

Attention-Driven Sequential Feature Fusion Framework for Effective Brain Tumor Diagnosis

ISSN: 2637-8078



***Corresponding author:** Ifza Shad, School of Computer Science and Engineering, Central South University, Changsha, China

Submission: 📅 April 09, 2025

Published: 📅 May 21, 2025

Volume 7 - Issue 3

How to cite this article: Ifza Shad*, Omair Bilal and Arash Hekmat. Attention-Driven Sequential Feature Fusion Framework for Effective Brain Tumor Diagnosis. Significances Bioeng Biosci. 7(3). SBB. 000662. 2025.

DOI: [10.31031/SBB.2025.07.000662](https://doi.org/10.31031/SBB.2025.07.000662)

Copyright@ Ifza Shad, This article is distributed under the terms of the Creative Commons Attribution 4.0 International License, which permits unrestricted use and redistribution provided that the original author and source are credited.

Ifza Shad*, Omair Bilal and Arash Hekmat

School of Computer Science and Engineering, Central South University, China

Abstract

Brain tumors are cells with abnormal growth patterns within the brain. They can appear in various forms and affect different areas of the brain. Brain tumors result from uncontrolled cell growth and are a leading cause of adult mortality. Early detection is crucial and significantly improves patient survival rates. MRI is effective for early detection, but traditional human inspection methods are inefficient for processing large datasets and time-consuming to handle large amounts of data. Our research proposes an advanced attention-based feature fusion method to address this issue. We utilized DenseNet201 and Xception as our base models due to their exceptional performance and proven efficacy in medical image classification tasks. These models excel at extracting robust and high-dimensional features, which are integrated into an advanced feature fusion framework to enhance diagnostic precision. This framework incorporates ConvLSTM layers and Convolutional Block Attention Modules (CBAM) to enhance feature attention and capture sequential dependencies. Our fusion technique leverages the strengths of these base models, further enhanced by CBAM attention mechanism. This approach represents a substantial advancement in medical imaging and provides an effective tool for diagnosing complex conditions. Our proposed model shows superior performance when evaluated on a widely recognized brain tumor dataset. It succeeds with an inspiring accuracy of 98.83%. This high accuracy demonstrates the potential of our method to advance early brain tumor detection significantly. By reducing misclassification rates, our model provides a reliable and efficient diagnostic tool that can improve patient outcomes. Additionally, we examined the generalizability of our proposed model on an additional Kaggle dataset consisting of 4 classes, which produced highly promising and effective results. In addition, visualization techniques such as Grad-CAM, feature maps, ROC curve and confusion matrix analysis are employed to interpret the model's decisions and validate its effectiveness.

Keywords: Brain tumor; Classification; Feature fusion; Attention mechanism; Deep learning

Introduction

Brain tumors present a substantial obstacle to medical diagnostics. They require timely and precise identification to facilitate effective therapy. This introduction provides background information. It begins with an outline of the symptoms of brain tumors and the need for early identification [1]. These areas include cognitive functioning and general conditions. Understanding this diversity is essential for creating effective diagnostic methods. Knowledge of the unique attributes of brain tumors is crucial for developing efficient detection methods [2]. An early diagnosis is vital for improving cure outcomes and patient prognosis. The best way to find brain tumors is with Magnetic Resonance Imaging (MRI), which is based on magnetism and provides high-resolution images of soft brain tissues without harming the individual. However, our research proposes advanced methods that can further enhance the efficiency of MRI [3]. The primary function of MRI is to identify and characterize brain tumors. This section delves into the specifics of this application [4]. Conventional approaches rely on the visual inspection of medical images, which is a time-consuming process prone to human error. The manual examination of medical images is the cornerstone of traditional methods, which can be laborious and potentially result in inaccuracies [5]. In recent years, the medical industry has undergone substantial transformations due to technological advancements. Artificial Intelligence (AI) has played a pivotal role in this evolution. AI has become a vital

component in healthcare's continuous digital transformation process. It contributes significantly to the rapid progress observed in this field. Brain tumors present a substantial obstacle in medicine. Accurate identification, treatment, and surveillance are crucial for tackling this problem [6]. AI can transform brain tumor care by enabling earlier detection and more efficient treatment methods [7]. Because of their location, brain tumors pose a threat to human health. AI systems help doctors determine tumor size, location, classification and aggressiveness. This information aids doctors in making more accurate diagnoses and treatments. It also allows patients to understand their health [8]. To a greater extent, Machine Learning (ML) and Deep Learning (DL) methods are being used to understand and grade medical images, notably brain tumor imaging. Advanced computational methods can improve diagnostic accuracy and lower healthcare costs. ML systems can detect patterns in medical photos without human input. DL is a promising way to automate brain tumor identification, enabling early diagnosis in resource-limited areas where brain tumors are common [9]. Deep learning is advancing rapidly in the field of the field of computer vision. This allows the creation of automatic and accurate brain tumor detection methods. Convolutional Neural Networks (CNNs) are one of the most powerful and efficient deep learning methods for image processing including object detection and segmentation. Researchers are testing CNNs for brain tumor classification and detection using MRI data. They use deep learning to improve diagnosis and therapy [10]. DL is capable of quickly learning complex and organized features from raw data, removing the need for rule-based or human feature engineering. Automation of medical results simplifies processes and may improve healthcare quality. Automatic feature extraction and pattern recognition make DL a transformational tool for many situations [11].

Motivation and major contributions

The key contributions of this study are:

- a. Our primary contribution includes the development of a novel feature fusion framework that combines features extracted from two pre-trained models such as Xception and DenseNet 201. This cutting-edge framework has proven to be beneficial in improving the accuracy of MRI-based brain tumor detection. Feature-level fusion can greatly enhance model performance and generalizability by optimizing feature representation.
- b. Features of the backbone model are enhanced using a Convolutional Block Attention Module (CBAM). We have integrated ConvLSTM layers and CBAM blocks into our base models. This allows the models to recognize sequential dependencies and concentrate on essential image areas, enhancing the accuracy of diagnostics.
- c. We employ Grad-CAM to visualize the model's attention, revealing the specific regions that significantly influence predictions. This enhances the model's interpretability by providing clear insights into its decision-making process.

The basic layout of this paper is as follows: Section 3 provides a detailed description of the materials and methods used in the

current research. We present our proposed model and framework components in detail. Section 4 examines and discusses the experimental outcomes. The ablation study also underwent extensive discussion. Section 5 addresses future work and limits extensively. Section 6 reveals the conclusion.

Literature Study

This section explores the various applications of ML and DL techniques in studying virulent brain tumors and interpreting medical ideas. Medical image processing has advanced significantly in the past two decades, saving significant attention and research interest due to its vast potential in healthcare, especially in patient examination and diagnosis. Numerous studies have proposed ML-based approaches for classifying brain images and analyzing brain structures [10]. Machine learning algorithms have proven invaluable in this domain, enabling researchers to develop advanced models for analyzing complex medical data. These techniques have facilitated the identification of patterns and features that may be difficult to discern through traditional methods. This ultimately contributes to the precise identification and management of tumors in the brain. Abd-Ellah et al. [12] proposed a novel semi-automatic segmentation method, showing their dedication to brain tumor analysis innovations. This research uses cutting-edge DL and ML to advance brain tumor diagnosis and treatment. Yushkevich et al. [13] Novel Convolutional Neural Network (CNN) architecture to classify brain cancer images. By modifying the ResNet50 architecture, they removed the last five layers and added eight new layers. This method aimed to diagnose tumors in brain MRI images and achieved an accuracy of 97.2% of the time when diagnosing brain tumors. Çinar et al. [14] evaluated other CNN models, including AlexNet, DenseNet201, InceptionV3, and GoogLeNet, and utilized the model with the highest performance for classifying brain tumor images. Another research Latif G [15] proposed a way to sort glioma tumors into numerous classes using DL attributes and a Support Vector Machine (SVM) classifier. A deep CNN extracted features from MR images and fed them into the SVM classifier. The proposed practice achieved 96.19% accurate for the HGG Glioma class using the FLAIR modality and 95.46% accuracy using the T2 mode to get accurate results for the LGG Glioma tumor type, classifying four Glioma classes.

A CNN was developed by Kumar et al. [16] with the aim of detecting brain tumors using Conv2D with Leaky ReLU, achieving a validation accuracy of 78.57%. This performance was higher than the conventional model's 99.20% training accuracy using Conv2D with ReLU. Both models had the same training accuracy score of 99.20%. Yahyaoui et al. [17] proposed a study on semantic classification and fusion of 2D and 3D brain images. The DenseNet model classified 2D images into Glioma, Meningioma, and Pituitary tumors with 92.06% accuracy. The 3D-CNN model graded gliomas as high or low-grade using 3D images, achieving 85% accuracy. Mahmud et al. [18] discuss and compare other models including Inception V3, ResNet-50, and VGG16 with the proposed novel Correlation Learning Mechanism (CLM) that integrates a CNN with a classic architecture. Using a dataset of 3264 MR images, the

proposed model achieved superior performance with an accuracy of 93.3%. Woźniak et al. [19] The support neural network optimizes the CNN by identifying the best filters for pooling and convolution layers, enhancing the primary classifier's learning speed and efficiency. The proposed CLM model achieves approximately 96% accuracy. Mandloi et al. [20] This work introduces an explainable brain tumor detection and classification model using pre-trained models. A conditional Generative Adversarial Network (cGAN) addresses data imbalance and overfitting, while models like MobileNet, InceptionResNet, EfficientNet, and VGGNet perform detection and classification. InceptionResNetV2 achieved detection accuracies of 99.6% while EfficientNet-B0 excelled in classification with accuracies of 99.3%.

Another study proposes a Parallel Deep Convolutional Neural Network (PDCNN) topology to capture global and local features simultaneously. It uses dropout regularization and batch normalization. Evaluated on three MRI datasets, the method achieved 97.33% accuracy on a binary tumor identification dataset, 97.60% on the Figshare dataset I and II, and 98.12% on the Multiclass Kaggle dataset [21]. Another approach Gayathri et al.

[22] The fine-tuned VGG-16 CNN model, trained on 3253 brain MRI images, achieved 94% accuracy after hyperparameter optimization. It demonstrated strong performance, often outperforming other techniques and showing high sensitivity and specificity. In Table 1, we present a summary of prior research [15,17-25]. After an extensive review of existing methods, we identified limitations such as suboptimal accuracy rates high rates of missed detections and lack of robustness and generalization in current brain tumor detection and classification approaches. We have developed a novel model to achieve the highest possible accuracy precise detection with minimal missed detections and robustness and generalization across diverse datasets and real-world scenarios to address these deficiencies. Our model uses advanced methods like data preprocessing, feature extraction, patio-temporal dependencies, attention mechanisms and ensemble learning to find and classify brain tumors better than others. The results demonstrate the superiority of our approach in terms of accuracy, precision, recall, and overall performance. The subsequent sections of this work will provide detailed evaluation metrics and comparisons with existing methods.

Table 1: Summary of prior research.

Methodology	Performance	Limitations	Reference	Year
Semi-automatic brain tumor using T1W configuration	Improved MRI scan accuracy and efficiency	Requires manual intervention	[15]	2023
Modified ResNet50, evaluated other CNN models	97.2% accuracy in brain tumor diagnosis	Limited generalization	[17]	2021
Multi-class Glioma classification with SVM classifier	96.19% (HGG), 95.46% (LGG) accuracy	Computationally intensive	[18]	2020
CNN with 'Leaky ReLU' activation	78.57% validation accuracy	Lower validation accuracy	[19]	2019
Semantic classification and fusion of 2D and 3D images	DenseNet: 92.06%, 3D-CNN: 85%,	Increased computational requirements	[20]	2021
Comparison of various models (ResNet-50, VGG16, Inception V3)	Proposed CNN: 93.3% accuracy	Performance may vary with different datasets	[21]	2022
Correlation Learning Mechanism (CLM) combining CNN with support neural network	96% accuracy	Increased model complexity	[22]	2020
Explainable brain tumor detection using pre-trained models	InceptionResNetV2: 99.2%	Dependence on pre-trained models	[23]	2023
Parallel Deep Convolutional Neural Network (PDCNN) topology	97.33% (binary), 97.60% (Figshare), 98.12% (Kaggle)	Increased model complexity	[24]	2021
Fine-tuned VGG-16 CNN model	94% accuracy after hyperparameter optimization	Limited generalization	[25]	2022

Materials and Methods

Dataset description

In this research, a publicly available brain tumor dataset from Kaggle [24], consisting of 3,000 image samples was utilized. The dataset includes two classes of brain tumors: Tumor and No-Tumor. Additionally, (Table 2) summarizes the number of samples within

each class. We split the dataset into training and testing by a ratio of 80:20. Figure 1 illustrates the distribution of these classes.

