Undergraduate Research Journal

Issue 16, Volume 2 April 2024

Diagnosis of Cancer Using AI Technology

Emily Yu

ACKNOWLEDGEMENTS

This paper was written for WRI 100: Advanced Writing with Yiran Xu.

Diagnosis of Cancer Using AI Technology

Emily Yu

Department of Natural Science, University of California, Merced

April 1, 2024

Abstract

The rapid growth of cancer is the second-leading cause of death in the U.S., and lung cancer has the highest mortality rate. This, in turn, creates a high demand for new technology to effectively treat this disease. This synthesis project delves into the capabilities of AI technologies that are specialized in diagnosing lung cancer at an early stage and their effectiveness in creating personalized treatment and risk prevention. By understanding more about the Explainable Artificial Intelligence (XAI) tool along with the Triplet Network method, the study reveals that it has a high accuracy rate (86.39%) on early diagnosis of lung cancer and can examine a patient's genome for cancer risk. The application of AI technology for the diagnosis of complex diseases like cancer is essential for maximizing patients' safety and recovery, along with the reassurance that medical professionals need. As AI technology continually improves over time, it may soon perform more complex tasks than simply scanning and analyzing data.

Keywords: AI technology, cancer, risk prevention

Diagnosis of Cancer Using AI Technology

Cancer is a group of diseases defined as the uncontrollable growth of defective cells in the body. These cells will continually grow and lump up as tumors. It can spread throughout the body if left untreated. This abnormal characteristic of the disease is why it is the second greatest cause of mortality in the US, which poses a serious global public health issue (Siegel et al., 2020). The five leading causes of cancer-related deaths in 2017 were lung and bronchus, prostate, colorectum, pancreas, and liver, which are around 78,600, 30,400, 27,700, 23,000, and 18,200, respectively. Siegel anticipated that there would be over 600,000 cancer-related deaths and 1.8 million new cancer cases in the U.S. in 2020. As such, identifying and getting it treated at an early stage is crucial to the patient's health. Once it becomes malignant, it will be exponentially hard to cure.

This is when the use of artificial intelligence in oncology – the study of cancer and tumors – becomes advantageous. Artificial intelligence has the capability to predict patient outcomes, determine the best treatment options, and accurately diagnose malignancies using pathology slides and radiological imaging (Chen et al., 2021). This is also where Explainable Artificial Intelligence (XAI) comes in. It was developed over the last decade and specified for tasks like cancer diagnosis. Upon data input, it can generate interpretable models, visualization, certainty scores, patient-specific explanations, and the ability to learn and apply knowledge from every previous observation. Understanding how well the XAI tool can detect and diagnose cancer is just as important.

Research Questions

How has AI technology changed the process of cancer diagnosis and risk prevention?

Capabilities of AI Technology

How Explainable Artificial Intelligence (XAI) Works

The Machine Learning (ML) method is a commonly used AI tool in healthcare. It is a subsection of AI technology that has built-in algorithms that can comprehend large amounts of data. With this data, they can identify patterns and adapt. This type of AI method requires minimal human interference, as it primarily runs on trial and error. However, it is incapable of forming complex interpretability as it is only limited to the data given and its experience. Erdaw & Tachbele (2021) delved into a few examples, like X-rays or MRI scans. These methods are commonly used to diagnose diseases like COVID-19, where they heavily focus on pattern recognition rather than looking out for actual diseases or infections. As a result, the amount of tasks that ML is able to perform is extremely limited, and it may not make any complex predictions other than the data it was given.

Newer AI technology, like the XAI tool, which utilizes the Deep Learning (DL) method, can surpass other AI technology in the healthcare field as it can overcome these obstacles. DL is another subsection of AI technology that derives from ML. Compared to ML's algorithm, DL is inspired by the human brain. The focus of the DL algorithm is to mimic the process of how the human brain learns and thinks. This allows DL to comprehend a larger amount of data. As such, it can recognize complex patterns in pictures, text, sounds, and other data to produce accurate insight and prediction. Chen et al. (2021) analyzes the key factor for the ability to conduct complex pattern recognition. DL has multiple hidden layers (deep neural networks) to automatically learn and extract certain features from the data. These neural networks are capable of automatically learning complex data, which can make them more adept for tasks like

image and speech recognition. Since the DL method can perform a wide variety of tasks with ease, this also means that there are different techniques used to identify and diagnose cancer. These can range from image-based to parameter data.

