*Electronic Letters on Computer Vision and Image Analysis 25(3):65-80, 2026*

# Brain Tumor Classification Accuracy Enhancement via DenseNet201 Compared to Four Alternative Models

Zahraa Thamer[1*] Noor S. Sagheer[2] Oğuz Ata[3] Ashwan A. Abdulmunem[4] and Hawraa Thamer[5]

[1*] *Department of IT, College of Computer Science and Information Technology, University of Karbala, Karbala 56001, Iraq*
[2,4] *Department of Computer Sciences, College of Computer Science and Information Technology, University of Karbala'a, Karbala, Iraq*
[3] *Department of Computer Sciences, College of Computer Engineering, University of Altınbaş, Mahmutbey Mahallesi, İstanbul, Turkey*
[5] *Department of Banking and Financial Sciences, College of Administration and Economics, University of Kerbala'a, Karbala, Iraq*

Received 1st of August 2025; accepted 1st of September 2026

**Abstract**

Early brain tumor staging is critical to the success of currently available treatment planning and modalities. Magnetic Resonance Imaging (MRI) is a necessary method for treatment planning and early diagnosis. Most current classification methods do not achieve a high level of classification accuracy, which may result in some tumor cases being misclassified or classified as suspicious. In our proposed work, we leverage the DenseNet201 deep learning architecture as the foundation for classifying brain tumors. A dataset in this methodology consisted of MRI scans, which included four unique types of tumors: meningioma, pituitary tumor, no tumor, and glioma. To improve classification precision, improve generalization, and reduce overfitting, we employed three strategies (a1, a2, and a3) during the classification process. When our model was evaluated on unseen test data, it achieved an impressive 99.77% accuracy. The proposed technique effectively classifies malignant brain tumors, despite the similarities within classes and the varying tumor morphologies. There are four state-of-the-art techniques of deep learning that we compared our results with: ResNetRS50, ResNet50, VGG19, and EfficientNetV2, which achieved accuracies of 99.77%, 98.71%, 96.88%, and 96.58%, respectively. In the field of brain tumor classification, our approach has outperformed others in terms of performance, highlighting its potential for clinical applications.

*Key Words*: Brain tumor, Convolutional Neural Network (CNN), DenseNet201, Magnetic Resonance Imaging, ResNetRS50, Deep Learning.

## 1 Introduction

Brain tumors are among the most dangerous diseases of the nervous system that pose a direct threat to human life today. The earlier and more accurate the diagnosis, the higher the success rate of treatment and the chance of survival. Due to the diverse functional and morphological characteristics of tumors and the complexity of the anatomical structure of the brain, diagnosis is a significant challenge that requires a high level of analytical precision and advanced medical expertise [1]. Magnetic resonance imaging (MRI)is among the

Correspondence to: noor.sabah@uokerbala.edu.iq

Recommended for acceptance by Angel D. Sappa
https://doi.org/10.5565/rev/elcvia.2305
ELCVIA ISSN: 1577-5097
Published by Computer Vision Center / Universitat Autònoma de Barcelona, Barcelona, Spain

most commonly used medical imaging techniques for diagnosing brain tumors because of its ability to provide high-resolution representations of soft tissues, which allows for precise determination of the tumor's location, characteristics, and size [2] [3]. Therefore, recent studies (2025) have utilized machine learning techniques to improve MRI image analysis, thereby reducing dependence on human assessment and enhancing clinical diagnostic accuracy [4]. Relying solely on human expertise is a laborious and time-consuming process that can lead to inconsistencies in diagnostic results. It is therefore crucial to develop Computer-Aided Diagnostic (CAD) systems that support medical decision-making and provide rapid and accurate results. Artificial intelligence plays a vital role in radiation therapy, improving treatment planning efficiency and tumor targeting accuracy. Preprocessing methods in medical image analysis systems are also essential, aiming to enhance the quality of input data and strengthen learning models. These methods include contrast enhancement, normalization, and image resizing, which help standardize image characteristics and reduce noise[5]. Preprocessing strategies have been successfully applied in numerous computer vision and medical imaging tasks. For instance, thermal imaging studies using convolutional neural networks have demonstrated the importance of noise reduction and image enhancement techniques in improving diagnostic accuracy [6]. In addition, understanding the physical properties is important for accurate data analysis. Several comparisons have recently been made between deterministic models of light propagation through biological tissues and Monte Carlo simulations [7]. This has provided a clear view of how energy interacts with heterogeneous structures, which is important for enhancing diagnostic reliability and improving imaging parameters.[8][9]. Feature extraction encompasses complex computer vision applications, such as omnidirectional imaging, and is not limited to the medical field. Comparative studies have also demonstrated the effectiveness of deep learning methods in detecting abnormalities in MRI images [10]. These results collectively underscore the pivotal role of physical modeling and preprocessing in improving the efficiency of deep learning across various imaging applications. Data augmentation techniques are used to address class imbalances and enhance data diversity. Methods such as rotation and reflection, along with advanced techniques like data augmentation during testing, have proven effective in reducing prediction uncertainty and improving model generalization [11] [9]. On the other hand, convolutional neural networks (CNNs) have demonstrated high efficiency in medical image interpretation due to their ability to extract features from raw data automatically [12][13]. The DenseNet201 model stands out among the advanced models. It connects each layer to all previous layers, facilitating feature reuse and improving information flow within the network. It also enhances training efficiency and model performance by reducing the number of parameters and mitigating the class imbalance caused by vanishing gradients [14].To enhance the reliability of results across heterogeneous MRI datasets, this work integrates advanced preprocessing with the DenseNet201 model within a unified framework, unlike previous studies.

- Contribution to the classification of brain tumors: Some previous studies have relied on deep learning architectures and various optimization methods to improve brain tumor classification. Our study contributes to the advancement of this field in several key ways. First, we proposed a method rely on DenseNet201, which establishes a new standard for classification accuracy across publicly available MRI datasets (FigShare, SARTAJ, and Br35H), demonstrating high generalizability across heterogeneous data sources and robustness. Second, we conduct a systematic comparative analysis with the latest models, involving ResNet50, ResNetRS50, VGG19, and EfficientNetV2, providing practical insights for selecting optimal models in medical imaging tasks. Finally, this work highlights the clinical applicability of the proposed method as a reliable decision-providing tool for the early detection of brain tumors, potentially improving diagnostic accuracy and patient outcomes.

