POLYTECHNIQUE MONTRÉAL
affiliée à l'Université de Montréal

Intraoperative confirmation of prostate cancer using AI-based Raman spectroscopy for tumor targeting

DAVID ORLANDO GRAJALES LOPERA
Institut de génie biomédical

Thèse présentée en vue de l'obtention du diplôme de *Philosophiæ Doctor*

Génie biomédical

Juin 2025

POLYTECHNIQUE MONTRÉAL
affiliée à l'Université de Montréal

Cette thèse intitulée :

Intraoperative confirmation of prostate cancer using AI-based Raman spectroscopy for tumor targeting

présentée par **David Orlando GRAJALES LOPERA**
en vue de l'obtention du diplôme de *Philosophiæ Doctor*
a été dûment acceptée par le jury d'examen constitué de :

Stephan REUTER, président
Samuel KADOURY, membre et directeur de recherche
Cynthia MÉNARD, membre et codirectrice de recherche
Benjamin DE LEENER, membre
Peter L. CHOYKE, membre externe

DEDICATION

To all volunteer patients
who enable science to advance . . .

ACKNOWLEDGEMENTS

Thinking back —from learning to read and play the flute to writing this doctoral thesis— many people have contributed to my personal and professional growth. Furthermore, if we consider that one also learns from challenges (and bad experiences), the list becomes even longer. For this reason, to whom it may concern... Thank you.

I would like to express my gratitude to my research director, Samuel Kadoury, and co-director, Cynthia Ménard, for their patience, support, and the learning opportunities offered throughout this long journey. I also wish to acknowledge the entire team of professors, physicians, technicians, and students who were part of the research team. I am grateful to Polytechnique Montréal and the Centre de recherche du Centre hospitalier de l'Université de Montréal (CRCHUM) for the environment and for the research and teaching opportunities they provided. Special thanks also to the Institut TransMedTech de Montréal, the Collaborative Health Research Program (NSERC and CIHR), and the Fonds de recherche du Québec (FRQ) for funding this project, which, as an international student, is especially valuable and essential.

In summary, to the patients, professors, research team members, family, family (this is not a typo), friends, colleagues, acquaintances, and my students...

Gracias!

RÉSUMÉ

Cibler précisément les tissus malins lors des interventions pour le cancer de la prostate (CaP) demeure un défi clinique majeur, en particulier lorsque les tumeurs ne sont pas discernables des tissus normaux dans les conditions standards de la salle d'intervention. Les procédures actuelles, par exemple pour un échantillonnage ou traitement focal ciblé, bénéficient de plusieurs systèmes de navigation, dont certains intègrent l'échographie transrectale (TRUS), l'enregistrement multimodale d'images et le suivi électromagnétique (EM) pour guider les interventions. Toutefois, ces approches reposent principalement sur des images préopératoires pour la planification et sur l'histopathologie postopératoire pour la confirmation. En conséquence, il manque un mécanisme de confirmation en temps réel et in-situ permettant d'ajuster avec précision le ciblage tumoral. L'accès à de telles informations peropératoires pourrait améliorer la prise de décision, accroître le rendement diagnostique des biopsies, personnaliser les traitements et, ultimement, améliorer les résultats pour les patients.

La spectroscopie Raman (RS) constitue une technique optique non destructive et sans agent de contraste, capable de contribuer à la résolution de ce défi. Elle repose sur la diffusion inélastique de la lumière, où une petite fraction de photons interagit avec les modes de vibrations des molécules de l'échantillon, produisant de subtils changements d'énergie connus sous le nom d'effet Raman, et fournissant des signatures moléculaires en temps réel. Cette information moléculaire peut être exploité par des méthodes de classification basées sur l'intelligence artificielle (IA) pour appuyer la caractérisation tissulaire et la détection de maladies. Associée à l'apprentissage automatique, la RS a montré des résultats prometteurs pour l'identification de plusieurs pathologies, notamment la maladie d'Alzheimer, les troubles cardiovasculaires et divers types de cancers.

Avec plus de 1,4 million de nouveaux cas diagnostiqués chaque année, le CaP représente la deuxième forme de cancer la plus fréquente chez l'homme à l'échelle mondiale, après le cancer du poumon. Comme pour tout cancer, une détection précoce et une localisation précise des tumeurs sont des facteurs déterminants pour le succès du traitement. Cette thèse explore l'intégration de la RS à des modèles de classification fondés sur l'IA afin de permettre une confirmation du CaP en temps réel, localisée et in-situ, dans le cadre d'interventions ciblées sur la tumeur. À cette fin, la recherche a été structurée autour de quatre objectifs, associés à quatre articles de journaux, au cours de deux phases méthodologiques : Phase I, axée sur des expériences in-vivo et des modèles d'apprentissage automatique (machine learning) pour la détection de CaP à l'aide de données multimodales ; Phase II, axée sur des expériences ex-

vivo assistées par robot, des modèles d'apprentissage profond (deep learning) et des stratégies d'apprentissage par transfert pour la classification spectrale.

Dans la partie initiale de la Phase I, un prototype de système de navigation intégrant la fusion d'images et le suivi EM a été évalué lors de sa mise en œuvre clinique pour la curiethérapie à haut débit (HD). Sa performance a été analysée sur 43 procédures en termes d'efficacité du flux de travail, de précision du recalage d'image et de précision de la reconstruction du cathéter basée sur le suivi EM. Cette évaluation a mis en évidence l'intégration transparente du système dans les flux de travail cliniques existants avec une courbe d'apprentissage minimale, la meilleure identification des cathéters grâce au suivi EM et les avantages du recalage déformable par rapport à l'alignement rigide. Cette contribution a été essentielle non seulement pour établir l'utilité clinique du système, mais aussi pour soutenir les phases ultérieures du projet.

L'étape suivante de cette phase a consisté en une étude de faisabilité prospective visant la valeur diagnostique de la RS in-vivo combinée à des caractéristiques radiomiques basées sur l'imagerie par résonance magnétique multiparamétrique (IRMmp). À l'aide d'un système de navigation développé par l'équipe, intégrant la fusion TRUS-IRM et le suivi EM, une sonde Raman a permis l'acquisition de spectres ciblés chez 18 patients, pendent curiethérapie HD. Ces spectres, combinés aux radiomiques et aux biopsies de confirmation, ont été utilisés pour entraîner des machines à vecteur de support (SVM) pour la détection du CaP. L'approche multimodale a révélé la complémentarité de l'information mésoscopique/macroscopique, avec une précision atteignant 0,83 pour la détection des cancers cliniquement significatifs, surpassant les modèles unimodaux. La stratégie de sélection de caractéristiques a également contribué à de bonnes performances tout en limitant la complexité des modèles.

Dans la Phase II, ex-vivo, nous avons d'abord mis en œuvre un système de RS assisté par robot pour la caractérisation automatique d'échantillons de tissus frais prélevés par biopsie. Un bras robotique à 6 degrés de liberté et une sonde optique statique ont permis une acquisition précise et rapide des spectres. Une erreur de calibration 3D moyenne de 0,20 mm a été obtenue, ce qui est particulièrement pertinent au vu du faible diamètre des biopsies ($\sim$0,8 mm). Cette précision facilite la manipulation délicate des tissus et l'assignation d'étiquettes (ground truth labels) spécifiques à chaque site, condition essentielle à l'entraînement rigoureux des modèles. Parallèlement, une première exploration des réseaux neuronaux convolutifs unidimensionnels (1D-CNN) a été menée pour la détection du cancer en utilisant des ensembles de données rétrospectives de RS provenant de tissus cérébraux, mammaires et prostatiques. Cette analyse préliminaire a mis en évidence des performances supérieures des modèles CNN, en particulier pour les tissus cérébraux et mam-

maires, comparativement aux classificateurs SVM traditionnels et a souligné la capacité du 1D-CNN à apprendre les caractéristiques spectrales automatiquement.

Pour conclure cette phase, le système de RS assisté par robot a été déployé dans la salle de procédure pour une étude de faisabilité ex-vivo prospective impliquant des échantillons de biopsie de 10 patients pendant curiethérapie HD. Les spectres Raman acquis et les annotations histopathologiques spécifiques au site ont constitué un ensemble de données prospectives sur le CaP, qui a servi de domaine cible. Un modèle 1D-CNN, adapté de l'architecture ResNet, a été développé pour classer les tissus cancéreux par rapport aux tissus bénins. Pour pallier le manque de données, typique des études à petite cohorte, le modèle a d'abord été préentraîné sur de amples ensembles de données RS rétrospectives (cerveau, sein, prostate), avant d'être adapté au domaine cible par fine-tuning supervisé et adaptation en temps réel non supervisée (TTA). Lorsque les annotations histologiques ne sont pas disponibles, la TTA a permis d'améliorer les performances par rapport au modèle SVM, soulignant sa pertinence dans des contextes cliniques contraints. Une fois les annotations disponibles, le fine-tuning du modèle préentraîné a atteint une précision moyenne de 0,79, surpassant les performances des modèles SVM et TTA. L'intégration du système d'acquisition et du modèle d'apprentissage profond au sein du flux clinique a permis des prédictions en moins de trois minutes après la prise de tissu, sans perturber la procédure thérapeutique, directement dans la salle d'intervention.

La comparaison des deux phases, une fois complétées, a mis en évidence des différences clés ainsi que des bénéfices pratiques. Le passage de la caractérisation in-vivo à l'approche ex-vivo a permis de réduire les perturbations du flux clinique et les délais. Étant donné que les biopsies sont toujours effectuées et que la RS est non destructive, plusieurs mesures peuvent être réalisées sur chaque échantillon, ce qui améliore la capacité à capter l'hétérogénéité tissulaire. En ce qui concerne la classification, bien que les caractéristiques sélectionnées pour les SVMs offre une certaine interprétabilité, les 1D-CNNs représentent une opportunité d'exploiter des ensembles de données externes provenant de domaines connexes.

Ces résultats démontrent la faisabilité et le potentiel clinique de l'association RS-IA pour une confirmation peropératoire du CaP. Les méthodes développées dans cette thèse, tant in-vivo qu'ex-vivo, ouvrent la voie à l'implémentation d'une caractérisation tissulaire optique en temps réel pour améliorer le ciblage tumoral lors des biopsies et traitements par curiethérapie HD. La méthodologie proposée ainsi que les connaissances acquises devraient soutenir les futures études cliniques et les efforts de transfert vers la pratique afin d'améliorer le diagnostic du cancer et les thérapies personnalisées.

ABSTRACT

Accurately targeting malignant tissue during prostate cancer (PCa) interventions remains a fundamental clinical challenge, particularly when tumors are not distinguishable from surrounding normal tissue under standard procedure room conditions. Current prostate intervention workflows, for example, for targeted biopsy sampling or focal treatment, benefit from different navigation systems options. In some cases, these systems integrate transrectal ultrasound (TRUS), image registration, and electromagnetic (EM) tracking to target regions of interest in the gland. Nonetheless, since these methods rely on preoperative imaging for planning and post-intervention histopathology for confirmation, there remains a lack of in-situ, real-time PCa confirmation to target tumors accurately. Such intraoperative insights could guide decision-making, improve biopsy yield, personalize focal therapy, and ultimately enhance patient outcomes.

Raman spectroscopy (RS) is a label-free, non-destructive optical technique with the potential to contribute to solving this challenge. It is based on the inelastic scattering of light, where a small fraction of photons interact with vibrational modes of the molecules in the sample, producing subtle energy shifts known as the Raman effect, and providing real-time molecular signatures. Artificial intelligence (AI)-based classification methods can exploit this molecular information to support tissue characterization and disease detection. When combined with machine learning, RS has shown promising results in identifying several conditions, including Alzheimer's disease, cardiovascular disorders, and various types of cancer.

With over 1.4 million new cases reported annually, PCa is the second most commonly diagnosed cancer among males worldwide, following lung cancer. As with all cancers, early detection and accurate tumor localization remain critical determinants of treatment success and patient outcomes. This thesis investigates the integration of RS with AI-based classification models to enable real-time, localized, in-situ PCa confirmation for supporting tumor-targeted interventions. To this end, the research has been structured around four objectives, associated with four journal articles, across two methodological phases: Phase I, with a focus on in-vivo experiments and machine learning models for PCa detection using multimodal data; Phase II, focused on ex-vivo robot-assisted experiments, deep learning models, and transfer learning strategies for spectral classification.

In the initial part of Phase I, a prototype navigation system integrating image fusion and EM tracking for high-dose-rate (HDR) brachytherapy interventions was evaluated during its early clinical deployment. We assessed its performance across 43 procedures in terms of

workflow efficiency, registration accuracy, and EM-based catheter reconstruction accuracy. This evaluation highlighted the system's seamless integration into existing clinical workflows with minimal learning curve, the improved catheter identification enabled by EM tracking, and the advantages of surface-based deformable registration over rigid alignment. This contribution was essential not only to establish the clinical utility of the prototype system but also to support the subsequent phases of the project.

The next step in this phase focused on a prospective feasibility study, which was conducted to assess the diagnostic value of in-vivo RS combined with image-based radiomics features from multiparametric magnetic resonance imaging (mpMRI). We used an in-house developed navigation system, integrating EM tracking and TRUS-MRI fusion, to guide an RS probe to specific regions of interest in the prostate. Spectra were acquired from 18 patients undergoing HDR brachytherapy. Co-located RS, mpMRI-derived radiomics features, and confirmation biopsies were used to train support vector machine (SVM) classifiers for PCa detection. These in-vivo measurements allowed to identify the benefit of complementary mesoscopic/macroscopic information of the multimodal characterization approach. Integrating RS with mpMRI radiomics achieved an accuracy of up to 0.83 for detecting clinically significant PCa, significantly outperforming models based on either modality alone. Also, the implemented feature selection strategy contributed to the good performance while limiting the model's complexity.

In Phase II, conducted in an ex-vivo context, we first implemented a robot-assisted RS system for the automatic characterization of fresh biopsy tissue samples. A 6-degree-of-freedom robotic arm and a static optical probe enabled precise and efficient spectral acquisition. The sub-millimeter precision achieved through the evaluated calibration methods (3D error of 0.20 mm) is particularly important given the small diameter of the biopsy samples ($\sim$0.8 mm). This level of precision not only facilitates delicate tissue handling but also supports the application of site-specific labels, an essential feature for accurate training of classification models. Additionally, at this point, we conducted a preliminary investigation of one-dimensional convolutional neural network (1D-CNN) models for cancer detection using retrospective brain, breast, and prostate RS datasets. This exploration showed better performance for brain and breast compared to traditional SVM and underscored the 1D-CNN's ability to learn spectral features automatically without hand-crafted input.

To complete this phase, the robot-assisted RS system was deployed in the procedure room for a prospective ex-vivo feasibility study involving biopsy samples from 10 patients undergoing HDR brachytherapy. The collected Raman spectra and site-specific histopathological labels formed a prospective PCa dataset serving as the target domain. A 1D-CNN model, adapted

from the ResNet architecture, was developed to classify cancerous versus benign tissue. To mitigate data scarcity typical in early-stage studies, the model was first pre-trained on large retrospective RS datasets from brain, breast, and prostate tissues, and subsequently adapted to the target domain using supervised fine-tuning and unsupervised test-time adaptation (TTA). In scenarios where labels are unavailable (unsupervised), TTA improved predictions over the SVM baseline, demonstrating its relevance for small-cohort studies and real-world constraints. Once histopathological ground truth labels were available, adapting the pre-trained 1D-CNN model to prospective RS data via fine-tuning reached an average accuracy of 0.79, outperforming the SVM baseline and TTA models. The integration of the acquisition system and deep learning models into the clinical workflow enabled PCa predictions directly in the procedure room, in under 3 minutes following tissue sampling, with minimal disruption of the therapy procedure.

Comparing the two phases upon completion revealed key differences and practical benefits. Transitioning from in-vivo to ex-vivo characterization minimized clinical workflow disruption and additional delays. Since biopsies are still collected and RS is non-destructive, multiple measurements can be performed on each core, improving the ability to capture tissue heterogeneity. Regarding classification, while the use of selected features in SVMs offers some interpretability, 1D-CNNs present a valuable opportunity to leverage external datasets from related domains.

These findings validate the feasibility and clinical promise of combining RS with AI-driven models for intraoperative PCa confirmation. The methods developed and evaluated throughout this thesis, both in-vivo and ex-vivo, represent a step toward implementing real-time optical tissue characterization to enhance tumor targeting during biopsy and brachytherapy procedures. The methodology and insights gained are expected to support future clinical studies and eventual translational efforts to improve cancer diagnostics and personalized therapy.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ACRONYMS

1D	One-dimensional
AP	Additional processes
ADC	Apparent diffusion coefficient
AI	Artificial intelligence
AUC	Area under the curve
BCS	Breast conserving surgery
CHUM	Centre Hospitalier de l'Université de Montréal
CI	Confidence interval
COM	Center of mass
CNN	Convolutional neural network
CRPC	Castration-resistant prostate cancer
DCE	Dynamic contrast-enhanced
DL	Deep learning
DOF	Degrees-of-freedom
DRS	Deep residual shrinkage
DSC	Dice similarity coefficient
DWI	Diffusion-weighted imaging
EBRT	External beam radiotherapy
EEG	Electroencephalograms
EM	Electromagnetic
FDA	Fisher Discriminant Analysis
FP	Fingerprint
GAN	Generative adversarial network
GBM	Glioblastoma multiforme
GLCM	Gray-level co-occurrence matrix
GS	Gleason score
GT	Ground truth
GTV	Gross tumor volume
HDR	High-dose-rate
H&E	Hematoxylin and eosin
HG	High-grade tumor percentage
HW	High wavenumber
IB	Integrated boost

IBSI	Image Biomarker Standardization Initiative
IDC-P	Intraductal carcinoma of the prostate
IJCARS	International Journal of Computer Assisted Radiology and Surgery
IPCAI	International Conference on Information Processing in Computer-Assisted Interventions
IR	Infrared
ISUP GG	Grade group according to the International Society of Urological Pathology
JBO	Journal of Biomedical Optics
KNN	k-Nearest neighbors
LOO	Leave-one-out
LOPO	Leave-one-patient-out
ML	Machine learning
MRI	Magnetic resonance imaging
mpMRI	Multiparametric magnetic resonance imaging
mTRE	Mean target registration error
NIR	Near-infrared
OR	Operating room
PCA	Principal component analysis
PCa	Prostate cancer
PET	Positron emission tomography
PI-RADS	Prostate Imaging Reporting and Data System
PR	Procedure room
PSMA	Prostate-specific membrane antigen
RBF	Radial basis function
RF	Random forest
RLM	Run-length matrix
RNN	Recurrent neural network
RO	Radiation oncologist
ROC	Receiver operating characteristic
ROI	Region of interest
RS	Raman spectroscopy
SVM	Support vector machine
T2	T2-weighted
TL	Transfer learning
TRE	Target registration error

TRF	Tool reference frame
TRUS	Transrectal ultrasound
TTA	Test-time adaptation
TZ	Transitional zone

LIST OF APPENDICES

CHAPTER 1 INTRODUCTION

Accurately targeting an area that is not visually identifiable is inherently challenging. This is particularly true for targeted prostate cancer (PCa) procedures —such as transperineal / transrectal biopsy, focal radiation, or tumor resection— where tumor and normal tissue are indistinguishable under standard procedure room (PR) conditions. The PCa diagnosis and focal treatments are limited by a lack of intraoperative information to accurately target tumors, for instance, with needles for biopsy or catheters for brachytherapy. Valuable techniques exist for navigation, tissue characterization, and classification —such as image registration or electromagnetic (EM) tracking for navigation; multiparametric magnetic resonance imaging (mpMRI) and Raman spectroscopy (RS) for tissue characterization; support vector machines (SVMs) or deep learning (DL) models for classification. When combined, these methods have great potential as an integrated tool to detect PCa directly in the PR. This intraoperative confirmation could support decision-making, treatment personalization, and tumor targeting, improving the diagnostic yield of biopsy and efficacy of focal radiotherapy.

1.1 Prostate cancer: context and challenges

It is estimated that there were more than 19.9 million new cancer cases worldwide in 2022, with lung, breast, colorectum, and prostate being the sites with the highest incidence. Among males around the world, PCa was the second most diagnosed cancer, second only to lung cancer, with more than 1.4 million new cases per year [17].

For all types of cancer, including PCa, early detection and accurate tumor localization are crucial for treatment success and patient prognosis. Over the years, different methods have been studied and developed to detect or confirm PCa, some in blood, some using medical images [18, 19]. Nevertheless, the histopathological analysis of biopsy samples remains the standard of care to confirm cancer diagnosis [20, 21]. Acquisition of these samples is usually performed by transrectal ultrasound (TRUS)-guided biopsy, which can yield up to 34% false-negative rates and does not allow in-situ characterization [22–24].

For histopathological analysis, biopsy cores are fixed and processed according to standard histopathologic procedures, then examined by an expert to identify cellular patterns on stained sample slides [25]. Based on the predominant and secondary patterns, a pathology report is generated, including the Gleason score (GS), the grade group according to

the International Society of Urological Pathology (ISUP GG), and the high-grade tumor percentage (HG) for each biopsy core [3, 26].

Currently, several effective treatment options exist for PCa, either alone or in combination, including radiotherapy, chemotherapy, immunotherapy, and surgical tumor excision. Their success largely depends on early and accurate diagnosis [3, 27, 28].

In the pursuit of improving the effectiveness and accuracy of targeted treatments and biopsies, medical imaging plays a fundamental role. Nuclear imaging modalities, such as positron emission tomography (PET), offer high specificity but come with the disadvantage of exposure to ionizing radiation [29]. Magnetic resonance imaging (MRI), particularly some mpMRI sequences, provides high sensitivity and is used by physicians to report on tumors [18, 30]; however, it requires long acquisition times and MRI-compatible instrumentation [31, 32]. Ultrasound, in contrast, is an affordable, fast, and widely available modality, but not all tumors are visible with this technique [33, 34]. To leverage the strengths of each imaging modality, preoperative and intraoperative image fusion strategies are employed to enhance tumor targeting [35, 36].

Despite the use of imaging for intervention guidance and histopathologic analysis for post-intervention confirmation, a critical gap remains: the absence of real-time intraoperative information to confirm the precise location of cancerous tissue. To address this limitation, a rapid, localized, and non-destructive characterization technique such as RS, when combined with advanced classification methods, can provide complementary information to enhance intraoperative decision-making.

1.2 Contributions

This thesis aims to advance real-time intraoperative PCa detection by combining RS with artificial intelligence (AI)-based classification methods. By leveraging knowledge from multiple disciplines, this approach facilitates minimally invasive, multimodal analysis. The main contributions are summarized as follows:

- Design and evaluation of a TRUS-MRI-EM 3D navigation system for guiding an optical probe to target regions within the prostate. This includes the clinical assessment of a prototype interventional platform (UroNav) used in brachytherapy treatments and the deployment of an in-house RS navigation system to enhance spatial localization of target sites.
- Development of a multimodal classification model that integrates in-vivo RS and mpMRI-

based radiomics features for PCa detection. To our knowledge, this study presents the first in-vivo RS measurements specifically aimed at distinguishing cancerous from non-cancerous prostate tissue. A feasibility study involving 18 patients was conducted to acquire both data types and train SVM-based models.

- Implementation of a robot-assisted RS system for real-time spectral acquisition from fresh biopsy cores in the PR. The system demonstrated high positioning accuracy, repeatability, and reduced acquisition time, facilitating more precise histopathological label assignment (site-specific labels vs. core-level diagnosis) and addressing a common limitation in RS-based tissue classification.

- Leveraging multi-organ retrospective RS datasets through transfer learning to enable DL-based real-time confirmation of PCa. Spectral data were acquired from biopsy samples of 10 patients using the robotic system, and pre-trained CNN models were adapted to this limited prostate dataset. The proposed supervised and unsupervised adaptation strategies were evaluated, with supervised methods achieving performance comparable to previously reported multimodal approaches. Tissue classification was achieved directly in the PR within 3 minutes of sample collection.

1.3 Thesis structure

This thesis is organized into ten chapters and follows an article-based format. Chapter 1 introduces the clinical and scientific motivations for this work and outlines its main contributions. Chapter 2 provides background information, covering prostate anatomy and the current standard of care for PCa diagnosis, relevant medical imaging modalities, and the physical principles of RS, to establish the foundation for the subsequent chapters. Chapter 3 offers a literature review aligned with the different phases of the project, including imaging and navigation systems for prostate interventions, biomedical applications of RS —particularly for cancer detection— and the use of AI for spectral and radiomics analysis. To complete this contextual framework, Chapter 4 defines the research question, hypotheses, objectives, and the general methodology that guided the development of this work.

Chapters 5 to 8 present the core research contributions of this thesis, each corresponding to a peer-reviewed journal publication. These contributions are structured around two main phases of a feasibility study for RS-based PCa detection: an initial in-vivo phase (Phase I: Chapters 5 and 6), followed by an ex-vivo phase (Phase II: Chapters 7 and 8). Chapter 5 presents the first article, published in *Radiotherapy and Oncology*. It reports on a prospective evaluation of a prototype navigation system (UroNav) that integrates EM tracking and MRI-

TRUS image registration. This study supports the subsequent clinical testing of an in-house navigation system. Chapter 6 presents the second article, published in the *Journal of Biomedical Optics (JBO)*. It describes a prospective in-vivo pilot study evaluating a novel approach for PCa detection by combining in-situ RS with radiomics features extracted from mpMRI. Moving to the ex-vivo phase, Chapter 7 presents the third article, published in the *International Journal of Computer Assisted Radiology and Surgery (IJCARS)*. It introduces a robotic system for automated RS characterization of fresh biopsy tissue and evaluates the performance of a DL model for RS-based cancer detection in datasets from multiple organs. Finally, Chapter 8 presents the fourth and last article submitted to the *Biophotonics Discovery (BIOS)*. It investigates strategies to enable RS-based PCa detection in small-cohort studies by leveraging retrospective multi-organ datasets through model pre-training and supervised/unsupervised adaptation. The proposed DL model, trained on multi-organ RS data, was evaluated in a real-time robotic workflow for prostate biopsy analysis.

The final part of this thesis synthesizes the scientific insights gained across the four studies. Chapter 9 presents a general discussion of the main contributions, limitations, and clinical implications. Finally, Chapter 10 summarizes the key findings and proposes future directions for advancing RS-based PCa detection and real-time clinical decision support.

CHAPTER 2 BACKGROUND

2.1 Prostate cancer: diagnosis and treatment

In North America, PCa showed the highest incidence (excluding non-melanoma skin cancer) and mortality among males, with more than 255,000 new cases and 78,000 deaths in 2022 [17]. The prostate is a small gland, approximately 15–20 cm^3 in volume, located in the pelvic cavity anterior to the rectum and inferior to the bladder. It surrounds the prostatic urethra and is closely associated with the seminal vesicles (Fig.2.1) [17].

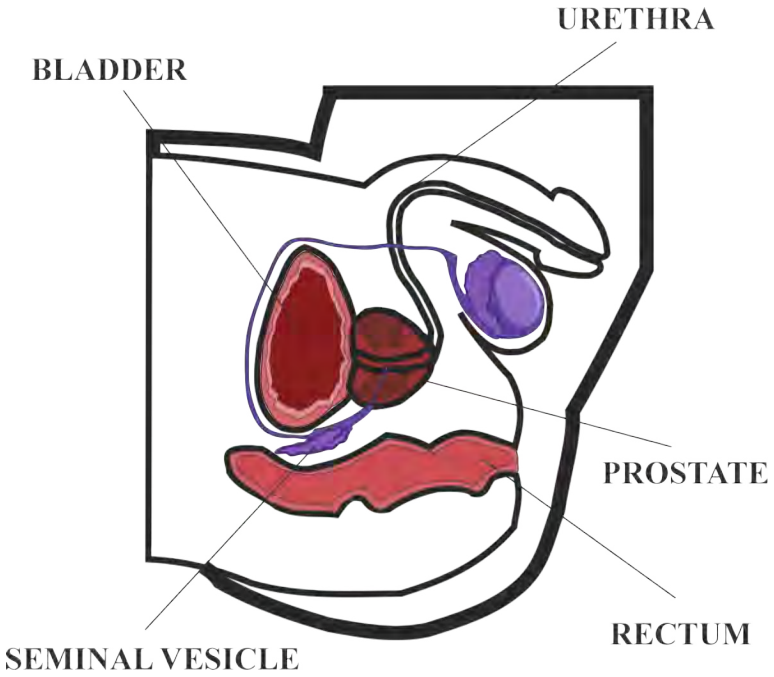


Figure 2.1 Anatomy of the pelvic region in lithotomy position.

Uncontrolled growth in the cells of the gland gives rise to cancer, and the resulting tumors are located mostly in the peripheral zone of the prostate, which makes up 70% of the gland's tissue. While various histological subtypes exist —including sarcoma, small cell carcinoma, carcinoid tumours, and squamous cell carcinoma— adenocarcinoma, often originating at multiple sites throughout the prostate gland, accounts for the vast majority of cases (95%) [3].

Accurate and timely diagnosis is critical for effective treatment and improved patient outcomes. In addition to digital rectal examination, blood-based tests such as prostate-specific

antigen (PSA) have long been used for early screening. However, PSA testing has been criticized for its limited specificity, as elevated PSA levels may occur in benign conditions. Advanced imaging modalities —such as mpMRI and PET— have enhanced the detection and localization of PCa. However, they are limited by factors such as acquisition time, cost, and, in the case of PET, exposure to ionizing radiation (see Section 2.2). Despite these advancements, histopathological analysis of biopsy specimens remains the gold standard for definitive diagnosis (Fig.2.2).

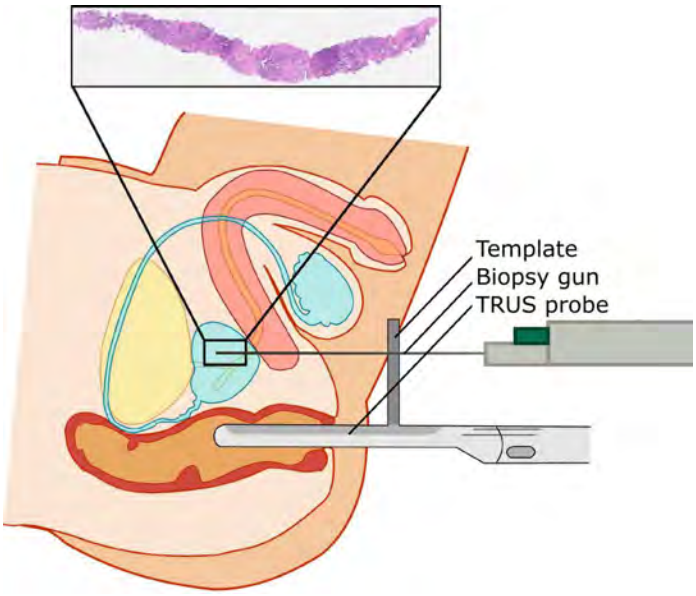


Figure 2.2 Transrectal ultrasound (TRUS)-guided transperineal biopsy.

Given the anatomical location of the prostate, tissue sampling can be performed through minimally invasive procedures using either a transrectal or transperineal approach. In clinical practice, TRUS-guided biopsy is the most common method (Fig.2.2). This technique involves the systematic extraction of 6 to 16 cores (up to 24 cores in saturation biopsy), distributed across different regions of the prostate to increase the likelihood of capturing cancerous tissue (Fig.2.3). However, due to the random nature of this sampling strategy, and since not all lesions are hypoechoic, some lesions may be missed, leading to false-negative rates as high as 34% [3, 18, 22, 37]. Once the biopsy cores are acquired, they are fixed, sectioned, and stained for histopathological analysis. Each slide is examined by an expert based on glandular architecture and assigns a GS, determined by the most prevalent and second most prevalent patterns. This score is then translated into an ISUP GG ranging from 1 (least aggressive) to 5 (most aggressive), along with an estimation of HG. For instance, a slide composed of 80% pattern 4 and 20% pattern 3 would be reported as GS: 4+3=7, ISUP GG: 3/5, and HG: 80% [3, 25, 26].

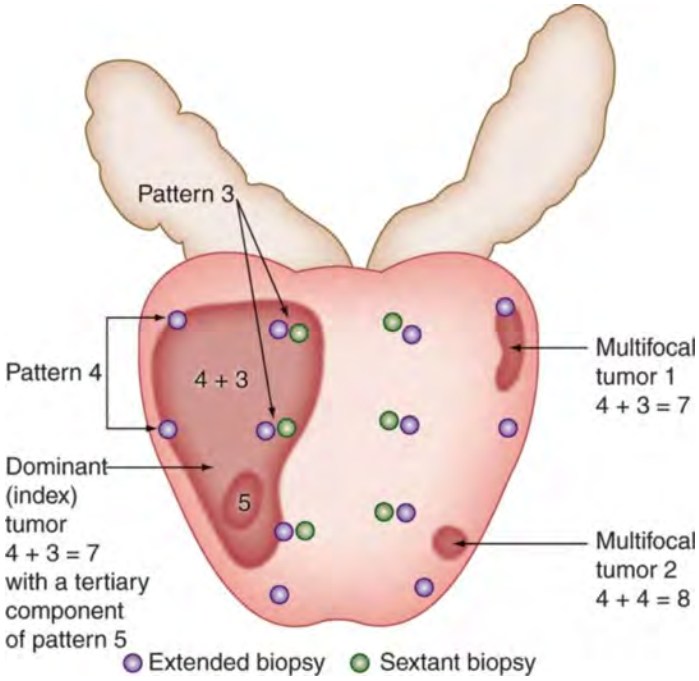


Figure 2.3 Effect of tumor multifocality and biopsy sampling techniques on biopsy/radical prostatectomy Gleason score (GS) correlation. In this example, areas of Gleason pattern 4 tumor are primarily identified by laterally directed extended biopsy cores, whereas sextant biopsy cores under sample these areas. Areas of tertiary pattern 5 within the index tumor and nondominant multifocal tumor 2 with Gleason score 4 + 4 remain unsampled by both sextant and extended biopsy cores. (CC BY) [3].

In addition to accurate diagnosis, appropriate therapeutic management is critical to improving outcomes in patients with PCa. Treatment options for localized PCa include radical prostatectomy, radiotherapy (with or without androgen deprivation therapy), thermal or cryoablation, and, in selected low-risk cases, active surveillance [3, 38]. Among non-surgical approaches, radiotherapy is one of the most commonly used and effective treatments for localized PCa. Techniques such as external beam radiotherapy (EBRT) or brachytherapy allow for high-dose delivery to the prostate while minimizing exposure to surrounding organs at risk. Recent advances in imaging, treatment planning, and radiation delivery have improved targeting accuracy and therapeutic outcomes [38, 39].

In standard EBRT, the entire prostate gland is treated with a uniform radiation dose. However, local recurrences after radiotherapy often originate from the primary tumor site, motivating the development of focal dose escalation strategies [39, 40]. This approach delivers a higher radiation dose to the macroscopic tumor visible on mpMRI, while maintaining strict dose constraints for surrounding organs at risk. The FLAME trial, a phase III multicenter randomized study, showed that adding such a focal boost significantly improved 5-year bio-

chemical disease-free survival without increasing toxicity or compromising quality of life [40]. This method, facilitated by advances in imaging and guidance, enhances tumor control while preserving whole-gland treatment coverage. A similar type of focal boost can also be achieved using brachytherapy [38, 41].

High-dose-rate (HDR) brachytherapy is a minimally invasive treatment for localized PCa, either used alone or combined with EBRT. This method is based on the temporary implementation of small radioactive sources, using transperineal delivery catheters and an afterloading system in a brachytherapy suite. This allows delivery of higher doses of radiation to the tumor, reducing tissue involvement. Due to its underlying physical principles, the safety and effectiveness of this modality depend on precise targeting [6, 36, 41].

Despite advances in both diagnosis and treatment, the success of needle-based interventions in PCa remains strongly tied to the ability to localize malignant tissue accurately. Therefore, there has been increasing interest in the integration of advanced imaging modalities and image-guided navigation systems to support such prostate interventions.

2.2 Multimodal imaging and guided interventions

Several medical imaging modalities are used in the management of PCa, each offering advantages and limitations. During interventions, TRUS remains the most widely used imaging technique due to its real-time capability, low cost, and accessibility. It is commonly employed to guide prostate biopsies and treatments such as brachytherapy. However, TRUS has limited soft tissue contrast and often fails to visualize tumors, especially those that are isoechoic or located in challenging anatomical zones [33–35]. In contrast, mpMRI provides superior soft tissue characterization and is widely used for PCa localization, staging, and treatment planning. The prostate imaging reporting and data system (PI-RADS) has become a standard framework to interpret mpMRI findings [18,30]; further details on the mpMRI sequences can be found in Section 3.1.1. Despite its important sensitivity, mpMRI requires long acquisition times, is not easily integrated into the operative workflow, and necessitates MRI-compatible instruments [31, 32]. Alternatively, PET images, particularly when using prostate-specific membrane antigen (PSMA)-targeted tracers, offer excellent specificity for detecting PCa metastases and recurrent disease [1,19,29]. However, PET has limited spatial resolution and involves exposure to ionizing radiation [29]. Given that no imaging modality offers optimal real-time anatomical detail and tumor visibility, image fusion and registration strategies are increasingly adopted to combine the complementary strengths of TRUS, mpMRI, and PET for improved guidance in prostate interventions.

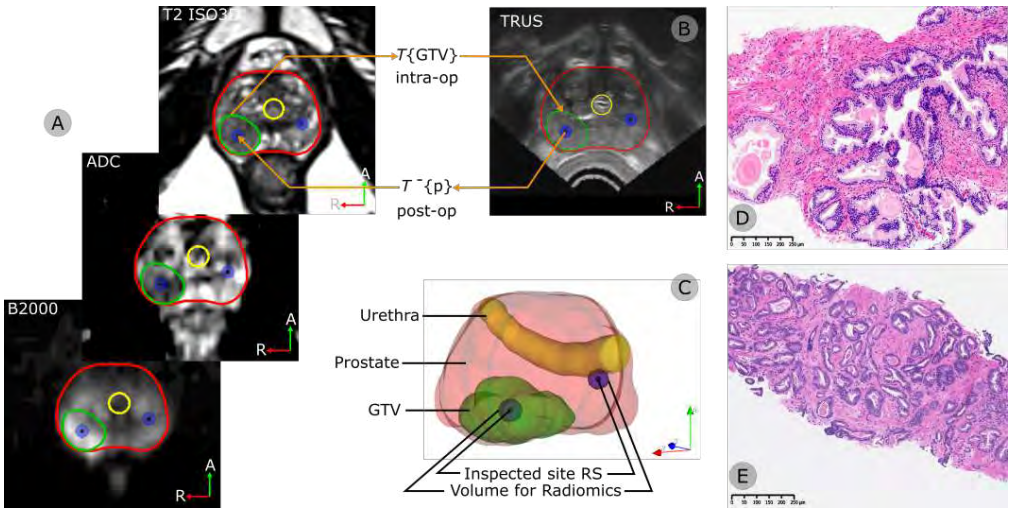


Figure 2.4 Sample image registration and co-location results. (A) Gross tumor volume (GTV) originally segmented on preoperative multiparametric magnetic resonance imaging (mpMRI) is (B) projected over the real-time transrectal ultrasound (TRUS) using a surface-based elastic registration algorithm; the inverse process allows identifying the MRI coordinates corresponding to the inspected sites. (C) A 3D model of the plan can be rendered. (D) Pathological image samples of benign prostatic parenchyma and (E) acinar adenocarcinoma GS: 3+4=7, ISUP GG=2, and HG: 40-50% [Reproduced from Chapter 6].

Among the most common fusion strategies in PCa procedures is the combination of pre-operative mpMRI with intraoperative TRUS. This TRUS-MRI fusion leverages the already analyzed mpMRI images, used for PI-RADS and including organ and tumor segmentations (Fig.2.4.A), with the real-time imaging of TRUS (Fig.2.4.B), providing enhanced guidance during biopsy or focal therapy. However, aligning these two modalities is not trivial, as they are often acquired under different conditions that result in anatomical deformations —such as changes in bladder filling, rectal probe pressure, or patient positioning. Image registration techniques are therefore employed to align the preoperative and intraoperative datasets. Rigid registration (e.g., center of mass matching or urethral alignment), which assumes fixed geometry, may be sufficient in some scenarios but fails to account for organ deformation. In contrast, elastic (or non-rigid) registration techniques can compensate for the deformation of the prostate by deforming the image volume, typically using surface-based methods, leveraging organ segmentations, or image intensity metrics [35,42,43]. These registration approaches are key to enabling accurate multimodal fusion. Some navigation systems incorporating such technologies are reviewed in the section 3.1.

2.2.1 Radiomics

Medical image fusion and multimodal registration are powerful tools for improving PCa diagnosis and treatment guidance. However, beyond the visual interpretation, these imaging modalities —especially mpMRI— contain valuable quantitative information that can be extracted through radiomics and computationally analyzed [44]. Radiomics involves converting standard-of-care medical images into high-dimensional data, offering image-based biomarkers that reflect tissue characteristics noninvasively. A wide variety of standardized radiomics features are primarily classified as intensity-, texture-, or shape-based. Intensity-based features describe pixel or voxel distributions (e.g., mean, median, standard deviation), while texture-based features assess the spatial arrangement of intensities using models like the gray-level co-occurrence matrix (GLCM), run-length matrix (RLM), or wavelet transforms. Extracted from a region of interest (ROI), shape-based features include descriptors such as volume, surface area, and sphericity, and are especially useful when organs or lesions are segmented. Nowadays, there are tools that allow the extraction of hundreds of features either from ROIs (structures or lesions segmented in the image) or in a voxel-wise manner (in this case, the ROI is every single voxel and its surroundings) [45, 46]. For the latter case, shape-based features are not applicable.

In PCa, most radiomics studies focus on mpMRI sequences —particularly T2-weighted (T2), diffusion-weighted imaging (DWI), apparent diffusion coefficient (ADC) maps, and dynamic contrast-enhanced (DCE)— due to their clinical relevance in PI-RADS-based evaluation [47, 48]. It is important to note that this feature extraction is done on preoperative images, which do not provide real-time information during the interventions. The radiomics pipeline typically includes segmentation of the ROI, feature extraction, normalization, and feature selection, followed by the development of classification or prediction models [8, 46]. While radiomics offers rich quantitative information from mpMRI, it is inherently limited by its reliance on preoperative images, which may not accurately represent the tissue state during the intervention. This underscores the need for real-time, intraoperative tools to characterize PCa tissue.

2.3 Principles of Raman spectroscopy

Light–tissue interactions play an important role in biomedical imaging and spectroscopy techniques. In its interaction with biological tissue, light may be absorbed, elastically scattered (Rayleigh scattering), inelastically scattered (Raman scattering), or re-emitted as fluorescence. These processes depend on the tissue's optical properties and the wavelength of the

incident light [49, 50]. Most optical diagnostic systems, including RS, operate within the near-infrared (NIR) therapeutic window (650—900nm), where absorption is relatively low, reducing the risk of tissue damage from laser exposure. Lasers at 670, 785, 830, and 1064nm are commonly employed for Raman tissue characterization [51].

