Using Hybrid Deep Learning to Classify and Detect Brain Tumors

Vikas Kumar
CS24DP003
Computer Science and Engineering

***Abstract*—Brain tumors, whether benign or malignant, are significant threats in human health, requiring that they be detected early and very accurately to be treated perfectly. This paper presents a deep learning hybrid approach using multi-convolutional neural architectures for improving the precision and accuracy of brain tumor's detection and classification from MRI scans. The proposed model, through the use of a dataset comprising 7,023 MRI images from the datasets Figshare, SARTAJ, and Br35H, classifies tumors into one of four categories: glioma, meningioma, pituitary tumor, and no tumor. This study has examined the hybrid model benefits from the strengths of diverse CNN architectures for multi-task classification, allowing the automatic detection of tumors, their classification based on grade, and segmentation-based location identification. The model thus had a good accuracy level at 88.76% on the training set and 86.19% on the validation set, proving its efficiency in offering a reliable diagnostic application for medical images. This illustrates that hybrid deep learning methodology could further help in the future advancement of healthcare diagnostics.**

Keywords: Brain tumor detection, MRI image classification, Hybrid deep learning, CNN, Medical imaging diagnostics, Tumor segmentation

I. Introduction

Brain tumors are the abnormal growths of cells in the brain and can either be benign, which means non-cancerous, or malignant, referring to the cancerous type. They can pose serious health threats because they can grow at a high rate and easily damage surrounding tissues, potentially spreading to other parts of the brain or body. Accurate early detection and classification of brain tumors are the keys to improving treatment outcomes and saving lives. The World Health Organization (WHO) considers brain tumors to be among the most common causes of death in the world. In the next years, there will be a rise in cases.

Medical imaging, in particular, Magnetic Resonance Imaging (MRI), plays a significant role in diagnosing brain tumors. However, analyzing these images manually is time-consuming, error-prone, and labor-intensive, especially because of the large volume of images produced in daily clinical practice. Traditional machine learning methods are applied to the detection and classification of tumors. However, they often involve handcrafted features, which can be inefficient and less accurate. In contrast, deep learning techniques such as Convolutional Neural Networks have proven effective in automatically extracting features, and better accuracy is seen in image-based medical diagnostics.

A deep learning-based model, called DeepTumorNet, will be designed for the classification of tumor detection from MRI images. We aim to work with three kinds of tumors, including glioma, meningioma, and pituitary, in addition to identifying no tumor cases. A hybrid approach that integrates multiple CNN architectures is envisioned in this hybrid model, where we focus on enhancing precision and accuracy in the classification and detection of tumors. It not only serves to enhance diagnostic capabilities but is also one of the demonstration projects of how artificial intelligence will revolutionize healthcare.

II. Background and Motivation

Brain tumors represent one of the most threatening health risks worldwide, regardless of their nature, being benign or malignant. Brain tumors are a great cause of mortality, with thousands of new cases recorded annually worldwide, according to the WHO. These tumors can be primary when they grow within the brain or are secondary when they grow within the brain and spread from other parts of the body. Gliomas, meningiomas, and pituitary tumors are some of the most common primary brain tumors, each with distinct features and treatment implications. The sooner and more accurately such a tumor is detected, the better the patient's chances of survival. Still, manual analysis of MRI scans, which have become one of the mainstays of diagnosis, remains an extremely complex, time-consuming, and error-prone process.

Medical Imaging and its Role in Brain Tumor Diagnosis Magnetic Resonance Imaging (MRI) is extensively used in brain tumor detection since it generates high-contrast grayscale images of soft tissues. Also, MRI is a non-ionizing radiological tool, so it does not pose significant risk in clinical diagnosis. The critical information regarding size, location, and type of tumors obtained through the use of MRI images, but the human interpretation process associated with the manual assessment is prone to human errors, and thus it becomes quite impractical in the era of vast clinical data generated each day.

