BRAIN TUMOR DETECTION USING MACHINE LEARNING

Md. Mehedi Hasan Polas
Dept. of Computer Science & Engineering
American International University, Bangladesh
22-46566-1@student.aiub.edu

Tridib Sarkar
Dept. of Computer Science & Engineering
American International University, Bangladesh
22-46444-1@student.aiub.edu

Sajin Mahmud Arpon, *Member, IEEE*,
Dept. of Computer Science & Engineering
American International University, Bangladesh
22-46629-1@student.aiub.edu

Talha Hossain Sifat
Dept. of Computer Science & Engineering
American International University, Bangladesh
22-46344-1@student.aiub.edu

***Abstract*—Brain tumor detection is a critical task in medical diagnostics that requires early and accurate identification to improve patient outcomes. This research presents a machine learning-based approach utilizing a Convolutional Neural Network (CNN) to classify brain tumors from MRI images efficiently. The proposed model is trained and evaluated on a publicly available dataset containing 253 brain MRI images, divided into tumor and non-tumor classes. The CNN architecture incorporates convolutional and max-pooling layers for feature extraction and uses a sigmoid activation function for binary classification. Experimental results demonstrate an accuracy of 89% and an F1-score of 0.89, indicating robust performance in distinguishing tumor presence. The model shows promise for deployment in clinical and resource-limited settings to support early diagnosis. Future work will focus on expanding the dataset, incorporating transfer learning, and integrating domain-specific features to enhance classification accuracy.**

***Index Terms*—Brain Tumor Detection, Machine Learning, Convolutional Neural Network, Medical Image Analysis, MRI Classification**

I. Introduction

A Brain tumor is an abnormal growth of cells inside the brain that can be either non-cancerous or cancerous. These tumors disrupt normal brain function by pressing against surrounding tissues, leading to symptoms like headaches, memory loss, seizures, and behavioral changes. Early detection through advanced diagnostic methods is crucial for effective treatment and improved patient survival rates, making this a critical area of research in modern healthcare. [1]

Several researchers have explored brain tumor detection using advanced machine learning and deep learning approaches. Patro et al. [2] presented a powerful ensemble model that combines multiple deep learning techniques to improve classification accuracy. Their system classified tumors with over 96% accuracy, showing how model fusion can enhance performance by capturing different feature representations. While effective, ensemble models often require significant computational resources, which may limit their use in low-resource settings. In contrast, Hassan and Boulila [3] developed a hybrid method that integrates fuzzy thresholding with deep learning. Their technique first segments the MRI images to separate tumor regions using fuzzy logic, which helps reduce irrelevant information. Then, deep learning is applied for classification, resulting in improved detection speed and accuracy. This fusion of traditional image processing with deep models adds both interpretability and efficiency to the system. Thokar et al. [4] emphasized the importance of precise image segmentation before classification. They used deep learning-based segmentation to identify anomalous regions in MRI images. This preprocessing step enhanced the clarity of tumor boundaries and helped improve detection accuracy. Their results highlight that deep learning models benefit significantly from clean and segmented input data. On the other hand, Anwar et al. [5] explored the power of transfer learning by fine-tuning pretrained models on MRI datasets. Their approach increased detection precision without the need for massive computational effort, making it more practical for real-time clinical use. This also reduces training time and enhances performance on smaller datasets, making it ideal for healthcare environments with limited data. Tariq et al. [6] introduced a system that uses Vision Transformers (ViT) and EfficientNetV2 for multi-class brain tumor detection. These models, although more complex, are highly effective in recognizing patterns in medical images due to their global attention mechanisms. They achieved strong results in classifying glioma, meningioma, pituitary tumors, and healthy scans. Their work shows the potential of transformer-based models for medical imaging tasks, although they also require advanced hardware. Finally, Chakrabarty [7] contributed by providing a publicly available MRI brain tumor dataset on Kaggle, which contains labeled images categorized into different tumor types. This dataset is frequently used in training and validating models, and it plays a vital role in benchmarking performance and enabling reproducibility in brain tumor research.

This research aims to develop a simple and fast machine learning system to detect and classify brain tumors from MRI images, providing quick and accurate results to assist doctors. It is designed for use in hospitals, clinics, rural health centers, and mobile or cloud platforms, improving patient outcomes through early detection and better treatment planning.