Table 2: Class-wise dataset distribution.

Category	Train	Test	Total Samples
Tumor	1200	300	1500
No-Tumor	1200	300	1500

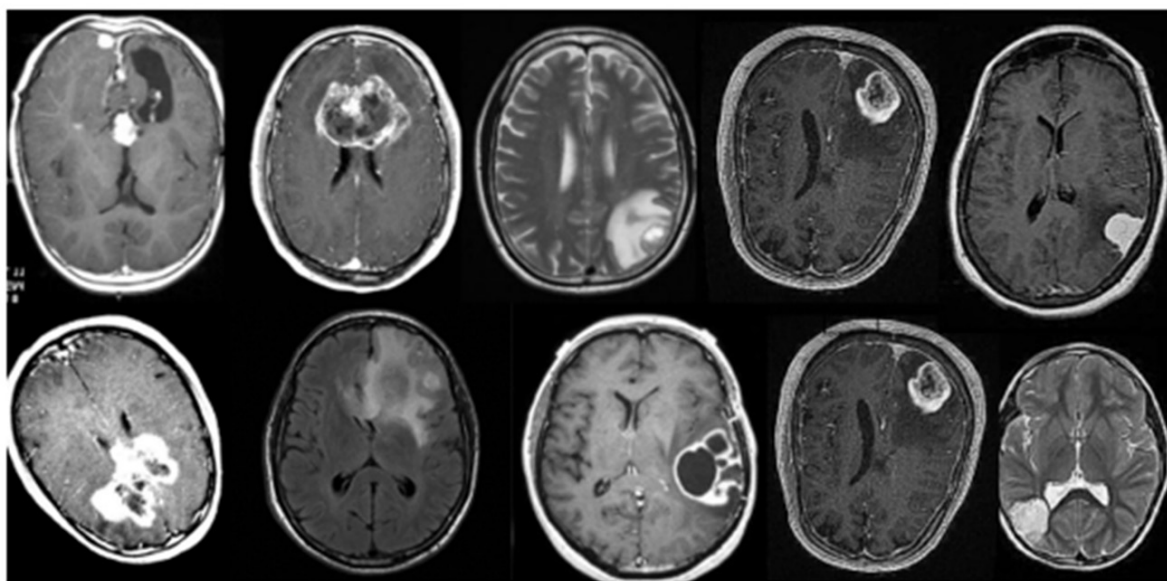


Figure 1: Brain tumor dataset sample images.

MRI dataset preprocessing

The BR35H dataset underwent preprocessing to enhance the performance of our proposed architecture. Brain MRI data often contains background noise, making normalization a critical step. Since deep learning models like DenseNet201 and Xception require input images of specific dimensions, all images were resized to 224x224 pixels to align with the input specifications of these models, which are optimized for square formats. This resizing step not only improves computational efficiency but also promotes generalization and facilitates the use of pre-trained models, thereby boosting overall model performance. Instead of collecting new data, our approach focused on refining the existing dataset to provide a more diverse and comprehensive training set, which aimed to enhance the classification capabilities of the deep learning model. During rescaling, each image was divided by 255 to ensure pixel values ranged from 0 to 1, standardizing the images for model processing. Normalization was employed to maintain data consistency across all images, justifying model bias toward specific intensity ranges and enabling the model to learn meaningful patterns more effectively. Shuffling was implemented to prevent the model from learning any sequence-related biases in the training data, reducing the risk of overfitting. By randomly shuffling the images prior to training, the model's learning process was made more robust and reliable, contributing to stable training outcomes.

Pre-trained CNN models

This study aims to develop a robust framework for accurately diagnosing brain tumors by integrating features from two pre-trained CNN models, DenseNet201 and Xception. The extracted features from these models are concatenated to form a comprehensive feature set. These concatenated features are then processed using a ConvLSTM layer, which combines convolutional and LSTM layers to capture both spatial and temporal dependencies in the data. This helps in retaining crucial information across

different frames, boosting the accuracy of tumor detection and classification using this model. Additionally, a CBAM (Convolutional Block Attention Module) is incorporated to refine the feature selection process further. The CBAM module applies both channel and spatial attention mechanisms to highlight the most important features while suppressing less significant ones. By focusing on the most relevant features, the model's performance is remarkably superior. Finally, these refined and fused features are used for the classification task, resulting in a more accurate and reliable brain tumor diagnosis.

DenseNet201: The DenseNet-201 convolutional neural network design broadens the original model and proposed by Salim F, et al. [25]. Its primary feature is dense feed-forward connectivity between layers. The architecture is made up of various dense blocks with direct-layer connections. Transition layers customize feature map sizes by convolution and pooling between dense blocks. A global average pooling and classification layer follows the final dense block. The number of layers in each dense block consists of batch normalization, ReLU activation, and 3x3 convolution. DenseNet-201 additionally includes bottleneck layers with 1x1 convolutions to reduce computational complexity. The growth rate (k) hyperparameter determines how many new feature mappings each layer creates. It promotes network feature propagation and reuse [26]. The transition layers between dense blocks reduce the number of feature maps using 1x1 convolution and pooling allowing more dense blocks to be stacked while keeping the model compact. Unlike ResNets, we concatenate features instead of summarizing them before passing them into a layer. Therefore, the 'layer' receives feature maps from all previous convolutional blocks. Its feature maps are available to all subsequent L-l layers. This equation adds $L(L+1)$ connections, not the usual L connections found in typical structures. Due to its dense connectivity pattern, the method uses a Dense Convolutional Network [27]. Figure 2 illustrates the architectural overview of the DenseNet201 model.

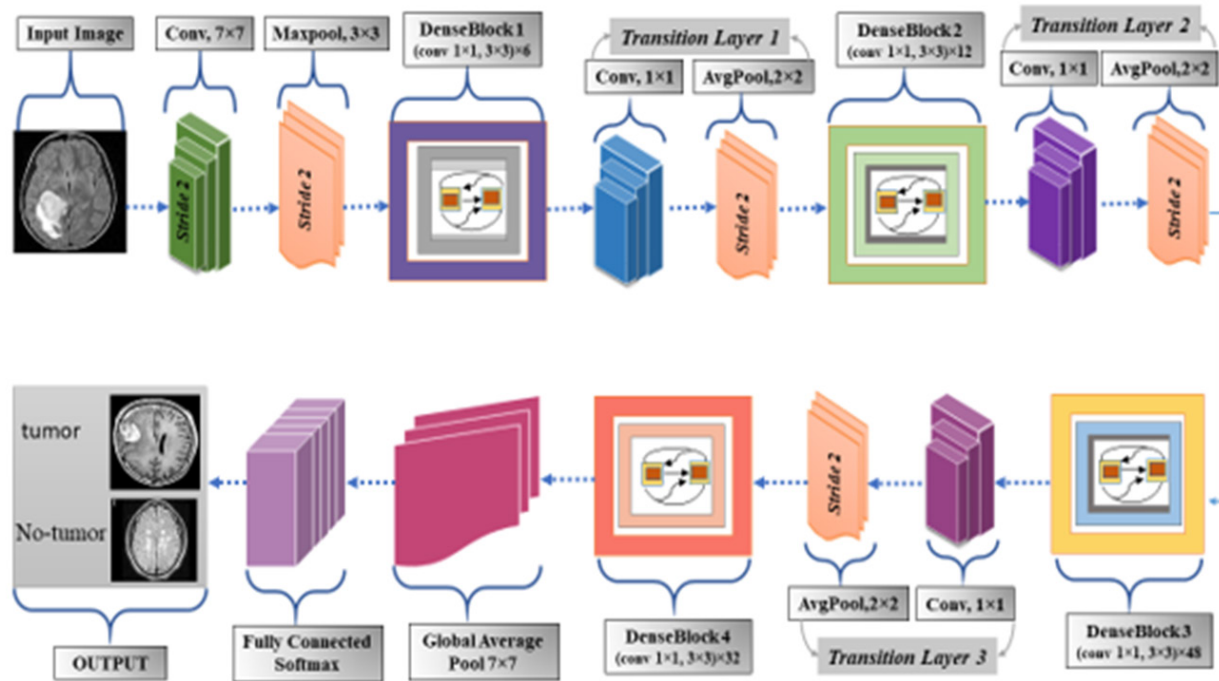


Figure 2: Architecture of DenseNet201.

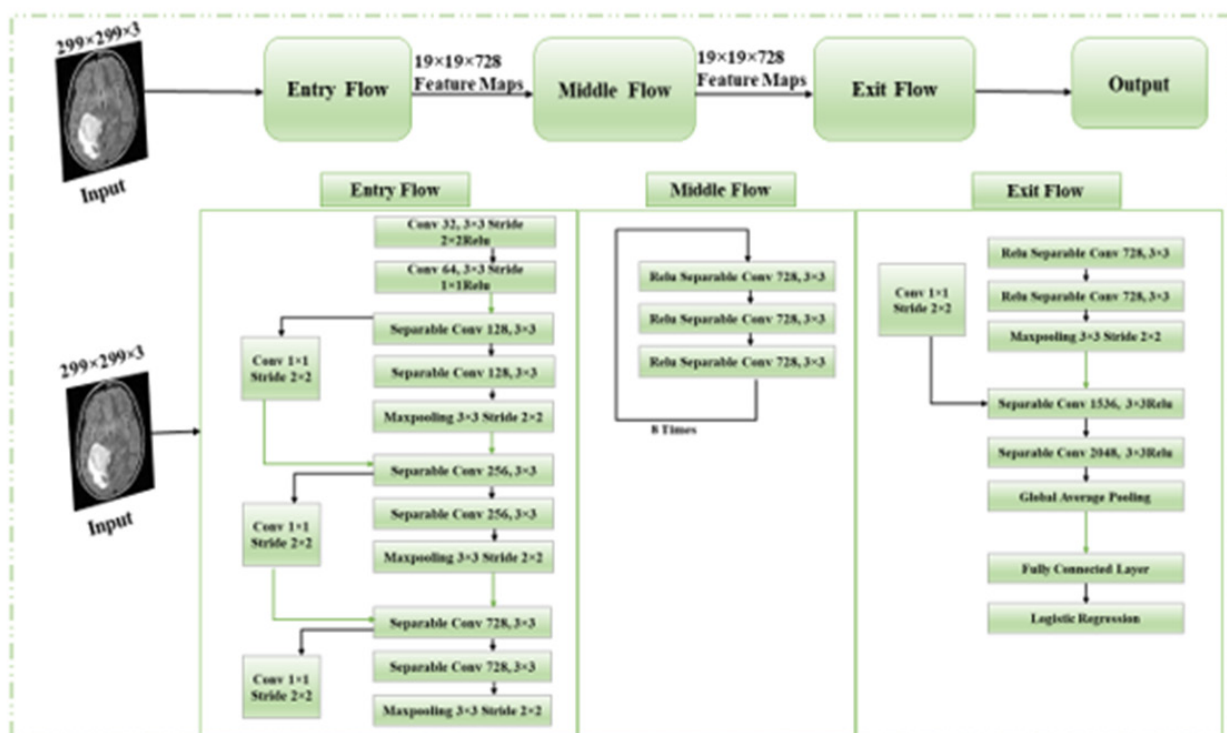


Figure 3: Architecture of Xception.

