

DEVELOPMENT OF A RESEARCH TESTBED FOR
INTRAOPERATIVE OPTICAL SPECTROSCOPY TUMOUR
MARGIN ASSESSMENT

by

DAVID J. MORTON

A thesis submitted to the
Department of Electrical and Computer Engineering
in conformity with the requirements for
the degree of Master of Applied Science

Queen's University
Kingston, Ontario, Canada

October 2023

Copyright © David J. Morton, 2023

Abstract

Cancer is a prevalent group of diseases, resulting in millions of deaths globally each year. For early-stage cancer, surgical intervention is a primary treatment option. However, this often results in incomplete tumour resection, necessitating a secondary revision surgery. The difficulty of intraoperative tumour margin assessment is a factor contributing to high rates of revision surgery. Despite remarkable improvements in perioperative care due to advances in medical technology, the intraoperative assessment of tumour margins remains limited. The objective of this thesis is the development and evaluation of a tracked tissue sensing testbed prototype for navigated tumour margin assessment. Two principal contributions are presented in this thesis.

The first contribution focuses on the design, implementation, and characterization of a testbed for real-time tissue assessment. This testbed employs diffuse reflection broadband optical spectroscopy for tissue characterization and electromagnetic tracking for navigation. The tissue sensor and a k-nearest-neighbour classifier are used to identify tissue types, while spatial tracking enables the predictions to be overlaid on the scanned tissue. The testbed was characterized using an ex vivo tissue phantom to evaluate its accuracy and real-time usability. The results indicate that our testbed is capable of accurate real-time navigated tissue inspection. The second contribution focuses on the translation of the testbed for clinically relevant tissue interrogation.

Ex vivo human kidney tissue was interrogated using the testbed, and a modelling pipeline was developed for the classification of cancerous and non-neoplastic tissue using linear discriminant analysis. The results indicate that the testbed is viable for rapid application to novel tissues and offers promising tissue classification capabilities.

The ability to assess tumour margin status intraoperatively has the potential to increase surgeon confidence in complete tumour resection, reducing the rates of revision surgeries. Future work will focus on the development of testbed support for combining multiple tissue sensing modalities and experimenting with state-of-the-art tissue classification networks to improve testbed capabilities. Accurate intraoperative tumour margin assessment tools are crucial in improving outcomes for early-stage cancer patients worldwide.

Acknowledgments

First and foremost, I would like to express my gratitude to my supervisors, Dr. Parvin Mousavi and Dr. Gabor Fichtinger. The opportunity to contribute to such a diverse and intellectually stimulating research group has been deeply rewarding. The mentorship and opportunities provided throughout my Master's degree have contributed significantly to both my personal growth and growth as a researcher. Dr. Mousavi, your guidance and succinct feedback throughout my Master's degree has been invaluable. Your instruction in CISC 881 introduced me to the field of medical informatics, revealing the remarkable potential of AI in addressing real-world medical challenges. Dr. Fichtinger, your mentorship and teachings in CISC 330 first captivated my interest in computer-assisted interventions. You were a true inspiration in my pursuit of a Master's degree in the Perk and Med-i labs.

In addition, I would like to extend my sincere appreciation to Dr. Amoon Jamzad and Dr. Tamas Ungi for their pivotal roles throughout my Master's degree. Dr. Jamzad, you were a true pillar of support. You always knew the right questions to ask and never hesitated to offer your knowledge and insight when I needed it most. Your joyful personality is contagious, I couldn't have done it without you. Dr. Ungi, I am grateful for your grounded perspective and for your readiness to help at a moment's notice. For the countless hours you spent guiding me through

the intricacies of 3D Slicer software development, dropping everything to help me through a seemingly endless supply of bugs. My love for software development truly blossomed under your guidance.

Furthermore, I wish to acknowledge each and every one of my lab mates who passed through the doors of the Perk and Med-i labs. Our countless hours of conversation and deliberation about every topic under the sun will be dearly missed. More broadly, I would like to express my deep appreciation for every teacher, coach, mentor, and friend who has played a part in shaping me into the person I am today. I am eternally grateful.

Finally, I would like to thank my family for their unwavering love and support throughout my life. Mom and Dad, thank you for always being there to lift me up in my times of need and to celebrate my times of triumph. Ali, thank you for being an exceptional role model and the best big sister I could have asked for. You all mean the world to me.

Contents

Abstract	i
Acknowledgments	iii
Chapter 1: Introduction	1
1.1 Motivation	1
1.2 Problem	3
1.3 Objective	4
1.4 Contributions	5
1.5 Organization of Thesis	6
Chapter 2: Background	8
2.1 Surgical treatment of cancer	8
2.2 Intraoperative tumour margin assessment techniques	10
2.2.1 Histopathology	11
2.2.2 Medical imaging	12
2.2.3 Mass spectrometry	13
2.2.4 Spectroscopy	15
2.3 Tracked diffuse reflection broadband optical spectroscopy	16
2.3.1 Optical tissue interrogation	16

2.3.2	Spatial tracking	17
2.4	System integration with open-source software	19
Chapter 3:	Development and characterization of research testbed	21
3.1	Design and implementation	21
3.1.1	An overview of testbed design	22
3.1.2	Hardware components and specifications	24
3.1.3	Software development and integration	27
3.2	Exploring spectroscope sensitivity to environmental parameters . . .	34
3.2.1	Probe-to-surface distance	35
3.2.2	Probe-to-surface angle	36
3.3	Characterizing the testbed performance	37
3.3.1	Experimental setup	37
3.3.2	Characterization experimentation	38
3.4	Results of spectroscope sensitivity analysis	41
3.4.1	Effects of probe-to-surface distance	41
3.4.2	Effects of probe-to-surface angle	44
3.5	Real-time navigation display characterization	45
3.5.1	Latency characterization	45
3.5.2	Spatial classification characterization	46
Chapter 4:	Application of research testbed for interrogation of hu-	
	man kidney tissue	49
4.1	Collection and analysis of ex vivo human kidney tissue	49
4.1.1	Data collection	50

4.1.2	Data distribution analysis	53
4.2	Tissue classification experiments	53
4.2.1	Modelling pipeline	54
4.2.2	Preprocessing experimentation	58
4.3	Results of ex vivo human kidney tissue analysis and experiments . . .	60
4.3.1	Data distribution analysis	60
4.3.2	Preprocessing experiments	64
Chapter 5:	Summary and Conclusions	71
5.1	Summary	71
5.2	Limitations	72
5.3	Future work	73
References		76
Appendix A:		87
A.1	Standard operative procedures	88
A.2	PLUS configuration file	89

List of Tables

3.1	The testbed latencies for the navigation (EMT) and classifier portions of the display visualization. The mean and standard deviation is reported in ms and is calculated over 20 trials.	45
3.2	The Dice similarity coefficient and accuracy metrics computed from our spatial classification visualization [61].	47
4.1	Example of an imported and formatted dataset	55
4.2	The results of our preprocessing experimentation using a scan-based cross-validation regimen. The mean balanced accuracy, sensitivity, specificity, F1, AUC, and accuracy are reported.	65
4.3	The results of our preprocessing experimentation using a sample-based cross-validation regimen. The mean balanced accuracy, sensitivity, specificity, F1, AUC, and accuracy are reported.	66

List of Figures

2.1	Perioperative workflow for the surgical treatment of early-stage cancer	9
2.2	Illustration of excised tumour cross-section: Negative margin (left) vs. Positive margin (right)	10
2.3	Technological landscape overview for intraoperative margin assessment	11
2.4	An example EM tracking system, the NDI Ascension Trakstar [60] . .	18
3.1	A diagram of the tracked tissue sensing testbed [61]	23
3.2	An overview of the hardware setup of the testbed	24
3.3	Theoretical spectrometer grating efficiency [63] (Left) and theoretical broadband light source output intensity [62] (Right)	25
3.4	The user interface of the developed software module. (Left) The graph- ical user interface, (Centre) the 3D navigation display, and (Right) the spectrum viewer	28
3.5	An overview of the real-time hardware communication pipeline	29
3.6	The developed data collection interface for our open-source software. (Left) The graphical user interface (Right) A real-time visualization of the optical spectra	30

3.7	Illustration of the 3D navigation display interface. Red and green points represent classes 0 and 1 respectively, while the blue model represents the optical probe.	32
3.8	Experimental setup for sensitivity analysis experiments	34
3.9	Experimental setup for the probe-to-surface distance experiments . .	35
3.10	Experimental setup for the probe-to-surface angle experiments shown at $\sim 90^\circ$, 70° , and 45°	36
3.11	A visual of the testbed latency characterization experiment	38
3.12	An illustration of the biological tissue phantom design used in this experiment	39
3.13	Experimental workspace for the characterization of testbed performance using an ex vivo biological tissue phantom	40
3.14	Resulting spectra from the probe-to-surface distance experiments. Spectra recorded closest to the tissue (Left), furthest from the tissue (Right), and an intermediary distance (Centre)	41
3.15	An example of ambient light in the room reflected off a white piece of paper	42
3.16	The average (red) and maximum (blue) signal differences as a function of measured peak intensity	43
3.17	The average (red) and maximum (blue) signal differences as a function of probe angle	44
3.18	The resultant spatial classification visualization for our ex vivo tissue phantom [61]	46

3.19	A confusion matrix for the spatial classification of bovine tissue (class 0) and porcine tissue (class 1) [61]	47
4.1	Images of the resected ex vivo human kidney tissue specimens	50
4.2	The experimental setup for the interrogation of tissue specimens . . .	51
4.3	Approximate regions of cancerous (red) non-neoplastic (green) and ambiguous (white) tissue determined visually by a trained pathologist .	52
4.4	Modelling pipeline for tissue classification	54
4.5	Visualization of a raw broadband spectrum	55
4.6	Light source compensation. (Left) The light source intensity curve is in black and the reflected intensity curve is in blue (Right) The compensated intensity curve is in blue)	56
4.7	PCA latent space visualization of metadata. Pathology colourized (Left) and Sample colourized (right)	61
4.8	PCA latent space visualization as a function of time. Cancer (Left) and Non-neoplastic (right)	63
4.9	Leave-one-sample-out confusion matrix. Non-neoplastic tissue is class 0 and cancerous tissue is class 1	67
A.1	Plus configuration script used to integrate the spectrometer and electromagnetic tracker into the 3D slicer testbed	89

Chapter 1

Introduction

1.1 Motivation

Cancer represents a pervasive and complex group of diseases that have profound effects on both individuals and society. With millions of annual cancer-related deaths, its consequences are far-reaching, leaving countless more without a close friend or loved one [1]. Significant progress in medical technology has been driven by the exponential growth in computational power and computer miniaturization. This revolution has empowered clinicians with unprecedented patient-specific data, leading to remarkable advances in medical informatics and preoperative diagnostics for enhanced surgical planning. However, translating these advancements into practical intraoperative applications poses substantial challenges. Intraoperative technologies must deliver rapid results within minutes rather than days while requiring additional trained personnel to prepare tissue, operate the device, and interpret the results. This urgency demands seamless hardware integration, user-friendly interfaces, and efficient algorithms to ensure accurate and timely outcomes in the dynamic surgical environment.

The development of deployable intraoperative technologies is a time-consuming and interdisciplinary process. It demands clinical expertise to assess the clinical relevance of potential solutions, software proficiency to design interfaces and algorithms to provide clinically useful information, and comprehensive hardware knowledge to effectively construct and calibrate a real-time system. The development of open-source medical software such as 3D Slicer [2], SlicerIGT [3], OpenIGTLink [4], and PLUS [5], has fostered a collaborative environment and eased the burden on individual research groups. PLUS and OpenIGTLink have lowered the barrier for hardware knowledge, facilitating the open distribution of hardware interfacing capabilities. Similarly, 3D Slicer has contributed to reducing the threshold for software development by enabling the sharing of software building blocks among researchers and developers. SlicerIGT has been crucial in enabling the integration of these open-source packages to seamlessly create deployable hardware systems with software interfaces.

The expanding reach of the internet has helped to allow communities of open-source collaborators to flourish [6]. Open-source medical software and their communities have the potential to open the doors to clinical translation and experimentation at unprecedented speed [7]. The reduction of required expertise allows researchers to work more closely and more efficiently, allowing for rapid ideation, exploration, and experimentation of solutions to clinically relevant problems. Developing a framework that enables easy experimentation with new ideas while minimizing technical requirements for producing and characterizing novel systems holds the promise of significantly enhancing the development of intraoperative technologies. The reduction in time from desktop to operating table could have profound impacts on patient outcomes.

1.2 Problem

Early-stage cancer diagnosis and treatment play a crucial role in improving patient outcomes and reducing mortality [8]. Surgical intervention is a common approach for treating early-stage cancers [9]. However, the visual distinction between healthy and cancerous tissue can be challenging, necessitating reliance on preoperative planning and imaging for interventional guidance. Unfortunately, these static images and models are limited in their ability to adapt to the dynamic nature of the surgical environment. For example, in breast cancer interventions the breast medium is highly deformable resulting in high tumour mobility [10]. This results in reduced registration accuracy between the preoperative tumour state and the actual tumour state, leading to errors in tumour boundaries [11]. Thus, cancerous tissue is often missed during the initial surgery [12]. This disparity between pre-operative and intraoperative knowledge contributes to a substantial rate of revision surgeries, exposing patients to additional risks, psychological distress, and increased financial burden [13].

The development of accurate and effective intraoperative margin assessment tools is paramount to help rectify this gap. While various tissue detection technologies show promise, there is currently no effective, accurate, and real-time standard of care for intraoperative tumour margin assessment. The process of developing novel tools for intraoperative margin assessment is interdisciplinary and time-consuming. Often requiring the development of a new software interface specific to each tool. The technical requirements and protracted development process hamper the timely translation of innovative ideas into practical applications. The development of a modular research testbed for intraoperative margin assessment serves as a crucial initial stride in expediting the design, implementation, experimentation, and translation of novel

solutions. By facilitating seamless hardware and software integration and offering an interface for clinical experimentation, this testbed could help minimize the barriers to entry for researchers. Such a testbed would enable rapid prototyping and evaluation of novel technologies, bridging the technical gap between clinical ideation, implementation, and deployment of intraoperative solutions.

1.3 Objective

The objective of this thesis was to develop and evaluate a prototype of a tracked tissue sensing testbed for navigated tumour margin assessment. The development of the testbed facilitates the application of intraoperative margin assessment technologies in an effort to bridge the gap in intraoperative knowledge. Using available technologies for tissue sensing, navigation, and classification, I first designed and implemented an open-source research testbed for tracked tissue sensing. Diffuse reflection broadband optical spectroscopy was employed for tissue sensing, while electromagnetic tracking was used for navigation. The classifiers that were experimented with included Linear Discriminant Analysis (LDA) and k-Nearest-Neighbour (kNN). I devised experimental methods to determine the sensitivity of the tissue sensor to various external parameters during signal acquisition. I then developed an experimental pipeline to characterize the performance of the testbed for real-time operation on an ex vivo biological tissue phantom. Finally, I applied the testbed for the interrogation of ex vivo human kidney tissue. I developed and experimented with a modelling pipeline for the classification of cancerous and non-neoplastic tissue. The testbed could empower users with intuitive and real-time margin status information. This capability enables surgeons to scan a tumour cavity post-resection, bolstering the surgeon's confidence

in the complete removal of the tumour.

1.4 Contributions

The objective of this thesis was addressed through the following contributions:

- Developed and characterized a tracked tissue sensing testbed for navigated margin assessment
 - Created a module in 3D Slicer to enable real-time data collection, classification, and visualization
 - Implemented a tracked tissue sensor using a diffuse reflection spectroscopy and an electromagnetic tracking system
 - Conducted a sensitivity analysis of the tracked tissue sensor to signal acquisition parameters
 - Evaluated our testbed using an experimental performance characterization pipeline
- Applied testbed to interrogate cancerous and non-neoplastic human kidney tissue ex vivo
 - Collected a dataset comprised of cancerous and non-neoplastic tissue using the testbed and a standard operating procedure.
 - Designed a machine-learning classification pipeline for the differentiation of cancerous and non-neoplastic tissue
 - Evaluated and analyzed various tissue properties and preprocessing methods using the classification pipeline

This work also appears in the following publication:

- D. Morton, L. Connolly, L. Groves, et al., “Tracked tissue sensing for tumor bed inspection,” in Medical Imaging 2023: Image-Guided Procedures, Robotic Interventions, and Modeling, C. A. Linte and J. H. Siewerdsen, Eds., International Society for Optics and Photonics, vol. 12466, SPIE, 2023, 124661K. doi:10.1117/12.2654217

1.5 Organization of Thesis

Chapter 2 begins with a background of the surgical treatment of cancer, detailing the challenges associated with intraoperative tumour margin assessment. It proceeds with an overview of the current technological landscape of intraoperative tumour margin assessment. It goes on to provide a background of the tissue sensor and tracker employed in this thesis, followed by a detailed review of the open-source software enabling our deployable tracked tissue sensor.

Chapter 3 outlines the first contribution: the development and characterization of our research testbed. It begins with a description of the design and implementation of our testbed for tracked tissue sensing. Two experiments are described, the first to assess the sensitivity of the tissue sensor during data collection, and the second to characterize the performance of the testbed during operation. The results of the experimentation are presented.

Chapter 4 outlines the second contribution: the application of the research testbed for the interrogation of human kidney tissue. It details the procedure used to collect data from the tissue samples, followed by an exploration of the distribution of the data with respect to its associated meta-data. This is followed by the description of a

machine-learning modelling pipeline and subsequent tissue classification experimentation. The results of the analysis and experimentation are presented.

Chapter 5 provides a summary of the key conclusions, limitations, and future directions for this research.

Chapter 2

Background

This chapter is organized into four sections. A background of the surgical treatment of early-stage cancers is provided (2.1), highlighting the importance of tumour margin assessment. This is followed by a technological review of current approaches to intraoperative tumour margin characterization (2.2), a background of the technologies employed in this thesis (2.3), and an overview of the open-source software landscape enabling the development of deployable tools (2.4).

2.1 Surgical treatment of cancer

Cancer is one of the most prevalent diseases worldwide [1]. Complications arising from cancer resulted in nearly 10 million deaths in 2020 which accounted for ~ 1 in 6 deaths globally [1]. Cancers are characterized by abnormal cell division resulting in uncontrolled growth [1] and are typically named for their organ or tissue of origin [14]. The surgical resection of cancerous tissue is a common treatment approach for patients with early-stage cancers [9]. Early-stage cancer typically refers to a localized tumour mass that has not metastasized to another area of the body [15]. Long-term surgical outcomes are linked to the achievement of complete tumour resection while

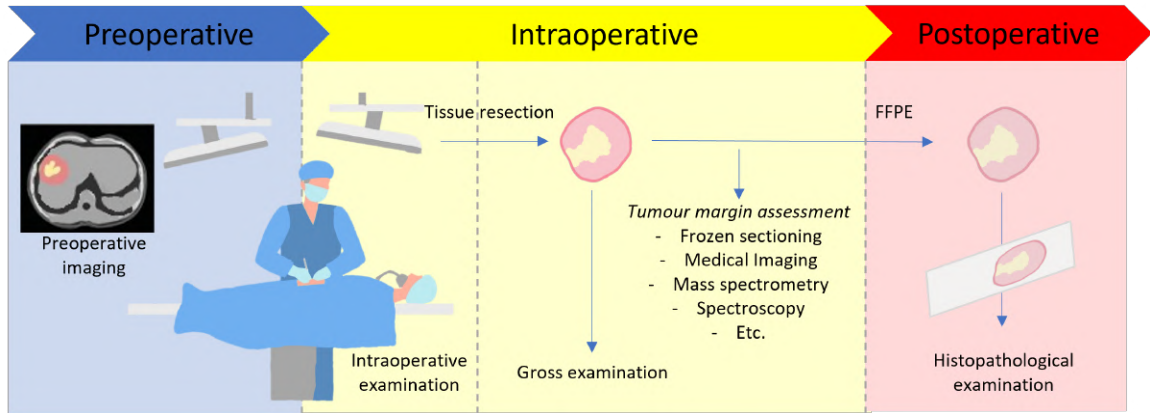


Figure 2.1: Perioperative workflow for the surgical treatment of early-stage cancer

maximally conserving healthy tissue [16]. Tissue-conserving surgery is a primary treatment option which aims to completely resect the cancerous tissue while resecting a minimal margin of healthy tissue surrounding the tumour [17]. The target size of this healthy tissue margin varies by procedure and is typically on the order of millimetres [18]. The presence or absence of cancerous tissue within the healthy tissue margin is known as the margin status and plays a significant role in surgical outcomes (Figure 2.2) [16]. A negative margin status indicates an intact healthy tissue margin and complete resection of the tumour. A positive margin status indicates cancerous tissue is present on the exterior of the excised tumour and is strongly associated with cancer recurrence [16].