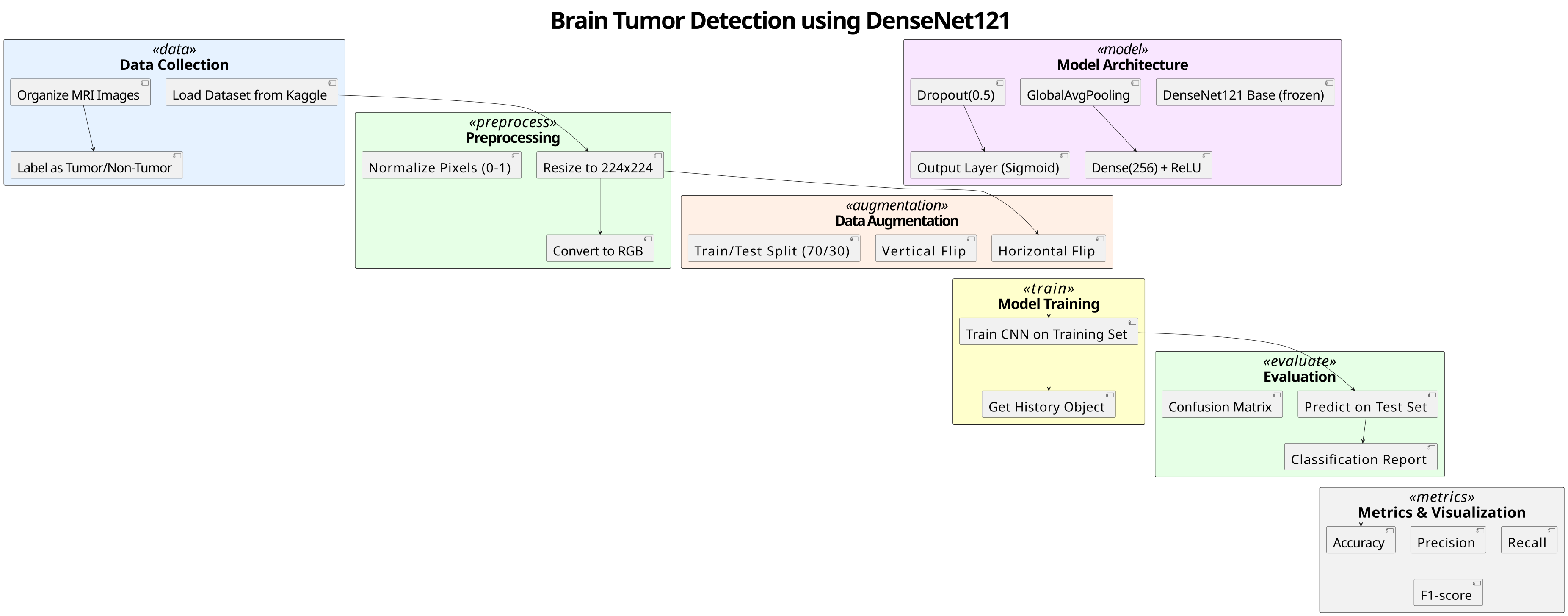

Fig. 1: Workflow of Brain Tumor Detection using DenseNet121

II. Methodology

The project methodology involves collecting and preprocessing brain tumor image datasets. After preprocessing the images the next step is to build a Convolutional Neural Network model using the Keras Sequential API. The model includes convolutional layers with ReLU activation max pooling layers and fully connected layers with a softmax activation function for multi-class classification. The model is trained using the preprocessed training dataset and validated on a separate testing dataset to check its general performance. Evaluation metrics such as accuracy loss precision recall and F1-score are used to measure the model performance. Visualization tools are used to better understand the learning process and output. The model that shows the highest accuracy and best performance is selected as the final model for brain tumor detection. This methodology uses machine learning to create a strong and reliable model that is able to help in the early detection of brain tumors and support doctors in making faster decisions.

A. DATA COLLECTION PROCEDURE

For this study the brain tumor image dataset was collected from the Kaggle repository [7]. The dataset contains a total of 253 MRI images of the brain. Out of these 155 images show the presence of brain tumor and 98 images show no tumor. The dataset includes MRI images in grayscale format which are easy to process. The images are clearly labeled into two categories such as yes for brain tumor detected and no for brain tumor not detected. All images were carefully selected and organized to make sure the data is clean and suitable for training and testing machine learning models. The dataset helps to build an effective model for detecting brain tumors using image classification. The collection process ensures the data is balanced and supports reliable performance of the detection model.

B. DATA VALIDATION PROCEDURE

The data validation procedure was used to check the reliability of the brain tumor image dataset. The process included checking the completeness and correctness of all images. It also checked if the data was organized properly and labeled correctly into yes and no categories. Any missing or unclear images were removed to make the dataset clean and useful. The dataset was also checked to make sure it followed ethical standards and had no sensitive or private information. These steps helped to make sure the dataset was ready and safe to use for training and testing in brain tumor detection.

