Jordanian Journal of Computers and Information Technology (JJCIT), Vol. 12, No. 01, March 2026.

BRAIN-TUMOR CLASSIFICATION USING RESNET50 ENHANCED WITH SE AND CBAM ATTENTION MECHANISMS

Nadia Shamsulddin Abdulsattar[1] and Fatimah S. Abdulsattar[2]

(Received: 5-Dec.-2025, Revised: 15-Feb.-2026, Accepted: 9-Mar.-2026)

ABSTRACT

MRI image classification of brain tumors is critical for accurate and early diagnosis. New developments in deep learning have revealed that inserting attention mechanisms into convolutional neural networks can greatly improve classification performance. The ECA attention mechanism is also introduced in this study. This work assesses the effectiveness of Squeeze-and-Excitation (SE) and Convolutional Block Attention Module (CBAM) sequentially integrated with the ResNet50 model, which increases classification accuracy, precision and recall when compared to the basic model, according to experimental results on two datasets for brain tumors. The suggested model employs attention mechanisms to focus valuable information selectively and suppress irrelevant information. The experiments are conducted on two datasets (Brain Tumor MRI and Brisc). The first dataset displays great improvements over basic CNN models, with precision, recall, accuracy, F1 score and AUC at 0.9914, 0.9903, 0.9945, 0.9908 and 0.9989, respectively. The second dataset gives the results for precision, recall, accuracy, F1 score and AUC at 0.9860, 0.9857, 0.9860, 0.9858 and 0.9985, respectively. From these results, the importance of attention mechanisms in deep-learning models for medical imaging is highlighted, which suggests that SE and CBAM modules can be available as more dependable and effective instruments for brain-tumor classification in clinical settings. Future studies should investigate transformer-based and hybrid attention techniques to enhance automated brain tumor categorization.

KEYWORDS

Brain Tumor, MRI, ResNet50, Squeeze-and-Excitation (SE), CBAM.

1. INTRODUCTION

Many studies have focused on the detection of brain tumors in recent years [1][2][3][4][5][6][7][8][9][10][11]. Researchers have suggested some techniques to reveal brain tumors in MRI scans of the brain [12]. The accuracy of each of these techniques has varied. Convolutional neural networks (CNNs), a common deep-learning technique, have an advantage over classical networks. These can learn from the input directly without needing human feature extraction [13]-[14]. The CNN model is composed of several layers, each with a distinct job. Features are estimated by the convolution layer, the size of the features from the previous layer is reduced by the pooling layer, the features of high-level are elicited and the output of the model is predicted by the fully connected layer, also known as the dense layer. The basic structure of CNN employs activation functions, like Tanh, Sigmoid and ReLU [15]-[16] and [17]. Brain tumors, like meningiomas, gliomas and pituitary adenomas, are among the brain tumors that pose serious health risks and need to be precisely identified to receive focused treatments. The clinical gold standard for non-invasive tumor visualization and evaluation is still magnetic resonance imaging (MRI). Recent advances in deep learning; namely, convolutional neural networks (CNNs), have revolutionized computer-aided diagnosis in neuroimaging by offering automated solutions that can match or surpass human expert performance when given sufficient, well-annotated data [18]. Optimizing feature extraction from intricate MRI data is a great challenge in the automated classification of brain tumors. Although standard CNNs function well, they can be improved by using attention techniques that direct the network to more noticeable aspects of the image [19]. Several studies have explored CNN architectures for brain-tumor classification. ResNet variants are widely employed due to their depth and performance. SE networks improve channel interdependencies by recalibrating feature maps. CBAM combines channel and spatial attention to further improve discriminative feature

1. N. S. Abdulsattar is with Department of Computer, College of Basic Education, Mustansiriyah University, Baghdad, Iraq. Email: nadiashamsaldeen@uomustansiriyah.edu.iq
2. F. S. Abdulsattar is with Department of Computer Engineering, College of Engineering, Mustansiriyah University, Baghdad, Iraq. Email: fsa.abdulsattar@uomustansiriyah.edu.iq

representation. To enhance the representational capacity of the baseline ResNet50 model, channel and spatial-attention mechanisms were incorporated into the residual blocks. Sequential integration of SE and CBAM was used to exploit their complementary channel and spatial-attention mechanisms while avoiding attention redundancy and improving model generalization.

2. RELATED WORK

A method for detecting and categorizing brain tumors using Deep Residual Networks (ResNet) was presented by Sahaai et al. [20]. In this method, a ResNet50 model based on transfer learning and the CNN architecture is employed to achieve multi-class classification of brain tumors. Through several brain-tumor dataset categories, this method obtains 95.3%, 94.6%, 92.2%, 93.7% and 87.8% for accuracy, F1 score, recall, precision and specificity, respectively.

Oladimeji and Ibitoye [21] employed the pretrained model ResNet50 with the CBAM. When comparing this method with other classification methods that use the same dataset (Brain Tumors dataset), the ResNet50-CBAM achieved good results compared with existing deep-learning classification methods, such as CNNs. It performed a recall, accuracy, precision and AUC at 99.01%, 99.43%, 98.7% and 99.25%, respectively. The study's experimental findings demonstrated the convolutional block attention mechanism scheme's better results in brain-tumor categorization.

To categorize brain tumors, Vinston et al. [22] employed the pretrained ResNet50 with the CBAM. This method outperformed previous techniques, like traditional convolutional neural networks. On the same dataset (Brain Tumors dataset), the ResNet50-CBAM model revealed good results, reaching 99.53%, 99.11%, 99.35% and 98.75% for area under the curve (AUC), recall, accuracy and precision, respectively. The convolutional block attention-mechanism architecture works exceptionally well for brain-tumor classification, according to experimental results.

Employing the dataset on Kaggle, which contains 3,096 MRI images. Huang and Prakash [23] trained a ResNet50V2-based model. They evaluated five models: Basic ResNet50V2, Squeeze-and-Excitation, Convolutional Block Attention Module, Self-Attention and Attention Gated Network. To compare the classification accuracy of the models, two proportion Z-tests were employed. The SE model surpassed the basic ResNet50V2 in classification, achieving an accuracy of about 98.4% and an AUC of 1.00, while the ResNet50V2 achieved about 92.6%. In this study, SE achieved the highest results compared to the other attention mechanisms, such as CBAM.

To categorize brain tumors independently, Rakesh et al. [24] employed a complex convolutional neural network. They used a set of methods to process the dataset, like cropping, splitting and uncropping. The ResNet-50 pretrained model is the primary model employed in this method. With 81.67%, 74.3%, 82.55% and 81.67% for accuracy, precision, recall and specificity, the suggested model performs admirably and exceeds early predictions.

To categorize brain tumors, Md and Ankit [25] created and evaluated custom convolutional neural networks. Our findings showed that these tailored models beat popular structures, like ResNet18 and VGG16, in accuracy and efficiency of computation while maintaining a simpler design. The custom CNN fulfilled 98.09% accuracy in multi-class classification, 98.67% on the Br35H dataset and 99.62% on the Brain Tumor MRI dataset in binary classification. The custom CNNs offered an additional computationally efficient option, while ResNet18 and VGG16 kept good performance levels in comparison.

Muhammad et al. [26] suggested CNNs, which are excellent at extracting features at the convolutional layer. The architectures of the outstanding CNNs employed in medical image processing, such as VGG16, ResNet-50, MobileNet, InceptionV3 and EfficientNetB7, are compared with the brain-tumor classification job. With the result, VGG16 gives good findings on other CNN architectures. For test-set data, VGG16 produces 100% accuracy, sensitivity, specificity, precision and F1-score. This study indicates the effectiveness of the strategy suggested by showing outstanding performance in categorizing brain tumors and no tumor from MRI scans.

