HIGH SCHOOL EDITION
Journal of Student Research

Volume 13 Issue 3 (2024)

Advancing Brain Tumor Diagnosis through Machine Learning: A Comparative Study

Dhruv Veda

Centennial High School, USA

ABSTRACT

Brain tumor is a devastating disease affecting thousands of Americans every year. The disease requires an early and accurate diagnosis. Machine learning could be a very powerful way to speed up the diagnosis. This study explores the efficacy of various machine learning models in diagnosing and classifying brain tumors using MRI scans. Convolutional Neural Network (CNN) models were compared with traditional machine learning algorithms, including Logistic Regression, K-Nearest Neighbors (KNN), Decision Tree, and Random Forest (RF), on a dataset containing MRI images of different brain tumor types. The study came to the conclusion that the CNN was more effective than other models, and all of these models would need larger datasets before considering them usable as a medical tool.

Introduction

In 2022, 72,360 Americans of age 40 plus were diagnosed with some form of brain tumor (*Brain Tumor Facts*. (n.d.).) Brain tumor is a disease in which a rapid cell growth (tumor) forms in the brain often due to a genetic mutation of a tumor suppressor gene which is responsible for regulating the cell cycle. There are multiple types of brain tumors and some of the most common types are glioma tumor (Mayo Clinic. (2019). *Glioma - Symptoms and causes*.), meningioma tumor (Mayo Clinic. (2024, March 29).), and pituitary tumor (Mayo Clinic. (2019). *Pituitary tumors - Symptoms and causes*.). The major difference in these is based on the location of the tumor itself. Depending on the location of the tumor the symptoms themselves may also vary. Glioma tumors can cause speech and memory impairments (Mayo Clinic. (2019). *Glioma - Symptoms and causes*.). Pituitary tumors can cause an underdosage of pituitary hormones (Mayo Clinic. (2019). *Pituitary tumors - Symptoms and causes*.), and meningioma tumors can cause memory loss, vision loss, and hearing loss (Mayo Clinic. (2024, March 29).). A key challenge of these diseases is their quick onset and the difficulty in their treatment. There are three major treatments for brain tumor - chemotherapy, radiation therapy, and surgery (Mayo Clinic. (2021, August 6).).

Traditional Diagnosis

Even after these treatments there is no guarantee that the tumor will subside. There is therefore a need for a quick diagnosis of the tumor in order to mitigate its effects before it develops. There are three major tests used to diagnose a brain tumor: Neurological exams, imaging tests, and a biopsy (Mayo Clinic. (2021, August 6).). Often multiple of these tests are used as a confirmation method with one of the first checks being a Magnetic Resonance Imaging (MRI) scan. The difficulty with an MRI scan is that it can be hard for physicians to analyze the image and spot the discrepancy associated with a tumor and distinguish it from images with no tumor. To assist medical professionals in this endeavor, research scientists have begun to develop artificial intelligence (AI) programs to help physicians diagnose tumors earlier from signs they would not have recognized in the past.

Artificial Intelligence (AI) for Diagnosis

There are multiple attempts at the creation of a highly accurate AI all of which have varying effectiveness. The first study was a review conducted by Omar Kouli and their main goal was to document the streamlining of the MRI examination process (Kouli, O., Hassane, A., Badran, D., Kouli, T., Hossain-Ibrahim, K., & Steele, J. D. (2022).). They decided to extract data from the Rayan systematic review website. The data for the papers included MRI brain scans from multiple different databases and sources using a set of 38 papers. The data specified MRIs with and without brain tumors but no specification of tumor types.

Many of the successful studies used a deep learning neural network model. The most successful studies have achieved an accuracy of about 78.6%. Despite the relative success of this method, few tests used a deep learning model. Some notable limitations include the lack of explanation about the models used. The selection criteria meant that only models trained on a tumor vs. no tumor dataset, with no differentiation between the tumors, were examined.

Another notable study was conducted by Javier Díaz-Pernas (Díaz-Pernas, F. J., Martínez-Zarzuela, M., Antón-Rodríguez, M., & González-Ortega, D. (2021).). The goal of this study was to aid in the identification of brain tumors due to the difficulty of the process thanks to minor variations in brain shapes and sizes. They used a convolutional neural network (CNN). The dataset that was collected was MRI scans collected from the Chinese general hospitals in Nanfang and Guangzhou. The dataset contained 3064 images taken from 233 patients They trained a 16-layer deep learning model to separate glioma tumors, meningioma tumors, and pituitary tumors. They could get an accuracy of up to 97%.

