Design of Computational Models to Identify Motion Primitive Workflows in Robotic Surgical Systems

Analysis of the Unsuccessful Robotic Hysterectomy of Laurie Featherstone

A Thesis Prospectus
In STS 4500
Presented to
The Faculty of the
School of Engineering and Applied Science
University of Virginia
In Partial Fulfillment of the Requirements for the Degree
Bachelor of Science in Computer Science

By
Sara Liu

December 9, 2022

On my honor as a University student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments.

ADVISORS

Benjamin Laugelli, Department of Engineering and Society

Homa Alemzadeh, Department of Electrical and Computer Engineering, Computer Science

Kay Hutchinson, Department of Electrical and Computer Engineering

Introduction

Surgical robots are complex Cyber-Physical Systems (CPS) that enhance the hand movement of surgeons and have been increasingly used to address a wide range of surgical operations with greater precision, better visualization, and smaller incisions. However, robotic surgery is expensive, requires special training, and can cause devastating complications (Cleveland Clinic, n.d.; Sheetz et al., 2020). Thus, it is critical for safety and efficiency to be top priorities in the field of robotic surgery.

Surgical robots can provide data to train computational models that give insights into the accuracy, quality, correctness, and efficiency of hand movements, enabling the improvement of robot-assisted procedures (van Amsterdam et al., 2021). To address the technical risks associated with robotic surgery, I will propose such models for identifying objective surgical actions in data from a surgical robot collected using two different methods. The results can then be used to analyze safety incidents and provide simulation-based training and automated assessments (van Amsterdam et al., 2021).

While technical improvements are essential, there are also important non-technical factors that affect the safety of robotic surgeries. Surgical robot manufacturers and hospitals should enforce extensive robotic training as well as support an environment in which surgeons do not feel pressured to take misaligned actions (Featherstone, 2017; Siegal et al., 2018). Likewise, surgeons should only decide to perform robotic surgeries when they are ready, and when robotic surgeries would actually be beneficial to the patient (Featherstone, 2017). To strictly impose these requirements, there needs to be changes in the legal system (Siegal et al., 2018). I will investigate the role that social, economic, and legal factors, such as inadequate training, statutory restrictions, and group dynamics, contributed to the many complications of patient Laurie Featherstone's hysterectomy alongside technical factors.

To effectively improve safety and efficiency in robotic surgery, the technical and non-technical aspects must both be addressed in order to provide a more comprehensive approach to this sociotechnical problem. In what follows, I elaborate on a technical project to analyze the performance of machine learning models in identifying fine-grained surgical actions in data collected using two different methods. I also apply Actor-Network Theory (ANT) to examine how human and non-human actors contributed to an unsuccessful robotic surgery for Featherstone to determine how safety can be improved in the future.

Technical Project Proposal

Surgical robots can provide both video and quantitative motion trajectories of instruments during surgical operations. Researchers can use this data to generate computational models that improve operational understanding (van Amsterdam et al., 2021). For example, researchers have analyzed kinematic data and segmented surgical operations at various levels to create models. Two such levels are tasks and gestures. In general, tasks are defined as short engagements such as suturing, needle passing, and knot tying. Tasks can be divided into gestures, which are surgical movements made with a specific purpose, such as reaching for a needle with the right hand (Ahmidi et al., 2017). Hutchinson, Li, et al. (2021) developed a rubric for identifying executional and procedural errors in tasks and gestures by analyzing public kinematic and video data collected by the da Vinci Surgical System (dVSS). Van Amsterdam et al. (2021) and Goldbraikh et al. (2022) also studied gesture recognition in different types of surgery; van Amsterdam et al. (2021) analyzed many different models of gesture recognition, while Goldbraikh et al. (2022) used a sensor system to detect gestures similar to the one I will use.

Gestures can be further decomposed into motion primitives (MPs), basic surgical actions restricted to a smaller set of modular gesture types that are more objective than gestures. MPs are defined to effect changes in surgical context, which describe contact and hold relations for the graspers and objects (Hutchinson, Reyes, et al., 2022). Hutchinson, Reyes, et al. (2022) developed an app to help label surgical context in tasks from three publicly available datasets, automatically translated the context data into MPs, and trained action recognition models based on MP workflows instead of gestures.

