Journal of Trends in Computer Science and Smart Technology (ISSN: 2582-4104)
www.irojournals.com/tcsst/

Hybrid Channel-Spatial Attention Enhanced EfficientNet for Multi-Class Brain Tumor Classification Using MRI

Kiran Kumar M.[1], Dheeraj Kumar K V.[2], Gunji. Sreenivasulu[3], Kusuma S.[4]

[1,4]Assistant Professor, Department of Computer Science and Engineering (Data Science), Madanapalle Institute of Technology and Science (MITS), Deemed to be University, Madanapalle, Andhra Pradesh, India.
[2]PG Student, [3]Assistant Professor, Department of Computer Science and Engineering, Madanapalle Institute of Technology and Science (MITS), Deemed to be University, Madanapalle, Andhra Pradesh, India.

E-mail: [1]kiran9sai@gmail.com, [2]konenidheerajkumar@gmail.com, [3]gsr1251@gmail.com, [4]kusu87@gmail.com

Orcid ID: [1]0000-0001-8383-67671, [2]0009-0008-1963-54082, [3]0002-2140-07033, [4]0000-0003-3420-53144

Abstract

Brain tumor classification from Magnetic Resonance Imaging (MRI) remains challenging because of tumor heterogeneity, the overlapping intensity distribution of tumors, and the limited availability of medical images. This study introduces a lightweight deep learning model by integrating a Hybrid Channel–Spatial Attention (HCSA) module into the EfficientNet-B0 architecture for multi-class brain tumor classification. The HCSA module processes channels and spatial features sequentially to enhance feature representation by focusing on tumor-relevant information and better localizing the tumor region while maintaining computational efficiency. The proposed framework was evaluated on a publicly available Kaggle MRI brain tumor dataset, comprising 5,600 MRI images across 4 classes – glioma, meningioma, pituitary tumor and no-tumor. Comparative experiments were conducted with ResNet50, DenseNet121, EfficientNet-B0, and attention-enhanced EfficientNet models with identical preprocessing, data augmentation and training conditions. The proposed HCSA-EfficientNet-B0 model achieved a classification accuracy of 98.57%, an F1-score of 0.99, and an ROC-AUC value of 0.999 on the internal hold-out test set. Ablation results indicate that the combination of channel attention and spatial attention consistently outperforms each of the two attention mechanisms in isolation. Moreover, the proposed framework effectively captures the tumor-relevant regions, as suggested by Grad-CAM visualizations, which provide qualitative evidence supporting the interpretability of the classification results. The proposed framework is effective for multi-class brain tumor classification using MRI images and is able to improve classification performance while achieving computational efficiency.

Keywords: EfficientNet-B0, Hybrid Attention, Brain Tumor Classification, Grad-CAM; Multi-Class Classification; Attention Mechanism.

DOI: https://doi.org/10.36548/jtcsst.2026.3.014
Received: 11.05.2026, received in revised form: 15.06.2026, accepted: 07.07.2026, published: 18.07.2026

1. Introduction

Brain tumors are among the most serious neurological disorders due to their growth rate and heterogeneity, which result in various clinical problems. The timely and accurate diagnosis of such diseases plays an important role in assisting clinicians in making decisions and increasing patients' chances of recovery. The preferred method for imaging brain tumors is Magnetic Resonance Imaging (MRI), which provides high-quality visualization of the tissue in a non-invasive way. However, the evaluation of MRI images is complicated by the long time needed for it and the possible variability of the diagnostic process caused by different levels of clinicians' experience. These limitations have prompted the development of artificial intelligence-based computer-aided diagnostic (CAD) systems that assist clinicians in assessing brain tumors faster, more reliably, and objectively.

Recent advances in CNNs have led to significant improvements in the automated classification of brain tumors. The cross-comparative analysis of deep CNN models has shown high resilience to conventional machine learning algorithms in tumor heterogeneity [1]. Hybrid deep learning models also enhance reliability in detection and accuracy of classification[2], [3]. Also, it was confirmed that transfer learning practices promote better performance in generalization especially when annotated medical data is scarce [4], [5].

Architectures based on EfficientNet have also attracted a significant share of attention because of their compound scaling approach that achieves a balance between depth, width, and resolution without compromising computational efficiency. EfficientNet and EfficientNetV2 have shown good performance in tumor classification using MRI images [6], [7], [8]. Optimized B0, B4, and B7 variants of EfficientNet have delivered competitive multi-class classification accuracy [9], [10][11]. Moreover, more comprehensive adaptations such as EfficientNet-B9 are suggested to enhance the high-precision tumor detection. Although such developments have been made, even pure EfficientNet-based classifiers may still encounter difficulties in accurately localizing subtle tumor regions, particularly when tumor boundaries are ambiguous.

Even though channel-spatial attention mechanisms have been studied in computer vision, integrating them into lightweight medical image classification remains challenging, mainly because of computational constraints and the way features need to be represented. The proposed HCSA module follows the channel-spatial attention scheme defined in the CBAM model. Rather than introducing a new attention formulation, this study focuses on the task-specific integration of channel-spatial attention within the EfficientNet-B0 architecture for MRI-based brain tumor classification.

Attention mechanisms are added to CNNs to overcome disadvantages of spatial discrimination. Channel attention enhances discriminative feature maps, highlighting tumor-relevant channels, whereas spatial attention enhances the localization of abnormal regions in MRIs. EfficientNet frameworks with hybrid attention mechanisms have been shown to achieve better predictive results and interpretability [12], [13]. Strategies involving ensemble attention and self-attention further strengthen these approaches, albeit with increased computational complexity the [14], [15]. Thus, balancing classification accuracy and computational efficiency remains a significant research problem.

In addition to the classification, a number of studies have investigated joint segmentation-classification models based on UNet-EfficientNet hybrids [16], [17], [18]. These models enhance the localization of the tumor, although they need more supervision and are more costly to compute. Deformable approaches to attention-based segmentation also exhibit promising results [19], but due to their architectural complexity, they might not be able to be deployed in clinical environments in real time.

Though these have been developed, there are research gaps. To begin with, successful models are based on deep or ensemble architecture, which raises the number of parameters and time to inference. Second, frameworks that are segmentation-integrated come with extra annotations and computational costs. Third, there are not many studies that have explored lightweight yet effective hybrid attention mechanisms in effective backbone architectures in order to classify tumors in multi-class MRI.

To overcome these problems, this work proposes a Hybrid Channel-Spatial Attention (HCSA) for automated multi-class brain tumor classification. The proposed HCSA module selectively refines channel-wise and spatial characteristics, augmenting tumor discrimination at the cost of a lightweight architecture. As compared to transformer-based or ensemble-heavy architectures, the suggested architecture provides minimal overhead in terms of parameters and better classification robustness. Moreover, Grad-CAM visualization is also added to increase the interpretability and clinical relevance.

The key contributions of this work are:

- A lightweight Hybrid Channel-Spatial Attention (HCSA) module embedded into EfficientNet-B0 is developed.
- Extensive comparison and benchmarking with EfficientNet, transfer learning, and attention-based baselines.
- High classification accuracy while maintaining low computational complexity.
- Improved interpretability with the help of visualization-based tumor localization.

The proposed framework aims to provide an accurate, practical, and computationally efficient framework for MRI-based classification of brain tumors that could be implemented in real-world clinical practice.

2. Related Work

The recent trend in MRI-based brain tumor classification is deep learning, as it can automatically learn hierarchies of representations with respect to complex medical images. Initial convolutional neural network (CNN) models exhibited significant enhancements relative to more conventional machine learning approaches by directly extracting discriminative features of MRI scans without feature engineering. Significant comparative experiments proved that the current deep CNN designs perform better than traditional classifiers when it comes to treating tumor heterogeneity and inter-class similarity [1]. Hybrid deep learning systems also enhanced robustness with the integration of optimized learning pipelines and multi-stage processing strategies [2], [3]. Nevertheless, most of the initial CNN models were prone to overfitting and poor generalization, when used on heterogeneous data.

Transfer learning methods have become common across the board to alleviate data scarcity and enhance the generalization models. In medical imaging, fine--tuning pretrained networks has been demonstrated to converge faster and exhibit a better stability in performance [4], [5]. EfficientNet-based models have proven to be a particularly successful architecture among the various architectures tested, due to their systematic compound scaling strategy, which balances network depth, width, and input resolution in a systematic manner. EfficientNetV2-based frameworks performed well in multi-class classification and were computationally efficient [6]. Correspondingly, EfficientNet-B7 and EfficientNet-B0 architecture refined and improved using an optimization algorithm delivered competitive results with comparatively few parameters [9]. Though more advanced variants like EfficientNet-B9 resulted in better accuracy in recent studies, they were more complex, which is why their use is also being questioned in terms of deployability in resource-limited clinical settings.

