HIGH SCHOOL EDITION
Journal of Student Research

Volume 12 Issue 2 (2023)

Bioinformatic Analysis of Cancerous Cells

Ayush Yavagal[1] and Jothsna Kethar[#]

[1]Gifted Gabber
[#]Advisor

ABSTRACT

Cancer is a ruthless disease that has no definite cure and it is very consequential to treat it. Chemotherapy and other cancer treatments have lasting negative effects on patients, like fatigue, diarrhea, nausea, and many other harmful side effects. To decrease the time period of vigorous cancer treatments like chemotherapy, cancer should be detected very early. Pathologists and clinicians have used many methods of cancer diagnosis over the years, but to do this, large amounts of data about a patient and their history are needed. This is known as the bioinformatic analysis of cancer cells. Bioinformatics is the act of using biological information to aid in the diagnosis of a disease, which in this case is cancer. When pathologists have diagnosed cancer in the past, the main component of diagnosis was the processing of this biological data, and it still is a major part of diagnosis today. With the rise of new technologies, Artificial Intelligence has been brought into the medical limelight and has been used extensively in the medical field to process the abundance of this data in recent years. With the use of Artificial Intelligence, the analysis speed of data used to diagnose cancer is higher than if pathologists alone analyzed the data. This paper focuses on the processing of bioinformatic data in pathologists' cancer diagnosis workflow.

Introduction

Cancer is a disease where abnormal cell growth (where cells grow without control) occurs and spreads to other parts of the body if it is not benign, meaning non-lethal. There are many types of cancer, including (but not limited to) breast cancer, lung cancer, and skin cancer, but all of them have a couple of things in common. The disease happens when the body's normal control mechanisms stop working for some cells, hence the uncapped growth. These abnormally growing cells come together to form a mass called a tumor, which can spread to other parts of the body. Cancer is a very important topic in the medical field, mainly because there isn't a viable and definite cure for it yet. It can be cured through processes like chemotherapy, radiation, and surgery, but these processes can be damaging to the human body. Unless cancer has been detected early, it cannot be effectively treated without some suffering on the patient's part. It is easier to diagnose and attempt to cure cancer during earlier stages, which is also called screening (the act of detecting or diagnosing cancer in very early stages). Despite the negative aspects of cancer detection, there are always many improvements being made to cancer diagnosis in general. This allowed more people to get better help with fighting cancer. For example, the overall survival rate for all cancers rose from 2011 to 2017, from 49% to 68%.

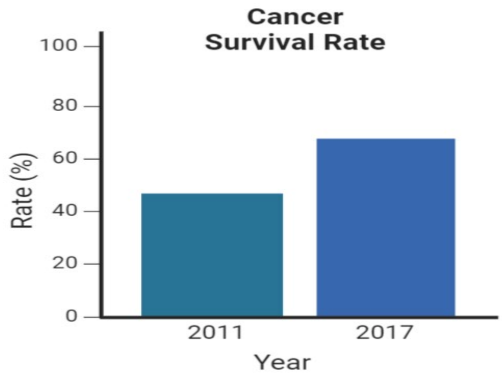

Figure 1. Shows the increase in survival rate for all cancer overall from 2011 to 2017. Created and copyrighted by Ayush Yavagal.

To diagnose cancer in an early stage, also known as screening, data analysis with pathologists use Artificial Intelligence to find and ultimately prevent the spread of cancer while saving time and preventing diagnosing mistakes. This branch of medical research and data is called Computational Pathology. Computational Pathology is the use of advances in technology like Artificial Intelligence and data collected from clinical pathology to improve patient care, and using it will help reduce errors that humans will otherwise make during diagnosis. Computational Pathology is the use of this technology to analyze medical images derived from microscopes or other medical devices.

