

Article

Efficient Uncertainty-Aware Dual-Attention Network for Brain Tumor Detection

Sitara Afzal and Jong Ha Lee *

Department of Biomedical Engineering, Keimyung University, Daegu 42601, Republic of Korea

* Correspondence: segeberg@kmu.ac.kr

Abstract

Brain tumor detection from magnetic resonance imaging (MRI) is fundamental to computer-aided diagnosis, yet automated models must remain robust to heterogeneous imaging conditions. Despite strong recent progress, many deep learning and transformer-based approaches primarily optimize performance accuracy without explicitly improving feature selectivity and spatial localization, and they typically produce deterministic output without uncertainty estimates, which limits reliability. To overcome these limitations, we introduce UA-EffNet-DA, an uncertainty-aware EfficientNet framework that addresses these limitations through three complementary components. First, EfficientNet-B4 serves as an efficient backbone for discriminative feature extraction. Second, lightweight dual attention modules, comprising channel and spatial attention in parallel, are applied to the model to emphasize what and where discriminative features to focus within MRI slices. Third, Monte Carlo dropout is employed during inference to quantify predictive uncertainty and enable confidence-aware decision. Experiments on two public benchmarks demonstrate strong performance, yielding accuracies of 98.73% on the Figshare dataset and 99.23% on the Kaggle dataset. In addition, explainable AI analysis using Gradient-weighted Class Activation Mapping (Grad-CAM) further indicates that the proposed model concentrates on diagnostically relevant tumor regions rather than background structures, supporting transparent decision-making. Ablation studies confirm the complementary contribution of dual attention refinement and uncertainty-aware inference. Overall, the proposed UA-EffNet-DA framework offers an efficient and interpretable approach for brain tumor detection that supports more reliable clinical decision support through uncertainty-aware predictions.

Keywords: brain tumor; magnetic resonance imaging; deep learning; EfficientNet; attention mechanism; uncertainty-aware inference; explainable AI

MSC: 68T07; 68U10



Academic Editor: Marjan Mernik

Received: 16 March 2026

Revised: 15 April 2026

Accepted: 18 April 2026

Published: 23 April 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Magnetic Resonance Imaging (MRI) is a major tool for the detection and evaluation of brain tumors because of its superior soft-tissue contrast and its ability to reveal structural abnormalities in neural tissue [1]. Accurate identification of tumor type is essential for treatment planning, prognosis assessment, and surgical decision-making [2]. However, manual analysis of MRI scans is time-consuming and often affected by inter-observer variability, particularly when tumor boundaries are unclear or imaging conditions vary [3]. These challenges have motivated the development of computer-aided diagnosis (CAD) systems that support automated and consistent brain tumor classification [4–6].

In recent years, deep learning approaches, especially convolutional neural networks (CNNs), have demonstrated significant performance in medical image analysis [7,8]. In addition, transfer learning with pretrained architectures such as ResNet [9], DenseNet [10], and EfficientNet [11] has further improved classification performance by reusing hierarchical features learned from large-scale datasets [12–14]. However, brain tumor MRI images often exhibit substantial intra-class variability, overlapping intensity patterns, and complex anatomical backgrounds, which can cause conventional CNN models to learn non-discriminative or irrelevant features [15,16]. As a result, predictions may be influenced by surrounding tissues. A key limitation of many existing CNN-based brain tumor classification methods is that they treat all channel responses and spatial regions in a largely uniform manner during feature extraction.

Standard convolutional operations do not explicitly distinguish tumor-relevant responses from non-relevant ones, leading to diluted feature representations and reduced interpretability [17]. To improve performance, several recent studies have adopted multi-branch or ensemble architectures; however, these approaches often increase computational complexity and reduce suitability for real-time or deployment-oriented settings [18,19]. Attention mechanisms, in contrast, have emerged as an effective way to improve feature quality by guiding neural networks toward more informative representations. Channel Attention (CA) emphasizes discriminative feature responses, whereas Spatial Attention (SA) highlights anatomically relevant regions. Although attention-based designs have shown promising results in medical imaging, many existing approaches either rely on computationally heavy multi-network structures or incorporate only a single type of attention [20].

In a recent study, Zarenia et al. [16] proposed a deformable-attention and saliency-based framework for automated multi-class brain tumor MRI analysis, demonstrating the value of attention-guided feature learning. However, their framework is oriented toward a broader classification-plus-segmentation pipeline and may not directly support lightweight single-backbone classification for deployment-focused settings. Similarly, Ishaq et al. [21] presented an EfficientNet-based brain tumor classification framework with Grad-CAM-based interpretability, demonstrating the practicality of EfficientNet for MRI analysis. Nevertheless, their model does not explicitly combine channel-wise and spatial feature refinement within the backbone feature stream.

To address these identified gaps, we propose UA-EffNet-DA, an Uncertainty-Aware EfficientNet with Dual Attention architecture that introduces three key methodological innovations built upon an EfficientNet-B4 backbone for brain tumor MRI detection. EfficientNet serves as a strong and computationally efficient feature extractor through compound scaling of network depth, width, and input resolution. On top of this backbone, the framework incorporates parallel dual attention mechanisms, consisting of Channel and Spatial Attention, to explicitly model both which features are important and where the model should focus within MRI slices. The Channel Attention module recalibrates feature channels to emphasize tumor-related representations, whereas the Spatial Attention module guides the network toward discriminative tumor regions. This design improves feature selectivity without introducing excessive architectural complexity, while maintaining a single-backbone structure that is more practical for deployment than multi-branch or transformer-based alternatives.

In addition, the proposed framework incorporates Monte Carlo (MC) dropout to estimate predictive uncertainty, which is important in clinically sensitive decision-support settings. By keeping dropout active at inference time and performing multiple stochastic forward passes ($T = 20$), the model produces a distribution of outputs whose variance reflects epistemic uncertainty. This allows the system to flag low-confidence predictions

where the model exhibits high predictive variance, signaling cases that may benefit from expert radiological review. Unlike deterministic approaches that provide only point predictions, UA-EffNet-DA reports predictive uncertainty alongside classification results, supporting more informed and confidence-aware interpretation.

Following are the key contributions of this study:

1. We develop an uncertainty-aware EfficientNet-B4-based framework for brain tumor MRI detection by integrating channel and spatial attention-based feature refinement within a single-backbone architecture.
2. We enhance feature selectivity and spatial focus through complementary channel and spatial attention mechanisms, enabling the model to emphasize tumor-relevant representations and anatomically meaningful regions.
3. We incorporate Monte Carlo dropout with 20 stochastic forward passes at inference time to estimate predictive uncertainty and identify low-confidence cases that may require additional expert review.
4. We employ Grad-CAM visualization to examine whether the proposed framework concentrates on diagnostically relevant tumor regions rather than irrelevant background structures, thereby improving interpretability.
5. We validate the proposed framework through quantitative evaluation and ablation analysis on publicly available brain tumor MRI datasets, demonstrating the contribution of attention-based refinement and uncertainty-aware inference.

The remainder of this paper is as follows: Section 2 reviews the related literature. Section 3 presents the proposed methodology. Section 4 presents the results of our experiments. Section 5 analyzes and contrasts the implications of our findings. Section 6 concludes the paper.

2. Literature Review

Brain tumor classification from MRI has become an important research area, with deep learning methods showing strong performance improvements over traditional approaches. Recent progress has been driven mainly by three methodological directions: transfer learning with efficient CNN backbones, attention-based feature refinement, and transformer-based architectures. In this section, we review these categories and highlight the remaining challenges that motivate the proposed framework.

2.1. Transfer Learning and EfficientNet-Based Approaches

EfficientNet architectures have attracted considerable interest for brain tumor classification because their compound scaling strategy provides an effective balance between network depth, width, and resolution. Islam et al. [22] introduced BrainNet, a comprehensive framework based on the EfficientNet family together with advanced preprocessing and augmentation techniques on 3064 T1-weighted contrast-enhanced MRI images. Their EfficientNetB3 model achieved strong performance, although the study also noted limitations related to computational complexity and dependence on high-quality preprocessing. Similarly, Zulfiqar et al. [23] employed EfficientNets for multi-class brain tumor classification from MR images and demonstrated their effectiveness across glioma, meningioma, and pituitary tumor categories. However, their approach remained sensitive to class imbalance and required extensive augmentation to achieve optimal performance.

A recent study by Ishaq et al. [24] presented an improved EfficientNet-based framework for multi-grade brain tumor detection and incorporated Grad-CAM to improve interpretability. Their results further confirmed the practicality of EfficientNet for MRI analysis, although the framework still faced challenges in handling heterogeneous tumor morphology and required substantial computational resources during inference. Likewise,

Alnashwan and Chelloug [25] combined custom CNN models with EfficientNet-B4 and reported strong classification performance across ten tumor categories, but their system required specialized hardware and showed reduced accuracy on smaller tumor regions. Vimala et al. [26] also explored transfer learning with EfficientNet B0–B4 models on the CE-MRI Figshare dataset and demonstrated the usefulness of Grad-CAM for interpretability. However, their approach showed overfitting tendencies and limited generalization beyond the benchmark dataset. Taken together, these studies confirm the effectiveness of EfficientNet-based transfer learning for brain tumor MRI classification. However, they also show that strong benchmark performance alone does not fully address issues related to feature selectivity, robustness, computational efficiency, and generalization.

2.2. Attention Mechanisms for Enhanced Feature Extraction

Attention mechanisms have emerged as an effective way to guide neural networks toward diagnostically relevant features and spatial regions. Nisha et al. [27] proposed a multi-scale attention U-Net with an EfficientNetB4 encoder for brain tumor segmentation. Their framework improved multi-scale feature extraction by capturing tumor boundaries at different spatial resolutions, although it required extensive hyperparameter tuning and showed reduced performance on irregular tumor boundaries. Nahiduzzaman et al. [28] combined VGG16 with attention mechanisms and optimized hyperparameters for four-class brain tumor classification. Their model achieved high test accuracy, but the framework remained computationally demanding and showed reduced robustness for atypical tumor presentations. Ahmed et al. [29] introduced a hybrid ViT-GRU model with explainable AI support for brain tumor detection. Although the model demonstrated strong performance, it also depended on sufficient annotated data and introduced higher training complexity.