The objective of the study of Mohammad and Muhammet [27] was to categorize brain cancers, like gliomas, meningiomas and pituitary tumors, using the images of brain MRI. The CNNs and CNN-dependent categorization methods included transfer learning, Inception-V3, EfficientNetB4 and VGG19. F1 score, recall, accuracy and imprinting were employed to assess these models. VGG16

yielded the best results, with an accuracy rate of 98%. The same transfer-learning model has an F1 score of 97%, an area under the curve of 99%, a recall of 98% and a precision of 98%. The CNN structure and CNN-dependent transfer-learning models are essential to the health of humans for the detection and timely medication of diseases early.

In contrast to the previous three suggested models (InceptionResNet, DenseNet121 and NasNet Large), Asif et al. [28] suggested a model of CNN that depends on the architecture of the Xception that utilizes the optimizer (ADAM). Sensitivity, accuracy, precision, specificity and F1-score values for the Xception model were 99.68%, 99.67%, 99.68%, 99.66% and 99.68% on the dataset (BR35H) and 96.55%, 91.94%, 87.50%, 87.88% and 91.80% on the dataset (Brain Tumor). The suggested approach performs well compared to those in the existing literature in detecting brain tumors.

To elicit features from the images of brain MRI, Deepak and Ameer [29] employed a pre-trained GoogLeNet and adopted the idea of deep transfer learning. The obtained features are classified using integrated, validated classifier models. The experiment utilized a five-fold cross-validation approach at the level of patient, using the MRI dataset from figshare. The proposed approach achieved a good result compared with the cutting-edge techniques, achieving an accuracy of 98%. The other performance measures employed in the work stand for recall, specificity, AUC, F-score and precision. The study's findings suggested that transfer learning is beneficial in situations where medical images are rare.

The CNN technique, known as en-CNN, was proposed by Hapsari et al. [30]. This approach depends on VGG-16, which has four ReLU layers, four max pooling layers and seven convolutional layers. The three steps of the novel method are augmentation, preprocessing and the application of an en-CNN. The 4 MRI sequences: FLAIR, T1, T2 and T1CE are also employed in our suggested approach for classification. Employing the ADAM optimizer, the suggested approach achieves an accuracy of the multi-sequence MRI dataset (BraTS 2018) with mini-batch size 128 and epoch 200 of 95.5% for T1, 94% for T2, 95.5% for T1CE and 97% for FLAIR. Compared to earlier studies using the same dataset, the accuracy was 4% higher.

Lin et al. [31] enhanced the ResNet50 model by preprocessing the image and employing fractional calculus; then, transfer learning and the attention mechanism (ECA) are applied. After that, the enhanced ResNet50 is collected with EfficientNetB0 to increase the accuracy. The enhanced ResNet50 increased the accuracy, precision, recall and F1 score to 98.78%, 98.82%, 98.68% and 98.75%, respectively and the value of Kappa was raised to 4.7%.

Yan et al. [32] suggested a new model that fuses (enhanced MobileNetV1 and EfficientNetB0). To improve the ability of feature extraction, add local-global attention after the top-level feature map. It uses transfer learning in EfficientNetB0 and a (3 × 3) convolution is added to the residual shortcut. The results showed that the new model achieved 94.58% in classification, 94.64% in precision and 93.22% in the coefficient of Kappa.

Prashantha and Prakash [33] suggested a model to identify the best handcrafted features employing a genetic algorithm and a finetuned CNN using three datasets: (TWB-HM, RD and TCIA-IXI). The findings of the suggested model performed well compared with state-of-the-art methods, achieving 99.40% accuracy with the DCA method on the RD dataset.

Benbakreti et al. [34] designed experiments to diagnose masses in breast images using three pretrained models (ResNet18, InceptionV3 and AlexNet) on three merged datasets (DDSM, MIAS and Inbreast). The ResNet18 achieved good results of 95% accuracy, 94.91% recall, 94.90% precision and 94.91% F1 Score.

3. METHOD

The architecture of the suggested deep-learning approach for classifying brain tumors is covered in this part. It combines a ResNet50 model with two attention mechanisms: Squeeze-and-Excitation (SE) blocks and Convolutional Block Attention Modules (CBAM).

The Attention mechanisms substantially improve (MRI) tumor categorization. This approach increases classification accuracy by improving the model's capacity to extract significant spatial and channel-wise characteristics of MRI data. These three strong deep-learning components are integrated in an

architecture called the ResNet50-SE-CBAM to significantly boost classification performance, feature extraction and model interpretability, particularly for difficult tasks, such as brain-tumor classification from MRI scans.

This work employs the Residual Network to extract features from preprocessed pictures employing pretrained ImageNet weights [35]. This model is composed of convolutional layers, identity blocks and a final softmax layer. The convolutional and max-pooling layer weights were locked for stability and to prevent changes during the experiment. ResNet has been chosen over pretrained networks because of its better performance and its ability to manage the vanishing gradient problem with skip connections. SE Blocks (Squeeze-and-Excitation) [36] are applied after residual blocks, which add channel attention: (Squeeze: Global Average Pooling, Excitation: Fully connected (FC) layers with ReLU + Sigmoid). Multiplies the original feature map by learned channel weights. That effect enhances important feature channels and suppresses irrelevant ones. CBAM (Convolutional Block Attention Module) [37] is applied after SE blocks to complement channel attention with spatial attention, consisting of two sub-modules: (Channel Attention and Spatial Attention). Channel Attention (similar to SE, but lighter) and Spatial Attention are applied to 2D convolution on concatenated average/max-pooled spatial descriptors, which produces a spatial attention map to focus on informative regions. That effect helps the model localize tumor regions in MRI scans.

In this work, the SE and CBAM attention modules were not integrated inside each residual bottleneck block, but rather were incorporated into the ResNet50 backbone at particular intermediate feature stages. The (conv3_block4_out) layer's output, which is the last block of the (Conv3_x) stage, was subjected to the SE module, yielding feature maps with dimensions of (28 x 28 x 512). In order to represent channel interdependencies and produce adaptive channel weights, the SE block now employs global average pooling for channel-wise recalibration, followed by a two-layer fully linked bottleneck structure. Before extracting deeper features, this improves mid-level semantic representations. The CBAM module was applied to the output of the (conv5_block3_out) layer, which yields high-level feature maps with dimensions of (7×7×2048). This layer is the final block of the (Conv5_x) stage. Channel attention and spatial attention are applied successively by CBAM in this position. Before shared fully connected layers, channel attention is calculated using both global average pooling and global max pooling. A convolutional layer with a (7×7) kernel is then used to produce spatial attention in order to highlight informative spatial locations. This allows the network to fine-tune high-level discriminative features before the final classification layers and Global Average Pooling. The integration method was depth-aware and sequential: CBAM was used at a deeper stage to improve both channel and spatial discrimination in high-level features, while SE was used at an intermediate stage to reinforce channel-feature learning in mid-level representations.

