53% of the attempts without image guidance in more challenging procedures such as cervical epidural injections [25, 26]. Moreover, complications such as pneumocephalus [27], nerve root compression [28], subcutaneous emphysema [29] and venous air embolism [30] have been proved to be related to the air or liquid injection while using LOR technique. To improve

the success rate of epidural puncture and decrease the number of puncture attempts, there is a strong demand for an effective imaging technique to guide the epidural needle insertion.

Currently, imaging modalities, such as ultrasound and fluoroscopy, have been utilized during the needle access [31, 32]. However, the complex and articulated encasement of bones allows only a narrow acoustic window for the ultrasound beam [26]. Fluoroscopy does not have soft tissue contrast and, thus, cannot differentiate critical soft tissues (such as blood vessels and nerve roots) that need to be avoided during the needle insertion. Moreover, the limited resolution and contrast in fluoroscopy make it difficult to distinguish different tissue layers in front of the needle tip, especially for the cervical and thoracic epidural anesthesia where the epidural space is as narrow as 1-4 mm [33]. To improve the needle placement accuracy, novel optical imaging systems have been designed and tested. A portable optical epidural needle system based on fiberoptic bundle was designed to identify the epidural space [34], but there are some limitations for the optical signal interpretation and needle trajectory identification due to the uncertain direction of needle bevel or the surrounding fluid [35]. Additionally, optical spectral analysis has been utilized for tissue differentiation during epidural space identification [36, 37]. However, the accuracy of measured spectral results can be compromised by the surrounding tissues and the diffused blood during the puncture.

Optical coherence tomography (OCT) is a non-invasive imaging modality that can visualize the cross-sections of tissue samples [38]. At 10–100 times higher resolution (~10 μm) than ultrasound and fluoroscopy, OCT can improve the efficacy of tissue imaging [39]. OCT has been integrated with fiber-optic catheters and endoscopes for numerous internal imaging applications [40], [41-43]. Fiber-optic based OCT probe systems have been proposed in epidural anesthesia needle guidance and it provided promising results in identifying

epidural space in pig models [44, 45]. In the previous study, our group has also reported a forward-imaging endoscopic OCT needle device for real-time epidural anesthesia placement guidance and demonstrated its feasibility in piglets *in vivo* [26]. By fitting the OCT needle inside the hollow bore of the epidural needle, no additional invasiveness is introduced from the OCT endoscope. The high scanning speed of OCT system allows real-time imaging of the tissue OCT images in front of the needle. The tissues in front of the needle tip can be recognized based on the distinct OCT imaging features of the different tissues.

Convolutional neural networks (CNN) has been widely used for classification of medical images [46, 47] and have been applied for OCT images in macular, retina and esophageal related research for automatic tissue segmentation [48-50]. To help improve the efficiency of tissue recognition, herein we proposed to use CNN to classify and recognize different epidural tissue types automatically. In this study, we developed a computer-aided diagnosis (CAD) system based on CNN to automatically locate the epidural needle tip based on the forward-view OCT images. To the best of our knowledge, this is the first attempt to combine forward-view OCT system with CNN for guiding the epidural anesthesia procedure. Five epidural layers (including fat, interspinous ligament, ligamentum flavum, epidural space and spinal cord) were imaged to train and test the CNN classifiers based on Inception [51], Residual Network 50 (ResNet50) [52] and Xception [53]. After the needle tip arrives the epidural space, the OCT images can then be used to estimate the distance of the needle tip from the dura mater to avoid spinal cord damage. We trained and tested regression models based on Inception, ResNet50 and Xception using OCT images with manually labeled distances. The Inception model achieved the best performance with a mean absolute

percentage error of 3.05%±0.55%. These results demonstrated the feasibility of this novel imaging strategy for guiding the epidural anesthesia needle placement.

6.2 Results

6.2.1 OCT images of five epidural layer categories

The schematic of the experiment using our endoscopic OCT system was shown in Figure 6.1(A). Cross-sectional 2D OCT image examples of fat, interspinous ligament, ligamentum flavum, epidural space and spinal cord were shown in Figure 6.1(B). Because of the gap between needle tip and dura mater, epidural space was the simplest to be recognized. Among the other four tissues, interspinous ligament showed the most obvious imaging features, including the maximum penetration depth and the clear transverse stripes due to the thick fiber structure. Compared to other tissue types, ligamentum flavum showed higher imaging brightness close to the surface and the shallowest imaging depth. Imaging depths of fat and spinal cord were similar, but the imaging intensity of fat was not as evenly distributed as spinal cord. The corresponding histology results were also included in Figure 6.1(B). These tissues presented different cellular structures and distributions and correlated well with their OCT results except for fat. The fat tissue was featured with pockets of adipocytes in the histology, while this feature was not clear in the OCT results. This may be caused by the tissue compression we applied to mimic the clinical insertion scenario.

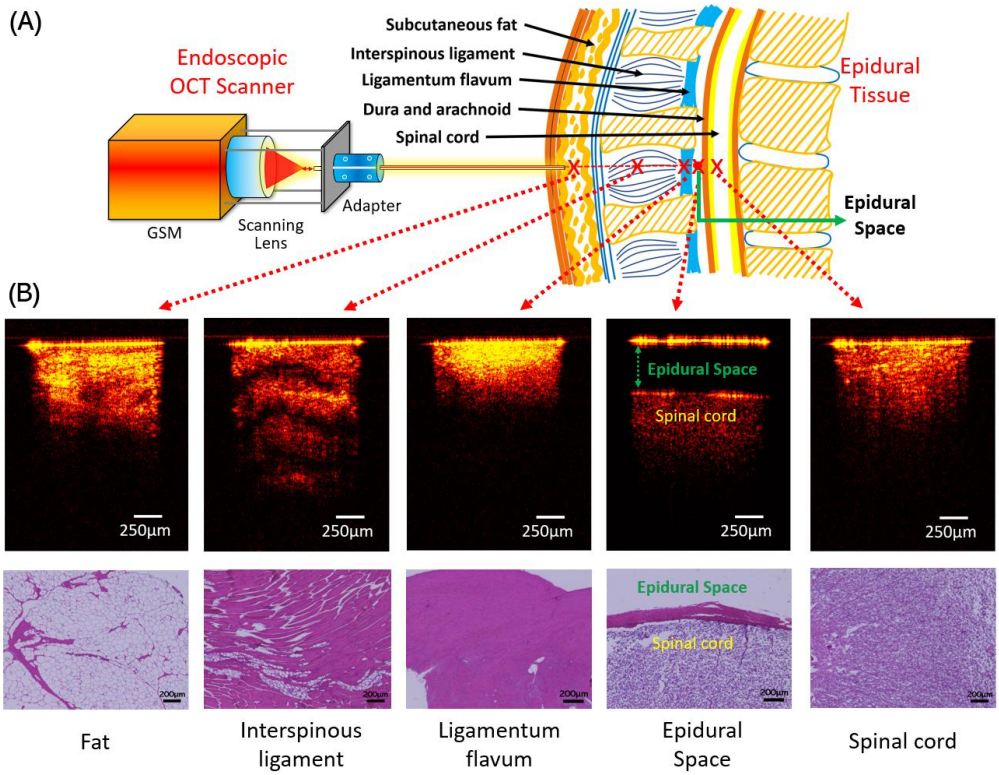


Figure 6.1. (A) Endoscopic OCT scanner setup and the representative OCT images of five epidural tissue layer categories. (B) Histology results of different tissue layers.

6.2.2 Multi-class classification of OCT images by tissue layers using sequential binary method

OCT images of the five tissue layers were classified using CNN models based on three architectures, including: ResNet50 [52], Xception [54] and Inception [55]. However, the overall accuracies of the multi-class classification models based on Inception reached ~66%. Although this was significantly higher than the accuracy of 20% by random guessing, further improvement was needed for clinical use.

Since the multi-class classification results were not satisfactory, herein we proposed to use sequential binary methods to improve the classification accuracies. During the needle

placement, the needle was inserted through fat, interspinous ligament, and ligamentum flavum until reaching the epidural space. Continuing the needle insertion beyond the epidural space can puncture the dura and damage the spinal cord. The classification process was thus divided into a sequential process of four binary classifications: (1) fat vs interspinous ligament; (2) interspinous ligament vs ligamentum flavum; (3) ligamentum flavum vs epidural space; and (4) epidural space vs spinal cord. The prediction results were shown in the Table 6.1.

Table 6.1. Average accuracies and standard error based on the practical tissue layer sequence during puncture for cross-validation.

	Fat vs Ligament			Ligament vs Flavum		
Testing Fold	ResNet50	Xception	Inception	ResNet50	Xception	Inception
S1	95.1% ± 2.2%	82.8% ± 8.5%	90.4% ± 2.2%	97.9% ± 1.7%	98.5% ± 1.2%	97.8% ± 2.0%
S2	93.6% ± 2.4%	88.2% ± 6.5%	92.4% ± 2.1%	98.7% ± 1.2%	98.9% ± 1.0%	98.9% ± 1.0%
S3	94.5% ± 2.5%	84.4% ± 6.5%	88.0% ± 2.4%	99.4% ± 0.5%	99.6% ± 0.3%	99.0% ± 0.5%
S4	90.9% ± 2.9%	84.7% ± 6.3%	89.7% ± 2.8%	98.5% ± 1.3%	98.7% ± 0.8%	97.7% ± 1.9%
S5	89.9% ± 2.3%	86.9% ± 4.7%	89.3% ± 2.5%	98.4% ± 1.4%	98.4% ± 0.9%	98.0% ± 1.5%
S6	90.6% ± 3.7%	82.3% ± 8.5%	89.9% ± 2.2%	98.8% ± 0.9%	98.2% ± 1.2%	97.8% ± 1.9%
S7	90.7% ± 3.5%	86.0% ± 3.3%	88.1% ± 3.0%	98.4% ± 1.4%	99.1% ± 0.7%	97.0% ± 2.7%
S8	88.7% ± 3.8%	82.7% ± 6.3%	86.3% ± 2.7%	98.7% ± 1.1%	98.8% ± 0.6%	98.7% ± 0.8%
Average	**91.8% ± 1.0%**	**84.7% ± 2.2%**	**89.3% ± 2.2%**	**98.6% ± 0.4%**	**98.8% ± 0.3%**	**98.1% ± 0.6%**
	Flavum vs Epidural Space			**Epidural Space vs Spinal Cord**		
Testing Fold	ResNet50	Xception	Inception	ResNet50	Xception	Inception
S1	100.0% ± 0.0%	100.0% ± 0.0%	100.0% ± 0.0%	100.0% ± 0.0%	100.0% ± 0.0%	100.0% ± 0.0%
S2	100.0% ± 0.0%	100.0% ± 0.0%	100.0% ± 0.0%	100.0% ± 0.0%	100.0% ± 0.0%	100.0% ± 0.0%
S3	100.0% ± 0.0%	100.0% ± 0.0%	100.0% ± 0.0%	100.0% ± 0.0%	100.0% ± 0.0%	100.0% ± 0.0%
S4	100.0% ± 0.0%	100.0% ± 0.0%	100.0% ± 0.0%	100.0% ± 0.0%	100.0% ± 0.0%	100.0% ± 0.0%
S5	100.0% ± 0.0%	100.0% ± 0.0%	100.0% ± 0.0%	100.0% ± 0.0%	100.0% ± 0.0%	100.0% ± 0.0%
S6	100.0% ± 0.0%	100.0% ± 0.0%	100.0% ± 0.0%	100.0% ± 0.0%	100.0% ± 0.0%	100.0% ± 0.0%
S7	100.0% ± 0.0%	100.0% ± 0.0%	100.0% ± 0.0%	100.0% ± 0.0%	100.0% ± 0.0%	100.0% ± 0.0%
S8	100.0% ± 0.0%	100.0% ± 0.0%	100.0% ± 0.0%	100.0% ± 0.0%	100.0% ± 0.0%	100.0% ± 0.0%
Average	**100.0% ± 0.0%**	**100.0% ± 0.0%**	**100.0% ± 0.0%**	**100.0% ± 0.0%**	**100.0% ± 0.0%**	**100.0% ± 0.0%**

Overall, ResNet50 showed the best prediction results. There was substantial variability in the test accuracy among different subjects especially for the prediction accuracy of "Fat vs Interspinous Ligament". While three subjects had test accuracies higher than 98.8%, the subjects in the S2 fold had the lowest test accuracy of 67.3%. This may be due to the tissue

variability among different back bone samples and the different tissue compression during imaging especially considering fat is subject to tissue compression. The areas under the ROC curve (AUC) differed among different samples.

Table 6.2. Average and standard error for cross-testing for the four binary comparisons for ResNet50.

