

Thomas Seli

Convolutional Neural Networks for Brain Tumor Detection in MRI Scans

Abstract

This report will evaluate the effectiveness of using convolutional neural networks (CNNs) for the detection and classification of brain tumors such as glioma, meningioma, and pituitary tumors. Using an open-source dataset of MRI scans, I aim to assess the potential of deep learning for brain image analysis and its broader implications for the future of healthcare. Overall, this research demonstrates the promise of CNNs and other deep learning techniques for more accurate and efficient brain tumor diagnosis, highlighting the emerging role of AI in advancing cancer research and opening new horizons in clinical practice.

Introduction

Brain tumors have presented themselves as a prevalent public health issue in recent years. In fact, according to the National Cancer Institute, it is estimated that in 2024 there will be 25,400 new cases and 18,760 new deaths caused by this aggressive disease. Furthermore, a low survival rate of only about 33.4% highlights the need for early and accurate detection for proper preventative measures (SEER, n.d.).

Currently, the most common method for detecting brain tumors is Magnetic Resonance Imaging (MRI), which allows us to scan and image the brain without the need to perform an invasive surgery. However, manually detecting brain tumors from MRIs is time-consuming, requires a high degree of expertise, and is prone to frequent inaccuracies due to human error. The most commonly observed types of brain tumors in current clinical practice are meningioma, glioma, and pituitary tumors. While meningioma and pituitary tumors are usually benign, glioma tumors are often malignant and are known to have the highest fatality rate among all kinds of brain tumors. Because there are a variety of different brain tumors, treatment options are highly dependent on a correct diagnosis, which can often be highly subjective (Abdusalomov, Mukhiddinov, & Whangbo, 2023). Thus, devising methods to make the diagnosis of brain cancer more efficient and accurate is crucial in the ongoing fight against this deadly disease.

Alongside the emerging prominence of MRI scans, there have been significant technological advancements in the field of object detection and image classification using deep learning frameworks. One such method is the convolutional neural network, which is a type of neural network that is specifically designed to extract meaningful features from complex visual data (Khaliki & Başarslan, 2024). The bridge between these technologies presents an opportunity for us to develop models that can help health professionals make more reliable diagnoses of brain cancer, ultimately improving patient outcomes and saving lives.

Methods

The dataset used for this report was taken from Kaggle and consists of 3264 MRI images consisting of four distinct classes: no tumor, glioma, meningioma, and pituitary tumor (Kaggle, 2020). The images are presented in the sagittal, axial, and coronal planes. With over 64,000 downloads and multiple citations in research papers, this dataset has proved to be a valuable resource for brain cancer research.

Before training the convolutional neural network, the MRI dataset was preprocessed using pixel normalization and cropping to remove a majority of unnecessary blank space from the image. This helps to improve training efficiency, emphasize the regions of interest, and remove bias from images of different scales and dimensions. Figure 1 presents a random sample of six preprocessed images from the training dataset, representing each of the four classes.