A separate study used Transfer Learning (Mehrotra, R., Ansari, M. A., Agrawal, R., & Anand, R. S. (2020).). The main goal was to assist in the severity of the tumor in order to help the professional ascertain the best treatment. They collected 696 MRI images of slices of the brain with either positive or negative tests from online databases. The dataset was applied to a convolutional neural network model and achieved results of up to 98% accuracy. However, this system did not distinguish types of tumors.

Finally, a study by E. Irmak focused on helping reduce the high number of errors made in brain tumor diagnosis (Irmak, E. (2021).). They combined the publicly available RIDER, REMBRANT, TCGA-LCG, and TCIA datasets and used a CNN with some models achieving up to 99% accuracy. The data used included information on diagnosis for each patient. It had 1640 positive samples and 1350 negative samples.

Need for a Model Comparison

Although there are multiple studies there is a clear information gap when it comes to the comparison of multiple models. This project compares different models including a Convolutional Neural Network, K Nearest Neighbors, Decision Tree, Logistic Regression, and Random Forest model. It focuses on their effectiveness at image classification as well as what scenarios (aiding in diagnosis of the tumor or aiding in classification of the tumor) each model might be most useful in.

Methods

Dataset Used

For this project, the dataset used was “Brain Tumor Classification (MRI)” (Chitnis, S., Hosseini, R., & Xie, P. (2022).), which contained 394 images. The images were MRI scans sorted into four classes; no tumor, pituitary tumor, Glioma Tumor, and meningioma tumor. As detailed in figure 1, there was an imbalance in the number of tumors. There were significantly less MRIs with no tumors (400 data points) than all the other classes (800 data points each). This imbalance was addressed with the metrics used in order to get an accurate assessment of the tested models.

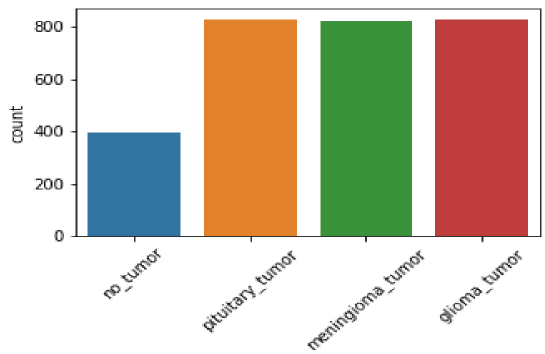

Figure 1. Distribution within Class Types

Each of the tumors are separated based on the location of the tumor within the brain (Figure 2) with each tumor being in the pituitary gland, glial area, and meninges layer of the brain. The data was gathered from patient MRI scans and any identifying information was removed. It is publicly available on Kaggle and was created by Sartaj Bhuvaji, Ankita Kadam, Prajakta Bhumkar, Sameer Dedge.

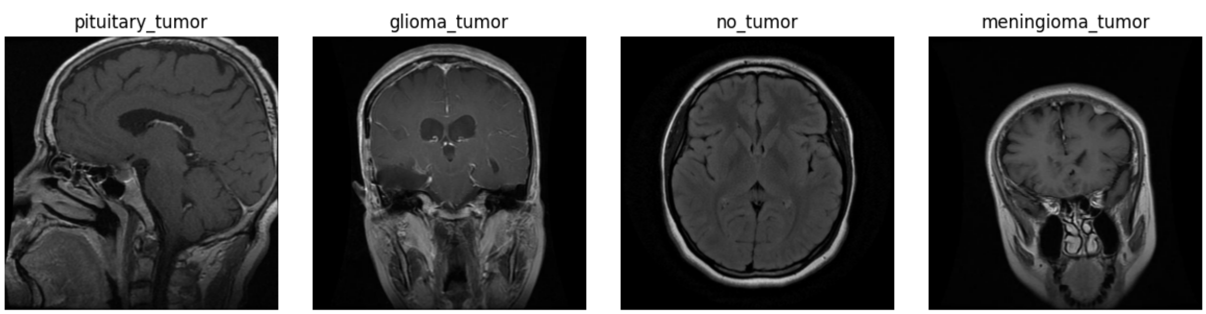

Figure 2. Types of Tumors

The data was separated into train and test data in a random 88:22 split. Two separate sets of data were created, one with the four main classes and one with two classes: tumor and no tumor. This would allow for analysis of the model in multiple situations such as diagnosis vs differentiation of tumors. The data was then flattened and one hot encoded so that it could be fed to the models.

