IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 8 March 2025, accepted 1 April 2025, date of publication 7 April 2025, date of current version 17 April 2025.

Digital Object Identifier 10.1109/ACCESS.2025.3558411

RESEARCH ARTICLE

Hybrid 3B Net and EfficientNetB2 Model for Multi-Class Brain Tumor Classification

R. PREETHA, M. JASMINE PEMEENA PRIYADARSINI, AND J. S. NISHA
School of Electronics Engineering, Vellore Institute of Technology, Vellore, Tamil Nadu 632014, India
Corresponding author: M. Jasmine Pemeena Priyadarsini (jasmin@vit.ac.in)

ABSTRACT Brain tumor classification and diagnosis are critical for timely and effective treatment, as brain tumors can severely impact patient health and survival. This paper introduces an enhanced method for the multiclass classification of brain tumors using a novel three-branch convolutional neural network integrated with EfficientNetB2 feature fusion. We systematically compared multi-branch architectures, ranging from single to six branches, to identify the optimal configuration for accurate tumor classification. The three-branch network (3B Net) consistently yielded superior performance across multiple classification tasks. Our experiments, conducted across three distinct datasets, involved binary classification (tumor vs. normal), three-class classification (meningioma, glioma, and pituitary), and four-class classification (non-tumor, meningioma, glioma, and pituitary). The proposed method achieved a binary classification accuracy of 99.50%, a three-class classification accuracy of 94.95% without augmentation, 98.72% with augmentation, and a four-class classification accuracy of 97.80%. Applying data augmentation techniques to the three-class classification dataset led to a marked improvement in classification accuracy. Blind test evaluations confirmed the robustness of the model for both binary and four-class classification, while five-fold cross-validation for the augmented three-class classification dataset further validated its performance. The 3B Net demonstrated the highest accuracy and robustness, significantly outperforming other network configurations. The integration of EfficientNetB2 for feature extraction further enhanced the model's classification capabilities. These results underline the effectiveness of our 3B Net architecture, suggesting its potential for deployment in automated brain tumor diagnosis and improvement in clinical decision-making processes.

INDEX TERMS Brain tumor, convolutional neural network (CNN), deep learning (DL), deep neural network (DNN), EfficientNet, magnetic resonance imaging (MRI), multi-class classification.

I. INTRODUCTION

The brain, a vital part of our central nervous system, regulates bodily functions through an intricate network of neurons. Any disruption or abnormal growth in brain cells can impact the connected organs, potentially leading to severe functional impairments [1]. Brain tumors, characterized by the uncontrolled proliferation of brain cells, are a major health concern. American Cancer Society estimates that approximately 25,400 malignant brain and spinal cord tumors will be diagnosed in the United States, with 14,420 cases in males and 10,980 in females in 2023 [2]. This statistic would be significantly higher if benign tumors were included. Additionally, around 18,760 people are expected to succumb to these malignancies, with 10,690 males and 8,070 females affected. These tumors rank as the 10th leading cause of cancer-related deaths in the U.S. [3].

Early and accurate detection of brain tumors is crucial for improving patient survival rates and expediting treatment [4]. However, manual detection is often laborious, time-consuming, and prone to errors, given the diversity in tumor types and sizes. Expert intervention is essential, yet even specialists can find complex cases challenging. This underscores the need for automated systems to assist in the precise detection and classification of brain tumors [5].

The associate editor coordinating the review of this manuscript and approving it for publication was Sandra Costanzo.

Medical imaging plays a vital role in the diagnosis of brain tumors. Magnetic resonance imaging (MRI) and Computed Tomography (CT) scans are crucial for detecting and assessing abnormal tissue in the brain, facilitating prompt follow-up and treatment decisions [6]. MRI is often favored over CT scans because it employs radio waves and a powerful magnetic field to generate images, as opposed to the radiation, such as X-rays, used in CT scans. This approach avoids the harmful radiation associated with CT scans. MRI produces highly detailed images of the body's internal structures, particularly soft tissues. It generates various types of high-contrast grayscale images, including FLAIR, Proton Density (PD), T2, T1 contrast-enhanced, and T1-weighted images, each providing unique insights into different tissue properties and abnormalities.

