Proposal
Brain Cancer Detection MRI Images

CS 375 - H01
Ewan Lee, el44; Kie Li, jl263; Faith Leverett, fml

I. OBJECTIVE

The primary objective of this project is to create a convolutional neural network (CNN) model that is capable of performing binary image classification on a data set consisting of MRI scans of the brain. The classification divides the images into two categories and so must be able to distinguish between images that display brain cancer or are non-cancerous. By the end of the project, this model should be able to successfully perform this classification on our dataset with over 90% accuracy.

II. BACKGROUND AND MOTIVATION

Brain tumors affect about 30 adults out of 100,000. Less than 1% of the time, these tumors will be diagnosed as malignant and eventually cause brain cancer. There are many different kinds of brain cancers; many are extremely dangerous and usually are highly difficult to detect, with some having up to a 95% fatality rate within the first five years after diagnosis. As such, it is extremely important to be able to quickly and accurately identify MRI scans which display the telltale signs of brain cancer in order to issue swift diagnoses. The motivation for this project is to create a CNN which can make these classifications and thus give those affected by brain cancer an accurate measure of information about their health.

III. DATA SOURCE

The chosen dataset is the "Brain Cancer Detection MRI Images" on kaggle. The dataset contains 800 images of the brain, 408 images being normal brains and 392 of them being of an abnormal brain. In terms of looking at the raw images, the difference between them at first seems minimal. However, upon closer examination, the brains that are abnormal look more deformed or warped than the supposedly normal brain.

IV. METHODOLOGY

Given that we have a dataset of images that are to be classified as either normal or abnormal brains, we must use a binary image classification

A. Data Preprocessing : We must make sure that the images are split into training and testing sets, ideally by shuffling them to mix up any order. We may also have to reshape the images in order to train them to a model.

B. Feature Engineering : We will use layers, also known as kernels, and apply them to the data. Then we will train the data in order to find the best weights to make an accurate classification model.

C. Model : The model will be a CNN (Convolutional Neural Network) since it is known to be used for preprocessing and image recognition

D. Evaluation Metrics : The evaluation metrics to be used will be accuracy and ability for the model to recall and be able to consistently predict and classify images.

V. EXPECTED OUTCOMES

The expected outcomes of this project include creating a machine learning model that can detect brain cancer through MRI scans. The project will generate predictions based on exploratory data analysis to provide insight into patterns and features that may be early indicators or precursors to cancer. The project will also include an evaluation on the model's accuracy and performance as well as an analysis of the differences between brain cancer and non-cancer images.

VI. TIMELINE

Week	Tasks	allocation
Week 1-2	Data collection and reprocessing	Faith
Week 2-3	EDA (Exploratory Data Analysis)	Faith
Week 2-3	Related works	Kie
Week 3-4	Model implementation	Ewan
Week 5-6	Model evaluation and refinement	Ewan
Week 7	Final model section and report preparation	Kie
Week 8	Project presentation and submission	All members

REFERENCES [optional]

[1] https://www.saberhealth.com/news/blog/brain-cancer-facts
[2] https://www.hopkinsmedicine.org/health/conditions-and-diseases/brain-tumor