Artificial Intelligence, a widely known field in computer science, is the simulation of fundamental aspects of human thinking and it uses problem-solving and learning to achieve a conclusion or goal. Although Artificial Intelligence is used in many different areas, like manufacturing and education, Scientists normally use Artificial Intelligence to process large amounts of data that humans cannot. Medical data usually comes in abundance to ensure specificity and exactness in diagnosis. Large amounts of this information are fed into the Artificial Intelligence program and outputs are made based on what the Artificial Intelligence was programmed to output. Artificial Intelligence is a better way of diagnosing cancer because the margin of error is reduced compared to traditional methods of cancer diagnosis. For instance, a traditional pathologist would have an accuracy of 0.966, compared to the increased accuracy of the diagnosis made by Artificial Intelligence at 0.995.

Artificial Intelligence based cancer cell detection is a more beneficial alternative to the traditional analysis done by pathologists because of many factors. As the use of Artificial Intelligence in the medical field becomes more common, the precision and accuracy of cancer cell diagnosis will be greatened.

The Use of Artificial Intelligence to Detect Cancer

Artificial Intelligence has the potential to transform the way cancer is detected and diagnosed for the better. Its ability to analyze large amounts of data quickly and accurately makes it a valuable tool for pathologists and other medical professionals. One of the key benefits of using Artificial Intelligence in cancer diagnosis is its efficiency. It can process data from images, such as CT scans and MRIs, much faster than a human could, freeing up valuable time for pathologists to focus on other tasks. Another advantage of Artificial Intelligence is its accuracy. By learning from thousands of examples of healthy and cancerous cells, an Artificial Intelligence program can accurately identify abnormalities in new images with a high degree of accuracy. In addition to these benefits, the use of Artificial Intelligence in cancer diagnosis also has the potential to improve patient care by providing more personalized treatment options based on the patient's specific medical history and needs. Overall, the integration of Artificial Intelligence into the cancer diagnosis workflow is likely to continue to grow in the coming years as more and more medical professionals realize its potential to improve patient recovery and results.

Collecting Data

Before creating an Artificial Intelligence algorithm or even screening using traditional methods, data needs to be collected. This data, if it is to be specifically processed by Artificial Intelligence, is usually brought to the digital world in many forms, one of them being image metadata. Image metadata is information about an image embedded into the image or stored in a separate file along with the image to help further specify and provide context for the given image. For example, image metadata in the medical field could include the patient ID (to assign the image to a patient), patient name, birth date, disease history, and other general information. More specific data could include information about the scanned sample slides; resolution (if working digitally), viewing bounds, and more image-specific data called metadata.

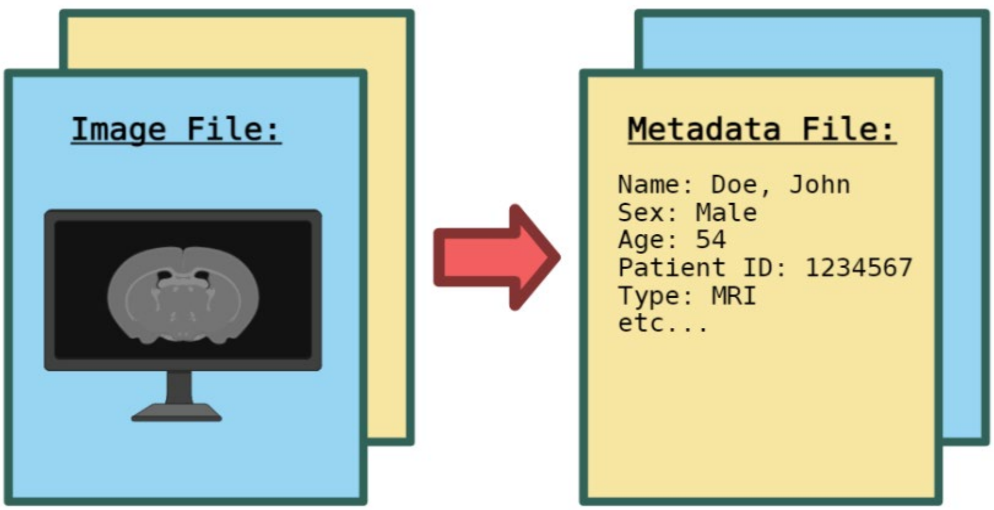

Figure 2. Shows the basic idea of image metadata. Metadata is the information embedded or paired with an image. Created and copyrighted by Ayush Yavagal.