Moreover, while EfficientNet backbones have strong representational capability, pure convolutional networks might fail to focus on tumor-relevant tissue, particularly when tumor margins are unclear or have overlapping intensities across different tumor types. To overcome this weakness, attention mechanisms have been included to improve feature discrimination. Channel attention mechanisms re-established the feature maps by highlighting informative tumor properties, and spatial attention modules enhance localization in MRI images. Hybrid assessment models coupled with EfficientNetV2 proved to be more predictive and interpreted more effectively [13], [20]. TEnsemble attention techniques further added to classification robustness [21] although in most cases such methods commonly add extra computational burden and inference time.

Besides classification, a number of studies have examined joint segmentation-classification paradigms to improve spatial awareness during tumor detection. UNet-EfficientNet hybrid networks made it possible to perform segmentation and classification simultaneously, enhancing localization performance [22], [23]. Adaptive attention-based segmentation models further improved multi-class tumor segmentation using saliency mapping processes. Although these combined models can be very high-performing, they require further pixel-level annotations and a significant amount of computation, which can constrain on scalability and real-time application.

In this work, we focus on the task of tumor classification, as classification-driven models require minimal labelling effort, decreased computational load, and better real-time implementation compared to their segmentation-driven counterparts.

In general, the available literature indicates that EfficientNet systems, transfer learning schemes, and attention systems can significantly enhance MRI-based brain tumor classification. Many high performing models, however, are based on deeper backbones, ensemble attention schemes, or segmentation-based designs, which add parameters and inference cost. The need still exists for lightweight but effective hybrid attention to improve channel-wise discrimination and spatial localization without impacting computational efficiency. Driven by these constraints, this paper proposes a Hybrid Channel-Spatial Attention (HCSA) module as a part of an EfficientNets-B0 backbone. The suggested framework will arrive at a trade-off between discriminative power, interpretability, and computational efficiency, thus allowing the brain tumor classification model to be applied to multiple classes in a real-world clinical setting.

3. Dataset and Preprocessing

This study utilized a publicly available dataset from Kaggle Brain Tumor MRI, which is frequently used for multiclass brain tumor classification. The dataset comprises 5,600 axial brain MRI images, classified into four diagnostic categories: glioma, meningioma, pituitary tumor, and no tumor. The images are realistic, presenting a broad spectrum of tumor sizes, shapes, locations, and intensity patterns, reflecting real clinical imaging conditions. The dataset does not provide MRI sequence labels (e.g., T1- or T2-weighted); therefore, no assumptions regarding MRI acquisition sequences are made in this study. Figure 1 demonstrates typical MRI images from the dataset, showcasing the diversity of tumor morphology and anatomical variation across the various tumor classes.

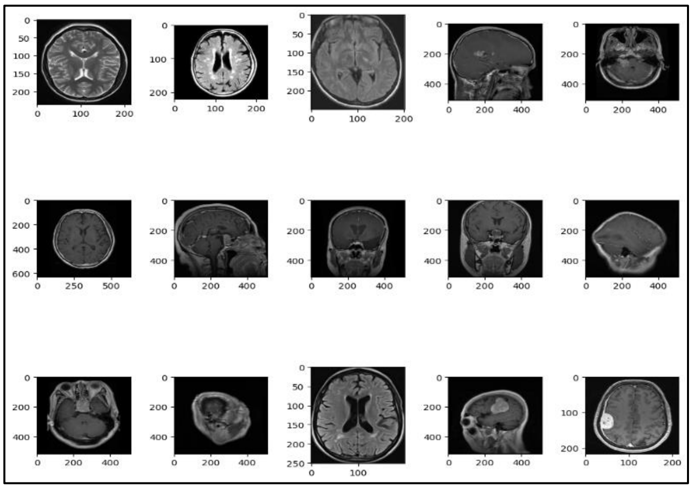

Figure 1. Representative brain MRI tumor images

The Kaggle Brain Tumor MRI Dataset [25] was selected due to its public availability and balanced representation across four diagnostic classes, providing consistent imaging samples that facilitate experimentation. The dataset contains an equal number of MRI images for each of the four categories: glioma, meningioma, pituitary tumor, and no-tumor, which enables unbiased testing across these distinct classifications. Furthermore, the dataset exhibits variation in tumor characteristics, including shape, size, and location, allowing for a comprehensive evaluation of lightweight deep learning methods in multi-class brain tumor classification.

An 80:10:10 stratified split was applied at the image level, resulting in 4,480 images for training, 560 for validation, and 560 for testing. This stratification method ensured an equal number of images per category within each of the three data sets, thereby maintaining a balanced class distribution during both the training and validation phases. Given the absence of patient-specific identifiers, image-level stratification was performed, consistent with common practice in medical imaging data analysis. Table 1 presents the distribution of MRI images by class.

Table 1. Distribution of MRI images across tumor classes and dataset splits

Class	Total Images	Training	Validation	Testing
Glioma	1,400	1,120	140	140
Meningioma	1,400	1,120	140	140
Pituitary	1,400	1,120	140	140
No-tumor	1,400	1,120	140	140
Total	5,600	4,480	560	560

3.1 Preprocessing

A standardized preprocessing pipeline was used to process all MRI images to enhance feature consistency and stability in training. As the dataset had pre-cropped brain areas, explicit skull stripping was not needed. To reduce inter-image intensity variations due to variations in acquisition conditions, Z-score intensity normalization was implemented to ensure the absence of such variations. All MRI scans were normalized to a resolution of 224 by 224 pixels and then formatted into three-channel RGB images (224 × 244 × 3). This resolution was adopted because it coincides with the typical input size of the pre-trained EfficientNet-B0 model while providing a good compromise between the amount of information retained from the tumors and the efficiency of computation during the training and prediction processes.

To prevent data leaks, dataset splitting was performed prior to the data augmentation step. Data augmentation was only performed on the training split, leaving the validation and testing sets untouched without any exposure during the model's training process. Additionally, each individual image was confined to being included in only one dataset split to avoid any overlapping between the three different data splits. Considering the publicly available dataset, which lacks patient IDs, it was impossible to confirm that no patients overlapped between the three data splits.

3.2 Data Augmentation

Data augmentation was used in training to increase model generalization and minimize overfitting. Random horizontal flipping, small rotations (±10%), changes in brightness and contrast, random zooming (±10%), and Gaussian noise were the augmentation operations. The simulated changes are realistic variations that can be met in the acquisition of clinical MRI. Only the training set received changes. The validation and test sets remained unchanged because it was crucial not to influence the performance evaluation.

4. Proposed Methodology

This section explains the suggested system for brain tumor classification with an EfficientNet-B0 backbone that would receive an additional Hybrid Channel-Spatial Attention (HCSA) process. The structure is meant to improve channel-wise discrimination of features and localization of tumor regions in brain MRI.

The dimensions of feature tensors for the architecture of the HCSA-EfficientNet-B0 model are illustrated in Figure 2. The diagram illustrates how the feature tensors propagate through the network, indicating the size of the spatial dimension and the channel depth of the feature tensors at each level. The HCSA blocks are used at Level 3 and Level 5, representing low and high-level feature extraction, respectively, allowing for both detailed anatomical features and high-level semantic features to be refined.

