

Brain Tumor Detection Using Convolutional Neural Networks

S. S. Turkane¹, R. R. Wani², A. R. Sonawane³, A. R. Gite⁴

^{1,2,3,4}*Department of Electronics and Telecommunication Engineering, College of Pravara Rural Engineering College Loni Maharashtra 413736,*

Abstract—This research focuses on accurate brain tumor detection using MRI images through deep learning and transfer learning techniques. The study classifies tumors into four types pituitary, meningioma, glioma, and no tumor using CNN and pre-trained models like MobileNetV2, Vision Transformer (ViT), and VGG16. The results show that all models achieved high accuracy above 95%. Among them, VGG16 performed the best with about 99% training accuracy and 98% testing accuracy, along with excellent precision, recall, and F1-score values. The study concludes that VGG16 is the most effective model for clinical use in supporting radiologists with accurate brain tumor analysis. The final classifications are obtained by merging these features and passing them through a linear classifier. This approach identifies strong and varied characteristics and provides a precise. Proficiency and expertise are required for radiologists to accurately detect brain tumors, a process that requires a significant amount of time. Deep learning technologies are being used and more to automate the diagnosis of brain tumors, resulting in outcomes that are more precise and efficient compared to previous methods. This research focuses on accurate brain tumor detection using MRI images through deep learning and transfer learning techniques. The study classifies tumors into four types — pituitary, meningioma, glioma, and no tumor using CNN and pre-trained models like MobileNetV2, Vision Transformer (ViT), and VGG16. The results show that all models achieved high accuracy above 95%. Among them, VGG16 performed the best with about 99% training accuracy and 98% testing accuracy, along with excellent precision, recall, and F1-score values. The study concludes that VGG16 is the most effective model for clinical use in supporting radiologists with accurate brain tumor analysis. The final classifications are obtained by merging these features and passing them through a linear classifier. This approach identifies strong and varied characteristics and provides a precise. Proficiency and expertise are required for radiologists to accurately detect brain tumors, a process that requires a significant amount of time. Deep learning technologies

are being used and more to automate the diagnosis of brain tumors, resulting in outcomes that are more precise and efficient compared to previous methods.

Index Terms—A rectified linear unit (ReLU), Convolutional neural networks (CNNs), Deep learning (DL), Magnetic resonance Imaging (MRI), Visual geometry group (VGG), Vision transformers (ViT)

I. INTRODUCTION

Brain tumors arise from the irregular progression of cells in brain, making them one of the most serious health issues because of their potential severity and influence on a person's lifespan. Brain tumors are usually characterized into two major types: primary tumors, cancerous brain tumors that originate from the brain, and secondary tumors, which develop due to cancer that may penetrate to the brain from remaining parts of the body [1] Brain tumor detection is crucial for improving patient survival but remains challenging due to variations in tumor shape, size, and location. Traditional methods like the Gamma Knife are time-consuming, so AI-based deep learning techniques are now preferred for accurate and faster diagnosis. Recent studies show high accuracy, but many lack prediction certainty, which is vital for clinical decision-making. This research focuses on developing a reliable AI-based system for precise brain tumor classification [2] Brain tumors are abnormal cell growths that can be benign or malignant, affecting brain function and health. MRI plays a key role in detecting and analyzing these tumors through image segmentation and classification. Deep learning provides an accurate, fast, and non-invasive way to diagnose and grade brain tumors. This study focuses on developing an automated system to improve early detection and treatment planning [3].

Early detection of malignant brain tumors is critical for improving patient survival, but manual diagnosis using MRI is time-consuming, costly, and prone to human error. To overcome these challenges, automated deep learning methods, especially Convolutional Neural Networks (CNNs), are increasingly used for accurate and efficient tumor detection. Various CNN-based models from 2017 to 2021 have achieved accuracies ranging from 91% to 100%, showing the strong potential of AI in brain tumor classification. This research builds on these advancements to develop a more precise and reliable diagnostic system [4]. Recent studies highlight the growing use of hybrid deep learning systems for improved medical image analysis. Transfer learning models like VGG16, ResNet50, and Inception V3 have shown strong performance in brain tumor classification. This research proposes a novel hybrid ensemble model combining Exception and Parallel Deep CNN (PDCNN) architectures to capture both global and fine-grained MRI features. [5]

This study explains brain tumor detection and classification using MRI images with deep learning. MRI scans are used to identify different tumor regions like necrosis, edema, and enhancing areas. The proposed model uses a complementary kernel-based CNN and a modified U-Net for accurate tumor segmentation and classification. It achieves high accuracy on benchmark datasets such as SARTAJ, Br35H, and Figshare using metrics like F1-score, precision, and sensitivity [6]. This study introduces a CBAM-SMK model that combines Convolution Block Attention Modules with Separable Multiresolution Kernels to improve brain tumor classification. It achieves high performance with 97% accuracy and is designed for clinical use, even in limited-resource settings. Overall, this work advances deep learning-based medical diagnostics for reliable brain tumor detection [7].

Artificial intelligence and deep learning are becoming important in medical diagnosis, especially for detecting cancers like brain tumors. Techniques such as CNN and Inception V3 help identify tumor types like glioma, meningioma, and pituitary tumors using MRI and histopathology images. [8]. A brain tumor is an abnormal growth of cells in the brain, which can be benign or malignant. Early detection is very important because brain tumors can be life-threatening. In India, thousands of new cases and deaths occur every year.

Modern Computer-Aided Diagnosis (CAD) systems using machine learning and deep learning help doctors detect brain tumors more accurately and efficiently from MRI images [9]. Recent healthcare advances are driven by new technologies, aging populations, and lessons from the COVID-19 pandemic. Smart healthcare systems and IoT devices help improve patient care and resource management. Deep learning has become a powerful tool in medical imaging, helping diagnose diseases like pneumonia and neurological disorders with greater accuracy and efficiency [10].

Brain tumors are one of the deadliest diseases worldwide, making early and accurate diagnosis essential. MRI is the main tool for detection, but manual analysis is slow and prone to errors. Deep learning, especially CNN-based models, helps automate tumor classification and improve accuracy, enabling faster and more reliable diagnosis for better treatment planning [11]. Brain tumors form due to abnormal cell growth in the brain, affecting normal functions and often becoming life-threatening. MRI is the best imaging method for early and accurate detection, but manual analysis is slow and inconsistent. Deep learning, especially YOLO-based models, helps automate detection quickly and precisely, though challenges remain in identifying small or low-contrast tumors [12]. Radiologists spend a lot of time manually labelling MRI scans, so technology can help by using machine learning and deep learning. Deep learning, especially CNNs, can classify MRI images of brain tumors more accurately. However, small medical datasets make training difficult. Data augmentation and Generative Adversarial Networks (GANs) can create more realistic training images to improve accuracy. DenseNet121 and GANs together can help detect different brain tumor types better and faster, while also keeping patient data private [13]

Brain tumors are a serious global health issue with increasing cases and deaths every year. Deep learning, especially CNN-based models, helps doctors detect and classify tumors more accurately using MRI scans. This study uses a hybrid model called SwinResDual (SwRD), combining Resnet and Swin Transformers, to improve brain tumor classification for both binary and multiclass cases using MRI datasets [14]. Because conventional algorithms are slow and inaccurate, an effective automatic method is required for accurate

brain tumor detection. Otsu binarization is used for picture improvement and GLCM for feature extraction in the suggested Hybrid Deep Convolutional Neural Network with Nature-Inspired ResNet-152 Transfer Learning (Hyb-DCNN-ResNet-152 TL) [15]. Although many models, such as CNN, Googlenet, and 3D CNN, have been employed to detect brain tumors, their accuracy is still quite low.[16]. DCNNs have shown promise, but the complexity of brain tumors necessitates a more sophisticated strategy. In order to evaluate MRI images, the suggested Multi-Path Convolutional Neural Network (M-CNN) introduces several paths, such as diffusion, spectroscopy, and perfusion imaging. This method improves classification beyond simple benign-malignant labelling by capturing a variety of tumor traits using varying kernel sizes. With little data, transfer learning with models like AlexNet, Resnet, and Inception V3 helps minimize overfitting. In brain tumor classification, the M-CNN reduces computational cost while improving feature variety, accuracy, and efficiency [17]. A brain tumor, which can be either benign or malignant, is an aberrant and unchecked development of brain cells. Malignant tumors can spread from other parts of the body or start in the brain. [18].

