Week 7: Brain Tumor Classification Challenge

Overview

📌 Overview

In this competition, participants will classify brain tumor images from a provided MRI dataset into distinct categories. The dataset consists of multiple MRI scans of brain tumors, and the goal is to develop a model that accurately classifies the tumor type or predicts the presence of a tumor based on MRI images.

📝 Task Description

Participants are tasked with developing a machine learning or deep learning model that classifies brain tumor images into the appropriate categories. You will work with the provided MRI dataset, which includes images of brain tumors with labeled categories.

The challenge involves:

- Preprocessing the MRI images to extract meaningful features for classification.
- Using the PyTorch framework, implement a classification CNN model to classify between **Glioma**, **Meningioma**, **Pituitary**, and **No Tumor**.
- Evaluating the model based on its accuracy, precision, and recall, and ensuring it generalizes well to unseen MRI scans.

Participants are expected to:

- Explore and preprocess the MRI dataset to improve model performance.
    - A custom dataset class to load the MRI dataset.
    - Proper transformations for image preprocessing.
    - Datasets and dataloaders for training and validation.
    - Implement a classification algorithm using PyTorch and CNNs for image classification.
    - A suitable loss function for multi-class classification.
    - An optimiser for training the model.
    - Save the trained model and make predictions on unseen MRI images.
- Submit your code notebook and saved model to GitHub for verification.

📊 Evaluation

Submissions will be evaluated using standard classification metrics, including:

- F1 Score

The model's ability to generalise to unseen MRI scans will be a key factor in the evaluation.

📤 Submitting

- Participants must submit a CSV file containing two columns:
    - `image_name`: The name of each test image file (e.g., `img_001.jpg`)
    - `label`: The predicted class label (`0` for glioma, `1` for meningioma, `2` for no tumor, `3` for pituitary)
- The file should be formatted as follows:
  ```csv
  image_name,label
  img_001.jpg,0
  img_002.jpg,3
  img_003.jpg,2
  ```

💾 Code Requirements

All teams must submit their notebooks and saved model to their **team's GitHub repository** for verification.

Start

Closes saturday

📊 Evaluation

Submissions in this competition will be evaluated based on their ability to accurately predict the presence or absence of cardiovascular disease.

🏆 Scoring Metric

The competition will use the **F1 Score** as the primary evaluation metric. The F1 Score is the harmonic mean of precision and recall, providing a balance between false positives and false negatives.

📌 Label Information

Each ground-truth label is a **numeric class value** representing a brain tumor type:

- `0` - Glioma
- `1` - Meningioma
- `2` - No Tumor

- `3` - Pituitary Tumor

🏅 Leaderboard Calculation

- The **Public Leaderboard** score is based on a subset of the test data.
- The **Private Leaderboard** score is computed on the full test dataset and will determine the final rankings at the end of the competition.

The team with the highest **F1 Score** on the **Private Leaderboard** will be declared the winner.

Good luck!

Dataset Description

🧬 What is a Brain Tumor?

A brain tumor is a collection, or mass, of abnormal cells in the brain. Your skull, which encloses the brain, is a very rigid structure. Any growth inside such a restricted space can cause serious problems.

Brain tumors can be:

- **Cancerous (malignant)**
- **Noncancerous (benign)**

As these tumors grow, whether benign or malignant, they increase the pressure inside the skull. This pressure can lead to brain damage and can become life-threatening.

🚨 Importance of the Subject

Early detection and accurate classification of brain tumors is a critical research domain in the field of **medical imaging**. It plays a significant role in:

- Identifying the correct tumor type
- Choosing the most effective treatment method
- Increasing survival rates
- Reducing complications and improving quality of life for patients

Automating this process with AI models helps radiologists make faster, more consistent decisions and may improve diagnosis in regions with limited medical expertise.

❓ What Will You Predict?

You are tasked with developing a model to classify **brain MRI images** into one of the following tumor categories:

- `0` - Glioma
- `1` - Meningioma
- `2` - No Tumor
- `3` - Pituitary Tumor

Your model will predict one of these numeric class labels for each image in the **test set**.

📦 Provided Files

- **`train.csv`** — Metadata for training images. Contains image filenames and corresponding labels (0–3).
- **`test.csv`** — Metadata for test images. Contains image filenames only (labels are hidden).
- **`sample_submission.csv`** — A template for your predictions. Submit a file in this format to be scored.

📁 Image Folders (Directory Structure)

- `train/` — Contains subfolders named `glioma`, `meningioma`, `notumor`, and `pituitary`, each with their respective images.
- `test/` — A flat directory containing **all test images** (mixed, no subfolders).

📊 CSV Columns

`train.csv`

Column	Description
`image_name`	Filename of the training image (e.g., `img001.jpg`)
`label`	Tumor class label (0, 1, 2, or 3)

`test.csv`

Column	Description
`image_name`	Filename of the test image (e.g., `img501.jpg`)

`sample_submission.csv`

Column	Description
`image_name`	Filename of the test image
`label`	Predicted tumor label (to be filled by model)

🗒️ Label Map

Label	Class Name
`0`	Glioma
`1`	Meningioma
`2`	No Tumor
`3`	Pituitary Tumor

🌐 Dataset Sources

- 📦 **Kaggle Dataset**: Brain Tumor MRI Dataset by Masoud Nickparvar
- ☁️ **Google Drive Mirror**: Access here