

HARNESSING EXPLAINABLE AI: HYBRID DEEP LEARNING AND MULTI-HEAD ATTENTION MECHANISMS FOR ENHANCED MRI ANALYSIS

Mirza Mumtaz Zahoor^{*1}, Mahnoor Shahzadi Hashmi², Maryam Zaffar³,
Maria Bibi⁴, Ammara Khan Niazi⁵, Usama Saqlain⁶

^{*1,2,3,4,5,6}Department of Computer Science, Faculty of Computer Sciences, IBADAT International University, Islamabad, 44000, Pakistan

^{*1}mumtazzahoor5@gmail.com

DOI: <http://doi.org/10.5281/zenodo.21531442>

Keywords

pandemic, anxiety, health care workers, hybrid feature selection, Chi-square, mRMR

Article History

Received: 25 January 2026

Accepted: 10 March 2026

Published: 24 March 2026

Copyright @Author

Corresponding Author: *
Mirza Mumtaz Zahoor

Brain tumors is a the most life-threatening neurological diseases, with malignant cases often resulting in low survival rates. Surgery, chemotherapy, and radiotherapy are common treatments, but their success strongly depends on accurate and suitable diagnosis. Magnetic Resonance Imaging (MRI) has emerged as the most promising modality for brain tumor analysis. However, analyzing MRI scans can be a tedious task, especially for large images, and is highly subjective, prone to errors, and challenging for radiologists. this research, we develop a novel deep hybrid learning model, **Res-MobAT-BRTC**, for early and precise classification of brain tumors using MRI images. The model integrates Mobile Net and ResNet-50 as backbones to effectively extract both low- and high-level features, while a multi-head attention mechanism enhances discriminative capability and provides explainable outputs for clinical experts. The study employs benchmark datasets such as BraTS and Figshare MRI collections, encompassing different tumor types as well as normal instances. Experimental results demonstrate that the proposed framework is superior in terms of accuracy, precision, recall, and F1-score to existing machine learning and deep learning models. Despite the high performance, there are still challenges, including imbalanced datasets, computational complexity, and generalizability. Future work will focus on growing the dataset, integrating explainable AI (XAI) mechanisms to improve clinician trust, and optimizing the framework for multi-modal imaging integration. By addressing diagnostic challenges through a robust hybrid deep learning-based model, this research contributes to advancing automated brain tumor analysis, facilitating earlier detection, reliable classification, and ultimately improving patient survival rates

INTRODUCTION

The brain is one of the most vital organs in the human body, controlling vital processes and performing numerous other critical functions.[1] The cerebrum, brainstem, and cerebellum are its components. The brain is the core of the body, controlling all of its functions. It is the center of

all human activity, including planning, analyzing, problem-solving, thinking, and communication, both within itself and with other organs through nerve messages. The brain is one of the strangest-looking organs in the human body, as its anatomy is exceptionally complex and participates in a variety of different functions. Brain tumor refers

to a mass or growth of abnormally formed cells in the brain. They are either non-malignant or malignant that is, arising within the brain or other body parts. Brain tumors represent one of the most dangerous malignancies, as they arise due to the uncontrolled growth of brain cells. The main feature of any tumor formation is the disruption of the balance between cell division and cell death rates. This leads to the unrestrained growth of abnormal cells, resulting in focal or diffuse damage to one or more regions of the brain, depending on where the tumor is located and its size.[2] Brain tumors, which are classified broadly into primary and secondary, also called metastatic tumors. The main difference between primary and secondary brain tumors is that primary brain tumors are those that originate from the brain or spinal cord, whereas secondary brain tumors are those that spread to the brain from other parts of the body.. A primary brain tumor is a growth that develops in the brain or spinal cord from healthy cells, while secondary brain tumors occur when cancerous cells from other parts of the body such as the lungs, breasts, or kidneys, spread to the brain. Moreover, a primary brain tumor can either be benign or malignant. In contrast, secondary brain tumors mostly arise from cancerous cells. The most common brain tumor is a primary brain tumor which accounts for approximately 3 out of 5 brain tumors. On the other hand, secondary brain tumors are rare but more common than primary brain tumors. Additionally, secondary brain tumors mainly indicate the presence of cancer somewhere else in the body and usually develops when cancer cells spread to the brain. A tumor is characterized by the abnormal and uncontrolled division of cells in the brain. When such aberrant cell growth occurs within brain tissue, it is referred to as a brain tumor. Brain tumors can either be benign, malignant, primary, or secondary brain tumors. Benign brain tumors are non-cancerous while malignant brain tumors are cancerous. Primary brain tumors are cancers that originate in the brain or the central nervous system. On the other hand, secondary brain tumors or metastatic brain tumors are cancers that spread to the brain from other body parts. Metastatic brain tumors are the ones that are

mostly malignant and originate from cancers in the lungs, breasts, and kidneys. Gliomas, a common type of brain tumor derived from glial cells, represent one of the primary tumor categories. When cancerous cells from other regions disseminate to the brain, they can cause metastatic brain tumors. Brain tumors can develop in any part of the brain, including neural tissue, meninges, bones, or surrounding structures. These tumors are particularly aggressive, exhibiting a mortality rate of approximately 70% [3]. The major types of brain tumors are gliomas, meningiomas, and pituitary tumors. Of these, gliomas are the tumors most commonly diagnosed in adults. While primarily occurring in the brain, they also sometimes develop in the spinal cord. Gliomas originate from glial cells, which are the supportive cells of the central nervous system. These cells perform various functions necessary for the proper functioning of neurons such as providing insulation and protection.[4]. Gliomas are tumors that originate from glial cells. These cells are essential for the normal functioning of the central nervous system, as they provide support and nutrition to neurons. Meningiomas are tumors that are formed by the meninges, which is a set of membranes that surround the brain and the spinal cord beneath the skull. According to the research data published by the Central Brain Tumor Registry of the United States (CBTRUS), meningiomas comprise “53.9% of all non-malignant brain tumors diagnosed between 2013 and 2017. According to the data published by the Central Brain Tumor Registry of the United States (CBTRUS), during the period between 2013 and 2017, meningiomas constituted 53.9% of all non-malignant brain tumors.[5]. Pituitary tumors are abnormal growths of cells that occur in the pituitary gland. The pituitary gland is a small, pea-shaped organ located at the base of the brain behind the nasal cavity. It controls hormone production and plays an important role in regulating various body functions. Although most pituitary tumors are benign, they often affect the normal functioning of the pituitary gland, causing it to produce an abnormally large amount of hormones. The pituitary gland controls many important processes in the body,

including growth, development, reproduction, and metabolism. Magnetic Resonance Imaging (MRI) is an important imaging technique that can be used to detect and evaluate brain tumors at an early stage. A complicated strategy is usually used to treat brain tumors, especially when they are still in their early stages. Surgery, chemotherapy, and radiotherapy are frequently used methods to treat the tumor directly. Accurate diagnosis is significant for effective treatment. Several types of brain scans, such as Single-Photon Emission Computed Tomography, also known as SPECT, computerized tomography or CT, and magnetic resonance imaging or MRI, can be used to study the brain structure and function. However, MRI is the most effective method for diagnosing brain tumors because it allows obtaining a more informative image of the brain than other methods and is safer for patients..[6] Brain tumors can be detected and diagnosed using a variety of imaging tests, including ultrasonography (US), computed tomography (CT) and magnetic resonance imaging (MRI). However, MRI is the most optimal method because it provides high-quality imaging without subjecting patients to dangerous ionizing radiation compared to CT and X-ray machines. At the same time, MRI is versatile in that it allows obtaining different types of scans, among which are T1- and T2-weighted images. The former is useful for diagnostic imaging because it highlights normal anatomical structures, whereas the latter is more sensitive and informative, thus being able to distinguish abnormal tissue from healthy ones. [7] A timely and correct diagnosis of brain tumors is an important factor in successful treatment. However, the task of classifying these tumors is not always an easy one, given their wide spectrum of sizes, shapes, locations, and appearances. To address the issue, researchers have begun employing both conventional machine learning (ML) and modern deep learning (DL) algorithms. In the case of ML, there is a need to convert an image into representative numerical features to be used as an input for training a classifier. This step is quite time-consuming and depends heavily on the quality of manual feature extraction. Multiple studies utilized ML algorithms like Support Vector