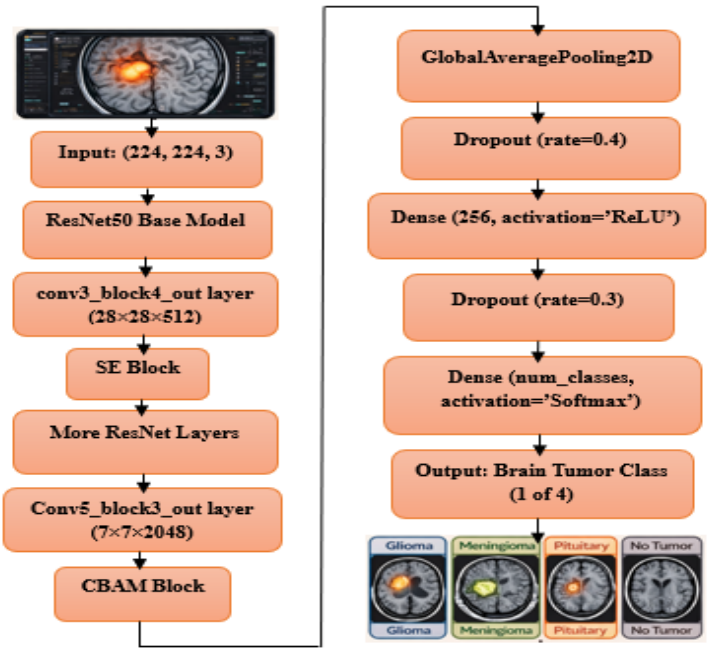


Figure 1. ResNet50-SE-CBAM architecture.

This selective stage-level integration reduces computing complexity and attention redundancy while preserving the stability of the original residual learning structure, in contrast to approaches that stack numerous attention modules within each residual block. (224 × 224) pixels make up the model's input, which is scaled with 1.0 / 255 to normalize the pixel values to the range [0, 1]. Figure 1 clearly illustrates the paradigm of the suggested approach for classifying brain tumors.

The framework of the proposed model is seen in the Figure above: A frozen, pretrained ResNet50 processes the input image. It applies the CBAM block at the deep feature layer (conv5 block3_out) and the SE block at the intermediate feature layer (conv3 block4_out). The subsequent layers are: dense, dropout, global average pooling and softmax classification. GlobalAveragePooling2D keeps spatially distributed information besides converting 3D feature maps into 1D vectors. Dropout lowers the overfitting effect. To better the decision boundaries, a dense layer incorporates learnable parameters.

4. RESULTS AND DISCUSSION

4.1 Datasets

Two datasets are employed in this work: the first dataset (Brain Tumor MRI) from Kaggle, which works with ResNet50 and other comparable CNN architectures. This collection consists of 7,023 MRI pictures of human brains classified into 4 tumor classes: “pituitary, meningioma, glioma and no tumor”. The “figshare, SARTAJ and Br35H” datasets were incorporated to make this dataset [38]. Because acquisition settings, institutions and many scanners are mixed, visual diversity is expanded, which contributes to robust model training. The dataset summary is provided in Table 1.

Table 1. Dataset summary of brain tumors.

Class	Training	Testing	Total
pituitary	1,457	300	1,757
meningioma	1,339	306	1,645
glioma	1,321	300	1,621
no tumor	1,595	405	2,000

The second dataset (Brisc) consists of 6,000 MRI pictures, also classified into 4 tumor classes: “pituitary, meningioma, glioma and no tumor”. This set consists of super-quality data that has been expertly annotated for classifying and segmenting the brain tumors. It addresses typical issues in current datasets (Figshare, BraTS), such as inconsistent annotation, narrow tumor focus and class imbalance. A realistic assessment of model generalization to unknown clinical data is made feasible by this dataset's entirely distinct distribution of brain MRI pictures that have never been seen in training sources [39]. The dataset summary is provided in Table 2.

Table 2. Dataset summary of Brisc.

Class	Training	Testing	Total
pituitary	1,457	300	1,757
meningioma	1,329	306	1,635
glioma	1,147	254	1,401
no tumor	1,067	140	1,207

This distribution is obviously skewed when compared with the initial dataset; in particular, the fraction of images that contain no tumor significantly decreased. This displays an additional obstacle for evaluating the reliability and resilience of classification models when real-world class imbalance occurs.

It applies ImageNet pre-processing for normalization and resizes each image to (224, 224) to meet the ResNet50 input specifications. During the training stage, it makes use of data-augmentation methods comprising rotation, flipping, translation and zooming. 20% of the dataset is employed for validation and 80% (remaining) is utilized for training. A collection of brain-image classes is demonstrated in Figure 2.

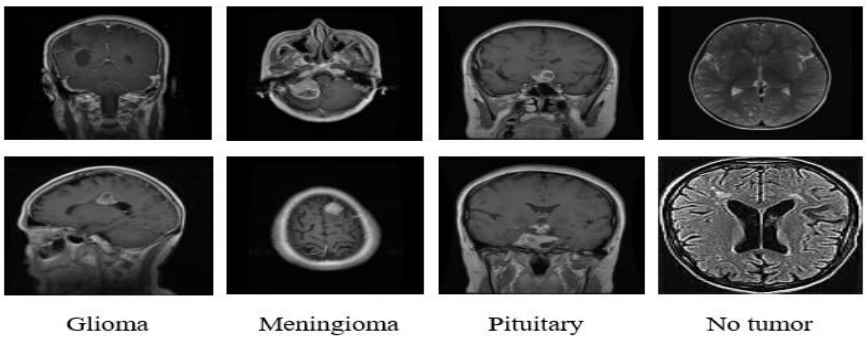


Figure 2. Several categories of brain images.

4.2 Experimental Setup

The proposed ResNet50 improved with "Squeeze-and-Excitation (SE) and Convolutional Block Attention Module (CBAM)", was assessed on two benchmark medical imaging datasets: "a Brain Tumor MRI dataset and the BRISC dataset". The tests employed transfer learning from pre-trained ImageNet weights and the model's upper layers were subsequently fine-tuned (the last residual block of ResNet50). A fair evaluation was ensured by dividing the dataset into a 20% validation set and an 80% training set. Generalization was enhanced by the use of data-augmentation methods, which stand for flipping, zooming and random rotations. The model's classification performance was outstanding across all evaluation metrics. The TensorFlow Python module was used to write all of the code. To run all the codes, the usual Google Colab platform with a GPU (T4) accelerator was used.

4.3 Training and Validation Performance

To train all models, we employed the AdamW optimizer. To raise stability and computational efficiency on GPU hardware, the training employed an "automatic mixed precision" method. In this work, the number of epochs and batch size were experimentally chosen depending on validation performance and convergence behaviour. Due to GPU memory limitations and its demonstrated efficacy in small-scale medical imaging datasets, a batch size of 16 was selected. The model converged within 35 epochs of training and additional training did not substantially increase validation accuracy, suggesting diminishing returns and possible overfitting. The additional parameters used during the training process are displayed in Table 3.

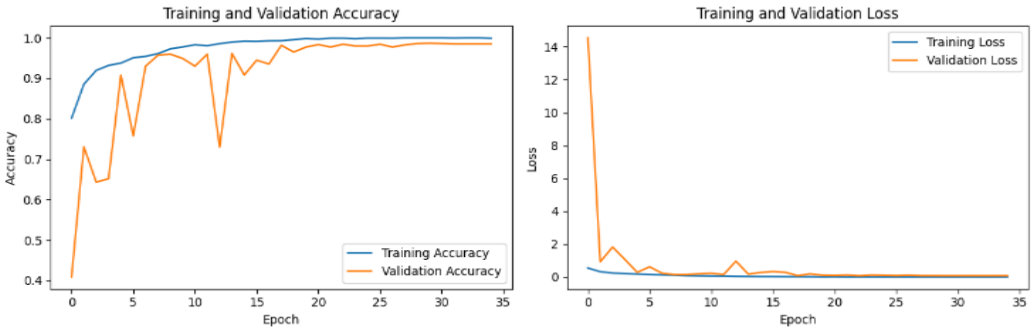
Table 3. The training parameters for the proposed model.