Other studies have explored attention-driven refinement in more specialized forms. Rasheed et al. [30] showed that integrating attention mechanisms into CNNs can improve feature representation in MRI-based brain tumor classification, while also noting the additional computational overhead and risk of overfitting. Zarenia et al. [16] proposed deformable attention and saliency mapping for automated multi-class MRI brain tumor classification and segmentation, demonstrating the ability to capture complex tumor patterns. However, their framework required substantial training data and showed reduced performance on smaller tumor regions. Gasmi et al. [31] used co-attention mechanisms in an ensemble of MobileNetV3 and EfficientNetB7 networks and achieved strong performance on both Figshare and BraTS 2019 datasets. Nevertheless, the multi-model design increased computational complexity and reduced suitability for resource-constrained deployment. Abdusalomov et al. [32] integrated CBAM into YOLOv7 for brain tumor detection and improved focus on salient regions, although the model still faced challenges related to computational efficiency and false positive responses in anatomically complex regions.

Overall, attention-based approaches improve lesion-focused representation learning and interpretability. However, many of them either rely on computationally heavy designs or introduce additional architectural complexity, which can limit their practical deployment.

2.3. Transformer-Based Approaches

Vision Transformers (ViTs) have further advanced brain tumor MRI analysis by modeling long-range dependencies and global contextual information. Chandraprabha et al. [33] proposed an end-to-end ViT-CNN architecture and reported strong classification performance across multiple tumor classes. However, the model required significant computational resources and extensive annotated data. Reddy et al. [34] compared several fine-tuned Vision Transformer variants for multi-class brain tumor classification and showed that transformer-based models can achieve high accuracy. At the same time, they also

reported high parameter counts, sensitivity to hyperparameter settings, and training instability on limited datasets. Krishnan et al. [35] introduced a rotation-invariant Vision Transformer and improved robustness to orientation variation, but the method added computational complexity and required careful tuning of augmentation parameters.

Recent work [36] on hierarchical multi-scale vision transformer models further demonstrated that transformer-based designs can improve brain tumor classification while reducing training time through multi-resolution patch embeddings. Nevertheless, despite this progress, important limitations remain across the broader literature. Many studies report high accuracy on benchmark datasets but provide limited evidence of generalization beyond those settings [17,37]. In addition, improved predictive performance is often achieved at the cost of increased model complexity, higher computational burden, or reduced deployment practicality. These observations motivate the need for a framework that maintains a stronger balance between efficiency, feature selectivity, interpretability, and prediction reliability. In response, our work adopts a single-backbone EfficientNet-based design with lightweight channel and spatial attention-based refinement, while also incorporating uncertainty-aware inference to indicate low-confidence predictions.

3. Proposed Methodology

This section presents the comprehensive methodological framework employed for developing our Uncertainty-Aware EfficientNet with Dual Attention (UA-EffNet-DA) architecture for brain tumor classification. We systematically describe the dataset preparation, the proposed architectural innovations, the mathematical formulations underlying our dual attention mechanisms, the uncertainty quantification strategy, and the rigorous experimental protocol designed to validate each component's contribution. Figure 1 represents the overall framework of the proposed study.

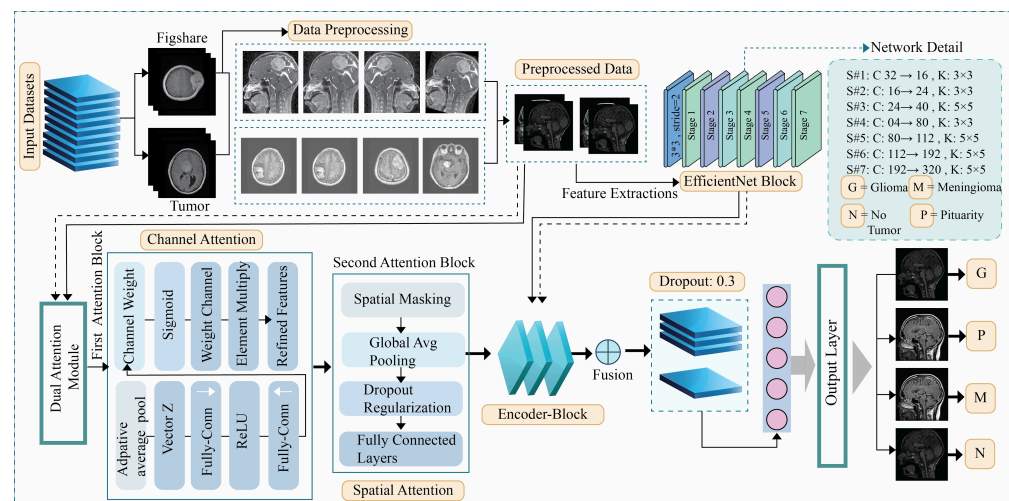


Figure 1. Overall workflow of the proposed UA-EffNet-DA framework for brain tumor detection.

3.1. Proposed Architecture: UA-EffNet-DA

Our proposed Uncertainty-Aware EfficientNet with Dual Attention (UA-EffNet-DA) architecture introduces three primary innovations over conventional transfer learning approaches: (1) the integration of complementary channel and SA mechanisms applied after feature extraction to enhance discriminative representation learning, (2) the incorporation of uncertainty quantification through Monte Carlo dropout to provide calibrated confidence estimates alongside predictions, and (3) architectural modifications enabling grayscale medical image processing with preserved transfer learning benefits from Im-

geNet pretraining. Figure 2 presents the comprehensive architectural diagram illustrating the sequential flow of feature transformations through the network.

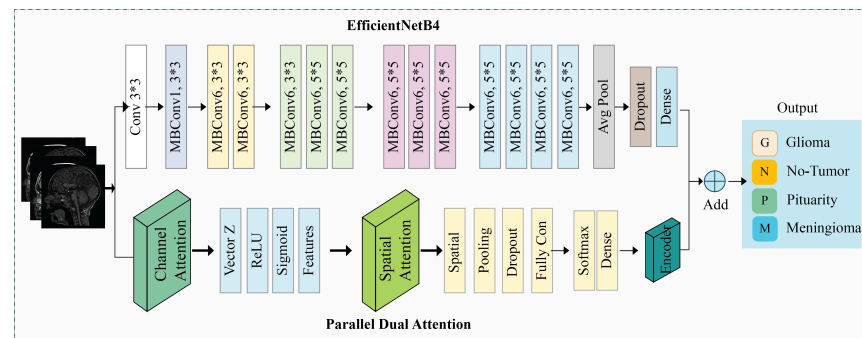


Figure 2. Graphical representation of the proposed framework showing the end-to-end pipeline.

3.1.1. EfficientNet-B4 Backbone with Grayscale Adaptation

The proposed architecture employs EfficientNet-B4 as the feature extraction backbone because of its strong parameter efficiency achieved through compound scaling of network depth, width, and input resolution. EfficientNet-B4 comprises approximately 19 million parameters organized into mobile inverted bottleneck convolution (MBConv) blocks, providing a favorable balance between representational capacity and computational efficiency for medical image analysis.

To accommodate grayscale MRI inputs while retaining the benefits of ImageNet pretraining, the first convolutional layer was modified to accept a single-channel input. Specifically, the input layer was changed from $\text{Conv2d}(3, 48, \text{kernel_size}=3, \text{stride}=2, \text{padding}=1)$ to $\text{Conv2d}(1, 48, \text{kernel_size}=3, \text{stride}=2, \text{padding}=1)$. The pretrained RGB weights were averaged across the channel dimension to initialize the grayscale kernel, as given in Equation (1).

$$W_{\text{gray}} = \frac{1}{3} \sum_{i \in \{R, G, B\}} W_{\text{RGB}, i} \quad (1)$$

This weight-averaging strategy preserves the statistical characteristics learned during ImageNet pretraining while enabling direct adaptation to grayscale inputs without extensive retraining of early-layer filters. The remaining convolutional blocks retain their pretrained parameters, allowing the network to exploit hierarchical feature representations learned from large-scale natural-image datasets.

The EfficientNet-B4 backbone progressively downsamples the spatial resolution while expanding the channel dimension, producing a final feature tensor $\mathbf{F} \in \mathbb{R}^{C \times H \times W}$ of size $1792 \times 7 \times 7$, where $C = 1792$ represents the number of channels, and $H = W = 7$ corresponds to the reduced spatial resolution.

3.1.2. Channel-Attention Mechanism

The channel attention module adaptively refines channel-wise responses by learning the dependencies between feature channels. This design allows the network to emphasize informative channels while suppressing less discriminative responses, thereby addressing a common limitation of conventional CNNs that treat all channels uniformly. Algorithm 1 represents the step-by-step formulation of the Channel Attention module, including its input, output, and feature recalibration process.

Algorithm 1: Channel Attention Module

Input	Feature tensor $\mathbf{F} \in \mathbb{R}^{C \times H \times W}$, reduction ratio r
Output	Channel-refined tensor $\mathbf{F}_{CA} \in \mathbb{R}^{C \times H \times W}$
1	Compute global channel descriptors: $z_c = \frac{1}{HW} \sum_{i=1}^H \sum_{j=1}^W F_c(i, j)$, for $c = 1, \dots, C$
2	Form channel descriptor vector: $\mathbf{z} = [z_1, z_2, \dots, z_C]^\top$
3	Apply bottleneck transformation: $\mathbf{h} = \text{ReLU}(\mathbf{W}_1 \mathbf{z})$
4	Compute channel attention weights: $\mathbf{s} = \sigma(\mathbf{W}_2 \mathbf{h})$
5	Broadcast $\mathbf{s}$ over the spatial dimensions
6	Recalibrate feature tensor: $\mathbf{F}_{CA} = \mathbf{s} \odot \mathbf{F}$

The channel attention mechanism follows a squeeze-and-excitation strategy consisting of global spatial pooling, nonlinear channel interaction modeling, and channel-wise rescaling. First, global average pooling is used to compress each feature map into a scalar descriptor:

$$z_c = \frac{1}{HW} \sum_{i=1}^H \sum_{j=1}^W F_c(i, j) \quad (2)$$

Here z_c denotes the global descriptor of the c -th channel and $F_c(i, j)$ represents the activation at spatial location (i, j) in the corresponding feature map.

Next, the channel descriptors are passed through a bottleneck multilayer perceptron with reduction ratio $r = 16$:

$$\mathbf{s} = \sigma(\mathbf{W}_2 \text{ReLU}(\mathbf{W}_1 \mathbf{z})) \quad (3)$$

Here $\mathbf{W}_1 \in \mathbb{R}^{\frac{C}{r} \times C}$ and $\mathbf{W}_2 \in \mathbb{R}^{C \times \frac{C}{r}}$ are learnable weight matrices, $\text{ReLU}(\cdot)$ denotes the rectified linear unit, and $\sigma(\cdot)$ is the sigmoid function that produces channel attention weights $\mathbf{s} \in [0, 1]^C$.

