SSRG International Journal of Electrical and Electronics Engineering
ISSN: 2348-8379/ https://doi.org/10.14445/23488379/IJEEE-V12I3P103

Volume 12 Issue 3, 21-35, March 2025

Original Article

Brain Tumor Classification of MRI Dataset Using Ensemble Learning with EfficientNetV2 and ViT-B16

Bora Pavani[1], Raghunath Mandipudi[2], Siva Suresh Kumar Gandi[3], Pothina Mohnish Sabarinath[4], Yada Sunitha[5]

[1]*Department of CSE, Vignan's Institute of Information Technology, Vishakhapatnam, Andhra Pradesh, India.*
[2]*Department of ECE, Adityha College of Engineering & Technology, Surampalem, Andhra Pradesh, India.*
[3]*Department of ECE, Vignan's Institute of Engineering for Women (A), Vishakhapatnam, Andhra Pradesh, India.*
[4]*Department of Life Sciences, GITAM (Deemed to be University), Vishakhapatnam, Andhra Pradesh, India.*
[5]*Department of CSE, CVR College of Engineering, Telangana, India.*

[1]*Corresponding Author : pavani.uniquecse@gmail.com*

Received: 21 November 2024 Revised: 08 February 2025 Accepted: 25 February 2025 Published: 31 March 2025

Abstract *- Unidentified and untreated brain cancer can be fatal. Radiologists routinely use pictures from MRIs and CT scans to make early diagnoses of brain disorders. Assessing the border of the brain tumor on MRI scans and figuring out its potential pathology are crucial stages in catching this dangerous condition early on. We categorize and segment brain tumors based on features such as consistency, uneven borders, and arrangement. Disparities between observers and such substantial deviations can lead to serious issues during neurosurgical procedures. On the other hand, it can be difficult at low-income medical facilities to not have radiologists to review medical images. Machine learning-based automatic analysis of medical images may be able to help with diagnosis in order to solve the problem at hand. The importance of Magnetic Resonance Imaging (MRI) in detecting and managing brain malignancies has increased exponentially. Given the complexity and diversity of tumor features, accurately classifying brain tumors from MRI images remains a challenging task. This article talks about how EfficientNet V2 and ViT-B16-powered ensemble models can be used to sort tumor cells into different groups. The Geometric Average Ensemble Model that was created was 95% accurate compared to other implementations. It was trained on data from 700 MRI images of brain tumors and then tested on 281 images. The study's results show a clearer enhancement in image classification of brain tumors than in previous studies.*

Keywords *- Brain cancer, Deep Learning, ViT-B16, Ensemble model, Machine Learning, EfficientNet V2, Brain tumor classification.*

1. Introduction

Diagnosing brain tumors has become complicated in the progression of medicine. Accurate and early identification of various brain tumor phenotypes has implications for treatment progression and improved long-term outcomes for brain tumor patients. The brain is known to be the body's most important part. It takes care of all the voluntary and involuntary actions. Brain tumors form when the fundamental process of cell division goes awry. Brain tumors can be benign or malignant. 80% of all brain tumors are benign, and 20% are malignant. To do this, researchers from a related study looked into how to combine clinical data, like a patient's demographics and medical history, into a model that could learn and help with personalized brain tumor diagnosis and treatment. A good way of doing this is by implementing hybrid models that integrate multiple other models or approaches to gain more performance. One study found that using multiple pre-trained VGG-16 Convolutional Neural Networks (CNNs) to classify three types of brain damage improved accuracy compared to a single model. They also explored the potential for more accurate classification of brain tumor types using longitudinal data, including image data collection at specific patient time points. Moreover, integrated models, clinical data, and longitudinal data should contribute to creating even more robust self-diagnostic tools while current brain tumor classification systems evolve. Brain tumors are an increasingly important health problem, and accurate diagnosis and grading are crucial for successful therapy. The standard way of classifying brain tumors - for example, requiring doctors to manually review MRI scans - is generally laborious and misleading. To overcome this challenge, researchers have investigated the application of hybrid models that integrate several machine-learning algorithms to achieve better accuracy and robustness in the classification of brain tumors. New methods are being developed to detect brain tumors more precisely. Many of these techniques rely on machine learning and artificial intelligence. A class of brain malformations can be identified by the field lines of their magnetic moment; in which researchers have started to explore the possibility of using a triple contrast study, which compares similar and contrast images of a brain MRI, to learn greater discrimination in tumor classification. An innovative deep learning-based model specifically designed for the classification of multi-modal data, including but not limited to MRI, genomics, and patient populations, is shown to enhance accuracy. Deep Neural Networks Show Promise for Predicting Adult and Pediatric Brain Tumors. Further studies always reveal that deep neural networks are capable of reaching comparability to doctors' accuracy in diagnosing brain disorders based on histopathological images. Using the power of Artificial Intelligence (AI), it creates intelligent algorithms to classify tumors based on their molecular profile. Triplet Contrast Learning encodes the idea of a more detailed diagnosis and enables a faster and more accurate solution. In particular, the study demonstrates the potential for the system to reduce misclassification, including tumor differentiation.

The main advantage of this field is the use of artificial neural networks that simulate the structure and function of the human brain. Our main focus is image analysis, especially in the field of computer vision, where machine intelligence relies on visual information. Convolutional neural networks are the best networks in this field because they are good at analyzing images such as matrix patterns. Recent advances in deep learning have led to the development of many models, such as hybrid methods, that have proven useful in brain tumor studies. Tuned for brain tumor dataset to improve performance. Additionally, new techniques such as triple contrast studies have been proposed to use data structures to learn more discriminatory methods for classifying brain tumors. Provide information and guide physicians to develop effective treatment strategies. Automated segmentation of brain tumors in Magnetic Resonance Imaging (MRI) data using deep learning, such as the U-Net model, provides good results in identifying tumors, which is an important step in treatment planning. -Trained CNNs such as VGG-16 with feature extraction techniques such as Grey-Scale Co-occurrence Matrix (GLCM) have also been trained for brain tumor classification.

Additionally, applying diagnostic products such as YOLO has demonstrated the ability to identify and describe brain tumors accurately, thus enabling computer-assisted diagnosis. Evaluating EfficientNetV2 with EfficientNetV1 is a superior approach. It admits that a lot of length metrics don't strike a balance between training speed and performance. Instead of utilizing its counterpart, EfficientNetV2 employs a non-parallel approach to assess the most crucial extra layers in later network stages. This method aids in preserving the harmony between estimated expenses and precise outcomes.

To reduce both memory-eager storage and training latency associated with larger image size, the image size scaling technique is limited to lower resolution levels in EfficientNetV2. Due to its transformer architecture, ViT B16 is capable of being highly performant, recording data in global contexts, and scaling. This pre-training on large datasets facilitates transfer learning. Though this approach has a high computation cost, it presents a good alternative to traditional CNNs in cases where finer precision and global perspective are needed. Researchers are incorporating deep learning and other ensemble modeling techniques to automate brain cancer identification, which will eliminate the errors in manually inspecting the image samples.

For instance, triplet contrastive learning can utilize neural networks to minimize/maximize the distance between tumor samples based on if they belong to the same (or different) tumor sub-type. Deep learning models can also benefit from integrating clinical data. Ensemble models are a powerful method in machine learning that significantly generate different classification results from the predictions of multiple individual (base) models, particularly in the area of brain tumor classification. This phenomenon contributes to the reduction of model overfitting and bias caused by using a single model, and these methods can be performed with multiple types of variables and datasets. These models have advantages such as improved effectiveness, efficiency, and precision, which can help the global categorization of brain tumors based on clinical data. Integrated technologies allow researchers and clinicians to make more accurate diagnoses and more effective treatment plans.

Whether for early detection, treatment, or management of several brain tumors, the classification of brain tumors is an essential task in diagnosis and treatment. Conventional approaches to brain tumor classification, such as having skilled doctors manually analyze MRI scans, can be laborious, prone to mistakes, and heavily reliant on individual experience. Classifying brain tumors using deep learning and machine learning. This method has demonstrated excellent performance in correctly identifying and categorizing various brain cancers.

When examining the efficacy of CNNs, our goal is to determine whether it can produce a technique that will aid in brain classification and diagnosis. It will be challenging to develop a model that can identify different types of tumors after first differentiating between tumor and non-tumor cells. In machine learning, models carry out tasks. The accuracy, training duration, and resource requirements of the aforementioned model are some variables that impact its success. The general architecture of the proposed work is represented in Figure 1. This article is structured in the following manner: The approach of the study is described in Section 2. In Section 3, a relevant work is examined, and a brief synopsis of the research is provided. The materials and procedures part is included in section 4. The computational analysis and results are presented in Section 5. Section 6 concludes the study and provides directions for future research.