Medical imaging techniques such as CT scans and MRIs are valuable sources of data for computational pathology. These methods provide pathologists with detailed images of the inside of a patient's body, allowing them to find any visual abnormalities that may indicate cancer or other conditions. Biopsies are another useful tool for computational pathology, as they allow pathologists to examine small samples of tissue or other cells under a microscope in order to make a more detailed diagnosis. By analyzing high-quality images of cells, pathologists can gain a deeper understanding of a patient's condition and develop more accurate treatment plans. In addition to using medical imaging techniques and biopsies, computational pathology also relies on other data sources such as electronic health records, lab test results, and genetic data. By combining these various sources of information, computational pathology can provide a more complete understanding of a patient's condition and help guide treatment methods. Overall, computational pathology is a powerful tool for improving the accuracy and efficiency of diagnosis and treatment in the field of pathology. By using advanced technologies such as medical imaging and biopsies, pathologists can gain a deeper understanding of a patient's condition and develop more effective treatment plans.

How Can Cancer Be Diagnosed with Patient Data?

Data in the medical field is extensive and very important in the diagnosis process, and as explained before, is usually found in CT scans or MRI scans. However, other types of data such as image metadata or patient data can also be extremely useful in the diagnosis process. For example, image metadata includes information about the image itself, such as the bounding area or the size of the image. This can be very helpful in determining which part of the slide was scanned and making the processing of the image easier and more convenient when using Artificial Intelligence. Patient data, on the other hand, includes more general medical-related information about the patient, such as their allergies, medical history, and other relevant details. This information can help pathologists personalize treatment and diagnosis by getting a better understanding of the patient and their specific needs. For instance, knowing if the patient has had a medical condition in the past can greatly help pathologists with the diagnosis process because it shows that the patient may have a higher chance of relapsing or getting diagnosed with cancer again. Before Artificial Intelligence was used in the medical field, pathologists used their expertise and experience to find tumors and abnormalities and diagnose cancer in CT scans and MRI scans. While it is undoubtedly possible to screen for cancer without Artificial Intelligence, it is not as efficient or accurate as using Artificial Intelligence for diagnosis. This is because Artificial Intelligence can analyze large amounts of data much more quickly and accurately than humans, making it possible to identify patterns and features that may indicate the presence of cancerous cells. Additionally, Artificial Intelligence can be trained to recognize these patterns and features using thousands of medical images, which helps to improve the accuracy of the diagnosis process.

Whole-Slide Imaging

Artificial Intelligence has the potential to completely change the way that we analyze and interpret patient data, particularly in the field of pathology. One way that Artificial Intelligence can be used in this context is through the use of Whole-Slide Imaging (WSI), which allows pathologists to produce high-quality images of biopsy samples and to interact with them in a digital format. Traditionally, Artificial Intelligence has been used to process and diagnose normal images in radiology, but WSI allows for the creation of much higher-quality images that can be analyzed in higher detail. To use WSI, a slide containing a sample obtained by using the method of biopsy is viewed with a special device that processes the image and sends it to a viewing software program. The pathologist can then interact with a digital version of the entire slide, allowing them to examine it in greater detail and identify specific features or changes that may help with the diagnosis. In addition to improving the accuracy and efficiency of diagnostic processes, WSI has the potential to significantly reduce the workload of pathologists by making certain tasks automated and allowing them to focus on more complex and interpretive aspects of their work. Because of this WSI and Artificial Intelligence in pathology will start to become more integrated in the coming years, and these technologies will play an increasingly important part in the diagnosis and treatment of cancer and other diseases.

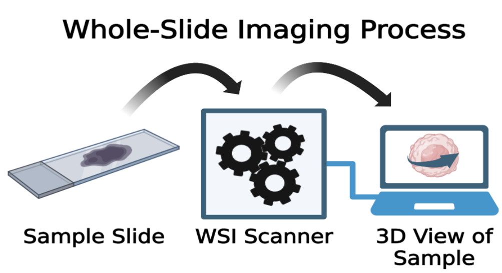

Figure 3. The Whole-Slide Imaging process is shown in an easy-to-understand manner. Created and copyrighted by Ayush Yavagal.