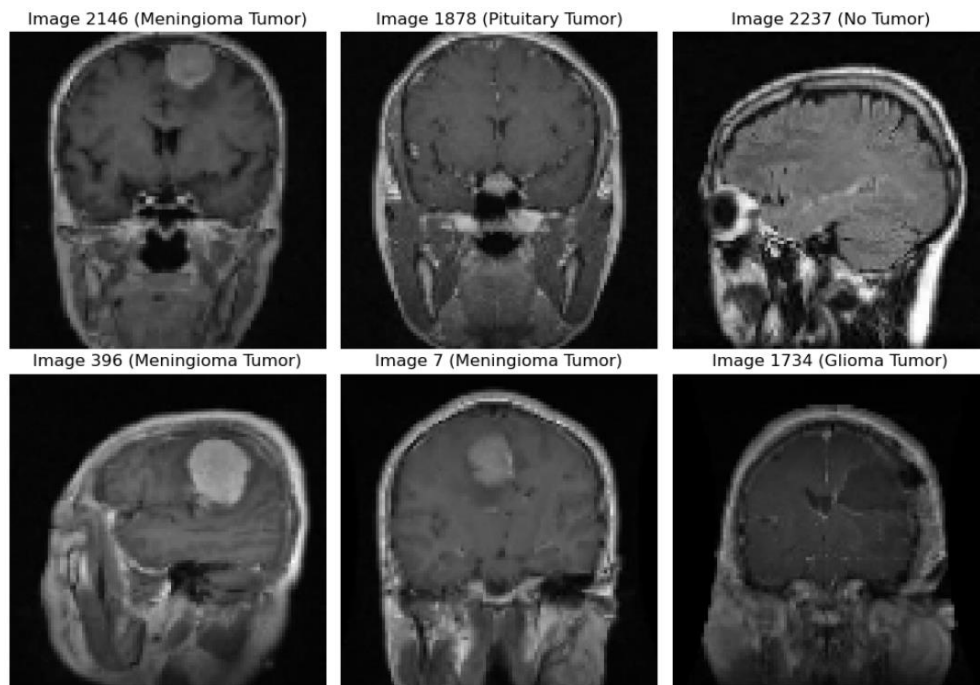


Figure 1: Sample of Preprocessed MRI Images

The dataset was also partitioned into training, validation, and testing sets. Figure 2 below shows the class balance and distribution of the splits. We can see that the class of tumor-free MRI images is underrepresented, however the three tumor classes show a relatively balanced distribution. Although the dataset is not perfectly balanced, it provides sufficient data for effective training in each class.