Xception: For deep convolutional neural networks, Xception uses only depth-wise separable convolution layers. François Chollet [28] suggested it in 2017. The Xception architecture represents a significant advancement in Convolutional Neural Networks (CNNs), primarily focused on convolutions that can be separated by depth to boost effectiveness and efficiency in deep learning models.

This design aims to improve the feature extraction process by distinctly separating spatial and channel-wise convolutions, which effectively reduces the cost of computation without compromising how expressive a model is. The architecture consists of several depth-based convolutions followed by point-based convolutions. Depthwise convolutions apply a single convolutional filter per

input channel, allowing the model to process spatial information more efficiently. Pointwise convolutions, on the other hand, use a 1x1 convolution to combine the outputs of the depthwise convolutions, enabling the network to learn complex feature representations across channels. This combination provides a robust mechanism for useful feature extraction and learning how to use hierarchical models. By separating spatial and channel-wise operations, Xception significantly enhances model efficiency. This design reduces the number of parameters and the amount of computation required, leading to faster inference and training times [29]. Despite the reduced complexity, the model maintains high performance and expressive power, making it highly effective for a wide range of computer vision tasks, such as image classification, object detection, and segmentation. The innovative approach of the Xception architecture in combining depthwise and pointwise convolutions contributes to its superior performance compared to traditional CNNs. The fact that it is simple to extract, and process features while maintaining high accuracy makes it a valuable tool in various applications, particularly where computational resources are limited or where rapid processing is crucial. Figure 3 provides a concise summary of the architectural structure of the Xception model.

ConvLSTM

Hochreiter and Schmidhuber first introduced the Long Short-Term Memory (LSTM) model in 1997 [30]. LSTMs are a specific kind of Recurrent Neural Network (RNN) that are specifically built to represent and comprehend long-term relationships in sequential data precisely. The network accomplishes this by employing a system of gates (namely, forget, input, and output) that govern the transmission of information allowing for the linking to store or discard information as needed. LSTMs play a vital role in tasks that involve the comprehension of temporal or sequential data, such as speech recognition, language modeling, and time-series prediction. LSTMs networks are involved in image processing for tasks such as image captioning video analysis and action detection. These applications require the ability to preserve temporal information. The LSTM can selectively incorporate or ignore data in or out of the cell state. Structures known as gates intentionally manage this data. The structure known as a gate controls the entry of data into the cell state. A gate is formed by combining a sigmoid function with a point-wise multiplication operation. The sigmoid function can produce any value between zero and one. In an LSTM model various gates are employed to transfer our most recent data from one cell to another cell [31]. The architecture overview of the LSTM model is illustrated in Figure 4.

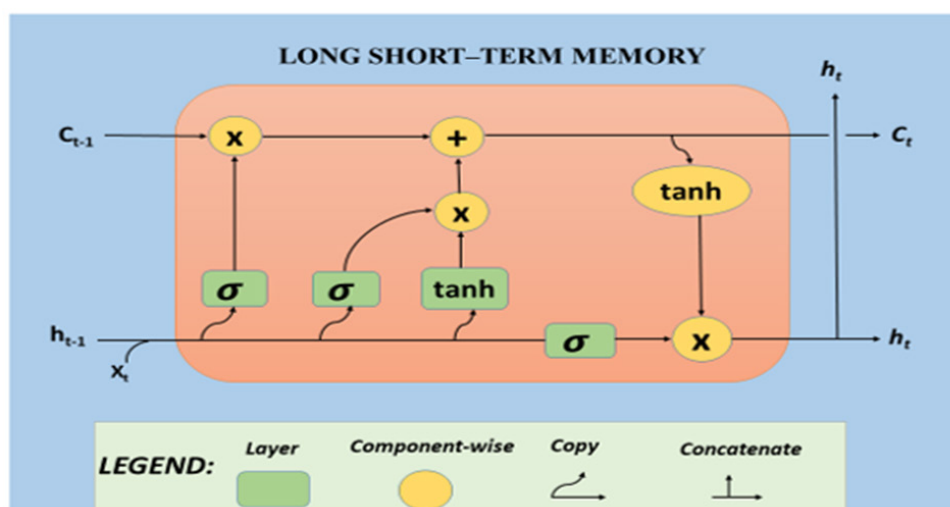


Figure 4: Workflow of LSTM architecture.

Convolutional block attention module (CBAM)

In 2018, Sanghyun Woo and his colleges presented Convolutional Block Attention Module (CBAM) at the European Conference on Computer Vision (ECCV) [32]. Channel and spatial attention are used in this module to improve CNN representation by focusing on key features. The CBAM module improves Convolutional Neural Networks (CNNs) by adding ways to pay attention to mechanisms in both the channel and spatial dimensions. The system comprises a Spatial and Channel Attention Module, which operates in sequence to enhance significant elements and diminish less significant ones. The incorporation of this dual attention method boosts the model's feature representation, resulting in improved performance spanning tasks such as picture classification, object identification,

and image segmentation. The Channel Attention Module employs global average and max pooling, which are then integrated using a shared MLP, and the sigmoid function is used to create a channel focus map: $M_c(F) = \sigma(\text{MLP}(\text{AvgPool}(F)) + \text{MLP}(\text{MaxPool}(F)))$.

The Spatial Attention Module utilizes channel-wise pooling, concatenates the outcomes, and employs a convolution layer to generate a spatial attention map: $M_s(F') = \sigma(\text{Conv}([\text{AvgPool}(F'); \text{MaxPool}(F')]))$. It possesses a low weight and may be seamlessly incorporated into pre-existing structures. CBAM effectively highlights crucial features, hence improving performance in applications such as picture classification and object recognition [33]. (Figure 5) depicts the architectural overview of the CBAM module.

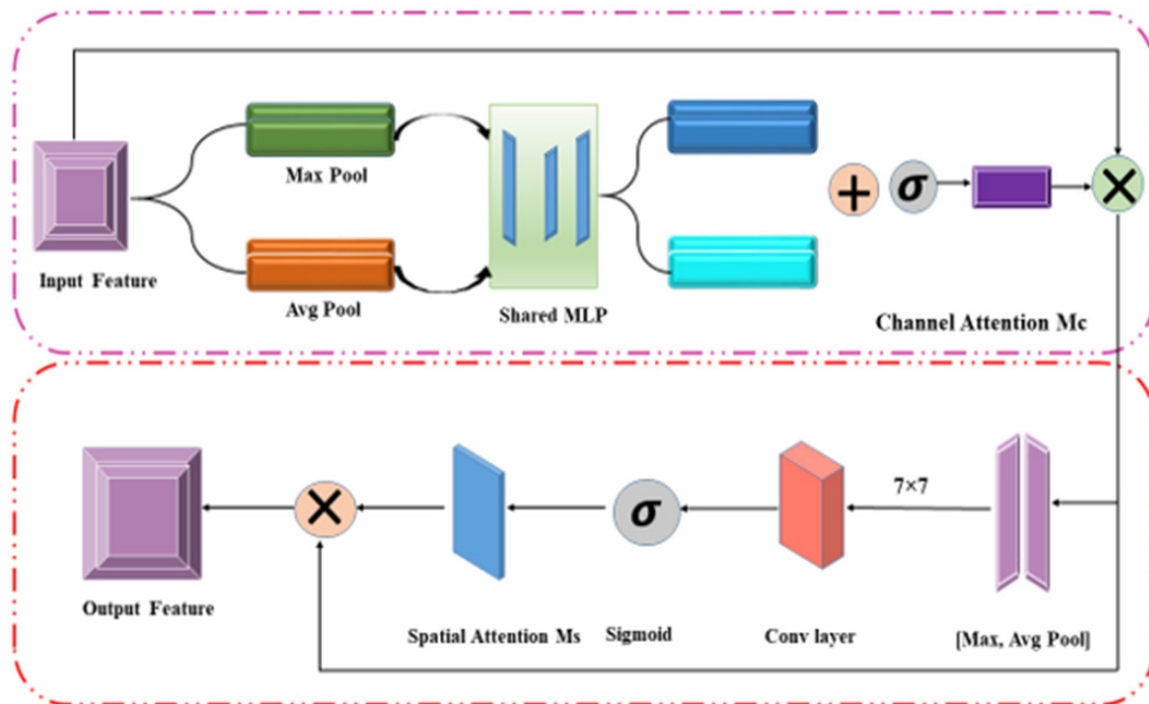


Figure 5: CBAM (Convolutional Block Attention Module) architecture.

Feature fusion

Feature fusion is a process of concatenating features from distinct models. It is frequently executed using basic operations such as addition or joining. However, this may not be the optimal selection [34]. Feature fusion is a widely used technique that has been independently created and utilized by multiple researchers in different domains for a long time. Feature fusion ideas have been used in various domains, including sensor fusion during the 1980s and 1990s and subsequently in computer vision and multimodal

learning during the 2000s and 2010s. Feature fusion is a notion that has developed throughout time in the domains of machine learning and data fusion. It lacks a singular founder but has been cultivated and improved by numerous scholars from different fields. The concept of integrating characteristics from several origins to enhance the effectiveness of a model has been intensively investigated in domains such as pattern recognition, computer vision, and multimodal data analysis [35]. Figure 6 illustrates the operational overview of the feature fusion.

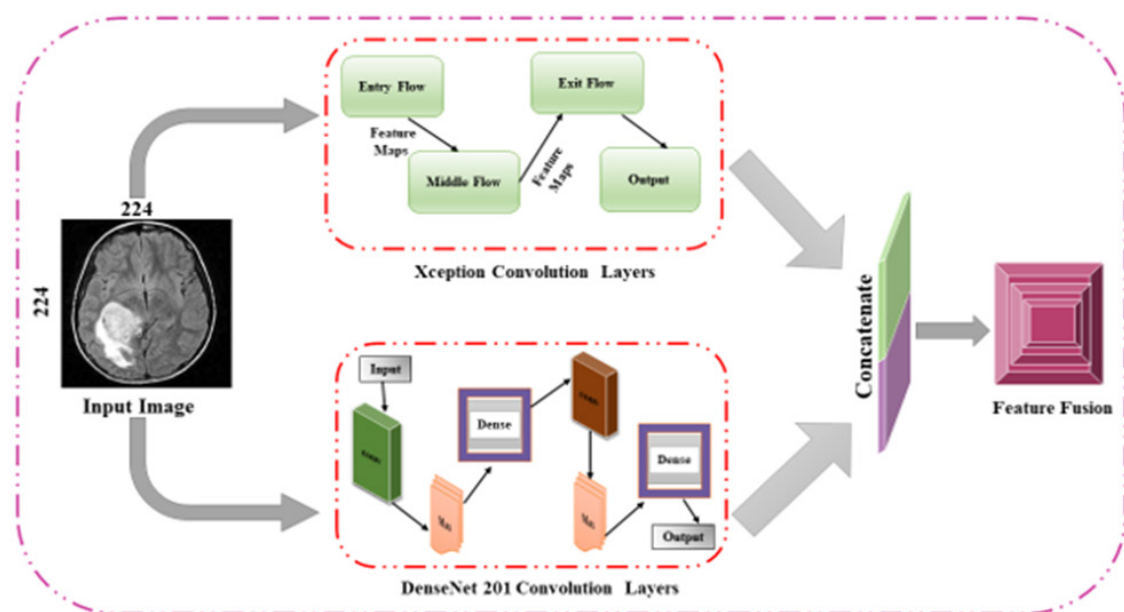


Figure 6: Overview of feature fusion approach.

