GLOBALAI

LEVERAGING THE POWER OF AI IN THE BATTLE AGAINST CANCER

AREEZ VISRAM
HANA MURJI

2020
INNOVATE 1Z03

EXECUTIVE SUMMARY

A resilient society in the context of the modern world is one that can turn to innovation in the face of adversity. As artificial intelligence continues to push and challenge the conventions of the world, humanity must work alongside this remarkable technology to pose solutions to ailments once considered incurable. This report outlines GlobalAI's proposal for a new application of artificial intelligence that assists medical professionals in diagnosing and treating cancer.

The report begins by illustrating the undiscriminating effects of cancer on humanity, including staggering statistics from across the globe. Next, a discussion on the proposed AI solution touches on the feasibility of implementation within the next five years and explains the desired outreach and target audience. The report details the specific AI techniques that will be used by the proposed algorithm, as well as measures that will be implemented to ensure patient privacy and minimize bias. Finally, similar projects will be explored in order to broaden understanding of the capabilities of deep learning in cellular biology and specifically cancer diagnosis, and a prototype model of the mechanism behind the proposed AI is introduced.

The goal of this report is to inform on the widespread effects of cancer, and demonstrate the potential of leveraging artificial intelligence to aid humanity in the battle against this deadly disease.

TABLE OF CONTENTS

INTRODUCTION

Cancer is the leading cause of mortality in the country, and doctors, researchers and families have racked their brains for decades trying to optimize the process of cancer diagnosis and treatment. According to the Canadian Cancer Society, nearly one in two individuals is expected to develop cancer in their lifetime (2016).

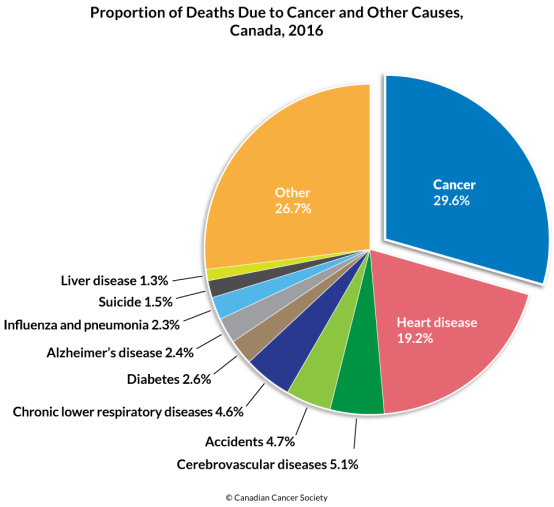

Figure 1: According to the Canadian Cancer Society, 29.6% of all deaths nationwide are due to cancer

The engineers at GlobalAI have developed an application of Artificial Intelligence that integrates nanotechnology, high resolution imaging, and years of scientific learnings to optimize cancer diagnosis, minimize human error, and suggest optimal treatment. This device creates an opportunity for affordable and accessible personalized healthcare while paving the way to better quality of life and care of patients and families affected by cancer.

DESCRIPTION

When a patient displays signs and symptoms of cancer, their physician will order a plethora of imaging tests and medical lab visits which include, MRIs, blood work, and endoscopies. Once diagnosed with cancer, further testing can determine the stage of cancer, grade of tumour, genetic markers, and the risk group of the patient. Although this process seems straightforward enough, it is estimated that nearly one in five cancer cases are misdiagnosed (BBC, 2020). In fact, one study found that 28% of the mistakes made in the process of diagnosis were life-threatening or life-altering (2020). A report in the Journal of the American Medical Association detailed the ten most prominent sources of error in misdiagnosing a cancer patient, and interestingly, each one related back to the negligence, malpractice, or simply misunderstanding of the physician (2018).