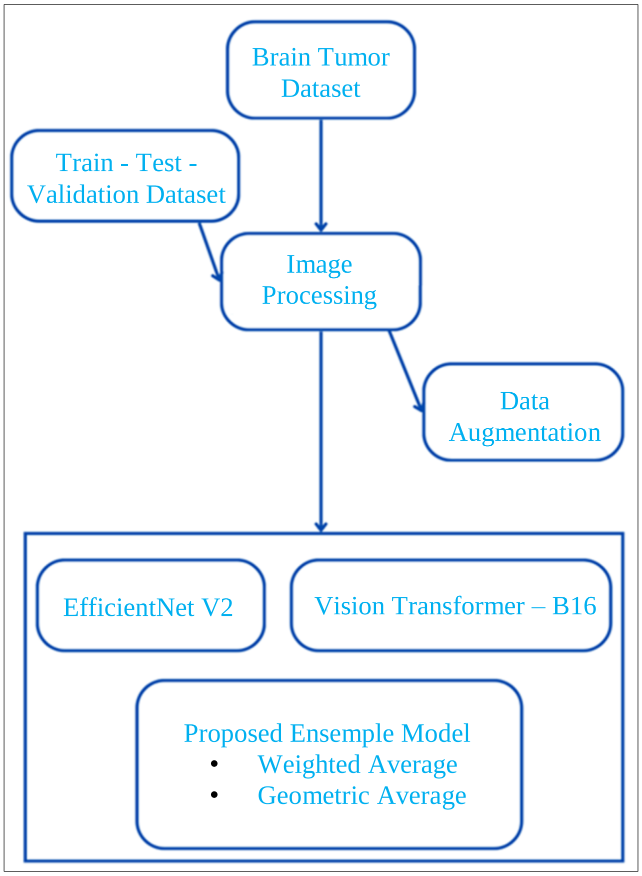

Fig. 1 Architecture overview of proposed work

2. Research Methodology

Machine learning algorithms for classifying brain tumor involve a systematic literature review, as depicted in Figure 2, dataset collection, image preprocessing, transfer learning and Ensemble techniques. The dataset used is the Kaggle Brain Tumor MRI Dataset. The chosen model, such as Transfer Learning and Ensemble Models, is trained to identify features associated with brain tumor. The algorithm's effectiveness is assessed by varying its parameters. Nevertheless, limitations, including dataset class conflicts, computational limitations, unique representation requirements, and data registration, impede its application. Cooperation between radiologists, oncologists, computer scientists, and other stakeholders is essential to address these obstacles.

3. Related Work

The paper [1] evaluated and compared ten advanced deep-learning models using an unbalanced dataset for three distinct types of brain tumours. The experimental results show that the Inception model outperforms all the other models for classifying the three groups. Unlike the other models, the performance of the EfficientNet model is relatively weak. In addition, this work provides deeper insight into the most common deep neural networks in application to MRI datasets to classify brain tumors.

Three approaches were proposed in this study [2]. The first approach is preprocessing the original images to improve their quality. The authors propose an alternative method to extract the preprocessed image's anti-noise interference capacity and image resolution. The third approach classifies the entry images into either tumors or non-tumors. Finally, the authors applied optimization to ensure they had the right categorization. Compared with existing approaches, the proposed method achieves accuracy improvement of 0.14, 0.09, and 0.08 for tumors and 0.09, 0.14, and 0.10 for non-tumors on MRI brain imaging datasets.

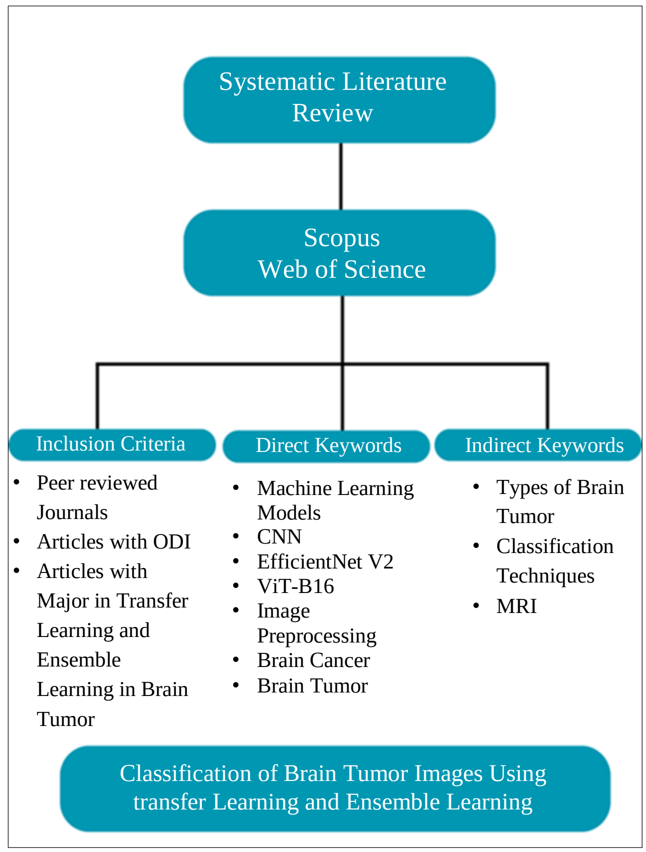

Fig. 2 Systematic literature review of proposed work

The usefulness of adaptive learning architectures for brain tumour diagnosis is assessed in this study [3]. This study employed six distinct learning algorithms: ResNet-50, MobileNet, VGG16, Inception-V3, DenseNet-121, and EfficientNet V2-M. Validation of the model utilizing publicly available MRI datasets and comparison with analogous research. Reliable data, like random rotation, is included in the data set to address the issue of the form of the clusters' inconsistency. Based on experiments, the EfficientNetV2-M model performs 98.01% more accurately than other models. Furthermore, by integrating the output of many models, this research seeks to develop new models with more sophisticated inference and generalization capabilities. It was designed in this context to combine the EfficientNetV2-M architecture with the DenseNet-121 and Inception-V3 architectures. The accuracy of the EfficientNetV2-M + Inception-V3 model with combining approaches is 0.98. It has been demonstrated that the tandem model performs better than the state-of-the-art in enhancing patient outcomes and medical imaging technologies. Important features for the classification models in the research work are extracted using deep learning techniques like Inception V3 and DenseNet201 [4]. To increase classification accuracy, the utility incorporates pre-classification in addition to the features gathered by deep learning models. Four groups of features were created using the suggested approach. The system assesses the categorization system's performance using two data metrics in addition to other performance measures. The dataset was outperformed by all three models. Furthermore, using dataset 1, four different optimization strategies that were already in use were compared, and using dataset 2, five different approaches were assessed independently. During training and testing, the accuracy of deep learning processes was enhanced by the suggested model.

The primary objective of this work [5] is to develop a sequential brain tumor that uses fully Convolutional neural networks and deep learning to identify and categorize patterns. Two steps make up sample preparation: first, separating the neoplastic from the non-neoplastic brain, and then figuring out what kind of tumor it is. The MRI dataset for brain tumours was used to examine two models. To get the best results, four optimizers were trained for three different classification tasks. The most recent deep modeling-based brain tumour classification and segmentation methods are applied to address current issues [6]. Initially, images are sourced from the internet. The newly suggested segmentation model in the paper is then applied to the gathered data. The updated algorithm in the work uses IMOA to optimize the deep learning method's limitations. Lastly, the collaborative mesh network handles tumour classification.

As a result, high-level estimates decide the ultimate outcomes. Different measurements and methods are used for performance and comparison. 51 related works were included in this paper's systematic review and evaluation [7]. In order to determine the features and benefits, the authors carefully gathered data related to dual diagnoses and produced a table. For glioblastoma-lymphoma and mild glioma compared to chronic glioma, the pooled area under the curve was 89% and 99%, respectively. For benign and malignant tumours, the total sensitivity and specificity were 90% and 93%, respectively. When comparing mild gliomas to chronic gliomas, the overall sensitivity and specificity were 99% and 94%, respectively. Tumour and metastatic tumour differentiation had a sensitivity and specificity of 89% and 87%, respectively. Gliomas had the highest disagreement among pituitary tumours among brain tumour classifications, with a sensitivity of 99% and a specificity of 99%. The authors of this work [8] used a multi-center brain tumour dataset that includes comments on division and classification as well as several types of brain tumour datasets. They suggest a contemporary division technique to advance the accuracy of brain division. This method uses a crossover thick expanded convolution module and a dispersion-weighted expanded convolution module to gather various information while using fewer parameters. Employing a variety of importance of care and quality of care models to effectively gather data pertaining to jobs. The authors presented a new computer-assisted symptomatic calculation that combines the suggested approach with an updated therapy calculation for brain cancer. This approach uses division covers as additional channel highlights in addition to advancing division accuracy to treat brain tumors.

The authors of this paper [9] suggest a care-based multi-residue CNN for the categorization of brain tumours in various spaces in order to address the current issues. The authors suggest using a lightweight residual multiscale CNN to capture high-level feature representation of various receptive fields in more detail. Furthermore, a model for choosing several discriminating themes is suggested. The suggested Classification Model, which has been demonstrated to capture broad patterns, is layered with the proposed model. Based on the outcomes of two benchmark data experiments, the suggested model performs better than current CNN architectures and techniques. On both the datasets used, the suggested model's accuracy is 0.97. Radiologists frequently diagnose brain tumours early on using MRI and CT scan imaging. However, a shortage of radiologists to evaluate medical pictures can be an issue in underdeveloped healthcare institutions. Deep learning-based automatic analysis of medical images may be able to help with diagnosis and solve this issue. Conventional approaches frequently concentrate on developing specialized algorithms that deal with a single issue, such as categorization or brain function. This work [10] proposes a new multitask network that integrates several U-Nets in sequence to replace the conventional VGG16. This network can do segmentation, distribution, and localization using the same model all at once.