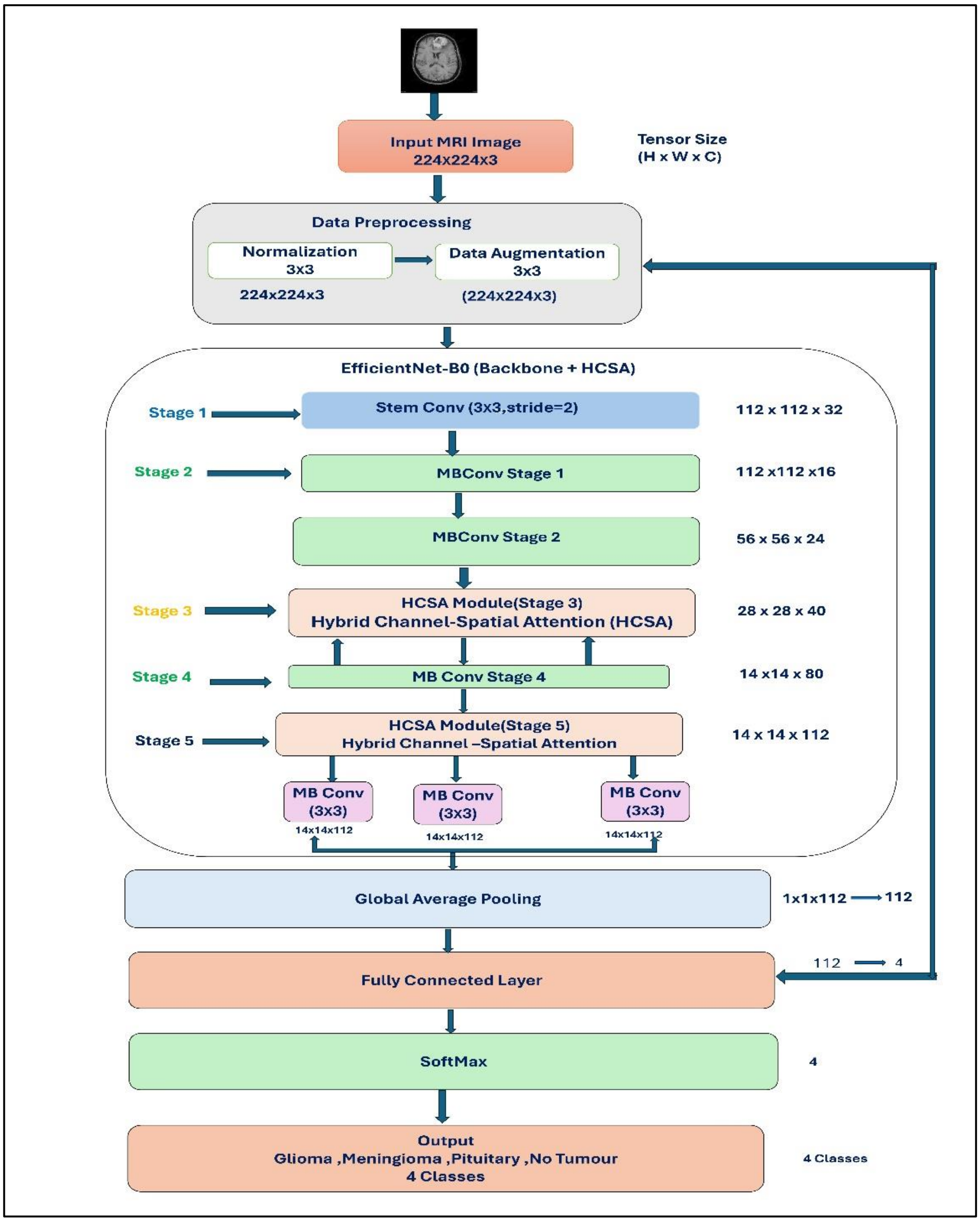

Figure 2. Architecture of the proposed brain tumor classification framework

4.1 Pipeline Overview

The proposed system's sequential steps in the processing pipeline are the entry of the MRI images, followed by preprocessing, feature extraction, attention refinement, and classification. Finally, the system outputs a prediction of the tumor classification. This general mapping can be stated as follows:

$$X \overset{P(\cdot)}{\rightarrow} \tilde{X} \overset{E(\cdot)}{\rightarrow} F \overset{A(\cdot)}{\rightarrow} F^{*} \overset{f(\cdot)}{\rightarrow} \hat{y}\ ) \tag{1}$$

Where:

- X denotes the input MRI slice,
- $P(\cdot)$ represents preprocessing operations,

- $E(\cdot)$ corresponds to the EfficientNet-B0 feature extractor,
- $A(\cdot)$ denotes the HCSA attention mechanism,
- $F$and F^* represent feature maps before and after attention refinement, respectively,
- $f(\cdot)$ is the classifier head,
- $\hat{y}$ denotes the predicted tumor class.

This pipeline facilitates step-wise refinement of feature representation, where information on features discriminative of tumors is highlighted and any irrelevant background feature is are suppressed.

4.2 EfficientNet-B0 Backbone

EfficientNet-B0 is selected for feature extraction due to its compound scaling, which balances network depth, width, and input resolution to achieve high accuracy rates and low resource costs. It consists of MBConv (Mobile Inverted Bottleneck Convolution) blocks, ideal for feature learning. The initial layers of the backbone are frozen to preserve low-level texture and edge information, while the subsequent layers are tuned to discover tumor-specific semantic patterns. To further improve feature representation, HCSA modules are selectively utilized in Stages 3 and 5 of the EfficientNet-B0 architecture. The features learned in Stage 3 include intermediate-level anatomical texture and edge details, along with the local boundaries of the tumor, while the features learned in Stage 5 include high-level semantic representations that play an essential role in distinguishing between various tumor classes. By using HCSA modules at these two particular stages, the proposed model can refine both intermediate and high-level features without increasing computational overhead.

4.3 Hybrid Channel-Spatial Attention (HCSA)

In an effort to enhance the quality of feature representation, this work presents a Hybrid Channel-Spatial Attention (HCSA) block, which integrates channel attention with spatial attention in a sequential order. The design of the HCSA module was inspired by the nature of brain MRI data, where differences in intensity and vague delimitations of tumors thighlight the deficiency of conventional attention mechanisms.

The proposed HCSA model follows the same sequential approach of channel-attention and spatial-attention as the Convolutional Block Attention Module (CBAM) [24]. However, the main distinction lies not in the definition of the attention mechanism itself, but in how this mechanism is incorporated into the overall architecture, namely the EfficientNet-B0. Attention refinement is only performed on the feature maps from Stage 3 and Stage 5, as they offer complementary intermediate and high-level feature representations, which are most useful for discriminating the brain tumor characteristics. Stage 3 extracts intermediate anatomical textures, edge information, and local tumor boundaries, whereas Stage 5 extracts high-level semantic representations that are essential for accurate tumor classification. By limiting the HCSA modules to these stages, the discriminative feature learning capability is improved while maintaining the light architecture and efficient computation of EfficientNet-B0, without needing to insert attention modules throughout the entire network.

Table 2. Architectural design differences between CBAM and the proposed HCSA integration strategy

Design Aspect	CBAM [24]	Proposed HCSA Integration Strategy
Attention Structure	Sequential Channel Attention followed by Spatial Attention	Sequential Channel Attention followed by Spatial Attention
Channel Descriptor Generation	Global Average Pooling and Global Max Pooling	Global Average Pooling and Global Max Pooling
Spatial Descriptor Generation	Channel-wise Average Pooling and Max Pooling with 7×7 Convolution	Channel-wise Average Pooling and Max Pooling with 7×7 Convolution
Attention Formulation	Generic CBAM attention mechanism	Same attention formulation adopted within EfficientNet-B0
Backbone Dependency	Applicable to various CNN architectures	Integrated specifically within EfficientNet-B0
Attention Placement	Typically inserted throughout CNN feature extraction stages	Selectively inserted at Stage 3 and Stage 5 feature representations
Feature Refinement Objective	General feature enhancement	Refinement of intermediate and high-level MRI tumor features
Application Domain	General computer vision tasks	Multi-class brain tumor classification from MRI images
Computational Objective	Improved feature representation	Lightweight attention integration with minimal computational overhead

A comparison of the original CBAM module and the proposed integration of the HCSA module is presented in Table 2. Although the proposed HCSA module uses the exact same channel-attention and spatial-attention formulation as CBAM, the novelty of the current study does not rest on the creation of another attention mechanism. Instead, the novelty in this case comes from the selective use of the channel-spatial attention module to refine the intermediate and high-level feature representations in the EfficientNet-B0 network, thus improving its classification performance.