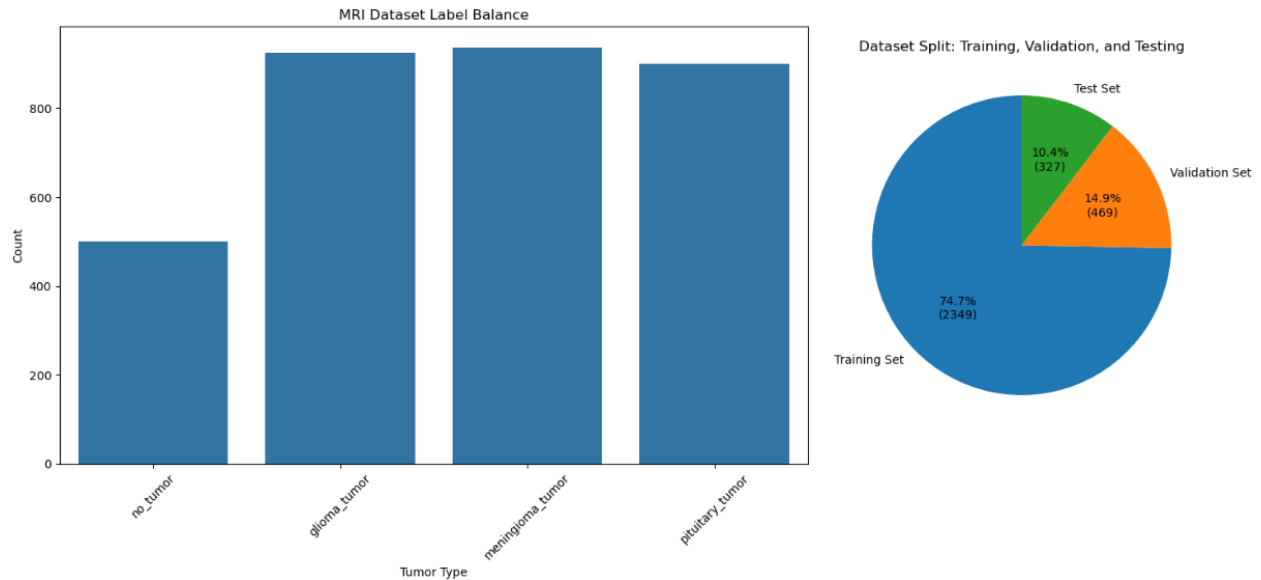


Figure 2: MRI Dataset Class Balance and Split Distribution

After the preprocessing and data exploration procedures, I performed a hyperparameter search with stratified 5-fold cross-validation to find an optimal network configuration. Initially, it proved to be quite difficult to balance training time and performance. For instance, some grid search runs could take upwards of 8 hours to return results that yielded suboptimal performance for each fold. To address this, a custom grid search strategy was implemented, dynamically adjusting the network structure based on hyperparameters such as the number of convolutional layers, filter size, number of filters, dropout rate, and batch size. Furthermore, to remedy the overfitting issue, I decided to implement l2-regularization with an alpha of 0.001, add dropout to the dense layer, and increase the class weight of the glioma class. This class weight adjustment was necessary because performance in the glioma class was quite poor when looking at the initial confusion matrices. After making these adjustments, I reran the grid search and chose the configuration with the majority vote for the best performance over the 5 folds. The optimal configuration identified through this process includes four convolutional layers with 3x3 filters, 8, 16, 32, and 64 filters, respectively. Further, there is a 128-neuron dense layer with a 0.25 dropout rate and a batch size of 8 for training.

The optimal convolutional neural network was trained for a maximum of 100 epochs and used the validation set for performance monitoring and early stopping with patience of 15. To ensure that I can yield unbiased results and properly assess the quality of the model, I reserved the test set for the final performance metrics after the training was completed.

Results and Discussion

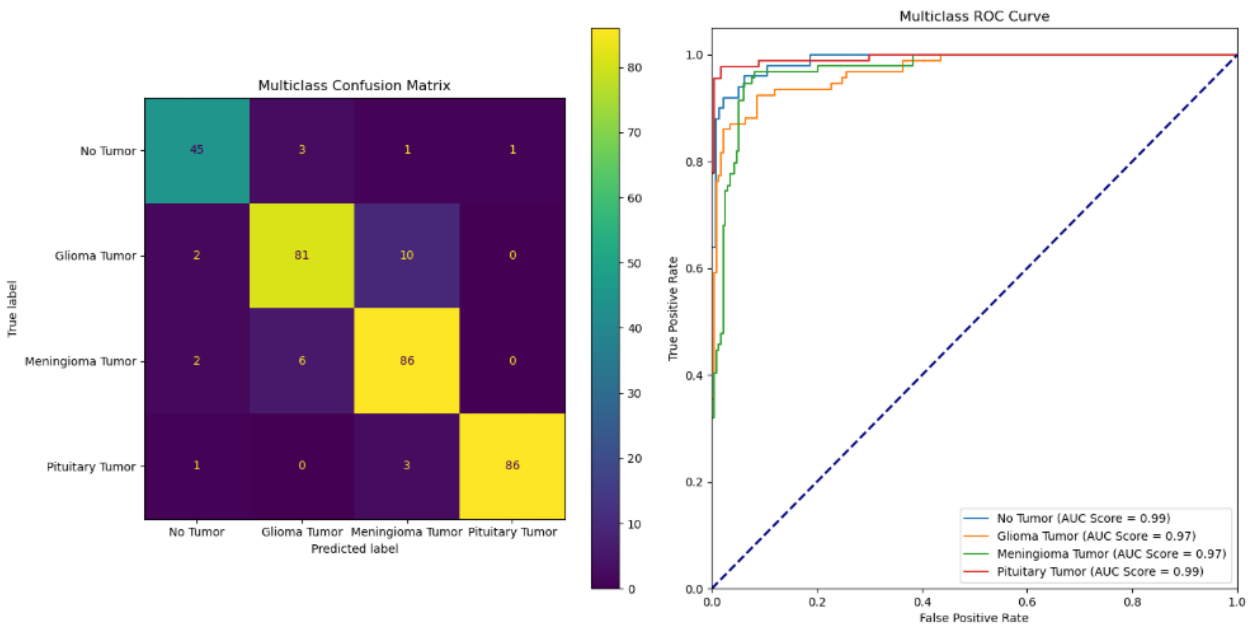
The model's performance was evaluated by a variety of metrics. Table 1 presents the accuracy, precision, recall, F1-score, and AUC across both multiclass and binary classification paradigms. In the multiclass scenario, the performance is assessed across the four distinct tumor