Histopathological examination is the current gold standard for determining margin status post-surgery [19]. The excised tissue is fixed in formalin and stained to enhance cell contrast [20]. The tissue is prepared on slides, and microscopic examination is used to assess margin status at a cellular level [20]. If a positive margin is identified revision surgery is performed to remove the remaining cancerous tissue if feasible. Notably, this is not always possible when working around vital structures.

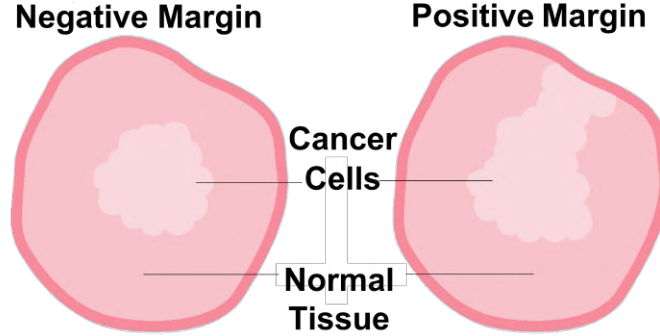


Figure 2.2: Illustration of excised tumour cross-section: Negative margin (left) vs. Positive margin (right)

The procedures with the highest rate of positive margins are ovarian, breast, prostate, thyroid, oral cavity and lung cancers [12]. Breast cancer is the most common surgically treated cancer in women globally [21]. The rate of revision surgery is estimated to be as high as 30% [22]. Revision surgeries pose several challenges and risks for patients, including an increased likelihood of postoperative complications, impaired cosmesis, heightened psychological distress, and increased healthcare costs [13].

2.2 Intraoperative tumour margin assessment techniques

The visual and tactile ambiguity of healthy and cancerous tissue is often a contributing factor to high rates of revision surgeries. Accurate intraoperative assessment of tumour margins plays an important role in reducing the chance of such revisions. An assortment of technologies show promise for intraoperative margin assessment, offering varying degrees of assessment speed and capability. These technologies employ techniques to characterize tissues based on properties such as chemical composition, acoustic interactions, and electromagnetic interactions. In this section, we provide

a review of the current methods for intraoperative tumour margin detection. We discuss the advantages and limitations of each technology, focusing on their practical applicability in the intraoperative setting.

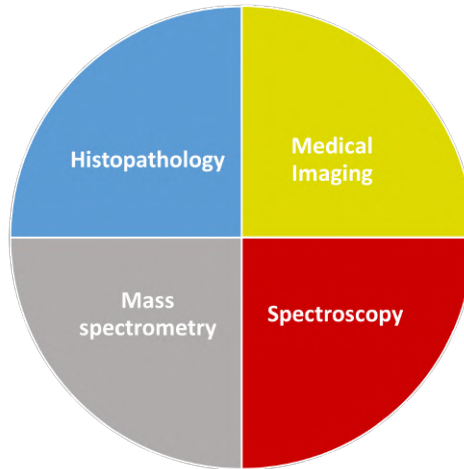


Figure 2.3: Technological landscape overview for intraoperative margin assessment

2.2.1 Histopathology

Histopathological examination is the gold standard for post-operative tumour margin assessment [19]. However standard histopathological examination requires substantial tissue preparation leading to multiple days before a margin status is determined [23]. Frozen sectioning and imprint cytology have been developed to hasten histopathological examination for use intraoperatively. Frozen sectioning is a technique that involves freezing portions of the resected tumour for rapid sectioning and staining of the specimen. This process typically takes around 30 minutes and provides an output similar to standard histopathology, although with some degradation [24]. However, it is impractical to examine the entire margin using this technique, and it requires

specialized equipment and staff to be available during the procedure [25]. Imprint cytology is another intraoperative histopathology technique. The resected tumour specimen is rubbed on a glass slide and viewed under a microscope to assess the presence of cancer. Imprint cytology relies on the principle that malignant cells will adhere to the glass [26]. This technique typically takes ~ 15 minutes and is most useful when a surgeon already has a high suspicion of a positive margin [27]. This technique also requires a specialized cytopathologist to be present.

Intraoperative histopathological examination techniques provide excellent capability for the accurate detection of margin status. However, these techniques are time-consuming and require the presence of specialized personnel to prepare and analyze the tissue during surgery.

2.2.2 Medical imaging

Medical imaging techniques are commonly used for preoperative planning and are increasingly used in intraoperative margin assessment [28]. We explore magnetic resonance imaging, micro CT, contrast-based imaging, and ultrasound-based techniques.

Magnetic resonance imaging is a noninvasive imaging technique offering excellent soft tissue differentiation capabilities [29]. However, MRI suffers from slow acquisition times, upwards of 50 minutes, and high costs reducing its intraoperative usability [30]. Micro CT provides high-resolution images capable of soft tissue differentiation. It has shown promise for enabling whole specimen margin assessment [31]. The acquisition time for micro CT is faster than MRI, typically taking ~ 15 minutes [31]. However, there are concerns about radiation exposure in addition to substantial acquisition

and interpretation times [32]. Contrast-based imaging techniques, such as fluorescence imaging and positron emission tomography (PET), involve the introduction of contrast agents prior to surgery. These techniques have shown great ability in detecting regions of residual cancer. However, additional time is required for the agents to manifest in the tumour and additional trained personnel are required to facilitate their admission [33]. Ultrasound (US) imaging techniques are highly prevalent in many interventions, such as prostate biopsy, for real-time intraoperative guidance. US-based techniques use acoustic waves to characterize tissues. Various techniques have been developed to use the US signal for tumour margin assessment. US elastography uses high-frequency acoustic waves to measure tissue stiffness and elasticity to infer tissue composition [34]. Temporally enhanced ultrasound (TeUS) is a signal-based technique that uses a multi-second time series of raw US data (RF) to characterize tissue [35]. Although US techniques show promise for intraoperative use, they require physical contact for signal acquisition and are not always suited for shallow depths of field. Physical contact can deform the tissue, increasing the difficulty of accurately registering the scan to the physical tissue volume or location in the tumour bed.

Overall, there are several medical imaging techniques which show promise for intraoperative margin assessment. However, many are burdened by limitations such as slow acquisition times, high cost, physical imaging constraints, and a requirement of trained personnel to interpret margin status.

2.2.3 Mass spectrometry

Mass spectrometry is a technology which uses the mass-to-charge ratio (m/z) of ions in a sample to chemically characterize tissue. They work under the principle that

cancerous and benign tissues carry a unique ion concentration profile. There are destructive and non-destructive techniques for signal acquisition, containing imaging and point-based variations.

Non-destructive mass spectrometry techniques typically involve the deposition of a medium to interrogate the ionic profile of a tissue surface. Imaging techniques such as MALDI (matrix-assisted laser desorption/ionization) [36], and DESI (desorption electrospray ionization) [37] use this concept to provide high-resolution readings in the m/z domain while preserving spatial information. However, these techniques require tissue to be prepared before scanning, contributing to increased total acquisition times. The MasSpecPen [38], and SpiderMass [39] offer handheld point-based designs, increasing acquisition speed at the cost of spatial information. Destructive mass spectrometry, such as REIMS (rapid evaporative ionization mass spectrometry) [40], rapidly vaporizes tissue using an electrocautery device. The ionic profile of the resulting smoke is used to characterize the tissue. The NaviKnife implements this technique in a point-based device to return the ionic profile intraoperatively. The delay of the NaviKnife readings is variable at ~ 2 seconds [41] and the signals have shown promise for differentiating tissue types.

Mass spectrometry-based techniques show promise for intraoperative margin assessment. However, the current state has some limitations. REIMS, for instance, is highly invasive and non-repeatable. Image-based techniques suffer from slow acquisition times and necessitate ex vivo tissue preparation, while point-based techniques suffer from non-locality due to variable time delays associated with ion travel.

2.2.4 Spectroscopy

Spectroscopy is a non-invasive technology which electromagnetically characterizes tissue by observing its interactions with various wavelengths of light. Tissues with different chemical compositions and physical structures will reflect, absorb, and transmit light differently, resulting in a unique electromagnetic profile for each tissue type. We review Raman spectroscopy, optical coherence tomography, radio frequency spectroscopy, bioimpedance spectroscopy, and diffuse reflection spectroscopy for use in intraoperative tumour margin assessment.

Raman spectroscopy is a technique that uses monochromatic light to probe the vibrational modes of molecules to characterize tissue composition. It takes advantage of the Raman effect, which involves the inelastic scattering of photons and the resulting frequency shift from molecular interactions [42]. Raman spectroscopy has shown promise as a noninvasive method for tissue differentiation [43]. However, Raman spectroscopy can suffer from slow acquisition times, requiring up to 10 seconds per point depending on the device [44]. Furthermore, Raman requires an experienced operator to be present during surgeries [45]. Optical coherence tomography (OCT) is a technique for cross-sectional tissue imaging. It functions on principles similar to ultrasound, producing an image based on light reflected from various depths. OCT uses a band of near-infrared light to image subsurface tissue structures up to 2 mm below the surface [46]. OCT is capable of real-time imaging and has shown promise in intraoperative tumour margin detection [47]. However, in practice, OCT has required up to 4 minutes to scan a 5 cm³ region, [47] and requires trained personnel to interpret images during surgery [48]. Radiofrequency spectroscopy is a technique which interrogates tissue in the radio wavelength band. The Margin Probe is a handheld, point-based

deployment of this technique [49]. A promising reduction in the re-excision rates of breast cancer has been observed with its use [50]. The acquisition time is ~ 1.5 seconds per point [51]. Bioimpedance spectroscopy is another interesting technique that characterizes tissue using its electrical impedance. Bioimpedance spectroscopy is sensitive to intracellular and extracellular changes in tissue based on alterations in dielectric properties. This technique has been employed by the ClearEdge [52] handheld probe and has shown the ability to discriminate between human breast carcinoma and benign tumours [53]. This device takes ~ 5 minutes to characterize an excised tumour specimen [52]. Diffuse reflection spectroscopy is a technique which uses the diffuse reflection of tissue at various wavelengths to characterize tissue. The characteristic reflection of tissue is dependent on the electromagnetic properties of the tissue. The reflection intensity of the scattered light is measured at various wavelengths, creating an optical fingerprint that can be used to differentiate between tissues [54].

Overall, many spectroscopy techniques offer non-invasive and repeatable tissue sensing. There are many examples of point-based probes with rapid acquisition times and promising tumour margin assessment capabilities. However many of these point-based methods do not leverage spatial location information. We see promise in the development of a tracked tissue sensing testbed for the rapid testing and deployment of intraoperative techniques.

2.3 Tracked diffuse reflection broadband optical spectroscopy

2.3.1 Optical tissue interrogation

The non-invasive, and repeatable nature of diffuse reflection broadband optical spectroscopy makes it an ideal candidate for both in vitro and ex vivo margin assessment.

Thus, we use this technique as our primary focus during testbed development. The diffuse reflection broadband optical spectroscopy technique relies on 3 main components: a broadband light source, an optical fibre reflection probe, and a spectrometer. A broadband light source is used to illuminate a tissue sample in a desired wavelength band. In optical spectroscopy, this is a subset of the optical spectrum, composed of the visible, near-ultraviolet and near-infrared bands. The optical fibre reflection probe is used to transmit the radiated light to the tissue surface. When the light interacts with the tissue, the reflected intensity at each wavelength is dependent on the electromagnetic properties of the tissue. In the case of cancerous and benign tissues, each will reflect and absorb light in varying proportions, depending on the wavelength. The light is collected by the optical fibre probe and transmitted to the spectrometer. The spectrometer is used to measure the intensity of the reflected light at each wavelength. Frequency-specific diffraction gratings are used to separate the light into component wavelengths [55]. Each wavelength is directed to an individual detector where the intensity is measured. The result is a broadband spectrum which electromagnetically characterizes the tissue.

2.3.2 Spatial tracking

Spatial tracking can be used in conjunction with point-based tissue sensing to allow for tissue characterization with spatial registration. Spatial tracking systems allow position and orientation data to be introduced into the system [56]. For application in surgery, sensors or markers are typically rigidly fixed to various surgical tools and the patient [57], [58]. The two main spatial tracking systems used are electromagnetic tracking and optical tracking [57]. These two technologies each offer differing

strengths and weaknesses. Optical tracking is characterized by superior tracking accuracy, however, it is limited to a direct line of sight between the tracker and the optical markers [57]. Electromagnetic tracking is not constrained by the line of sight and offers good accuracy. However, it is susceptible to metallic and ferromagnetic materials in the workspace [59]. We focus on the use of electromagnetic tracking.

An electromagnetic (EM) tracking system is composed of an EM field generator, EM sensors, and a controller (Figure 2.4). The generator contains coils of wire which produce time-varying EM fields [57]. The EM tracking sensors typically contain small coils which interact with the magnetic field. The variable change in magnetic flux flowing through the coils induces a current which is returned to the controller [57]. The magnitudes of the currents are directly related to the sensor position and orientation. The controller aids in field generation, acquires sensor data, and calculates the position and orientation of the EM sensor [57]. This allows the pose of each tool in the workspace to be tracked relative to a reference coordinate system.

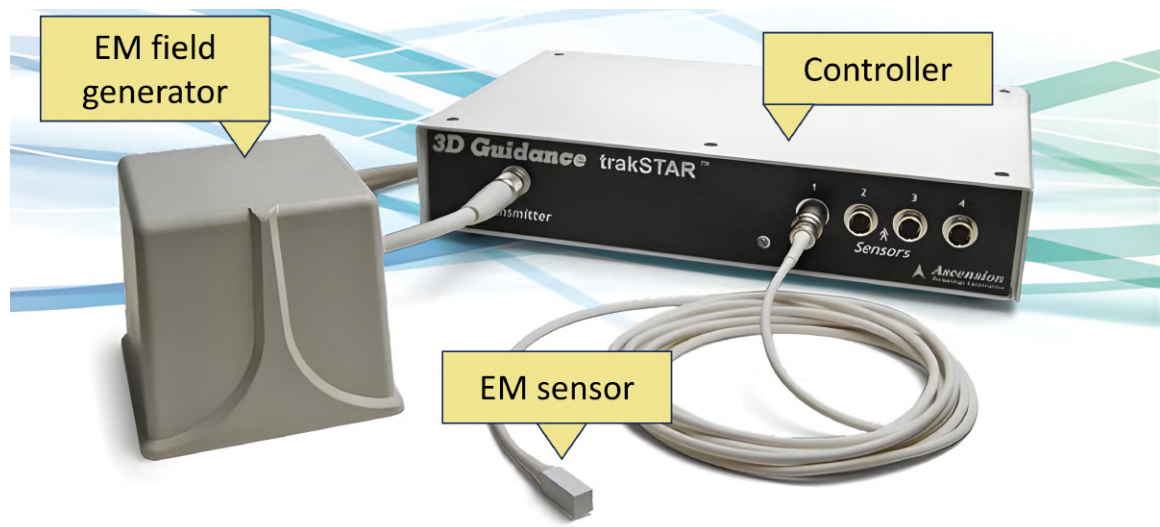


Figure 2.4: An example EM tracking system, the NDI Ascension Trakstar [60]

2.4 System integration with open-source software

For a tracked tissue sensor to be used effectively in the operating theatre, it necessitates the development of a deployable system. The basic requirement of a deployable system is to extract sensor information and communicate the underlying domain knowledge to the user in an intuitive and clinically useful manner. Some common forms of communication include the production of auditory cues, visualizations, and tactile feedback. The production of such a deployable system requires real-time hardware communication to acquire relevant information, data processing to extract domain knowledge, and an intuitive and easily interpretable interface to communicate with the clinician.

We focus on the use of open-source software and communication standards such as 3D Slicer, SlicerIGT, OpenIGTLink, and PLUS [2]–[5]. 3D Slicer is an open-source software which was created for the development of surgical navigation software [2]. Since its conception, it has blossomed into a multifaceted software development platform with a worldwide community of contributors. 3D Slicer is central to the development of deployable systems due to its modular and extendable design. It handles the tissue sensor data processing, navigation, and visualization to produce intuitive feedback for the clinician. SlicerIGT is a 3D Slicer extension which offers a suite of tools and modules to assist with the navigation, visualization, and manipulation of medical images [3]. It enables tool and patient registration to a common coordinate frame, and real-time updating of tracked images and models. OpenIGTLink is an open-source communication protocol and interface for image-guided interventions[4]. OpenIGTLink facilitates communication between hardware from different vendors provided they are interfaced with the OpenIGTLink specification. PLUS (Public

software Library for Ultrasound) is an open-source platform that facilitates the development of computer-assisted intervention systems [5]. It does this by integrating real-time imaging, sensing, and navigation components with data processing and visualization software. PLUS supports efficient data acquisition from a wide range of devices and leverages OpenIGTLink to effectively make hardware information accessible to 3D Slicer. Using a standardized script, PLUS provides additional features such as synchronizing hardware, computing tracking transforms between tools, and adjusting data acquisition parameters. PLUS enables our system to be agnostic to the chosen hardware, thereby enhancing modularity within the system and facilitating rapid prototyping.

In this work, we leverage open-source software to develop a modular testbed architecture for the rapid deployment and characterization of tracked tissue sensors. The testbed offers adaptability to various PLUS-compatible devices, allowing for simplified experimentation with novel preprocessing, classification, and visualization techniques. This framework aims to facilitate the development and implementation of innovative margin assessment technologies.

Chapter 3

Development and characterization of research testbed

This chapter presents a comprehensive overview of the design, implementation, and characterization of our tracked tissue sensing testbed. This chapter is organized into several sections to address different aspects of the research. Section 3.1 focuses on the design and implementation of the testbed, providing a detailed account of the hardware components, the software developed, and the calibration required for our tracked tissue sensing testbed. Section 3.2 describes experiments that assess the tissue sensor sensitivity to various data collection parameters. Section 3.3 details experiments to characterize the real-time performance of the testbed using an ex vivo tissue phantom. Section 3.4 and Section 3.5 detail the results of our experimentation.