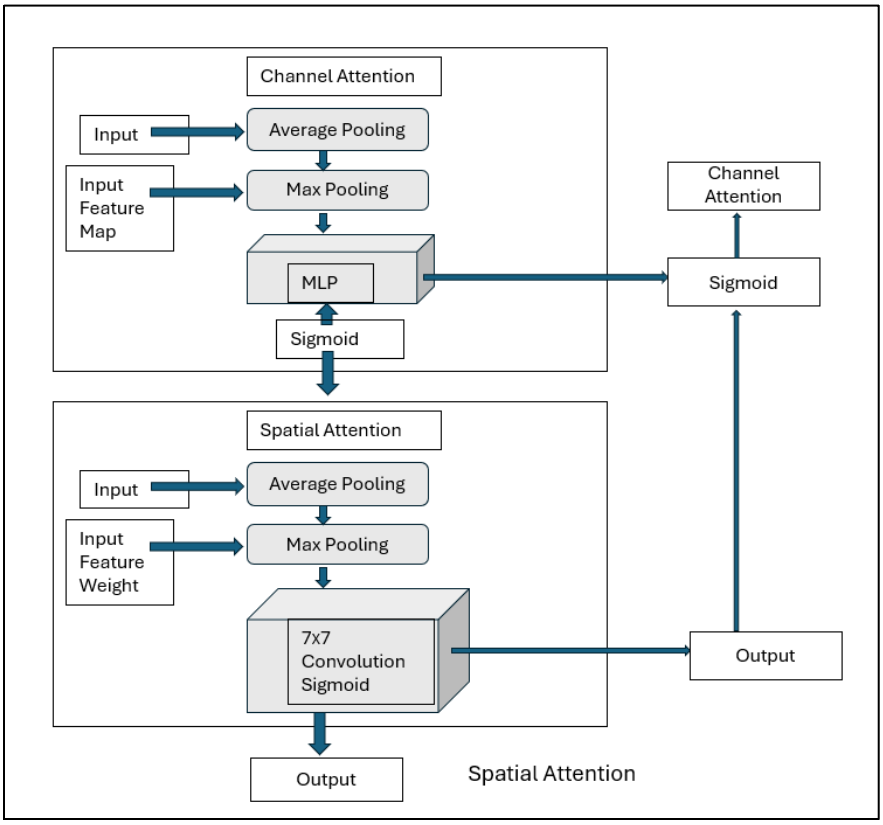

Figure 3. Hybrid channel–spatial attention (HCSA) module architecture

The HCSA module that we have proposed differs from CBAM in that it is designed explicitly to optimize tumor classification using light-weight MRI images with the EfficientNet-B0 framework. It is introduced selectively within the layers of EfficientNet, at levels where feature extraction can be enhanced with little impact on computational requirements.

The HCSA block is able to direct the network to improve channel-wise feature importance in conjunction with spatial localization; therefore, bringing out tumor-relevant patterns by deemphasizing the effect of background structures. Such a focused optimization of features contributes to better classification reliability and model interpretability. The architecture of the proposed HCSA module framework is depicted in Figure 3. Despite being developed based on the CBAM methodology, the suggested model applies an integration technique in a task-oriented manner by including HCSA modules only in stages 3 and 5 of the EfficientNet-B0 model architecture. As a result, the model benefits from intermediate and high-level features necessary for brain tumor detection without increasing its complexity.

4.3.1 Channel Attention Mechanism

Given a feature map: $F \in \mathbb{R}^{C\times H\times W}$ the channel attention mechanism aims to identify and enhance the informative feature channels. Two global context descriptors are computed using global average pooling and global max pooling:

$$F_{\text{avg}} = \text{AvgPool}(F), F_{\text{max}} = \text{MaxPool}(F) \tag{2}$$

where both outputs lie in $\mathbb{R}^{C}$. These descriptors are passed through a shared two-layer multilayer perceptron (MLP) with a reduction ratio r, followed by ReLU and sigmoid activations

$$M(F) = W_2\,\delta(W_1 F) \tag{3}$$

Where $W_1 \in \mathbb{R}^{\frac{C}{r}\times C}$, $W_2 \in \mathbb{R}^{C\times\frac{C}{r}}$, $\delta(\cdot)$ denotes the ReLU activation, and a sigmoid activation produces the final channel attention weights:

$$C_{\text{att}} = \sigma(M(F_{\text{avg}}) + M(F_{\text{max}})) \tag{4}$$

The channel-refined feature map is the following:

$$F' = C_{\text{att}} \odot F \tag{5}$$

It is a selective process since it enhances channels that are related to tumor-relevant textures and eliminates non-discriminative channels.

4.3.2 Spatial Attention-Based Feature Refinement

Channel attention enhances feature importance across channels. Spatial attention, in contrast, captures their spatial distribution within the MRI slice. Starting from the channel-refined feature map F' ,two spatial descriptors are obtained by applying average pooling and max pooling operations along the channel dimension.

$$F'_{\text{avg}} = \frac{1}{C}\sum_{c=1}^{C} F'_c \, , F'_{\text{max}} = \max_{1\le c\le C} F'_c \tag{6}$$

These descriptors are concatenated and passed through a convolutional layer with a kernel size of 7×7, followed by a sigmoid activation:

$$S_{att} = \sigma\left(Conv_{7\times 7}\left(\left[F'_{avg}, F'_{max}\right]\right)\right) \tag{7}$$

The spatially refined feature map is computed as:

$$F'' = S_{\text{att}} \odot F' \tag{8}$$

This mechanism enables effective localization of tumor-relevant regions within the MRI image.

4.3.3 Sequential Fusion Strategy

Channel attention is also used in the HCSA module, with spatial attention coming second. The resulting smooth representation of features is as:

$$F^* = S_{\text{att}} \odot (C_{\text{att}} \odot F) \tag{9}$$

The sequential multiplicative fusion would be more appropriate in MRI-based tumor classification since, at the channel level, salient textures and contrasts can be fixed, which allows faithful spatial localization of tumor regions. The multiplicative strategy has the same characteristics of reducing irrelevant anatomical structures while imposing minimal computational load.

4.3.4 Model Complexity

The computational efficiency of the designed HCSA-EfficientNet-B0 was architecture assessed by conducting a complexity analysis considering the number of trainable parameters, model size, inference time, and inference throughput. The corresponding complexity metrics are summarized in Table 3. As shown in Table 3, the proposed architecture contains 3.61 million trainable parameters and has a size of 13.77 MB, implying that the architecture is still lightweight after using HCSA blocks. On average, the inference process takes 25.20 ms to classify one MRI image and achieves 39.68 frames per second (FPS). This means that the designed architecture is computationally efficient but still has the capacity to perform the task of classifying brain tumors accurately.

Table 3. Computational complexity of the proposed HCSA-EfficientNet-B0 model

Metric	Value
Trainable Parameters	3.61 M
Model Size	13.77 MB
Average Inference Time	25.20 ms/image
Throughput	39.68 FPS

4.4 Classifier Head and Loss Function

After attention-based feature extraction, it is followed by global average pooling:

$$z = \frac{1}{HW}\sum_{i=1}^{H} \sum_{j=1}^{W} F^*_{:,i,j} \tag{10}$$

Where $z \in \mathbb{R}^C$ represents the compact feature vector. This vector is passed through a fully connected layer:

$$o = Wz + b \tag{11}$$

where $o \in \mathbb{R}^4$ corresponds to the four tumor classes. Class probabilities are obtained using the SoftMax function:

$$\hat{y}_i = \frac{\exp(o_i)}{\sum_{j=1}^{4} \exp(o_j)} \tag{12}$$

Categorical cross-entropy loss is used to optimize the network:

$$\mathcal{L} = -\sum_{i=1}^{4} y_i \log(\hat{y}_i) \tag{13}$$

where y_i represents the ground-truth one-hot encoded label.

4.5 Training Protocol

The network used AdamW for optimization. A single GPU was utilized, and the learning rate was set to 1×10^{-4} with weight decay of the same value for regularization. A cosine annealing schedule of the learning rate for optimization was used, defined as the following:

$$\mathrm{LR}(t) = \mathrm{LR}_{\min} + \frac{1}{2}(\mathrm{LR}_{\max} - \mathrm{LR}_{\min})(1 + \cos(\frac{t\pi}{T})) \tag{14}$$

The network was trained with the AdamW optimization technique and a cosine annealing learning rate schedule for a maximum of 100 epochs with a batch size equal to 16. The proposed method was trained and tested with the same fixed random seed under the experimental setup discussed in Section 5. The performance results obtained were based on one experimental setup.

The learning rate was dynamically adjusted with cosine annealing scheduling during the entire training process. This approach provided more stable convergence while gradually decreasing the learning rate, which helped the optimization routine run smoother and lowered the chance of overfitting. The identical scheduling policy was used for all compared architectures to ensurefair experimental comparison.

5. Experimental Configuration

A diverse range of baseline architectures was selected as a first step to measure the performance of the HCSA-enhanced EfficientNet-B0 model with as much rigor as possible. Such baselines contain a variety of architectural families, such as lightweight convolutional networks, traditional deep CNNs, and attention-augmented ones. They were all trained and tested under the same experimental conditions, using the same preprocessing policies, data augmentation policies, dataset splits, and training policies, which allowed for a productive and reproducible comparison.

5.1 Baseline Architectures

5.1.1 EfficientNet-B0 (Without Attention)

The standard EfficientNet-B0 architecture without any attention mechanisms serves as the primary baseline. This model establishes the reference performance of the backbone network and enables direct assessment of the performance gains achieved through the proposed HCSA module. ImageNet pretraining and the same training configuration as the proposed model were employed.