Machine Learning (ML) Models Used in this Project

The models used were a CNN from Keras, a K-nearest neighbor (KNN) algorithm from sklearn, a Logistical Regression from sklearn, a decision tree from sklearn, and a Random Forest (RF) Classifier from sklearn. Each model was controlled with a set of hyperparameters being changed over a range of values in order to produce the most effective model possible.

- For the logistical regression, the maximum number of iterations the algorithm would perform was tuned.

- For the KNN, the number of neighbors was iterated over, kept low to avoid overfitting, and tested along a range of 1-20. The model, taken from the sklearn library, used simple flattened versions of the image data.
- For the RF Classifier, the number of trees used in the algorithm was tuned.
- For the Decision tree, the Entropy model was used, and the maximum depth of the model was tuned as the independent variable in the study.

The CNN layers were adjusted to avoid overfitting while also achieving the best results possible. The layers of the CNN are detailed in figure 3.

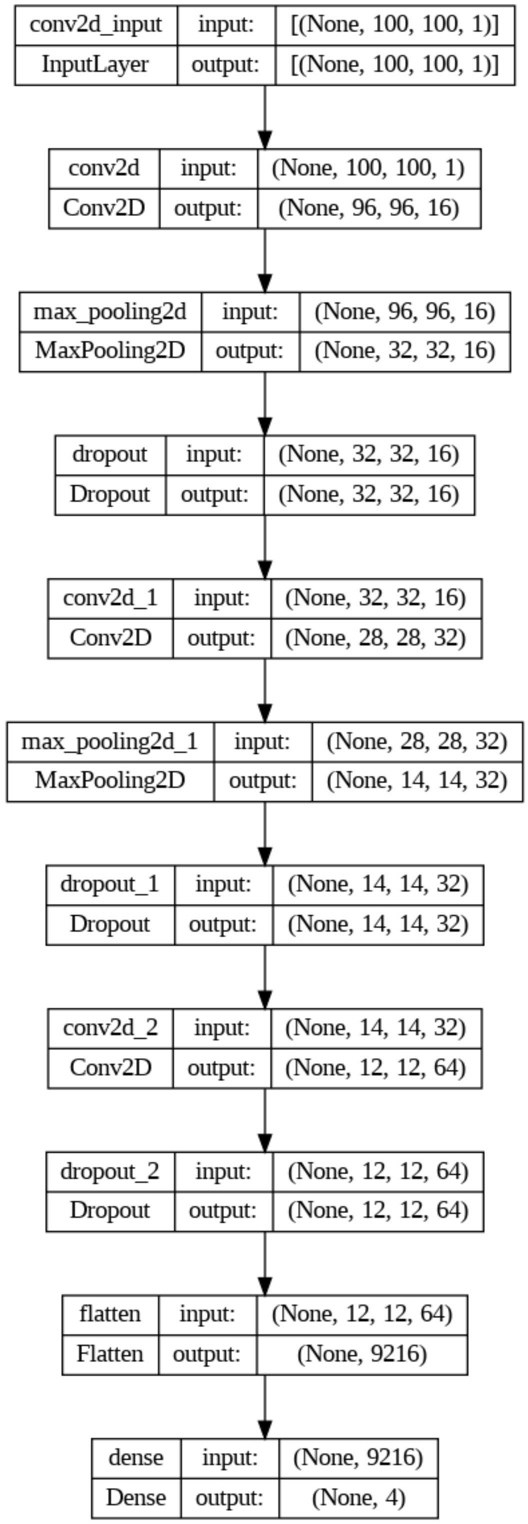

Figure 3. Convolutional Neural Network Layers

Results

The models were tested by iterating through a series of variables to see how it would affect various scores. These scores were based on the dataset that was used, two classes or four classes, and the metrics. Both f1 score and accuracy were used. The f1 score is much better with small and uneven amounts of data between classes, whereas accuracy was used in order to allow for comparison with past studies that had used the metric of accuracy.