Testing Fold	Fat vs Ligament	Ligament vs Flavum	Flavum vs Epidural Space	Epidural Space vs Spinal Cord	Average
S1	81.3%	98.8%	100.0%	100.0%	**95.0% ± 4.6%**
S2	67.3%	98.0%	100.0%	100.0%	**91.3% ± 8.0%**
S3	92.0%	87.7%	98.8%	100.0%	**94.6% ± 2.9%**
S4	99.9%	99.4%	100.0%	100.0%	**99.8% ± 0.1%**
S5	98.8%	99.4%	100.0%	100.0%	**99.6% ± 0.3%**
S6	79.5%	99.0%	100.0%	100.0%	**94.6% ± 5.1%**
S7	94.2%	99.8%	100.0%	100.0%	**98.5% ± 1.4%**
S8	98.8%	100.0%	100.0%	100.0%	**99.7% ± 0.3%**
Average	**89.0% ± 4.2%**	**97.8% ± 1.5%**	**99.8% ± 0.2%**	**100.0% ± 0.0%**	**96.65% ± 1.32%**

Class activation heatmaps of ResNet50 models were created for representative images to show the salient features used for classification (Figure 6.2A). Each binary classification model paid attention to different regions of the images. For example, the black empty space was important for the models to recognize the epidural space images. A video stream of the OCT images was used to demonstrate the sequential binary models. The number of images was 100, 700, 100, 100, and 150 for fat, interspinous ligament, ligamentum flavum, epidural space and spinal cord, respectively, which was proportional to the width of these tissue layers [56-60]. After the binary classifier of fat vs. ligament detects 35 interspinous ligament images in the last 50 images, the needle was considered to be in the interspinous ligament and the next binary classifier of interspinous ligament vs. ligamentum flavum was activated to detect the upcoming ligamentum flavum. This simple logic was used to switch all the subsequent classifiers. Figure 6.2B showed some images from a video that can be found in the Github

repository. Scenes from the video showing the switch from classifier 2 to classifier 3 and its arrival to epidural space. Each image showed three important pieces of information. First, the proportion of the last 50 images that were predicted to belong to Class 1, e.g., Class 1 was interspinous ligament in the first Classifier and was ligamentum flavum in the second Classifier. Initially, when the number of images was less than 50, the denominator shows the total number of images. Additionally, the color of fraction followed traffic lights colors. It changed from green to yellow at 26 and from yellow to red at 35. The second information was the current classifier. The last information was the truth and predicted label. The switch of binary classifier occurred when the number of images predicted as Class 1 reached 35. The fraction did not appear anymore when the last classifier was reached.

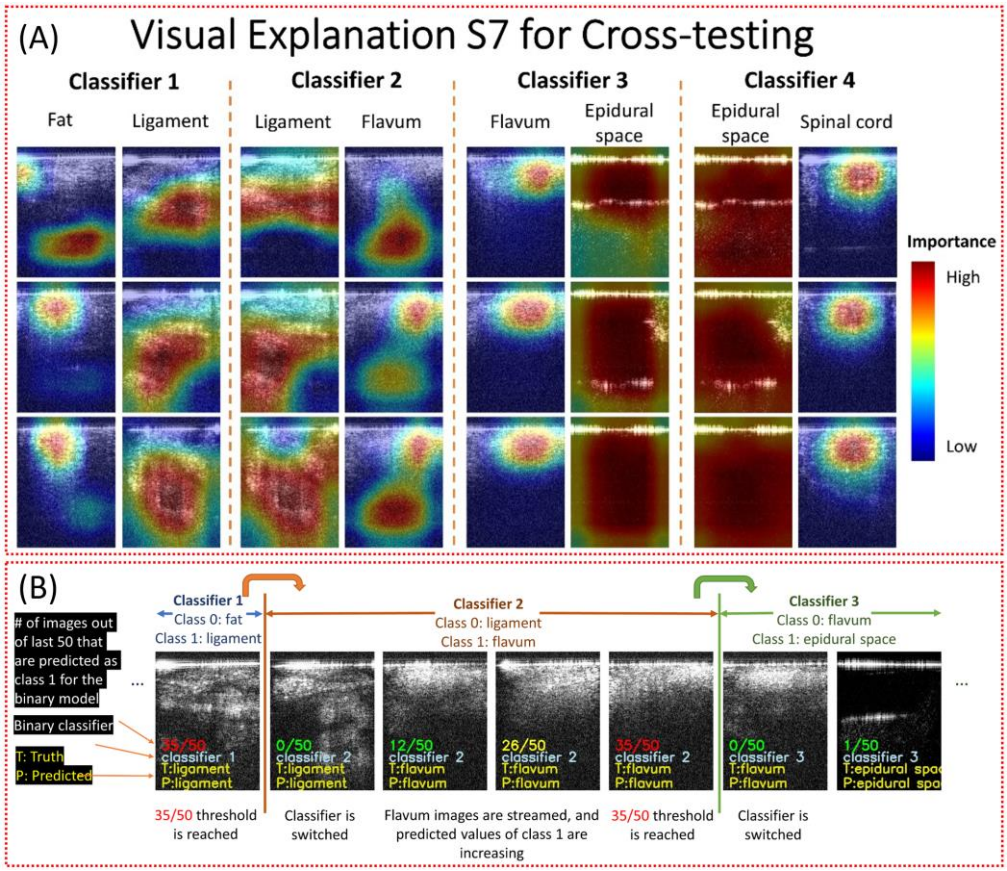


Figure 6.2. Class activation heatmaps for Subject 7 using ResNet50 in cross-testing (A) and video captures of the insertion process (B).

6.2.3 Estimation of the distance between the needle tip and dura mater by regression

Inception, ResNet50, and Xception were compared for the regression task of estimating the distance of the needle tip to the dura mater. In Table 6.3, the mean and standard error of the cross-validation mean absolute percentage error (MAPE) for ResNet50, Xception, and Inception in all testing folds were shown. In every fold, the Inception model outperformed the ResNet50 and Xception models, indicated by the lowest MAPE.

Table 6.3. The average loss for each model type in cross-validation for each testing fold.

Testing folds	ResNet50	Xception	Inception
S1	3.71% ± 0.91%	3.99% ± 1.14%	**3.13% ± 0.60%**
S2	4.04% ± 0.99%	3.84% ± 0.98%	**3.42% ± 0.82%**
S3	3.82% ± 0.88%	3.62% ± 0.78%	**3.09% ± 0.69%**
S4	3.97% ± 1.34%	4.67% ± 1.68%	**3.31% ± 0.82%**
S5	3.88% ± 0.84%	4.15% ± 1.08%	**3.23% ± 0.60%**
S6	3.89% ± 0.82%	4.71% ± 1.31%	**3.41% ± 0.83%**
S7	3.88% ± 1.31%	4.51% ± 1.61%	**3.81% ± 1.25%**
S8	2.77% ± 0.32%	3.33% ± 0.39%	**2.61% ± 0.37%**
Average	**3.74% ± 0.14%**	**4.10% ± 0.18%**	**3.25% ± 0.12%**

In each testing rotation, a new Inception model was trained using all the images in the seven cross-validation folds and then evaluated on the unseen testing images in the one testing fold. Examples of OCT images with different distances between needle tip and tissue were shown in Figure 6.3(A). A model was trained on 21,000 images belonging to subjects 1, 2, 3, 4, 5, 6, and 8, and tested on 3,000 images belonging to subject 7. The distribution of the errors from the Inception model during the seventh testing fold (i.e., testing images belong to subject 7) can be visualized with the violin plots in Figure 6.3(B). The MAPE on this testing set was 3.626%, and the mean absolute error (MAE) was 34.093 μm. From the testing results on the

Inception architecture, it was evident that the regression model can accurately estimate the distance to the dura mater in most of the OCT images.

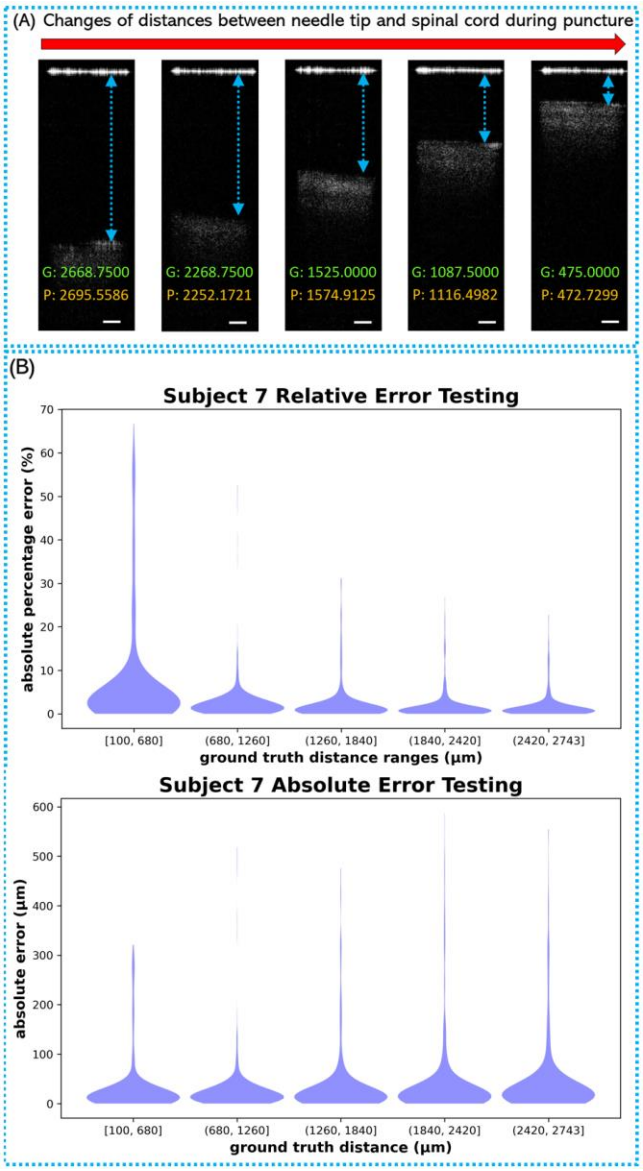


Figure 6.3. (A) Examples of epidural space images with different distances between needle tip and spinal cord surface. G: labeled ground truth value (μm); P: prediction value (μm); Scale bar: 250μm. (B) The distribution of the predicted absolute percentage errors and absolute error in testing fold 7 with 3000 testing images.

6.3 Discussion

In this study, we validated our endoscopic OCT system for epidural anesthesia surgery guidance. The OCT endoscope can provide 10-100 times higher resolution than conventional medical imaging modalities. Moreover, this proposed endoscopic OCT system is compatible with the clinical-used epidural guiding methods (*e.g.*, ultrasound, fluoroscopy, and CT), and will complement these macroscopic methods by providing the detailed images in front of the epidural needle.

Five different tissue layers including fat, interspinous ligament, ligamentum flavum, epidural space and spinal cord were imaged. To assist the OCT image interpretation, a deep learning-based CAD platform was developed to automatically differentiate the tissue layers at the epidural needle tip and predict the distance from the needle tip to dura mater.

Three convolutional neural network architectures, including ResNet50, Xception and Inception, were tested for image classification and distance regression. The best classification accuracy of the five tissue layers were 60-65% from a multi-class Inception classifier. The main challenge was the differentiation between fat and spinal cord. because they had similar feature in OCT images (Figure 6.1). Based on the needle puncture sequence, we divided the overall classification into four sequential binary classifications: Fat vs Interspinous Ligament; Interspinous Ligament vs Ligamentum Flavum; Ligamentum Flavum vs Epidural Space, and Epidural Space vs Spinal Cord. The overall prediction accuracies of all four classifications reached to more than 90%. ResNet50 presented the best overall performance compared to Xception and Inception. Due to the unique features of epidural space in OCT images, it was possible to achieve >99% precision when the needle arrived the epidural space. Table 6.2 showed the accuracies of ~99.8% and 100% when classifying Epidural Space vs Ligamentum

Flavum and Epidural Space vs Spinal Cord. This will allow accurate detection of the epidural space for injection of the anesthetic during epidural anesthesia. The sequential transition from one binary classifier to the next was controlled accurately using a simple logic, which was demonstrated in a video simulating the insertion of a needle through the five tissue layers (Figure 6.2). In the future, this can be improved by combining CNN with Recurrent Neural Network (RNN) to handle the temporal dimension of video streaming data [61]. Additionally, we developed a CNN regression model to estimate the needle distance to the dura mater upon entry of the epidural space. For the regression task, Inception provided better performance compared to Xception and ResNet50. The mean relative error was 3.05%, which was able to track the accurate location of the needle tip in the epidural space.

