Introduction of Artificial Intelligence in the Healthcare Industry

A Research Paper submitted to the Department of Engineering and Society

Presented to the Faculty of the School of Engineering and Applied Science

University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements for the Degree

Bachelor of Science, School of Engineering

Quinn Lewis

Spring 2021

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

Advisor
Prof. S. Travis Elliott, Department of Engineering and Society

Introduction

The purpose of this paper is to analyze the role Artificial Intelligence will have in various medical fields. Specifically, the analysis focus on fields in which physicians rely on medical imagery and scans to diagnose patients such as radiology, cardiology, and more. As advancements have been made in Artificial Intelligence, Machine Learning, and Deep Learning algorithms, applications for automated diagnosis have increased. From 2007 to 2017, research publications about AI in radiology increased from 100-150 to 700-800 papers (Pesapane, F., Codari, M. & Sardanelli, F, 2018). While these papers have shown very promising results, there has not been a widespread clinical acceptance of AI in many fields. This is due to a combination of human and technical factors which I plan to address in this paper. I conclude with insights on how the many factors in this complex system could best work together in future clinical settings to provide best results for patients. I will be utilizing the Actor Network Theory (ANT) framework because of its focus on interactions between human and non-human factors within a system (Crawford, 2020).

Background

Developments in AI, Machine Learning, and Deep Learning

Artificial Intelligence is a popular, general term used to discuss future developments in many topics such as driverless cars, automated warehouses, and social media ads. While discussion and research around AI has been rapidly increasing within the last decade, AI has been around since the 1950s (Pesapane, F., Codari, M. & Sardanelli, F, 2018). The idea of algorithms learning as they receive more and better data is called Machine Learning and has long been used for statistical analysis. This recent fascination in AI is due to major developments in Deep Learning algorithms.

Most of these new developments have been in the field of Convolution Neural Networks (CNNS). In CNNs, every pixel of an image is passed as input data through multiple layers of operators before outputting classifications. These layers, or kernels, interact with each other, mimicking the neurons of our brains. At every kernel, mathematical operations like Convolution and Transformations are applied to images to recognize hidden patterns that may not be seen by the naked eye. After seeing many images, the network learns features that are frequent in unhealthy patients and can begin identifying them in others. Neural Networks are named after our brains because they learn as they experience more and more data like humans.

In the same way CNNs help driverless cars distinguish between stop signs, other cars, and children in the road, CNNs can also spot tumors and abnormalities in human organs and tissues. In a 2019 Stanford competition, a start-up group's AI was able to more accurately diagnose lung disease than a team of Stanford radiologists (Pennic, 2019). Similarly, in 2020, Google established that their AI could detect breast cancer with greater accuracy than experts (McKinney, 2020). Many studies have found that while these technologies show promise, they are a product of their small sample sizes. Since CNNs learn only from the data they are given, lack of diversity in a training set can lead to inconsistent results when applied to a broader population. This is the same phenomenon that leads to facial recognition algorithms to less accurately recognize people of color's faces because the training data includes predominately white faces (Garvie, Clare, and Jonathan Frankle, 2016). This exemplifies why diverse data is a key factor in the development of this technology.

STS Framework

I will use the Actor Network Theory (ANT) to analyze the many factors that interact in the use of Deep Learning in the medical industry. ANT looks at a complex system and attempts to understand how all the agents work together. Notably, ANT differs from other frameworks as it places a large emphasis on how non-human factors interact with human factors. This framework considers these agents as equals in their importance to the network. This framework also differentiates between intermediators and mediators. Mediators are agents that are integral to the system and its function while intermediators simply flow through the system. This important distinction allows us to acknowledge all factors in a system while focusing on the differential ones.

ANT is perfect for my analysis of AI in medical imagery fields because of its emphasis on human as well as non-human factors. Understanding how these factors interact to produce effective outcomes for patients and providers will be the majority of this analysis. Since Deep Learning has not been applied in most fields clinically, a portion of this discussion will be devoted to how these networks should interact to perform effectively. The primary mediators in this network that I will focus on are the algorithms, data, medical professionals, and the general health organizations that set the standards and laws. As this technology becomes more prevalent in all fields, these four factors will shape each other. The intermediators are primarily the patients who seek medical care.