This review, which covers research from 2016 to 2024, focuses on deep learning-based brain tumor detection and classification. It evaluates datasets and hybrid models, contrasts MRI, CT, and PET methods, and pinpoints important research gaps like inadequate generalization and data bias.[19] The study offers a concise summary to direct future investigations and enhance the precision of clinical diagnoses. Section 1 presents the background, motivation, and main contributions of the study. Section 2 surveys the related works on brain tumor classification using DL techniques. Section 3 deliberates the MRI database, the pre-processing, and the augmentation practices applied to the data. The methodology approach is presented in Section 4, with the custom CNN model and fine-tuned VGG-16 architecture. [20] Section 5 accounts the experimental results with the corresponding comparison of the performance to contemporary methods. Finally, Section 6 completes the work by summarizing the results and signifying possible information for forthcoming research.

Novelty and contributions:

- Developed and evaluated four distinct models for brain tumor classification: one custom ensemble CNN, and three transfer learning models (VGG16, MobileNetV2, Vision Transformer).
- Conducted a comparative study of traditional CNNs, lightweight architectures, and transformer-based networks on the same MRI dataset.
- Demonstrated that VGG16 achieved the best clinical balance, with ~98.78 % accuracy and specificity above 0.98.
- Reported complete evaluation metrics (accuracy, F1-score, precision, specificity, recall) to ensure reliable performance for medical applications.

II. LITERATURE REVIEW

Brain tumors can significantly impair cognitive abilities, and the best course of therapy depends on early and precise diagnosis. However, manual diagnosis is ineffective and prone to mistakes. In order to address this, the paper [21] presents three transfer learning-based deep learning models that efficiently classify brain tumors without the need for preprocessing, segmentation, hybrid techniques, or data augmentation. These models conserve computational resources and are thoroughly tested using the Fig share T1-weighted contrast-enhanced MRI dataset, which comprises 3064 images from three different types of tumors. Five-fold cross-validation yielded an accuracy of 98.36% for InceptionV3, and the exception model achieved 99.19% using batch normalization and an optimized Adam optimizer. 98.85% was attained using a more refined version of exception that used PSO with the SGD optimizer. [22] to classify brain tumors on MRI images using refined deep neural networks. To ensure realistic input quality, the expected model is evaluated using datasets from Nickparvar and Cheng along with pre-processing steps like resize, normalize, data augmentation, and noise reduction. Notable architectures, including DenseNet-121, InceptionV3, ResNet-50, VGG16, and the TTL model, showed exceptional performance, producing precision, F1-scores, and recall all close to 93% as well as accuracy (94.5%), sensitivity (92.8%), and specificity (93.1%) scores. The observed results demonstrate good performance above traditional methods, providing radiologists with reliable decision-

support tools that can improve treatment planning, speed up diagnosis, and lower costs. Because deep learning-based brain tumor detection can produce an automated and precise diagnosis, it has garnered special attention. In order to assess the practicality of three superior deep learning techniques for the classification of brain cancers from MRI scans, this proposed work [23] looks at Conventional CNN, Optimized CNN configuration, and Transfer Learning with VGG-16. The MRI dataset from the Kaggle platform, which consists of 1755 pictures, was used to train (80%) and test (20%) the DL models. The models' performances were assessed using accuracy and loss criteria. According to the results, the Optimized CNN demonstrated greater feature extraction capabilities with 90.0% accuracy, despite the Conventional CNN setting the baseline of 85% accuracy. With an accuracy of 97.4%, the VGG-16 model demonstrates its. Additionally, the study highlights the wider advantages of early and accurate detection for patient outcomes and overall care quality. Future study suggests that algorithms be validated in actual clinical settings and additional diverse datasets for generalization, which may raise issues with patient confidentiality and data security. Larger datasets and deep learning models will be used in future research to improve accuracy. By offering guidance on selecting the best configuration for brain tumor detection, our work contributes significantly to the advancement of AI-based diagnostic medicine.

Human health and life expectancy are significantly impacted by brain tumors. Because magnetic resonance imaging (MRI) offers precise structural information about the brain, it is frequently employed for tumor detection. Deep transfer learning was investigated for brain tumor categorization by Prabhpreet Kaur et al. [24]. An enhanced deep learning system for automatic brain tumor diagnosis utilizing MRI images was created by Mohammed Amin Almaiah et al. [25]. The suggested approach uses sophisticated preprocessing techniques like contrast enhancement and normalization to improve image quality and consistency. Brain tumors are accurately extracted and classified into glioma, meningioma, pituitary, and normal categories using a hybrid CNN model that blends VGG19 and EfficientNet architectures. This method outperforms conventional single-model frameworks with an astounding accuracy of 99.1%. The model is also appropriate for real-time

medical use due to its quicker convergence and reduced processing cost.[26] used MRI images to classify tumors using a variety of transfer learning architectures, including AlexNet, GoogleNet, ResNet-50, and VGG-16. The suggested hybrid GN-AlexNet model outperformed the other models in terms of accuracy, F1-score, sensitivity, and specificity. The achieved accuracies AlexNet (95.60%), GoogleNet (98.99%), ResNet-50 (96.90%), and VGG-16 (97.66%) showcase these models' resilience and dependability in resolving class imbalance and guaranteeing nearly flawless categorization. These results imply that early tumor diagnosis and treatment planning can be greatly enhanced by combining transfer learning with efficient data augmentation. Future studies might concentrate on improving this framework, testing it on bigger and more varied datasets, and creating an approachable clinical tool to help doctors diagnose brain tumors. A deep learning model to identify brain tumors from MRI images and categorize meningioma, pituitary, and glioma types was proposed by Md. Ridwan et al. With 99.08% accuracy, the Vision Transformer outperformed all other models, followed by CNN-based and hybrid models.[27] Future work will emphasize on improving adaptability for real-world clinical use, ensuring early detection and better treatment outcomes. Sadia Islam Tonni et al. [28] introduce an advanced hybrid transfer learning model to classify brain tumors from MRI images. The proposed model combines advanced preprocessing, ensemble deep learning, and explainable AI (XAI) to deliver accuracy and transparency. Image quality is first improved through noise reduction and contrast enhancement, after which VGG16 and ResNet152V2 models are integrated, achieving an impressive 99.47 % accuracy on a dataset with four-classes. To make the results interpretable, Grad-CAM highlights tumor regions directly on the scans, while SHAP quantifies the importance of different features, building clinical confidence in the system. An ablation study further demonstrates how different hyperparameters affect performance. Overall, the framework proves to be highly accurate, robust, and explainable well-suited for early tumor detection. The BRATS 2019–2021 dataset containing MRI images is reinforced and preprocessed for reducing overfitting, then segmented through VGG-19 layers, followed by feature extraction using AWA and feature selection via a Model Agnostic Concept

Extractor (MACE). A. Maria Nancy et al. [29] recommended a model for highly accurate classification of tumors into tumor core, enhanced tumor, and whole tumor, achieving 96.20 %, 98.20 %, and 99.40 % on the 2019, 2020, and 2021 datasets. Segmentation performance also showed strong Dice Similarity Coefficient (DSC) scores of 91.06 %, 91.40 %, and 94.86%, outperforming traditional CNNs. The proposed model is validated through accuracy, specificity, sensitivity, DSC, and HD confirms the robustness of the approach, and comparisons demonstrate its superiority over existing models such as MAC-ARUNet3+, BTSC-TNAS, Mask R-CNN, SwinBTS, and Residual U-Net. [30]. The ResNet18 model was first trained on one dataset, with 20 % of

the data reserved for testing, and later validated on a second dataset to evaluate generalizability. The framework achieved impressive results—99.65 % accuracy for four tumor classes and 98.05 % for seven classes on the first dataset. With the proposed preprocessing methods, results further improved, reaching 99.31 % (four-class) and 98.90 % (seven-class) on the second dataset, and boosting the first dataset to 99.84 % and 99.03 % respectively. The proposed work achieves good accuracy across both datasets and extends to seven tumor categories. This work demonstrates high accuracy and strong robustness with generalization, offering valuable improvements for brain tumor diagnosis.