CNNs have shown to be a valuable tool in biomedical imaging. Manually configuring CNN architectures for an imaging modality can be a tedious trial-and-error process. ResNet, Inception, and Xception are commonly used architectures for general image classification tasks. Here, we showed that the architectures can be easily adapted for both classification and regression tasks in biomedical imaging applications. The best performance was obtained by ResNet50 for the binary classifications and by Inception for the distance regression.

The nested-cross validation and testing procedure was computationally expensive, but it provided the uncertainty quantification of the test performance across subjects. The wall-clock time for training the binary classification models on NVIDIA Volta GPUs were ~11 min per validation fold for ResNet50, ~32 min per validation fold for Xception, and ~11 min per validation fold for Inception. The wall-clock time for training the regression models on NVIDIA RTX 3090 GPUs were ~50 min per validation fold for ResNet50, ~145 min per validation fold for Xception, and ~36 min per validation fold for Inception. The inferencing

for the binary classifications on NVIDIA Volta GPUs took 13 millisecond per image on average. The inferencing for the distance regression on NVIDIA RTX 3090 GPUs took 2.1 millisecond per image on average. In future, the inferencing by these large CNN models can be further accelerated by weight pruning and knowledge distillation [62].

In the next study, we will use the GRIN lens with a suitable diameter for practical 16-gauge Tuohy needle used in epidural anesthesia in our future hardware design [63, 64]. Furthermore, we will miniaturize the size of our OCT scanner to make our system more portable and convenient for anesthesiologists to use in clinical applications. Finally, we will test the performance of our endoscopic OCT system together with the deep learning-based CAD platform in the *in-vivo* pig experiments. Difference of OCT images from *in-vivo* and *ex-vivo* condition may deteriorate the *in-vivo* testing results. In that case, we will re-train our model using *in vivo* pig data. Additionally, during the *in-vivo* experiments, there will be blood vessels surrounding the spinal cord [65]. To address this, we plan to further use Doppler OCT method for the blood vessel detection to avoid the rupture of blood vessels during epidural needle insertion.

6.4 Method

6.4.1 Experiment setup

The schematic of our forward-view OCT endoscope was shown in Figure 6.4. Its working principle was based on a Michaelson interferometer with a reference arm and a sample arm [38]. The endoscopic system was built on a swept-source OCT (SS-OCT). The light source was a wavelength-swept laser with 1300 nm central wavelength and 100 nm bandwidth [66]. The laser had the maximum scanning speed at 200 kHz A-Scan rate. The light from the laser was first unevenly split by a fiber coupler (FC). 97% power was split into

the circulator and transmitted into the interferometer, and the other 3% was input to the Mach-Zehnder interferometer (MZI) which provided the triggering signal for data sampling. The 97% power was further split by another 50:50 FC to the reference arm and the sample arm. The reflected signal from the reference arm and the backscattered signal from the sample arm interfered with each other and were collected by a balanced detector (BD) for noise reduction. The signal was then sent to data acquisition board (DAQ) and computer for post-processing based on Fourier transform [67]. While imaging the samples in the air, the axial resolution reached to 10.6 μm and the lateral resolution was 20 μm.

To achieve the endoscopic imaging, a gradient-index (GRIN) rod lens was added in the sample arm. It was fixed in front of the scanning lens of the galvanometer scanning mirror (GSM). The GRIN lens used in this study had a total length of 138 mm, an inner diameter of 1.3 mm, and a view angle of 11.0°. It was protected by a thin-wall steel tubing. For dispersion compensation, a second set of identical GRIN lens was stabilized in front of the reflector (mirror) of the reference arm. In addition, two polarization controllers (PC) were placed in each arm to reduce the noise level.

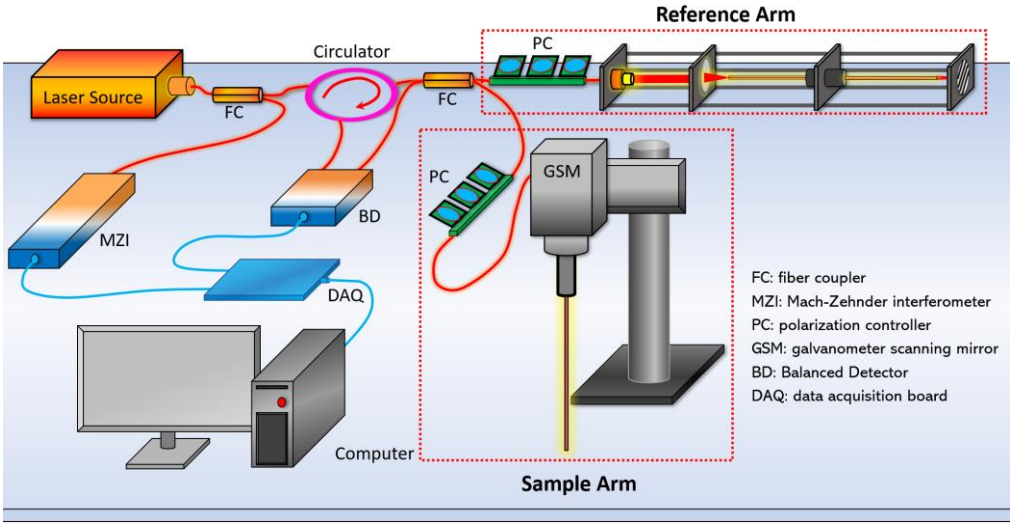


Figure 6.4. Schematic of forward-view OCT endoscope system.

The GRIN lens utilized in the sample arm was assembled in front of the OCT scanning lens of the GSM. To decrease the reflection from the proximal end surface of the GRIN lens that significantly degraded the imaging quality, the proximal surface of the GRIN lens was aligned ~1.5 mm off the focus of the scanning lens. The GRIN lens had four integer pitch length to relay the images from the distal end to its proximal surface [68]. In the sample arm, the proximal GRIN lens surface was adjusted close to the focus point of the objective after the OCT scanner. Thus, the spatial information from the distal surface (tissue sample) of the GRIN lens transmitted to the proximal surface was further collected by the OCT scanner. Therefore, OCT images of the epidural tissues in front of the GRIN lens can be successfully obtained. Our endoscopic system provided ~1.25 mm field of view (FOV) with sensitivity of 92 dB.

6.4.2 Data acquisition

Backbones from eight pigs were acquired from local slaughterhouses and cut at the middle before imaging to expose different tissue layers. From the cross-section of the sample, different tissue types could be clearly distinguished through the tissue anatomic features and their positions as shown in Figure 6.5. To further limit the number of misclassified results, two lab members confirmed the tissue types before imaging started. In Figure 6.5, five tissue layers including fat, interspinous ligament, ligamentum flavum, epidural space and spinal cord can be distinguished from their anatomic appearance. The OCT needle was placed against these confirmed tissue layers to obtain their OCT structural images. Following the practice of epidural needle placement, we mimicked the puncturing process by inserting the OCT endoscope through fat, interspinous ligament, ligamentum flavum and epidural space of our sample. Since the targeted position of the anesthetic injection is the epidural space with

width ~ 1-6 mm [69], we also obtained OCT images of epidural space by positioning the needle tip in front of the spinal cord at different distances. To mimic the condition of accidental puncture into spinal cord, we took OCT images while inserting the endoscope into the spinal cord. Some force was applied during imaging the four tissue types (fat, interspinous ligament, ligamentum flavum, and spinal cord) to generate compression to better represent the actual *in-vivo* clinical situation.

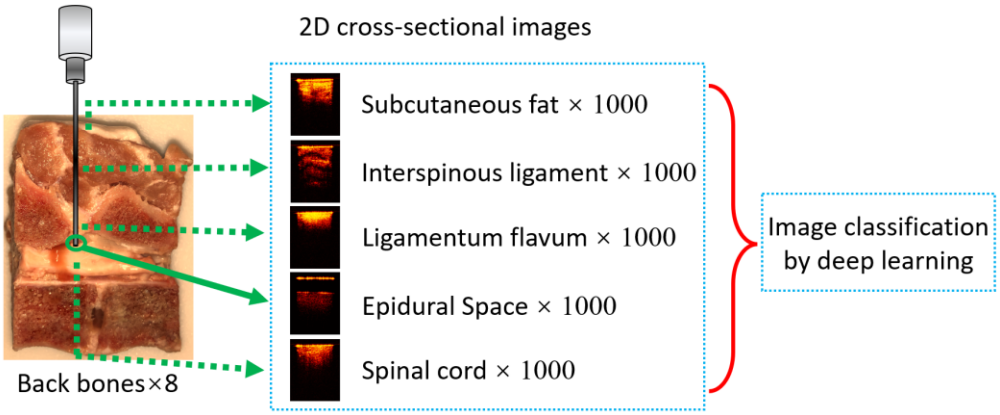


Figure 6.5. Data acquisition process

For each backbone sample, 1,000 cross-sectional OCT images were obtained from each tissue layer. To decrease noise and increase the deep-learning processing speed, the original images were further cropped to smaller sizes that only contained the effective tissue information. Imaged were cropped to 181×241 pixels for the tissue classification. The data was uploaded to Zenodo (http://doi.org/10.5281/zenodo.5018581) [70].

At the end of imaging, tissues of fat, interspinous ligament, ligamentum flavum and spinal cord with dura mater of the porcine back bones were excised and processed for histology following the same orientation of OCT endoscope imaging to compare with corresponding OCT results. The tissues were fixed with 10% formalin, embedded in paraffin, sectioned (4 μm thick) and stained with hematoxylin and eosin (H & E) for histological

analysis. Images were analyzed by Keyence Microscope BZ-X800. Sectioning and H & E staining was carried out by the Tissue Pathology Shared Resource, Stephenson Cancer Center (SCC), University of Oklahoma Health Sciences Center. The Hematoxylin (cat#3801571) and Eosin (cat# 3801616) were purchased from Leica biosystems, and the staining was performed utilizing Leica ST5020 Automated Multistainer following the HE staining protocol at the SCC Tissue Pathology core.

6.4.3 Convolutional neural networks

Convolutional Neural Networks (CNN) were used to classify OCT images by epidural layers. Three CNN architectures, including ResNet50 [52], Inception [51] and Xception [53], were imported from the Keras library [71]. The output layer of the models was a dense layer that represented the number of categories. The images were centered by subtracting training mean pixel value. The SGD with Nesterov momentum optimizer was used with a learning rate of 0.01, a momentum of 0.9, and a decay of 0.01. The batch size was 32. Early stopping was used with a patience of 10. The loss function used was sparse categorical cross entropy.

Nested cross-validation and testing [72, 73] were used for model selection and benchmarking as described previously [74]. This evaluation strategy provided an unbiased estimation of model performance with uncertainty quantification using two nested loops for cross-validation and cross-testing. Images were acquired from eight subjects in this dataset. The images were divided to 8 folds by subjects to account for the subject-to-subject variability. An eight-fold cross-testing loop was performed by rotating through every subject for testing and using the remaining seven subjects (7,000 images) for cross-validation. In the cross-validation, six subjects were used for training and one subject for validation in each rotation. The 7-fold cross-validation loop was used to compare the performance of three

architecture models: ResNet50, Xception and Inception. The model with the best cross-validation performance was automatically selected for performance benchmarking in the corresponding testing fold. The performance of this overall procedure was evaluated by aggregating the testing performance from all 8 testing folds. Grad-CAM [75] was used to generate instance-wise explanation of selected models [76, 77].

The computation was performed using the Schooner supercomputer at the University of Oklahoma and the Summit supercomputer at Oak Ridge National Laboratory. The computation on Schooner used five computational nodes, each of which had 40 CPU cores (Intel Xeon Cascade Lake) and 200 GB of RAM. The computation on Summited used up to 10 nodes, each of which had 2 IBM POWER9 processors and 6 NVIDIA Volta Graphic cards. The complete code for the classification models can be found at https://github.com/thepanlab/Endoscopic_OCT_Epidural.

The classification accuracy of the models was computed as:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad \textbf{(1)}$$

where TP was True Positives, TN was True Negatives, FP was False Positives, and FN was False Negatives.

Receiver Operating Characteristic (ROC) curves were used to visualize the relationship between sensitivity and specificity. The area under the curve (AUC) of ROC was also used to assess the overall performance of the models.

6.4.4 Epidural distance prediction using deep learning

OCT images of epidural space were obtained at a range of distances between approximately 0.2 mm and 2.5 mm from the needle tip to the spinal cord surface (dura mater). A total of 24,000 images from eight subjects were used for this task. For each image taken in

the epidural space for the distance estimation task, the distance in micrometers (μm) from the epidural needle to the dura mater was manually calculated and labeled. This distance label served as the ground truth for computing the loss during the training process in the regression model. All images were of 241x681 pixels on X and Z (depth) axes with pixel size of 6.25 μm. The pixel values for each image were scaled in the range of 0-255.