The digital freedom and heightened interactivity and control that the WSI allows give pathologists an easier-to-use slide to work with. This digital slide can be zoomed into one step higher than the zoom level that the slide was scanned at, and it can also be zoomed out to show a larger part of the slide or even the whole slide. This level of interactivity allows flexibility and gives pathologists a better view of their slides than with traditional microscopes. To create WSI images, a WSI scanner is used. A sample slide is processed with the WSI scanner hardware and the data is sent into compatible software where different parts of the sample are stitched together to make one big image; a whole slide. This image can be viewed and interacted with in a high-quality 2-dimensional viewing software. The pixels of the high-resolution image can also be used in an Artificial Intelligence program to find patterns or detect certain abnormalities, which could mean that there are cancerous cells in the area. The Whole Slide Imaging scanner itself is essentially a microscope controlled by robotic means. With the digital freedom and peak quality that WSI scanners provide, there comes a price; WSI scans take up extremely large amounts of data. These scans produce at least 4 terabytes (TB) of data, and pathologists need a maximum of 100 TB of extra storage ready on the side (the amount of data depends on the scanner and the image properties). To put this into perspective, a normal PNG image is only around 4 kilobytes in size, while a Whole Slide Imaging scan is at least 4 billion kilobytes (4,000,000,000 KB).

After the WSI scanner produces the high-quality digital slide, an Artificial Intelligence program can be used to find patterns or abnormalities in the image. To do this, the Artificial Intelligence model is fed hundreds of thousands of healthy slide scans and cancerous slide scans (the more scans fed into the model, the more accurate it will be and the more experience it will have) and will process them using a certain method and find any abnormalities that it thinks looks like cancer, based on the images it was previously fed.

Advantages of Artificial Intelligence in Digital Pathology

Using Artificial Intelligence to detect tumors from cancer images has an abundance of advantages as it is a technology that is very rapidly being improved and is meant to assist humans with work that otherwise would have low efficiency due to certain limitations. The versatility and efficiency of Digital Pathology using Artificial Intelligence have far surpassed previous methods of cancer detection. An Artificial Intelligence program is less prone to making mistakes compared to humans because it follows a strict set of rules without any major factor preventing it from doing so, faster than humans. Human error can be caused by mistakes or rule violations, both of which a certain level of a well-programmed Artificial Intelligence program will not and should not commit.

The Artificial Intelligence algorithms themselves, unlike Whole Slide Imaging results, used to be very large in size, but have been developed and simplified into algorithms and programs that can run on even a laptop. The large amounts of data that have to be processed for cancer screening are more efficiently processed by Artificial Intelligence than by humans. For instance, a pathologist may experience fatigue after long hours of processing data, while an Artificial Intelligence program will continuously do what it is programmed to do without being drained of energy (not considering device specifications, status, and other hardware and software limitations). For a general example, some patients who were evaluated for cataracts were assessed by a pathologist with and without the help of Artificial Intelligence. In a study of cataract diagnosis, Artificial Intelligence had lower accuracy than experienced pathologists (87% vs 99%) but was able to produce a diagnosis more quickly (2.8 minutes vs 8.5 minutes). Even though the accuracy slightly decreased with the use of Artificial Intelligence, the speed increased more than three times the speed of a pathologist not assisted with Artificial Intelligence. The use of Artificial Intelligence can free up huge amounts of time spent finding abnormalities for pathologists so they can focus on other important items, speeding up the diagnosing process.