Machine (SVM) to classify brain tumor images and detect different tumor types and patterns in MRI scans[8], k-nearest neighbors [9], and decision trees [10]. Unlike conventional machine learning approaches, using deep learning (DL) algorithms allows for extracting relevant features from medical images automatically. The convolutional neural network (CNN) model is considered one of the most common DL architectures for analyzing medical images due to its ability to learn lower- and higher-level features. The CNNs' high accuracy in image classification tasks has motivated researchers to apply these networks for solving medical image analysis problems, including classifying brain tumors.[11], [12], [13], [14], [15], [16], [17], [18]In [19] To develop a machine learning based system for brain tumor detection and segmentation, a set of 304 brain MRI images were taken from 120 patients at Tiantan Hospital. Their approach involves image preprocessing, GLCM-based feature extraction, EKF-SVM classification, and segmentation using k-means clustering and region-growing algorithms. The skull stripping process significantly improves the overall detection performance. The classification accuracy achieved was 96.05%, while tumor borders were effectively delineated using k-means clustering. However, the study's limitation was that the high classification performance might be influenced by the relatively small number of negative (non-tumor) cases in the dataset. In [20] conducted research on 180 histopathologically confirmed brain malignancies (77 lymphomas, 45 glioblastomas, and 58 metastases), employing Classification and Regression Tree (CART) models with perfusion and spectroscopy MRI data for multiclass tumor classification. Their decision tree model reached a high accuracy of 0.96, with an R^2 of 0.887 and impressive AUC values of 0.98 for PCNSLs, 0.98 for GBMs, and 1.00 for METs, supported by five decision rules with expected probabilities ranging from 0.907 to 0.989. In [21] adopted a deep learning framework, evaluating object detection algorithms on the Brain Tumor Figshare (BTF) dataset (3064 slices from 233 subjects) and the BRATS 2018 dataset (285 glioma cases). Their best-performing network, YOLOv5, was paired

with a 2D U-Net for pixel-wise segmentation, achieving a mean Average Precision (mAP) of 89.5% for detection and a Dice Score Coefficient (DSC) of 88.1% for segmentation, outperforming Mask R-CNN. The study suggests that future research should focus on developing a single framework that can perform both tumor detection and segmentation.[22] The study also recommends that the approach be tested on larger and more diverse data sets. In [23] also explored deep learning techniques, leveraging an improved ResNet architecture for glioma segmentation using the BraTS2020 dataset. By incorporating long skip connections and optimized bottleneck blocks, the model overcame gradient vanishing issues and outperformed conventional CNNs and fully convolutional networks. The model's applicability is limited by its dependence on single data sets and its susceptibility to noise in MRI scans. The ways different researchers attempted to overcome these barriers have demonstrated how hybrid learning can provide an adequate solution for these challenges. In [24] The authors proposed a hybrid network which was constructed using YOLO V5, K-means + , Spatial Pyramid Pooling Fast (SPPF +) and Bidirectional Feature Pyramid Network (Bi-FPN) for the improved detection and segmentation of tumorous tissues from the set of MRI images considered. Sahoo et al. (2023) presented a novel hybrid framework integrating CNNs with GANs, achieving 98.85% accuracy. Their model enhanced diagnostic precision through GAN-based data augmentation, increasing variation and resilience, thereby supporting early-stage tumor identification. [25] In this study, the authors evaluated the performance of the deep transfer learning method on the basis of a data base comprising 7023 brain MRI images from 4 classes, namely, pituitary tumour, meningioma, glioma, and normal brain. Fine-tuning pre-trained architectures such as ResNet152, VGG19, DenseNet169, and MobileNetv3, they achieved a peak accuracy of 99.75% with MobileNetv3. To address data imbalance, image enhancement strategies and five-fold cross-validation were employed, ensuring robustness and generalizability. Considered together, the studies

demonstrate the shift in the field from traditional machine learning approaches to the use of deep learning and hybrid algorithms. At the same time, the reviewed research indicates the necessity to expand the scope of data used for training algorithms and pay attention to the combination of different methods for more accurate detection and segmentation of brain tumors. Artificial Intelligence (AI) is a crucial technology in numerous sectors, and healthcare is one of its most significant applications. [26]. However, even though deep learning (DL) tools are extremely successful in disease diagnosis, medical image analysis, and treatment recommendation, such models are often black-boxes since they do not show how specific results were obtained. This circumstance decreases the level of trust that professionals have in AI-powered systems, which, consequently, limits DL implementation in medicine. Therefore, it is essential to make such models interpretable to ensure that the final results are understandable by humans. Interpreting deep learning models is a critical step toward fulfilling the promise of AI to revolutionize healthcare. Since those models are typically not transparent, they are often called black-boxes.[27]. Explainable Artificial Intelligence (XAI) develops approaches that enable machines to explain their predictions and decisions, making AI more transparent to its users. As applied to Convolutional Neural Networks (CNNs), the explainable AI methods allow one to visualize and understand which parts of the input image contribute the most to the final image classification by means of weighted gradients. In particular, CNN explainability methods locate the regions of the input image that most influence the model prediction. They do so by combining the gradients of the output at the last convolutional layer into an activation heatmap, which highlights the regions of the input that are the most influential for the model prediction. This approach helps to increase the interpretability and trustworthiness of CNN-based medical image classification models, which is why most recent research studies attempt to combine deep learning approaches with explainable AI [28], [29], [30] works have proposed utilizing the concept of artificial

intelligence (AI) to identify and classify brain tumors using different methods. Some works applied the concept of transfer learning using one model while others designed various Convolutional Neural Networks (CNNs) or hybrid and ensemble learning techniques.. Each of the approaches used has its strengths and weaknesses. The methods used in transfer learning utilized only one pre-trained network to identify features and therefore could be limited by the performance of the chosen network. On the other hand, custom CNNs require huge amounts of annotated data and high-end hardware to obtain good results. Hybrid and ensemble models suggest a different approach, which is based on combining the properties of several networks. The essence of the method lies in the fact that the resulting model demonstrates higher accuracy, better generalizability, and a lower error probability inherent in decisions made by one algorithm. However, in any case, be it hybrid or ensemble models, explainability is vital since it allows one to significantly increase the accuracy of the results' implementation. It is especially important in medicine, where the ability of the algorithm to explain its decision-making process reduces the likelihood of errors during timely treatment. To overcome the limitations described above, the approach presented classifies the three most common brain tumor categories (glioma, meningioma, and pituitary) using an ensemble of models based on transfer learning. The proposed approach was trained and tested on the publicly available data from Figshare. The study utilized the publicly accessible data in order to benefit from using several pre-trained convolutional neural networks. Those networks were able to learn visual features, improve the accuracy of image classification tasks, and decrease the time spent on training due to the pretraining process on big data sets. (MobileNet and ResNet-50) The pre-trained models were fine-tuned by replacing their classification heads with a task-specific classifier. The visualization of the image regions that contribute to the classification results was done using Gradient-weighted Class Activation Mapping. This helped to understand which parts of the image were used for classification and increased the overall interpretability of the model

performance..[31] An innovative approach based on Gradient-weighted Class Activation Mapping was applied toward identify the image regions that are the most influential for the model. This allowed the model to demonstrate its predictions' accuracy better than conventional approaches while providing a more transparent description of the process. [32] [33]The combination of explainable artificial intelligence and an ensemble model resulted in a higher classification quality than could be demonstrated by the models separately. The significant contributions of this work have been primarily focused on the following aspects:

- An improved Grad-CAM technique has been incorporated into the proposed ensemble model to identify discriminative regions for the predicted class, which has also been extended to visualize the prediction from individual convolutional layers and multiple convolutional layers.
- An ensemble framework of MobileNet and ResNet-50 by transfer learning has been introduced for classifying the three types of brain tumors, namely, glioma, meningioma, and pituitary tumor, which attains state-of-the-art results on a benchmark dataset.
- The proposed model has been evaluated on the basis of two strategies, namely, the train-test partition and stratified-k fold cross-validation, which ascertains the robustness of the designed framework.
- The proposed design has been benchmarked with the state-of-the-art approaches like MobileNet, Xception, ResNet-50, DenseNet121, and InceptionV3 for classifying the three types of tumors, which depicts that the attained results are superior to the existing models.
- A comprehensive comparison has been made with the existing state-of-the-art methods to highlight the efficiency of the proposed design.

The remaining part of the paper has been structured as follows: Section 2 sheds light upon

the dataset and the preprocessing performed for the experiments. Section 3 describes the methodology implemented for the experiments. The attained results have been presented in Section 4 while Section 5 concludes the paper and brings out the contributions of the paper.

