HIGH SCHOOL EDITION
Journal of Student Research
Volume 11 Issue 4 (2022)

Optimization of Pediatric Cancer Diagnosis with Convolutional Neural Networks (CNNs)

Simran Saluja [1], Rajagopal Appavu [#], and Jothsna Kethar [#]

[1] Rocklin High School, Rocklin, CA, USA
[#] Advisor

ABSTRACT

In today's world, technology has become much more prevalent in the world of medicine. Growing fields like biotechnology and artificial intelligence are helping save and improve lives in ways that we couldn't have imagined just 40, 50 years ago. Some of the most common examples of this today are prosthetics and using artificial intelligence in radiology. In the past few years, artificial intelligence technology has been advancing and scientists have begun to research whether deep learning algorithms like convolutional neural networks can be used to help detect signs of and diagnose cancer. Specifically, a growing research field refers to using deep learning and CNN models to detect pediatric cancer, one of the hardest cancers to detect based on symptoms. In this paper, it will be discussed whether deep learning algorithms are effective in use for the detection or diagnosis of pediatric cancers.

Introduction

Pediatric cancer is found in children between the birth and the age of 14. It is rare compared to adult cancers, and is difficult to detect as the symptoms are similar to regular illnesses which most children face. In today's world, the development of artificial intelligence has given health professionals the opportunity to advance the diagnosis of pediatric cancers using technologies like artificial intelligence. AI algorithms have to be trained first in order for them to accurately come to conclusions when in use. Oftentimes due to the lack of variety and image noise in the available pediatric cancer imaging data, it can be difficult to properly train and use an AI algorithm to detect pediatric cancers. This paper has put together knowledge about pediatric cancer detection and artificial intelligence algorithms to determine whether deep learning can be accurately used to detect pediatric cancer.

Literature Review

Background of Pediatric Cancer

Pediatric cancer is a cancer that is developed in the body between birth and 14 years of age (*NCI Dictionary of Cancer Terms*, n.d.). Pediatric and adult cancers differ from each other in many ways including how they grow, how they spread, and how they are treated (*NCI Dictionary of Cancer Terms*, n.d.). Pediatric cancer is also rarer in comparison to adult cancer (*NCI Dictionary of Cancer Terms*, n.d.). Some of the most common types of pediatric cancer include leukemia, brain and spinal cord tumors, lymphoma, neuroblastoma, Wilms tumor, retinoblastoma, and cancers of the bone and soft tissue (*NCI Dictionary of Cancer Terms*, n.d.).

Pediatric cancer is also more difficult to detect than adult cancers (*Finding Cancer in Children*, n.d.). Cancer in children is often misdiagnosed or not even thought about as the showing symptoms are similar to illness and hardship faced on a more regular basis (*Finding Cancer in Children*, n.d.). This includes, but is not limited to: An unusual lump or swelling, unexplained paleness and loss of energy, easy bruising or bleeding, an ongoing pain in one area of the body, limping, unexplained fever or illness that doesn't go away, frequent headaches, often with vomiting, sudden eye or vision changes, or sudden unexplained weight loss (*Finding Cancer in Children*, n.d.).

Pediatric cancer comes in many forms, and affects children all across the world. Many times, due to the growing nature of children's bodies, more issues come to play in the detection, diagnosis, and recovery processes of those with pediatric cancers (*Finding Cancer in Children*, n.d.).

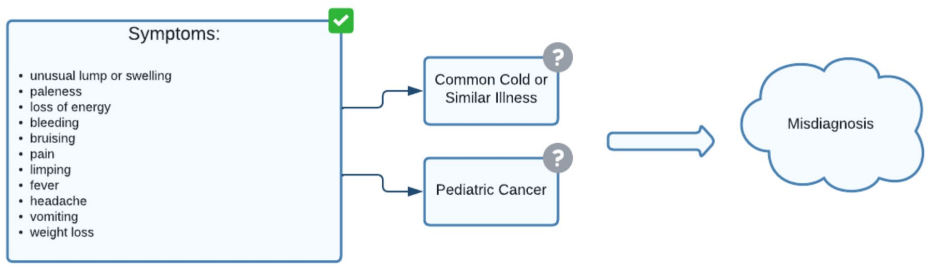

Figure 1. Misdiagnosis pathways.