Proposed architecture

We first employed two top-performing pre-trained models, DenseNet201 and Xception, to extract features from the input data. DenseNet201 was chosen for its dense connectivity pattern, which ensures maximum information flow between layers and improves gradient propagation. Xception was selected for its efficient depthwise separable convolutions, which enhance feature extraction by separating spatial and channel-wise operations, reducing computational complexity while preserving high expressive power. These models were fine-tuned on our specific dataset to ensure optimal feature extraction tailored to our task. After extracting features from DenseNet201 and Xception, we utilized a ConvLSTM layer to process these features. The ConvLSTM layer combines convolutional operations with Long Short-Term Memory (LSTM) networks, making it particularly effective for tasks requiring both spatial and temporal feature extraction. This layer captures spatial dependencies within each frame through convolutional operations and temporal dependencies across frames through LSTM units. By preserving spatio-temporal dependencies, the ConvLSTM layer enhances the model's ability to understand complex patterns and sequences in the data, which is crucial for accurately diagnosing brain tumors. Subsequently, we incorporated the CBAM attention block to refine the extracted features further. The CBAM applies to both channel and spatial attention mechanisms to emphasize the most relevant features

and suppress the fewer vital ones. The channel attention module focuses on the inter-channel relationships, highlighting significant feature maps, while the spatial attention module concentrates on the spatial relationships within each feature map.

By integrating these two attention mechanisms, CBAM enhances the representational power of the features, ensuring that the model focuses on the most informative aspects of the input data. This mechanism improves model's accuracy and robustness by effectively capturing essential features for the classification task. Finally, the refined features obtained from the CBAM module are used for classification. The classification layer, typically a fully connected layer followed by a SoftMax activation function, processes these features to predict the presence and type of brain tumor. This step converts the high-dimensional feature representations into class probabilities, allowing for accurate and reliable diagnosis. By leveraging the strengths of DenseNet201, Xception, ConvLSTM, and CBAM, our proposed methodology achieves high performance in brain tumor classification, ensuring precise and efficient diagnostic capabilities. Figure 7 illustrates the overall methodology framework of proposed model. The improved characteristics are then used to categorize the MRI images into two distinct classes such as tumor and no-tumor. This stage uses DL models that feature fusion and attention mechanisms to accurately detect if there is a tumor in the brain resulting in a binary classification. Figure 8 is illustrated as the proposed pipeline for brain tumor detection.

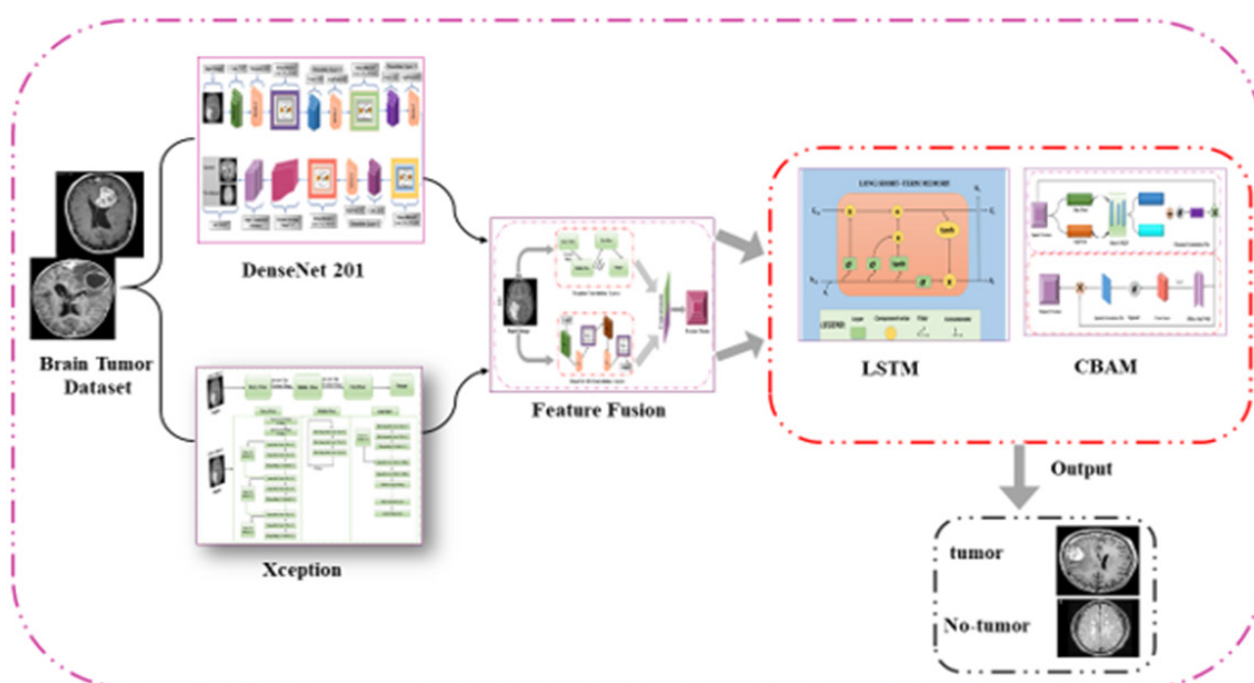


Figure 7: Framework of proposed methodology.

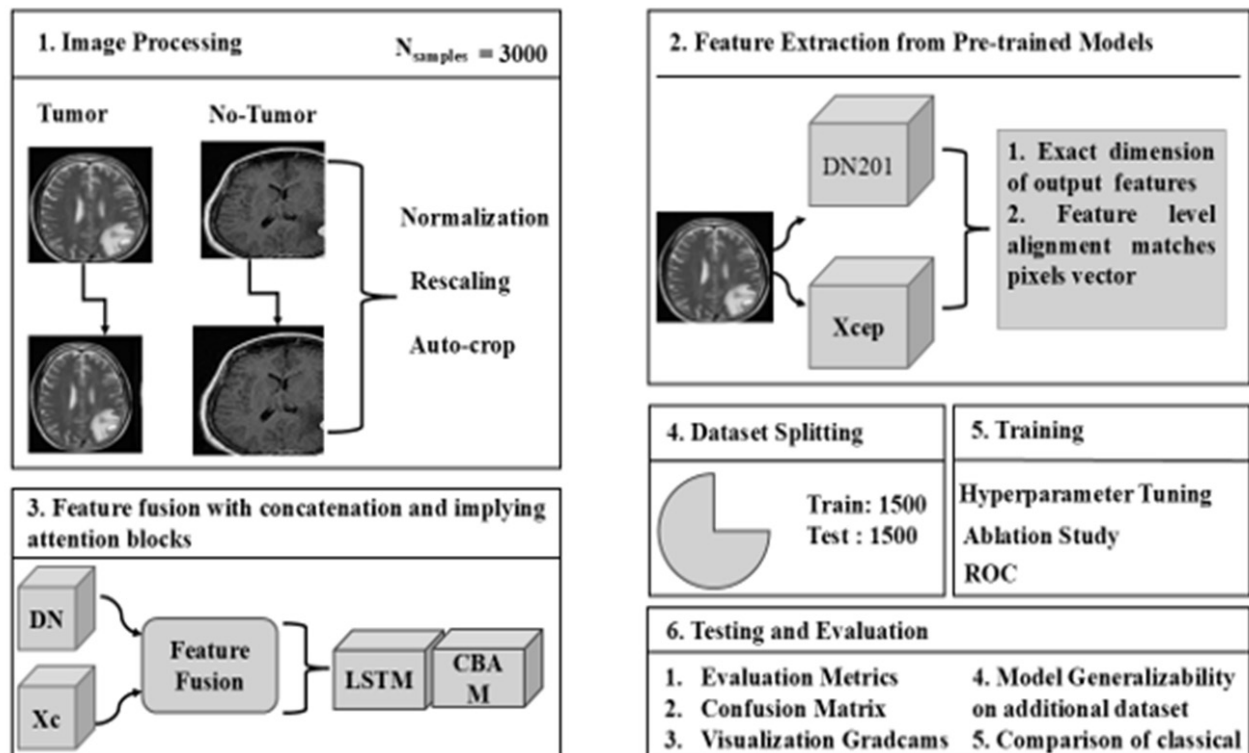


Figure 8: Overview of the proposed pipeline for the brain tumor detection, including image processing, feature extraction (DenseNet201, Xception), feature fusion, dataset splitting, training, and evaluation with metrics and Grad-CAM visualization.

Algorithm 1 presents an expanded view of the proposed methodology

Algorithm 1: Proposed Architecture for Brain Tumor Disease Detection

- Load pre-processed MRI images of brain tumors with enhanced quality.
- Import the two models, DenseNet 201 and Xception.
- Upload the weight values of the train model
- Features are combined using a concatenation layer in feature fusion
- Integrating LSTM and CBAM block
- Load the test dataset;

Output: entire dataset predictions; Classification Report, Assess Score of the Feature Fusion Architecture, Ablation Study, Confusion Matrix,

Result and Discussion

The following section describes the experimental setup utilized in current research. Subsequently, it examines the specifics of hyper-parameters and performance measures employed for model assessment and an examination of feature maps. It then presents the results of the employed technique and compares them with existing methods to emphasize its far superior outcomes.

Experimental setup

The study utilized the Keras framework and the Python programming language to formulate the feature fusion model. The studies were conducted in a popular Python environment, making use of the GPU runtime to leverage its superior processing capabilities. The model performed testing and training on a system with 16 GB of RAM and an NVIDIA GeForce MX350 GPU.

Hyper parameters of the model

To confirm the consistency of the proposed model, we trained it using the optimal hyper-parameters. The approach we employ for training involves using batch sizes of 64 for 20 epochs and 0.001 as the learning rate. We also employed top layers with global average pooling and dropout of 0.4, followed by a dense layer with 1024 units, and finally, a dense layer with neurons regarding the number of classes. The Adam optimizer achieved higher accuracy and lower memory utilization compared to other optimizers. The effectiveness of our training procedure has to do with the use of cross-entropy loss function for categorical, which highlights plays an essential role in evaluating and improving overall performance.

Performance metrics

Closely align our findings with the results of prior research and evaluate the effectiveness of our proposed model by using the most common evaluation criteria used in DL challenges. We classified the diagnosis of each diagnosed sample as True Positive (TP), True

Negative (TN), False Positive (FP), or False Negative (FN). TP and TN signify the precise classifications, although FP and FN specify the inaccurate evaluations. The formulas for this performance are given in the mathematical expressions below:

$$Accuracy(AC) = \frac{TP + FP}{TP + FP + TN + FN} \quad (1)$$

$$Precision(PR) = \frac{TP}{TP + FP} \quad (2)$$

$$Recall(RC) = \frac{TP}{TP + FN} \quad (3)$$

$$F1-Score(FS) = \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

Classification results

Table 3 provides a performance evaluation of different models for classifying tumor and no-tumor classes based on Precision, Recall, F1 Score, and Accuracy. The Proposed Model exhibits the peak performance in all metrics, achieving a Precision, Recall, and F1 Score of 0.99 for both tumor and no-tumor classes, along with an overall Accuracy of 0.9883. This performance indicates the model's superior capability in accurately identifying both classes with minimal errors. In contrast, the Xception model shows the lowest performance among the evaluated models. The tumor class achieved a recall of 0.92 and an overall accuracy of 0.9483. These results suggest that while Xception can identify tumors with reasonable accuracy, it is less effective than the Proposed Model in reducing misclassifications. This highlights the effectiveness of the Proposed Model, emphasizing its potential for high-precision applications in medical diagnostics. (Figure 9) presents the confusion matrix of the proposed and base model. DenseNet201 (base model 1) misclassified 26 samples when tested on a dataset of 1,500 images. Xception (base model 2)

was evaluated on the same test set and misclassified 31 samples, which is higher than DenseNet201. In alteration, the same dataset was employed to test our proposed model and resulted in only 7 misclassifications. These results demonstrate the robustness and effectiveness of the proposed method. This approach significantly reduces misclassification errors compared to the predictions of the individual baseline models. The Receiver Operating Characteristic (ROC) curve [36] is a critical tool for evaluating the performance of classification models by plotting the True Positive Rate (TPR) against the False Positive Rate (FPR) across various thresholds. It visualizes the model's effectiveness in distinguishing between positive and negative classes. The Area Under the Curve (AUC) serves as a comprehensive metric, with values approaching 1 signifying exceptional classification performance. (Figure 10) displayed ROC curve, the DenseNet201 (Base-1) model achieves an AUC of 0.9906, reflecting robust classification with high sensitivity and low rates of false positives. The Xception (Base-2) model yields an AUC of 0.9776, indicating strong yet slightly lower performance compared to DenseNet201. The proposed model excels with an AUC of 0.9989, demonstrating near-perfect class separation and minimal misclassification.

