Deep Learning for Automated Segmentation of Coronary Artery CTs, Hypertrophic Cardiomyopathic MRIs, and Hydrocephalic CTs
(Technical Paper)

Healthcare Wearable Data Privacy Concerns and Future Solutions for More Accountable Data Handling
(STS Paper)

A Thesis Prospectus Submitted to the

Faculty of the School of Engineering and Applied Science
University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements of the Degree
Bachelor of Science, School of Engineering

Ramiz Akhtar
Fall, 2020

Technical Project Team Members
Rohan Patel
Vignesh Valaboju
Sharon Zheng

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

Introduction

With a forecasted $78 billion increase in market value between 2018 and 2025, smart healthcare wearables are becoming commonplace technologies that collect biometric data to provide users with analytical health insights, but where does the collected data actually go and how much control does the user have over its post-collection use (Ugalmugle & Swain, 2019)? Considering that the Health Insurance Portability and Accountability Act (HIPAA) does not apply to voluntary purchases from wearable device manufacturers, manufacturers are left largely unchecked regarding their handling of user health data (Banerjee et al., 2018). In recent years, healthcare wearables have seen a gradual expansion in their data collection capabilities by pushing features such as pulse oximetry, sleep tracking, and even tone analysis to predict the user's emotional state (Amazon.com, 2020). As wearable technologies continue to collect more data, it is imperative for users to understand the privacy implications of sending their health data to wearable manufacturers. The future use of transmitted health data is currently unknown, and the door is open for data originators to be negatively impacted by their own health data through scenarios such as increased insurance rates for sleeping less than the recommended healthy standard (Allen, 2018). Therefore, the proposed topic for the formal STS research paper is an in-depth analysis of healthcare wearable privacy issues.

While health wearable manufacturers expand their capability to provide insights into user health, empowering doctors to more efficiently derive insights from medical images will allow for improved patient care. In modern healthcare, medical images produced from modalities such as Computed Tomography (CT) and Magnetic Resonance Imaging (MRI) have become ubiquitous for preventative disease screening, disease diagnosis, and treatment selection for patients; they allow clinicians to non-invasively analyze potential disease states of organs (Bercovich & Javitt, 2018). The images contain insurmountable qualitative and quantitative information for clinicians to analyze. However, physicians must perform manual segmentation of medical images, which is an arduous and time-consuming process for differentiating anatomical regions within the images, to derive insights from them. By utilizing deep learning, this technical research project aims to automate segmentation for three disease states: coronary artery disease, hypertrophic cardiomyopathy, and hydrocephalus. Innovation through the creation of an automated process is required for improved segmentation efficiency, decreased inter-observer bias, and higher quality diagnoses (Merjulah & Chandra, 2019, p. 10). Therefore, the proposed topic for the formal technical research report is a discussion about developing automated segmentation algorithms for the three aforementioned disease states.

Deep Learning for Automated Segmentation of Medical Images

Considering medical professionals must undertake a multitude of tasks, maintaining high efficiency when completing any given task is of utmost priority. Most current image-based diagnoses rely on radiologists to qualitatively interpret and manually segment CT and MRI images, inherently resulting in time-consuming and arduous endeavors. On the contrary, quantitative evaluation and automated image segmentation can improve the efficiency, accuracy, reliability of diagnosis by extracting imaging biomarkers, which are measurable indicators of a disease state, to provide insights such as probability for disease progression, intensity of disease presence, and other numerical features about the physiology and anatomy of the image (Kim et al., 2015). Identified below are three disease states of interest that require medical image analysis and will benefit from automated image segmentation.

Coronary artery disease is classified by the narrowing of the arteries via a buildup of plaque, limiting the blood flow to the heart, and often leading to a heart attack (Mayo Clinic, 2020).

Coronary artery disease is the most common type of heart disease, killing 365,914 people in 2017, and affecting 18.2 million adults above the age of 20 in the United States (CDC, 2020). Current identification of coronary artery disease mainly consists of manual CTA (Computed Tomography Angiography) segmentation to identify calcification, plaque residues, and the narrowing of coronary arteries (Arnett et al., 2019).