- Study Organization: The research is structured as follows: Section 2 presents previous studies; Section 3 presents the methodology used; Section 4 contains the results and their discussion; and the final section, Section 5, presents the conclusions and a discussion of future work.

## 2 Literature Reviews

Accurate and primary diagnosis of tumor kinds is crucial for improving patient outcomes and increasing the chances of recovery. In recent years, artificial intelligence, particularly machine learning methods, has seen important advancements, revolutionizing clinical practice. A CNN model has been established to classify brain MRI images: no tumor, glioma, meningioma, and pituitary tumor. It was trained and validated utilizing a publicly accessible dataset including 980 images of healthy brains and 3,064 brain MRI scans. Also, they make a comparison of AlexNet, GoogleNet, DenseNet201, and ResNet101 architectures, and they note that the ResNet101 model reached the highest accuracy, at 98.76% [15]. The researchers developed a PSO-KELM model that is used to classify tumors into four categories: glioma, meningioma, non-glioma, and pituitary tumor. Two reference datasets were utilized for evaluating the system's performance, and it demonstrated its superiority over existing methods. The accuracy after applying the model to both datasets 1, 2 was 96.17% and 97.92%, respectively [16]. The researchers developed a U-Net model utilizing the ResNet50 network on the Figshare dataset, achieving a high IoU of 95.04%. They explored various CNN models for tumor classification, including MobileNet V2, Inception V3, ResNet50, DenseNet201, and NASNet, with accuracy rates of 91.8%, 92.8%, 92.9%, 93.1%, and 99.6%, respectively, with NASNet demonstrating the highest accuracy [3]. In another study, researchers proposed the MBINet model, containing eight layers utilizing a Self-ONN to classify Reconstructed Microwave Brain (RMB) images into six categories. Using an SMBI system with an antenna sensor to collect 1,320 real RMB images and applying pre-processing steps such as augmentation, rescaling, and normalization. The proposed model achieved high performance with an accuracy of up to 98%, representing its potential for tumor classification utilizing RMB images [17]. In a related study, researchers proposed a deep transfer learning model using five architectures (DenseNet121, Xception, InceptionResNetV2, ResNet152V2, and DenseNet201), modifying their final layers to enhance accuracy. The Xception-based model achieved 99.67% accuracy for three tumor types (pituitary, glioma, meningioma) and 95.87% for four classes, including healthy cases [14]. Researchers proposed an optimization of hyperparameters based on the Bayesian method for CNNs in a separate study. The enhanced CNN model was evaluated on 3064 T1-weighted CE-MRI images and outperformed several established architectures. Notably, the optimized CNN reached 98.70% of validation accuracy [18]. Researchers present a model trained on imbalanced data (BRATS2019 and BRATS2018) utilizing deep transfer learning. By utilizing the EKbHFV and MGA methods to optimize feature extraction, which are classified using a multi-class SVM classifier with a cubic kernel, an accuracy of 95% [19]. In this paper, researchers introduced a CNN model that was evaluated on benchmark data for classifying brain tumors. The model accuracy was 98% outperforming pre-trained models, including VGG16, VGG19, ResNet50, InceptionV3, and MobileNetV2 [12]. Researchers developed a computer-aided system. In this system, an algorithm for classifying objects that depends on deep learning networks was utilized with the policy of augmentation of Auto-Function ImageNet to improve performance. The EfficientNetV2-L (21 K) and ViT-L/16(21 K) models were used to evaluate the system, where the system achieved 94.13% accuracy with ViT-L/16 (21 K), while EfficientNetV2-L (21 K) achieved 95.90% [11]. One study introduced a technique to improve image retrieval in brain tumor analysis. The technique involves three steps: removing the brain from the skull, applying a threshold to divide the tumor, and rapidly guessing the tumor size utilizing MRI images [20]. Recent studies have applied DenseNet201 with transfer learning for Alzheimer's disease classification using MRI images across four classes. Resizing and normalization techniques were used to ensure data consistency, achieving high performance with 95% training accuracy and 92% validation accuracy. Evaluation metrics showed strong classification capability, especially in distinguishing between Mild and Moderate Dementia cases[21]. Also, recent studies have developed deep learning models for both binary and multiclass brain tumor classification using MRI datasets such as Figshare, Br35H, and Harvard Medical. The architecture of a deep CNN was initially applied, while transfer learning with fine-tuned VGG16 and Xception models was later used to improve performance and address overfitting. The results achieved high accuracy, reaching 100% on the Harvard Medical dataset, and up to 97–98% accuracy on larger datasets[13].

# 3 Methodology

## 3.1 Dataset

The brain tumor MRI dataset was compiled from three sources: Br35H, SARTAJ, and Figshare, which contained 7,023 images from Kaggle. The SARTAJ dataset had difficulty accurately classifying glioma images. Figshare provided a wide range of medical images. The Br35H dataset contributed additional "no tumor" images. Finally, it contained a dataset of four sets of images: "meningioma", "pituitary tumor", "glioma", and "no tumor" [22]. The data was saved in CSV format to facilitate efficient processing during training. The dataset was divided into two sets: a training set containing 5618 images (80%) and a test set containing 1404 images (20%). Preprocessing techniques were applied to the training dataset to improve image quality and ensure consistency. Then, data augmentation techniques were applied, producing an expanded dataset of 6384 images. This expanded dataset was then divided into two sets: a training set containing 5108 images and a validation set containing 1276 images. This division was used to get better model performance and prevent over-customization during training.

## 3.2 Pre-processing

The brain tumor MRI dataset was preprocessed to improve its quality for analysis and ease of use. Initially, the image sizes were unequal, posing a challenge to accurate analysis [5]. The dataset included images with 382 varying dimensions. The average height is 453.88 pixels ( max 1446, min 168), and the average width is 451.56 pixels. To standardize the data and avoid information loss, all images were resized to predefined dimensions using interpolation, which preserved image quality by estimating pixel values from neighboring pixels. Eq. (1) describes interpolation-based image scaling in both horizontal and vertical directions. The scaled image ($I'$) was obtained by applying horizontal and vertical scaling to the original image ($I$) with scaling factors ($rx$) and ($ry$), respectively. The process can be performed sequentially in either order, such that

$$I' = S(I, rx, ry) = Sy(Sx(I, rx), ry) = Sx(Sy(I, ry), rx) \tag{1}$$

Here, ($Sx$) and ($Sy$) denote the scaling operations along the horizontal and vertical axes, respectively. Also, to enhance image quality for analysis, contrast and brightness were improved utilizing functions of the OpenCV library [23] [24]. For each image, brightness and contrast have been increased by 20, as in Eq. (2).