categories. On the other hand, the binary performance evaluates the model's ability to distinguish strictly between tumor and non-tumor cases. The model achieved a strong multiclass accuracy of 91% and a binary tumor detection accuracy of about 97%.

Table 1: Results Table

	Multiclass Tumor Detection	Binary Tumor Detection
Accuracy	0.912	0.969
Precision	0.913	0.982
Recall	0.911	0.982
F1-Score	0.912	0.982
AUC Score	0.981	0.995

Analyzing the confusion matrices and ROC curves in Figure 3 below, we can see that while the model can effectively discern between the healthy and pituitary tumor classes, it still struggles more to distinguish between glioma and meningioma tumors. This is critical to note given the significantly higher malignancy of gliomas. In medical practice, a misclassification in this case could have serious consequences. Thus, further research into adjusting class weights or refining the preprocessing methods might be necessary to improve performance in this area.



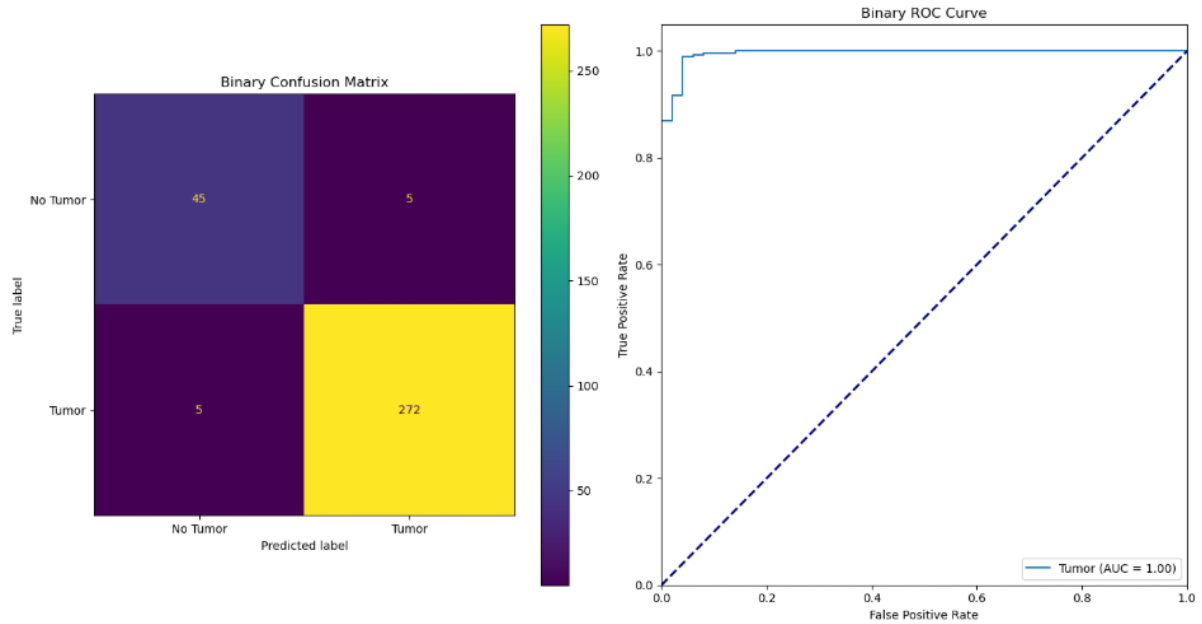


Figure 3: Confusion Matrices and ROC Curves for Multiclass and Binary Classification

Also, the learning curve in Figure 4 below shows that while the training loss steadily decreases, the validation loss is much more volatile. This indicates that the model can learn patterns in the training data well but does not generalize to the validation set as effectively. Thus, there is likely still a slight overfitting issue despite the regularization techniques used, further highlighting areas for model improvement such as more advanced regularization techniques or architectural adjustments. Despite this, the results are promising overall and show that convolutional neural networks are effective for brain tumor detection and diagnosis.