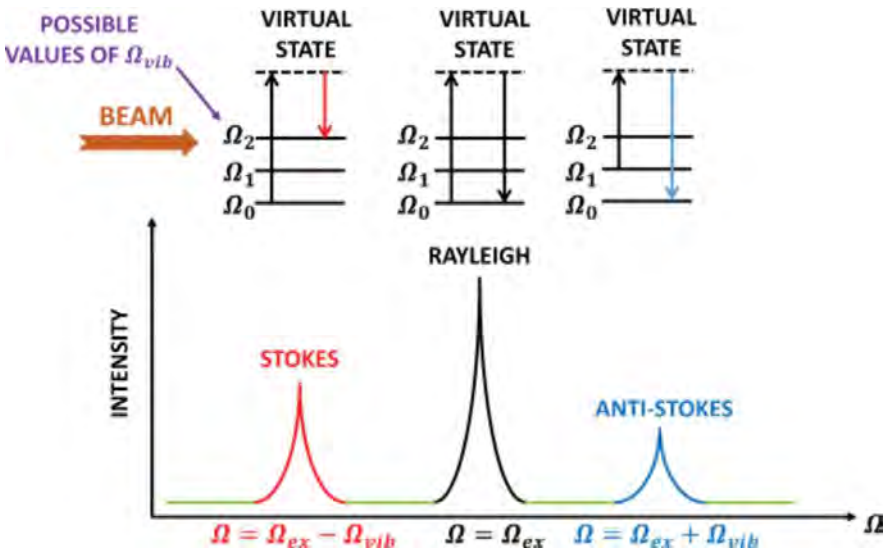


Figure 2.5 Schematic diagram illustrating the electronic levels involving Stokes, Rayleigh and Anti-Stokes scattering and corresponding representative Raman spectra; the symbol Ω represents photon frequencies. (CC BY).

As a label-free, non-destructive technique, RS provides molecular-level information based on inelastic scattering and has been applied in several fields, including material science and medicine. In the context of scattering, when monochromatic laser light interacts with a molecule, the majority of photons undergo Rayleigh scattering, retaining their original energy (elastic). However, a small fraction —approximately 1 in every 10^6 to 10^8 photons— undergoes inelastic scattering, exchanging energy with the molecule's vibrational modes (Raman effect). This involves excitation of the molecule to a short-lived virtual state, followed by re-emission of a photon with either lower (Stokes) or higher (anti-Stokes) energy than the incident light. The energy difference between the incident and scattered photons, known as the Raman shift, corresponds to the molecule's vibrational transitions and serves as a molecular signature of the sample. Raman shifts (in cm^{-1}) are calculated as $\nu\ (\mathrm{cm}^{-1}) = \left(\frac{1}{\lambda_0} - \frac{1}{\lambda}\right)$, where λ_0 and λ are the excitation and scattered wavelengths (in cm), respectively, which provides a wavelength-independent spectral axis for analysis. This is illustrated in the Jablonski diagram (Fig. 2.5), which shows the electronic transitions associated with Rayleigh, Stokes, and anti-Stokes scattering, alongside representative Raman spectra. Because most molecules are in their ground vibrational state at physiological temperatures, Stokes scattering is more

probable than anti-Stokes [3]. Although Raman scattering occurs rarely and yields inherently weak signals, it offers highly specific information about molecular composition —unlike elastic scattering— making it particularly valuable for tissue characterization and disease detection in complex biological environments.

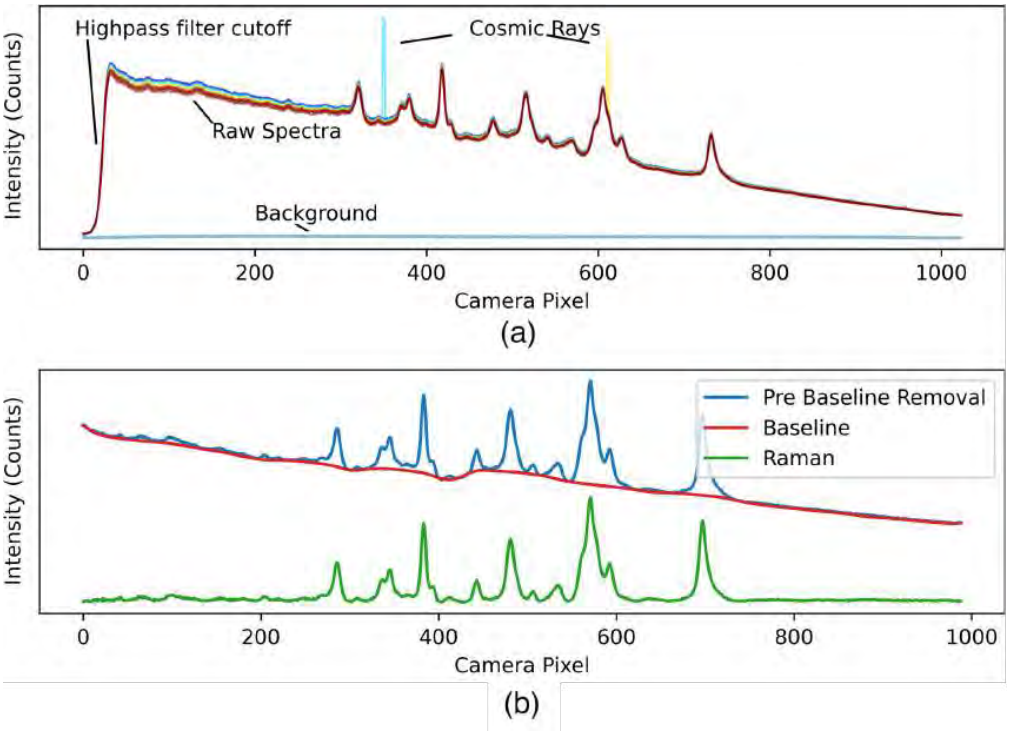


Figure 2.6 Overview of processing steps demonstrated on a signal measured from nylon with a point-probe system. (a) The raw accumulations and background; (b) after baseline removal.[Adapted from [4]] (CC BY).

Robust processing strategies are essential for obtaining clean and meaningful spectral information (spectra with defined peaks) from the raw measurements. Common preprocessing steps include averaging, background and cosmic ray subtraction, system response correction (e.g., using the NIST standard SRM2214), autofluorescence baseline removal (e.g., using BubbleFill [4]), spectral normalization, and x-axis calibration to express spectra in Raman shift units (cm^{-1}). Figure 2.6 presents an example of the raw measurements on nylons and the final result after processing [4].

These corrections enable more reliable downstream analyses, such as peak identification or multivariate classification. In biological tissues, Raman spectra reveal distinct peaks corresponding to molecular vibrations in proteins (e.g., amide at $\sim$ 1659cm^{-1}, or tryptophan 1338-1340cm^{-1}), lipids (e.g., 1299-1301 cm^{-1} CH_2 deformation, or 1441-1453 cm^{-1} C-H bending), nucleic acids (e.g., $\sim$ 785 cm^{-1}). At 1004 cm^{-1}, a peak associated with phenylalanine

is even used as a qualitative sign of quality in biological samples [3, 10, 52]. These molecular markers provide critical insight into tissue composition and pathological state, supporting tissue discrimination in downstream classification tasks.

CHAPTER 3 LITERATURE REVIEW

3.1 Navigation systems for prostate interventions

Intraoperative imaging is fundamental as a guidance technique for needle-based PCa procedures, such as TRUS-guided biopsy sampling. In recent years, the development of multimodal navigation systems has significantly improved spatial accuracy during tumor-targeted interventions. These systems integrate information from sensitive preoperative imaging with intraoperative imaging through image registration. Commercial platforms such as BiopSee (Pi Medical), PredictiveFusion (MIM Software), and SmartFusion (Canon Medical Systems) rely on rigid image registration to fuse preoperative mpMRI with real-time TRUS, enabling physicians to visualize MRI-detected lesions during TRUS-guided procedures [5, 32, 53, 54]. However, one of the challenges of multimodal image registration for prostate applications is the deformation of structures due to the rotation of the pelvis, the presence of an endorectal probe, the placement of a foley catheter, and the insertion of brachytherapy catheters; hence the need to investigate elastic registration and to evaluate its performance PR conditions [35, 53]. Figure 3.1 shows a comparison of rigid and elastic registration for TRUS and MRI from the prostate [5].

To address these limitations, elastic registration methods have been developed, using anatomical landmarks, organ surface contours, or intensity-based methods to compensate for these intermodal variations —reporting target registration error (TRE) of about 2 mm— [5,24,35, 55,56]. Systems such as Artemis (Eigen) and Urostation (Koelis) integrate elastic registration into clinical workflows to support targeted biopsies with improved spatial accuracy [18,32,57]. Traditional TRUS-guided systematic biopsies exhibit high false-negative rates, reported between 20–47% [22, 37]; leveraging complementary multimodal fusion techniques can significantly reduce targeting errors, bringing false-negative rates closer to 15% in some clinical studies [32, 37]. Image registration strategies are also employed to leverage complementary information from modalities such as MRI and PET to enhance diagnosis and aid surgical planning for prostate intervention, particularly since integrated PET-MRI systems are not widely available. [58].

In the HDR brachytherapy workflow, accurate placement and localization of catheters are as important as the localization of treatment targets. Current TRUS-guided systems allow manual catheter reconstruction, where depth is inferred from measured catheter free-lengths (i.e., the segment of known-length catheters extending from the template), which is a time-consuming process. When multiple catheters are in place, interference can affect localization

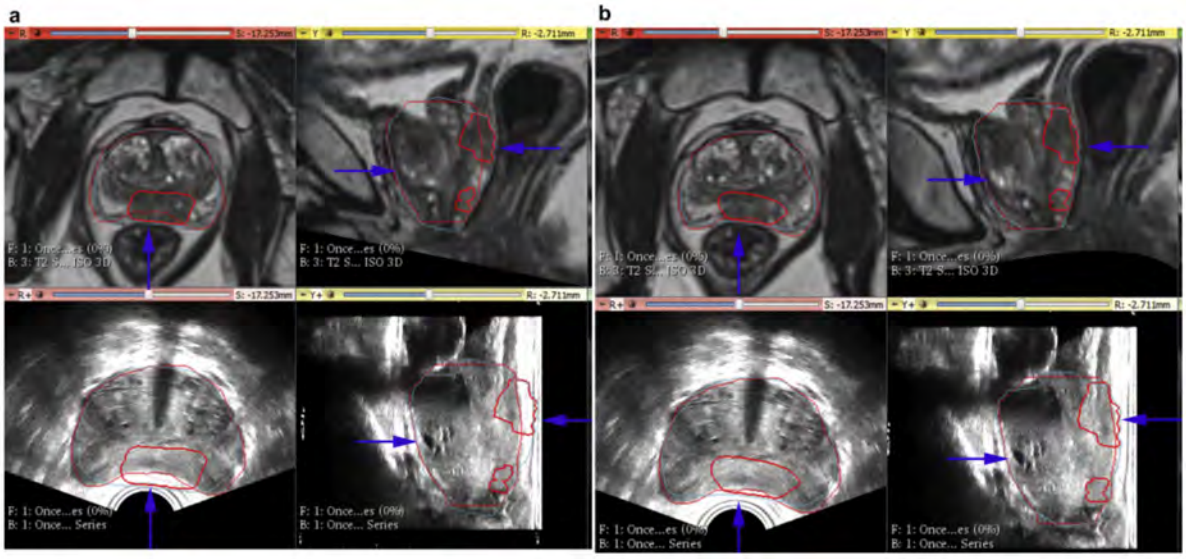


Figure 3.1 Representation of the (a) rigidly transformed and (b) deformed MRI volume on the top row and the TRUS volume on the bottom row. Arrows show regions of interest (a cyst and a rectum), and the reference TRUS prostate volume is outlined in red, whereas the MRI-registered gross tumor volume (GTV) and prostate are in dark red and blue, respectively. (CC BY). [Adapted from [5]].

accuracy, significantly impacting dose delivery and treatment outcomes [6, 59]. The EM tracking technology is used during minimally invasive interventions to localize rigid or flexible instruments inside the body, even when they are outside the direct line of sight (unlike commonly used infrared tracking systems, which require visual contact between the camera and infrared markers). Although EM tracking has a limited working volume, this is not a significant constraint for interventions targeting confined regions such as the pelvis. This approach can be used to replace manual segmentation during catheter reconstruction [59,60]. The EM system consists of a field generator, a set of coils or sensors tracked within the generator's region of influence (the coils are usually located at the tip of a stylet, or attached to the template or TRUS probe), and a control unit that processes sensor signals to determine their location and orientation (Fig.3.2). In this case, catheter reconstruction is performed by introducing an EM-tracked stylet into each catheter, offering submillimetric precision (<0.7mm) and improved time efficiency [6, 60, 61].

3.1.1 Image-based prostate cancer detection

Non-invasive medical imaging modalities, such as mpMRI, play a critical role in the detection, characterization, and management of PCa, beyond navigation and guidance. A

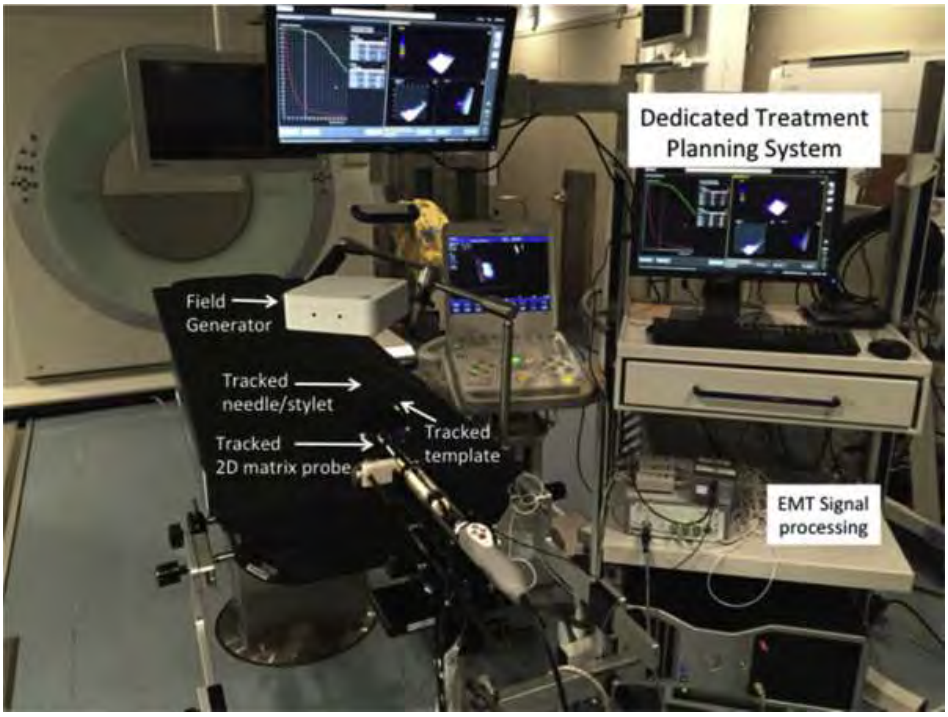


Figure 3.2 Components of the proposed new system deployed in a shielded brachytherapy-dedicated operating room. Key components such as the field generator, EM-tracked stylet, EMT signal processing unit, and treatment planning systems are shown. EMT = electromagnetic tracking. (CC BY) [6].

typical mpMRI analysis combines T2 imaging, DWI, ADC maps, and DCE imaging to provide complementary anatomical and functional information about prostate tissue [32]. Table 3.1 summarizes key features of MRI sequences in PCa assessment. High-resolution T2-weighted imaging offers excellent anatomical detail, facilitating the assessment of the anatomical zones, lesion localization, and detection of extraprostatic extension. DWI evaluates tissue microstructure and cellularity that may indicate the presence of a malignancy, with ADC maps providing a quantitative measure of diffusion properties of the tissue; when different diffusion weightings (b-values) are evaluated, high b-values images (e.g., 1000–1500 s/mm^2) emphasize restricted diffusion, which is often associated with cellular structures. Although benign lesions can also appear hypointense (which can be addressed using complementary DWI images), ADC values have shown an inverse correlation with tumor aggressiveness. The DCE images assess tumor vascularity and perfusion by tracking gadolinium-based contrast kinetics, adding valuable information for lesion detection, staging, and treatment monitoring; given the time constraint on blood tracking, DCE typically uses rapid acquisition techniques. Each imaging modality presents specific limitations: T2 imaging has limited sensitivity for small or low-grade lesions; DWI faces technical challenges such as b-value selection, signal-to-

noise ratio issues, and artifacts; and DCE is constrained by variability and artifacts, limiting its diagnostic value. Nevertheless, the integration of complementary data enhances the diagnostic utility of mpMRI [7, 32, 48]. Figure 3.3 illustrates representative prostate images from each mpMRI modality, where lesions exhibited hypointense signals in T2 and ADC map, while hyperintense signals DWI b-values (b = 50) [7].

Table 3.1 Summary of key multiparametric MRI (mpMRI) modalities, the biological or structural information they provide, and their respective clinical roles in prostate cancer (PCa) assessment.

Imaging Modality	**Biological/Structural Information**	**Clinical Role in PCa Assessment**
T2-weighted (T2)	Anatomy and zonal anatomy; tissue morphology	Localizes lesions, especially in the peripheral zone; baseline for mpMRI assessment.
Diffusion-weighted (DWI/ADC)	Tissue microstructure and cellularity via water diffusion	Detects restricted diffusion in malignancies; ADC inversely correlates with Gleason score.
Dynamic contrast-enhanced (DCE)	Tumor perfusion, vascularity, and permeability dynamics	Highlights early enhancement in aggressive tumors; used when T2w/DWI are inconclusive.

As previously mentioned, these sequences are valuable for expert visual assessment (e.g., following PI-RADS guidelines), but their discriminative potential can also be leveraged by DL models. Some DL architectures based on convolutional neural networks (CNNs), such as ResNet, U-Net, or YOLO, have been widely applied for 2D and 3D mpMRI analysis. These models have demonstrated competitive performance, reducing interobserver variability and assisting radiologists in tasks such as prostate segmentation, malignancy detection (with area under the receiver operating characteristic curve values reaching up to 0.90) [7, 62], and staging (accuracy up to 0.85) [62, 63]. Beyond mpMRI applications, DL strategies can also be used for image processing, for example, in attenuation correction in full-body PSMA-PET, using generative adversarial network (GAN) models [64]. Although full 3D DL-based computer-aided diagnosis strategies continue to be developed, challenges such as the lack of imaging standardization, limited annotated data, and high computational costs remain [48]. Applications that use whole images (e.g., MRI 3D volume) as input to the classification models are beyond the scope of the present project; instead, radiomics-based strategies are explored.

Several platforms now enable standardized radiomics feature extraction from medical images. For instance, PyRadiomics offers more than 120 features across multiple categories, with

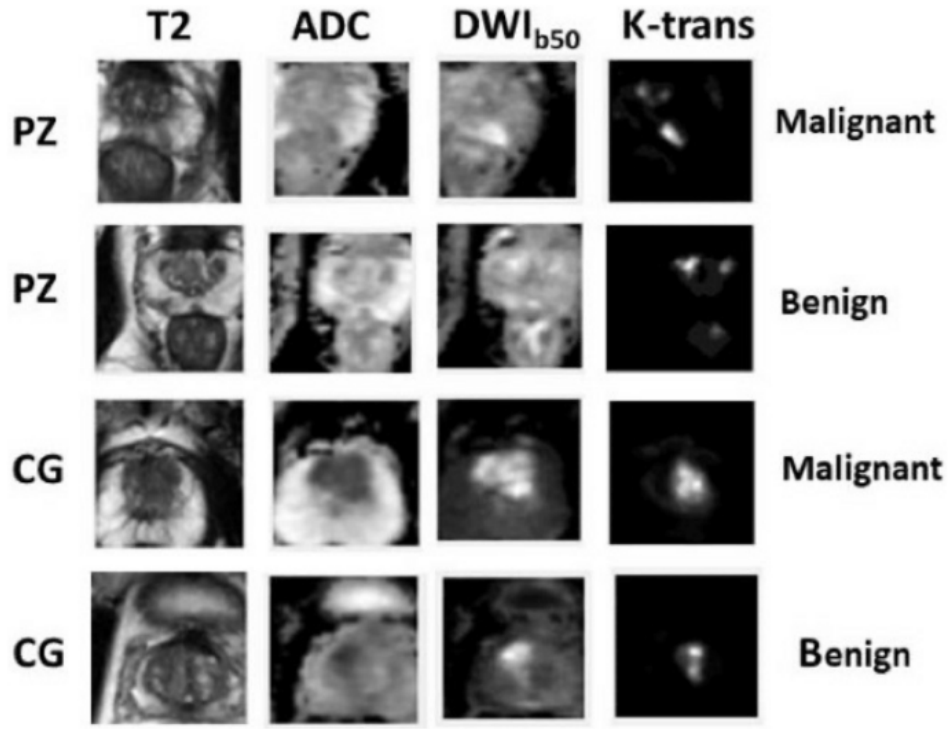


Figure 3.3 Sample images of the cropped 64 × 64 pixel rectangle from all four modalities: T2, ADC, DWI, and K-trans after resampling and registration. The lesions are malignant in peripheral zone (PZ), benign in PZ, malignant in central gland (CG), and benign in CG from row 1 to 4 respectively. (CC BY) [7].

additional features extractable through image preprocessing techniques such as filtering and wavelet transforms [45]. Moreover, the Image Biomarker Standardization Initiative (IBSI) has further established standardized definitions and reference values for 169 radiomics features, ensuring reproducibility and comparability across different software implementations [65]. In the context of PCa detection, first-order and GLCM-based (texture-based) features extracted from mpMRI sequences such as ADC, b1000, b2000, and T2 have been identified as promising image-based biomarkers [44,47,48,66]. As an interpretable and straightforward example, the radiomics feature *Energy* quantifies the magnitude of voxel intensities as the sum of squared values within the segmented region. Consequently, lower image intensities yield lower Energy values. In ADC images, lesions typically appear hypointense, resulting in lower Energy values, whereas normal tissue exhibits higher values for this feature [7, 46, 48]. Classification models based on radiomics have demonstrated strong performance in detecting clinically significant PCa [47,66]. For example, a model trained with nine radiomics features achieved a sensitivity of 0.84 and specificity of 0.73 in detecting clinically significant PCa [66]; in this study, the authors used different filters on the images and shape-based features, which

require the tumor's segmentation. Another study, which enrolled over 300 patients, reported an area under the curve (AUC) with values greater than 0.90 in the receiver operating characteristic (ROC). This was achieved using features derived from T2 and ADC sequences, performing comparably or even outperforming PI-RADS v2 in distinguishing cancerous from non-cancerous tissue [47]. The potential of radiomics as a complement for PI-RADS was also studied, showing that the performance of PI-RADS was significantly improved when combined with radiomics across both transitional and peripheral zones of the prostate (Fig. 3.4) [8].

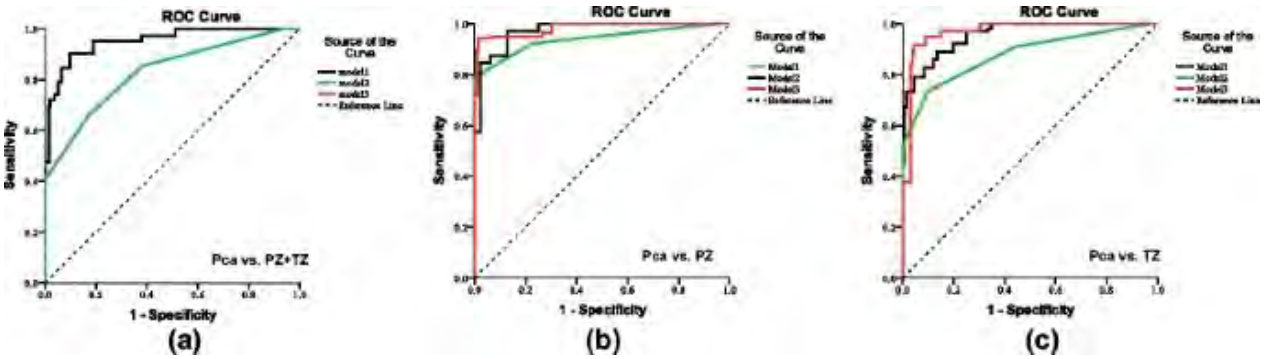


Figure 3.4 Receiver operating characteristic (ROC) plots of the three models for PCa versus normal PZ + TZ (a), PCa versus normal PZ (b), and PCa versus normal TZ tissue (c), respectively. Model 1, trained with radiomics (c) (black); Model 2, trained with PI-RADS (green); Model 3, trained with radiomics and PI-RADS scores. PCa prostate cancer, PZ peripheral zone, TZ transitional zone. (CC BY) [8].

Models trained on quantitative features rather than complete image volumes may offer greater computational efficiency, and radiomics has demonstrated potential to enhance the diagnostic capabilities of standard mpMRI evaluation, reinforcing its growing role as a complementary tool for PCa detection.

3.2 Raman spectroscopy in biomedical applications

As a label-free, non-destructive characterization technique, RS has been widely used in materials science, biology, and biomedical fields. The interaction of light with molecular vibrational states generates a molecular signature of the sample with the potential to detect several diseases, such as Alzheimer's disease, cardiovascular conditions, skin lesions, and cancer [51, 67–70]. Gaba et al. identified several strengths and weaknesses of SR in biomedical applications, which are presented in Table 3.2 [1]. The development of optical fibers and miniaturized probes —instead of Raman microscopes, for instance— has further expanded

the clinical potential of RS, enabling its application not only in ex-vivo tissue analysis but also in in-vivo studies during open surgical settings or minimally invasive procedures, such as endoscopic or needle-based interventions [9, 70]. These probes can be used on accessible surfaces, like the skin, but also inside the body, with their design parameters (e.g., diameter and flexibility) adjusted to meet the requirements of specific clinical applications. Figure 3.5 illustrates examples of fiber-optic Raman probes designed for various clinical applications, highlighting configurations for forward- and side-view detection, as well as options for enhanced focusing [9].

Table 3.2 Advantages and disadvantages of Raman spectroscopy for biomedical diagnostics. (CC BY) [1].

Advantages	**Disadvantages**
Non-destructive, non-invasive High specificity Simultaneous detection of biomolecules Compatible with physiological measurements due to minimal water interaction In vivo applications Suitable for chemical analysis, quantification, classification and imaging of biological materials	Weak Raman signals can lead to long acquisition times Not widely incorporated into current clinical workflows Sophisticated data analysis Autofluorescence can overwhelm the Raman signal (sample dependent)

Raman spectra of biological tissues are commonly divided into the fingerprint (FP) region ($\sim$600–1800cm^{-1}) and the high-wavenumber (HW) region ($\sim$2500–3400cm^{-1}). The FP region contains most of the molecular vibrational signatures —especially from proteins, lipids, and nucleic acids— making it the primary range for biochemical identification (Fig.3.6). The HW region provides complementary information, capturing stretching vibrations (e.g., C–H, N–H, O–H) that are useful for assessing lipid and water content. Recent advances have combined both regions (FP and HW), usually using dual light sources, to enhance the sensitivity and specificity of RS, particularly in detecting cancers of the oral cavity, skin, brain, and prostate [11, 70, 71].

Given the characteristics of RS signals discussed in previous sections, the processing of raw spectra is fundamental for extracting meaningful and reproducible information. Although general steps such as averaging of multiple accumulations, filtering, and baseline removal are implemented in most studies, no standardized protocols or parameters have been universally adopted [1, 72, 73]. Tools such as the Open Raman Processing Library (ORPL) [4], an open-source platform, have been developed to standardize and streamline these corrections, improving reproducibility and reliability across different Raman studies and devices. Figure 3.7 illustrates how this tool can be applied to spectra acquired from organs/samples using

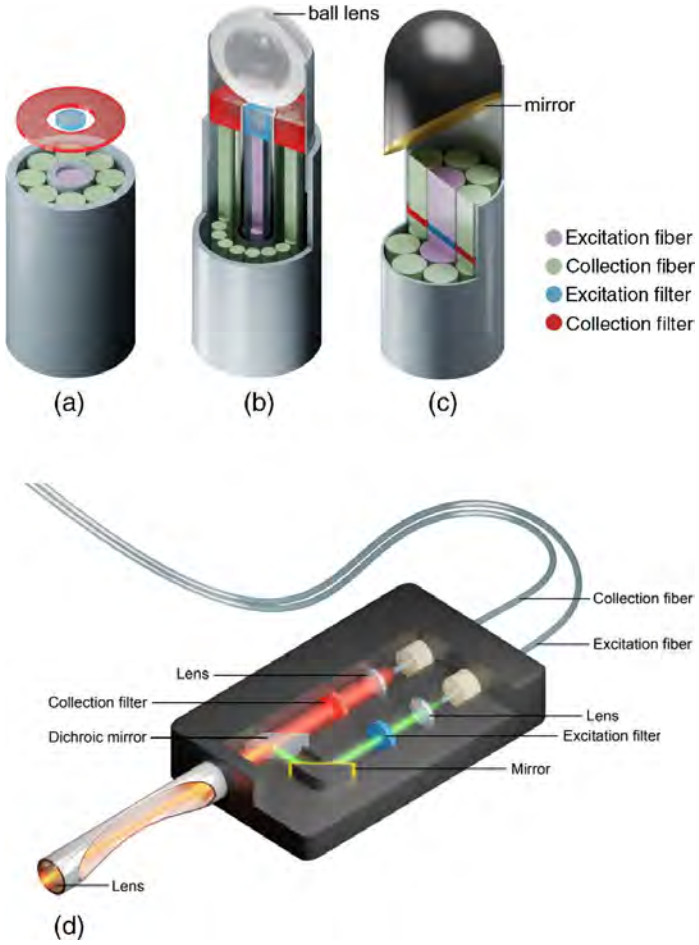


Figure 3.5 Different fiber optical Raman probes. (a) Basic endoscopic probe, (b) with ball lens for focusing, (c) with side-view option, and (d) handheld Raman probe. (CC BY) [9].

different acquisition systems [4].

3.2.1 Diagnosis of cancer and other diseases

Raman spectroscopy has demonstrated considerable potential for diagnosing various diseases beyond oncology. In neurodegenerative disorders, RS has been used to detect molecular alterations associated with Alzheimer's disease, distinguishing affected brain tissue from healthy controls [74, 75]. In cardiovascular applications, RS has been employed to characterize atherosclerotic plaques and assess the biochemical composition of myocardial tissue [76]. Beyond these fields, several studies have evaluated RS for diagnosing skin, colon, lung, breast, and brain cancers [51]. In dermatology, Schleusener et al. studied RS-based differentiation of healthy skin from basal cell carcinoma and melanoma using in-vivo Raman measurements on over 104 patients [70]. Petersen et al. used RS to analyze ex-vivo biopsy samples for col-

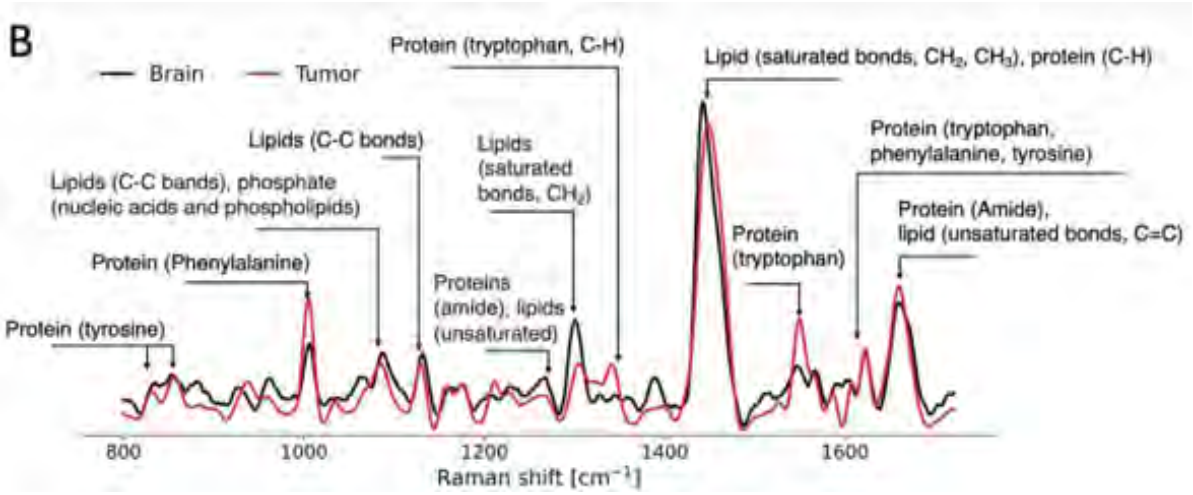


Figure 3.6 Mean spectral fingerprint measurements from 67 patients showing key spectral peaks used for tumor detection. Spectral fingerprints were taken from tumor (red) and non-tumoral brain (black). C-H, carbon-hydrogen single bonds; C=C, carbon–carbon double bonds (unsaturated); C–C, carbon–carbon bonds; CH_2, ethyl group; CH_3, methyl group. (CC BY) [Adapted from [10]].

orectal cancer detection, achieving high sensitivity and specificity in distinguishing normal, pre-malignant, and malignant tissues in a cohort of 151 patients [69]. Other applications include minimally invasive diagnosis of lung cancers [76], bladder cancer detection using superficial and nonsuperficial Raman probes [77], observation of skin changes after breast cancer treatment [78], and others [79–81].

Among the many cancer applications, breast and brain tumors have been particularly well-studied with RS in recent years [13,82–85]. In breast cancer, Haka et al. performed ex-vivo RS measurements on 58 patients, attaining 0.94 sensitivity and 0.96 specificity for distinguishing cancerous tissues from normal and benign tissues [82]; subsequently, a prospective application of their algorithm on an ex-vivo clinical dataset (21 patients) resulted in a sensitivity of 0.83, a specificity of 0.93, highlighting its potential for in-vivo applications and margin detection [68]. Using cultured breast cancer cell lines, Zhang et al. explored the use of RS combined with machine learning for breast cancer subtype classification [86]. Approximately 4,500 Raman spectra for the 6 cell lines were acquired, then classified using principal component analysis (PCA)-discriminant function analysis and PCA-SVM classifiers. Both methods achieved subtype classification accuracy exceeding 0.92. This approach demonstrated a rapid and effective method for breast cancer subtype analysis for cellular-level investigations and recognized the challenges associated with triple-negative breast cancer. David et al. used RS to interrogate breast tissue specimens from 20 patients undergoing surgery, acquiring

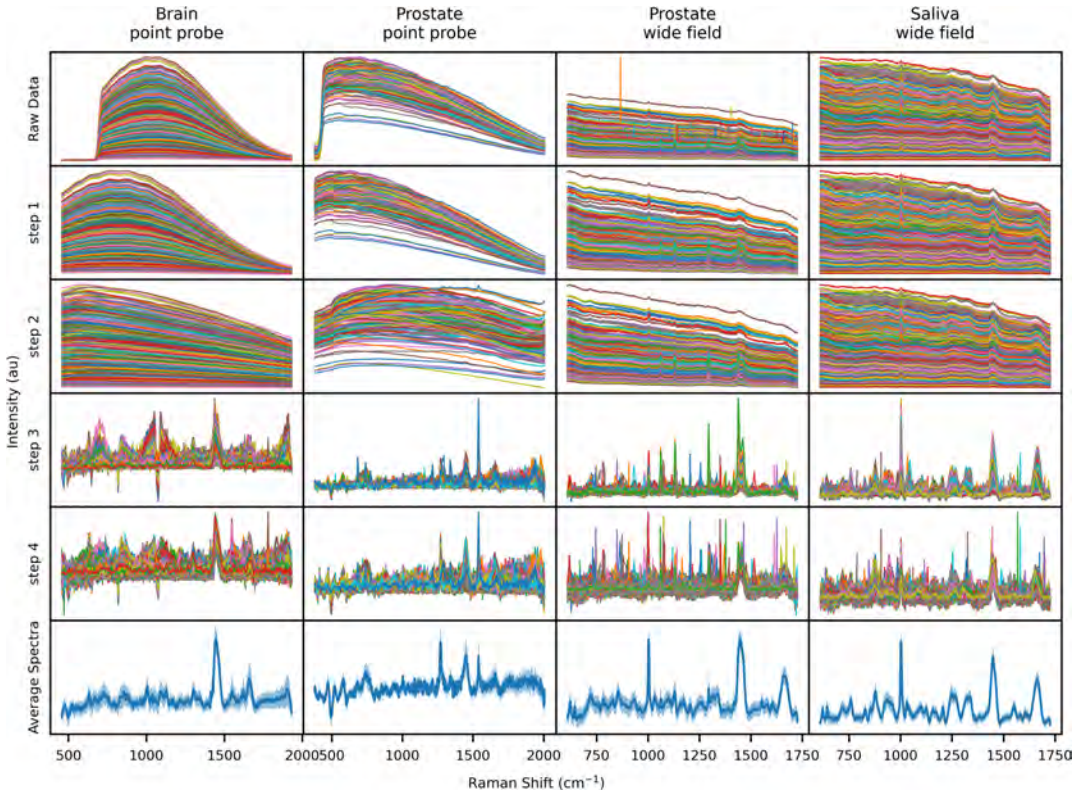


Figure 3.7 Raw data: measured raw spectra from instruments with accumulations combined. Step 1: Truncation (for datasets 1 and 2) and cosmic ray removal. Step 2: Background removal (for datasets 1 and 2) and calibration of x and y axis. Step 3: Baseline removal with BubbleFill. Step 4: SNV normalization. Average Spectra: average spectra computed from step 4 results with standard deviation zone represented as shadow. (CC BY) [4].

238 ex-vivo measurements that were spatially matched to histopathology [87]. The authors used a four-step methodology to assign histopathological labels to the inspected sites; they identified that tissue compression, rotation, and sectioning can alter specimen shape, limiting the reliability of site-specific ground truth (GT). Nevertheless, their method (using an SVM classifier) achieved a sensitivity of 0.93 and a specificity of 0.95 for distinguishing normal tissue from invasive ductal or lobular carcinoma. Notably, these results were obtained using only two spectral bands: the 940 cm^{-1} C-C protein stretch and the 1004 cm^{-1} phenylalanine peak.

Similarly, in brain cancer, Aguiar et al. used RS combined with PCA and Euclidean distance to distinguish brain tumors (glioblastoma, medulloblastoma, meningioma) from normal tissues in-vitro [88]. Analyzing 172 spectra, their model achieved 0.97 sensitivity, highlighting lipid, cholesterol, and protein signatures relevant for tumor discrimination. In 2018, Desroches et al. conducted an in-vivo study involving 19 patients undergoing neurosurgical

resection procedures. Using a handheld fiber-optic RS probe, they collected 280 spectra samples and showed that RS could be used to differentiate among normal brain, infiltrated brain, and dense cancer tissue, with a sensitivity and specificity of 0.80 and 0.90 [67]. An in-vitro study by Iturrioz-Rodríguez et al. focused on distinguishing glioblastoma multiforme (GBM) cells from healthy astrocytes using RS. By analyzing the spectral region between 1000 and 1300 cm^{-1}, which includes peaks related to DNA/RNA and cytochrome c, the model achieved an average discrimination accuracy of 0.92, highlighting key molecular signatures relevant to brain cancer diagnosis [89]. More recently, Ember et al. conducted a multicenter study to differentiate tumor from non-tumoral brain tissue in-situ [10]. Using 976 measurements from 67 patients with glioblastoma, brain metastases, or meningioma, the system achieved diagnostic accuracies of 0.91, 0.97, and 0.96, respectively, demonstrating the potential of RS for guiding surgical resections in the brain.

3.2.2 Raman spectroscopy for prostate

Following its success in detecting various diseases, RS has also been explored for the characterization of prostate tissues. However, it faces challenges such as lower signal quality and higher autofluorescence compared to applications in brain or breast tissue. Some studies have focused on distinguishing prostate from extraprostatic tissue, using RS probes integrated into a da Vinci surgical robot and a dual excitation RS system to enable real-time tissue classification during minimally invasive prostatectomy [90]. In addition to tissue discrimination, RS has also been investigated as a tool for monitoring treatment response, with proof-of-concept studies showing its ability to detect biochemical changes following HDR brachytherapy in prostate tissue [91].

Several studies have investigated the use of RS for the detection and characterization of PCa [1]. In early work by Crow et al., snap-frozen samples from 27 patients were scanned using a Raman microscope; PCA and linear discriminant analysis were implemented to classify samples with GS < 7, GS $= 7$, or GS > 7, achieving an overall classification accuracy of 0.89 [92]. Wang et al. extended the use of RS to castration-resistant prostate cancer (CRPC), analyzing Raman spectra from cell lines and patient tissues (50 patients in total) using PCA combined with SVM classification. Their model achieved sensitivity and specificity of 0.88 for identifying CRPC tissues [93]. In another study, Aubertin et al. collected 477 Raman spectra from 18 post-prostatectomy specimens, using a dual-excitation Raman system and a single-point probe to capture both FP and HW regions. Classification using SVMs showed that combining the two spectral regions improved cancer detection performance, with an AUC of 0.91. Figure 3.8 shows the average Raman spectra of the study, covering both the fingerprint

and high wavenumber regions. The authors also compared different cross-validation strategies and obtained similar classification results when combining both regions, suggesting minimal patient-specific bias in the spectra [11]. In a diagnostic accuracy study by Grosset et al., formalin-fixed paraffin-embedded prostate tissue specimens from 483 patients were analyzed using Raman micro-spectroscopy and SVMs to differentiate intraductal carcinoma of the prostate (IDC-P), PCa, and benign tissue, identifying spectral similarities between IDC-P and the adjacent PCa. [94]. More recently, van Breugel et al. conducted a large clinical trial on fresh ex-vivo biopsy cores from 84 patients, acquiring 2,395 spectra from 599 cores. Their classification model demonstrated a sensitivity of 0.90 and a specificity of 0.80 for detecting clinically significant PCa [21]. Across these studies, either a leave-one-out (LOO) or a leave-one-patient-out (LOPO) cross-validation strategy was employed, and all emphasized the importance of precise histopathological mapping to ensure high-quality label assignment to the characterized sites.