Evolution of Techniques in Tumor Detection Traditionally, machine learning techniques like Support Vector Machines (SVMs), k-Nearest Neighbors (kNN), and Decision Trees have been used for brain tumor classification. Although effective, these methods rely on handcrafted features, which are both labor-intensive and prone to inaccuracies. Deep learning, particularly Convolutional Neural Networks (CNNs), has revolutionized the landscape of medical image analysis. CNNs automate feature extraction, thus making tumor classification more accurate and scalable. Successful models such as AlexNet, VGG16, ResNet50, and EfficientNets have been pre-trained on this domain by utilizing transfer learning to address the limited labeled data challenge.

The above approaches, despite the mentioned successes, have a number of challenges:

Data Dependency: High-performing deep learning models require a lot of labeled datasets, which are usually scarce in medical imaging. High Computational Costs: The training of complex models necessitates a lot of computational resources, including GPUs and large memory capacities. Limited Task Integration: Most models are designed for single tasks such as classification or segmentation, and they cannot perform multi-task operations like tumor detection, grading, and localization simultaneously. Motivation for the Proposed Hybrid Model The limitations above point to a need for a robust model that serves multiple purposes - high in accuracy and practicably usable. It presents the hybrid deep learning model that integrates CNN architectures to fully exploit all their strengths at once, supporting multi-task classification from tumor detection down to grade classification or even precise location through segmentation.

The primary motivations for taking up this project are that

Advancing Diagnostic Accuracy: The hybrid CNN architecture-based proposed model is able to achieve high accuracy, with training and validation accuracies of 88.76% and 86.19%, respectively. This can directly improve the reliability of brain tumor diagnosis. Multi-task Requirements: The model is able to detect the presence of tumors, classify them as either glioma, meningioma, or pituitary, and identify the exact location of the tumor, thus providing a holistic solution for diagnosis. Improving Accessibility: The project integrates a Streamlit-based user interface, ensuring that healthcare professionals can easily use the model without requiring technical expertise in deep learning. Manual Challenges Mitigated: Automation of feature extraction and classification reduces reliance on manual processes, minimizing human error and enabling efficient handling of large datasets. Bridging Technological Gaps The hybrid model bridges gaps found in current research, focusing on the integration of multiple diagnostic tasks into one framework, especially when this integration involves combining the best features of advanced CNNs with segmentation capabilities. Real-World Relevance With such growing demand for computer-assisted diagnostic tools in the field of medical imaging, this project aligns all together in enhancing their model's performance on the 7,023 dataset MRI images, which reflects the potential of revolutionizing the clinical practices by providing efficient diagnostic help for doctors. This invention does not only improve patient care quality but also tells us about the transformative character of artificial intelligence in health sciences.

This project aims to contribute meaningfully to the field of medical diagnostics in addressing gaps in existing methods and combining state-of-the-art deep learning techniques to make a foundation for future advancements in brain tumor detection and classification.

III. Related Work

The field of brain tumor classification has witnessed significant advancements, particularly with the application of machine learning (ML) and deep learning (DL) techniques. Earlier methods primarily relied on traditional ML algorithms like Support Vector Machines (SVM), k-Nearest Neighbors (kNN), and Principal Component Analysis (PCA) for feature extraction and classification. However, these approaches depend heavily on handcrafted features, which are not only time-consuming to engineer but also less robust, especially for large and complex datasets [1], [2].

In recent years, DL-based models have demonstrated unparalleled capabilities in automatic feature extraction and classification. Convolutional Neural Networks (CNNs) have emerged as the most widely used architecture in brain tumor classification tasks, owing to their ability to automatically learn hierarchical features from raw data [1], [2].

A. Key Contributions in Literature

DeepTumorNet Model: The DeepTumorNet model utilizes GoogLeNet as a base architecture but extends its capabilities by replacing the last five layers with 15 new layers, incorporating leaky ReLU activation to address the dying ReLU problem. This model achieves remarkable accuracy of 99.67% on the CE-MRI dataset, outperforming traditional and other DL models like ResNet, AlexNet, and MobileNet [1].