Parameters	**The value**
Optimizer	AdamW
Batch size	16
Epochs	35
Loss function	categorical_crossentropy
Learning rate	1e-4
Hidden-layer activation function	Relu
Classification-activation function	Softmax

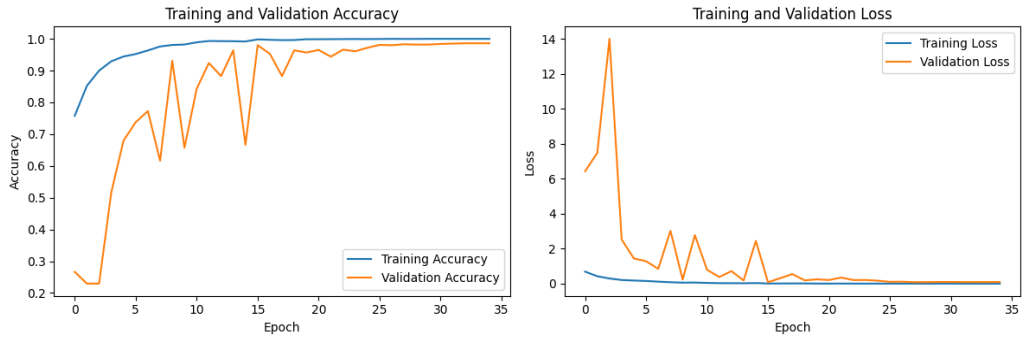
All models were trained on two datasets (Brain Tumor and Brisc). The suggested model can be seen in the accuracy and loss behaviour in Figure 3.

Each model was then developed and validated using a five-fold cross-validation method. Use a 5-fold to prevent overfitting and demonstrate the model's robustness. For each fold, the metrics report the mean ± standard deviation over three runs (See Table 4). Both datasets were subjected to 5-fold cross-validation in order to assess model robustness and minimize overfitting bias. The suggested ResNet50+SE+CBAM model demonstrated solid generalization performance by achieving the highest mean accuracy with the lowest standard deviation.

Strong baselines, CNN models, such as InceptionV3 and DenseNet121, were included to contextualize the gains (See Table 5). This demonstrates that the suggested model is competitive, even when compared to CNN baselines.



(a)



(b)

Figure 3. The proposed model's changing performance over time in terms of accuracy and loss on two datasets. (a) Brain Tumor dataset. (b) Brisc dataset.

Table 4. Cross-validation performance for each model on two datasets.

Dataset	Model	K-Fold	Accuracy (Mean ± Std)	Precision (Mean±Std)	Recall (Mean±Std)	F1-score (Mean±Std)
Brain Tumor MRI	ResNet50	5-Fold	0.9501±0.0038	0.9509±0.0150	0.9484±0.0226	0.9496±0.0087
	ResNet50+SE	5-Fold	0.9764±0.0024	0.9750±0.0098	0.9702±0.0164	0.9726±0.0103
	ResNet50+CBAM	5-Fold	0.9809±0.0019	0.9800±0.0064	0.9796±0.0077	0.9798±0.0072
	ResNet50+SE+CBAM	5-Fold	0.9945±0.0012	0.9914±0.0076	0.9903±0.0080	0.9908±0.0080
Brisc	ResNet50	5-Fold	0.9540±0.0035	0.9535±0.0085	0.9533±0.0092	0.9534±0.0092
	ResNet50+SE	5-Fold	0.9639±0.0029	0.9639±0.0081	0.9629±0.0279	0.9634±0.0106
	ResNet50+CBAM	5-Fold	0.9830±0.0018	0.9830±0.0107	0.9825±0.0165	0.9827±0.0158
	ResNet50+SE+CBAM	5-Fold	0.9860±0.0015	0.9860±0.0074	0.9857±0.0093	0.9858±0.0086

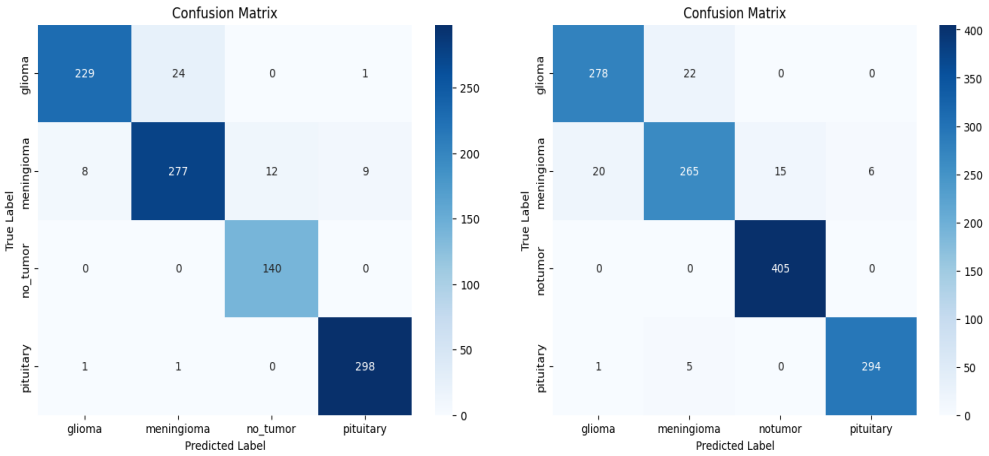
Table 5. The table of baseline comparison.

Model	MRI Accuracy	Brisc Accuracy
InceptionV3	0.9856	0.9845
DeseNet121	0.9849	0.9840
ResNet50+SE+CBAM	0.9945	0.9860

4.4 Performance Evaluation

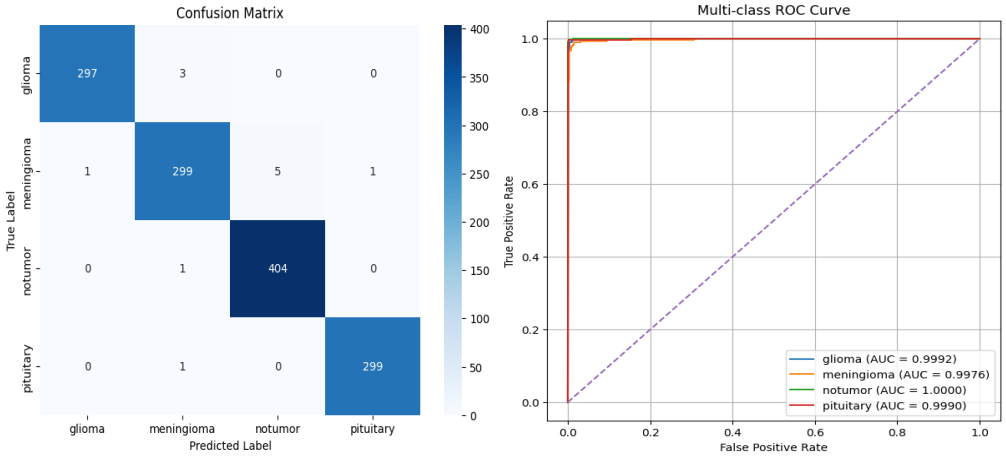
A confusion matrix and a Receiver Operating Characteristic (ROC) curve were employed for additional analysis of the classification performance. The confusion matrix exhibits how well the model can recognize various types of brain tumors, such as glioma, meningioma, pituitary and normal instances. The robustness of the suggested approach is confirmed by high “true positive and true negative” rates. Moreover, the model's great discriminative capacity across all classes is demonstrated by ROC curves and Area Under the Curve (AUC) values, underscoring its dependability in clinical-diagnosis settings.