2. Proposed methodology

In this work, we have developed a deep hybrid learning-based model designed for the early and accurate classification of brain tumors into their respective types. The proposed framework, Res-MobAT-BRTC, adopts a hybrid deep learning approach that integrates spatial and residual learning-based convolutional neural networks (CNNs), augmented by a multi-head attention mechanism to enhance the model's discriminative ability. The methodology comprises three stages, each aimed at optimizing classification performance. The first stage involves the implementation of standard deep CNN architectures, which serve as the foundation of the proposed model. As a result of the analysis, the MobileNet and ResNet-50 algorithms were selected as the basic models for solving the given problem. Subsequently, according to the second stage, the residual part was added to the models. It can alleviate gradient exploding/vanishing problems in deep networks to some extent, thus enabling the network to learn more deep features. As the final step, a multi-head attention is introduced before the classifier to enhance the model's ability of capturing discriminative regions in the MRIs, which facilitates more accurate and robust predictions. This attention-based block contributes to the model's improved classification performance and interpretability. To reduce the

risk of overfitting, a **dropout layer** is incorporated into the model during training. [34] and L2 regularization [35]. The dropout layer is a technique that reduces overfitting in neural networks by randomly deactivating some neurons during training. Another technique that helps reduce overfitting is L2 regularization. This method penalizes the loss function leading to smaller weights in the model. The TensorFlow API was used to augment training data improving the model's performance on test sets. [36]

2.1 Dataset and preprocessing

In this work we have evaluated our models on a benchmark dataset [37] this data set has gathered the from publically available open sources, Fig share, and the BR35H. There were two hospitals in China that provided the MRI scans used for the study, and they are referred to as Nanfang Hospital and the General Hospital of Tianjiang Medical University. Such datasets are typically compiled by researchers who intend to train algorithms to perform computer vision tasks, which has become a common practice in the medical and machine learning fields.

Masoud Nickparvar was the one who organized all of these disparate sets, making it easier for other researchers to utilize the data. As a result, he was able to create an extensive library of 7023 labeled MRI photographs that have been categorized into four types of tumors: glioma, meningioma, pituitary, and no tumor. Thus, providing an excellent resource for testing advanced techniques in deep learning-based brain tumor classification systems.[38]

Proposed Model Diagram (Developed Res-MobAT-BRTC)