EfficientNets for Transfer Learning: EfficientNet-based transfer learning methods have proven effective in classifying glioma, meningioma, and pituitary tumors. By fine-tuning pre-trained EfficientNet models (B0 to B4) with a CE-MRI dataset, these approaches achieved a classification accuracy of 99.06% along with excellent precision, recall, and F1-scores. Grad-CAM visualization further enhanced interpretability by highlighting tumor regions [2].

Hybrid Models with Feature Fusion: Some studies propose hybrid models combining handcrafted and deep features. For instance, one approach fused features extracted from a discrete wavelet transform with deep CNN features for better classification accuracy. These models utilize a combination of feature extraction, data augmentation, and transfer learning [1], [2]. **Pre-trained CNNs:** Studies have shown that fine-tuned pre-trained networks like AlexNet, ResNet50, and VGG16 achieve high accuracies (up to 98.93%) for MRI-based brain tumor classification. These models leverage transfer learning to overcome the challenge of limited medical imaging datasets [2].

B. Performance Comparison

While traditional ML methods often achieve moderate accuracy and require intensive manual feature engineering, DL-based approaches consistently outperform them due to their ability to learn complex features. Hybrid models and transfer learning techniques further enhance performance, with accuracy improvements ranging from 95% to 99.67% across various datasets and architectures [1], [2].

IV. Mathematical Expressions

A. Convolution Operation in CNNs

The convolution operation extracts features from input images by applying a kernel or filter. It is mathematically expressed as:

$$S(i,j) = (I * K)(i,j) = \sum_m \sum_n I(i+m, j+n) \cdot K(m,n)$$

Where:

- $I(i,j)$: Input image matrix
- $K(m,n)$: Kernel or filter matrix
- $S(i,j)$: Output feature map

B. Activation Functions

ReLU (Rectified Linear Unit):

$$f(x) = \begin{cases} 0, & \text{if } x < 0 \\ x, & \text{if } x \geq 0 \end{cases}$$

Leaky ReLU:

$$f(x) = \begin{cases} x, & \text{if } x \geq 0 \\ \alpha x, & \text{if } x < 0 \end{cases}$$

Softmax for Classification:

$$P(y = j | x, W, b) = \frac{\exp(W_j^T x + b_j)}{\sum_{k=1}^{K} \exp(W_k^T x + b_k)}$$

Where W and b are the weights and biases, and K is the number of classes.

C. Pooling Operation

Pooling reduces the spatial dimensions of feature maps. For max pooling with a window size $n \times n$:

$$S(i,j) = \max_{(m,n) \in \text{window}} I(i+m, j+n)$$

D. Performance Metrics

Accuracy:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

Precision:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

Recall (Sensitivity):

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

F1-Score:

$$\text{F1-Score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

Cross-Entropy Loss: For multi-class classification:

$$L = -\sum_{i=1}^{N} \sum_{j=1}^{K} y_{ij} \log(\hat{y}_{ij})$$

Where:

- y_{ij}: True label
- $\hat{y}_{ij}$: Predicted probability
- N: Number of samples
- K: Number of classes

V. Problem Statement

Brain tumors, whether benign or malignant, are among the most critical medical conditions, posing significant risks to human health and life. Accurate and early detection of brain tumors is crucial for effective treatment and improving survival rates. Magnetic Resonance Imaging (MRI) is a commonly used diagnostic tool that provides detailed images of brain structures. However, the manual analysis of MRI scans is a time-consuming, error-prone process that heavily relies on the expertise of radiologists. The growing volume of MRI data generated daily in clinical settings exacerbates these challenges, making manual diagnosis infeasible at scale. Recent advancements in deep learning, particularly Convolutional Neural Networks (CNNs), have revolutionized medical imaging by automating feature extraction and improving classification accuracy. Despite their success, existing CNN-based models are often limited in scope, focusing solely on classification without integrating multiple diagnostic tasks, such as tumor detection, grading, and segmentation. This project aims to design and implement a hybrid deep learning model that detects and classifies brain tumors from MRI images. The model will integrate multi-task capabilities, including tumor detection, multi class classification task, and will be complemented by a user-friendly interface to facilitate practical application in medical environments. By addressing the limitations of existing approaches, this project seeks to improve the efficiency, accuracy, and accessibility of brain tumor diagnostics.