The authors use the brain tumour segmentation dataset and the brain tumour MRI dataset to train and branch segmentation, respectively. Four different forms of brain tumours can be concurrently identified, segmented, and localized in MRI scans by combining the information from the three approaches. With a Dice coefficient of 0.86 and a classification rate of 0.97, the multitasking strategy performed well. Furthermore, compared to other techniques, this one exhibits improved computing efficiency. Our approach might be a useful tool for diagnosing illnesses that radiologists are unable to identify in hospitals with limited resources. Unchecked cell proliferation is known as brain cancer, and it is a global health concern. However, due to variations in tumour size, form, and location, as well as constraints in clinical practice, classifying brain tumours using computer systems has proven challenging. Raising the accuracy of the diagnosis of brain tumours is crucial since even an insignificant error in a person's decision-making might result in an increased mortality rate. A novel strategy that harnesses learning to enhance the early identification and decision-making of major mental illnesses is presented in this work [11]. Background images of tumour cells were gathered from clinical data, preprocessed, amplified using various communication approaches, and categorized using an adaptive hierarchical ResNet optimized for Border Collie Optimisation. The DeepLabV3 model was used to classify abnormal photos before being fed into the suggested model for ultimate classification. In brain tumour classification and weight prediction, the suggested model performs better than existing heuristic classification techniques, with gains ranging from 1.3% to 4.4% above other models tested in this investigation.

Because Feature Selection (FS) enhances classification and lessens information bias, it is a crucial stage in processing images relevant to radiology. However, the advancement of radiology-based brain tumour research is constrained by the absence of universal techniques. The authors of this study [12] provide a selection method based on 3-factor Cascade Selection and discuss the features of FS approaches utilized in related research in order to address these issues. The suggested approach splits FS into two phases. Initially, it selects less informative and task-relevant traits based on genetic information. Recursive feature extraction is then carried out to find the matching features with the most effective classification function. The authors tested seven files with thirteen different brain tumour classifications in order to determine whether the suggested model is easily extrapolated. They then used a five-point evaluation method to assess the overall performance. The proposed model has proved to be overall effective in all tasks. It is faster, a little bit simpler, more flexible, safe, and has a nicer appearance in the world than the 13 procedures. Our research illustrates how applying multifactorial approaches might enhance FS performance and generate fresh concepts for future advancements.

The primary results demonstrate that the suggested model [13] performs optimally, obtaining an astounding 0.96 classification accuracy and a seven percentage point reduction in false positives on three thousand sample data points. From the comparison with the other models, the authors infer that the proposed model is the most efficient and accurate and has the least misclassified records. This model's robustness and scalabilities are confirmed by the leading feature effects from the figure share dataset, as well as the model precision and hyperparameter values. The proposed model has emerged as a novel and effective solution to support the classification of medical images. The outcomes of the investigation confirm that the proposed model is an ideal tool for medical image review and provides a new solution for long-sought tools that can further explore the potential use of hyperspectral imaging. This study [14] aims to look into CNNs' potential for classifying brain MR images, which is crucial for quickly identifying brain tumours and expediting the course of therapy. The goal of many researchers is to develop neural networks that are more precise and less complicated. In this case, the authors reduced the number of filters in our convolution layer to 4, 8, and 16 and based the data on the single-pass CNN model. under the CNN framework. The adoption of the Support Vector Machine (SVM) classifier in place of the conventional SoftMax classifier yielded the greatest gains in comprehension and communication. When comparing the suggested model to the research of other researchers, it can be seen that the accuracy of classifying brain tumours in the data increases by more than 0.1 when the complexity of the neural network model is decreased. In this work, two MRI documents are used to implement the suggested two-way CNN with the final SVM classifier network to validate our approach's efficacy and performance.

The accuracy of the proposed model is 0.98 for primary data, 0.98 for secondary data, and 0.99 for primary and secondary data combined. Because tumour features are diverse and varied, classifying brain images from MRI scans is still a difficult pursuit. The categorization of tumour cells using support vector machines is introduced in this article [15]. Their proposal involved a novel method specifically for segmentation, noise reduction, and feature extraction. In the end, a support vector machine was used for the classification stage; a total of 24 MRI images were used with a full 0.999 accuracy, including 11 benign and 13 malignant brain tumours for training and 16 images for testing. Recent years have seen the proposal of several automated techniques for brain segmentation or classification to address issues with MRI scans; nevertheless, no clever technique for recognizing the different types of tumours in MRI pictures has been put up yet. Here, the authors describe a novel automated tumour classification model [16] that combines a decision maker for segmentation, a decoder for three distinct categories of brain tumours, and global transformation for particular representation in brain MR imaging.

The suggested approach was examined using datasets pertaining to brain tumours, and the outcomes were assessed in a number of projects and solitary investigations. With an accuracy rate of 0.97 needed for every assignment, the multiple learning model significantly improved the simultaneous categorization and segmentation of brain tumours. Two MRI records with over 3000 images of three different types of tumours and non-tumor images served as the basis for this study [17]. The authors find the optimal network hyperparameters using training and validation sets. The model's performance is assessed using various training and testing techniques, comprising correction, transfer learning, augmentation of data, and training from scratch. The models' accuracy is compared, and their complexity in terms of network resources, training time, and prior images have been investigated by the authors. Many networks achieved the accuracy of two datasets; the best model achieved 98.7% accuracy. This is on par with research work standards. The paper's authors [18] stress how essential it is to develop measures that satisfy system objectives and lessen unfavourable effects that can taint data or give people undue influence over their motivations. They offer a comprehensive procedure for setting measures that are utilized for important needs, error prevention, and design considerations. Indicator design solutions are given in this article, along with examples of indicators that don't function in different domains. The contributors stress the significance of comprehending the relationship between knowledge and goals and the necessity of consistency and clarity when there is a conflict between objectives.

The paper [26] proposes the EfficientNet, a light-tuned Convolutional neural network, for brain tumour detection. The authors of this research suggest splitting brain growth into three different files with different distributions using a variable metric. The effectiveness of the proposed model is assessed through performance evaluation, and the results are contrasted with those acquired by state-of-the-art methods. The average scores obtained were 0.98 for all performance criteria. At first, a novel DTA technique [28] was presented to find tumours in healthy people's MR imagery. DTA is specially engineered and optimized to provide a highly powerful feature space fed into a collection of Machine Learning (ML) classifiers. The second stage of the new brain classification method uses machine learning to categorize various tumour kinds using a hybrid feature that combines static and dynamic information. The ability of the suggested algorithm to recognize heteromorphism and the various behaviours of various tumours is a positive indicator. A variety of segmentation techniques are available for diagnosing medical photos. Image features, including the capacity to discern between similarities and differences, are frequently employed in segmentation techniques.

3.1. Research Gaps

Although there is a strong body of literature to date in the area of automated brain tumor classification, there are several specific gaps in this literature. This includes the failure of strong generalization across datasets, the influence of time criticality on computational constraints, and the scarcity of work on multi-class tumor classification. Recent models are trained on limited datasets, resulting in overfitting and, thus, inadequate generalization behaviour on new MRI scans. Unlike previous studies reporting only one class of tumor (tumor versus non-tumor), this study is focused on differential classification among different tumor types.

4. Materials and Methods

The goal of this work is to examine approaches to classify different forms of brain tumors in patients using ensemble learning models, specifically with EfficientNet V2 B0(Model#1) and ViT-B16(Model#2). This work also includes assessing model performances following training and the selection approach. Technological advancements in medical imaging have revolutionized brain tumor diagnosis and treatment, enabling precise information acquisition and personalized treatment planning through automated segmentation and analysis of MRI data.

4.1. Dataset

This dataset is a private compilation of T1 and T2 magnetic resonance images and contrast-enhanced T1 and T2 imagery sorted by type of brain tumor. Radiologists interpreted the images, which were obtained without any marking or patient identification, and made them available for research. Meningiomas, which are usually benign tumors arising from the membranes surrounding the brain and spinal cord; meningiomas, which are the most aggressive and common type of malignant brain tumor; and metastatic brain cancers, are among the several tumor types that are divided by the images provided. The number of images in the sample is 1253. The dataset provided has 39 classifications. By utilizing cutting-edge deep learning techniques, significant progress has been made in the creation of automated methods for the segmentation and classification of brain tumors.

Astrocitoma T1	Astrocitoma T1C+	Astrocitoma T2
Ependimoma T1C+	Ependimoma T2	Ganglioglioma T1
Ganglioglioma T1C+	Ganglioglioma T2	Germinoma T2
Meduloblastoma T2	Meningioma T1	Meningioma T1C+
Meningioma T2	Neurocitoma T1	Neurocitoma T1C+
Neurocitoma T2	Papiloma T1	Papiloma T1C+
Papiloma T2	Schwannoma T1	Schwannoma T1C+
Schwannoma T2	_NORMAL T1	_NORMAL T2

Fig. 3 The classes in the brain tumor dataset considered for training

Even though the dataset is large, there are classes with small sample sizes that might not have enough data for reliable model testing and training. In order to ensure consistent outcomes and concentrate on the most common forms of tumors, classes with no more than ten samples were decided to be excluded from the analysis. Following the application of this filter, the final dataset consisted of 24 classes encompassing 1168 multi-parametric MRI scans that included T2-weighted, T1-weighted, and contrast-enhanced T1-weighted images. The list of these classes is illustrated in Figure 3.

