Deep Learning for Brain Tumor Detection Using Magnetic Resonance Imaging (MRI)

Leo Yao yy3959@nyu.edu
Tianyu Zhao tz2263@nyu.edu
Yumo Li yl10192@nyu.edu

Milestone 1

1. Motivation

Medical imaging is a high-impact application of deep learning, offering the potential to automate diagnostics and reduce human error. Disease prediction from scans is both clinically relevant and well supported by open data. For this project, we are choosing to focus on Magnetic Resonance Imaging (MRI) for disease detection tasks. Specifically, this project focuses on enhancing the diagnostic accuracy and interpretability of different deep learning models for brain tumor detection in MRI scans. The core question we are trying to answer is:

> *Can integrating advanced preprocessing pipelines, attention mechanisms, and hyperparameter optimization improve both performance and interpretability across different models such as YOLOv7, U-Net, and R-CNN? (Model choices are subject to change.)*

2. Related Works

We studied several recent, related studies and aim to build upon some state-of-the-art models from their works. The major paper we studied is *Brain Tumor Detection Based on Deep Learning Approaches and Magnetic Resonance Imaging* paper (Abdusalomov et al., 2023), which addresses the critical challenge of brain tumor detection by introducing an enhanced deep learning framework based on the YOLOv7 model to automate and improve the accuracy of tumor diagnosis in MRI scans. Their experimental results demonstrate that the proposed model achieves an impressive 99.5% accuracy and outperforms several state-of-the-art CNN architectures in detecting brain tumors. This paper demonstrates a strong performance by refining YOLOv7 with advanced modules such as CBAM, SPPF+, and BiFPN to achieve such an impressive accuracy. However, it shows limitations in consistently identifying small tumors and requires further validation across diverse datasets and clinical conditions to fully establish its robustness.

In addition, we studied previous works on the models we plan to implement. The *YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors* paper (Wang et al., 2022) introduces YOLOv7, a real-time object detector that establishes a new state-of-the-art by achieving the highest accuracy and fastest inference speeds among existing detectors. The *U-Net: Convolutional Networks for Biomedical Image Segmentation* paper (Ronneberger et al., 2015) introduces a special convolutional network with a symmetric encoder-decoder structure and skip connections that efficiently achieves precise biomedical image segmentation even with limited training data. It can achieve high accuracy even with limited labeled data by leveraging aggressive data augmentation with precise localization. However, its original implementation used fixed hyperparameters, which limits adaptability to heterogeneous MRI datasets. The *Mask R-CNN* paper (He et al., 2018) introduces Mask R-CNN, a simple yet flexible framework that extends Faster R-CNN by adding a parallel branch for predicting high-quality segmentation masks for each detected object, thus unifying object detection and instance segmentation under a single model. Its intuitive design extends Faster R-CNN with a dedicated mask branch and the innovative RoIAlign operation, yielding high-quality instance segmentation and state-of-the-art performance on benchmarks. However, its two-stage framework introduces additional computational overhead, and its reliance on precise region proposals can limit real-time applicability and robustness in scenarios with significant object overlap.

3. Methodology

Our methodology is divided into two primary parts: new data preprocessing technique and comparative model training with hyperparameter tuning.

3.1 Data Preprocessing

We will apply a series of image processing techniques such as blurring, cropping, and resizing. For example, using Gaussian filters to blur the image not only helps reduce noise but also makes the edges more prominent, which is crucial for highlighting anatomical structures in medical images. As discussed in Section 4, our image data does not share the same dimensions. Therefore, cropping and resizing will standardize image dimensions and focus on regions of interest, ensuring that irrelevant background information is minimized.

Furthermore, given the importance of clear boundaries in medical diagnostics, we plan to implement edge enhancement techniques so that the models can detect and delineate critical structures more easily. We may also use data augmentation to simulate various realistic deformations, effectively enlarging our training dataset and improving the model's robustness to variations.

3.2 Model Training

In this part, we will implement several deep learning techniques including both those covered in the course and more advanced ones. Specifically, we will train and compare three models: YOLOv7, U-Net, and Mask R-CNN. While U-Net and Mask R-CNN are convolutional neural networks (CNNs), a core concept from our studies, YOLOv7 incorporates advanced computer vision and object detection techniques that extend beyond the course.