Analysis

Algorithms and Data

The use of clean, accurate, and representative data is the most important factor in the success of CNNs. Because CNNs learn from experience like human brains do, the best neural networks will be trained by diverse and representative datasets containing millions of images. Currently, these datasets do not exist at this scale. This is why many critics disregard the results of competitions between CNNs and experts like the Stanford and Google cases. In a majority of these competitions, software developers did not show their algorithms could work on populations not represented in their training data (Liu X, Faes L, Kale AU, Wagner SK, Fu DJ, Bruynseels A, et al, 2019). In order to demonstrate a high success rate of their software, developers design experiments showing effectiveness on a very small and precise population. This problem is known as overfitting. Just like with the racial bias of facial recognition, the CNN expects every image of a scan to look like the scans it was trained with. This causes the algorithm to fail in recognizing tumors and differences in other source populations and to expect features of the training population in other source populations. The answer to this problem is increasing access to diverse, uniform, and large datasets.

The need for collaboration to construct large datasets has become more apparent. The 21st Century Cures Act has attempted to encourage the sharing of valuable image data (US Congress, 2016). Stanford University makes large datasets available for multiple scans including chest X-rays and CT scans (Stanford AIMA, 2021). While these datasets are a good start, they are nowhere near the scale needed for representative training data. These also highlight an important shortfall of CNNs. Because of this need for large datasets, only the most frequently performed scans will become accurate (Hosny, Parmar, Quackenbush, Schwartz, & Aerts , 2018). For this reason, Chest X-rays, mammograms, and CT scans will likely be the first scans to use widespread AI diagnosis.

Another key factor is the need for clean and standardized images. Noisy images are images that have markings caused by dust, scanning, or other conditions. Noise in an image is like static in sound quality and prohibits CNNs from recognizing patterns. AI experts can use techniques like augmentation and deformation to denoise images. Additionally, there is a need for standardization in labeling and sharing datasets. It is important to know which images contain tumors or abnormalities so the CNN can look for these patterns in other images.

Medical Professionals, Organizations, and Standards

The primary human agents in this network include the medical professionals who will use this technology in their careers. Additionally, organizations composed of these individuals have the collective power to foster or hinder the emergence of this technology. Although the president of the Radiology Society of North America gave a positive outlook on the future of AI in Radiology, she acknowledged the “fear that has caused anxiety within radiological societies” (Neri, de Souza, &Brady, 2019). While there is a history of Artificial Intelligence driven automation replacing workers in many industries, there is no significant evidence indicating that AI will replace Radiology. Rather, new developments in technology can be seen as a way to supplement Radiologists. Artificial Intelligence can greatly reduce the time Radiologists spend in image analysis for certain abnormalities while giving doctors more time for other clinical functions. Cooperation among experts and technology has the potential to be the most beneficial outcome for patients.

Currently, organizations as well as individual clinicians have great power in deciding how this technology is used. Even after researchers have proven its effectiveness and gotten it approved by the FDA for clinical use, Deep Learning algorithms need to prove “clinical utility” to be recommended practice by experts, academic societies, and third-party organizations (Morra, Delsanto, & Correale, 2019). These groups must agree that the benefits of new technologies improve outcomes for patients. Better outcomes can be realized in costs and accuracy. The fear that many Radiologists have about this emerging technology could motivate industry leaders to withhold their recommendations for Deep Learning to be used in clinical workplaces. While concerns about algorithm overfitting are real, these organizations are going to need to be forward thinking in how they interact with this technology.

Discussion

When AI engines were capable of beating international grandmasters in chess, many people thought it was the end of chess. In reality, these chess computers have empowered chess players to reanalyze their matches and become substantially more skilled than previous generations. The advancement in these chess engines have evolved the game rather than destroy it. Teams comprised of human and AI partners have outperformed humans or AI individually (Duca Iliescu, Delia Monica, 2020).