Pediatric cancer comes in many forms, and affects children all across the world. Many times, due to the growing nature of children's bodies, more issues come to play in the detection, diagnosis, and recovery processes of those with pediatric cancers (*Finding Cancer in Children*, n.d.).

Background of AI Algorithms & CNN's

Before an artificial intelligence algorithm can be used for making real decisions, it must be trained with relevant data in order for it to develop a logical method to come to a conclusion (Pessach, D., & Shmueli, E., 2020). Oftentimes, issues that are seen due to an algorithm's decision are due to errors from its' training (Pessach, D., & Shmueli, E., 2020).

Figure 2. Basic AI algorithm process.

One example of this would be when an algorithm is not trained with enough data to represent all parts of the topic (Cowgill, B., & Tucker, C. E., 2020). It would result in the algorithm making decision flaws during its use in the real world (Cowgill, B., & Tucker, C. E., 2020). This is because the algorithm doesn't have enough experience with other sides of the matter to be able to take them into account (Cowgill, B., & Tucker, C. E., 2019).

AI algorithms are used in many parts of our lives, but most people do not immediately recognize where they leave their mark. For example, in the world of advertisement, biotechnology, product development, and more, AI has become quite useful (Pessach, D., & Shmueli, E., 2020).

Deep Learning is a subfield of Machine Learning, concerned with algorithms that are inspired by the brain. This subfield can handle large amounts of data, and uses hidden layering rather than traditional techniques for pattern recognition. Specifically, the most popular and useful type of Deep Learning is Convolutional Neural Networks. They are most often used for analyzing visual imagery. The role of CNNs is to get images in a form which is easier to process while not losing features needed to make an accurate conclusion (Mandal, 2021).

There are two main types of images that can be processed through a CNN: RGB Image and Grayscale Image. RGB Images are images with color, and are processed using a matrix of pixel values with 3 planes. It has 3 color channels (Red, Green, Blue), each with a pixel x pixel number for whether there is one of that "colored" picture in that spot, and they are then layered on top of one another creating the image. Grayscale images on the other hand uses the same method as RGB except that it is only one plane layer. This is due to the fact that RGB images use layering to create color (Mandal, 2021).

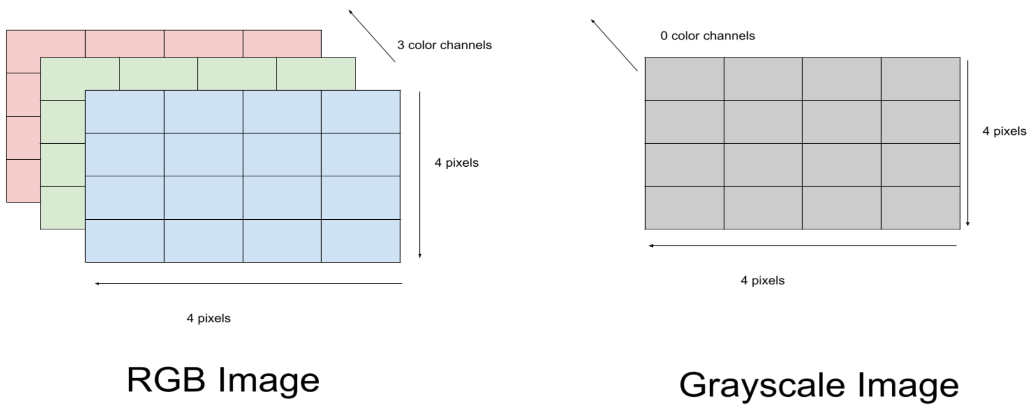

Figure 3. RGB image and Grayscale Images.

But how do Grayscale/RGB images work? How are the images simplified?