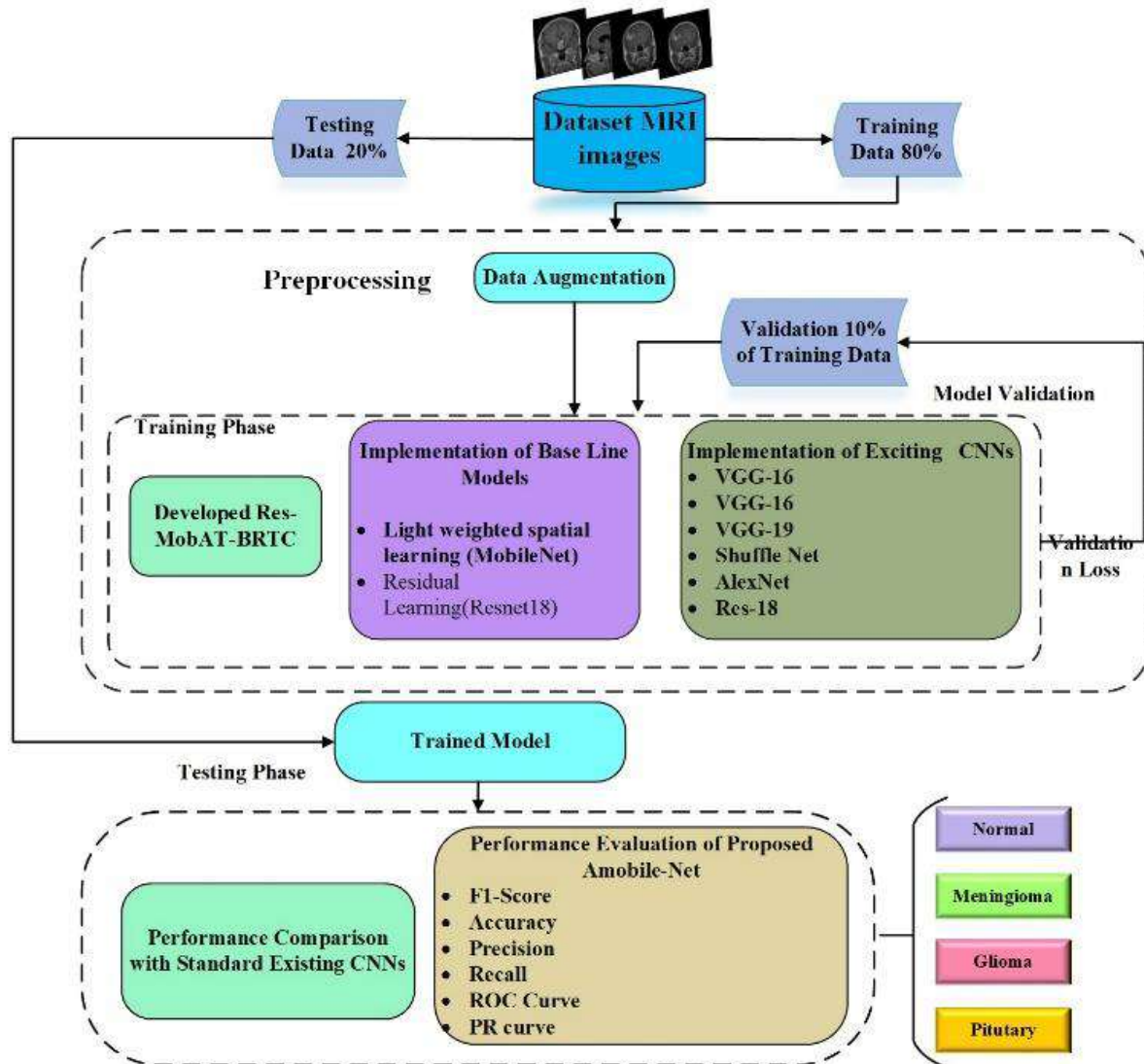


Figure 1. Res-MobAT-BRTC Model

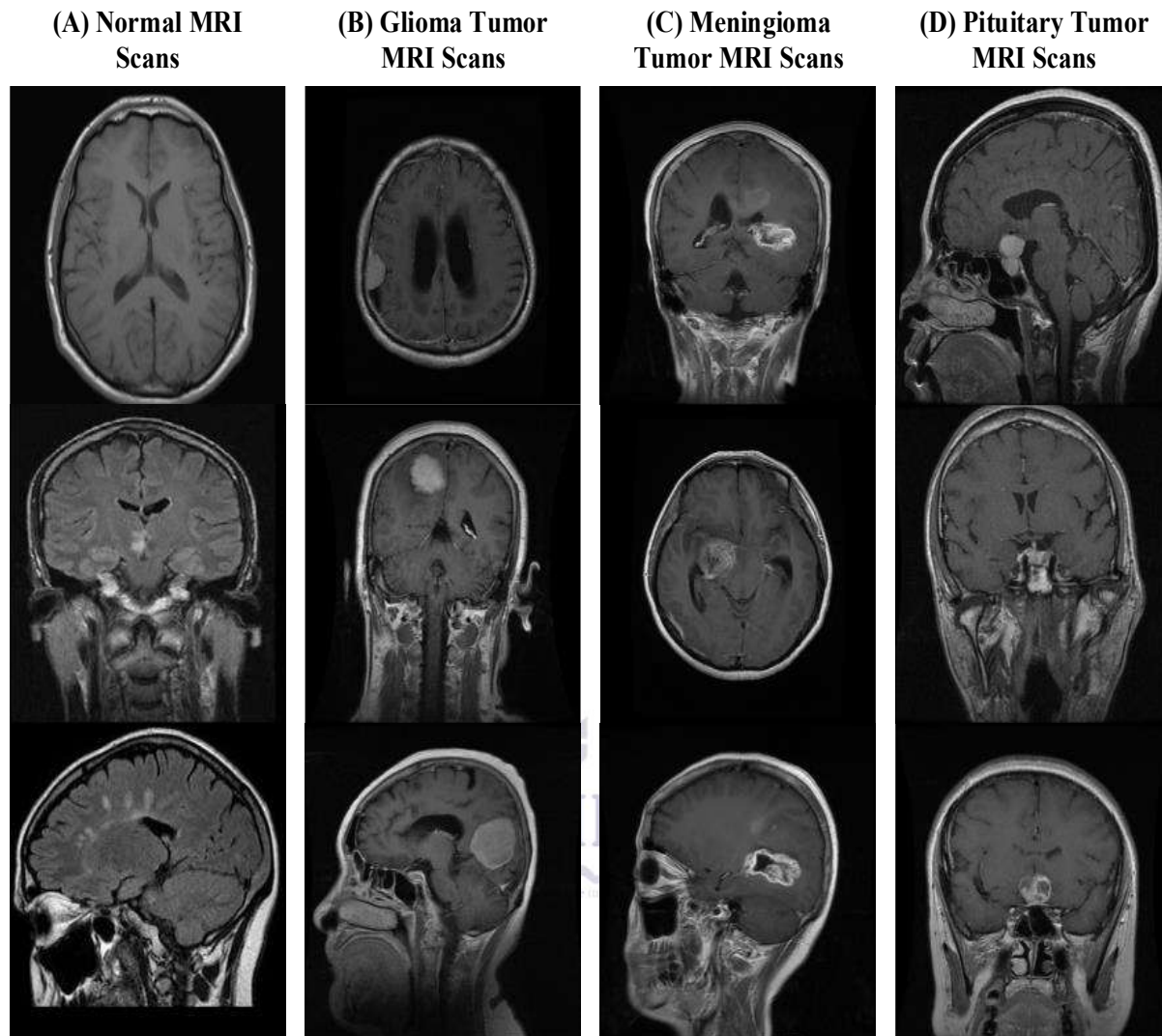


Figure 2. Sample Images from Dataset

The dataset consists of four classes of brain tumors (Glioma, Meningioma, Non-tumor, and Pituitary) with a fairly balanced representation of images for each class. To increase the accuracy of my model, I used a preprocessing step to adjust the input image size and augmentation steps to improve model performance. To ensure that all the MR images were in the same input size, I performed resizing at a standard resolution (224×224). For image augmentation, I applied multiple techniques, such as rotation, flipping, and scaling, among others. Standardization allows the model to converge more quickly. During training, we use a dynamic set of transformations for augmentation, such as

rotation, zoom, shear, shifts in height and width, and horizontal flipping. Adding variability in image orientation and scaling improves the model's generalization. Replicating changes in lighting, brightness adjustments (0.8 – 1.2) are also used. Guaranteeing objective model evaluation, a 10% validation split is also applied. This preprocessing technique goal to improve model performance and correctly classify brain tumor images.

Performance evaluation of our proposed model is compared with several existing standard CNN architectures, such as VGG-16 [39], VGG-19 [40], Alex Net [41], Squeeze Net [42], and ResNet-

50 [43] which have been widely used in various image classification applications.

3.1 Develop Model Architecture (Res-MobAT-BRTC)

The Res-MobAT-BRTC model introduces a hybrid architecture that combines the strengths of spatial and residual learning based convolutional neural networks (CNNs) with customized attention mechanisms, for enhanced model

performance in brain tumor classification.[44] Developed model have following three key components Mobile Net backbone, Residual Blocks and Multi-Head Attention. These components collectively enhance the model's performance, Enabling more accurate classification and improved learning capacity for complex medical imaging tasks.

Proposed Model Architecture Diagram

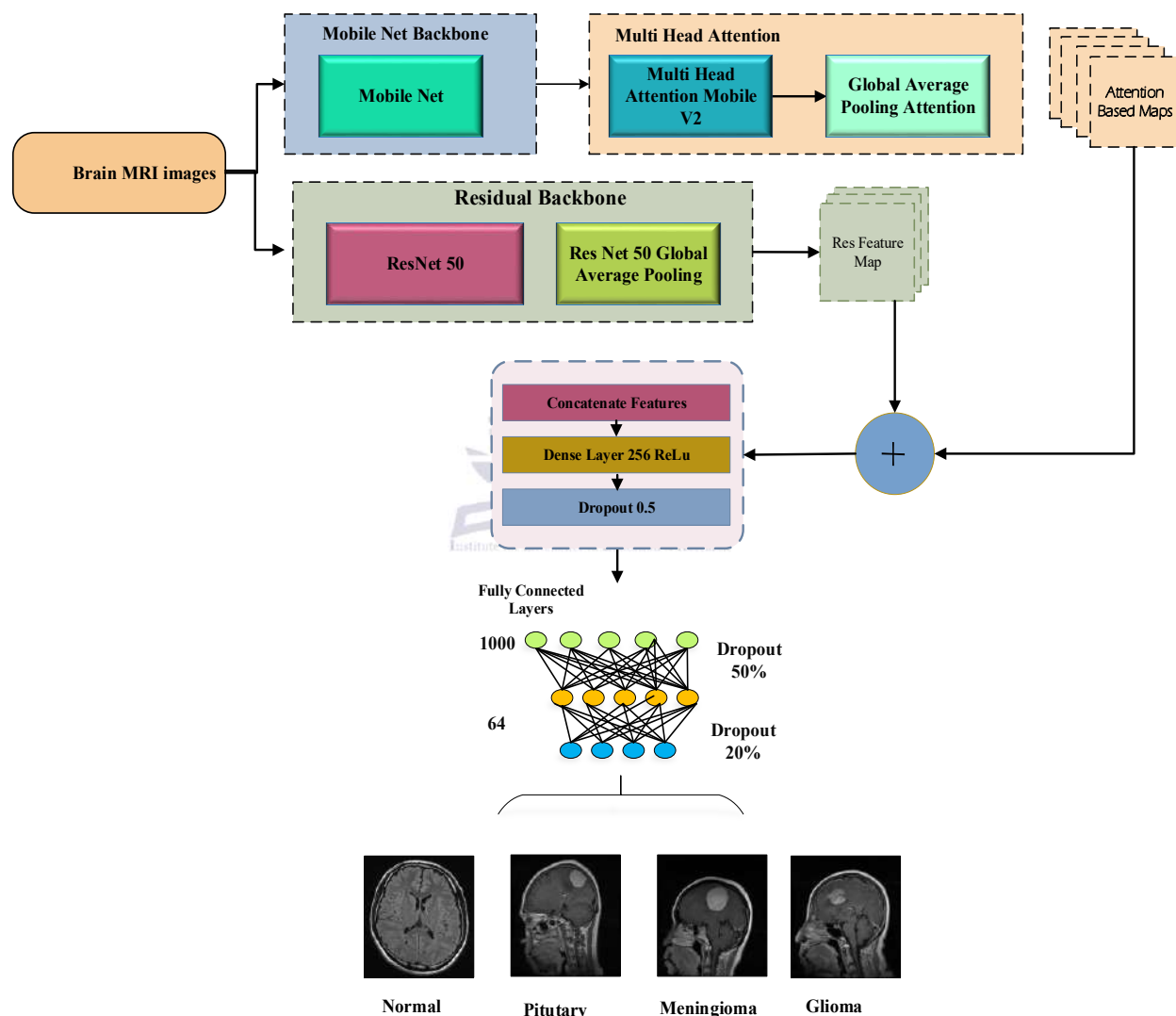


Figure 3. Proposed Model Architecture

3.1.2 Mobile Net Backbone

Res-MobAT-BRTC model is core deep learning architecture. Mobile Net an efficient and light weighted CNNs. It employs depth-wise separable convolutions, and famous for its lower

computational complexity, lowers the number of parameters and processing requirements without compromising on performance. It offers the ideal basis for processing medical images, particularly in the context of brain tumor classification, where

the capture of rich hierarchical information is crucial. Mobile Net serves as the backbone, extracting essential features from MR images and

3.1.3 Residual learning block

Res-MobAT incorporates residual blocks to improve the model's learning capacity. This block tackles the vanishing gradient problem that often occurs in deeper neural networks by introducing shortcut connections that facilitate residual mapping learning. Residual connection allowing

providing a powerful representation for further processing. [45]

gradients to avoid particular layers, allow the network to acquire more complex and discriminative properties. It improves the model's performance to identify more complex problems and generate more accurate tumor classification. [31]

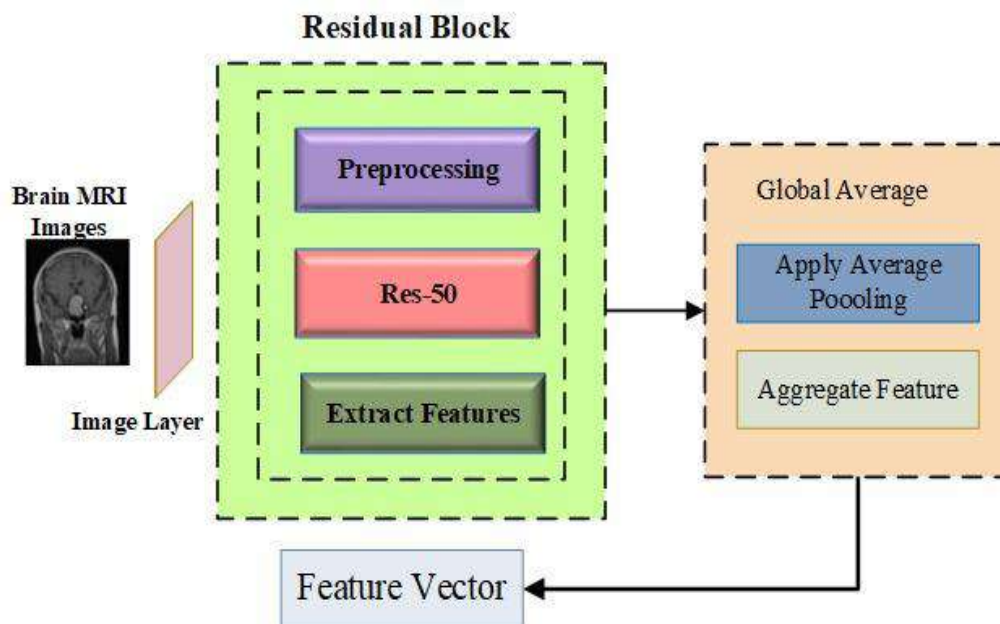


Figure 4. Residual Block

3.1.4 Res-Mobile Net_Attention

Multi-Head Attention at third stage of the model's development, enhanced the discrimination ability of Mobile Net architecture. Developed model become capable to focusing on the most relevant parts of the input image by considering multiple attributes simultaneously, allowing it to capture a variety of visual information. Multi-Head Attention for improved the model's capacity to detect variation in tumor properties. The model's overall classification ability is much enhanced by this approach,