Much like in the game of chess, the success of Convolution Neural Networks in clinical practices relies heavily on the discussed actors working together. Much of the dialogue around this topic pits the medical professions against the developing technology; however, this idea of conventional radiology vs. AI is not accurate (Pesapane, F., Codari, M. & Sardanelli, 2018). This dichotomy harbors the fear of Radiologists who see diagnosis competitions as attempts to replace their jobs. As is the case with Computer Aided Detection (CAD), which has long been used for breast cancer detection, the medical field should utilize evolving technology. Many researchers in the field have embraced the changing landscape of technology by recognizing the need for a more cooperative approach.

New appreciations for the future of Deep Learning in diagnostics should lead to differences in curriculum for Radiologists. These students should be exposed to more computer science and algorithm topics so that they understand how Digital Process Imaging works. Having advanced knowledge of the subject will make them better partners with AI and allow them to interact more beneficially with the software. This will also give them an understanding of the strengths and weaknesses of algorithms. It is important to recognize that CNNs are not perfect and have shortcomings.

There remains to be people against the acceptance of CNNs in the medical fields. They remain skeptical about the shortfalls of current algorithms' ability to generalize over populations. It is fair to say that Radiology cannot be significantly improved with existing technologies today; however, these technologies are rapidly improving. The clinicians who get on board with these changes today will be better suited to be competitive when AI offers major improvements to efficiency.

Conclusion

Deep Learning algorithms have shown promise in their ability to diagnose diseases and abnormalities across various medical fields. Lack of diverse, accurate data and inadequate designs have produced unfavorable results in some studies. As the algorithms and datasets improve, clinicians will have to make a choice of whether they want to partner with AI in their workplace. The support and recommendations of clinical professionals as well as medical organizations will determine how quickly this technology can be utilized in clinical practice.

References

Pesapane, F., Codari, M. & Sardanelli, F. Artificial intelligence in medical imaging: threat or opportunity? Radiologists again at the forefront of innovation in medicine. Eur Radiol Exp 2, 35 (2018). https://doi.org/10.1186/s41747-018-0061-6

Crawford, T. (2020, September 28). Actor-Network Theory. Oxford Research Encyclopedia of Literature. Retrieved 10 May. 2021, from https://oxfordre.com/literature/view/10.1093/acrefore/9780190201098.001.0001/acrefore-9780190201098-e-965.

Pennic, Jasmine. “AI vs. Humans: AI Solution Beats Stanford Radiologists in Chest X-Ray Diagnostics Competition.” Healthcare IT News, 23 Aug. 2019, hitconsultant.net/2019/08/22/ai-tech-beats-radiologists-in-stanford-chest-x-ray-diagnostic-competition/.

McKinney, S.M., Sieniek, M., Godbole, V. et al. International evaluation of an AI system for breast cancer screening. Nature 577, 89–94 (2020). https://doi.org/10.1038/s41586-019-1799-6

Garvie, Clare, and Jonathan Frankle. "Facial-recognition software might have a racial bias problem." The Atlantic 7 (2016).

Liu X, Faes L, Kale AU, Wagner SK, Fu DJ, Bruynseels A, et al. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. Lancet Digit Health. (2019) 1:e271–97. doi: 10.1016/S2589-7500(19)30123-2

Commissioner, Office of the. “21st Century Cures Act.” U.S. Food and Drug Administration, FDA, www.fda.gov/regulatory-information/selected-amendments-fdc-act/21st-century-cures-act.

Center for Artificial Intelligence in Medicine and Imaging, Shared Datasets, https://aimi.stanford.edu/research/public-datasets

Hosny, A., Parmar, C., Quackenbush, J., Schwartz, L. H., & Aerts, H. (2018). Artificial intelligence in radiology. Nature reviews. Cancer, 18(8), 500–510. https://doi.org/10.1038/s41568-018-0016-5

Neri, E., de Souza, N., Brady, A. et al. What the radiologist should know about artificial intelligence – an ESR white paper. Insights Imaging 10, 44 (2019). https://doi.org/10.1186/s13244-019-0738-2

Morra, L., Delsanto, S., & Correale, L. (2019). ARTIFICIAL INTELLIGENCE in MEDICAL IMAGING: From Theory to Clinical Practice. CRC Press: Place of publication not identified.

Duca Iliescu, Delia Monica. “The Impact of Artificial Intelligence on the Chess World.” JMIR serious games vol. 8,4 e24049. 10 Dec. 2020, doi:10.2196/24049