Recent advancements in deep learning(DL) and convolutional neural networks (CNNs) have shown significant promise in automating medical image analysis [7]. These technologies have revolutionized various fields, including medical image processing, by providing tools for efficient and accurate diagnosis. The automation of brain tumor segmentation [8] and classification is a critical area of research, aiming to enhance diagnostic accuracy and streamline the treatment process. Numerous studies have explored various DL techniques, each presenting unique strengths and limitations.

Rastogi et al. [9] developed a multi-branch network incorporating Inception blocks and five-fold cross-validation, demonstrating impressive accuracy for multi-class brain tumor classification. Despite advancements in diagnostic accuracy and model robustness, the research highlighted limitations in image quality, demographic variability, and interpretability, suggesting a need for more consistent performance across diverse real-world scenarios. The research paper [10] provides a comprehensive comparative analysis of ten DL models to reveal the superior performance of Inception models over EfficientNet models on an unbalanced MRI dataset. However, the study's reliance on a single unbalanced dataset and lack of cross-dataset evaluation limited its generalizability to diverse clinical settings. Sandhiya and Raja [11] present a novel hybrid optimization technique combining DL architectures (InceptionV3 and DenseNet201) with radiomic features and a PSO-KELM classifier, this study achieved high classification accuracy. Nonetheless, its reliance on specific datasets limited its generalizability, highlighting the need for validation across diverse clinical datasets. Deepak and Ameer [12] effectively demonstrate a high-accuracy brain tumor classification system using transfer learning but highlights gaps like overfitting with limited data and misclassification of meningioma, suggesting a need for improved data augmentation and further tuning of the transfer-learned model for better generalization.

Sahin et al. [13] applied Bayesian Multi-Objective (BMO) optimization to the Vision Transformer (ViT) architecture, significantly enhancing its efficiency and accuracy. However, the study's findings, tailored to specific optimization techniques and parameters, necessitate further research to validate these improvements across other transformer-based models and diverse medical imaging datasets. By integrating a multi-path CNN structure with an SVM classifier [14], achieved a significant accuracy improvement of up to 99.1%. However, the reliance on limited datasets called into question the model's generalizability in diverse clinical settings, indicating the need for validation on larger and more varied datasets. The research paper [15] effectively enhances multi-class brain tumor classification accuracy by utilizing deep transfer learning models, particularly highlighting the exceptional performance of the Xception architecture. Despite achieving impressive results, this study highlighted the need for validation across diverse clinical environments to ensure model robustness.

Zulfiqar et al. [16] effectively demonstrated fine-tuned EfficientNet models, particularly EfficientNetB2, achieving high accuracy and robustness. However, the exclusion of the normal class during cross-validation indicated a potential gap in assessing the model's generalizability across all clinical scenarios, necessitating comprehensive validation on more diverse datasets. The paper [17] presents a model that incorporates ResNet-50 architecture combined with global average pooling. This approach is designed to tackle challenges such as vanishing gradients and overfitting when classifying brain tumors from MRI scans. The proposed method achieves high classification accuracy, however, it lacked extensive cross-dataset validation to confirm the model's robustness across different imaging conditions and patient populations. Aurna et al. [18] proposed a novel two-stage deep CNN ensemble that achieves high accuracy in brain tumor classification, but it is limited by the inclusion of only three tumor types due to data scarcity, leaving a gap for future studies to explore a wider variety of brain tumors and incorporate additional imaging modalities such as CT and PET scans for enhanced prediction accuracy. Kazemi et al. [19] introduced an efficient parallel deep learning architecture for brain tumor classification, achieving superior accuracy with fewer resources, but the study is limited by the use of only two image databases, highlighting the need for broader dataset incorporation and adaptation of the model for real-time surgical applications and 3D systems for enhanced generalizability.

Vidyarthi et al. [20] introduced a robust methodology for multi-class classification of malignant brain tumors using a novel feature selection algorithm (Cumulative Variance method) and machine learning classifiers, achieving high accuracy with diversified features; however, it lacks validation on larger, gold-standard datasets and could benefit from exploring DL models for enhanced multi-class classification. The BTC-fCNN [21] model offered a fast and efficient method for multi-class classification with impressive accuracy and reduced computational cost. The study was limited to validation on a single dataset, indicating a need

for further testing on diverse datasets to ensure broader applicability and robustness. The BMRI-NET [22] model leveraged three pre-trained models to enhance generalization, achieving high accuracy and recall. However, it primarily focused on performance evaluation on a single dataset, leaving room for further validation on diverse datasets to ensure broader applicability. The approach [23] combined EfficientNet's feature extraction capabilities with attention mechanisms for enhanced performance. However, the study primarily evaluated the model on a newly generated dataset, suggesting the need for validation on larger and more diverse datasets.