$$dst(x, y) = \text{saturate}_{\text{cast}} < uchar > (\alpha \cdot src(x, y) + \beta) \tag{2}$$

Where $dst(x, y)$ represents the pixel value in the processed image, and $src(x, y)$ represents the corresponding pixel value in the original image. The parameter $\alpha$ controls image brightness (default $\alpha = 1$, indicating no change), while $\beta$ controls image contrast (default $\beta = 0$, indicating no change). The imbalanced class distribution issue was tackled through data augmentation methods such as flipping and rotation, guaranteeing sufficient illustration for all tumor groups [5]. Rotation is expressed mathematically in Eq. (3) as:

$$R_\theta(I) = R(I, \theta) \tag{3}$$

Where $R_\theta(I)$ denotes the image rotated by an angle $\theta$, $I$ represents the original image, and $\theta$ is the rotation angle applied to the image. In this work, $\theta$ varies from $-10^\circ$ to $+10^\circ$ with a step size of $5^\circ$. The image flipping operations in the horizontal and vertical directions can be expressed by Eqs. (4) (5).

$$F_h(I) = F(I, h) \tag{4}$$

$$F_v(I) = F(I, v) \tag{5}$$

Where $F_h(I)$ represented the image flipped horizontally, $F_v(I)$ represented the image flipped vertically, $I$ is the original image, $h$ denoted horizontal flipping, and $v$ denoted vertical flipping.

To further develop image quality and organize for model training, images were normalized to a 0-1 range utilizing Eq. (6):

$$normalized_image = \frac{image}{255.0} \tag{6}$$

The parameters, like contrast difference and peak signal-to-noise ratio (PSNR), are used to evaluate the effectiveness of the pre-processing steps. The difference between the enhanced and original images of a pituitary tumor is visually demonstrated in Figure 1.

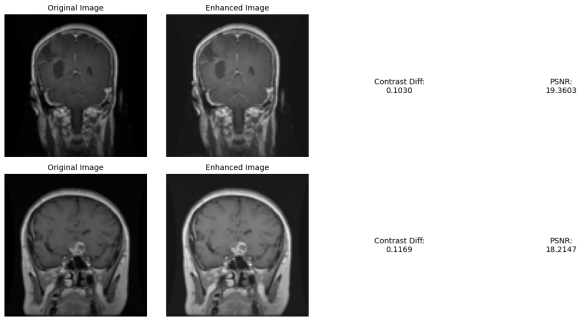



Figure 1: (a, c), Before enhancement processing and (b, d) after it.

The class imbalance in the dataset before the augmentation process is shown in Figure 2. The distribution of images over categories after augmentation is shown in Figure 3, highlighting the improved class balance achieved through this technique.

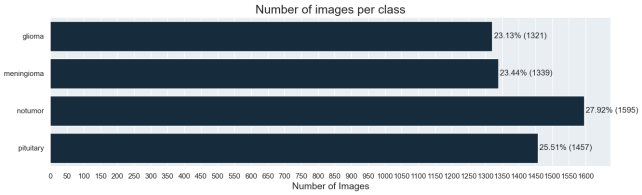



Figure 2: Number of images per class before augmentation.

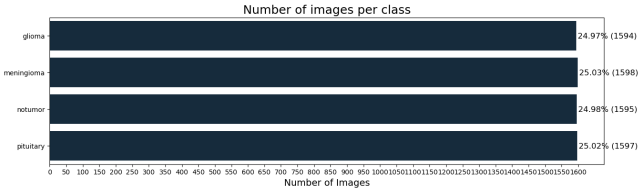



Figure 3: Number of images per class after augmentation.

A feed-forward network architecture has dense connections between layers, resulting in accuracy comparable to ResNet while using fewer parameters. DenseNet-201, a 201-layer CNN, was particularly good at recognizing images. It employed sequential concatenation rather than summing the output features, inspired by ResNet's skip connections [25]. In the architecture, thick squares (dense blocks) were separated by transition layers that contained 1×1 convolution, batch normalization, and 2×2 average pooling for down-sampling. The proposed model introduced a custom CNN architecture based on DenseNet201. It involved additional layers such as Dense layers and GlobalAveragePooling2D. GlobalAveragePooling2D aggregates feature maps from the final output layer by computing the global average. The Dense layer included dot product computation

with weights, addition of bias, and application of an activation for nonlinearity. A dropout layer was used to prevent overfitting by randomly setting input units to zero during training. The output layer uses a softmax activation, producing probability distributions for classification. During training, weights are updated using an optimizer with a base learning rate of 0.0001, which is dynamically modified via the ReduceLROnPlateau callback. The categorical cross-entropy loss function was evaluated by comparing the predicted and true outputs difference, while accuracy evaluated model performance. Thoroughly, the model combined DenseNet201 with custom layers, followed by an output layer configuration and softmax classification, optimizing convergence and accuracy during training. It was important to note that the data preprocessing pipeline shown in Figure 4 was not limited to the proposed model, but was applied to all deep learning models used in this study, such as DenseNet201, ResNet50, VGG16, EfficientNetV2, and ResNetRS50 uniformly. This ensured an unbiased and fair comparison between the models under the processing conditions and the same input.

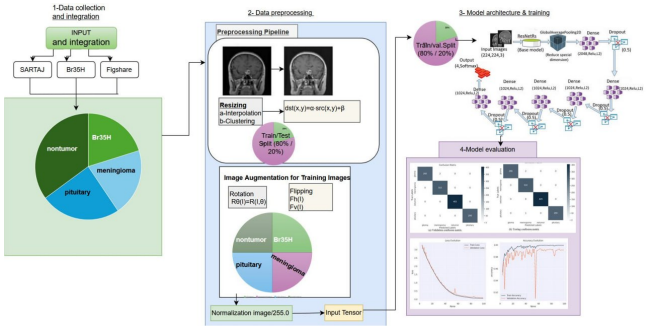



Figure 4: General framework for data processing and model training.