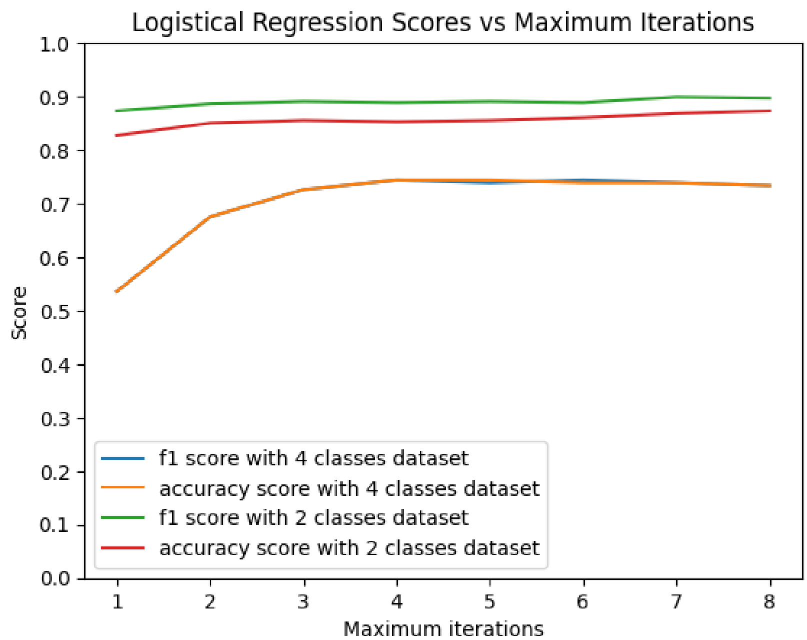

Figure 4. Logistical Regression Scores VS Maximum Iterations

The Logistical regression model achieved high scores in the two classes dataset with f1 scores around 90% and accuracy scores around 5% less than that while on the two classes data both the f1 score and accuracy score seem to be close to 70%.

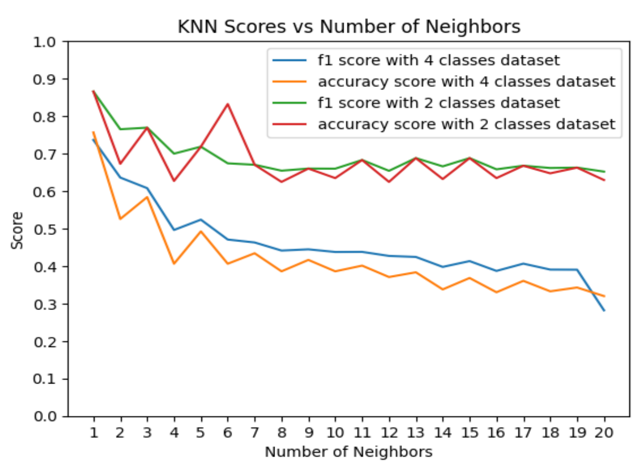

Figure 5. KNN Scores VS Number of Neighbors

The KNN model was more accurate when there were 1 to 3 neighbors. The highest f1 score obtained for the 4-class score was at 75% and for the 2 class score it was 87%. The accuracy score followed similar trends and seemed to be generally less than the f1 score on average.

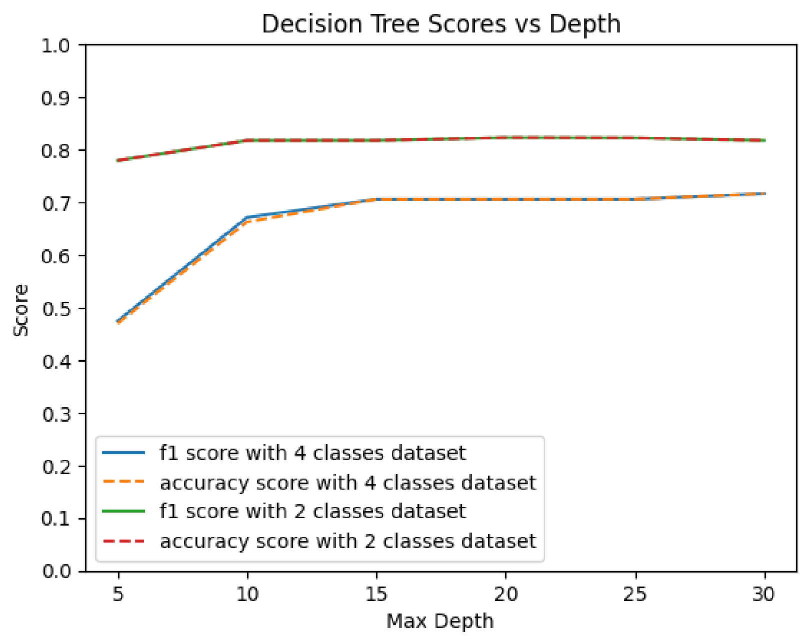

Figure 6 Decision Tree Scores VS Depth

The Decision Tree model showed consistency between the accuracy scores and f1 scores showing that it was fairly accurate at identifying each type of tumor. On the 2 classes data it scored around 80% but it could only reach around 70% with 4 classes data.

