Teleoperation Data Collection for Imitation Learning in Autonomous Robotic Tumor Resection

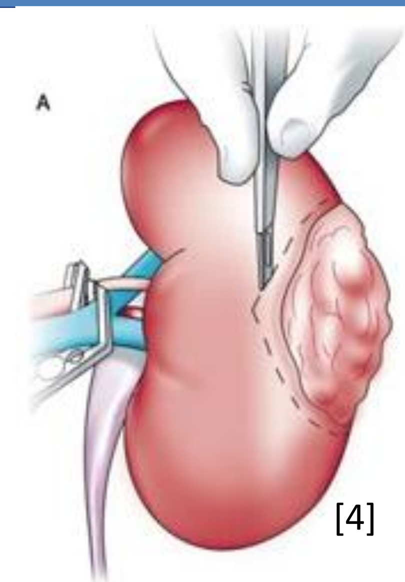
Raquel Susko, Ethan Kilmer, Jiawei Ge, Axel Krieger

Laboratory for Computational Sensing and Robotics, Johns Hopkins University, Baltimore, MD, USA

Introduction

Clinical Background

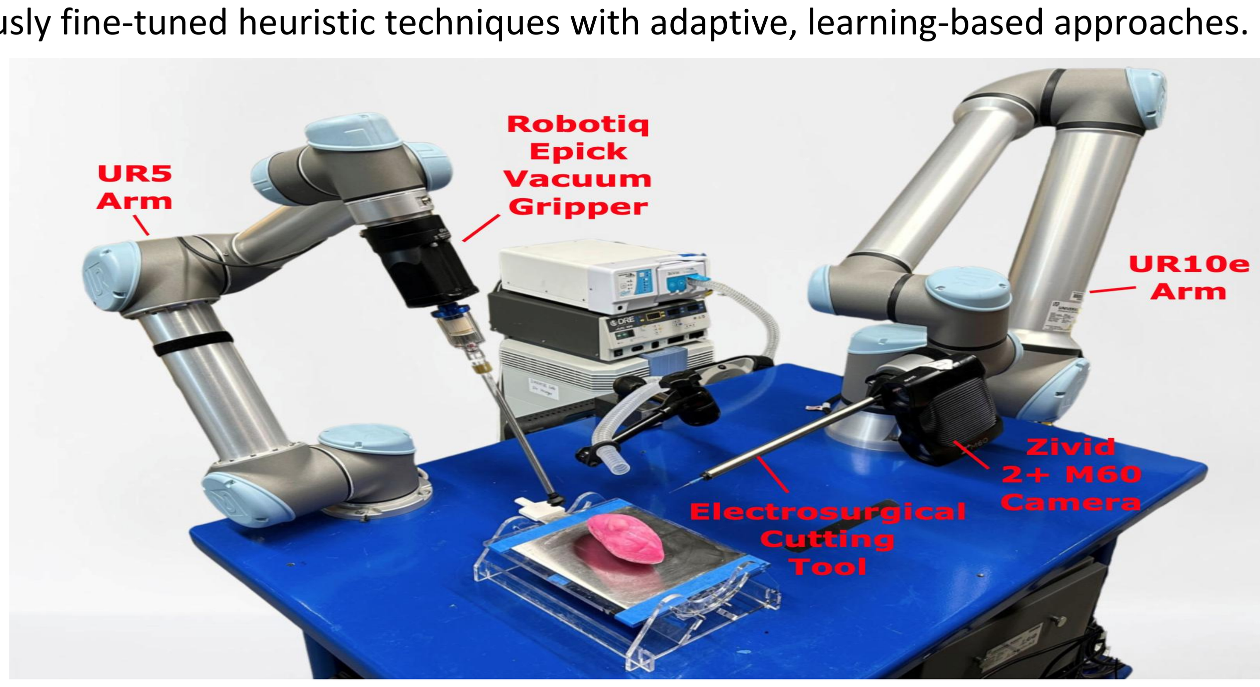
- More than 65,000 patients are diagnosed with kidney and renal pelvis cancers annually [2].
- Tumors in kidneys are often treated through partial nephrectomy - removal of a portion of the kidney - to completely excise unhealthy tissue.
- Traditional manual surgical approaches face inherent limitations in consistency and precision, leading to potential complications.
- Advanced automation in the operating room offers potential to reduce these complications.
- This work contributes to the development of an experimental autonomous surgical robot designed for tumor resection.