The fact that CNN models operate independently of specific datasets allowed for their wider application across diverse data sources. [26]. When you train a new model, transfer learning keeps the settings of a model that has already been trained. This method makes the process much more efficient by utilizing what is already known to cut down on the time and money needed to train a model from scratch. [25]. The DenseNet201 model was built using weights that had already been trained on the ImageNet dataset (weights='imagenet'). The top layer, which had several connected layers, was not included (include_top=False). The model's inputs were (224, 224, 3). To solve the specific classification problem, this model was given more layers. It was helpful to use patterns that were gotten from the model's lower layers to build our own instead of starting from scratch with a deep neural network. In this research, we concentrated on employing a function explicitly tailored to the DenseNet201 architecture. CNNs have demonstrated exceptional efficacy in image classification owing to their capacity to learn hierarchical features from unprocessed pixel data. We set up our method to train the model by using customizable limits that were specific to the tasks and datasets. A pre-compiled Keras model was used for the task version, and it had a pre-trained DenseNet201 base with the traditional layers. Training and validation data were organized as Pandas data frames, which included information about the input images and labels. To make sure that image management worked well, a pre-processing task was done, such as resizing and straightening. The input images are 224 x 224 pixels in size for DenseNet201 performance and compatibility, and there are 64 of them in a batch. The model goes through the entire dataset 100 times during training, which is what the number of epochs means. Data generators prepare groups of training and validation data ahead of time. Callback functions like "reduce_lr" and "EarlyStopping" keep an eye on and control training, making it more efficient and stopping overfitting. The three main things that were returned were a log object with metrics, information about the validation data generator, and the trained model with adjusted weights and learned features. The "train" function gave a complete framework for training deep learning models for image classification tasks, allowing for efficient training that was specific to each dataset and goal. The model reached an accuracy of 99.22% on validation data after 100 epochs of training, with a loss of 0.0970. These results show that the DenseNet201 model can improve image classification performance, which gives ideas for further

testing and improvement.

### 3.3 Ablation Study of the Proposed Framework

An ablation study was conducted to investigate the individual and combined contributions of the proposed strategies to the performance of DenseNet201. To ensure a fair comparison, the model architecture, hyperparameters, and training settings were kept identical across all experiments, while only the proposed strategies were varied. The evaluated strategies included: (a1) image pre-processing and enhancement, (a2) data augmentation, and (a3) transfer learning with regularization techniques. The regularization techniques consisted of Dropout, EarlyStopping, and ReduceLROnPlateau. The effectiveness of these strategies was assessed by comparing the model performance under different experimental configurations.

### 3.4 Alternative Architectural Approaches

The DenseNet201 model was developed in a structured way and is always used in the same way, demonstrating that the methodology is careful and reliable. Some important parts of this method are careful dataset selection, precise pre-processing methods, and careful tuning of extra layers. The additional layers' proven effectiveness shows how important it was to carefully plan and integrate deep learning architectures. DenseNet201, VGG19, EfficientNetV2, ResNet50, and ResNetRS50 were all made using the same systematic method. The selection of datasets and pre-processing techniques, along with the number of supplementary layers, remained uniform across all models. Figure 5 shows how the model is built. It has an input layer for image acquisition, ResNetRS50, and custom layers added to DenseNet201 to improve pattern recognition. The unified output layer for all models is responsible for classifying the images.

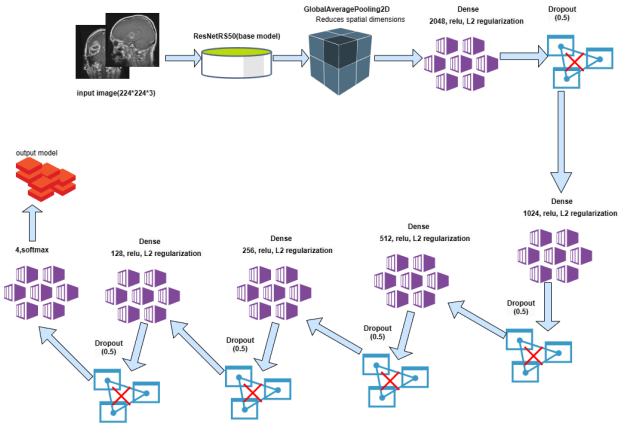



Figure 5: Common transfer learning architecture used in this study.

## 4 Result

In this section, the model's performance was evaluated using metrics such as accuracy scores, loss functions, and other metrics associated with validation and test data. Confusion matrices are used to determine classification behavior, providing deeper insights, revealing misclassification patterns, and potential biases. A learning rate of 0.001 was adopted with the Adam optimizer to achieve a balance between convergence speed and training stability. Additionally, a ReduceLROnPlateau strategy with a factor of 0.3 and a patience of 2 was implemented, allowing the model to automatically decrease the learning rate when performance stagnates, which justifies the chosen initial value. A dropout rate of 0.5 was applied in the final layer to reduce overfitting by preventing reliance on specific neurons and encouraging the model to learn more generalized tumor features.

Furthermore, the number of training epochs was controlled using an EarlyStopping mechanism with a patience of 5 based on validation loss, ensuring that training stops at the optimal point and avoids overfitting. This section also includes a comparative analysis of the performance of pre-trained models such as ResNet RS50, ResNet50, EfficientNetV2, and DenseNet201.

### 4.1 DenseNet201Training Performance

A vital illustration of how loss and accuracy measures interact during the training of a brain tumor classification model (DenseNet201) was provided in Figure 6. It includes two crucial graphics: "Loss" and "Accuracy". A comprehensive analysis of the two graphs provides deep insight into the behavior of the model during training, its generalizability, and the areas where performance can be improved. Figure 7 illustrates two confusion matrices that represent how well the DenseNet201 models can sort different types of brain tumors on test and validation sets that they haven't seen before. Each matrix indicates how well the model's predictions match the real labels. The rows show the real classes ( pituitary, glioma, non-tumor, and meningioma), and the columns show the predicted classes. Precision Eq. (7) [27], which measures the percentage of correctly classified examples inside each anticipated category, was a crucial parameter that was emphasized by the confusion matrices. High precision meant that the model was likely to be correct when it predicted a certain type of tumor.

$$precision = \frac{TruePositives}{TruePositives + FalsePositives} \tag{7}$$

While Recall, illustrated by Eq. (8) [27], indicated the identification of actual tumors in each class.

$$Recall = \frac{TruePositives}{TruePositives + FalseNegative} \tag{8}$$

Eq. (9) [27] represents the F1-score.

$$F1_score = \frac{2 * (Positives * Recall)}{Positives + Recalls} \tag{9}$$

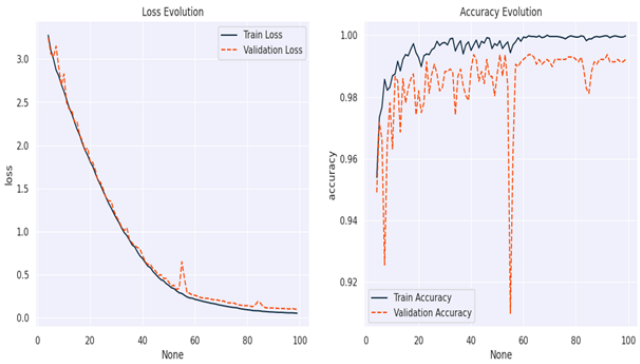



Figure 6: Loss and accuracy for DenseNet201 model.

