Classification of Brain Tumor Using MRI Images and Artificial Intelligence

Research Questions

1. Can MRI images be utilized to classify brain tumors using machine learning models?
2. How effective are machine learning models in detecting brain tumors from MRI images?
3. How do various machine learning models compare accuracy in classifying brain tumors?

Objectives

1. To gather and preprocess (resize) MRI images of brain tumors.
2. To perform feature engineering on the image dataset.
3. To train and evaluate multiple deep learning models for brain tumor classification using MRI images.
4. To compare the performance of different models and select the most accurate one for tumor detection.

Background

Brain tumours present a severe medical challenge that necessitates precise and early diagnosis for effective treatment. MRI (Magnetic Resonance Imaging) is a non-invasive technique extensively utilized to detect brain abnormalities. This project aims to develop machine learning models to classify MRI images into categories of tumor and non-tumor. Four classes in the data set include no tumor, meningioma tumor, glioma tumor, and pituitary tumor. The project will use labelled MRI images from a Kaggle dataset to assess the performance of convolutional neural networks and other deep learning architectures in identifying brain tumors. The ultimate objective is to create a generalized machine learning model to aid medical professionals in early diagnosis and treatment planning.

Recent advancements in deep learning have shown considerable promise in medical image analysis, with CNNs being notably effective in feature extraction and accurate predictions. (Huang et al., 2020) have highlighted the efficacy of using DFNN for brain tumor segmentation, demonstrating the model's capability to manage complex medical imaging data. (Dong et al., 2017) explored multimodal brain tumor segmentation techniques, emphasizing the need to address data variability using advanced U-Net based neural networks. The diagnostic process for brain tumors is time-consuming and heavily dependent on radiologists' expertise. The growing volume of data has made traditional methods expensive and inefficient. (Swati et al., 2019) highlights deep learning approaches like the Brain Tumor Classification Model based on CNN, showcasing the potential for precise and efficient diagnosis.

References

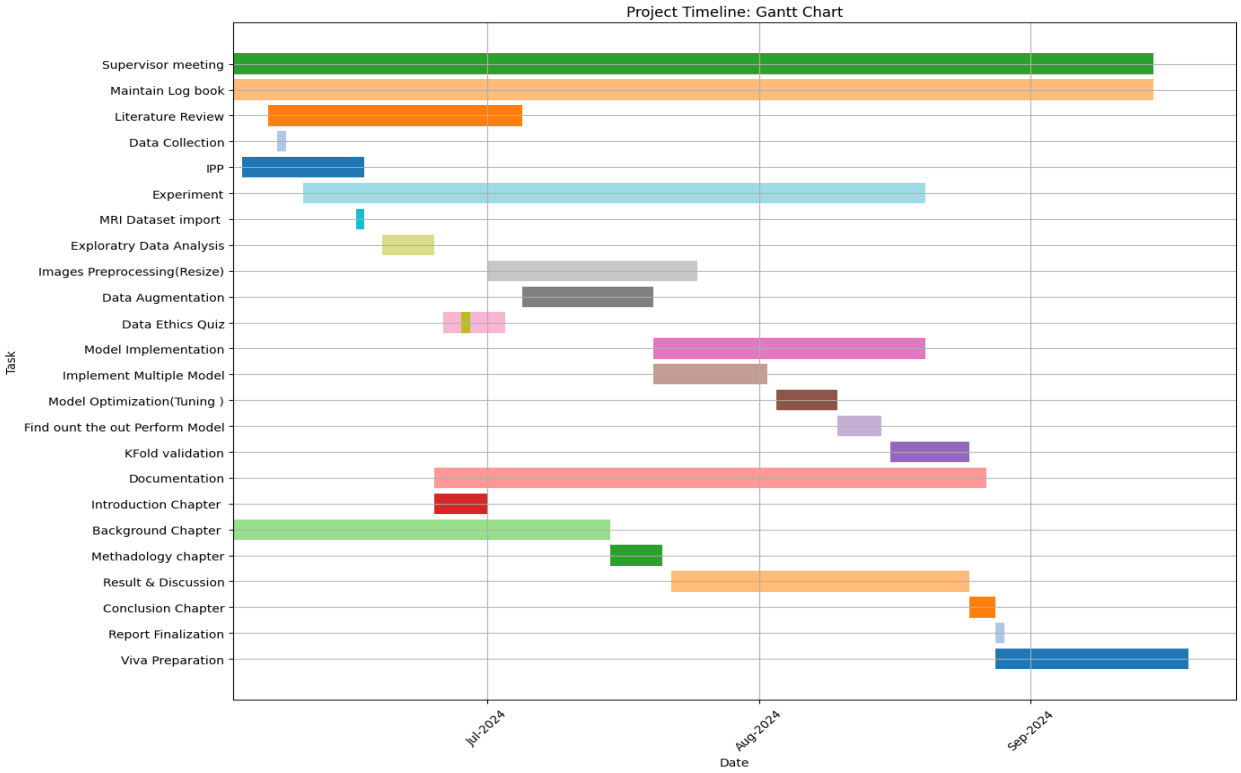
Dong, H., Yang, G., Liu, F., Mo, Y., & Guo, Y. (2017). Automatic Brain Tumor Detection and Segmentation Using U-Net Based Fully Convolutional Networks. *Communications in Computer and Information Science*, *723*, 506–517. https://doi.org/10.1007/978-3-319-60964-5_44

