# Automating and Improving Brain Tumor Classification Using Novel Convolutional Neural Network Model

Lagnajeet Panigrahi

## SUMMARY

With each second that passes patients who suffer from brain cancer become more at risk. Delays and misinterpretations from radiologists of brain tumor MRIs (magnetic resonance images) will affect the administration of treatment, endangering patients. It is necessary to consider methods that minimize fallacies present while maximizing time efficiency. This project aims to address the problem using AI classification models, in particular, convolutional neural networks (CNNs) for the purpose of distinguishing between tumor types. This project successfully develops a novel model that is capable of distinguishing—using MRI scans—between glioma, meningioma, pituitary tumors, and unaffected brains. Analysis of the model concluded that it was accurate and consistent: 91 percent accurate and F1-score of 0.92. This displays that CNNs can be used for brain tumor classification and other medical environments for diagnosis. At minimum, they can be used as an aid to experts in the field for more conclusive diagnosis.

## KEYWORDS

**Convolutional Neural Network (CNN), Glioma Tumor, Magnetic Resonance Imaging (MRI), Meningioma Tumor, Pituitary Tumor**

## INTRODUCTION

Brain tumors are masses or the growth of abnormal cells inside an individual's brain: they can be either light-threatening and malignant or harmless and benign (Mayo Clinic, 2021). Various different types of brain tumors can occur including meningioma, glioma, and pituitary tumors. Delays and misinterpretations of data will slow the speed of the administration of treatment, which could have disastrous consequences on a patient, making it all the more important to reach conclusions as fast and accurately as possible (Onder et al., 2021). Currently, brain tumors are being classified by radiologists by performing an MRI (Magnetic Resonance Imaging) scan and then manually analyzing the resulting image. Although this method is viable, it is susceptible to many errors. In fact, in a study, it was found that in the practice of radiology, errors are more common than they should be: it is estimated that the day-to-day rate of error is 3-5% of reported cases (Brady, 2017). Not only is the traditional method prone to mistakes, but it also creates delays in diagnosis due to the length of the process radiologists perform (NCI Staff, 2020). Consequently, these problems result in a system that fails to perform in a manner that is most beneficial to the lives of the patients. To overcome this, a new system is required: one that can diagnose life-threatening brain tumors earlier to ensure the timeline for treatment is not compromised whilst also maintaining accuracy.

Studies show that certain AI models can diagnose brain tumors in significantly less time than radiologists (NCI Staff, 2020). Ultimately, this project strives to use AI, in particular, convolutional neural networks created in Python, to accurately diagnose and categorize brain tumors in an effort to remove human error and shorten delays. There are many types of brain tumors and various subcategories,but this project will focus on identifying and differentiating between healthy brains, meningioma, glioma, and pituitary tumors as they are the most common amongst patients (John Hopkins Medicine, 2022). Such a project will greatly impact the scientific community as it will not only provide a method to automate brain tumor classification, but also pave the way for the use of AI for the diagnosis of other health-related problems.

## METHODS

### Data Collection & Data Preprocessing

The creation of a classification model that leverages MRI scans to identify brain tumors requires a dataset containing target brain tumor MRI scans as well as MRIs of healthy, unaffected brains. The dataset used in this project contains 3264 brain MRIs of healthy brains as well as brains that have the tumors that this experiment is targeting; additionally, this dataset was adapted from Kaggle (Sartaj et al., 2020). After the data is collected, the MRI scans are converted into grayscale, resized, normalized, and converted to arrays.

### Model Architecture Creation

The data will be split into training and testing with eighty percent of the data being contained within the training set - this is to ensure there is ample data for training while also preserving an adequate quantity for testing purposes. The model will be created using, primarily, the Tensorflow library functions: convolutional layers, dense layers, dropout layers, and max-pooling layers. Along with this a data generator will be utilized to randomize the data for better results.

### Model Training

Upon creation of the model, the model will be fit to have a batch size of 32 and be run for 50 epochs to provide substantial time for training and development of methodology for speedy and accurate classification.

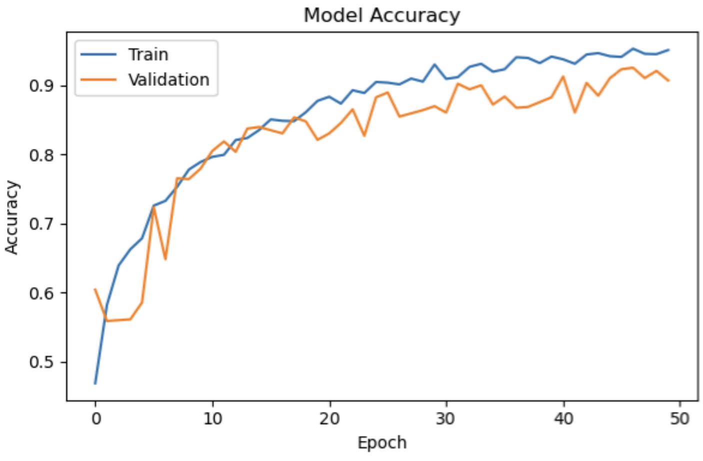
### Model Testing & Data Analysis

The model will be tested on the testing dataset created by the aforementioned split of original data. The model accuracy will be measured using simple binary accuracy and loss as well as an F1-Score (calculated using precision and recall metrics from the model).

