HealthcAIre: An Examination of the Current State of Data Science Applications in Disease Detection

Dhiraj Bankar
Bellevue University

Grant Parr
Bellevue University

Sarah Holdmn
Bellevue University

Christopher Salinas
Bellevue University

Abstract

As the largest sector of employees in the United States, health care workers may get nervous when they think about artificial intelligence (AI) arriving in their industry. When many of us hear the term artificial intelligence we imagine robots doing our jobs, rendering us obsolete. Since AI-driven computers are programmed to make decisions with little human intervention, there is a real sense that objective machines may soon make the difficult decisions we now entrust to the subjective humanity of doctors, nurses and medical professionals.

The reality is that AI already exists in empowering disease diagnosis in the health care industry and it has for some time. In this paper we will focus on the status of AI driven data science analysis in the diagnostic treatment of medical patients. We will research how data scientists analyze vast amount of patient data sets to create algorithms to identify skin cancer and retinopathy in their nascent stages. Specifically, we will analyze existing research on how data science is used in preventative treatments. We will look at the percentage of current preventative treatments that use AI technology and test their effectiveness against more traditional methods.

Author Keywords

Artificial intelligence; health care; big data; patient data; cancer diagnosis; skin cancer

Introduction

With the advancement in Artificial Intelligence and Machine Learning within the past few years, it is only reasonable that companies find opportunities for these techniques to empower treatment methods in the health care industry. These tools can enable earlier screening, treatment and diagnosis of disease, and bring us closer towards preventative and proactive healthcare. Today, two of the leading causes of preventable blindness in adults [25] are diabetic retinopathy (DR) and diabetic macular edema (DME). More than 2 in 5 Americans [24] with diabetes have some stage of diabetic retinopathy. Both of these diseases are serious and can lead to blindness and

retinal detachment. In recent years, Google and Verily (Alphabet's healthcare and life sciences arm) have been researching the use of machine learning to screen for these diseases. Similarly, 25% of the worlds medical examinations deal with skin diseases. Skin cancer is the leading cause of death from these skin diseases. Within the last 5 years, researchers at Google and Stanford have created image processing algorithms to detect skin cancer early through nothing more than images. In our research, we will take a look at the development of a machine learning algorithm to power automated screening with the intention of detecting disease earlier and expanding access to screening for DR, DME, and skin cancer. We will also look at the history of AI, its historical origins, how it works, and how it will can and will change the way eye care is provided.

AI's Use in Diagnosing Skin Cancer

For the last four years AI has been battling the most common human malignant: skin cancer. It is estimated that one quarter of all treatments provided to patients around the world are for skin conditions. A shortage of dermatologists and the high cost associated with seeing a skin specialist has create a need for another means to address skin conditions. AI and data science is in a unique position to help those suffering from skin conditions. Researchers at Google and Stanford University created AI solutions to meet patient needs. The Google and Stanford researchers used deep convolutional neural networks (CNNs) to auto classify skin lesions using images. The Stanford researchers trained a CNN using a dataset of 129,450clinical consisting of 2,032 different diseases [1]. They tested the CNN's performance against 21 board-certified dermatologists on biopsy-proven clinical images. The Stanford CNN's performance was on par with all tested experts. In fact, CNN achieved 72.1% overall accuracy while the average accuracy for the 21 board-certified dermatologists was 65.56%.

Google researchers underwent a similar test in 2019 [2]. The research team evaluated the model with 17,777 de-identified cases from 17 primary care clinics. They bifurcated the cases and used the portion of records dated between 2010 and 2017 to train the AI system, reserving the portion from 2017 to 2018 for evaluation. During training, the model leveraged over 50,000 differential diagnoses provided by over 40 dermatologists.

In a test of the system's diagnostic accuracy, the researchers compiled diagnoses from three U.S. board-certified dermatologists. Just over 3,750 cases were aggregated to derive the ground truth labels, and the AI system's ranked list of skin conditions achieved 71% and 93% top-1 and top-3 accuracies, respectively. Furthermore, when the system was compared against three categories of clinicians (dermatologists, primary care physicians, and nurse practitioners) on a subset of the validation data set, the team reports that its top three predictions demonstrated a top-3 diagnostic accuracy of 90%, or comparable to dermatologists (75%) and "substantially higher" than primary care physicians (60%) and nurse practitioners (55%).