Table 1: Advantages and limitations of existing works.

Study / Reference	Method / Model(s) Used	Dataset Size & Classes	Reported Accuracy	Remarks
Raja Waseem Anwar et al. [22]	Transformative Transfer Learning (VGG16, ResNet-50, InceptionV3, DenseNet-121)	Nickparvar & Cheng datasets	94.50 %	Good sensitivity/specificity (~93 %), but overall accuracy lower than ours.
Ilyyasu, U et al. [23]	Conventional CNN, Optimized CNN, VGG-16 (transfer learning)	Kaggle MRI dataset (1755 images, 3 classes)	CNN: 85 %, Optimized CNN: 90 %, VGG-16: 97.4 %	Limited dataset; VGG-16 strong but still below ours.
Prabhpreet Kaur et al. [24]	ResNet152, GoogleNet + ML classifiers (SVM, KNN, GNB, etc.)	MRI dataset (train 80 %, test 20 %)	ResNet152: 98.53 %, GoogleNet Lower	Strong performance but still under us 98.78 % VGG16.
Rizal Dwi Prayogo et al. [25]	Hybrid CNN (ResNet50V2, MobileNetV2, DenseNet121)	MRI dataset (7023 images, 4 classes)	98.75 %	Very competitive, but slightly below us VGG16.
Mohammed Amin Almaiah et al. [26]	Hybrid GN-AlexNet, VGG-16, ResNet-50, GoogleNet	MRI dataset	VGG-16: 97.66 %, GN-AlexNet Better	Their hybrid GN-AlexNet outperforms, but their VGG16 baseline is lower than ours.
Md. Ridwan et al. [27]	VGG16 transfer learning (compared with ViT & others)	MRI dataset (Glioma, Pituitary, Meningioma)	97.08 %	VGG16 baseline is weaker than ours.
Proposed Work (This Study)	Ensemble CNN, VGG16, MobileNetV2, Vision Transformer	MRI dataset (7023 images, 4 classes)	VGG16: 98.78 %, MobileNetV2: 97.48 %, ViT: 96.72 %, CNN: 96.72%	Outperforms baselines with balanced metrics and >0.98 specificity.

III. DATA SET

The combined brain tumor classification multi-centre, multi- sequence MRI dataset is a curated collection of pre-operative MRI exams with segmentation and radiological reports. In order to address the issues related to class imbalance as well as the missing

healthy samples in each dataset, three publicly available datasets have been combined, which include the Figshare Brain Tumor Dataset [31], the SARTAJ Brain Tumor Classification Dataset [32], and the Br35H Brain Tumor Detection Dataset [33]. The resulting dataset contains 7023 T1-weighted contrast-enhanced MRI scans [34] grouped into four

categories: glioma (1621 images), meningioma (1645 images), pituitary tumor (1757 images), and no tumor (2000 images). These images, all in JPEG format with a resolution of 512×512 pixels, are split into 5712 training samples (81 %) and 1311 testing samples (19 %), ensuring sufficient data for supervised learning. Preprocessing steps, such as augmentation, normalization, and then resizing, are typically applied to account for differences in acquisition protocols and to improve generalization across patient variations in tumor morphology, size, and location. The training and testing images are classes are shown in Table 2.

Table 2: Image counts of all classes in brain tumor classification.

Class	Testing Images	Training Images	Total Images
No Tumor	405	1595	2000
Glioma	300	1321	1621
Pituitary	300	1457	1757
Meningioma	306	1339	1645
Total	1311	5712	7023

Kokz et al. [35] demonstrates the utilization of computational modeling and numerical simulations to scrutinize biomedical structures. Such modeling methods overlay with systems-based examination and soft computing methods (e. g., neural networks, fuzzy logic) for medical decision support and stress-strain prediction in biological classifications. Al-Refai et al. [36] worked on systematic modeling and pattern recognition methods. Such modeling and techniques can be further explored employing soft computing methods for example machine learning for molecular design, property prediction or automated structural classification-making it applicable to computational aspects. The component datasets each bring unique strengths. The Br35H dataset includes 3000 scans (with or without tumor) for binary scans included; contains glioma, meningioma, and pituitary tumor; Together, these datasets form a competitive benchmark that is largely employed by CNN and transformer-based methods. The curated dataset is shown in Fig. 1