The regression model was developed to estimate the distance from the epidural needle to the dura upon entry into the epidural space automatically. Three architectures, including ResNet50, Inception, and Xception, were compared using nested cross-validation and testing as described above. The final output layer consisted of a single neuron with an identity activation function for regression on the continuous distance values [78]. The SGD algorithm with Nesterov momentum optimization was used with a learning rate of 0.01, momentum of 0.9, and a decay rate of 0.01. Training took place with a batch size of 32 over 20 epochs. The mean absolute percentage error (MAPE) and mean absolute error (MAE) were the metrics used to evaluate the regression performance due to their intuitive interpretability in relation to the relative error. The MAPE and MAE performance metrics are defined in equation (2) and (3), respectively. Model training and testing for the regression task was performed on a workstation equipped with dual NVIDIA RTX 3090 GPUs. The complete code for the regression models can be found at: https://github.com/thepanlab/Endoscopic_OCT_Epidural.

The classification accuracy of the models was computed as:

$$MAPE(\%) = \frac{100\%}{n}\sum_{i=1}^{n}\frac{|Y_i - X_i|}{Y_i} \quad \textbf{(2)}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|Y_i - X_i| \quad \textbf{(3)}$$

where Y_i was the ground truth distance, X_i was the predicted distance, and n was the number of OCT images.

Reference

1. Moraca, R.J., D.G. Sheldon, and R.C. Thirlby, *The role of epidural anesthesia and analgesia in surgical practice.* Annals of Surgery, 2003. **238**(5): p. 663-673.
2. Svircevic, V., et al., *Meta-analysis of Thoracic Epidural Anesthesia versus General Anesthesia for Cardiac Surgery.* Anesthesiology, 2011. **114**(2): p. 271-282.
3. Hollmann, M.W., et al., *Epidural anesthesia prevents hypercoagulation in patients undergoing major orthopedic surgery.* Regional Anesthesia and Pain Medicine, 2001. **26**(3): p. 215-222.
4. Hadimioglu, N., et al., *Combination of Epidural Anesthesia and General Anesthesia Attenuates Stress Response to Renal Transplantation Surgery.* Transplantation Proceedings, 2012. **44**(10): p. 2949-2954.
5. Carli, F., J.L. Trudel, and P. Belliveau, *The effect of intraoperative thoracic epidural anesthesia and postoperative analgesia on bowel function after colorectal surgery - A prospective, randomized trial.* Diseases of the Colon & Rectum, 2001. **44**(8): p. 1083-1089.
6. Li, Y., et al., *Effects of thoracic epidural anesthesia/analgesia on the stress response, pain relief, hospital stay, and treatment costs of patients with esophageal carcinoma undergoing thoracic surgery: A single-center, randomized controlled trial.* Medicine (Baltimore), 2019. **98**(7): p. e14362.
7. Pesteie, M., et al., *Automatic Localization of the Needle Target for Ultrasound-Guided Epidural Injections.* IEEE Trans Med Imaging, 2018. **37**(1): p. 81-92.
8. Scott, D.B., *Identification of the epidural space: Loss of resistance to air or saline?* Regional Anesthesia, 1997. **22**(1): p. 1-2.
9. Turnbull, D.K. and D.B. Shepherd, *Post-dural puncture headache: pathogenesis, prevention and treatment.* Br J Anaesth, 2003. **91**(5): p. 718-29.
10. Horlocker, T.T., *Complications of spinal and epidural anesthesia.* Anesthesiol Clin North Am, 2000. **18**(2): p. 461-85.
11. Horlocker, T.T. and D.J. Wedel, *Neurologic complications of spinal and epidural anesthesia.* Reg Anesth Pain Med, 2000. **25**(1): p. 83-98.
12. Apfel, C.C., et al., *Prevention of postdural puncture headache after accidental dural puncture: a quantitative systematic review.* Br J Anaesth, 2010. **105**(3): p. 255-63.
13. Wu, C.L., et al., *Gender and post-dural puncture headache.* Anesthesiology, 2006. **105**(3): p. 613-8.
14. Hodges, S.D., et al., *Cervical epidural steroid injection with intrinsic spinal cord damage. Two case reports.* Spine (Phila Pa 1976), 1998. **23**(19): p. 2137-42; discussion 2141-2.
15. Tripathi, M., S.S. Nath, and R.K. Gupta, *Paraplegia after intracord injection during attempted epidural steroid injection in an awake-patient.* Anesth Analg, 2005. **101**(4): p. 1209-11, table of contents.
16. Horlocker, T.T., et al., *A retrospective review of 4767 consecutive spinal anesthetics: central nervous system complications. Perioperative Outcomes Group.* Anesth Analg, 1997. **84**(3): p. 578-84.

17. Keohane, M., *Patient comfort: spinal versus epidural anesthesia for cesarean section*. Anesth Analg, 1996. **82**(1): p. 219.
18. Rigg, J.R.A., et al., *Epidural anaesthesia and analgesia and outcome of major surgery: a randomised trial*. Lancet, 2002. **359**(9314): p. 1276-1282.
19. Hoffmann, V.L.H., et al., *Posterior epidural space depth: safety of the loss of resistance and hanging drop techniques*. British Journal of Anaesthesia, 1999. **83**(5): p. 807-809.
20. Frumin, M.J., et al., *Sites of sensory blockade during segmental spinal and segmental peridural anesthesia in man*. Anesthesiology, 1953. **14**(6): p. 576-83.
21. Grondin, L.S., et al., *Success of Spinal and Epidural Labor Analgesia Comparison of Loss of Resistance Technique Using Air versus Saline in Combined Spinal-Epidural Labor Analgesia Technique*. Anesthesiology, 2009. **111**(1): p. 165-172.
22. Honigmann, S., et al. *EpiGuide 2D: visibility assessment of a novel multi-channel out-of-plane needle guide for 2D point-of-care ultrasound*. in *Medical Imaging 2019: Image-Guided Procedures, Robotic Interventions, and Modeling*. 2019. International Society for Optics and Photonics.
23. Hermanides, J., et al., *Failed epidural: causes and management*. British Journal of Anaesthesia, 2012. **109**(2): p. 144-154.
24. McLeod, A., A. Roche, and M. Fennelly, *Case series: Ultrasonography may assist epidural insertion in scoliosis patients*. Canadian Journal of Anaesthesia-Journal Canadien D Anesthesie, 2005. **52**(7): p. 717-720.
25. Stojanovic, M.P., et al., *The role of fluoroscopy in cervical epidural steroid injections: an analysis of contrast dispersal patterns*. Spine (Phila Pa 1976), 2002. **27**(5): p. 509-14.
26. Tang, Q., et al., *Real-time epidural anesthesia guidance using optical coherence tomography needle probe*. Quant Imaging Med Surg, 2015. **5(1)**: p. 118-124.
27. Saidman, L.J. and E.I. Eger, 2nd, *Change in Cerebrospinal Fluid Pressure during Pneumoencephalography under Nitrous Oxide Anesthesia*. Anesthesiology, 1965. **26**: p. 67-72.
28. Cheng, A.C.K., *Intended Epidural-Anesthesia as Possible Cause of Cauda-Equina Syndrome*. Anesthesia and Analgesia, 1994. **78**(1): p. 157-159.
29. Carter, M.I., *Cervical surgical emphysema following extradural analgesia*. Anaesthesia, 1984. **39**(11): p. 1115-6.
30. Mehta, M., M.D. Sokoll, and S.D. Gergis, *Effects of venous air embolism on the cardiovascular system and acid base balance in the presence and absence of nitrous oxide*. Acta Anaesthesiol Scand, 1984. **28**(2): p. 226-31.
31. Grau, T., et al., *Real-time ultrasonic observation of combined spinal-epidural anaesthesia*. European Journal of Anaesthesiology, 2004. **21**(1): p. 25-31.
32. Kang, S.Y., et al., *Advantages of the combination of conscious sedation epidural anesthesia under fluoroscopy guidance in lumbar spine surgery*. Journal of pain research, 2020. **13**: p. 211.
33. Richardson, J. and G.J. Groen, *Applied epidural anatomy*. Continuing Education in Anaesthesia, Critical Care & Pain, 2005. **5**(3): p. 98-100.

34. Ting, C.K. and Y. Chang, *Technique of fiber optics used to localize epidural space in piglets.* Optics Express, 2010. **18**(11): p. 11138-11147.
35. Lin, S.P., et al., *Discriminant analysis for anaesthetic decision-making: an intelligent recognition system for epidural needle insertion.* British Journal of Anaesthesia, 2012. **108**(2): p. 302-307.
36. Rathmell, J.P., et al., *Identification of the Epidural Space with Optical Spectroscopy An In Vivo Swine Study.* Anesthesiology, 2010. **113**(6): p. 1406-1418.
37. Anderson, T.A., et al., *Raman Spectroscopy Differentiates Each Tissue from the Skin to the Spinal Cord: A Novel Method for Epidural Needle Placement?* Anesthesiology, 2016. **125**(4): p. 793-804.
38. Huang, D., et al., *Optical Coherence Tomography.* Science, 1991. **254**(5035): p. 1178-1181.
39. Li, X.D., et al., *Optical coherence tomography: Advanced technology for the endoscopic imaging of Barrett's esophagus.* Endoscopy, 2000. **32**(12): p. 921-930.
40. Carotenuto, B., et al., *Optical Guidance Systems for Epidural Space Identification.* IEEE Journal of Selected Topics in Quantum Electronics, 2017. **23**(2): p. 371-379.
41. Sharma, G.K., et al., *Long-Range Optical Coherence Tomography of the Neonatal Upper Airway for Early Diagnosis of Intubation-related Subglottic Injury.* American Journal Of Respiratory And Critical Care Medicine, 2015. **192**(12): p. 1504-1513.
42. Chen, Y., et al., *Ultrahigh resolution optical coherence tomography of Barrett's esophagus: preliminary descriptive clinical study correlating images with histology.* Endoscopy, 2007. **39**(7): p. 599-605.
43. Sergeev, A.M., et al., *In vivo endoscopic OCT imaging of precancer and cancer states of human mucosa.* Optics Express, 1997. **1**(13): p. 432-440.
44. Kao, M.C., et al., *Intelligent epidural needle placement using fiber-probe optical coherence tomography in a piglet model.* Biomedical Optics Express, 2018. **9**(8).
45. Kuo, W.C., et al., *Fiber-needle swept-source optical coherence tomography system for the identification of the epidural space in piglets.* Anesthesiology, 2015. **122**(3): p. 585-94.
46. Litjens, G., et al., *A survey on deep learning in medical image analysis.* Medical Image Analysis, 2017. **42**: p. 60-88.
47. Shen, D., G. Wu, and H.-I. Suk, *Deep Learning in Medical Image Analysis.* Annual Review of Biomedical Engineering, 2017. **19**(1): p. 221-248.
48. Wang, C., et al., *Adversarial convolutional network for esophageal tissue segmentation on OCT images.* Biomedical Optics Express, 2020. **11**(6): p. 3095-3110.
49. Rong, Y., et al., *Surrogate-Assisted Retinal OCT Image Classification Based on Convolutional Neural Networks.* IEEE Journal of Biomedical and Health Informatics, 2019. **23**(1): p. 253-263.

50. Rasti, R., et al., *Macular OCT Classification Using a Multi-Scale Convolutional Neural Network Ensemble.* IEEE Transactions on Medical Imaging, 2018. **37**(4): p. 1024-1034.
51. Szegedy, C., et al. *Rethinking the inception architecture for computer vision*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.
52. He, K., et al. *Deep residual learning for image recognition*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.
53. Chollet, F. *Xception: Deep learning with depthwise separable convolutions*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.
54. Kassani, S.H., et al. *Diabetic retinopathy classification using a modified xception architecture*. in *2019 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*. 2019. IEEE.
55. Xia, X., C. Xu, and B. Nan. *Inception-v3 for flower classification*. in *2017 2nd International Conference on Image, Vision and Computing (ICIVC)*. 2017. IEEE.
56. Rajput, D., A.K. Srivastava, and R. Kumar, *Spinal epidural lipomatosis: An unusual cause of relapsing and remitting paraparesis.* J Pediatr Neurosci, 2010. **5**(2): p. 150-2.
57. Frostell, A., et al., *A Review of the Segmental Diameter of the Healthy Human Spinal Cord.* Front Neurol, 2016. **7**: p. 238.
58. Waldman, S.D., *Preface,* in *Pain Management*, S.D. Waldman and J.I. Bloch, Editors. 2007, W.B. Saunders: Philadelphia. p. xiii.
59. Cramer, G.D., S.A. Darby, and G.D. Cramer, *Clinical anatomy of the spine, spinal cord, and ANS*. 3rd ed. 2014, St. Louis, Mo.: Elsevier. xv, 672 p.
60. Jang, D. and S. Park, *A morphometric study of the lumbar interspinous space in 100 stanford university medical center patients.* J Korean Neurosurg Soc, 2014. **55**(5): p. 261-6.
61. Masood, S., et al. *Real-Time Sign Language Gesture (Word) Recognition from Video Sequences Using CNN and RNN*. 2018. Singapore: Springer Singapore.
62. Aghli, N. and E. Ribeiro. *Combining Weight Pruning and Knowledge Distillation for CNN Compression*. in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2021.
63. Joshi, G. and S. McCarroll, *Evaluation of combined spinal-epidural anesthesia using two different techniques.* Regional Anesthesia: The Journal of Neural Blockade in Obstetrics, Surgery, & Pain Control, 1994. **19**(3): p. 169-174.
64. Magill, J.C., et al., *A novel actuator for simulation of epidural anesthesia and other needle insertion procedures.* Simulation in healthcare: journal of the Society for Simulation in Healthcare, 2010. **5**(3): p. 179.
65. Hansen, J.T., F.H. Netter, and C.A.G. Machado, *Netter's clinical anatomy*, in *Netter basic science*. 2019, Elsevier,: Philadelphia, PA. p. 1 online resource (xviii, 555 pages).