3.1 Design and implementation

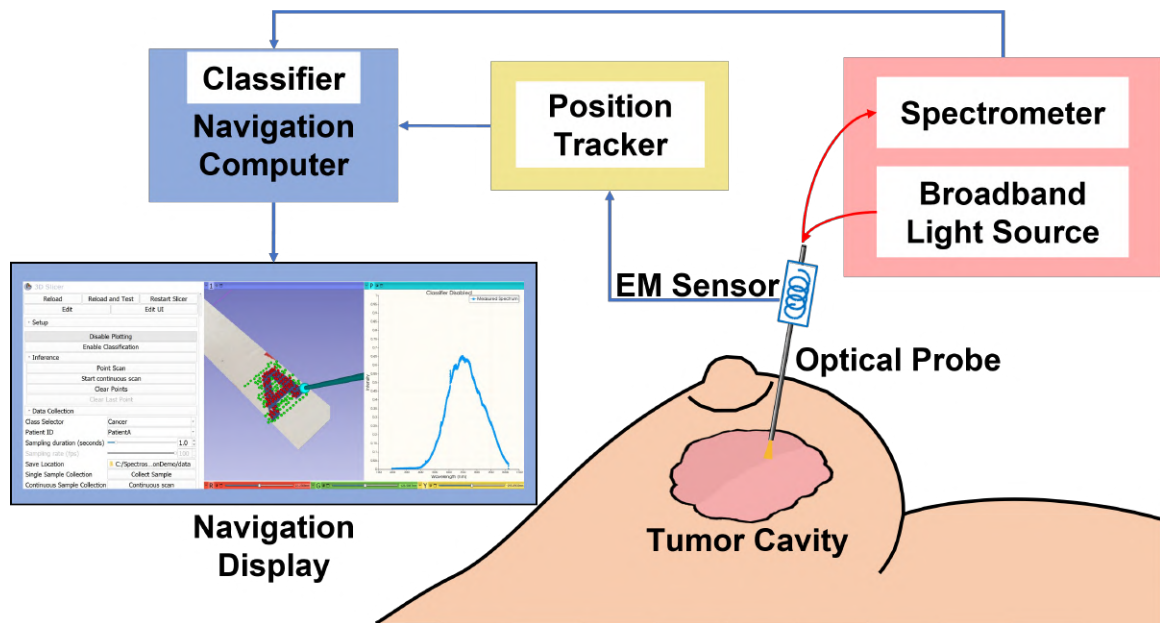
This section presents the design and implementation of the tracked tissue sensing testbed. We provide an overview of the testbed design, followed by a review of the constituent hardware components. We subsequently discuss the development of

software to enable real-time tumour margin assessment. This includes software to integrate the hardware components, enable data collection, and facilitate an intuitive navigation display. We finally overview a calibration workflow for the system.

3.1.1 An overview of testbed design

Our tracked tissue sensing testbed is designed for intraoperative tumour margin assessment. Enabling a surgeon to perform a freehand or assisted scan of the tissue surface post-resection. This includes the ex vivo inspection of tumour margins, as well as the in vitro inspection of the tumour bed. The latter is visualized in Figure 3.1.

The testbed design consists of four main subsystems: a tissue sensor for tissue characterization, a position tracker for recording sensor pose, a classifier for predicting tissue type, and a navigation computer for processing and presenting information in an intuitive graphical representation. Figure 3.1 provides an overview of the testbed design.



The tissue sensor, represented in red in Figure 3.1, is a diffuse reflection optical spectroscopy probe. Light is reflected off the tissue, and measured using a spectrometer. The light is separated into individual wavelengths, characterizing the tissue. The position tracker, depicted in yellow in Figure 3.1, is an electromagnetic tracking system. An electromagnetic sensor is rigidly attached to the spectroscope, allowing for precise determination of the probe position and orientation in 3D space. A trained machine learning classifier is employed within the software to predict the tissue type based on the acquired spectrum. The navigation software, depicted in blue in Figure 3.1, is responsible for communicating sensor information to the user in an intuitive manner. It enables the collection, storage, processing, classification and display of hardware information.

3.1.2 Hardware components and specifications

The testbed consists of two subgroups of hardware: the tissue sensor and the position tracker. A diagram of the workspace is shown in Figure 3.2. The tissue sensor components are shown in red and the position tracker components are shown in yellow. The controller for the position tracker is located below the workspace.

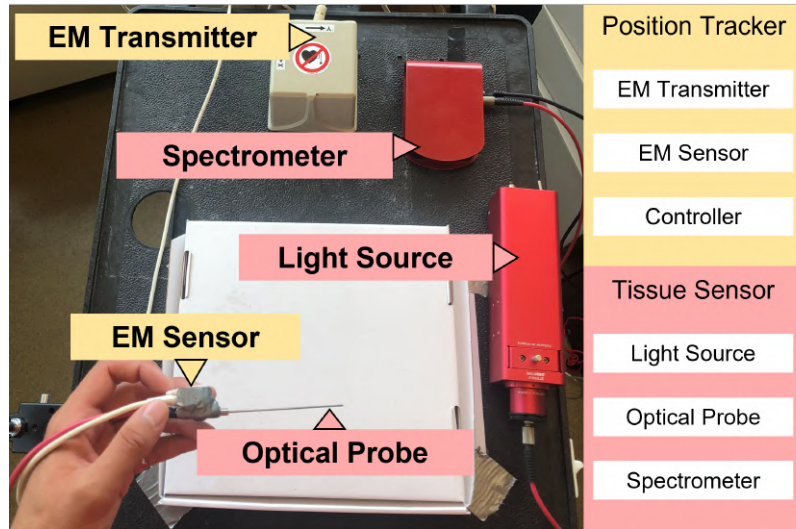


Figure 3.2: An overview of the hardware setup of the testbed

Tissue sensor

The tissue sensor consists of several components, including a broadband light source (Thorlabs, USA) [62], a spectrometer (Thorlabs, USA) [63], and an optical fibre reflection probe (Aventes, NL) [64].

For the broadband light source, we use the SLS201L stabilized tungsten light source [62]. This incandescent bulb operates by passing current through a tungsten filament, causing it to heat and radiate light in the range of 360 to 2600 nm. Given that our light source emits in the near and mid-IR range, the power of the light source

was chosen to limit tissue heating. The SLS201L light source operates at 10 mW over the wavelength range, providing minimal tissue heating during expected use [65]. The colour temperature of the bulb is 2796 K with stability of $\pm 15\text{K}$. The bulb has an output power drift of 0.1% per $^{\circ}\text{C}$. To ensure stable operation, a warm-up period of 40 minutes is required [62]. The radiated light spectrum is depicted in Figure 3.3. The light source does not have a uniform output, gradually increasing from 360 to a peak at 1000 nm.

The spectrometer employed is the CCS200 compact charge-couple-device spectrometer [63]. It operates within a wavelength range of 200 nm to 1000 nm with a resolution of $\sim 1/4$ nm. The spectrometer features a grating of 600 lines/mm, and its efficiency can be observed in Figure 3.3. The spectroscopy integration time setting ranges from $10\mu\text{s}$ to 60s with a scan rate of 100 Hz. The spectrometer is amplitude-corrected for wavelengths greater than 380 nm, thus we assume the spectrometer receiver curve is flat for this range.

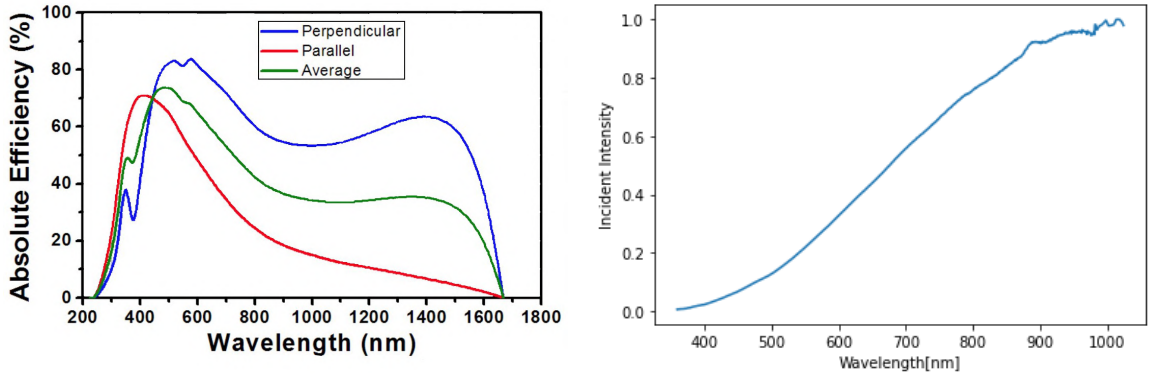


Figure 3.3: Theoretical spectrometer grating efficiency [63] (Left) and theoretical broadband light source output intensity [62] (Right)

The optical fibre reflection probe used is the Aventes FCR-7UVIR200-2 [64]. This probe consists of six illumination fibres that deliver light from the source to the sample. A seventh fibre, located along the centre of the probe, collects the reflected light and delivers it to the spectrometer. Each fibre has a diameter of 200 μm , and the probe has a numerical aperture of 0.22.

In summary, the tissue sensor operates in an effective wavelength range of 360 nm to 1000 nm, encompassing the visible spectrum and near-infrared range. This wavelength was chosen for a variety of reasons. We aimed to extend the work in Connolly et al. (2022), which interrogated a wavelength range from 330 to 500 nm [66]. In our prototype, we extend this range to encompass the visible and near IR range to allow for more features to be interrogated. This enables us to incorporate features that would not be visible to a clinician. Moreover, this range includes wavelengths that have shown promise for differentiating cancerous and benign breast tissue [67].

Position tracker

The position tracker used is the Trakstar 3D Guidance system (NDI, USA), which comprises an EM transmitter, an EM sensor, and a control module. We use the mid-range transmitter from NDI which offers a workspace range with a diameter of 75 cm. The angular range is $\pm 180^\circ$ azimuth and roll with $\pm 90^\circ$ elevation. We use the Model 800 6-degree of freedom (DoF) 8mm sensor which allows for a maximum tracking distance of 66 cm. It achieves an accuracy of 1.4 mm for position and 0.50° RMS for orientation [60]. Cernelev et al. (2023) analyzed our position tracker and found the accuracy to range from 1 mm near the transmitter to 4mm at the fringes of the workspace [68]. Cernelev et al. (2023) concluded that the 8mm sensor was able

to maintain clinically acceptable accuracy for distances up to 50 cm [68]. The study also found that the accuracy of the position tracker tends to decrease with distance, and therefore, the testbed was operated in close proximity to the transmitter, within a range of 10cm - 30cm.

3.1.3 Software development and integration

A 3D Slicer module was developed to facilitate three major software objectives:

1. Enable real-time interaction with hardware
2. Enable collection, labelling, and storage of dataset
3. Support real-time spatial classification and provide intuitive data visualization

The user interface of the software module is depicted in Figure 3.4. The graphical user interface, located on the left-hand side of Figure 3.4, consists of three tabs. The setup tab allows for the initialization of the module, including establishing hardware connections and selecting a classifier. The data collection tab facilitates the live collection, labelling, and storage of datasets. The inference tab enables the real-time operation of the testbed for margin assessment.

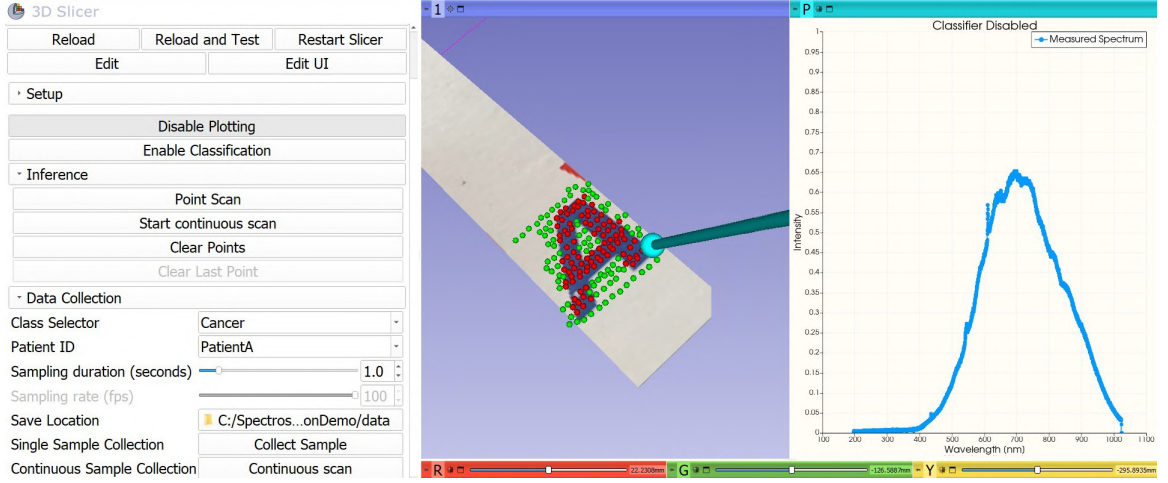


Figure 3.4: The user interface of the developed software module. (Left) The graphical user interface, (Centre) the 3D navigation display, and (Right) the spectrum viewer

Enabling real-time hardware communication

The first software objective in developing the testbed is to enable real-time interaction with the hardware. To facilitate hardware communication, we use the open-source software PLUS and OpenIGTLink protocol [4], [5]. The communication pipeline is visualized in Figure 3.5.

PLUS provides a device interface that simplifies hardware communications through a standardized configuration script. This script allows for the selection of hardware-specific parameters, synchronization, and preprocessing. Appendix A.2 contains the configuration script used to integrate the spectrometer and electromagnetic tracker.

The configuration file allows for the manipulation of spectrometer parameters, including `AcquisitionRate`, `IntegrationTimeSec`, and `AveragedItemsForFiltering`. The `AcquisitionRate` parameter controls the sampling rate of the spectrometer and is set to the default value of 30 Hz. `AveragedItemsForFiltering` determines the width of

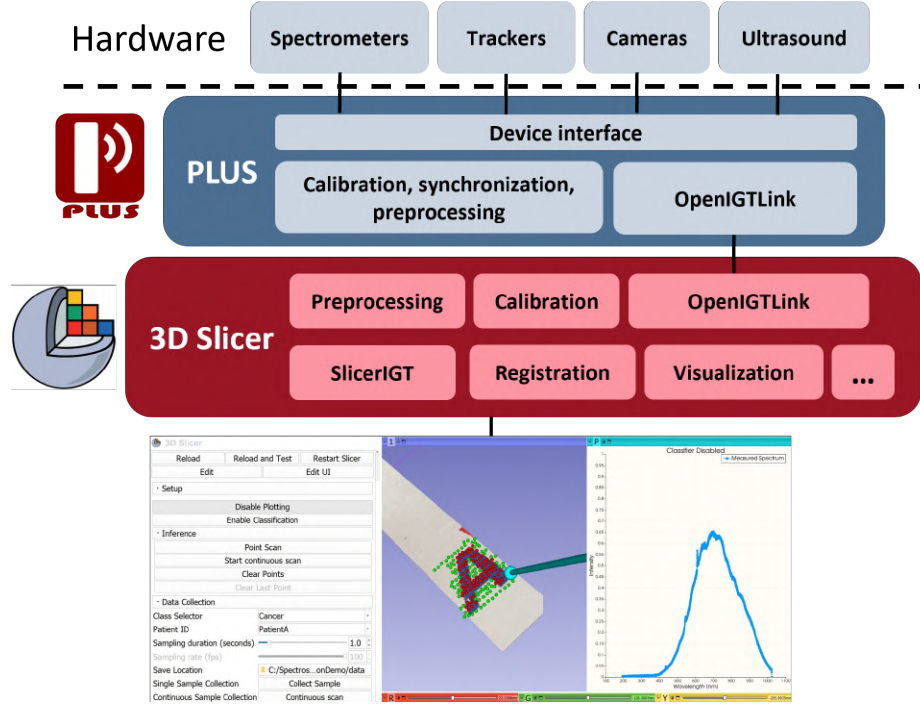


Figure 3.5: An overview of the real-time hardware communication pipeline

the filtering window and is also set to its default value of 20. `IntegrationTimeSec` modifies the length of each signal acquisition. To reduce the probability of signal saturation, this parameter was set to 0.01 seconds. This parameter is optimized for each tissue type based on the absolute reflectivity of the tissues. A lower integration time prevents signal saturation while external signal integration maintains a strong signal-to-noise ratio. The spectrometer data is loaded as a 2D matrix, where the first row represents the wavelengths and the second row denotes the corresponding observed intensity at each wavelength. The tracking data is loaded as a homogeneous transformation matrix, providing information about the location and orientation relative to a reference sensor.

Enabling data collection

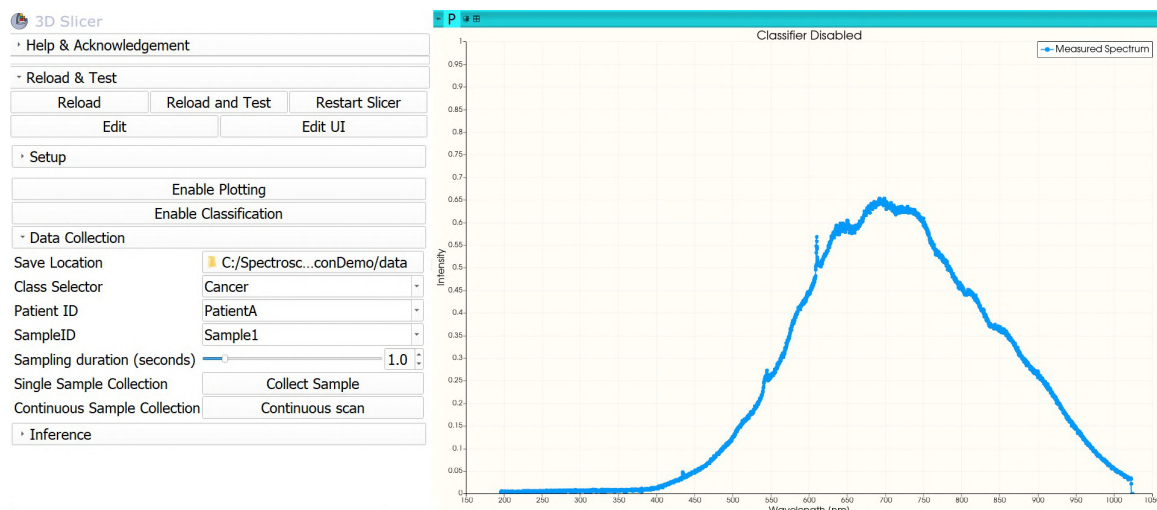


Figure 3.6: The developed data collection interface for our open-source software. (Left) The graphical user interface (Right) A real-time visualization of the optical spectra

The second software objective is to develop a software interface for efficient data collection, labelling, and storage. The data collection interface, depicted in Figure 3.6, includes the following features:

1. Class selector: Enables the user to specify the histopathology or tissue type being scanned.
2. Patient ID: Allows assignment of the scan to a specific patient, with anonymization according to a reference sheet.
3. Sample ID: Differentiates between samples when multiple tissue segments are obtained from a single patient.
4. Continuous sample collection: Allows continuous collection of spectra from a single class, useful for scanning large, homogenous tissue samples.

5. Single sample collection: In conjunction with sampling duration, enables scanning of a single location for a consistent duration, useful for heterogeneous samples.
6. Spectrum viewer: Displays a real-time visualization of the collected optical spectrum

From the interface on the left, the user inputs the class label and metadata, selects the appropriate collection mode, and initiates the data collection process. To record the data, we create a `vtkMRMLSequenceBrowserNode` and `vtkMRMLSequenceNode`. The data is recorded for the specified duration using a `QTimer` object dictated by the sampling duration parameter or user input. The data is formatted and converted to CSV format for storage in the dataset. Each spectrum is stored in a 2D array with the wavelength values in the first row and the spectral intensities in subsequent rows. Metadata including the class label, patient ID, and sample ID is encoded in the file name and filesystem structure. File names are structured by date, PatientID, SampleID, and Class; for example `Mar3-PatientA-Sample1-Cancer`.

The spectrum viewer, located on the right-hand side of Figure 3.6, has been developed to provide real-time visualization of the incoming spectrum. This feature enables the assessment of signal quality and ensures the acquisition of a strong and clear signal, minimizing noise and artifacts. Additionally, the spectrum viewer offers a preliminary exploration of new tissue features, allowing for early insight for later analysis. This is essential for exploratory sensitivity analysis performed later in section 3.2. The spectrum viewer operates as follows. First, the incoming spectrum is retrieved from the 3D Slicer OpenIGTLink module, and stored in a `vtkMRMLTableNode`. A `vtkMRMLPlotSeriesNode` is used to customize the display properties of the table,

including line style and colour. Live display of the incoming spectra is enabled using a `vtkMRMLPlotChartNode` which allows for parameters such as chart labels and axis ranges to be adjusted to enhance visualization. Furthermore, if a machine learning model is uploaded and active, the title of the chart is dynamically set based on the predicted classification, offering immediate context and interpretation of the investigated tissue type.

