**Brain tumor prediction using Convolutional Neural Networks (CNN)**

**1.Project Overview:**

Brian tumors, a major health concern around the world, are caused by the uncontrolled development of cells within brain tissue. Brain tumors range from benign(non-cancerous) to malignant(cancerous). Detecting brain tumors is a critical medical approach that involves identifying the presence and type of tumors in the brain. Traditionally, this process has been carried out through various medical imaging techniques such as MRI (Magnetic Resonance Imaging) and CT (Computed Tomography) scans (Cheng et al. 2015). However, deep learning, a subset of artificial intelligence, has emerged as a promising technology for automating and enhancing the accuracy of brain tumor detection. Traditional methods rely on manual analysis of medical images, which is time-consuming and prone to error. Deep learning offers a creative solution to this challenge. Deep learning is a good choice for this task to analyze large and complete datasets, detect subtle patterns, and make predictions based on these patterns. Convolutional Neural Network (CNN) is a type of deep learning algorithm that can learn to recognize image patterns. By training a CNN on a large dataset of medical images, we can teach it to detect brain tumors with perfect accuracy (Litjens et al. 2017).

**Research Questions:**

* How can a convolutional neural network (CNN) model be developed to accurately detect brain tumor from MRI images?
* How does the performance of the proposed CNN model compare to current state-of-the-art methods for brain tumor detection?

**Objectives:**

* Development: Design and train a CNN model capable of accurately predicting brain tumour from MRI scans.
* Evaluation: Compare the performance of the developed model with existing state-of-the-art methods.

**Scope:**

This project will cover the CNN model's development, testing, evaluation, and validation. The dataset will include annotated MRI scans from publicly available sources.

**2.Project Plan:**

**Data Collection:**

The study obtained a significant number of MRI brain scans from public databases and classified them as pituitary tumors, gliomas, meningiomas, and noncancerous tumors. Convolutional Neural Network (CNN) models will be trained and tested on selected images per category to effectively utilize resources and accurately represent each group. Grayscale images will be processed to highlight important anatomical features, with OpenCV handling the loading and filtering. Bilateral filtering will be used to reduce the noise, while cv2.COLORMAP\_BONE and added pseudo-colours to improve detail recognition.

**Model Development:**

The CNN model will be built using the "Brain Tumor MRI Dataset," which includes ‘testing and training’ folders with 1131 and 5712 data images respectively. The images were divided into 4 groups ‘glioma, meningioma, notumor, and pituitary. A Convolutional Neural Network (CNN) will be created using Keras and TensorFlow. Normalization and data augmentation will be key preprocessing processes to address dataset variability (Cheng et al. 2015). The intention is to use a modified VGG16 architecture because of its depth and demonstrated effectiveness in medical imaging tasks. To reduce overfitting, training will use a Stochastic Gradient Descent (SGD) optimizer with cross-validation to minimize the given loss function by iteratively updating model parameters. This strategy intends to improve the model's accuracy and robustness in detecting brain tumors.

**Model Evaluation and Validation:**

The model will be tested using the "Brain Tumor MRI Dataset" and the cross-validation technique will be applied to confirm its robustness. To minimize overfitting, the dataset will be organized into training, validation, and test sets (Cheng et al. 2015). The model's effectiveness will be assessed using performance indicators such as accuracy, sensitivity, specificity, and AUC-ROC. I will also compare the model's performance against existing pre-trained methods to identify areas for improvement. Statistical analysis will be conducted to validate the findings, ensuring that the model is both reliable and accurate in predicting brain tumors.

**Documentation and Reporting:**

I plan to properly document each stage of my project, from data preprocessing to model building and evaluation. Detailed records of techniques, hyperparameters, and outcomes will be kept ensuring consistency. I'll integrate these results into a detailed final report that highlights important accomplishments and areas for improvement. In addition, I will create a presentation to effectively communicate the project's results. This extensive documentation and reporting will serve as a clear and useful reference for future research and practical applications.

