Using Computer Vision and Vision Artificial Intelligence to Improve the Fan Experience at Sports Stadiums
(Technical Project)

The Socio-Technical Implications of Artificial Intelligence in Medical Image Analysis
(STS Project)

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
Kaihil Patel

December 15, 2023

Technical Team Members: Kaihil Patel

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

Rider Foley, Department of Engineering and Society

Rosanne Vrugtman and Briana Morrison, Department of Computer Science

Introduction

In the field of Computer Science, there has been a rapidly growing interest in the subdisciplines of machine learning and artificial intelligence. These specialty fields empower computers to learn and make decisions via programmatic training. They operate on the principle that machines can analyze and comprehend vast amounts of data, recognizing complex patterns and relationships within it. This is achieved through the utilization of algorithms that iteratively learn from the data, allowing them to improve their performance over time (El Naqa & Murphy, 2015). From healthcare and finance to the entertainment and automotive industries, one of the key strengths of machine learning is its adaptability across a variety of domains. By extracting meaningful insights and predictions from data, machine learning algorithms have revolutionized automating tasks and optimizing processes.

Within the realm of machine learning, computer vision emerges as a specialized branch dedicated to enabling computers to comprehend and interpret visual information, much like the human visual system. Machine learning models in computer vision often utilize Convolutional Neural Networks (CNNs). CNNs employ multiple layers of image analysis, each learning to detect distinct features in an input image by applying filters that refine and enhance the details in subsequent layers. This complex model opens the door to a plethora of applications such as object detection and tracking in medical imaging (Voulodimos et al., 2018).

Like any technology, there are moral and ethical implications that come with the use of the technology (Winner, 1980). Personal privacy concerns and unintended bias can have varied consequences depending on the implications and stakes of the vision model. For instance, medical professionals' burnout has become an internationally recognized issue in recent years, and implementing technologies to assist in their administrative tasks could help solve the issue

(West et al., 2018). However, due to the lack of information on machine learning and computer vision, its use risks grave consequences if not built and utilized properly. Not only is there a gap in knowledge in the healthcare industry when trying to use computer vision to help automate different processes (Flah et al., 2021) but also in the public when trying to understand where they might be biased.

I will be studying two specific aspects of computer vision: how it can be used to improve attendee experiences in sports stadiums and what its implications are in the healthcare industry.

Computer Vision for Improving Stadium Experience

For the technical portion of this paper, I will be discussing my summer 2023 internship experience and project. The company's current full-time team was working on using computer vision to improve the fan experience at sports stadiums. Stadiums and venues across the world experience operating inefficiency issues causing revenue loss and unfavorable attendee experiences. The use of Vision Artificial Intelligence as an effective solution is directly supported by alwaysAI's claim that "Computer vision can directly grow businesses' ROI (Return on Investment). Whether it's improving productivity, streamlining processes and tracking downtime, reducing operating costs, or monitoring safety compliance, computer vision improves business efficiency." (Siddell, 2023, p. 4). This technique is not currently in use at sports venues today, so the idea and implementation are new and unique.

My team and I were some of the first employees to have our hands on this project, so there was no existing software or direction other than a basic plan that had been outlined. We were tasked with creating a proof of concept, or minimum viable product, to showcase to the client in pursuit of a long-term contract for this project.

An initial step in accomplishing this task was to identify which areas of concern to target for improvement. Stadium data submitted by the fans on gamedays was collected and shared with us by the team and my employer for us to analyze. An in-depth analysis of the extensive data was conducted. Along with personal experiences, a plethora of problems that all contribute to the stadium's dissatisfactory rating were discovered from the data. However, it is best to solve the overall issue incrementally, so three leading areas of improvement were identified: foot traffic time, long lines, and empty trash bin availability.

After researching computer vision approaches to this matter, a software package utilizing CNNs was decided as the best path to our goal. We built a prototype using the new, state-of-the-art computer vision model developed by Ultralytics, YOLOv8 (Solawetz, 2023), and Roboflow's Supervision. A model was designed to identify people and track their movements across a venue using live data.