Navigation display development

The final software objective is the development of a navigation display to facilitate an intuitive, real-time visual representation of the class predictions in 3D space. An example navigation display is shown on the right of Figure 3.7.

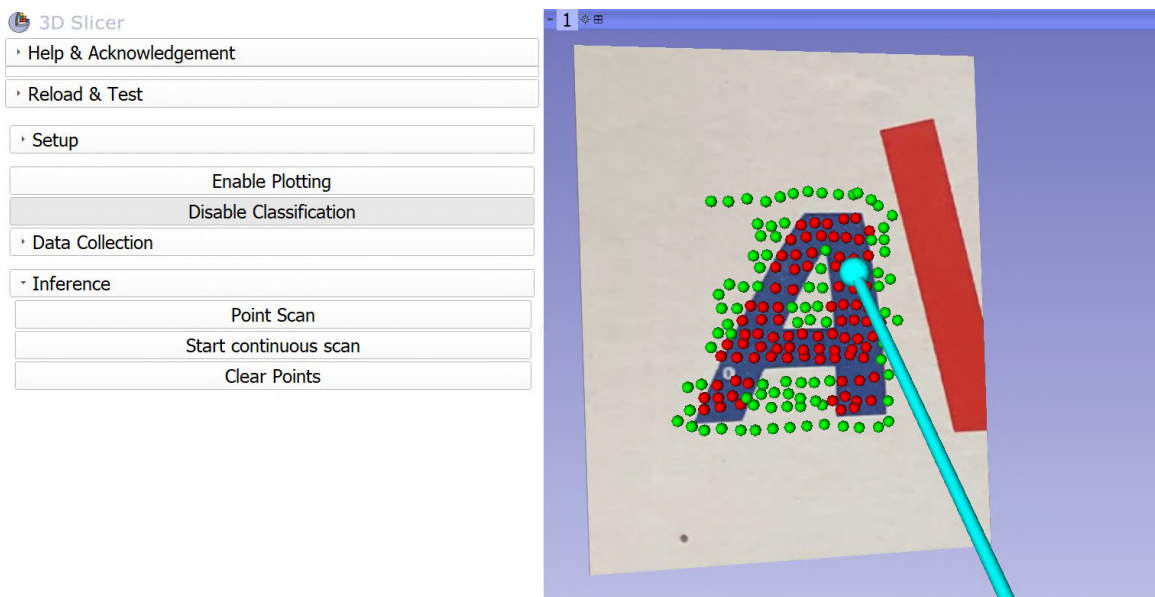


Figure 3.7: Illustration of the 3D navigation display interface. Red and green points represent classes 0 and 1 respectively, while the blue model represents the optical probe.

A blue cylindrical model represents the probe location and orientation. A static image of the scanned area is registered to the scene. The location and classification data are combined to generate a point with colour denoting class. For example, locations predicted as cancer are coloured red, and locations predicted as benign are coloured green. The user can toggle between scanning a single point or performing a continuous scan. The user can input the path to a trained machine-learning classifier. The trained classifier should be packaged into a pickle file using the joblib library before importation into the module. When classification is enabled, the module observes incoming spectra, applies relevant data processing, and inputs each spectrum into the model for prediction. The model prediction is passed to the navigation display and the spectrum viewer for visualization.

Navigation calibration workflow

Prior to each scan, the navigation display requires calibration to ensure accurate visualization. The EM sensor tracks the handle of the probe, rather than the tip of the probe directly. The probe tip is tracked indirectly through a rigid transform from the actual sensor location. The location of the probe tip is calibrated using the Pivot Calibration module in 3D Slicer. The tip is fixed in a small divot and the probe is pivoted about this location in a semi-spherical motion for ~ 5 seconds. The transform is calculated and the process is repeated until a transformation error of < 1 mm is achieved. An image of the region being scanned is registered to the 3D Slicer scene. Four reference points are marked on the region of interest. The image is imported into 3D Slicer and four markers are overlaid on the reference points. The calibrated probe is placed on each reference point to record the position in 3D space.

The Fiducial Registration Wizard module is used to compute a transform between the two sets of points. This transform is applied to the image for spatial alignment to the scene. This registration is repeated until the accuracy reaches $\sim 1\text{mm}$.

3.2 Exploring spectroscope sensitivity to environmental parameters

This section explores the sensitivity of the spectroscope signal to various environmental parameters. The goal was to explore the parameters to which the spectroscope was robust or sensitive. We experimented with the effect of the probe-to-surface distance, the ambient light of the room, and the probe-to-surface angle.

The experimental setup is shown in Figure 3.8. A 2-DoF stage allows for translation in the x (blue) and y (yellow) directions. A 2-DoF stand allows for translation in the z-direction (red) and rotation about the y-axis. A 3D-printed adapter was developed to fix the probe to the stand.

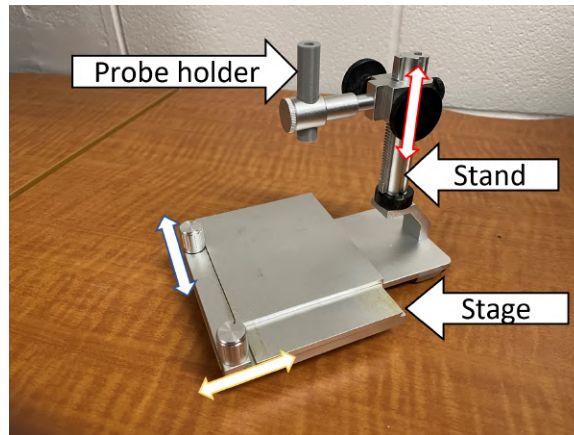


Figure 3.8: Experimental setup for sensitivity analysis experiments

3.2.1 Probe-to-surface distance

The signal sensitivity to the probe-to-surface distance was experimentally evaluated. We used signal intensity as a proxy for distance in this experiment as signal intensity and distance are inversely correlated. A white piece of paper was used to reduce the effects of inhomogeneities in the scanned area to help isolate for distance. The experimental setup is shown in Figure 3.9

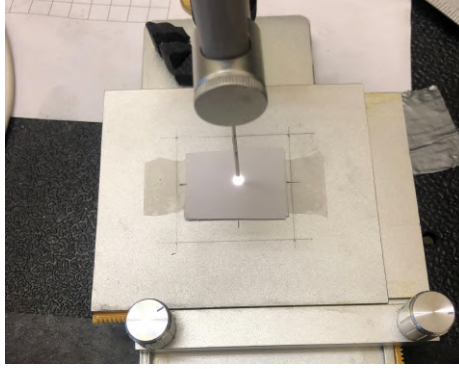


Figure 3.9: Experimental setup for the probe-to-surface distance experiments

The testbed was evaluated at varying heights to assess the effect of distance on the received signal intensities. To perform this experiment the tissue sensor was fixed vertically above a white piece of paper using the calibration setup in Figure 3.8. The intensity of the signal was monitored using the spectrum viewer. The probe was lowered until the peak signal intensity was just under the maximum of 1 (without saturation). The spectrum was saved and an optical image was taken. The probe was raised, and this process was repeated for peak signal intensities of 0.8, 0.6, 0.4, 0.2, 0.1, 0.05, and 0.02. We compared the differences between each probe-to-surface distance with the maximum signal intensity. We calculated the mean and standard deviation within each differential signal.

3.2.2 Probe-to-surface angle

The signal sensitivity to the probe-to-surface angle was experimentally evaluated. A white piece of paper was used to help isolate the effects of angle. The probe was placed in the calibration setup seen in Figure 3.8. The probe position was set perpendicular to the paper with a signal intensity of approximately 0.9. An image was taken and the spectrum was recorded. The probe was rotated approximately 10° . The probe tip position was adjusted to ensure approximately the same scan location and signal intensity were maintained. The process was repeated until the probe was at approximately a 45° angle. Three example images can be seen in Figure 3.10. We compared the differences between each recorded angle with the angle obtained 90° to the paper. We calculated the mean and standard deviation within each differential signal.

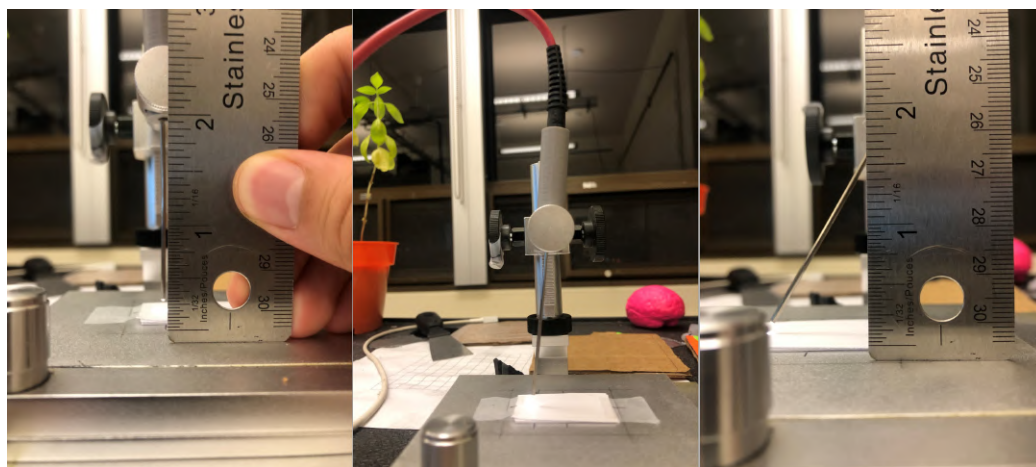


Figure 3.10: Experimental setup for the probe-to-surface angle experiments shown at $\sim 90^\circ$, 70° , and 45°

3.3 Characterizing the testbed performance

This section details the process of characterizing the performance of our tracked tissue sensing testbed. We characterized the latency of the navigation and classification portions of the navigation display. We then characterized the accuracy of the testbed’s spatial classification using an ex vivo biological tissue phantom consisting of porcine and bovine tissue.

3.3.1 Experimental setup

To characterize the navigation display, we required a trained tissue classifier. Since the classification was not the primary focus of this experiment, a simplified classifier was trained and deployed. The classifier used in this experiment was a k-nearest-neighbour (kNN) model provided by the scikit-learn library [69]. The kNN model hyperparameters included `n_neighbors`, `metric`, `weights`, `algorithm`, and `leaf size`. `n_neighbors` indicated the number of neighbouring points used to make a prediction during inference. The default value of 3 was selected. The Euclidean distance was used as the metric to determine point-to-point distance. A uniform weight distribution was used as the default during inference. The `auto` parameter was used to optimally select between the `BallTree`, `KDTree`, or brute-force search algorithms. A default leaf size of 30 was used for the `BallTree` and `KDTree` algorithms. To train the kNN, bovine and porcine tissue was acquired and prepared for sampling. The surface of each tissue was scanned and labelled, which resulted in a dataset of 200 spectra, 100 bovine and 100 porcine. Each spectrum was cropped to 360-1000 nm and a minimax normalization was performed. A binary label of 0 and 1 was assigned to bovine and porcine tissue spectra respectively. The trained kNN classifier was imported into the testbed.

3.3.2 Characterization experimentation

Testbed latencies

Before characterizing the spatial classification performance of the testbed, the latency of the navigation and classification was experimentally characterized. The navigation display was calibrated in accordance with the workflow described in section 3.1. We determined its suitability for real-time use and ensured temporal consistency between the classification and location data. The classification and navigation latencies were

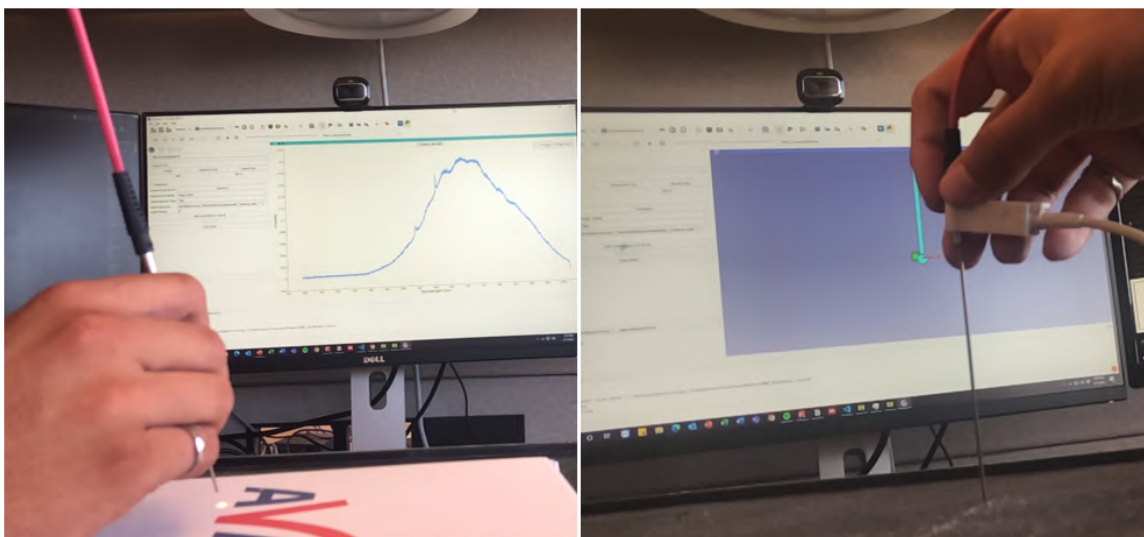


Figure 3.11: A visual of the testbed latency characterization experiment

determined using a 240-frame-per-second slow-motion camera to record the experimental workspace and navigation display (Figure 3.11). To determine the classification latency, the probe was then repeatedly transitioned between two solid colours for 20 cycles. The slow motion camera was analyzed frame by frame to approximate the delay between the probe reaching the colour, and the display displaying the class. To determine the electromagnetic tracking latency the probe was repeatedly lifted from

a rigid surface and returned for 20 cycles. The slow-motion video was analyzed and the frames were counted between when the physical probe and virtual probe model stopped moving.

Spatial classification performance

To characterize the spatial classification performance of the testbed, an ex vivo tissue phantom was created using bovine and porcine tissue to simulate a tumour cavity. The phantom layout is visible in Figure 3.12. Porcine tissue is used as the main phantom body with bovine tissue inserted to represent a tumour. The phantom was prepared fresh and frozen immediately. The tissue phantom measured approximately 10 cm in diameter and 3 cm in depth. Before experimentation, the phantom was removed from the freezer and allowed to thaw completely.

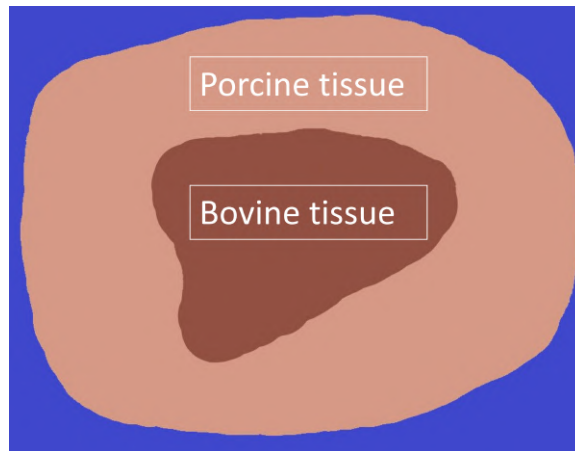


Figure 3.12: An illustration of the biological tissue phantom design used in this experiment

The thawed tissue phantom was placed in the workspace as shown in Figure 3.13. An overhead optical image of the tissue phantom was taken prior to data collection for use as the ground truth image. The optical probe was calibrated and the ground truth

image was registered to the coordinate system. The trained classifier was inputted into the testbed, and a continuous data collection mode was selected. The tissue phantom was scanned freehand in a rough grid pattern with 1 mm spacing between displayed points. The spatial classification was overlaid on the ground truth image of the tissue phantom. The number of correctly and incorrectly classified points was recorded for each class. The testbed visualization was evaluated using the accuracy and the Dice similarity coefficient.

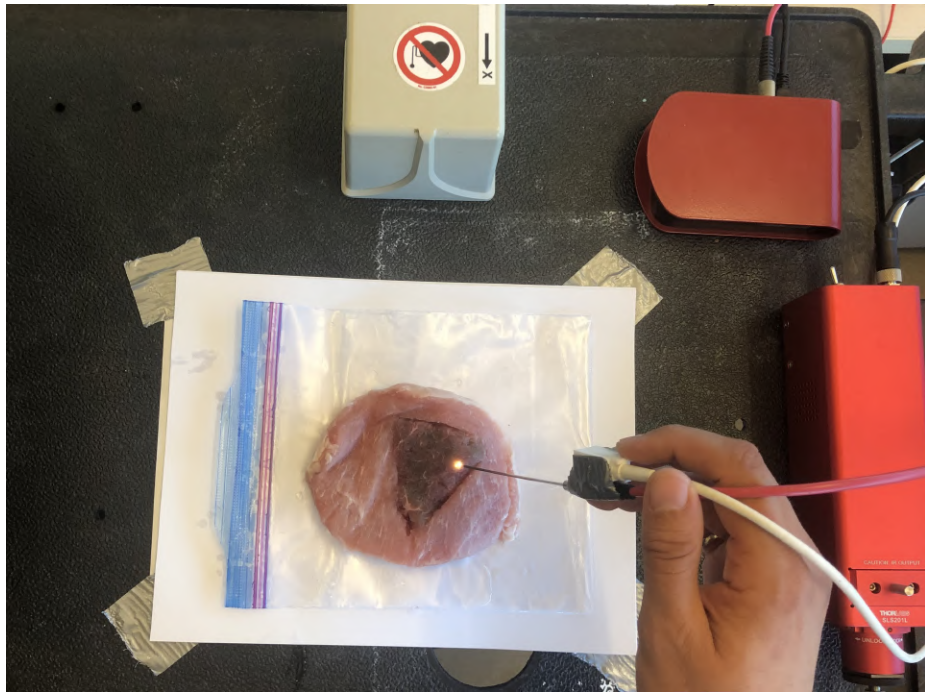


Figure 3.13: Experimental workspace for the characterization of testbed performance using an ex vivo biological tissue phantom

3.4 Results of spectroscope sensitivity analysis

3.4.1 Effects of probe-to-surface distance

The behaviour of our spectroscope signal was evaluated as a function of the distance between the probe and the measured surface. Representative measurements of the received signal at various distances are displayed in Figure 3.14. We observed that

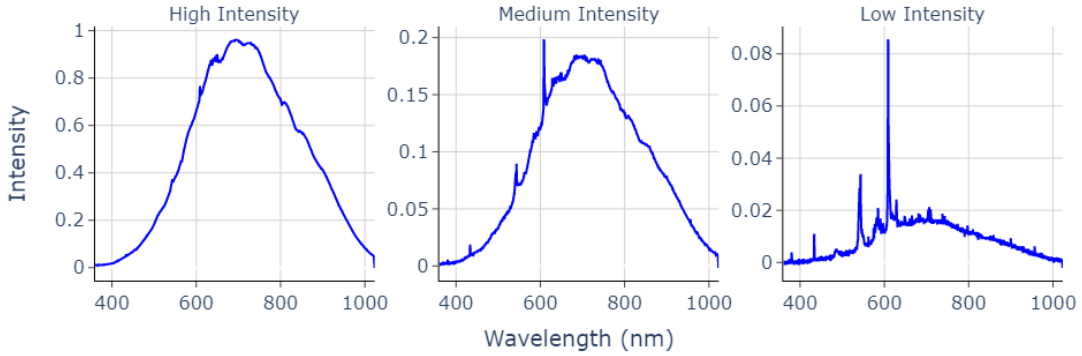


Figure 3.14: Resulting spectra from the probe-to-surface distance experiments. Spectra recorded closest to the tissue (Left), furthest from the tissue (Right), and an intermediary distance (Centre)

the overall signal intensity decreased with distance. However, two distinct signal behaviours were superimposed in the signal. This was the result of two distinct light sources, the broadband light source emitted from the probe, and the ambient light emitted in the room. The effect of each light source attenuated with distance at different rates. The intensity per unit area decreases as light propagates to and from the surface. In this experiment, the broadband light was moved away from the surface while the ambient light distance remained constant. The intensity per unit area decreased for the broadband light source while remaining constant for the

ambient light source. Thus the ratio of ambient light to broadband light increased with distance. As a result, distance information was encoded into the signal. The ambient light was isolated from the signal and displayed in Figure 3.15.