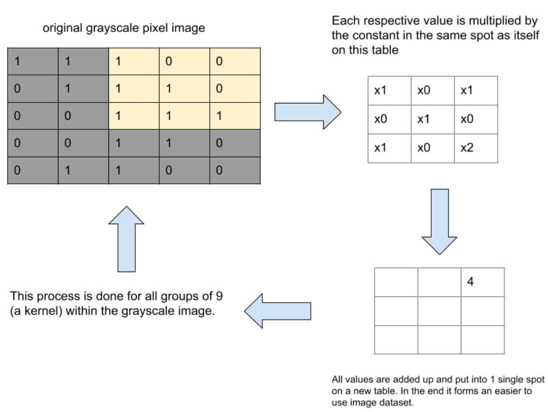

Figure 4. Grayscale image simplification process.

The original grayscale image is split into kernels - respectively groups of 9 values from the original pixelated image. These values within the kernel are multiplied by a constant based on position. The nine values are added up and put into a single spot to simplify the image in order to properly classify them within an experiment process (Mandal, 2021).

Overall, testing different deep learning models including CNN's specifically is the way to go when analyzing an image-based dataset. CNN's use RGB and Grayscale imaging processes to simplify and analyze the images.

Combining Pediatric Cancer and CNNs/AI

Detection and diagnosis can be done using a variety of methods (*Childhood Cancer - Diagnosis*, 2021). Some of the modern ways include: CT and CAT scans, PET scans, MRI scans, ultrasound imaging, blood tests, biopsies, and lumbar puncture (*Childhood Cancer - Diagnosis*, 2021).

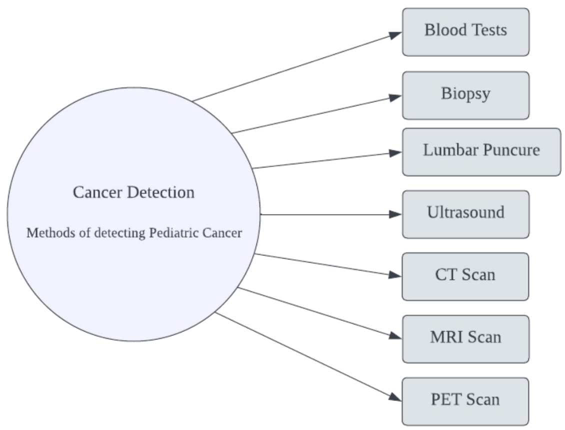

Figure 5. Methods of pediatric cancer detection & diagnosis.

1. Blood tests measure the number of different types of cells in a person's blood. Levels of certain cells that are too high or too low can indicate the presence of certain types of cancer. (*Childhood Cancer - Diagnosis*, 2021)
2. A biopsy is the removal of tissue for examination under a microscope. Other tests can suggest that cancer is present, only a biopsy can make a definite diagnosis (*Childhood Cancer - Diagnosis*, 2021).
3. A lumbar puncture is a procedure in which a needle is used to take a sample of cerebral spinal fluid to look for cancer cells or tumor markers (*Childhood Cancer - Diagnosis*, 2021).
4. An ultrasound uses sound waves to create a picture of the internal organs (*Childhood Cancer - Diagnosis*, 2021).
5. CT and CAT scans take pictures of the inside of the body using x-rays then combine these pictures into an image that shows any abnormalities (*Childhood Cancer - Diagnosis*, 2021).
6. An MRI scan produces detailed images of the body using magnetic fields and can also be used to measure the tumor's size (*Childhood Cancer - Diagnosis*, 2021).
7. "A PET scan is a way to create pictures of organs and tissues inside the body. A small amount of a radioactive sugar substance is injected into the patient's body. This sugar substance is taken up by cells that use the most energy. Because cancer tends to use energy actively, it absorbs more of the radioactive substance. However, the radioactivity in the substance is too low to be harmful to patients. A scanner then detects this substance to produce images of the inside of the body" (*Childhood Cancer - Diagnosis*, 2021).

There are four general steps during the diagnosis and treatment of cancer - in this case, pediatric cancer. First, imaging/initial detection. If suspicious activity is determined during this step, a tissue sample will be taken and a biopsy will be done. From there, based on further observation and clinical judgement, an official diagnosis can be made. Lastly, a treatment plan is administered in order to help the patient recover (Bi, W. L., et al., 2019).