particularly when it comes to identifying complex patterns that are essential for precise brain tumor diagnosis. Combined these components form a robust architecture can recognize the complex details of brain tumor studies. Combining Mobile Net with attention mechanisms and residual learning blocks, the Res-MobAT-BRTC model performs better than convolutional CNN-based techniques, offering increased classification accuracy and resilience in distinguishing between healthy and affected brain tissue.[46]

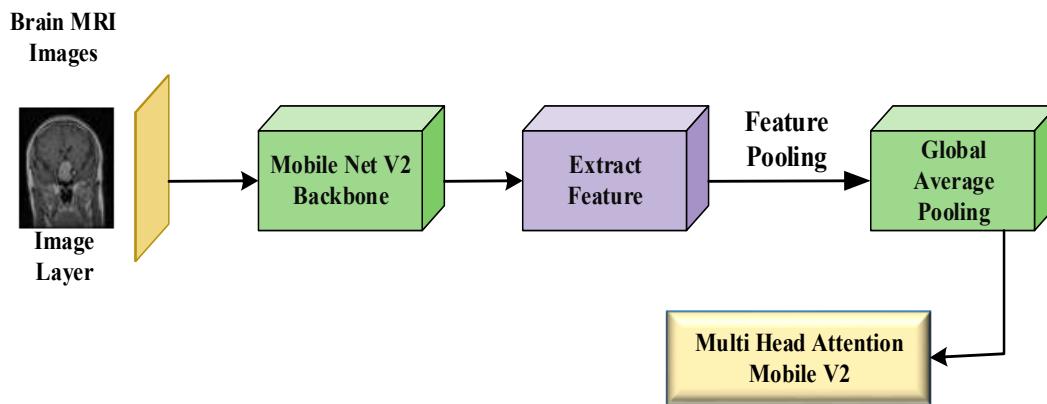


Figure 5. Mobile Net with Multi-Head Attention

3.2 Hyperparameters

In order to compare the networks accurately, each network was trained under equal conditions for the same number of epochs, while recording their training and validation accuracies over time. Adaptive Moment Estimation was used as an optimizer for all runs of the algorithm [47]. Each network was trained using the Adam optimizer and a batch size of 10, an initial learning rate of 0.0001, and 25 training epochs for both networks. In addition, networks were optimized using Reduce-on-Plateau that reduced the learning rate by 50% if the validation accuracy did not improve for three consecutive epochs. All hyperparameters were selected after performing a set of preliminary experiments in which the networks' behavior was analyzed. The values presented above were chosen based on the highest classification accuracy and optimization results.

3.3 Explainable Model

To visualize the results, it was proposed to use Grad-CAM as an explainability method for image recognition models. Grad-CAM is the algorithm that highlights the input image regions that are the most influential for the model prediction. The main benefit of Grad-CAM is that it can be

applied to different convolutional neural networks (CNNs). It allows to visualize the CNN's decisions without changing it and retraining, which is one of the main benefits of Grad-CAM algorithm. First, the last layers of MobileNet and ResNet-50 CNN were chosen, as they contain the most significant spatial features. The lower layer would have fewer spatial information since it was already processed by several layers of CNN. Secondly, these layers were selected because they have a sufficient amount of detail. The fully connected layers would have high-level features, but their complexity would be too high to visualize them effectively.[31] Grad-CAM is an algorithm that highlights the most relevant regions in an input image by calculating the importance weights of gradients that are passed through the activation map. It is possible to highlight the regions that are "responsible" for an image prediction made by a neural network with this technique. Several modifications of Grad-CAM allow to visualize activations of different convolutional layers, not only a single convolutional layer.

4. Results and Discussion

We developed a deep hybrid learning-based model, Res-MobAT-BRTC, for early and accurate

brain tumor classification. The model integrates spatial and residual learning using deep CNNs, followed by a multi-head attention mechanism to enhance discrimination. It combines ensemble learning with attention and regional residual learning to improve performance. The methodology involves three stages: implementing standard CNNs, selecting the best-performing models (Mobile Net and ResNet-50), and enhancing them by integrating a residual block into Mobile Net and adding an attention module before classification. Each enhancement in the architecture leads to improved performance. In this paper, we have applied various performance measures for evaluating the proposed computer-aided diagnosis system for developing a robust model. For the experiment, the CE-MRI dataset was used, wherein 70% of the data was allocated for training while the rest 30% was used for testing the performance of the model. The experiments were conducted using the TensorFlow framework on an NVIDIA RTX 3050 GPU with 16 GB memory. The performance of models was evaluated using the confusion matrix, which demonstrated model prediction results compared to actual data results. Additionally, several performance assessment metrics, namely accuracy, precision, recall, and F1 score, were computed based on the confusion matrix to identify the optimal model for this particular task.

5. Comparative analysis with existing CNNs

When comparing the proposed Res-MobAT-BRTC model with existing Convolutional Neural Networks (CNNs), the results show a significant improvement in both accuracy and performance. As illustrated in the table, traditional models such as VGG-16, Squeeze Net, VGG-19, Alex Net, and ResNet-50 achieve accuracy rates ranging from 59% to 92%. Specifically, Squeeze Net attains a relatively higher accuracy of 81%, while VGG-16 records the lowest accuracy at 59%. VGG-19 and ResNet-50 display improved performance with accuracies of 84% and 92%, respectively, and Alex Net achieves an accuracy of 86%. In contrast, the Res-MobAT-BRTC model exhibits a remarkable accuracy of 99%, alongside consistently high precision, recall, and F1-score metrics. These results demonstrate the ability of the Res-MobAT-BRTC model to perform the classification task better than the traditional CNNs. Previous studies have investigated the use of machine learning (ML) and deep learning (DL) algorithms in the diagnosis and identification of various types of brain tumors on the same dataset. To enhance performance prediction, numerous studies have implemented ensemble learning, a strategy that combines different models' power to improve performance prediction. This method achieves better results by overcoming weaknesses of isolated ML models.

Table 2. Based CNNs Performance Analysis

Model Name	Accuracy	Precision	Recall	F1-Score
VGG-16	59%	76%	0.58	55%
Squeeze Net	81%	82%	0.81	81%
VGG-19	84%	87%	0.83	83%
Alex Net	86%	87%	0.86	85%
ResNet-50	92%	93%	0.93	93%
Developed model (Res-MobAT-BRTC)	99%	99%	0.99	99%

5.1 Confusion Matrix

Confusion matrix is a learning tool used to determine the performance of a classification algorithm. It is a tabular layout that compares the models predicted class labels with the dataset true or actual class labels. Rows represents the actual classes and columns represents the anticipated classes. The matrix clearly shows accurate predictions along its diagonal elements and classifies errors in the off-diagonal columns. In classification, the confusion matrix identifies four

primary components of predictions: False Positives, False Negatives, True positive and True Negative. By accessing these figures, one may calculate key performance measures like as accuracy, precision, recall, and F1-score. Confusion matrix provides information on the model's reliability and possible areas for improvement by not only counting the number of correct and incorrect predictions but also displaying the types of errors the model generates.