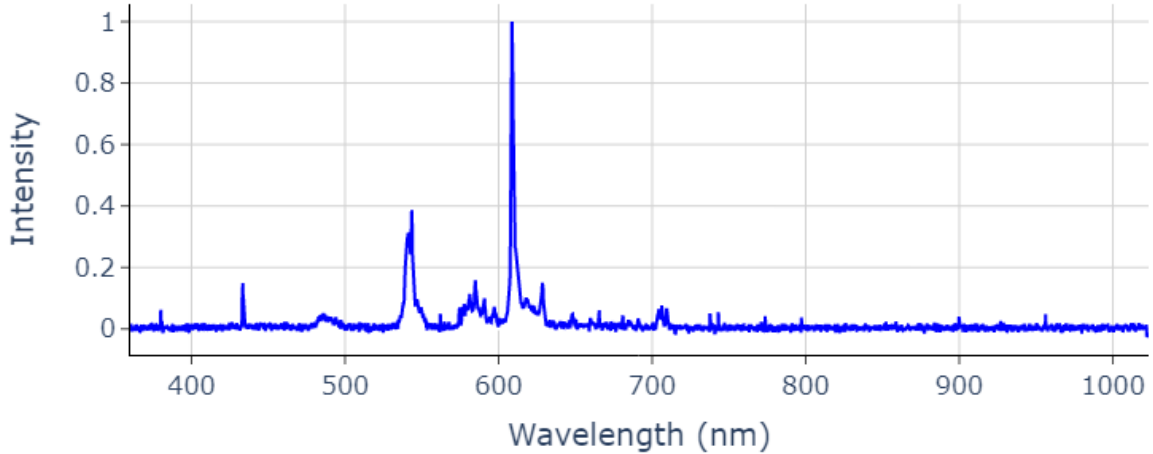


Figure 3.15: An example of ambient light in the room reflected off a white piece of paper

The ambient light is highly dependent on environmental variables. These include factors such as the type, age, and location of the light source, as well as dynamic obstructions such as personnel. The ambient light may introduce features specific to a location or data collection procedure. These features may influence model training. A clinic that produces mostly cancerous data may have its light source used to classify cancerous tissue rather than tissue properties alone. The ambient light should be removed as part of the data preprocessing pipeline.

After the ambient light had been removed, we observed how the received broadband light varied with probe-to-surface distance. The spectrum recorded at each distance was compared to the highest-intensity signal. The mean and maximum differences are displayed in Figure 3.16 as a function of signal intensity.

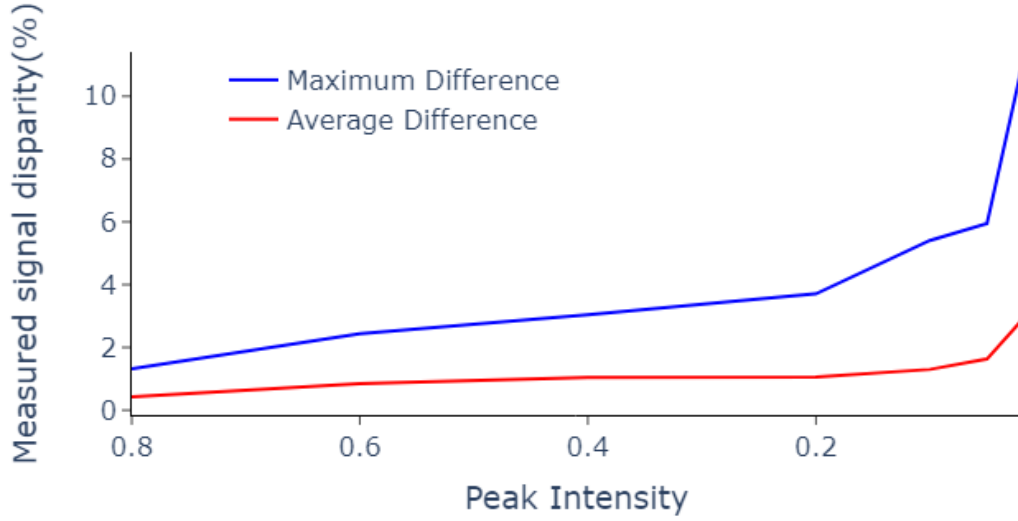


Figure 3.16: The average (red) and maximum (blue) signal differences as a function of measured peak intensity

We observed that the probe-to-surface distance had a minimal impact on the shape of the received signal. The scaling factor of the signal appeared to be constant for each wavelength in the range. When comparing the signal at each distance to the initial reading, the mean difference remained relatively low. The mean difference was less than 1% until the signal intensity decreased below 0.1. The signal difference remained less than 3% at an intensity of 0.02. We observed that the standard deviation increased for low-intensity signals which coincided with a decreased signal-to-noise ratio.

From this experiment, we infer an approximate minimum intensity threshold of ~ 0.1 for the received signal to be maximally beneficial for real-time use. This threshold ensures a strong signal-to-noise ratio, potentially reducing the chance of false predictions as a result of noise. In addition to a minimum intensity threshold,

the addition of uncertainty estimation would be beneficial, communicating classifier prediction confidence, or filtering uncertain predictions altogether. These results indicate that minimax normalization is a suitable normalization technique once the ambient light has been removed from the signal.

3.4.2 Effects of probe-to-surface angle

The effect of the probe-to-surface angle was evaluated through observation of the spectroscope signal. The recorded spectrum for each angle was compared to the vertical as a baseline. The mean and maximum differences are displayed in Figure 3.17 as a function of signal intensity. The average difference across all angles was found to be constant at $\sim 0.18\%$ with a maximum difference of $\sim 0.8\%$. This difference was not significantly different in magnitude from the background noise.

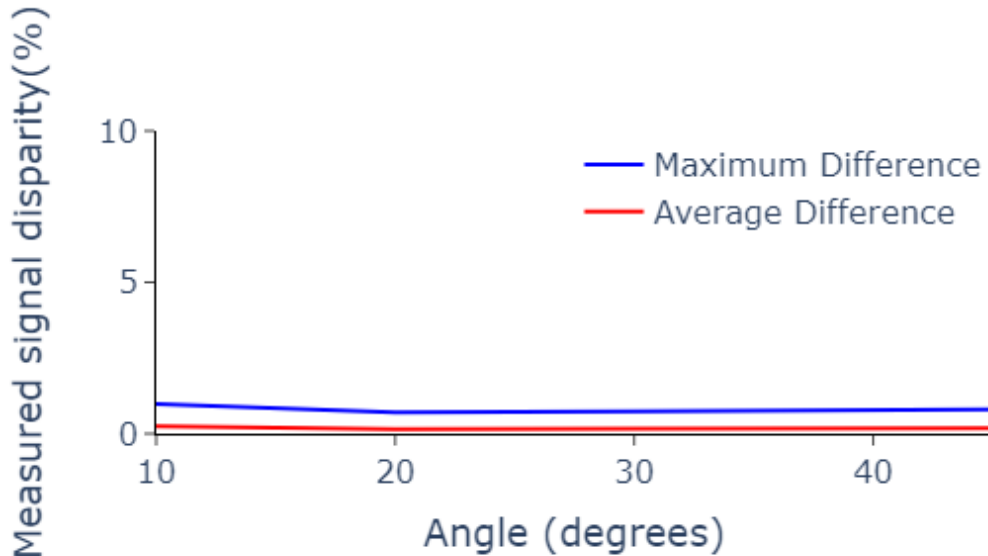


Figure 3.17: The average (red) and maximum (blue) signal differences as a function of probe angle

We observed that the angle of the probe had a minimal effect on the received spectrum. This result indicates that the variation of probe angle resulting from freehand tissue scanning should have a minimal impact on the acquired signal.

3.5 Real-time navigation display characterization

This section details the results of our testbed characterization process in terms of spatial classification performance and testbed latency.

3.5.1 Latency characterization

The latency of the tracking and classification components of the navigation display were experimentally determined to be $125\text{ms} \pm 11\text{ms}$ and $167\text{ms} \pm 26\text{ms}$ respectively (Table 3.1).

Table 3.1: The testbed latencies for the navigation (EMT) and classifier portions of the display visualization. The mean and standard deviation is reported in ms and is calculated over 20 trials.

EMT latency [ms]	125 ± 11
Classification latency [ms]	167 ± 26

A study from Xu et al. (2013) assessed the effects of latency on real-time tool manipulation in a dV-Trainer simulator environment [70]. System latencies were progressively increased in 100ms increments from ~ 0 to 1000ms. They concluded that latencies of $<200\text{ms}$ facilitated easy tool manipulation, with latencies $<300\text{ms}$ deemed safe by all participants [70]. Our testbed achieved this threshold for both the classification network and the tracking latencies. This result also allows for improved temporal alignment of the position and classification data streams through an

estimated offset, increasing the accuracy of the navigation display. The effect of this increased alignment will be particularly prevalent when either the probe location or tissue class is changing rapidly. Overall, the results indicate that our testbed architecture is capable of real-time classification, localization, and visualization. Facilitating quick and accurate assessment of tissue margins.

3.5.2 Spatial classification characterization

The output of our spatial classification characterization is displayed in Figure 3.18. The red points are classified as bovine tissue and the green points are classified as porcine tissue. The spatial classification is overlaid on the ground truth image of the tissue phantom. Our testbed achieved a Dice similarity coefficient of 0.93, and an accuracy of 0.94 (Table 3.2). The confusion matrix can be seen in Figure 3.19.

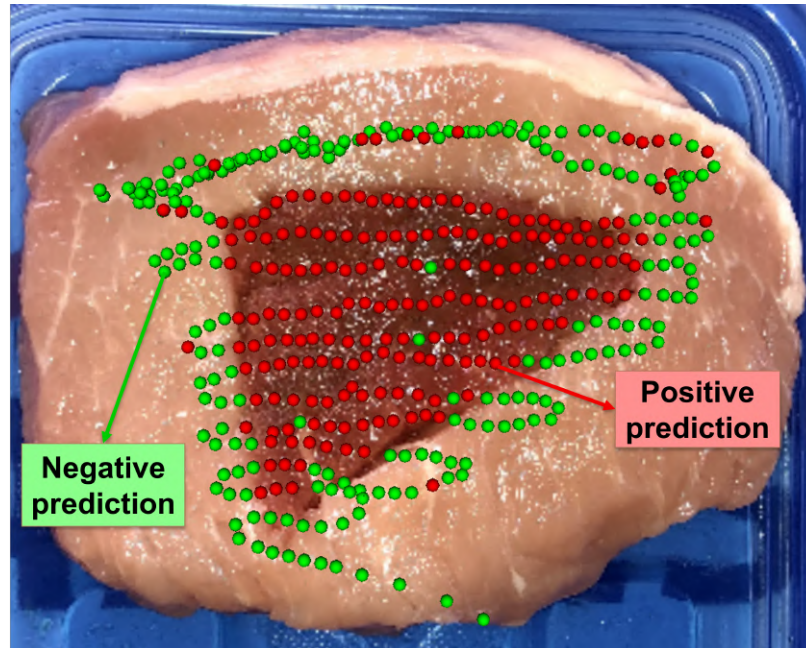


Figure 3.18: The resultant spatial classification visualization for our ex vivo tissue phantom [61]

Table 3.2: The Dice similarity coefficient and accuracy metrics computed from our spatial classification visualization [61].

Dice similarity coefficient	0.93
Accuracy	0.94

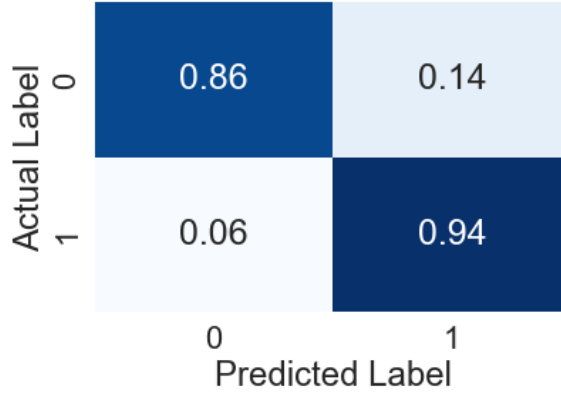


Figure 3.19: A confusion matrix for the spatial classification of bovine tissue (class 0) and porcine tissue (class 1) [61]

Our spatial classification results demonstrate potential for the real-time classification and visualization of a 3D tissue surface. The visualization in Figure 3.18 met basic qualitative metrics, and the system’s modularity enables task-specific enhancements. The testbed effectively generated an intuitive visualization of the predicted classes, with the navigation display aligning well qualitatively. Red and green points were primarily located in their respective ground truth regions, although minor discrepancies were observed along tissue boundaries and throughout the navigation display. The presence of false negatives and positives within bulk tissue regions were most likely due to the simplified classifier and training regimen used. Discrepancies at the tissue boundaries were most likely a result of various tracking and registration errors within the testbed. The modularity of the testbed allows it to adapt to specific

task requirements. The tissue sensor, tracking system, and classifier can be modified according to the procedure requirements. For instance, if a procedure requires higher navigation accuracy, the TrakSTAR (Ascension, USA) electromagnetic tracker could be replaced with a PLUS-compatible optical tracker. Additionally, the tissue classifier is interchangeable and can be tailored for a specific application, allowing for rapid experimentation.

Overall, the results show that our tracked tissue sensing testbed is able to accurately interrogate tissue samples with real-time navigation and classification. Further work focuses on the adoption of more comprehensive classification pipelines for the application of the testbed to clinically relevant tissue types.

Chapter 4

Application of research testbed for interrogation of human kidney tissue

In the previous chapter, it was shown that the tracked tissue sensing testbed is capable of accurate real-time navigation and visualization of ex vivo tissue. This chapter focuses on the application of our testbed to interrogate cancerous and non-neoplastic human kidney tissue. This chapter is organized into several sections. Section 4.1 details the procedure used for the collection and analysis of ex vivo human kidney data. Section 4.2 presents a modelling pipeline designed for the classification of cancerous and non-neoplastic tissue. It follows with an overview of experiments evaluating pre-processing methods on classification performance. Section 4.3 details the results of our experimentation and analysis.

4.1 Collection and analysis of ex vivo human kidney tissue

In this section, we review the data collection procedures used to interrogate ex vivo human kidney tissue. We outline the exploratory data analysis techniques used to gain preliminary information about relationships within the dataset. The ex vivo

tissue samples were collected by Dr. Kevin Ren and interrogated under the guidance of Dr. Martin Kaufmann.

4.1.1 Data collection

Three fresh human kidney tissue samples were collected from a single patient, denoted Patient A. The tissue samples remained unfixed and were stored at -80°C . The samples were heterogeneous and were composed of cancerous and non-neoplastic regions of an excised kidney. The front and back of the tissue samples are shown in Figure 4.1. Each tissue sample was approximately 25mm by 25mm by 5mm in size. The experiment was performed in a biosafety cabinet to minimize contamination.

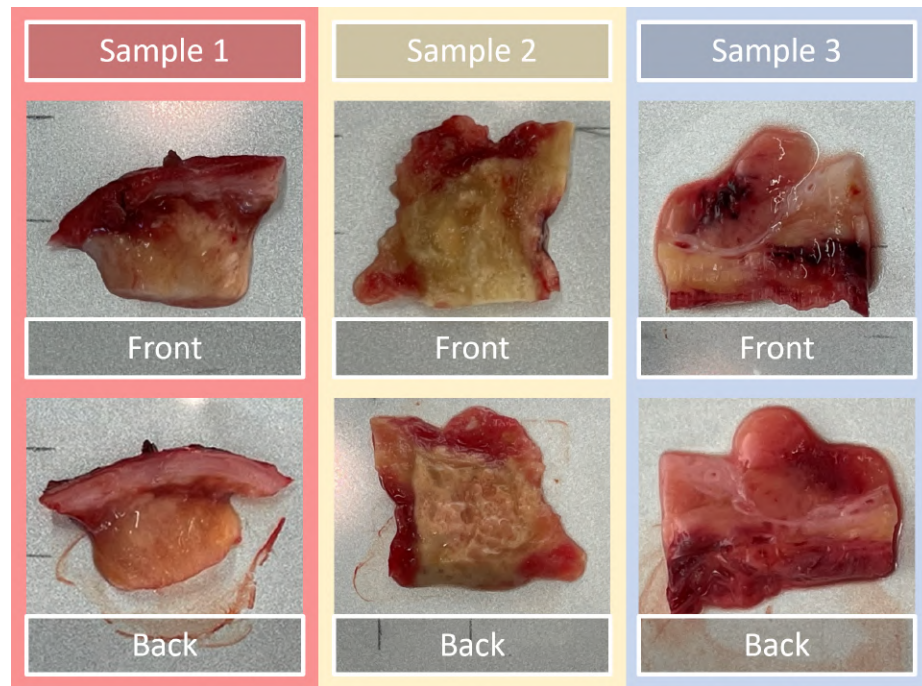


Figure 4.1: Images of the resected ex vivo human kidney tissue specimens

The experimental setup can be seen in Figure 4.2. The setup consisted of the tissue sensor, a 2-DoF stage, and a 1-DoF stand. The stage offered a resolution of ~ 23 mm translation per full knob rotation. The stand offered a resolution of ~ 19 mm translation per full knob rotation.

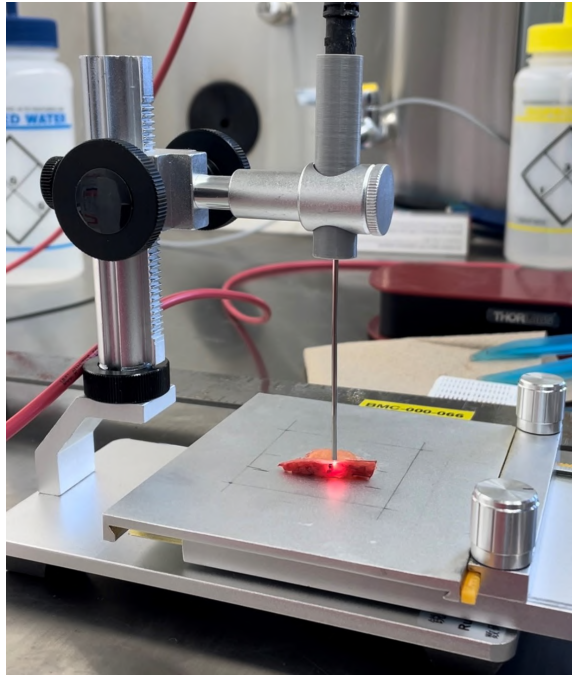


Figure 4.2: The experimental setup for the interrogation of tissue specimens

A detailed description of the standard operating procedure can be found in Appendix A.1. Sample 1 was removed from the freezer and thawed completely for 20 minutes. The sample was placed in the centre of the stage. An overhead image was taken of the specimen. The image was manually segmented by a clinician. The data collection module was used in single-point collection mode with an acquisition time of 1 second. Due to time constraints associated with drying tissue, the light source was not manually toggled before each reading. Reflected ambient light was instead recorded at multiple locations on each pathological region before the scan to estimate

the ambient light at a later time. Each pathological region of the tissue sample was scanned in a grid pattern. The sample was allowed to rehydrate in saline before the second side was scanned. This process was repeated for Sample 2 and Sample 3. Figure 4.3 displays the resulting ground truth segmentations of the cancerous and

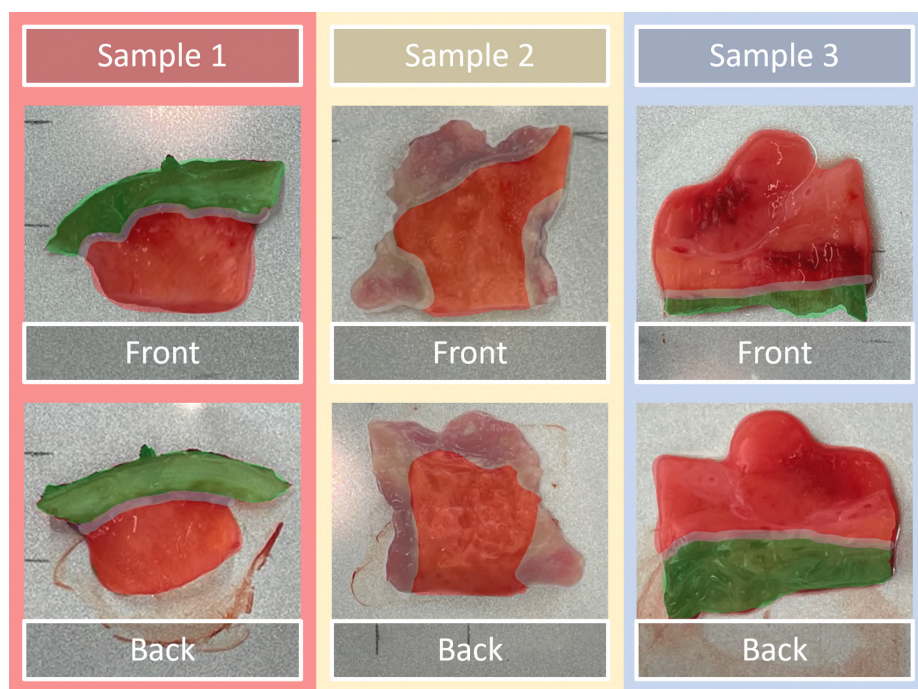


Figure 4.3: Approximate regions of cancerous (red) non-neoplastic (green) and ambiguous (white) tissue determined visually by a trained pathologist

non-neoplastic tissue, as visually determined by a trained pathologist. The green and red overlays represent non-neoplastic and cancerous tissue respectively. The white overlay denotes regions that were not scanned due to ambiguity. The resulting dataset contained 315 cancerous and 84 non-neoplastic unique spectra recordings. The associated patient ID, sample ID, scan side, and histopathological metadata were recorded.