A leading concern for stadium attendees is personal privacy. This became apparent during data collection after many attendees asked why we were video recording their movements. Because all the data is collected via cameras, public attendees' bodies are a critical part of how the models train and operate. Nonetheless, as stated in a Forbes article, "Computer vision is just a tool to help humans do their jobs" (Ciabarra, 2019) and should not be considered a personal privacy concern. I agree with the authors of "Enabling Video Privacy through Computer Vision" that if personal data is not being stored or distributed, the use of such data to improve the lives of citizens is ethical (Senor et al., 2005). However, the people being used as data in the film have a right to be aware of what is happening and how their bodies and movements are being used in a stadium setting. As Andrews (2006) stated, citizens not only have rights but also obligations, such as being informed, engaging in public matters, and being respectful of the common welfare. As technological developers, these must be considered as it is our duty to assist the public in exercising technological citizenship.

Socio-Tech of Automated Medical Imaging

One potentially impactful use of computer vision is in the medical and healthcare fields. This use gained popularity in 2012 (Sarvamangala and Kulkarni, 2022), and it has already had widespread success in the industry. However, that success has not come without its drawbacks. There is growing concern about the potential bias of the predictions the machine learning models make (Neuroscience News, 2023).

In addition to alleviating professional burnout by automating administrative tasks, automating the imaging analysis process for breast cancer radiologists obtained a high classification accuracy and showed that a multistage classification scheme is more effective than the single-stage classification in establishing a new strong benchmark for image analysis (Gao et al., 2018). I will be studying the implications of computer vision models in healthcare using the Actor-Network Theory (ANT) framework defined by Latour (1992) in *Where Are the Missing Masses? The Sociology of a Few Mundane Artifacts*. To accomplish this, it is important to understand the machine learning models in use and see how the training of these models is concerning for impacted social groups.

Medical Image Computing (MIC) can be systematically biased against certain subgroups with protected attributes like age, race, and gender. This bias is considered by many as against the principles of bioethics and autonomy as it increases the probability of patients being mistreated or not treated at all (Lara et al., 2022). Treating this problem requires diving into the roots of the data. Data is the lifeblood of a machine learning model, serving as the foundation from which it learns patterns and makes predictions. The quality, quantity, and diversity of data directly impact a model's accuracy and capabilities, making it crucial for effective performance. As outlined in Figure 1, one foundational step in the machine learning lifecycle is training data feature extraction. If this data is biased in any way, like a disproportionate representation of subgroups, the final evaluations will be as well.

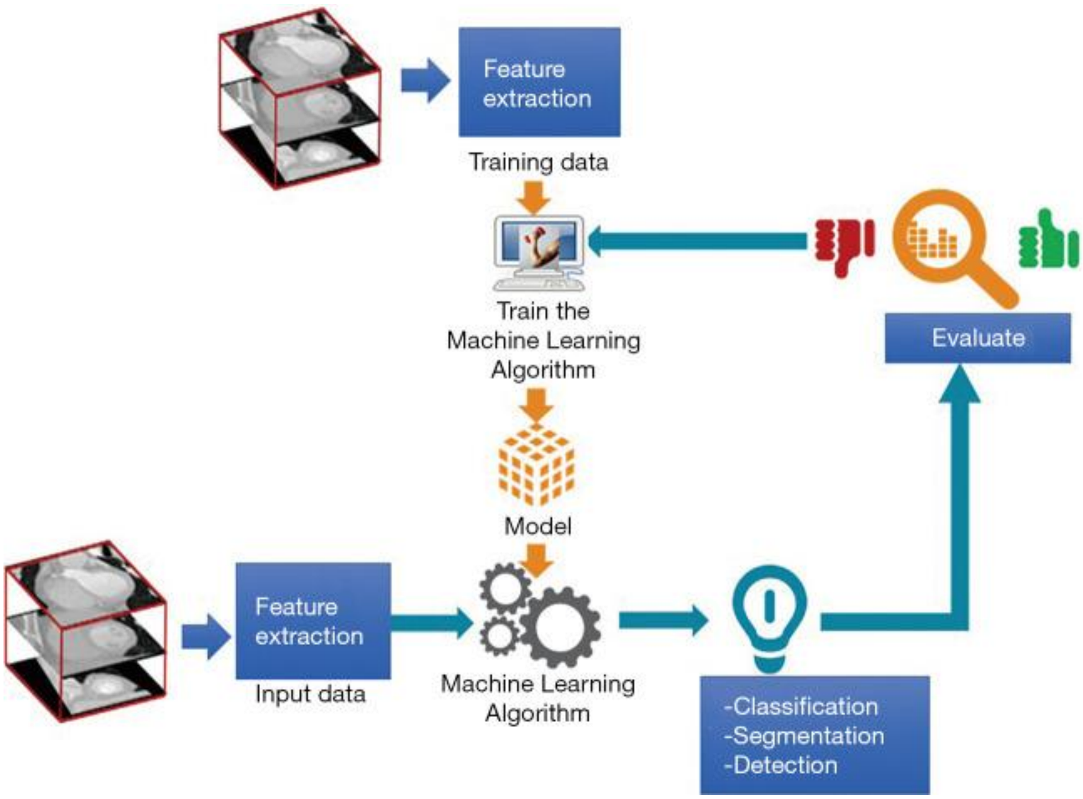

Figure 1: The flow of computer vision and machine learning interaction in medical imaging (Olveres et al., 2021).