66. Wang, C., et al., *Deep-learning-aided forward optical coherence tomography endoscope for percutaneous nephrostomy guidance.* Biomedical Optics Express, 2021. **12**(4): p. 2404-2418.
67. Tang, Q.G., et al., *Real-time epidural anesthesia guidance using optical coherence tomography needle probe.* Quantitative Imaging in Medicine and Surgery, 2015. **5**(1): p. 118-124.
68. Jung, J.C., et al., *In Vivo Mammalian Brain Imaging Using One- and Two-Photon Fluorescence Microendoscopy.* Journal of Neurophysiology, 2004. **92**(5): p. 3121-3133.
69. Manchikanti, L. and S. Atluri, *chapter 152 - Lumbar Epidural Nerve Block*, in *Pain Management*, S.D. Waldman and J.I. Bloch, Editors. 2007, W.B. Saunders: Philadelphia. p. 1281-1293.
70. Wang, C., et al., *Endoscopic OCT for epidural anesthesia*. 2021: Dataset on Zenodo.
71. Chollet, F., *Keras*. 2015.
72. Iizuka, N., et al., *Oligonucleotide microarray for prediction of early intrahepatic recurrence of hepatocellular carcinoma after curative resection.* The Lancet, 2003. **361**(9361): p. 923-929.
73. Krstajic, D., et al., *Cross-validation pitfalls when selecting and assessing regression and classification models.* Journal of Cheminformatics, 2014. **6**(1): p. 10.
74. Wang, C., et al., *Deep-learning-aided forward optical coherence tomography endoscope for percutaneous nephrostomy guidance.* Biomedical optics express, 2021. **12**(4): p. 2404-2418.
75. Selvaraju, R.R., et al. *Grad-cam: Visual explanations from deep networks via gradient-based localization*. in *Proceedings of the IEEE international conference on computer vision*. 2017.
76. Chollet, F., *Deep learning with Python*. Vol. 361. 2018: Manning New York.
77. Badre, A. and C. Pan, *LINA: A Linearizing Neural Network Architecture for Accurate First-Order and Second-Order Interpretations.* IEEE Access, 2022. **10**: p. 36166-36176.
78. Xia, P., J. Hu, and Y. Peng, *EMG-based estimation of limb movement using deep learning with recurrent convolutional neural networks.* Artificial organs, 2018. **42**(5): p. E67-E77.

CHAPTER 7: ENHANCING EPIDURAL NEEDLE GUIDANCE USING A POLARIZATION-SENSITIVE OPTICAL COHERENCE TOMOGRAPHY PROBE WITH CONVOLUTIONAL NEURAL NETWORK

This chapter is reproduced from "Chen Wang, Paul Calle, Justin C. Reynolds, Nu Bao Tran Ton, Feng Yan, Anthony M. Donaldson, Avery D. Ladymon, Pamela R. Roberts, Alberto J. De Armendi, Kar-Ming Fung, Chongle Pan, Qinggong Tang, *Epidural anesthesia needle guidance by forward endoscopic optical coherence tomography and ensemble deep learning*, Sci. Rep., 2022" and "Chen Wang, Yunlong Liu, Paul Calle, Xinwei Li, Ronghao Liu, Qinghao Zhang, Feng Yan, Kar-ming Fung, Andrew K. Conner, Sixia Chen, Chongle Pan, Qinggong Tang, *Enhancing Epidural Needle Guidance using a Polarization-Sensitive Optical Coherence Tomography Probe with Convolutional Neural Networks*, **J. Biophotonics**, 2023."

7.1 Introduction

Epidural anesthesia is a versatile and transformative anesthetic technique, which has improved pain relief and heightened patient satisfaction. The procedure finds application across the cervical, thoracic, and lumbar positions for a variety of operations, including as cardiovascular surgery [1], stomach surgery [2], painless delivery [3]. It helps relieve perioperative stress responses and symptoms. Epidural anesthesia involves the insertion of a needle into the epidural space, followed by the injection of local anesthetic around the spine. This procedure effectively blocks the transmission of pain signals from the spinal nerves to the brain. The epidural needle insertion generally aims to the vertebral spinous spaces or paramedian vertebral interspace [4]. During the insertion, the needle passes through subcutaneous fat, ligaments, and ligamentum flavum and then arrives at the epidural space where the spinal cord is located. Ensuring the precise placement of the epidural needle is critically important to achieve both efficacy and safety in the administration of epidural

anesthesia. Any error in needle placement can lead to consequential complications. Post-dural puncture headache (PDPH) is a frequently observed iatrogenic complication arising from inadvertent injury to the dura mater, with the overall incidence rate from 6% to 36% [5]. Epidural anesthesia is widely used in delivery when the estrogen secretion of the pregnant dramatically increases. Elevated estrogen levels can impact the tone of cerebral vessels, potentially leading to an increased vascular distension response to cerebrospinal fluid (CSF) hypotension. As a consequence, pregnant mothers are at a heightened risk for PDPH [6]. Additionally, if the epidural needle is misplaced, it may lead to the occurrence of epidural hematomas, spinal cord damage, and abscess [7].

To sense the position of the epidural needle tip during its insertion, anesthesiologists generally employ a loss of resistance (LOR) technique. This involves continuously injecting saline or air into the needle via a syringe and monitoring the resulting feedback force. By doing so, they can sense the occurrence of the LOR, which indicates the successful traversal of various spinal tissues and the arrival of the needle to the epidural space [8]. However, LOR remains challenging in clinical practice as it lacks visual feedback. Needle insertion guided solely by LOR can lead to failure rate of up to 20% [9]. To improve the precision of needle guidance in the administration of epidural anesthesia, various imaging techniques have been utilized. Notably, ultrasound has become a widely employed imaging modality for needle guidance during epidural anesthesia. [10]. Real-time ultrasound has been demonstrated to increase the epidural needle placement accuracy [11]. Fluoroscopy has been employed to guide the epidural needle placement, particularly in non-obstetrical situations [12]. It was shown to enhance the success rate of needle insertion in cases involving adhesion and stenosis [13]. However, the resolutions of ultrasound and fluoroscopy are usually at millimeter level.

Thus, ultrasound and fluoroscopy encounter challenges when attempting to discern the epidural space, which exhibits a width spanning from 2 to 3 mm in the cervical region and from 5 to 6 mm in the lumbar region. [14, 15]. Furthermore, acoustic window of ultrasound is limited by the complex back bones. Fluoroscopy lacks soft tissue contrast, making it difficult to distinguish different tissues [9]. More recently, novel techniques such as fiberoptic bundle or optical spectral analysis have been reported [16, 17], but the movement of surrounding fluids or tissues in front of the imaging sites is likely to impact their imaging results.

We demonstrated the use of optical coherence tomography (OCT) for accurate identification of the epidural needle position by taking the advantages of its micrometer-level resolution in tissue visualization [18]. In our previous studies, we designed a forward-view swept-source OCT (SS-OCT) endoscopic system to visualize the tissues in front of the epidural needle [9, 19, 20]. To help anesthesiologists interpret the real-time imaging during the operation, we developed deep-learning models to automate the identification of tissue types from the OCT images [19]. The overall classification accuracy for identifying the correct tissue out of five tissue types was only approximately 66% [19]. To improve the tissue recognition accuracy, a series of binary classification models were developed by assuming a fixed sequence of spinal tissues to be encountered during the needle insertion. The average accuracy of 96.65% was achieved by the binary model, while the tissue was assumed always manifest in a specific sequence corresponding to the epidural anatomy. The situations can be complicated as tissue distribution is not always ideal, hindering the clinical translation of the binary model.

In this study, we developed a polarization-sensitive OCT (PS-OCT) probe to improve the tissue recognition accuracy for epidural needle placement. PS-OCT represents an

advancement over traditional intensity OCT as it enables the acquisition of polarization-related features through birefringence and diattenuation optical signals [21]. Because tissues such as nerve fibers exhibit strong polarization contrast signals, these additional polarization features may enhance the accuracy of spinal tissue recognition. The PS-OCT used in this study was equipped with much higher axial resolution (~5.5μm in air) compared to the previous SS-OCT system (~14μm in air). Here, we developed a PS-OCT endoscope specifically designed to be fit inside the clinically used epidural needle and tested it on porcine backbone samples. The performance of PS-OCT for distinguishing the porcine backbone tissues were benchmarked in this research.

7.2 Method

7.2.1 Experimental setup

The experimental setup of our PS-OCT probe is shown in Figure 7.1. It is a single-input device with a unique polarization-sensitive detector unit, enabling simultaneous acquisition of two orthogonal polarization states at full imaging speed. A linear polarized fiber laser operating at a center wavelength of 1,300 nm with a bandwidth of 170 nm served as the light source. The laser light was initially directed through a polarization-maintaining (PM) circulator, ensuring its transmission in a specific direction to facilitate the routing of laser signals among various components. Upon traversing the circulator, the light was collimated within the space and divided into a reference arm and a sample arm using a beam splitter. The light directed into the sample arm was illuminated to the tissue, while the light routed into the reference arm served as the reference beam. A quarter-wave plate (QWP) rotated by 22.5° orientation was put in the reference arm and exited with a 45° linear polarization after passing through the QWP twice. Another QWP rotated by 45° was put in the sample arm to obtain

circularly polarized light, which made the system sensitive to all polarization effects from the sample and independent to the orientation of the sample in plane. Polarization-maintaining fiber was employed to ensure polarization preservation across the entire system. The two lights, which were either backscattered or reflected from the sample arm and reflected within the reference arm, converged at the beam splitter before re-entering the collimator. Subsequently, the resulting interference light was transmitted to the optical subsystem for post-processing. The optical sub-system split the interference light into the vertical linearly polarized signal and horizontal linearly polarized signal by two polarization-sensitive beam splitters (PBS), which are further detected and processed by two polarization-sensitive channel sensors [22-24]. The two individual OCT images from the unique polarization-sensitive detector unit can be shown separately or combined in a total intensity image.

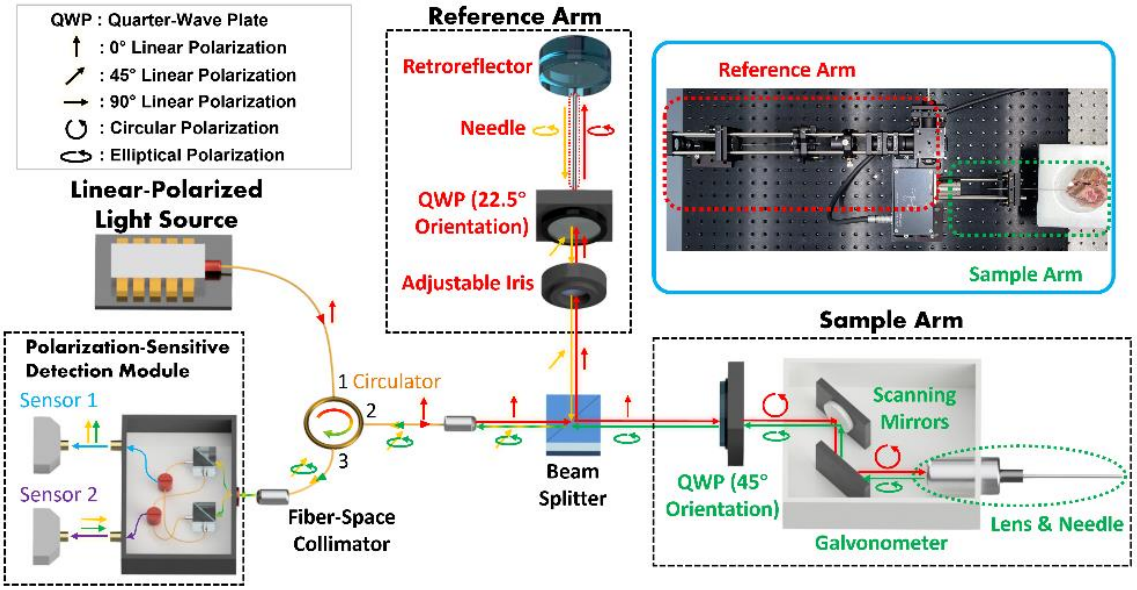


Figure 7.1. Experimental setup of the PS-OCT probe.