The proposed technology will be ingested via a capsule and programmed to investigate a region of interest within the body. The neural network will be trained on a dataset of chemical makeup to determine if there are cancerous cells present, the extent of the damage, the severity and stage of the cancer, and can be further extended to recommend the best course of action for treatment. With this revolutionary technology, there will no longer be a need for expensive, time-consuming, and potentially harmful tests, and the rate of misdiagnoses will plummet as the most common sources of human error are eliminated. This device has the capacity to build an entire pathology report with exceptional accuracy, impartiality, and efficiency. Amalgamating this information along with the chemical makeup of the individual, their physiology and lifestyle, it can propose the ideal plan of action that is unique to the patient and accommodating to their needs. By integrating the incredible capacity of deep learning and neural networks with the wealth of anonymous data at the disposal of medical researchers, there lies an opportunity to revolutionize the conventional method of cancer diagnosis and treatment and change millions of lives.

TARGET MARKET

As mentioned previously, nearly one in two individuals is expected to develop cancer in their lifetime (Canadian Cancer Society, 2016). As such, this AI has an extremely large target market. According to the World Health Organization, cancer is the second leading cause of death globally, and every one in six deaths is caused by cancer (2018). The AI being proposed will help doctors confirm the presence of cancer cells in a patient as well as identify the location and severity of the tumour. As shown in Figure 2 below, the probability of survival from cancer decreases significantly with time (Han, Lee and Park, 2019).

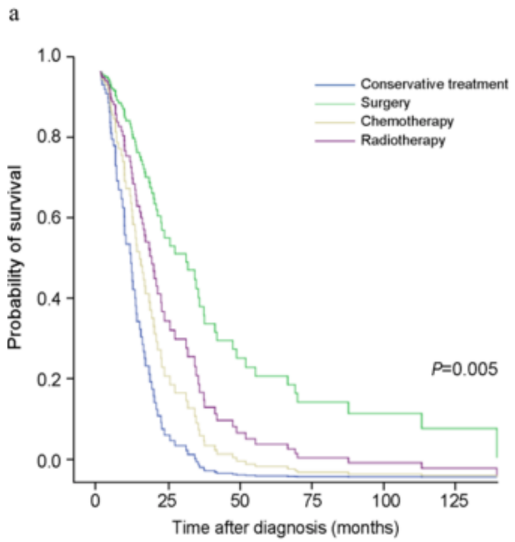

Figure 2: A graph showing a person's probability of survival vs the number of months after cancer diagnosis

As explained by Research UK, "Cancer that's diagnosed at an early stage, when it isn't too large and hasn't spread, is more likely to be treated successfully"(2018). The AI being proposed will assist doctors in diagnosing cancer at an earlier stage, allowing for a higher chance of survival with successful treatment. The AI will be utilized by cancer doctors worldwide, allowing them to effectively treat the many patients that suffer from cancer.

BIAS AND DATA PROTECTION

As discussed by Matthew Jordan in lecture six of Innovate 1Z03, two of the biggest concerns around AI and its applications in medicine are data protection and bias prevention. As shown in Figure 3 below, more than 70% of people listed “protect my personal info” as their biggest concern with AI today (Marketing Charts, 2017).