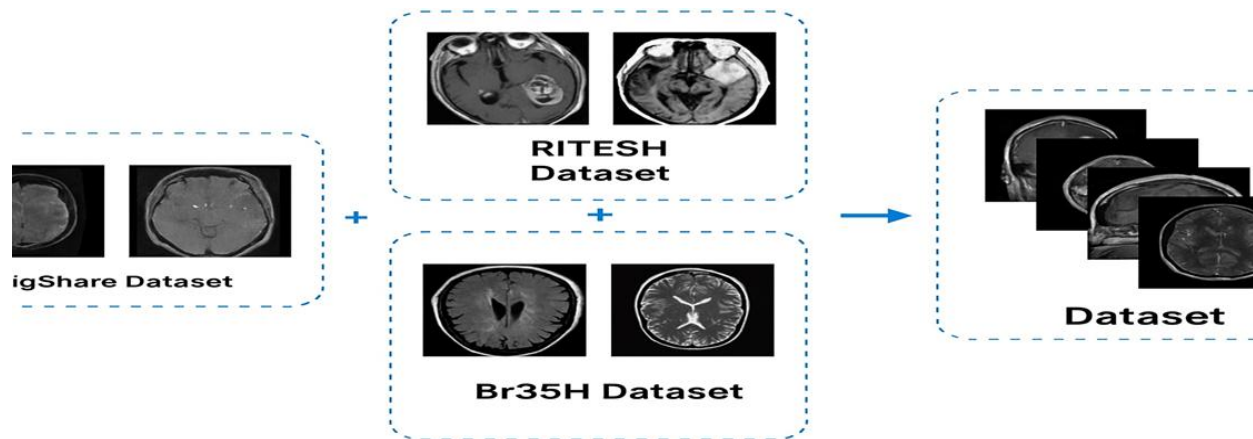


Fig. 1. Curation of dataset from FigShare+RITESH+Br35H

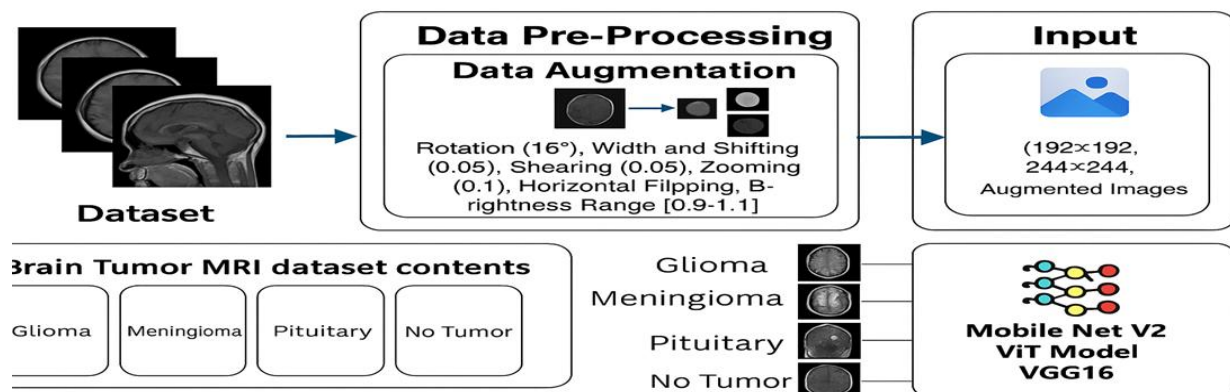


Fig. 2. The Proposed architecture for brain tumor classification

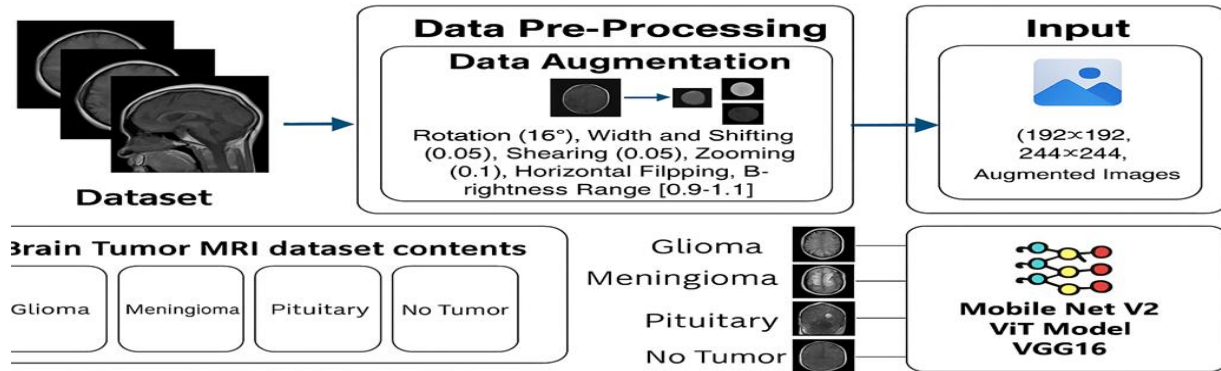


Fig. 3. The Proposed architecture for brain tumor classification

IV. METHODOLOGY

The brain tumor classification methodology consists of preparing the Brain Tumor MRI data set, which is organized into four different categories: Glioma, Meningioma, Pituitary tumor, and No Tumor. Since medical imaging data is often limited in size, we performed data augmentation practices to increase variability and increase model generalization. These augmentations consist of controlled perturbations such as rotation (up to 15 degrees), height and width shift (0.05), shear (0.05), zooming (0.1), horizontal flip, and brightness transform in the range [0.9, 1.1]. Some of these steps, which simulated real imaging variations, make the dataset more challenging and representative of possible clinical cases. After augmentation, the MRI scans were resized to standard input dimensions of 192×192, 244×244, and 256×256 pixels, creating uniform input formats suitable for training different

neural network architectures. Once the pre-processed data were ready, it was fed into multiple DL models for classification. To make the best use of distinctive strengths from the model architectures, we used the MobileNetV2 (a lightweight CNN targeting efficient computation), VGG16 (a deeper CNN with strong feature extraction), and Vision Transformer (ViT) that captured the global image dependencies using attention mechanisms. An Ensemble CNN approach was also implemented to integrate predictions from numerous models for improving the classification accuracy. These models were subsequently fine-tuned to distinguish among the four classes from the augmented MRI images. The rigorous preprocessing, diverse input sizes, and multiple advanced architectures provided a comprehensive framework to achieve accurate and reliable brain tumor categorization. The overall brain tumor classification architecture is shown in the Fig. 2.

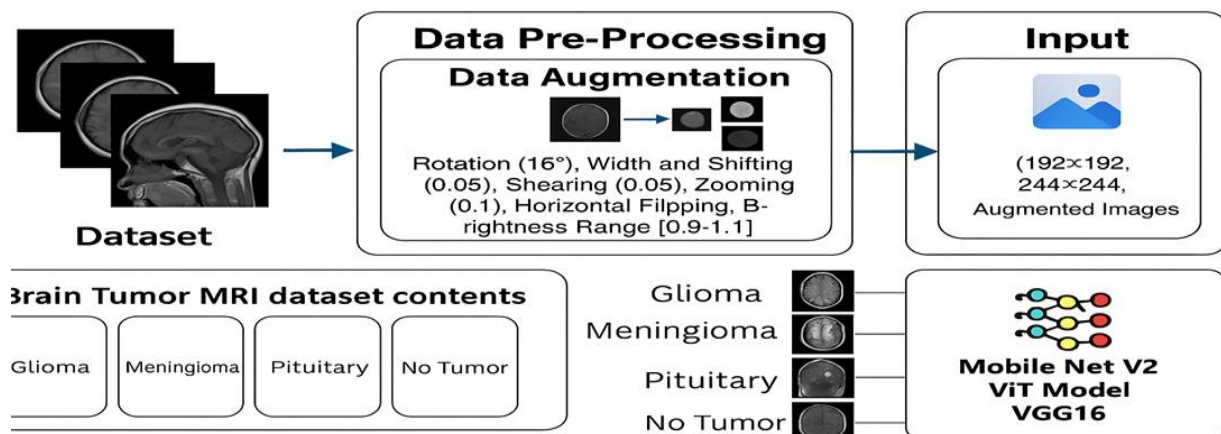


Fig. 4. The Proposed architecture for the ViT model.

Ensemble CNN Architecture

The architecture of the ensemble CNN shown in Fig. 3 is based on a well-defined pipeline for preprocessing and a classification pipeline for the brain tumor MRI dataset. The workflow begins with a dataset of MRI images characterized into four classes: glioma, meningioma, pituitary tumor, and no tumor. This is the dataset that will be used as the base for future analysis. The first is preprocessing the data, in particular, data augmentation to enlarge the dataset with more varied data and improve the model for generalization. During augmentation, several transformations are applied to the original images. These involve rotation by 15 degrees, shifts of up to 0.05 of image size in width or height, shearing at an angle of 0.05, and zooming by 0.1. Moreover, horizontal flipping is utilized for mirror images, and brightness is varied between 0.9 and 1.1 for different illuminations. The resulting augmented images are resized to standardized dimensions, such as 192×192 , 244×244 , or 256×256 pixels, to ensure homogeneity and compatibility with the subsequent models. These augmented images are used as inputs to an ensemble of CNNs. A CNN [37] ensemble is a collection of CNNs that can be trained independently from each other or can be designed to specialize in different features or subsets of the data. This variety of the ensemble CNN leads to the enhancement of classification accuracy and robustness, not only by decreasing the overfitting problem, but also by expanding the variety of patterns that they capture.

Vision Transformer (ViT) Architecture

The Vision Transformer (ViT) architecture [38] shown in Fig. 4 is based on a well-defined pipeline for preprocessing and a classification pipeline of brain tumor MRI data set. The workflow begins with a dataset of MRI images characterized into four classes: glioma, meningioma, pituitary tumor, and no tumor. This dataset serves as the foundational input and undergoes an initial data preprocessing phase focused

on data augmentation to increase diversity and improve the model's generalization competence. The augmentations are applied to the images by rotating them up to 15 degrees to account for different angles, shifting them by width and height up to 0.05 of the original height and width to model varying positions, shearing the image angle by up to 0.05 to mimic slight distortions, and zooming in or out by a factor of 0.1 to make the field of view smaller or larger. Furthermore, the horizontal flip operation produces invariance with respect to the image reverse and brightness and contrast, which are adjusted within a range of 0.9 to 1.1 to address differences in image illumination. The augmented images are then resized to a standardized resolution of 256×256 pixels to confirm consistency and compatibility with the subsequent model stages. These preprocessed images are then passed to the ViT model, which uses a sequence of dense neural network layers for the extraction and classification of features. The network starts with an input dense layer applied to the input data, and then the output from the layer is normalized at training time using batch normalization for increased stability and faster training. This is succeeded by a second 256 units dense layer again, with batch normalization and ReLU activation, and a second dropout layer with a rate of 0.3 to further control model complexity. The second dense layer of 512 units with 0.5 rate dropout and ReLU activation further expands the scale of the network to capture complex features. The architecture ends with a final thick layer with 16 units in addition to a ReLU activation, which gives the probabilities of classification, to recognize the kind of brain tumor (Glioma, Meningioma, Pituitary tumor, or No Tumor) of the input MRI images. This dual-level behavior and an appropriate regularization methodology allow for robust and accurate performance in tumor classification.