In this system, we integrated a gradient index (GRIN) lens into the sample arm to attain a forward view of the endoscope, as depicted in the bottom right of Figure 7.1. We specifically selected a GRIN lens with a diameter of 1.3 mm, closely resembling the

dimensions of a typical epidural needle. The GRIN lens measures 138.0 mm in length and offers a viewing angle of 11.0°. The proximal end of this GRIN lens is positioned near the focal plane of the scanner lens to ensure sufficient light power passes through the GRIN lens body. To counteract potential dispersion, which can affect image quality, we introduced an additional GRIN lens with identical specifications into the light path of the reference arm. This served as a compensatory element, guaranteeing precise and accurate imaging outcomes from the system. The axial imaging resolution of the PS-OCT system can reach up to 5.5μm in air, allowing us to capture subtle structures within the samples. The scanning rate ranges from 5.5 to 76.0 kHz, and the sensitivity is 109 dB (at 5.5 kHz). The GRIN lenses are from GoFoton Corporation, and the laser and all other the optical components are from Thorlabs Inc.

In addition to conventional intensity OCT imaging, PS-OCT offers supplementary tissue contrast derived from tissue birefringence characteristics. As illustrated in Figure 7.1, our PS-OCT endoscope incorporates two sensors designed to capture light with distinct polarization states. Here we consider the individual intensity of sensor 1 (receiving the vertical linear polarized light) as I_1 , and the sensor 2 (receiving the horizontal linear polarized light) as I_2. OCT intensity results are calculated with the total intensity of $I = I_1 + I_2$, where $I_1 = E_1 \times E_1^*$ and $I_2 = E_2 \times E_2^*$ (E represents the complex amplitude of the electric field). In our system, PS imaging results include phase retardation, optic axis, and degree of polarization uniformity (DOPU) [25]. Stokes parameters can be calculated as [25]:

$$Q = E_1 E_1^* - E_2 E_2^* \tag{1}$$

$$U = E_{+45^o} E_{+45^o}^* - E_{-45^o} E_{-45^o}^* = E_1 E_2^* + E_2 E_1^* \tag{2}$$

$$V = E_R E_R^* - E_L E_L^* = i\left(E_x E_y^* - E_y E_x^*\right) \tag{3}$$

The three PS-OCT imaging modes can be calculated following the formula below:

$$Phase\ retardation = tan^{-1}\left(\sqrt{I_1/I_2}\right) \in [0, \pi/2] \quad (4)$$

$$Optic\ Axis = \frac{1}{2} atan2(U, Q) \in [-\pi/2, \pi/2] \quad (5)$$

$$DOPU = \sqrt{\langle Q \rangle^2 + \langle U \rangle^2 + \langle V \rangle^2} \in [0, 1] \quad (6)$$

The incorporation of PS modes resulted in enhanced tissue contrast within the epidural samples. The assessment of phase retardation was utilized to quantitatively measure the accumulated phase change experienced by light within the tissue across the two distinct channels. Likewise, the assessment of the optic axis measured the relative angular orientation between the fast and slow axes. It is noteworthy that in instances where the sample tissue displayed a uniform birefringence property, both phase retardation and optic axis values exhibited a proportional increase with depth. Furthermore, DOPU metric facilitated a novel perspective. DOPU involved the assessment of average Stokes parameters, which in turn correlated with the extent of incident light depolarization within the sample. This comprehensive utilization of PS modes contributed to an enriched understanding of tissue properties and interactions. [25].

7.2.2 Deep learning procedure

Six porcine backbone samples were imaged using our PS-OCT probe. We scanned five different epidural tissue layers: 1) subcutaneous fat; 2) interspinous ligament; 3) ligamentum flavum; 4) epidural space; and 5) spinal cord, as shown in Figure 7.2. The image size was set at 430 × 950 pixels with a pixel size of 3 μm. For each backbone sample, we obtained one thousand images in each imaging mode from every tissue type, resulting in a total of 24,000 images (1,000 images × 4 imaging modes × 6 samples). CNN was employed for the

classification of tissue layers, enabling the automation of the epidural needle tip positioning procedure. Specifically, the ResNet50 architecture from the Keras library was utilized for this purpose [26]. The SGD optimizer was set with 50 epochs, with the learning rate of 0.01, momentum of 0.9, decay of 0.01. The batch size was set to be 24. We employed cross-validation to optimize the number of training epochs, while retaining the model architecture and all other hyperparameters consistent with our prior studies [19, 27, 28]. We used nested cross-validation and cross-testing to evaluate the prediction results [19, 27, 28]. Images were separated into six folds based on the imaging results. The dataset encompassed images from all six samples. These configurations were maintained consistently throughout all training sessions. This uniformity facilitated the comparison of performance among different imaging modes and ensured an unbiased assessment of classification capabilities. The cross-validation procedure encompassed 30 folds for each OCT imaging mode, resulting in a total of 120 folds. In addition, we conducted cross-testing with 6 folds. The computation time for each fold was approximately 130 minutes, using one NVIDIA RTX 3090 graphic card.

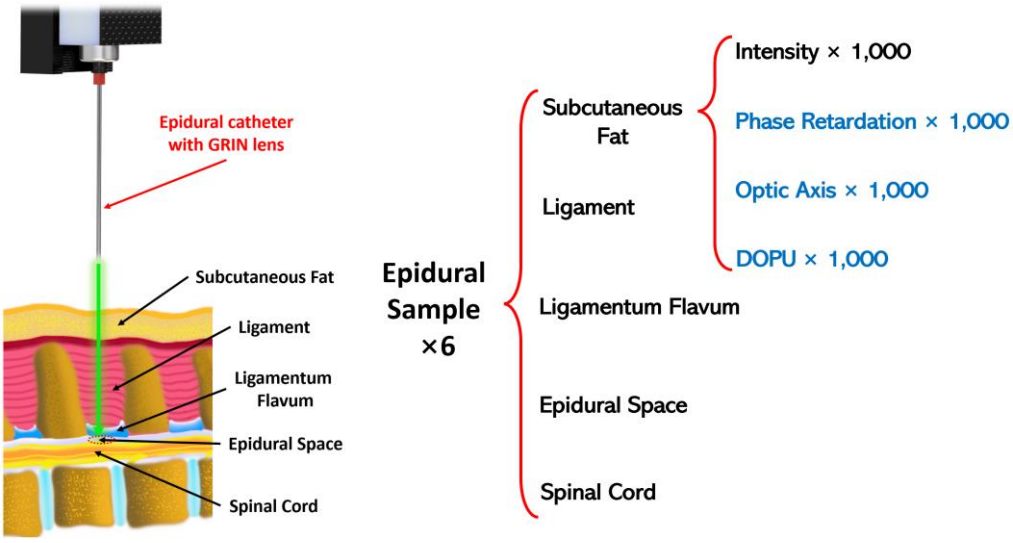


Figure 7.2: Data acquisition of the epidural tissues with the PS-OCT probe.

We used three parameters: 1) Precision; 2) Recall; and 3) F_1 to evaluate the prediction performance of our model. These metrics are widely used in deep learning to assess classification performance and the capabilities of the modes. Precision quantifies the accuracy level, recall indicates the proportion of relevant retrieved instances, and the F_1 score represents the harmonic mean between precision and recall. These metrics are computed as follows (where TP stands for true positive, FP for false positive, and FN for false negative):

$$Precision = \frac{TP}{TP+FP} \tag{7}$$

$$Recall = \frac{TP}{TP+FN} \tag{8}$$

$$F_1 = 2 \cdot \frac{Precision \cdot Recall}{Precision+Recall} = \frac{TP}{TP+\frac{1}{2}(FP+FN)} \tag{9}$$

7.3 Results

7.3.1 PS-OCT imaging results

PS-OCT images of the five epidural tissues in the four OCT modes and corresponding histology were shown in Figure 7.3. Regarding the OCT intensity outcomes, the five tissues exhibited distinct patterns in the OCT imaging results, which align with findings from our prior studies [9, 19]. Upon reaching the epidural space with the needle, a noticeable gap became evident between a prominent line (indicated by yellow double arrows in Figure 7.3) that aligned with the surface of the GRIN lens and the dura mater of the spinal cord. The tissues that the needle passed through during insertion were also distinguishable. The subcutaneous fat exhibited low imaging intensity and depth, characterized by an irregular distribution of bright bars or dots representing adipocytes. In contrast, the ligament tissue displayed the greatest imaging depth among all the tissues. Additionally, the ligament displayed horizontal lines attributed to the fibrous nature, mirroring the characteristics

observed in the histology results. The flavum exhibited remarkably high brightness in the OCT intensity output, a consequence of its dense elastic fibers. Notably, when the needle penetrated the spinal cord, the tissue region appeared exceedingly bright, indicative of the densely packed nerve fibers.

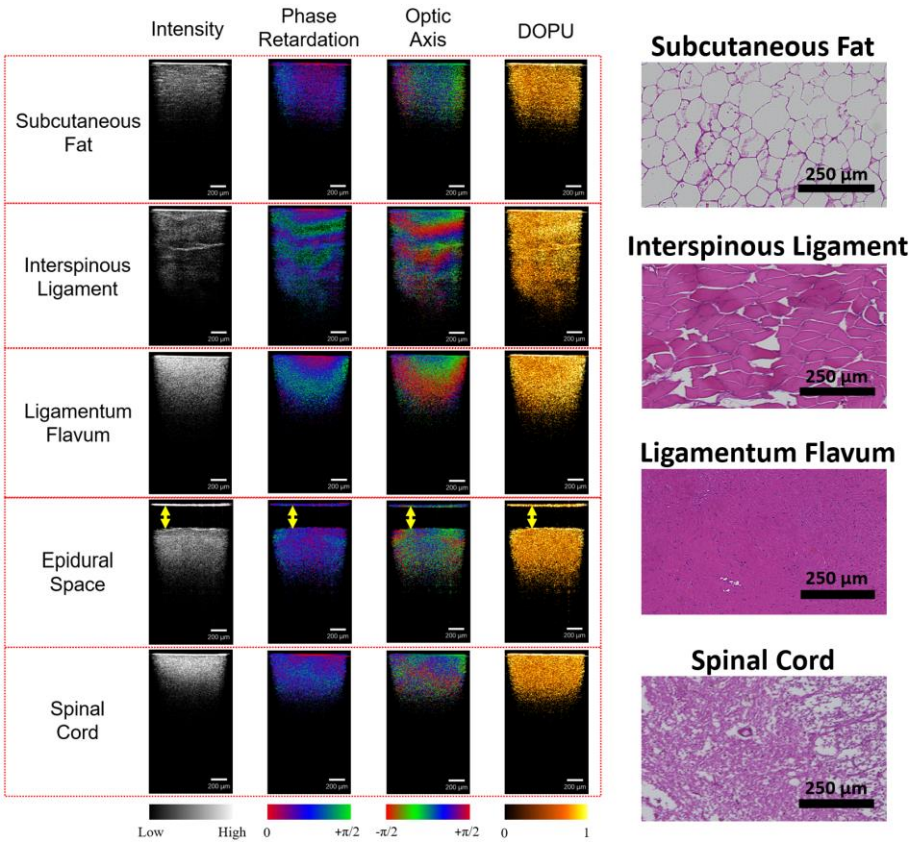


Figure 7.3. PS-OCT imaging results of epidural tissues and corresponding histology.

In the PS imaging modes depicted in Figure 7.3, fiber-based tissues are notably discernible, particularly in the phase retardation and optic axis modes. Clear transverse layers were evident within the interspinous ligament and ligamentum flavum tissues, albeit exhibiting varying layer thicknesses and depths. While the spinal cord also displayed transverse layer characteristics, they were not as pronounced as those observed in the ligament and flavum tissues, which consist predominantly of elastin and collagen fibers [29, 30]. Conversely, in the case of subcutaneous fat tissue, which lacks a fiber structure, there

existed lengthwise polarization contrast, a phenomenon previously documented [31]. This observation suggests a promising avenue of incorporating polarization-sensitive contrast into intensity results, with the potential to enhance the recognition of epidural tissues.