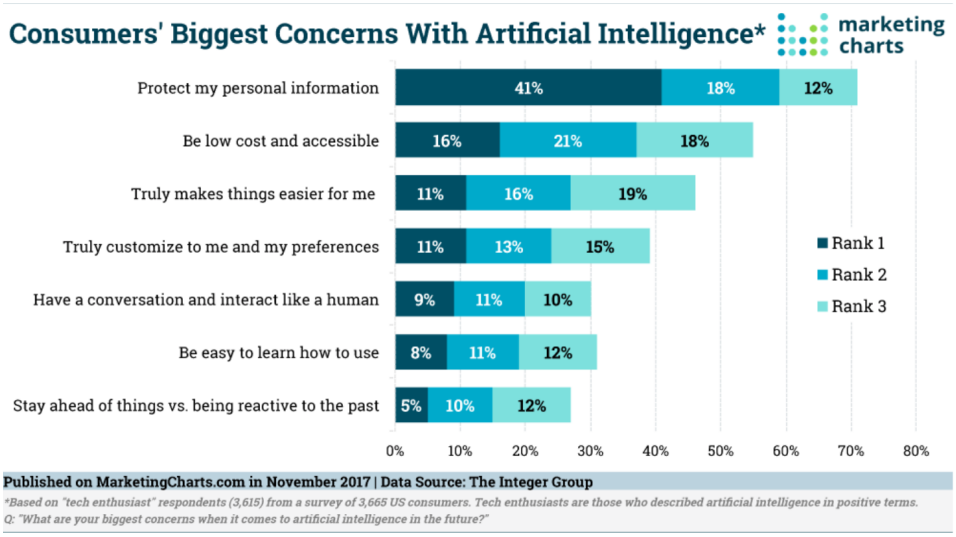

Figure 3: A chart showing consumers’ biggest concerns with AI in a ranking format

It can also be argued that the healthcare industry has the most personal information on citizens of any industry. As Matthew Jordan stated, “99.98% of Americans could be identified using just 15 attributes” (2020). It is not far-fetched to think that doctors and hospitals have access to all 15 of these attributes, and possibly more, needed to personally identify patients. It is for this reason that the AI being proposed will be trained on anonymous data, and will also have various privacy techniques implemented to allow for patient anonymity. The data that will be used to train the AI will be images of cells from random individuals including an equal amount of cancer and non-cancer patients, whose personal information will be protected from the parties developing the algorithm.

The AI will be taught using federated learning, which means the data will be accessed without it being seen at all (Jordan, 2020). The necessary data (cell images), will go straight from the labs to the neural network without being seen by anyone. This federated learning technique will ensure the utmost anonymity of patient data. Furthermore, the AI will be trained using differential privacy techniques. As explained by Matthew Jordan, differential privacy refers to the process by which it can be ensured that the dataset remains the exact same, even with one person's data removed (2020). These differential privacy and federated learning implementations will ease concerns of patient privacy and confidentiality with the proposed AI.

Another large concern regarding AI in medicine is bias in the dataset. As stated by James Manyika, Jake Silberg and Brittany Presten, "Bias can creep into algorithms in several ways. AI systems learn to make decisions based on training data, which can include biased human decisions…" (2019). In the AI being proposed, there are many examples of bias that can be found in the dataset such as race, gender, age, etc. Preventing bias in machine learning is a human responsibility, and the researchers responsible for training this neural network must ensure that there exists as little bias as possible in the dataset. This means that the data must be equally distributed in aspects such as age, gender, race, and patient type. There must be an equal amount of men as women, young people as old people, and all races must be represented proportionally. Furthermore, there must be an equal amount of data from cancer patients and non-cancer patients in the dataset. Ensuring these balances is the most important step that will be undertaken to ensure there exists as little bias as possible in the AI.

AI TECHNIQUES

The cancer-cell identification AI being proposed for development within the next five years relies heavily on AI techniques that have been used in AI applications for many years. The two most simple, yet arguably most important techniques, are computer vision and image recognition. The success of this AI is dependent on the computer's ability to recognize the cell patterns from a dataset of images. As said by Tristan Greene, "image recognition,...typically, involves the creation of a neural network that processes the individual pixels of an image. Researchers feed these networks as many pre-labelled images as they can, in order to "teach" them how to recognize similar images"(2018). Image recognition and computer vision will be needed to allow the AI to identify that the picture being looked at is indeed a cell. Next, the AI will use deep learning to be able to identify the patterns in the cell images. As explained by Diana Kwon, a biologist that was part of a research team that used deep learning to identify structures in cells, "deep learning, an artificial intelligence (AI) approach where algorithms learn to identify patterns in datasets"(2019). Deep learning would be used to train a convolutional neural network to analyze and classify the cell images as cancer cells, and be able to determine the body part, as well as the severity of any potential cancer. The final AI technique will be a technique used to ensure privacy and data protection, differential privacy. As said by Matthew Jordan, "we can use differential techniques to make sure that this compiled dataset is roughly the same even if we end up removing one person's brain scan from the pool."(2020). Differential privacy techniques will be used to ensure that removing one person's cell images will not affect the dataset, ensuring privacy and protection of every patient's data.

