Brain Tumor Multimodal Image Classification Using CT and MRI Scans

Jonish Bishwakarma
MSDS692 X40: Data Science Practicum
Regis University
jbishwakarma@regis.edu

Abstract

This practicum project presents a deep learning–based multimodal classification system for detecting brain tumors using Computed Tomography (CT) and Magnetic Resonance Imaging (MRI). The project was restructured after encountering feasibility issues with a prior text-image sentiment analysis approach on Reddit. In this revised framework, high-resolution medical images from a publicly available dataset are utilized to develop two parallel convolutional neural networks, which are then fused into a joint classification model. The system is trained and evaluated on real-world multimodal image data, and the goal is to improve tumor detection performance through modality fusion. This work combines computer vision, multimodal learning, and healthcare informatics to support early detection of brain tumors.

I. Introduction / Background

Brain tumors represent a severe neurological condition that often requires early detection and precise diagnosis for effective treatment. Medical imaging plays a crucial role in identifying tumors, with CT scans offering clear visualization of bone structures and MRI scans providing superior soft-tissue contrast. While several classification models exist for single modalities, integrating both CT and MRI data may yield more robust and accurate diagnostic tools.

The original proposal intended to conduct sentiment analysis on Reddit posts combining text and image data. However, due to limited availability of emotional image data on Reddit and scraping limitations, the project faced early setbacks. To ensure data sufficiency and alignment with machine learning objectives, the project was redefined to focus on a well-curated multimodal medical image dataset for brain tumor detection. This shift maintains the spirit of multimodal analysis while adopting a more scientifically feasible and impactful application.

II. Problem Statement

New Core Question:

Can a multimodal deep learning model that fuses CT and MRI scans improve the classification accuracy of brain tumor detection compared to single-modality models?

The primary goal is to explore whether combining features from CT and MRI imaging modalities enhances the ability of neural networks to classify brain images as “Tumor” or “Healthy.” The system is intended to serve as a potential diagnostic aid in clinical settings.

III. Related Work

Recent advances in computer vision have led to the successful application of CNNs in medical image analysis. Studies have shown high accuracy in single-modality tumor classification using either CT [1] or MRI scans [2]. However, multimodal fusion remains underexplored, particularly in accessible public datasets. Other research suggests that modality fusion using attention mechanisms or feature concatenation can significantly enhance classification robustness in biomedical imaging tasks [3].

This project builds on such foundations by applying a multimodal CNN fusion architecture to a real-world medical image dataset and evaluating it against unimodal baselines.

IV. Methodology / Approach

The methodology follows a standard machine learning pipeline with multimodal integration:

- **Data Acquisition:** Use a Kaggle-hosted dataset containing labeled CT and MRI brain tumor scans categorized into “Healthy” and “Tumor.”
- **Preprocessing:** Resize all images to 224x224 pixels and normalize them. Use data augmentation techniques such as rotation, flipping, and brightness adjustment to reduce overfitting.
- **Model Design:**
  - Two independent CNNs extract features from CT and MRI modalities.
  - Feature vectors from both networks are concatenated.
  - A fusion layer classifies the joint representation into binary classes.
- **Training:** Use cross-entropy loss and the Adam optimizer to train the model over multiple epochs with an 80/20 train-test split.
- **Evaluation:** Use accuracy, precision, recall, F1-score, and a confusion matrix to evaluate performance. Also visualize prediction results and training curves.

All code will be version-controlled using Git and developed in a Jupyter/Colab environment using PyTorch and scikit-learn.

V. Data Description

The dataset consists of:

- **CT Scan Images:** Around 4,600 images labeled as either “Tumor” or “Healthy.”
- **MRI Images:** Around 5,000 images labeled similarly.
- Each modality is divided into folders by class and processed using torchvision’s `ImageFolder` API.

The images are in .jpg/.png format and vary in name and source, making this a real-world scenario with diverse data quality and distribution. Proper EDA (Exploratory Data Analysis) was used to assess label balance, modality characteristics, and image integrity.

VI. Expected Outcomes

- A fully functional multimodal classification model capable of distinguishing brain tumors using both CT and MRI images.
- Performance metrics significantly better than unimodal baselines (expected accuracy ¿90%).
- Visual outputs such as:
  - Image-level predictions with true/predicted labels
  - Confusion matrix and classification report
  - Augmented image samples and intensity histograms
- A complete, reproducible codebase with potential for extension to other medical datasets or 3D modalities.

VII. Conclusion

This revised practicum project focuses on a highly practical and clinically relevant challenge—automated brain tumor detection using multimodal imaging. Unlike the initial proposal, which encountered data limitations, this version ensures a robust dataset and a clearly defined scope. Through deep learning and modality fusion, the project aspires to demonstrate that combining CT and MRI scans improves diagnostic accuracy. The model’s results and visualizations may offer valuable insights for both machine learning research and medical imaging applications.

REFERENCES

[1] T. El-Gazzar, K. Amin, and A. Shalaby, "Deep learning for classifying brain tumors using CT images," *IEEE EMBS Conference on Biomedical Engineering and Sciences*, 2019.
[2] J. Cheng et al., "Enhanced performance of brain tumor classification via tumor region augmentation and partition," *PloS one*, vol. 10, no. 10, 2015.
[3] Y. Huang et al., "Fusion of multimodal medical imaging data based on deep learning: A review," *Information Fusion*, vol. 72, pp. 71–85, 2021.