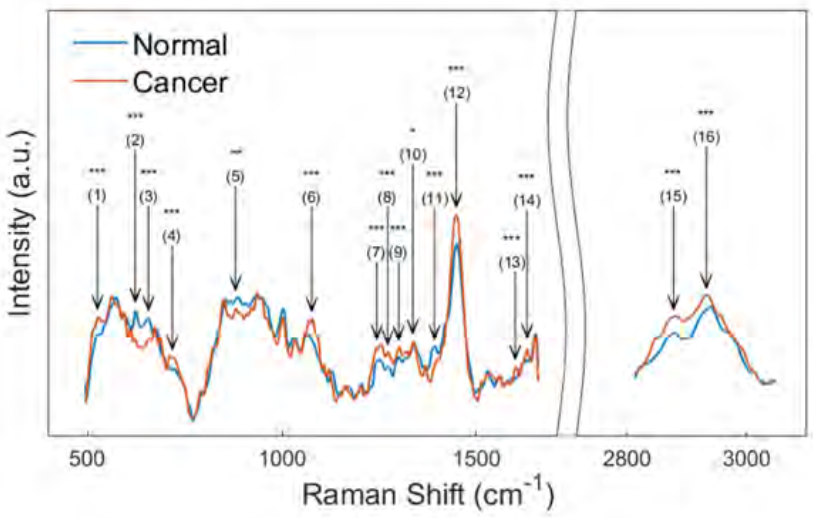


Figure 3.8 Fingerprint (490–1660 cm^{-1}) and high wavenumber (2800–3050 cm^{-1}) average Raman spectra for normal (N = 393) and cancer (N = 84) prostatic tissues. The arrows identify prominent Raman peaks for which the results of univariate statistical tests are identified as follows: *** for $p<0.001$, ** for $p<0.01$, and * for $p<0.05$. (CC BY) [11].

3.3 Artificial intelligence in biomedical applications

Building upon the advances in imaging and spectroscopy presented in previous sections, AI techniques have emerged as essential tools for the analysis and interpretation of biomedical data. Machine learning (ML) and deep learning (DL) models are widely used to extract

meaningful patterns from high-dimensional datasets, enabling tasks such as classification, clustering, and regression. ML models, such as decision trees, SVMs, and logistic regression, typically rely on a predefined set of numerical features and vary in their complexity and interpretability. The DL approaches, including CNNs and recurrent neural networks (RNNs), offer the ability to learn hierarchical representations directly from complete data automatically. While increasing model complexity can improve performance on large datasets, it also elevates the risk of overfitting, particularly in experiments with high-dimensional data (e.g., RS and mpMRI) with limited sample sizes and cohorts, inherent in early-stage studies. Careful management of model complexity and the use of robust validation strategies are therefore critical to ensure generalizability and prevent performance degradation on new, unseen data [13, 14, 95].

3.3.1 Classical machine learning for spectral classification

One of the ML models commonly used for supervised classification, random forest (RF), consists of an ensemble of decision trees, offers robustness to overfitting, and provides internal measures of feature importance [50, 62, 91]. Across both medical (in-vivo and ex-vivo) and non-medical applications, SVM models are among the most widely used for Raman spectra classification [10, 23, 46, 47, 90, 91, 96]. The SVMs construct an optimal hyperplane that maximizes the margin between classes in a transformed feature space, making them particularly effective in high-dimensional settings (Fig.3.9) [12, 21, 50, 62]. Given the large number of features derived from RS (where each datapoint —wavenumber cm^{-1}— can be treated as an individual feature) and mpMRI radiomics, dimensionality reduction and feature selection are critical steps to prevent overfitting and improve model interpretability. Techniques such as PCA, Lasso regression, handcrafted feature engineering, and peak selection based on known molecular vibration states are frequently employed [66, 69, 97]. Selecting peaks or radiomics features enhances classification performance and provides biologically meaningful insights into tissue composition and pathology.

The SVMs have been widely used in both spectroscopy and radiomics studies. For instance, Wang et al. employed an SVM classifier with a radial basis function (RBF) kernel to detect PCa based on mpMRI radiomics features [8]. Feature selection was performed using SVM-recursive feature elimination, a method that iteratively removes less informative features to optimize classification performance. Using radiomics features alone, their SVM model achieved an AUC of 0.95 for discriminating PCa from normal transitional zone (TZ) tissue (Fig.3.4) [8]. In breast cancer cell classification, Zhang et al. implemented a PCA–SVM pipeline to analyze Raman spectra acquired from cultured cells. First, PCA was used for

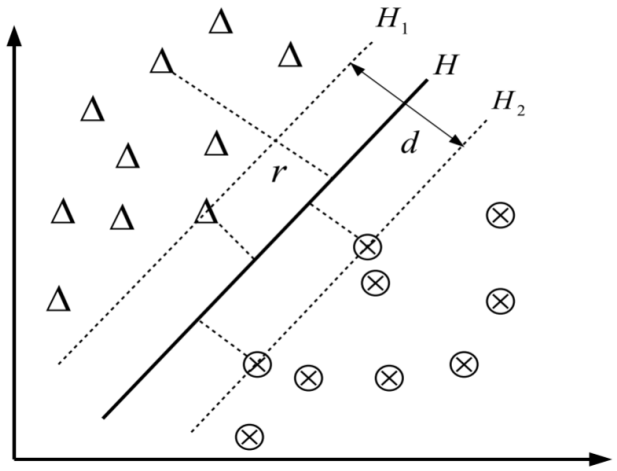


Figure 3.9 Linearly separable samples indicated in a hyperplane. H: hyper planes; d: maximal margin (CC BY) [12].

dimensionality reduction, followed by SVM classification using a one-versus-one strategy, where binary SVMs were trained to distinguish among subtypes, achieving accuracies greater than 0.97 for cancer detection and greater than 0.92 for subtype classification [86]. Similarly, for the diagnosis and prognosis of CRPC, Wang et al. applied a PCA–SVM framework to Raman spectra acquired from prostate tissues and cell lines. The authors used PCA to reduce spectral dimensionality, and an SVM with an RBF kernel and particle swarm optimization solver was trained and evaluated using LOO cross-validation, achieving a sensitivity and specificity of 0.88 for CRPC detection [93].

However, classical ML approaches often rely on manual feature engineering and struggle with scalability when faced with large, unstructured datasets. Although dimensionality reduction techniques such as PCA help summarize variance, they may discard subtle yet informative features critical for classification, and some of them further reduce model interpretability. Moreover, the overall performance of classical methods remains highly dependent on the quality and relevance of the selected features. These limitations have motivated the adoption of DL approaches, which can perform automatic feature extraction directly from high-dimensional inputs when sufficient data are available.

3.3.2 Deep learning for cancer detection

Diverse DL methods have been successfully applied to complete mpMRI images for PCa classification, detection, and segmentation tasks (as discussed in Section 3.1.1), bypassing the need for handcrafted or radiomics feature extraction, trading off interpretability. In a broader

context, multimodal DL approaches such as ARTERA (ResNet+CatBoost) have been developed to integrate clinical, molecular, and imaging data from digitized histopathological slides to improve risk stratification and treatment decision-making in PCa [98]; However, ARTERA does not leverage mpMRI or RS data. Beyond histopathology, AI tools are increasingly used to guide therapy by integrating immunohistochemical, genomic, proteomic, and clinical data. These models help identify biomarkers, predict treatment response, personalize therapies, and explore factors affecting efficacy, such as molecular features or patient demographics [99, 100]. Despite these advances, many models remain limited by the size and heterogeneity of training datasets, underscoring the need for robust, large-scale validation [100, 101]. In this project, we specifically focus on applying DL to 1D RS signals and to 1D radiomics feature vectors extracted from mpMRI, rather than developing models based on full medical images or predictive clinical models.

Several studies have demonstrated the utility of DL methods for cancer detection based on 1D signals. In 2020, Santilli et al. employed an autoencoder architecture (encoder-decoder based on fully connected layers) for basal cell carcinoma detection using mass spectrometry technology (iKnife), achieving 0.96 accuracy after signal augmentation [102]. Among DL approaches, CNNs offer key advantages over classical ML methods by enabling task-specific feature extraction directly from raw, high-dimensional inputs, thus eliminating the need for manual feature engineering. In a CNN model, hierarchical convolutional filters automatically learn local patterns and progressively abstract higher-level features, allowing robust classification even in complex datasets. Although CNN architectures were initially designed for 2D data and have achieved high performance in well-known computer vision applications (e.g., AlexNet, VGG-Net, and ResNet), they can be adapted into one-dimensional (1D) architectures for processing 1D signals such as temporal sequences, electroencephalograms (EEG), or one-channel spectra. This adaptation provides significant computational and data efficiency gains compared to transforming 1D signals into 2D representations (using field transformation such as Gramian angular fields), which have also been explored for spectral data analysis [103].

In contrast, Ma et al. combined RS with a 1D-CNN architecture for the binary classification of healthy and cancerous breast tissues using Raman spectra collected from 20 patients [13]. Their model consisted of a convolutional layer for local spectral feature extraction, followed by a pooling layer and two dense layers for global representation and classification (Fig.3.10). After training with data augmentation, the 1D-CNN achieved an overall accuracy of 0.92, with a sensitivity of 0.98 and a specificity of 0.86. For comparison, they also trained classical ML models, including SVMs with various kernels (the linear kernel achieving the best performance with an accuracy of 0.86) and Fisher Discriminant Analysis (FDA), both

using PCA for dimensionality reduction. The 1D-CNN consistently outperformed the SVM and FDA classifiers, demonstrating superior sensitivity and specificity, and highlighting the advantages of automatic hierarchical feature extraction from raw spectral inputs enabled by DL approaches.

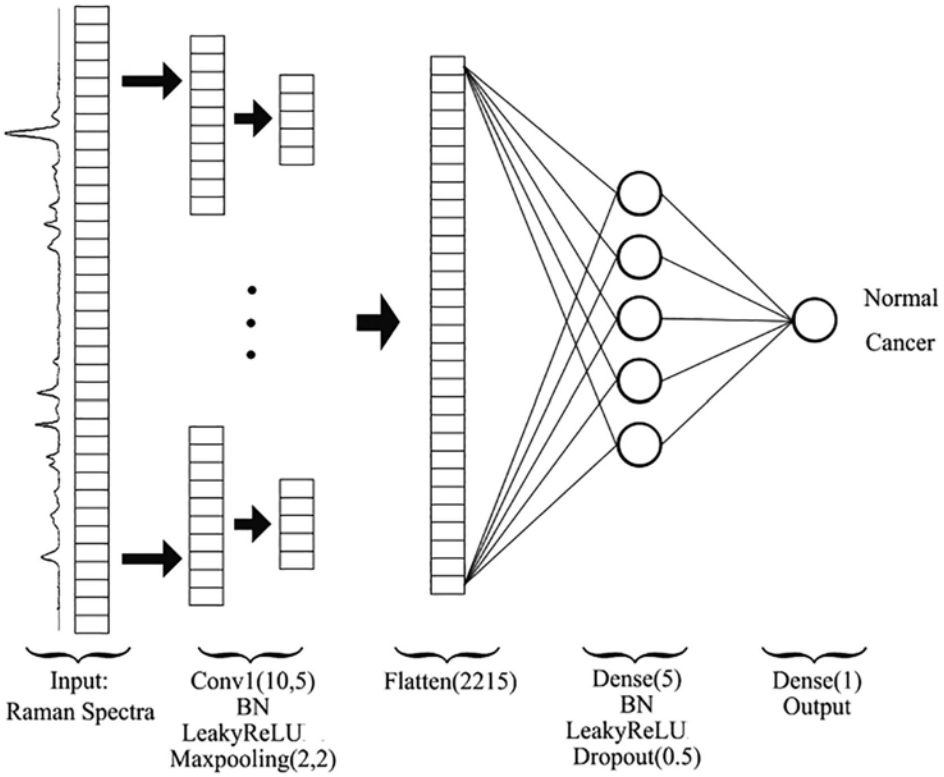


Figure 3.10 The structural diagram of the 1D-CNN classification model for Raman spectra. (CC BY) [13].

In domains with larger datasets, such as bacterial and mineral classification, Zhou et al. proposed a deeper CNN architecture, adapted from VGG architecture with deep residual shrinkage (DRS) blocks —called DRS-VGG— (Fig.3.11). This model integrates residual connections with soft thresholding mechanisms to eliminate redundant spectral signals and enhance feature extraction, and it was compared against ML approaches (SVM, PCA-SVM, RF, and k-Nearest Neighbors (KNN)) [14]. Their model, built for multiclass classification, achieved top-1 accuracies of 0.98, 0.93, and 0.95 across three Raman datasets —compared to SVM accuracy results of 0.96, 0.55, 0.92 on the same three subsets. Trained on over 10,000 mineral and 60,000 bacterial spectra without the need for data augmentation, the DRS-VGG network leveraged its depth and regularization strategies (e.g., dropout and batch normalization) to achieve both high accuracy and stable predictions with minimal variance. Although SVM performed competitively on cleaner mineral and bacterial datasets, its accu-

racy dropped to nearly 0.55 on more challenging data subsets, highlighting its sensitivity to signal heterogeneity.

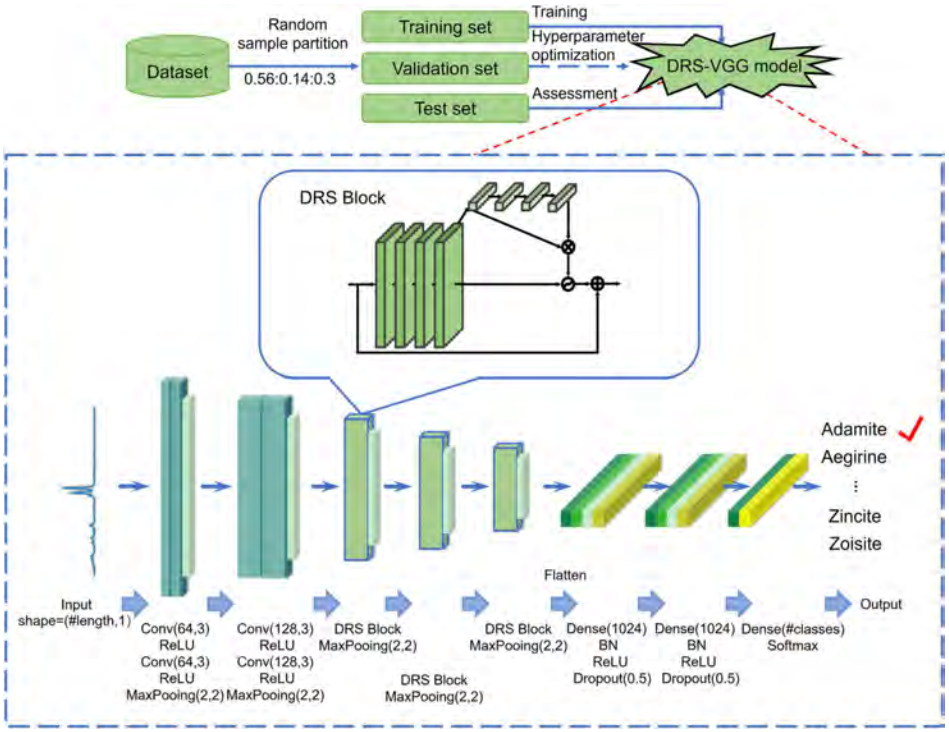


Figure 3.11 Diagram of the proposed DRS-VGG network for Raman spectral identification process. The DRS-VGG network is a one-dimensional neural network stacked by convolutional layer, activation function layer, maxpooling, DRS block, BN, and so on. DRS: deep residual shrinkage (CC BY) [14].

The availability of large volumes of data opens the possibility of using more complex models (e.g., deeper architectures, recurrent blocks, and more robust validation strategies), as evident when comparing [13] and [14]. In general, CNNs require large and diverse datasets to achieve good generalization, a requirement that often becomes a major limitation in clinical applications and specialized domains like RS-based tissue classification.

Transfer learning and domain Adaptation

A common challenge in DL medical applications is the limited availability of labeled data specific to the task of interest. This scarcity poses a significant obstacle to developing robust DL models, which typically require large, diverse datasets to generalize well. Transfer learning (TL) has emerged as an effective strategy to address this limitation by leveraging knowledge acquired from related tasks. Instead of training models from scratch, TL enables the initialization of models with features learned from large datasets in similar or broader

domains. To overcome domain shifts, the model is subsequently adapted to the specific target task through fine-tuning, test-time adaptation (TTA), domain adaptation, or related techniques. TL not only addresses data scarcity but also reduces training time and computational requirements. In medical imaging, models pre-trained on large-scale natural image datasets (e.g., ImageNet with 1.28M images) are frequently adapted to medical imaging tasks with far fewer samples (often fewer than 600 images) [104]. Even foundation models like MedSAM, developed on over 1M medical images for universal segmentation across diverse modalities and cancer types, are initialized with weights from SAM, a model trained initially on 11M natural images [105].

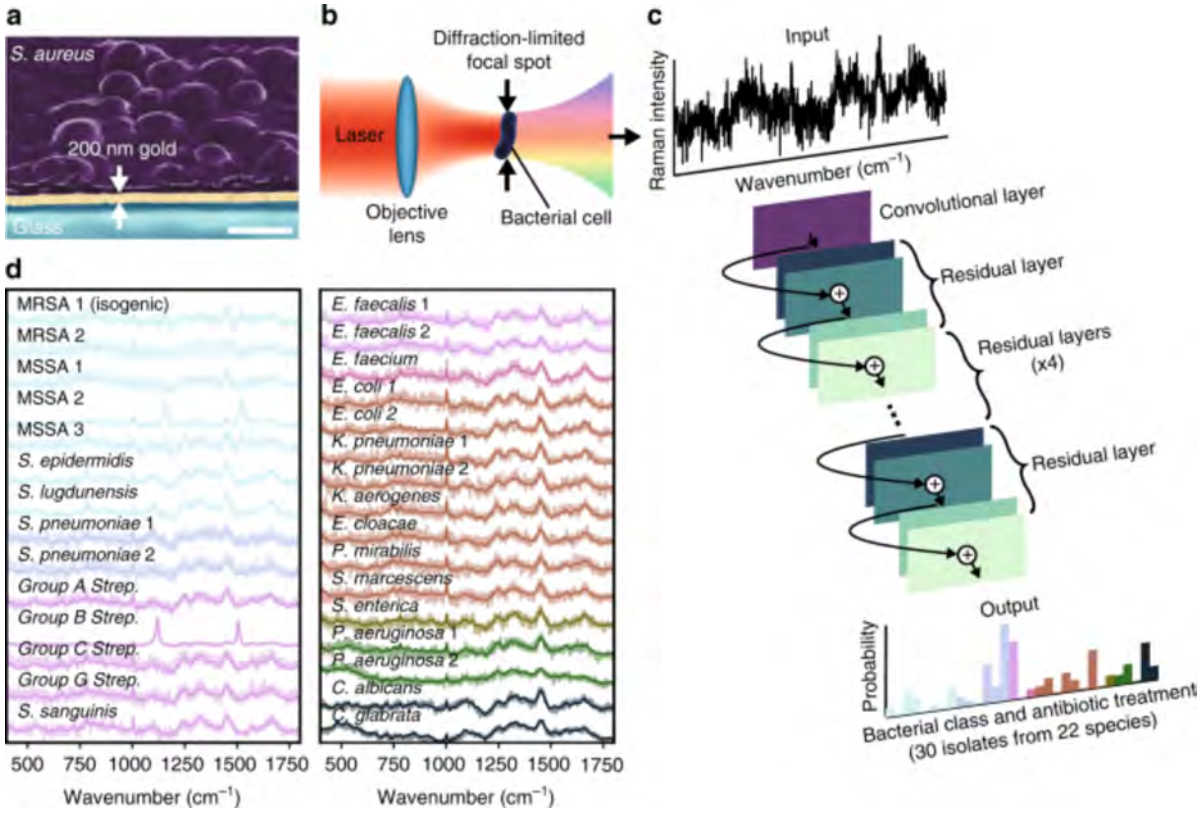


Figure 3.12 A convolutional neural network (CNN) can be used to identify bacteria from Raman spectra. a) To build a training dataset of Raman spectra, we deposit bacterial cells onto gold-coated silica substrates and collect spectra from 2,000 bacteria over monolayer regions for each strain. b) Conceptual measurement schematic for Raman signal from single cells. c) Using a one-dimensional residual network with 25 total convolutional layers, low-signal Raman spectra are classified as one of 30 isolates. d) Raman spectra of bacterial species can be difficult to distinguish, and short integration times (1 s) lead to noisy spectra (SNR = 4.1). (CC BY) [15].

One demonstration of TL with fine-tuning (supervised learning) in RS was presented by Ho

et al. [15]. To overcome challenges associated with the low signal-to-noise ratio of bacterial Raman spectra and the limited number of clinical samples, the authors developed a ResNet-inspired CNN architecture with 26 layers (Fig.3.12). The model was initially pre-trained on a large reference dataset comprising 60,000 spectra from 30 bacterial and yeast isolates under controlled conditions, achieving isolate-level classification accuracies of 0.82. Fine-tuning of all CNN parameters was then performed using clinical patient isolates (10 spectra per patient), enabling the model to generalize from standardized laboratory data to the more heterogeneous clinical domain and achieving isolate-level classification accuracies exceeding 0.90. Similarly, in the field of brain–computer interfaces, Zhang et al. addressed inter-subject variability in EEG signals by fine-tuning a deep CNN model pre-trained on data from multiple subjects [16]. They proposed five adaptation schemes —in each scheme, a different number of layers were frozen (non-adapted) and the remaining layers were adapted during fine-tuning— and adjusted the learning rate, improving the classification accuracy of motor imagery tasks to 0.84, significantly outperforming previous subject-independent approaches (Fig.3.13).

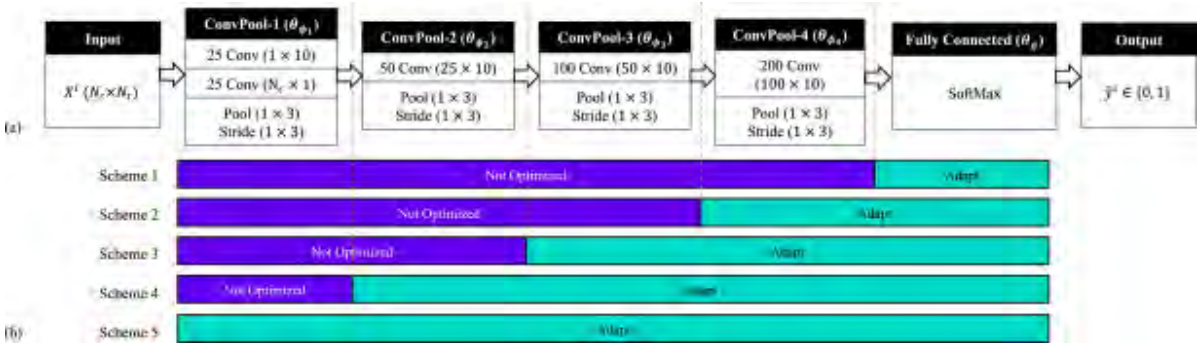


Figure 3.13 Illustrations of (a) Network architecture and (b) Adaptation schemes: optimizing a subset of model parameters [16]. (CC BY).

Unsupervised learning methods, particularly those based on TTA, have become increasingly important when labeled data are unavailable during model deployment. In TL, where models are pre-trained on broad source data, TTA enables a model to adjust its parameters based only on incoming unlabeled test data (target data), improving generalization without requiring additional annotated datasets [106]. In the context of computer vision, Wang et al. proposed TENT, a pioneering TTA method based on entropy minimization [107]. Their approach updates only the affine parameters of normalization layers during testing, while keeping the rest of the network fixed. Moving to signal-based applications, Mao et al. introduced a TTA framework for online epilepsy seizure prediction from EEG signals [108]. Their method addresses the privacy constraints of source-free domain adaptation (i.e., the absence of labeled source data used for pre-training the model) and tackles the challenges of

online learning without labeled data. The proposed framework combines a brief fine-tuning phase on a single seizure event with continuous TTA during deployment. Specifically, only the task-specific heads are updated, while the shared feature extractor remains frozen to preserve source domain knowledge and enhance computational efficiency. Given the success of TTA in imaging [107] and EEG-based applications [108], there is an interest that expands into other fields. To the best of my knowledge, the application of TTA to RS for cancer detection has not yet been reported.

3.4 Summary

This literature review explored different technologies and methodologies that play a fundamental role in detecting, treating, and managing PCa. Despite significant progress in imaging, spectroscopy, and machine learning and deep learning, several challenges persist in the accurate and timely detection of PCa, especially in the context of targeted prostate interventions. Diverse navigation systems have enhanced tumor targeting by integrating multimodal image fusion and real-time tracking capabilities. Nonetheless, small spatial inaccuracies remain and must be considered during treatment planning (e.g., when defining the planning target volume). Moreover, the reliance of current systems on cancer-specific information from preoperative imaging limits their capacity to confirm malignancy intraoperatively. Beyond visual interpretation, mpMRI can also be analyzed quantitatively by extracting radiomics features, enabling new types of analysis and facilitating integration with other diagnostic modalities.

This chapter also provided a detailed overview of RS biomedical applications, particularly in cancer detection. As a label-free optical modality, RS offers a promising solution by enabling real-time acquisition of molecular signatures from biological samples. However, the clinical integration of this technology remains limited by several challenges. The inherently weak Raman signal and the high dimensionality of spectral data require robust preprocessing, feature selection, and feature engineering techniques. Additionally, developing reliable classification models can be limited by factors such as inaccuracies in ground truth annotations (e.g., site-specific vs. core-level labels) and the need for disease- and organ-specific data acquisition.

Finally, the AI-based approaches used to analyze this rich information (spectral features or image-based radiomics) were also discussed in this section. These methods must account for the high dimensionality of the input data and the typically small sample sizes available for training. Classical ML methods, such as SVMs, are widely used in cancer detection based on RS or radiomics and often achieve good classification performance, but they typically require

extensive feature engineering/selection. On the other hand, DL models, especially 1D-CNNs, enable automatic feature extraction and typically achieve superior performance, though they require large annotated datasets to be effective.

Together, these limitations highlight a critical gap: the absence of real-time, in-situ diagnostic tools that can integrate navigation, molecular characterization, and robust classification. Bridging this gap is essential to support intraoperative decision-making, enhance targeting accuracy, and enable more personalized therapy in PCa interventions.

CHAPTER 4 RESEARCH METHODOLOGY

Despite advances in imaging for intervention planning and navigation, as well as histopathologic analysis for post-procedural validation, a crucial need remains for intraoperative cancer confirmation. A solution that provides real-time, localized answers to this challenge —and integrates seamlessly into current PR workflows— could improve the efficacy of targeted treatments and offer timely information to support decision-making or procedure adaptation directly in the PR.

The research question that provides a guide and motivation for the present project is: **Does the combination of Raman spectroscopy and artificial intelligence-based classification models enable real-time, localized in-situ confirmation of prostate cancer to support tumor-targeted interventions?** The hypotheses and research objectives on the path to answering this question are described below.

4.1 Hypotheses and objectives

The following three hypotheses served as the foundation for this study:

Hypothesis 1: A system integrating real-time 2D TRUS, EM tracking, and 3D imaging enables accurate navigation of an optical probe to predefined targets within the prostate.

Hypothesis 2: The integration of in-situ RS with mpMRI-based radiomics features enables highly accurate PCa detection.

Hypothesis 3: Accurate real-time confirmation of PCa in the PR can be achieved using DL models applied to RS from fresh biopsy samples.

Based on the three hypotheses, four specific objectives (and their sub-objectives) were defined, each aligning with one of the studies presented in Chapters 5 to 8.

Objective 1: Evaluate a navigation system that integrates TRUS, EM tracking, and 3D imaging for accurate tumor targeting.

1.1 Evaluate the performance of a prototype interventional system (UroNav) in the early phase of clinical evaluation in 43 consecutive interventions, by assessing workflow efficacy, EM catheter reconstruction accuracy, and MRI-TRUS registration accuracy.

Objective 2: Develop and optimize a multimodal classification model for PCa by integrating in-vivo RS and mpMRI-based radiomics features.

2.1 Acquire RS and mpMRI-based radiomics data from 18 patients undergoing HDR brachytherapy, using an in-house developed navigation system.

2.2 Train SVM models using features independently extracted from RS and radiomics to evaluate their individual and combined discriminative capabilities.

2.3 Optimize model complexity by reducing the number of input features to 10 while maintaining a classification accuracy of at least 0.80.

Objective 3: Implement and evaluate a robotic system for real-time RS data acquisition from biopsy cores in the PR.

3.1 Integrate a robot-assisted RS system for automated spectral acquisition from fresh biopsy cores, targeting a positioning accuracy <0.4 mm, and evaluate its repeatability and acquisition time.

3.2 Conduct preliminary evaluations of CNN models in retrospective RS datasets (prostate, breast, and brain cancer), assessing their versatility and automatic feature extraction capabilities through binary classification performance, aiming to achieve accuracies >0.72, >0.90, and >0.78, respectively.

Objective 4: Develop and adapt deep learning models using transfer learning for intraoperative PCa classification.

4.1 Acquire RS data from biopsy samples of 10 patients undergoing HDR brachytherapy using the robot-assisted RS system.

4.2 Implement and optimize transfer learning strategies leveraging retrospective multi-organ RS datasets, with the aim of achieving classification accuracy of over 0.72 after adaptation to a limited intraoperative prostate dataset.

4.2 General methodology

The research activities conducted in this feasibility study were structured into two phases, as illustrated in Figure 4.1: Phase I focuses on in-vivo experiments (Objectives 1 and 2) and Phase II focuses on ex-vivo experiments (Objectives 3 and 4). In both phases, experimental

procedures were carried out during HDR brachytherapy interventions at the radiation oncology department, Centre Hospitalier de l'Université de Montréal (CHUM). To contextualize the integration of the proposed methods, it is first necessary to present the clinical workflow of a prostate HDR brachytherapy procedure.

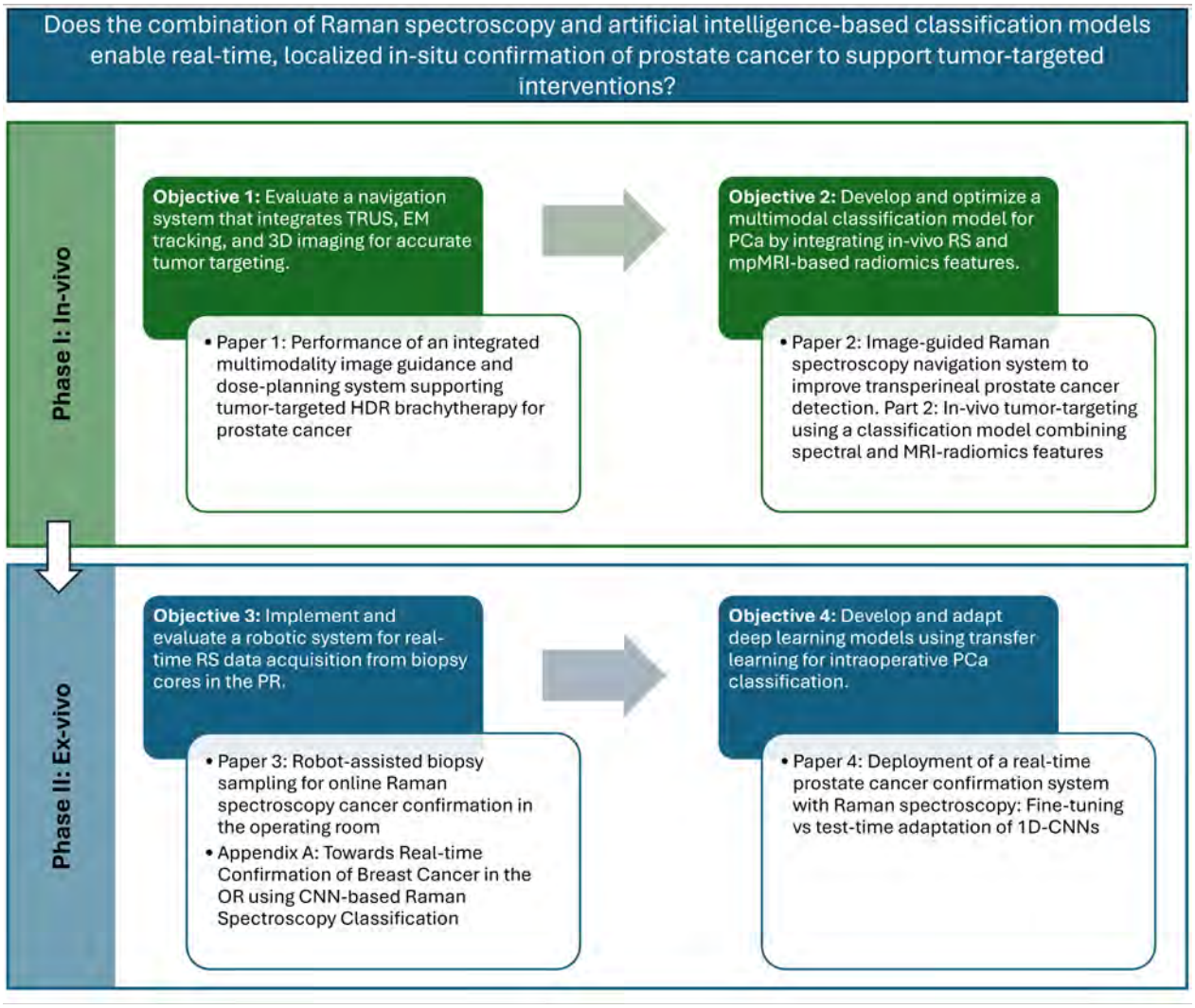


Figure 4.1 Methodological organization of the thesis.

The typical workflow of prostate HDR brachytherapy includes the following steps: (1) preoperative planning using mpMRI, (2) intraoperative guidance with TRUS, (3) image fusion, (4) catheter implantation, (5) dose planning and delivery, and (6) postoperative histopathological confirmation when indicated (Fig.4.2). The integration of RS-based tissue characterization impacts this workflow differently depending on whether it is applied in-vivo or ex-vivo on excised biopsy samples. These distinctions, along with the design of each research phase, are detailed in the following sections.

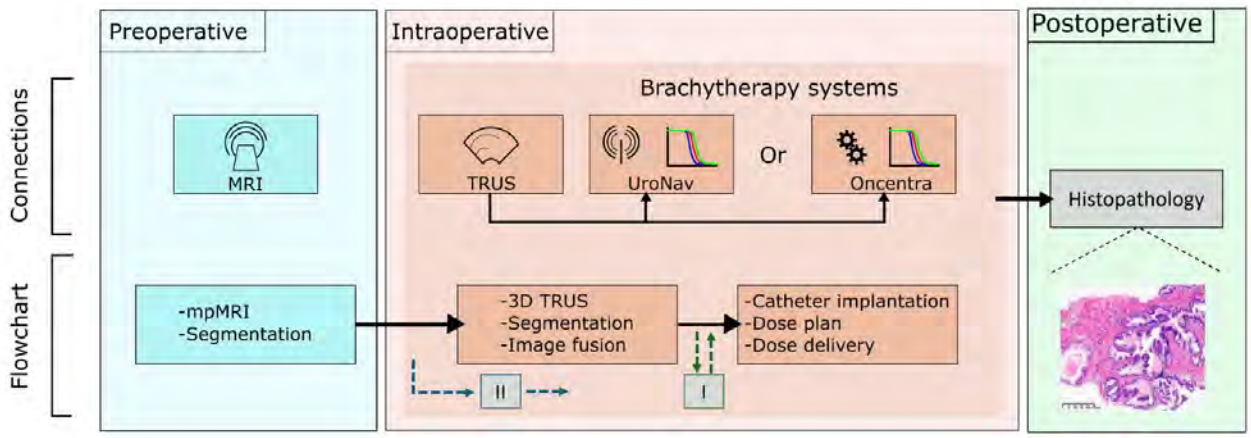


Figure 4.2 Flowchart illustrating the typical workflow of prostate High-dose-rate (HDR) brachytherapy. The diagram is organized into three phases: preoperative planning, intraoperative guidance, and postoperative confirmation. Two of the systems that can be used to support the intervention (UroNav and Oncentra) are illustrated; blocks "I" and "II" represent the place where the activities related to each research phase are introduced.

Phase I —In-vivo measurements

Phase I of the study focuses on acquiring in-vivo RS measurements during prostate HDR brachytherapy. A prototype interventional system (Invivo/UroNav, Philips Disease Management Solutions, Gainesville) was used to assist the brachytherapy procedure, specifically supporting tumor-targeting. This system integrates TRUS-MRI fusion and EM tracking to guide catheter implantation. The evaluation of the performance of this prototype interventional system (Objective 1) is critical for two main reasons. First, as an early-phase clinical device, its assessment provides insights into the practical benefits and limitations of image fusion and tracking technologies in prostate interventions. Second, the fused images and EM tracking hardware from the prototype system are also used by the in-house RS navigation system to guide intraoperative probe placement. The system's performance —regarding workflow efficiency, catheter reconstruction accuracy, and registration accuracy— is detailed in Chapter 5.

For RS acquisition, the standard brachytherapy workflow was momentarily paused before catheter insertion (block "I" in Fig.4.2). During this interval, the RS probe was guided to pre-identified suspicious regions using the in-house navigation system, enabling spectral acquisition, radiomics feature extraction, and biopsy collection (Fig.4.3). After data acquisition, the brachytherapy procedure is resumed as planned. The models, experimental design, and results derived from the acquired multimodal data (Objective 2) are presented in Chapter 6. While the RS navigation system is described in that chapter, its pre-clinical validation

was previously published in a co-authored paper not included as a core contribution of this thesis [109].

Figure 4.3 Clinical setup —Phase I: (a) ultrasound system; (b) near-infrared laser and spectrometer; (c) EM tracking system; (d) EM field generator; (e) 3D Slicer navigation system. Right: Photograph of the system deployed during a brachytherapy procedure at CHUM.

Phase II —Ex-vivo measurements

Phase II focuses on acquiring ex-vivo RS measurements from biopsy cores collected with a 23-gauge biopsy gun. Due to the small size and fragility of the samples, precise probe positioning is critical. To ensure reproducibility and accuracy, a robotic arm was integrated for sample manipulation in preparation for data acquisition in PR (Objective 3). The design, calibration, and clinical deployment of this system are described in Chapter 7. Chapter 7 also presents the CNN-based cancer detection results on retrospective RS datasets (prostate, breast, and brain cancer). This exploration builds upon prior work on breast cancer detection using CNNs, as presented in a workshop proceedings paper [110].

Prospective measurements in Phase II were also acquired during HDR brachytherapy procedures assisted by the Oncentra platform, with ex-vivo RS acquisition running in parallel to the treatment and causing minimal workflow interruption (block "II" in Fig.4.2). The experimental setup (Fig.4.4), the CNN-based PCa detection models, and the robotic RS system (Objective 4) are detailed in Chapter 8. Given the limited patient cohort typical of feasibility studies, this chapter also explores transfer learning approaches —such as fine-tuning and TTA— to leverage retrospective multi-organ spectra for pre-training, followed by adaptation to prospective prostate data.

In summary, the specific methods, experiments, and results related to each objective are detailed in the subsequent chapters: Chapter 5 (Objective 1), Chapter 6 (Objective 2),

Figure 4.4 Clinical setup —Phase II: (a) optical probe; (b) robotic arm; (c) emergency stop; (d) optical system; (e) 3D Slicer navigation system. Right: Photograph of a biopsy sample on the platform, during a brachytherapy procedure at CHUM.

Chapter 7 (Objective 3), and Chapter 8 (Objective 4). The overall discussion and conclusions are presented in Chapters 9 and 10.

CHAPTER 5 ARTICLE 1: PERFORMANCE OF AN INTEGRATED MULTIMODALITY IMAGE GUIDANCE AND DOSE-PLANNING SYSTEM SUPPORTING TUMOR-TARGETED HDR BRACHYTHERAPY FOR PROSTATE CANCER

This article was published in the journal *Radiotherapy and Oncology*, November 2021.
Remarks: The first author contributed to the experimental measurements, data processing, bibliographic research, and manuscript writing, representing approximately 90% of the work.

Performance of an integrated multimodality image guidance and dose-planning system supporting tumor-targeted HDR brachytherapy for prostate cancer

[1,2]David Grajales, [1,2]Samuel Kadoury, [2]Roozbeh Shams, [2,3]Maroie Barkati, [2,3]Guila Delouya, [2]Dominic Béliveau-Nadeau, [2]Benedicte Nicolas, [1]William T. Le, [2]Mustafa-Karim Benhacene-Boudam, [2,3]Daniel Juneau, [2,3]Jean N. DaSilva, [2,3]Jean-Francois Carrier, [2]Gilion Hautvast, [2,3]Cynthia Ménard,

[1]Polytechnique Montréal, Montréal, Québec, Canada
[2]Centre de recherche du Centre Hospitalier de l'Université de Montréal, Montréal, Québec, Canada
[3]Université de Montréal, Montréal, Québec, Canada
[4]Philips Healthcare, Imaging Systems, Netherlands

Abstract

Background and purpose: Advances in high-dose-rate brachytherapy to treat prostate cancer hinge on improved accuracy in navigation and targeting while optimizing a streamlined workflow. Multimodal image registration and electromagnetic (EM) tracking are two technologies integrated into a prototype system in the early phase of clinical evaluation. We aim to report on the system's accuracy and workflow performance in support of tumor-targeted procedures. **Materials and methods:** In a prospective study, we evaluated the system in 43 consecutive procedures after clinical deployment. We measured workflow efficiency and EM catheter reconstruction accuracy. We also evaluated the system's MRI-TRUS registration accuracy with/without deformation, and with/ without y-axis rotation for urethral alignment at initialization. **Results:** The cohort included 32 focal brachytherapy and 11 integrated boost whole-gland implants. Mean procedure time excluding dose delivery was 38 min (range: 21–83) for focal, and 56 min (range: 38–89) for whole-gland implants; stable

over time. EM catheter reconstructions achieved a mean differ- ence between computed and measured free-length of 0.8 mm (SD 0.8, no corrections performed), and mean axial manual corrections 1.3 mm (SD 0.7). EM also enabled the clinical use of a non or partially visible catheter in 21% of procedures. Registration accuracy improved with y-axis rotation for urethral align- ment at initialization and with the elastic registration (mTRE 3.42 mm, SD 1.49). **Conclusion:** The system supported tumor-targeting and was implemented with no demonstrable learning curve. EM reconstruction errors were small, correctable, and improved with calibration and control of external distortion sources; increasing confidence in the use of partially visible catheters. Image registra- tion errors remained despite rotational alignment and deformation, and should be carefully considered.

Keywords: Prostate cancer, Brachytherapy, Image registration, Electromagnetic tracking, Navigation system.