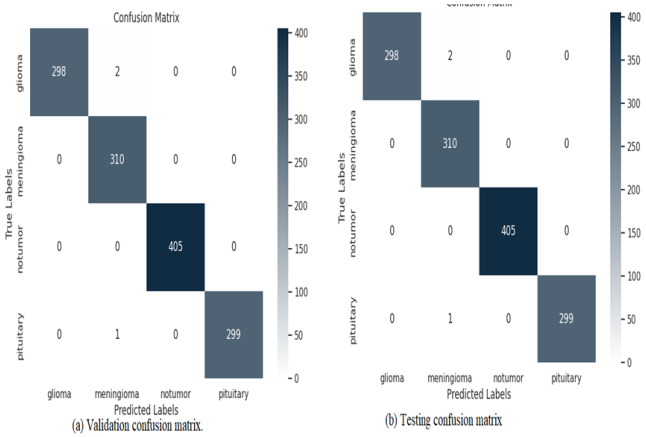



Figure 7: Validation and testing confusion matrices for DenseNet201.

According to the validation confusion matrix analysis, DenseNet201 did remarkably well. The non-tumor class had an F1-score of 99.84%, a recall of 99.69%, and flawless precision (100%). It achieved 99.53%, 99.38%, and 99.69% for F1-score, precision, and recall, respectively, for gliomas. The model showed outstanding performance in pituitary tumor classification, achieving 99.06% for recall, F1-score, and precision. In meningioma classification, the recall, F1-score, and precision are reaching 98.44%. Whole validation accuracy was 99.22%, reflecting consistent results across categories with balanced macro and weighted averages. The testing confusion matrix confirmed the model's reliability. The model also performed very well across all tumor types in the test set, with almost perfect metrics. Recall, F1-score, and precision for the classifications of glioma, pituitary, non-tumor, and meningioma were all greater than 99%, and some were even 100%. The model's consistent and excellent ability to classify was shown by both the macro and weighted averages, which showed that the overall accuracy was 99.77%. The results of the confusion matrix show that DenseNet201 is very good at accurately classifying different types of brain tumors in unseen data.

## 4.2 ResNet50 Training Performance

Figure 8 shows how the ResNet50 model learns over time. It has two important subplots that show how its training has changed. The fact that the validation loss dropped while the accuracy rose indicates that the model is learning important brain tumor features, which shows that it can generalize well. The representation of validation loss (loss on unseen data) also decreases. This simultaneous decrease in validation and training losses suggests that the model was not simply memorizing the training data (overfitting), but acquiring generalizable features applicable to novel, unseen instances.

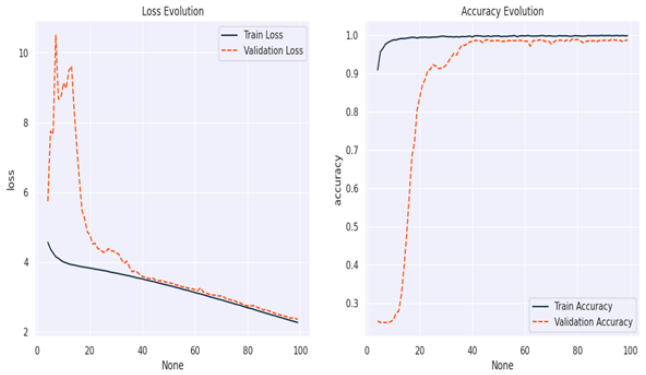



Figure 8: Loss and accuracy for ResNet50 model.

Figure 9 shows the validation and testing confusion matrices, highlighting ResNet50's efficiency.

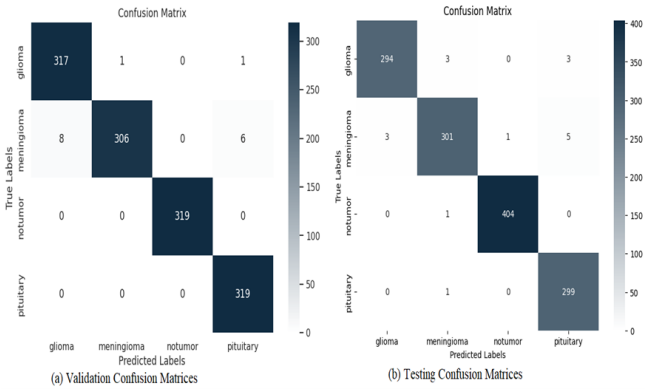



Figure 9: Validation and testing confusion matrices for ResNet50.

## 4.3 ResNetRS50 Training Performance

The progression of the ResNetRS50 model in brain tumor classification is shown in Figure 10, revealing two key viewpoints: Loss and Accuracy evolution.

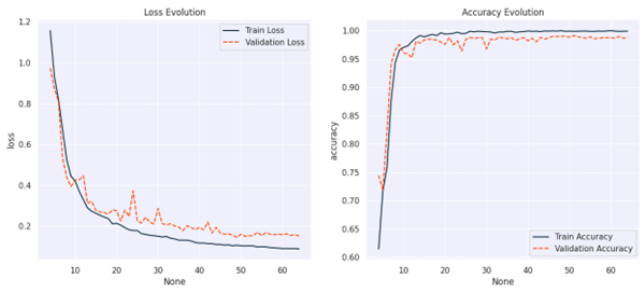



Figure 10: Loss and accuracy for ResNetRS50 model.

Figure 11 illustrates the validation and testing confusion matrices. These matrices highlight the model's strengths while also revealing potential areas for improvement.