4.2. Data Preprocessing

In the fields of machine learning and data analysis, data preprocessing is essential since it makes sure the data is prepared for further processing and analysis. This is especially crucial when it comes to the identification of brain tumors, as the performance of the prediction models can be greatly impacted by the nature and quality of the input data.

We will construct an input data pipeline using TensorFlow to manage the loading and passing of the picture data to the model to train models with this dataset. We'll ensure the training data is batched and prefetched while the model is training on a previously passed sample to obtain a faster training time. Three dataset sets are created: train, test, and validation. The list in each of them is displayed in Figure 4.

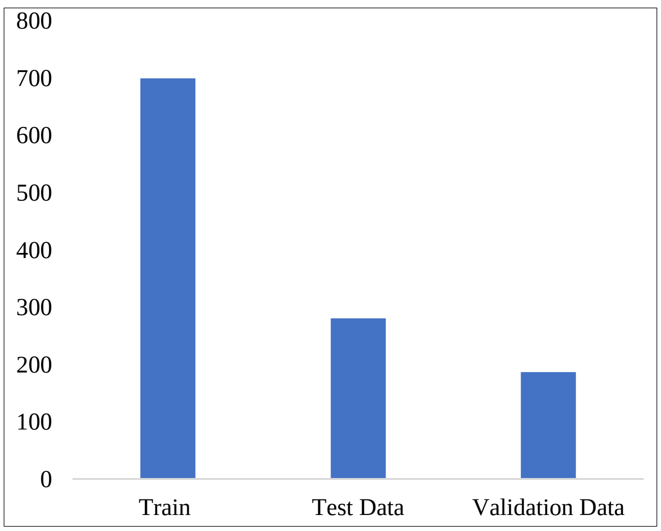

Fig. 4 The dataset details

4.3. Data Augmentation

Data augmentation, or creating more synthetic data samples to increase the training dataset, is a crucial component of data preprocessing in this field. This can be extremely helpful in medical imaging activities where labeled data availability is frequently limited. Convolutional Neural Networks (CNNs) have been shown to perform better on medical imaging classification tasks when differentiating between different types of data augmentation. Research has specifically demonstrated that model performance is determined by how well an augmented training set preserves the characteristics of the original medical images. The input image and the augmented image are shown in Figure 5.

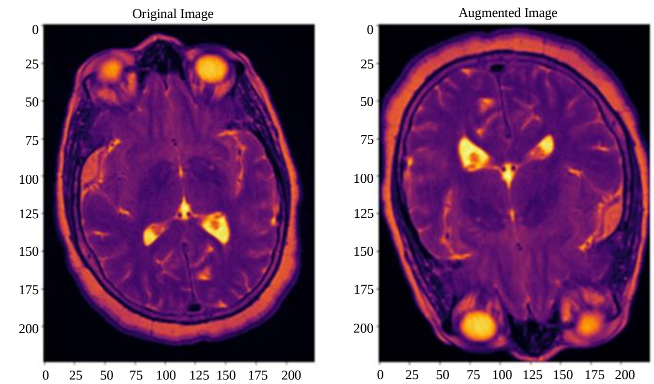

Fig. 5 The augmented image of input image

4.4. EfficientNetV2

Work designs have been a key factor in the deep learning landscape's ongoing evolution, propelling advancement in a number of domains, such as computer vision and natural language processing. The advent of the EfficientNet architecture, which has shown outstanding performance and efficiency, is one such ground-breaking development. A methodical approach to optimizing neural network architecture design choices for object detection is demonstrated by the EfficientNet family of models. Of particular importance is the suggested weighted bi-directional feature pyramid network; this facilitates smooth and effective multiscale feature fusion. Moreover, the compound scaling technique concurrently and consistently scales the backbone, feature network, and prediction networks' resolution, depth, and width.

The principle behind the compound scaling approach is to scale with a constant ratio in order to balance the width, depth, and resolution parameters. The mathematical method is demonstrated by the equations below.

$$\mathrm{Depth(d)} = \alpha^{\phi} \tag{1}$$

$$\mathrm{Width(w)} = \beta^{\phi} \tag{2}$$

$$\mathrm{Resolution\ (r)} = \gamma^{\phi}, (1) \tag{3}$$

$$\mathrm{Suchthat}\ \alpha\beta^{\phi}\gamma^{\phi} \approx 2 \tag{4}$$

$$\mathrm{Where}\ \alpha, \beta, \gamma \geq 1$$

The EfficientNet family of object detectors is the product of these fundamental advancements, and it has continuously beaten earlier state-of-the-art models in terms of efficiency as determined by parameters and FLOPS. The goal of EfficiekntNet V2 is to increase speed and efficiency through architectural enhancements. As this work tackles the topic as a classification problem with many labels, we will use Categorical Crossentropy as the loss function to train this model. Only the positive class retains its term in the loss in the particular (and typical) situation of multi-class classification because the labels are one-hot vectors. The target vector consists of a single element that is not zero. After eliminating the summation's items that are 0 because of the target labels, we can write:

$$\text{Loss} = -\sum_{j=1}^{\text{output Size}} y_j \log p_j \quad (5)$$

Where,
y_j represents the true class label for all j's
p_j represents the probability for all j's

This work will employ the Adam optimizer with 0.001 as the learning rate in terms of the optimizer.

4.5. ViT-B16

For image recognition tasks like task recognition, object detection, and image categorization, ViT is utilized. ViT is based on the transformer architecture, which is used to convert text into level tokens and generate text embeddings in Natural Language Processing (NLP). However, ViT first enters the image into the patch, reminiscent of the NPL converter's word tokens. The graphs are produced by combining these "patches" with encoder manipulation. There are the following components to the transformer encoder block, Normalization of Layers, and Multi-Layer Perceptrons with Multi-head Attention (MLP).

Vision Transformer outperforms traditional Convolutional Neural Network (CNN) architectures in terms of performance while requiring a smaller pre-training consumption. However, Vision Transformers exhibit less bias, making the process of data augmentation and conversion more straightforward for smaller files. As we observe, CNNs typically do better on small datasets, but Transformers perform well on huge datasets. Utilize the Transformer and CNN architectures to complete the assignment as efficiently as possible. For tasks like object detection, where it's crucial to identify minute details, the hybrid invisible switch has shown to be effective.

4.6. The Proposed Ensemble Models

Ensemble modeling has developed into an impressive machine-learning technique that offers an effective way of raising a model's accuracy. Collaborative learning is based on the idea that a team of specialists can solve a problem more precisely than one expert working alone. Supporting this, benchmark studies have shown that hybrid models can outperform their single-based classifier counterparts.

The basic idea of the integrated model is to combine the forecasts of multiple modelers and utilize their strengths and weaknesses to return a better general result. For instance, it is particularly useful in domains such as medical imaging applications where one model cannot possibly capture all the complexities. The computation complexity of the attention mechanism prevents intrinsic transformer models from incorporating high-resolution inputs, the most fundamental requirement of complex tasks such as detection and segmentation. To address this, we proposed novel hybrid ensemble architectures leveraging the strengths of transformers and CNNs.

Due to hardware constraints, EfficientNetV2 suffers from potential flaws like slow training speed or large model size. Ensemble methods can usually do a better job of classifying things than any of the individual models because they train more than one base model and then combine the results. This simple approach of an integrated model should embed the combined estimate of several models weighted by consideration of strengths and weaknesses to find a better one. This approach proves particularly helpful in fields like remote sensing applications, where it would be impossible for a single model to account for all the intricacies. A combination approach, which trains multiple models and aggregates their output, can frequently outperform a single model in classification.The mathematical representation of the Ensemble Model is,

$$E_M(S_M, D_T, M_D, R_A, C_A) = S_M\left(\sum_1^n I_L(D_T, M_D, R_A, C_A)\right) \quad (6)$$

Where,
E_M - Ensemble Model
S_M - Strategy for integrating the models
D_T - The input dataset
M_D - Models considered for the study
R_A - Result Analysis
C_A - Computational Analysis
I_L - Individual Learning Models

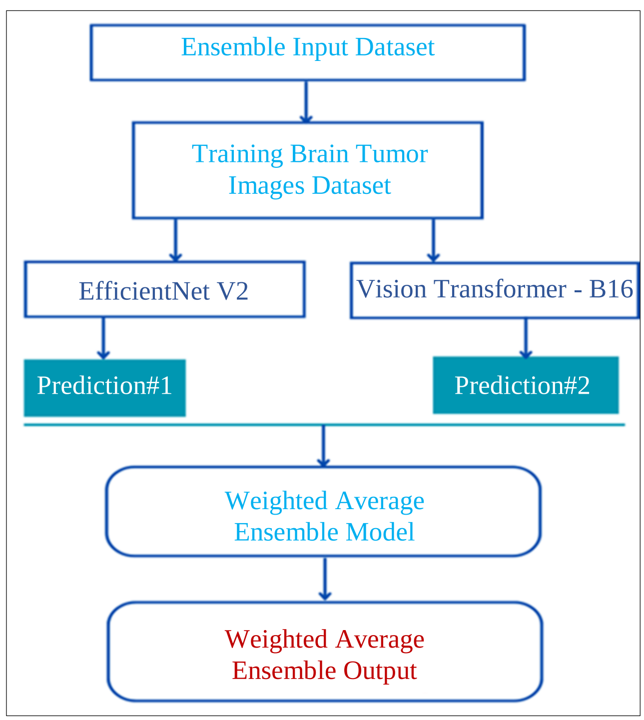

Fig. 6 Architecture overview of weighted average ensemble model

Ensemble Model#1- We adopt the Weighted Average of two individual prediction models, EfficientNet V2 and ViT-B16. The proposed model is depicted in Figure 6.