[4]

Proposed Robotic Solution

- The Autonomous System for Tumor Resection (ASTR) is a dual-arm robotic platform with UR5 and UR10e robots equipped with vacuum gripper, electrosurgical tools, and 3D vision for autonomous partial nephrectomy [3].
- Previous studies achieved successful imitation learning on da Vinci systems using frameworks like SRT and SRT-H, but ASTR's UR platform lacks an intuitive teleoperation interface, creating a barrier for collecting high-quality training datasets [1].
- **Project Goal: Integrate a Force Dimension haptic device for intuitive UR5 teleoperation, enabling researchers to generate high-quality datasets for imitation learning applications.**
- Ultimate Goal: Apply imitation learning to tissue retraction tasks in partial nephrectomy, replacing previously fine-tuned heuristic techniques with adaptive, learning-based approaches.



Imitation Learning Pipeline

- **Multi-modal data** collection through **Force Dimension haptic teleoperation** of a UR5 surgical robot.
- **Temporal sequence modeling** incorporates previous states to learn sequential dependencies and motion patterns in surgical tasks.
- Multi-modal fusion combines **2D RGB images**, **3D point clouds**, **robot pose**, and **vacuum pressure** through specialized encoders (CNN/ViT, PointNet++, MLP).
- Outputs **relative displacement**, **task completion flags**, and **suction flags** for vacuum grasping applications.
- **Real-time deployment** enables autonomous surgical robot control through trained policy inference with continuous feedback operation.

Acknowledgements

This research was funded through the NSF Grant as a part of the Research Experiences for Undergraduates program at the Laboratory for Computational Robotics at Johns Hopkins Whiting School of Engineering (NSF Grant No. EEC-1852155).
- Justin Chen, Johns Hopkins Whiting School of Engineering, consulting on force dimension implementation.
- Nural Yilmaz, Johns Hopkins Whiting School of Engineering, consulting on Imitation Learning pipeline
- Anton Deguet, Johns Hopkins Whiting School of Engineering, consulting on CISST and FD libraries
- Idris Summola, Johns Hopkins Whiting School of Engineering, consulting on resources
- Viet Pham, Johns Hopkins Whiting School of Engineering, consulting on ROS2