5.1.2 EfficientNet-B0 with Squeeze-and-Excitation Attention

Squeeze-and-Excitation (SE) modules are also added to the EfficientNet-B0 backbone to analyse the effect of channel-wise attention. The SE blocks were positioned at the respective network levels as the HCSA modules suggested, which maintains architectural similarity and allows for a direct comparison of channel-only attention to the hybrid one.

5.1.3 EfficientNet-B0 + Spatial Attention

A spatial-only attention variant was implemented by introducing a convolution-based spatial attention mechanism. This model emphasizes spatial localization of tumor regions without performing channel recalibration, allowing isolation of the spatial attention effect and comparison with the proposed hybrid design.

5.1.4 ResNet50

Medical practitioners have used ResNet50, a standard example of a residual network, in medical imaging for a long time. The residual unity facilitates the training of deep models. In this study, ResNet50 serves as a benchmark for comparing traditional architectures against those based on EfficientNet.

5.1.5 DenseNet121

DenseNet121 has a dense architecture whereby the feature maps of previous layers are reused throughout the entire network. Such interconnectivity is effective for using features efficiently and enhance gradient transmission in training. DenseNet121 is added to the group of models because its small size and demonstrated functionality with small medical imaging datasets make it a good model to compare alternative methods of feature propagation with the EfficientNet models.

5.2 Experimental Setup

5.2.1 Data Splitting and Evaluation Protocol

Each model was trained and tested on the Brain Tumor MRI dataset [25] with a stratified image-level split 80% for training, 10% for validation, and 10% for testing. All four tumor classes were proportionately represented in each subset through stratification. To ensure consistency in evaluations the same test set was used for all models. The test set was held out from the training and validation sets but originated from the same publicly available dataset [25]; therefore, the evaluation reflects internal hold-out testing rather than external dataset validation. Dataset partitioning was completed before augmentation, and all augmentation operations were restricted to the training subset to prevent information leakage into the validation and testing sets.

The proposed model was trained and evaluated using a fixed random seed and an internal hold-out test set. The reported performance metrics correspond to a single experimental run under the described experimental settings.

5.2.2 Preprocessing and Augmentation

Training was done on all the models using the same preprocessing pipeline containing z-score intensity normalization and 224x 224 pixel resizing. The same data augmentation strategies, such as random horizontal flipping, small-angle rotations, brightness and contrast manipulations, zooming, and Gaussian noise, were used. A single preprocessing and augmentation system will ensure that the observed variations in performance are due at least in large part to differences in architecture.

5.2.3 Training Configuration

Optimization of all the models was set to the AdamW optimizer with an initial learning rate of 1×10^{-4}cosine annealing learning-rate scheduling and a batch size of 16. Training was done for a maximum of 100 epochs with early stopping guided by the validation performances. This is a standard training arrangement which makes all of the assessed models comparable.

EfficientNet-B0 pretrained ImageNet weights were used for this transfer learning start, essentially for initialization purposes. During the optimization stage we fine-tuned the higher-level feature extraction layers along with the classification heads, while the more basic layers still kept their generic visual representations, in order to help convergence feel steadier, and also lower the overfitting risk.

5.2.4 Hardware Environment

All experiments were conducted on a single NVIDIA RTX 3070 with 8GB of memory. The EfficientNet-B0 was lightweight, preventing the HCSA module from imposing significant overhead; therefore, training and inference could be efficiently run on moderate computational hardware. Model complexity was assessed primarily through parameter count analysis and hardware resource requirements.

6. Evaluation Metrics for Model Performance Assessment

To fully evaluate the proposed HCSA-improved EfficientNet-B0 model and compare it with baseline architectures, multiple evaluation metrics were employed. This approach was chosen because the brain MRI data exhibited class imbalance and heterogeneous tumor features, and it combined global, class-wise, and balanced measures to provide a dependable and clinically significant assessment.

6.1 Primary Evaluation Metrics

In addition to accuracy, precision, recall, F1-score, and ROC-AUC evaluation, future analysis may also incorporate a Matthews Correlation Coefficient (MCC) and Cohen's Kappa coefficient. This could offer a more complementary assessment for multiclass classification reliability and prediction consistency simultaneously.

Accuracy: Accuracy is the percentage of correctly sized MRI images across all tumor types. Although this is a common metric, in cases of class imbalance, accuracy alone may not be sufficient. It is defined as:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \tag{15}$$

Sensitivity (Recall): Sensitivity is a measure of how accurately the model would diagnose cases of tumors. The sensitivity of the tests should be very high, since failure to detect tumors can be life-threatening in clinical practice:

$$\text{Sensitivity} = \frac{TP}{TP+FN} \tag{16}$$

Specificity: Specificity is a measure of the capability of the model to classify many non-tumors correctly, and therefore reduce false-positive predictions:

$$\text{Specificity} = \frac{TN}{TN+FP} \tag{17}$$

F1-Score: The harmonic mean of precision and recall is more appropriate with imbalanced dataset, which is specifically true with F1-score:

$$\text{F1} = 2 \cdot \frac{\text{Precision}\cdot\text{Recall}}{\text{Precision}+\text{Recall}} \tag{18}$$

Matthews Correlation Coefficient (MCC): The four elements of a confusion matrix are all included in the MCC, giving it a steady and balanced measure of classification quality. It has the status of being one of the most reliable measures of medical diagnostic assessment:

$$\text{MCC} = \frac{TP\cdot TN - FP\cdot FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \tag{19}$$

ROC-AUC: ROC-AUC is used to measure the capability of the classifier to differentiate between tumor and non-tumor classes at different decision thresholds, which gives information on the sensitivity-specificity trade-off.

Multiclass ROC-AUC performance was computed using a one-versus-rest approach, where each tumor category was separately assessed against the other categories and then aggregated with macro averaging overall.

6.2 Secondary Evaluation Metrics

Precision-Recall AUC (PR-AUC): PR-AUC is particularly informative in situations with class imbalance, as it focuses on the model's capability to correctly identify all tumor cases with a minimum number of false positives.

Balanced Accuracy: Balanced accuracy compensates for unequal class distributions by averaging sensitivity and specificity, providing a fairer assessment across tumor categories.

$$\text{Balanced Accuracy} = \frac{1}{2}(\text{Sensitivity} + \text{Specificity}) \tag{20}$$

This metric ensures equitable evaluation across all four tumor classes.

7. Results

7.1 Training and Validation Performance

The trends of accuracy for training and validation for 25 epochs are shown in Figure 4. It can be seen that there is fast convergence of the model at the beginning of training, where the accuracy for training increases significantly from about 87% initially to above 95% within the first few epochs. The trend continues until a maximum accuracy level of around 98.57% in the last epoch. Similar trends are noted for the validation accuracy, where the increase from around 94% at the beginning to near 99% in the later stages of training shows the consistent behaviour of the validation curve. Thus, both curves show a stable and increasing nature, indicating good generalization of the model.

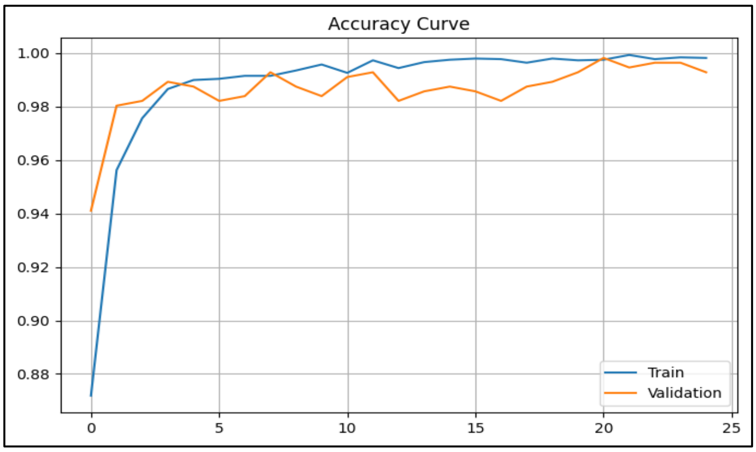

Figure 4. Training and validation accuracy curves of the proposed model

Minor fluctuations in validation accuracy happen during some epochs. This is common in brain MRI classification tasks because of similarities between classes and differences in tumor appearance. The close match between training and validation curves shows stable optimization and effective generalization. There is no sign of serious overfitting.

7.2 Test Set Performance and Confusion Matrix Analysis

The performance of the proposed approach was evaluated by achieving 98.57% accuracy for classification on an internal hold-out test dataset containing 560 MRI images. Based on the confusion matrix (shown in Figure 5), the excellent classification of tumors is evident, with MRI samples mainly falling into their corresponding classes.

The confusion matrix, when using the internal hold-out test set (560 MRI images) is shown in Figure 5. Overall classification accuracy was 98.57% with 552 images correctly classified. The confusion matrix shows very good diagonal dominancemeaning only eight images were misclassified.