Finally, the learned attention weights are broadcast across the spatial dimensions and applied to the original feature tensor:

$$\mathbf{F}_{CA} = \mathbf{s} \odot \mathbf{F} \quad (4)$$

Here $\odot$ denotes element-wise multiplication with broadcasting across all spatial locations.

3.1.3. Spatial Attention Mechanism

Complementing channel attention, the spatial attention (SA) module highlights diagnostically relevant spatial regions while suppressing background responses. This is particularly important in medical imaging because tumor regions usually occupy only a limited portion of the image, whereas surrounding tissues often contribute little to the final decision.

The SA module first computes two complementary channel-compressed descriptors from the channel-refined tensor $\mathbf{F}_{CA}$: an average-pooled map and a max-pooled map along the channel dimension:

$$F_{\text{avg}}(i, j) = \frac{1}{C} \sum_{c=1}^C F_{CA,c}(i, j) \quad (5)$$

$$F_{\text{max}}(i, j) = \max_{c \in \{1, \dots, C\}} F_{CA,c}(i, j) \quad (6)$$

Here F_{avg} captures average channel activity and F_{max} highlights the most salient activation at each spatial location.

These two descriptors are concatenated to form a two-channel spatial representation:

$$\mathbf{F}_{\text{concat}} = [F_{\text{avg}}; F_{\text{max}}] \in \mathbb{R}^{2 \times H \times W} \quad (7)$$

A convolution with a 7×7 kernel is then applied to generate the spatial attention map:

$$\mathbf{M}_s = \sigma(f^{7 \times 7}(\mathbf{F}_{\text{concat}})) \quad (8)$$

Here $f^{7 \times 7}(\cdot)$ denotes the convolution operation and $\sigma(\cdot)$ constrains the resulting map to $\mathbf{M}_s \in [0, 1]^{H \times W}$.

The spatial attention map is subsequently broadcast across channels and applied to the channel-attended feature tensor:

$$\mathbf{F}_{\text{SA}} = \mathbf{M}_s \odot \mathbf{F}_{\text{CA}} \quad (9)$$

This consecutive use of channel attention and spatial attention creates a hierarchical refinement process in which channel attention identifies what features are important, whereas spatial attention determines where the model should focus. Together, they improve discriminative representation learning for brain tumor classification. Algorithm 2 represents the step-by-step formulation of the Spatial Attention module, including its input, output, and feature recalibration process.

Algorithm 2: Spatial Attention Module

Input	Channel-refined tensor $\mathbf{F}_{\text{CA}} \in \mathbb{R}^{C \times H \times W}$
Output	Spatially refined tensor $\mathbf{F}_{\text{SA}} \in \mathbb{R}^{C \times H \times W}$
1	Compute average-pooled spatial map: $F_{\text{avg}}(i, j) = \frac{1}{C} \sum_{c=1}^C F_{\text{CA},c}(i, j)$
2	Compute max-pooled spatial map: $F_{\text{max}}(i, j) = \max_{c \in \{1, \dots, C\}} F_{\text{CA},c}(i, j)$
3	Concatenate pooled maps: $\mathbf{F}_{\text{concat}} = [F_{\text{avg}}; F_{\text{max}}]$
4	Compute spatial attention map: $\mathbf{M}_s = \sigma(f^{7 \times 7}(\mathbf{F}_{\text{concat}}))$
5	Broadcast $\mathbf{M}_s$ across channels
6	Refine feature tensor: $\mathbf{F}_{\text{SA}} = \mathbf{M}_s \odot \mathbf{F}_{\text{CA}}$

3.1.4. Uncertainty Quantification Through Monte Carlo Dropout

A limitation of deterministic neural network models is that they provide point predictions without explicitly quantifying predictive uncertainty. This is particularly problematic in medical diagnosis, where confidence calibration is important for safe decision support. To address this limitation, Monte Carlo (MC) dropout was incorporated as an approximate Bayesian inference technique.

A dropout layer with probability $p = 0.4$ was placed before the final classification layer and retained during both training and inference. At test time, the model performs $T = 20$ stochastic forward passes using different dropout masks, thereby generating a distribution of predictions:

$$\hat{\mathbf{y}}^{(t)} = \text{softmax}(f_{\theta^{(t)}}(\mathbf{x})), \quad t = 1, \dots, T \quad (10)$$

Here $\theta^{(t)}$ denotes the effective model parameters under the t -th dropout realization, $\mathbf{x}$ is the input image, and $\hat{\mathbf{y}}^{(t)}$ is the predicted class-probability vector.

The predictive mean is computed as:

$$\bar{\mathbf{y}} = \frac{1}{T} \sum_{t=1}^T \hat{\mathbf{y}}^{(t)} \quad (11)$$

Predictive uncertainty is then estimated from the variance across stochastic predictions:

$$\sigma^2 = \frac{1}{T} \sum_{t=1}^T \left(\hat{\mathbf{y}}^{(t)} - \bar{\mathbf{y}} \right)^2 \quad (12)$$

Here the square is applied element-wise. High predictive variance indicates elevated epistemic uncertainty, which is commonly observed in ambiguous samples or cases that are underrepresented in the training distribution. In contrast, low variance combined with high confidence suggests a more reliable prediction. This uncertainty estimation mechanism enables the identification of low-confidence cases that may require additional expert review. Algorithm 3 represents the step-by-step formulation of the dropout module, including its input, output, and feature recalibration process.

Algorithm 3: MC-Dropout-Based Uncertainty-Aware Inference

Input	Input image $\mathbf{x}$, trained model $f_{\theta}(\cdot)$ with dropout active at inference, number of stochastic passes T
Output	Predictive mean $\bar{\mathbf{y}}$ and predictive variance σ^2
1	For each stochastic pass $t = 1, \dots, T$, sample a dropout mask and compute $\hat{\mathbf{y}}^{(t)} = \text{softmax}(f_{\theta^{(t)}}(\mathbf{x}))$
2	Compute predictive mean: $\bar{\mathbf{y}} = \frac{1}{T} \sum_{t=1}^T \hat{\mathbf{y}}^{(t)}$
3	Compute predictive variance: $\sigma^2 = \frac{1}{T} \sum_{t=1}^T \left(\hat{\mathbf{y}}^{(t)} - \bar{\mathbf{y}} \right)^2$
4	Return $\bar{\mathbf{y}}$ and σ^2

3.2. Evaluation Metrics

Model performance was evaluated using multiple evaluation metrics that capture different aspects of multi-class detection quality and are widely adopted in medical image analysis. Overall accuracy was used to measure the proportion of correctly classified samples:

$$\text{AccuracyMeasure} = \frac{T_P + T_N}{T_P + T_N + F_P + F_N} \quad (13)$$

Class-wise precision and recall were computed to characterize false-positive and false-negative behavior:

$$\text{PrecisionMeasure} = \frac{T_P}{T_P + F_P} \quad (14)$$

$$\text{Recall} = \frac{T_P}{T_P + F_N} \quad (15)$$

The F1-score was used as the harmonic mean of precision and recall:

$$\text{F1-Measure} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (16)$$

Here, T_P are true positive cases, F_P are false positive cases, F_N are false negative cases and T_N are true negative cases. For multi-class evaluation, AUC-ROC was computed in a one-vs-rest manner for each class, and the macro-averaged AUC was reported to summarize the overall discriminative performance across classes. Confusion matrices were created as well to visualize the class-wise predictions and identify systematic misclassification

trends among glioma-tumor, meningioma-tumor, pituitary-tumor, and no-tumor categories, where applicable.

Explainability via Gradient-Weighted Class Activation Mapping

A Gradient-Weighted Class Activated Map (Grad-CAM) was used to increase clinical interpretability and ensure that UA-EffNet-DA targets diagnostically important regions. It calculates the gradient of the target class scores with respect to the feature maps of the final convolutional layer and uses these gradients to construct class-specific significance weights.

$$\alpha_k^c = \frac{1}{HW} \sum_{i=1}^H \sum_{j=1}^W \frac{\partial y^c}{\partial A_k^{ij}} \quad (17)$$

where α_k^c denotes the importance weight of the k -th feature, and A_k^{ij} is the activation at the spatial location (i, j) of features A_k . The class-discriminative localization map is then obtained by a weighted combination of feature maps followed by the $\text{ReLU}(\cdot)$ function to retain only positive contributions:

$$L_{\text{Grad-CAM}}^c = \text{ReLU}\left(\sum_k \alpha_k^c A_k\right) \quad (18)$$

Grad-CAM visualizations were used for qualitative analysis to confirm that the proposed model emphasizes tumor-relevant regions rather than background structures or imaging artifacts, thereby improving the transparency and interpretability of the model's decision-making process.

4. Experimentation and Result

This section describes the experimental setup adopted to evaluate the proposed UA-EffNet-DA framework and presents the corresponding performance results. Specifically, we detail the datasets used for training and evaluation, the preprocessing strategies applied to MRI slices, and the procedures followed to ensure consistent performance assessment across different imaging conditions.

4.1. Datasets

To evaluate the generalization capability and robustness of the proposed architecture, experiments were conducted on two publicly available brain tumor MRI benchmark datasets. These datasets include T1-weighted MRI scans covering three major tumor categories: glioma, meningioma, and pituitary tumors.

The primary dataset used in this study was obtained from the Figshare [38] repository and consists of 3064 axial CE-MRI slices belonging to glioma, meningioma, and pituitary tumor classes. The images were acquired from multiple patients under varying scanning protocols and institutional settings, thereby introducing realistic inter-patient variability and imaging heterogeneity representative of clinical deployment scenarios.

The second dataset used in this study was obtained from the Kaggle [39] brain tumor MRI dataset. In contrast to Figshare, the Kaggle dataset includes an additional non-tumor category, providing a broader classification setting that requires discrimination between tumor and no-tumor cases as well as among tumor classes. This dataset also exhibits variations in intensity distributions, spatial resolution, and acquisition artifacts compared to the Figshare dataset, thereby providing a more stringent evaluation of the model's ability to generalize beyond the training distribution. Table 1 provides the summary of both the datasets utilized in this study.