5.1 Introduction

High-dose-rate (HDR) brachytherapy is an established treatment modality for localized prostate cancer (PCa), the most common cancer in men in North America (excluding non-melanoma skin cancer) with more than 235,000 new cases in 2020 [36, 111]. By virtue of underlying physics, the safety and effectiveness of this modality hinge on accurate targeting [43]. Transrectal ultrasound (TRUS) is commonly used to guide brachytherapy, biopsies, and other prostate interventions due to its real-time nature, ease of use, and low cost [35]; nevertheless, most cancers are invisible in TRUS [33–35]. On the other hand, multiparametric MRI and prostate-specific membrane antigen (PSMA)-PET/CT have greater sensitivity and specificity (respectively) for the localized characterization of PCa [29,31,32,35]. However, these are not optimal modalities to guide interventions due to longer acquisition time, limited compatibility with surgical instruments, and cost [35, 36].

Multimodal image registration is currently studied and used to take advantage of complementary information to assist HDR brachytherapy and other tumor-targeted prostate interventions. Some of these interventional systems, such as BiopSee (Pi Medical), PredictiveFusion (MIM software) and SmartFusion (Canon medical Systems), work by rigidly fusing preoperative MRI with intra-operative TRUS [5, 32, 53, 54]. However, rigid registration accuracy is not ideal due to gland deformation caused by an altered patient position, TRUS-probe placement, or brachytherapy catheter insertion; these accumulated uncertainties can lead to target registration errors over 5mm [35, 53]. To compensate for these variations, several elastic registration methods have been developed based on target landmarks, organ surfaces, and image intensity [5, 24, 35, 55, 56]. Fusion systems like Artemis (Eigen) and Urostation

(Koelis) are already used to guide prostate biopsies, applying multimodal elastic image registration [18, 32, 57].

In HDR brachytherapy workflow, the accurate placement and localization of catheters are as important as the localization of treatment targets. Current TRUS-guided systems allow manual catheter reconstruction where depth is inferred from measured catheter free-lengths (defined as the section of the known-length catheters extending from the template), which is a time-consuming process; when multiple catheters are in place, interference could affect localization accuracy, significantly impacting dose delivery and treatment outcomes [6, 59]. Electromagnetic (EM) tracking is a technology used during minimally invasive interventions for catheter localization and reconstruction [59, 60]. The EM system consists of a field generator, a set of coils or sensors tracked within the generator's region of influence, and a control unit that processes sensors' signals to determine their location and orientation. The catheter reconstruction is done by introducing an EM tracked stylus into each catheter [6, 60, 61].

Given the complexity required to achieve high precision and accuracy during HDR brachytherapy procedures, systems that integrate all steps in a streamlined workflow are needed. In this study, we evaluate the performance of a prototype navigation system for HDR brachytherapy, which is in the early phase of clinical deployment. Our analysis of workflow performance, registration, and catheter reconstruction accuracy intends to determine the potential of a system that combines multimodal elastic image registration with EM tracking for HDR brachytherapy guidance.

5.2 Materials and methods

5.2.1 Clinical data and equipment

Between September 2019 and December 2020, 34 patients with histological diagnosis of PCa were enrolled on a prospective clinical trial approved by the Research Ethics Board (NCT03378856), and underwent HDR brachytherapy using the prototype system. Nine of these patients underwent two separate interventions. In total, 43 procedures were included in the analysis of the system's performance.

Prior to the procedure, a planning MRI (3D T2-weighted FSE, b2000 DWI, +/- DCE) was obtained on a 1.5 T Siemens Aera Magnetom (Siemens Healthineers, Erlangen, Germany) using surface coils. PSMA-PET/CT (18F-DFCPyL [19]) images were also acquired in a subset. For intraoperative imaging, two different ultrasound systems were used: BK Flex-Focus and bk3000 using a BK endocavity biplane transducer (BK Ultrasound, Herlev, Denmark).

The prototype system (Invivo/UroNav, Philips Disease Management Solutions, Gainesville,

USA) was designed to support prostate tumor-targeting with the integration of multimodal image registration, while improving procedural workflow efficiencies and catheter reconstruction accuracy through EM tracking. The integrated EM system (Aurora NDI, Waterloo, Canada) consists of an EM field generator placed over the patient's pelvic region, EM sensors placed on the template and TRUS-probe holder, and an EM stylus to reconstruct the catheters (Figure 1).

Figure 5.1 Clinical setup: A) prototype navigation and dose-planning system, B) ultrasound system, C) EM field generator-tracking system.

For accurate EM tracking, the system was calibrated twice: before the first case (calibration 1: including template and TRUS-probe sensors) and before intervention number 27 (calibration 2: including template and TRUS-probe sensors, this time with the metallic leg holders and revised probe holder in place) due to observed systematic shifts. The EM systems could be susceptible to distorting equipment present in the operation room, including cellphones. For the last 6 cases, the radiation oncologist (RO) performing the procedure was instructed not to carry a cellphone, removing a potential source of error.

5.2.2 Workflow

Pre-procedure: Planning images (MRI and PET) were co-registered for precise tumor localization, and manual segmentation of the prostate, GTV, PTV, and visible urethra was performed by the RO (Eclipse, Varian Medical Systems, Palo Alto, USA). T2-weighted images and corresponding segmentations were then transferred to the prototype system.

Intra-procedure: After set-up in the brachytherapy suite, baseline interventional 3D-TRUS images were reconstructed through manual sagittal sweep, followed by manual segmentation of the prostate and urethra/indwelling foley catheter. MRI and TRUS images and contours were first initialized with rigid registration based on prostate segmentations' center of mass (COM) and manual y-axis rotation (around right-left axis) for urethral alignment. Using the contour-based algorithm of the prototype system, deformable registration was then performed, enabling visualization of deformed structures projected over real-time TRUS within the navigation system. Manual segmentation of GTV and PTV on TRUS images was then performed based on these projections (propagated MRI contours) and underlying TRUS contrast. A pre-plan was generated based on a predefined prescription template, and proposed catheter sites were displayed.

Catheter implantation proceeded with attempts to maintain visibility on TRUS, and a final 3D-TRUS volume was acquired for treatment planning. The EM stylus was fully inserted in each catheter then slowly retracted to create a 3D reconstruction of dwell paths; the reconstruction was manually adjusted (in the axial plane) to align with the visible catheter in the TRUS images. Catheter depths were then verified against measured free-lengths. Segmented structures (TRUS prostate segmentation and the segmentations anchored to it) were subsequently adjusted as needed based on prostate displacements incurred with needle insertion, and a final treatment plan was generated with an inverse planning algorithm. All data (images, segmentations, transformation matrices, RT plan) were stored locally.

5.2.3 Performance evaluation

The open-source 3D-Slicer platform [112], version 4.11.0-2020-02-12, was used to visualize and process DICOM RTData. The system's evaluation focused on three aspects: workflow performance, catheter reconstruction accuracy, and elastic registration accuracy.

Related to workflow efficiency, we evaluated the duration of the intervention, calculated between the first 3D-TRUS acquisition and the last approved RTPlan. For catheter reconstruction accuracy, the measured free-length was compared to computed free-length as a

measure of depth accuracy; and the distance between catheter reconstruction corresponding points (considering only catheters' segment inside the prostate) before and after the manual adjustments.

For registration accuracy, MRI segmentations were rigidly registered to the TRUS volume using the initialization transformation matrix used and stored during the procedure, including the manual rotation when applied. The elastic deformation field was retrospectively calculated based on MRI and TRUS prostate surface contours, using the prototype system's algorithm. This deformation field was then applied to the MRI segmentations.

The following metrics were calculated before and after deformation to assess the rigid and deformable (elastic) registration accuracy:

Target Registration Error (TRE): Up to 5 landmarks (e.g., cyst, calcification) were screened and selected by an expert RO on both TRUS and MRI. The distances between the corresponding landmarks were calculated and compared using the SlicerRegistrationQA module. The mean TRE (mTRE) was then calculated for all the cases.

Urethral midline – 3D distance: Two sets of points (range: 16-25) were placed on the midline of MRI and TRUS urethra segmentation (in the center of equidistant cross-section). The MRI points were used to create a 3D model of the MRI urethral midline. The distances between the TRUS set of points and the closest point in the MRI urethral midline model were calculated and compared. This process was made using the SlicerCurveMaker extension.

Dice Similarity Coefficient (DSC) and 95%-Hausdorff distance: These two similarity metrics were used to verify that the surface-based deformation was working properly. They were calculated using the SlicerSegmentComparison module.

5.2.4 Impact of initialization procedures

In this application, the initialization point for the elastic registration is the position of the prostate segmentation in both modalities before starting the elastic deformation. Three different initializations were evaluated:

Centered: The volumes were centered using a translation based on the COM of the prostate segmentation ROIs from both modalities.

Autoaligned: The volumes were automatically aligned based on the urethral segmentation. In this method, the superior and inferior urethral points (center of urethra cross-section at the intersection of the prostate boundary) were connected by a line in both modalities. Then, the MRI volume was rotated, such that the lines would be oriented along a common axis. Finally, the MRI volume was translated to match the urethra inferior point (point of highest confidence) in both modalities.

Manual: After the COM translation applied by the system, the MRI volumes were rotated manually to align the urethra in y-axis, when the treating team considered it needed.

A 3D-Slicer extension (SegRegUrethra) was developed to calculate the transformation matrix for the centered and autoaligned initializations. For each initialization point, the deformation field was retrospectively calculated using the prototype system's algorithm. The same metrics described in the previous section were computed to compare the final effect of pre-alignment.

5.2.5 Statistical analysis

The differences were evaluated using a paired and two-samples Student's t-test (two-sided). A p-value <0.01 was considered as a statistically significant difference. Boxplots show the median (central mark), 25th-75th percentiles (box), extreme data not including outliers (whiskers) and outliers (+). Mean values are reported as mean(SD).

5.3 Results

Of the 43 cases, 32 consisted of focal brachytherapy implants, and 11 consisted of integrated boost (IB) implants. After automatic centering, 29 cases were manually rotated about the y-axis to correct the alignment of the urethra in sagittal images (caused by rotation of the pelvis in lithotomy position) prior to deformable registration.

The procedure time is shown in Figure 2.A. The mean duration was 43.1 minutes (range: 21.4-88.7); when discriminating by focal and IB, the duration was 38.6 minutes (range: 21.4-83.1) and 56.0 minutes (range: 37.5-88.7), respectively. A learning curve was not demonstrated.

In total, 387 catheters were implanted (Figure 2.B). The mean number of catheters per case was 9 (range: 3-18); for IB procedures, the mean was 15 (range: 13-18), and for focal 7 (range: 3-14). In 21% of the procedures, at least one (range: 1-3) partially-visible catheter was enabled for use in treatment planning due to the added information provided by EM reconstruction, where the offset was minimal or systematic for all other visible catheter reconstructions in the implant (internal validation).

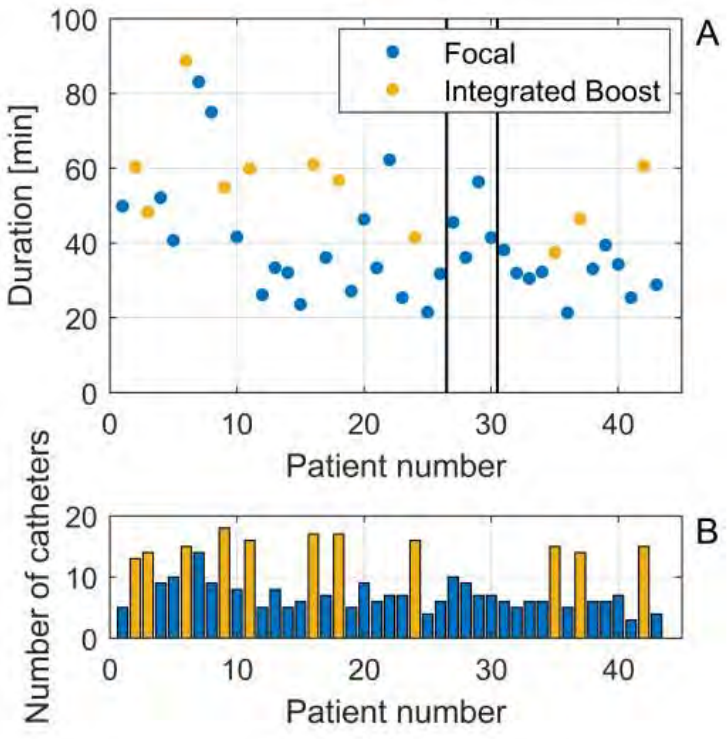


Figure 5.2 A) Procedure duration and B) number of catheters for each intervention. The duration was calculated between the first 3D-TRUS reconstruction and the last approved RT plan. Each black line indicates a 3-month pause due to COVID-19.

On the superior-inferior axis, computed and measured catheter free-lengths presented a sub-millimeter systematic offset with a mean of 0.85(0.81) mm.

In the axial view, 64% of EM catheter reconstructions required millimetric manual in-plane adjustments, based on visible catheters in TRUS (Figure 3.A). The 3D distances between corresponding points in the original and adjusted reconstruction were calculated to find each catheter's mean and maximum distance. These results can be divided into three subsets: after the first calibration of the system, after the second calibration, and after the second calibration but with the cell phone removed from the field, as shown in Table 1. The maximum applied manual adjustments are presented in Figure 3.B.

Comparing subsets I and II, no significant difference was found for mean distance (p=0.415) nor maximum distance (p=0.762). The differences between subsets II and III were significant for both mean distance (p=2.6E-7) and maximum distance (p=0.002). As shown in Figure3.B and by P95% on Table1, the distance of extreme cases decreased from I to II, and even more for III.

The calculated metrics for elastic registration accuracy are presented in Figure 4, with the

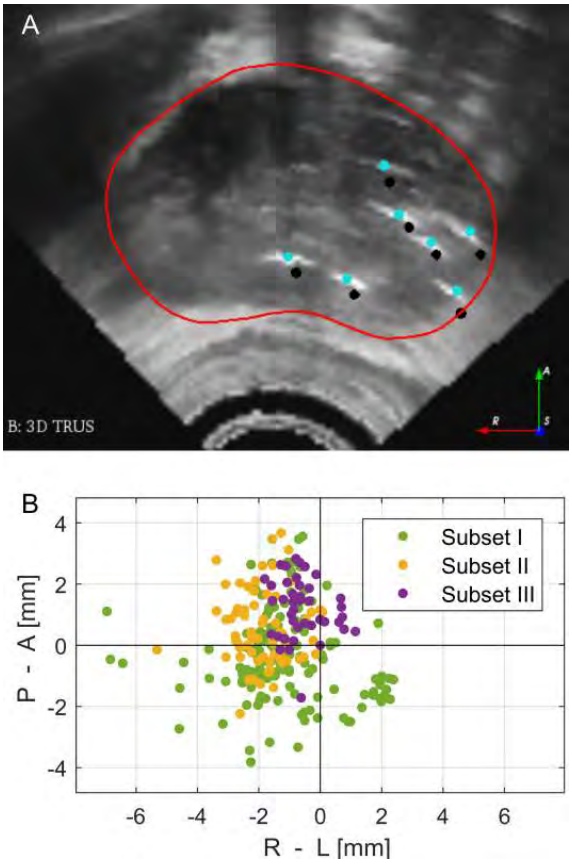


Figure 5.3 A) Visualization of original (black) and adjusted (cyan) catheter reconstructions, and prostate segmentation (red) for a sample case, in axial view. B) Maximum applied manual adjustments per catheter based on 3D distance between corresponding points. The three subsets correspond to I) calibration 1 with external interference (256 catheters), II) calibration 2 with external interference (90 catheters), and III) calibration 2 without external interference (41 catheters).

corresponding p-value.

Six cases were excluded (0 landmarks identified) from the TRE analysis. Nine cases had 1 landmark, 15 cases had 2 landmarks, 8 cases had 3 landmarks, 8 cases had 4 landmarks, and only 1 case had 5 landmarks. Of the 37 evaluated cases (Figure 4.A), the mTRE was 3.98(1.99) mm after rigid and before elastic registration. After applying the elastic registration, the mTRE was 3.41(1.48) mm.

To compare urethral midline distance (Figure 4.B), 2 cases were excluded due to poor visibility of the urethra on MRI, which did not allow reliable segmentation. The mean distance before and after deformation was 3.26(1.99) mm and 2.89(1.24) mm, respectively.

Prostate contours agreement was observed after deformation as expected given the use of a

Table 5.1 3D distances between corresponding points in the original and adjusted catheter reconstructions in the prostate region. "P95%" represents the 95 percentile of the data, and "-" denotes the cases where the radiation oncologist was not carrying a cellphone.

Subset	**External interference**	**Calibration**	**No. Catheter**	**Mean Distance [mm]**		**Maximum Distance [mm]**	
				Mean (SD)	P95%	Mean (SD)	P95%
I	+	1	256	1.95 (1.22)	3.90	2.29 (1.40)	4.37
II	+	2	90	2.06 (1.09)	3.29	2.24 (1.23)	3.79
III	-	2	41	1.26 (0.65)	2.22	1.69 (0.71)	2.87
All cases			387	1.87 (1.17)	3.41	2.18 (1.33)	3.82

surface-based elastic registration algorithm (Figure 4.C-D). Mean DSC was 0.82(0.08) and 0.94(0.02) after rigid and after elastic registration, respectively. The mean 95%-Hausdorff distance was 5.08(1.96) mm before deformation and 1.92(0.73) mm after deformation.

Figure 5 illustrates an example of the three initialization points. Since the proposed method for automatic alignment is based on urethral segmentations, 2 cases were excluded from this analysis (all metrics) due to the lack of urethra segmentation on MRI. Four more cases were excluded from mTRE assessment (0 landmarks identified). Figure 6 presents the results for the three initializations and the p-values comparing the autoaligned method to the others.

The mTRE after deformation was 4.39(2.03) mm for the autoaligned initialization and 4.57(2.13) mm for centered volumes (Figure 6.A). For the manual initialization point (the one used during the intervention), the mTRE was 3.42(1.49) mm, but the difference was not significant ($p>0.01$).

The mean distance between urethral midlines of 2.16(0.89) mm for the autoaligned initialization (Figure 6.B) was statistically smaller than both centered at 3.25(1.16) mm and manual at 2.89(1.24) mm.

Finally, the mean 95%-Hausdorff distance was 1.73(0.55) mm for centered, 2.01(0.81) mm for autoaligned, and 1.93(0.75) mm for manual initialization; mean DSC was 0.94(0.02) for the 3 initializations. No significant difference was found for both metrics (Figure 6.C-D).

5.4 Discussion

In this study, we evaluated the clinical performance of a prototype system for HDR brachytherapy with regards to workflow efficiencies, EM catheter reconstruction accuracy, and multi-

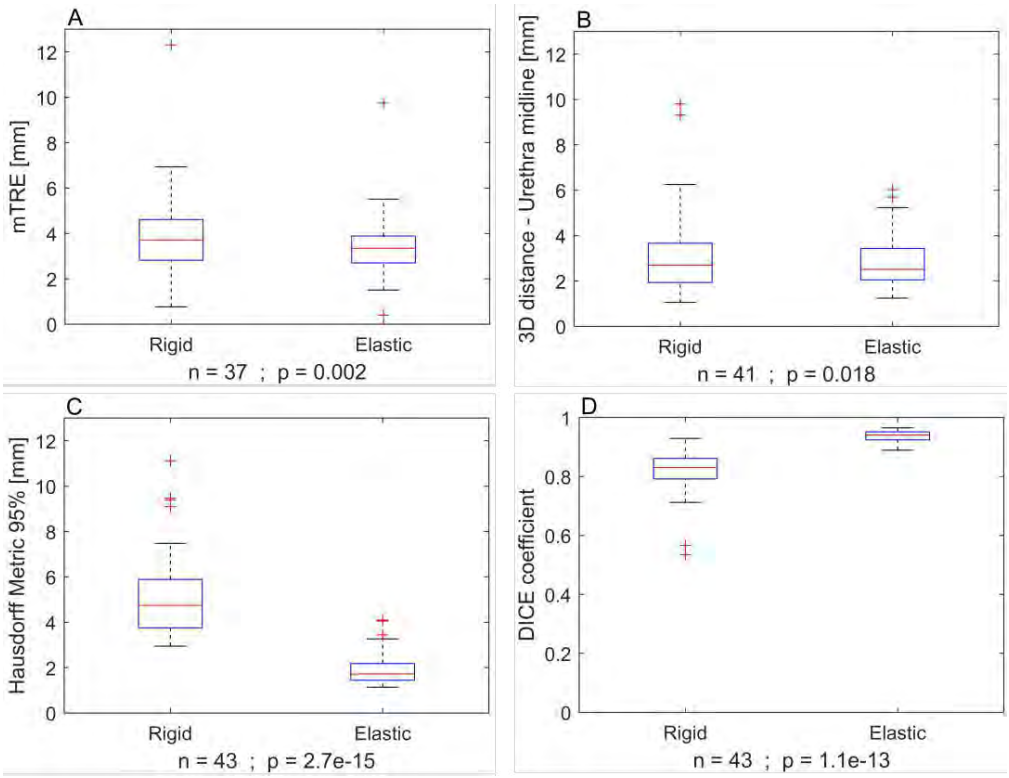


Figure 5.4 Performance comparison for rigid and elastic registration methods. A) mTRE, B) Urethral midline 3D distance, C) 95%-Hausdorff, and D) DSC.

modality image registration accuracy.

Average procedure time for conventional (no EM tracked) TRUS-guided prostate HDR brachytherapy is reported from 85 to 120 minutes, including 15 to 20 minutes of set-up [6, 38, 113]. The procedure workflow for the prototype system compared favorably with current standard-of-care practice and was integrated with little, if any, learning curve as observed in our timeline of procedural efficiencies.

In traditional non-tracking systems, catheter reconstruction is based on the image signature and measures of catheter free-length. This can be challenging due to inherent artifacts of images, especially in the context of implanted devices [114]. In focal salvage treatments, the high number of catheters in a small volume (presence of shadowing or strong reflection) can exacerbate this problem, such that some catheters only be partially visible on TRUS [36, 114, 115]. EM tracking can increase confidence in catheter reconstruction (especially in depth). The residual error in free-length is less than the tolerance for non-tracking systems with high-quality images (1mm) [116]. As demonstrated in 9 procedures, EM tracking also increases confidence in use of partially visible catheters in an otherwise accurately

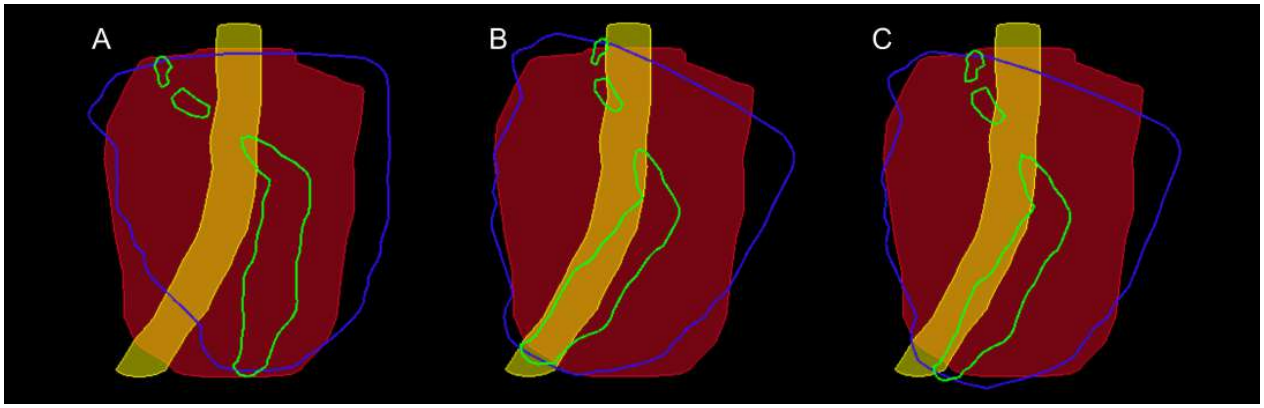


Figure 5.5 Representation of A) centered, B) autoaligned, and C) manual initialization points for a sample case. Filled slices are from TRUS segmentation and contours from MRI segmentation: prostate (red and blue) and urethra (yellow and green).

reconstructed implant. However based on our results, the use of EM tracking alone without global image validation is not recommended.

Flexible catheters can be bent by anterior pressure from the TRUS-probe. Given that the EM reconstruction is made using a rigid stylus, we observe a systematic anterior-posterior displacement, more pronounced at the distal tip [61, 114]. Seventy-two % of manual adjustments were smaller than 2mm, which is the reported tolerance for non-tracking systems [116], and 95% were smaller than 3mm, which is a displacement reported to impact dose plans [117]. Nonetheless, we recognized that such errors could be clinically significant thereby motivating calibration 2. Calibration and control of external interference sources (cellphone) significantly improved the accuracy of catheter reconstruction, with remaining displacements due to catheter flexibility. Image-based corrections of EM reconstructions remain essential in this workflow. Other possible causes of interference (even distant to the EM field, e.g., US system) should be studied to identify possible causes of the remaining systematic right shift.

One of the challenges of multimodal image registration for prostate applications is the deformation of structures due to rotation of the pelvis, the presence of an endorectal probe, the placement of a foley catheter, and the insertion of brachytherapy catheters; hence the need to investigate elastic registration and to evaluate its performance in real conditions [35, 53].

Surface-based elastic registration displayed slightly better results (mTRE) than rigid registration, but uncertainties remain. The mTRE value after deformation is within the reported range (2.1-3.7 mm) for similar systems [5, 43, 56]. Although mTRE is one of the most widely used metrics for assessing image registration performance [35], 1 or 2 landmarks in 56% of the cases make these results very localized and dependent on the specific location of the

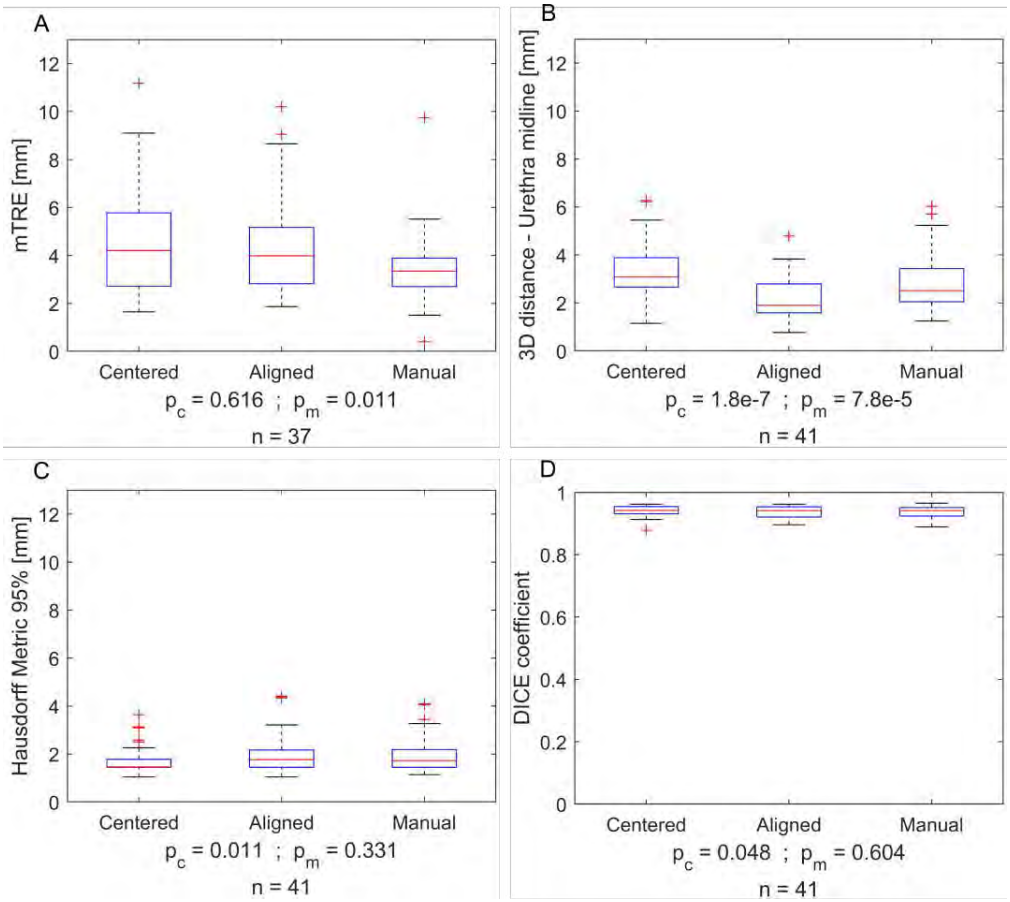


Figure 5.6 Evaluation of the elastic registration method for three different initialization points; pc was used to compare autoaligned-centered start points and pm autoaligned-manual. A) mTRE, B) Urethral midline 3D distance, C) 95%-Hausdorff, and D) DSC.

landmark (e.g., distance to the surface).

The urethra is an important reference structure in prostate interventions because of its visibility in different modalities. Besides, it is crucial to consider the urethra location during HDR brachytherapy to minimize radiation damage [36, 43]. The distance between urethral midlines after deformation was reduced by 0.36 mm (not statistically significant), but the achieved distance is within the accuracy range reported for the system (2-3 mm) [56]. When using point-based metrics, the verification validity for the overall image registration improves by increasing the number of points [118]. Urethra visibility allowed the use of more points (range: 16-25) distributed along the gland, it provides a clearer depiction of the algorithm's performance in the prostate, within the central region. Since the urethra is easier to identify than landmarks, usage of more reference points reduced the impact of the error added by manual segmentation [119].

Cognitive or rigid computational registration, typically performed when physicians wish to use MRI information during the brachytherapy procedure, are known to carry errors. Although we demonstrate that important uncertainties remain with elastic registration, we were able to demonstrate that this strategy improves accuracy without compromising workflow performance [57, 120].

It is well established that MRI and TRUS images are acquired at different pelvic rotations around the right-left axis. Such rotations can be large and lead to errors in registration if not corrected a priori; surface-based registration algorithms do not normally include automatic rotation in the basic workflow (additional segmentations or fiducial selection would be needed) but usually include tools to do it manually. We found that both an automated and a manual approach to align the urethral angle prior to applying deformable registration improved registration accuracy (33% and 11%, respectively) as measured by the urethral midline distance. As expected for a surface-based algorithm, Hausdorff and DSC did not show significant differences for the three initializations since the similarity of the contours is optimized during the elastic registration [35], showing that the initialization does not affect the performance of the algorithm on the surface. Finally, regarding mTRE, no significant results were obtained when comparing the manual initialization and our automated method at the specific sites of the landmarks.

In future work, we aim to further advance the technique to minimize requirements for manual catheter reconstruction adjustments (e.g., flexible EM tracked stylus), and to refine software tools for such manual adjustments. Furthermore, the integration of autosegmentation on TRUS (prostate, urethra, catheters) could greatly improve time efficiencies. Finally, attempts to better approach pre-planning geometry at set-up may further improve registration performance. Each step will contribute to our vision of a highly streamlined procedure that can be performed under 30 minutes with minimal anesthesia requirements, and with high accuracy and confidence. This vision can only be achieved with the full integration of a suite of technical innovations.

5.5 Conclusion

We found that workflow efficiency and technical performance of a prototype navigation system supporting tumor-targeted HDR prostate brachytherapy aligned with our current clinical practice and was not disruptive of our clinical throughput despite its added underlying complexity. EM tracking is accurate enough to complement catheter reconstruction combined with TRUS, and occasionally enabled the use of catheters that were clinically well-positioned within an accurately reconstructed implant, but that were partially visible in TRUS. Al-

though gains were demonstrated with the system's deformable image registration tools compared with rigid alignment, uncertainties remain and must be considered when defining the PTV. Finally, an automated tool to align the urethral angle before applying deformable registration, as proposed in this evaluation, is justified.

5.6 Funding, sources of support and acknowledgements

- Funded in part with in kind support from Philips InVivo, and Grant awards from Varian Medical systems and NSERC/CIHR.

- This research was undertaken thanks, in part, to funding from the Canada First Research Excellence Fund through the Trans- MedTech Institute.

5.7 Conflict of interest statement

None.

CHAPTER 6 ARTICLE 2: IMAGE-GUIDED RAMAN SPECTROSCOPY NAVIGATION SYSTEM TO IMPROVE TRANSPERINEAL PROSTATE CANCER DETECTION. PART 2: IN-VIVO TUMOR-TARGETING USING A CLASSIFICATION MODEL COMBINING SPECTRAL AND MRI-RADIOMICS FEATURES

This article was published in the *Journal of Biomedical Optics (JBO)*, September 2022.
Remarks: The first author contributed to the experimental measurements, data processing, bibliographic research, and manuscript writing, representing approximately 90% of the work.

Image-guided Raman spectroscopy navigation system to improve transperineal prostate cancer detection. Part 2: In-vivo tumor-targeting using a classification model combining spectral and MRI-radiomics features

[1,2]David Grajales, [1,2]Fabien Picot, [1,2]Roozbeh Shams, [1,2]Frédérick Dallaire, [1,2]Guillaume Sheehy, [1,2]Stephanie Alley, [2]Maroie Barkati, [2]Guila Delouya, [2]Jean-Francois Carrier, [2]Mirela Birlea, [2]Dominique Trudel, [1,2,3]Frédéric Leblond, [2]Cynthia Ménard, [1,2]Samuel Kadoury

[1]Polytechnique Montréal, Montréal, Québec, Canada
[2]Centre de recherche du Centre Hospitalier de l'Université de Montréal, Montréal, Québec, Canada
[3]Institut du Cancer de Montréal, Montréal, Québec, Canada

Abstract

Significance: The. prostate cancer (PCa) diagnosis and treatment of PCa prostate cancer (PCa) are limited by a lack of intraoperative information to accurately target tumors with needles for biopsy and brachytherapy. Novel image-guidance techniques using optical devices could improve the diagnostic yield of biopsy and efficacy of radiotherapy. **Aim:** To evaluate the performance of multimodal PCa detection using biomolecular features from in situ Raman spectroscopy (RS) combined with image-based (radiomics) features from multiparametric magnetic resonance images (mpMRI). Approach: In a prospective pilot clinical study, 18 patients were recruited and underwent high-dose-rate (HDR) brachytherapy. Multimodality image fusion (preoperative mpMRI with intraoperative transrectal ultrasound) combined with electromagnetic tracking was used to navigate an RS needle in the prostate prior to brachytherapy. This resulting dataset consisted of Raman spectra and co-located radiomics

features from mpMRI. Feature selection was performed with the constraint that no more than 10 features were retained overall from a combination of inelastic scattering spectra and radiomics. These features were used to train support vector machine (SVM) classifiers for PCa detection based on leave-one-patient-out cross-validation (LOPOCV). **Results:** RS along with biopsy samples were acquired from 47 sites along the insertion trajectory of the fiber-optics needle: 26 were confirmed as benign or grade group=1, and 21 as grade group>1, according to histopathological reports. The combination of the fingerprint region of the RS and radiomics showed an accuracy of 83% (sensitivity=81% and a specificity= 85%), outperforming by more than 9% models trained with either spectroscopic or mpMRI data alone. An optimal number of features was identified between 6 and 8 features, which have good potential for discriminating grade group$\geq$ 1 / grade group<1 (accuracy=87%) or grade group>1 / grade group$\leq$ 1 (accuracy=91%). **Conclusions:** In situ Raman spectroscopy combined with mpMRI radiomics features can lead to highly accurate PCa detection for improved in-vivo targeting of biopsy sample collection and radiotherapy seed placement.

Keywords: Raman spectroscopy, prostate cancer, tissue optics, multimodal imaging, machine learning, support vector machines, magnetic resonance imaging, ultrasound imaging.

6.1 Introduction

Transperineal biopsy and high-dose-rate (HDR) brachytherapy are two needle-based procedures for diagnosing and treating prostate cancer (PCa), respectively. HDR brachytherapy allows the delivery of a considerable dose of radiation to the tumor site using temporary implants while reducing the surrounding tissue involvement [36, 120]. Although using transrectal ultrasound (TRUS)-guided biopsy is the standard of care, it can have up to 30% false-negative rates and does not allow in situ characterization [22, 24]. It is clear that the efficiency of both techniques relies significantly on accurate localization of the tumor and needle.

Image-guidance is one of the strategies to support these localized procedures, using TRUS, which provides real-time anatomical information of the prostate and neighbor structures. It therefore provides navigation support for biopsies or brachytherapy catheter implantation, but since not all lesions are hypoechoic, it does not provide information on tumor location [18, 121]. On the other hand, magnetic resonance imaging (MRI), more specifically multiparametric MRI (mpMRI) sequences based on diffusion, offer higher sensitivity, allowing the visualization of certain lesions. Currently, mpMRI is used by physicians to report tumors (PIRADSv2.1 [18,30]) and plan interventions, but it also allows the extraction of quantitative features (radiomics), which could be used as biomarkers [44,120,122]. However, MRI presents

certain limitations to guide interventions due to longer acquisition time, limited compatibility with surgical instruments, and cost [35, 36]. Multimodal image registration is currently studied and used to take advantage of complementary information of TRUS and MRI to assist HDR brachytherapy and other tumor-targeted prostate interventions [5, 35, 42, 123].

This approach, combined with electromagnetic (EM) tracking, helps provide navigation capabilities for targeting tumors [60, 61, 124]. However, neither of these modalities allows in situ, real-time tissue characterization, which could significantly impact the diagnosis and treatment efficacy, reducing false-negative rates and boosting personalized treatments for more than 1.4 million PCa new cases diagnosed every year worldwide [72, 90, 109, 111].

Raman spectroscopy (RS), on the other hand, characterizes microscopic information of prostate tissue, providing real-time molecular signatures and taking advantage of the tissue's highly sensitive and specific optical properties [72, 94, 125–127]. Based on inelastic light scattering, RS has been used for years for ex vivo sample characterization, producing spectra with molecular vibrational states information, showing great potential for detecting several diseases [51, 67–69]. Furthermore, with the development of optical fiber RS probes, this technique is moving to clinical applications [9]; different optical probe designs have been used for in vivo tissue characterization (in human and animal models) for targeting skin cancer in open surgeries [70], minimally invasive diagnosis of lung cancers [76], bladder cancer detection using a superficial and nonsuperficial Raman probes [77], observation of skin changes after breast cancer treatment [78], and others [79–81, 85]. In prostate applications, it has been used for ex vivo characterization and in vivo margin detection [11, 23, 90, 109, 126], but, to the best of our knowledge, so far, not for real-time in vivo prostate tumor burden confirmation, which can provide great benefit for clinical procedures.

As previously described, mpMRI has remarkable tumor-related information on a larger scale, especially for diffusion-weighted sequences, where radiomics can extract this information quantitatively [44, 48, 123]. There is a wide variety of standardized radiomics features primarily classified as intensity-, texture-, or shape-based. Shape-based are especially useful when lesions are segmented; the other two classes have been studied, identifying potential on some first order and some gray-level-correlation-matrix (GLCM) features [45, 47].

Multimodal and multiscale characterization is advantageous for tissue characterization purposes given the complementary information it provides, as single modalities may not capture all critical elements of the interrogated sample [11, 127–129].

This pilot clinical study aims to evaluate the feasibility of a multimodal and multiscale characterization approach for in vivo PCa classification during clinical procedures. We combined real-time mesoscopic characterization provided by RS and macroscopic characterization from

preoperative mpMRI, co-localized with multimodal image registration and EM tracking, as input for a support vector machine (SVM), for assessing the classification potential of such characterization. This paper is the second part of a joint work carried out in the framework of the same pilot clinical study; part 1 [96] presents details of the optical system, results of ex vivo experiments, and their comparison with in vivo results.

6.2 Materials and methods

6.2.1 Clinical Data and Equipment

This pilot clinical study was conducted between September 2020 and August 2021, with 18 patients with histological diagnosis of PCa, enrolled on a prospective clinical trial approved by the Research Ethics Board (NCT03378856).

Planning mpMRI (3D T2-weighted FSE, b2000 DWI, +/- DCE) were obtained on a 1.5T Siemens Aera Magnetom (Siemens Healthineers, Erlangen, Germany) using surface coils. Voxels on acquired T2 images were 1x1x1 mm3, b2000 images (diffusion-weighted images b-value of 2000 s/mm2) consisted of 2.6x2.6x5 mm3 voxels, and voxels on calculated ADC maps were 1.8x1.8x4 mm3. PSMA-PET/CT (18F-DFCPyL [94]) images were also acquired in a subset. For intraoperative imaging, a bk3000 ultrasound system was used with a BK endocavity biplane transducer (BK Ultrasound, Herlev, Denmark). Brachytherapy procedures were assisted by a prototype interventional system (Invivo/UroNav, Philips Disease Management Solutions, Gainesville, USA) in the early phase of clinical deployment.

We used a custom system consisting of navigation and optical components (Fig. 1). The optical component contains a dual source (671nm and 785nm, Semrock, New Tork, USA), a spectrometer (EmVision LLC, Fl, USA), and the custom EM tracked optical probe (EmVision LLC, Fl, USA), designed to perform in situ, minimally invasive characterization [96]. The use of this subsystem, controlled by customized MATLAB R2017b (Mathworks, Ma, USA) software, allows to stimulate the tissue and detect energetical shifts due to inelastic scattering from light-tissue interactions, which is correlated to molecular vibration modes [90].

The navigation component is connected to the ultrasound system and uses MRI-TRUS fusion to project structures segmented on MRI over the real-time TRUS images. It is also connected to the EM tracking platform (Aurora NDI, Waterloo, Canada) consisting of a control unit, an EM field generator placed over the patient's pelvic region, and three six-degrees-of-freedom EM sensors: the first one placed on the template as a fix reference, the second fixed to the TRUS-probe holder to track the field of view of the current 2D image, and the third one, small enough to be integrated to the custom optical probe lumen, to reconstruct the probe

while navigating to pre-identified targets. A 3D Slicer module was created for visualization and control of the subsystems [112]. More details on the system could be found in previous works [109, 130].