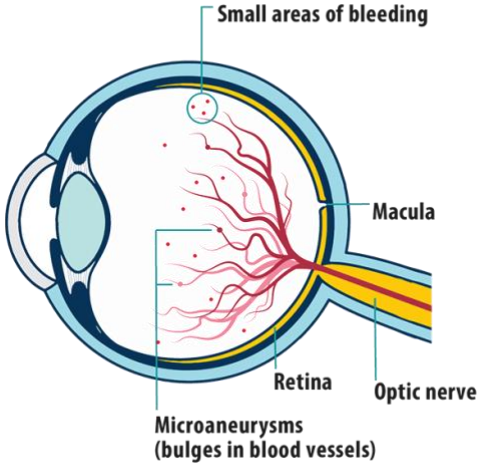
This demonstrates an artificial intelligence capable of classifying skin cancer with a level of competence comparable to dermatologists. Outfitted with deep neural networks, mobile devices can potentially extend the reach of dermatologists outside of the clinic. With the prevalence of mobile devices there is great

potential to provide low-cost universal access to vital diagnostic care.

Detecting Diabetic Retinopathy

Referable diabetic retinopathy is defined as moderate or severe diabetic retinopathy or referable diabetic macular edema. [4] If diabetic retinopathy is caught early enough, it can be treated without leading to irreversible blindness. The process of detecting this disease involved having the specialist examine pictures of the back of the eye of the patient. They then rate them for disease severity based on the type of lesions present, such as hemorrhages and microaneurysms. These lesions indicate bleeding in the eye. Examining these pictures require specialized training, and there aren't enough of these specialists to screen everyone at risk in many parts of the world.

Google researchers worked with doctors in the US and India to compile a dataset of 128,000 images which were looked at by 3-7 ophthalmologists out of a panel of 54. [4] This dataset was then used to train a deep neural network to detect diabetic retinopathy. The parameters of the neural network are initially set to random values. Then, the severity grade given by the function is compared with the known grade from the training set. The parameters are then modified to decrease the error on that image. [3] This process is done for every image in the training set many times through. This is how the function "learns" how to compute the severity of diabetic retinopathy from the pixel intensities of the images in the training set. The resulting function can calculate the DR severity on new images with the right training data.



Researchers tested this algorithm in two clinical validation sets of roughly 12,000 images against 7-8 board-certified ophthalmologists as the reference standard. They found that the algorithm (at the operating point selected for high specificity) had 90.3% and 87.0% sensitivity and 98.1% and 98.5% specificity for detecting referable diabetic retinopathy. [3] Google researchers concluded that deep learning algorithms had high specificity and sensitivity for detecting diabetic retinopathy and macular edema in retinal fundus photographs. [3] While this is a great start, further research is still needing to determine feasibility of applying this algorithm in the clinical setting. It also needs to be determined whether it would lead to improved care and outcomes compared with the current ophthalmologic assessment. Automated diabetic retinopathy screening methods have a strong potential to help doctors evaluate patients more quickly

and can get those who need help sent to a specialist. Given the advances in deep learning recently, this is just one example to demonstrate the ability of machine learning to help solve important problems in healthcare.

Ethics of AI in Healthcare

The ethics of machine learning within the medical field can be often overlooked in search of potentially saving lives, however, there are real ethical dilemmas when implementing machine learning in a person's medical decisions. In a recent study conducted in the United Kingdom, 83% of the people survey understood the need for AI in improving medical decisions, however, 63% of those surveyed were uncomfortable having their personal information used to improve medical decision. [6].

Many factors must be considered with the implementation of AI, such as, how accurate is the information (inputs and outputs), are there biases within the data (overt or disparate), is the information transparent (proper disclosures of use), and having the patient understand how much of the AI is used for a final determination to ensure the doctor's experience is factored in. *D*octors also shared concerns with the use and implementation of AI. 50% of the doctor's surveyed cited the following biggest concerns: AI will be responsible for a fatal error, AI will be poorly implemented or won't work properly, and AI will be overhyped and not meet expectations [7].

Conclusion

In years to come, artificial intelligence will continue to play a significant role in the way healthcare providers care for their patients. There is a clear and imminent potential for earlier, more accurate, and more uniform diagnosis, as we have seen in the detection of diabetic retinopathy, diabetic macular edema, and skin cancer. AI can improve the classification of disease and will open doors to unseen avenues for developing therapeutic interventions. It will also allow us to discover latent features in ophthalmic images, clinical profiles, and genetics, [25] which will improve our understanding of eye disease. We can venture to guess that in the near future, eyecare providers will have devices that, take an image of a patient's retina, and tell of the exact diagnosis at the touch of a button. Likewise, we can expect dermatologists to use image classifying AI to help patients detect skin cancer at early stages. Artificial intelligence and machine learning systems will undoubtedly reshape the eyecare and skincare industry in the coming years, and open doors for other applications in healthcare.