7.3.2 Prediction results by deep learning

The classification of the five epidural tissues was conducted using the ResNet50 architecture based on its strong performance demonstrated in our prior studies [19]. For this research, a total of 1,000 cross-sectional images were captured from each tissue layer within every mode. In total, six backbone samples were utilized. Here we evaluated the classification of the five epidural groups in the four modes, respectively. The aggregated accuracy for both cross-validation and cross-testing outcomes is shown in Table 7.1. Notably, the intensity mode exhibited weaker predictive performance than the other polarization-sensitive (PS) modes. Collectively, DOPU was the most effective among the three polarization-sensitive (PS) modes. In comparison to the OCT intensity mode, which achieved an accuracy of 81.60%, all three PS-OCT modes exhibited superior average predictive capabilities. Specifically, phase retardation attained an accuracy of 87.39%, optic axis yielded 85.97%, and DOPU outperformed with the highest result of 91.61%. Cross-testing accuracy was subsequently computed for each of the six testing folds. Like the cross-validation findings, DOPU exhibited the highest accuracy at 91.53%, while the OCT intensity mode recorded the lowest accuracy of 81.89%. Interestingly, in contrast to the cross-validation outcomes, the optic axis mode surpassed the performance of the phase retardation mode. Considerable variability in testing accuracy was observed across different subjects, particularly pronounced in the case of Subject 1. Subject 1 displayed the lowest performance across all four modes. This variability could be attributed to differences in tissue characteristics among various backbone samples.

Furthermore, the subjects achieving the highest performance varied across different modes in the testing results. Specifically, Subject 4 provided the best performance in the intensity and phase retardation modes, while Subject 5 generated the optimal results in the optic axis and DOPU modes. This variability in performance across different imaging modes suggests a potential complementary effect that can be harnessed.

Grad-CAM heatmaps were utilized to visually show the explanations of the ResNet50 model classification as shown in Figure 7.4. A conspicuous disparity between the PS modes and the intensity mode lay in their distinct focal areas. In the case of the intensity mode, the model predominantly emphasized the luminous surface of the needle itself. In the three PS modes, the model exhibited a discernible shift in focus, prioritizing the tissues situated beneath the epidural space. An advantage of PS modes is that the cumulative change increases in the phase retardation and optic axis as the depth increases. The value for DOPU is not affected by depth change since it measures the Stokes parameters average level around each pixel. While in OCT intensity mode, the signal decreases as the depth increases due to light scattering. In other words, the signal from the deep region provides less useful information in the OCT intensity mode while less affected in the PS modes. This is clearly shown in the heatmap for ligaments, which only covered the shallower part of the intensity image but spanned all the depth in the PS modes.

Table 7.1: Accuracies for cross-validation and cross-testing results in different imaging modes.

Cross-Validation	Intensity	Retardation	Optic Axis	DOPU
E1	84.72%	92.00%	90.80%	94.01%
E2	82.48%	84.21%	84.93%	91.18%
E3	84.62%	87.78%	82.14%	91.74%
E4	75.42%	85.95%	85.70%	92.61%
E5	80.17%	84.74%	84.11%	89.34%
E6	82.22%	89.64%	88.13%	90.80%
Average	81.60±3.47%	87.39±3.02%	85.97±3.07%	91.61±1.60%

Cross-Testing	Intensity	Retardation	Optic Axis	DOPU
E1	66.34%	58.16%	76.84%	75.86%
E2	81.70%	81.82%	84.46%	87.48%
E3	73.06%	89.48%	91.42%	90.88%
E4	94.16%	99.54%	92.56%	98.52%
E5	93.58%	96.66%	93.94%	97.28%
E6	82.52%	90.38%	86.06%	99.18%
Average	81.89±11.02%	86.01±14.97%	87.55±6.44%	91.53±8.98%

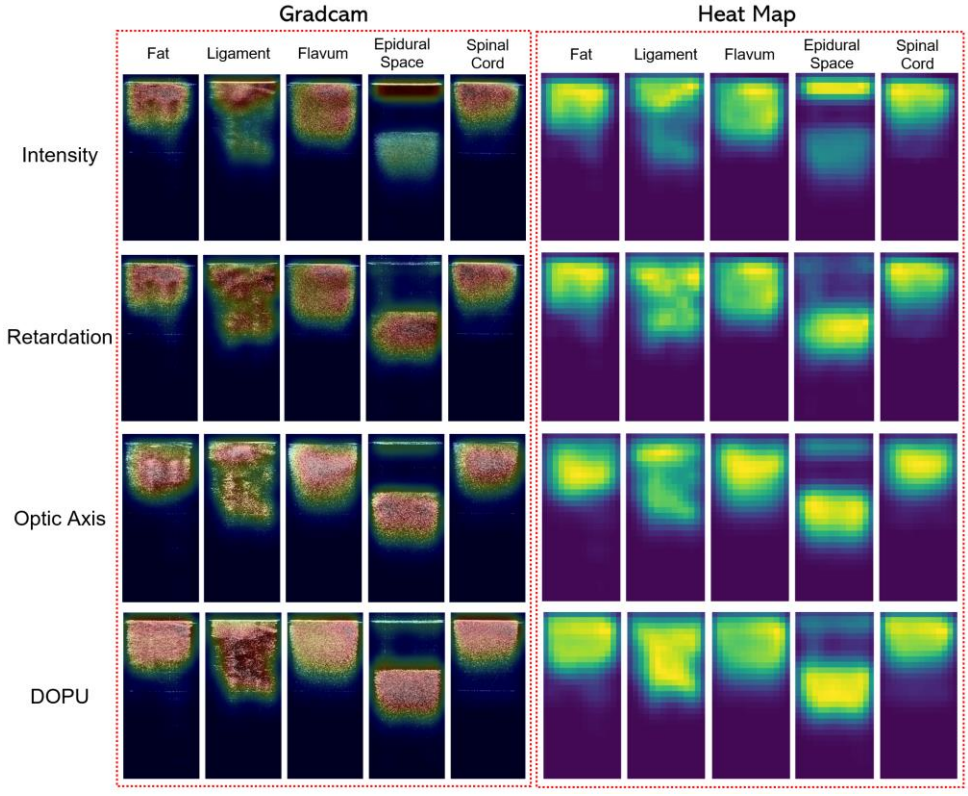


Figure 7.4. Grad-CAM and heatmap examples of different imaging results in four modes.

Table 7.2: Precision, Recall, and F1 of different tissue types in each model (Red font for the highest value and green for the lowest value).

Precision	Fat	Ligament	Flavum	Epidural Space	Spinal Cord
Intensity	81.91±1.28%	86.20±4.92%	76.44±5.91%	99.89±2.52%	79.97±4.72%
Retardation	86.73±3.76%	91.31±3.81%	87.19±3.13%	99.97±2.31%	84.13±3.58%
Optic Axis	87.59±3.86%	90.36±3.71%	84.13±3.87%	97.72±1.96%	83.02±2.33%
DOPU	94.80±2.76%	96.03±4.01%	87.11±3.32%	99.69±0.32%	88.94±1.65%

Recall	Fat	Ligament	Flavum	Epidural Space	Spinal Cord
Intensity	92.85±3.05%	82.17±2.54%	64.10±3.77%	94.38±0.11%	74.52±2.81%
Retardation	83.78±3.52%	86.63±2.25%	82.94±2.57%	95.35±0.02%	88.26±2.82%
Optic Axis	83.50±1.66%	87.20±2.31%	74.41±2.90%	94.73±1.15%	90.00±2.74%
DOPU	88.73±0.98%	86.27±0.75%	88.57±3.56%	99.46±0.31%	95.03±2.91%

F_1	Fat	Ligament	Flavum	Epidural Space	Spinal Cord
Intensity	85.82±1.93%	80.17±3.96%	64.07±4.37%	96.47±1.66%	73.38±3.44%
Retardation	82.90±3.39%	87.65±3.17%	84.13±2.54%	97.07±1.56%	83.97±3.28%
Optic Axis	83.88±2.60%	86.40±2.57%	76.13±2.79%	95.71±1.31%	84.74±1.72%
DOPU	90.85±1.87%	89.22±3.00%	86.87±3.22%	99.56±0.23%	90.70±1.87%

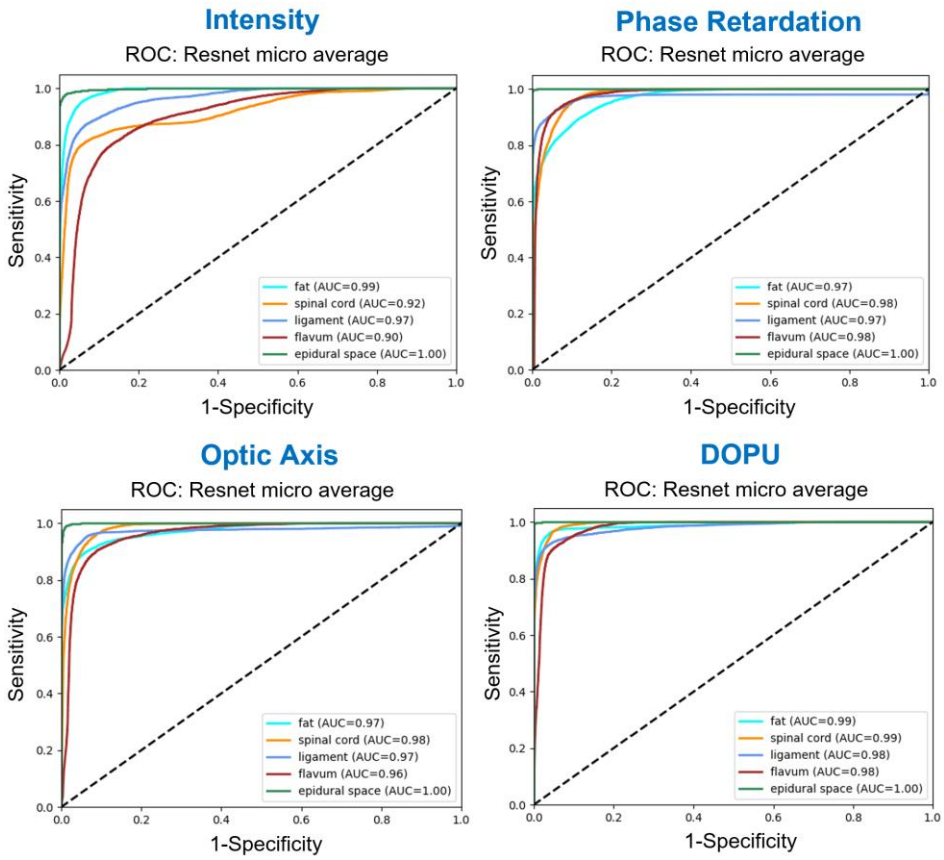


Figure 7.5. ROC curves of different imaging modes.

Table 7.2 displays the classification outcomes for various tissue types, while the associated receiver operating characteristic (ROC) curves are illustrated in Figure 7.5. From the averaged values for each tissue layer, epidural space was the most easily recognizable due to the gap between needle tip and the tissue surface and DOPU presented the best F_1 score of 99.56. Given that the epidural space is the target destination for the epidural needle, accurate identification of this site ensures precise needle placement. Compared to other tissues, flavum is most difficult to identify but the results in retardation and DOPU are significantly improved (>84%) compared to 64% from intensity mode. Fat, ligament, and spinal cord all had over 80% accuracy results in the three PS modes with the highest F_1 score (>89%) from DOPU mode. Therefore, in comparison to conventional structural OCT imaging, PS-OCT imaging can provide better classification accuracy using more tissue information.

7.4 Discussion

In this study, we assessed the performance of a forward-view PS-OCT probe for guiding epidural anesthesia needle placement. Our innovative system offers significantly improved resolution compared to conventional imaging techniques such as ultrasound and fluoroscopy, providing a clearer front view of the epidural needle. By employing PS-OCT, our system can acquire both intensity and PS signals and achieved superior axial resolution of 5.5μm, outperforming our previous research on SS-OCT with an axial resolution of ~10μm. The high-resolution imaging results offered intricate details, allowing medical professionals to discern the epidural tissue type and accurately determine the needle tip location, which explained the significant increase in cross-validation accuracy from 66% using SS-OCT to 81.60% using intensity mode alone. Furthermore, the imaging results obtained in PS modes revealed distinct features of the five epidural tissue layers, enabling more effective differentiation between them compared to intensity mode alone. This enhanced understanding of tissue characteristics could aid in optimizing the needle guidance procedure for epidural anesthesia.