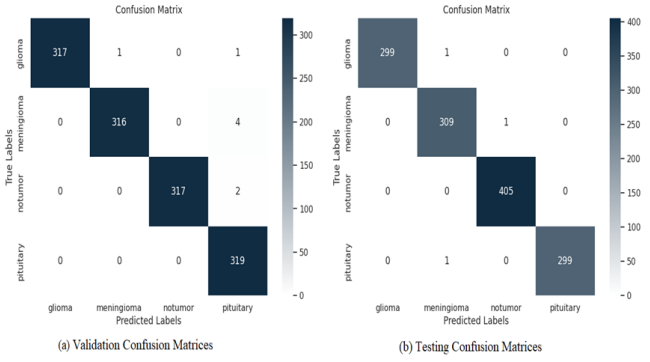



Figure 11: Validation and testing confusion matrices for ResNetRS50.

### 4.4 VGG19 Training Performance

Two important graphs were shown in Figure 12, which provide information about how the VGG-19 model for brain tumor classification is trained. These graphs show how accuracy and loss metrics change across training epochs, or iterations.

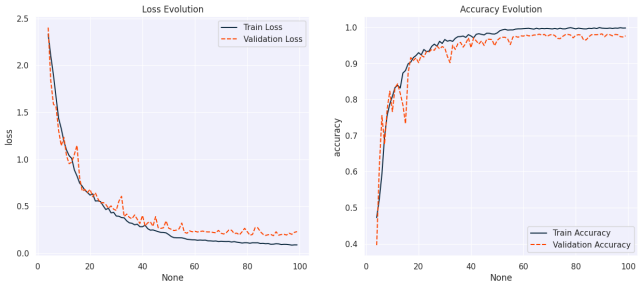



Figure 12: Loss and accuracy for VGG19 model.

Figure 13 introduces the VGG19 diagnostic summary of its performance on the same dataset, emphasizing both its advantages and disadvantages through the confusion matrices for both testing and validation.

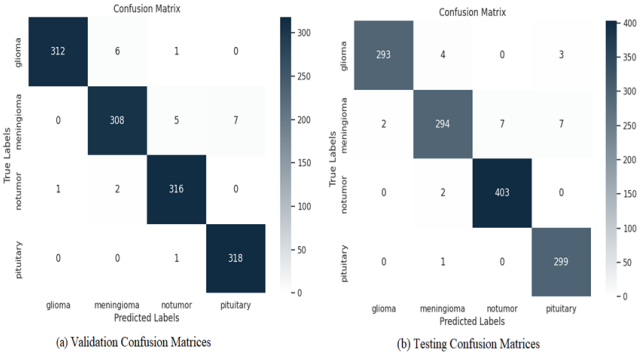



Figure 13: Validation and testing confusion matrices for VGG19.

### 4.5 Training Performance for EfficientNetV2

Figure 14 presents two line graphs explaining the training progression of the EfficientNetV2 model, showing variations in loss and accuracy across training epochs.

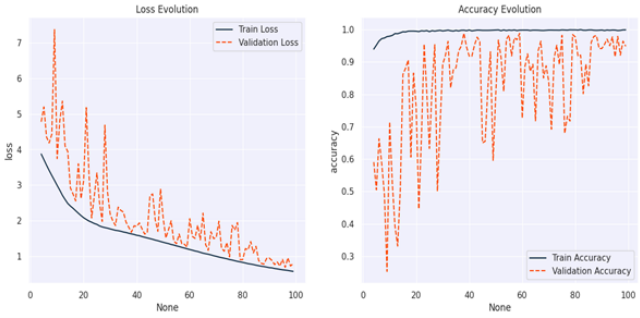



Figure 14: Loss and accuracy for EfficientNetV2 model.

Using Validation and Testing Confusion Matrices, Figure 15 presents EfficientNetV2's comprehensive evaluation of its performance on unseen data, emphasizing both its advantages and disadvantages through confusion matrices for both testing and validation.

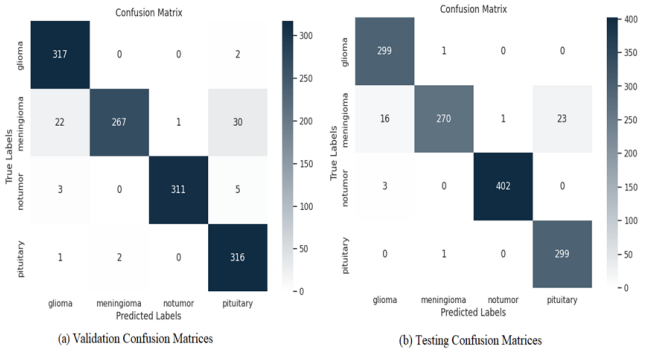



Figure 15: Validation and testing confusion matrices for EfficientNetV2.

## 4.6 Models Performance Comparison

A performance of all advanced deep learning models in brain tumor classification is evaluated in this study. The generation of a complete MRI dataset by combining images from the SARTAJ, Br35H, and Figshare datasets ensured the accuracy of the evaluation and addressed the data quality limitations identified in previous studies. Initially, DenseNet201 was trained using the original MRI images without any preprocessing or data augmentation. Under these conditions, the model achieved a maximum accuracy of 95.13% and exhibited noticeable oscillations in both the training and validation curves, indicating unstable learning behavior and a tendency toward overfitting. After incorporating the proposed framework, which combines image pre-processing and enhancement, data augmentation, and transfer learning with regularization techniques, the model performance improved substantially, achieving an accuracy of 99.77%. Furthermore, the training process became more stable, with reduced fluctuations and a smaller gap between the training and validation curves. These results demonstrate the effectiveness of the proposed strategies in improving the model's generalization capability while reducing overfitting. Table 1 shows the stability and improvement of training after applying the proposed framework.

| **Configuration** | **Accuracy (%)** | **Training Stability** |
|---|---|---|
| DenseNet201 without preprocessing and data augmentation | 95.13 | Unstable (oscillating training and validation curves) |
| DenseNet201 with the proposed framework (a1+a2+a3) | 99.77 | Stable |
| Performance Improvement | +4.64 | Improved |

Table 1: Ablation study of the proposed framework using DenseNet201

Table 2 presents the relative performance analysis of the models, while Table 3 presents the relative performance analysis of the validated models. The DenseNet201 and ResNetRS50 models achieved the highest performance and accuracy. Two significant outcomes are obtained from the results:

- The organization of the dataset influences: For a model to perform well, the dataset must be of high quality, especially in the tasks of medical imaging.
- Model selection significance: When it comes to brain tumor identification, the selection of a deep learning model has a great impact on the accuracy of the classification.