C. DATA PREPROCESSING AND NORMALIZATION

In the data preprocessing and normalization stage the brain tumor images were loaded and resized to a fixed size of 224x224 pixels. The images were converted to RGB format to keep the color information important for image classification. Each image was normalized by dividing the pixel values by 255 so that the values were between 0 and 1. This helped to improve the model training and make the learning process stable. The labels yes and no were converted into numerical form using a label encoding method. This step helped the model to understand the classes. The images and labels were then changed into NumPy arrays for faster processing. This preprocessing and normalization process helped to prepare the data in the correct format so that the machine learning model was able to train effectively for brain tumor detection.

D. FEATURE EXTRACTION

The feature extraction process used deep learning techniques to help the model learn and identify important patterns from the brain tumor images during training. The convolutional layers of the Convolutional Neural Network extracted features like edges, textures and shapes that are commonly found in brain tumor regions. These layers helped the model to understand different levels of detail in the images. Max pooling layers were used to reduce the size of the feature maps while keeping the most important information. This made the training process faster and less complex. After that the features were flattened into a single vector using dense layers. This allowed the model to use the extracted features for final classification. This method helped the model to focus on key characteristics that showed the presence of a brain tumor and improved the accuracy of the classification process.

E. CLASSIFICATION ALGORITHMS

In this project a Convolutional Neural Network model was used for the classification task of brain tumor detection. CNN is a type of deep learning model that is designed to work well with image data. The CNN architecture included several convolutional layers followed by max pooling layers that helped to reduce the image size while keeping the most important features. These layers allowed the model to automatically learn and extract features from brain MRI images such as edges shapes and textures. After feature extraction the data was flattened and passed through dense layers which helped in classifying the images based on the features learned. The final layer used a sigmoid activation function which gave the probability of brain tumor presence. The model was trained using the Adam optimizer and binary cross entropy loss function to achieve accurate and efficient classification results.

F. DATA ANALYSIS TECHNIQUES

Data analysis techniques used in this project focused on checking the performance of the trained Convolutional Neural Network model for brain tumor detection. The analysis included accuracy loss precision recall and F1-score to measure how well the model was able to detect brain tumors from MRI images. A confusion matrix was also created to show the classification results by displaying true positive true negative false positive and false negative values. These metrics and visual tools gave clear understanding of how correctly the model was able to classify brain tumor and non-tumor cases. This analysis helped to confirm that the model was reliable and strong for use in real medical environments.

$$\text{Accuracy} = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}}$$

$$F_1 \text{ score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

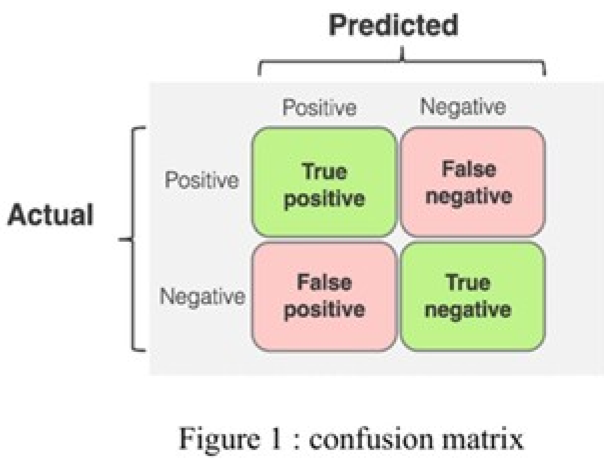

Fig. 2: Confusion Matrix

III. RESULTS AND DISCUSSION

The Convolutional Neural Network CNN model was used for brain tumor detection and classification. The dataset was divided into 70 percent for training and 30 percent for testing. The CNN model was trained on the training data and then tested on the validation data to evaluate its performance. The model showed good results with high accuracy and F1-score. These metrics indicate that the CNN was able to correctly identify brain tumor images and distinguish them from non-tumor images. Overall the CNN model demonstrated reliable performance in detecting brain tumors using MRI images.

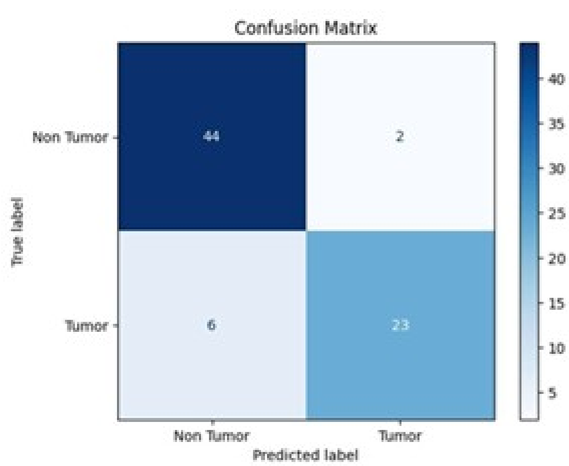

Fig. 3: Confusion Matrix Analysis of CNN Model

A. CONFUSION MATRIX ANALYSIS

The confusion matrix shows the performance of a brain tumor detection system. It correctly predicted 44 non tumor cases and 23 tumor cases. It wrongly predicted 2 non tumor cases as tumor and 6 tumor cases as non tumor. The system is able to detect non tumor cases with high accuracy.