4.1.2 Data distribution analysis

We performed a preliminary visualization of the dataset to explore the separability of the cancerous and non-neoplastic tissue. To facilitate this, we employed Principal Component Analysis (PCA), a dimensionality reduction technique known for its effectiveness in transforming high-dimensional data into a more manageable, lower-dimensional form [71]. This transformation is accomplished while maximally preserving the variance within the dataset, thereby enhancing data interpretability and minimizing information loss [71]. This enables a visual analysis of data distribution trends. The output of the PCA technique is a set of principal components that capture the directions of maximum variance within the dataset. Each principal component represents a linear combination of the original wavelength bins, with weights indicating their importance. As our visualization was limited by 3D space, we chose the first 3 principal components to represent the data. Thus, the dataset was transformed from size 399×2578 to size 399×3 . We used the PCA implementation from the scikit-learn library with the single value decomposition solver set to the default of `scipy.linalg.svd` [72]. The dataset was visualized in the 3D latent feature space. We generated 3D interactive scatter plots where each point in the plot represented a spectrum. We interrogated the data using the collected metadata. This included pathological labels, sample numbers, and the timestamp of data acquisition.

4.2 Tissue classification experiments

In this section, we present the architecture of a machine-learning classification pipeline for the differentiation of cancerous and non-neoplastic tissue. We experimented with

a variety of preprocessing techniques and assessed their impact on classification performance.

4.2.1 Modelling pipeline

We present a baseline machine learning pipeline for tissue differentiation experimentation.

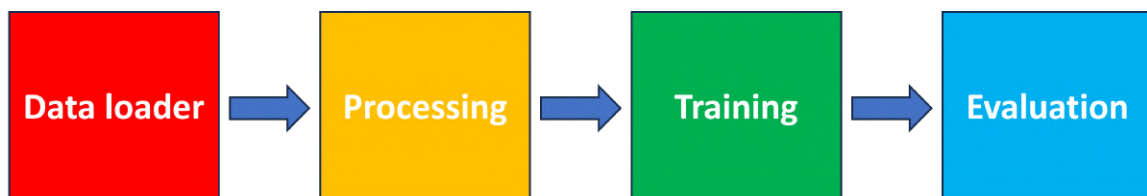


Figure 4.4: Modelling pipeline for tissue classification

The pipeline (Figure 4.4) comprises a data loader, preprocessor, model trainer, and evaluator.

Data loader

The data loader is tasked with importing and formatting the raw dataset into a pandas data frame for convenient manipulation. An example of a raw input spectrum is visualized in Figure 4.5. A single spectrum is a 2D matrix where the first row contains the wavelengths and the second row contains the associated signal intensities. An example of the formatted dataset is shown in Table 4.1. Metadata is parsed from the file name of each spectrum and organized into columns such as patient ID, sample number, scan side, numeric class label (0 or 1), and descriptive class label (cancer, normal, fatty cancer, hemorrhaged cancer etc). The data loader calculates the mean of each 1-second scan to produce a single spectrum for each scan location. This

spectrum is saved in the data column of the data frame. The timestamp of data acquisition is recorded in the Time column. This is repeated for each acquired point.

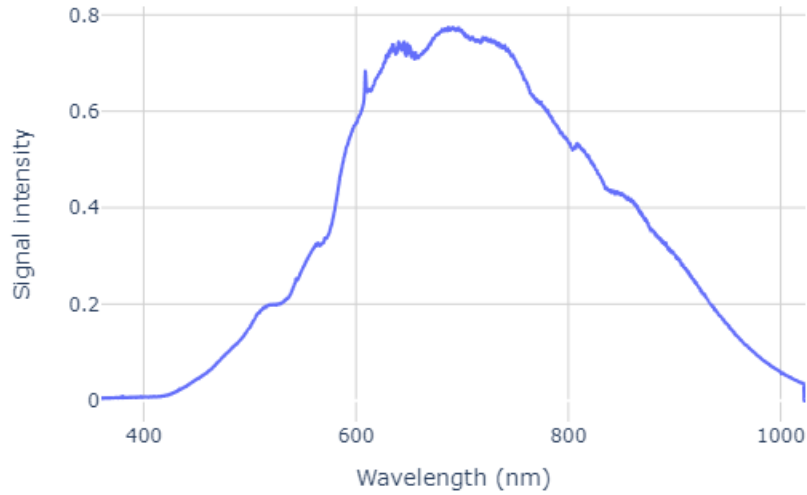


Figure 4.5: Visualization of a raw broadband spectrum

Table 4.1: Example of an imported and formatted dataset

PatientID	Sample	Scan	Label	Data	Time
PatientA	Sample1	back	1,Cancer	[360, 0.05]...[1000, 0.02]	11:22:23
PatientA	Sample1	back	1,Cancer	[360, 0.04]...[1000, 0.03]	11:22:31
...
PatientA	Sample3	front	0,Normal	[360, 0.15]...[1000, 0.04]	16:13:43

Default preprocessing

The default preprocessing pipeline primarily addresses signal compensation based on the light source intensity curve. The received signal is not a direct measure of the tissue properties but rather contains information about the light source used, and each medium the light interacts with during acquisition. Each light source and spectrometer will have different output and receiving curves which should be compensated

for to isolate the signal. The spectrometer manual specifies that the receiving curve is amplitude corrected for our operating band [63]. Thus, the theoretical receiving curve is assumed to be flat. The broadband light source manual specifies a theoretical output intensity curve (Figure 3.3). The received spectrum is divided by the theoretical light source intensity curve. Figure 4.6 presents an example spectrum pre and post-compensation.

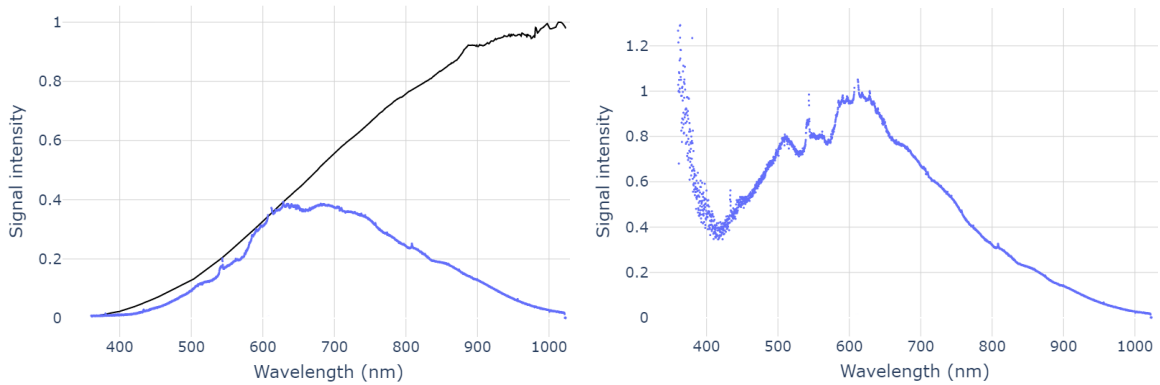


Figure 4.6: Light source compensation. (Left) The light source intensity curve is in black and the reflected intensity curve is in blue (Right) The compensated intensity curve is in blue)

On the left, an uncompensated spectrum is shown in blue with the theoretical light source intensity curve shown in black. On the right, we see a compensated spectrum shown in blue. We observe that the compensated signal tends to infinity in the range of 360 nm to 420 nm. At these low wavelengths, the light source intensity curve approaches zero. Thus these are regions of low signal-to-noise ratio. As a result, the dataset was cropped to 420 nm for all subsequent experiments.

Model training

The machine learning model we used was the Linear Discriminant Analysis (LDA) classifier. The limited size of our dataset was a confounding factor in this selection. LDA is an approach which can be used with less data compared to traditional deep learning methods, which contain a large number of trainable parameters. LDA aims to find a linear combination of features that maximizes the separation between different classes while minimizing the variance within each class. The output dimensionality is limited to the lesser of the number of features and one less than the number of classes. Given two classes we are limited to 1 dimensional output. We employed the LDA model from the scikit-learn library [73]. The LDA used an SVD solver with the threshold for determining singular value significance set to the default of $1.0e^{-4}$. To train and evaluate the model the data was split into a training and testing set. The dataset showed a significant class unbalance. Sample 1 contained 77 cancerous spectra and 58 non-neoplastic spectra. Sample 2 contained 112 cancerous spectra and 0 non-neoplastic spectra. Sample 3 contained 126 cancerous spectra and just 26 non-neoplastic spectra. To alleviate this class unbalance, we randomly oversampled the training set. We employed the RandomOverSampler from the imbalanced-learn library [74]. The minority class was randomly oversampled until it reached parity with the majority class.

Model evaluation

The model was evaluated using various train-test splits and k-fold cross-validation. We performed leave-one-out cross-validation based on the sample number and scan side. We created 6 folds which consisted of sample 1 front, sample 1 back, sample 2

front, sample 2 back, sample 3 front, and sample 3 back. We trained and evaluated the model six times, where 5 folds were used for training and the remainder was used for testing. We repeated this cross-validation process based on sample number alone, using 3 folds which consisted of sample 1, sample 2, and sample 3. We trained and evaluated the model three times, where 2 folds were used for training and the remainder was used for testing. Each cross-validation was performed until a total of 30 trials were conducted. We evaluated the model in terms of balanced accuracy, sensitivity, specificity, F1 score, AUC, and accuracy. The mean and standard deviation over the trials were recorded for each train-test split.

4.2.2 Preprocessing experimentation

We used the modelling pipeline to explore various preprocessing methods. We analyzed the effect of ambient light compensation, probe-to-surface distance normalization, and dimensionality reduction methods on classification performance.

Ambient light compensation

The ambient light contains information specific to the data collection environment, as well as probe-to-tissue distance information. The ambient light within a signal should be removed to improve the robustness of the dataset to these factors. The ideal method of ambient light compensation would be to directly measure the ambient light reflecting off a tissue before each reading is taken. However, due to hardware limitations, this was not feasible. Instead, we interrogated two methods of ambient light compensation: ambient light estimation (ALE) and ambient peak zeroing (APZ). We first attempted to estimate the ambient light contained within a spectrum using

data collected prior to each scan with the light source on and off. We isolated the effect of the ambient light from each of these signals and averaged them to get a baseline ambient light measurement. We then estimated the proportion of ambient light contained within a given spectrum using the ratio of the ambient light peak to the broadband signal peak. The intensity of the ambient light peak was measured by isolating the signal around a consistent location and calculating the maximum value. The peak broadband signal was found by zeroing the known ambient light peaks and finding the resulting maximum signal intensity. We denote the ratio of ambient light to broadband light the AL-BL ratio. The AL-BL ratio was computed for the spectra collected prior to each scan and averaged to attain a baseline AL-BL ratio. To estimate the ambient light in each spectrum, the AL-BL ratio is calculated and used to scale the baseline ambient light according to the baseline AL-BL ratio. The scaled ambient light is then subtracted from the spectrum. To limit overprocessing the data, we calculated a general baseline ambient light and baseline AL-BL ratio using all 6 scans. Therefore each individual spectrum did not contain embedded pathology or scan-specific information in the form of a specific baseline ambient light spectrum. A second, simpler, technique looked to estimate the location of the peak ambient light and to eliminate it from each spectrum. The prerecorded spectra were analyzed and the ambient light spectra were extracted. The location of the peak ambient light was determined, and the associated wavelengths were zeroed with a buffer of $\sim \pm 3$ nm around the peak. We evaluated the effect of the various ambient light compensation methods on the classification metrics.

Feature dimensionality reduction

We examined the effect of various feature reduction techniques on classifier performance. We evaluated the classification performance after PCA dimensionality reduction and signal binning were applied. For consistency, all methods reduced the feature space from 2578 features, down to 100 features. For PCA, we used the PCA implementation from the scikit-learn library as seen previously with the number of components set to 100 [72]. For signal binning, we calculated 100 wavelength bins by dividing the length of the signal by 100. For each bin, we found the mean signal intensity for the wavelength range. We repeated this process for all remaining bins to produce a lower-resolution signal. We performed an ablation study which included a baseline with no preprocessing, each preprocessing method in isolation, and every combination of preprocessing methods.

4.3 Results of ex vivo human kidney tissue analysis and experiments

In this section, we discuss the results of our human kidney collection and analysis. We begin with our 3D PCA analysis of the dataset with respect to tissue metadata. We follow this with an analysis of preprocessing techniques and finally a discussion of train-test split experiments.

4.3.1 Data distribution analysis

Pathological and sample-based analysis

The results of our latent analysis of pathological class, and tissue sample metadata are shown in Figure 4.7. Colour, shape, and transparency are used as visual aids. On the left, we use orange and blue to highlight the cancerous and non-neoplastic tissue

respectively. The sample ID is visualized with the shape of the marker where samples 1, 2, and 3 are represented by circles, squares, and diamonds respectively. Shape fill is used to differentiate scans of the front and back of the sample. On the right, we use blue, red, and green to highlight samples 1, 2, and 3 respectively. Slightly lighter colours are used to differentiate scans of the front and back of the sample. The pathology is visualized with the shape of the marker where circles and crosses represent cancerous and non-neoplastic tissue respectively.

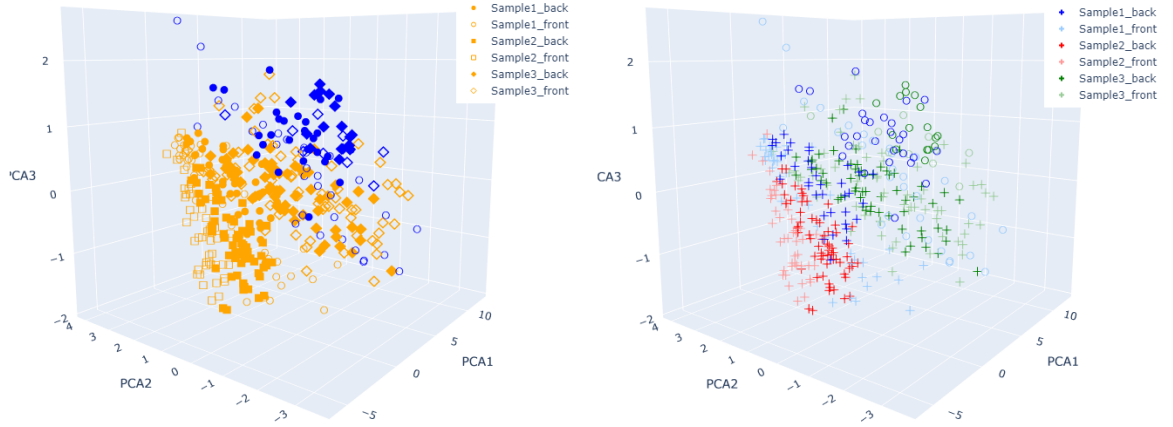


Figure 4.7: PCA latent space visualization of metadata. Pathology colourized (Left) and Sample colourized (right)

In Figure 4.7 (left) we observe loose clustering of pathologies, with cancerous spectra (orange) clustering in the bottom-left and non-neoplastic (blue) clustering in the top-right. These are promising observations that support the potential for successful differentiation based on pathology. However significant mixing of the pathologies is still present. Sample 3 in particular has cancerous regions which appear much more similar to the normal tissue than the cancerous tissue in other samples. Figure 4.7 (right) should ideally be perfectly spread out with no clustering. However, this is not

what we observe. We do observe good mixing between Sample 1 (blue) and Sample 3 (green), but some clustering is still visible. Sample 2 (red) contains only cancerous spectra, explaining the strong clustering. However, Sample 2 maintains visible clusters when compared to the cancerous components of Sample 1 and Sample 3. We also note some separation between the front and back scans of the same sample. This separation is most noticeable in Sample 1 and decreases with Sample 2 and Sample 3. Sample 1 was the first scan and took significantly longer than Sample 2 and Sample 3 due to inexperience with the standard operating procedure (SOP).

From these observations, we conclude that the recorded spectrum is impacted by the data collection process which contributes to variation between the scans. A rigorous and consistent standard operating procedure may help alleviate interscan variance, while future work should focus on normalizing their effects.

Analysis of pathological latent shift over time

The results of our analysis of latent shift over time are shown in Figure 4.8. The colour of the point is used to indicate the time elapsed since the start of each scan. A continuum of colours from purple to green to yellow is used to indicate chronology from the earliest to the latest point. Cancerous and non-neoplastic scans are plotted separately on the left and right graphs respectively.

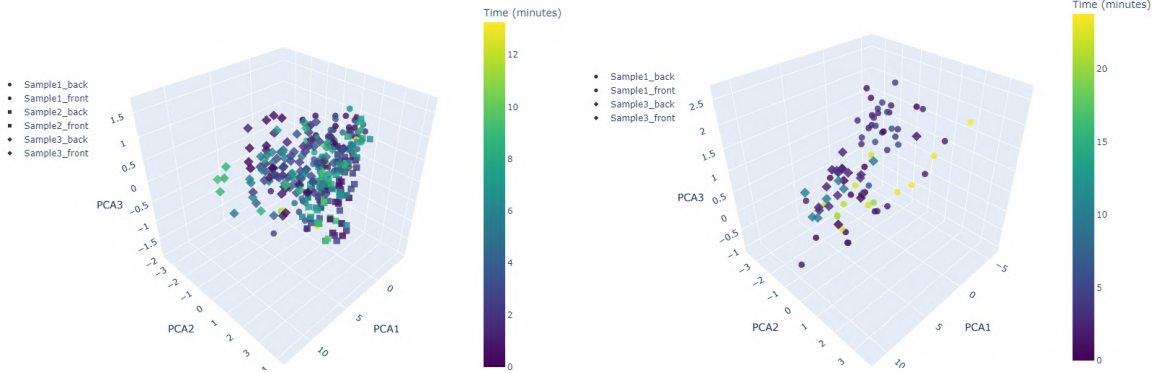


Figure 4.8: PCA latent space visualization as a function of time. Cancer (Left) and Non-neoplastic (right)

The cancerous tissue scans show no sign of a clear latent shift as a function of time. The latent space is observed to be randomly distributed. This trend holds when observing individual tissue scans. We make similar observations for the non-neoplastic tissue scans. However, an anomaly is observed in the scan of the front of sample 1. This was the first scan performed, which introduced significantly longer scan times. This long scan time resulted in significant moisture loss in the tissue. The tissue was partially scanned, allowed to rehydrate and surface dry before finishing the scan. We observe clustering between the two halves of the scan. We hypothesize that moisture loss is an additional factor to control for in future work. However, these results are limited as they are not able to efficiently isolate the time variables. Over time, different areas of the tissue are scanned. This results in changing spectra which adds noise and uncertainty to the observations. Future work should experiment with observing a single area of tissue as a function of time as well as controlling for moisture loss. Overall, the qualitative analysis of the distribution of the data through latent space visualization showed promising results.