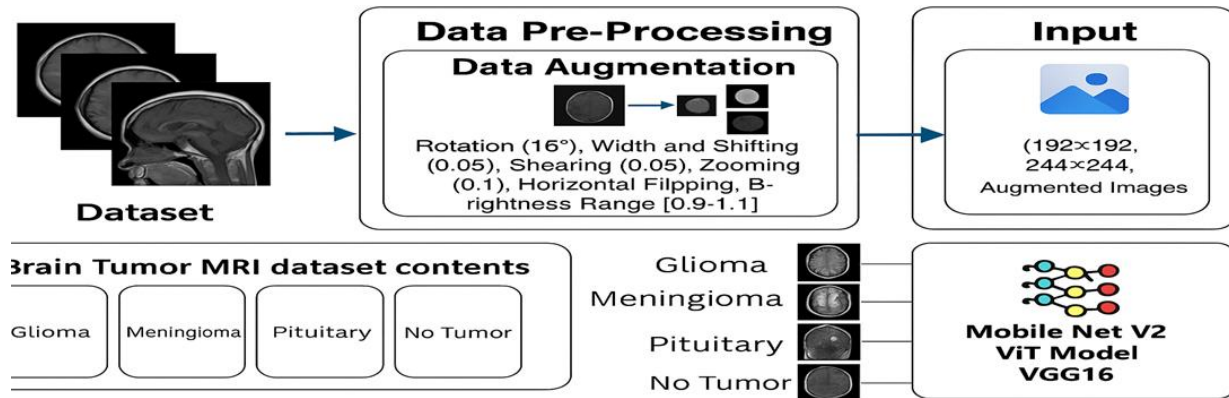


Fig. 5. The Proposed architecture for the MobileNetV2 model

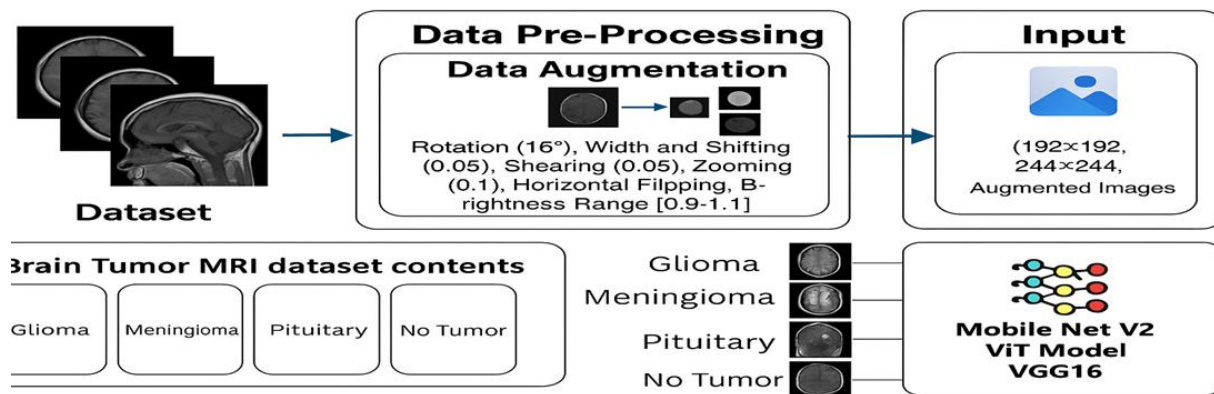


Fig. 6. The proposed architecture for the VGG16 model.

MobileNetV2 Architecture

The MobileNetV2 architecture [39] shown in Fig. 5 is developed to analyze and classify the brain tumor MRI images. The pipeline begins with a dataset of MRI images classified into four classes: Glioma, Meningioma, Pituitary tumor, and No Tumor. This dataset goes through an initial data preprocessing stage that revolves around data augmentation to enrich the dataset diversity and the model's generalization capabilities. The augmentation involves several operations: rotations up to 15 degrees, height and width, and shifts up to a factor of 0.05 to simulate positional variations, shear up to 0.05 angle, which is used to mimic slight distortion, and zooming by a factor of 0.1, to change the field of view. Furthermore, horizontal flip leads to mirror image instances, while brightness is normalized to be between 0.9 and 1.1 to account for variance in illumination of the image. The augmented images are then resized to 256×256 pixels to be compatible with the model stages for the subsequent processes. These preprocessed images are

fed to the MobileNetV2 model, which contains various dense layers and is designed for lightweight and efficiency. The processing of data begins with a preliminary dense layer that transforms the input information, with a softmax activation function to manage the multi-class classification task by converting raw scores into probabilities. Batch Normalization is utilized to normalize the layer inputs in order to increase stability and improve the training process. At the same time, a ReLU activation function induces the non-linearity required by the network to capture complex features. A dropout layer (rate = 0.4) is employed to reduce overfitting by randomly dropping 40% of neurons in the training process for every iteration for better generalization. This is followed by a second dense layer of neurons in which the rectified linear activation function (ReLU) is applied, and the dropout rate is set to 0.3 to further control the complexity and bias of the model towards any one feature. The MobileNetV2 backbone contributes to the model's efficiency by reducing

computational complexity while maintaining high performance on the classification task.

VGG16 Architecture

The VGG16-based model [40] shown in Fig. 6 is designed to classify brain tumor MRI scans into four categories: Glioma, Meningioma, Pituitary, and No Tumor. The process begins with MRI images that undergo a preprocessing stage with data augmentation to improve diversity and robustness. Images can be rotated up to 15°, shifted up to 5% in width and height, sheared somewhat (up to 0.05), zoomed in by 10%, flipped horizontally, and adjusted brightness between 0.9 and 1.1. Before being fed into the model, all images are scaled to a standard 256×256-pixel size after augmentation to ensure consistency. After preprocessing, the images are passed into the VGG16 architecture which provides a strong convolutional backbone for feature extraction. On top of this base, additional fully connected (dense) layers are added to

specialize the network for tumor classification. Using a ReLU activation adds non-linearity and serves as a feature transformer for the first dense layer, enabling the network to identify intricate patterns. In order to improve stability and convergence, batch normalization is used here to normalize inputs to each layer. A dropout layer with a rate of 0.4 is then used to deactivate 40 % of neurons randomly during the training process, which helps avoid overfitting. To balance learning and regularization, the model is further deepened with a second dense layer of 256 units, which is once more combined with batch normalization, ReLU, and a lower dropout rate of 0.3. These carefully designed dense, normalization, and dropout layers, combined with the deep convolutional structure of VGG16, give the model an outstanding balance between feature extraction, stability, and generalization, making it ideal for precise brain tumor classification.

Table 3: The performance evaluation for the ensemble CNN model.

#	Class	Precision	Recall (Sensitivity)	F1 Score	Specificity	Train Accuracy	Train Loss	Test Accuracy	Test Loss
1	Glioma	98.94 %	93.67 %	96.23 %	99.70 %	95.62 %	12.16 %	96.72 %	11.28 %
2	Meningioma	95.99 %	93.79 %	94.88 %	98.81 %	95.62 %	12.16 %	96.72 %	11.28 %
3	No Tumor	96.19 %	99.75 %	97.94 %	98.23 %	95.62 %	12.16 %	96.72 %	11.28 %
4	Pituitary	96.10 %	98.67 %	97.37 %	98.81 %	95.62 %	12.16 %	96.72 %	11.28 %
5	Average	96.81 %	96.47 %	96.60 %	98.89 %	95.62 %	12.16 %	96.72 %	11.28 %

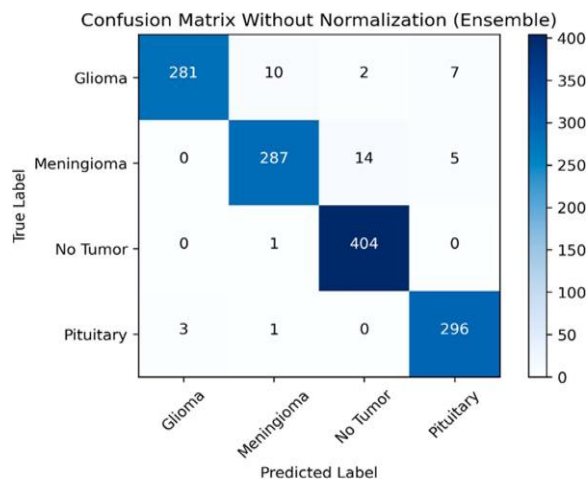


Fig. 7. The confusion matrix for the ensemble CNN model.