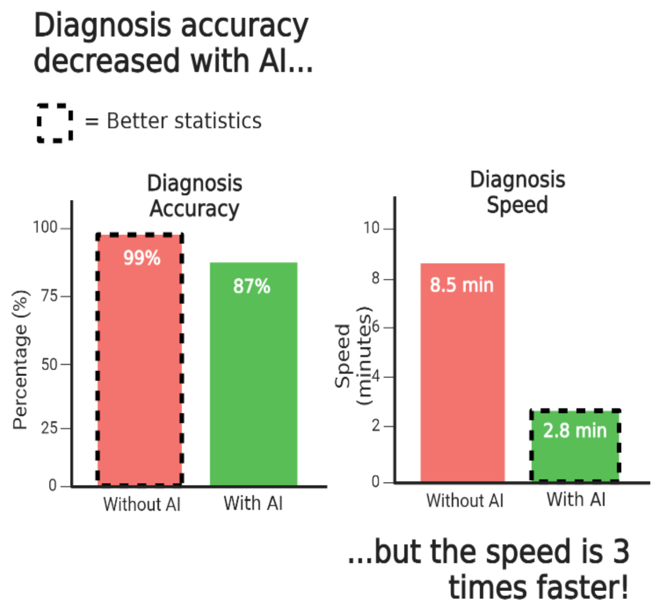

Figure 4. Comparison of cataract diagnosis accuracy and speed with and without the use of Artificial Intelligence. Created and copyrighted by Ayush Yavagal.

Considering these statistics, it should be kept in mind that Artificial Intelligence currently exists in the medical field (and in many other fields) to provide quick suggestions and helpful paths toward certain goals for pathologists and clinicians, and the final disease diagnosis almost always needs to be done by an experienced pathologist, with or without Artificial Intelligence. To further explain, a pathologist is more capable of making authoritative decisions than an Artificial Intelligence program. After the Artificial Intelligence processes the data and comes up with results, a pathologist or data analyst has to go through the data and results again to make sure the Artificial Intelligence had produced an accurate diagnosis. This will prevent any errors that may have occurred during the Artificial Intelligence diagnosing process.

How Artificial Intelligence in Cancer Diagnosis Works

Artificial intelligence is a very versatile and powerful tool that is used in many industries, including the medical field. In the medical field, Artificial Intelligence is often used for image and data analysis, including the analysis of MRI and CT scans. These types of scans produce detailed images of the inside of the body, which can be used to detect abnormalities or anomalies that may show the presence of cancer or other conditions. Before an Artificial Intelligence model can be used to analyze medical images for abnormalities, it should first be trained on a large dataset of labeled images, so that it can “learn” from experiences. This typically involves providing the model with thousands of images showing healthy cells and tissue with no abnormalities, as well as thousands of images showing cancerous tumors or other abnormalities. Each image is labeled as either healthy or showing an abnormal condition, and the Artificial Intelligence model uses this information to learn about the characteristics and patterns that are related to each type of image. Once the Artificial Intelligence model has been trained on this dataset, it can then be used to analyze new images in order to determine whether or not they show healthy cells and tissue or whether they contain abnormalities. To do this, the model looks for patterns and features that it has learned to associate with either healthy or abnormal sample conditions. If the model finds an adequate number of these patterns and features in a given image, it will classify the image as either healthy or abnormal. For example, an image may include an unusual bump in the body part, one that looks similar to images of bumps that have been fed into the model. In the case of cancer diagnosis, the Artificial Intelligence model can be used to identify abnormalities in medical images that may indicate the presence of cancerous tumors. It can also provide information on the location of these abnormalities, which can help pathologists to better understand the diagnosis and potentially reduce the time it would take to identify an abnormality without the assistance of the model. By using Artificial Intelligence to analyze medical images, pathologists can possibly detect cancer at an earlier stage, which can improve patient results and reduce the overall cost of treatment.

The Future of Computational Pathology

As explained in the previous sections, using computational pathology for cancer screening from an efficiency standpoint is better than traditional cancer screening methods. Since all technology is not perfect, this raises some questions about medical workplaces as the new tool, Artificial Intelligence, would be introduced. Just like in the times before computational pathology, a variety of experienced and skillful medical professionals are needed to understand the work. Artificial Intelligence has not reached a point where it can be used as a standalone tool that does not need the input of pathologists, and it will not be any time soon. This means scientists that specialize in certain Artificial Intelligence fields, such as data analysts and data scientists, are also needed to help with the implementation of Artificial Intelligence into the diagnosis workflow. The data analysts should have some kind of professional medical education or background to manage the complex medical data encounters that could exist. Data analysts are scientists who work with data to process or decipher it, usually in complex situations. They are an important necessity in the medical workplace because they help with handling medical data and the creation of the Artificial Intelligence program itself.