Huang, Z., Xu, H., Su, S., Wang, T., Luo, Y., Zhao, X., Liu, Y., Song, G., & Zhao, Y. (2020). A computer-aided diagnosis system for brain magnetic resonance imaging images using a novel differential feature neural network. *Computers in Biology and Medicine*, *121*, 103818. https://doi.org/10.1016/J.COMPBIOMED.2020.103818

Swati, Z. N. K., Zhao, Q., Kabir, M., Ali, F., Ali, Z., Ahmed, S., & Lu, J. (2019). Brain tumor classification for MR images using transfer learning and fine-tuning. *Computerized Medical Imaging and Graphics*, *75*, 34–46. https://doi.org/10.1016/J.COMPMEDIMAG.2019.05.001

Task Timeline

Task	Description	Start Date	End Date
Supervisor Meeting	Update the supervisor about the project's progress	02/06/2024	19/09/2024
Log Book	Maintain the meeting and lecture records.	02/06/2024	19/09/2024
Literature Review	A comprehensive review of work related to Brain Tumor classification.	03/06/2024	07/06/2024
Data Acquisition	Download the MRI dataset from Kaggle and preprocess the images (resize).	04/06/2024	20/06/2024
IPP	Prepare and submit the Initial Project Plan (IPP).	03/06/2024	17/06/2024
Data Analysis	Perform Data Analysis on Data.	13/06/2024	28/06/2024
Data Ethics Quiz	Prepare and take the data ethics quiz.	26/06/2024	03/07/2024
Model Training	Train Multiple deep-learning models using Keras.	25/07/2024	15/08/2024
Model Optimization	Optimize the layers of the model to get accurate results.	15/08/2024	21/08/2024
Documentation	Write the complete report.	21/08/2024	25/08/2024
Final Project Report	Check the final draft and submit the final Documentation.	25/06/2024	29/08/2024
Viva Preparation	Prepare for oral Test.	29/08/2024	19/09/2024



Overview of the Dataset

The dataset used in this project is the Brain Tumor MRI dataset from Kaggle. The dataset is sourced from various contributors, including Figshare, SARTAJ, and Br35H. The dataset is approximately 500MB and divided into training and testing sets.

Data Collection

Data will be collected from Kaggle, an online platform providing datasets for machine learning and data science projects. The specific dataset, "Brain Tumor MRI Dataset" by Masoud Nickparvar, can be accessed here. This dataset includes images in various formats, which will require preprocessing steps such as normalization and resizing to ensure compatibility with machine learning models.

Summary of Data

The dataset consists of MRI images in JPEG format, divided into four categories: no tumor, meningioma tumor, glioma tumor, and pituitary tumor. The dataset contains a total of 7023 MRI images, which are further divided into training and testing sets.

Document Control

GitHub will be employed for version control. The repository address for the project is at this link. All code and data files will be systematically organized with clear naming conventions like model_training.py, data_preprocessing.py, and final_report. Each commit will include descriptive messages to track changes and updates throughout the project.

Metadata

A ReadMe file will be included in the GitHub repository. This file will provide an overview of the project, instructions on how to run the code, dependencies, and a brief description of each script. It will be regularly updated to reflect any changes or additions to the project.

Ethical Considerations

GDPR Compliance: The data from Kaggle adheres to GDPR requirements by anonymizing patient information. MRI images do not contain personally identifiable information, ensuring compliance with GDPR.

UH Ethical Policies: The project conforms to UH ethical policies by using publicly available anonymized datasets from Kaggle for research.

Data Usage Permission: The data is publicly available on Kaggle under a license that permits its use for academic and research purposes.

Ethical Data Collection: The data was collected ethically by the original researchers and made available on Kaggle, ensuring ethical standards were met during data collection and collation.