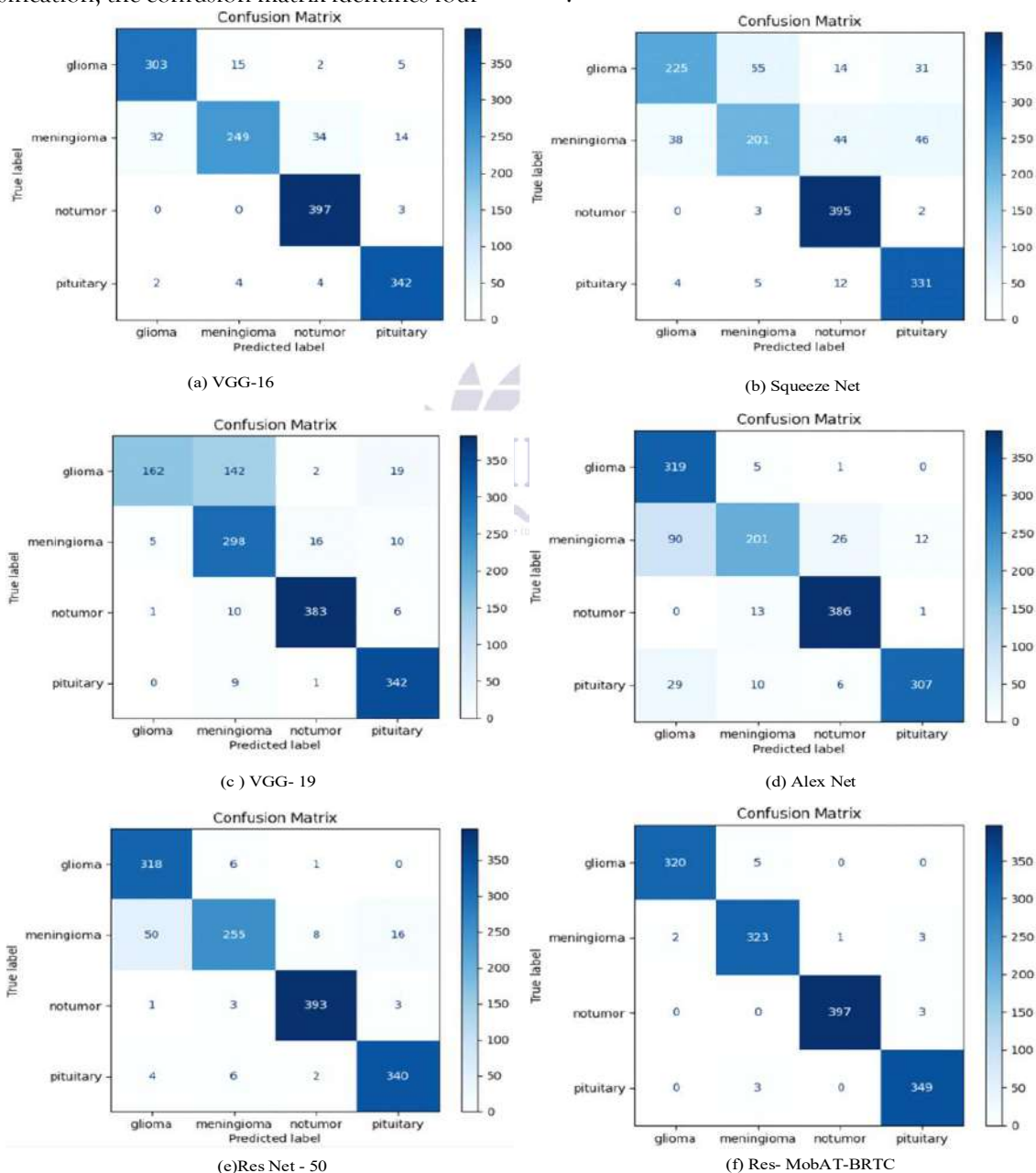
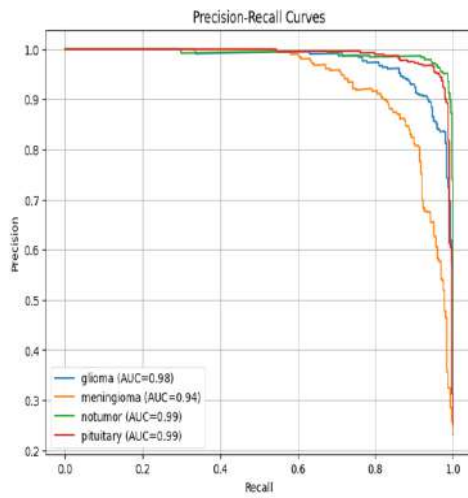
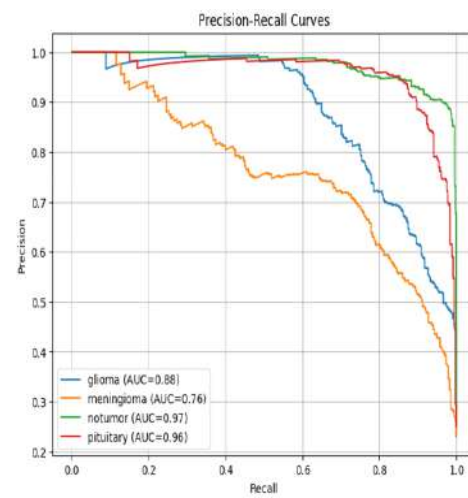


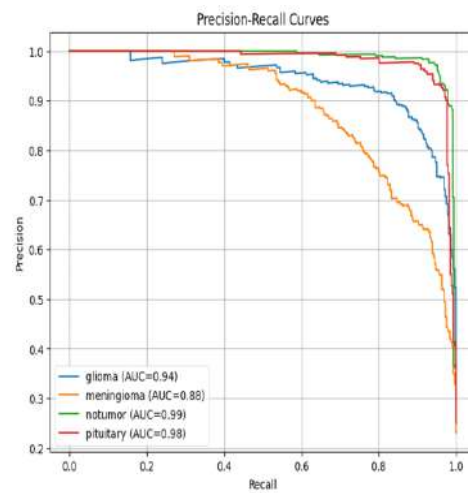
Figure 6. Confusion Matrix Based Performance Analysis



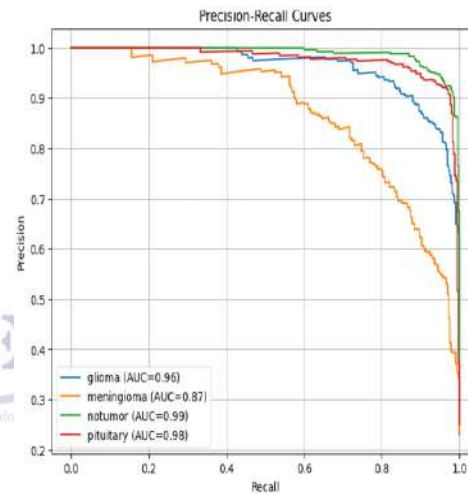
(a)VGG-16



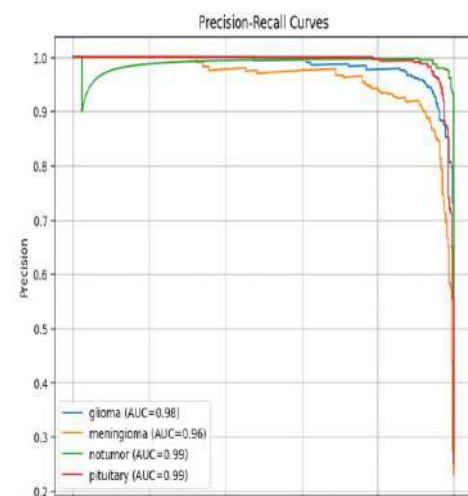
(b)Squeeze Net



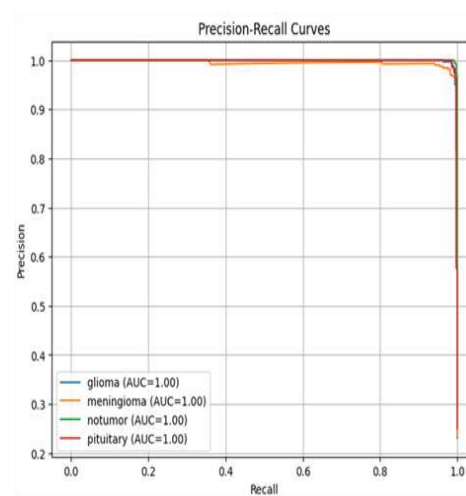
(c)VGG-19



(d)Alex Net



(e) ResNet -50



(f) Res-MobAT-BRTC

Figure 7.PR Curve based Performance Analysis

5.2 PR Curve

A Precision-Recall curve is a graphical method used for evaluating the performance of a binary classification model. By plotting Recall on the x-axis and displaying Precision on the y-axis.

5.3 Evaluation Using Cross-Validation

In order to evaluate the accuracy of the developed model, a stratified five-fold cross-validation was undertaken. Stratified five-fold cross-validation implies dividing a data set into 5 equal folds and, successively taking one fold as a validation set and the remaining four as training data. Thus, all five subsets of the sample are used for both training and testing, which reduces the variance of the results. Unlike standard k-fold cross-validation, the stratified version keeps the ratio of the classes in each fold the same as in the original dataset. Thus, for classification tasks, when the number of instances in the classes is not proportional to their representation in the sample, stratified cross-validation is used. In the present research, 5-fold stratified cross-validation is applied during the training of the model and the training set division into training and validation samples for both classes. The test set was used for the final evaluation of the trained model's performance.