Besides the intensity and penetration depth contrast in the intensity OCT mode, PS modes enhanced the feature of the tissues with fiber structures such as ligament, flavum, and spinal cord. We conducted a thorough assessment of our models' generalization performance using a nested cross-validation and cross-testing procedure. This allowed us to gain valuable insights into their robustness and potential for clinical applications. The cross-testing results showed an excellent average accuracy of 91.53% in DOPU mode, with phase retardation and optic axis also yielding accuracies at 86.01% and 87.55%, respectively. In contrast, intensity modes only exhibited a prediction accuracy of 81.89% even with the improved imaging

resolution. Further evaluation using precision, recall, and F1 scores for each tissue type revealed that the epidural space was consistently the easiest tissue layer to recognize in both PS modes and intensity mode. This finding is particularly promising as the epidural needle needs to be placed inside the epidural space, making its recognition a critical step. The recognition accuracies were also improved for the other tissues in all the three PS modes. In general, the three PS modes provided better predictions than intensity, and DOPU yielded the highest F_1 score across all five epidural tissue types. Importantly, the diverse imaging modes demonstrated unique capabilities in identifying specific tissues. By integrating the strengths of each mode, we anticipate significant improvements in tissue classification accuracy. The inference time on NVIDIA RTX 3090 GPUs was 0.0314s per image and 0.1256s for all the 4 imaging modes on average, thus real-time prediction using all four modes can still be achieved in our system. When multiple GPUs are available, the inference time can be further reduced by parallelizing the inferencing across GPUs.

The endoscopic PS-OCT system developed in this study was based on a single input polarization state, thus the polarization parameters (phase retardation, optic axis, and DOPU) were all cumulative values [32]. The cumulative phase retardation only indicated the phase retardation between the principal polarization states along the complete optical path through the tissue rather than the phase retardation effect at a single depth location [33, 34]. Therefore, the ‘local’ polarization information was not able to be quantitatively provided. On the other hand, due to the direct proportionality between the local birefringence mode and the actual birefringent signal per pixel in the OCT images, the local polarization signal suffers from diminished sensitivity [35, 36]. Compared to the local polarization mode, the cumulative polarization mode exhibited higher sensitivity and demanded a considerably simplified

system, so it remained the preferred choice for image classification purposes [35, 37, 38]. Considering the key to the epidural guidance application is the differentiation between different spinal tissues, cumulative polarization parameters were applied in this study. Additionally, we plan to upgrade our system to enhance its usability and convenience in clinical settings. To validate the clinical translatability of our PS-OCT probe, we will continue to improve our technology using *in-vivo* pig models. Furthermore, considering the anatomical variations and tissue properties differences between pig and human epidural tissues, we will continue this project on human samples in the future. To improve the automatic tissue prediction in our future work, we will combine PS-OCT and intensity OCT as an integrated input into a multi-channel CNN [39] or an ensemble learning system [40]. By leveraging the strengths of individual modes for identifying specific tissues, this integrated approach is expected to achieve an enhanced overall tissue recognition accuracy, albeit with the trade-off of increased computational costs and model complexity. Furthermore, by employing various model interpretation methods [41, 42], we can gain insights into which imaging features and modalities contribute most to accurate classifications of different tissues. This will help us further optimize our OCT imaging technique and enhance its diagnostic capabilities. By pursuing these steps, we aim to advance the capabilities of our PS-OCT probe and deep learning model, ultimately contributing to improved epidural anesthesia needle guidance procedures and enhancing clinical outcomes.

Reference

1. Svircevic, V., et al., *Meta-analysis of thoracic epidural anesthesia versus general anesthesia for cardiac surgery.* The Journal of the American Society of Anesthesiologists, 2011. **114**(2): p. 271-282.
2. Hayashi, H., et al., *Prospective randomized study of open versus laparoscopy-assisted distal gastrectomy with extraperigastric lymph node dissection for early*

gastric cancer. Surgical Endoscopy And Other Interventional Techniques, 2005. **19**: p. 1172-1176.
3. Moraca, R.J., D.G. Sheldon, and R.C. Thirlby, *The role of epidural anesthesia and analgesia in surgical practice*. Annals of Surgery, 2003. **238**(5): p. 663-673.
4. Hernandez, A.N.A. and P. Singh, *Epidural Anesthesia*, in *StatPearls [Internet]*. 2022, StatPearls Publishing.
5. Patel, R., et al., *A comprehensive update on the treatment and management of postdural puncture headache*. Current pain and headache reports, 2020. **24**: p. 1-9.
6. Chekol, B., T. Yetneberk, and D. Teshome, *Prevalence and associated factors of post dural puncture headache among parturients who underwent cesarean section with spinal anesthesia: A systemic review and meta-analysis, 2021*. Annals of Medicine and Surgery, 2021. **66**: p. 102456.
7. Tripathi, M., S.S. Nath, and R.K. Gupta, *Paraplegia after intracord injection during attempted epidural steroid injection in an awake-patient*. Anesth Analg, 2005. **101**(4): p. 1209-11, table of contents.
8. Frumin, M.J., et al., *Sites of sensory blockade during segmental spinal and segmental peridural anesthesia in man*. Anesthesiology, 1953. **14**(6): p. 576-83.
9. Tang, Q., et al., *Real-time epidural anesthesia guidance using optical coherence tomography needle probe*. Quant Imaging Med Surg, 2015. **5(1)**: p. 118-124.
10. Kang, S.Y., et al., *Advantages of the combination of conscious sedation epidural anesthesia under fluoroscopy guidance in lumbar spine surgery*. Journal of pain research, 2020. **13**: p. 211.
11. Elsharkawy, H., et al., *Real-time ultrasound-guided lumbar epidural with transverse interlaminar view: evaluation of an in-plane technique*. Pain Medicine, 2019. **20**(9): p. 1750-1755.
12. Yeager, M., et al., *Fluoroscopy-assisted epidural catheter placement: an exploratory analysis of 303 pre-operative epidurograms*. Acta Anaesthesiologica Scandinavica, 2016. **60**(4): p. 513-519.
13. Kang, S.Y., et al., *The Comparison of Fluoroscopy-Guided Epidural Anesthesia with Conscious Sedation and General Anesthesia for Endoscopic Lumbar Decompression Surgery: A Retrospective Analysis*. World Neurosurgery, 2022. **159**: p. e103-e112.
14. Richardson, J. and G.J. Groen, *Applied epidural anatomy*. Continuing Education in Anaesthesia, Critical Care & Pain, 2005. **5**(3): p. 98-100.
15. Waldman, S.D., *Pain management E-book*. 2011: Elsevier Health Sciences.
16. Lin, S.P., et al., *Discriminant analysis for anaesthetic decision-making: an intelligent recognition system for epidural needle insertion*. British Journal of Anaesthesia, 2012. **108**(2): p. 302-307.
17. Anderson, T.A., et al., *Raman Spectroscopy Differentiates Each Tissue from the Skin to the Spinal Cord: A Novel Method for Epidural Needle Placement?* Anesthesiology, 2016. **125**(4): p. 793-804.

18. Huang, D., et al., *Optical Coherence Tomography.* Science, 1991. **254**(5035): p. 1178-1181.
19. Wang, C., et al., *Epidural anesthesia needle guidance by forward-view endoscopic optical coherence tomography and deep learning.* Sci Rep, 2022. **12**(1): p. 9057.
20. Wang, C., et al., *Endoscopic optical coherence tomography for epidural anesthesia guidance (Conference Presentation)*. SPIE BiOS. Vol. PC11949. 2022: SPIE.
21. Hee, M.R., et al., *Polarization-sensitive low-coherence reflectometer for birefringence characterization and ranging.* JOSA B, 1992. **9**(6): p. 903-908.
22. de Boer, J.F., et al., *Two-dimensional birefringence imaging in biological tissue by polarization-sensitive optical coherence tomography.* Optics Letters, 1997. **22**(12): p. 934-936.
23. Hitzenberger, C.K., et al., *Measurement and imaging of birefringence and optic axis orientation by phase resolved polarization sensitive optical coherence tomography.* Optics express, 2001. **9**(13): p. 780-790.
24. Roth, J.E., et al., *Simplified method for polarization-sensitive optical coherence tomography.* Optics Letters, 2001. **26**(14): p. 1069-1071.
25. de Boer, J.F., C.K. Hitzenberger, and Y. Yasuno, *Polarization sensitive optical coherence tomography - a review [Invited].* Biomedical Optics Express, 2017. **8**(3): p. 1838-1873.
26. He, K., et al. *Deep residual learning for image recognition*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.
27. Wang, C., et al., *Computer-aided Veress needle guidance using endoscopic optical coherence tomography and convolutional neural networks.* J Biophotonics, 2022. **15**(5): p. e202100347.
28. Wang, C., et al., *Deep-learning-aided forward optical coherence tomography endoscope for percutaneous nephrostomy guidance.* Biomed Opt Express, 2021. **12**(4): p. 2404-2418.
29. Hirabayashi, S., *Ossification of the ligamentum flavum.* Spine Surg Relat Res, 2017. **1**(4): p. 158-163.
30. Iwanaga, J., et al., *Anatomical and Biomechanical Study of the Lumbar Interspinous Ligament.* Asian J Neurosurg, 2019. **14**(4): p. 1203-1206.
31. Ding, Z., et al., *Imaging Spinal Structures With Polarization-Sensitive Optical Coherence Tomography.* IEEE Photonics Journal, 2016. **8**(5): p. 1-8.
32. Ding, Z., C.-P. Liang, and Y. Chen, *Technology developments and biomedical applications of polarization-sensitive optical coherence tomography.* Frontiers of Optoelectronics, 2015. **8**(2): p. 128-140.
33. Gräfe, M.G., et al., *Subretinal Fibrosis Detection using Polarization Sensitive Optical Coherence Tomography.* Translational Vision Science & Technology, 2020. **9**(4): p. 13-13.

34. Braaf, B., et al., *Fiber-based polarization-sensitive OCT of the human retina with correction of system polarization distortions.* Biomedical optics express, 2014. **5**(8): p. 2736-2758.
35. Feroldi, F., et al., *In vivo multifunctional optical coherence tomography at the periphery of the lungs.* Biomedical Optics Express, 2019. **10**(6): p. 3070-3091.
36. Villiger, M., et al., *Deep tissue volume imaging of birefringence through fibre-optic needle probes for the delineation of breast tumour.* Scientific Reports, 2016. **6**.
37. Willemse, J., et al., *Optic axis uniformity as a metric to improve the contrast of birefringent structures and analyze the retinal nerve fiber layer in polarization-sensitive optical coherence tomography.* Optics Letters, 2019. **44**(15): p. 3893-3896.
38. Grafe, M.G.O., et al., *Subretinal Fibrosis Detection Using Polarization Sensitive Optical Coherence Tomography.* Transl Vis Sci Technol, 2020. **9**(4): p. 13.
39. Zhao, M.Y., et al., *Cloud Shape Classification System Based on Multi-Channel CNN and Improved FDM.* Ieee Access, 2020. **8**: p. 44111-44124.
40. Kumar, A., et al., *An Ensemble of Fine-Tuned Convolutional Neural Networks for Medical Image Classification.* Ieee Journal of Biomedical and Health Informatics, 2017. **21**(1): p. 31-40.
41. Badre, A. and C.L. Pan, *LINA: A Linearizing Neural Network Architecture for Accurate First-Order and Second-Order Interpretations.* Ieee Access, 2022. **10**: p. 36166-36176.
42. Badre, A. and C.L. Pan, *Explainable multi-task learning improves the parallel estimation of polygenic risk scores for many diseases through shared genetic basis.* Plos Computational Biology, 2023. **19**(7).

CHAPTER 8: CONCLUSION AND FUTURE WORK

In this dissertation, we have explored the feasibility of using conventional structural OCT and PS-OCT to establish various endoscopic imaging navigation systems. We have demonstrated the applicability across a range of medical imaging procedures, including percutaneous nephrostomy (PCN), Veress needle guidance, bladder and renal cancer diagnosis, and epidural anesthesia surgical guidance. Our systems are capable of recognizing different tissue types in real-time during these procedures. In addition, we have integrated convolutional neural network (CNN) methods to automate the tissue recognition process. The tissue identification accuracies in different medical procedures are overall promising, demonstrating the potential for practical implementation.

In the future, we will continue our current work in following aspects:

1. Based on our previous research, we intend to initially implement endoscopic PS-OCT system in PCN, Veress needle, and epidural anesthesia. Currently we have already obtained imaging results using PS-OCT endoscopic OCT probe on human epidural samples. Relative results are shown as follows:

2. We will further make our endoscopic OCT platform more portable and convenient to use in clinical situations. We intend to translate our system to be a product which can be adopted in clinical settings in the future.

3. We plan to improve our CNN models to achieve higher accuracy and efficiency in tissue recognition. We will apply more compact models to reduce the training and inference time. Moreover, we will validate our model in more samples to further enhance the model applicability.