As the assistance of Artificial Intelligence becomes more widely used, the room for improvement will grow, just as it does with any other game-changing technology. The impact of Artificial Intelligence on cancer diagnosis and the medical workplace as a whole will be significant, as there are many factors that could be improved with the integration of Artificial Intelligence. For example, the development of smarter and more efficient Artificial Intelligence programs could lead to more accurate and efficient diagnoses, as well as a reduction in the need for manual corrections by pathologists. Additionally, the use of higher-quality Whole Slide Imaging scanners and the images they produce could improve the overall quality of slides and make them easier to analyze. Improved data management and efficient data manipulation could also simplify the diagnosis workflow and make it more efficient. These are just a few of the many potential improvements that could be made to the cancer diagnosis process with the help of Artificial Intelligence. As Artificial Intelligence technology continues to advance, it is likely that it will be used in an even broader range of situations, further improving the efficiency and accuracy of a cancer diagnosis.

Challenges with Artificial Intelligence

Artificial Intelligence is always being improved, and like many other developing technologies, it is definitely not perfect. Using Artificial Intelligence in clinical pathology and the detection of cancer can come with many benefits, such as increasing accuracy and efficiency, but there are also many challenges to consider. One of these challenges is the creation of carefully hand-selected data sets for Artificial Intelligence models to process. This process involves labeling the images with key attributes, segregating them, and annotating them so that the models can learn from them. Creating these data sets can be very time-consuming and costly for pathologists, as there need to be hundreds of thousands of images that are high in quality for the Artificial Intelligence model to learn from. It is also expensive since the images must be genuine and pathologists must use expensive medical equipment to create them. Obtaining high-quality medical scans can be tedious because the equipment can take long amounts of time to process a sample. As a result, there may be a lack of data available for Artificial Intelligence models to learn from, which can limit the effectiveness of the technology. Sharing this data and creating algorithms to process it can also be challenging, as the data being transferred may be sensitive and subject to strict security and privacy regulations. There may also be technical issues related to data compatibility that need to be addressed. To lessen these challenges, some actions can be taken, such as enforcing standardized data formats to guarantee easier data sharing. However, these solutions may also have unintended consequences, such as reducing the security of the data or increasing the risk of data leaks. Despite these challenges, Artificial Intelligence is still an efficient and helpful tool for aiding pathologists in the diagnosis of cancer. However, it is important to realize that Artificial Intelligence should not be used as a standalone diagnostic tool and should always be used with the expertise of medical professionals and experienced pathologists to minimize the risk of errors or complications.

Further Research Topics

Artificial Intelligence is a topic that is being widely looked into in the scientific community, which means there will still be many topics and fields in which Artificial Intelligence can be researched upon. The reliability of some Artificial Intelligence models that are in the process of being developed can be low at times, raising questions about when an Artificial Intelligence model can actually be used to assist pathologists in the real world. Some models are pretty reliable in the medical workplace, but almost all of them are not perfected in terms of efficacy because they can always be improved. For example, the accuracy of a certain model could be low, so Data Scientists, people who work with and improve data and data management, could come up with new algorithms and models to get more accurate results in cancer diagnosis and image analysis in general. Another topic that should be looked upon is the costs of using Artificial Intelligence in the medical field in general, whether they are hardware, software, or overall costs. This will help in determining if Artificial Intelligence is worth using or if it is too expensive to integrate into the medical field everywhere. There are a lot of topics that can be discussed and discovered. These are only some of the suggestions on areas of discussion that can be looked into by other researchers and clinicians to better the understanding of Artificial Intelligence and its role in the bioinformatic analysis of cancerous cells.