Figure 4 shows the confusion matrix for the baseline ResNet50 model. Figures 5 and 6 show the confusion matrix and ROC curve for the proposed model. The results display a significant reduction in misclassification for all tumor classes, suggesting that the attention mechanisms improve discriminative feature learning and classification reliability.



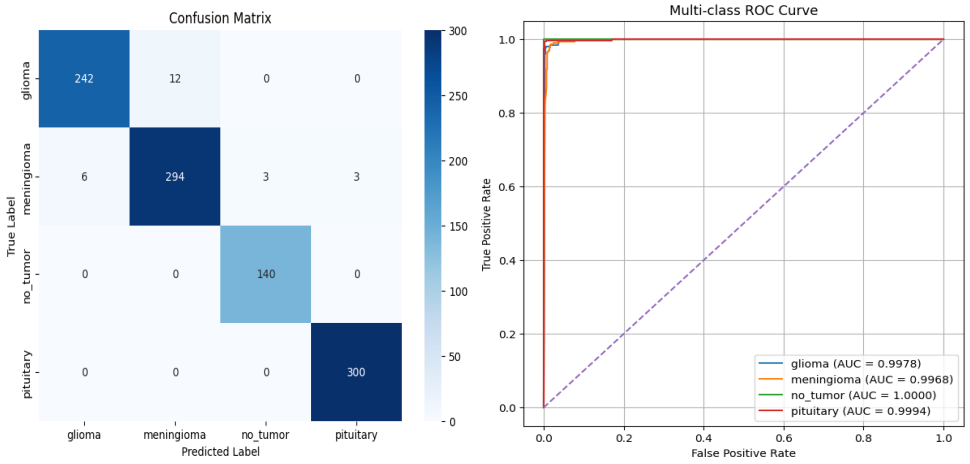
(a) (b)

Figure 4. The confusion matrix for the baseline ResNet50 model on two datasets. (a) Brain Tumor dataset. (b) Brisc dataset.



(a) (b)

Figure 5. The proposed model's performance on the brain-tumor dataset. (a) Confusion matrix. (b) ROC curve.



(a) (b)

Figure 6. The proposed model's performance on the Brisc dataset. (a) Confusion matrix. (b) ROC curve.

The suggested approach performs better on all classification tasks, as indicated by the confusion matrix. In particular, a greater number of correctly identified samples is shown by the confusion matrix's diagonal elements, which have significantly increased. There are fewer misclassifications when the values at other locations in the matrix decrease. Diagonal values indicate accurate

classifications, but off-diagonal values indicate inaccurate classifications [40], [41]. The suggested model (ResNet50+SE+CBAM) presented outstanding classification performance on the two datasets, demonstrating its capacity to capture distinct tumor-related characteristics efficiently (See Table 6).

Table 6. The metrics per class for the suggested model on two datasets.

Dataset	Model	Class	Pre.	Recall	F1-Sco.	AUC	Acc.	Loss
Brain Tumor MRI	ResNet50+SE +CBAM	Glioma	1.00	0.99	0.99	0.9992	0.9945	0.0366
		Meningioma	0.98	0.98	0.98	0.9976		
		No tumor	0.99	1.00	0.99	1.0000		
		Pituitary	1.00	1.00	1.00	0.9990		
Brisc	ResNet50+SE +CBAM	Glioma	0.98	0.95	0.96	0.9978	0.9860	0.0918
		Meningioma	0.96	0.96	0.96	0.9968		
		No tumor	0.98	1.00	0.99	1.0000		
		Pituitary	0.99	1.00	1.00	0.9994		

4.5 Error Analysis

The suggested model (ResNet50+SE+CBAM) was submitted to an error analysis, which revealed that the majority of misclassifications happen when there are imaging artifacts, low contrast, or small lesion regions. Confusion can result from benign lesions that visually are similar to malignant patterns (see Table 7).

Table 7. The table of class-wise performance and error analysis for the proposed model.

Dataset	Model	Class	Pre.	Recall	F1-Sco.	Major Error Source
Brain Tumor MRI	ResNet50+SE+CBAM	Meningioma	0.98	0.98	0.98	Small lesion size
	ResNet50+SE+CBAM	No tumor	0.99	1.00	0.99	Motion artifacts
Brisc	ResNet50+SE+CBAM	Glioma	0.98	0.95	0.96	Low contrast regions
	ResNet50+SE+CBAM	No tumor	0.98	1.00	0.99	Overlapping features

Table 7 displays class-wise performance and error analysis of the proposed model. High recall and precision are found in all classes. The majority of misclassifications happened in patients with motion artifacts, low-contrast imaging and small-lesion regions—all of which are frequent problems in clinical MRI acquisition. These results show that the suggested model is reliable, but they also point out several drawbacks that might be fixed with higher-resolution scans or multi-modal imaging.

4.6 Ablation Experiments

An ablation study was performed to assess the influence of SE, CBAM and ECA on classification performance, thereby investigating the contribution of each architectural component. This work measures the impact of each attention mechanism on enhancing feature representation on the Brain Tumor and BRISC datasets. Five models evaluated on two datasets for brain tumors are displayed in Table 8. ResNet50, ResNet50+SE, ResNet50+CBAM, ResNet50+ECA and ResNet50+SE+CBAM. The macro-average method was employed to calculate the ROC AUC in order to guarantee that each tumor class contributed equally and to reduce the impact of class imbalance. Five models were initialized with ImageNet pretrained weights and trained under the same conditions.

While the ResNet50 model had a strong basis, performance was greatly enhanced by including attention mechanisms. Although CBAM substantially increased performance by integrating channel and spatial attention, enabling the network to focus on discriminative problematic regions, the SE module allowed channel-wise feature recalibration.

By accurately imitating local cross-channel interactions without dimensionality reduction, the "Efficient Channel Attention (ECA)" mechanism further enhanced classification performance, revealing its efficacy with no processing overhead. ECA relies on a one-dimensional (1D) convolutional process instead of the fully connected layers used in traditional attention mechanisms, like SE-Net, thereby reducing information loss and achieving higher learning efficiency. When ECA is combined with the ResNet50 model, the representation of features extracted from images is enhanced, leading to improved performance in classification and computer-vision tasks at a lower computational cost [31], [42]. Yet, the best results were obtained with the combined SE + CBAM setup, suggesting

that simultaneous modeling of channel and spatial attention offers more thorough feature refinement than channel-only methods.

Table 8. Ablation-study results on brain tumor and brisc datasets.