Ensemble Model#2 - We adopt the Geometric Average of two individual prediction models, EfficientNet V2 and ViT-B16.

This work additionally looks at methods to combine things to get the best results possible for brain tumor classification. The proposed model is depicted in Figure 7.

4.7. Novelty of Proposed Ensemble Models

The proposed Ensemble Models using EfficientNetV2 & ViT-B16 strives to provide several new contributions to the prediction of brain tumors.

1. Hybridization of CNNs and Transformers: CNN-based models traditionally show weak performance at capturing long-range dependencies in images. The Vision Transformer (ViT-B16) performs very well in this space, so we opted to combine it later with EfficientNetV2, aiming to balance out both computational efficiency and accuracy.
2. Transfer Learning Pipeline: To fine-tune EfficientNetV2 and ViT-B16 instead of re-training, MRI Brain Tumor images are used, which optimizes the extraction of images, leading to better detection with faster speed.
3. Improved Generalization Via Data Augmentation: The model is trained on a wide variety of data through an aggressive data augmentation strategy, contributing to its robustness and ability to accurately classify brain tumors with different MRI scans.

4.8. Benefits of Proposed Ensemble Models

The following are the benefits of using an ensemble model between EfficientNetV2 and ViT-B16:

1. Improved Performance: Combining models with different architectures can lead to an ensemble of models that performs better than any single model on its own, such as combining the transformer-based ViT-B16 with the efficient EfficientNetV2. The different capabilities of these models can work together and deliver forecasts that are more reliable and accurate.
2. Robustness and Generalization: We believe the diversity of architectural choices between ViT-B16 and EfficientNetV2 strengthens the ensemble model's ability to generalize. This model is expected to generalize well on unseen data and to be more robust to overfitting.
3. Trade-off Balance: The ViT-B16 model spends a lot of effort achieving good performance, and the EfficientNetV2 model focuses more on the efficiency of the network. Ensemble modeling gives you the advantage of creating an arbitrary number of high-performing outcomes without overburdening your computer power, thus creating a fair compromise.

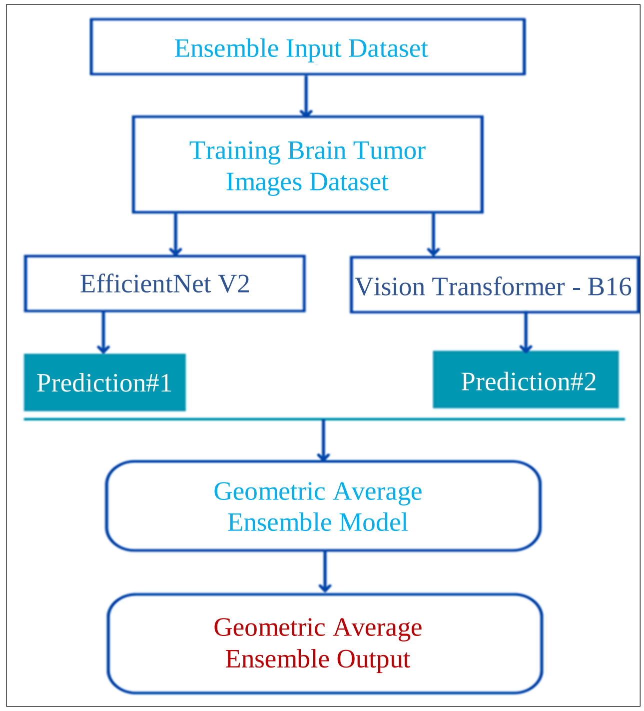

Fig. 7 Architecture overview of geometric average ensemble model

5. Computational Analysis

The experimental findings for the classification of brain tumors are shown in this section. We put the recommended strategy into practice using Python. When the results of the proposed model are averaged over several trials, they show a significant improvement over the current cutting models. Across a broad variety of indicators, the new model's results show a significant improvement. The results of two Ensemble models are described.

i. The model was trained using EfficientNet V2 and ViT B16 as stand-alone models.
ii. The model's training outcomes using the weighted average ensemble model with EfficientNet V2 and ViT B16
iii. The model's training outcomes using the geometric average ensemble model with EfficientNet V2 and ViT B16

The hybrid neural network integration process enhances the evaluation and classification processes, and all system outcomes and repercussions are enhanced in terms of the assessment metrics. The performance of the ensemble model was assessed and contrasted with the ViT B16 and EfficientNet V2 models. This model simulation's assessment indicators were assessed and contrasted. The hybrid neural network ensemble model yields superior outcomes with fewer errors compared to other models. The ensemble model is compared to currently used models, such as ViT B16 and EfficientNet V2. The features of the training and analysis processes for extraction and classification are improved by the hybrid EfficientNet V2 and ViT B16ensemble model.

5.1. Evaluation Metrics

We documented assessment metrics and assessed the models' performance on the test data to do this analysis. We'll apply several well-known classification metrics because this is a categorical classification task.

1. Classification Report
2. Accuracy Score
3. Precision
4. Recall
5. F1-score
6. Matthews Correlation Coefficient

5.2. Model Histories

The Training_loss and val_loss of the EfficientNet V2 model is as shown in Figure 8.

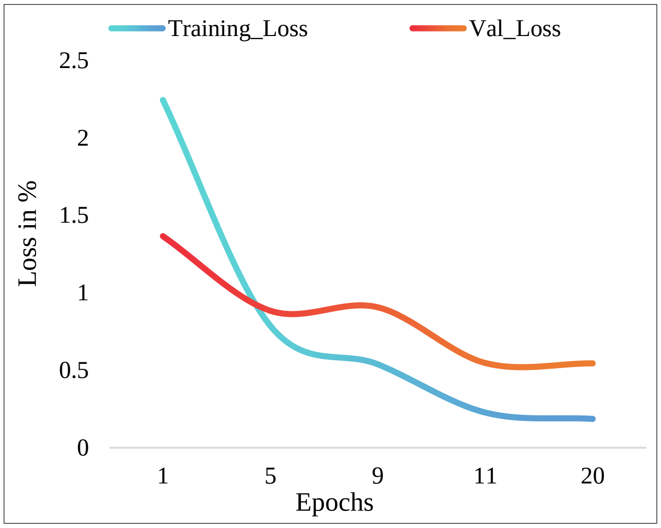

Fig. 8 The training_loss vs Val_loss with EfficientNet V2

The training_accuracy and val_accuracy of the EfficientNet V2 model is as shown in the Figure 9.

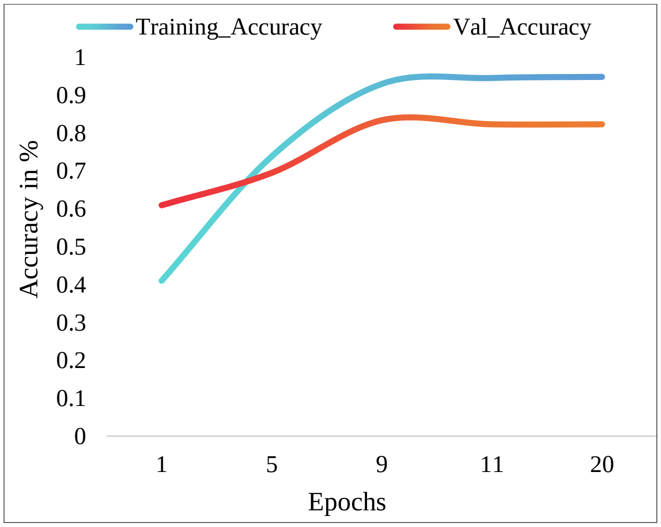

Fig. 9 The training_accuracy vs Val_accuracy with EfficientNet V2

The following observation can be drawn based on the results about loss and accuracy.

1. The minimum loss is with epoch 22
2. The highest accuracy is with epoch 22
3. The loss of the model is 37.98%
4. The accuracy of the model is 91.46%

The Training_loss and val_loss of the ViT-B16 model is as shown in Figure 10.

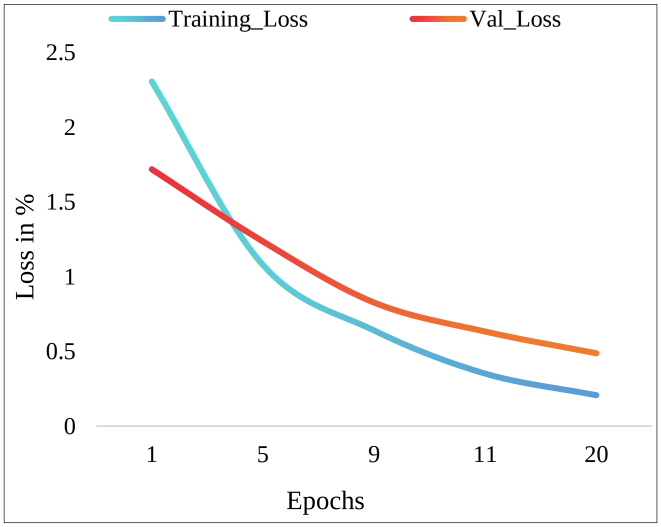

Fig. 10 The training_loss vs Val_loss with ViT-B16

The training_accuracy and val_accuracy of the ViT-B16 model are shown in Figure 11.