5.4 Grad-CAM Results

The effectiveness of Grad-CAM as an explanation method was validated on images from the test set, which were not used during training. The activation maps produced by Grad-CAM correctly identified the regions of interest that contributed to the prediction results of both networks for axial, coronal, and sagittal MRI views. This indicates that the model can highlight the areas of the image that are most relevant for its decisions.

5.5 Computational Complexity Analysis

The proposed approach is based on the concept of transfer learning, which makes use of already trained models. Such models allow to reduce the time spent on training since their parameters were optimized on large data sets, enabling them to capture general patterns specific to the input data type. The accuracy of the final result is sufficient because of using these pre-trained

models. On the other hand, the increased accuracy requires more computational power since all of the networks need to be run in order to make predictions. Also, it should be noted that the current architecture does not allow to benefit from interaction between models because all of them are trained independently. Therefore, there is an opportunity to improve the model by considering hybrid approaches. They could enable to replace some of the convolutional layers with the handcrafted features extracted from the images using traditional machine learning algorithms.[48] Thus, it would allow both to reduce the computational complexity and to provide more high-level features to the model, improving its performance.

6. Conclusion

The proposed Res-MobAT-BRTC framework, combining MobileNet, ResNet-50, and a multi-head attention mechanism, achieved 99% accuracy in brain tumor classification using MRI images, outperforming conventional models. Preprocessing was done on the dataset by extracting ROI through applying a threshold, cropping, and resizing to remove the unnecessary black parts of the image. To reduce overfitting, dropout layers and L2 regularization were added to the model during training. To increase the amount of data and improve generalization and address the class imbalance problem, data augmentation was performed[49]. The proposed methodology was tested on a publicly available database using a train/test approach and five-fold stratified cross-validation to evaluate the proposed solution appropriately. Grad-CAM was applied to visualize the image regions that influence the model decisions, and the results were validated on MRI scans in axial, coronal, and sagittal views. The results demonstrated the effectiveness of the hybrid CNN-based ensemble concept for brain tumor classification. The proposed approach achieved better results than a single network due to the combination of different CNNs' features. Moreover, Grad-CAM provided gradient-based visualizations of MRI scans, which increases the trustworthiness of the obtained results. Thus, the created prototype can be beneficial for neurologists and neurosurgeons

in their practice and research. Future work will be concentrated on developing explainable artificial intelligence methods, including utilizing multi-modal scans, clinical data, and tumor segmentation, to increase the efficacy of the presented approach. Additionally, the design of efficient hybrid deep learning and machine learning models that reduce computational complexity while providing higher-level features for better classification will be investigated.[50]

REFERENCES

- [1] S. S. Yadav and S. M. Jadhav, "Deep convolutional neural network based medical image classification for disease diagnosis," *J. Big data*, vol. 6, no. 1, pp. 1–18, 2019.
- [2] J. R. McFaline-Figueroa and E. Q. Lee, "Brain Tumors," *Am. J. Med.*, vol. 131, no. 8, pp. 874–882, Aug. 2018, doi: 10.1016/j.amjmed.2017.12.039.
- [3] P. K. Mishro, S. Agrawal, R. Panda, and A. Abraham, "A Survey on State-of-the-Art Denoising Techniques for Brain Magnetic Resonance Images.," *IEEE Rev. Biomed. Eng.*, vol. 15, pp. 184–199, 2022, doi: 10.1109/RBME.2021.3055556.
- [4] L. M. Wang, Z. K. Englander, M. L. Miller, and J. N. Bruce, "Malignant Glioma," *Adv. Exp. Med. Biol.*, vol. 1405, pp. 1–30, 2023, doi: 10.1007/978-3-031-23705-8_1.
- [5] K. D. Miller *et al.*, "Brain and other central nervous system tumor statistics, 2021," *CA. Cancer J. Clin.*, vol. 71, no. 5, pp. 381–406, Sep. 2021, doi: 10.3322/CAAC.21693.
- [6] R. Bhargava *et al.*, "Contrast-enhanced magnetic resonance imaging in pediatric patients: review and recommendations for current practice.," *Magn. Reson. Insights*, vol. 6, pp. 95–111, 2013, doi: 10.4137/MRI.S12561.
- [7] S. M. Rezaei, N. Chegeni, F. Baghaei Naeini, D. Makris, and S. Bakas, "Within-Modality Synthesis and Novel Radiomic Evaluation of Brain MRI Scans," *Cancers (Basel)*, vol. 15, no. 14, Jul. 2023, doi: 10.3390/CANCERS15143565.
- [8] M. O. Khairandish, M. Sharma, V. Jain, J. M. Chatterjee, and N. Z. Jhanjhi, "A Hybrid CNN-SVM Threshold Segmentation Approach for Tumor Detection and Classification of MRI Brain Images," *IRBM*, vol. 43, no. 4, pp. 290–299, Aug. 2022, doi: 10.1016/J.IRBM.2021.06.003.
- [9] H. Zerouaoui and A. Idri, "Deep hybrid architectures for binary classification of medical breast cancer images," *Biomed. Signal Process. Control*, vol. 71, p. 103226, Jan. 2022, doi: 10.1016/J.BSPC.2021.103226.
- [10] T. V. Moga *et al.*, "Multiparametric Ultrasound Approach Using a Tree-Based Decision Classifier for Inconclusive Focal Liver Lesions Evaluated by Contrast Enhanced Ultrasound," *J. Pers. Med.*, vol. 11, no. 12, Dec. 2021, doi: 10.3390/JPM11121388.
- [11] P. Dutande, U. Baid, and S. Talbar, "Deep Residual Separable Convolutional Neural Network for lung tumor segmentation," *Comput. Biol. Med.*, vol. 141, Feb. 2022, doi: 10.1016/J.COMPBIOMED.2021.105161.
- [12] M. A. Kassem, S. M. Naguib, H. M. Hamza, M. M. Fouda, M. K. Saleh, and K. M. Hosny, "Explainable Transfer Learning-Based Deep Learning Model for Pelvis Fracture Detection," *Int. J. Intell. Syst.*, vol. 2023, 2023, doi: 10.1155/2023/3281998.
- [13] A. Naseer, T. Yasir, A. Azhar, T. Shakeel, and K. Zafar, "Computer-Aided Brain Tumor Diagnosis: Performance Evaluation of Deep Learner CNN Using Augmented Brain MRI," *Int. J. Biomed. Imaging*, vol. 2021, 2021, doi: 10.1155/2021/5513500.
- [14] S. H. Khan, "A Novel Feature Map Enhancement Technique Integrating Residual CNN and Transformer for Alzheimer Diseases Diagnosis," Mar. 2024, Accessed: Jul. 23, 2026. [Online]. Available: <https://arxiv.org/pdf/2405.12986>