Table 3: Performance evaluation of the proposed model by class.

Models	Class	Precision	Recall	F1 Score	Accuracy
Dense Net201	tumor	0.97	0.94	0.96	0.9566
	no-tumor	0.94	0.97	0.96	
Xception	tumor	0.97	0.92	0.95	0.9483
	no-tumor	0.93	0.97	0.95	
Feature Fusion	tumor	0.98	0.97	0.98	0.9812
	no-tumor	0.97	0.98	0.98	
Proposed Model	tumor	0.99	0.98	0.99	0.9883
	No-tumor	0.99	0.98	0.99	

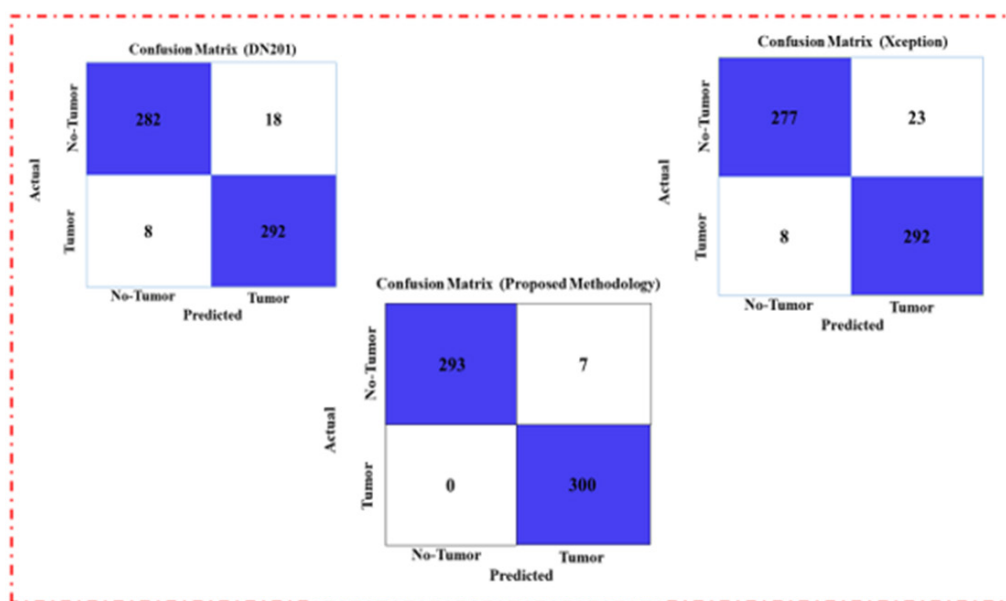


Figure 9: Comparison of the confusion matrix between the base models and the proposed one.

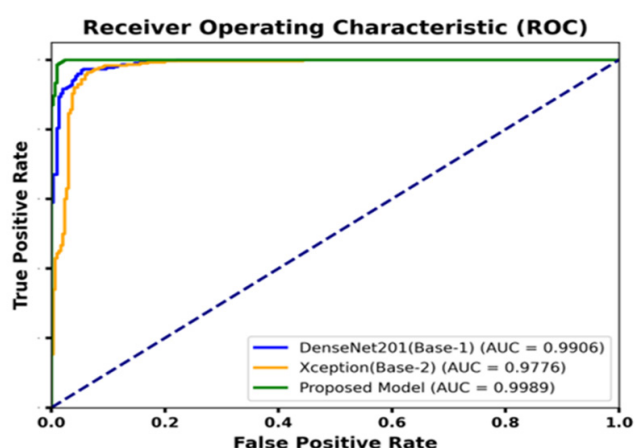


Figure 10: ROC curves for base models (1,2) and the proposed fusion model.

This superior performance is largely attributed to the model's integration of feature fusion techniques enhanced by the CBAM and ConvLSTM. Feature fusion aggregates diverse features from the base models, allowing the proposed approach to harness complementary information for improved feature representation. CBAM enhances this process by applying spatial and channel-wise attention, guiding the model to prioritize critical features while minimizing irrelevant data, thereby boosting its discriminative capabilities. ConvLSTM further enhances the architecture by capturing both spatial and sequential relationships, making the model highly adept at tasks involving temporal patterns. These architectural advancements collectively elevate the proposed model's AUC, underscoring its superior performance in complex classification

challenges compared to conventional approaches. Cohen's Kappa is a statistical measure that evaluates the agreement between models while accounting for chance, making it essential for reliable classification in medical diagnostics. Figure 11 presents an analysis of the Kappa coefficient and error values for many models on the test set. The DenseNet201 model achieved a Kappa value of 91.33, indicating a high level of reliability. In comparison, the Xception model achieved a slightly lower Kappa value of 89.66, reflecting substantial agreement. The proposed model achieved the highest Kappa value of 97.66, demonstrating superior performance. This high Kappa value underscores the proposed model's effectiveness in accurately classifying brain MRI images, highlighting its potential for practical application in medical diagnostics. The proposed model is robust and capable of measuring errors during the training of each batch, making it both easy to implement and understand. Figure 12 compares the Mean Squared Error (MSE) and Mean Squared Logarithmic Error (MSLE) of the proposed framework with the base model. The base model evaluates probabilistic predictions, providing a detailed performance metric, especially for values close to zero. This capability highlights the proposed model's effectiveness in improving error measurement and performance assessment compared to the base model. The graph illustrates that the proposed model significantly outperforms DenseNet201 and Xception in terms of Mean Squared Error (MSE) and Mean Squared Logarithmic Error (MSLE). The proposed model achieves an MSE of approximately 0.01 and an MSLE of around 0.005. In contrast, DenseNet201 and Xception have MSE values slightly above 0.05 and MSLE values near 0.03. These results highlight the proposed model's enhanced predictive accuracy and its ability to minimize errors more effectively than the baseline models.

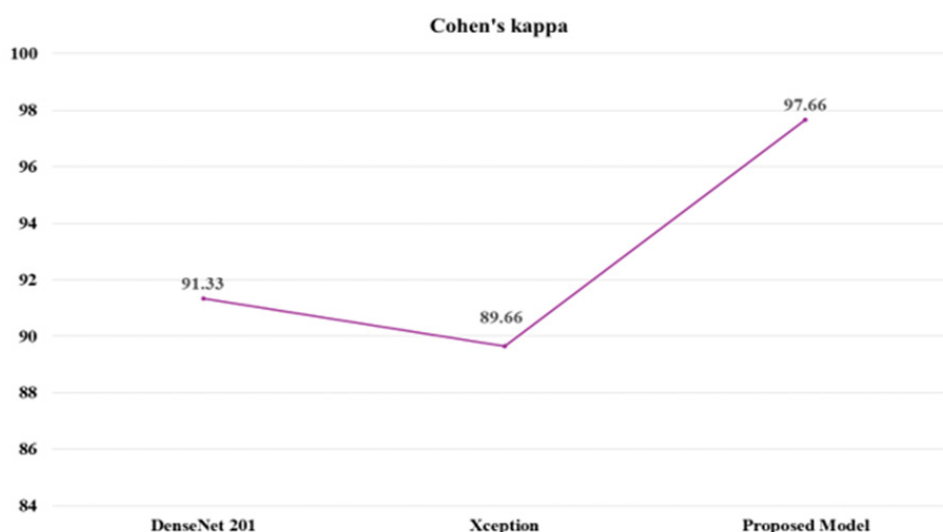


Figure 11: Cohen's Kappa Coefficient assessing model reliability alongside baseline models.

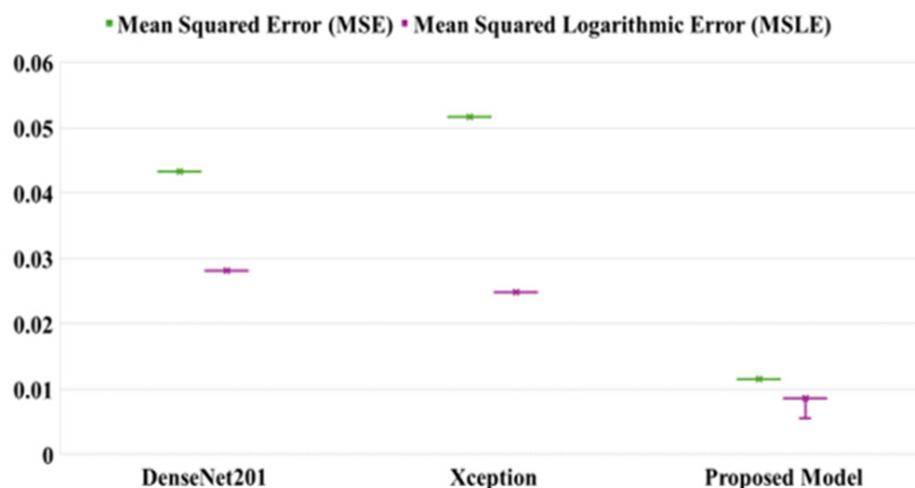


Figure 12: MSE and MSLE of the proposed framework in comparison of base models.

Visualization grad-CAMs

Gradient-weighted Class Activation Mapping (Grad-CAM) is a DL technique that visualizes the specific regions in an image that are essential for a model's prediction. It calculates the gradients of the predicted class concerning the last convolutional layer of a neural network. Grad-CAM creates a heat map by assigning weights to the activations of this layer based on the gradients, highlighting important areas in the input image. This capability aids in understanding which features the model focuses on and the reasoning behind its predictions. Consequently, Grad-CAM is crucial for interpreting and explaining the decisions of Convolutional Neural Networks (CNNs) in tasks like image classification and object detection [37]. Figure 13 presents the results of applying

Grad-CAM to brain MRI images from the test dataset, highlighting the detection of tumors using various deep-learning models. The first row shows the original MRI images. In the second row, the DenseNet201 models generate heat maps indicating regions of focus, but these maps may also highlight irrelevant areas, potentially leading to false positive results. Similarly, the third row shows the Xception model's activation regions, which may not accurately locate the tumor. The final row uses the FF-CBAM (Feature Fusion with Convolutional Block Attention Module) technique, which significantly enhances the model's focus on the specific tumor area, resulting in more accurate and reliable visualizations. This approach leverages feature fusion and attention mechanisms to improve the precision of tumor detection by concentrating on the most relevant features in the images.

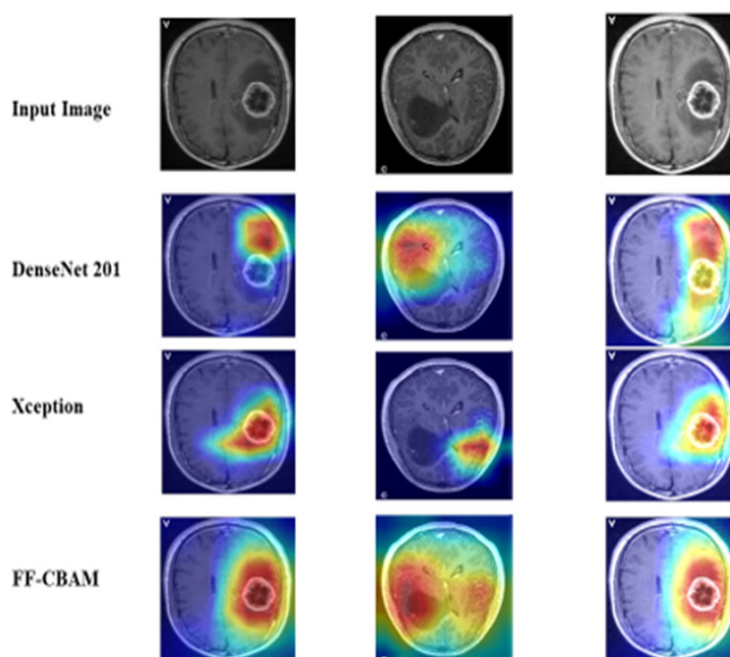


Figure 13: Grad-CAMs result.