Hypertrophic cardiomyopathy (HCM), a disease with a prevalence of 1 in 200, is characterized by unexplained enlargement of the heart's left ventricle (LV), misalignment of cardiac fibers, and cardiac scarring, all of which can lead to heart dysfunction (Kramer et al., 2015; Olivotto et al., 2011; Semsarian et al., 2015). Most HCM cases are genetically inherited and can gradually progress towards heart failure and cardiac arrest (Watkins et al., 2011). Current identification methods for HCM include genetic testing and manual segmentation of cardiac magnetic resonance (CMR) images for unexplained left ventricular wall thickness and scarring (Wigle et al., 1985).

Hydrocephalus affects 1 million people in the United States and is very common in children as 1.1 in every 1000 infants will develop hydrocephalus (Tully & Dobyns, 2014). As a result of increased pressure deep within the brain, hydrocephalus is characterized by fluid buildup in the brain's ventricles which can damage brain tissues and cause severe disabilities (Hamilton et al., 2016). Manual image segmentation of brain CTs is essential for hydrocephalus diagnosis and is used to distinguish areas of brain ventricles, identify ventricular abnormality, and determine the severity of a given hydrocephalus case.

Innovation in the form of automating each disease state's segmentation process is required to achieve improved segmentation efficiency, subsequent reduction in diagnosis times, and overall improvements in condition monitoring capabilities. There have been several recent advances in using artificial intelligence to automatically segment medical images. Used in the context of segmentation, deep learning convolutional neural networks (CNNs) allow for extraction of image features for classification. The first aim of the project is to develop and validate deep learning-based segmentation algorithms with a CNN backbone for each of the aforementioned disease states using the TensorFlow python package. For coronary artery segmentation, innovative multi-channel inputs about the prior shape of blood vessels will be fed into the CNN. For HCM scar and LV wall segmentation, the objective will be to evaluate the patient population generalizability of an improved segmentation algorithm. For hydrocephalic brain ventricle segmentation, CT images will be segmented for increased spatial resolution compared to the current algorithmic standard of segmenting MRI images. Automated segmentation results will be compared against manual image segmentation, the current industry standard, as a metric of success.

Presenting the aforementioned automated segmentation algorithms to physicians in their raw form would create potential for unintended modification of the algorithms, thus creating technical issues, reducing efficiency, and compromising patient care. To alleviate this concern, the second aim of the project is to create a universal software application using the React Native framework to house and black-box the underlying automated segmentation algorithms. Developing a user interface (UI) that prioritizes simplicity, intuitive use, and extensive user control would grant physicians the power to see biomarker information in as much depth as needed while avoiding common problems, such as electronic medical record click fatigue, that plague physician interactions with software (Collier, 2018). Therefore, implementing the automated segmentation algorithms in a complete software package would make them more likely to see use in clinical and clinical training environments.

The resources used to conduct research for the project are primarily medical and computer science journals that describe cutting-edge computational approaches used in the field of medical imaging. In regards to human resources, the project will be completed with three other biomedical engineering students who are either pursuing a computer science minor or have a generally strong computational medical imaging background. The final project report will be structured to describe the three disease states of focus, discuss limitations with current manual segmentation approaches, explain the project's automated segmentation algorithms with corresponding validation metrics, and lastly identify areas for future work.

Data Privacy Concerns in Healthcare Wearables

Imagine a future in which smart contact lenses are able to continuously monitor blood glucose levels for diabetic patients. Alphabet, the parent company of Google, has indefinitely postponed such a project after investing millions of dollars in research and development in an attempt to create wearable technologies of the future (Farr, 2018). While the glucose detecting contact lens has not yet come to fruition, Alphabet's expenditure on the project emphasizes that big technology companies are expecting significant growth in the wearable health technology sector, and thus, are spending heavily to stay ahead of competitors. New wearable technologies present technology companies with the opportunity to collect additional categories of user health data. However, does the rate of advancement in privacy standards match the rapid advancement of new health data collection methods?

Currently, wearable device data collection can be explored in terms of user identification, GPS location tracking, sensor-based data, and Internet-of-Things (IoT) data transmission (Banerjee et al., 2018). User identification includes data supplied by the user to the service that corresponds to a particular wearable device, including name, age, height, and weight. GPS location tracking is often used to record the path taken for outdoor exercises such as running. Sensor-based data sources include heart rate, blood oxygen levels, and a single-lead electrocardiogram. IoT data transmission involves data encryption and transmission to the parent device of the wearable, which is usually a smartphone that contains the wearable manufacturer's proprietary software. Once health data is collected by the wearable manufacturer's software, users are provided analytical insights into their health. However, beyond the provided health insights, users cannot follow data submitted to a service because data handling post-collection has effectively been black boxed by wearable manufacturers (Purcell & Rommelfanger, 2017). In exchange for insights, the user has effectively given up ownership of personal data. Users must also consider that malicious entities are capable of stealing the data at any moment throughout the health data collection process. In a recently discovered exploit of Apple's iPhone operating system, iOS, malicious parties were able to covertly steal text messages, GPS location, and private health data from users' phones after they visited certain web pages (Beer, 2019). While many wearable manufacturers use a cloud-based service to manage health data collected by the wearable, local copies of the data still exist on the smartphone for efficient access.

Considering the aforementioned privacy issues surrounding the use of wearable devices and their corresponding service, 60% of people interviewed in a study (n=20) were unconcerned about the privacy of their health data after submitting the data to a health wearable service (Lowens et al., 2017). Users who were concerned about health data privacy cited the potential for insurance companies to pay for access to an aggregate of health data and use it to determine insurance rates for customers. In contrast, users who were unconcerned about health data privacy stated that most health data collected, such as steps taken, is trivial and does not require extensive thought about privacy. While metrics similar to steps taken are arguably trivial, the need for greater transparency in data collection practices and data security is apparent considering the rapid advancement of wearable technologies and new categories of data they can collect. A sustainable and modern privacy infrastructure needs to be in progress to prepare for the inevitable time when health wearables begin collecting even more sensitive health data such as blood glucose and blood pressure. A motivation of the full paper is to raise awareness among the public about the importance for prioritizing privacy at the current moment.

To analyze the intricate problem of health data privacy, the framework of Actor Network Theory (ANT) will be utilized. ANT provides an analytical basis for a complex network of relationships between technologies, governments, and people by identifying how each relevant stakeholder, physical artifact, and/or non-physical artifact interacts with each other (Cressman, 2009). Key stakeholders and artifacts for a discussion of health data privacy include: wearable technologies, wearable technology manufacturers, wearable technology services, smartphones, consumers, governments, malicious entities, medical professionals, and insurance companies. The technicalities of communication between wearable technologies, smartphones, and wearable technology services will be black boxed to prevent the ANT analysis from becoming convoluted. A primary criticism of ANT is that it ignores "intangible" elements such as values and norms and instead focuses on empirical observation that does not fully encompass a given topic (Cressman, 2009). Specifically related to health data privacy, ANT would ignore topics such as a wearable manufacturer's potential ethical decision to refrain from selling health data even after owning the data per the manufacturer's privacy policy. To compensate for ANT's shortcoming, the framework of technological momentum will supplement ANT.

Technological momentum, introduced by historian of technology Thomas Hughes, emphasizes that a technological system must align with the needs of society in its early stages of development. However, as the technology matures and gains momentum, it becomes difficult to alter its trajectory because it has begun simultaneously influencing society at-large (Hughes, 1994). A criticism of technological momentum is that it is time-dependent, meaning that it is not necessarily clear when in time a technology has reached the critical point of becoming deterministic. In regards to data privacy in wearables, it would be beneficial to quantify the urgency for data privacy reform. Both theories and their respective criticisms will be further analyzed in context of health data privacy within the analysis of the full STS research paper.

Research Question and Methods

The STS question of focus is: As wearable technologies progress, what are the previous, current, and forthcoming data privacy issues with the new generation of wearable technologies? Research will be conducted using the documentary research method and interviews. The documentary research method will utilize a variety of sources ranging from technology media articles to peer-reviewed data privacy journals. There are four categories in which research will be organized and analyzed: previous and present wearable data collection practices, technological security concerns, user awareness of privacy, and future implementations for improved health data security. The key words to research include "wearable data privacy," "biometric security concerns," and "reading privacy policies." The interview method will reach out to people of all ages and backgrounds who use health wearables, such as Fitbit, Android Wear devices, and Apple Watches, to better understand their stance on health data privacy before and after presenting them with current and future health data privacy concerns. Questions for the interview will be adapted from Lowens et al. and will be similar to the following: How long have you been wearing a health wearable? Do you read the privacy policy before beginning use of the device? How concerned are you with the data privacy of the collected information from your wearable? The interview will provide information about people's level of awareness regarding data privacy issues and reasons for that level of awareness.

Conclusion

For the STS research question, the objective is to analyze shifts in privacy policy and data security over time as healthcare wearables have continued to advance. Understanding the public's current perception regarding data privacy is essential to forcing a grassroots shift towards more proactive privacy regulations. Therefore, interviews will be conducted to understand individuals' opinions on data privacy before and after presenting them with information about future privacy risks under current regulations. Lastly, potential solutions that acknowledge the previously discussed privacy concerns will be analyzed to determine their impact on creating a more privacy-centric and secure health data environment, ultimately moving forward alongside the rapidly advancing health wearable technologies. The formal STS research paper will synthesize the findings of the aforementioned objectives by late April 2021.

For the technical topic, the objective is to create automated deep learning segmentation algorithms for coronary artery disease, hypertrophic cardiomyopathy, and hydrocephalus as well as integrating each of these algorithms into a user-friendly software package that allows for ease of use in the clinical environment. Compared to manual segmentation, automated segmentation allows for increased efficiency while maintaining similar segmentation accuracy to manual segmentation by an experienced physician. Ultimately, automated segmentation will decrease diagnosis time for the three disease states and allow for improved condition monitoring capabilities. The formal technical research paper will synthesize the design technicalities of the aforementioned objectives, provide discussion into algorithm's effectiveness in terms of segmentation accuracy, and announce the availability of the automated segmentation package as a free-to-use software. Research and algorithm development will be completed by early March 2021 and writing will be completed by late April 2021.

References

Allen, M. (2018, July 18). *Health Insurers Are Vacuuming Up Details About You—And It Could Raise Your Rates*. ProPublica. https://www.propublica.org/article/health-insurers-are-vacuuming-up-details-about-you-and-it-could-raise-your-rates?token=pc1OkzviktiLco0hJY5BDRPpI44H_bKV

Amazon.com. (2020). *Amazon Halo Privacy*. Amazon.Com Help. https://www.amazon.com/gp/help/customer/display.html?nodeId=GL99TQL4B7ADPBDH

Arnett, D. K., Blumenthal, R. S., Albert, M. A., Buroker, A. B., Goldberger, Z. D., Hahn, E. J., Himmelfarb, C. D., Khera, A., Lloyd-Jones, D., McEvoy, J. W., Michos, E. D., Miedema, M. D., Muñoz, D., Smith, S. C., Virani, S. S., Williams, K. A., Yeboah, J., & Ziaeian, B. (2019). 2019 ACC/AHA Guideline on the Primary Prevention of Cardiovascular Disease: Executive Summary: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. *Circulation*, *140*(11). https://doi.org/10.1161/CIR.0000000000000677

Banerjee, S. (Sy), Hemphill, T., & Longstreet, P. (2018). Wearable devices and healthcare: Data sharing and privacy. *The Information Society*, *34*(1), 49–57. https://doi.org/10.1080/01972243.2017.1391912

Beer, I. (2019, August 29). Project Zero: A very deep dive into iOS Exploit chains found in the wild. *Project Zero*. https://googleprojectzero.blogspot.com/2019/08/a-very-deep-dive-into-ios-exploit.html

Bercovich, E., & Javitt, M. C. (2018). Medical Imaging: From Roentgen to the Digital Revolution, and Beyond. *Rambam Maimonides Medical Journal*, *9*(4). https://doi.org/10.5041/RMMJ.10355

CDC. (2020, September 8). *Heart Disease Facts | cdc.gov*. Centers for Disease Control and Prevention. https://www.cdc.gov/heartdisease/facts.htm

Collier, R. (2018). Rethinking EHR interfaces to reduce click fatigue and physician burnout. *CMAJ : Canadian Medical Association Journal*, *190*(33), E994–E995. https://doi.org/10.1503/cmaj.109-5644

Cressman, D. (2009). *A Brief Overview of Actor-Network Theory: Punctualization, Heterogeneous Engineering & Translation*. School of Communication, Simon Fraser University.

Farr, C. (2018, November 16). *Alphabet stops its project to create a glucose-measuring contact lens for diabetes patients*. CNBC. https://www.cnbc.com/2018/11/16/alphabet-verily-stops-smart-lens-glucose-measuring-contact-lens.html

Hamilton, M., Gruen, J. P., & Luciano, M. G. (2016). Introduction: Adult hydrocephalus. *Neurosurgical Focus*, *41*(3), E1. https://doi.org/10.3171/2016.6.FOCUS16272

Hughes, T. (1994). *Technological Momentum*. MIT Press.

Kim, H., Park, C. M., Goo, J. M., Wildberger, J. E., & Kauczor, H.-U. (2015). Quantitative Computed Tomography Imaging Biomarkers in the Diagnosis and Management of Lung Cancer. *Investigative Radiology*, *50*(9), 571–583. https://doi.org/10.1097/RLI.0000000000000152

Kramer, C. M., Appelbaum, E., Desai, M. Y., Desvigne-Nickens, P., DiMarco, J. P., Friedrich, M. G., Geller, N., Heckler, S., Ho, C. Y., Jerosch-Herold, M., Ivey, E. A., Keleti, J., Kim, D.-Y., Kolm, P., Kwong, R. Y., Maron, M. S., Schulz-Menger, J., Piechnik, S., Watkins, H., … Neubauer, S. (2015). Hypertrophic Cardiomyopathy Registry (HCMR): The rationale and design of an international, observational study of hypertrophic cardiomyopathy. *American Heart Journal*, *170*(2), 223–230. https://doi.org/10.1016/j.ahj.2015.05.013

Lowens, B., Motti, V. G., & Caine, K. (2017). Wearable Privacy: Skeletons in The Data Closet. *2017 IEEE International Conference on Healthcare Informatics (ICHI)*, 295–304. https://doi.org/10.1109/ICHI.2017.29

Mayo Clinic. (2020, June 5). *Coronary artery disease—Symptoms and causes*. Mayo Clinic. https://www.mayoclinic.org/diseases-conditions/coronary-artery-disease/symptoms-causes/syc-20350613

Merjulah, R., & Chandra, J. (2019). Chapter 10—Classification of Myocardial Ischemia in Delayed Contrast Enhancement Using Machine Learning. In D. J. Hemanth, D. Gupta, & V. Emilia Balas (Eds.), *Intelligent Data Analysis for Biomedical Applications* (pp. 209–235). Academic Press. https://doi.org/10.1016/B978-0-12-815553-0.00011-2

Olivotto, I., Girolami, F., Sciagrà, R., Ackerman, M. J., Sotgia, B., Bos, J. M., Nistri, S., Sgalambro, A., Grifoni, C., Torricelli, F., Camici, P. G., & Cecchi, F. (2011). Microvascular Function Is Selectively Impaired in Patients With Hypertrophic Cardiomyopathy and Sarcomere Myofilament Gene Mutations. *Journal of the American College of Cardiology*, *58*(8), 839–848. https://doi.org/10.1016/j.jacc.2011.05.018

Purcell, R. H., & Rommelfanger, K. S. (2017). Biometric Tracking From Professional Athletes to Consumers. *The American Journal of Bioethics*, *17*(1), 72–74. https://doi.org/10.1080/15265161.2016.1251652

Semsarian, C., Ingles, J., Maron, M. S., & Maron, B. J. (2015). New Perspectives on the Prevalence of Hypertrophic Cardiomyopathy. *Journal of the American College of Cardiology*, *65*(12), 1249–1254. https://doi.org/10.1016/j.jacc.2015.01.019

Tully, H. M., & Dobyns, W. B. (2014). Infantile hydrocephalus: A review of epidemiology, classification and causes. *European Journal of Medical Genetics*, *57*(8), 359–368. https://doi.org/10.1016/j.ejmg.2014.06.002

Ugalmugle, S., & Swain, R. (2019). *Wearable Medical Devices Market Share: 2025 Growth Statistics*. Global Market Insights. https://www.gminsights.com/industry-analysis/wearable-medical-devices-market

Watkins, H., Ashrafian, H., & Redwood, C. (2011, April 27). *Inherited Cardiomyopathies* (world) [Review-article]. Http://Dx.Doi.Org/10.1056/NEJMra0902923; Massachusetts Medical Society. https://doi.org/10.1056/NEJMra0902923

Wigle, E. D., Sasson, Z., Henderson, M. A., Ruddy, T. D., Fulop, J., Rakowski, H., & Williams, W. G. (1985). Hypertrophic cardiomyopathy. The importance of the site and the extent of hypertrophy. A review. *Progress in Cardiovascular Diseases*, *28*(1), 1–83. https://doi.org/10.1016/0033-0620(85)90024-6