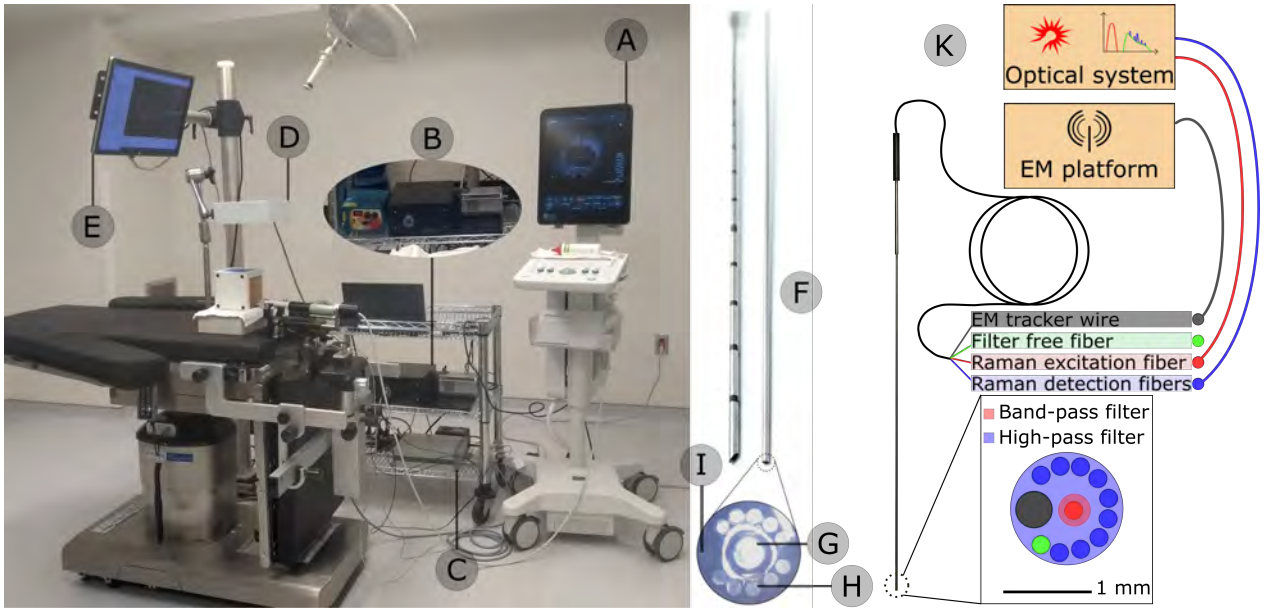


Figure 6.1 Clinical setup: A) ultrasound system, B) near-infrared laser and spectrometer, C) EM tracking system, D) EM field generator, E) 3D Slicer navigation system, F) closeup of the RS probe fiber bundle next to cannula, G) Raman excitation fiber, H) Raman detection fibers, I) EM sensor, and K) schematic illustration of optical setup.

6.2.2 Workflow and Data Acquisition

HDR brachytherapy procedures include the following steps: i) importing the preoperative images and predetermined contours into the interventional system, ii) performing a 3D reconstruction from 2D TRUS image and segmenting the prostate, iii) registering images and propagating the MRI contours to TRUS, iv) inserting and reconstructing the catheters, and calculating the dose plan, and finally v) delivering the dose. All these steps were performed with the patient under general anesthesia and in lithotomy position. Details of the intervention (brachytherapy workflow), image registration, and EM tracking can be found in [131].

The optical acquisitions were performed during HDR brachytherapy procedures after the elastic image registration (defined below); once the optical measurements were completed, the brachytherapy procedure resumed. The general diagram of the connections and the flow diagram can be seen in Fig. 2.

Gross tumor volumes (GTV) were segmented on T2 MRI, based on mpMRI and PSMA-PET

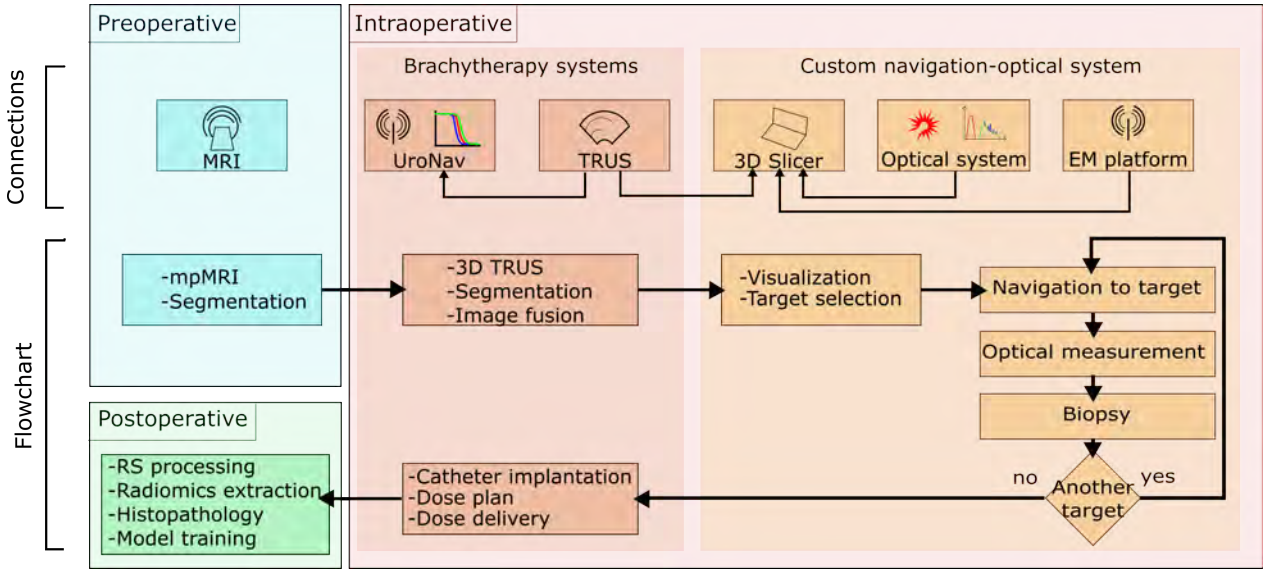


Figure 6.2 Flowchart and connections diagram of the systems involved in the intervention. The larger boxes divide the pre-, intra-, and post-operative phases of the procedure. The left column in the intraoperative section contains the equipment and activities inherent to brachytherapy. The right column contains the steps and systems added to the interventions to carry out the present study.

images; prostate and urethra were segmented in preoperative images (using Eclipse, Varian Medical Systems, Palo Alto, USA) and intraoperative images (using the interventional system). Additionally, T2 images and contours were elastically registered to the intraoperative TRUS, using a surface-based algorithm integrated into the prototype interventional system; for this process, the prostate contours (from MRI and TRUS) were automatically centered, then manually aligned following the urethra angle, and then automatically deformed looking for an optimal surface correspondence. TRUS images and propagated contours were exported to our custom navigation system, enabling visualization of deformed structures (e.g., GTV) projected over real-time TRUS.

Targets (e.g., GTV center of mass, healthy tissue far from the GTV) were pre-identified based on the preoperative MRI, and the optical probe was navigated to them supported by a coaxial needle (cannula). Once at the site of interest, the dual optical source stimulated the tissue, and the spectrometer captured the response signal (from 50 to 100 RS spectra per site). The coordinates of the inspected sites in the TRUS reference system were recorded, and a confirmation biopsy was taken at the same location.

Biopsy cores were fixed and processed according to standard histopathologic procedures observed by an expert to identify patterns on the stained sample slides [25]. According to the

predominant and secondary patterns, a report was generated presenting the Gleason score (GS), the grade group according to the International Society of Urological Pathology (ISUP GG), and the high-grade tumor percentage (HG) for each biopsy core [26]. For instance, a slide presenting tumor tissue, consisting of 80% pattern 4 and 20% pattern 3, will be reported as GS: 4+3=7, ISUP GG: 3/5, and HG:80% [132]. Due to the impact over treatment planning, classification algorithms for PCa are usually trained for detecting tumors with ISUP GG>1, so initially, we set our "ISUP GG>1" prediction task by labeling the observations with ISUP GG$\leq$ 1 (including ISUP GG=1 and benign tissue) as false, and sites with ISUP GG>1 as true [36, 47, 66].

The RS signal processing consisted of averaging all the spectra from one site, removing autofluorescence and cosmic rays, standard normal variate normalization, and finally assigning each pixel of the spectrometer to a Raman shift (Fig. 3) [96]. We obtained a single spectrum per site at the end of this process. This was applied for the fingerprint (FP) and the high wavenumber (HW) region of the RS, according to the source used to stimulate the tissue. For this study, every single Raman shift on the spectrum was a feature for the next steps [130].

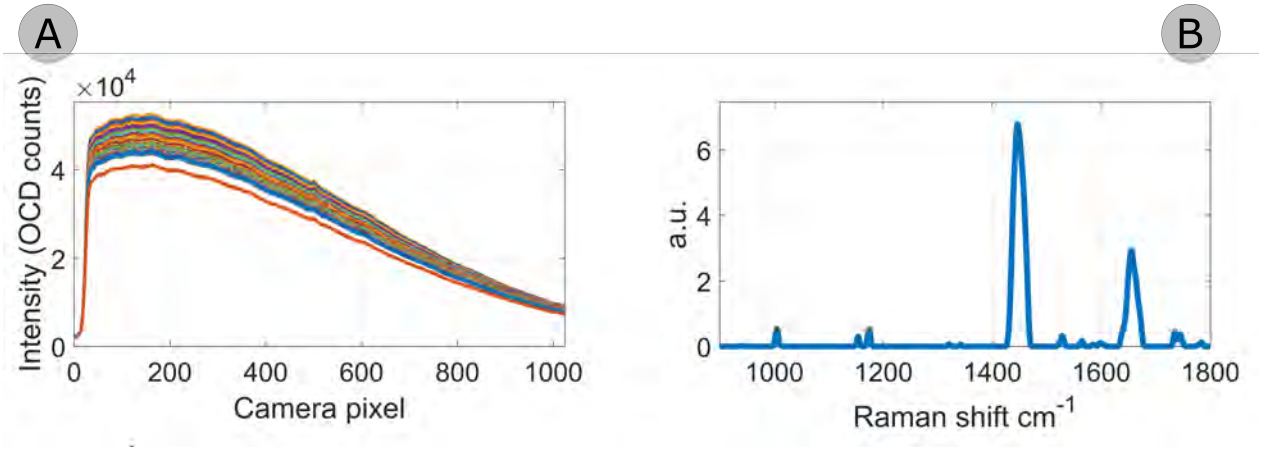


Figure 6.3 Sample signal processing. (A) 50 raw fingerprint acquisitions from one inspected site and (B) the resulting processed Raman spectrum.

To extract MRI-based features (Rad), the PyRadiomics platform was used [45], which offers up to 120 features from different categories; based on literature, we selected 8 first order and 8 GLCM features (Table 1) to be calculated on T2, ADC, and b2000 mpMRI [44, 47, 48, 66]. Using the transformation matrix employed to register MRI and TRUS, we applied the inverse process to find the corresponding coordinates on the mpMRI for each inspected site. We calculated the radiomics features in a 5 mm radius spherical volume at each position.

Table 6.1 Radiomics features extracted from the three mpMRI sequences, and their identifiers.

Group	**Name**	**Identifier**		
		T2	**ADC**	**b2000**
First order	Energy	r1	r17	r33
	TotalEnergy	r2	r18	r34
	Entropy	r3	r19	r35
	Mean	r4	r20	r36
	Median	r5	r21	r37
	StandardDeviation	r6	r22	r38
	MeanAbsoluteDeviation	r7	r23	r39
	Uniformity	r8	r24	r40
GLCM	Autocorrelation	r9	r25	r41
	ClusterShade	r10	r26	r42
	Contrast	r11	r27	r43
	Correlation	r12	r28	r44
	Id (Inverse Difference)	r13	r29	r45
	DifferenceEntropy	r14	r30	r46
	JointEntropy	r15	r31	r47
	JointEnergy	r16	r32	r48

6.2.3 Classification Model Training

Given the low number of patients, which is inherent to a pilot clinical study, we decided to limit the complexity of the classification model as a strategy to avoid overfitting issues. Thus, we trained SVM models for binary classification, with a linear kernel and a cost matrix that doubles the penalty to false negatives (C = [0,1;2,0]), using MATLAB R2017b. For training and validation, we followed a leave-one-patient-out cross-validation (LOPOCV) scheme (i.e., models were trained with data from all patients but one, which was used for validation).

In a high-dimensional dataset, especially with a limited number of observations and patients, feature selection is crucial. To perform this step, we set a maximum number of features to be selected (max_nf) and applied a 3-step selection approach:

i. Amount of variation: The principle is to discard the features that have almost the same value for all the observations. We calculated the variance for each feature and only retained the features with var(X)>0.03 for the next selection step.

ii. Correlation with the target: We calculated the correlation coefficients between each feature and the assigned label (ground truth), then discarded the features with an absolute value of correlation $< 10\%$.

iii. Lasso regression: This method assigns a weight to each feature within an optimization function and gives a non-zero weight only to features that contribute significantly to establishing a decision boundary, which means that the number of selected features could be equal to or less than the maximum number of non-zero coefficients we set (max_nf).

This 3-step process was applied after dividing training/validation sets, i.e., the observations of the patient left out were not considered during the feature selection. As a result, the number of features selected for each fold may vary slightly, so the value reported in the results is the average of all folds.

We describe below three sets of experiments focused on the classification potential of collected features (RS and Rad) to train the predictive model.

Feature combination experiment

We trained different models using independent sets of features (FP, HW, or Rad), as well as the different combinations between them (FP + HW, FP + Rad, FP + HW + Rad, HW + Rad). For this experiment, we arbitrarily set a max_nf = 10, trying to limit the number of features, close to 50% of the number of patients. Using the posterior probability for each measurement used for validation, we plotted the receiver operating characteristic (ROC) curve for each model to describe the general performance comparing the area under the curve (AUC). We then calculated the accuracy, sensitivity, and specificity to evaluate the performance at the optimal threshold.

Number of features experiment

Using the combination of features demonstrating the best performance in the previous experiment (Sec. 2.3.1), we trained a model increasing the max_nf value for the selection process (from 2 to 18) to find the optimal number of features for the classification task. Then we identified the features that, among the 18 iterations of the LOPOCV, were selected more than 10 times (most frequently selected features).

Prediction task experiment

The experiments presented in Sec. 2.3.1 and 2.3.2 were focused on "ISUP GG>1" prediction using the labels we described in Sec. 2.2. For this experiment, we applied different criteria to the histopathological results (the limit for class selection) to assign labels for two other prediction tasks. For the first one, we labeled the observations with ISUP GG$\geq$ 1 as true and benign tissue as false ("ISUP GG$\geq$ 1" prediction). For the second one, "high grade"

prediction, we used the same labels from the previous one, but excluded the observations with an HG<20%. We used the most frequently selected features identified in the previous experiment (Sec. 2.3.2) to train the models, with no other feature selection.

6.3 Results

6.3.1 Clinical data

In total, the dataset consisted of 47 inspected sites, with the corresponding histopathological report for ground truth, and 4650 features (Table 2). Sample images of two histopathological slides are presented in Fig. 4.

Elastic registration allowed the projection of the deformed GTV contours on real-time TRUS, which was used to set up to 5 sites of interest and for navigation support during the intervention; the inverse process allowed the co-location of the inspected sites on the mpMRI, post-intervention (Fig. 4.A-B). According to the voxel size of mpMRI and the volume defined for radiomics extraction (blue sphere on Fig. 4), ranges of 428-482, 34-38, and 24-32 voxels were used from T2, ADC, and b2000, respectively.

Table 6.2 Clinical data.

Number of patients	18
Median age [years]	68 (range:60-74)
Median inspected sites/patient	2 (range:2-5)
Total inspected sites	47
Benign	23
GS: 3+3=6 / ISUP GG: 1/5	3 (3/3)[a]
GS: 3+4=7 / ISUP GG: 2/5	10 (3/10)[a]
GS: 4+3=7 / ISUP GG: 3/5	8
GS: 4+4=8 / ISUP GG: 4/5	3
Total number of features	4650
FP (RS)	1801
HW (RS)	2801
Rad	48
Mean added time [min]	18.4 (SD=4.9)

[a] Number of samples with HG<20%.

6.3.2 Feature combination experiment

Following a dichotomization of the collected data during the brachytherapy procedures ("ISUP GG>1" criteria), these experiments consisted of 21 ISUP GG>1 and 26 ISUP GG$\leq$ 1 sam-

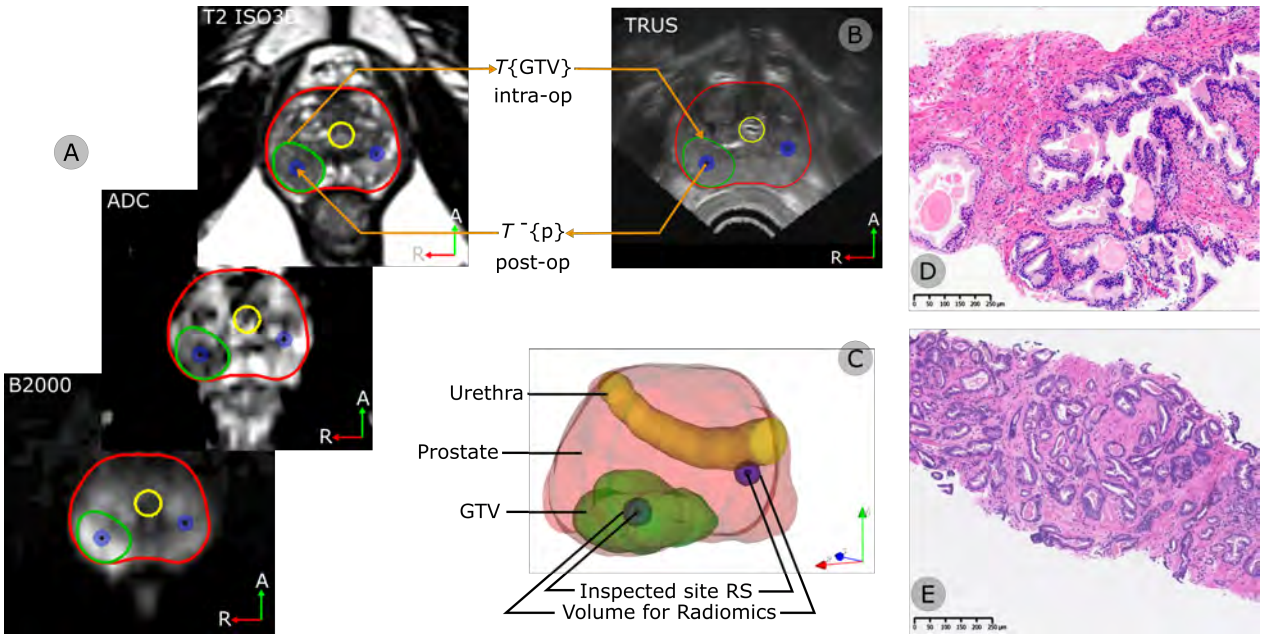


Figure 6.4 Sample image registration and co-location results. (A) GTV originally segmented on preoperative mpMRI is (B) projected over the real-time TRUS using a surface-based elastic registration algorithm; the inverse process allows identifying the MRI coordinates corresponding to the inspected sites. (C) A 3D model of the plan can be rendered. (D) Pathological image samples of benign prostatic parenchyma and (E) acinar adenocarcinoma GS: 3+4=7, ISUP GG=2, and HG: 40-50%.

ples. The AUC results for the different combinations of features are shown in Table 3, and the ROC curves for the groups of features involved in the model with the best performance are presented in Fig. 5.

The FP + Rad model showed the best overall performance, based on AUC, followed closely by three other combinations: FP + HW, Rad, and FP + HW + Rad. From these four models, FP + Rad and FP + HW + Rad yielded a low mean support vectors ratio (number of support vectors needed to train the model, divided by the total number of observations) compared to the other two models.

Prediction accuracy, sensitivity, and specificity were calculated at the optimal point on the ROC curve, which is the closest point to the (0,1) corner of the plot. These metrics also showed the superior performance of FP + Rad (Table 3). Regarding the prediction accuracy, FP + Rad outperformed all other models but one for more than 9%; FP + HW + Rad, that also combined optical and image-based features, was the only model close to this accuracy (only 2% below), but also 10% below in sensitivity.

Table 6.3 Classification performance of SVM models (LOPOCV) for discriminating ISUP GG>1 and ISUP GG≤ 1. SV rate: number of support vectors/total of observations.

Features	**Number Features**	**AUC**	**Prediction Accuracy**	**Sensitivity**	**Specificity**	**SV Rate**
FP	8.3	0.74	0.72	0.67	0.77	0.64
HW	8.3	0.55	0.53	0.48	0.58	0.76
Rad	8.4	0.79	0.74	0.76	0.73	0.70
FP + HW	8.2	0.80	0.70	0.78	0.64	0.62
FP + Rad	8.3	0.82	0.83	0.81	0.85	0.55
FP + HW + Rad	8.2	0.79	0.81	0.71	0.88	0.52
HW + Rad	8.6	0.58	0.62	0.67	0.58	0.64

6.3.3 Number of features experiment

Using the combination of features with the best overall performance (FP + Rad) identified in the previous experiment (Sec. 3.2), we trained models varying max_nf for the feature selection process. As explained in Sec. 2.3, the selected features for each fold (LOPOCV) could vary slightly, up to the set limit (max_nf), so the actual number of features reported is the mean number of features selected for all the folds. According to AUC results and metrics calculated at the optimal point (Fig. 6), there is an optimal region between 6.1 and 8.3 used features, resulting when setting the max_nf at 7 and 10, respectively. Limiting the number of features used to train the classification model aimed to reduce overfitting given the small datasets and improve generalization capabilities; but the remarkable drop in specificity observed in the graph for more than 9 features emphasizes the importance of the feature selection process.

The features selected on each fold for the model that used on average 8.3 features were not always the same, so Fig. 7.B-D shows how many times (number of folds) each feature was selected. FP Raman spectra were averaged by class for visualization (Fig. 7.A), and from those 1801 features (Raman shifts), 5 features were selected more than 10 times (i.e., during the feature selection applied over each fold): 994, 1007, 1334, 1766, 1772 cm^{-1}. Similarly, radiomics features were averaged by class (Fig. 7.C), and 2 features were selected more than 10 times: r17 (Energy from ADC) and r46 (DifferenceEntropy from B2000), both selected on all the folds (18 times). This total of 7 most frequently selected features could make a set of features with great potential for further studies.

It can be observed, as expected, that the valleys in the Raman spectrum were not selected, except for some features just at the base of a significant peak, which contained information on the width of the peak (e.g., 1108 or 1575 cm^{-1}). As for the radiomics features, it can

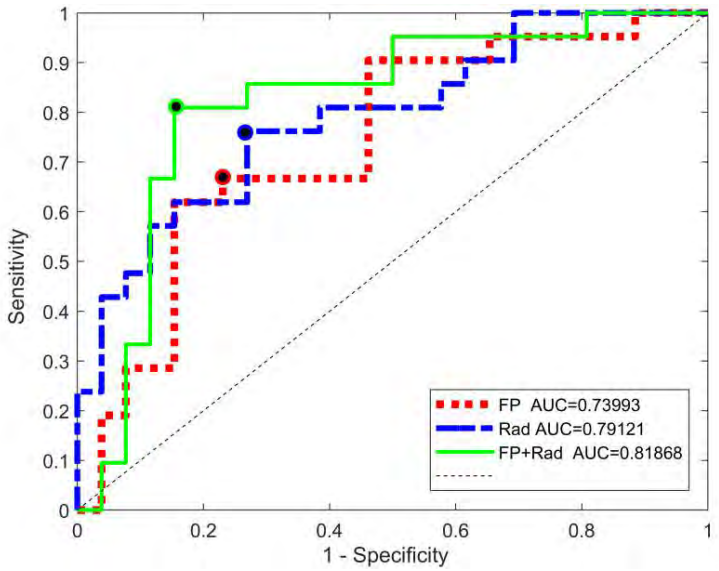


Figure 6.5 ROC curve and optimal point for discriminating ISUP GG>1 and ISUP GG$\leq$ 1, for models including feature selection (max_nf=10). AUC: Area under the curve.

be seen that, apart from the 2 most frequently selected features, only a few were selected at least once (r18, r20, r39, r44); according to the selection method, none of the T2 features, GLCM from ADC and first order from b2000 contributed to the classification task.

6.3.4 Prediction task experiment

We can observe that the different criteria to set the classes for the prediction tasks resulted in different numbers of labels and observations (Table 4). The three observations with ISUP GG=1 (Table 2) were HG<20%, so the last set of labels ("high grade") did not include these samples.

The models trained for each prediction task were compared in terms of general performance (ROC) and the performance at the optima point, and the results are presented in Fig. 8 and Table 4.

6.4 Discussion

This study evaluated the potential of using macroscopic and mesoscopic tissue characterization for PCa classification purposes, combining in vivo RS and radiomics from preoperative mpMRI, for detection of ISUP GG>1 assisted by a custom navigation system. As in different commercial devices for other applications, our navigation system allows to guide the physician during the insertion of the optical probe and track it in real-time in correspondence to

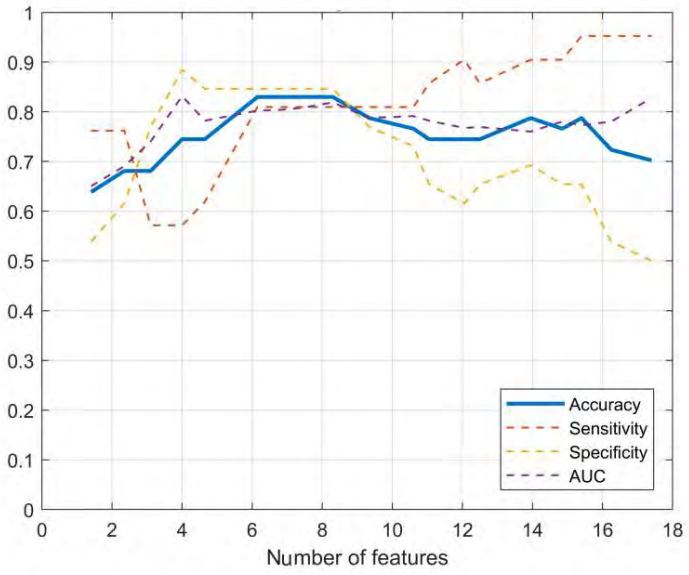


Figure 6.6 Classification performance of the FP + Rad model in function of the number of features selected.

Table 6.4 Classification performance using the most frequently selected features, based on different prediction tasks.

Prediction task	**"ISUP GG>1"**	**"ISUP GG≥ 1"**	**"High grade"**
Total Inspected sites	47	47	41
True class (+1)	21 ISUP GG>1	24 ISUP GG≥ 1	18 ISUP GG≥ 1, HG>20%
False class (-1)	26 ISUP GG≤ 1	23 ISUP GG<1	23 ISUP GG<1
Accuracy	0.91	0.87	0.88
Sensitivity	0.90	0.83	0.89
Specificity	0.92	0.91	0.87

pre-operative MRI, which contributes to a time-efficient procedure, by adding an acceptable time to the procedure.

Previous studies showed the important potential of RS for PCa detection; for instance, ex vivo studies [23,92,109] presented a classification accuracy ranging between 83% to 89%. For some of these studies, in addition to not being carried out under in vivo conditions, models were trained with at least 10 times more ex vivo data; this prepares the model to overcome data variability and widens the range of features to be used (no feature selection in some cases). These differences (ex vivo/ in vivo, size of the dataset, and number of features) may be the reason why the RS models alone are not as performant, especially using HW; however, it should be noted that FP and FP + HW have accuracy larger or equal to 70% using just

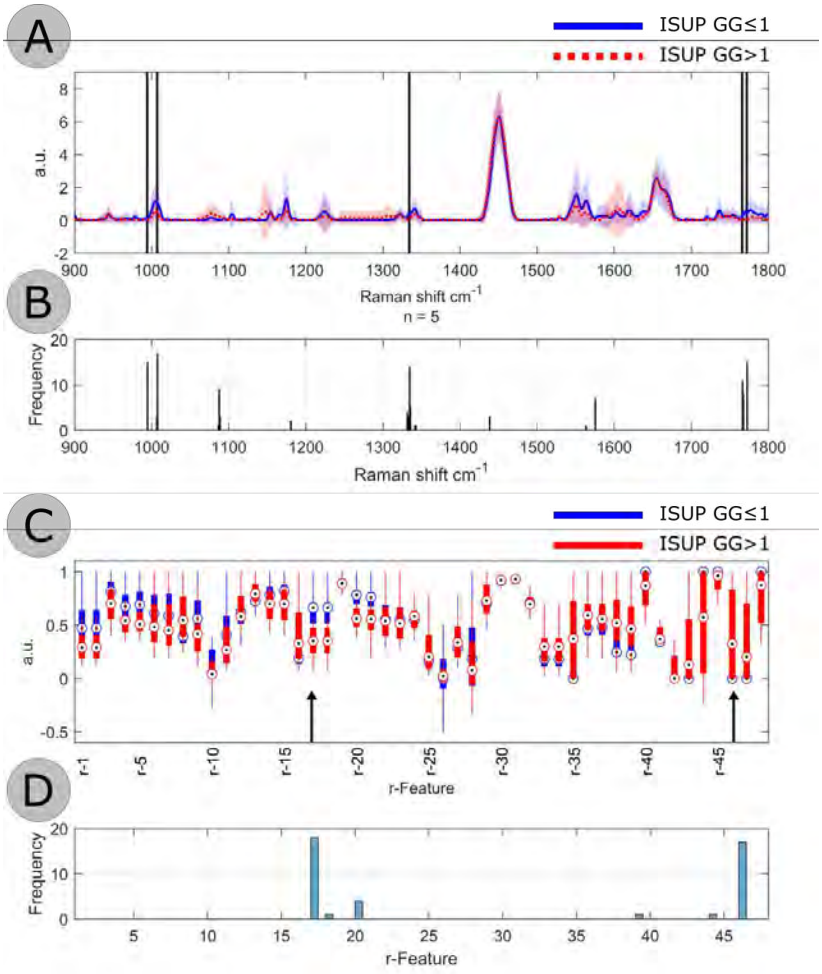


Figure 6.7 Average and standard deviation (A) of processed FP Raman spectra and (B) histograms of selected features for RS; average and standard deviation (C) of mpMRI radiomics and (B) histograms of selected radiomics features. The radiomics features corresponding to each identifier are listed in Table 1.

10 features for in vivo conditions.

There are models trained only with MRI-based features that achieve AUC above 80% in the literature [47, 66]. For instance, a model trained with 9 radiomics features achieved a sensitivity of 84% and a specificity of 73%, detecting clinically significant PCa better than the Rad model [66]. Nevertheless, they used some shape-based features that need the tumor's segmentation and some features extracted after using different filters on the images. The relatively simpler features used in this study simplify the processing of the images, do not require the segmentation of the GTV, and allow the extraction of the features at the place where the probe is placed; this is essential for potential real-time applications.

As multimodal information provides significant advantages for navigation, different char-

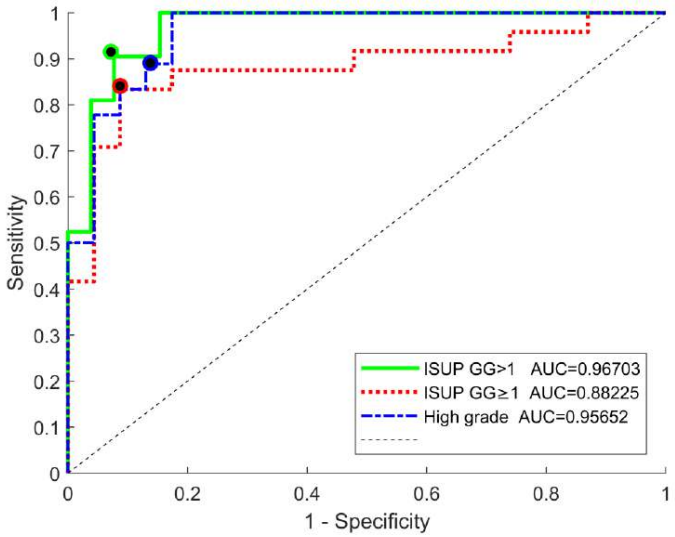


Figure 6.8 ROC curve and optimal point for models trained with the 7 most frequently selected features, for the three different prediction tasks ("ISUP GG>1", "ISUP GG≥ 1", "high grade").

acterization modalities are also advantageous for classification purposes [128]. One of the models that combined optical and image-based properties exhibited the best performance: FP + Rad. The complementary information of mesoscopic and macroscopic characterization allowed FP + Rad to outperform the models that only used a single modality (FP or Rad) on all metrics. Data variability is difficult to model in small datasets; features coming from different modalities, principles, or equipment could contribute to model generalization. The performance of this model, using 8 features from two different modalities, is comparable to the studies mentioned before that use more features from a single modality in ex vivo conditions.

The combination of FP + Rad was the only model which yielded a sensitivity and specificity over 80%, which is very important since prostate-specific antigen (PSA) is known for its high false positives rates, and TRUS-guided biopsies can lead to a high number of false negatives. Some mpMRI- or PET-based methods have greater sensitivity and specificity (respectively) for the localized characterization of PCa, but require time-consuming processing by experts or could have potential side effects for some tracers [19, 47, 131].

The most frequently selected Raman shifts were related to three important spectra bands that are usually more predominant in healthy (or ISUP GG≤1) tissue. The first one, from 990 to 1015 cm^{-1}, has a central peak usually related to proteins (phenylalanine) [23, 52, 94, 133]. From 1330 to 1350 cm^{-1}, the second band is commonly associated with collagen, DNA, or

RNA [52,94]. Finally, the band from 1760 to 1790 cm^{-1}, influenced by DNA/RNA, proteins and phospholipids, is also present on benign prostatic hyperplasia [52,126].

Energy, as a radiomics feature, is the measure of the magnitude of the value (square intensity of segmented voxels), so the lower the intensity is on the image, the lower metric. Low intensity in ADC images is generally associated with tumors, which is consistent with the higher mean value for ISUP GG≤1 tissue for this feature calculated on ADC images (r17) [32,48].

The loss in performance when using only the samples with HG>20% instead of using all the acquisitions ("high grade" vs. "ISUP GG>1") was very limited ($\sim$ 1% on AUC and 3% on accuracy). This indicates that, despite having limited datasets, by eliminating 6 observations (and using the same features), the model still performed adequately, which can be explained by the pooling of features from macro- and mesoscopic levels and the fact that the model uses less than 60% of the observations as support vectors. Comparing the "ISUP GG>1" vs. "ISUP GG≥1" prediction tasks, the performance at the optimal point is similar ($\sim$ 4% difference in accuracy), projecting the potential of the approach for both classification tasks. The slightly better performance in this experiment (prediction task experiment) compared to the initial one (feature combination experiment) may stem from the fact that the 7 features were chosen based on the selection of all folds, not on our LOPOCV scheme.

This pilot clinical study has some limitations, such as the limited number of patients enrolled, making the feature selection process critical. The selection approach was useful to exclude many silent features, but also to choose a small number that is comparable, ideally smaller, than the number of observations. Other limitations are related to some elements that may compromise the correct co-localization of the optical acquisitions and histopathological results; although the navigation system helped and the performance for "high grade" was correct (when removing the samples that could be less certain), the co-location of the ground truth would be an aspect to improve for the following stages.

The implemented approach has facilitating elements for performing a minimally invasive classification in real-time. However, for future work, the signal/image processing and the pre-trained classification model need to be integrated into the navigation system to get a complete system for in vivo real-time studies.

6.5 Conclusion

We demonstrated that complementary information from in situ RS and mpMRI radiomics features allowed to accurately stratify the ISUP GG>1/ISUP GG≤1, as well as discrimi-

nate ISUP GG≥1/ISUP GG<1 sites using SVM classifiers. This classification performance, combined with our custom navigation system, can lead to an accurate PCa detection and localization, improving tumor targeting in minimally invasive interventions.

6.6 Disclosures

F. Leblond is cofounder of ODS Medical Inc., a medical device company that seeks to commercialize the Raman spectroscopy system for real-time detection of tissue abnormalities.

6.7 Acknowledgements

This research was undertaken thanks, in part, to funding from the Canada First Research Excellence Fund through the TransMedTech Institute. We would also like to thank all the medical team at the radio-oncology department of CHUM.

CHAPTER 7 ARTICLE 3: ROBOT-ASSISTED BIOPSY SAMPLING FOR ONLINE RAMAN SPECTROSCOPY CANCER CONFIRMATION IN THE OPERATING ROOM

This article was published in the fast-track issue of the *International Journal of Computer Assisted Radiology and Surgery (IJCARS)*, April 2024, and presented at the IPCAI2024 conference, where it received the second-place award for Best Paper.

Remarks: The first author contributed to the experimental measurements, data processing, bibliographic research, and manuscript writing, representing approximately 90% of the work.

Robot-assisted biopsy sampling for online Raman spectroscopy cancer confirmation in the operating room

[1,2]David Grajales, [1,2]William T. Le, [1,2]Trang Tran, [1,2]Sandryne David, [1,2]Frédérick Dallaire, [1,2]Katherine Ember, [1,2,3]Frédéric Leblond, [2]Cynthia Ménard, [1,2]Samuel Kadoury

[1]Polytechnique Montréal, Montréal, Québec, Canada

[2]Centre de recherche du Centre Hospitalier de l'Université de Montréal, Montréal, Québec, Canada

[3]Institut du Cancer de Montréal, Montréal, Québec, Canada

Abstract

Purpose: Cancer confirmation in the operating room (OR) is crucial to improve local control in cancer therapies. Histopathological analysis remains the gold standard, but there is a lack of real-time in-situ cancer confirmation to support margin confirmation or remnant tissue. Raman spectroscopy (RS), as a label-free optical technique, has proven its power in cancer detection and, when integrated into a robotic assistance system, can positively impact the efficiency of procedures and the quality of life of patients, avoiding potential recurrence.

Methods: A workflow is proposed where a 6 DOF robotic system (optical camera + MECA500 robotic arm) assists the characterization of fresh tissue samples using RS. Three calibration methods are compared for the robot, and the temporal efficiency is compared with standard hand-held analysis. For healthy/cancerous tissue discrimination, a 1D-convolutional neural network is proposed and tested on three ex-vivo datasets (brain, breast, and prostate) containing processed RS and histopathology ground truth.

Results: The robot achieves a minimum error of 0.20mm (0.12) on a set of 30 test landmarks

and demonstrates significant time reduction in 4 of the 5 proposed tasks. The proposed classification model can identify brain, breast, and prostate cancer with an accuracy of 0.83 (0.02), 0.93 (0.01), and 0.71 (0.01), respectively.

Conclusion: Automated RS analysis with deep learning demonstrates promising classification performance compared to commonly used support vector machines. Robotic assistance in tissue characterization can contribute to highly accurate, rapid, and robust biopsy analysis in the OR. These two elements are an important step towards real-time cancer confirmation using RS and OR integration.

keywords: Raman spectroscopy; Robotic guidance; Biopsy cancer analysis; Convolutional neural networks.

7.1 Introduction

In targeted cancer therapies, cancer confirmation remains crucial to improve tumor control. Side effects of poor detection of cancerous tissue, such as an increased risk of recurrence, the necessity of a second surgery, or sub-optimal dose delivery, can lower the prognosis. Currently, the gold standard is histopathological evaluation. However, the multiple steps (formalin fixation, paraffin embedding, hematoxylin and eosin (H&E) staining) require time and equipment that prevent its real-time usage [134]. [134]

Attempts to confirm cancer in the operating room (OR) include visual inspection, which lacks specificity, and preoperative rapid frozen section analysis, which increases procedure time while requiring the presence of a pathologist in the OR [135]. Mass spectrometry has also been used for this purpose by analyzing the smoke generated during electrocautery incisions, however the technique is destructive and the system, complex and resource-intensive [136]. Fluorescence imaging has also been studied in the near-infrared (IR) range with good results, but tissue specific optical properties could attenuate the signal while still requiring the use of tumor-specific FDA-approved inks [137]. Consequently, no proven method currently exists for rapid and accurate cancer confirmation, while being easily integrated into existing workflows.

Raman spectroscopy (RS) is a powerful fluorophore- and staining-free optical technique providing real-time molecular signatures by taking advantage of the tissue's highly sensitive optical properties. By using a laser with a specific wavelength, one can stimulate the tissue and detect energy shifts which has proven useful for characterization of several diseases, including Alzheimer's, cardiovascular diseases [138], and cancer [139]. The development of optical fiber RS probes, as well as the imaging modality's non-destructive nature, makes it particularly well-suited for OR deployment [97].

Several studies have been recently conducted using RS in an attempt to achieve real-time cancer confirmation. In ex-vivo, RS was used to detect breast [87] and prostate [21] cancer with an accuracy of 0.91 and 0.82, respectively. For in-vivo, brain cancer [67] was studied with accuracy reaching 0.84. Recently, Pinto et al. proposed the integration of an RS probe to a DaVinci surgical robot to distinguish prostate from extra prostatic tissue during radical prostatectomy [90]. To the best of our knowledge, Grajales et al. [2] is the only clinical workflow for in-vivo RS collection, seamlessly integrating the acquisition process into brachytherapy workflows. However, the need to pause the procedure during RS acquisitions and restricted manual handling of the probe by only approved specialists could all be addressed in an ex-vivo parallel workflow. Indeed, designing a robotic-enabled process should allow simplifying non-invasive tissue sample analysis without affecting the surgical procedure.

Typically, RS spectra are classified based on the distinctive manually selected peaks, using machine learning methods such as logistic regression or support vector machines (SVM). Due to the high number of features, dimensionality reduction based on principal component analysis (PCA) [97] or feature selection using Lasso regression are commonly used as well [2]. Zhang et al., for example, used SVM and PCA with over 0.96 accuracy in breast cancer cell line differentiation [86]. Still, these methods lack the task-specific automatic feature extraction properties that deep learning (DL) methods offer. In particular, convolutional neural networks (CNN) have a less resource-intensive 1D adaptation, with direct application to high-dimensional signals (e.g., time series or spectra) for silent peak detection and feature extraction [95].

In 2020, Santilli et al. explored the use of an autoencoder (5 fully-connected layer encoder and 5 fully-connected layer decoder) for basal cell cancer detection, achieving an accuracy of 0.96 in binary classification [102]. Similarly, Ma et al. proposed a binary classification model, in this context for breast cancer, using a single 1D-CNN layer and two dense layers achieving 0.92 in accuracy [13]; in a cohort of 20 patients, they applied signal processing and data augmentation before the model training to obtain 0.92 in accuracy. In their study, Fisher Discrimination Analysis and SVM classifiers were trained and tested on the same data for comparison.

In this work, we propose a novel workflow for cancer confirmation based on RS acquisitions assisted by an industrial-grade robotic arm with a 1D-CNN-based spectra classification model tested retrospectively on 3 in-house datasets for brain, breast, and prostate cancer. The objective is to improve over existing methods both in terms of prediction accuracy, but as well in terms of workflow efficiency in the OR to lower timing delays, improving precision of sample localization and automating the acquisition process, thus making it suitable to be

incorporated into existing treatments with minimal effort.

7.2 Materials and methods

Targeted treatments requiring pathological cancer confirmation (breast-conserving surgery or biopsy during radiation therapy, e.g., HDR brachytherapy) include in their protocol a step to fix samples from excision or biopsy in formalin, which is sensitive to delays potentially degrading the fresh tissue. This limited window represents the time frame that can accept our proposed robot-assisted workflow (Fig. 7.1) while minimizing disruption of the ongoing therapy, where a custom-made box is used to contain the light and ensure it does not interfere with the OR interventions.