Datasets	Models	Acc.	Loss	Pre.	Recall	F1 Sco.	AUC
Brain Tumor MRI	ResNet50	0.9501	0.2182	0.9509	0.9484	0.9496	0.9937
	ResNet50+SE	0.9764	0.1881	0.9750	0.9702	0.9726	0.9981
	ResNet50+CBAM	0.9809	0.1114	0.9800	0.9796	0.9798	0.9985
	ResNet50+ECA	0.9725	0.1457	0.9782	0.9781	0.9781	0.9962
	ResNet50+SE+CBAM	0.9945	0.0366	0.9914	0.9903	0.9908	0.9989
Brisc	ResNet50	0.9540	0.2571	0.9535	0.9533	0.9534	0.9935
	ResNet50+SE	0.9639	0.1728	0.9639	0.9629	0.9634	0.9936
	ResNet50+CBAM	0.9830	0.0800	0.9830	0.9825	0.9827	0.9973
	ResNet50+ECA	0.9795	0.1492	0.9795	0.9793	0.9794	0.9967
	ResNet50+SE+CBAM	0.9860	0.0918	0.9860	0.9857	0.9858	0.9985

4.7 Computational Complexity and Inference Speed

In this sub-section, we examined the computational complexity and inference speed of each model to assess their effectiveness. The theoretical complexity was estimated by calculating the number of parameters (in millions (M)), training time (in minutes (m)), floating-point operations (FLOPs) (in Giga (G)), Multiply–Accumulate operations (MACs) (in Giga (G)), memory usage in Gigabytes (GB) and inference speed (in milliseconds per image) on GPU (T4) (See Table 9). For clinical and real-world applications, this kind of efficiency assessment is essential.

The NVIDIA Tesla T4 GPU's highest memory consumption during training was used to calculate memory utilization. The findings demonstrate that although attention techniques add a few feature recalibration layers to memory requirements, the overall memory footprint is still manageable for real-world use.

Table 9. Computational complexity and inference speed for five models on two datasets.

Datasets	Model	Params(M)	Training Time(m)	FLOPs(G)	MACs(G)	Memory Usage (GB)	Inference Speed (ms/img)
Brain Tumor MRI	ResNet50	25.6	14	4.1	2.05	2.10	18.3
	ResNet50+SE	28.2	18	4.3	2.15	2.28	19.8
	ResNet50+CBAM	28.8	25	4.4	2.2	2.35	20.5
	ResNet50+ECA	25.8	16	4.15	2.075	2.15	18.6
	ResNet50+SE+CBAM	30.0	30	4.6	2.3	2.60	22.1
Brisc	ResNet50	25.6	12	4.1	2.05	2.05	17.9
	ResNet50+SE	28.2	14	4.3	2.15	2.22	19.3
	ResNet50+CBAM	28.8	18	4.4	2.2	2.30	20.0
	ResNet50+ECA	25.8	14	4.15	2.075	2.10	18.2
	ResNet50+SE+CBAM	30.0	23	4.6	2.3	2.55	21.6

The GPU used for the runtime evaluation was an NVIDIA Tesla T4. Training, validation and inference overhead are included in the total runtime, whereas the average batch processing time during training is represented by the runtime per iteration. Due to additional attention procedures, the integration of SE and CBAM resulted in a minor rise in computational cost, as predicted (See Table 10).

Table 10. The runtime evaluation for the suggested model on two datasets.

Dataset	Model	Acc.	Training Time (m)	Total Runtime (m)	Runtime/ Iteration (ms)
Brain Tumor MRI	ResNet50+SE+CBAM	0.9945	30	32.4	82
Brisc	ResNet50+SE+CBAM	0.9860	23	24.8	65

4.8 Explainability Method (Grad-CAM)

To improve the transparency and reliability of the suggested model, explainability analysis was conducted using Grad-CAM, "Gradient-weighted Class Activation Mapping." The suggested model learns clinically relevant features rather than focusing on irrelevant brain regions, as evidenced by the activation maps, which reveal that the model mainly concentrated on tumor regions (highlighted in red and yellow). For instance, the regions of interest are focused around diseased areas in cases of meningioma, glioma and pituitary tumor, demonstrating the model's capacity to identify the unique clinical characteristics of each tumor type.

The findings also show that the model has a high degree of interpretability, which is an important characteristic in medical applications, since it enables physicians to comprehend the logic underlying the model's conclusions. Consequently, the Grad-CAM maps verify that the suggested model offers a trustworthy visual representation of classification accuracy rather than operating as a black box [43]. Figure 7 illustrates Grad-CAM visualization on two datasets.

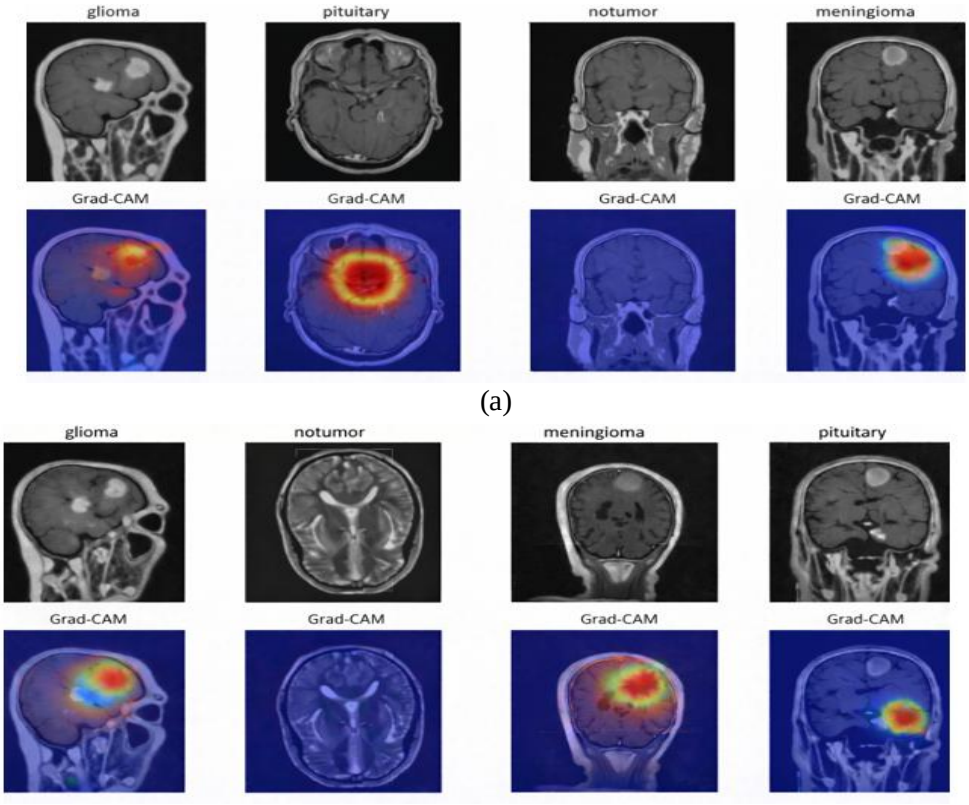


Figure 7. Grad-CAM visualization of the proposed model on two datasets.
(a) Brain tumor MRI dataset. (b) Brisc dataset.

4.9 Comparison of Results with Related Current Studies

The results of our investigation are now compared to those of other recent, related studies (See Table 11). All classification parameters from comparable studies, including "accuracy, precision, recall and F1-score", were outperformed by the results of our suggested model for accurate Brain Tumor MRI classification.

The proposed ResNet50-SE-CBAM model had an F1 score of 99.08%, accuracy of 99.45%, recall (sensitivity) of 99.03%, precision of 99.14% and AUC of 99.89% for the Brain Tumor MRI dataset and had an F1 score of 98.58%, accuracy of 98.60%, recall (sensitivity) of 98.57%, precision of 98.60% and AUC of 99.85% for the Brisc dataset.

The model further improves channel-wise feature recalibration by including Squeeze-and-Excitation (SE) blocks in addition to CBAM, which allows it to give priority to the most discriminative tumor features.

Table 11. The findings are in contrast to those of other recent studies.