Ensemble CNN Model

The train accuracy of the Ensemble CNN is 95.62 % and the test accuracy is 96.72 %, with a test loss of only 11.28 %, indicating good generalization. The

results in the performance Table 3 show consistently high values for most of the evaluation metrics; the average precision being 96.81 %, recall 96.47 %, F1-score 96.6 %, and specificity 98.89 %. Class-level outcomes show that glioma had the highest precision (98.94%), no tumor achieved near-perfect recall (99.75 %), and pituitary tumors demonstrated excellent balance between precision (96.10 %) and recall (98.70 %). The confusion matrix in Fig. 7 further validates these findings, clearly showing the distribution of correct and misclassified predictions across the four tumor classes. Although most cases were accurately classified, there were slight misclassifications between glioma and meningioma because of the overlap in imaging features. However, the outcomes demonstrate that the proposed Ensemble CNN is robust and well-designed for brain tumor classification; therefore, it is highly competitive for clinical decision support to enhance medical systems. The training and validation plots shown in Fig. 8 verify

an increasing accuracy over epochs and a consistent diminution in training and validation loss, indicating that the model is learning in a stable manner.

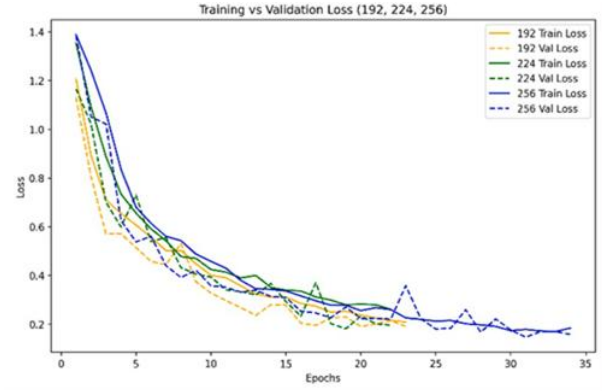
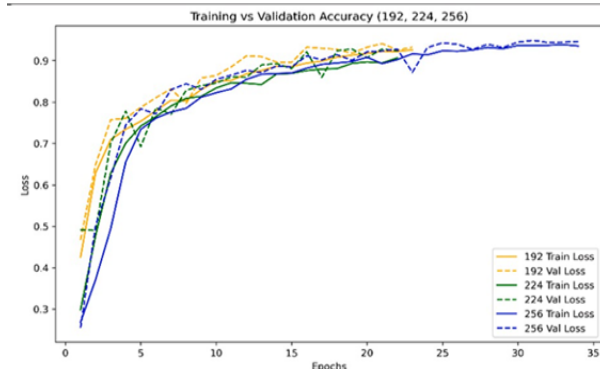


Fig. 8. The accuracy and loss plots for the ensemble CNN model.

Table 4: The performance evaluation for the ViT model.

#	Class	Precision	Recall (Sensitivity)	F1 Score	Specificity	Train Accuracy	Train Loss	Test Accuracy	Test Loss
1	Glioma	97.97 %	96.33 %	97.14 %	99.41 %	98.42 %	3.77 %	96.72 %	8.78 %
2	Meningioma	97.20 %	90.85 %	93.92 %	99.20 %	98.42 %	3.77 %	96.72 %	8.78 %
3	No Tumor	97.35 %	99.75 %	98.54 %	98.79 %	98.42 %	3.77 %	96.72 %	8.78 %
4	Pituitary	94.29 %	99.00 %	96.59 %	98.22 %	98.42 %	3.77 %	96.72 %	8.78 %
5	Average	96.70 %	96.48 %	96.55 %	98.90 %	98.42 %	3.77 %	96.72 %	8.78 %

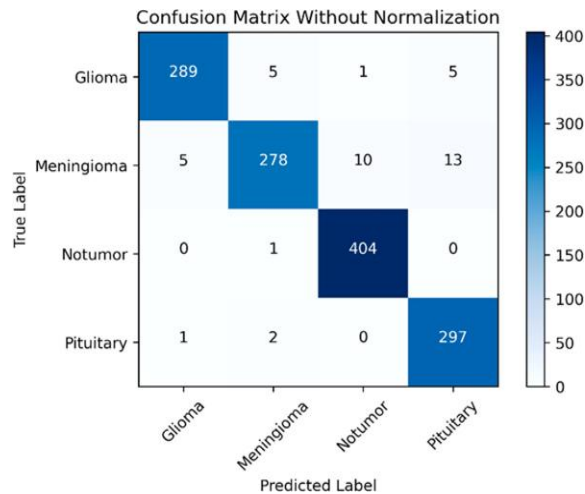


Fig. 9. The confusion matrix for the ViT model.

ViT Model:

The train accuracy of the Vision Transformer (ViT) was 98.42 %, and the test accuracy was as high as 96.72 %, with a test loss of 8.78 %. Results are presented analysing the performance in Table 4, showing the achievement of performance balanced in all classes with an average precision of 96.70 %, recall of 96.48 %, F1-score of 96.55 % and specificity of 98.90 %. At the class level, the model performed best

in detecting no tumor cases, reaching a recall of 99.75 % and an F1-score of 98.54 %, showing strong capability to differentiate healthy patients from tumor cases. High precision is achieved by glioma (97.97 %) with slightly reduced recall (96.33 %), while meningioma showed the lowest recall (90.85 %), indicating that this class was comparatively more challenging for the ViT model. The confusion matrix in Fig. 9 shows that most samples across all four categories were correctly classified, with few misclassifications between glioma and meningioma and between meningioma and pituitary tumors, which share overlapping features in MRI scans. The training and validation plots shown in Fig. 10 verify an increasing accuracy over epochs and a consistent decrease in training and validation loss, indicating that the model is learning in a stable manner. Globally, the ViT model showed a strong and meaningful performance. Nevertheless, with slightly lower sensitivity for meningioma, others will also need further fine-tuning or a class-balancing approach to reach their highest diagnostic reliability. However, its slightly lower sensitivity for meningioma suggests that additional fine-tuning or class-balancing strategies could further improve its diagnostic reliability.

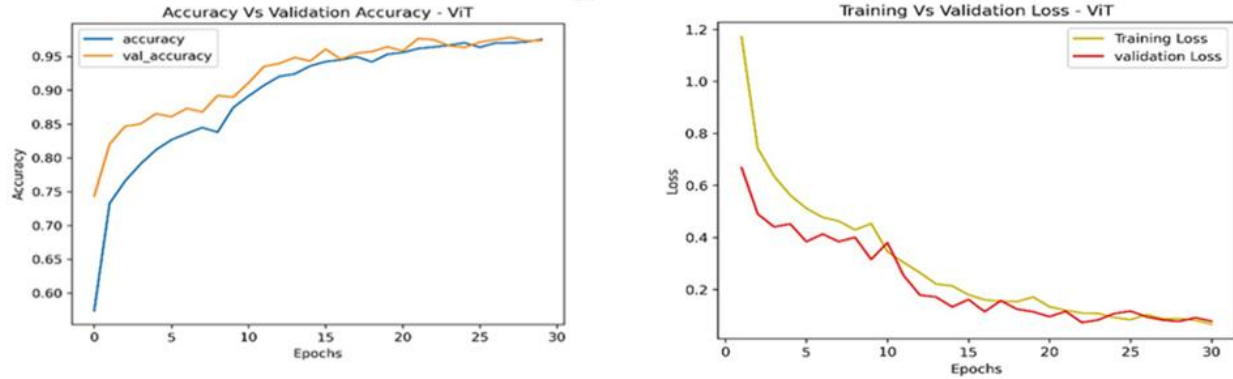


Fig. 10. The accuracy and loss plots for the ViT model.

Table 5: The performance evaluation for the MobileNetV2 model.