4.3.2 Preprocessing experiments

We performed an ablation study of the ambient light compensation, min-max normalization, and dimensionality reduction methods. We present the balanced accuracy, sensitivity, specificity, F1 score, AUC, and accuracy for each combination of preprocessing methods. We repeat the process for scan-based and sample-based train-test splits. The results of our preprocessing ablation study are shown in Tables 4.2 and 4.3. The baseline is the LDA model with no additional preprocessing. The ambient light compensation methods are labelled ALE (Ambient light estimation) and APZ (Ambient peak zeroing). The Min-max normalization is denoted MM. The dimensionality reduction methods are labelled PCA (Principal component analysis) and Bin (wavelength binning). The results of the baseline model with no additional preprocessing methods are presented. Followed by each preprocessing technique in isolation. We analyzed the effect of using ambient light compensation techniques in addition to min-max normalization to compensate for probe-to-tissue distance. Finally, we analyzed the effects of dimensionality reduction on the performance of the classifier.

Table 4.2: The results of our preprocessing experimentation using a scan-based cross-validation regimen. The mean balanced accuracy, sensitivity, specificity, F1, AUC, and accuracy are reported.

Experiment	Bal Acc	Sens	Spec	F1	AUC	Acc
Baseline	85 ± 14	92 ± 14	89 ± 16	89 ± 10	78 ± 12	84 ± 14
ALE	87 ± 14	93 ± 10	89 ± 15	90 ± 9	80 ± 13	86 ± 13
APZ	86 ± 14	93 ± 10	89 ± 15	89 ± 9	79 ± 12	85 ± 12
MinMax (MM)	85 ± 15	92 ± 14	88 ± 16	88 ± 10	77 ± 13	84 ± 14
PCA	89 ± 11	95 ± 10	90 ± 13	91 ± 8	83 ± 11	88 ± 11
Binning (Bin)	90 ± 10	91 ± 10	94 ± 11	92 ± 7	85 ± 8	89 ± 9
ALE+MM	86 ± 13	94 ± 10	89 ± 15	90 ± 8	80 ± 11	86 ± 12
APZ+MM	86 ± 14	92 ± 11	89 ± 15	89 ± 9	79 ± 12	85 ± 13
PCA+ALE+MM	89 ± 11	95 ± 8	90 ± 13	92 ± 7	84 ± 11	88 ± 10
PCA+APZ+MM	89 ± 11	95 ± 9	91 ± 13	92 ± 7	84 ± 10	89 ± 10
Bin+ALE+MM	90 ± 9	93 ± 8	93 ± 10	92 ± 7	86 ± 7	90 ± 8
Bin+APZ+MM	91 ± 9	91 ± 10	94 ± 10	92 ± 7	86 ± 8	90 ± 9

Table 4.3: The results of our preprocessing experimentation using a sample-based cross-validation regimen. The mean balanced accuracy, sensitivity, specificity, F1, AUC, and accuracy are reported.

Experiment	Bal Acc	Sens	Spec	F1	AUC	Acc
Baseline	71 ± 20	73 ± 37	86 ± 19	69 ± 27	57 ± 3	64 ± 27
ALE	71 ± 19	75 ± 33	86 ± 20	72 ± 22	58 ± 6	66 ± 23
APZ	69 ± 20	71 ± 39	86 ± 20	67 ± 30	55 ± 2	63 ± 28
MinMax (MM)	71 ± 20	73 ± 37	86 ± 19	69 ± 27	57 ± 3	64 ± 27
PCA	70 ± 17	74 ± 32	86 ± 20	72 ± 21	58 ± 6	64 ± 21
Binning (Bin)	72 ± 20	79 ± 30	84 ± 17	76 ± 19	58 ± 2	69 ± 23
ALE+MM	72 ± 17	75 ± 32	86 ± 19	73 ± 21	60 ± 5	66 ± 22
APZ+MM	69 ± 20	71 ± 39	86 ± 20	67 ± 29	55 ± 2	62 ± 28
PCA+ALE+MM	65 ± 10	70 ± 28	86 ± 20	70 ± 15	60 ± 7	60 ± 13
PCA+APZ+MM	69 ± 14	73 ± 29	86 ± 20	72 ± 17	60 ± 7	64 ± 17
Bin+ALE+MM	76 ± 17	83 ± 23	86 ± 18	81 ± 14	65 ± 7	74 ± 19
Bin+APZ+MM	71 ± 21	80 ± 28	83 ± 18	77 ± 18	57 ± 5	69 ± 22

We observe that the train-test split has a significant impact on the model performance. Each train-test split preserves different statistical variations which can increase or decrease the difficulty of the classification problem. Splitting the data by scan number (front and back of each sample) preserves the statistical variation between scans caused by differences in the scanning procedure. Spectra collected regionally are separated into different folds which reduces signal overlap. Due to the thin tissue samples, some signal mixing may appear for longer wavelength ranges, limiting the absolute reliability of the performance metrics. Splitting the data strictly

by the sample number theoretically provides the most robust results. This method maximally preserves statistical variation between the training and testing sets while minimizing signal contamination. However, the size of our dataset limits its application. The intersample variation in pathological presentation greatly hinders the classification performances. The confusion matrices for the 3-fold cross-validation are shown in Figure 4.9.



Figure 4.9: Leave-one-sample-out confusion matrix. Non-neoplastic tissue is class 0 and cancerous tissue is class 1

These results can be explained by visual inspection of the tissue samples and the subsequent class unbalanced of the dataset (Figure 4.1). Sample 1 contains 2 visually distinct pathological presentations. The cancerous tissue presents with a yellow hue while the non-neoplastic tissue presents with a dark red hue. Sample 2 contains only cancerous tissue with a presentation similar to that of Sample 1. Sample 3 shows significant variation in pathological representation. The non-neoplastic tissue presents with a similar dark red hue as in Sample 1. However, the cancerous tissue presents quite differently from the other two samples and is primarily pink and red with some yellow mixed in. Furthermore, cancer is over-represented in the dataset. Sample 1 contains 77 cancerous and 56 non-neoplastic spectra, Sample 2 contains 112 cancerous spectra, and Sample 3 contains 126 cancerous and only 26 non-neoplastic

spectra. We oversample the training set to compensate for cancerous bias, however, this does not help with variation in pathological presentation. As a result, when the model trains on Sample 1+2 and tests on Sample 3, the model has not seen cancerous tissue with a red hue and thus classifies those spectra predominantly as normal tissue. On the contrary, When the model trains on Sample 2+3 and tests on Sample 1, the model is shown an abundance of cancerous representations but very limited non-neoplastic representations. Thus the model predicts primarily cancer.

Overall our experimentation shows promising results for the ability to differentiate between cancerous and non-neoplastic human kidney tissue ex vivo. For our limited dataset, the scan-based train-test split is the most representative method for the comparison of preprocessing techniques. It preserves an abundance of variation while still allowing the model to learn. Comparison between the sample-based and scan-based train-test splits highlights the limitations of our dataset. In particular, we do not have a large enough dataset to create a representative training and testing set. Further data collection is required to draw more robust conclusions.

Ambient light compensation

The results of the ambient light compensation techniques are shown in the 2nd and 3rd portions of Table 4.2. Overall, both ambient peak zeroing (APZ) and ambient light estimation (ALE) showed promising results. Each was able to successfully reduce the effects of ambient light within the dataset without hindering classifier performance. Ambient light compensation techniques improved the effectiveness of min-max normalization by significantly reducing or eliminating the primary ambient

light peak. This allowed for more effective signal-peak normalization. Both techniques also showed slight improvements in the performance of the classifier. APZ was the most effective in eliminating the peak ambient light, allowing for optimal results from min-max normalization. However, this technique does not robustly remove the ambient light. The remainder of the ambient light remains in the signal, potentially influencing future prediction. ALE does a much better job of compensating for the entire ambient light signal. However, it often struggled to entirely remove the largest peak. The estimation technique was performed broadly, using a single average ambient light spectrum for the entire dataset. This was done to remove the possibility of contaminating the dataset with class and scan-specific information. Since different tissues interacted with the ambient light differently, this technique was less effective on a dataset scale. The ideal ambient light compensation would record the ambient light directly before each measurement to allow for accurate compensation. A light source which can be rapidly digitally toggled would be ideal for this application.

Dimensionality reduction

The results of the dimensionality reduction are shown in the 2nd and 4th portions of Table 4.2. We observe that both dimensionality reduction techniques offered substantial improvements in classification performance over the baseline method. The PCA technique improved the balanced accuracy by 4% and the binning technique offered an increase of 5%. We observe that the PCA and binning dimensionality reduction techniques were both beneficial to the performance of the classifier. Both techniques reduce the complexity of the input data which is significant for small datasets. Our dataset contains just 399 spectra whereas each spectrum contains 2578 features. This

introduces significant variance and noise into the data which is difficult to generalize for many machine learning models. By reducing the dimensionality, we simplify the data, reducing high variance noise that can distract the model, and allowing it to focus on relevant features.

Chapter 5

Summary and Conclusions

This chapter presents a summary of the key conclusions, limitations, and future directions of our work.

5.1 Summary

The objective of this thesis was the development and evaluation of a tracked tissue sensing testbed prototype for navigated tumour margin assessment. The results presented in this work demonstrated the capabilities of our tracked tissue sensing testbed for accurate real-time navigated tissue inspection. The first contribution focused on the development and characterization of the testbed for real-time tissue assessment. We showed that our tracked diffuse reflection spectroscope was capable of real-time spatial classification and navigation of an arbitrary tissue volume using handheld interrogation methods. The second contribution focused on the application of the testbed for a clinically relevant tissue type. The results of our interrogation of human kidney tissue demonstrated the viability of the testbed for the rapid application to novel tissue types. We showed our tracked tissue sensor was capable of real-time data collection and exploration while offering promising potential for the classification of

cancerous and non-neoplastic kidney tissues.

A substantial gap exists between preoperative and intraoperative information for the surgical treatment of early-stage cancer. The capacity to intraoperatively assess tumour margin status has the potential to increase surgeon confidence in complete tumour resection, mitigating the probability of revision surgery. The development of accurate intraoperative tumour margin assessment tools has a crucial role to play in improving outcomes for early-stage cancer patients around the world.

5.2 Limitations

The testbed characterization was performed on a large and deep tissue phantom. This is not necessarily representative of the performance on a small and thin tissue sample such as in the case of residual cancer and margin detection. Challenges associated with accurately determining the scan location with a non-contact probe, accurately determining the system latencies, and accurately compensating for inherent position error will all influence the performance of the testbed. The sensitivity analysis was performed on thin opaque white paper rather than tissue. We cannot be sure that translucent, complex tissue structures are also invariant to the distance and angle of the probe. The results of our preliminary ex vivo human kidney experiment are limited in their generalizability. The dataset was limited to 3 tissue samples from the kidney of a single patient, with non-neoplastic tissue significantly underrepresented in the dataset. The variance between scans added a barrier to drawing meaningful conclusions about pathological separability. Increasing the size and quality of the dataset are key considerations for future studies. Furthermore, this study was performed ex vivo on frozen and thawed tissue. For ex vivo tumour margin assessment

this may cause some discrepancy. We also did not account for in vitro specifics such as physiological conditions, blood flow, and cauterized tissue. Finally, the labels for this preliminary trial were noisy. They were determined by a clinician through visual inspection and domain knowledge but did not have histopathological confirmation. The diffuse reflection may have also caused some signal mixing between the two classes given the small form factor of the sample.

5.3 Future work

The immediate next steps of this research should address the limitations of the current testbed implementation. A larger and more representative dataset should be acquired to improve the generalizability of the results and to allow for more comprehensive model training. A robustly trained classification model would enable the testbed to be characterized using human kidney tissue samples, providing valuable insights into testbed performance in a more representative environment. The ambient light in the data collection room was determined to contaminate the received signal with time and location information which could compromise the integrity of the dataset. Using a digitally toggleable light source to collect ambient light directly before each reading could greatly improve ambient light compensation. This information could also be potentially used for probe distance estimation, as we have demonstrated that the ratio of peak ambient light intensity to peak signal intensity depends on the probe's distance from the tissue surface. This metric could enhance the accuracy of 3D spatial classification by correcting for discrepancies between the assumed measurement location at the probe tip and the actual measurement location on the tissue surface. Real-time latency measurements could further improve

the temporal calibration of the testbed and consequently, the spatial classification accuracy. Other potential improvements include registering spatial classification with a real-time video feed of the scanned area, exploring augmented reality to overlay classifications directly on patients, and considering visual representations like surface heat maps as an alternative to point clouds for improved visual cues.

The next phase of our research will concentrate on three primary areas: expanding the testbed to incorporate additional devices and modalities, investigating advanced classification models, and examining potential applications for surgical training. The flexibility and modularity of our testbed enables the implementation and combination of various margin assessment technologies. We are particularly interested in combining the broadband spectroscopy with temporally enhanced ultrasound. Previous studies, such as those by Connolly et al. (2022) [66] and Wilson et al. (2022) [75] have demonstrated the potential of multi-modal combinations. Connolly et al. (2022) [66] experimented with using throughput broadband spectroscopy and temporally enhanced ultrasound for multilayer tissue differentiation. And S. Wilson et al. (2022) [75] created a device to focus point-based electromagnetic techniques within a 2D ultrasound slice. Both may provide insight into potential future directions for our testbed. In addition to hardware enhancements, we aim to explore state-of-the-art classification models. Our work primarily focused on basic machine learning models, with limited exploration of deep learning models and complex training pipelines. Enhancing the testbed by integrating attention CNNs or transformer-based models with self-supervised learning and uncertainty estimation would greatly improve its usability and interoperability. Lastly, we see potential for our system in the domain of surgical training. Apart from its direct application in the operating theatre, our

system could function as a training apparatus for surgical procedures. The training environment permits substantial simplifications. For instance, UV-reflective ink could be injected into a tissue phantom. This allows the light source to be replaced with a narrow-band UV light source, significantly reducing the classification complexity and leading to robust classification. The training apparatus could then be used with high confidence to assess simulated tumour resections on the phantom.

References

- [1] World health organization (WHO), *Cancer*, 2021. [Online]. Available: <https://www.who.int/news-room/fact-sheets/detail/cancer>.
- [2] S. Pieper, M. Halle, and R. Kikinis, “3D Slicer,” *2004 2nd IEEE International Symposium on Biomedical Imaging: Macro to Nano (IEEE Cat No. 04EX821)*, vol. 2, pp. 632–635, DOI: 10.1109/ISBI.2004.1398617.
- [3] T. Ungi, A. Lasso, and G. Fichtinger, “Open-source platforms for navigated image-guided interventions,” *Medical Image Analysis*, vol. 33, pp. 181–186, 2016. DOI: 10.1016/j.media.2016.06.011.
- [4] J. Tokuda, G. S. Fischer, X. Papademetris, *et al.*, “OpenIGTLink: an open network protocol for image-guided therapy environment,” *The International Journal of Medical Robotics and Computer Assisted Surgery*, vol. 5, no. 4, pp. 423–434, 2009. DOI: 10.1002/RCS.274.
- [5] A. Lasso, T. Heffter, A. Rankin, C. Pinter, T. Ungi, and G. Fichtinger, “PLUS: Open-source toolkit for ultrasound-guided intervention systems,” *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 10, pp. 2527–2537, 2014. DOI: 10.1109/TBME.2014.2322864.

- [6] T. Kim and J. E. Zuckerman, “Realizing the potential of telemedicine in global health,” *Journal of Global Health*, vol. 9, no. 2, p. 20307, 2019. DOI: 10.7189/jogh.09.020307.
- [7] A. Gropper, “Open-source health care software,” *American Medical Association Journal of Ethics*, vol. 13, no. 9, pp. 632–636, 2011. DOI: 10.1001/VIRTUALMENTOR.2011.13.9.STAS1-1109.
- [8] World health organization (WHO), *Promoting cancer early diagnosis*. [Online]. Available: <https://www.who.int/activities/promoting-cancer-early-diagnosis>.
- [9] American Cancer Society, *Cancer Treatment & Survivorship Facts & Figures 2019-2021*, 2019. [Online]. Available: <https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/cancer-treatment-and-survivorship-facts-and-figures/cancer-treatment-and-survivorship-facts-and-figures-2019-2021.pdf>.
- [10] J.-H. Chen, S. Chan, Y. Zhang, S. Li, R.-F. Chang, and M.-Y. Su, “Evaluation of breast stiffness measured by ultrasound and breast density measured by MRI using a prone-supine deformation model,” *Biomarker Research*, vol. 7, no. 1, p. 20, 2019. DOI: 10.1186/s40364-019-0171-1.
- [11] Y. Guo, R. Sivaramakrishna, C.-C. Lu, J. S. Suri, and S. Laxminarayan, “Breast image registration techniques: a survey,” *Medical and Biological Engineering and Computing*, vol. 44, no. 1, pp. 15–26, 2006. DOI: 10.1007/s11517-005-0016-y.
- [12] R. Orosco, V. Tapia, and J. Califano, *et al.*, “Positive Surgical Margins in the 10 Most Common Solid Cancers,” *Scientific Reports 2018 8:1*, vol. 8, no. 1, pp. 1–9, 2018. DOI: 10.1038/s41598-018-23403-5.

- [13] J. L. Oh, “Multifocal or Multicentric Breast Cancer: Understanding Its Impact on Management and Treatment Outcomes,” *Methods of Cancer Diagnosis, Therapy and Prognosis*, pp. 583–587, 2008. DOI: 10.1007/978-1-4020-8369-3_40.
- [14] National Cancer Institute, *What Is Cancer?* 2021. [Online]. Available: <https://www.cancer.gov/about-cancer/understanding/what-is-cancer>.
- [15] National Cancer Institute, *Cancer Staging*, 2022. [Online]. Available: <https://www.cancer.gov/about-cancer/diagnosis-staging/staging>.
- [16] M. S. Moran, S. J. Schnitt, A. E. Giuliano, *et al.*, “Society of surgical oncology–american society for radiation oncology consensus guideline on margins for breast-conserving surgery with whole-breast irradiation in stages i and ii invasive breast cancer,” *International Journal of Radiation Oncology*Biology*Physics*, vol. 88, no. 3, pp. 553–564, 2014. DOI: 10.1016/j.ijrobp.2013.11.012.
- [17] A. C. Hargreaves, M. Mohamed, and R. A. Audisio, “Intra-operative guidance: Methods for achieving negative margins in breast conserving surgery,” *Journal of Surgical Oncology*, vol. 110, no. 1, pp. 21–25, 2014. DOI: 10.1002/JSO.23645.
- [18] A. J. C. on Cancer, *AJCC Cancer Staging Manual*, 7th, S. B. Edge, D. R. Byrd, C. C. Compton, A. G. Fritz, F. L. Greene, and I. Andy Trotti, Eds. 2010, Page 152, isbn:978-0-387-88440-0.
- [19] L.-J. Tseng, A. Matsuyama, and V. MacDonald-Dickinson, “Histology: The gold standard for diagnosis?” *Can Vet J*, vol. 64, no. 4, pp. 389–391, 2023.
- [20] R. Thavarajah, V. K. Mudimbaimannar, J. Elizabeth, U. K. Rao, and K. Ranganathan, “Chemical and physical basics of routine formaldehyde fixation,” *J Oral Maxillofac Pathol*, vol. 16, no. 3, pp. 400–405, 2012.