SIMILAR PRODUCTS

As stated by scientific writer Jef Akst, it is undeniable that "sophisticated machine learning models are changing how researchers interact with visual data" (2019). Developing such algorithms involves acquiring massive amounts of data to support the training of the model, which, when optimized, demonstrate extraordinary potential with applications ranging from cellular biology, economics, climate change, to virtually any other sector of human activity. There exist many similar products that GlobalAI can take inspiration from when developing the proposed AI.

Firstly, in 2018, Steve Finkbeiner published the results of a study conducted using his robotic microscopy system in conjunction with Google's approach to deep learning (Akst, 2019). The AI was successfully able to differentiate cell types and identify subcellular structures that traditionally, are only visible to the human eye after undergoing a staining process (2019). Fittingly, the article discussing this study is entitled, "AI sees more in microscopy than humans do," a testament to the profound capabilities that this technology has when paired with an abundance of high quality data.

Another study discusses a neural network that was trained to identify and discern the variations of cellular components in a manner more efficient and less costly than the scope of traditional technologies (Kwan, 2019). Biomedical engineer Greg Johnson and his team created their algorithm using a mere dozen images and their results were extraordinary. The team pioneered a more efficient method to perform 3D imaging and fluorescent labelling in the world of cellular and molecular biology (2019).

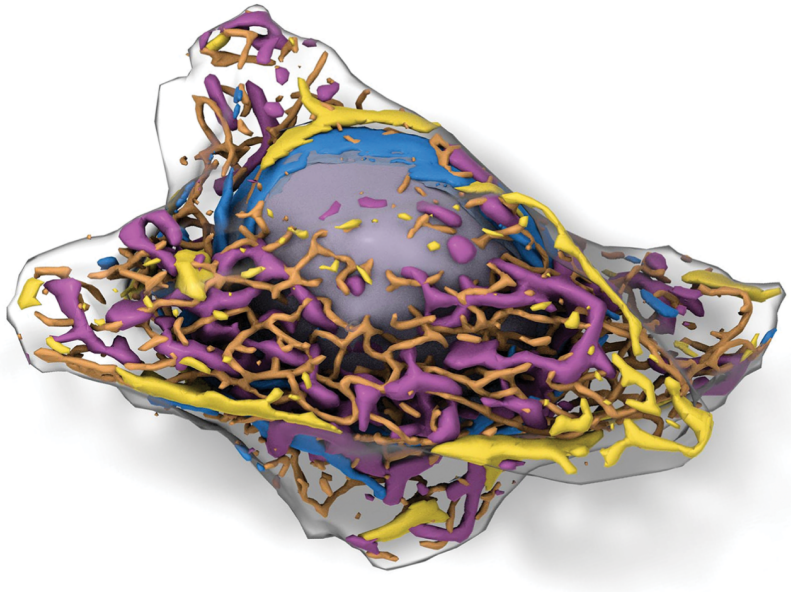

Figure 4: The 3D model of the cell created by Greg Johnson's algorithm

Finally, a group of doctoral candidates in Florida developed an AI that could detect cancerous cells in CT scans of lungs with an astonishing accuracy rate of 95%, compared to 65% when reviewed by humans (Washington, 2018). Their algorithm shared the approach of facial recognition software, and was trained by an initial one thousand CT scans depicting different features and stages of tumours. Along with learning what to identify, the AI had to learn what to ignore as well. The article explains that the system was taught to ignore other tissues, nerves, and masses it encountered in the scans (2018)

The above applications each boast impressive features that demonstrate the applications of deep learning in the field of cellular imaging and specifically, cancer diagnosis.

SAMPLE AI

Developed by Areez Visram

A sample algorithm that models the AI being proposed can be found at the following link (Please open from a Google Account):

https://colab.research.google.com/drive/1cisgJWxGkgEYJXXZyST0Znu5fSa-VEHB?usp=sharing

The instructions to run the AI are found in the Comments at the top of the cell in the notebook.