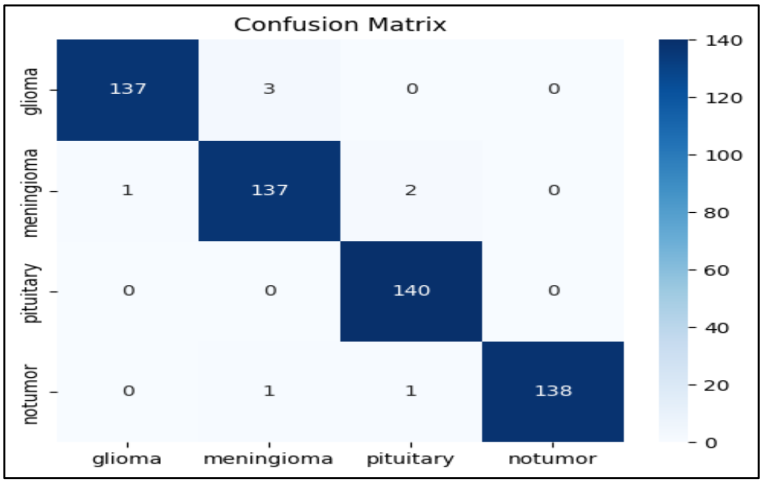

Figure 5. Confusion matrix of the proposed HCSA-EfficientNet-B0 model on the internal hold-out test set, demonstrating class-wise classification performance across the four tumor categories

For the glioma class, 137 of 140 images were correctly classified, with 3 misclassified as meningiomas. In the meningioma class, 137 images were correctly classified, images, 1 was incorrectly categorized as glioma, and two were misclassified as pituitary tumors. The pituitary tumor class achieved flawless classification, as all 140 images were correctly classified. In the no-tumor class, 138 images were correctly classified, with 1 image misclassified as meningioma and 1 as a pituitary tumor.

In summary, the confusion matrix demonstrates the suggested HCSA-EfficientNet-B0 model's capability to accurately categorize four diagnostic groups. There were only a few misclassifications between the glioma and meningioma types, which have similar MRI appearances and intensity patterns.

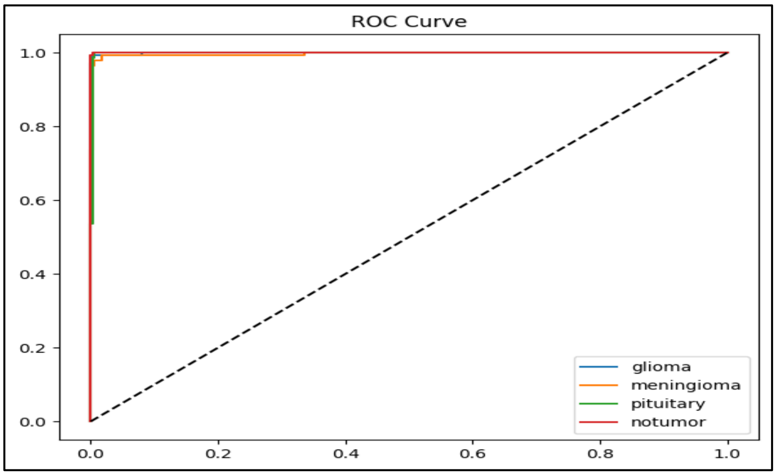

Figure 6. One-vs-rest receiver operating characteristic (ROC) curves for multiclass brain tumor classification

Figure 6 shows the ROC curve of the one-vs-rest strategy for the four tumor types. The curves stay close to the upper left area of the plot, suggesting high sensitivity and specificity regardless of the different decision levels. The ROC-AUC values generated are near 1.0 for all cases, showing the excellent discrimination ability of the suggested model with HCSA-based EfficientNet-B0.

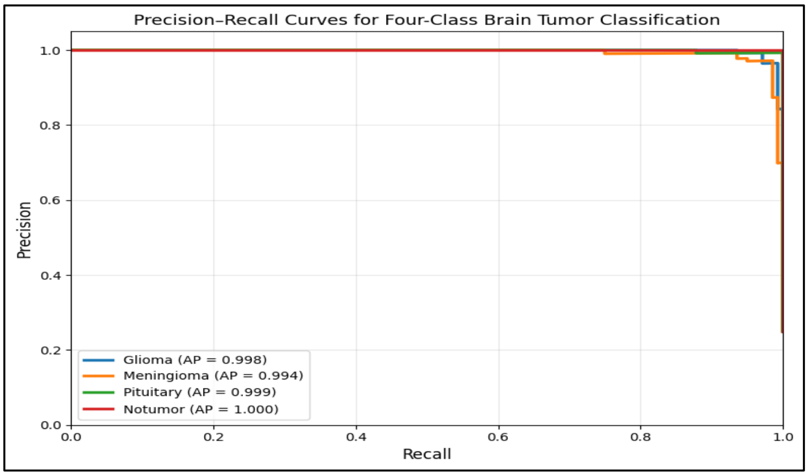

Figure 7. One-vs-rest precision–recall curves for four-class brain tumor classification

The one-vs-rest method provides Precision-Recall (PR) curves for each of the four classes, as shown in Figure 7, and the respective AP is provided in Table 4. PR curve visualization is especially important for measuring performance when working with imbalanced data, as it shows the relationship between precision and recall. From Table 4, the APs obtained from the suggested framework are 0.998 for the glioma class, 0.994 for the meningioma class, 0.999 for the pituitary tumor class, and 1.000 for the no-tumor class. This gives an average AP of 0.998.

Table 4. Precision–recall performance

Class	AP (PR-AUC)
Glioma	0.998
Meningioma	0.994
Pituitary Tumor	0.999
No-Tumor	1.000
Macro Average	0.998

7.3 Attention Visualization Analysis

In addition to analyzing the performance metrics of the proposed Hybrid Channel-Spatial Attention (HCSA) module, visualization analysis was also performed on representative MRI images. Figure 8 provides the attention maps generated by the proposed HCSA module, comprising the input MRI image, channel attention weights, the spatial attention map, and the output hybrid attention.

The purpose of the channel attention is to allocate adaptive weights to each feature channel based on its relative contribution to characterizing the tumor. As shown in Figure 8, only some channels have high values in the attention map, implying that the model learns to weight tumor-relevant discriminative features and suppress less relevant information channels. This allows for improving feature discrimination by focusing on tumor-specific patterns.

In addition, the spatial attention mechanism further analyzes and refines the obtained feature representation, assigning a specific weight to each location based on its clinical relevance to the tumor. As can be seen in Figure 8, the generated spatial attention map is highly activated near the tumor location, but there is low spatial attention in other parts of the image. This indicates that the proposed attention module can identify important regions of interest while suppressing noninformative parts of the image.

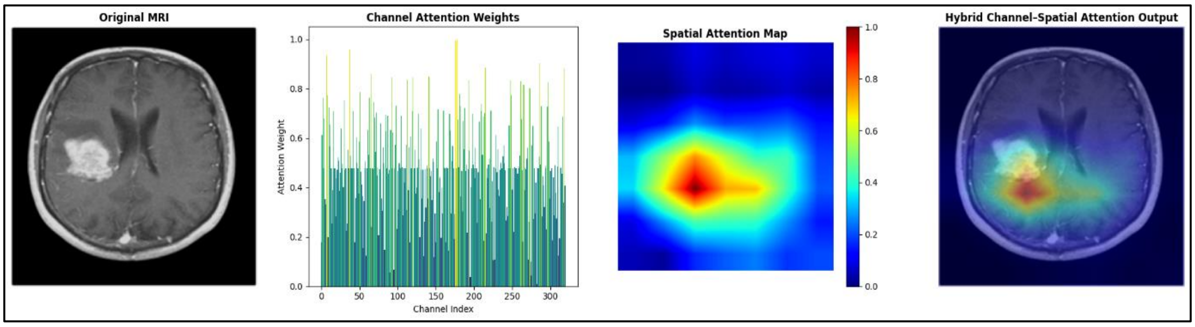

Figure 8. Visualization of the proposed hybrid channel–spatial attention (HCSA) module

Finally, as a result of fusing the outputs of the two attention modules, the Hybrid Channel-Spatial Attention generates an output attention map, which combines the advantages of both mechanisms. In this case, the activation of the output map corresponds to the tumor location depicted in the MRI image. Thus, the presented qualitative results serve as additional evidence of the effectiveness of the proposed HCSA mechanism, alongside its quantitatively proven performance.