Table 1. Summary of datasets and preprocessing configurations used for brain tumor classification.

Parameter	Figshare Dataset	Kaggle Dataset
Imaging Modality	T1-weighted CE-MRI	T1-weighted CE-MRI
Clinical Categories	Tumor-only	Tumor and Healthy Controls
Number of Classes	3	4
Glioma Samples	1426	1621
Meningioma Samples	708	1645
Pituitary Samples	930	1757
No-Tumor Samples	–	2000
Total MRI Slices	3064	3264
Patient Variability	Multi-patient	Multi-patient
Acquisition Protocols	Multi-institutional	Multi-institutional
Scanner Heterogeneity	Present	Present
Input Representation	Grayscale (single-channel)	Grayscale (single-channel)
Resized Resolution	224×224	224×224
Interpolation Method	Bicubic	Bicubic
Normalization	$\mu = 0.5, \sigma = 0.5$	$\mu = 0.5, \sigma = 0.5$
Slice Selection Strategy	All axial slices retained	All axial slices retained
Annotation Level	Slice-level detection	Slice-level detection
Data Split (Train/Test)	80%/20%	80%/20%

4.2. Preprocessing

All MRI slices were converted to single-channel grayscale format to reduce computational redundancy and eliminate potential color-space inconsistencies introduced by scanner-specific acquisition parameters. The images were subsequently resized to 224×224 pixels using bicubic interpolation to match the input requirements of the EfficientNet-B4 backbone.

Intensity normalization was applied using zero-mean unit-variance standardization:

$$I_{\text{norm}} = \frac{I_{\text{raw}} - \mu}{\sigma} \quad (19)$$

where I_{raw} denotes the original pixel intensity values in the range $[0, 255]$, $\mu = 0.5$, and $\sigma = 0.5$. This normalization ensures a consistent input dynamic range, which facilitates stable gradient propagation during training.

To improve robustness and reduce overfitting, online data augmentation was applied during training using geometric and intensity-based transformations that mimic natural variations in MRI acquisition. The augmentation pipeline included random horizontal flipping (probability 0.5), random rotation ($\pm 15^\circ$), brightness and contrast adjustment ($\pm 20\%$), and Gaussian noise injection with standard deviation $\sigma = 0.02$.

All augmentation operations were applied stochastically during training and were disabled during validation and testing to ensure consistent evaluation conditions.

4.3. Results

This section describes the training protocol used to ensure stable convergence, mitigate overfitting, and enable reproducible comparison across baselines, ablations, and datasets. The same optimization settings were used for all experiments unless otherwise stated, and the corresponding learning dynamics and performance trends.

The proposed UA-EffNet-DA model was optimized using AdamW, which decouples weight decay from the gradient-based parameter update to improve generalization. The initial learning rate was set to $\eta = 1 \times 10^{-4}$, which provided stable convergence for our

training regime. Training was conducted for 20 epochs with a batch size of 16, selected to balance gradient stability and GPU memory constraints on an NVIDIA RTX 3060 (12 GB). The default AdamW momentum parameters were used ($\beta_1 = 0.9$, $\beta_2 = 0.999$), and the weight decay coefficient was set to 0.01. The proposed model was trained and evaluated separately on both datasets, rather than by merging them into a single training pool.

Across datasets, the architecture remained unchanged except for the final classifier, whose output dimension was configured to reflect the dataset-specific number of classes. To reduce overfitting on limited medical imaging data, multiple complementary regularization mechanisms were employed. First, dropout with probability $p = 0.4$ was applied before the final classification layer. This dropout serves as standard regularization during training and is retained at inference time for uncertainty estimation. Figure 3 presents the confusion matrix obtained on the Figshare dataset, illustrating the class-wise prediction performance of the proposed UA-EffNet-DA framework across glioma, meningioma, and pituitary tumor categories. The matrix demonstrates strong diagonal dominance, indicating high classification consistency with minimal inter-class confusion among tumor types.

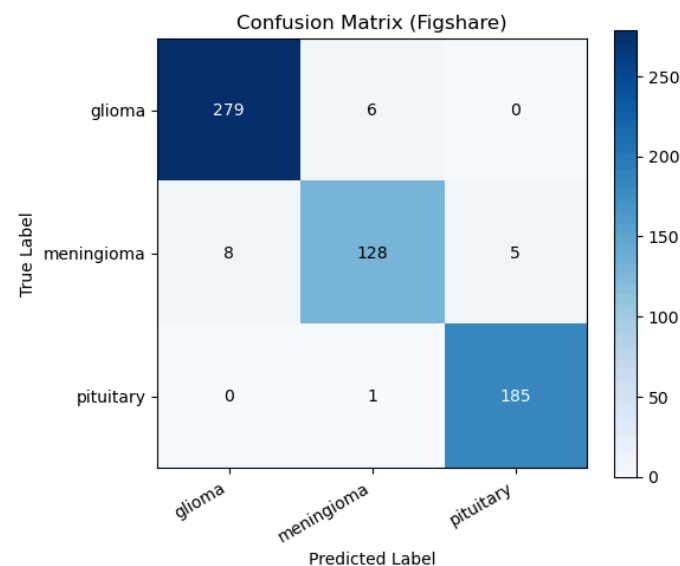


Figure 3. Confusion matrix of the proposed UA-EffNet-DA model on the Figshare brain tumor MRI dataset.

Similarly, Figure 4 depicts the confusion matrix for the Kaggle dataset, which includes an additional no-tumor class. The results show accurate discrimination between tumor and non-tumor samples, with limited misclassification between anatomically similar tumor categories such as glioma and meningioma. Figure 5 presents the ROC curves for the Figshare dataset, demonstrating the class-wise discriminative capability of the proposed UA-EffNet-DA framework across glioma, meningioma, and pituitary tumor categories. The curves remain close to the upper-left corner, indicating strong sensitivity–specificity trade-offs and high separability among tumor classes over varying decision thresholds. Similarly, Figure 6 depicts the ROC curves for the Kaggle dataset, which includes an additional no-tumor class. The results show consistently strong discrimination across both tumor and non-tumor categories, with high true positive rates achieved at low false positive rates, confirming the robustness of the proposed model under multi-class classification settings.

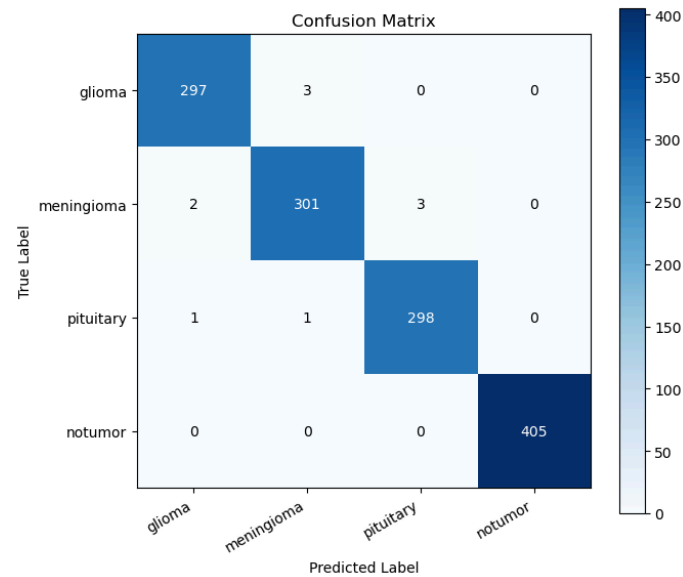


Figure 4. Confusion matrix of the proposed UA-EffNet-DA model on the Kaggle brain tumor MRI dataset.

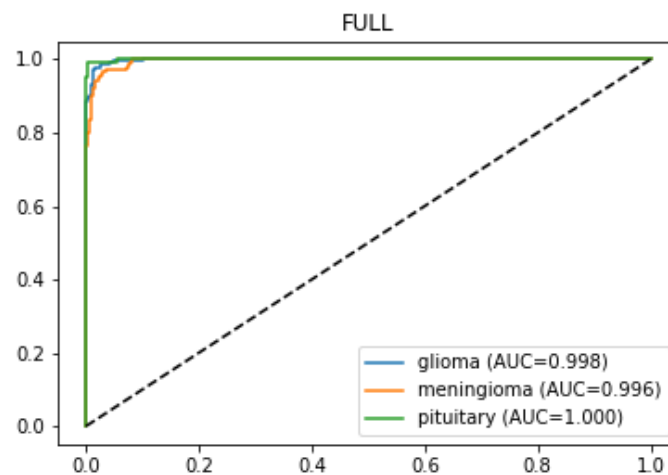


Figure 5. ROC curves of the proposed UA-EffNet-DA model on the Figshare brain tumor MRI dataset.

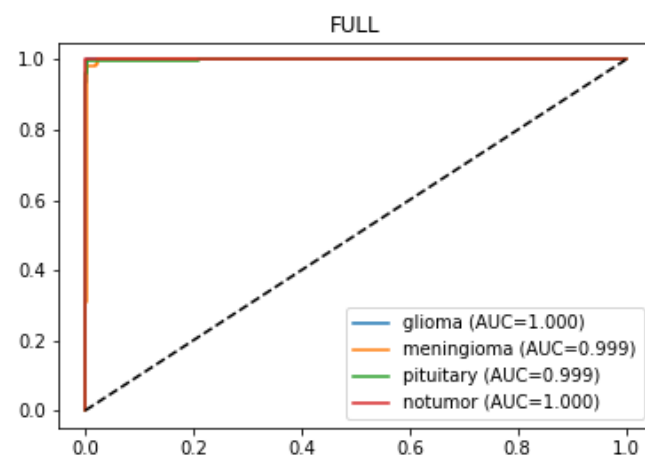


Figure 6. ROC curves of the proposed UA-EffNet-DA model on the Kaggle brain tumor MRI dataset.

Figure 7 shows representative qualitative prediction results of the proposed UA-EffNet-DA model on the Figshare brain tumor MRI dataset. Each sample is annotated with the predicted tumor class and the corresponding confidence score, enabling visual assessment of model behavior. The displayed examples indicate robust discrimination

across glioma, meningioma, and pituitary tumor categories under varying anatomical appearances and tumor morphologies. Similarly, Figure 8 shows representative qualitative prediction results of the proposed UA-EffNet-DA model on the Kaggle brain tumor MRI dataset. The grid includes samples from glioma, meningioma, pituitary, and no-tumor classes, together with the predicted labels and confidence scores. The qualitative results demonstrate reliable tumor and non-tumor discrimination across diverse MRI appearances and imaging conditions.