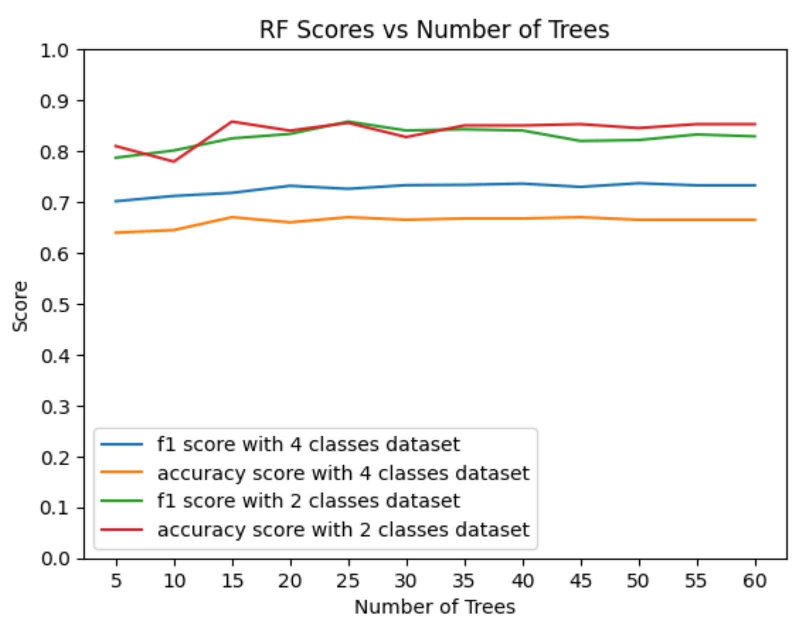

Figure 7 RF Scores VS Number of Trees

The Random Forest model obtained a 72% f1 score in the 4 classes dataset and 73% f1 score in the two-class dataset. The accuracies tended to be consistently lower on the 4 classes dataset but about the same as the f1 score on 2 classes dataset and stayed fairly consistent.

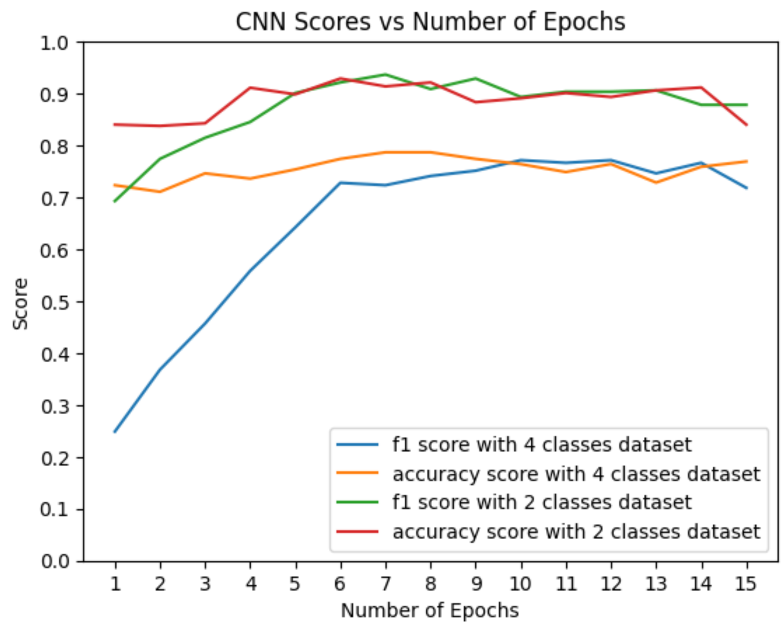

Figure 8 CNN Scores VS Number of Epochs

The CNN model tended to be more effective as more epochs were used but after around 10 epochs the model began to achieve similar results. This resulted in the f1 score for the four classes getting around a 78% f1 score whereas the 2 classes data allowed the model to reach around 93% f1 score. The accuracies seem to closely match with the f1 scores at higher epochs by accuracies were much greater than f1 scores when the models were trained less.