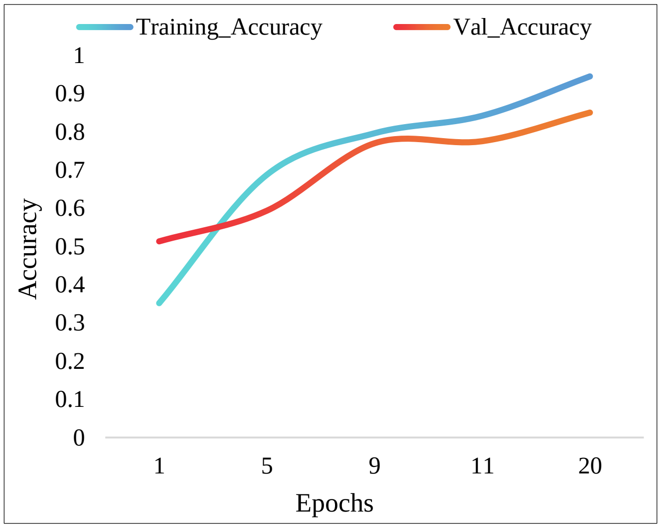

Fig. 11 The Training_accuracy vs Val_accuracy with ViT-B16

The following observation can be drawn based on the results of loss and accuracy.

1. The minimum loss is with epoch 16
2. The highest accuracy is with epoch 17
3. The loss of the model is 46.428%
4. The accuracy of the model is 85.05%

5.3. Comparative Analysis

This section presents a comparison between the proposed models and the existing conventional system. The performance of these models on diverse data sets, computational effectiveness, and task generalization will all be compared. When this model is compared to the previous way, the performance of this tool is superior.

As depicted in Figure 16. At 0.95, the results demonstrated a high level of accuracy. Compared to the current model, the ensemble model with geometric mean offers more accuracy. As a result, the suggested tactics outperform the current ones in terms of effectiveness. A number of designs have been developed as a result of the deep learning community's motivation for more accurate and efficient models, such as Vision Transformer (ViT-B16) and EfficientNet V2.

Furthermore, integration techniques like geometric mean integration and weighted average integration show potential for enhancing model performance. This article compares these models with their benefits, drawbacks, and real-world uses. A model of Convolutional neural network models called EfficientNet V2, whose confusion Matrix is represented in Figure 12, is intended to offer a trade-off between accuracy and efficiency, which makes them appropriate for a wide range of applications.

The V2 version of EfficientNet provides improvements in network scalability that reduce processing load and enhance performance compared to the previous version. On the other hand, the Vision Transformer (ViT-B16), a transformer-based architecture to deep learning (wherein the confusion matrix is shown in Figure 13), performs well in image classification tasks, but it has poor regularisation of neural network connection.

To combine the benefits of multiple models as an accuracy improvement strategy, the experiment of ensemble approaches being implemented, namely the Weighted Average Ensemble with its confusion matrix in Figure 14 and the Geometric Average Ensemble with its confusion matrix in Figure 15, is investigated. Unlike the Geometric Mean Ensemble, which computes the geometric mean of the predictions, the Weighted Average Ensemble utilizes predictions from different models and the weight for each model is assigned based on the performance.