VI. Methodology

A. Dataset

The dataset used for this project consists of MRI images of the human brain, specifically aimed at detecting and classifying brain tumors. This dataset is a combination of publicly available datasets, including Figshare, SARTAJ, and Br35H. It contains a total of 7,023 images categorized into four classes: glioma, meningioma, pituitary tumor, and no tumor.

B. Details of Dataset

- **Training Data:** 5,712 images
- **Testing Data:** 1,311 images
- **Categories:**
  - No Tumor
  - Glioma
  - Meningioma
  - Pituitary Tumor
- **Image Format:** 2D grayscale MRI images
- **Resolution:** Each image is resized to 224×224 pixels for compatibility with the input layer of the model.

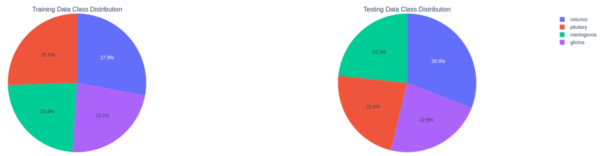

Fig. 1: data distribution

C. Preprocessing

To ensure compatibility with the deep learning model and improve performance, the following preprocessing steps were applied:

1) **Normalization:** All pixel values were normalized to a range of $[0, 1]$.
2) **Resizing:** Images were resized to 224×224 pixels to match the input size requirements of the CNN models.
3) **Data Augmentation:** Techniques such as rotation, flipping, and scaling were applied to artificially increase the dataset size and improve model generalization.

D. Dataset Source

The images in this dataset were sourced as follows:

- Glioma, meningioma, and pituitary tumor images were taken from the Figshare and SARTAJ datasets.
- No tumor class images were sourced from the Br35H dataset.

This dataset provides a balanced and diverse collection of MRI images, ensuring robust training and evaluation of the proposed hybrid deep learning model.

E. CNN Model Architecture

A hybrid CNN model for tumor detection using MRI images. It integrates features from two pre-trained architectures, VGG16 and InceptionV3, and combines them into a unified model. Below is the summary of the architecture and its parameters:

- **Pre-trained Model 1:**
  - Base model: **VGG16**
  - Output layers:
    - Global Average Pooling 2D
    - Dense layer with 256 units (ReLU activation)
    - Dropout (rate: 0.5)
- **Pre-trained Model 2:**
  - Base model: **InceptionV3**
  - Output layers:
    - Global Average Pooling 2D
    - Dense layer with 256 units (ReLU activation)
    - Batch Normalization
    - Dropout (rate: 0.6)
- **Hybrid Layer:**
  - Concatenates outputs from both pre-trained models.
  - Dense layer with 256 units (ReLU activation).
  - Dropout (rate: 0.6).
  - Dense output layer with 4 units (softmax activation).

This hybrid approach combines the hierarchical feature extraction of VGG16 and the diverse feature aggregation of InceptionV3. The use of pre-trained models significantly reduces the computational burden and training time, while fine-tuning enhances task-specific performance. The architecture is tailored for medical image classification tasks, demonstrating its potential for accurate tumor detection.

[gray]0.9 **Parameter**	**Value**
Image Size	(150, 150, 3)
Pre-trained Models	VGG16, InceptionV3
Number of Classes	4 (glioma, meningioma, notumor, pituitary)
Dropout Rates	0.5 (VGG16 branch), 0.6 (InceptionV3 branch & hybrid)
Dense Units (Hidden)	256 (both branches and hybrid layer)
Optimizer	Adam
Loss Function	Categorical Crossentropy
Metrics	Accuracy
Epochs	16
Batch Size	32
Data Augmentation	Rescaling, rotation, width/height shifts, zoom, flips
Trainable Layers (VGG16)	First 10 layers
Trainable Layers (InceptionV3)	First 10 layers

Fig. 2: model parameters for an image illustrating the proposed methodology for brain tumor detection.