Model	4-class f1	4-class accuracy	2-class f1	2-class accuracy
Logistical	0.74365	0.74365	0.89905	0.87310
KNN	0.73635	0.75635	0.86548	0.86548
RF	0.85787	0.67005	0.73669	0.85786
Decision Tree	0.71622	0.71622	0.82277	0.82227
CNN	0.77157	0.78680	0.92893	0.92132

Figure 9 Comparison of Accuracy by Model Types

Discussion

In this project, multiple different model types were studied in order to determine the most useful model for the diagnosis and classification of various forms of brain tumors. In this process, a model was built that can differentiate meningioma tumor, pituitary tumor, and glioma tumor from no tumor with an f1-score of 0.77157.

The CNN model was found to be more effective than basic ML models at this classification task. The implementation of this algorithm may empower physicians by providing a likelihood score calculated from the patient's MRI scan, assisting in identifying patients at risk for having a tumor.

The model could also be more effectively used with larger datasets or some other similar techniques.

Limitations

Several limitations exist for this project including but not limited to low amounts of data and a lack of tested models. The data used contained a low number of samples and a more effective version of this algorithm could be created by training the same CNN models but producing a higher f1 score.

Future plans for this project include, testing more models on the same data to see which model achieves a higher f1 score. Additionally, finding a larger dataset could help determine if more training data affects the model's accuracy.

Acknowledgments

The dataset was made publicly available on Kaggle by Sartaj Bhuvaji, Ankita Kadam, Prajakta Bhumkar, Sameer Dedge. This was the best sample of data publicly available for use.

References

Brain Tumor Facts. (n.d.). National Brain Tumor Society. https://braintumor.org/brain-tumors/about-brain-tumors/brain-tumor-facts/#:~:text=An%20estimated%2072%2C360%20adults%20age

Mayo Clinic. (2019). *Glioma - Symptoms and causes*. Mayo Clinic. https://www.mayoclinic.org/diseases-conditions/glioma/symptoms-causes/syc-20350251

Mayo Clinic. (2024, March 29). *Meningioma - Symptoms and causes*. Mayo Clinic. https://www.mayoclinic.org/diseases-conditions/meningioma/symptoms-causes/syc-20355643

Mayo Clinic. (2019). *Pituitary tumors - Symptoms and causes*. Mayo Clinic. https://www.mayoclinic.org/diseases-conditions/pituitary-tumors/symptoms-causes/syc-20350548

Mayo Clinic. (2021, August 6). *Brain tumor - Symptoms and causes*. Mayo Clinic. https://www.mayoclinic.org/diseases-conditions/brain-tumor/symptoms-causes/syc-20350084

Kouli, O., Hassane, A., Badran, D., Kouli, T., Hossain-Ibrahim, K., & Steele, J. D. (2022). Automated brain tumor identification using magnetic resonance imaging: A systematic review and meta-analysis. *Neuro-Oncology Advances*, *4*(1). https://doi.org/10.1093/noajnl/vdac081

Díaz-Pernas, F. J., Martínez-Zarzuela, M., Antón-Rodríguez, M., & González-Ortega, D. (2021). A Deep Learning Approach for Brain Tumor Classification and Segmentation Using a Multiscale Convolutional Neural Network. *Healthcare*, *9*(2), 153. https://doi.org/10.3390/healthcare9020153

Mehrotra, R., Ansari, M. A., Agrawal, R., & Anand, R. S. (2020). A Transfer Learning approach for AI-based classification of brain tumors. *Machine Learning with Applications*, *2*, 100003. https://doi.org/10.1016/j.mlwa.2020.100003

Irmak, E. (2021). Multi-Classification of Brain Tumor MRI Images Using Deep Convolutional Neural Network with Fully Optimized Framework. *Iranian Journal of Science and Technology, Transactions of Electrical Engineering*, *45*(3), 1015–1036. https://doi.org/10.1007/s40998-021-00426-9

Chitnis, S., Hosseini, R., & Xie, P. (2022). Brain tumor classification based on neural architecture search. *Scientific Reports*, *12*(1). https://doi.org/10.1038/s41598-022-22172-6