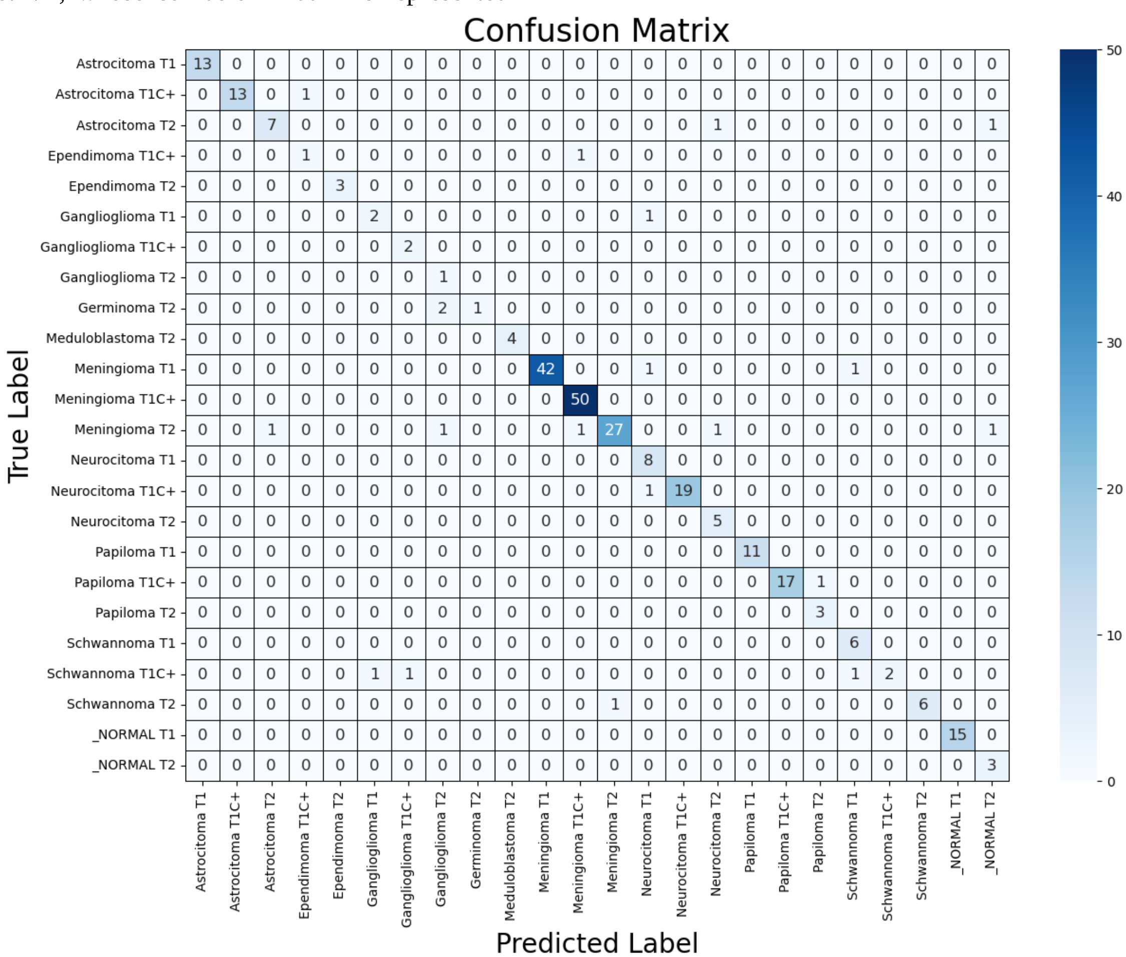

Fig. 12 Confusion matrix for EfficientNet V2 model

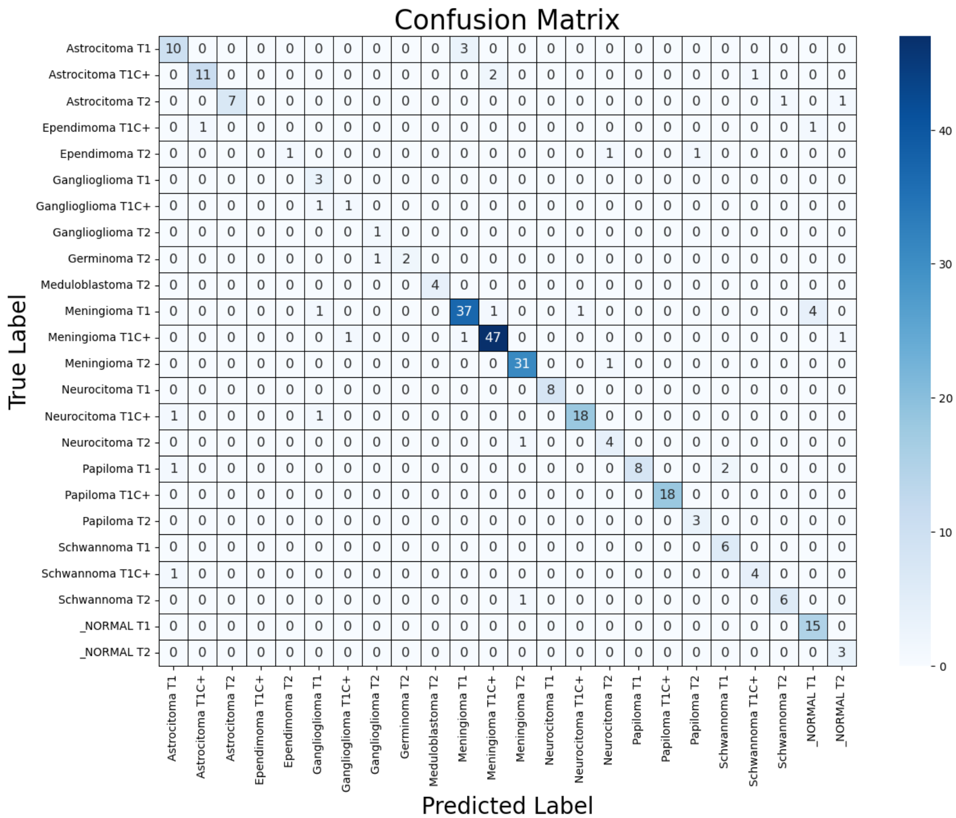

Fig. 13 Confusion matrix for ViT model

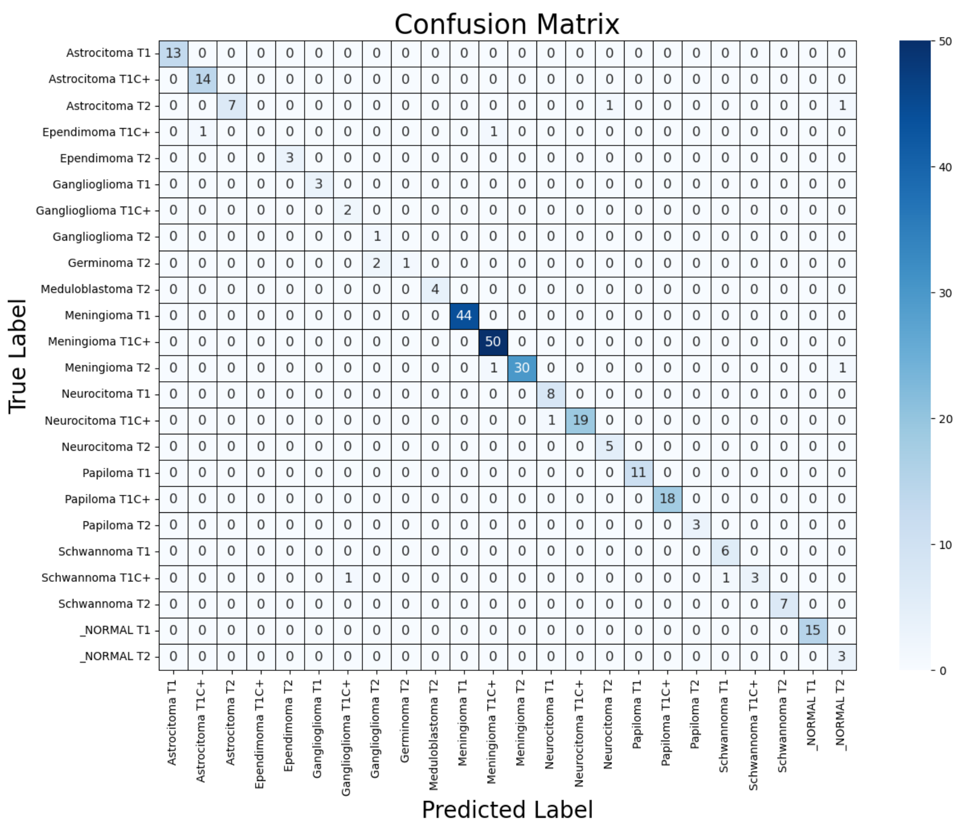

Fig. 14 Confusion matrix for weighted average ensemble model EfficientNet V2-ViT

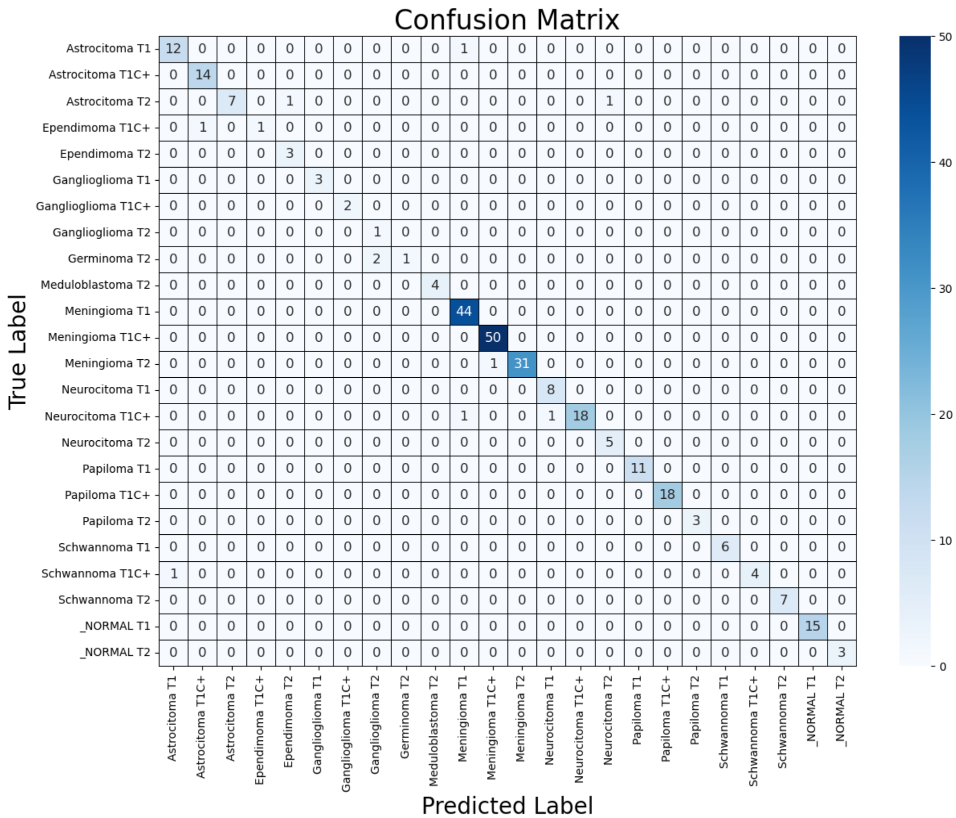

Fig. 15 Confusion matrix for geometric average ensemble model EfficientNet V2-ViT

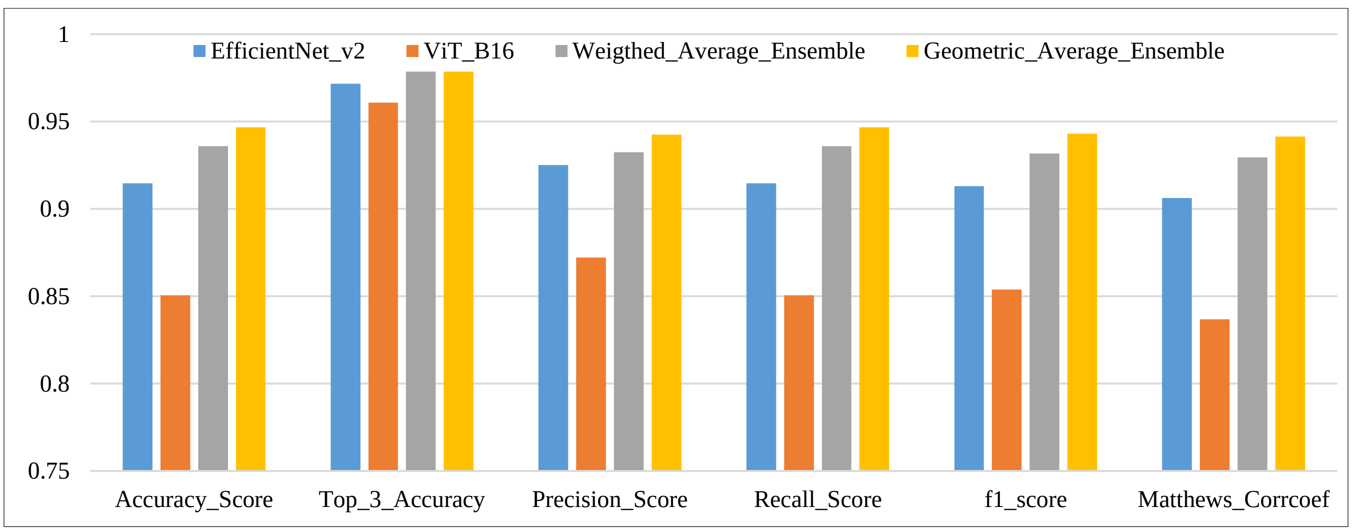

Fig. 16 The comparative analysis implemented models

Additionally, the knowledge gathered from this study will help practitioners and researchers choose the right architecture for the particular use cases they are working on. The EfficientNet model performed better than ViT-b16 when comparing performance measures. Regarding performance measures, the Geometric Average Ensemble outperformed the ViT-b16 and outperformed the Weighted Average Ensembles, EfficientNet, by a small margin. Previous studies have provided ample evidence of the advantages of using ensemble approaches. It’s interesting to note that a single huge model can perform worse than even a basic ensemble of identical structures. This can be explained by averaging the predictions from multiple models, which reduces the impact of noisy tree contributions. Additionally, it has been demonstrated that the testing accuracy increases monotonically with the size of the forest.

6. Conclusion

At the forefront of computer-aided diagnosis technology, deep learning algorithms are helping to quickly tackle the problems associated with brain tumor classification. This article demonstrates how conventional methods of diagnosing brain cancer cells are insignificant. The way a novel hybrid neural network assesses the tumor efficiently. The model was determined to be more effective after being assessed using a variety of performance metrics. This work suggests a sophisticated ensemble model combining ViT-B16 and EfficientNetV2 as a brain cancer classification technique. The CT scan images are used in this work.