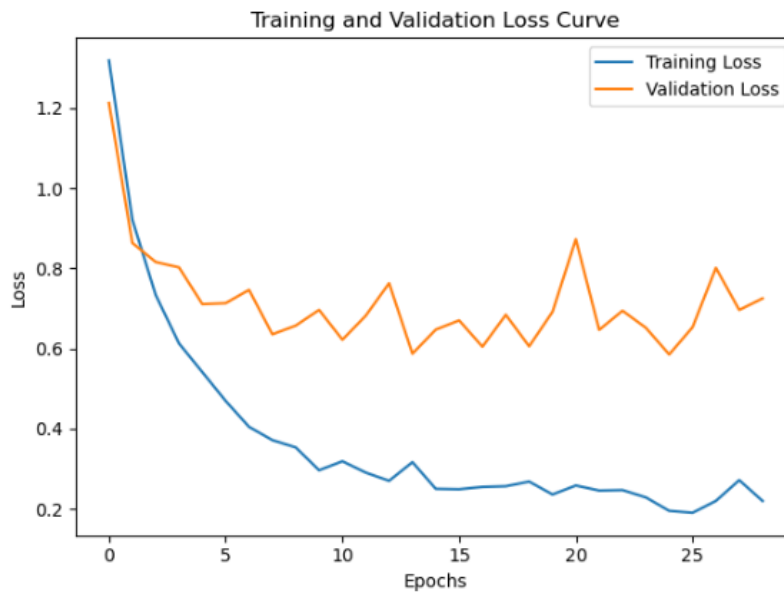


Figure 4: Training and Validation Learning Curve

To gain deeper insights into the model's decision-making process, the first 8 feature maps from each convolutional layer are visualized in Figure 5 below. We can see that the early layers are focused on capturing lower-level features such as the different edges and shapes in the brain. For example, filter 4 in the first layer appears to be looking for the silhouette of the MRI scan. This could help the network determine whether the image is in the sagittal, axial, or coronal plane. In general, the early layers seem to be focused on creating different low-level representations of the MRI scan and beginning to highlight possible tumors.

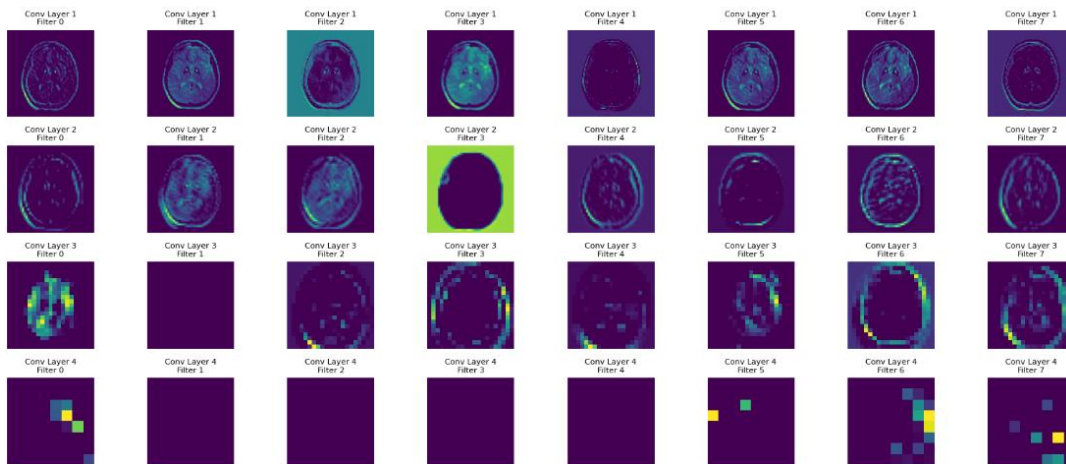
As such, analyzing these lower-level feature maps can give deeper insights into the structural and textural characteristics of different types of tumors. Looking at the first layer for the tumor classes, we can see how the feature maps can highlight the intensity of the pixels in the tumor area. In the second layer, we can get a more granular look at the margin between the tumor and the surrounding brain tissue. These feature maps from the earlier layers do a great job highlighting the shape, texture, and margin of potential tumors from different geometric perspectives. This can help a medical practitioner to better determine the malignancy of a tumor and proceed with proper treatment.