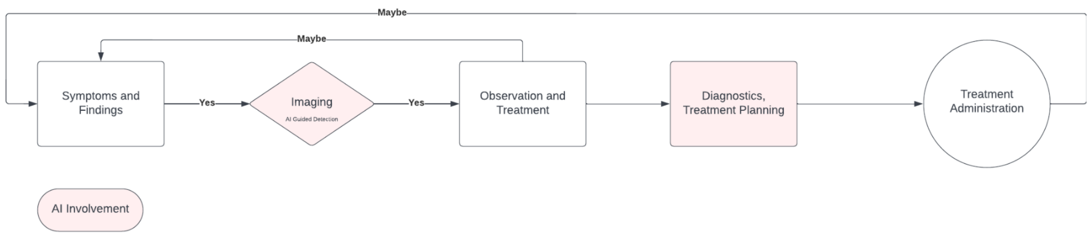

Figure 6. Where AI presents itself in the process of Detection, Diagnosis, and Treatment of Pediatric Cancer.

Artificial intelligence can help in a few parts of this process, as shown in the diagram above. First, it can help in risk calculation and initial detection based on patient history, radiographic imaging, and symptoms. Second, it can help with the official diagnosis using biopsy. The only way any cancer can be officially diagnosed is when a tissue sample is taken, and analyzed by a doctor/scientist. Here, if an algorithm is properly trained, it may be used to help analyze the tissue samples and optimize official diagnosis (Bi, W. L., et al., 2019).

Issues usually arrive in step two for pediatric cancer patients because cells from the biopsy can have SRBTC's (Liu, Y. & Sun, S.. , 2021). SRBTC's can create disruptions making it hard to determine whether a cell really is cancerous or not (Liu, Y. & Sun, S.. , 2021). In addition, it is hard to store SRBTC's and alike for long periods of time, making it difficult to put together viable data in order to train convolutional neural networks (Liu, Y. & Sun, S.. , 2021). Algorithms have to be trained using ready data first to be able to deal with situations encountering abnormalities in real practice (Liu, Y. & Sun, S.. , 2021).

Some of the other issues that come forward regarding deep learning algorithms when in training or use involve image noise (Liu, Y. & Sun, S.. , 2021). Many times the first issue seen in data used to train algorithms is image noise, caused by various cell tissues or impurities in the biopsy sample (Liu, Y. & Sun, S.. , 2021). Another problem is the diversity of patterns themselves. If massive training data are available, the model can be taught to put together similar categorical patterns (Liu, Y. & Sun, S.. , 2021). Smaller sets of data can look different from one another, and have diversity when put together, but alone small sets of data can make it hard for the model to put patterns together to come to a diagnosis (Liu, Y. & Sun, S.. , 2021).

Oftentimes, it is easier to use deep learning for adults diagnosis of cancer because there are much larger data sets, whereas for children and pediatric cancer, there aren't large data sets with a variety of images containing foreign objects/image noise commonly seen in biopsy (Liu, Y. & Sun, S.. , 2021). This is important because in order for an algorithm/artificial intelligence to be used and properly diagnose a disease like cancer, it first has to go through many rounds of training (Liu, Y. & Sun, S.. , 2021). Algorithms tend to make errors when they are trained for many reasons, but one of the most potent ones is if there is not enough data for the algorithm to analyze it can lead to misdiagnosis (Liu, Y. & Sun, S.. , 2021).

So far, deep learning and convolutional neural networks have been used to diagnose diseases in adults by detecting unknown objects in scans of the body (Liu, Y. & Sun, S.. , 2021). This is done by combining multi-layer and contextual features (Liu, Y. & Sun, S.. , 2021). This method has also been used for diagnosing neurological events/diseases in adults, and also diabetes (Liu, Y. & Sun, S.. , 2021). Not much has been done to use deep learning convolutional neural network platforms on children (Liu, Y. & Sun, S.. , 2021).

Current AI Cancer Diagnosis Research

In an experiment done as presented in the paper "Prognostic system for early diagnosis of pediatric lung disease using artificial intelligence", patient history records and factors that affected them were analyzed using an AI Algorithm they designed (Rajan, J. R., 2017). The records they used included 70 patients: 40 patient data used during training, 30 patients after to test the AI algorithm (Rajan, J. R., 2017).