VII. EXPERIMENTAL RESULTS

A. Model Architecture

The project utilizes a hybrid model combining two popular pre-trained deep learning models: VGG16 and InceptionV3. The architecture is designed to perform tumor detection from MRI images with the following components:

- **VGG16 Model:** The VGG16 model is used as the first base model, pre-trained on ImageNet. The initial layers of the model are frozen, while the last few layers are fine-tuned. It processes input images of size (150, 150, 3).
- **InceptionV3 Model:** The second base model, InceptionV3, is also pre-trained on ImageNet. Similar to VGG16, its initial layers are frozen, and the later layers are fine-tuned. The output of both models is combined, and fully connected layers are added for classification.
- **Hybrid Model:** The outputs from both the VGG16 and InceptionV3 models are concatenated, passed through dense layers, and the final layer outputs a softmax classification for four classes.

B. Training Setup

The training process used the following configurations:

- **Optimizer:** Adam optimizer was used for training.
- **Loss Function:** Categorical Crossentropy was used to calculate the loss.
- **Metrics:** Accuracy was used as the evaluation metric.

- **Callbacks:** Early stopping and model checkpoint callbacks were utilized to save the best model based on validation loss and prevent overfitting.

The model was trained for 16 epochs, using a batch size that was adjusted according to the available resources. The training utilized the multi-input generator to handle both VGG16 and InceptionV3 models simultaneously during training.

C. Model Training Results

The model showed steady improvement across the epochs in terms of both training loss and accuracy:

- **Epoch 1:** Training accuracy: 65.58%, Validation accuracy: 73.61%
- **Epoch 16:** Training accuracy: 88.76%, Validation accuracy: 86.19%

Over time, the training loss decreased, and the validation loss followed a similar trend, indicating the model's ability to learn the patterns in the MRI data. Some fluctuations in validation accuracy were observed, especially in later epochs, which could be attributed to overfitting or the complexity of the model.

D. Model Evaluation

The evaluation metrics were computed to assess the model's performance:

The classification report for each class is as follows:

TABLE I: Classification Report for Each Class

Class	Precision	Recall	F1-Score
Glioma	0.97	0.79	0.87
Meningioma	0.84	0.65	0.73
No Tumor	0.91	0.98	0.95
Pituitary	0.76	0.98	0.86

E. Confusion Matrix

The confusion matrix of the model on the validation data is shown below:

TABLE II: Confusion Matrix

	Glioma	Meningioma	No Tumor	Pituitary
Glioma	237	35	2	26
Meningioma	5	200	34	67
No Tumor	1	4	398	2
Pituitary	2	0	3	295

F. ROC Curve and AUC

The Receiver Operating Characteristic (ROC) curve and the corresponding AUC for each class were computed to evaluate the model's ability to distinguish between different tumor types. The model showed high AUC values, particularly for the "No Tumor" and "Pituitary" classes.

G. Random Sample Predictions

To further assess the model's performance, a set of random samples was selected from the validation data. The true and predicted labels for these samples are shown below. This helps visualize how well the model is performing on individual cases.