References

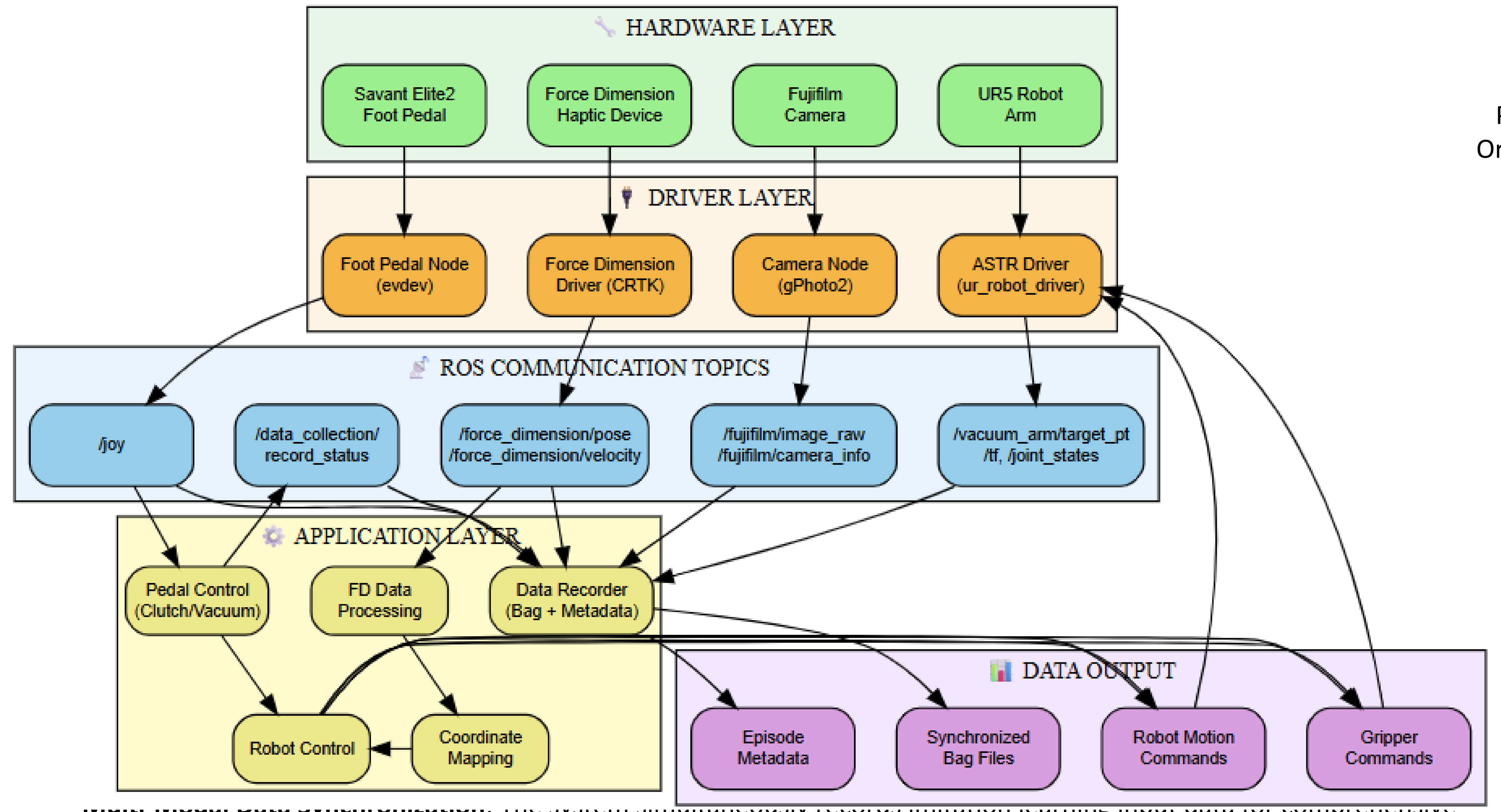
[1] "SRT-H: A hierarchical framework for autonomous surgery via language-conditioned imitation learning," J. W. Kim, J. Chen, P. Hansen, L. X. Shi, A. Goldenberg, S. Schmidgall, P. M. Scheikl, A. Deguet, B. M. White, D. R. Tsai, R. J. Cha, J. Jopling, C. Finn, and A. Krieger, Science Robotics, vol. 10, no. 104, pp. eadt5254, 2025.

[2] "Towards Fluorescence-Guided Autonomous Robotic Partial Nephrectomy on Novel Tissue-Mimicking Hydrogel Phantoms," E. Kilmer, J. Chen, J. Ge, P. Sarda, R. Cha, K. Cleary, L. Shepard, A. E. Ghazi, P. M. Scheikl, and A. Krieger, arXiv preprint arXiv:2503.02265, 2025.

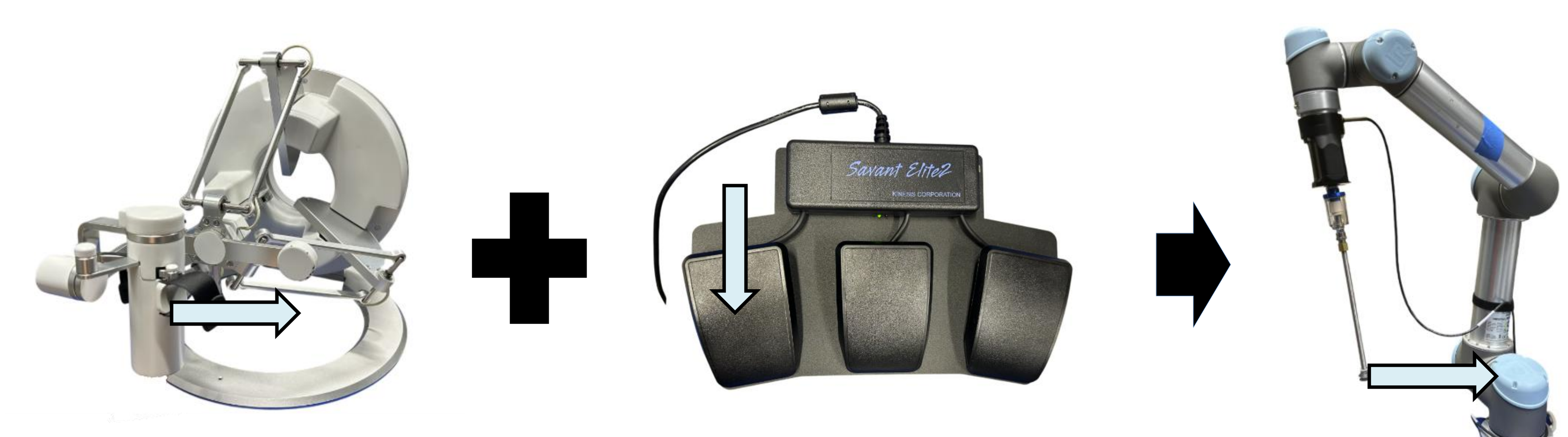
[3] J. Ge et al., "Autonomous System for Tumor Resection (ASTR) - Dual-Arm Robotic Midline Partial Glossectomy," IEEE Robotics and Automation Letters, vol. 9, no. 2, pp. 1166-1173, Feb. 2024, doi: 10.1109/LRA.2023.3341773.