Triplet Network Accuracy Rate on Lung Cancer

Over the last five years, there has been an increasing amount of research on the application of DL in cancer diagnosis, precision medicine, radiotherapy, and cancer research. Tran et al. (2021) noted that breast, lung, and brain cancer are the top three researched cancer diagnoses by AI, which equates to 16.4%, 14.0%, and 9.4%, respectively. This is in light of the most common types of cancer in breast, lung, and prostate. The accuracy rate of the XAI tool will vary a lot depending on the specific model, dataset, and task.

An example of a technique used is called the Triplet network. It is a specific type of neural network algorithm that was designed to improve the comprehensibility of models (van der Velden et al., 2022). It is especially used for tasks related to image retrieval and learning similarities. This task could be completed by incorporating three identical networks and parameters of data points: attention mechanism, activation visualization, and feature attribution methods. The attention mechanism is used to highlight important regions like discoloration, tissue, and cell abnormalities. This can be used to assist XAI in understanding which part of the image is contributing most to the similarities and differences between samples. Activation visualization and feature attribution methods are used to identify key features and process this information.

Upon analyzing the samples, the networks would be able to calculate values and compare them to the data points. As a result, XAI would be able to come up with a decision.

LaLonde et al. (2020) conducted a study using the Triplet Network to detect lung cancer in patients. It was able to detect the malignancy of lung nodules by its technique algorithm. As compared to other sample pictures, it can pick up subtlety, sphericity, margin, lobulation, speculation, and texture. In the end, it has an accuracy rate of 86.39%. Nonetheless, other techniques that target other types of cancer will display varying amounts of accuracy rates.

Risk Prevention and Possible Concerns

Although accurate cancer diagnosis, classification, and grading are crucial for the early detection of cancer and risk prevention, predicting gene mutation in cancer is also crucial. It can generate a model that can explain the probability of a genetic mutation that will arise in cancer cells (Chen et al., 2021). This can be done by gathering a comprehensive dataset of genetic information from the cancer patient. Fu et al. (2020) pointed out that characteristics from photos of tissue structures can predict genetic mutations when tumor specimens cannot be obtained for conventional mutation studies. Compared to direct sequencing, utilizing AI for tissue analysis is a lot more affordable. Collaboration with the XAI tool and the oncologist profession will allow a more personalized treatment plan, as they will know the risk and have an accurate insight into the genetic basis of cancer.

Although the XAI tool can perform a wide variety of tasks, it is also not a perfect tool in every situation. One of the concerns is input sensitivity. XAI can comprehend large amounts of data; however, if the input is relatively small, XAI is especially sensitive to small variations in the data (van der Velden et al., 2022). This can lead to inconsistent explanations. Thus, affecting the whole diagnosis process and its verdict, which reduces the trustworthiness of the explanation.

Another limitation is the lack of unified standards. Currently, there is no standardized practice of XAI techniques or metrics to evaluate the quality of explanations. This is due to the complexity of the XAI tool and the different techniques that could be created to cater to a specific niche task. Such as Triplet Network, which only looks out for malignant tumors of lung cancer through its shape and characteristics. As a result, there is some level of challenge when it comes to assessing and establishing difficulty in benchmarking. Without a common benchmark, measuring progress and effectively comparing methods will become challenging.