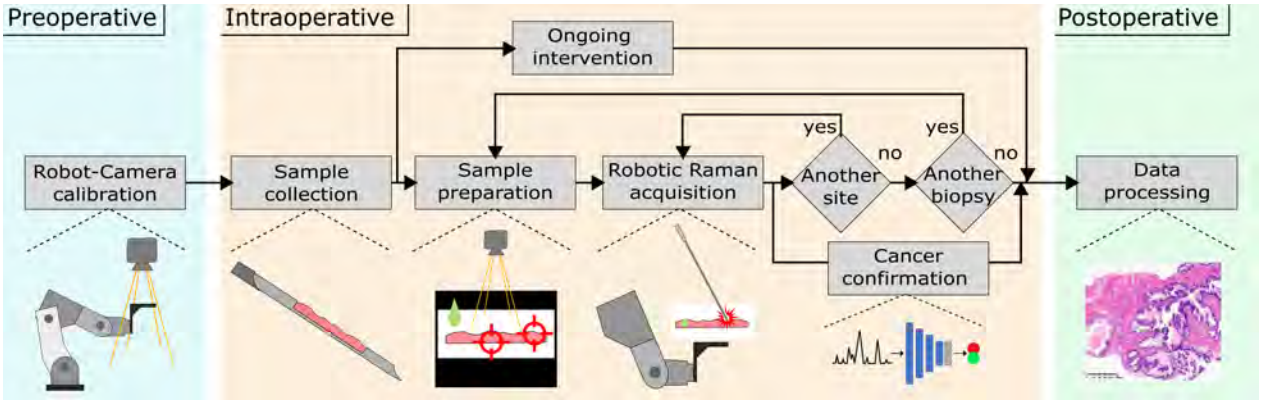


Figure 7.1 Tissue analysis workflow for robot operation and RS acquisitions. The process starts with the one-time calibration of the robot and includes sample handling, characterization (RS), and processing of the collected data and histopathological analysis.

Before the intervention, the proposed system, deployed in the OR, must be calibrated using fiducial markers, which was shown to impact targeting accuracy [140]. Then, the excised tissue is placed on the platform and histopathology-compatible ink is used to identify the orientation of the specimen (just for ground truth purposes) [134]. The platform is placed in the camera field of view, and the image is displayed on a screen, allowing the selection of multiple sites on the tissue sample for RS measurements, clicking on the image based on visual macroscopic evaluation of the tissue integrity. The semi-automated process pipeline then begins with the robot moving the platform so that the sample is under the fixed RS probe for acquisition at each selected site (i.e., just pressing enter after completing the measurement to go to the next site). Finally, formalin fixation can proceed as normal.

This robotic system accuracy has been tested in laboratory conditions with acetaminophen and collagen phantoms, while the training and evaluation of the ex-vivo classifiers were

performed on existing data (with ground truth histopathology). The complete workflow and system will be tested in future stages.

7.2.1 Robot acquisition system

The proposed system (Fig. 7.2) consists of two main components: (1) an optical device performing RS and (2) a 6 DOF robotic arm for biopsy sample handling.

The RS probe (EmVision LLC, Fl) contains a light source (785nm), a spectrometer, and a custom optical probe designed for in-situ characterization (1.22mm diameter). The use of this subsystem, controlled by custom MATLAB R2017b (Mathworks, MA) script, allows RS measurement of the sample's molecular signature.

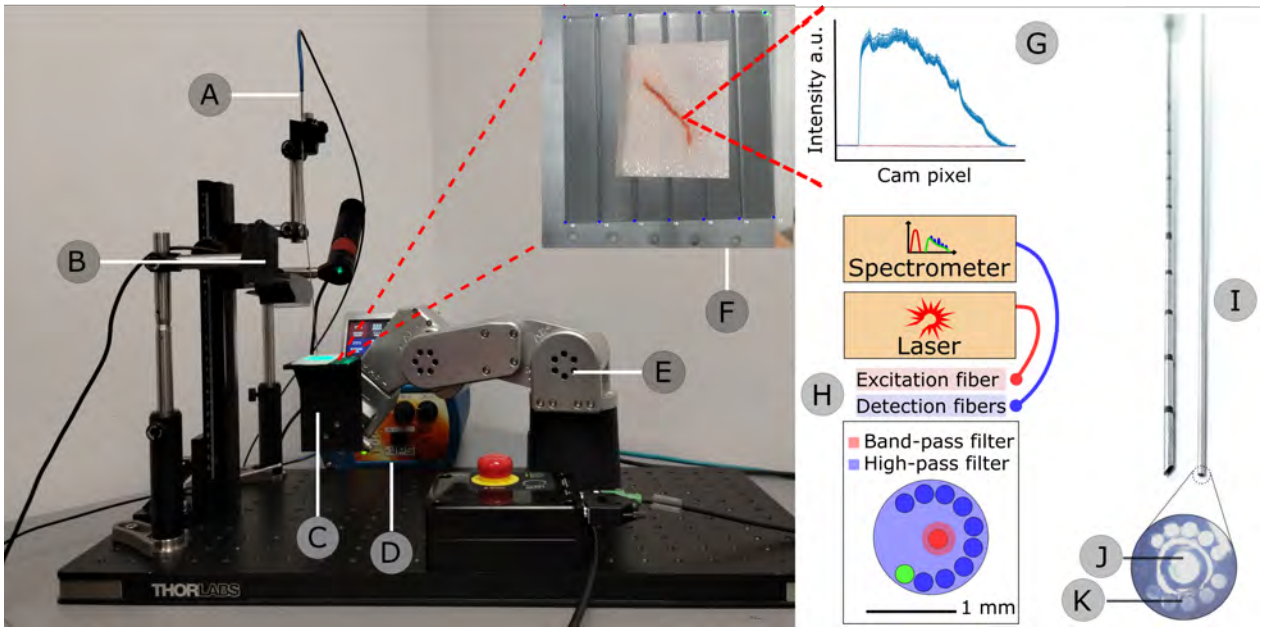


Figure 7.2 (A) Single-point RS probe; (B) Camera; (C) 3D printed platform; (D) Near-IR laser; (E) MECA500 robot arm; (F) Biopsy core on the platform; (G) Raw Raman acquisition; (H) Schematic illustration of optical setup; (I) Closeup of the RS probe fiber bundle next to cannula; (J) Excitation fiber; (K) Detection fibers.

The MECA500 robotic arm consists of a 6-axis compact industrial robot arm (Mecademic, Montreal, Canada), with a built-in controller and a Python API giving access to the robot's motion features. A custom 3D printed platform is added to the robot's tool plate to support the ex-vivo tissue sample. A Logitech C920 web camera (Logitech, Lausanne, Switzerland) is used to calculate the platform's pixel-to-real-world coordinates correspondence.

The robot is used to manipulate small tissue samples ($<$ 1cm) taken with a biopsy gun (23

gauge, 17mm in length, and 0.84mm in diameter). The customized platform was printed in PLA and with 19 reference points (fiducials) on a 3D printer with a resolution of ~ 0.2mm. The robot offers high precision and accuracy with a path and a position repeatability of 0.005mm.

The Mecademic Python API for the robot control offers, among others, commands to set a tool reference frame (TRF) defined in the center of the platform surface. It also has commands to move the TRF to a specific position (MoveLin$(x, y, z, \alpha, \beta, \gamma)$), define the angle for each of the 6 joints (MoveJoints$(\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6)$), define linear velocity (SetCartLinVel(v)), and error handling. All these commands were used to program methods to place the platform in the correct position with the desired velocity, avoiding collisions (position constraints) with the other objects in the workspace. The trajectories between the known positions of interest (e.g., camera and RS probe positions) were defined by linear paths, without blending, within the constrained workspace.

Calibration method

For the proposed application, where the RS probe and camera are in a fixed position, and the standard diameter sample is on the platform referenced by the robot, the depth does not need to be accounted for during calibration. Thus, a 2D-2D correspondence between the pixels of the camera image (m, n) and the coordinates on the robot platform (x, y) needs to be determined. We compared two mapping methods solved with the least squares method and one method using open-source tools from OpenCV.

Linear mapping: In this over-determined system, the least squares method can be used, which in this case requires at least three known points to generate the six equations to solve the system. Using Eq.7.1 and 7.2, a mapping of the position of a pixel and the actual position can be performed:

$$D_x m + E_x n + F_x = x \tag{7.1}$$

$$D_y m + E_y n + F_y = y \tag{7.2}$$

Conic mapping: For the conic approach, the principle of the linear method is used, comparing Eq.7.1–7.2 along with 7.3–7.4, of which the linear method is a particular case. However, by mapping to the equation of a conic, typical effects of lens distortions commonly found in the periphery can be effectively mitigated. This system can be solved with a minimum of six

known points.

$$A_x m^2 + B_x mn + C_x n^2 + D_x m + E_x n + F_x = x \tag{7.3}$$

$$A_y m^2 + B_y mn + C_y n^2 + D_y m + E_y n + F_y = y \tag{7.4}$$

The mapping equation $M \cdot U = R$, where M contains information from (m, n), U contains the 12 unknown coefficients and R contains the real (x, y) coordinates, can be solved for U by trigonometric handling.

Homography: We finally evaluated the homography calibration module from the OpenCV library which allows finding the perspective transformation (homography matrix) from at least 4 known points, a 3x3 matrix that considers rotation and translation in 2D (H). Once H is determined, (x, y) could be determined using the perspective transformation.

7.2.2 RS classification model

In contrast to the widely used SVM methods for RS cancer detection [86], DL approaches have only begun to be explored in recent years for RS [13, 95]. We propose a CNN-based model, fed with complete processed RS signals, to be compared with the baseline SVM.

Proposed 1D-CNN classification model

In the proposed approach, the selected architecture consists of 4 x 1D convolutional layers, each followed by a batch normalization and ReLU activation function for automated feature extraction. Specifically, 1D convolutions are well suited to detect salient peak signals and automatically extract features from complete signals [95, 102]. The number of layers and features (120, 60, 30) were selected to maximize classification performance while minimizing model capacity to better adapt to smaller datasets, inherent to pilot studies. For the discrimination component, a fully-connected layer (30 features) allowed training for supervised binary classification. It should be noted that using processed RS as inputs still requires manual intervention to correctly tune the preprocessing steps, once per patient surgery. This preprocessing (which is a considerable challenge) consists of averaging the accumulations, background and cosmic rays subtraction, normalization (NIST Raman standard - SRM2214), autofluorescence removal (BubbleFill [4]), standard normal variate normalization, and assigning each pixel of the spectrometer to a Raman shift [87].

Implementation details: The CNN was trained for 30 epochs with Adam optimizer and

learning rate of 0.001. All convolutional layers used a kernel size of 3 and a stride of 2 as a downsampling strategy. Models were trained on an NVIDIA GPU with 12GB RAM, optimizing the cross-entropy loss, implemented in PyTorch.

Baseline SVM model

In the literature, studies have identified specific peaks in the fingerprint region of the RS signal that have significant classification power for brain (1004, 1300, 1339, 1443, and 1659 cm^{-1}) [97], breast (940, 1004, 1129, and 1155 cm^{-1}) [87], and prostate (994, 1007, 1334, 1766, and 1772 cm^{-1}) [2] cancers. These peaks were used to establish the baseline for comparison. The SVM model was implemented using Scikit-learn and trained using a linear kernel, which is commonly used in RS classification and which, in specific applications (e.g., breast cancer and hepatitis), provides greater accuracy and specificity [13,133]. To reduce bias from training on an unbalanced dataset (breast dataset), the cost function penalized false positives by doubling the weight.

7.2.3 Datasets

Three retrospective datasets (Table 8.1) containing processed RS and histopathological analysis as ground truth were used to evaluate the proposed classification models integrated with the robotic system: an open-brain surgery study at Mount Sinai and Montreal Neurological Hospital, a breast conserving surgery (BCS) study at McGill University Health Centre (samples with $> 80\%$ cancer cells were labeled cancer; samples with $> 80\%$ normal cells were labeled healthy; the rest were discarded), and a high dose rate (HDR) brachytherapy study at the Centre Hospitalier de l'Université de Montréal (cancer labels were assigned to samples with a grade group according to the International Society of Urological Pathology (ISUP-GG) > 1).

The spatial correspondence between the site inspected with RS and the histopathological analysis is crucial. In the three studies mentioned above, some strategies were used to limit the uncertainty in the correspondence, such as averaging of cell counts in different regions of the specimen, thresholds of the percentage of predominant tissue, grids and ink for macroscopic picture/H&E image registration, border measurements exclusion, or TRUS+EM-tracking navigation system [2, 87].

Table 7.1 Patient characteristics of the available brain, breast and prostate cancer patient cohorts. Figures are presented either as counts or as median (range).

	Brain (n=67)	Breast (n=20)	Prostate (n=20)
Cohort			
Sex (female)	31	20	0
Median age at diagnosis	64 (52–74)	67 (54–77)	68 (64–74)
Acquisitions			
Measurements per patient	18 (1–80)	6 (1–25)	2 (2–5)
Total RS measurements	985	169	47
Accumulations per spectrum	20	10	50
Points per spectrum	1012	631	1801
Site label distribution			
Normal	420	59	26
Cancer	565	110	21

7.2.4 Experimental methodology and evaluation metrics

Robot positioning and timing: The three calibration methods were evaluated by varying the number of fiducials; the root of the mean squares of the Euclidean distance between the actual coordinates and the calculated coordinates of each fiducial used in the calibration process was calculated to define the optimal number of fiducials. Then, using the determined optimal fiducial count, the 3 methods were compared using 30 known test points for positioning error.

To evaluate manipulation time, 5 tasks were established to compare the length of baseline human workflow versus the robotic system. In the first (1-Any) and third (3-Any) tasks, RS measurements were to be acquired for 1 or 3 points anywhere on the tissue sample after (simulated) excision. Tasks two (1-Spec) and four (3-Spec) were performed similarly, but with predetermined challenging measurement locations on the sample. The final task requires performing RS acquisition at an initial point, a displacement to a secondary point to apply additional processes (AP), then returning to the initial point for a second post-AP acquisition (RS-AP-RS). This type of procedure may be of great value in future studies to test the effect of AP on the sample (e.g., photobleaching, inking, saline hydration, *etc.*) or the use of two probes for multimodal analysis. For each timing test, experiments were performed in triplicate and the measured times were reported as mean and standard deviation (SD).

Classifier training and evaluation: A 5-fold patient cross-validation was performed for each dataset (i.e., for each fold, the model was trained with data from $\sim 80\%$ of the patients; the rest was used for validation once). Given the small prostate dataset size ($n = 47$), $5\times$ offline data augmentation was performed based on Santilli et al. [102], before train/validation

split (i.e., augmented data were also included in the validation set). Performance metrics included area under the receiver-operating-characteristic (ROC) curve (AUC), accuracy, sensitivity, and specificity. The complete training-validation process was repeated 5 times for significance testing, with figures reporting the mean and SD. Statistical analysis was evaluated using a paired Student's t-test at $p < 0.05$ to show significant improvement over SVM baseline.

The proposed model's hyperparameters were determined (optimizing AUC) from a search of different learning rates (0.001, 0.005, 0.01), epochs (20, 30, 60), number of layers (2, 4), and features ([120, 60, 30], [120, 120, 30]) in a separate study [110].

7.3 Results

The accuracy of each calibration method was evaluated based on the number of fiducials used as shown in Fig. 7.3A. With the minimum number of calibration points, the linear and conic methods both present errors larger than the acceptable minimum sample size (1.52mm and 0.98mm $>$ 0.84mm). The worst performance shown by the homography method is 0.34mm with 5 points used for calibration. Above $n = 14$, no improvements can be obtained for all methods (linear: 0.40mm; conic: 0.21mm; homography: 0.22mm).

Landmarks (30 known test points) were evaluated for each method calibrated using 14 fiducials (determined in the previous step) in Fig. 7.3B. Both the conic and homography methods present a mean improvement over the linear method (0.14mm [95%CI 0.062, 0.22] and 0.11mm [95%CI 0.028, 0.19]) of 50% and 36% respectively of the linear method's average landmark distance (0.30mm $\pm$0.17SD). No significant difference was observed between the positioning accuracy of the conic and homography method (0.032mm [95%CI -0.25, 0.095]).

Fig. 7.3C shows the timing results for the 5 experimental setups, 3 repetitions. In all but one setup, the proposed robot-assisted workflow presents significantly faster results than the 3 human users (co-authors), with the largest improvements of 61.7s [95%CI 51.7, 66.7] against user A, 51.3s [95%CI 49.7, 53.0] against user B, and 54.7s [95%CI 45.7, 57.9] against user C in the RS+AP+RS setup, representing a 73%, 69% and 70% speedup respectively. The users were relatively familiar with the use of the optical stands and sample handling, and, at the time of measurement, they did not know the measured time for the robot or the other users. In the 1-Spec setup, no significant difference was observed between user B and the proposed method (1.67s [95%CI -0.33, 3.0]). Overall, significant gains are less noticeable in single-point acquisition (1-Any, 1-Spec) and more important in repetitive multi-point acquisitions (3-Any, 3-Spec, RS+AP+RS).

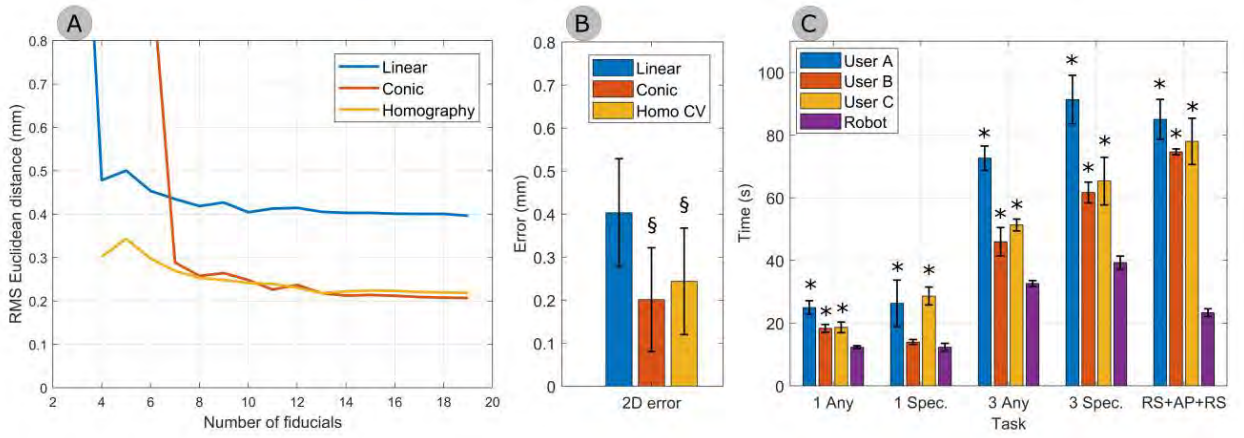


Figure 7.3 Accuracy and timing efficiency of the robotic workflow. (A) Calibration performance of the 3 methods as a function of the number of fiducials; (B) Mean Euclidean distance on 30 known landmarks. § indicates statistically significant difference compared to the linear method; (C) Time comparison by task. Asterisks indicate statistically significant difference compared to the robot.

Fig. 7.4 presents a classification ROC curve for each dataset. Statistically significant difference in mean ROC AUC was found for brain cancer (0.82 ± 0.03SD$< 0.90 \pm 0.02$SD, $p = 0.02$), but not for breast cancer (0.95 ± 0.01SD$= 0.95 \pm 0.03$SD, $p = 0.60$). For prostate cancer, hypothesis testing could not be performed from the reference study, however AUC was lower in the proposed method with 0.72 ± 0.04SD< 0.74.

Table 8.2 shows the thresholded metrics for comparative performance. Accuracy was significantly better in brain ($0.83 > 0.78$) and breast ($0.93 > 0.90$) while remaining comparative in the prostate dataset. Sensitivity was improved as well for breast cancer ($0.94 > 0.89$), and specificity for brain cancer ($0.94 > 0.81$). SD are not reported for SVM in prostate due to limited sample size and distribution of samples.

7.4 Discussion

In this work, a novel workflow was proposed for cancer confirmation on ex-vivo tissue that could be integrated with minimal disruptions into the OR for any targeted tumor therapy, combining CNN-based RS tissue analysis with robot-driven support. A semi-automated pipeline is presented using the MECA500 for robot-assisted non-destructive tissue spectral acquisition in parallel with the treatment, with both high positional precision and rapid process that would minimally impact subsequent delays to formalin fixation.

For robot positioning accuracy, homography and conic methods remain comparable when the

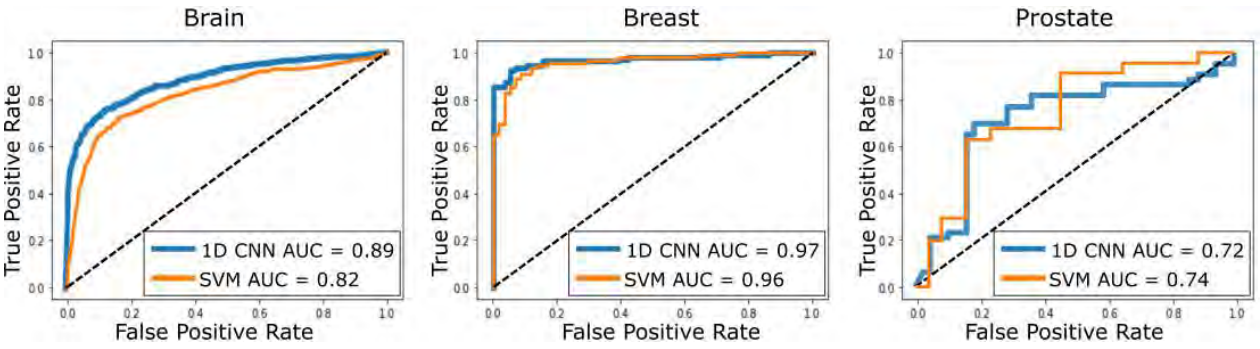


Figure 7.4 The ROC curve for the proposed model (1D-CNN) and the baseline (SVM) for brain, breast, and prostate cancer.

Table 7.2 Comparative results of 5-fold cross-validation for discriminating cancer from normal tissue. Note that baseline prostate results are obtained from [2] which uses single repetition leave-one-out cross-validation. Asterisks indicate statistically significant improvement against SVM baseline ($p < 0.05$). All figures are presented as: mean (SD) of 5 repetitions.

Model	Metric	Brain	Breast [110]	Prostate [2]
SVM	Accuracy	0.78 (0.03)	0.90 (0.01)	**0.72**
	Sensitivity	**0.76 (0.02)**	0.89 (0.01)	0.67
	Specificity	0.81 (0.05)	**0.91 (0.02)**	**0.77**
1D-CNN	Accuracy	**0.83 (0.02)***	**0.93 (0.01)***	0.71 (0.01)
	Sensitivity	0.74 (0.02)	**0.94 (0.02)***	0.67 (0.06)
	Specificity	**0.94 (0.04)***	0.88 (0.04)	0.76 (0.07)

recommended number of fiducials can be used (n=14). However, when calibration fiducial numbers are limited (4–6), the homography method is recommended. For both methods, the expected error rate represents $\sim 25\%$ of the biopsy sample size, with the potential to be better than the handheld protocol in addition to decreasing involuntary movements during RS measurement and not being limited to a grid, as presented in [2, 87]. Furthermore, the use of enhancing techniques, such as photobleaching, has recently been studied to reduce the effects of autofluorescence (via exposure of the analyzed sample to a secondary light source), which could improve downstream analyses [141]. To efficiently apply this at the same place where the RS is measured, high calibration is required, which favors the proposed robot system.

The delays incurred by any new workflow to be used in the OR remain a strong limiting factor for clinical adoption. In the more complex task (RS+AP+RS), where an AP could be applied between repeated measurements, the delays of our proposed pipeline average 23s, more than a 3-fold reduction compared with human users. Over the course of brachytherapy procedures

with up to 5 specimen analyses as shown in [2], this represents a modest $\sim$ 2min of impact over a 40–90min operation. Furthermore, unlike previous in-vivo workflows [2], the current robot system allows non-blocking RS acquisitions on excised samples, thus minimizing practical delays on the surgeon's workflow. Nevertheless, that lack of intraoperative validation remains a limitation for potential clinical deployment, an aspect which is planned in the next stages.

For downstream cancer confirmation tasks, state-of-the-art performance was demonstrated for both brain and breast, but with little effect on prostate cancer, which may be due in part to limited dataset size ($n = 47$), though likely that intrinsic prostate tissue properties remain challenging. It has been previously noted that high autofluorescence would require protocols with longer acquisition times to compensate for low signal intensity [1,21]. In terms of cancer confirmation performance of the proposed method based on CNN, both brain and breast datasets demonstrated improvements over the current state-of-the-art approach for RS classification (SVM). Still, the primary advantage this particular deep learning approach provides is automatic feature selection for the available spectra, compared with the baseline method requiring a separate study to determine the optimal RS fingerprint per cancer site, which may vary based on the nature of the cancer. Furthermore, this could also partly justify the comparative performance between CNN and SVM for prostate cancer: the baseline handcrafted features are optimally informative.

One major limitation of this study is the limited dataset sizes, totaling 107 patients across 3 cancer sites. Outside of gathering a larger cohort, data augmentation is a technique that has not been extensively used on all the datasets and could positively impact classification performance. Recent studies have also compared SVM and CNN and obtained an accuracy of 0.92 and 0.75 when using 1D-CNNs with and without data augmentation, respectively [13]. Given their cohort's similarity to the current study, improvements are expected with the use of such a technique.

7.5 Conclusion

The current study represents an important step towards real-time cancer confirmation using RS by streamlining the tissue analysis process while pushing the detection performance on multiple cancer sites. The proposed robotic process not only simplifies its intra-operative use, but also demonstrates its potential for minimally invasive procedures in order to accelerate its clinical translation.

7.6 Acknowledgements

This research was undertaken thanks, in part, to funding from the Canada First Research Excellence Fund through the TransMedTech Institute.

7.7 Declarations

Competing interests: F. Leblond is co-founder of ODS Medical (now Reveal Surgical) formed in 2015 to commercialize a Raman spectroscopy system for neurosurgical and prostate surgery applications. The other authors declare no conflicts of interest.

Ethics approval: All procedures performed in studies involving human participants, from which data were used in the present research, were in accordance with the ethical standards of the institutional and/or national research committee: (ODS Sentry System-1000/2019-5313), (HS#: STUDY-20-01371), (EAN 2021-5997), (NCT03378856).

Consent to participate: Informed consent was obtained from all individual participants included in the studies.

CHAPTER 8 ARTICLE 4: DEPLOYMENT OF A REAL-TIME PROSTATE CANCER CONFIRMATION SYSTEM WITH RAMAN SPECTROSCOPY: FINE-TUNING VS TEST-TIME ADAPTATION OF 1D-CNNs

This article was submitted to the journal *Biophotonics Discovery (BIOS)*, May 23rd, 2025.
Remarks: The first author contributed to the experimental measurements, data processing, bibliographic research, and manuscript writing, representing approximately 90% of the work.

Deployment of a real-time prostate cancer confirmation system with Raman spectroscopy: Fine-tuning vs test-time adaptation of 1D-CNNs

[1,2]David Grajales, [1,2]William T. Le, [1,2]Victor Blanquez-Yeste, [1,2]Frédérick Dallaire, [2]Feryel Azzi, [2]Guila Delouya, [2]Dominique Trudel, [1,2,3]Frédéric Leblond, [2]Cynthia Ménard, [1,2,3]Samuel Kadoury

[1]Polytechnique Montréal, Montréal, Québec, Canada
[2]Centre de recherche du Centre Hospitalier de l'Université de Montréal, Montréal, Québec, Canada
[3]Institut du Cancer de Montréal, Montréal, Québec, Canada

Abstract

Significance: Prostate cancer (PCa) confirmation during needle-based procedures is limited by the lack of intraoperative diagnostic tools. Raman spectroscopy (RS), combined with classification models, offers a promising solution for real-time tissue characterization, potentially improving sampling accuracy and therapy guidance. However, such models require tissue- and organ-specific data, making deployment in new studies challenging due to limited data availability.

Aim: To develop a 1D convolutional neural network (1D-CNN) for real-time PCa detection using RS on prospectively collected ex-vivo data, leveraging multi-organ pre-training and evaluating two domain adaptation strategies.

Methods: A ResNet-based 1D-CNN was trained for binary cancer/normal tissue classification. We implemented a pre-training strategy using retrospective RS data from brain, breast, and prostate (202 patients), along with pre-trained bacterial models, followed by efficient fine-tuning and test-time adaptation (TTA) to adapt to unseen domains.

Results: Prospective RS data were acquired using a robotic system from 10 PCa patients

(2–5 biopsies each). The fine-tuned model achieved 0.76 AUC, 0.79 accuracy, 0.83 sensitivity, and 0.72 specificity, outperforming support vector machines. TTA improved predictions when labels were unavailable.

Conclusion: Pre-trained 1D-CNNs combined with efficient fine-tuning or TTA enable accurate PCa detection in small-cohort settings using real-time RS.

Discovery statement: This work utilizes Raman spectroscopy and transfer learning strategies in a robot-assisted workflow to enable real-time prostate cancer detection. Fine-tuning and test-time adaptation of 1D-CNN models enable effective classification in small-cohort settings, supporting the clinical translation of RS-based tools for intraoperative decision-making.

keywords: Prostate cancer analysis; Raman spectroscopy; Convolutional neural networks; Test-time adaptation; Robotic guidance.

8.1 Introduction

In targeted cancer therapies for prostate cancer (PCa), spatially resolved biopsy sampling with rapid online cancerous tissue confirmation stands to improve diagnosis, prognostication, and guided therapies [36]. While several methods have been explored for intraoperative cancer confirmation, including mass spectrometry for breast cancer margin detection, [103] micro-ultrasound for PCa, [142], and fast freezing techniques, these tests remain invasive, resource-intensive, time-consuming, or expertise-dependent. As a result, the quest for a reliable, non-destructive method for immediate cancer confirmation —that can be used, for instance, to guide biopsy sampling, radioactive seed implantation in brachytherapy, or surgical margin characterization during tumor resection— remains a critical objective in enhancing the precision of targeted therapies.

Raman spectroscopy (RS) is a label-free optical imaging technique that provides real-time molecular signatures based on the inelastic scattering of light. When tissue is illuminated by a laser, most photons are elastically scattered, but a small fraction undergoes energy shifts due to interactions with molecular vibrational modes (Raman effect). These energy shifts generate a spectrum that reflects the molecular composition of the tissue. This information can be exploited to detect various diseases, including Alzheimer's, cardiovascular conditions, and cancer [1, 50].

Particularly in PCa, RS has been used to differentiate prostatic and extraprostatic tissue (with a probe integrated into a DaVinci surgical robot during radical prostatectomy) [90] and for the detection and grading of cancer. An in-vivo PCa detection study reported an accuracy

of 0.72 using RS alone and 0.83 in a multi-modal approach (RS + multiparametric MRI-based features) [2]. On the other hand, in a large cohort ex-vivo study with 84 patients, RS allowed differentiating cancer grades with a sensitivity of 0.90 and specificity of 0.80 [21]. However, prostate RS faces significant challenges that hinder its effectiveness in clinical applications [50]. One major issue is the significant autofluorescence in prostate tissue compared to other organs, which can limit the quality of an already weak signal [96, 143]. Another challenge is the accurate assignment of site-specific ground truth (GT) labels, as opposed to core-level diagnosis, which directly affects the performance of classification models and sometimes complicates data acquisition (e.g., requiring ink marking steps) [96, 144]. Finally, small cohorts constitute a significant obstacle to obtaining high-performance models, as they limit the complexity of the models that can be used and their generalizability [50,142]. Addressing these challenges is crucial for enhancing the diagnostic potential of RS in PCa.

Across medical (in-vivo and ex-vivo) and non-medical applications, RS-based sample classification is commonly performed using machine learning methods such as support vector machine (SVM) or logistic regression, often combined with dimensionality reduction (e.g., principal component analysis) or feature selection (e.g., Lasso regression) due to the signal's high dimensionality [21, 50]. Although successful applications have been reported, including Zhang et al.'s breast cancer cell line differentiation with 0.96 accuracy using SVM and PCA [86], these approaches lack automated task-specific feature extraction, which can be time-consuming or experience-dependent [14, 145].

In the past decade, deep learning (DL)-based methods have emerged as a promising alternative for classification tasks. While convolutional neural networks (CNNs) are traditionally applied to 2D data such as images and videos for classification, detection, or segmentation, 1D-CNN adaptations are less resource-intensive and well suited for silent peak detection and automated feature extraction on high-dimensional 1D signals (e.g., time series or spectra) [95,145]. These capabilities have facilitated the adoption of 1D-CNNs in various medical applications, including electroencephalography (EEG) [16]. While their use in RS remains limited, promising results have been reported. Ho et al. achieved 82% accuracy in identifying 30 bacterial pathogens in patient isolates (i.e., clinical strains), even with low-quality RS signals, using pre-training and fine-tuning strategies [15].

Recent advancements in DL have explored the use of data from related domains to mitigate data scarcity in specialized tasks. In image analysis, models pre-trained on large-scale natural image datasets (e.g., ImageNet with 1.28M images) are frequently adapted to medical imaging tasks with far fewer samples (often fewer than 600 images) [104]. Even foundation models like MedSAM, trained on over 1M medical images for universal segmentation, are initialized with

weights from SAM, a model trained on natural images [105]. This transfer learning approach enables models first to learn generalizable representations from abundant high-quality data and then adapt —through fine-tuning, domain adaptation, or test-time adaptation (TTA)— to specific downstream tasks [16, 103, 107, 108]. Given the success of this strategy in imaging, we explore its potential for RS-based tissue classification by leveraging data from related domains that share the same spectroscopic technique and biological features.

In this study, we propose an approach leveraging previously acquired multi-organ cancer spectra (202 retrospective cases) for pre-training a 1D-CNN-based classification model, followed by an adaptation step to prospective prostate spectra. We compare an unsupervised (TTA) and a supervised (efficient fine-tuning) adaptation strategy, tailored to different scenarios based on GT availability. A robot-assisted RS framework was deployed in the procedure room (PR) to collect prospective PCa data (target dataset). Our complete approach aims to address two of the challenges of PCa detection: accurate labels and limited specific data from 10 prospective patients, which ultimately contribute to improved diagnostic capabilities and precision in targeted therapies.

8.2 Materials and Methods

8.2.1 Retrospective datasets

To enable the use of recent pre-training methods in DL and to increase the available development datasets, three retrospective cohorts (Table 8.1) from in-house biobanks —containing processed RS and histopathological analysis from different anatomical sites— were combined into a single multi-organ cancer confirmation dataset. The three retrospective datasets are:

Brain: An open-brain surgery dataset where samples containing $> 90\%$ cancer cell burden were labeled as cancer, and as normal if no tumor cells were present; the rest were discarded [10], resulting in 985 measurement samples available for analysis from 67 patients.

Breast: A breast-conserving surgery dataset where samples with $> 80\%$ cancer cells were labeled as cancer and samples with $> 80\%$ normal cells were labeled as healthy; the rest were discarded [87], resulting in 169 measurement samples available for analysis from 20 patients.

Prostate: An in-vivo dataset acquired during high-dose-rate brachytherapy (19 patients), and an ex-vivo dataset on radical prostatectomy (96 patients). Cancer labels were assigned to samples with a grade group according to the International Society of Urological Pathology (ISUP GG) ≥ 1 [2]. While these data come from the same organ as our prospective target dataset, the instruments, protocols, and sample characteristics (e.g., specimen thickness, inspected depth) are different.

For these retrospective datasets, methods were employed to reduce alignment uncertainty between inspected site and GT, such as averaging cell counts across different regions of the specimen, using grids and ink for aligning macroscopic images with hematoxylin and eosin staining (H&E) images, or employing an electromagnetic/ultrasound navigation system [2, 10, 87, 96]. Overall, 4,655 RS measurements were available for analysis from 202 patients.

Table 8.1 Characteristics of the available brain, breast and prostate cancer patient cohorts. Figures are presented either as counts (percentage) or as median (range).

	Retrospective			Prospective
	Brain (n=67)	Breast (n=20)	Prostate (n=115)	Target Prostate (n=10)
Cohort				
Sex (female)	31 (46.3%)	20 (100%)	0 (0%)	0 (0%)
Median age at diagnosis	64 (52–74)	67 (54–77)	68 (64–74)	66 (63–72)
Acquisitions				
Measurements per patient	18 (1–80)	6 (1–25)	2 (2–75)	10 (7–18)
Accumulations per spectrum	20 (20–20)	10 (10–10)	35 (20–50)	10 (10–10)
Site label distribution				
Normal	420 (42.6%)	59 (34.9%)	3066 (87.6%)	45 (43.3%)
Cancer	565 (57.4%)	110 (65.1%)	435 (12.4%)	59 (56.7%)

8.2.2 Prospective dataset

The target prostate dataset was acquired prospectively in 10 patients with histologically confirmed PCa diagnosis who underwent high-dose-rate brachytherapy. They were enrolled in a prospective clinical trial approved by the Research Ethics Board, and recruited between April and December 2024. The system used for data acquisition in the PR consists of a robotic arm and an optical component (Fig.8.1.a). The optical measurements were performed on 2–5 biopsy cores per patient following the workflow evaluated in previous stages of this project [145], and summarized in the following sub-section.

8.2.3 Instruments and workflow

Prior to every intervention in the PR, the robot+camera subsystem was calibrated with fiducial markers to ensure targeting accuracy during sample handling. The manually excised tissue (biopsy core) was placed on the platform attached to the robot, and histopathology-compatible ink marks were used to determine the specimen's orientation (for GT purposes). Once in the camera's field of view, an image of the biopsy sample is displayed, enabling the

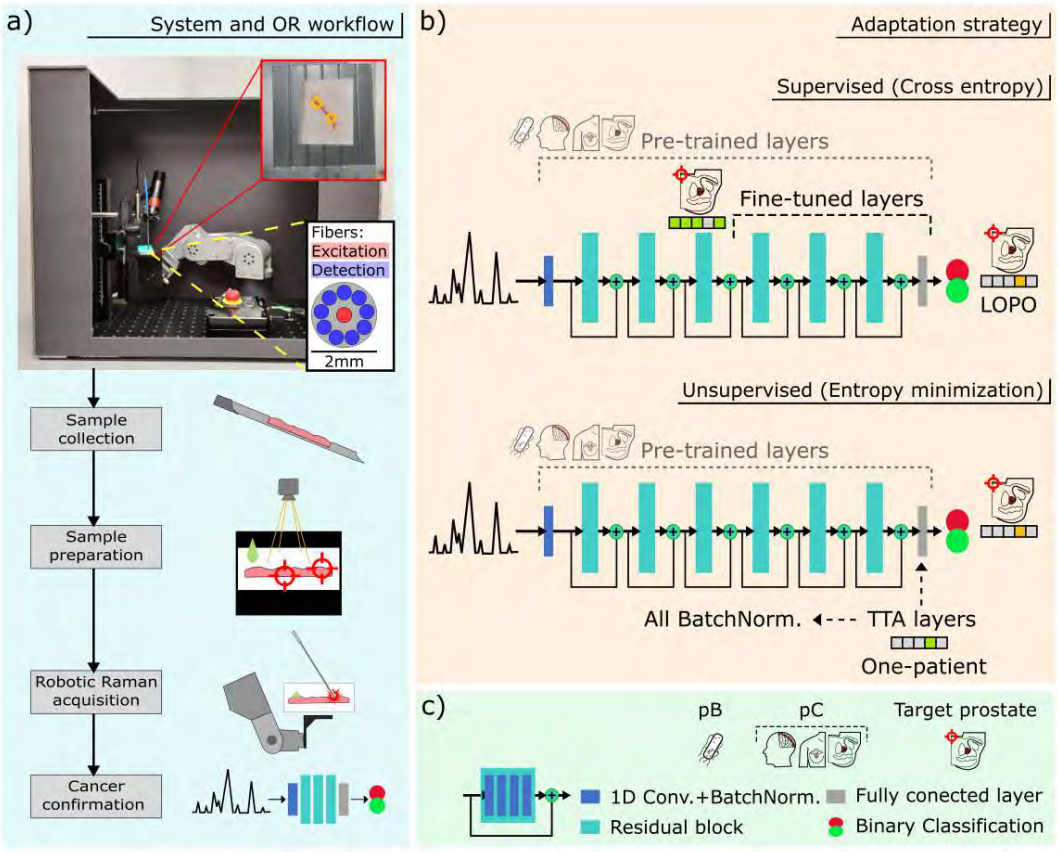


Figure 8.1 System and model overview. a) Robot-assisted Raman acquisition system, including the robotic arm, camera, and optical probe; insets show a front-view schematic of the probe tip (yellow), a close-up of the tissue positioning platform (red), and the intraoperative PR workflow. b) Schematic illustration of the 1D-CNN (ResNet-based) model used for PCa detection, showing the 2 adaptation strategies; leave-one-patient-out (LOPO). c) Symbol legend. pB: Bacteria pre-training; pC: Cancers pre-training (retrospective data from brain, breast, and prostate).

selection of RS measurement sites. The semi-automated pipeline started when the robot moved the platform to place each sample's site under the fixed RS probe. Finally, since RS is a non-destructive technique, the tissue was subsequently fixed in formalin for standard histopathological procedures. After processing, each spectrum consisted of 1,174 Raman shift feature points, presenting a high-dimensional data vector with a sample size of 104 spectra.

Robotic system and calibration: A 6-DOF robotic arm (Mecademic, Montreal, Canada), which integrates a built-in controller and a Python API enabling high precision and accuracy in motion control, was used for biopsy core handling. Images acquired with a Logitech-C920 camera were used to select the sample's sites to be characterized and to find the correspondence between the pixel coordinates in the image (m, n) and the coordinates on the robot

platform (x, y), based on known fiducial, using a conic 2D–2D mapping strategy. Details on mapping strategy selection and robot control can be found in [145]. By leveraging the robot's precise positioning, the distance between each characterized site and a reference ink mark is accurately determined, enabling the assignment of site-specific GT to Raman spectra. This approach reduces the dependency on core-level diagnoses and mitigates potential mislabeling caused by tissue heterogeneity [144].

Raman optical system and measurements: A near-infrared RS instrument (Sentry 1000-R, Reveal Surgical Inc., Montreal) was used to obtain spectral fingerprints from the samples. The complete system consisted of a hand-held Raman probe (EmVision, Florida), a $785nm$ light source (Innovative Photonic Solutions, New Jersey), a spectrometer (Newton model, Andor Technology, Belfast, Ireland), and a control laptop with custom software packages ORAS and ORPL [4]. The $12cm$ probe contains a central excitation fiber surrounded by nine collection fibers of $500\mu m$ core diameter, and optical filters in the probe tip to minimize signals from fiber materials and tissue autofluorescence and to ensure contact to interrogate $0.5mm$ diameter spots [10]. All raw measurements were acquired, in contact, inside of a protection box, covering a range from $400cm^{-1}$ to $2000cm^{-1}$, with auto-exposure control set to a range from $100ms$ to $1000ms$ for 10 accumulations; the detector was pre-cooled at -70°C. The applied preprocessing consists of averaging the accumulations, background and cosmic rays subtraction, instrument response correction (NIST Raman standard - SRM2214), autofluorescence removal (BubbleFill [4]), normalization, and x-axis calibration (Raman shift).