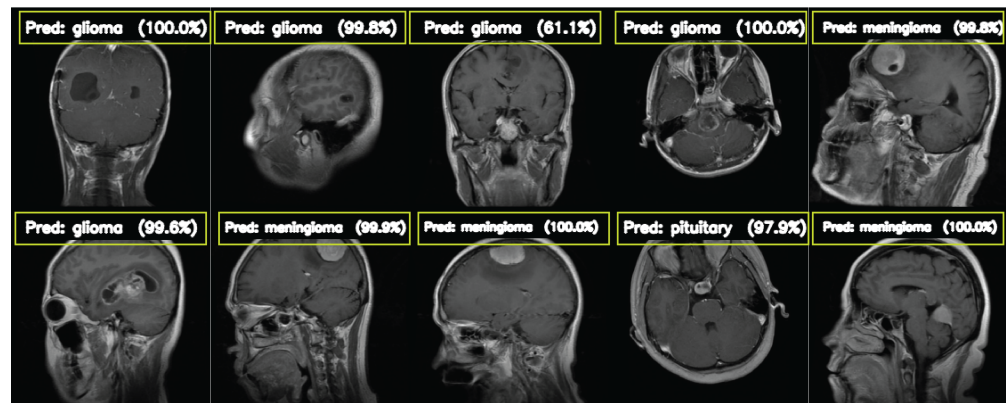


Figure 7. Qualitative prediction results of UA-EffNet-DA on the Figshare brain tumor dataset.

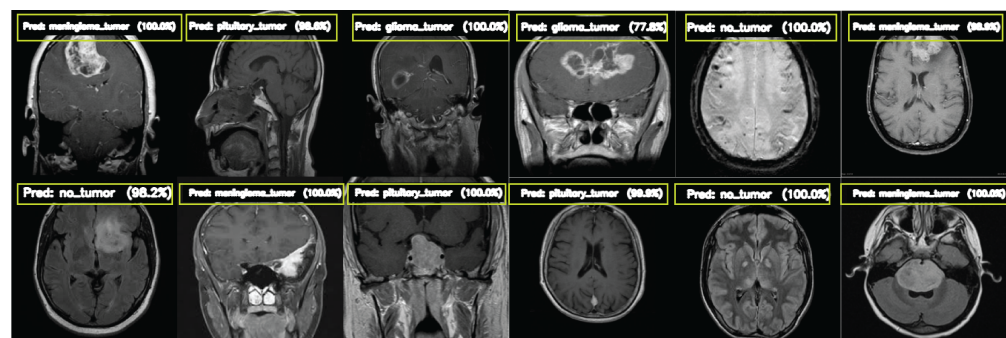


Figure 8. Qualitative prediction results of UA-EffNet-DA on the Kaggle brain tumor dataset.

On the Figshare dataset, the proposed UA-EffNet-DA achieved a precision, recall, and F1-score of 0.9647, 0.9605, and 0.9624, respectively. These results indicate strong and balanced multi-class discrimination across the glioma, meningioma, and pituitary tumor categories. Class-wise analysis further shows the highest recall for the pituitary class (0.9946), indicating highly reliable identification of pituitary tumor samples.

On the Kaggle dataset, the proposed model achieved an overall classification accuracy of 99.23% with a macro-averaged F1-score of 0.9917. Class-wise evaluation indicates consistently high precision and recall across all categories, with the no-tumor class achieving perfect classification performance. To further examine the decision-making behavior of the network, Grad-CAM visualizations are presented in Figures 9 and 10. The activation maps highlight diagnostically relevant tumor regions, demonstrating that the model focuses on meaningful anatomical structures rather than background artifacts.

These Grad-CAM results provide qualitative support for the proposed framework by showing that the network generally concentrates on lesion-relevant anatomical regions rather than irrelevant background content. This interpretability analysis complements the quantitative performance metrics and helps clarify the model's decision-making behavior across different lesion categories. For glioma cases, the model consistently attends to irregular infiltrative regions, whereas meningioma samples exhibit localized activation around well-circumscribed extra-axial masses. Pituitary adenoma predictions show fo-

cused responses within sellar regions, while non-tumor samples demonstrate minimal activation outside normal anatomical boundaries. Second, transfer learning from ImageNet-pretrained weights was used to initialize the EfficientNet-B4 backbone. Rather than freezing early layers, the entire backbone was fine-tuned end-to-end to adapt pretrained feature representations to the MRI domain while preserving the benefits of pretraining.

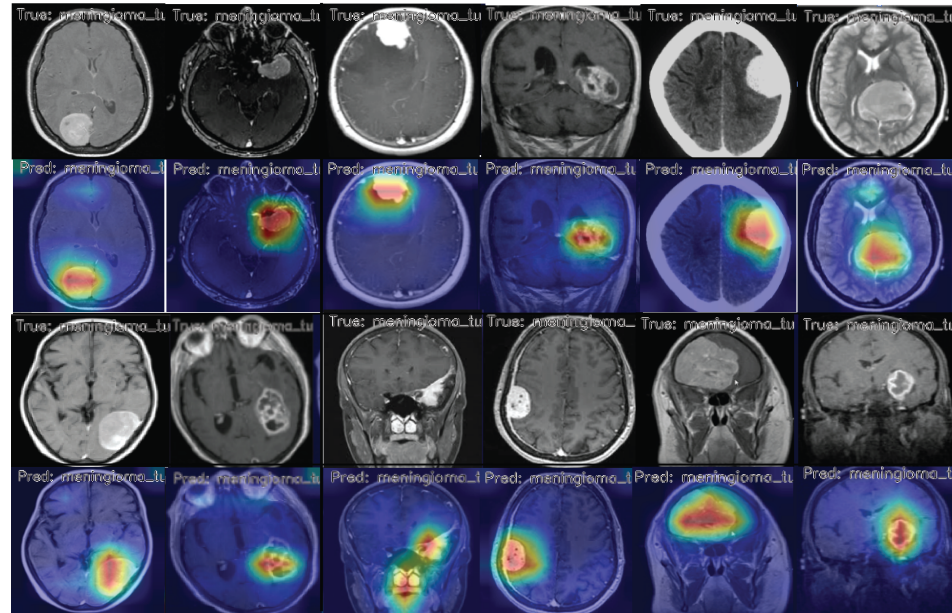


Figure 9. Grad-CAM results of UA-EffNet-DA on the Kaggle brain tumor dataset.

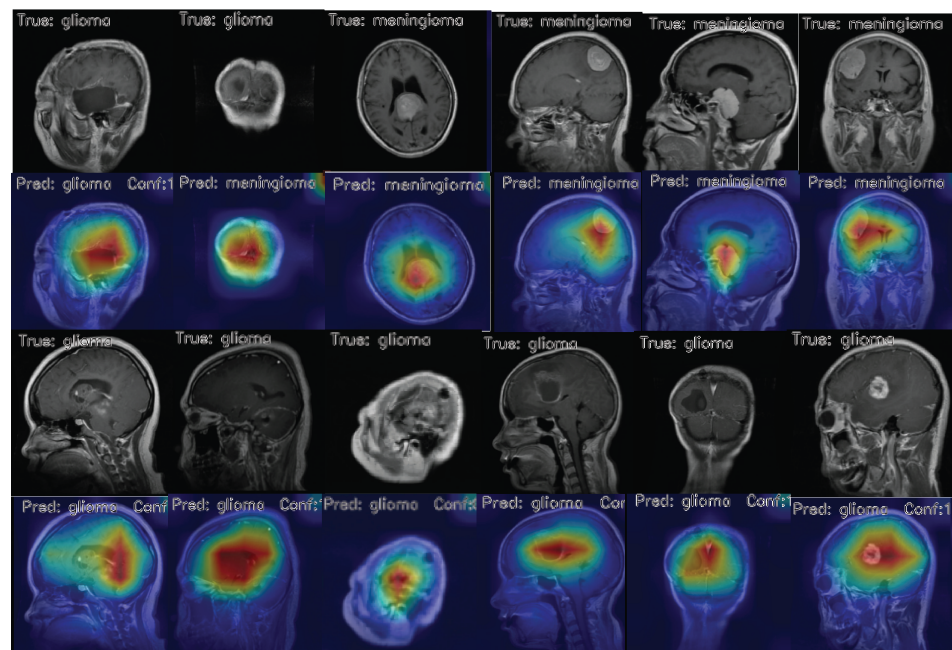


Figure 10. Grad-CAM results of UA-EffNet-DA on the Figshare brain tumor dataset.

4.4. Ablation Study

To assess the contribution of each component, ablation experiments were conducted on both the Figshare and Kaggle brain tumor MRI datasets. Four variants were evaluated under the same training and testing settings: (i) a baseline EfficientNet-B4 without attention or MC dropout, (ii) EfficientNet-B4 with channel attention (CA) only, (iii) EfficientNet-B4 with channel and spatial attention (CA + SA), and (iv) the full UA-EffNet-DA model

with CA, SA, and MC dropout. Accuracy and mean AUC were used as the primary comparison metrics.

Table 2 shows the results on the Figshare dataset. Adding channel attention improves the baseline performance, indicating the benefit of channel-wise feature recalibration. The CA + SA variant remains competitive, suggesting that spatial attention preserves strong discriminative performance while enhancing spatial refinement. The full model achieves the highest accuracy, indicating that the complete framework is effective in this setting.

Table 2. Ablation study of UA-EffNet-DA on the Figshare dataset.

Variant	CA	SA	MC Dropout	Accuracy	Mean AUC
Baseline	No	No	No	0.94894	0.997867
CA only	Yes	No	No	0.95632	0.998236
CA + SA	Yes	Yes	No	0.95632	0.998230
Full model	Yes	Yes	Yes	0.98732	0.997769

Table 3 presents the results on the Kaggle dataset. Here, CA provides the clearest gain over the baseline, while the additional effect of SA is not consistently reflected in top-line accuracy. The full model remains highly competitive and, importantly, extends the framework with uncertainty estimation. This suggests that SA contributes more to spatial focus and interpretability than to a guaranteed accuracy increase in every setting. Likewise, MC dropout is included primarily to support uncertainty-aware inference and regularization rather than to consistently improve deterministic classification accuracy.

Table 3. Ablation study of UA-EffNet-DA on the Kaggle dataset.