The current designs described are limited in their ability to objectively assess skills and lack the data needed for research using accessible surgical systems. Gesture descriptions are open to interpretation, and gesture labeling is thus often conflicting among trials. Even though Hutchinson, Reyes, et al. (2022) created a model that could predict MPs instead, it has not been applied to data other than publicly available datasets. If the problem of using subjective descriptors and limited training data is not addressed, models cannot consistently analyze surgical operations across multiple subjects and trials.

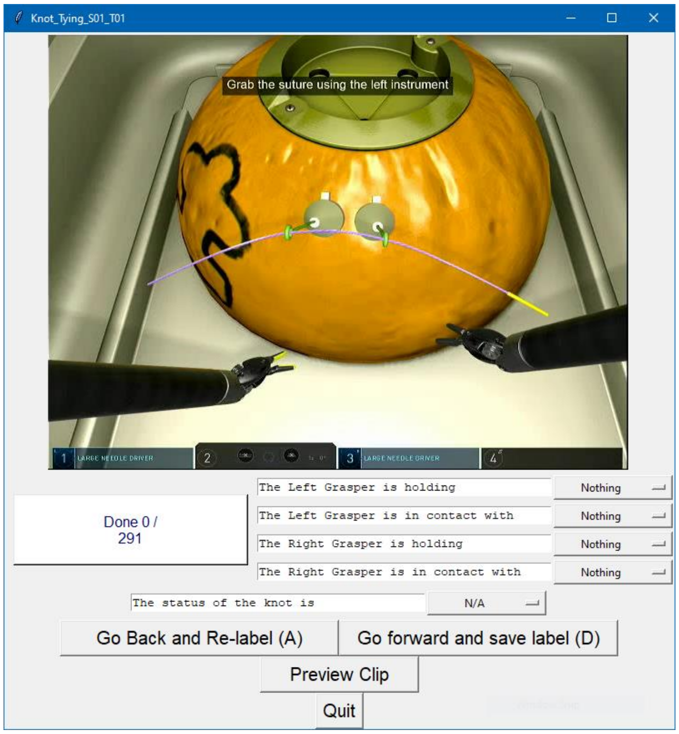
The goal of this technical project is to evaluate the performance of surgical activity segmentation models on data collected from a surgical robot using two different methods in an objective manner. I will use this information to assess the feasibility of using an independent system to collect data from surgical robots and the applicability of using that data for activity recognition as part of future research in safety monitors. One robotic platform I will use is the Raven, an open-architecture surgical robot which mimics the dVSS and can collect kinematic, video, and system log files (Li et al., 2019). The kinematic data from the Raven is based on motor encoder values and forward kinematic equations. The robotic surgery research community has been using the Raven since 2009, and medical students also resort to systems like the Raven to achieve adequate training due to access limitations with the dVSS (Glassman et al., 2014; Li et al., 2019). Thus, it is important to evaluate action prediction models based on the Raven due to its heavy use in the research and medical training communities.

The other apparatus is the Data Collection System (DCS), which is an independent system developed by UVA researchers. The DCS can also collect kinematic and video data, but the kinematic data is from a set of sensors attached to the controllers of a surgical robot such as the Raven or dVSS. In this technical project, I collect data from these two systems instead of the dVSS in the University Hospital because the dVSS is proprietary and requires heavy paperwork to directly gather kinematic data.

The primary task to achieve the goal of the technical project is to compare how well different models perform for data collected simultaneously from the Raven and the DCS attached to the Raven to see if data from a platform-independent collection method can be used to train models with comparable performance. In order to label the video data objectively, I will use the tool that Hutchinson, Reyes, et al. (2022) developed to perform context labeling, the most fine-grained method of segmentation which achieved almost perfect agreement between crowd-sourced and expert surgeon labels (Figure 1). I will then transform the context labels into MP labels using the framework from Hutchinson, Reyes, et al. (2022). Afterwards, I will train the models using the kinematic data from the two systems and evaluate the performance based on the MP labels generated from the framework. Finally, I will compare the outcomes of the models between the Raven and DCS by analyzing how well each system accurately predicted the MPs using the kinematic input data. Time permitting, my secondary task is to test the generalizability of the Temporal Convolutional Network (TCN) model on public kinematic data from the dVSS to data collected from the DCS attached to the dVSS. This comparison depends on the progress of another researcher working on a kinematic transform that homogenizes the kinematic variables from the dVSS and the DCS.