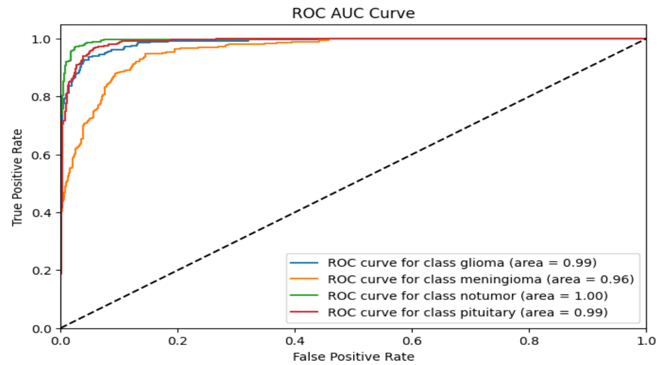

Fig. 3: ROC AUC Curve

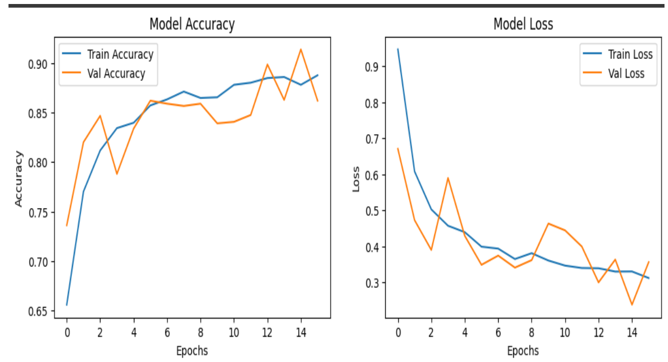

Fig. 4: model accuracy and model loss

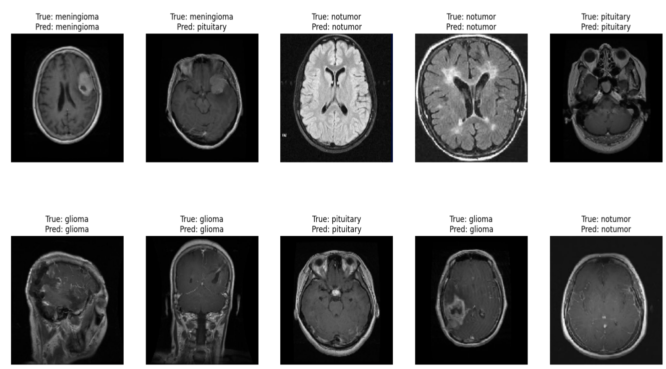

Fig. 5: random samples with actual and predicted labels

VIII. Conclusion

The hybrid model combining VGG16 and InceptionV3 demonstrated strong performance in MRI tumor detection. With an accuracy of 86.19%, the model showed good precision, recall, and F1 scores. The confusion matrix and ROC analysis revealed that the model performs particularly well in detecting "No Tumor" and "Pituitary" cases. However, further improvements could be made for better classification between "Glioma" and "Meningioma" classes.

IX. Future Scope

The future of tumor detection from MRI images using deep learning is highly promising, with significant potential for both academic research and real-world applications. Continued advancements in model architectures, such as incorporating Vision Transformers or attention mechanisms, could improve accuracy and robustness. Expanding datasets and exploring multi-modal approaches, including the integration of clinical data, would enhance the generalizability and applicability of the system.

Real-world deployment in clinical settings as a diagnostic assistant could revolutionize early detection and automated workflows, especially in resource-constrained environments. By optimizing lightweight models for edge AI deployment, this technology could reach underserved areas, addressing healthcare inequities.

Additionally, incorporating explainability techniques, like Grad-CAM, will increase trust among medical professionals by providing insights into model predictions. With these developments, the project is well-positioned to make impactful contributions to AI-driven healthcare, improving outcomes and accessibility for patients worldwide.

REFERENCES

[1] A. Raza, H. Ayub, J. A. Khan, I. Ahmad, A. S. Salama, Y. I. Daradkeh, D. Javeed, A. Ur Rehman, and H. Hamam, "A hybrid deep learning-based approach for brain tumor classification," *Electronics*, vol. 11, no. 7, p. 1146, 2022.

[2] B. Babu Vimala, S. Srinivasan, S. K. Mathivanan, Mahalakshmi, P. Jayagopal, and G. T. Dalu, "Detection and classification of brain tumor using hybrid deep learning models," *Scientific Reports*, vol. 13, no. 1, p. 23029, 2023.

[3] C. Rajeshkumar, K. R. Soundar, M. Sneha, S. S. Maheswari, M. S. Lakshmi, and R. Priyanka, "Convolutional neural networks (cnn) based brain tumor detection in mri images," in *2023 5th International Conference on Smart Systems and Inventive Technology (ICSSIT)*. IEEE, 2023, pp. 976–979.