Variant	CA	SA	MC Dropout	Accuracy	Mean AUC
Baseline	No	No	No	0.987796	0.999448
CA only	Yes	No	No	0.993898	0.999859
CA + SA	Yes	Yes	No	0.987796	0.999693
Full model	Yes	Yes	Yes	0.992372	0.999511

Overall, the ablation study supports the design rationale of UA-EffNet-DA, while also indicating that the proposed components play different functional roles across datasets rather than producing strictly monotonic gains in every case.

5. Comparison with State of the Art Approaches

This section presents a comparative evaluation of the proposed UA-EffNet-DA model against existing CNN-based transfer learning approaches and transformer-based architectures for brain tumor MRI classification. The comparison aims to assess the effectiveness of the proposed dual-attention and uncertainty-aware design in relation to widely used baseline families with different feature learning characteristics.

5.1. Comparison with Existing CNN-Based Transfer Learning Approaches

To further evaluate the proposed framework, we compared its performance with representative CNN-based transfer learning models, including ResNet152, DenseNet121, EfficientNetB0, and EDNet. Table 4 summarizes the class-wise and average results in terms of accuracy, precision, recall, specificity, and F1-score for the four classes: glioma, no tumor, meningioma, and pituitary.

Table 4. Performance comparison of CNN-based and transfer-learning approaches for brain tumor MRI classification. Here: G, glioma; N, no tumor; M, meningioma; P, pituitary; and A, average.

Model	Class	Accuracy (%)	Precision (%)	Recall (%)	Specificity (%)	F1-Score (%)
ResNet152 [9]	G	89.18	75.84	78.61	93.40	76.80
	N	94.45	87.69	75.77	97.97	75.90
	M	83.77	84.41	76.55	86.45	74.72
	P	92.22	98.20	86.23	94.26	82.16
	A	89.91	86.54	79.29	93.02	77.40
DenseNet121 [10]	G	96.45	94.77	92.80	97.80	93.57
	N	98.25	93.81	95.77	98.74	94.52
	M	96.38	94.26	93.44	97.58	93.67
	P	98.74	97.32	98.05	99.00	97.65
	A	97.46	95.04	95.02	98.28	94.85
EfficientNetB0 [11]	G	97.24	95.63	94.60	99.19	95.86
	N	98.65	96.41	95.62	98.63	96.76
	M	96.35	95.90	91.51	98.41	93.57
	P	97.95	93.70	99.36	97.43	96.38
	A	97.55	95.41	95.27	98.42	95.64
EDNet [40]	G	98.25	98.11	95.73	99.22	96.87
	N	99.57	98.80	98.38	99.78	98.58
	M	97.70	96.01	95.97	98.37	95.93
	P	98.96	97.02	99.40	98.81	98.18
	A	98.62	97.49	97.37	99.05	97.39
Swin Transformer [41]	G	83.49	75.32	62.37	91.88	68.24
	N	91.44	73.47	70.59	95.29	72.00
	M	77.98	67.65	47.92	90.48	56.10
	P	84.10	63.16	96.55	79.58	76.36
	A	84.25	69.90	69.36	89.31	68.18
Vision Transformer [42]	G	86.06	78.74	70.20	92.49	73.85
	N	91.44	76.93	78.14	95.68	77.33
	M	80.97	67.75	65.50	87.29	66.35
	P	93.66	83.83	95.38	93.06	89.14
	A	88.42	76.81	77.31	92.13	76.67

As shown in Table 4, ResNet152 [9] yields the weakest overall performance, particularly in precision, recall, and F1-score, indicating comparatively limited discriminative capability for complex brain tumor MRI patterns. DenseNet121 [10] improves all major metrics, reflecting the benefit of dense feature reuse and stronger transfer learning capability. EfficientNetB0 [11] provides a further improvement in average accuracy and F1-score, confirming the effectiveness of efficient compound scaling for medical image classification.

5.2. Comparison with State-of-the-Art Approaches

This section compares the proposed UA-EffNet-DA model with existing state-of-the-art approaches for brain tumor MRI detection. Table 5 presents the comparative performance of prior methods and the proposed framework on the Figshare and Kaggle datasets. The results show that UA-EffNet-DA achieves strong performance on both benchmarks, indicating that the proposed combination of an EfficientNet-B4 backbone, attention-based feature refinement, and uncertainty-aware inference is effective across datasets with different class composition and imaging variability.

The proposed framework was designed to address several limitations observed in earlier methods. First, EfficientNet-B4 provides a strong and computationally efficient backbone for feature extraction. Second, channel and spatial attention refinement improves feature selectivity by jointly modeling discriminative channel responses and anatomically relevant spatial regions. Third, MC dropout enables uncertainty-aware inference, allowing the framework to provide predictive confidence information in addition to class labels.

To quantitatively assess the uncertainty-aware behavior of the proposed framework, we extended the evaluation beyond standard classification metrics. Since dropout is retained during inference, predictive uncertainty was estimated using Monte Carlo dropout

with $T = 20$ stochastic forward passes. For each test image, the predictive mean probability vector was used for calibration analysis, while predictive variance was used as an uncertainty measure. Expected Calibration Error (ECE), Negative Log-Likelihood (NLL), and Brier score were computed to assess the quality of the predicted probabilities. In addition, reliability-oriented analysis, uncertainty-versus-error comparison, and uncertainty-based rejection analysis were performed to examine whether the model's confidence estimates were informative and practically useful. Beyond the detection performance reported in Table 5, the additional analyses in Tables 6 and 7 provide further evidence regarding the reliability of the proposed framework by reporting calibration metrics, predictive uncertainty, per-class Brier scores, and uncertainty-based rejection behavior. Moreover, Table 8 summarizes the computational characteristics of the compared models and the proposed framework, which is important for assessing deployment-oriented practicality. Taken together, these results suggest that the proposed model is not only competitive in classification performance but also supports confidence-aware interpretation while maintaining a lightweight single-backbone design.

Table 5. Comparative evaluation of the proposed UA-EffNet-DA model against leading approaches. FD = Figshare Dataset and KD = Kaggle Dataset.

Authors	Method	Dataset	Accuracy (%)	Precision (%)	Recall (%)	Specificity (%)	F1-Score (%)
Rahman et al. [43]	Parallel deep CNN	FD	97.60	97.00	97.00	–	97.00
Hammad et al. [44]	CNN	FD	96.86	97.00	96.00	–	97.00
Chen et al. [45]	ResNet101 + DenseNet121	KD	97.24	97.06	97.58	–	97.28
Demir and Akbulut [46]	R-CNN	KD	96.60	96.25	96.50	98.75	96.75
Saurav et al. [47]	AG-CNN	KD	95.71	95.77	96.33	98.52	95.98
Kumar et al. [48]	RNGAP	KD	97.08	95.20	95.20	–	95.50
Deepak et al. [49]	CNN and SVM	KD	95.82	92.00	92.00	–	90.21
Wang et al. [50]	Transformer	KD	95.31	95.29	95.17	95.25	94.22
Hossain et al. [51]	IVX16	KD	96.94	77.75	78.75	–	75.75
Zhu et al. [52]	Biformer	KD	95.36	95.67	–	96.93	94.82
Zebari et al. [53]	ResNet50 + VGG16	KD	88.19	87.14	87.26	88.16	87.65
Wang et al. [54]	RanMerFormer	FD	96.90	97.40	95.63	98.21	–
Sreedevi et al. [55]	ResNet50 + Bi-GRU + dual attention	KD	92.02	91.29	91.28	82.49	91.35
Chaki et al. [56]	BTSCNet	FD	96.60	93.00	95.03	95.50	–
Akter et al. [57]	Deep CNN	FD	96.70	100.00	93.60	–	96.70
Vijayalakshmi et al. [58]	CB-CIRL Net	FD	97.06	96.88	97.16	–	97.01
Kumar et al. [59]	Dolphin-SCA-based deep CNN	KD	95.30	–	97.70	95.30	–
Our proposed	UA-EffNet-DA	FD	98.73	97.21	97.89	97.55	98.01
Our proposed	UA-EffNet-DA	KD	99.23	99.23	99.15	99.16	97.70

Table 6. Calibration, predictive uncertainty, per-class Brier score, and uncertainty-based rejection analysis for the Figshare dataset.

Section	Metric	Value
Calibration	ECE	0.019708
	NLL	0.092271
	Overall Brier Score	0.048438
Uncertainty	Mean Uncertainty (Correct)	0.000263
	Mean Uncertainty (Incorrect)	0.007472
Per-class Brier	Glioma	0.016566
	Meningioma	0.023716
	Pituitary	0.008328
Rejection Analysis	50% retention (306 samples)	1.000000
	60% retention (367 samples)	1.000000
	70% retention (428 samples)	1.000000
	80% retention (489 samples)	1.000000
	90% retention (550 samples)	0.992727
	100% retention (612 samples)	0.967320

Table 7. Calibration, predictive uncertainty, per-class Brier score, and uncertainty-based rejection analysis for the Kaggle dataset.

Section	Metric	Value
Calibration	ECE	0.011842
	NLL	0.041386
	Overall Brier Score	0.021754
Uncertainty	Mean Uncertainty (Correct)	0.000148
	Mean Uncertainty (Incorrect)	0.004913
Per-class Brier	Glioma	0.024381
	Meningioma	0.019764
	Pituitary	0.011925
	No-tumor	0.013508
Rejection Analysis	50% retention (197 samples)	1.000000
	60% retention (236 samples)	1.000000
	70% retention (275 samples)	0.996364
	80% retention (315 samples)	0.993651
	90% retention (354 samples)	0.992938
	100% retention (394 samples)	0.992300

Table 8. Comparison of model complexity, memory consumption, training time, inference time, and computational cost of baseline models and the proposed UA-EffNet-DA model.