Analysis of feature map

Visual analysis of DL models is crucial for enhancing model interpretability, especially in practical applications. In this study, we investigate how our proposed core model leverages insights from the base models DenseNet201 and Xception. These insights enable the identification of regions with distinctive features. To achieve this, we perform a feature map analysis on brain tumor

MRI images, which helps us understand the functionality of the base models. Figure 14 illustrates that our approach improves the clarity of feature representation and accurately identifies critical regions necessary for classification. The figure also displays in-between feature maps from several convolutional blocks within an underlying model. The feature maps reveal just how the model processes its input through multiple filters, demonstrating its capability to extract features extracted from the same input image.

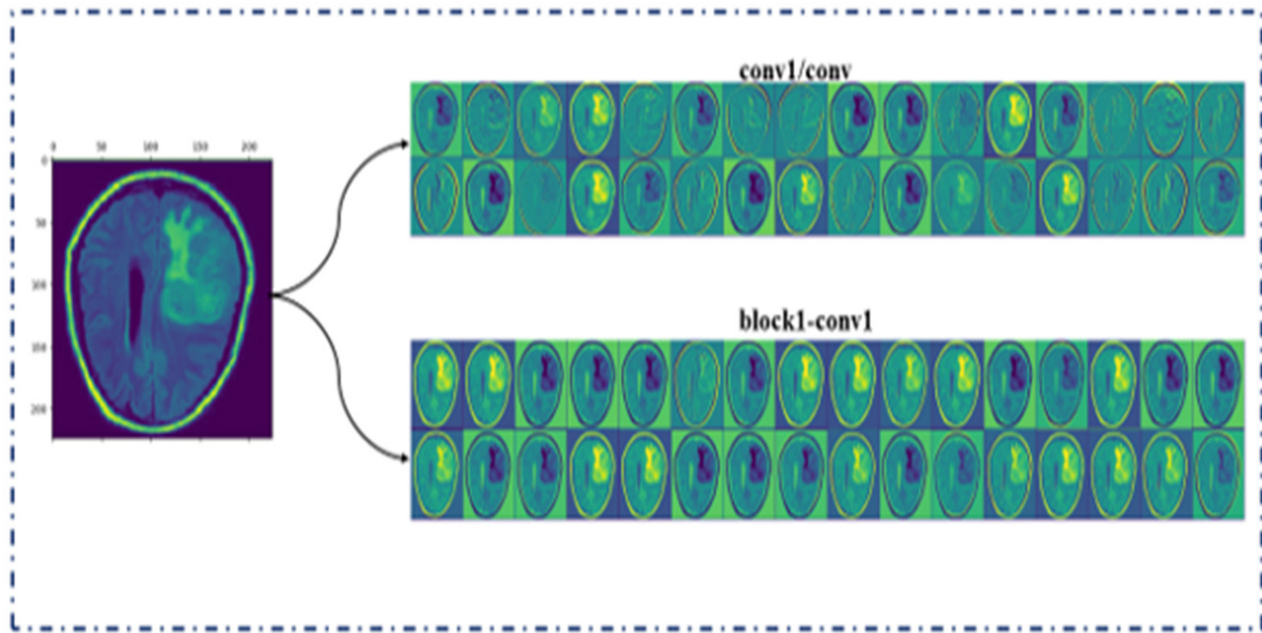


Figure 14: Feature map analysis of base models.

Model generalizability on additional dataset

In this study, the model's generalizability is assessed by its performance on an additional dataset. This approach highlights how well our proposed model can handle new data that was not part of the original training set. To ensure robust performance, we evaluated our model on an additional publicly available dataset from Kaggle [38], which contains four classes: Glioma, Meningioma, No Tumor, and Pituitary. The training and testing dataset details for each class are provided in (Table 4). The data set consists of 7,023 total samples, with 5,712 samples used for training and 1,311 samples used for testing. Each class is well represented, allowing for comprehensive testing of the model's classification capabilities across different tumor types. (Figure 15-17) presents the confusion matrix of the proposed and base model. The DN201 (base model 1) model shows strong performance in classifying the No Tumor and Pituitary classes, with minimal misclassifications, but faces challenges in distinguishing between Glioma and Meningioma, where 47 Glioma cases were misclassified as Meningioma and 29 Meningioma cases as No Tumor. The Xception (base model 2)

model offers improved generalization, especially for No Tumor and Pituitary, yet it also struggles with Glioma and Meningioma, misclassifying 56 Glioma cases as Meningioma and 14 Meningioma cases as No Tumor. The Proposed Model demonstrates the best generalization, with fewer misclassifications across all classes. It almost perfectly classifies No Tumor and Pituitary, with only one Meningioma case misclassified as Pituitary, and shows better distinction between Glioma and Meningioma, misclassifying 18 Glioma and 6 Meningioma cases. The findings demonstrate that the Proposed Model is highly robust and generalizes well to unseen data, making it a reliable choice for brain tumor classification tasks.

Table 4: Training and testing dataset details for each class.

Classes	Training Set	Test Set	Total Samples
Glioma	1321	300	1621
Meningioma	1339	309	1645
No Tumor	1595	405	2000
Pituitary	1457	300	1757
Total Samples	5712	1311	7023

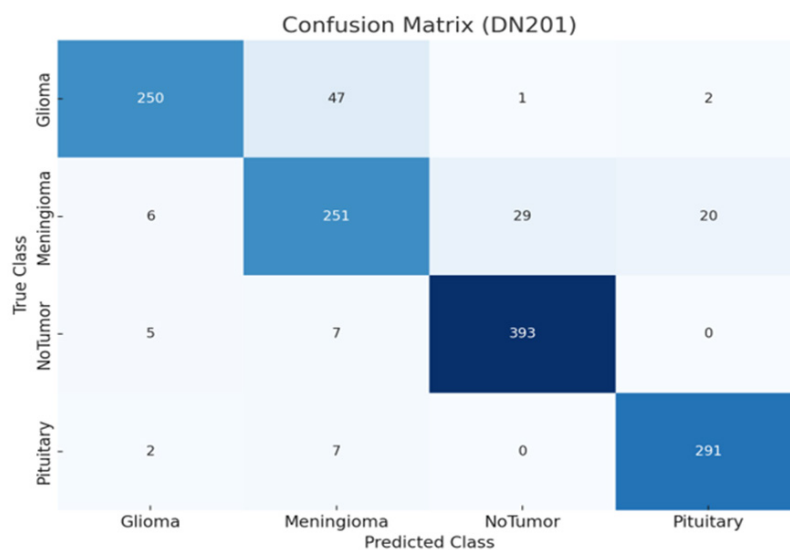


Figure 15: Confusion matrixes illustrate DN201 (base model 1).

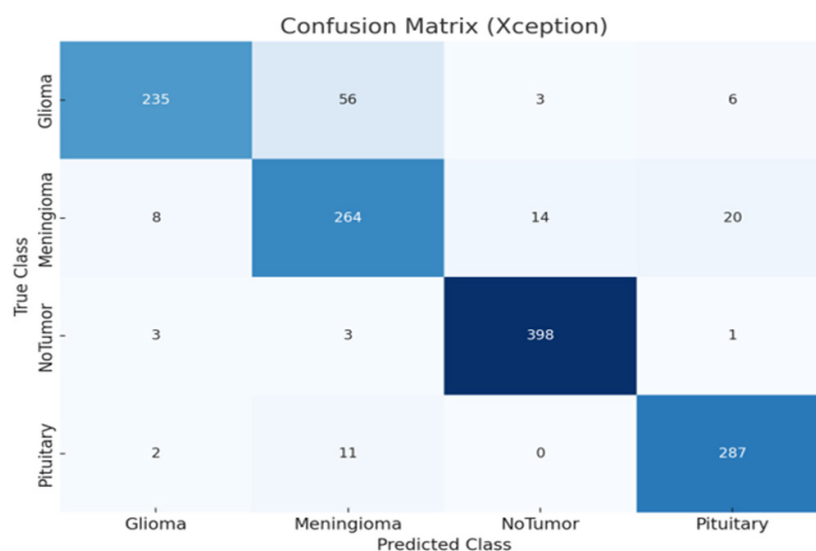


Figure 16: Confusion matrix illustrates Xception (base model 2).

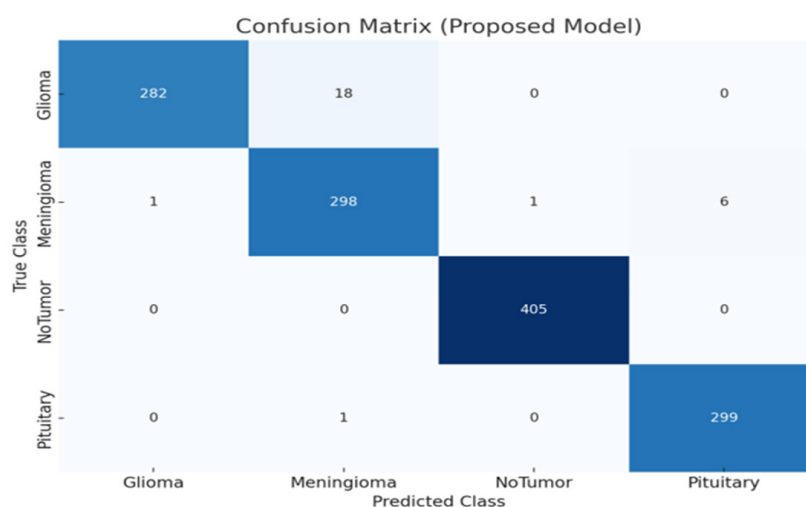


Figure 17: Confusion matrixes illustrate proposed model generalizability on the additional dataset.

Further model's generalizability performance on additional dataset was compared with two baseline models: DenseNet201 (base model 1) and Xception (base model 2). The results, depicted in Figure 18, indicate that DenseNet201 achieved an AUC value of 0.9874, while Xception reached an AUC value of 0.9864. In contrast, our proposed model demonstrated a significantly higher AUC of 0.9993. The near-perfect AUC value indicates that the model maintains both high accuracy and a low false-positive rate when exposed to previously unseen data. The results provide strong evidence of the robustness and generalizability of the proposed model. The evaluation of model generalizability on an additional dataset was conducted using various pretrained models as base architectures. Figure 19 illustrates the accuracy performance of

each model. The proposed model achieved the highest accuracy of 97.94%, demonstrating its superior ability to generalize unseen data. MobileNetV1 (MNV1) reached an accuracy of 94.66%, while MobileNetV2 (MNV2) achieved 91.23%. ResNet50V2 (RN50V2) performed slightly better, with an accuracy of 91.38%. Among the DenseNet architectures, DenseNet121 obtained an accuracy of 88.41%, DenseNet169 reached 90.54%, and DenseNet201 attained 90.39%. The proposed model significantly outperformed the other architectures, indicating its strong capacity for generalization across datasets. This highlights the effectiveness of using advanced feature fusion techniques and attention mechanisms in enhancing the model's robustness and performance in challenging scenarios.