## RESULTS AND DISCUSSION

When evaluating convolutional neural network effectiveness, it is essential to screen for overfitting. Overfitting is the principle of the model being trained to the point that it fits too closely to the training data, causing it to be able to only classify the training data, hindering performance on new testing data. Figures 1.1 and 1.2 represent the validation accuracy and training accuracy and validation loss and training loss, respectively. When the validation accuracy— the predicted accuracy on new data— is near the training accuracy, the model fits the data well because it will likely perform similarly to the training accuracy. If there is a significant decrease in validation accuracy, it can be concluded that the model is overfitting. This means that the model is too familiar with the training data and will have difficulty interpreting new samples. For both validation loss and training loss, the lower they are, the better. This is a metric to see the number of errors present during the training and testing phases and these values should also be close together. Since the validation and training accuracies are both high and close to each other and the loss is low, it can be concluded that this model properly fits its purpose and is able to classify new and unfamiliar data.

After establishing that the model is not overfitting, the model's accuracy must be analyzed. Figure 2, a confusion matrix, and Table 1, a classification report, can be utilized to determine the overall accuracy. In both the confusion matrix and the classification report, the numbers zero, one, two, and three represent the different classes being predicted: glioma, meningioma, pituitary, and healthy brain, respectively. The confusion matrix and classification report indicate a strong overall performance for this brain tumor classification model, which achieves a binary accuracy of 91% (Total correct predictions divided by total predictions). Analyzing the precision, recall, and F1-scores for each class, balanced metrics can be observed, with only minor variations, indicating consistency. The model performs uniformly across classes, with the macro average and weighted average F1-scores at 0.91(combination of precision and recall scores on a scale 0-1, where higher value indicates higher accuracy and consistency). Overall, the accuracy of 91% and the F1-Score is a robust indicator of this model's reliability across multiple categories.



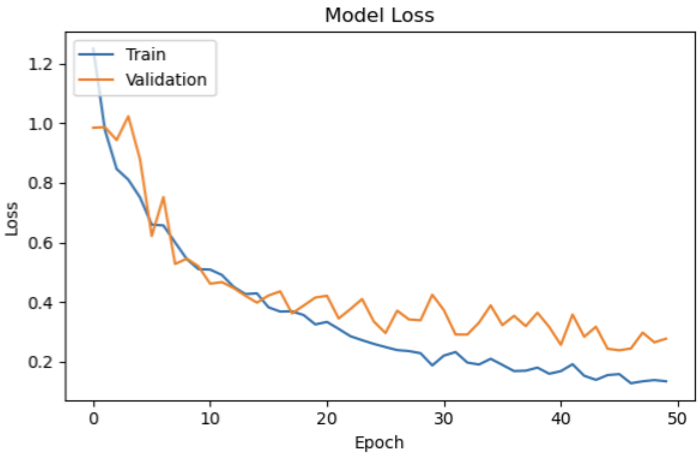


**Figure 1.1 Model Accuracy:** Outlines both the accuracy of the model during training and the predicted accuracy on new data (validation accuracy) over the course of the 50 epochs of training.

**Figure 1.2 Model Loss:** Outlines both the loss of the model during training and the predicted loss on new data (validation loss) over the course of the 50 epochs of training.