The above AI is just a model of what the proposed AI would do. It does not identify cancer cells. It simply uses image recognition to analyze and predict the composition of images. The AI works best for real-world objects such as cars or soccer balls.

CONCLUSION

As we race towards the future, artificial intelligence will continue to emulate human cognition in analyzing, interpreting, and comprehending medical data. The proposed application of AI would unite state-of-the-art deep learning technology as well as other AI techniques, respect and advocacy for patient privacy, sustained measures to avoid data bias, and be at the forefront of innovation. The AI would spark a genuine feeling of opportunity and hope for individuals and families who are affected by cancer. By optimizing the cost and accuracy of diagnosis, this device will revolutionize current conventions of cancer treatment and provide personalized healthcare to give patients important information and accessible options to support their fight against the disease. This device can be implemented at healthcare institutions across the globe in order to provide equitable access to this cutting edge, life saving technology. The integration of differential privacy and federated learning features will uphold the integrity of GlobalAI's commitment to patient privacy and confidentiality. Capitalizing on AI techniques such as image recognition, computer vision, deep learning, and differential privacy, this system will combine optimal functionality and accuracy to produce the highest quality output possible with the technology of the era. Finally, by building off the incredible research of cell biologists and computer engineers in the field, the engineers are able to borrow ideas and experiences in order to optimize the creation, training, and maintenance of this powerful AI. Given the aforementioned features, policies, and strategies, the cancer-detecting neural network is a new and incredible application of AI and will be a revolutionary addition to the scope of medical technology that exists today.

WORKS CITED

Akst, J. (2019, April 30). Artificial Intelligence Sees More in Microscopy than Humans Do. Retrieved from https://www.the-scientist.com/features/artificial-intelligence-sees-more-in-microscopy-than-humans-do-65746

Canadian Cancer Society. (n.d.). Cancer statistics at a glance - Canadian Cancer Society. Retrieved fromhttps://www.cancer.ca/en/cancer-information/cancer-101/cancer-statistics-at-a-glance/?region=on

CancerResearchUK. (2019, March 4). Why is early diagnosis important? Retrieved fromhttps://www.cancerresearchuk.org/about-cancer/cancer-symptoms/why-is-early-diagnosis-importnt#:~:text=Cancer that's diagnosed at an,of surviving are much lower.

Greene. (2020, May 19). A beginner's guide to AI: Computer vision and image recognition. Retrievedfrom https://thenextweb.com/artificial-intelligence/2018/07/18/a-beginners-guide-to-ai-computer-vision-and-image-recognition/

Han, S., Lee, Y.J., Park, J.S. et al. Prognosis of non-small-cell lung cancer in patients with idiopathicpulmonary fibrosis. Sci Rep 9, 12561 (2019).

Kwon, D. (2019, April 30). Deep Learning Algorithms Identify Structures in Living Cells. Retrieved fromhttps://www.the-scientist.com/notebook/deep-learning-algorithms-identify-structures-in-living-cels-65778?_ga=2.244697029.1360120774.1591996279-1723465584.1591996279

Manyika, J., Silberg, J., & Presten, B. (2019, October 25). What Do We Do About the Biases in AI?Retrieved from https://hbr.org/2019/10/what-do-we-do-about-the-biases-in-ai

One in five NI cancer diagnoses via emergency routes. (2020, January 15). Retrieved fromhttps://www.bbc.com/news/uk-northern-ireland-51124188

PTI. (2018, August 25). This AI system can identify cancer tumours better than humans - Latest News:Gadgets Now. Retrieved fromhttps://www.gadgetsnow.com/tech-news/this-ai-system-can-identify-cancer-tumours-better-than-umans/articleshow/65544540.cms

Tarver, T. (2018). Cancer Facts & Figures 2012. American Cancer Society (ACS). Journal of ConsumerHealth On the Internet, 16(3), 366–367. doi: 10.1080/15398285.2012.701177

WHO. (2018, September 12). Cancer. Retrieved fromhttps://www.who.int/news-room/fact-sheets/detail/cancer