Reference	Year	Model	Accuracy	Precision	Recall	F1 Sco.	AUC
Sahaai et al. [20]	2022	ResNet50	95.3%	93.7%	92.2%	94.6%	—
Oladimeji and Ibitoye [21]	2023	ResNet50+CBAM	99.43%	98.7%	99.01%	99.0%	99.25%
Vinston et al. [22]	2024	ResNet50+CBAM	99.35%	98.75%	99.11%	98.92%	99.53%
Huang and Prakash [23]	2025	ResNet50 SE CBAM	92.6% 98.4% 93.5%	—	—	—	98.7% 99.9% 99.3%
Lin et al. [31]	2025	ResNet50 Fusion ResNet50	95.27% 98.78%	95.40% 98.82%	95.32% 98.68%	95.36% 98.75%	—
Proposed Model (Brain Tumor MRI dataset)	2026	ResNet50+SE+CBAM	99.45%	99.14%	99.03%	99.08%	99.89%
Proposed Model (Brisc dataset)	2026	ResNet50+SE+CBAM	98.60%	98.60%	98.57%	98.58%	99.85%

The modest gain over earlier CBAM-only models, especially in F1 score and accuracy, suggests a better trade-off between “sensitivity and specificity”. Given the significant clinical hazards associated with both “false positives and false negatives”, this is especially important for medical applications.

The trend across the table confirms that attention mechanisms, like CBAM and SE blocks significantly boost CNN-based Brain Tumor MRI classification performance.

The proposed method sets a new benchmark with marginal, but consistent, improvements across all metrics. Even small percentage gains are meaningful in medical imaging, as they translate to fewer misclassifications in real-world clinical settings.

5. CONCLUSIONS

This work introduces a unique architecture for classifying Brain Tumor MRIs that combines ResNet50 with SE and CBAM on two datasets. With an F1-score of 99.08%, accuracy of 99.45%, recall (sensitivity) of 99.03%, AUC of 99.89% and precision of 99.14% on the Brain Tumor MRI dataset. And with an F1-score of 98.58%, accuracy of 98.60%, recall (sensitivity) of 98.57%, AUC of 99.85% and precision of 98.60% on the Brisc dataset.

The suggested ResNet50–SE–CBAM model outperformed a number of cutting-edge methods. By combining spatial–channel attention from CBAM with channel-wise recalibration from SE, the network was able to suppress irrelevant background input and concentrate more efficiently on tumor-relevant features, thereby improving its discriminative capacity. In addition, a sequential integration strategy was employed, where SE modules were applied in the early and intermediate stages of the network and CBAM modules were applied in the deeper layers, enabling hierarchical feature refinement and reducing attention redundancy. This sequential attention mechanism enhanced feature representation learning and contributed to the model’s superior classification performance.

The study's hopeful findings suggest that attention-enhanced deep-learning models, such as SE, CBAM and ECA and the integration between these mechanisms, have a lot of promise for strengthening the precision and effectiveness of medical image-based models used in research and development.

For future work, we will extend the research to larger, more varied datasets to ensure the model's generalizability. Finally, trying hybrid deep learning models, such as combining CNNs with Vision Transformers, could further enhance classification precision and adaptability in practical Brain Tumor MRI detection applications.

ACKNOWLEDGEMENTS

The authors would like to thank “the efforts of Mustansiriyah University in supporting this work.”

REFERENCES

[1] P. Sapra, R. Singh and S. Khurana, "Brain Tumor MRI Detection Using Neural Network," Int. J. of Innovative Science and Modern Engineering (IJISME), vol. 1, no. 3, 2013.

[2] S. Zhang and G. Xu, "A Novel Approach for Brain Tumor MRI Detection Using MRI Images," Journal of Biomedical Science and Engineering, vol. 9, no. 10, pp. 44–52, 2016.

[3] M. Alfonse and A. B. M. Salem, "An Automatic Classification of Brain Tumor MRIs through MRI Using Support Vector Machine," Egyptian Computer Science Journal, vol. 40, no. 3, pp. 11-21, 2016.

[4] J. Amin, M. Sharif, M. Yasmin and S. L. Fernandes, "A Distinctive Approach in Brain Tumor MRI Detection and Classification Using MRI," Pattern Recognition Letters, vol. 139, pp. 118-127, 2017.

[5] H. Dong et al., "Automatic Brain Tumor MRI Detection and Segmentation Using U-Net Based Fully Convolutional Networks," Proc. Med. Image Understand. Anal., pp. 506–517, Cham: Springer, 2017.

[6] S. K. V. Rao and B. Lingappa, "Image Analysis for MRI-based Brain Tumor Detection Using Hybrid Segmentation and Deep Learning Classification Technique," Int. Journal of Intelligent Engineering and Systems, vol. 12, no. 5, pp. 170–179, DOI: 10.22266/IJIES2019.1031.06, 2019.

[7] M. Sajjad et al., "Multi-grade Brain Tumor MRI Classification Using Deep CNN with Extensive Data Augmentation," J. of Computational Science, vol. 30, pp. 174–182, 2019.

[8] N. Abiwinanda et al., "Brain Tumor MRI Classification Using Convolutional Neural Network," World Congr. Med. Phys. Biomed. Eng., pp. 183–189, Singapore: Springer, 2018.

[9] A. Samreen et al., "Brain Tumor MRI Detection by Using Convolution Neural Network," Int. J. Online Biomed. Eng., vol. 16, no. 13, DOI: 10.3991/ijoe.v16i13.18545, 2020.

[10] H. Alsaif et al., "A Novel Data Augmentation-based Brain Tumor MRI Detection Using Convolutional Neural Network," Applied Sciences, vol. 12, no. 8, DOI: 10.3390/app12083773, 2022.

[11] M. S. I. Khan et al., "Accurate Brain Tumor MRI Detection Using Deep Convolutional Neural Network," Computational and Structural Biotech. J., vol. 20, DOI: 10.1016/j.csbj.2022.08.039, 2022.

[12] A. B. Abdusalomov, M. Mukhiddinov and T. K. Whangbo, "Brain Tumor MRI Detection Based on Deep Learning Approaches and MRI," Cancers, vol. 15, no. 16, DOI: 10.3390/cancers15164172, 2023.

[13] S. Albawi et al., "Understanding of a Convolutional Neural Network," Proc. of the 2017 IEEE Int. Conf. on Engineering and Technology (ICET), pp. 1–6, Antalya, Turkey, Aug. 2017.

[14] I. H. Sarker, "Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions," SN Computer Sci., vol. 2, no. 6, DOI: 10.1007/s42979-021-00815-1, 2021.

[15] R. Yamashita et al., "Convolutional Neural Networks: An Overview and Application in Radiology," Insights Imaging, vol. 9, no. 4, DOI: 10.1007/s13244-018-0639-9, 2018.

[16] Z. Li et al., "A Survey of Convolutional Neural Networks: Analysis, Applications and Prospects," IEEE Trans. Neural Netw. Learn. Syst., vol. 33, no. 12, DOI: 10.1109/TNNLS.2021.3084827, 2022.

[17] L. Alzubaidi et al., "Review of Deep Learning: Concepts, CNN Architectures, Challenges, Applications, Future Directions," J. Big Data, vol. 8, no. 1, DOI: 10.1186/s40537-021-00444-8, 2021.

[18] S. Anantharajan et al., "MRI Brain Tumor MRI Detection Using Deep Learning and Machine Learning Approaches," Measurement: Sensors, vol. 31, DOI: 10.1016/j.measen.2024.101026, 2024.

[19] M. Toğaçar, B. Ergen and Z. Cömert, "Tumor Type Detection in Brain MR Images Using a Deep Model Developed with Hyper-column Technique, Attention Modules and Residual Blocks," Medical & Biological Engineering & Computing, vol. 59, no. 1, DOI: 10.1007/s11517-020-02290-x, 2021.