4. Dataset

Category: notumor

Total images: 1595

Sample dimensions (height, width, channels):

```
[[244 206   3]
 [222 227   3]
 [248 208   3]
 [340 339   3]
 [225 225   3]]
```

Category: glioma

Total images: 1321

Sample dimensions (height, width, channels):

```
[[512 512   3]
 [512 512   3]
 [512 512   3]
 [512 512   3]
 [512 512   3]]
```

Category: meningioma

Total images: 1339

Sample dimensions (height, width, channels):

```
[[512 512   3]
 [512 512   3]
 [512 512   3]
 [512 512   3]
 [512 512   3]]
```

Category: pituitary

Total images: 1457

Sample dimensions (height, width, channels):

```
[[512 512   3]
 [512 512   3]
 [512 512   3]
 [512 512   3]
 [512 512   3]]
```

This dataset is one of the two datasets cited in the Brain Tumor Detection Based on Deep Learning Approaches and Magnetic Resonance Imaging original paper. It is a combination of the following three datasets: figshare, SARTAJ dataset, and Br35H. It contains different images of human brain MRI images which are classified into 4 classes: no tumor, glioma, meningioma, and pituitary. The reason we choose this dataset over the other is that the other dataset (SARTAJ) has a problem. The glioma class images are not categorized correctly. This dataset replaced the images in the glioma folder with images on the figshare site instead.

We can see that the dimensions, or the size of the images, in this dataset are different. The "no tumor" category has varied dimensions, while the other three categories have the same size. As a result, data preprocessing techniques such as image resizing is necessary to standardize inputs prior to additional analysis. Additionally, the RGB pixel distribution reveals a significant difference between images with and without tumors. Although the three types of tumor images have very similar overall RGB distributions, there are subtle differences among them that could be leveraged for a more detailed classification.

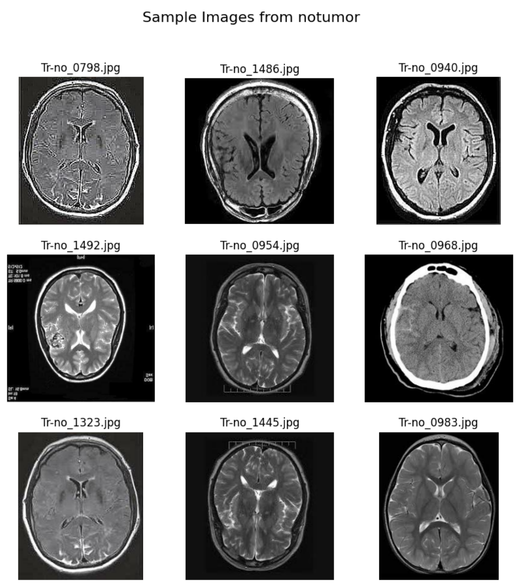
5. Work Plan

Motivation – Yumo Li
Related Works – Tianyu Zhao and Yumo Li
Methodology – Leo Yao and Yumo Li
Dataset – Leo Yao
Work Plan – Tianyu Zhao
Writing the Report - Leo Yao

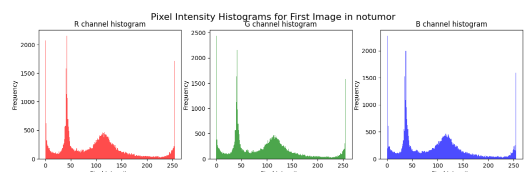
For the following weeks (3-8):
All team members will contribute either through team discussions or splitting the work.

Appendix A.

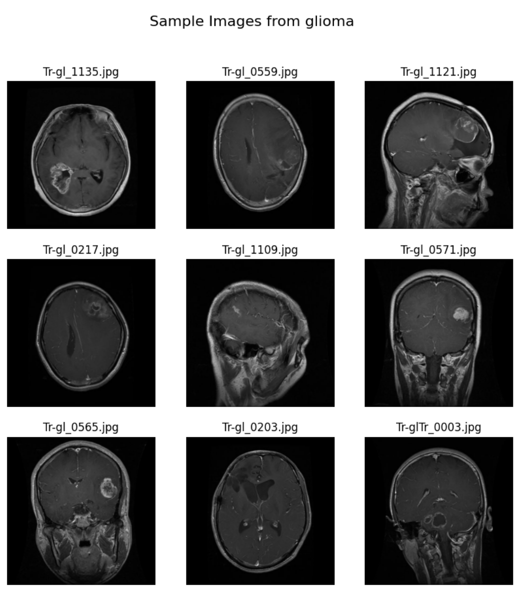
This appendix includes plots generated for examining characteristics of the dataset and exploratory data analysis in Python.