#	Class	Precision	Recall (Sensitivity)	F1 Score	Specificity	Train Accuracy	Train Loss	Test Accuracy	Test Loss
1	Glioma	98.94 %	93.33 %	96.05 %	99.70 %	99.80 %	0.73 %	97.48 %	10.28 %
2	Meningioma	93.69 %	97.06 %	95.35 %	98.01 %	99.80 %	0.73 %	97.48 %	10.28 %
3	No Tumor	99.51 %	99.51 %	99.51 %	99.78 %	99.80 %	0.73 %	97.48 %	10.28 %
4	Pituitary	97.39 %	99.33 %	98.35 %	99.21 %	99.80 %	0.73 %	97.48 %	10.28 %
5	Average	97.38 %	97.31 %	97.31 %	99.18 %	99.80 %	0.73 %	97.48 %	10.28 %

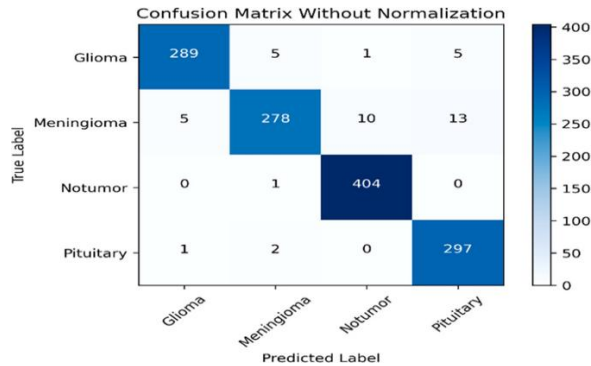


Fig. 11. The confusion matrix for the MobileNetV2 model.

MobileNetV2 Model

The MobileNetV2 model achieved an impressive train accuracy of 99.80 % and a test accuracy of 97.48 %, with a relatively low-test loss of 10.28 %. The results are summarized in the performance Table 5, showing strong and balanced outcomes with an average precision of 97.38 %, recall of 97.31 %, F1-score of 97.31 %, and specificity of 99.18%. The model excelled in detecting classes with no tumor for each class, getting a perfect precision/recall/F1-score of 99.51 %. High recall and F1-score (98.35 %) are also achieved for pituitary tumor detection (99.33 %). However, Glioma classification demonstrated slightly less recall (93.33 %) than precision (98.94 %),

meaning only a few cases were missed. Similarly, meningioma detection resulted in high recall (97.06 %) but was not so high precision (93.69 %), indicating some false positives. The confusion matrix shown in Fig. 11 validates these observations, showing excellent classification performance with minimal misclassifications. A few glioma cases are misidentified as meningioma and vice versa, highlighting the overlap between these tumor types in MRI imaging. The corresponding training and validation loss curves pair in Fig. 12 also revealed that MobileNetV2 learned well with smooth learning, which fluctuated slightly in the beginning epochs. Overall, MobileNetV2 achieved the best generalization results.

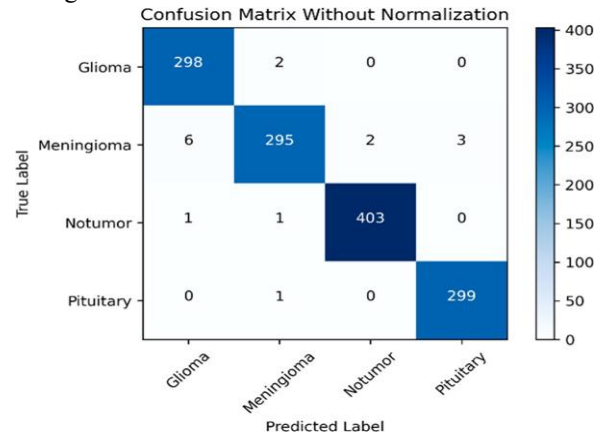


Fig. 13. The confusion matrix for the VGG16 model.

Table 6: The performance evaluation for the VGG16 model.

#	Class	Precision	Recall (Sensitivity)	F1 Score	Specificity	Train Accuracy	Train Loss	Test Accuracy	Test Loss
1	Glioma	97.70 %	99.33 %	98.51 %	99.31 %	99.37 %	1.91 %	98.78 %	4.26 %
2	Meningioma	98.66 %	96.41 %	97.52 %	99.60 %	99.37 %	1.91 %	98.78 %	4.26 %
3	No Tumor	99.51 %	99.51 %	99.51 %	99.78 %	99.37 %	1.91 %	98.78 %	4.26 %
4	Pituitary	99.01 %	99.67 %	99.34 %	99.70 %	99.37 %	1.91 %	98.78 %	4.26 %
5	Average	98.72 %	98.73 %	98.72 %	99.60 %	99.37 %	1.91 %	98.78 %	4.26 %

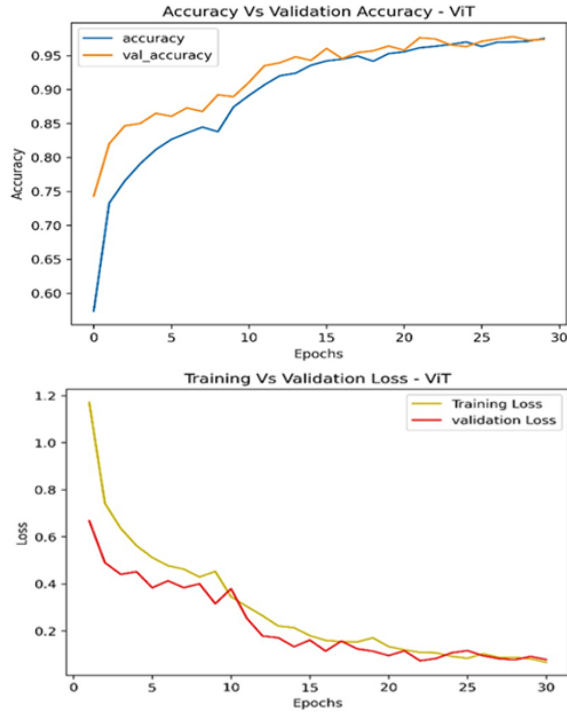


Fig. 12. The accuracy and loss plots for the MobileNetV2 model.

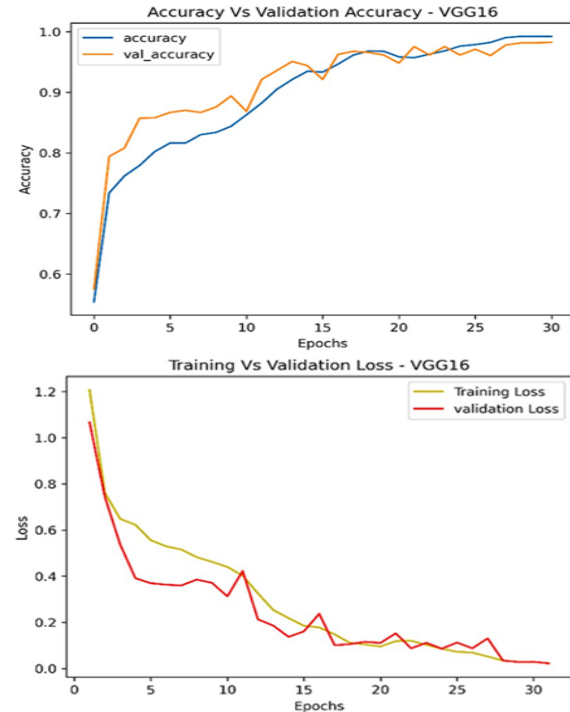


Fig. 14. The accuracy and loss plots for the VGG16 model.