The design of the Hybrid Channel-Spatial Attention (HCSA) module is shown in Figure 8. The channel attention part provides adaptive weighting for discriminatory feature channels, while the spatial attention part identifies regions relevant to tumors in MRI images. The hybrid output generated from the HCSA module shows the results after the two types of attention modules are successively utilized on the input features.

7.4 Quantitative Performance Summary

Precision and recall produced by the classes, and F1-scores, were generally high across the tumor categories. Though slightly lower recall was revealed in the meningioma class than in the other classes, the macro-averaged and weighted F1-scores were nearly the same (0.99) thus confirming the balanced and consistent classification performance. The obtained accuracy (98.57) is an indicator of the effectiveness of the suggested architecture in processing multiclass brain tumor classification.

In order to place the results into context, the developed model was compared to other popular models in terms of its accuracy. A performance comparison between all architectures is summarized in Table 5 below. For a fair comparison, all models were trained and tested under same data split, preprocessing and augmentation, optimizer settings, learning rate schedules

Table 5. Comparative performance of baseline architectures evaluated under identical experimental settings

Model	Accuracy (%)	F1-Score	AUC
ResNet50	96.8	0.96	0.972
DenseNet121	97.3	0.97	0.978
EfficientNet-B0	97.8	0.97	0.982
EfficientNet-B0 + SE	98.4	0.98	0.989
EfficientNet-B0 + HCSA (Proposed)	98.57	0.99	0.999

Although traditional deep networks such as ResNet50 and DenseNet121 worked, the EfficientNet-based networks performed better. The proposed framework, the HCSA-EfficientNet-B0 network, demonstrated the best classification accuracy (98.57%), F1-score (0.99), and ROC-AUC score (0.999), being at the same time lightweight in terms of parameters (3.61 million) and model size (13.77 MB), and had an average inference time of 25.20 ms per

MRI image (39.68 FPS). This means that it has a fine trade-off between classification and computation.

The baseline architecture models were all independently developed, trained, and tested using the same preprocessing and data augmentation techniques, data splits, optimizations, and evaluation metrics to ensure an even comparison with the HCSA-based EfficientNet-B0 model.

Apart from accuracy and F1-score, the proposed model achieved an MCC of 0.992 and a Cohen's Kappa value of 0.991, indicating that the model exhibits excellent consistency in multiclass classification.

7.5 Visual Interpretability Analysis

Representative MRI samples were used in Grad-CAM visualizations, and the results are provided in Figure 9 along with the original images. The activation maps also show that the proposed model consistently concentrates on regions relevant to tumors and inhibits non-informative background regions. Clear localization is observed in tumors of different sizes, shapes, and intensity appearances, as well as in irregular and heterogeneous lesions.

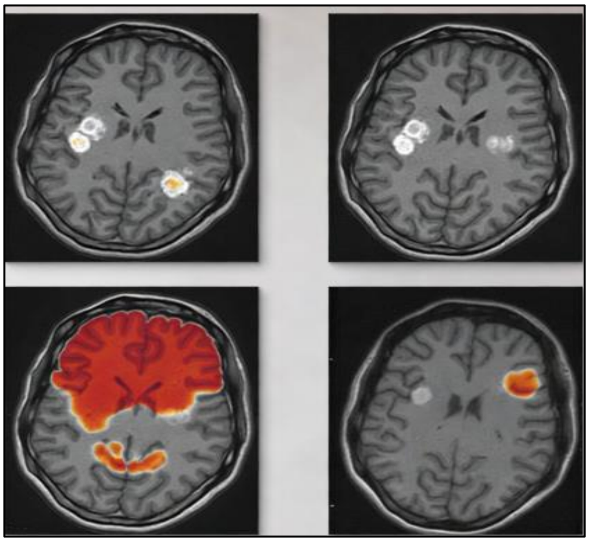

Figure 9. Grad-CAM heatmaps identify tumor-relevant areas in a representative brain MRI image, which indicates that the proposed model is effective at localizing its spatial features

These qualitative results support the quantitative ones and indicate that the given model can provide better interpretability, which is key to clinical decision-support systems and makes it possible to trust automated diagnostic models more.

Even with strong overall classification results, there were still a small number of wrong predictions, mainly for MRI samples with textures that seemed to overlap and tumor boundaries that were hard to separate clearly. In addition, low-contrast regions sometimes reduced the discriminative power. For future work, localization-guided refinement mechanisms could be blended in to push performance further in these more difficult imaging conditions.

8. Discussion

The experimental results demonstrate that the Hybrid Channel-Spatial Attention (HCSA) improved version of the EfficientNet-B0 model achieved a high accuracy of 98.57% on the Brain Tumor MRI Dataset for the classification of brain tumors. The proposed HCSA, combined with the EfficientNet-B0 framework, enhances the quality of feature representation by improving the ability to discriminate between tumor regions and the image background.

This approach is effective due to the complementary relationship between channel and spatial attention. Channel attention focuses on feature maps obtained from the textures, intensities, and structural features of tumors, while spatial attention identifies malignant regions in MRI images. The sequential combination of these two mechanisms makes them more informative and context-sensitive than single-attention strategies in feature learning. This is particularly advantageous in differentiating tumor types with similar visual appearances, such as glioma and meningioma.

Compared to current methods such as EfficientNet-based models, classical CNNs, and ensemble or hybrid models, the proposed model proves to be competitive or even superior in performance with a low computational footprint. Most previous studies report accuracies between 94% and 98% on comparable datasets, often at the expense of more complex models. In contrast, the suggested HCSA-EfficientNet-B0 framework is characterized by a low computational load, including 3.61 million trainable parameters, a 13.77 MB model size, and a processing time of 25.20 ms per MRI image (39.68 FPS). These factors confirm that the suggested framework provides better classification results without compromising computational efficiency.

The test was conducted on a single publicly available dataset, which might restrict its application to other imaging protocols, scanners, and clinical contexts. Furthermore, the existing research addresses classification and lacks explicit tumor segmentation and uncertainty estimation, which limits its applicability where accurate boundary description or confidence in decision-making is crucial. However, the suggested framework has the potential for clinical relevance. Although the hypothesis has not been clearly demonstrated, the high classification accuracy and interpretable Grad-CAM visualizations indicate that it can be used as a decision-support system to assist radiologists in tumor identification and diagnostic triage.

Though the proposed HCSA-EfficientNet-B0 framework achieved a test accuracy of 98.57%, all experiments were conducted using data from the same publicly available Brain Tumor MRI dataset. Although the test set was separated from the training and validation datasets, it originated from the same source and imaging process. Hence, the obtained accuracy values indicate generalization ability within the same dataset. However, the proposed framework was evaluated only on an internal hold-out test set. Consequently, its generalization performance on independent external datasets remains to be investigated.

While image-based stratification was done to ensure no overlap among train, validation, and test datasets, the absence of patient IDs in the public data makes it impossible to confirm patient-based stratification. Therefore, some overlap among patients cannot be ruled out. The inclusion of an internal dataset or a patient-based validation protocol must be considered in future studies.

Grad-CAM analysis provided qualitative interpretability evidence by highlighting diagnostically relevant image regions that contribute to the prediction outcomes. However, quantitative explainability evaluation metrics were not included in the present investigation, and this will be an important next step for future work. The present study provides only qualitative explainability analysis, while quantitative localization evaluation metrics such as insertion-deletion analysis and localization accuracy were not included. In addition, the EfficientNet-B0 backbone is computationally highly efficient and can be run in resource-limited settings. Further research includes validation through the use of multi-center datasets, generalization to multiple modalities of MRI datasets, and incorporating modules for segmentation and uncertainty quantification.

However, in spite of multiple tests carried out at different random initializations in order to increase the reliability of results, the current research did not conduct any tests for statistical significance. As a result, there is no statistical significance of the proposed HCSA-EfficientNet-B0 method over other baseline methods. Further research should include statistical hypothesis testing, including t-tests and non-parametric tests.

Even though the evaluation experiments produced encouraging classification results on the chosen benchmark dataset, validation with different datasets should be conducted. Some future investigations will evaluate model generalization using extra publicly available brain tumor MRI datasets in order to further assess deployment feasibility.

The proposed HCSA-EfficientNet-B0 architecture also exhibits good computational efficiency. The trained model has 3.61 million trainable parameters, and the model size is 13.77 MB. The average inference time of 25.20 ms per image, or 39.68 FPS, shows that the model has the potential to be efficient in the MRI image classification process and maintain the lightweight characteristics of practical computer-aided diagnosis systems.