A comparative summary of several previous studies on brain tumor classification, methodologies, highlighting datasets, performance metrics, and deep learning models, is illustrated in Table 4. This comparison provides context for evaluating the improvements and contributions of our work. ResNetRS50 reached comparable accuracy. DenseNet201 was selected as the reference model because of its dense feature reuse mechanism and its consistently competitive performance throughout the experiments.

| Model | Precision | Recall |
|---|---|---|
| ResNetRS50 | 0.9978 | 0.9975 |
| DenseNet201 | 0.9976 | 0.9975 |
| ResNet50 | 0.9863 | 0.9863 |
| VGG19 | 0.9801 | 0.9792 |
| EfficientNetV2 | 0.9647 | 0.9642 |

Table 2: Comparison of model performance metrics in Testing

| Model | Precision | Recall | F1_score |
|---|---|---|---|
| ResNetRS50 | 99.38 | 99.37 | 99.38 |
| DenseNet201 | 99.22 | 99.22 | 99.22 |
| ResNet50 | 98.77 | 98.75 | 98.74 |
| VGG19 | 97.67 | 97.65 | 97.65 |
| EfficientNetV2 | 95.22 | 94.84 | 94.76 |

Table 3: Comparison of Model Performance Metrics in Validation

| References | Model | Dataset | Accuracy |
|---|---|---|---|
| [28] | DenseNet201, ResNet50 | Figshare, SARTAJ, Br35H | 96.8%,96.3% |
| [29] | DenseNet201+SVM RBF | Brain MRI, Segmentation-LGG | 93.10% |
| [30] | ResNet50 | CBT Dataset(D2) | 75.63% |
| [14] | DenseNet201+DDB | Unspecified MRI | 97.06% |
| [31] | VGG19 (CNN) | Figshare MRI | 96% |

Table 4: Comparison of Model Results in Other Works

## 5 Conclusion

This study introduces a beneficial method of classification for brain tumours using DenseNet201, reaching over 99% accuracy on an individualistic test set. It outperforms models like VGG19, EfficientNetV2, and ResNet50, displaying DenseNet201's adaptability and the deep learning transformational possibilities in brain tumor diagnosis. As maintained by our research, clinicians may benefit from AI-powered systems' exceptional accuracy and dependability, which could result in earlier diagnosis, better patient outcomes, and more individualized treatment plans. For individuals with brain tumors and their families, continuous research and interdisciplinary collaboration will usher in a new era of enhanced diagnostic precision and hopefulness. Two key ideas for creating precise medical image analysis systems are highlighted in this study:

- Model Significance: For categorization to be successful, the architecture of deep learning must be carefully chosen. The remarkable potential of DenseNet 201 for brain tumor classification is demonstrated by its excellent performance in this study.

- Dataset Excellence: The carefully selected dataset, which contains images from FigShare, SARTAJ, and Br35H, was crucial to obtaining the best model performance. This highlights the importance of having large, high-quality datasets to develop reliable deep learning models for the medical field.

This study establishes a new benchmark relative to previous research and markedly enhances the classification of brain tumors. By comparing the results with those of other deep learning models like EfficientNetV2, VGG19, ResNet50, and ResNetRS50, we have learned important things that will help us choose the best models for future research. In addition, this work shows how important dataset quality and model selection are for getting high classification accuracy and better generalization in tumor diagnosis. Another contribution of this work is the unified framework evaluated on a combined multi-source MRI dataset and systematically compared to several state-of-the-art deep learning architectures under the same experimental conditions. This comprehensive evaluation provides a reliable assessment of model performance and helps in selecting appropriate architectures for brain tumor classification tasks. This study has a number of problems, even though the results are promising. The model is primarily trained and evaluated on publicly available datasets, which may not fully capture the variability and complexity of real-world clinical data, potentially constraining its generalizability in practical healthcare contexts. Also, the fact that hospitals have different imaging protocols and patient demographics was not taken into account. In the future, work on making sure that the model works with real clinical datasets and with medical institutions to ensure it works in real life. Additionally, making the model easier for healthcare professionals to use by integrating it into hospital workflows and decision-support systems.

## References

[1] S.A. Abdelaziz Ismael, A. Mohammed, H. Hefny, "An enhanced deep learning approach for brain cancer MRI images classification using residual networks", *Artificial Intelligence in Medicine* 102:101779, 2020. https://doi.org/10.1016/j.artmed.2019.101779.

[2] A. Tiwari, S. Srivastava, M. Pant, "Brain tumor segmentation and classification from magnetic resonance images: Review of selected methods from 2014 to 2019", *Pattern Recognition Letters* 131:244-260, 2020. https://doi.org/10.1016/j.patrec.2019.11.020.

[3] T. Sadad et al., "Brain tumor detection and multi-classification using advanced deep learning techniques", *Microscopy Research and Technique* 84(6):1296–1308, 2021. https://doi.org/10.1002/jemt.23688.

[4] M. Ottoni, A. Kasperczuk, and L. M. N. Tavora, "Machine Learning in MRI Brain Imaging: A Review of Methods, Challenges, and Future Directions", *Diagnostics* 15(21):2692, 2025. https://doi.org/10.3390/diagnostics15212692.

[5] N. Jiang and L. Wang, "Quantum image scaling using nearest neighbor interpolation", *Quantum Information Processing* 14(5):1559–1571, 2015. https://doi.org/10.1007/s11128-014-0841-8.

[6] R. Thukral, A. S. Arora, A. K. Aggarwal, T. Dora, and S. Sancheti, "Effectiveness of Passive Thermography in Predicting Mucositis in Head and Neck Cancer Patients Using Light-Weight Convolutional Neural Networks Approach", *Proceedings of the International Conference on Frontiers of Intelligent Computing: Theory and Applications* pp. 1–13, 2024. https://doi.org/10.1007/978-981-96-2124-8.

[7] S. S. Roy, A. S. Arora, and R. Thukral, "Light propagation through biological tissue: comparison of Monte Carlo simulation with deterministic models", *Biophysical Reviews* 17(1), 2025. https://doi.org/10.1007/s12551-025-01403-w.

[8] N. Netshamutshedzi, R. Netshikweta, J. C. Ndogmo, and I. C. Obagbuwa, "A systematic review of the hybrid machine learning models for brain tumour segmentation and detection in medical images", *Frontiers in Artificial Intelligence* 8:1615550, 2025. https://doi.org/10.3389/frai.2025.1615550.

[9] N. H. Lu, Y. H. Huang, K. Y. Liu, and T. B. Chen, "Deep learning-driven brain tumor classification and segmentation using non-contrast MRI", *Scientific Reports* 15(1):27831, 2025. https://doi.org/10.1038/s41598-025-13591-2.