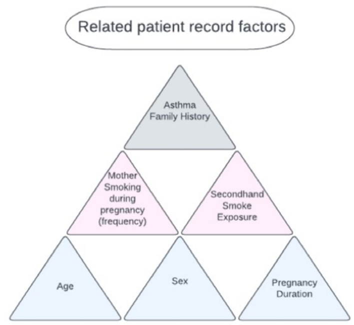

Figure 7. Related Patient Record Factors.

The result of their experiment led them to the conclusion that the AI Algorithm would be able to tell the likelihood of having lung cancer and asthma based on the information patients have on the factors listed in figure 5 (Rajan, J. R., 2017). In the end they were able to 100% match those with pediatric asthma using the algorithm and came to a conclusion that it could be used for pre-diagnosis and prevention of the disease (Rajan, J. R., 2017).

Conclusion

AI Algorithms are used to improve many parts of our lives. In healthcare and cancer research, they have been used in relation to adults with cancer, but not as often with pediatric cancer. Sadly, pediatric cancer is hard to diagnose due to the similarity of symptoms between it and day-to-day illnesses like the flu. In addition, deep learning algorithms can find it difficult to properly analyze training data for pediatric cancer, as many times there is image noise that blurs and makes it more difficult for an algorithm to make an accurate decision. This image noise is due to Small Round Blue Cell Tumors, often found in the cells of children who have cancer. This image noise can often times be reduced using convolutional neural networks, and simplifying images to an easier to analyze. In addition, image datasets are also hard to store and analyze, making it difficult to include images of these varied scenarios in training data for algorithms, resulting in less accuracy. On the other hand, AI Algorithms are able to accurately predict the % chance of an individual having cancer/other diseases based on symptoms and family history. Overall, deep learning CNN models can accurately predict the future possibility of an individual having a disease like cancer and are also able to accurately diagnose an adult with cancer. Nevertheless, because pediatric cancer has a variety of changing factors such as SRBTC's which cannot always be fairly included in a training dataset, it is possible but would be quite difficult to accurately use deep learning CNN models to diagnose pediatric cancer.

Acknowledgments

I would like to thank Coach Jo and Mr. Rajgopal for guiding me while writing this paper and for helping me polish it to the best it could be. I would also like to thank Mr. Ram Prasad for mentoring and guiding me through the process of coding the algorithm and designing my experiment.

References

Finding Cancer in Children. (n.d.). American Cancer Society. https://www.cancer.org/cancer/cancer-in-children/finding-childhood-cancers-early.html#:%7E:text=Possible%20signs%20and%20symptoms%20of%20cancer%20in%20children&text=But%20cancers%20in%20children%20can,the%20early%20signs%20of%20cancer.

NCI Dictionary of Cancer Terms. (n.d.). National Cancer Institute. https://www.cancer.gov/publications/dictionaries/cancer-terms/def/pediatric-cancer.

Bi, W. L., Hosny, A., Schabath, M. B., et al. (2019). Artificial intelligence in cancer imaging: Clinical challenges and applications. CA: a cancer journal for clinicians, 69(2), 127–157. https://doi.org/10.3322/caac.21552.

Childhood Cancer - Diagnosis. (2021, August 24). Cancer.Net. https://www.cancer.net/cancer-types/childhood-cancer/diagnosis.

Liu, Y. & Sun, S.. (2021). SagaNet: A Small Sample Gated Network for Pediatric Cancer Diagnosis. <i>Proceedings of the 38th International Conference on Machine Learning</i>, in <i>Proceedings of Machine Learning Research</i> 139:6947-6956 Available from https://proceedings.mlr.press/v139/liu21u.html.

Cowgill, B., & Tucker, C. E. (2019). Economics, fairness and algorithmic bias. preparation for: Journal of Economic Perspectives.

Pessach, D., & Shmueli, E. (2020). Algorithmic fairness. arXiv preprint arXiv:2001.09784. https://doi.org/10.48550/arXiv.2001.09784.

Rajan, J. R. (2017). Prognostic system for early diagnosis of pediatric lung disease using artificial intelligence. Current Pediatric Research.

Manav, Mandal (2021). CNN for deep learning: Convolutional Neural Networks. Analytics Vidhya. https://www.analyticsvidhya.com/blog/2021/05/convolutional-neural-networks-cnn/