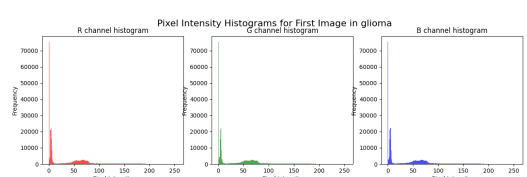
(a) No Tumor Sample Images



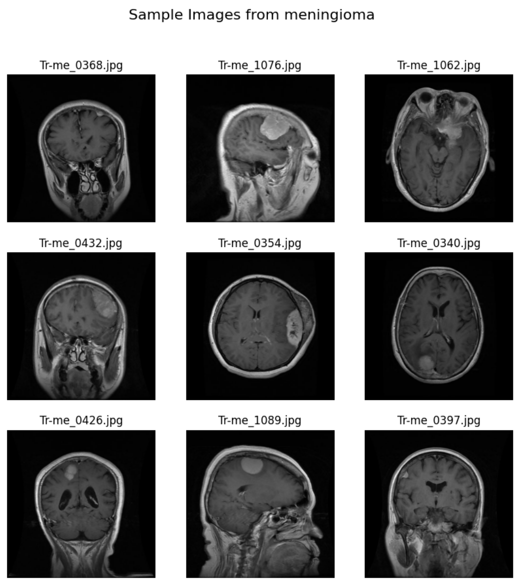
(b) No Tumor RGB Pixel Intensity Distribution



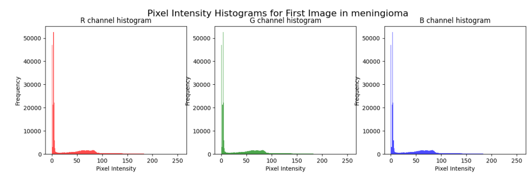
(c) Glioma Sample Images



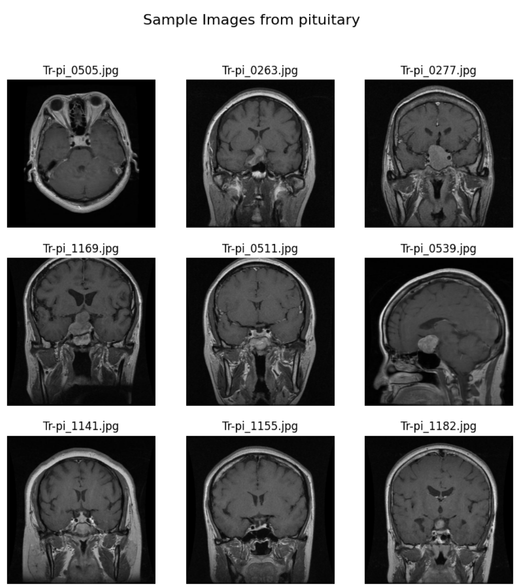
(d) Glioma RGB Pixel Intensity Distribution



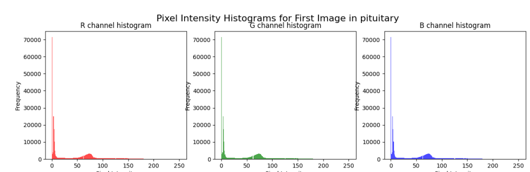
(e) Meningioma Sample Images



(f) Meningioma RGB Pixel Intensity Distribution



(g) Pituitary Sample Images



(h) Pituitary RGB Pixel Intensity Distribution

References

Akmalbek Bobomirzaevich Abdusalomov, Mukhriddin Mukhiddinov, and Taeg Keun Whangbo. Brain tumor detection based on deep learning approaches and magnetic resonance imaging. *Cancers*, 15(16), 2023. ISSN 2072-6694. doi: 10.3390/cancers15164172. URL `https://www.mdpi.com/2072-6694/15/16/4172`.

Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross Girshick. Mask r-cnn, 2018. URL `https://arxiv.org/abs/1703.06870`.

Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. *CoRR*, abs/1505.04597, 2015. URL `http://arxiv.org/abs/1505.04597`.

Chien-Yao Wang, Alexey Bochkovskiy, and Hong-Yuan Mark Liao. Yolov7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors, 2022. URL `https://arxiv.org/abs/2207.02696`.