Ethical concerns over the usage of AI technology also raise uneasiness in patients. As these technologies utilize sensitive data of the patients, it is only fair that they become hesitant. These concerns can range from bias in data and the transparency of these tools to the privacy of patient information and security. As the XAI tool learns over time, it can propagate bias from the data presented (Al-Antari, 2023). These can reinforce bias toward certain groups of people based on race, gender, age, or other factors. It is crucial to correct these mistakes to ensure fairness and equity. Transparency in the decision-making process of the XAI tool will allow accountability to those who created and used it. Security and privacy for patient information are also crucial, as they may inadvertently reveal sensitive information.

Conclusion

Explainable Artificial Intelligence (XAI) is a tool that utilizes Deep Learning (DL) algorithmic methods. It has the foundation to surpass any other previous method as it mimics the way the human brain learns and adapts for certain tasks. As a result, it can identify and comprehend larger amounts of data to recognize complex patterns with different techniques.

The accuracy rate varies significantly depending on the technique being used. Triplet Network has an accuracy rate of 86.39% on lung cancer. This is especially useful when it comes to diagnosing cancer at an early stage and determining if the tumor is malignant or not. Early treatment detection and risk prevention are also possible with the assistance of AI technology. Procedures like predicting gene mutations can characterize cancer and its potential to become malignant.

To answer the research question of how AI technology has changed the process of cancer diagnosis and risk prevention, it has made cancer more easily detectable at an earlier stage compared to traditional methods. This allows time for the patient to establish a treatment plan and risk prevention. Personalized treatment plans can be generated by AI tools. However, relying purely on AI technology like the XAI tool to diagnose cancer is by no means negating medical oncologists' decisions and their ability to properly treat cancer. Instead, it is rather an instrument for them to understand more about the type of cancer properly. Ensuring that the patient is getting the best possible treatment. Although AI tools like XAI can perform any task, there are still limitations and concerns over these technologies. Data bias, transparency, and privacy are the biggest concerns. Nonetheless, there is still room to improve.

References

Chen, Z., Lin, L., Wu, C., Li, C., Xu, R., & Sun, Y. (2021). Artificial intelligence for assisting cancer diagnosis and treatment in the era of precision medicine. *Cancer Communications*, 41(11), 1100–1115. https://doi.org/10.1002/cac2.12215

Erdaw, Y., & Tachbele, E. (2021). Machine Learning Model Applied on Chest X-Ray Images Enables Automatic Detection of COVID-19 Cases with High Accuracy. *International Journal of General Medicine*, 14, 4923–4931. https://doi.org/10.2147/IJGM.S325609

Fu, Y., Jung, A. W., Torne, R. V., Gonzalez, S., Vöhringer, H., Shmatko, A., Yates, L. R., Jimenez-Linan, M., Moore, L., & Gerstung, M. (2020). Pan-cancer computational histopathology reveals mutations, tumor composition and prognosis. *Nature Cancer*, 1(8), Article 8. https://doi.org/10.1038/s43018-020-0085-8

LaLonde, R., Torigian, D., & Bagci, U. (2020). Encoding Visual Attributes in Capsules for Explainable Medical Diagnoses. In A. L. Martel, P. Abolmaesumi, D. Stoyanov, D. Mateus, M. A. Zuluaga, S. K. Zhou, D. Racoceanu, & L. Joskowicz (Eds.), *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020* (pp. 294–304). Springer International Publishing. https://doi.org/10.1007/978-3-030-59710-8_29

Siegel, R. L., Miller, K. D., & Jemal, A. (2020). Cancer statistics, 2020. CA: *A Cancer Journal for Clinicians*, 70(1), 7–30. https://doi.org/10.3322/caac.21590

Tran, K. A., Kondrashova, O., Bradley, A., Williams, E. D., Pearson, J. V., & Waddell, N. (2021). Deep learning in cancer diagnosis, prognosis and treatment selection. *Genome Medicine*, 13(1), 152. https://doi.org/10.1186/s13073-021-00968-x

Van der Velden, B. H. M., Kuijf, H. J., Gilhuijs, K. G. A., & Viergever, M. A. (2022). Explainable artificial intelligence (XAI) in deep learning-based medical image analysis. *Medical Image Analysis*, 79, 102470. https://doi.org/10.1016/j.media.2022.102470