9. Conclusion

This study proposed a lightweight Hybrid Channel-Spatial Attention (HCSA)-based EfficientNet-B0 model for multi-class brain tumor classification using MRI. The proposed HCSA module was designed to improve channel-wise feature discrimination and spatial localization while maintaining computational efficiency. Experimental evaluation demonstrated that the proposed model achieved a 98.57% classification accuracy, a 0.99 macro F1-score, and a 0.999 multi-class ROC-AUC on the internal hold-out test set. Furthermore, the Precision-Recall plot indicates that AP scores achieved for glioma, meningioma, pituitary tumor, and no-tumor classes are 0.998, 0.994, 0.999, and 1.000, respectively. Attention mechanism visualization and Grad-CAM analysis further confirm the effectiveness of the proposed method in focusing on the tumor areas. Even though image-based splitting was implemented, patient separation cannot be validated due to the lack of patient-level information for each sample. Therefore, this is the main limitation of the present paper. Future work includes validation of the framework on independent multi-center data, patient-level experiment protocol design, and extending the architecture to jointly localize and classify the tumor. To conclude, a more efficient and interpretable framework called HCSA-EfficientNet-B0 is introduced and analyzed in the context of brain tumor classification based on MRI images. The HCSA-EfficientNet-B0 framework provides accurate, computationally effective, and transparent means of automatically classifying brain tumors using MRI images. Moreover, the proposed HCSA-EfficientNet-B0 architecture is shown to be computationally efficient with only 3.61 million trainable parameters, a network size of 13.77 MB, and an average inference time of 25.20 ms per MRI image (39.68 FPS), which is suitable for practical clinical use.

Ethics and Data Availability

The Brain Tumor MRI dataset, publicly accessible on Kaggle, is completely anonymous. Given that no discernible patient information is provided, institutional review board approval was not necessary. All experiments were a secondary analysis of open-access data.

References

[1] Sachdeva, Jainy, Deepanshu Sharma, and Chirag Kamal Ahuja. "Comparative Analysis of Different Deep Convolutional Neural Network Architectures for Classification of Brain Tumor on Magnetic Resonance Images." Archives of Computational Methods in Engineering 31, no. 4 (2024): 1959-1978.

[2] Babu Vimala, Baiju, Saravanan Srinivasan, Sandeep Kumar Mathivanan, Mahalakshmi, Prabhu Jayagopal, and Gemmachis Teshite Dalu. "Detection and Classification of Brain Tumor Using Hybrid Deep Learning Models." Scientific reports 13, no. 1 (2023): 23029.

[3] Pedada, Kameswara Rao, Bhujanga Rao, Kiran Kumar Patro, Jaya Prakash Allam, Mona M. Jamjoom, and Nagwan Abdel Samee. "A Novel Approach for Brain Tumour Detection Using Deep Learning Based Technique." Biomedical Signal Processing and Control 82 (2023): 104549.

[4] İncir, Ramazan, and Ferhat Bozkurt. "Improving Brain Tumor Classification with Combined Convolutional Neural Networks and Transfer Learning." Knowledge-Based Systems 299 (2024): 111981.

[5] Amarnath, Amarnath, Ali Al Bataineh, and Jeremy A. Hansen. "Transfer-Learning Approach for Enhanced Brain Tumor Classification in MRI Imaging." BioMedInformatics 4, no. 3 (2024): 1745-1756.

[6] Hassan, Elaheh, and Hamid Ghadiri. "Advancing brain tumor classification: A Robust Framework Using EfficientNetV2 Transfer Learning and Statistical Analysis." Computers in Biology and Medicine 185 (2025): 109542.

[7] Waskita, A. A., Julfa Muhammad Amda, Dwi Seno Kuncoro Sihono, and Heru Prasetio. "EfficientNetV2 based for MRI Brain Tumor Image Classification." In 2023 International Conference on Computer, Control, Informatics and its Applications (IC3INA), IEEE, 2023, 171-176.

[8] Zulfiqar, Fatima, Usama Ijaz Bajwa, and Yasar Mehmood. "Multi-Class Classification of Brain Tumor Types from MR Images Using EfficientNets." Biomedical Signal Processing and Control 84 (2023): 104777.

[9] Ghosh, Arpita, Badal Soni, and Ujwala Baruah. "Transfer Learning-Based Deep Feature Extraction Framework Using Fine-Tuned Efficientnet b7 for Multiclass Brain Tumor Classification." Arabian Journal for Science and Engineering 49, no. 9 (2024): 12027-12048.

[10] Anwar, Raja Waseem, Mohammad Abrar, and Faizan Ullah. "Transformative Transfer Learning for MRI Brain Tumor Precision: Innovative Insights." IEEe Access 13 (2025): 31749-31761.

[11] R. Preetha, M. J. P. Priyadarsini and J. S. Nisha, "Hybrid 3B Net and EfficientNetB2 Model for Multi-Class Brain Tumor Classification," in IEEE Access, vol. 13, 2025, 63465-63485.

[12] Pacal, Ishak, Omer Celik, Bilal Bayram, and Antonio Cunha. "Enhancing EfficientNetv2 with Global and Efficient Channel Attention Mechanisms for Accurate MRI-Based Brain Tumor Classification." Cluster Computing 27, no. 8 (2024): 11187-11212.

[13] Ferdousi, Sanjida, Utsha Das, and Tanvir Mahtab Zihan. "An Explainable Multi-Class Brain Tumor Classifier Using EfficientNetV2 with Hybrid Attention Mechanisms." In 2025 IEEE 2nd International Conference on Computing, Applications and Systems (COMPAS), IEEE, 2025, 1-6.

[14] Panigrahi, Soumyarashmi, Dibya Ranjan Das Adhikary, and Binod Kumar Pattanayak. "Hybrid Transfer Learning and Self-Attention Framework for Robust MRI-Based Brain Tumor Classification." Scientific Reports 15, no. 1 (2025): 21343.

[15] Celik, Fatih, Kemal Celik, and Ayse Celik. "Enhancing Brain Tumor Classification Through Ensemble Attention Mechanism." Scientific Reports 14, no. 1 (2024): 22260.

[16] Rai, Hari Mohan, Joon Yoo, and Serhii Dashkevych. "Two-headed UNetEfficientNets for Parallel Execution of Segmentation and Classification of Brain Tumors: Incorporating Postprocessing Techniques with Connected Component Labelling." Journal of Cancer Research and Clinical Oncology 150, no. 4 (2024): 220.

[17] P. Kumar Tiwary, P. Johri, A. Katiyar and M. K. Chhipa, "Deep Learning-Based MRI Brain Tumor Segmentation with EfficientNet-Enhanced UNet," in IEEE Access, vol. 13, 2025, 54920-54937.

[18] Preetha R, Jasmine Pemeena Priyadarsini M and Nisha J S, “Brain Tumor Segmentation Using Multi-Scale Attention U-Net with EfficientNetB4 Encoder for Enhanced MRI Analysis,” Sci. Rep., vol. 15, no. 1, Mar. 2025, 9914.

[19] Zarenia, Erfan, Amirhossein Akhlaghi Far, and Khosro Rezaee. "Automated Multi-Class MRI Brain Tumor Classification and Segmentation Using Deformable Attention and Saliency Mapping." Scientific Reports 15, no. 1 (2025): 8114.

[20] Priyadarshini, Pallavi, Priyadarshi Kanungo, and Tejaswini Kar. "Multigrade Brain Tumor Classification in MRI Images Using Fine Tuned EfficientNet." e-Prime-Advances in Electrical Engineering, Electronics and Energy 8 (2024): 100498.

[21] Irfani, Muhammad Daffa, and Untari Novia Wisesty. "Brain Tumor Classification Using EfficientNet-Based CNN Architecture in MRI Images." In 2025 International Conference on Data Science and Its Applications (ICoDSA), IEEE, 2025, 200-205.

[22] Isunuri, B. Venkateswarlu, and Jagadeesh Kakarla. "EfficientNet and Multi-Path Convolution with Multi-Head Attention Network for Brain Tumor Grade Classification." Computers and Electrical Engineering 108 (2023): 108700.

[23] Ishfaq, Qurat Ul Ain, Rozi Bibi, Abid Ali, Faisal Jamil, Yousaf Saeed, Rana Othman Alnashwan, Samia Allaoua Chelloug, and Mohammed Saleh Ali Muthanna. "Automatic Smart Brain Tumor Classification and Prediction System Using Deep Learning." Scientific reports 15, no. 1 (2025): 14876.

[24] Woo, Sanghyun, Jongchan Park, Joon-Young Lee, and In So Kweon. "Cbam: Convolutional Block Attention Module." In Proceedings of the European conference on computer vision (ECCV), 2018, 3-19.

[25] M. Nickparvar, “Brain Tumor MRI Dataset,” Kaggle, 2023. Available: https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset