Brain Tumor Detection

Background

Over 700,000 Americans are living with a brain tumor today and studies show that more than 84,000 people will be diagnosed with a primary brain tumor in 2021[1]. Diagnosing a brain tumor usually begins with a Magnetic Resonance Imaging (MRI) scan. The results are then reviewed by a neurologist to see if there is a tumor in the brain.

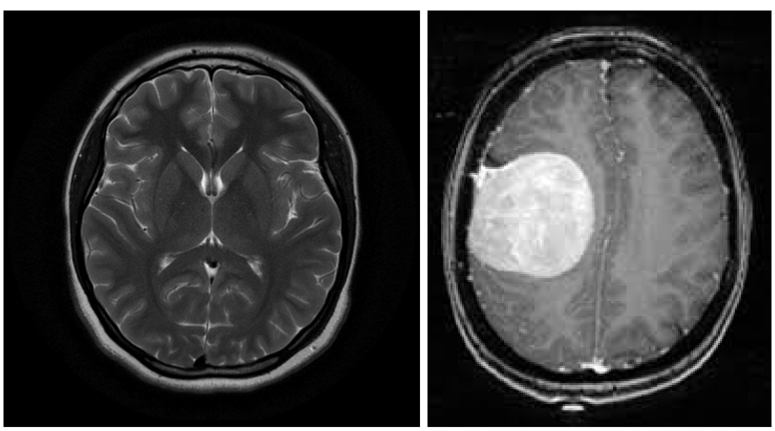
Problem

Using artificial intelligence to detect a brain tumor from the MRI scan would save money and most importantly time. Not only that, this could also possibly reduce human error in the tumor detection. With today’s ever-growing population, it is imperative that doctor’s use technology to determine brain scan results in a timely fashion.

Dataset

For the task in hand, we chose to use the [brain MRI images for brain tumor detection] from Kaggle. This dataset contains 155 images of MRI scans with brain tumors and 98 without any tumor.

Below are sample scans showing a brain without tumor (left) and a brain with tumor(right).



All images are grayscale, thus when imported will have the same value for R, G and B channels. We can use any one channel for our model. However, all images do not have the same dimensions and hence will have to be resized or padded before passing them into the neural network model.

Solution

The objective of this project is to build a **neural network model** that accurately classifies MRI brain scans as having a tumor or not. The input data images will be used to train the model to classify MRI scan images as having a tumor or not.

Benchmark

As a benchmark, we will build a **Support Vector Classifier** to compare against our neural network model. The contributor to this dataset on Kaggle does not specify the data source and hence we will not be able to compare our model performance to a published benchmark.

Evaluation metrics

Accuracy will be the primary metric used to evaluate the models.

$$Accuracy \ = \ \frac{TP + TN}{TP + TN + FP + FN}$$

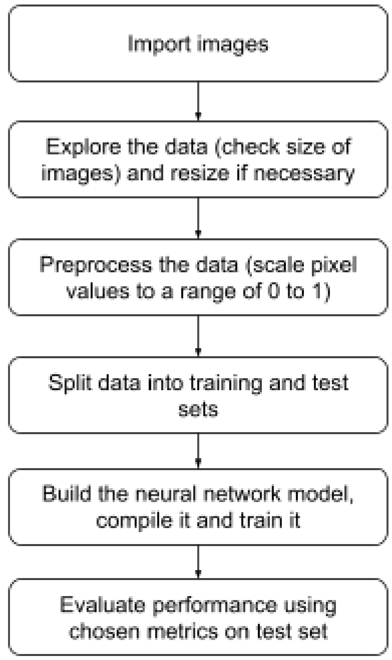
The secondary evaluation metric used will be **recall**. This is an important metric as it penalizes the false negative classifications. If a person has a tumor and our model predicts that the scan shows no tumor, that would be a terrible mistake!

$$Recall \ = \ \frac{TP}{TP + FN}$$

where TP: $True\ Positive,\ TN$: $True\ Negative,\ FP$: $False\ Positive,\ FN$: $False\ Negative$

Project design

The workflow to approach the problem by building a **neural network model** to classify brain scans is summarized in the flowchart below. Note, that the best parameters to configure the layers of the model will be obtained by **hyperparameter tuning**. Once we are satisfied with the neural network model, we will build a **Support Vector Machine classifier** to compare its performance.



References

1. Porter KR, McCarthy BJ, Freels S,Kim Y, Davis FG. Prevalence estimates for primary brain tumors in the United States by age, gender, behavior, and histology. Neuro-Oncology 12(6):520-527, 2010.