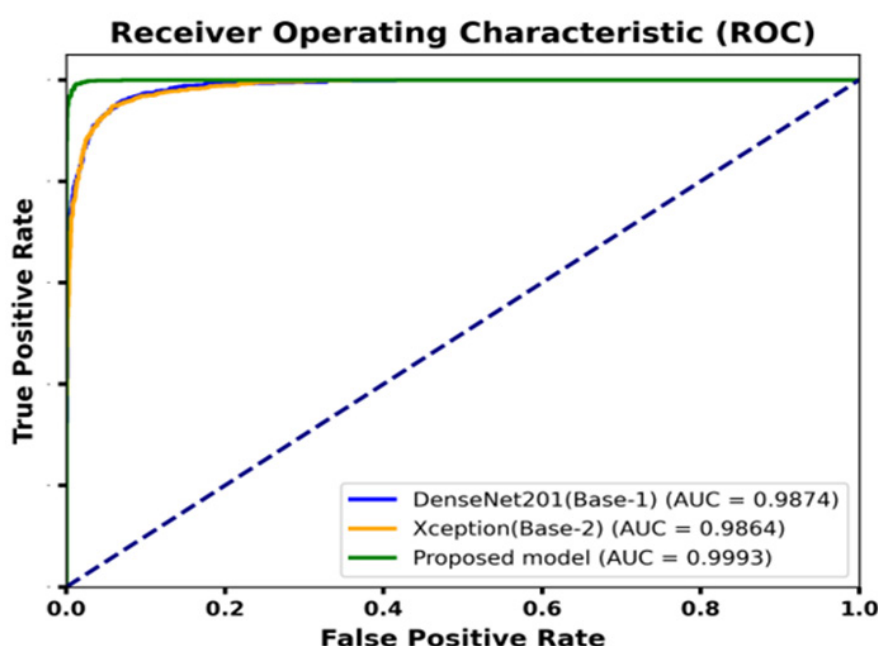


Figure 18: The ROC curve illustrates model generalizability on the additional dataset.

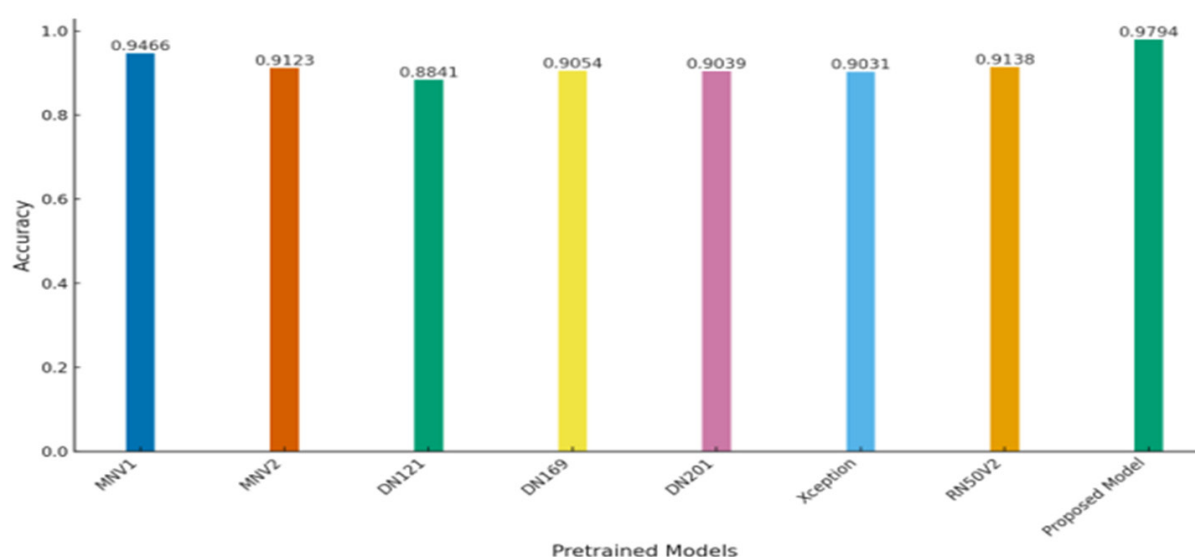


Figure 19: Accuracy comparison of pretrained models to assess model generalizability on the additional dataset.

Comparison of the proposed method with existing studies

Table 5: Performance of current approaches with the proposed model

Method	Accuracy %	Reference
Densenet201 with MGA-feature	95	[39]
Multiscale DCNN Tumor Classification	97.3	[40]
CNN-based VGG 16 Architecture	95.17	[41]
Feature Fusion Attention Method	98.83	Proposed Model

Table 5 presents a comparison of the accuracy results obtained from four distinct models. Densenet201 with MGA-feature [37] achieved an accuracy of 95%, while the Multiscale DCNN [38] for the tumor classification approach attained an accuracy of 97.3%. The VGG-16 architecture, based on CNN achieved an accuracy of 95.17%. In contrast, the Feature Fusion Attention Method, which is the proposed model, obtained a significantly higher accuracy of 98.83%. This performance indicates that the suggested model outperforms existing methods. The suggested model's attention mechanism and innovative feature integration drives its high accuracy. This method integrates and improves brain MRI

characteristics, reducing misdiagnosis and enhancing brain tumor categorization. The high precision provides valuable insights and supports the methods and reasons used to achieve these findings. Emphasizing the model's ability to explain outcomes clarifies its predictions. In fields where model interpretability is important for making decisions and building trust, this is very important [39,40].

Comparison with pre-trained models

Figure 20 compares performance measures for various deep learning models that consider a certain task as a classification task. The models evaluated in this study include MobileNet, MobileNetV2, DenseNet121, DenseNet169, Inception V3, and the Proposed Model. The evaluation process utilizes performance criteria such as Precision, Recall, F1-Score, and Accuracy. The Proposed Model achieves exceptional performance in all aspects, with a Precision of approximately 97.95%, Recall of 99.01%, F1-Score of 98.84%, and Accuracy rate of 98.83%. These results show the efficacy of the Proposed Model's high accuracy and reliability in data classification tasks. Although MobileNet demonstrates strong performance with high metric scores, MobileNetV2 and DenseNet121 exhibit lower values in Precision, Recall, F1-Score, and Accuracy. While DenseNet169 and InceptionV3 are effective, their performance does not match that of the Proposed Model. The comparative analysis clearly indicates that the Proposed Model is consistently better than the rest of the models in overall performance.

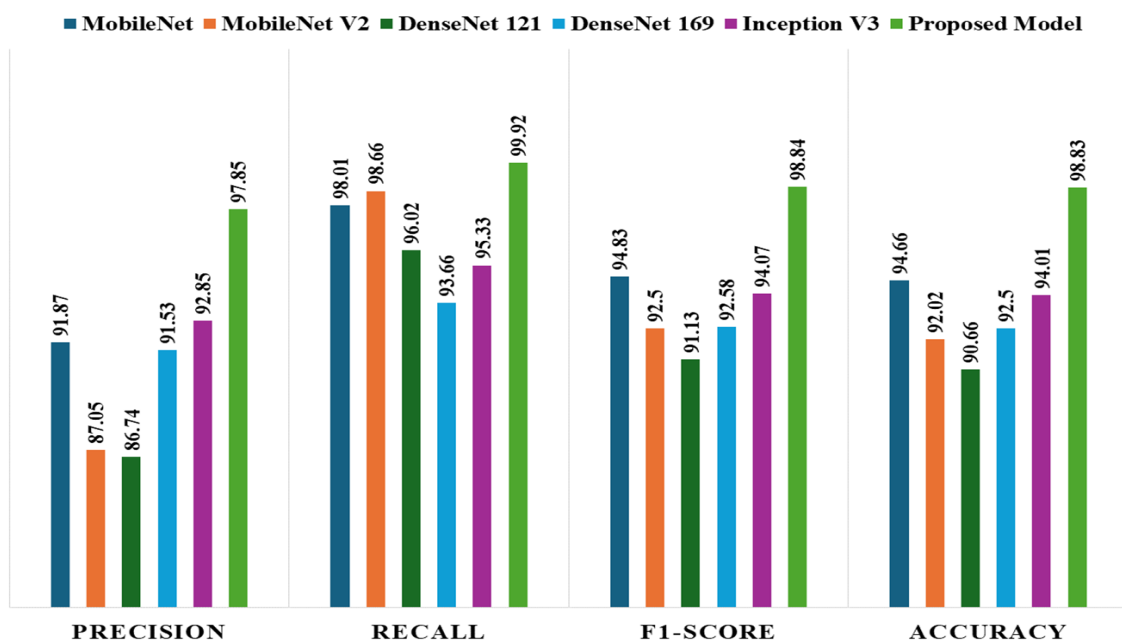


Figure 20: Comparative analysis of the proposed model with pre-trained models.

Ablation study

In this section, we will discuss the impact of each component within our proposed feature fusion framework. (Table 6) compares the initial models using various performance indicators for an exact task within the domain of deep learning or data analysis. Five models are evaluated in total and DenseNet201 and Xception are

selected as baseline models due to their effective performance. This analysis highlights the strengths and weaknesses of each model in relation to the proposed framework. In our ablation study, we assessed the impact of various components on model performance. We measured each model's performance based on precision, recall, F1-score, and accuracy. The proposed model consistently delivered the maximum scores through all metrics, achieving an F1-score

of 98.84%, an accuracy of 98.83%, a precision of 97.95%, and a memory of 99.01%. This indicates that incorporating the proposed model's components significantly enhances overall effectiveness, surpassing the performance of the other models in the study. (Figure 21) provides a comprehensive comparative analysis of the real and predicted values for our proposed method across various scenarios. We specifically evaluate the performance comparison of the model with and without feature fusion and the inclusion of an attention module. This analysis highlights the effectiveness of our

method, which combines ConvLSTM layers for capturing spatio-temporal dependencies with feature fusion and the CBAM attention module to enhance feature refinement. It is crucial to acknowledge the model's performance significantly decreased when feature fusion with the attention module was excluded. The feature fusion and attention module play a critical role in improving the predictive capabilities of the model. Their absence led to decreased accuracy as a whole and weakened the model's capability to forecast across various image categories.

Table 6: Impact of LSTM and CBAM attention module in proposed methodology.

Model	Precision %	Recall %	F1-Score %	Accuracy %
DenseNet201	94.19	97.33	95.73	95.66
Xception	92.69	97.33	94.95	94.83
Feature Fusion	96.42	99.01	97.69	97.66
FF+LSTM	97.06	99.03	98.18	98.16
Proposed Model	97.95	99.01	98.84	98.83

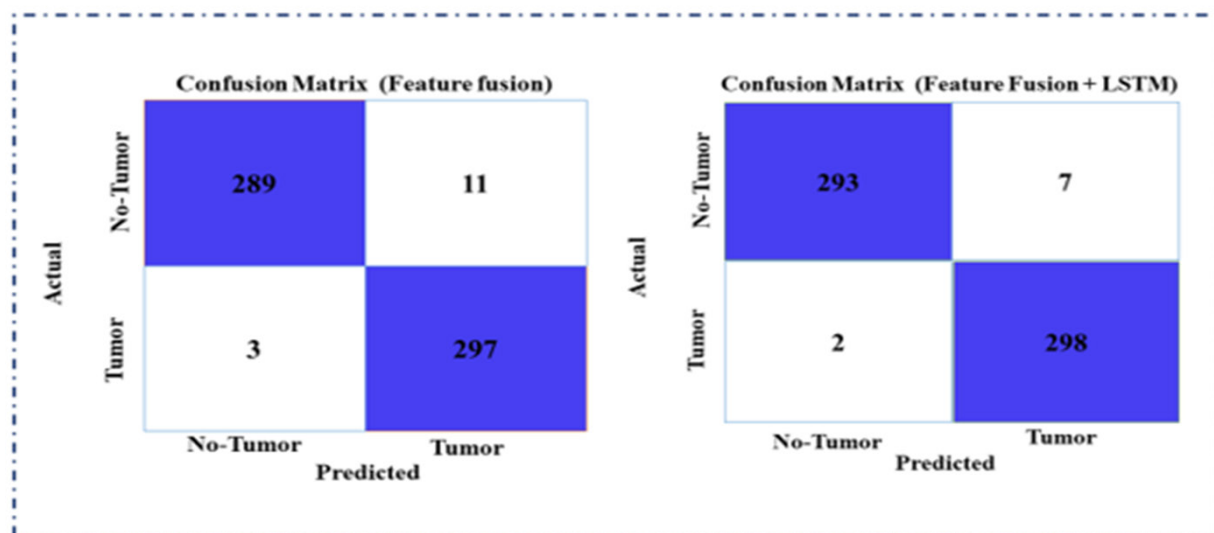


Figure 21: Confusion comparisons incorporating single feature fusion and with LSTM block.