As we move into deeper layers, the feature maps appear to be primarily focused on locating and identifying the tumors themselves. For instance, filters 0 and 3 in the fourth layer of the pituitary tumor feature maps seem to pin the circular tumor it has detected in the center of the brain. Further, we can also see that in all the tumor classes, filter 3 of the fourth layer seems to highlight the general presence of a tumor. On the other hand, filter 3 is empty in the tumor-free case. This kind of information could provide a physician with crucial information on the locations of any potential tumors and help them to make more informed and higher confidence diagnoses.

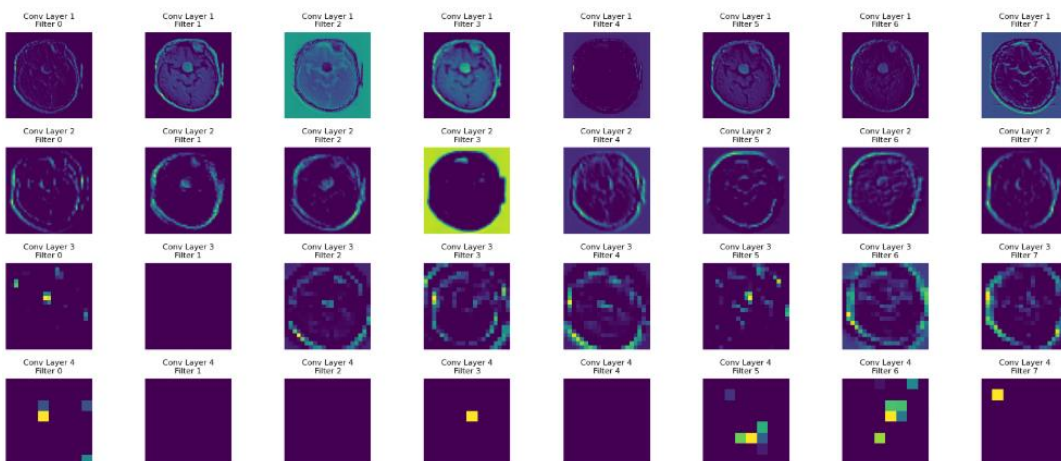
Further, the observation of empty filters, particularly in the deeper layers, is also important to cover. These could indicate overfitting as the network might have some filters that are too specific to the patterns in the training data and therefore cannot generalize on the test image. This notion is also consistent with the instability seen in the validation learning curve. However, it is also possible that the empty feature maps are not problematic. It might indicate efficiency in training through model sparsity, which can have the opposite effect of overfitting and lead to better generalization on unseen images. It is difficult to know which of the two phenomena we are observing, but overfitting is a likely explanation considering prior observations.

Overall, these feature maps provide insight into the hierarchical nature of how convolutional neural networks learn and can deepen our knowledge of how to distinguish between different types of tumors based on their distinctive features such as location, shape, size, and margin.

