Brain Tumor Detection

Question 1: Elevator Pitch

Our project focuses on brain tumor detection using advanced deep learning techniques. Leveraging MRI imaging data, we aim to develop a robust Convolutional Neural Network (CNN) model for classifying and identifying brain tumors. This will assist in faster, accurate diagnosis, supporting medical professionals in improving patient outcomes.

Question 2: Dataset Details

1) Collector(s):
Public datasets such as The Cancer Imaging Archive (TCIA) or Kaggle contributors.

2) Year:
The dataset may include MRI scans collected from 2015–2024, depending on the source.

3) Title of Dataset:
"Brain Tumor MRI Images Dataset"

4) Version Number (if any):
Version 1.0 or any version provided by the dataset source.

5) Publisher:
The Cancer Imaging Archive (TCIA) or Kaggle contributors.

6) DOI or URL:

- TCIA: https://www.cancerimagingarchive.net/
- Example Kaggle Dataset: https://www.kaggle.com/

7) Study/Paper/Reason:
The dataset was collected to aid in the early detection and classification of brain tumors, supporting research in medical imaging and oncology.

Question 3: Language and Libraries

Language:

- Python 3.13.1

Libraries:

- **Data Preprocessing & Augmentation:** OpenCV, Pillow
- **Model Building:** TensorFlow/Keras, PyTorch
- **Data Management:** Pandas, NumPy

- **Visualization:** Matplotlib, Seaborn
- **Evaluation:** Scikit-learn

Question 4: Code will write our own

- **Data Preprocessing:** Handling MRI image formats, resizing, normalization, and augmenting images (rotation, flipping, etc.).
- **Model Architecture:** Defining a CNN model to classify MRI images as tumorous or non-tumorous.
- **Training and Validation:** Writing scripts for model training, including loss and accuracy tracking.
- **Hyperparameter Tuning:** Implementing custom functions for hyperparameter search.
- **Performance Analysis:** Generating confusion matrices, ROC curves, and other evaluation metrics.
- **Visualization:** Displaying predictions with highlighted tumor regions (if segmentation is involved).

Question 5: Best Choice of Model

Model Choice:

- **Convolutional Neural Networks (CNNs):** Specifically, architectures like ResNet, VGG, or a custom-built CNN.
- **Why:** CNNs are highly effective at identifying spatial features in image data, making them ideal for medical image classification tasks.

Question 6: Hyperparameters and Optimization

Key Hyperparameters:

1. Learning Rate
2. Batch Size
3. Number of Filters
4. Kernel Size
5. Dropout Rate
6. Number of Layers

Optimization Strategy:

- Use grid search or random search for initial tuning.
- Employ learning rate schedulers or adaptive optimizers like Adam or RMSprop.
- Validate hyperparameters using a stratified k-fold cross-validation approach.

Question 7: Performance Evaluation

Metrics:

1. **Accuracy:** To measure overall classification performance.
2. **Precision and Recall:** To evaluate the model's ability to correctly identify tumors and avoid false negatives.
3. **F1-Score:** To provide a balance between precision and recall.
4. **AUC-ROC:** To assess model performance across different classification thresholds.

Techniques:

- Test the model on a separate test set to evaluate generalization.
- Visualize predictions with overlays on MRI scans (if segmentation is performed).
- Compare results with baseline models or published research.