- [21] H. Sung, J. Ferlay, R. L. Siegel, *et al.*, “Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries,” *CA: A Cancer Journal for Clinicians*, vol. 71, no. 3, pp. 209–249, 2021. DOI: 10.3322/CAAC.21660.
- [22] D. M. Black, K. K. Hunt, and E. A. Mittendorf, “Long term outcomes reporting the safety of breast conserving therapy compared to mastectomy: 20-year results of EORTC 10801,” *Gland Surgery*, vol. 2, no. 3, p. 120, 2013. DOI: 10.3978/J.ISSN.2227-684X.2013.06.01.
- [23] N. C. Institute, *Pathology reports*, 2022. [Online]. Available: <https://www.cancer.gov/about-cancer/diagnosis-staging/diagnosis/pathology-reports-fact-sheet#how-is-tissue-obtained-for-examination-by-a-pathologist>.
- [24] Y. K. Tao, D. Shen, Y. Sheikine, *et al.*, “Assessment of breast pathologies using nonlinear microscopy,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 111, no. 43, pp. 15 304–15 309, 2014. DOI: 10.1073/PNAS.1416955111/-/DCSUPPLEMENTAL.
- [25] J. M. Jorns, D. Visscher, M. Sabel, *et al.*, “Intraoperative frozen section analysis of margins in breast conserving surgery significantly decreases reoperative rates: one-year experience at an ambulatory surgical center,” *American Journal of Clinical Pathology*, vol. 138, no. 5, pp. 657–669, 2012. DOI: 10.1309/AJCP4IEMXCJ1GDTS.
- [26] R. Verma, M. Kumar, S. S. Khan, and A. S. Sachdev, “Imprint cytology: an appraisal,” *Journal of Oral Medicine, Oral Surgery, Oral Pathology and Oral Radiology*, vol. 2, no. 2, p. 110, 2016. DOI: 10.5958/2395-6194.2016.00016.3.
- [27] K. Esbona, Z. Li, and L. G. Wilke, “Intraoperative Imprint Cytology and Frozen Section Pathology for Margin Assessment in Breast Conservation Surgery: A

- Systematic Review,” *Annals of Surgical Oncology*, vol. 19, no. 10, pp. 3236–3245, 2012. DOI: 10.1245/s10434-012-2492-2.
- [28] J. Heidkamp, M. Scholte, C. Rosman, S. Manohar, J. J. Fütterer, and M. M. Rovers, “Novel imaging techniques for intraoperative margin assessment in surgical oncology: A systematic review,” *International Journal of Cancer*, vol. 149, no. 3, pp. 635–645, 2021. DOI: 10.1002/ijc.33570.
- [29] A. M. De Schepper, L. De Beuckeleer, J. Vandevenne, and J. Somville, “Magnetic resonance imaging of soft tissue tumors,” *European Radiology*, vol. 10, no. 2, pp. 213–223, 2000. DOI: 10.1007/s003300050037.
- [30] B. W. Maloney et al., “Review of methods for intraoperative margin detection for breast conserving surgery,” *Journal of Biomedical Optics*, vol. 23, no. 10, 2018. DOI: 10.1117/1.JBO.23.10.100901.
- [31] S. Q. Qiu, M. D. Dorrius, S. J. de Jongh, *et al.*, “Micro-computed tomography (micro-CT) for intraoperative surgical margin assessment of breast cancer: A feasibility study in breast conserving surgery,” *European Journal of Surgical Oncology*, vol. 44, no. 11, pp. 1708–1713, 2018. DOI: 10.1016/J.EJSO.2018.06.022.
- [32] R. Tang, J. M. Buckley, L. Fernandez, *et al.*, “Micro-computed tomography (Micro-CT): a novel approach for intraoperative breast cancer specimen imaging,” *Breast Cancer Research and Treatment*, vol. 139, no. 2, pp. 311–316, 2013. DOI: 10.1007/S10549-013-2554-6.
- [33] J. E. Kalinyak, W. A. Berg, K. Schilling, K. S. Madsen, D. Narayanan, and M. Tartar, “Breast cancer detection using high-resolution breast PET compared to whole-body PET or PET/CT,” *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 41, no. 2, pp. 260–275, 2014. DOI: 10.1007/s00259-013-2553-1.

- [34] R. M. Sigrist, J. Liao, A. E. Kaffas, M. C. Chammas, and J. K. Willmann, "Ultrasound Elastography: Review of Techniques and Clinical Applications," *Theranostics*, vol. 7, no. 5, p. 1303, 2017. DOI: 10.7150/THNO.18650.
- [35] S. Azizi, S. Bayat, P. Yan, *et al.*, "Deep recurrent neural networks for prostate cancer detection: Analysis of temporal enhanced ultrasound," *IEEE Transactions on Medical Imaging*, vol. 37, no. 12, pp. 2695–2703, 2018. DOI: 10.1109/TMI.2018.2849959.
- [36] Z. Takáts *et al.*, "Mass spectrometry sampling under ambient conditions with desorption electrospray ionization," *Science*, vol. 306, no. 5695, pp. 471–473, 2004. DOI: 10.1126/SCIENCE.1104404.
- [37] H. S. Kang, S. C. Lee, Y. S. Park, *et al.*, "Protein and lipid MALDI profiles classify breast cancers according to the intrinsic subtype.," *BMC Cancer*, vol. 11, p. 465, 2011. DOI: 10.1186/1471-2407-11-465.
- [38] J. Zhang, J. Rector, J. Q. Lin, *et al.*, *Science Translational Medicine*, vol. 9, no. 406, eaan3968, 2017. DOI: 10.1126/scitranslmed.aan3968.
- [39] P. Saudemont, J. Quanico, Y.-M. Robin, Z. Takats, M. Salzet, and I. Fournier Correspondence, "Real-Time Molecular Diagnosis of Tumors Using Water-Assisted Laser Desorption/Ionization Mass Spectrometry Technology," DOI: 10.1016/j.ccell.2018.09.009.
- [40] J. Balog, L. Sasi-Szabó, J. Kinross, *et al.*, "Intraoperative tissue identification using rapid evaporative ionization mass spectrometry," *Science Translational Medicine*, vol. 5, no. 194, 2013. DOI: 10.1126/SCITRANSLMED.3005623.
- [41] J. Ehrlich, C. Yeung, M. Kaufman, *et al.*, "Determining the time-delay of a mass spectrometry-based tissue sensor," in *Medical Imaging 2023: Image-Guided Procedures, Robotic Interventions, and Modeling*, C. A. Linte and J. H. Siewerdsen,

- Eds., International Society for Optics and Photonics, vol. 12466, SPIE, 2023, p. 124661D. DOI: 10.1117/12.2654359.
- [42] G. W. Auner, S. K. Koya, C. Huang, *et al.*, “Applications of Raman spectroscopy in cancer diagnosis,” *Cancer and Metastasis Reviews*, vol. 37, no. 4, pp. 691–717, 2018. DOI: 10.1007/s10555-018-9770-9.
 - [43] P. Gao, B. Han, Y. Du, *et al.*, “The Clinical Application of Raman Spectroscopy for Breast Cancer Detection,” *Journal of Spectroscopy*, vol. 2017, C. Krafft, Ed., p. 5383948, 2017. DOI: 10.1155/2017/5383948.
 - [44] I. A. Birtoiu, C. Rizea, D. Togoe, *et al.*, “Diagnosing clean margins through Raman spectroscopy in human and animal mammary tumour surgery: a short review,” *Interface Focus*, vol. 6, no. 6, 2016. DOI: 10.1098/RSFS.2016.0067.
 - [45] A. S. Haka, Z. Volynskaya, J. A. Gardecki, *et al.*, “In vivo Margin Assessment during Partial Mastectomy Breast Surgery Using Raman Spectroscopy,” *Cancer Research*, vol. 66, no. 6, pp. 3317–3322, 2006. DOI: 10.1158/0008-5472.CAN-05-2815.
 - [46] S. Aumann, S. Donner, J. F., and F. Müller, “Optical Coherence Tomography (OCT): Principle and Technical Realization,” in *High Resolution Imaging in Microscopy and Ophthalmology: New Frontiers in Biomedical Optics*, J. F. Bille, Ed. Cham: Springer International Publishing, 2019, pp. 59–85. DOI: 10.1007/978-3-030-16638-0_3.
 - [47] D. Mojahed, R. S. Ha, P. Chang, *et al.*, “Fully Automated Postlumpectomy Breast Margin Assessment Utilizing Convolutional Neural Network Based Optical Coherence Tomography Image Classification Method,” *Academic Radiology*, vol. 27, no. 5, e81–e86, 2020. DOI: 10.1016/j.acra.2019.06.018.

- [48] F. T. Nguyen, A. M. Zysk, E. J. Chaney, *et al.*, “Intraoperative Evaluation of Breast Tumor Margins with Optical Coherence Tomography,” *Systems Biology and Emerging Technologies*, 2009. DOI: 10.1158/0008-5472.CAN-08-4340.
- [49] M. Thill, “MarginProbe®: Intraoperative margin assessment during breast conserving surgery by using radiofrequency spectroscopy,” *Expert Review of Medical Devices*, vol. 10, no. 3, pp. 301–315, 2013. DOI: 10.1586/erd.13.5.
- [50] J. Wang, L. Zhang, and Z. Pan, “Evaluating the impact of radiofrequency spectroscopy on reducing reoperations after breast conserving surgery: A meta-analysis,” *Thoracic Cancer*, vol. 14, no. 16, pp. 1413–1419, 2023. DOI: 10.1111/1759-7714.14890.
- [51] T. Allweis, Z. Kaufman, S. Lelcuk, *et al.*, “A prospective, randomized, controlled, multicenter study of a real-time, intraoperative probe for positive margin detection in breast-conserving surgery,” *The American Journal of Surgery*, vol. 196, no. 4, pp. 483–489, 2008. DOI: 10.1016/j.amjsurg.2008.06.024.
- [52] J. M. Dixon, L. Renshaw, O. Young, *et al.*, “Intra-operative assessment of excised breast tumour margins using clearedge imaging device,” *European Journal of Surgical Oncology (EJSO)*, vol. 42, no. 12, pp. 1834–1840, 2016. DOI: 10.1016/j.ejso.2016.07.141.
- [53] Z. Du, H. Wan, Y. Chen, Y. Pu, and X. Wang, “Bioimpedance spectroscopy can precisely discriminate human breast carcinoma from benign tumors,” *Medicine*, vol. 96, no. 4, e5970, 2017. DOI: 10.1097/MD.0000000000005970.
- [54] L. De Boer, B. Molenkamp, T. Bydlon, *et al.*, “Fat/water ratios measured with diffuse reflectance spectroscopy to detect breast tumor boundaries,” *Breast Cancer Research and Treatment*, vol. 152, DOI: 10.1007/s10549-015-3487-z.

- [55] C. Palmer, *Diffraction Grating Handbook (8th edition)*. Richardson Gratings, Newport Corporation, 2014, ch. 2.1.
- [56] A. M. Franz, T. Haidegger, W. Birkfellner, K. Cleary, T. M. Peters, and L. Maier-Hein, “Electromagnetic tracking in medicine—a review of technology, validation, and applications,” *IEEE Transactions on Medical Imaging*, vol. 33, no. 8, pp. 1702–1725, 2014. DOI: 10.1109/TMI.2014.2321777.
- [57] A. Sorriento, M. B. Porfido, S. Mazzoleni, *et al.*, “Optical and Electromagnetic Tracking Systems for Biomedical Applications: A Critical Review on Potentialities and Limitations,” *IEEE Reviews in Biomedical Engineering*, vol. 13, pp. 212–232, 2020. DOI: 10.1109/RBME.2019.2939091.
- [58] K. Cleary and T. M. Peters, “Image-guided interventions: Technology review and clinical applications,” *Annual Review of Biomedical Engineering*, vol. 12, no. 1, pp. 119–142, 2010, PMID: 20415592. DOI: 10.1146/annurev-bioeng-070909-105249.
- [59] F. Poulin and L.-P. Amiot, “Interference during the use of an electromagnetic tracking system under OR conditions,” *Journal of biomechanics*, vol. 35, no. 6, pp. 733–737, 2002. DOI: 10.1016/s0021-9290(02)00036-2.
- [60] NDI, *3D Guidance [®] Tracking Technology for Your Most Realistic Training or Simulation System*. [Online]. Available: <https://ndicorpstg.wpengine.com/wp-content/uploads/2020/09/10005584-3D-Guidance-Simulation.pdf>.
- [61] D. Morton, L. Connolly, L. Groves, *et al.*, “Tracked tissue sensing for tumor bed inspection,” in *Medical Imaging 2023: Image-Guided Procedures, Robotic Interventions, and Modeling*, C. A. Linte and J. H. Siewerdsen, Eds., International Society for Optics and Photonics, vol. 12466, SPIE, 2023, 124661K. DOI: 10.1117/12.2654217.

- [62] Thorlabs, *Compact Stabilized Broadband Light Sources*, Thorlabs. [Online]. Available:
https://www.thorlabs.com/newgrouppage9.cfm?objectgroup_id=7269.
- [63] Thorlabs, *Compact CCD Spectrometers*, Thorlabs. [Online]. Available: https://www.thorlabs.com/newgrouppage9.cfm?objectgroup_id=3482&pn=CCS100.
- [64] Aventes, *Fcr-7uvir200-2*. [Online]. Available: https://avantesusa.com/wp-content/uploads/2019/01/Datasheet_Reflection_Probes.pdf.
- [65] Y. Ito, R. P. Kennan, E. Watanabe, and H. Koizumi, “Assessment of heating effects in skin during continuous wave near-infrared spectroscopy,” *Journal of Biomedical Optics*, vol. 5, no. 4, pp. 383–390, 2000. DOI: 10.1117/1.1287730.
- [66] L. Connolly, A. Jamzad, A. Nikniazi, *et al.*, “Feasibility of combined optical and acoustic imaging for surgical cavity scanning,” in *Medical Imaging 2022: Image-Guided Procedures, Robotic Interventions, and Modeling*, International Society for Optics and Photonics, vol. 12034, SPIE, 2022, 120341H. DOI: 10.1117/12.2611964.
- [67] C. J. Frank, D. C. Redd, T. S. Gansler, and R. L. McCreery, “Characterization of human breast biopsy specimens with near-ir raman spectroscopy,” *Analytical chemistry*, vol. 66, no. 3, pp. 319–326, 1994.
- [68] P.-D. Cernelev, K. Moga, L. Groves, *et al.*, “Determining boundaries of accurate tracking for electromagnetic sensors,” in *Medical Imaging 2023: Image-Guided Procedures, Robotic Interventions, and Modeling*, C. A. Linte and J. H. Siewerdsen, Eds., International Society for Optics and Photonics, vol. 12466, SPIE, 2023, pp. 386–391. DOI: 10.1117/12.2654428.

- [69] Scikit-learn Library, *k-Nearest-Neighbours Classifier*. [Online]. Available: <https://scikit-learn.org/stable/modules/generated/sklearn.neighbors.KNeighborsClassifier.html>.
- [70] S. Xu, M. Perez, K. Yang, C. Perrenot, J. Felblinger, and J. Hubert, "Determination of the latency effects on surgical performance and the acceptable latency levels in telesurgery using the dV-Trainer® simulator," *Surgical Endoscopy*, vol. 28, no. 9, pp. 2569–2576, 2014. DOI: 10.1007/s00464-014-3504-z.
- [71] I. T. Jolliffe and J. Cadima, "Principal component analysis: A review and recent developments," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 374, no. 2065, p. 20150202, 2016. DOI: 10.1098/rsta.2015.0202.
- [72] Scikit-learn Library, *Principle Component Analysis*. [Online]. Available: <https://scikit-learn.org/stable/modules/generated/sklearn.decomposition.PCA.html>.
- [73] Scikit-learn Library, *Linear Discriminant Analysis*. [Online]. Available: https://scikit-learn.org/stable/modules/generated/sklearn.discriminant_analysis.LinearDiscriminantAnalysis.html.
- [74] Imbalanced learn Library, *Random Over Sampler*. [Online]. Available: https://imbalanced-learn.org/stable/references/generated/imblearn.over_sampling.RandomOverSampler.html.
- [75] S. Wilson, C. Park, K. Barker, *et al.*, "Development of a novel, dual-modality image guidance system by combining a focused gamma probe with ultrasound imaging," in *Medical Imaging 2022: Image-Guided Procedures, Robotic Interventions, and Modeling*, C. A. Linte and J. H. Siewerdsen, Eds., International Society for Optics and Photonics, vol. 12034, SPIE, 2022, 120341F. DOI: 10.1117/12.2612730.

Appendix A

A.1 Standard operative procedures

Purpose

The purpose of this document is to provide insight into the desired procedure to collect tissue data using a spectroscopic probe.

Procedure

1. Tissue paper is placed in the center of the stage.
2. The skin specimen is placed atop the tissue paper in the center of the stage.
3. Optical imaging (photography) of the specimen. Please ensure the specimen and any measurement references are fully visible.
4. Position the region of interest (ROI) under the probe using the two knobs on the stage. The laser should directly illuminate the ROI.
5. Once in position:
 - (a) Lower the probe until it is within 5mm of the tissue:
 - Closer = smaller collection area = cleaner readings.
 - The probe should not touch the tissue surface.
 - (b) Tell the computer operator the tissue type (eg. cancer, healthy).
 - (c) The operator will request the probe be lowered or raised to achieve optimal signal strength.
6. Once confirmed, the operator will record the region of interest for 1 second.
7. Repeat steps 4-6 for all regions of interest. Repeat steps 1-6 for each tissue sample to be scanned.

A.2 PLUS configuration file

```
<PlusConfiguration version="2.0">
  <DataCollection StartupDelaySec="2.0" >
    <DeviceSet
      Name="PlusServer: ThorLabs spectrometer"
      Description="Broadcasting acquired spectrum through
        OpenIGTLink. First image row contains wavelength value,
        second image row contains intensities." />
    <Device
      Id="VideoDevice"
      Type="ThorLabsVideo"
      AcquisitionRate="30"
      IntegrationTimeSec="0.1"
      AveragedItemsForFiltering="20"
      > <!-- IntegrationTimeSec default="0.05" -->
      <!-- AcquisitionRate Default="30" -->
      <!-- AveragedItemsForFiltering Default="20"-->
      <DataSources>
        <DataSource Type="Video" Id="Video" />
      </DataSources>
      <OutputChannels>
        <OutputChannel Id="VideoStream" VideoDataSourceId="Video" />
      </OutputChannels>
    </Device>
    <Device
      Id="CaptureDevice"
      Type="VirtualCapture"
      BaseFilename="RecordingTest.igs.mha"
      EnableCapturingOnStart="FALSE" >
      <InputChannels>
        <InputChannel Id="VideoStream" />
      </InputChannels>
    </Device>
  </DataCollection>
  <PlusOpenIGTLinkServer
    MaxNumberOfIgtlMessagesToSend="1"
    MaxTimeSpentWithProcessingMs="50"
    ListeningPort="18944"
    SendValidTransformsOnly="true"
    OutputChannelId="VideoStream" >
    <DefaultClientInfo>
      <MessageTypes>
        <Message Type="IMAGE" />
      </MessageTypes>
      <ImageNames>
        <Image Name="Image" EmbeddedTransformToFrame="Image" />
      </ImageNames>
    </DefaultClientInfo>
  </PlusOpenIGTLinkServer>
</PlusConfiguration>
```

Figure A.1: Plus configuration script used to integrate the spectrometer and electromagnetic tracker into the 3D slicer testbed