Feature Maps for Class: no_tumor
Image Index: 246



Feature Maps for Class: pituitary_tumor
Image Index: 162



Feature Maps for Class: meningioma_tumor
Image Index: 211

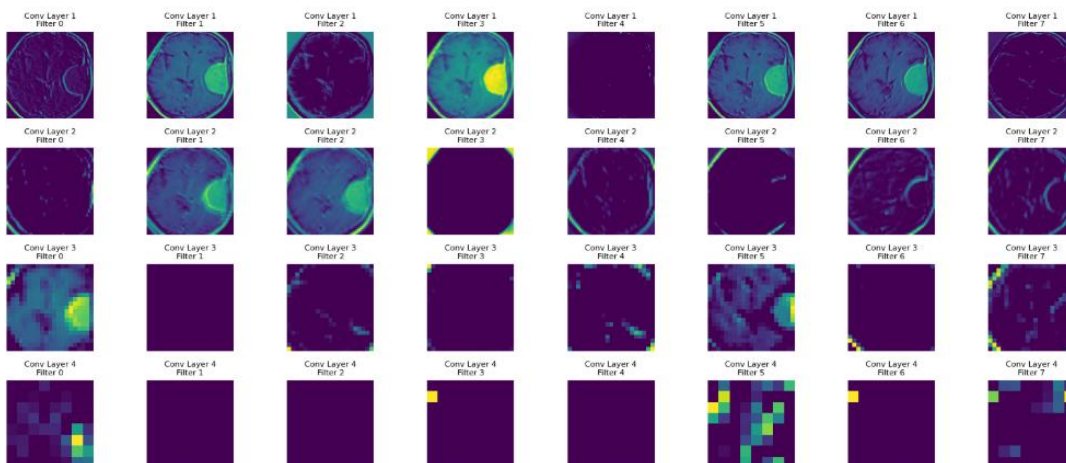




Figure 5: Feature Maps of Different Tumor Classes

Conclusions

My goal with this report was to not only determine the effectiveness of convolutional neural networks for brain tumor detection and diagnosis but also to gain deeper insights into how the model can classify different types of brain tumors and see what knowledge about brain tumors can be gained from the model itself. In general, the performance is quite promising and shows that AI-driven solutions to some of today's biggest health challenges are becoming increasingly effective and feasible. The use of these deep learning methods paves the way for more accurate identification of tumors and can aid physicians in making more informed decisions in a shorter time window without the need for invasive biopsies (Saeedi, Rezayi, Keshavarz, et al., 2023). While there is still much room for further improvement and research within this field, I believe this could mark a huge milestone in the fight against brain cancer.

Moreover, the applications of this technology extend beyond brain cancer. Deep learning methods might also hold promise for diagnosing other brain diseases, such as Alzheimer's. By enabling early detection of these deadly diseases in an accurate and timely manner, we can revolutionize patient outcomes and save lives.

References

Abdusalomov, A. B., Mukhiddinov, M., & Whangbo, T. K. (2023). Brain tumor detection based on deep learning approaches and Magnetic Resonance Imaging. *Cancers*, 15(16), 4172. <https://doi.org/10.3390/cancers15164172>

Cancer of the brain and other nervous system - cancer stat facts. SEER. (n.d.). <https://seer.cancer.gov/statfacts/html/brain.html>

Kaggle. (2020). *Brain Tumor Classification (MRI)*. Retrieved December 7, 2024, from <https://www.kaggle.com/datasets/sartajbhuvaji/brain-tumor-classification-mri>

Khaliki, M.Z., Başarslan, M.S. Brain tumor detection from images and comparison with transfer learning methods and 3-layer CNN. *Sci Rep* **14**, 2664 (2024). <https://doi.org/10.1038/s41598-024-52823-9>

Saeedi, S., Rezayi, S., Keshavarz, H. *et al.* MRI-based brain tumor detection using convolutional deep learning methods and chosen machine learning techniques. *BMC Med Inform Decis Mak* **23**, 16 (2023). <https://doi.org/10.1186/s12911-023-02114-6>