8.2.4 Proposed RS classification models

Given the recently demonstrated importance of pre-training DL models when extracting relevant features on target data, we leverage the publicly available 1D-CNN weights from the works of Ho et al. [15] for bacterial pathogens identification, trained on ∼60,000 RS samples. As their approach used a similar RS acquisition protocol as in this study, we hypothesized that the early feature extraction portion of the model trained on biological samples would directly benefit the adapted PCa confirmation task.

This study used a ResNet-based model for PCa confirmation, capitalizing on the improved gradient propagation and training stability attributed to residual blocks [15,108]. The model, built following the architecture used by Ho et al. [15], consists of an initial 1D convolutional layer (64 convolutional filters), six residual blocks, and a fully connected layer at the end for binary classification (binary cross-entropy loss). Residual blocks have four 1D convolutional layers and a residual connection (shortcut) between the input and output of the block (Fig.8.1.b–c).

The search for hyperparameters such as epochs and learning rate was conducted using one training-validation (60%/20%) split optimized for accuracy. The CNN underwent training for 25 epochs (selected after searching five values in the range 25–100), with early stopping after 8 consecutive epochs without improvement in validation accuracy; the Adam optimizer was used with a learning rate set at 0.001 (selected after searching four values in the range 0.0001–0.01). Each convolutional layer utilized a kernel size of 5 and a stride of 2 for downsampling purposes. The models (1.25M parameters) were implemented in PyTorch and trained on an NVIDIA GPU with 15GB of RAM.

A two-stage pre-training framework was used to address the challenges of adapting DL models to limited data: (i) the above-mentioned weights initialization (Bacteria pre-training or pB), and (ii) pre-training on the combined retrospective heterogeneous cancer dataset from diverse RS systems, organs, and tissue size (Cancers pre-training, or pC). The pre-trained model was adapted to the target prostate dataset under two different scenarios.

Supervised adaptation —Spectra and GT available for fine-tuning the specific classification task: Building on Zhang et al.'s adaptive transfer learning strategies [16], our approach selectively fine-tuned the pre-trained model. Specifically, the first half of the model remained unchanged, while the last three residual blocks and the fully connected layer were adapted to the target data, as shown in Figure 8.1.b, using a binary cross-entropy loss. This approach of optimizing transfer learning reduces trainable parameters by half, and naturally mitigates overfitting without relying on dropout layers [16, 104].

Unsupervised adaptation (TTA) —Spectra available at the moment of acquisitions, with no GT: Here, moving towards unsupervised learning, TTA was applied based on entropy minimization of the model's prediction ($\hat{y}$) of every new measurement with the objective function $H(\hat{y}) = -\sum_c p(\hat{y_c}) \log p(\hat{y_c})$, for the probability $\hat{y_c}$ of class c [107]. In this process, specific layers of the model (final fully connected layer and BatchNorms) were adapted in a limited number of steps (5) looking to reduce the unsupervised loss, then produce a classification prediction [106, 107].

8.2.5 Comparative methods

Baseline SVM: A baseline SVM was implemented for comparative purposes. Feature selection is an important part of using SVMs with limited number of high-dimensional samples, such as RS. For the prostate target dataset, a three-step method (amount of variation > 0.03; correlation with the target > 0.9; Lasso regression for a maximum of 10 features) was applied to the training set of each fold, considering every datapoint in the spectra as a feature [2].

Previous works have shown the potential of specific Raman peaks for PCa detection. Peaks associated with phenylalanine (994cm^{-1}, 1007cm^{-1}), with collagen, DNA, or RNA (1334cm^{-1}), and with DNA/RNA, proteins and phospholipids (1766cm^{-1}, and 1772cm^{-1}) were selected in an in-vivo study [2]. In an ex-vivo study, the peaks at 881cm^{-1} (Tryptophan), 1307cm^{-1}–1310cm^{-1} (Collagen), 1396cm^{-1} (β-Carotene), 1583cm^{-1} (Phenylalanine), and 1602cm^{-1} (Phenylalanine, tyrosine, tryptophan) were identified as biomarkers [21]. These two groups of selected features were also used for the comparative SVM model.

Random forest: A random forest (RF) classifier was also implemented and trained using scikit-learn with 100 trees (n_estimators=100), maximum depth of 5 (max_depth=5), considering all features at each split, and no correction for class imbalance.

8.2.6 Experimental methodology and evaluation metrics

Pre-training strategy selection: For the Bacteria pre-training strategy (pB), the clinical pre-trained model from [15] was used for initializing all network layers. Cancers pre-training (pC) involved all retrospective datasets (i.e., excluding prostate target dataset) divided into $60\% - 20\% - 20\%$ for training-validation-testing. The classification performance of both strategies, and their combination (pB+pC), was tested using the target prostate data (prospective dataset), with and without fine-tuning.

Comparative analysis of adaptation strategies: The proposed adaptation strategies (supervised and unsupervised) were tested and compared with SVM and RF. For fine-tuning adaptation (F), a 10-fold leave-one-patient-out cross-validation was performed on the prostate target dataset, where 9 patients' data were used for training in each fold, while the remaining patient was reserved for validation once. Equivalently, in TTA, one patient at a time was used to adapt the pre-trained model (the adaptation of the model based on one patient does not affect the other patients' results).

Evaluation and metrics: A leave-one-patient-out scheme was used to train the SVM and RF models and to adapt the 1D-CNN models via fine-tuning. Performance was assessed using the area under the receiver operating characteristic (ROC) curve (AUC), accuracy, sensitivity, and specificity. To ensure statistical reliability, all experiments were repeated five times, with results reported as mean and standard deviation (mean (SD)). Statistical significance was evaluated using a Student's t-test ($p < 0.05$), and confidence intervals (CI) were included to quantify result variability.

8.3 Results

8.3.1 Data acquisition

The system was deployed in the PR and collected data from 10 PCa patients (Table 8.1), with 2–5 biopsies obtained per patient. The entire process to characterize one biopsy core (including 3–6 RS measurements per sample) took between 4–5min, excluding biopsy acquisition time. The probe was positioned in direct contact with the tissue to optimize the signal-to-baseline ratio while minimizing interference from the support material. The DL models were deployed to assess inference timing. The time between tissue collection and the first prediction was <3min, including sample preparation, spectral acquisition, processing, and inference —performed in 584ms on average. The time between biopsy and formalin fixation was less than 12min for all samples. The histopathology results of each core, obtained subsequently, were used as GT (cancer labels were assigned to sites with ISUP GG ≥ 1). Images and sample predictions are presented in Figure 8.2.

Figure 8.3.a shows the post-processed Raman spectra data for the 10 prospective cases (target prostate dataset) and the features that were selected for SVM training. In an univariate analysis, statistical values and classification metrics were calculated to evaluate the discriminatory potential of each peak separately. Based on the p-value and mean difference CI, both classes mostly overlap (i.e., all $p > 0.05$ and all CIs include 0). In terms of classification, the best AUC was obtained with the intensities of the phenylalanine-associated peak (997cm^{-1}).

8.3.2 Pre-training strategy selection

The following experiment assessed the impact of pre-training strategies (bacterial spectra initialization and cross-organ transfer) with or without fine-tuning, compared with the base model without pre-training (no_p). Results are shown in Figure 8.4. Overall, fine-tuning not only increased the mean AUC, but also reduced the high inter-run variability.
The model combining both pre-training strategies and fine-tuning (F+pB+pC) and the non-initialized version (F+pC) exhibited superior overall performance. The difference in AUC between them was significant (F+pC - F+pB+pC = -0.06 AUC, $p = 0.012$, $CI = [-0.111, -0.037]$); further analysis showed that F+pB+pC achieved better accuracy (0.79 (0.03)) and specificity (0.72 (0.04)) than F+pC (0.69(0.02), 0.51(0.06)) with both differences being statistically significant ($p < 0.001$ and $p < 0.001$). No significant difference in sensitivity was observed between the models (F+pC - F+pB+pC = -0.01, $CI = [-0.041, 0.027]$).

Another key element of the initialization (pB) is the reduction in pre-training execution time. When calculating the mean number of epochs before the early stop, a value of 14

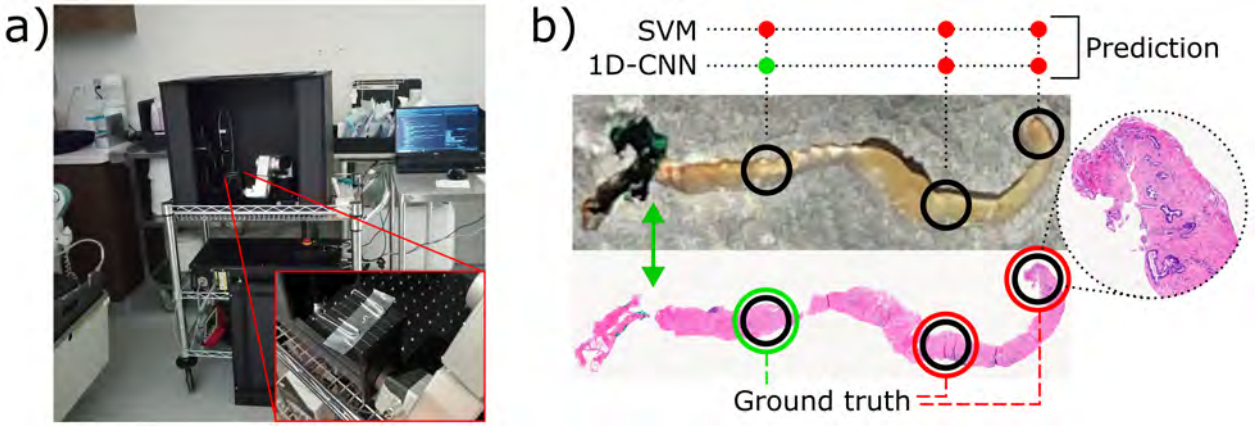


Figure 8.2 System deployed in the procedure room for real-time tissue characterization and classification. a) Robot-assisted RS system (protection box, robotic arm, and optical components) in the procedure room, with a close-up of a biopsy sample on the platform. b) Sample prediction results of real-time biopsy tissue classification. Photograph and corresponding histopathological slide of a biopsy core. Black circumferences: characterized sites; colored circumferences: diagnosis in the pathology report (red: cancer); colored circle: prediction of each model (red: cancer); green arrow: beginning of the ink mark.

(range: 9–17) was found for F+pB+pC, versus 18 (range: 15–23) for F+pC, representing a 26% reduction in execution time during pre-training, approximately −54 minutes per test (5 repetitions).

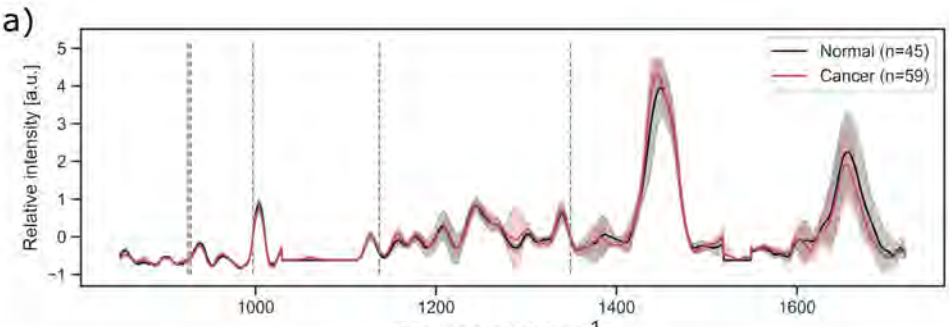
An additional test using only prostate retrospective data for pre-training (pRP+F) resulted in lower average performance and higher variability compared to the pC+F model (pRP+F - pC+F = −0.07 AUC, $CI = [-0.122, -0.012,]$), which was pre-trained on the same data plus brain and breast data. Despite the closer domain similarity, pRP+F showed −0.13 AUC ($p < 0.001$), −0.15 accuracy ($p < 0.002$), −0.07 sensitivity ($p < 0.256$), and −0.25 sensitivity ($p << 0.001$), compared to pB+pC+F, the approach that includes greater diversity of source data.

Based on these findings, subsequent downstream tests adopted the combined strategy of Bacteria pre-training followed by Cancers pre-training (p).

8.3.3 Quantitative analysis of adaptation strategies

Following the selection of the combined pre-training strategy (p), which yielded a residual misclassification rate of 28% during internal validation, the two adaptation strategies were finally evaluated on the prospective prostate dataset (10 patients). Table 8.2 presents the classification results across all metrics, along with those of the comparative models. The SVM

a)



b)

Wave number (cm−1)	Main vibrational modes	Biomolecular assignment	Intensity distribution Distance to mean value	P-value	Mean Difference CI	AUC
925	C--C stretch band	Amino acids		0.981	[-0.068, 0.072]	0.506
928	C--C stretch band	Amino acids		0.515	[-0.078, 0.038]	0.494
997	Symmetric aromaticring	Protein (phenylalanine)		0.075	[-0.001, 0.169]	**0.594**
1137	C--C bond	Lipids		0.123	[-0.119, 0.016]	0.389
1349	C--H	protein (tryptophan)		0.468	[-0.150, 0.082]	0.495

1.0 -0.5 0.0 0.5 1.0

Figure 8.3 a) Average Raman spectra with their variance from the prospective target prostate ex-vivo dataset (10 patients). The vertical dotted lines indicate the features selected for training the SVM models. b) Peak locations and corresponding biomolecular assignments are listed. Statistical and classification metrics, including intensity distribution distance, p-value (cancer vs. normal), mean difference with confidence interval (CI), and AUC, to assess the discriminatory potential of individual Raman shifts for prostate cancer detection.

results were obtained using fold-specific selected features (Figure 8.3) and literature-derived feature sets [2, 21], and the averaged values enable comparison with 1D-CNN performance.

The model fine-tuned after pre-training (1D-CNN+p+F) achieved +0.18 AUC, compared to the SVM baseline, with $p << 0.001$ and a 95% confidence interval ($CI = [0.154, 0.200]$) excluding the origin, suggesting a statistically significant performance improvement. Similarly, it showed a significant improvement of +0.18 ($p << 0.001$) in accuracy and +0.31 ($p = 0.003$) in specificity. Despite a +0.07 difference in sensitivity in favor of 1D-CNN+p+F, this was not significant ($p = 0.204$).

The slight improvements in AUC and accuracy of the pre-trained model adapted via TTA (1D-CNN+p+TTA) over SVM (+0.02, $p = 0.060$, $p = 0.081$) were not statistically significant, nor was the −0.09 difference in sensitivity ($p = 0.113$). However, the improvement in specificity was significant (+0.14, $p = 0.040$). This highlights the potential of TTA as an unsupervised adaptation strategy, achieving performance comparable to a supervised baseline trained with labeled data from 9 patients.

The AUC of the RF model was slightly better than SVM, but the difference was not significant

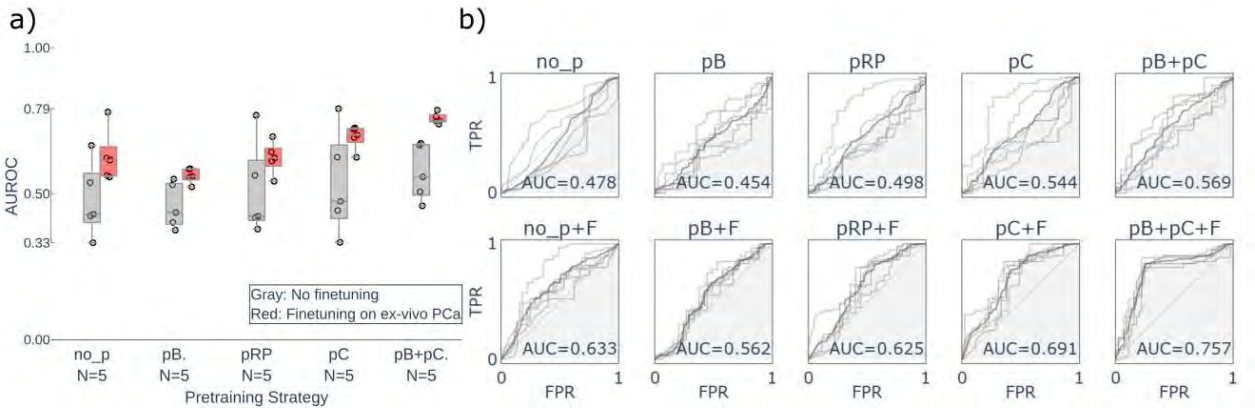


Figure 8.4 Comparison of pre-training strategies for the proposed 1D-CNN model for PCa ex-vivo confirmation using the target dataset (10 patients, leave-one-patient-out), with 5 repetitions. a) Boxes show the interquartile range (25%–75%), while lines show minimum and maximum. b) ROC curve for all model variants. F: fine-tuning; no_p: no pre-training; pB: Bacteria pre-training; pC: Cancers pre-training (retrospective brain, breast, and prostate); pRP: pre-training using only retrospective prostate data.

(+0.04 AUC, $p > 0.124$); sensitivity was the only metric where the difference between SVM and RF was significant (+0.13 in favor of SVM, $p < 0.040$).

To better illustrate the effect of supervised (fine-tuning), and unsupervised (TTA) adaptation, pairwise comparisons are presented in Figure 8.5.a. When using a pre-trained —without fine-tuning— model (1D-CNN+p) with a considerable inter-run variability, applying TTA (1D-CNN+p+TTA) increased AUC by +0.12 ($p = 0.070$), accuracy by +0.13 ($p = 0.017$), sensitivity by +0.13 ($p = 0.192$, $CI = [0.007, 0.329]$), and specificity by +0.11 ($p = 0.325$). In contrast, a previously fine-tuned model (1D-CNN+p+F) decreased the performance after unsupervised adaptation (1D-CNN+p+F+TTA) by −0.02 ($p = 0.102$) in AUC, −0.07 ($p = 0.004$) in accuracy, +0.10 ($p = 0.008$) in sensitivity, and +0.02 ($p = 0.579$) in specificity.

8.4 Discussion

This study explored pre-training and model adaptation methodologies for RS-based PCa confirmation in the PR, with prospective data from biopsy samples acquired using a novel robot-assisted workflow allowing site-specific GT assignment. The findings highlight the potential of leveraging multi-organ RS datasets to enhance real-time diagnostic accuracy in limited-cohort settings.

The significant variations observed in the baseline SVM results underscore the crucial impact of feature selection on model performance, revealing a strong dependence on both organ-

Table 8.2 Comparative results of 10-fold cross-validation for discriminating cancer from normal tissue. Asterisks indicate statistically significant difference against SVM baseline ($p < 0.05$). All figures are presented as mean (SD) of 5 repetitions. F: fine-tuning; p: Bacteria pre-training + Cancers pre-training.

Model	Variation	AUC	Accuracy	Sensitivity	Specificity
SVM	Feat. Sel.	0.57	0.60	0.88	0.26
	Feat. from [2]	0.56	0.60	0.66	0.53
	Feat. from [21]	0.60	0.62	0.75	0.45
	Average	0.58 (0.02)	0.61 (0.01)	0.76 (0.09)	0.41 (0.11)
RF	-	0.61 (0.03)	0.57 (0.03)	0.63 (0.03)*	0.51 (0.07)
1D-CNN	p+TTA	0.60 (0.01)	0.63 (0.01)	0.67 (0.04)	0.55 (0.04)*
	p+F	**0.76 (0.01)***	**0.79 (0.03)***	**0.83 (0.03)**	**0.72 (0.04)***
	p+F+TTA	0.74 (0.01)*	0.72 (0.03)*	0.73 (0.06)	0.70 (0.04)*

specific and dataset-specific characteristics. The 997 cm^{-1} peak, identified via the three-step selection, is notable for its presence across biological tissues (bacteria, brain, prostate) [10,13,15,146] and for achieving an AUC comparable to SVM, despite a non-significant class intensity difference ($p = 0.075$).

The limited number of samples —especially healthy tissue— compromises specificity and underscores the need for robust sampling. The 0.61 accuracy of SVM models is lower than the accuracy reported in prior studies (0.72 [2], 0.82 [21]), likely due to this study's smaller cohort ($n = 10$ vs. $n = 19$ and $n = 84$). These results highlight the need for accurate diagnostic models using limited datasets, with automatic feature extraction.

Initializing the model with bacterial spectra (pB) reduced convergence time and inter-run variability. Multi-organ pre-training (pC) better prepared the model for the molecular signatures present in the target dataset, giving access to > 200 patients and $> 3,000$ healthy spectra. High-quality brain and breast spectra, with lower autofluorescence and superior cancer detection rates [10, 87], further improved robustness. Training on diverse datasets also helps prevent overfitting [104, 147].

Once the source model was pre-trained (p), adaptation to prospective PCa data can follow two distinct strategies. If data collection is ongoing and GTs are unavailable, unsupervised adaptation via TTA is the preferred option. Alternatively, if a small labeled dataset is available, fine-tuning becomes feasible. In both cases, without adequate pre-training data, adaptation would not be possible.

The proposed model for the first scenario (CNN+p+TTA) achieves comparable accuracy and sensitivity, while outperforming the baseline SVM in specificity (0.55(0.04) vs. 0.41(0.11),

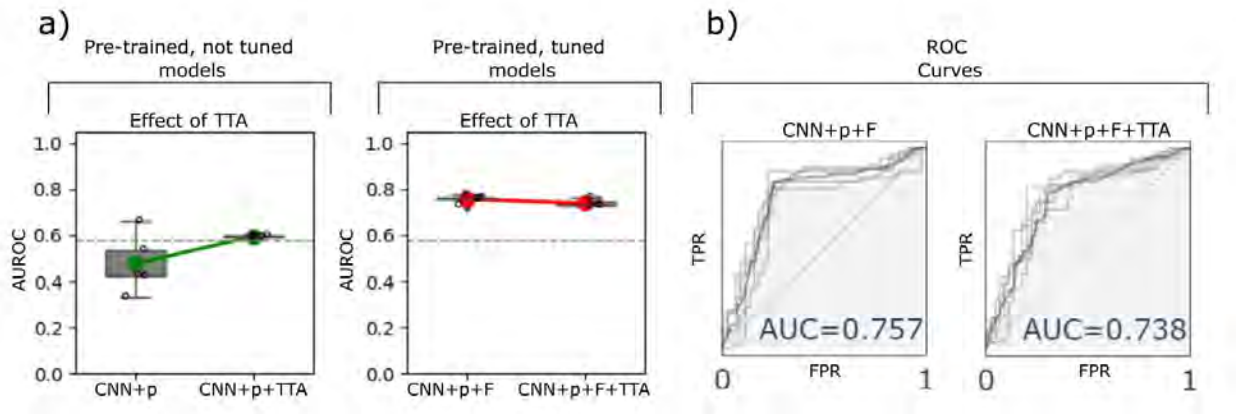


Figure 8.5 Comparison of AUC plots for the proposed models. a) Pairwise comparison on adaptation, 5 repetitions for all models. Boxes show the interquartile range (25%–75%), while Box-whiskers show minimum and maximum values within 1.5 times the interquartile range. The horizontal dotted lines indicate the mean AUC value for the SVM baseline model. b) ROC curve for the proposed models. F: fine-tuning; p: Bacteria pre-training + Cancers pre-training.

$CI = [0.022, 0.286)$. Notably, the SVM required training on data from 9 patients, whereas CNN+p+TTA adapted without prior exposure to the target dataset and without manual feature selection. This unsupervised adaptation is computationally efficient, enabling real-time implementation during data collection, with 5 fitting steps and final prediction completed in less than 5s. While TTA has been successfully applied to imaging [107] and EEG signals [108], to our knowledge, its application in Raman Spectroscopy (RS) for cancer detection is novel.

For cases where labeled data are available, the fine-tuned model (CNN+p+F) demonstrates a significant improvement in AUC (0.76 (0.01) vs. 0.58 (0.02) $p << 0.001$), accuracy ($p << 0.001$), and specificity ($p = 0.003$). It also reduces inter-run variability (lower SD in all metrics except accuracy compared to the SVM). Specificity was particularly problematic for models without fine-tuning, with SVM exhibiting good sensitivity at the cost of specificity. These results yield a positive predictive value of 0.76 (0.03), which in intraoperative settings could support decision-making and personalized treatment adjustments.

Our fine-tuning approach was based on the hypothesis that the initial 1D-CNN layers act as robust feature extractors, initialized from bacterial (∼60,000 samples) and cancer (∼4,600 samples) RS datasets. By freezing these lower layers and adapting only the later layers to the limited target prostate dataset, this strategy preserves the general spectral features, mitigates overfitting, and reduces computational time (−54min per 5 repetitions) [16], while effectively adapting to the PCa detection task.

In computer vision, pre-training on different domains is primarily used for feature extraction rather than classification [107, 147]. Similarly, in this study, despite differences in biological origin, bacterial and cancerous tissue spectra share fundamental Raman characteristics. This enables the CNN to learn relevant spectral representations during pre-training, which can be effectively refined for PCa detection.

The CNN+p+F model achieved a sensitivity of 0.83, representing a 13% improvement over the false-negative rates of MRI-guided biopsies [21]. Notably, despite being implemented with only 10 patients, its performance is comparable to or better than SVM-based studies using eight times larger cohorts (e.g., 0.72 sensitivity with SVM in [21]). Furthermore, the model's sensitivity and specificity align with those reported in other organ studies, such as breast cancer, where CNN models have achieved a sensitivity of 0.86 and specificity of 0.75, despite the use of significantly larger and higher-quality datasets [13].

The TTA approach improved performance when applied to a pre-trained —but not fine-tuned— model (8.5), helping mitigate distribution shifts. However, after fine-tuning, TTA reduced the overall performance. Since fine-tuning had already aligned the model to the target PCa data using a supervised objective (binary cross-entropy), applying TTA modified the feature space without supervision (entropy minimization), potentially causing the model to discard relevant learned features [106, 148]. When labels were available, supervised adaptation such as fine-tuning is superior [149]; TTA is most useful only when GT is unavailable, otherwise it may be unnecessary or even detrimental.

Limitations in this study relate to the small cohort size and controlled acquisition conditions. While using 10 patients allowed for the evaluation of adaptation strategies, the dataset may still under-represent interpatient PCa variability. All spectra were acquired using a single RS system at a single center, limiting the generalizability of the results. Although pre-training on multiple biobanks helps reduce overfitting to a specific cohort, future work will include multi-center data collection and system-level variability assessment, as well as integration of real-time classification into the acquisition workflow to improve signal quality and intraoperative efficiency.

8.5 Conclusion

This study demonstrated the potential of pre-trained 1D-CNN models, leveraging multi-organ cancer spectra, to enhance RS-based PCa detection accuracy despite the limited target domain sample size. When GT is unavailable, TTA provides results comparable to supervised SVM's; however, when GT is available, supervised fine-tuning improves cancer confirma-

tion. By integrating robot-assisted RS and efficient fine-tuning, our approach addresses critical challenges in PCa diagnosis, including accurate labeling and limited specific data. The proposed robotic-assisted workflow streamlines tissue analysis and supports the clinical translation of RS-based diagnostic tools for intraoperative decision-making.

8.6 Acknowledgements

This research was funded by Fonds de Recherche du Quebec, the Canada First Research Excellence Fund (TransMedTech Institute), and the Collaborative Health Research Program (NSERC and CIHR).

8.7 Declarations

Competing interests: F. Leblond is co-founder of ODS Medical (now Reveal Surgical) formed in 2015 to commercialize a RS system for neurosurgical and prostate surgery applications. The other authors declare no conflicts of interest.

Ethics approval: All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee (NCT03378856).

Consent to participate: Informed consent was obtained from all individual participants included in the studies.

Code and Data Availability: Data supporting the findings of this study are available at: https://github.com/DGdaorgralo/raman-1dcnn-pca-confirmation.git. Additional information may be obtained from the authors upon reasonable request and subject to institutional ethics approval.

CHAPTER 9 GENERAL DISCUSSION

The aim of this thesis was to develop and evaluate an AI-driven RS system for real-time, in-situ PCa confirmation to support tumor-targeted interventions during brachytherapy procedures. To achieve this, four research objectives were defined and grouped into two sequential phases. In Phase I, a prototype interventional system was evaluated, and in-vivo multimodal data (RS and radiomics) were acquired to train ML models for PCa detection. In Phase II, a robot-assisted system was implemented to enable ex-vivo spectral measurements from biopsy samples collected intraoperatively. These prospective data were used to adapt DL models, leveraging transfer learning from retrospectively acquired datasets from other organs and optical systems to compensate for the limited sample size. This chapter discusses the contributions, challenges, and implications of each phase of this feasibility study, along with the overall impact and future directions of this work.

9.1 Phase I: In-vivo experiments

Prototype navigation system evaluation

The initial part of this phase focused on evaluating a prototype interventional system developed by Philips Disease Management Solutions, intended to support HDR brachytherapy for PCa. Clinical evaluation under standard operating conditions demonstrated successful integration into routine HDR brachytherapy workflows, with procedure times (21–88 minutes) comparable to conventional TRUS-guided treatments (85–120 minutes, including up to 20 min of set-up). The system required minimal training, facilitating clinical adoption despite its underlying complexity. Despite the risk of interference with equipment in the PR, including cell phones, the EM-based catheter reconstruction improved confidence, especially for catheters that were only partially visible on TRUS. Although systematic anterior-posterior displacements were observed —likely due to catheter bending— most manual corrections remained within clinically acceptable tolerances. Nonetheless, image-based validation and system calibration remained necessary. Surface-based deformable registration showed slightly better accuracy than rigid alignment, with mTRE values within previous reported ranges (2.1–3.7 mm); residual uncertainties remain and must be taken into account when defining the PTV. Urethral midline 3D distance served as a robust internal validation metric, enabling anatomically meaningful performance assessments; achieved distances were within the accuracy range reported for the system (2–3 mm). The urethra is centrally located, visible across imaging modalities, and clinically relevant in HDR brachytherapy. Its continuous and

elongated structure allows the use of multiple reference points along the prostate, improving spatial coverage and reducing the variability associated with difficult-to-identify single-point landmarks.

Beyond standard performance evaluation, we also examined the influence of volume initialization for the deformable registration. Pelvic rotation differences between MRI and TRUS acquisitions can introduce alignment errors if uncorrected. While no significant improvements were observed in surface-based metrics (DSC and Hausdorff distances, consistent with the expectations for a surface-based registration method), pre-aligning the urethral angle —either manually or automatically— improved urethral midline agreement by up to 33% using the same deformable algorithm. These image fusion methods typically perform automatic volume centering but do not include alignment, which is therefore done manually. These findings support the implementation of an automated urethral alignment step prior to deformable registration, as proposed in Chapter 5 With this, we fulfilled *sub-objective 1.1: "Evaluate the performance of a prototype interventional system (UroNav) in the early phase of clinical evaluation in 43 consecutive interventions, by assessing workflow efficacy, EM catheter reconstruction accuracy, and MRI-TRUS registration accuracy"*. This evaluation was crucial not only to assess the system's clinical integration and technical reliability but also because its fused images and EM-tracked hardware form the foundation of the RS navigation workflow used later in this study.

Multimodal classification model for PCa

The second component of Phase I assessed the feasibility of in-vivo RS for PCa detection, integrated into a multimodal classification framework alongside radiomics from preoperative mpMRI. Between September 2020 and August 2021, optical measurements were acquired using a custom navigation system that guided the RS probe to specific intra-prostatic sites and tracked its position in real time relative to both the patient and preoperative medical images. The optical measurements added less than 20 minutes to the standard brachytherapy procedure. This enabled the completion of *sub-objective 2.1: "Acquire RS and mpMRI-based radiomics data from 18 patients undergoing HDR brachytherapy, using an in-house developed navigation system"*, resulting in a dataset with RS spectra, radiomics features, and histopathological confirmation.

Despite a limited dataset (18 patients), RS models based on FP or FP + HW regions exceeded 0.70 accuracy using fewer than 10 features. Radiomics-based models also performed competitively using simple, intensity- and texture-based metrics, avoiding complex preprocessing such as lesion segmentation —making them suitable for intraoperative use. These results addressed *sub-objective 2.2: "Train SVM models using features independently extracted from*

RS and radiomics to evaluate their individual and combined discriminative capabilities". The best results were achieved when combining optical (mesoscopic) and radiomics (macroscopic) features. The FP + Rad model was the only one to exceed 0.80 in both sensitivity and specificity, satisfying *sub-objective 2.3: "Optimize model complexity by reducing the number of input features to 10 while maintaining a classification accuracy of at least 0.80"*. These results underline the complementarity of RS and MRI-based information, which may enhance generalizability in small datasets by capturing distinct biological and imaging characteristics of PCa.

Most frequently selected Raman shifts were linked to benign tissue markers such as phenylalanine (990–1015 cm^{-1}), nucleic acids and collagen (1330–1350 cm^{-1}), and phospholipids (1760–1790 cm^{-1}). Radiomics features like energy from ADC images showed expected contrasts between benign and malignant tissues, supporting the biological relevance of selected features. The applied three-step feature selection strategy (amount of variation > 0.03; correlation with the target > 0.9, Lasso regression for a maximum of 10 features) yields a subset of features that allows for a degree of interpretability of the results. While our study primarily emphasized ADC-derived radiomics features (based on feature selection), mpMRI provides complementary information through its other sequences, each reflecting distinct aspects of tumor biology. As described in Section 3.1.1, T2-weighted imaging captures fine anatomical and organizational detail, DWI reflects microstructural and cellular density changes, and DCE highlights vascular alterations associated with angiogenesis. Integrating radiomics features from these complementary sequences, beyond ADC, with RS could allow a more comprehensive characterization of prostate cancer, capturing both microstructural and biochemical alterations in future studies. This part of the study demonstrated the feasibility of combining in-vivo RS and mpMRI-derived radiomics for PCa classification and laid the foundation for real-time, minimally invasive diagnostic tools. It is important to note that, in this case, the GT (histopathology) corresponds to the core-level diagnosis, with only one RS measurement associated with each core. Future work should aim to improve GT alignment and integrate classification into the navigation system.

9.2 Phase II: Ex-vivo experiments

Robot-assisted RS acquisition system and CNN exploration

The first component of this phase developed a semi-automated ex-vivo workflow for RS-based PCa confirmation, integrating the MECA500 robotic arm with DL-based tissue classification. The robot enabled fast, repeatable, and spatially precise Raman measurements on fresh biopsy samples, with minimal disruption to the intervention and no significant delay to

formalin fixation —addressing clinical concerns over pathology timelines.

To ensure targeting accuracy, different camera-robot calibration strategies were evaluated. In constrained cases, where few fiducials are available/visible for calibration, homography was more robust even with 4–6 fiducials. Both homography and conic calibration performed well with $\geq$14 fiducials. The $\sim$25% sample-size spatial error surpassed manual handling and enabled precise applications. In certain scenarios (not covered in this thesis), there may be research interest in applying additional procedures at the exact location where RS is acquired. These may include signal-enhancing techniques (e.g., photobleaching to reduce autofluorescence) or steps necessary for ground truth assignment (e.g., ink marking the location of the measurement). The robot enabled flexible acquisition strategies —ranging from single-site measurements to more complex workflows, e.g., RS + photobleaching + RS— with notable time savings when used repetitively. In such scenarios, delays were reduced by over two minutes, helping preserve tissue integrity for subsequent histopathology. With these tests, we achieved *sub-objective 3.1. "Integrate a robot-assisted RS system for automated spectral acquisition from fresh biopsy cores, targeting a positioning accuracy <0.4 mm, and evaluate its repeatability and acquisition time".*

Additionally, in Chapter 7, a preliminary investigation of CNN models for cancer detection using RS was also conducted. Classification using 1D-CNNs across retrospective brain, breast, and prostate RS datasets showed better performance for brain and breast compared to traditional SVM. While CNN and SVM performed similarly for PCa, likely due to the small cohort and challenging tissue properties (e.g., high autofluorescence, weak signal), the CNN's ability to learn spectral features automatically —without hand-crafted input— was a key advantage. The use of a single architecture across all tissue types also demonstrated the model's versatility when sufficient data is available. At this point, *sub-objective 3.2 "Conduct preliminary evaluations of CNN models in retrospective RS datasets (prostate, breast, and brain cancer), assessing their versatility and automatic feature extraction capabilities through binary classification performance, aiming to achieve accuracies >0.72, >0.90, and >0.78, respectively."* is completed, which serves as an inspiration for the following tests.

Transfer learning strategies for intraoperative PCa classification

The second component of Phase II focused on ex-vivo data acquisition and evaluating transfer learning strategies to enable accurate RS-based cancer detection from fresh prostate biopsies. Between April and December 2024, a prospective dataset was collected in the PR using the robot-assisted RS system, completing *sub-objective 4.1 "Acquire RS data from biopsy samples of 10 patients undergoing HDR brachytherapy using the robot-assisted RS system".* The robotic arm enabled consistent, high-precision measurements across biopsy cores from

10 patients, enabling the assignment of site-specific labels (instead of core-level diagnosis) to Raman spectra with minimal disruption of therapy workflow.

A general limitation in DL applications is data availability —104 spectra from 10 patients are insufficient to train a deep model with six residual blocks. To address this, transfer learning was employed by pre-training on large, multi-organ RS datasets (collected from different organs, optical systems, and centers), including prostate, brain, breast, and bacterial spectra. Pre-training improved convergence, robustness, and generalizability: cancerous tissue data provided relevant spectral features, while bacterial spectra contributed contextual initialization and diversity. Models were then adapted to the prospective prostate dataset using supervised fine-tuning and unsupervised TTA. Fine-tuned models achieved the highest performance, surpassing the 0.72 classification accuracy achieved in previous studies (with more patients and simpler models) and significantly outperforming baseline SVMs in specificity and AUC.

TTA was evaluated as a real-time, unsupervised adaptation strategy when ground truth labels are unavailable. While it improved just pre-trained models, TTA was less effective when applied after fine-tuning, indicating that supervised objectives better align the model to the target domain when labels are present. With these results, *sub-objective 4.2* was fulfilled: *"Implement and optimize transfer learning strategies leveraging retrospective multi-organ RS datasets, with the aim of achieving classification accuracy of over 0.72 after adaptation to a limited intraoperative prostate dataset"*. Importantly, the fine-tuned CNN model matched or exceeded prior PCa results obtained with larger cohorts, underscoring the feasibility of accurate, real-time RS-based cancer detection through effective transfer learning and robotic standardization.

9.3 In-vivo vs ex-vivo approaches for RS-based PCa detection

Besides the performance evaluation of a prototype navigation system (Phase I, Chapter 5) and the implementation of a robot-assisted RS system (Phase II, Chapter 7), Phases I and II of this study explored two different approaches for in-situ PCa confirmation using RS: in-vivo measurements acquired directly inside the prostate during HDR brachytherapy (Phase I, Chapter 6), and ex-vivo measurements performed on fresh biopsy cores collected during the same type of treatment (Phase II, Chapter 8). Both approaches have distinct advantages and limitations related to clinical integration, data quality, and process flexibility.

In-vivo (Phase I)

In-vivo measurements offer immediate access to the prostate tissue without the need for

sample handling or preparation. This direct, intraprostatic contact provides cleaner spectral signals, with reduced risk of contamination from sample holders or air gaps (although bleeding can hinder). Additionally, real-time co-registration with ultrasound and preoperative MRI allows precise localization of the probe within the prostate gland, which is not possible in ex-vivo measurements. Probe stability is less critical in this setting, as the relatively large organ and surrounding tissue buffer slight movements. Moreover, measurements are performed at the exact anatomical site of interest, which enhances spatial accuracy. However, in-vivo inspection introduces significant logistical and procedural constraints. Sterile conditions must be maintained, limiting who can handle the equipment. The measurement process requires temporarily pausing the treatment workflow, introducing potential delays. Due to sterility and safety constraints, additional processing steps such as photobleaching are not feasible. Only one measurement can be acquired per confirmation biopsy site (limited to core-level diagnosis), reducing the ability to explore intra-tumoral heterogeneity.

Ex-vivo (Phase II)

The ex-vivo approach offers greater flexibility and is more suitable for early-stage development, data collection, and model validation. Measurements can be performed in parallel with ongoing treatment, minimizing disruption to clinical workflow. Sterility is not required, allowing the RS system to be operated by any trained personnel or robotic platforms. Most importantly, multiple measurements can be acquired from each biopsy sample, enabling better sampling of tissue heterogeneity, allowing site-specific GT, and improving training data for DL models. Additional processing steps, such as photobleaching or inking, can also be applied without risk to the patient. However, ex-vivo measurements are sensitive to sample handling and positioning. Probe-to-sample contact must be precise to avoid spectral contamination from the support surface. The biopsy site is identifiable in the TRUS images, but the RS probe location relative to the prostate anatomy is not directly visible. Furthermore, a short waiting time (under 3 minutes) is required between sampling and classification, although model inference itself takes less than 1 second.

Synthesis and Recommendation

From a technical and procedural standpoint, ex-vivo RS offers the most practical and scalable framework for data acquisition and model training. Its minimal impact on treatment flow, flexibility in measurement conditions, and potential for rich, heterogeneous data make it ideal for iterative development and validation of RS-based diagnostic tools. In needle-based interventions, there is typically no clinical incentive to avoid tissue sampling. As a non-destructive technique, RS can be applied directly to biopsy cores ex-vivo, providing rapid cancer confirmation while preserving the sample for comprehensive histopathological analysis, which

remains essential for detailed characterization. Once robust models have been developed and validated using ex-vivo data, the in-vivo workflow becomes a viable option for seamless, localized tumor confirmation during treatment planning. For example, RS confirmation — whether through biopsy or direct measurement— could be performed at catheter insertion sites, offering a faster (even with sample handling) alternative to current histopathological timelines.

CHAPTER 10 CONCLUSION

10.1 Summary of work and advancement of knowledge

This thesis explored the integration of RS and DL methods for real-time PCa detection in the PR. The research was structured into two main phases: an in-vivo phase focused on real-time guidance, multimodal data acquisition during brachytherapy, and ML-based classification (Chapters 5 and 6); and an ex-vivo phase centered on DL-based classification of RS signals acquired from fresh prostate biopsy cores using a robot-assisted system (Chapters 7 and 8).

In Phase I, we evaluated a prototype intervention system developed by Philips in its early clinical testing stages and deployed an in-house navigation platform enabling co-registered RS acquisition during prostate HDR brachytherapy. This work pioneered the combination of EM tracking with TRUS-MRI fusion to navigate the RS probe. A key milestone was the execution of the first in-vivo RS measurements for PCa detection (distinct from prior studies focused on prostatic/extraprostatic tissue discrimination), laying the foundation for real-time optical feedback during minimally invasive PCa therapy. Although sample numbers were limited —as expected in early-phase studies— SVM models achieved promising results in multiple binary classification tasks by combining RS and mpMRI-radiomics features.