Acknowledgements

We thank Prof. Shankar Parajulee for providing us with insight to tackle this project. We thank Bellevue University's DSC500-T303 class for giving us useful, insightful feedback to improve our research.

References

1. Dermatologist-level classification of skin cancer with deep neural networks Andre Esteva1*, Brett Kuprel1*, Roberto A. Novoa2,3, Justin Ko2, Susan M. Swetter2,4, Helen M. Blau5 & Sebastian Thrun
2. A Deep Learning system for differential diagnosis of skin diseases, Yuan Liu, Ayush Jain, Clara Eng, David H. Way, Kang Lee, Peggy Bui, Kimberly Kanada, Guilherme de Oliveira Marinho, Jessica Gallegos, Sara Gabriele, Vishakha Gupta, Nalini Singh, Vivek Natarajan, Rainer Hofmann-Wellenhof, Greg S. Corrado, Lily H. Peng, Dale R.

Webster, Dennis Ai, Susan Huang, Yun Liu, R. Carter Dunn, David Coz

3. Gulshan V, Peng L, Coram M, et al. Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. JAMA. 2016;316(22):2402–2410. doi:10.1001/jama.2016.17216
4. Abràmoff MD, Folk JC, Han DP, et al. Automated analysis of retinal images for detection of referable diabetic retinopathy. JAMA Ophthalmol. 2013;131(3):351-357. [2]Anna Cavender, Shari Trewin, Vicki Hanson. 2014. Accessible Writing Guide. Retrieved August 22, 2014 from http://www.sigaccess.org/welcome-to-sigaccess/resources/accessible-writing-guide/
5. On the ethics of algorithmic decision-making in healthcare – Grote & Berens
6. Machine learning in medicine: Addressing ethical challenges – Vayena, Blasimme & Cohen
7. Overcoming Barriers in AI Adoption in Healthcare - https://newsroom.intel.com/wp-content/uploads/sites/11/2018/07/healthcare-iot-infographic.pdf
8. Predictive Analytics Solutions: Defining Predictive Analytics in Healthcare
9. 5 Reasons Healthcare Data Is Unique and Difficult to Measure - Dan LeSueur
10. Data Science for Healthcare: What Today's Leaders Must Know - Michael Levy Scott R. Klemmer, Michael Thomsen, Ethan Phelps-Goodman, Robert Lee, and James A. Landay. 2002. Where do web sites come from?: capturing and interacting with design history. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '02), 1-8. http://doi.acm.org/10.1145/503376.503378
11. Artificial Intelligence As a Medical Device in Radiology: Ethical and Regulatory Issues in Europe and the United States - Filippo Pesapane
12. Medical Ethics Considerations on Artificial Intelligence - Kadircan H.Keskinbora Marilyn Schwartz. 1995. *Guidelines for Bias-Free Writing.* Indiana University Press, Bloomington, IN.
13. Ethical Dimensions of Using Artificial Intelligence in Health Care - Michael J. Rigby
14. Dermatologist-Level Classification of Skin Cancer with Deep Neural Networks - Andre Esteva
15. Machine Learning and Artificial Intelligence Research for Patient Benefit: 20
16. Critical Questions on Transparency, Replicability, Ethics, and Effectiveness
17. Artificial Intelligence in Medical Imaging: Threat or Opportunity? Radiologists Again at the Forefront of Innovation in Medicine - Filippo Pesapane, Marina Codari
18. Applied Data Science in Patient-Centric Healthcare: Adaptive Analytic Systems for Empowering Physicians and Patients - Marco Spruit, Miltiadis Lytras
19. HEALTH BANK – A Workbench for Data Science Applications in Healthcare
20. Opportunities and Challenges of Artificial Intelligence in Healthcare - Oksana Iliashenko
21. Artificial Intelligence and Healthcare - Dina Ross
22. Artificial Intelligence in Healthcare: Past, Present and Future - Fei Jiang
23. Artificial intelligence-enabled healthcare delivery - Sandeep Reddy
24. Deep learning in ophthalmology: The Technical and Clinical Considerations- Daniel S.W. Ting

25. Zhang X, Saaddine JB, Chou C-F, et al. Prevalence of Diabetic Retinopathy in the United States
26. Raman R, Rani PK, Reddi Rachepalle S, et al. Prevalence of diabetic retinopathy in India