Conclusion

The use of computational pathology techniques for the bioinformatic analysis of cancerous cells has great potential for improving the efficiency and accuracy of traditional cancer detection and treatment. With the help of artificial intelligence and learning algorithms, it is possible to analyze large amounts of medical data, including images and genetic information, to identify patterns and make more accurate diagnoses. The integration of computational pathology is likely to continue growing in the coming years, as more healthcare providers see the many benefits it has to offer. The future of Computational Pathology looks bright, considering the new improvements that are constantly being made to Artificial Intelligence. While there are definitely challenges to be overcome, such as the creation of curated data sets and the development of algorithms that can handle complex data sets, the potential benefits of AI in pathology are massive. Along with Artificial Intelligence, the amount of new medical data and discoveries that are made is also being increased, calling for the much-needed use of Artificial Intelligence in digital pathology to help manage new data. Deep Learning and Artificial Intelligence have greatly improved the efficiency and accuracy of the detection of cancer. As Artificial Intelligence is normally built for speed and efficiency, it may lag behind in accuracy on some occasions. But, being a constantly improving technology, it makes the cancer screening workflow faster and smoother. Ultimately, the use of Artificial Intelligence to process bioinformatics for the diagnosis of cancer is widely recognized and will continue to be recognized for years to come.

Acknowledgments

I would like to thank Coach Jo, Kaviya Ganesan, and Smrutha Vikram for aiding in the creation of this research paper. I would also like to thank Dr. Rajagopal Appavu for providing any needed references for specific information and guidance in writing the research paper.

References

Cui, M., & Zhang, D. Y. (2021). Artificial intelligence and computational pathology. Laboratory Investigation, 101(4), 412–422. https://doi.org/10.1038/s41374-020-00514-0

Singh, S. (2022, February 12). Computer-Aided Diagnosis of Gliomas Using Machine LearningClassification Algorithms https://static.wixstatic.com/media/1db4ae_c98a4638656645168586aff767a5f6d6~mv2.jpg/v1/fit/w_185%2Ch_415%2Cal_c%2Cq_80/file.jpg Shreya Singhhttps://static.wixstatic.com/media/1db4ae_c98a4638656645168586aff767a5f6d6~mv2.jpg/v1/fit/w_185%2Ch_415%2Cal_c%2Cq_80/file.jpg Shreya Singh. Gifted Gabber. https://www.giftedgabber.com/post/computer-aided-diagnosis-of-gliomas-using-machine-learningclassification-algorithms

Bi, W. L., Hosny, A., Schabath, M. B., Giger, M. L., Birkbak, N. J., Mehrtash, A., Allison, T., Arnaout, O., Abbosh, C., Dunn, I. F., Mak, R. H., Tamimi, R. M., Tempany, C. M., Swanton, C., Hoffmann, U., Schwartz, L. H., Gillies, R. J., Huang, R. Y., & Aerts, H. J. W. L. (2019). Artificial intelligence in cancer imaging: Clinical challenges and applications. CA: A Cancer Journal for Clinicians. https://doi.org/10.3322/caac.21552

Niazi, M. K. K., Parwani, A. V., & Gurcan, M. N. (2019). Digital pathology and artificial intelligence. The Lancet Oncology, 20(5), e253–e261. https://doi.org/10.1016/s1470-2045(19)30154-8

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

Liu, Y., Gadepalli, K., Norouzi, M., Dahl, G. E., Kohlberger, T., Boyko, A., Venugopalan, S., Timofeev, A., Nelson, P. Q., Corrado, G. S., Hipp, J. D., Peng, L., & Stumpe, M. C. (2017). Detecting Cancer Metastases on Gigapixel Pathology Images. ArXiv.org. https://arxiv.org/abs/1703.02442

Schiffman, J. D., Fisher, P. G., & Gibbs, P. (2015). Early Detection of Cancer: Past, Present, and Future. American Society of Clinical Oncology Educational Book, 35(35), 57–65. https://doi.org/10.14694/edbook_am.2015.35.57

Raab, S. S., & Grzybicki, D. M. (2010). Quality in Cancer Diagnosis. CA: A Cancer Journal for Clinicians, 60(3), 139–165. https://doi.org/10.3322/caac.20068

Iyengar, J. N. (2021, January 1). Whole slide imaging: The futurescape of histopathology Iyengar JN - Indian J Pathol Microbiol. https://www.ijpmonline.org/article.asp?issn=0377-4929;year=2021;volume=64;issue=1;spage=8;epage=13;aulast=Iyengar