Figure 1

App to Perform Context Labeling on Video Data



STS Project Proposal

In January 2015, Laurie Featherstone visited a gynecologist with complaints of a small fibroid, a growth in her uterus, and an irregular menstrual cycle. During her sub-ten-minute appointment, the doctor told her "If I were you, I'd choose a hysterectomy, and I'd elect the robot. Less down time, little scarring, and less than a 3% complication rate" (Featherstone, 2017, para. 1). On the night of the surgery, she learned that the original surgeon was no longer with the group, but another surgeon with more experience wanted to proceed with the operation. Featherstone took both surgeons' advice and chose to have a complete hysterectomy with the dVSS, without knowing that robotic hysterectomies have high rates of serious complications; in one study published in 2015, the complication rate was 18% (Featherstone, 2017; Marrs, 2022).

Unfortunately, Featherstone encountered numerous complications. After the operation, Featherstone complained of intense pain to the doctor, who dismissed the concern. After a week, Featherstone went in for an emergency room visit to find that her ureter had been burned and she had severe hydronephrosis, the swelling of the kidney due to urine buildup. The doctor was supposed to perform a dye test following the surgery that could have detected this, but it was never ordered. At this point, a urologist had to dissect and reimplant Featherstone's ureter, but needed to attach by feel instead of using the typical method due to the numerous adhesions from the robotic surgery. In addition to the ureter issues, Featherstone also had problems with her bladder, right diaphragm, and pelvic wall (Featherstone, 2017).

Featherstone, attorneys, scholars, and news reporters often attribute the complications in this case to systemic robot failure as well as negligence and greediness on the part of the doctors (Bisnar Chase Personal Injury Attorneys, LLP, n.d.; Featherstone, 2017; Fombu, 2018; Siegel et al., 2018; Villanueva, 2020). For example, Featherstone (2017) refers to her first doctor as a "young female" and directly puts shame on doctors who cause harm by performing unnecessarily major operations. Fombu (2018) uses the Featherstone case to demonstrate the risks associated with robotic technology.

However, while these previous writers have argued that technical design and dereliction of duty are the main factors associated with Featherstone's complications, they have not adequately addressed other factors such as inadequate training, legal limitations, and group dynamics. If we attribute the robotic system's failure only to the current factors considered, then we will not have a more comprehensive account of the range of factors that contributed to the failure of the hysterectomy. By considering those other factors, various social groups can take action to reduce the chances of future robot-assisted hysterectomy failures.

In the STS research paper, I argue that the lack of comprehensive robotic training, legal restrictions, and team dynamics not only had a major impact on the negative outcomes of Featherstone’s surgery, but had a greater influence than that of technical aspects or neglect. To make this argument, I will apply ANT, a conceptual framework that considers how human and non-human elements serve as actors in a heterogeneous, sociotechnical network. A network builder (NB) assembles this heterogeneous network of actors to accomplish a goal by translating or realigning their interests to serve those of the network. These actors constantly define and redefine a sociotechnical arena by combining together to impact technology. The process of the technical development that then occurs over time is called translation (Cressman, 2009). In my argument, the actors behind the training, legal, and teamwork shortfalls include Intuitive Surgical, the manufacturer of the dVSS; the Food and Drug Administration (FDA); the American legal system; surgeons; and hospitals (Siegal et al., 2018). In this case, the NB who recruited these actors into a network which failed to accomplish its goal is the surgical department in the hospital. To support my argument, I will analyze evidence from primary and secondary sources such as Featherstone’s personal documentation, news coverage, and journal articles, which will shed insight on various aspects related to the case and provide concrete statistics.

Conclusion

The technical project will deliver and analyze the outcomes of several models that can identify MPs using kinematic data collected using two different methods from a surgical robot. The results can then be used to enhance the safety and efficiency of robot-assisted surgical operations through resilience assessments, context-aware monitoring, and simulation-based trainings. The STS research paper will deliver new insights into how sociotechnical, economic, and legal factors contributed to the numerous complications resulting from a robotic hysterectomy for Laurie Featherstone. I will accomplish this analysis by applying ANT to characterize how human and non-human actors contributed to unsafe measures in the robotic surgery. The STS project can then inform the technical project by providing insights on non-technical aspects that can shape the future development of trainings and integration of surgical robots into hospitals. The combined results of the technical and STS research will provide a new understanding into how a variety of factors can play a role in improving the safety and effectiveness of robotic surgery.

