Final Report: Brain Tumor Classification Using CNN

---

1. Introduction

This project aims to develop a Convolutional Neural Network (CNN) model for classifying brain tumors from MRI images into four categories: glioma, meningioma, pituitary tumor, and no tumor. The dataset used consists of grayscale MRI scans labeled accordingly, and the goal is to automate early tumor detection to assist radiologists and improve diagnostic accuracy.

---

2. Data Preparation

Preprocessing Steps: - All images were resized to 128x128 pixels. - Since the images are grayscale, a single color channel was retained. - Labels were one-hot encoded to match the softmax output. - Pixel values were normalized by dividing by 255.

Augmentation Techniques Used: - Rotation - Zoom - Width/height shift - Horizontal flipping

Sample Visualizations: Random samples from each class were visualized to verify image quality and label consistency.

---

3. Model Building & Training

Model Architecture (for all experiments):

- Conv2D(32 filters) + MaxPooling
- Conv2D(64 filters) + MaxPooling
- Conv2D(128 filters) + MaxPooling
- Flatten
- Dropout (0.5)
- Dense(64) + Dense(4, softmax)

Optimizer and Hyperparameters:

- Optimizer: Adam
- Learning rate: 0.001
- Loss: Categorical Crossentropy
- Batch size: 32
- Epochs: 10

Hardware:

- Kaggle environment (GPU-accelerated)

---

4. Results & Comparisons

Training & Validation Accuracy Curves:

Despite several adjustments (such as data augmentation and dropout), the model's validation accuracy showed a **declining trend** in later experiments.

Confusion Matrix:

A confusion matrix was generated on the validation set to analyze class-wise performance.

Summary Table:

Experiment	Data Augmentation	Dropout	Validation Accuracy
Baseline	No	No	~84%
+ Dropout	No	0.5	~82%
+ Aug	Yes	0.5	~80%

---

5. Insights & Observations

- Adding dropout slightly reduced overfitting but did not improve validation accuracy.
- Data augmentation made the model more robust to image variance but led to lower accuracy, possibly due to increased noise.
- Overall, attempts to improve generalization **did not yield better results**; in fact, the baseline model performed best.
- This might be due to the small dataset size, short training duration, or insufficient tuning.

---

6. Conclusion

The CNN model was able to classify brain tumors with reasonable accuracy. However, further improvements like transfer learning, deeper architectures, or more epochs could be explored. Fine-tuning and using larger, balanced datasets might help overcome current limitations.

---

7. Code & Implementation Reference

All code, models, and experiments are included in the attached Jupyter Notebook file: `cnn-final-project.ipynb`, which was implemented using the Kaggle environment.