**Project Timeline:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Start Date** | **Task** | **Milestone** | **No.Of Days** |
| **22/10/24** | Data Collection | Confirmation on Dataset | 10 |
| **1/11/24** | Literature Review | Researching past and present research on CNN brain tumor prediction. | 10 |
| **1/11/24** | PDM plan preparation | Preparing the plan | 5 |
| **5/11/24** | PDM plan submission |  | 1 |
| **7/11/24** | Data Analysis and Pre-processing | Analysis using Keras and image data generator. Filtering and scaling the data. | 5 |
| **11/11/24** | Model Training, Feature Engineering | Training models using CNN pipelines | 7 |
| **18/11/24** | Quiz preparation |  | 3 |
| **21/11/24** | Quiz exam |  | 1 |
| **22/11/24** | Results Analysis and Evaluation | Evaluation based on training | 3 |
| **26/11/24** | Final Results Analysis | Results analysis and visualizations | 5 |
| **1/12/24** | Report Writing | First draft of report completed | 20 |
| **21/12/24** | Changes for report after review | Check with supervisor for any changes to be made | 10 |
| **06/01/25** | Final Review and Submission | Report submitted |  |

**3. Data Management Plan:**

* **Data Collection**: The "Brain Tumor MRI (<https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset>) " dataset on Kaggle is critical for brain tumor diagnosis research. It contains a variety of MRI images from healthy people as well as tumor patients.
* **Metadata:** The dataset is of Version 1 (158.6 MB) of 7022 high-resolution photos which are grouped into 1,131 testing and 5,712 training images and are categorized as gliomas, meningiomas, pituitary tumors, and non-tumor scans.
* **Storage:** A dedicated project folder created in my computer in which the dataset is stored. The pre-processed data will be separated to ensure proper organization.
* **Code:** The project code, which includes data processing, model training, and evaluation routines, will be kept on a GitHub repository. This will ensure that the code is well-organized and easy to track. A README file will be included, with an overview of the project, setup instructions, and code explanations.
* **Backup:** Cloud storage(Google Drive) and GitHub will regularly backup data and code. The GitHub platform is the primary source and Google drive will be used as backup source.
* **Github link :** <https://github.com/vineeth7858/Brain-Tumour-Prediction-Project.git>
* **Version Control:** GitHub will serve as the version control system to monitor code development. Consistent commits will document updates within the codebase, facilitating collaboration and enabling restoration of prior versions when needed. Each iteration will be thoroughly documented to clarify the modifications implemented.

**Ethical Considerations:**

This study highlights ethical aspects, including patient data confidentiality, GDPR compliance, and the appropriate use of public MRI datasets. The findings will be provided in a transparent manner to aid repeatability and clinical application.

The study does not involve interaction with other people or the collecting of new data, thus there is no need for ethical approval from the University of Hertfordshire Ethics Committee. Furthermore, all code will be tested responsibly, assuring transparency, repeatability, and adherence to data usage restrictions. The findings will be given without any potential harm or bias, with an emphasis on impartial social analysis.

**Reference:**

Cheng, J., Huang, W., Cao, S., Yang, R., Yang, W., Yun, Z., Wang, Z. and Feng, Q. (2015). Enhanced Performance of Brain Tumor Classification via Tumor Region Augmentation and Partition. PLOS ONE, 10(10), p.e0140381. doi: <https://doi.org/10.1371/journal.pone.0140381>.

Litjens, G., Kooi, T., Bejnordi, B.E., Setio, A.A.A., Ciompi, F., Ghafoorian, M., van der Laak, J.A.W.M., van Ginneken, B. and Sánchez, C.I. (2017). A Survey on Deep Learning in Medical Image Analysis. *Medical Image Analysis*, 42(1), pp.60–88. doi: <https://doi.org/10.1016/j.media.2017.07.005>.

Pereira, S., Pinto, A., Alves, V. and Silva, C.A. (2016). Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images. *IEEE Transactions on Medical Imaging*, 35(5), pp.1240–1251. doi: <https://doi.org/10.1109/tmi.2016.2538465>.