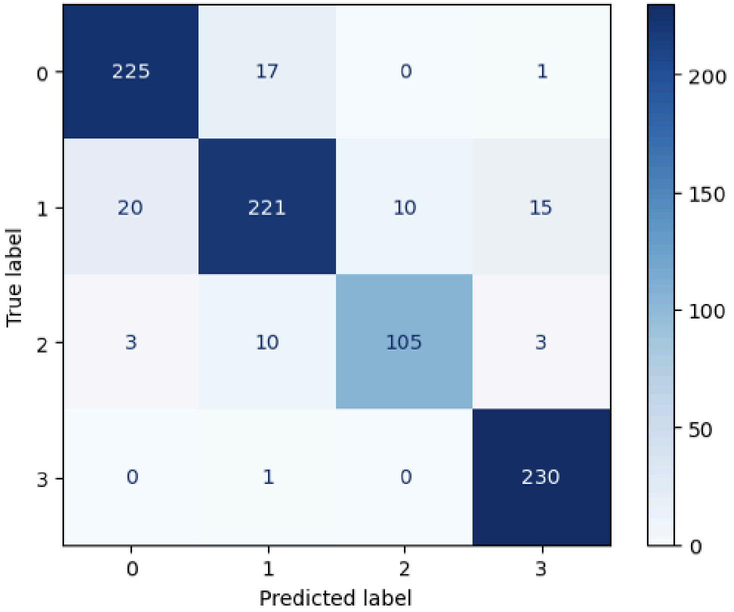
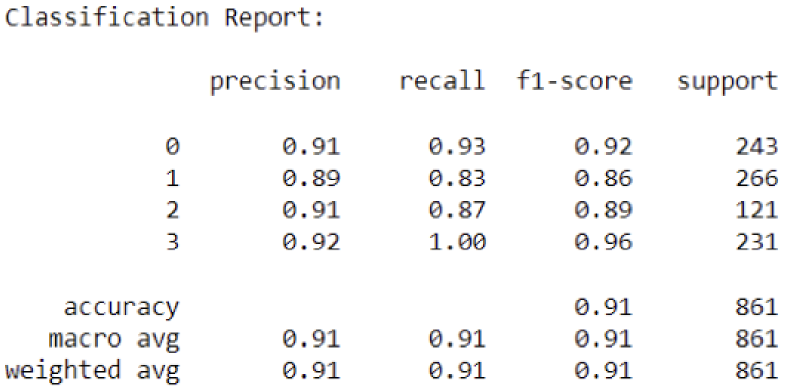


**Figure 2 Confusion Matrix:** Displays the quantity of correct classifications during testing (where predicted label and true label are equivalent) and incorrect classifications (where predicted label and true label are not equivalent). The labels zero, one, two, and three represent the different classes being predicted: glioma, meningioma, pituitary, and healthy brain, respectively.



```
Classification Report:

              precision    recall  f1-score   support

           0       0.91      0.93      0.92       243
           1       0.89      0.83      0.86       266
           2       0.91      0.87      0.89       121
           3       0.92      1.00      0.96       231

    accuracy                           0.91       861
   macro avg       0.91      0.91      0.91       861
weighted avg       0.91      0.91      0.91       861
```

**Table 1 Classification Report:** Displays the data from Figure 2 converted to metrics (F1-score, recall, and precision) as well as the binary accuracy. The numbers zero, one, two, and three represent the different classes being predicted: glioma, meningioma, pituitary, and healthy brain, respectively.

## CONCLUSIONS

The primary objective of this project was to use artificial intelligence, in particular convolutional neural networks within deep learning, to create a model that can classify brain tumors quickly and with a minimum of 90 percent accuracy. This was achieved as the model reached an accuracy above 90 percent while remaining consistent. This project has the capability to save countless lives due to earlier diagnosis, allowing for more time for treatment. These models, if trained accordingly, have the potential to be more accurate than radiologists, saving even more lives. If not trusted with full diagnosis, at the very minimum, models such as this can be used as a safeguard or as a second set of eyes for diagnosis. Radiologists can reach their own conclusions about diagnosis and then use classification models to verify their findings.

## REFERENCES

1. Brady, A. P. (2017, February). Error and discrepancy in radiology: Inevitable or avoidable? Insights into imaging. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5265198/
2. John Hopkins Medicine. (2021, November 8). Brain tumor types. Johns Hopkins Medicine. Retrieved from https://www.hopkinsmedicine.org/health/conditions-and-diseases/brain-tumor/brain-tumor-types
3. Mayo Foundation for Medical Education and Research. (2021, August 6). Brain Tumor. Mayo Clinic. Retrieved from https://www.mayoclinic.org/diseases-conditions/brain-tumor/symptoms-causes/syc-20350084
4. NCI Staff. (2020, February 12). Artificial Intelligence Expedites Brain Tumor Diagnosis. National Cancer Institute. Retrieved from https://www.cancer.gov/news-events/cancer-currents-blog/2020/artificial-intelligence-brain-tumor-diagnosis-surgery
5. Onder, O., Yarasir, Y., Azizova, A., Durhan, G., Onur, M. R., &amp; Ariyurek, O. M. (2021, April 20). Errors, discrepancies and underlying bias in radiology with case examples: A pictorial review - insights into imaging. SpringerOpen. Retrieved from https://insightsimaging.springeropen.com/articles/10.1186/s13244-021-00986-8
6. Sartaj Bhuvaji, Ankita Kadam, Prajakta Bhumkar, Sameer Dedge, and Swati Kanchan. (2020). Brain Tumor Classification (MRI) [Data set]. Kaggle. https://doi.org/10.34740/KAGGLE/DSV/1183165