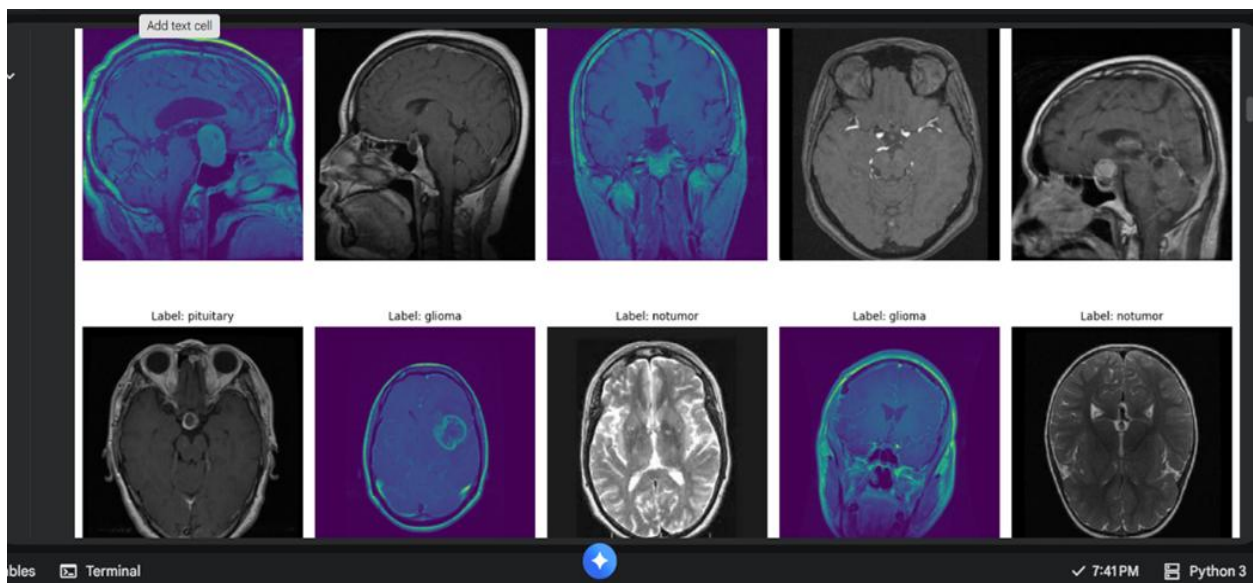


Fig. 15. The accuracy and loss plots for the VGG16 model

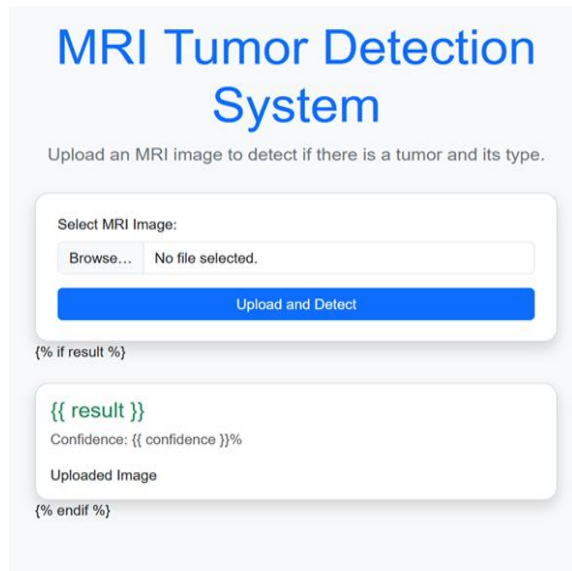


Fig.15.Web Interface of the MRI Tumor Detection System

VGG16 model:

The VGG16 model achieved a strong training accuracy of 99.37 % and a test accuracy of 98.78 %, with a very low-test loss of 4.26 %. The performance metrics presented in Table 6 show highly consistent results, with an average precision of 98.72 %, recall of 98.73 %, F1-score of 98.72 %, and specificity of 99.60 %. The model can distinguish the “No Tumor” class, achieving a perfect precision, recall, and F1-score balance at 99.72 %. Likewise, pituitary tumor detection reached high recall precision and recall across most classes, the VIT model yielded comparable findings, although with a slightly lower sensitivity for meningioma. The MobileNetV2 model delivered 97.48 % test accuracy, with excellent performance in “no tumor” and pituitary classification, showcasing its efficiency and suitability for lightweight deployment. The VGG16-based transfer learning model achieved the highest performance, with 98.78 % test accuracy, an average precision and recall of 98.7 %, and specificity of 0.996, making it the most reliable and well-balanced approach. These results demonstrate how ensemble-based approaches and transfer learning significantly increase diagnosis accuracy, lower misclassification rates, and give radiologists reliable decision support. Overall, the findings support VGG16 as the best model for clinical use, with VIT and MobileNetV2 acting as competitive and computationally efficient counterparts.

V. FUTURE SCOPE

Future studies on the identification and categorization of brain tumors may concentrate on developing more sophisticated and useful AI-based diagnostic tools. The generalization and dependability of deep learning models will be enhanced by extending research to incorporate vast, varied, and multi-institutional MRI datasets. Combining various imaging modalities, such as CT, PET, and fMRI, can improve diagnostic accuracy and offer more thorough tumor characterisation. Predictions will be more transparent and reliable for clinical usage when explainable AI techniques like Grad-CAM and SHAP are incorporated. These models could be integrated in real time into computer-aided diagnostic (CAD) systems to help physicians make decisions and arrange treatments. Furthermore, by integrating CNNs, transformers, and graph-based models, it is possible to more accurately and comprehensively detect brain tumors in future medical applications by capturing both spatial and contextual data.

VI. CONCLUSION

The research for the appropriate classification of brain tumor MRI images into four categories (glioma, pituitary tumor, meningioma, and no tumor) shows the power of DL and transfer learning models. The four architectures assessed were a custom Ensemble CNN, Vision Transformer (ViT), MobileNetV2, and VGG16. The Ensemble CNN’s robustness and generalisability were confirmed by its 96.72 % test accuracy with balanced metrics. With a test accuracy of 96.72 %

ACKNOWLEDGMENTS

The authors sincerely appreciate the referees’ positive feedback on the revised version of the manuscript.

REFERENCES

- [1] R. Hashemzahi, S. J. Mahdavi, M. Kheirabadi, and S. R. Kamel, “Detection of brain tumors from MRI images based on deep learning using hybrid model CNN and NADE,” *Biocybernetics and Biomedical Engineering*, vol. 40, no. 3, pp. 1225–1232, 2020.

- [2] Z. U. Nisa, S. M. Bhatti, A. Jaffar, T. Mazhar, T. Shahzad, Y. Y. Ghadi, A. Almogren, and H. Hamam, "Beyond accuracy: Evaluating certainty of AI models for brain tumour detection," *Computers in Biology and Medicine*, vol. 193, Art. no. 110375, Jul. 2025.
- [3] J. Yang, M. A. Siddique, H. Ullah, G. Gilanie, L. Y. Por, S. Alshathri, W. El-Shafai, H. Aldossary, and T. R. Gadekallu, "BrainCNN: Automated brain tumor grading from magnetic resonance images using a convolutional neural network-based customized model," *SLAS Technology*, Art. no. 100334, Jul. 2025.
- [4] S. Sinha, S. Mali, A. K. Pathak, and S. Rajkhowa, "BRAIN-SCN-PRO: A machine learning model for the classification of brain tumors using a convolutional neural network architecture," *Biomedical Signal Processing and Control*, vol. 108, Art. no. 107942, Oct. 2025.
- [5] S. Yoon, "Brain tumor classification using a hybrid ensemble of Exception and parallel deep CNN models," *Informatics in Medicine Unlocked*, vol. 54, Art. no. 101629, Jan. 2025.
- [6] B. M. Jebin, S. I. Shyla, K. N. Bel, and C. J. Sheela, "Brain tumor detection and classification using U-Net and CNN with brain texture pattern analysis," *Biomedical Signal Processing and Control*, vol. 110, Art. no. 108156, Dec. 2025.
- [7] M. C. Binish, S. R. R. S. R., and V. Thomas, "CBAM-SMK: Integrating Convolution Block Attention Module with separable multi-resolution kernels in deep neural networks for brain tumor classification," *Biomedical Signal Processing and Control*, vol. 112, Art. no. 108483, Feb. 2026.
- [8] U. Amjad, A. Raza, M. Fahad, D. Farid, A. Akhunzada, M. Abubakar, and H. Beenish, "Context aware machine learning techniques for brain tumor classification and detection—A review," *Heliyon*, vol. 11, no. 2, Jan. 2025.
- [9] R. Rekha, P. Shruti, M. Deekshitha, and J. Akash, "DCSwin-UNet: Dual Encoder U-Net based on CNN and Swin Transformer with Trainable Multiplication Layer for brain tumor segmentation from MRI images," *Biomedical Signal Processing and Control*, vol. 110, Art. no. 108325, Dec. 2025.
- [10] B. B. Gupta, A. Gaurav, and V. Arya, "Deep CNN based brain tumor detection in intelligent systems," *International Journal of Intelligent Networks*, vol. 5, pp. 30–37, 2024.