We conducted a performance comparison between our proposed model and baseline models that were developed independently. First, we incorporated ConvLSTM layers into the base models to extract robust features while maintaining spatio-temporal dependencies. Then, we performed feature fusion by combining the features from these two base models. Afterward, we integrated the CBAM attention module in the feature fusion to refine the features further and enhance focus on channel and spatial information. The baseline models demonstrated strong predictive capabilities and outperformed other pre-trained models. Our architecture surpasses the performance of existing models, showing a steady improvement in accuracy and reliability. These outcomes highlight the efficiency of our methods and then suggest the possibility for further advances to be achieved in the classification of image tasks. Using a test set of 1,500 images, our model consistently performed well and correctly classified 1,493

images. This outcome demonstrates the strength and reliability of our proposed framework.

Limitations and Future Work

During our research on brain tumor classification, we identified several margins that exceed the scope of our current research. We recognize these limitations and propose areas for future improvement.

Limitations

- The brain tumor dataset collection is one of the few publicly available resources for MRI images, which limits our ability to evaluate model efficacy thoroughly. Comprehensive datasets are essential for improving clinical testing accuracy.
- Incorporating datasets that include patient symptoms, medical histories, and age profiles can significantly enhance model

accuracy and precision. However, collecting comprehensive patient data poses a challenge. Collaborating with hospitals to facilitate data accessibility for researchers is crucial. This partnership will improve medical insights and model performance over time.

Future work

- a. Future research will focus on leveraging larger and more diverse datasets as they become available. This will enable a more comprehensive assessment of our models' performance, overcoming current limitations related to data scarcity.
- b. Developing a smartphone application for real-time diagnostics in resource-constrained areas could provide immediate insights and potentially save lives. Future work will include optimizing the model for multi-classification to enhance its utility and ensure accessibility in diverse healthcare settings.

Conclusion

Brain tumors pose a severe challenge to healthcare due to their complexity and the critical need for initial detection. A precise diagnosis is necessary to improve outcomes for patients and enable timely treatment. Traditional diagnostic methods often prove inadequate in providing the precision required for effective intervention. Therefore, there is a critical requirement for advanced computational techniques that can enhance the accuracy of brain tumor detection. This study presents a new approach for predicting brain tumor growth through feature-level fusion. We selected DenseNet201 and Xception as our base models because of their superior ability to effectively extract features from medical images. These features are further enhanced using the CBAM module, which refines spatial and channel-wise attention to focus on the most critical areas of input images. To capture sequential dependencies and improve prediction accuracy, we integrate ConvLSTM layers at the feature-fusion level. This combination allows the model to process temporal information effectively, leading to more precise outputs. Our comprehensive evaluation of a publicly available brain tumor dataset with MRI scans demonstrates that our model outperforms existing deep learning methods, achieving an impressive accuracy of 98.83%. We use visualization techniques such as Grad-CAM, Receiver Operating Characteristics (ROC) curve feature maps, and confusion matrix analysis to interpret the model's decisions. These tools validate the model's reliability and effectiveness in highlighting important regions that influence its predictions. This approach represents a significant advancement in diagnostic accuracy, offering a powerful tool for early brain tumor detection. Our method is likely to significantly expand clinical effects by providing a more precise and reliable diagnostic process. In future work, we aim to expand this framework to include larger datasets and explore real-time applications in clinical settings.

Authorship Contributions

Ifza Shad: Conceptualization, Methodology, Data Curation, Writing-Original draft, Validation; Omair Bilal: Methodology, Investigation, Writing-review & editing; Arash Hekmat: Data Curation, Visualization.

Data Availability

Data will be made available on request.

References

1. Jadhav RN, Sudhaghar G (2024) Deep learning approaches for brain tumor detection in MRI images: A comprehensive survey. *International Journal of Intelligent Systems and Applications in Engineering* 12(13s): 586-602.
2. Islam MN, Azam SN, Islam SM, Kanchan HM, Parvez SAHM, et al. (2024) An improved deep learning-based hybrid model with ensemble techniques for brain tumor detection from MRI image. *Informatics in Medicine Unlocked* 47: 101483.
3. Zhou L, Wang M, Zhou N (2024) Distributed federated learning-based deep learning model for privacy MRI brain tumor detection. *Electrical Engineering and Systems Science* p. 10026.
4. Bhuyan R, Nandi G (2023) Systematic study and design of multimodal MRI image augmentation for brain tumor detection with loss aware exchange and residual networks. *International Journal of Imaging Systems and Technology* 34(2): e22989.
5. Anantharajan S, Gunasekaran S, Subramanian T, Venkatesh R (2024) MRI brain tumor detection using deep learning and machine learning approaches. *Measurement Sensors* 31: 101026.
6. Khan MA, Park H (2024) A convolutional block base architecture for multiclass brain tumor detection using magnetic resonance imaging. *Electronics* 13(2): 364.
7. Seere SKH, Karibasappa K (2020) Threshold segmentation and watershed segmentation algorithm for brain tumor detection using support vector machine. *European Journal of Engineering and Technology Research* 5(4): 516-519.
8. Khaliki MZ, Başarslan MS (2024) Brain tumor detection from images and comparison with transfer learning methods and 3-layer CNN. *Scientific Reports* 14(1): 2664.
9. Raghavendra U, Gudigar A, Paul A, Goutham TS, Inamdar MA, et al. (2023) Brain tumor detection and screening using artificial intelligence techniques: Current trends and future perspectives. *Comput Biol Med* 163: 107063.
10. Solanki S, Singh PU, Chouhan SS, Jain S (2023) Brain tumor detection and classification using intelligence techniques: An overview. *IEEE Access* 11: 12870-12886.
11. Abdusalomov AB, Mukhiddinov M, Whangbo TK (2023) Brain tumor detection based on deep learning approaches and magnetic resonance imaging. *Cancers (Basel)* 15(16): 4172.
12. Abd EMK, Awad AI, Khalaf AAM, Hamed HFA (2019) A review on brain tumor diagnosis from MRI images: Practical implications, key achievements, and lessons learned. *Magnetic Resonance Imaging* 61: 300-318.
13. Yushkevich PA, Gao Y, Gerig G (2016) ITK-SNAP: An interactive tool for semi-automatic segmentation of multi-modality biomedical images. 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Orlando, USA.
14. Cinar A, Yildirim M (2020) Detection of tumors on brain MRI images using the hybrid convolutional neural network architecture. *Medical Hypotheses* 139: 109684.
15. Latif G, Brahim BG, Iskandar DNFA, Bashar A, Alghazo J (2022) Glioma tumors' classification using deep-neural-network-based features with SVM classifier. *Diagnostics (Basel)* 12(4): 1018.
16. Kumar G, Kumar P, Kumar D (2021) Brain tumor detection using convolutional neural network. *IEEE International Conference on Mobile Networks and Wireless Communications (ICMNC)*, Karnataka, India.
17. Yahyaoui H, Ghazouani F, Farah IR (2021) Deep learning guided by an ontology for medical images classification using a multimodal fusion.

- International Congress of Advanced Technology and Engineering (ICOTEN), Taiz, Yemen.
18. Mahmud MI, Mamun M, Abdelgawad A (2023) A deep analysis of brain tumor detection from MR images using deep learning networks. *Algorithms* 16(4): 176.
 19. Wozniak M, Silka J, Wieczorek M (2023) Deep neural network correlation learning mechanism for CT brain tumour detection. *Neural Computing and Applications* 35: 14611-14626.
 20. Mandloi S, Zuber M, Gupta RK (2024) An explainable brain tumor detection and classification model using deep learning and layer-wise relevance propagation. *Multimedia Tools and Applications* 83: 33753-33783.
 21. Rahman T, Islam MS (2023) MRI brain tumor detection and classification using parallel deep convolutional neural networks. *Measurement: Sensors* 26: 100694.
 22. Gayathri P, Dhavileswarapu A, Ibrahim S, Paul R, Gupta R (2023) Exploring the potential of VGG-16 architecture for accurate brain tumor detection using deep learning. *Journal of Computers Mechanical and Management* 2(2): 13-22.
 23. Naseer A, Yasir T, Azhar A, Shakeel T, Zafar K (2021) Computer-aided brain tumor diagnosis: Performance evaluation of deep learner CNN using augmented brain MRI. *International Journal of Biomedical Imaging* 2021(1): 1-11.
 24. Huang G, Liu Z, Der MLV, Weinberger QK (2017) Densely connected convolutional networks. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, USA.
 25. Salim F, Saeed F, Saeed S, Qasem NS, Al HT (2023) DenseNet-201 and Xception pre-trained deep learning models for fruit recognition. *Electronics* 12(14): 3132.
 26. Jaiswal A, Gianchandani N, Singh D, Kumar V, Kaur M (2021) Classification of the covid-19 infected patients using DenseNet201 based deep transfer learning. *Journal of Biomolecular Structure and Dynamics* 39(15): 5682-5689.
 27. Rismiyati, Nur ES, Khadijah, Shiddiq NI (2020) Xception architecture transfer learning for garbage classification. In 4th International Conference on Informatics and Computational Sciences (ICoICS), Indonesia.
 28. Chollet F (2017) Deep learning with depthwise separable convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, USA.
 29. Hochreiter S, Schmidhuber J (1997) Long short-term memory. *Neural Comput* 9(8):1735-1780.
 30. Amin J, Ali Shad S, Sial R, Saba T, Raza M, et al., (2020) Brain tumor detection: A long short-term memory (LSTM)-based learning model. *Neural Computing & Applications* 32: 15965-15973.
 31. Woo S, Park J, Young LJ, Kweon SI (2018) CBAM: Convolutional block attention module. *Computer Vision-ECCV* pp. 3-19.
 32. Wang W, Tan X, Zhang P, Wang X (2022) A CBAM based multiscale transformer fusion approach for remote sensing image change detection. *Journals and Magazines* 15: 6817-6825.
 33. Zhang M, Xu S, Song W, He Q, Wei Q (2021) Lightweight underwater object detection based on yolo v4 and multi-scale attentional feature fusion. *Remote Sensing* 13(22): 4706.
 34. Song W, Li S, Fang L, Lu T (2018) Hyperspectral image classification with deep feature fusion network. *IEEE Transactions on Geoscience and Remote Sensing* 56(6): 3173-3184.
 35. Hoo ZH, Candlish J, Teare D (2017) What is an ROC curve? *Emergency Medicine Journal* 34(6): 357-359.
 36. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, et al. (2020) Grad-CAM: Visual explanations from deep networks via gradient-based localization. *International Journal of Computer Vision* 128: 336-359.
 37. Osman Ö, Okan İB, Hüseyin G, Khan F, Hussain J, et al. (2023) Multiple brain tumor classification with dense CNN architecture using brain MRI images. *Life (Basel)* 13(2): 349.
 38. Sharif MI, Khan AM, Alhussein M, Aurangzeb K, Raza M (2021) A decision support system for multimodal brain tumor classification using deep learning. *Complex & Intelligent Systems* 8: 3007-3020.
 39. Díaz PFJ, Martínez ZM, Antón RM, González OD (2021) A deep learning approach for brain tumor classification and segmentation using a multiscale convolutional neural network. *Healthcare* 9(2): 153.
 40. Waghmare VK, Kolekar HMK (2020) Brain tumor classification using deep learning. *Internet of Things for Healthcare Technologies* 73: 155-175.