- [15] S. H. Khan and R. Iqbal, "A Comprehensive Survey on Architectural Advances in Deep CNNs: Challenges, Applications, and Emerging Research Directions," Mar. 2025, Accessed: Jul. 23, 2026. [Online]. Available: <https://arxiv.org/pdf/2503.16546>
- [16] S. H. Khan, R. Iqbal, and S. Naz, "A Recent Survey of the Advancements in Deep Learning Techniques for Monkeypox Disease Detection," Nov. 2023, Accessed: Jul. 23, 2026. [Online]. Available: <https://arxiv.org/pdf/2311.10754>
- [17] S. H. Khan *et al.*, "COVID-19 infection analysis framework using novel boosted CNNs and radiological images," *Sci. Rep.*, vol. 13, no. 1, Dec. 2023, doi: 10.1038/S41598-023-49218-7.
- [18] S. H. Khan and R. Iqbal, "RS-FME-SwinT: A Novel Feature Map Enhancement Framework Integrating Customized SwinT with Residual and Spatial CNN for Monkeypox Diagnosis," Oct. 2024, Accessed: Jul. 23, 2026. [Online]. Available: <https://arxiv.org/pdf/2410.01216>
- [19] S. Sahoo, S. Mishra, B. Panda, A. K. Bhoi, and P. Barsocchi, "An Augmented Modulated Deep Learning Based Intelligent Predictive Model for Brain Tumor Detection Using GAN Ensemble," *Sensors (Basel)*, vol. 23, no. 15, p. 6930, Aug. 2023, doi: 10.3390/S23156930.
- [20] R. Vallée *et al.*, "Machine learning decision tree models for multiclass classification of common malignant brain tumors using perfusion and spectroscopy MRI data," *Front. Oncol.*, vol. 13, no. August, pp. 1–12, 2023, doi: 10.3389/fonc.2023.1089998.
- [21] R. Ahsan, I. Shahzadi, F. Najeeb, and H. Omer, "Brain tumor detection and segmentation using deep learning," *Magn. Reson. Mater. Physics, Biol. Med.*, no. September, 2024, doi: 10.1007/s10334-024-01203-5.
- [22] S. Naz and S. H. Khan, "Residual-SwinCA-Net: A Channel-Aware Integrated Residual CNN-Swin Transformer for Malignant Lesion Segmentation in BUSI," Dec. 2025, Accessed: Jul. 23, 2026. [Online]. Available: <https://arxiv.org/pdf/2512.08243>
- [23] M. Aggarwal, A. K. Tiwari, M. P. Sarathi, and A. Bijalwan, "An early detection and segmentation of Brain Tumor using Deep Neural Network," *BMC Med. Inform. Decis. Mak.*, vol. 23, no. 1, pp. 1–12, 2023, doi: 10.1186/s12911-023-02174-8.
- [24] A. Abdusalomov, M. Rakhimov, J. Karimberdiyev, G. Belalova, and Y. I. Cho, "Enhancing Automated Brain Tumor Detection Accuracy Using Artificial Intelligence Approaches for Healthcare Environments," *Bioengineering*, vol. 11, no. 6, 2024, doi: 10.3390/bioengineering11060627.
- [25] S. K. Mathivanan, S. Sonaimuthu, S. Murugesan, H. Rajadurai, B. D. Shivahare, and M. A. Shah, "Employing deep learning and transfer learning for accurate brain tumor detection," *Sci. Rep.*, vol. 14, no. 1, pp. 1–15, 2024, doi: 10.1038/s41598-024-57970-7.
- [26] M. U. Javeed *et al.*, "MindMate: An Emotion-Aware Generative AI System for Personalized Mental Health Support," *J. Comput. Biomed. Informatics*, vol. 10, no. 1, Dec. 2025, doi: 10.56979/1001/2025/1167.
- [27] W. J. von Eschenbach, "Transparency and the Black Box Problem: Why We Do Not Trust AI," *Philos. Technol.*, vol. 34, no. 4, pp. 1607–1622, Dec. 2021, doi: 10.1007/S13347-021-00477-0.
- [28] C. Zhang, J. He, and L. Shang, "An X-ray image classification method with fine-grained features for explainable diagnosis of pneumoconiosis," *Pers. Ubiquitous Comput.*, vol. 28, no. 2, pp. 403–415, Apr. 2024, doi: 10.1007/S00779-023-01730-3.

- [29] T. Kim, N. H. Moon, T. S. Goh, and I. D. Jung, "Detection of incomplete atypical femoral fracture on anteroposterior radiographs via explainable artificial intelligence," *Sci. Reports* 2023 131, vol. 13, no. 1, pp. 10415-, Jun. 2023, doi: 10.1038/s41598-023-37560-9.
- [30] T. Hussain, S. H. Khan, U. B. Ejaz, and M. M. Zahoor, "ENHANCED DEBERTA-BASED HADITH UNDERSTANDING FOR ISLAMIC ASSISTANCE," *Spectr. Eng. Sci.*, vol. 3, no. 12, pp. 1736-1747, Dec. 2025, doi: 10.5281/zenodo.18266199.
- [31] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization," *Int. J. Comput. Vis.*, vol. 128, no. 2, pp. 336-359, Oct. 2016, doi: 10.1007/s11263-019-01228-7.
- [32] "(PDF) Impacts of Cloudburst Events on Biodiversity in Pakistan's Northern Areas and Development of a Machine Learning Model for Cloudburst Prediction Article Info." Accessed: Jul. 23, 2026. [Online]. Available: https://www.researchgate.net/publication/396803587_Impacts_of_Cloudburst_Events_on_Biodiversity_in_Pakistan's_Northern_Areas_and_Development_of_a_Machine_Learning_Model_for_Cloudburst_Prediction_Article_Info
- [33] M. U. Javeed, S. M. Aslam, W. Y. Ramay, M. M. Zahoor, Z. Raza, and M. M. Iqbal, "The Role of Artificial Intelligence in Academic Writing: A Literature Review," *Annu. Methodol. Arch. Res. Rev.*, vol. 3, no. 11, pp. 242-258, Nov. 2025, doi: 10.63075/QK7N8H28.
- [34] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A Simple Way to Prevent Neural Networks from Overfitting," *J. Mach. Learn. Res.*, vol. 15, no. 56, pp. 1929-1958, 2014, Accessed: Jun. 06, 2026. [Online]. Available: <http://jmlr.org/papers/v15/srivastava14a.html>
- [35] C. Zhang, S. Bengio, M. Hardt, B. Recht, and O. Vinyals, "Understanding deep learning (still) requires rethinking generalization," *Commun. ACM*, vol. 64, no. 3, pp. 107-115, Mar. 2021, doi: 10.1145/3446776;GROUPTOPIC:TOPI C:ACM-PUBTYPE.
- [36] M. Y. Wani, M. Zaffar, M. M. Zahoor, S. A. Rathore, and T. Abbas, "AUTOMATED RETINAL BLOOD VESSEL SEGMENTATION VIA U-NET AND VGG-BASED MODELS," *Spectr. Eng. Sci.*, pp. 490-498, Apr. 2025, doi: 10.5281/zenodo.15227615.
- [37] "Brain Tumor MRI Dataset."
- [38] M. M. Zahoor and S. H. Khan, "CE-RS-SBCIT A Novel Channel Enhanced Hybrid CNN Transformer with Residual, Spatial, and Boundary-Aware Learning for Brain Tumor MRI Analysis," Aug. 2025, Accessed: Jul. 23, 2026. [Online]. Available: <https://arxiv.org/pdf/2508.17128>
- [39] "Understanding VGG: A Deep Dive into Visual Geometry Group's Convolutional Neural Network | SERP AI."
- [40] Y. Chen, L. Li, W. Li, Q. Guo, Z. Du, and Z. Xu, "Deep learning processors," *AI Comput. Syst.*, pp. 207-245, 2024, doi: 10.1016/B978-0-32-395399-3.00012-3.
- [41] "AlexNet: The Architecture that Revolutionized Computer Vision | by Pranay Rishith | DataDrivenInvestor."
- [42] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size," Feb. 2016.

- [43] "What Is ResNet-18? How to Use the Lightweight CNN Model."
- [44] M. M. Zahoor *et al.*, "Brain Tumor MRI Classification Using a Novel Deep Residual and Regional CNN," *Biomed.* 2024, Vol. 12, Page 1395, vol. 12, no. 7, p. 1395, Jun. 2024, doi: 10.3390/BIMEDICINES12071395.
- [45] "(PDF) Deep Learning in Hematology: Automated Counting of Blood Cells Using YOLOv5 Object Detection Article Info." Accessed: Jul. 23, 2026. [Online]. Available: https://www.researchgate.net/publication/403061607_Deep_Learning_in_Hematology_Automated_Counting_of_Blood_Cells_Using_YOLOv5_Object_Detection_Article_Info
- [46] M. Zaffar *et al.*, "A NOVEL FEATURE SELECTION METHOD FOR PREDICTION OF FACTORS AFFECTING ANXIETY IN HEALTH CARE WORKERS DURING PANDEMIC," *Spectr. Eng. Sci.*, vol. 4, no. 2, pp. 331-353, Feb. 2026, doi: 10.5281/zenodo.18710341.
- [47] D. P. Kingma and J. L. Ba, "Adam: A Method for Stochastic Optimization," *3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc.*, Dec. 2014, Accessed: Jun. 06, 2026. [Online]. Available: <https://arxiv.org/pdf/1412.6980>
- [48] R. Iqbal and S. H. Khan, "RSwinV2-MD: An Enhanced Residual SwinV2 Transformer for Monkeypox Detection from Skin Images," Jan. 2026, Accessed: Jul. 23, 2026. [Online]. Available: <https://arxiv.org/pdf/2601.01835>
- [49] S. H. Khan, "Advanced Hybrid Transformer LSTM Technique with Attention and TS Mixer for Drilling Rate of Penetration Prediction," Aug. 2025, Accessed: Jul. 23, 2026. [Online]. Available: <https://arxiv.org/pdf/2508.05210>
- [50] S. Naz, S. H. Khan, R. Iqbal, and M. Safiullah, "SWINUNET: A HYBRID SWIN TRANSFORMER-CNN ARCHITECTURE WITH ADAPTIVE FEATURE FUSION FOR BREAST ULTRASOUND LESION SEGMENTATION," *Spectr. Eng. Sci.*, vol. 4, no. 1, pp. 664-676, Jan. 2026, doi: 10.5281/zenodo.18383193.