[10] J. Jeong, S. Bang, Y. Jung, and J. Jo, "A Study on the Performance Comparison of Brain MRI Image-Based Abnormality Classification Models", *Life* 15(10):1614, 2025. https://doi.org/10.3390/life15101614.

[11] J. hyeon Lee, J. woo Chae, and H. chong Cho, "Improved Classification of Brain-Tumor MRI Images Through Data Augmentation and Filter Application", *Journal of Electrical Engineering and Technology* 18(4):3135–3142, 2023. https://doi.org/10.1007/s42835-023-01542-8.

[12] Z. Rasheed et al., "Automated Classification of Brain Tumors from Magnetic Resonance Imaging Using Deep Learning", *Brain Sciences* 13(4), 2023. https://doi.org/10.3390/brainsci13040602.

[13] H. Sadr, M. Nazari, S. Yousefzadeh-Chabok, H. Emami, R. Rabiei, and A. Ashraf, "Enhancing brain tumor classification in MRI images: a deep learning-based approach for accurate diagnosis", *Image and Vision Computing* 159:105555, 2025. https://doi.org10.1016/j.imavis.2025.105555.

[14] S. Asif, M. Zhao, F. Tang, and Y. Zhu, "An enhanced deep learning method for multi-class brain tumor classification using deep transfer learning", *Multimedia Tools and Applications* 82(20):31709–31736, 2023. https://doi.org/10.1007/s11042-023-14828-w.

[15] M.C. Daniel, L.M. Ruxandra, "Brain Tumor Classification Using Pretrained Convolutional Neural Networks", *Proc. of 2021 Intl. Conf. on Engineering of Modern Electric Systems (EMES)*, 2021. https://doi.org/10.1109/EMES52337.2021.9484102.

[16] B. Sandhiya, S. Kanaga Suba Raja, "Deep Learning and Optimized Learning Machine for Brain Tumor Classification", *Biomedical Signal Processing and Control* 89:105778, 2024. https://doi.org/10.1016/j.bspc.2023.105778.

[17] A. Hossain et al., "A Lightweight Deep Learning-Based Microwave Brain Image Network Model for Brain Tumor Classification Using Reconstructed Microwave Brain (RMB) Images", *Biosensors* 13(2), 2023. https://doi.org/10.3390/bios13020238.

[18] M. A. Amou, K. Xia, S. Kamhi, and M. Mouhafid, "A Novel MRI Diagnosis Method for Brain Tumor Classification Based on CNN and Bayesian Optimization", *Healthcare* 10(3), 2022. https://doi.org/10.3390/healthcare10030494.

[19] M. I. Sharif, M. A. Khan, M. Alhussein, K. Aurangzeb, and M. Raza, "A decision support system for multimodal brain tumor classification using deep learning", *Complex and Intelligent Systems* 8(4):3007–3020, 2022. https://doi.org/10.1007/s40747-021-00321-0.

[20] A.H. Ali et al., "Provably Efficient and Fast Technique for Determining the Size of a Brain Tumor in T1 MRI Images", *Proc. of 2024 Intl. Conf. on Artificial Intelligence in Information and Communication (ICAIIC)*:608-613, 2024. https://doi.org/10.1109/ICAIIC60209.2024.10463420.

[21] S. Singla and R. Gupta, "Deep Learning-Based Alzheimer's Disease Classification Using MRI Images and DenseNet201 Architecture", *Proceedings of the 2025 International Conference on Next Generation Communication & Information Processing (INCIP)* pp. 801–806, 2025. https://doi.org10.1109/INCIP64058.2025.11019380.

[22] "Brain Tumor MRI Dataset", *Kaggle*, Accessed: Dec. 08, 2024. [Online]. Available: https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset

[23] T. Moraes, P. Amorim, J. V. Da Silva, and H. Pedrini, "Medical image interpolation based on 3D Lanczos filtering", *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization* 8(3):294–300, 2020. https://doi.org/10.1080/21681163.2019.1683469.

[24] A. Kaur and G. Kaur, "A review on image enhancement with deep learning approach", *ACCENTS Transactions on Image Processing and Computer Vision* 4(11):16–20, 2018. https://doi.org/10.19101/TIPCV.2018.411002.

[25] M. A. Khan et al., "Computer-Aided Gastrointestinal Diseases Analysis from Wireless Capsule Endoscopy: A Framework of Best Features Selection", *IEEE Access* 8:132850–132859, 2020.https://doi.org/10.1109/ACCESS.2020.3010448.

[26] S. Das and R. S. Goswami, "Advancements in brain tumor analysis: a comprehensive review of machine learning, hybrid deep learning, and transfer learning approaches for MRI-based classification and segmentation", *Multimedia Tools and Applications* 84(23):26645–26682, 2025. https://doi.org/10.1007/s11042-024-20203-0.

[27] S. Safwat, A. Mahmoud, I. Eldesouky Fattoh, and F. Ali, "Hybrid Deep Learning Model Based on GAN and RESNET for Detecting Fake Faces", *IEEE Access* 12:86391–86402, 2024. https://doi.org/10.1109/ACCESS.2024.3416910.

[28] M. Celik and O. Inik, "Development of hybrid models based on deep learning and optimized machine learning algorithms for brain tumor multi-classification", *Expert Systems with Applications* 238:122159, 2024. https://doi.org/10.1016/j.eswa.2023.122159.

[29] J. J. Da Costa Nascimento et al., "New Health of Things Approach to Classification and Detection of Brain Tumors Using Transfer Learning for Segmentation in MRI Images", *Proceedings of the International Joint Conference on Neural Networks (IJCNN)*, 2023. https://doi.org/10.1109/IJCNN54540.2023.10191399.

[30] F. Alam, F. C. Tisha, S. A. Rahman, S. Sultana, M. A. M. Chowdhury, W. R. Ahmed, and M. S. Arefin, "Automated brain disease classification using transfer learning based deep learning models", *International Journal of Advanced Computer Science and Applications* 13(9), 2022. https://doi.org/10.14569/IJACSA.2022.01309109.

[31] Y. P. Singh and D. K. Lobiyal, "A comparative analysis and classification of cancerous brain tumors detection based on classical machine learning and deep transfer learning models", *Multimedia Tools and Applications* 83(13):39537–39562, 2024. https://doi.org/10.1007/s11042-023-16637-7.