Models	Total Parameters	Trainable Parameters	Memory Size (MB)	Training Time per Fold (min)	Inference Time (s)	FLOPs (G)
ResNet152 [9]	84,023,044	83,879,300	321	42	13	11.51
DenseNet121 [10]	7,041,604	6,957,956	27.8	22	6	2.83
EfficientNetB0 [11]	4,054,695	4,012,672	15.9	19	4	0.39
EDNet [40]	11,096,319	10,970,648	43.7	35	8	–
Vision Transformer [42]	9,392,388	9,392,388	36	9	7	17.56
Swin Transformer [41]	349,812	337,236	1.38	8	4	4.49
UA-EffNet-DA	19,747,637	19,747,637	75.33	0.5862	0.032078	1.5665

6. Conclusions

In this study, we propose UA-EffNet-DA, an uncertainty-aware EfficientNet-B4 framework with dual attention refinement for brain tumor MRI classification, targeting a key limitation of existing CNN- and transformer-based methods: insufficient feature selectivity and spatial localization, coupled with deterministic predictions that do not convey reliability for clinical decision support. Our approach adopts an efficient single-backbone design and introduces a lightweight channel and SA to explicitly learn both what discriminative features to emphasize and where to focus within MRI slices, while MC dropout enables confidence-aware inference by estimating predictive uncertainty. This combination constitutes the core novelty of UA-EffNet-DA, improving discriminative representation learning without the computational burden of multi-branch or ensemble architectures and supporting safer decision-making through uncertainty estimates. Extensive experiments on Figshare and Kaggle datasets demonstrate strong performance relative to existing approaches, and ablation studies validate the complementary contribution of dual attention and uncertainty-aware inference under controlled settings. Future work will extend the framework to patient-level evaluation and broader multi-institutional validation, and will investigate improved calibration and uncertainty modeling to further enhance clinical reliability.

Author Contributions: Conceptualization, S.A. and J.H.L.; methodology, S.A.; software, S.A.; formal analysis, S.A.; investigation, S.A.; data curation, S.A.; visualization, S.A.; writing original draft preparation, S.A.; writing review and editing, S.A. and J.H.L.; supervision, J.H.L.; project administration, J.H.L.; and funding acquisition, J.H.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Commercialization Promotion Agency for R&D Outcomes (COMPA) funded by the Ministry of Science and ICT (MSIT). (RS-2025-02413013 IP Enhancement and Commercialization to Promote the Market Adoption of Vision Intelligence-Based Multi-Biosignal Measurement Technology) and this research was supported by the Regional Innovation System & Education(RISE) program through the Daegu RISE Center, funded by the Ministry of Education(MOE) and the Daegu, Republic of Korea (2025-RISE-03-002) and this research was supported by a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (grant number : HI23C0942, RS-2024-00433896).

Data Availability Statement: The datasets used in this study are publicly available. The Figshare brain tumor dataset is available at <https://www.kaggle.com/datasets/ashkhagan/figshare-brain-tumor-dataset> (accessed on 10 April 2026), and the Brain Tumor MRI dataset is available at <https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset> (accessed on 10 April 2026). These datasets were derived from public-domain resources and are also described in the corresponding references [38,39].

Acknowledgments: During the preparation of this work the authors used AI in order to improve language clarity and readability. After using this service, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AI	Artificial Intelligence
AUC	Area Under the Curve
AUC–ROC	Area Under the Receiver Operating Characteristic Curve
CA	Channel Attention
CAD	Computer-Aided Diagnosis
CE-MRI	Contrast-Enhanced Magnetic Resonance Imaging
CNN	Convolutional Neural Network
Grad-CAM	Gradient-weighted Class Activation Mapping
MC Dropout	Monte Carlo Dropout
MRI	Magnetic Resonance Imaging
ReLU	Rectified Linear Unit
ROC	Receiver Operating Characteristic
SA	Spatial Attention
UA-EffNet-DA	Uncertainty-Aware EfficientNet with Dual Attention
ViT	Vision Transformer

References

1. Dorfner, F.J.; Patel, J.B.; Kalpathy-Cramer, J.; Gerstner, E.R.; Bridge, C.P. A Review of Deep Learning for Brain Tumor Analysis in MRI. *npj Precis. Oncol.* **2025**, *9*, 2. [[CrossRef](#)]
2. Vollmuth, P.; Karschnia, P.; Sahm, F.; Park, Y.W.; Ahn, S.S.; Jain, R. A Radiologist’s Guide to IDH-Wildtype Glioblastoma for Efficient Communication with Clinicians: Part I—Essential Information on Preoperative and Immediate Postoperative Imaging. *Korean J. Radiol.* **2025**, *26*, 246–268. [[CrossRef](#)]

3. Ghadimi, D.J.; Vahdani, A.M.; Karimi, H.; Ebrahimi, P.; Fathi, M.; Moodi, F.; Khodadadi Shoushtari, F.; Valizadeh, G.; Mobarak Salari, H.; Saligheh Rad, H. Deep Learning-Based Techniques in Glioma Brain Tumor Segmentation Using Multi-Parametric MRI: A Review on Clinical Applications and Future Outlooks. *J. Magn. Reson. Imaging* **2025**, *61*, 1094–1109. [\[CrossRef\]](#)
4. van der Loo, I.; Bucho, T.M.T.; Hanley, J.A.; Beets-Tan, R.G.H.; Imholz, A.L.T.; Trebeschi, S. Measurement Variability of Radiologists When Measuring Brain Tumors. *Eur. J. Radiol.* **2025**, *183*, 111874. [\[CrossRef\]](#)
5. Iftikhar, S.; Anjum, N.; Siddiqui, A.B.; Ur Rehman, M.; Ramzan, N. Explainable CNN for Brain Tumor Detection and Classification Through XAI Based Key Features Identification. *Brain Inform.* **2025**, *12*, 10. [\[CrossRef\]](#)
6. Afzal, S.; Khan, I.U.; Lee, J.W. A transfer learning-based approach to detect cerebral microbleeds. *Comput. Mater. Contin.* **2022**, *71*, 1903. [\[CrossRef\]](#)
7. Shi, T.; Kujawa, A.; Linares, C.; Vercauteren, T.; Booth, T.C. Automated Longitudinal Treatment Response Assessment of Brain Tumors: A Systematic Review. *Neuro-Oncology* **2025**, *27*, 1946–1971. [\[CrossRef\]](#)
8. Chen, C.; Mat Isa, N.A.; Liu, X. A Review of Convolutional Neural Network Based Methods for Medical Image Classification. *Comput. Med. Imaging Graph.* **2025**, *120*, 109507. [\[CrossRef\]](#) [\[PubMed\]](#)
9. Zhang, L.; Li, H.; Zhu, R.; Du, P. An infrared and visible image fusion algorithm based on ResNet-152. *Multimed. Tools Appl.* **2022**, *81*, 9277–9287. [\[CrossRef\]](#)
10. Albelwi, S.A. Deep architecture based on DenseNet-121 model for weather image recognition. *Int. J. Adv. Comput. Sci. Appl.* **2022**, *13*, 559–656. [\[CrossRef\]](#)
11. Kumar, A.; Nelson, L.; Arumugam, D. Deep Learning-Based Classification of Brain Tumours on MRI Images Using EfficientNetB0. In Proceedings of the 2024 4th International Conference on Technological Advancements in Computational Sciences (ICTACS), Tashkent, Uzbekistan, 13–15 November 2024; pp. 219–225.
12. Hosny, K.M.; Tarawneh, A. Explainable AI and Vision Transformers for Detection and Classification of Brain Tumor: A Comprehensive Survey. *Artif. Intell. Rev.* **2025**, *58*, 259. [\[CrossRef\]](#)
13. Hasan, M.J.; Hasan, M.; Akter, S.; Mahi, A.B.S.; Uddin, M.P. Enhancing Brain Tumor Classification with a Novel Attention Based Explainable Deep Learning Framework. *Biomed. Signal Process. Control* **2026**, *112*, 108636. [\[CrossRef\]](#)
14. Afzal, S.; Maqsood, M.; Mehmood, I.; Niaz, M.T.; Seo, S. An efficient false-positive reduction system for cerebral microbleeds detection. *Comput. Mater. Contin.* **2021**, *66*, 2301. [\[CrossRef\]](#)
15. Lu, N.H.; Huang, Y.H.; Liu, K.Y.; Chen, T.B. Deep Learning-Driven Brain Tumor Classification and Segmentation Using Non-Contrast MRI. *Sci. Rep.* **2025**, *15*, 27831. [\[CrossRef\]](#)
16. Zarenia, E.; Far, A.A.; Rezaee, K. Automated Multi-Class MRI Brain Tumor Classification and Segmentation Using Deformable Attention and Saliency Mapping. *Sci. Rep.* **2025**, *15*, 8114. [\[CrossRef\]](#) [\[PubMed\]](#)
17. Patel, R.; El Bahi, H. Deep Learning Approaches for Brain Tumor Detection and Classification Using MRI Images (2020 to 2024): A Systematic Review. *J. Imaging Inform. Med.* **2025**, *38*, 1403–1433. [\[CrossRef\]](#)
18. Abdelhaliem, I.; Dixon, J.; Abdelhamid, A.; Saleh, G.A.; Khalifa, F. A Multimodal Adaptive Inter-Region Attention-Guided Network for Brain Tumor Classification. *IEEE Access* **2025**, *13*, 187964–187975. [\[CrossRef\]](#) [\[PubMed\]](#)
19. Liu, Y.; Wen, Y.; Yang, L.; He, L.; Zhou, M. A General Framework for Efficient Medical Image Analysis via Shared Attention Vision Transformer. *IEEE Trans. Med. Imaging* **2025**. [\[CrossRef\]](#) [\[PubMed\]](#)
20. Zhang, R.; Luo, X.; Lv, J.; Cao, J.; Zhu, Y.; Wang, J.; Zheng, B. Enhancing medical image classification with context modulated attention and multi-scale feature fusion. *IEEE Access* **2025**, *13*, 15226–15243. [\[CrossRef\]](#)
21. Ullah, Z.; Hong, M.; Mahmood, T.; Kim, J. Systematic Integration of Attention Modules into CNNs for Accurate and Generalizable Medical Image Classification. *Mathematics* **2025**, *13*, 3728. [\[CrossRef\]](#)
22. Islam, M.R.; Talukder, M.A.; Uddin, M.A.; Akhter, A.; Khalid, M. BrainNet: Precision Brain Tumor Classification with Optimized EfficientNet Architecture. *Int. J. Intell. Syst.* **2024**, *2024*, 3583612. [\[CrossRef\]](#)
23. Zulfiqar, F.; Bajwa, U.I.; Mehmood, Y. Multi-class classification of brain tumor types from MR images using EfficientNets. *Biomed. Signal Process. Control* **2023**, *84*, 104777. [\[CrossRef\]](#)
24. Ishaq, A.; Ullah, F.U.M.; Hamandawana, P.; Cho, D.-J.; Chung, T.-S. Improved EfficientNet Architecture for Multi-Grade Brain Tumor Detection. *Electronics* **2025**, *14*, 710. [\[CrossRef\]](#)
25. Alnashwan, R.O.; Bibi, R.; Ali, A.; Jamil, F.; Saeed, Y.; Alnashwan, R.O.; Chelloug, S.A.; Muthanna, M.S.A. Automatic smart brain tumor classification and prediction system using deep learning. *Sci. Rep.* **2025**, *15*, 14876. [\[CrossRef\]](#)
26. Vimala, B.B.; Mathivanan, S.K.; Shah, M.A. Efficient brain tumor grade classification using ensemble deep learning models. *BMC Med. Imaging* **2024**, *24*, 297. [\[CrossRef\]](#)
27. R, P.; M, J.P.P.; Nisha, J.S. Brain tumor segmentation using multi-scale attention U-Net with EfficientNetB4 encoder for enhanced MRI analysis. *Sci. Rep.* **2025**, *15*, 9914. [\[CrossRef\]](#) [\[PubMed\]](#)
28. Nahiduzzaman, M.; Wani, N.; Ramani, M.; Kumar, A.; Pant, S.; Kotecha, K.; Kulkarni, A. Optimized deep learning for brain tumor detection: A hybrid approach with attention mechanisms and clinical explainability. *Sci. Rep.* **2025**, *15*, 31386. [\[CrossRef\]](#)