The respective accuracy of the four implemented models-EfficientNet V2, ViT-B16, Weighted Average Ensemble, and Geometric Average Ensemble Models-is (0.91, 0.85, 0.93, and 0.95). The Geometric Average Ensemble Model is the most effective of these. Four implemented models-EfficientNet V2, ViT-B16, Weighted Average Ensemble, and Geometric Average Ensemble Models-were also examined for analysis based on their Top 3 Accuracy values. The accuracy of the four models is, respectively, 0.97, 0.96, 0.98, and 0.98. The Geometric Average Ensemble Model, in conjunction with the Weighted Average Method, is the best model out of these. The recommended transfer learning architecture and ensemble models demonstrate the best brain tumour classification abilities, which outperform the most recent models.

These advancements could result in better early detection of brain cancer, more efficient treatment planning, and, ultimately, better patient care. It is suggested that combining PET and CT scans enhances the diagnosis of brain cancer once it has been made. It has been demonstrated that early brain tumor diagnosis lowers the risk of brain cancer. We classified the depiction of tomography images. We present an efficient hybrid ensemble method for brain cancer classification.

It is envisaged that features from cancer datasets will be extracted by deep feature extraction. An ensemble technique has been devised to provide a robust detection model. The proposed architectures have shown increased generalization and robustness. This marks a major step forward in the field of deep learning-based brain tumor classification and opens the door to better, scalable, and practical diagnostic tools. Future work will increase the integration of multi-modal data (histopathology, genomics, radiology, clinical measures), enhanced capabilities for explainability and interpretationability, federated learning for development of privacy-preserving AI, lightweight deployment of models for real-time diagnosis at the point-of-care, improved data augmentation and GAN-based approaches for generating synthetic data, and true clinical validation in the general population of patients. Future models ought to integrate MRI scans with more clinical data, utilize Grad-CAM and SHAP to visualize which features aid in classification decisions and adopt federated learning to prevent privacy leaks in AI. In addition, optimizations for edge computing and tackling data insufficiency with GANs would also increase the robustness of the model.

6.1. Ethical Considerations

ML/DL Based Brain Tumor Diagnosis has to depend on high sensitivity and specificity to eliminate false negatives and false positives. Monitoring and re-calibrating are necessary for real-world reliability. To prevent biases in ML/DL systems, diverse datasets should be used for training ML/DL models; MRI scans include sensitive personal health information, giving rise to privacy and security concerns. It is important to comply with HIPAA and GDPR to protect patient data. ML/DL should serve as a decision-support tool, but ultimately, decision-making responsibility for diagnosis and treatment planning should remain with physicians. Grad-CAM visualizations can be performed to understand ML/DL decision making with techniques of explanation.

References

[1] Tarun Agrawal et al., "A Comparative Study of Brain Tumor Classification on Unbalanced Dataset Using Deep Neural Networks," *Biomedical Signal Processing and Control*, vol. 94, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[2] Satyavati Jaga, and K. Rama Devi, "Brain Tumor Classification Utilizing Triple Memristor Hopfield Neural Network Optimized with Northern Goshawk Optimization for MRI Image," *Biomedical Signal Processing and Control*, vol. 95, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[3] Ramazan İncir, and Ferhat Bozkurt, "Improving Brain Tumor Classification with Combined Convolutional Neural Networks and Transfer Learning," *Knowledge-Based Systems*, vol. 299, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[4] B. Sandhiya, and S. Kanaga Suba Raja, "Deep Learning and Optimized Learning Machine for Brain Tumor Classification," *Biomedical Signal Processing and Control*, vol. 89, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[5] Alain Marcel Dikande Simo et al., "Introducing a Deep Learning Method for Brain Tumor Classification Using MRI Data Towards Better Performance," *Informatics in Medicine Unlocked*, vol. 44, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[6] Baireddy Sreenivasa Reddy, and Anchula Sathish, "A Multiscale Atrous Convolution-based Adaptive ResUNet3 + with Attention-based Ensemble Convolution Networks for Brain Tumour Segmentation and Classification Using Heuristic Improvement," *Biomedical Signal Processing and Control*, vol. 91, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[7] Laís Silva Santana et al., "Application of Machine Learning for Classification of Brain Tumors: A Systematic Review and Meta-Analysis," *World Neurosurgery*, vol. 186, pp. 204-218, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[8] Zekuan Yu et al., "HSA-Net with a Novel CAD Pipeline Boosts Both Clinical Brain Tumor MR Image Classification and Segmentation," *Computers in Biology and Medicine*, vol. 170, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[9] Tapas Kumar Dutta, Deepak Ranjan Nayak, and Yu-Dong Zhang, "ARM-Net: Attention-Guided Residual Multiscale CNN for Multiclass Brain Tumor Classification Using MR Images," *Biomedical Signal Processing and Control*, vol. 87, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[10] Salman Fazle Rabby, Muhammad Abdullah Arafat, and Taufiq Hasan, "BT-Net: An End-to-End Multi-Task Architecture for Brain Tumor Classification, Segmentation, and Localization from MRI Images," *Array*, vol. 22, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[11] Muhammad Yaqub et al., "DeepLabV3, IBCO-based ALCResNet: A Fully Automated Classification, and Grading System for Brain Tumor," *Alexandria Engineering Journal*, vol. 76, pp. 609-627, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[12] Longfei Li et al., "Universal Multi-Factor Feature Selection Method for Radiomics-Based Brain Tumor Classification," *Computers in Biology and Medicine*, vol.164, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[13] K. Ramalakshmi et al., "A Hyperdimensional Framework: Unveiling the Interplay of RBP and GSN within CNNs for Ultra-Precise Brain Tumor Classification," *Biomedical Signal Processing and Control*, vol. 96, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[14] Sahar Khoramipour, Mojtaba Gandomkar, and Mohsen Shakiba, "Enhancement of Brain Tumor Classification from MRI Images Using Multi-Path Convolutional Neural Network with SVM Classifier," *Biomedical Signal Processing and Control*, vol. 93, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[15] Belayneh Sisay Alemu et al., "Magnetic Resonance Imaging-Based Brain Tumor Image Classification Performance Enhancement," *Scientific African*, vol. 22, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[16] Shirin Kordnoori et al., "Deep Multi-Task Learning Structure for Segmentation and Classification of Supratentorial Brain Tumors in MR Images," *Interdisciplinary Neurosurgery*, vol. 36, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[17] Daniel Reyes, and Javier Sánchez, "Performance of Convolutional Neural Networks for the Classification of Brain Tumors Using Magnetic Resonance Imaging," *Heliyon*, vol. 10, no. 3, 2024. [Google Scholar] [Publisher Link]

[18] Monika Agarwale et al., Deep Learning for Enhanced Brain Tumor Detection and Classification," *Results in Engineering*, vol. 22, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[19] N. Nagarani et al., "Self-Attention Based Progressive Generative Adversarial Network Optimized with Momentum Search Optimization Algorithm for Classification of Brain Tumor on MRI Image," *Biomedical Signal Processing and Control*, vol. 88, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[20] Salha M. Alzahrani, "ConvAttenMixer: Brain Tumor Detection and Type Classification Using Convolutional Mixer with External and Self-Attention Mechanisms," *Journal of King Saud University-Computer and Information Sciences*, vol. 35, no. 10, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[21] Sandhya Sandeep Waghere, and Jayashri Prashant Shinde, "A Robust Classification of Brain Tumor Disease in MRI Using Twin-Attention Based Dense Convolutional Auto-Encoder," *Biomedical Signal Processing and Control*, vol. 92, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[22] P. Sreedevi et al., "Da-resbigru-Brain Tumor Classification Using Dual Attention Residual Bi Directional Gated Recurrent Unit Using MRI Images," *Biomedical Signal Processing and Control*, vol. 88, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[23] Deependra Rastogi et al., "Multi-Class Classification of Brain Tumour Magnetic Resonance Images Using Multi-Branch Network with Inception Block and Five-Fold Cross Validation Deep Learning Framework," *Biomedical Signal Processing and Control*, vol. 88, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[24] Tripty Singh et al., "BrainNet: A Deep Learning Approach for Brain Tumor Classification," *Procedia Computer Science*, vol. 235, pp. 3283-3292, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[25] Fatima Zulfiqar, Usama Ijaz Bajwa, and Yasar Mehmood, "Multi-Class Classification of Brain Tumor Types from MR Images Using EfficientNets," *Biomedical Signal Processing and Control*, vol. 84, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[26] Pallavi Priyadarshini, Priyadarshi Kanungo, and Tejaswini Kar, "Multigrade Brain Tumor Classification in MRI Images Using Fine Tuned Efficientnet," *e-Prime-Advances in Electrical Engineering, Electronics and Energy*, vol. 8, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[27] Pendela Kanchanamala, K.G. Revathi, and M. Belsam Jeba Ananth, "Optimization-Enabled Hybrid Deep Learning for Brain Tumor Detection and Classification from MRI," *Biomedical Signal Processing and Control*, vol. 84, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[28] S. Ganesh, S. Kannadhasan, and A. Jayachandran, "Multi Class Robust Brain Tumor with Hybrid Classification Using DTA Algorithm," *Heliyon*, vol. 10, no. 1, 2024. [Google Scholar] [Publisher Link]

[29] Hossein Mehnatkesh et al., "An Intelligent Driven Deep Residual Learning Framework for Brain Tumor Classification Using MRI Images," *Expert Systems with Applications*, vol. 213, 2023. [CrossRef] [Google Scholar] [Publisher Link]