Word Count: 2,077

References

Ahmidi, N., Tao, L., Sefati, S., Gao, Y., Lea, C., Haro, B. B., Zapella, L., Khudanpur, S., Vidal, R., & Hager, G. D. (2017). A dataset and benchmarks for segmentation and recognition of gestures in robotic surgery. *IEEE Transactions on Biomedical Engineering*, *64*(9), 2025-2041. https://doi.org/10.1109/TBME.2016.2647680

Bisnar Chase Personal Injury Attorneys, LLP. (n.d.). *Da vinci surgical robot poses serious risks for patients*. https://www.bestattorney.com/blog/da-vinci-surgical-robot-poses-serious-risks-for-patients/

Cleveland Clinic. (2021, December 15). *Robotic Surgery.* https://my.clevelandclinic.org/health/treatments/22178-robotic-surgery

Cressman, D. (2009). A brief overview of actor-network theory: Punctualization, heterogeneous engineering & translation. *Simon Fraser University Institutional Repository*, 1-17. https://core.ac.uk/download/pdf/56377732.pdf

Featherstone, L. (2017, November 8). *The business of hysterectomy shatters lives*. Hormones Matter. https://www.hormonesmatter.com/business-hysterectomy-shatters-lives/

Fombu, E. (2018, December 27). *Precision medicine, the risks of robotics and how Susan Potter will live forever*. Dr. Emmanuel "Manny" Fombu. https://www.emmanuelfombu.com/blog-source/2018/9/10/the-power-of-healthcare-data-and-why-we-need-decentralization-nb2jt-2gtzp-gpz9y-yyx7e-n3cdg-wlhl7-gkfrg-w6ebk

Glassman, D., White, L., Lewis, A., King, H. H., Clarke, A., Glassman, T., Comstock, B., Hannaford, B., & Lendvay, T. S. (2014). Raven surgical robot training in preparation for da Vinci® use: A randomized prospective trial. *Studies in Health Technology and Informatics, 196,* 135-141. https://doi.org/10.3233/978-1-61499-375-9-135

Goldbraikh, A., Volk, T., Pugh, C. M., & Laufer, S. (2022). Using open surgery simulation kinematic data for tool and gesture recognition. *International Journal of Computer Assisted Radiology and Surgery*, *17*, 965-979. https://doi.org/10.1007/s11548-022-02615-1

Hutchinson, K., Li, Z., Cantrell, L. A., Schenkman, N. S., & Alemzadeh, H. (2022). Analysis of executional and procedural errors in dry-lab robotic surgery experiments. *The International Journal of Medical Robotics and Computer Assisted Surgery*, *18*(3), 1-15. https://doi.org/10.1002/rcs.2375

Hutchinson, K., Reyes, I., Li, Z., & Alemzadeh, H. (2022). *COMPASS: A formal framework and aggregate dataset for generalized surgical procedure modeling* [Unpublished manuscript]. https://arxiv.org/pdf/2209.06424.pdf

Li, Y., Hannaford, B., & Rosen, J. (2019). *Raven: Open surgical robotic platforms* [Unpublished manuscript]. https://arxiv.org/pdf/1906.11747.pdf

Marrs, C. (2022, October 26). *Conquering the uterus - trends in hysterectomy*. Hormones Matter. https://www.hormonesmatter.com/conquering-the-uterus-trends-in-hysterectomy/

Sheetz, K. H., Claflin, J., & Dimick, J. B. (2020). Trends in the adoption of robotic surgery for common surgical procedures. *JAMA network open*, *3*(1), 1-9. https://doi.org/10.1001/jamanetworkopen.2019.18911

Siegel, E. R., McFadden, C., Monahan, K., Lehren, A. W., & Siniauer, P. (2018, December 19). The da Vinci surgical robot: A medical breakthrough with risks for patients. *NBC News*. https://www.nbcnews.com/health/health-news/da-vinci-surgical-robot-medical-breakthrough-risks-patients-n949341 van Amsterdam, B., Clarkson, M. J., & Stoyanov, D. (2021). Gesture recognition in robotic surgery: A review. *IEEE Transactions on Biomedical Engineering*, *68*(6). https://doi.org/10.1109/TBME.2021.3054828