[4] "LiteratureWatch, July–December 2004," Journal of Endourology, vol. 19, no. 2, pp. 253-263, 2005.

Data Collection Pipeline



- **Multi-Modal Data Synchronization**: The system simultaneously records imitation learning input data for comprehensive capture and easy use.
- **Intuitive Haptic Teleoperation**: Smooth control of the UR5 robot through the Force Dimension allows for natural and precise surgical manipulation demonstrations.
- **Episode-Based Recording Management**: Foot pedal controls enable hands-free start/stop of data recording sessions, allowing operators to focus on surgical tasks while capturing discrete demonstration episodes.
- **Real-Time Quality Assurance**: The system provides live feedback on data collection status, sensor synchronization, and recording quality to ensure high-fidelity datasets suitable for robust imitation learning model training.



Devices Used:
- The **Force Dimension** haptic device provides 6DOF teleoperation control with force feedback for intuitive human demonstration.
- A **UR5 robot** equipped with vacuum gripper serves as both the teleoperated platform and autonomous deployment system.
- The **Fujifilm camera** captures high-resolution 2D RGB images of the surgical workspace for vision-based learning.
- The **Savant Elite2 foot pedal** provides hands-free control for clutch engagement, vacuum activation, and data recording operations.
- A **ROS2-based recording system** ensures synchronized data collection across all sensor modalities for consistent training datasets.

Conclusion & Future Work

This project successfully established a teleoperation system that intuitively integrates a Force Dimension haptic device with a Universal Robot arm, enabling precise and responsive real-time manipulation. This work provides a strong foundation for advanced human-robot collaboration in tasks requiring fine motor control and remote operation.

Future work will focus on:
- **Enhanced Control Framework:** Implementation of PID controller architecture for smoother teleoperation control.
- **Data Collection for Imitation Learning:** Collect vacuum grasper data using the force dimension for training.
- **Imitation Learning Model Architecture Evaluation:** Determine the best features for vacuum grasping performance.
- **Dual-Arm Teleoperation Expansion:** Integration of a second Force Dimension device to enable control of both robotic arms for comprehensive surgical procedure data collection.

Results

Vacuum Retraction Performance

Kidney Phantom Orientations

Vacuum Grasper Initial Positions

- The experiment consisted of four orientations of the kidney phantom and three variations in the initial position of the vacuum grasper for a total of 12 trials.
- **Objective: Teleoperate the vacuum grasper to successfully latch onto and retract the tumor.**
- Purpose: Evaluate performance and usability of force dimension teleoperation for its intended use of tissue retraction in tumor resection tasks

Kidney Phantom Orientation	Vacuum Grasper Initial Position: 1	2	3	Success Rate (%)
1	✓	✓	✓	100
2	✓	✗	✓	66.7
3	✓	✓	✗	66.7
4	✓	✓	✓	100
Success Rate (%)	100	75	75	83.3

- Teleoperation performance decreased when kidney phantoms were oriented with tumors positioned farther from the vacuum grasper's initial starting position, indicating distance-dependent manipulation challenges.
- System limitations were observed in clutching mechanics and vacuum movement smoothness, resulting in less precise control during extended reach maneuvers.