B. RESULTS VALIDATION BY GRAPHICAL REPRESENTATION

Fig. 4: Accuracy graph of training and testing

The graph shows the training and validation accuracy over 10 epochs. Training accuracy improves steadily and reaches above 90 percent by the final epoch. Validation accuracy increases quickly in the beginning and stays around 88 to 89 percent for most of the training. However it drops slightly after epoch 6 while training accuracy continues to rise. This indicates that the model may start overfitting after some epochs. Overall the model shows strong performance on both training and validation data.

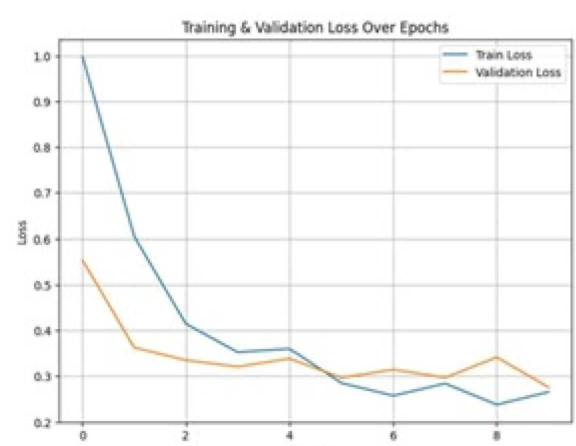

Fig. 5: Accuracy graph of training and testing

The graph shows the training and validation loss over 10 epochs. Training loss decreases steadily from 1.0 to around 0.26. Validation loss also decreases at first and remains low between 0.28 and 0.36. After epoch 4 validation loss shows small fluctuations while training loss keeps improving. This suggests that the model is learning well but may begin to overfit slightly. Overall both losses are low which indicates good model performance.

IV. CONCLUSION AND FUTURE RECOMMENDATIONS

In conclusion the Convolutional Neural Network CNN model showed strong performance in detecting brain tumors using MRI images. The model achieved an accuracy of 89 percent and an F1 score of 0.89. These results prove that CNN is able to classify brain tumor and non-tumor images in a reliable way. The model learned key patterns from the data and provided consistent results during testing. For future work the dataset should be expanded to include more images with different types of brain tumors. Using advanced methods like transfer learning with pre-trained models will be helpful to improve performance. Including domain-based knowledge such as tumor size and location may also increase accuracy. These steps will support the use of machine learning models in real medical systems and help doctors detect brain tumors at an early stage.

REFERENCES

[1] Paromita Kundu, Chandana Mohanty, and Sanjeeb Sahoo. Retraction. retraction notice to "antiglioma activity of curcumin-loaded lipid nanoparticles and its enhanced bioavailability in brain tissue for effective glioblastoma therapy" [acta biomaterialia 8 (2012) 2670-2687]. *Acta Biomater*, 9:7074, 06 2013.

[2] S Gopal Krishna Patro, Nikhil Govil, Surabhi Saxena, Brojo Kishore Mishra, Abu Taha Zamani, Achraf Ben Miled, Nikhat Parveen, Hashim Elshafie, and Mosab Hamdan. Brain tumor classification using an ensemble of deep learning techniques. *IEEE Access*, 12:162094–162106, 2024.

[3] Nashaat Hassan and Wadii Boulila. Efficient approach for brain tumor detection and classification using fuzzy thresholding and deep learning algorithms. *IEEE Access*, PP:1–1, 01 2025.

[4] Bibat Thokar, Binod Sapkota, Babu R. Dawadi, and Shashidhar R. Joshi. Medical image segmentation for anomaly detection using deep learning techniques. *IEEE Access*, 12:185460–185481, 2024.

[5] Raja Anwar, Mohammad Abrar, and Faizan Ullah. Transformative transfer learning for mri brain tumor precision: Innovative insights. *IEEE Access*, PP:1–1, 01 2025.

[6] Anees Tariq, Muhammad Munwar Iqbal, Muhammad Javed Iqbal, and and Iftikhar Ahmad. Transforming brain tumor detection empowering multi-class classification with vision transformers and efficientnetv2. *IEEE Access*, 13:63857–63876, 2025.

[7] Navoneel Chakrabarty. Brain mri images for brain tumor detection. https://www.kaggle.com/datasets/navoneel/brain-mri-images-for-brain-tumor-detection, 2023. Accessed: 2025-05-24.