29. Ahmed, M.; Hossain, M.M.; Islam, M.R.; Ali, M.S.; Nafi, A.A.N.; Ahmed, M.F.; Ahmed, K.M.; Miah, S.; Rahman, M.; Niu, M.; et al. Brain tumor detection and classification in MRI using hybrid ViT and GRU model with explainable AI in Southern Bangladesh. *Sci. Rep.* **2024**, *14*, 22797. [\[CrossRef\]](#)
30. Rasheed, Z.; Ma, Y.K.; Ullah, I.; Al-Khasawneh, M.; Almutairi, S.S.; Abohashrh, M. Integrating convolutional neural networks with attention mechanisms for magnetic resonance imaging-based classification of brain tumors. *Bioengineering* **2024**, *11*, 701. [\[CrossRef\]](#)
31. Celik, F.; Celik, K.; Celik, A. Enhancing brain tumor classification through ensemble attention mechanism. *Sci. Rep.* **2024**, *14*, 22260. [\[CrossRef\]](#) [\[PubMed\]](#)
32. Abdusalomov, A.B.; Mukhiddinov, M.; Whangbo, T.K. Brain Tumor Detection Based on Deep Learning Approaches and Magnetic Resonance Imaging. *Cancers* **2023**, *15*, 4172. [\[CrossRef\]](#)
33. Chandraprabha, K.; Ganesan, L.; Baskaran, K. A novel approach for the detection of brain tumor and its classification via end-to-end vision transformer-CNN architecture. *Front. Oncol.* **2025**, *15*, 1508451. [\[CrossRef\]](#) [\[PubMed\]](#)
34. Reddy, C.K.K.; Reddy, P.A.; Janapati, H.; Assiri, B.; Shuaib, M.; Alam, S.; Sheneamer, A. A fine-tuned vision transformer based enhanced multi-class brain tumor classification using MRI scan imagery. *Front. Oncol.* **2024**, *14*, 1400341. [\[CrossRef\]](#) [\[PubMed\]](#)
35. Krishnan, P.T.; Krishnadoss, P.; Khandelwal, M.; Gupta, D.; Nihaal, A.; Kumar, T.S. Enhancing brain tumor detection in MRI with a rotation invariant Vision Transformer. *Front. Neuroinform.* **2024**, *18*, 1414925. [\[CrossRef\]](#)
36. Sankari, C.; Jamuna, V.; Kavitha, A.R. Hierarchical multi-scale vision transformer model for accurate detection and classification of brain tumors in MRI-based medical imaging. *Sci. Rep.* **2025**, *15*, 38275. [\[CrossRef\]](#) [\[PubMed\]](#)
37. Shahin, A. Fine-tuned ResNet34 for efficient brain tumor classification. *Sci. Rep.* **2025**, *15*, 36910. [\[CrossRef\]](#)
38. Singh, J. FigShare. *J. Pharmacol. Pharmacother.* **2011**, *2*, 138. [\[CrossRef\]](#)
39. Nickparvar, M. *Brain Tumor MRI Dataset*; Kaggle: San Francisco, CA, USA, 2021.
40. Choi, Y.; Lee, Y.; Shin, D.; Cho, J.; Park, S.; Lee, S.; Baek, J.; Bae, C.; Kim, B.; Heo, J. Ednet: A large-scale hierarchical dataset in education. In *Proceedings of the International Conference on Artificial Intelligence in Education*; Springer: Cham, Switzerland, 2020; pp. 69–73.
41. Liu, Z.; Lin, Y.; Cao, Y.; Hu, H.; Wei, Y.; Zhang, Z.; Lin, S.; Guo, B. Swin transformer: Hierarchical vision transformer using shifted windows. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, Montreal, QC, Canada, 10–17 October 2021; pp. 10012–10022.
42. Dosovitskiy, A.; Beyer, L.; Kolesnikov, A.; Weissenborn, D.; Zhai, X.; Unterthiner, T.; Dehghani, M.; Minderer, M.; Heigold, G.; Gelly, S.; et al. An image is worth 16 × 16 words: Transformers for image recognition at scale. *arXiv* **2020**, arXiv:2010.11929.
43. Rahman, T.; Islam, M.S. MRI Brain Tumor Detection and Classification Using Parallel Deep Convolutional Neural Networks. *Meas. Sens.* **2023**, *26*, 100694. [\[CrossRef\]](#)
44. Hammad, M.; ElAffendi, M.; Ateya, A.A.; Abd El-Latif, A.A. Efficient Brain Tumor Detection with Lightweight End-to-End Deep Learning Model. *Cancers* **2023**, *15*, 2837. [\[CrossRef\]](#)
45. Chen, W.; Tan, X.; Zhang, J.; Du, G.; Fu, Q.; Jiang, H. A Robust Approach for Multi-Type Classification of Brain Tumor Using Deep Feature Fusion. *Front. Neurosci.* **2024**, *18*, 1288274. [\[CrossRef\]](#)
46. Demir, F.; Akbulut, Y. A New Deep Technique Using R-CNN Model and L1NSR Feature Selection for Brain MRI Classification. *Biomed. Signal Process. Control* **2022**, *75*, 103625. [\[CrossRef\]](#)
47. Saurav, S.; Sharma, A.; Saini, R.; Singh, S. An Attention-Guided Convolutional Neural Network for Automated Classification of Brain Tumor from MRI. *Neural Comput. Appl.* **2023**, *35*, 2541–2560. [\[CrossRef\]](#)
48. Kumar, R.L.; Kakarla, J.; Isunuri, B.V.; Singh, M. Multi-Class Brain Tumor Classification Using Residual Network and Global Average Pooling. *Multimed. Tools Appl.* **2021**, *80*, 13429–13438. [\[CrossRef\]](#)
49. Deepak, S.; Ameer, P.M. Automated Categorization of Brain Tumor from MRI Using CNN Features and SVM. *J. Ambient. Intell. Humaniz. Comput.* **2021**, *12*, 8357–8369. [\[CrossRef\]](#)
50. Wang, S.; Wu, J.; Chen, M.; Huang, S.; Huang, Q. Balanced Transformer: Efficient Classification of Glioblastoma and Primary Central Nervous System Lymphoma. *Phys. Med. Biol.* **2024**, *69*, 045032. [\[CrossRef\]](#)
51. Hossain, S.; Chakrabarty, A.; Gadekallu, T.R.; Alazab, M.; Piran, M.J. Vision Transformers, Ensemble Model, and Transfer Learning Leveraging Explainable AI for Brain Tumor Detection and Classification. *IEEE J. Biomed. Health Inform.* **2024**, *28*, 1261–1272. [\[CrossRef\]](#) [\[PubMed\]](#)
52. Zhu, L.; Wang, X.; Ke, Z.; Zhang, W.; Lau, R.W.H. Biformer: Vision Transformer with Bi-Level Routing Attention. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Vancouver, BC, Canada, 18–22 June 2023; pp. 10323–10333.
53. Zebari, N.A.; Mohammed, C.N.; Zebari, D.A.; Mohammed, M.A.; Zeebaree, D.Q.; Marhoon, H.A.; Abdulkareem, K.H.; Kadry, S.; Viriyasitavat, W.; Nedoma, J.; et al. A Deep Learning Fusion Model for Accurate Classification of Brain Tumours in Magnetic Resonance Images. *CAAI Trans. Intell. Technol.* **2024**, *9*, 790–804. [\[CrossRef\]](#)

54. Wang, J.; Lu, S.Y.; Wang, S.H.; Zhang, Y.D. RanMerFormer: Randomized Vision Transformer with Token Merging for Brain Tumor Classification. *Neurocomputing* **2024**, *573*, 127216. [[CrossRef](#)]
55. Sreedevi, P.; Kiran, A.; Sri, T.S.; Poornima, E.; Polepaka, S.; Reddy, Y.S. Da-resbigru-brain tumor classification using Dual attention residual bi-directional gated recurrent unit using MRI images. *Biomed. Signal Process. Control* **2024**, *88*, 105596. [[CrossRef](#)]
56. Chaki, J.; Woźniak, M. A Deep Learning Based Four-Fold Approach to Classify Brain MRI: BTSCNet. *Biomed. Signal Process. Control* **2023**, *85*, 104902. [[CrossRef](#)]
57. Akter, A.; Nosheen, N.; Ahmed, S.; Hossain, M.; Yousuf, M.A.; Almoyad, M.A.A.; Hasan, K.F.; Moni, M.A. Robust Clinical Applicable CNN and U-Net Based Algorithm for MRI Classification and Segmentation for Brain Tumor. *Expert Syst. Appl.* **2024**, *238*, 122347. [[CrossRef](#)]
58. Vijayalakshmi, B.; Anand, S. Cross Prior Bayesian Attention with Correlated Inception and Residual Learning for Brain Tumor Classification Using MR Images (CB-CIRL Net). *J. Neurosci. Methods* **2025**, *417*, 110392. [[CrossRef](#)] [[PubMed](#)]
59. Kumar, S.; Mankame, D.P. Optimization Driven Deep Convolution Neural Network for Brain Tumor Classification. *Biocybern. Biomed. Eng.* **2020**, *40*, 1190–1204. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.