There is a strong argument that doctors should be equal participants in the developmental process of image-computing technology. To gain the confidence of the primary users (radiologists), detailed explanations of the decisions are needed, especially when vision algorithms are used for predicting diseases. Analysis algorithms should be fully interpretable to decision-makers to ensure that they can be regarded as trustworthy for influencing decision-making. If this transparency is present, radiologists will know what to watch out for in biopsy preparation to decrease the chances of false diagnoses (Shah et al., 2022). This process of having formal involvement in the design process of advanced, high-stakes technology by those who will be using it is referred to as Co-Design (Donia et al., 2021).

Latour (1992) describes actors, such as humans and computer models, in his ANT framework as equal parts of a socio-technical relationship. He emphasizes the importance of studying the relationships and associations between these actors, rather than focusing solely on individual entities or overarching social structures. In the case of computer vision in medical image analysis, the key actor relationship is the data. By understanding the importance of data variety, quality, and quantity, doctors can assist the models in making better predictions and reducing discrimination against specific groups. Furthermore, this increases the knowledge medical professionals have about the technology they are using, allowing them to understand their limitations and trustworthiness given specific circumstances, also defined by Latour as *interessement*. A collaboration of decision-makers with the medical vision system can shape and transform the social and material world. Latour's power dynamics, which implies that no one actor group holds hierarchical power over another, was expanded by Greenhalgh and Stones (2010) to discuss the sharing of power between patients, doctors, hospitals, and even government entities due to their laws and restrictions on data privacy in healthcare. As actor groups influence each other via data, power is shared to create a single, fully informed decision-maker.

Research and Methods

As exemplified, there are varied concerns with the use of computer vision and machine learning models to automate medical imaging. This includes the accuracy and representation of the training data itself as well as the power they hold in decision-making for the patients. This leads me to ask: to what extent does implicit bias impact the use of computer vision by healthcare imaging centers in the United States?

In addition to conducting literature reviews of computer vision models employed in medical imaging centers and examining recent legislative actions, I plan on interviewing radiologists at the University of Virginia Health System. UVA Health recently implemented the Syngo Virtual Cockpit in its imaging to improve prediction accuracy and productivity (Lin-David, 2023), which opens new doors in the medical automation field, however, it may pose concerns to the doctors or patients. Understanding the primary users' views and procedures regarding this technology will provide insights into the current interessement and what steps can be taken to improve it. Potential questions include but are not limited to: what are the accuracy trends by demographic, if they are aware of potential bias or limitations, what their views are on data privacy, and any examples of past situations of disparity or influence by the automated system.

Conclusion

I developed a Vision Artificial Intelligence model to identify attendees and track their movements to increase operating efficiencies in sports stadiums, such as those rooted in concessions and foot traffic. The STS research outlined in this prospectus discusses the social implications of the increasing use of computer vision in the healthcare industry. These insights provide a comprehensive review of the current use cases, specifically medical image analysis, and their development-use cycle in hopes of identifying improvements to the data collection and decision-making process.

Both topics utilize machine learning for computer vision models for diverse reasons. However, each introduces issues that must be considered in the developmental and use processes to align with the overall goals of improving societal functions.

References

Andrews, C. J. (2006). Practicing technological citizenship. *IEEE Technology and Society Magazine*, *25*(1), 4.

Ciabarra , C. (2019, December 18). Council Post: AI Vs. Privacy: Does Computer Vision Violate Your Right To Privacy? *Forbes*. https://www.forbes.com/sites/forbestechcouncil/2019/12/18/ai-vs-privacy-does-computer-vision-violate-your-right-to-privacy/?sh=7d753dfd1327

Donia, J., & Shaw, J. A. (2021). Co-design and ethical artificial intelligence for health: An agenda for critical research and practice. *Big Data & Society*, 8(2). https://doi.org/10.1177/20539517211065248

El Naqa, I., Murphy, M.J. (2015). What Is Machine Learning? *Machine Learning in Radiation Oncology*. https://doi.org/10.1007/978-3-319-18305-3_1

Flah, M., Nunez, I., Chaabene, W.B., Nehdi, M.L. (2021). Machine Learning Algorithms in Civil Structural Health Monitoring: A Systematic Review. *Archives of Computational Methods in Engineering* 28, 2621–2643. https://doi.org/10.1007/s11831-020-09471-9

Gao, J., Yang, Y., Lin, P., & Park, D. S. (2018). Computer Vision in Healthcare Applications. *Journal of Healthcare Engineering*, 5. https://doi.org/10.1155/2018/5157020

Greenhalgh, T., Stones, R., (2010). Theorising big IT programmes in healthcare: Strong structuration theory meets actor-network theory. *Social Science & Medicine*, 70(9), 1285-1294, https://doi.org/10.1016/j.socscimed.2009.12.034.

Lara, R., Echeveste, R. & Ferrante, E. (2022). Addressing Fairness in Artificial Intelligence for Medical Imaging. *Nature Communications*, 13, 4581 https://doi.org/10.1038/s41467-022-32186-3

Latour, B. (1992). Where Are the Missing Masses? The Sociology of a Few Mundane Artifacts. In W. E. Bijker & J. Law (ed.), *Shaping Technology / Building Society: Studies in Sociotechnical Change*, 152-179.

Lin-David, H., (2023, August 8). How UVA Radiology's Virtual Cockpit Expands Care. *Radiology and Medical Imaging*, https://med.virginia.edu/radiology/2023/08/08/how-uva-radiologys-virtual-cockpit-expands-care/.

Neuroscience News. (2023, April 26). Unveiling the Hidden Biases in Medical AI: Paving the Way for Fairer and More Accurate Imaging Diagnoses. *Neuroscience News*, https://neurosciencenews.com/ai-medical-imaging-23117/#:~:text=One%20of%20the%20main%20sources

Olveres, J., González, G., Torres, F., Moreno-Tagle, J. C., Carbajal-Degante, E., Valencia-Rodríguez, A., Méndez-Sánchez, N., & Escalante-Ramírez, B. (2021). What Is New In Computer Vision and Artificial Intelligence in Medical Image Analysis Applications. *Quantitative Imaging in Medicine and Surgery*, 11(8), 3830-3853. https://doi.org/10.21037/qims-20-1151

Sarvamangala, D. R., & Kulkarni, R. V. (2022). Convolutional Neural Networks in Medical Image Understanding: a Survey. *Evolutionary Intelligence, 15*(1), 1–22. https://doi.org/10.1007/s12065-020-00540-3

Senior, A., Pankanti, S., Hampapur, A., Brown, L., Ying-Li, T., Ekin, A., Connell, J., Fe Shu, C., Lu, M., (2005). Enabling Video Privacy Through Computer Vision. *IEEE Security & Privacy*, 3(3), 50-57. https://ieeexplore.ieee.org/document/1439502

Shah, S. M., Khan, R. A., Arif, S., Sajid, U. (2002). Artificial Intelligence for Breast Cancer

Analysis: Trends & Directions. *Computers in Biology and Medicine*, 142, https://doi.org/10.1016/j.compbiomed.2022.105221.

Siddell, K. (2023, May 1). Computer Vision is Driving Huge ROI for Businesses. *AlwaysAI*, https://alwaysai.co/blog/computer-vision-drives-huge-roi

Solawetz, J., & Francesco. (2023, January 11). What is YOLOv8? The Ultimate Guide. *Roboflow Blog*. https://blog.roboflow.com/whats-new-in-yolov8/

Voulodimos, Athanasios; Doulamis, Nikolaos; Doulamis, Anastasios; Protopapadakis, Eftychios. (2018). Deep Learning for Computer Vision: A Brief Review. *Computational Intelligence and Neuroscience*, 1-13. https://doi.org/10.1155/2018/7068349

West, C. P., Dyrbye, L. N., & Shanafelt, T. D. (2018). Physician Burnout: Contributors, Consequences and Solutions. *Journal of Internal Medicine*, 283(6), 516-529. https://doi.org/10.1111/joim.12752

Winner, L. (1980). Do Artifacts Have Politics? *Daedalus*, 109(1), 121–136. http://www.jstor.org/stable/20024652