[20] M. B. Sahaai et al., "ResNet-50 Based Deep Neural Network Using Transfer Learning for Brain Tumor MRI Classification," AIP Conf. Proc., vol. 2463, DOI: 10.1063/5.0082328, 2022.

[21] O. O. Oladimeji et al., "Brain Tumor MRI Classification Using ResNet50-convolutional Block Attention Module," Applied Computing and Informatics, DOI: 10.1108/ACI-09-2023-0022, 2023.

[22] V. R. Raja et al., "Enhanced Brain Tumor MRI Analysis Integrating ResNet50 with CBAM," J. of Electrical Systems, vol. 20, no. 6s, DOI: 10.52783/jes.3272, May 2024.

[23] K. A. Huang et al., "Evaluating the Impact of Attention Mechanisms on a Fine-tuned Neural Network for MRI Tumor Classification," Cureus, vol. 17, no. 3, DOI: 10.7759/cureus.80872, Mar. 2025.

[24] R. K. Yadav et al., "A Model for Brain Tumor MRI Detection Using a Modified Convolution Layer ResNet-50," Indian J. of Information Sources and Services, vol. 14, no. 1, DOI: 10.51983/ijiss-2024.14.1.3753, 2024.

[25] M. A. Khan and R. B. Z. Auvee, "Comparative Analysis of Resource-efficient CNN Architectures for Brain Tumor MRI Classification," arXiv preprint, arXiv: 2411.15596, Nov. 2024.

[26] M. Fachrurrozi et al., "Improving the Performance for Automated Brain Tumor MRI Classification on Magnetic Resonance Imaging," IAES Int. J. of Artificial Intell., vol. 13, no. 2, pp. 1679–1686, 2024.

[27] M. Z. Khaliki et al., "Brain Tumor MRI Detection from Images and Comparison with Transfer Learning Methods," Scientific Reports, vol. 14, no. 1, DOI: 10.1038/s41598-024-52823-9, 2024.

[28] S. Asif et al., "Improving Effectiveness of Different Deep Transfer Learning-based Models for Detecting Brain Tumor MRIs from MR Images," IEEE Access, vol. 10, DOI: 10.1109/ACCESS.2022.3153306, 2022.

[29] S. Deepak et al., "Brain Tumor MRI Classification Using Deep CNN Features *via* Transfer Learning," Computers in Biology and Medicine, vol. 111, DOI: 10.1016/j.compbiomed.2019.103345, 2019.
[30] H. P. A. Tjahyaningtijas et al., "Brain Tumor MRI Classification in MRI Images Using En-CNN," Int. J. of Intelligent Engineering and Systems, vol. 14, no. 4, DOI: 10.22266/ijies2021.0831.38, 2021.
[31] J. Lin, L. Huang, L. Ding and S. Yan, "Deep Brain Tumor MRI Lesion Classification Network: A Hybrid Method Optimizing ResNet50 and EfficientNetB0 for Enhanced Feature Extraction," Fractal and Fractional, vol. 9, no. 9, Art. no. 614, DOI: 10.3390/fractalfract9090614, 2025.
[32] S. Yan, M. Feng and Y. Cai, "An Integrated Deep Learning Model with Enhanced EfficientNetB0 and MobileNetV1 for Diabetic Retinopathy Grading," Biomedical Signal Processing and Control, vol. 113, Art. no. 108915, DOI: 10.1016/j.bspc.2025.108915, 2026.
[33] S. J. Prashantha and H. N. Prakash, "Feature Level Fusion Framework for Brain MR Image Classification Using Supervised Deep Learning and Hand Crafted Features," Jordanian Journal of Computers and Information Technology (JJCIT), vol. 8, no. 4, pp. 314–330, Dec. 2022.
[34] S. Benbakreti et al., "Using ResNet18 in a Deep-learning Framework and Assessing the Effects of Adaptive Learning Rates in the Identification of Malignant Breast Masses in Mammograms," Jordanian Journal of Computers and Information Technology (JJCIT), vol. 10, no. 1, pp. 93–107, Mar. 2024.
[35] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," arXiv preprint, arXiv: 1512.03385, 2015.
[36] J. Hu, L. Shen and G. Sun, "Squeeze-and-excitation Networks," Proc. of IEEE CVPR, DOI: 10.1109/CVPR.2018.00745, 2018.
[37] S. Woo et al., "CBAM: Convolutional Block Attention Module," Proc. ECCV, [Online], Available: https://arxiv.org/abs/1807.06521, 2018.
[38] M. Nickparvar, "Brain Tumor MRI Dataset," Kaggle, DOI: 10.34740/KAGGLE/DSV/2645886, 2021.
[39] A. Fateh et al., "BRISC: Annotated Dataset for Brain Tumor MRI Segmentation and Classification with Swin-HAFNet," arXiv preprint, arXiv: 2506.14318, 2025.
[40] M. Hossin and M. N. Sulaiman, "A Review on Evaluation Metrics for Data Classification Evaluations," Int. J. of Data Mining & Knowledge Management Process, vol. 5, no. 2, pp. 1–11, 2015.
[41] E. F. Gomes and R. S. Barbosa, "Deep Learning Approaches for Brain Tumor MRI Classification in MRI Scans: An Analysis of Model Interpretability," Applied Sciences, vol. 16, no. 2, Art. no. 831, DOI: 10.3390/app16020831, 2026.
[42] Q. Wang, B. Wu, P. Zhu, P. Li, W. Zuo and Q. Hu, "ECA-Net: Efficient Channel Attention for Deep Convolutional Neural Networks," Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 11534–11542, Seattle, USA, DOI: 10.1109/CVPR42600.2020.01155, 2020.
[43] R. R. Selvaraju et al., "Grad-CAM: Visual Explanations from Deep Networks *via* Gradient-based Localization," Int. J. of Computer Vision, vol. 128, no. 2, pp. 336–359, Feb. 2020.

ملخص البحث:

يُعدّ تصنيف صور الرّنين المغناطيسي لأورام الدّماغ أمراً بالغ الأهمّية للتّشخيص الدّقيق والمبكّر. وقد كشفت التطورات الحديثة في مجال التّعلّم العميق أنّ إدخال آليات الانتباه في الشّبكات العصبية الالتفافية يُحسّن أداء التّصنيف بشكلٍ ملحوظ. تقدّم هذه الدّراسة نموذجاً مبتكراً لتصنيف صور الرّنين المغناطيسي لأورام الدّماغ باستخدام RESNET50، تمّ تحسين أداءه بإدخال آلية الانتباه SE وآلية الانتباه CBAM. وقد جرى تجريبه على مجموعتي بيانات تحتويان على صور رنينٍ مغناطيسي لأورام الدّماغ. وأثبتت التّجارب حدوث تحسُّنٍ ملموسٍ في مختلف مؤشّرات الأداء مقارنةً بالنّموذج القائم على الشبكة العصبية الالتفافية الأساسية. وتبرز هذه النّتائج أهمية آليات الانتباه في نماذج التّعلُّم العميق للتّصوير الطّبي، الأمر الّذي يشير إلى إمكانية استخدام وحدات SE و CBAM كأدواتٍ موثوقةٍ وفعالة لتصنيف أورام الدّماغ في البيئات السّريرية. وينبغي أن تعمل الدّراسات المستقبلية آليات الانتباه القائمة على المحوّلات والتّقنيات الهجينة لتعزيز التّصنيف الآلي لأورام الدّماغ.