In Phase II, we investigated DL-based classification of ex-vivo RS spectra acquired from fresh prostate biopsy specimens. To enable this, we implemented a robot-assisted RS system directly in the PR. This system demonstrated high positioning accuracy, repeatability, and reduced characterization time, allowing for more precise, site-specific histopathological labeling. By overcoming the well-known limitation of core-level ground truth, this represents a key advancement in ground-truth reliability, directly impacting model training and the clinical relevance of RS-based classification.

We explored 1D-CNN models, including both shallow architectures and deeper residual block-based networks. Although the size of the prospective dataset was limited, we employed transfer learning strategies using retrospective multi-organ RS datasets for model pre-training. Through supervised and unsupervised adaptation, we achieved performance surpassing traditional SVMs in the prospective PCa dataset, with supervised methods reaching accuracy levels comparable to prior multimodal techniques. The deployment of these models with the robotic system enabled real-time tissue classification in the PR within 3 minutes of sample collection. This demonstrated not only the feasibility of rapid intraoperative DL-based PCa

confirmation, but also, to our knowledge, one of the first uses of test-time adaptation (TTA) and real-time classification directly in the PR.

10.2 Limitations

While this thesis demonstrates the feasibility of RS and DL integration for real-time PCa detection, several limitations must be acknowledged:

- **Small cohort sizes:** Both phases involved limited patient numbers (n=18 in-vivo, n=10 ex-vivo), restricting model complexity, statistical power, and generalizability. In IA applications, the models are only as good as the data available, and the scarcity of data is a known problem. A broad data acquisition, with more patients and more samples, is fundamental to extend the validation of the proposed methods.
- **Single-center data and system variability:** All ex-vivo spectra were collected using the same RS system in a controlled setting, which may limit generalizability. Future work should assess system-level variability across sites.
- **Ground-truth co-localization:** In-vivo RS acquisitions relied on EM tracking and TRUS-MRI fusion, but spatial uncertainties may have affected the accuracy of histopathological labels. Improved co-registration methods or 3D histology mapping could enhance reliability. Although progress has been made in this aspect for the ex-vivo measurements, this is a limitation regarding in-vivo measurements requiring biopsy confirmation.
- **Model interpretability:** While 1D-CNNs outperformed SVMs, their reduced transparency may hinder clinical adoption. Explainability techniques could help improve trust and usability.

Despite these limitations, the findings lay essential groundwork for further development of real-time, AI-assisted optical biopsy systems in PCa care.

10.3 Future Research

To conclude this thesis, the following research directions are recommended for advancing the clinical translation of RS-based, DL-assisted PCa detection systems:

Recommendation 1: Expand the PCa-specific RS dataset across institutions and systems. The limited size and single-center origin of the current database restrict model

generalizability. Future efforts should prioritize the acquisition of larger, multi-center datasets using different RS systems and biopsy workflows, ensuring quality labeling. This would improve robustness to inter-patient and inter-system variability and strengthen the clinical applicability of the proposed framework.

Recommendation 2: Explore new adaptation strategies leveraging existing RS datasets. Building on this thesis' use of transfer learning, future studies should investigate task-specific adaptation techniques —such as few-shot learning— to exploit RS datasets from other organs or acquisition conditions, in a more efficient way. These strategies, commonly used in computer vision, could reduce the need for large prostate-specific datasets and support broader deployment, mainly in early-stage studies.

Recommendation 3: Improve model interpretability to support clinical translation. The limited transparency of CNN-based classifiers remains a challenge for clinical deployment. Future research should explore techniques such as class activation mapping (CAM), attention mechanisms, or post hoc explanation tools to better understand what the models learn. Hybrid approaches combining handcrafted features with deep learning may also help bridge the gap between performance and interpretability.

These future directions aim to reinforce the clinical impact of this work and address the technical and translational challenges of deploying AI-assisted optical biopsy systems in real-world PCa interventions.

REFERENCES

[1] F. Gaba *et al.*, "Raman spectroscopy in prostate cancer: techniques, applications and advancements," *Cancers*, vol. 14, no. 6, p. 1535, 2022.

[2] D. Grajales *et al.*, "Image-guided raman spectroscopy navigation system to improve transperineal prostate cancer detection. part 2: In-vivo tumor-targeting using a classification model combining spectral and mri-radiomics features," *Journal of Biomedical Optics*, vol. 27, no. 9, pp. 095 004–095 004, 2022.

[3] R. B. Shah and M. Zhou, *Prostate Needle Biopsy Sampling Techniques: Impact on Pathological Diagnosis*. Springer Berlin Heidelberg, 2012, pp. 11–14. [Online]. Available: https://doi.org/10.1007/978-3-642-21369-4_2

[4] G. Sheehy *et al.*, "Open-sourced Raman spectroscopy data processing package implementing a novel baseline removal algorithm validated from multiple datasets acquired in human tissue and biofluids," *Journal of Biomedical Optics*, vol. 28, no. February, pp. 1–20, 2023.

[5] E. Poulin *et al.*, "Validation of mri to trus registration for high-dose-rate prostate brachytherapy," *Brachytherapy*, vol. 17, pp. 283–290, 2018, recently, several articles (31e35) were pub- lished on deformable registration with TREs ranging from 2.4 to 3.4 mm,.

[6] L. Beaulieu *et al.*, "Real-time electromagnetic tracking–based treatment platform for high-dose-rate prostate brachytherapy: Clinical workflows and end-to-end validation," *Brachytherapy*, vol. 17, pp. 103–110, 2018. [Online]. Available: https://doi.org/10.1016/j.brachy.2017.04.247

[7] W. Zong *et al.*, "Prostate cancer malignancy detection and localization from mpmri using auto-deep learning as one step closer to clinical utilization," *Scientific Reports*, vol. 12, 12 2022.

[8] J. Wang *et al.*, "Machine learning-based analysis of mr radiomics can help to improve the diagnostic performance of pi-rads v2 in clinically relevant prostate cancer," *European Radiology*, vol. 27, pp. 4082–4090, 2017.

[9] E. Cordero *et al.*, "In-vivo Raman spectroscopy: from basics to applications," *Journal of Biomedical Optics*, vol. 23, no. 07, p. 1, 2018.

[10] K. Ember *et al.*, "In situ brain tumor detection using a raman spectroscopy system—results of a multicenter study," *Scientific Reports*, vol. 14, pp. 1–12, 2024. [Online]. Available: https://doi.org/10.1038/s41598-024-62543-9

[11] K. E. A. Aubertin *et al.*, "Combining high wavenumber and fingerprint raman spectroscopy for the detection of prostate cancer during radical prostatectomy," *Biomedical Optics Express*, vol. 9, pp. 4484–4489, 2018.

[12] J. Nayak, B. Naik, and H. S. Behera, "A comprehensive survey on support vector machine in data mining tasks: Applications challenges," *International Journal of Database Theory and Application*, vol. 8, pp. 169–186, 2015.

[13] D. Ma *et al.*, "Classifying breast cancer tissue by Raman spectroscopy with one-dimensional convolutional neural network," *Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy*, vol. 256, 2021.

[14] M. Zhou *et al.*, "An end-to-end deep learning approach for raman spectroscopy classification," *Journal of Chemometrics*, vol. 37, no. 2, p. e3464, 2023.

[15] C. S. Ho *et al.*, "Rapid identification of pathogenic bacteria using Raman spectroscopy and deep learning," *Nature Communications*, vol. 10, no. 1, 2019.

[16] K. Zhang *et al.*, "Adaptive transfer learning for EEG motor imagery classification with deep Convolutional Neural Network," *Neural Networks*, vol. 136, pp. 1–10, 2021.

[17] F. Bray *et al.*, "Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: A Cancer Journal for Clinicians*, vol. 74, no. 3, pp. 229–263, 2024.

[18] J. Streicher *et al.*, "A review of optimal prostate biopsy: indications and techniques," *Therapeutic Advances in Urology*, vol. 11, pp. 1–8, 2019.

[19] M. H. Dornan *et al.*, "Simplified and robust one-step radiosynthesis of [18f]dcfpyl via direct radiofluorination and cartridge-based purification," *Journal of Labelled Compounds and Radiopharmaceuticals*, vol. 61, pp. 757–763, 2018.

[20] P. Rawla, "Epidemiology of Prostate Cancer," *World journal of oncology*, vol. 10, no. 2, pp. 63–89, apr 2019. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/31068988https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6497009/

[21] S. J. van Breugel *et al.*, "Raman spectroscopy system for real-time diagnosis of clinically significant prostate cancer tissue," *Journal of Biophotonics*, vol. 16, no. 5, p. e202200334, 2023.

[22] M. A. Bjurlin, J. S. Wysock, and S. S. Taneja, "Optimization of prostate biopsy: Review of technique and complications," *Urologic Clinics of North America*, vol. 41, pp. 299–313, 2014, false-negative.

[23] K. Aubertin *et al.*, "Mesoscopic characterization of prostate cancer using raman spectroscopy: potential for diagnostics and therapeutics," *BJU International*, vol. 122, pp. 326–336, 8 2018, doi: 10.1111/bju.14199. [Online]. Available: https://doi.org/10.1111/bju.14199

[24] S. Kadoury *et al.*, "Realtime trus/mri fusion targeted-biopsy for prostate cancer: A clinical demonstration of increased positive biopsy rates," *Prostate Cancer Imaging. Computer-Aided Diagnosis, Prognosis, and Intervention*, vol. 6367, pp. 52–62, 2010.

[25] G. Paner *et al.*, "Essential updates in grading, morphotyping, reporting, and staging of prostate carcinoma for general surgical pathologists," *Archives of Pathology Laboratory Medicine*, vol. 143, 3 2019.

[26] D. Milonas *et al.*, "Impact of grade groups on prostate cancer-specific and other-cause mortality: Competing risk analysis from a large single institution series," 2021.

[27] T. Mee *et al.*, "The use of radiotherapy, surgery and chemotherapy in the curative treatment of cancer: results from the forty (favourable outcomes from radiotherapy) project," *British Journal of Radiology*, vol. 96, 11 2023.

[28] S. N. Pentyala *et al.*, "Prostate cancer: a comprehensive review," *REVIEW Medical Oncology*, vol. 17, pp. 85–105, 2000.

[29] K. Bouchelouche and P. L. Choyke, "Psma pet in prostate cancer – a step towards personalized medicine." *Current opinion in oncology*, vol. 28, pp. 216–221, 2016.

[30] R. Chelluri *et al.*, "Prostate cancer diagnosis on repeat magnetic resonance imaging-transrectal ultrasound fusion biopsy of benign lesions: Recommendations for repeat sampling," *Journal of Urology*, vol. 196, pp. 62–67, 2016, method. [Online]. Available: http://dx.doi.org/10.1016/j.juro.2016.02.066

[31] P. Blanchard, C. Ménard, and S. J. Frank, "Clinical use of magnetic resonance imaging across the prostate brachytherapy workflow," *Brachytherapy*, vol. 16, pp. 734–742, 2017.

[32] W. Venderink *et al.*, "Multiparametric magnetic resonance imaging for the detection of clinically significant prostate cancer: What urologists need to know. part 3: Targeted biopsy," *European Urology*, vol. 77, pp. 481–490, 2020. [Online]. Available: https://doi.org/10.1016/j.eururo.2019.10.009

[33] S. Kadoury *et al.*, "Accuracy assessment of an automatic image-based pet/ct registration for ultrasound-guided biopsies and ablations," in *Medical Imaging 2011: Visualization, Image-Guided Procedures, and Modeling*, vol. 7964, 2011, pp. 818–826.

[34] T. Yang *et al.*, "The predictive efficacy of hypoechoic lesion in ultrasound for prostate cancer in chinese people: Five-year experience in a moderated 10-core transperineal prostate biopsy procedure," *Oncotarget*, vol. 8, pp. 79 433–79 440, 2017, hypoechoic.

[35] Y. Zhang *et al.*, "Recent advances in registration methods for mri-trus fusion image-guided interventions of prostate," *Recent Patents on Engineering*, vol. 11, pp. 115–124, 2017, edwing. [Online]. Available: http://dx.doi.org/10.2174/1872212110666161201115248

[36] S. W. Dutta *et al.*, "Prostate cancer high dose-rate brachytherapy: review of evidence and current perspectives," *Expert Review of Medical Devices*, vol. 15, pp. 71–79, 2018. [Online]. Available: https://doi.org/10.1080/17434440.2018.1419058

[37] P. Dell'Oglio *et al.*, "There is no way to avoid systematic prostate biopsies in addition to multiparametric magnetic resonance imaging targeted biopsies," *European urology oncology*, vol. 3, pp. 112–118, 2 2020.

[38] H. Zhang *et al.*, "Building a high-dose-rate prostate brachytherapy program with real-time ultrasound-based planning: Initial safety, quality, and outcome results," *Advances in Radiation Oncology*, vol. 5, pp. 388–395, 2020. [Online]. Available: https://doi.org/10.1016/j.adro.2020.02.002

[39] L. Burgess *et al.*, "A review on the current treatment paradigm in high-risk prostate cancer," 9 2021.

[40] L. G. Kerkmeijer *et al.*, "Focal boost to the intraprostatic tumor in external beam radiotherapy for patients with localized prostate cancer: Results from the flame randomized phase iii trial," *Journal of Clinical Oncology*, vol. 39, pp. 787–796, 3 2021.

[41] O. Lauche *et al.*, "Single-fraction high-dose-rate brachytherapy using real-time transrectal ultrasound based planning in combination with external beam radiotherapy

for prostate cancer: Dosimetrics and early clinical results," *Journal of Contemporary Brachytherapy*, vol. 8, pp. 104–109, 2016.

[42] S. Kadoury *et al.*, "Pp-0158 registration accuracy of an integrated mr -trus navigation system for prostate hdr brachytherapy," *Radiotherapy and Oncology*, vol. 158, pp. S122–S123, 2021. [Online]. Available: https://doi.org/10.1016/S0167-8140(21)06450-1

[43] A. Gomez-Iturriaga *et al.*, "Impact of intraoperative mri/trus fusion on dosimetric parameters in ct3a prostate cancer patients treated with high-dose-rate real-time brachytherapy," *Journal of Contemporary Brachytherapy*, vol. 6, pp. 154–160, 2014.

[44] R. Stoyanova *et al.*, "Prostate cancer radiomics and the promise of radiogenomics," *Translational Cancer Research*, vol. 5, pp. 432–447, 2016, useful for ground truth discussion.

[45] J. Griethuysen *et al.*, "Computational radiomics system to decode the radiographic phenotype," *Cancer Research*, vol. 77, pp. 1–8, 2018.

[46] S. Yao, H. Jiang, and B. Song, "Radiomics in prostate cancer: basic concepts and current state-of-the-art," *Chinese Journal of Academic Radiology*, vol. 2, pp. 47–55, 2020. [Online]. Available: https://doi.org/10.1007/s42058-019-00020-3

[47] T. Chen *et al.*, "Prostate cancer differentiation and aggressiveness: Assessment with a radiomic-based model vs. pi-rads v2," *Journal of Magnetic Resonance Imaging*, vol. 49, pp. 875–884, 2019.

[48] M. Schwier *et al.*, "Repeatability of multiparametric prostate mri radiomics features," *Scientific Reports*, vol. 9, pp. 1–16, 2019.

[49] H. McGregor *et al.*, "Clinical utility of raman spectroscopy: current applications and ongoing developments," *Advanced Health Care Technologies*, p. 13, 2016, quiza para raman theory.

[50] P. M. Conforti *et al.*, "Raman Spectroscopy and AI Applications in Cancer Grading: An Overview," *IEEE Access*, vol. 12, no. March, pp. 54 816–54 852, 2024.

[51] S. Devpura *et al.*, "Vision 20/20: The role of Raman spectroscopy in early stage cancer detection and feasibility for application in radiation therapy response assessment," *Medical Physics*, vol. 41, no. 5, 2014.

[52] A. C. S. Talari *et al.*, "Raman spectroscopy of biological tissues," *Applied Spectroscopy Reviews*, vol. 50, pp. 46–111, 2015.

[53] W. J. van de Ven *et al.*, "Mr-targeted trus prostate biopsy using local reference augmentation: initial experience," *International Urology and Nephrology*, vol. 48, pp. 1037–1045, 2016.

[54] P. Zogal *et al.*, "Biopsee® - transperineal stereotactic navigated prostate biopsy," *Journal of Contemporary Brachytherapy*, vol. 3, pp. 91–95, 2011.

[55] Y. Hu *et al.*, "Mr to ultrasound image registration for guiding prostate biopsy and interventions," *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2009*, vol. 5761, pp. 787–794, 2009.

[56] R. Sparks *et al.*, "Fully automated prostate magnetic resonance imaging and transrectal ultrasound fusion via a probabilistic registration metric," in *Medical Imaging 2013: Image-Guided Procedures, Robotic Interventions, and Modeling*, vol. 8671, 3 2013, p. 86710A, From Duplicate 1 (Fully automated prostate magnetic resonance imaging and transrectal ultrasound fusion via a probabilistic registration metric - Sparks, Rachel; Bloch, B. Nicholas; Feleppa, Ernest; Barratt, Dean; Madabhushi, Anant)

RMSE
texture features were. [Online]. Available: https://doi.org/10.1117/12.2007610

[57] J. K. Logan *et al.*, "Current status of magnetic resonance imaging (mri) and ultrasonography fusion software platforms for guidance of prostate biopsies," *BJU International*, vol. 114, pp. 641–652, 2014.

[58] J. M. Bugeja *et al.*, "Prostate volume analysis in image registration for prostate cancer care: a verification study," *Physical and Engineering Sciences in Medicine*, vol. 46, pp. 1791–1802, 12 2023.

[59] C. Kirisits *et al.*, "Review of clinical brachytherapy uncertainties: Analysis guidelines of gec-estro and the aapm," *Radiotherapy and Oncology*, vol. 110, pp. 199–212, 2014. [Online]. Available: http://dx.doi.org/10.1016/j.radonc.2013.11.002

[60] A. M. Franz *et al.*, "Electromagnetic tracking in medicine -a review of technology, validation, and applications," *IEEE Transactions on Medical Imaging*, vol. 33, pp. 1702–1725, 2014.

[61] S. Bharat *et al.*, "Electromagnetic tracking for catheter reconstruction in ultrasound-guided high-dose-rate brachytherapy of the prostate," *Brachytherapy*, vol. 13, pp. 640–650, 2014. [Online]. Available: http://dx.doi.org/10.1016/j.brachy.2014.05.012

[62] M. He *et al.*, "Research progress on deep learning in magnetic resonance imaging–based diagnosis and treatment of prostate cancer: a review on the current status and perspectives," 2023.

[63] H. Sherafatmandjoo *et al.*, "Prostate cancer diagnosis based on multi-parametric mri, clinical and pathological factors using deep learning," *Scientific Reports*, vol. 14, 12 2024.

[64] K. C. Ma *et al.*, "Deep learning-based whole-body psma pet/ct attenuation correction utilizing pix-2-pix gan," Tech. Rep., 2024.

[65] A. Zwanenburg *et al.*, "The image biomarker standardization initiative: Standardized quantitative radiomics for high-throughput image-based phenotyping," *Radiology*, vol. 295, pp. 328–338, 5 2020.

[66] X. Min *et al.*, "Multi-parametric mri-based radiomics signature for discriminating between clinically significant and insignificant prostate cancer : Cross-validation of a machine learning method," *European Journal of Radiology*, vol. 115, pp. 16–21, 2019, clinicaly signifiant.

[67] J. Desroches *et al.*, "A new method using raman spectroscopy for in-vivo targeted brain cancer tissue biopsy," *Scientific reports*, vol. 8, no. 1, p. 1792, 2018.

[68] A. S. Haka *et al.*, "Diagnosing breast cancer using Raman spectroscopy: prospective analysis," *Journal of Biomedical Optics*, vol. 14, no. 5, p. 054023, 2009.

[69] D. Petersen *et al.*, "Raman fiber-optical method for colon cancer detection: Cross-validation and outlier identification approach," *Spectrochimica Acta - Part A*, vol. 181, pp. 270–275, 2017.

[70] J. Schleusener *et al.*, "In-vivo study for the discrimination of cancerous and normal skin using fibre probe-based Raman spectroscopy," *Experimental Dermatology*, vol. 24, no. 10, pp. 767–772, 2015.

[71] J. Desroches *et al.*, "Development and first in-human use of a Raman spectroscopy guidance system integrated with a brain biopsy needle," *Journal of Biophotonics*, vol. 12, no. 3, pp. 1–7, 2019.

[72] P. Gao *et al.*, "The Clinical Application of Raman Spectroscopy for Breast Cancer Detection," *Journal of Spectroscopy*, vol. 2017, no. 1, 2017.

[73] S. Elumalai, S. Managó, and A. C. De Luca, "Raman microscopy: Progress in research on cancer cell sensing," *Sensors (Switzerland)*, vol. 20, no. 19, pp. 1–19, 2020.

[74] M. Lin *et al.*, "Laser tweezers raman spectroscopy combined with machine learning for diagnosis of alzheimer's disease," *Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy*, vol. 280, 11 2022.

[75] H. A. Esawii *et al.*, "Investigation into alzheimer's-related amyloid- conformational transformations and stability influenced by green iron oxide nanoparticles (gionp)," *International Journal of Biological Macromolecules*, vol. 298, 4 2025.

[76] H. C. McGregor *et al.*, "Development and in vivo test of a miniature raman probe for early cancer detection in the peripheral lung," *Journal of Biophotonics*, vol. 11, p. e201800055, 11 2018, doi: 10.1002/jbio.201800055. [Online]. Available: https://doi.org/10.1002/jbio.201800055

[77] M. Stomp-Agenant *et al.*, "In vivo raman spectroscopy for bladder cancer detection using a superficial raman probe compared to a nonsuperficial raman probe," *Journal of Biophotonics*, vol. 15, pp. 1–9, 2022.

[78] M. Seong *et al.*, "Longitudinal raman spectroscopic observation of skin biochemical changes due to chemotherapeutic treatment for breast cancer in small animal model," *Journal of Spectroscopy*, vol. 2017, 2017.

[79] T. Bhattacharjee *et al.*, "Transcutaneous in vivo raman spectroscopy: Detection of age-related changes in mouse breast," *Vibrational Spectroscopy*, vol. 67, pp. 80–86, 2013. [Online]. Available: http://dx.doi.org/10.1016/j.vibspec.2013.04.004

[80] A. Verkentin *et al.*, "Trimodal system for in vivo skin cancer screening with combined oct-raman and co-localized optoacoustic measurements," *Journal of Biophotonics*, vol. 11, 2018.

[81] W. Wang *et al.*, "Real-time in vivo cancer diagnosis using raman spectroscopy," *Journal of Biophotonics*, vol. 8, pp. 527–545, 2015.

[82] A. S. Haka *et al.*, "Diagnosing breast cancer by using Raman spectroscopy," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 35, pp. 12 371–12 376, 2005.

[83] D. Lazaro-Pacheco *et al.*, "Raman spectroscopy of breast cancer," *Applied Spectroscopy Reviews*, vol. 55, no. 6, pp. 439–475, 2020.

[84] P. Kumar, "Raman spectroscopy as a promising noninvasive tool in brain cancer detection," *Journal of Innovative Optical Health Sciences*, vol. 10, pp. 1–8, 2017.

[85] J. Desroches *et al.*, "Characterization of a Raman spectroscopy probe system for intraoperative brain tissue classification," *Biomedical Optics Express*, vol. 6, no. 7, p. 2380, 2015.

[86] L. Zhang *et al.*, "Raman spectroscopy and machine learning for the classification of breast cancers," *Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy*, vol. 264, 2022.

[87] S. David *et al.*, "In-situ raman spectroscopy and machine learning unveil biomolecular alterations in invasive breast cancer," *Journal of Biomedical Optics*, vol. 28, no. 3, pp. 036 009–036 009, 2023.

[88] R. P. Aguiar *et al.*, "Discriminating neoplastic and normal brain tissues in vitro through raman spectroscopy: A principal components analysis classification model," *Photomedicine and Laser Surgery*, vol. 31, pp. 595–604, 2013.

[89] N. Iturrioz-Rodríguez *et al.*, "Discrimination of glioma patient-derived cells from healthy astrocytes by exploiting raman spectroscopy," *Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy*, vol. 269, 3 2022.

[90] M. Pinto *et al.*, "Integration of a raman spectroscopy system to a robotic-assisted surgical system for real-time tissue characterization during radical prostatectomy procedures," *Journal of Biomedical Optics*, vol. 24, no. 2, pp. 025 001–025 001, 2019.

[91] K. Milligan *et al.*, "Raman spectroscopy and supervised learning as a potential tool to identify high-dose-rate-brachytherapy induced biochemical profiles of prostate cancer," *Journal of Biophotonics*, vol. 15, pp. 1–16, 2022.

[92] P. Crow *et al.*, "The use of raman spectroscopy to identify and grade prostatic adenocarcinoma in vitro," *British Journal of Cancer*, vol. 89, pp. 106–108, 2003.

[93] L. Wang *et al.*, "Raman spectroscopy, a potential tool in diagnosis and prognosis of castration-resistant prostate cancer," *Journal of Biomedical Optics*, vol. 18, p. 087001, 2013.

[94] A. A. Grosset *et al.*, "Identification of intraductal carcinoma of the prostate on tissue specimens using raman micro-spectroscopy: A diagnostic accuracy case–control study with multicohort validation," *PLoS Medicine*, vol. 17, pp. 1–20, 2020.

[95] S. Kiranyaz *et al.*, "1d convolutional neural networks and applications: A survey," *Mechanical systems and signal processing*, vol. 151, p. 107398, 2021.

[96] F. Picot *et al.*, "Image-guided raman spectroscopy navigation system to improve transperineal prostate cancer detection. part 1: Raman spectroscopy fiber-optics system and in situ tissue characterization," *Journal of Biomedical Optics*, 2022.

[97] É. Lemoine *et al.*, "Feature engineering applied to intraoperative in-vivo raman spectroscopy sheds light on molecular processes in brain cancer: a retrospective study of 65 patients," *Analyst*, vol. 144, no. 22, pp. 6517–6532, 2019.

[98] A. Esteva *et al.*, "Prostate cancer therapy personalization via multi-modal deep learning on randomized phase iii clinical trials," *npj Digital Medicine*, vol. 5, 12 2022.

[99] Y. Arita *et al.*, "Advancements in artificial intelligence for prostate cancer: Optimizing diagnosis, treatment, and prognostic assessment," *Asian Journal of Urology*, 2 2025. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S2214388225000074

[100] E. Compérat and R. Grobholz, "Artificial intelligence: redefining the future of prostate cancer diagnostics," *Diagnostic Histopathology*, 5 2025. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S1756231725000684

[101] O. O'Connor and T. P. McVeigh, "Increasing use of artificial intelligence in genomic medicine for cancer care- the promise and potential pitfalls," *BJC Reports*, vol. 3, p. 20, 4 2025. [Online]. Available: https://www.nature.com/articles/s44276-025-00135-4

[102] A. M. Santilli *et al.*, "Perioperative margin detection in basal cell carcinoma using a deep learning framework," *International Journal of Computer Assisted Radiology and Surgery*, vol. 15, no. 5, pp. 887–896, 2020.

[103] L. Connolly *et al.*, "Imspect: Image-driven self-supervised learning for surgical margin evaluation with mass spectrometry," *International Journal of Computer Assisted Radiology and Surgery*, vol. 19, pp. 1129–1136, 2024. [Online]. Available: https://doi.org/10.1007/s11548-024-03106-1

[104] H. E. Kim *et al.*, "Transfer learning for medical image classification: a literature review," *BMC Medical Imaging*, vol. 22, pp. 1–13, 2022. [Online]. Available: https://doi.org/10.1186/s12880-022-00793-7

[105] J. Ma *et al.*, "Segment anything in medical images," *Nature Communications*, vol. 15, no. 1, 2024.

[106] S. Mounsaveng *et al.*, "Bag of Tricks for Fully Test-Time Adaptation," *Proceedings - 2024 IEEE Winter Conference on Applications of Computer Vision, WACV 2024*, pp. 1925–1934, 2024.

[107] D. Wang *et al.*, "Tent: Fully Test-Time Adaptation By Entropy Minimization," *ICLR 2021 - 9th International Conference on Learning Representations*, 2021.

[108] T. Mao *et al.*, "Online seizure prediction via fine-tuning and test-time adaptation," *IEEE Internet of Things Journal*, vol. 11, pp. 20 784–20 796, 2024.

[109] R. Shams *et al.*, "Pre-clinical evaluation of an image-guided in-situ raman spectroscopy navigation system for targeted prostate cancer interventions," *International Journal of Computer Assisted Radiology and Surgery*, vol. 15, pp. 867–876, 2020. [Online]. Available: https://doi.org/10.1007/s11548-020-02136-9

[110] D. Grajales *et al.*, "Towards Real-Time Confirmation of Breast Cancer in the OR Using CNN-Based Raman Spectroscopy," *Cancer Prevention Through Early Detection (CaPTion 2023-MICCAI)*, vol. 2, pp. 17–28, 2023.

[111] H. Sung *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.

[112] A. Fedorov *et al.*, "3d slicer as an image computing platform for the quantitative imaging network," *Magnetic Resonance Imaging*, vol. 30, pp. 1323–1341, 2012. [Online]. Available: http://dx.doi.org/10.1016/j.mri.2012.05.001

[113] S. Banerjee *et al.*, "Use of ultrasound in image-guided high-dose-rate brachytherapy: Enumerations and arguments," *Journal of Contemporary Brachytherapy*, vol. 9, pp. 146–150, 2017.

[114] J. Zhou, L. Zamdborg, and E. Sebastian, "Review of advanced catheter technologies in radiation oncology brachytherapy procedures," *Cancer Management and Research*, vol. 7, pp. 199–211, 2015.

[115] D. L. Batchelar *et al.*, "Intraoperative ultrasound-based planning can effectively replace postoperative ct-based planning for high-dose-rate brachytherapy for prostate cancer," *Brachytherapy*, vol. 15, pp. 399–405, 2016. [Online]. Available: http://dx.doi.org/10.1016/j.brachy.2016.04.003

[116] B. R. Smith *et al.*, "Implementation of a real-time ultrasound-guided prostate hdr brachytherapy program," *Journal of Applied Clinical Medical Physics*, pp. 189–214, 2021.

[117] I.-K. K. Kolkman-Deurloo, M. A. Roos, and S. Aluwini, "Hdr monotherapy for prostate cancer: A simulation study to determine the effect of catheter displacement on target coverage and normal tissue irradiation," *Radiotherapy and Oncology*, vol. 98, pp. 192–197, 2011. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S016781401100003X

[118] K. K. Brock *et al.*, "Use of image registration and fusion algorithms and techniques in radiotherapy: Report of the aapm radiation therapy committee task group no. 132," *Medical Physics*, vol. 44, pp. e43–e76, 7 2017, https://doi.org/10.1002/mp.12256. [Online]. Available: https://doi.org/10.1002/mp.12256

[119] J. M. Fitzpatrick, *Detecting failure, assessing success.* Biomedial Engineering Series, 2001, pp. 117–136.

[120] P. McGeachy *et al.*, "Mri-trus registration methodology for trus-guided hdr prostate brachytherapy," *Journal of Applied Medical Physics (in-review)*, pp. 1–11, 2021.

[121] R. B. Shah and M. Zhou, *Prostate Needle Biopsy Sampling Techniques: Impact on Pathological Diagnosis BT - Prostate Biopsy Interpretation: An Illustrated Guide.* Springer Berlin Heidelberg, 2012, pp. 11–14. [Online]. Available: https://doi.org/10.1007/978-3-642-21369-4_2

[122] E. Cordelli *et al.*, "Radiomics-based non-invasive lymph node metastases prediction in breast cancer," *Proceedings - IEEE Symposium on Computer-Based Medical Systems*, vol. 2020-July, pp. 486–491, 2020.

[123] J. D. Le, J. Huang, and L. S. Marks, "Targeted prostate biopsy: Value of multiparametric magnetic resonance imaging in detection of localized cancer," *Asian Journal of Andrology*, vol. 16, pp. 522–529, 2014.

[124] M. C. Lavallée *et al.*, "Commissioning of an intra-operative us guided prostate hdr system integrating an em tracking technology," *Brachytherapy*, pp. 1–9, 2021.

[125] S. Cui, S. Zhang, and S. Yue, "Raman Spectroscopy and Imaging for Cancer Diagnosis," *Journal of Healthcare Engineering*, vol. 2018, 2018.

[126] R. M. Lopes *et al.*, "Diagnostic model based on Raman spectra of normal, hyperplasia and prostate adenocarcinoma tissues in vitro," *Spectroscopy*, vol. 25, no. 2, pp. 89–102, 2011.

[127] I. P. Santos *et al.*, "Raman spectroscopy for cancer detection and cancer surgery guidance: Translation to the clinics," *Analyst*, vol. 142, pp. 3025–3047, 2017.

[128] C. Chirizzi *et al.*, "A bioorthogonal probe for multiscale imaging by 19f-mri and raman microscopy: From whole body to single cells," *Journal of the American Chemical Society*, vol. 143, 2021, quiz[a util
MRI Macro
Raman Micro.

[129] M. Unal *et al.*, "Raman spectroscopy–based water measurements identify the origin of mri t2 signal in human articular cartilage zones and predict histopathologic score," *Journal of Biophotonics*, vol. 15, pp. 15–25, 2022, useful
Shift for collagen, and water.

[130] F. Picot *et al.*, "Data consistency and classification model transferability across biomedical raman spectroscopy systems," *Translational Biophotonics*, pp. 1–11, 2020.

[131] D. Grajales *et al.*, "Performance of an integrated multimodality image guidance and dose-planning system supporting tumor-targeted HDR brachytherapy for prostate cancer," *Radiotherapy and Oncology*, vol. 166, pp. 154–161, 2021.

[132] E. Short, A. Y. Warren, and M. Varma, "Gleason grading of prostate cancer: a pragmatic approach," *Diagnostic Histopathology*, vol. 25, pp. 371–378, 2019. [Online]. Available: https://doi.org/10.1016/j.mpdhp.2019.07.001

[133] X. Wang *et al.*, "Rapid screening of hepatitis B using Raman spectroscopy and long short-term memory neural network," *Lasers in Medical Science*, vol. 35, oct 2020.

[134] S. C. Lester *et al.*, "Protocol for the examination of specimens from patients with invasive carcinoma of the breast," *Archives of pathology & laboratory medicine*, vol. 133, no. 10, pp. 1515–1538, 2009.

[135] A. Nunez *et al.*, "Accuracy of gross intraoperative margin assessment for breast cancer: experience since the sso-astro margin consensus guidelines," *Scientific Reports*, vol. 10, no. 1, p. 17344, 2020.

[136] D. L. Phelps *et al.*, "The surgical intelligent knife distinguishes normal, borderline and malignant gynaecological tissues using rapid evaporative ionisation mass spectrometry (reims)," *British journal of cancer*, vol. 118, no. 10, pp. 1349–1358, 2018.

[137] K.-L. Lou *et al.*, "Fabrication of tumor targeting rare-earth nanocrystals for real-time nir-iib fluorescence imaging-guided breast cancer precise surgery," *Nanomedicine: Nanotechnology, Biology and Medicine*, vol. 43, p. 102555, 2022.

[138] F. Parent *et al.*, "Intra-arterial image guidance with optical frequency domain reflectometry shape sensing," *IEEE transactions on medical imaging*, vol. 38, no. 2, pp. 482–492, 2018.

[139] Y. Qi *et al.*, "Applications of raman spectroscopy in clinical medicine," *arXiv preprint arXiv:2304.07740*, 2023.

[140] Y. Sun *et al.*, "Image guidance of intracardiac ultrasound with fusion of pre-operative images," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2007: 10th International Conference, Brisbane, Australia, October, 2007, Proceedings, Part I 10*. Springer, 2007, pp. 60–67.

[141] J. Schleusener, J. Lademann, and M. E. Darvin, " Depth-dependent autofluorescence photobleaching using 325, 473, 633, and 785nm of porcine ear skin ex-vivo," *Journal of Biomedical Optics*, vol. 22, no. 9, p. 091503, 2017.

[142] P. F. R. Wilson *et al.*, "ProstNFound: Integrating Foundation Models with Ultrasound Domain Knowledge and Clinical Context for Robust Prostate Cancer Detection," in *Medical Image Computing and Computer Assisted Intervention*. Springer Nature, 2024, pp. 499–509.

[143] F. Dallaire *et al.*, "Quantitative spectral quality assessment technique validated using intraoperative in-vivo Raman spectroscopy measurements," *Journal of Biomedical Optics*, vol. 25, no. 04, p. 1, 2020.

[144] M. Jermyn *et al.*, "A review of Raman spectroscopy advances with an emphasis on clinical translation challenges in oncology," *Physics in Medicine and Biology*, vol. 61, no. 23, pp. R370–R400, 2016.

[145] D. Grajales *et al.*, "Robot-assisted biopsy sampling for online Raman spectroscopy cancer confirmation in the operating room," *International Journal of Computer Assisted Radiology and Surgery*, vol. 19, no. 6, pp. 3–11, 2024.

[146] G. Van Rossum and F. Drake, *Python 3 Reference Manual*, CreateSpace, Ed. Scotts Valley, CA: CreateSpace, 2009.

[147] J. R. Zech *et al.*, "Variable generalization performance of a deep learning model to detect pneumonia in chest radiographs: A cross-sectional study," *PLoS Medicine*, vol. 15, no. 11, pp. 1–17, 2018.

[148] S. Niu *et al.*, "Efficient Test-Time Model Adaptation without Forgetting," *Proceedings of Machine Learning Research*, vol. 162, pp. 16 888–16 905, 2022.

[149] A. Sedghi *et al.*, "Deep neural maps for unsupervised visualization of high-grade cancer in prostate biopsies," *International journal of computer assisted radiology and surgery*, vol. 14, pp. 1009–1016, 2019.

APPENDIX A LIST OF SCIENTIFIC CONTRIBUTIONS & AWARDS

Journal papers

1. D. Grajales et al., «Robot-assisted biopsy sampling for online Raman spectroscopy cancer confirmation in the operating room», International Journal Of Computer Assisted Radiology And Surgery, vol. 19, n.o 6, pp. 1103-1111, abr. 2024, doi: 10.1007/s11548-024-03100-7.

2. D. Grajales et al., «Image-guided Raman spectroscopy navigation system to improve transperineal prostate cancer detection. Part 2: in-vivo tumor-targeting using a classification model combining spectral and MRI-radiomics features», Journal Of Biomedical Optics, vol. 27, n.o 09, sep. 2022, doi: 10.1117/1.jbo.27.9.095004.

3. F. Picot et al., «Image-guided Raman spectroscopy navigation system to improve transperineal prostate cancer detection. Part 1: Raman spectroscopy fiber-optics system and in situ tissue characterization», Journal Of Biomedical Optics, vol. 27, n.o 09, sep. 2022, doi: 10.1117/1.jbo.27.9.095003.

4. D. Grajales et al., «In Vivo Confirmation of Prostate Tumor Burden During HDR Brachytherapy With a Combined Raman Spectroscopy and MRI Radiomics Approach», International Journal Of Radiation Oncology*Biology*Physics, vol. 111, n.o 3, p. e279, oct. 2021, doi: 10.1016/j.ijrobp.2021.07.898.

5. D. Grajales et al., «Performance of an integrated multimodality image guidance and dose-planning system supporting tumor-targeted HDR brachytherapy for prostate cancer», Radiotherapy And Oncology, vol. 166, pp. 154-161, nov.2021,
doi: 10.1016/j.radonc.2021.11.026.

6. R. Shams et al., «Pre-clinical evaluation of an image-guided in-situ Raman spectroscopy navigation system for targeted prostate cancer interventions», International Journal Of Computer Assisted Radiology And Surgery, vol. 15, n.o 5, pp. 867-876, mar. 2020, doi: 10.1007/s11548-020-02136-9.

Conferences

1. D. Grajales*, W. Le, F. Dallaire, G. Sheehy, S. David, T. Tran, F. Leblond, C. Ménard, and S. Kadoury. "Towards Real-time Confirmation of Breast Cancer in the OR using

CNN-based Raman Spectroscopy Classification". CaPTion 2023 (MICCAI Workshop) Lecture Notes in Computer Science. Singapore, 22 October 2023.

2. David Grajales, Fabien Picot, Roozbeh Shams, Maroie Barkati, Guila Delouya, Dominic Béliveau-Nadeau, Benedicte Nicolas, Frederic Leblond, Cynthia Menard, Samuel Kadoury. In Vivo Confirmation of Prostate Tumor Burden during HDR Brachytherapy with a Combined Raman Spectroscopy and MRI Radiomics Approach. POSTER VIEWING Q&A - Annual Meeting of the American Society for Radiation Oncology (ASTRO), October 2021, Annual Meeting Proceedings, Int. J. Radiat. Oncol., vol. 111, no. 3, p. e279, 2021, doi:10.1016/j.ijrobp.2021.07.898.

3. Shams R, Menard C, Grajales D, Béliveau-Nadeau D, Delouya G, Boudam K, Carrier JF, et Kadoury S. Registration accuracy of an integrated MR-TRUS navigation system for prostate HDR brachytherapy. World Congress of Brachytherapy (ESTRO). Radiotherapy and Oncology, Volume 158, Supplement 1, (2021) Pages S122-S123, ISSN 0167-8140, https://doi.org/10.1016/S0167-8140(21)06450-1.

4. Cynthia Ménard, David Grajales, Roozbeh Shams, Maroie Barkati, Guila Delouya, Dominic Béliveau-Nadeau, Benedicte Nicolas, Karim Boudam, Daniel Juneau, Jean DaSilva, Jean-Francois Carrier, Samuel Kadoury. MRI/PET directed and TRUS/EM Guided Prostate Tumor Targeted HDR Brachytherapy: Performance of a Prototype System. ORAL SCIENTIFIC SESSION - International Journal of Radiation Oncology, Biology, Physics VOLUME 108, ISSUE 3, SUPPLEMENT, S66-S67, NOVEMBER 01, (2020) https://doi.org/10.1016/j.ijrobp.2020.07.2202.

Awards & Scholarships

1. 2nd place Best oral presentation (PolyCongRE Congress) [2025]
2. Quebec Research Fund excellence scholarship (FRQ) [2024-2025]
3. 2nd place The Monarch, J&J MedTech Best Paper Award (IPCAI2024) [2024]
4. Outstanding Reviewer Award (14th IPCAI Conference) [2023]
5. TransMedTech institute excellence scholarship (Ph.D.) [2021-2023]
6. 2nd place English version of the 3 Minute Thesis (3MT) Science Dissemination Contest (Polytechnique Montreal) [2023]
7. TransMedTech institute excellence scholarship (M.Sc.A.) [2019-2021]