Rajput et al. [24] introduced a transfer learning-based approach for brain tumor classification, leveraging pre-trained CNN models to achieve high accuracy; however, while demonstrating notable advantages such as reduced training time and improved performance, it may be limited by the computational resources required for fine-tuning DNNs and the potential need for expertise. The gap lies in exploring the integration of patients' health attributes with image features to further enhance classification accuracy and in validating the proposed approach on larger and more diverse datasets to ensure its robustness across different clinical scenarios. Cheng et al. [25] improved classification accuracy by augmenting the region of interest and using ring-form subregion partitioning. While leveraging surrounding tissues and advanced partition schemes, the study could benefit from integrating more sophisticated preprocessing and supervised dictionary learning techniques for further refinement. The research gap lies in validating diverse tumor types and exploring additional preprocessing and feature coding enhancements. Sultan et al. [26] demonstrated the efficacy of a custom DNN for multi-classification of brain tumors and glioma grading using MRI images, achieving high accuracy. However, the research was limited by the relatively small dataset, necessitating further validation on larger and more diverse datasets to ensure robustness and generalizability.

The reviewed studies illustrate significant advancements in brain tumor classification using various DL and hybrid techniques. Despite significant advancements in medical imaging and DL, current methods for brain tumor classification face several limitations that hinder their practical applicability in real-world medical settings. Existing models often struggle with achieving high accuracy across diverse datasets, particularly when distinguishing between multiple tumor types. A common challenge is the reliance on single-branch network architectures, which may not fully capture the complex, multi-scale features of brain tumors. Additionally, although pre-trained models like EfficientNet have shown promise, there is a lack of comprehensive studies exploring their integration with custom network architectures to enhance performance. Moreover, there is a pressing need for larger and more diverse datasets, cross-dataset validation, and ensuring model robustness and generalizability across different clinical scenarios. Addressing these gaps is essential for developing robust, generalizable models that can handle variations in tumor appearance and improve diagnostic accuracy. Future studies should address these areas to improve the practical applicability of DL models for brain tumor classification.

The key research contributions of this study are as follows:

- Extensive Network Architecture Comparison: We conducted a comprehensive comparison of network architectures ranging from single-branch to six-branch and demonstrated that the 3B Net exhibited superior performance for brain tumor classification.
- Multi-Class Classification Enhancement: Our study encompassed binary, three-class, and four-class classification tasks, rigorously testing the model across diverse scenarios to highlight its versatility and robustness.
- Data Augmentation for Improved Accuracy: Applied advanced data augmentation techniques to address dataset imbalances.
- EfficientNetB2 Feature Fusion: Combined EfficientNetB2's feature extraction with a custom 3B Net for enhanced performance.
- Comprehensive Performance Evaluation: Our study offers a detailed evaluation of various network configurations and classification tasks, providing valuable insights into their strengths and weaknesses and establishing a benchmark for future research in brain tumor classification with DL.
- Robust Diagnostic Tool Development: Contributed to developing an automated diagnostic tool for accurate brain tumor classification.
- Public Dataset Validation: By validating our approach with publicly available datasets, we ensure the reproducibility and generalizability of our findings, fostering transparency and enabling further research and comparison with other methods to advance brain tumor classification technologies.

The remainder of this paper is structured as follows: Section II details the proposed workflow, including preprocessing techniques, data augmentation, and the design of the hybrid 3B Net and EfficientNetB2 model. In Section III, we describe the experimental setup, covering dataset, system and software requirements, hyperparameter tuning, and performance evaluation metrics. Section IV presents the results obtained from the experiments, followed by a comprehensive discussion of the findings, their implications, and a comparison with state-of-the-art methods. Finally, Section V concludes the paper and outlines potential directions for future research.

II. PROPOSED WORKFLOW

This section outlines the comprehensive workflow employed in this study for the classification of brain tumors using MRI. The proposed workflow in the form of a block diagram is shown in Figure 1 is designed to explore the performance of various neural network architectures systematically and to identify the optimal configuration for

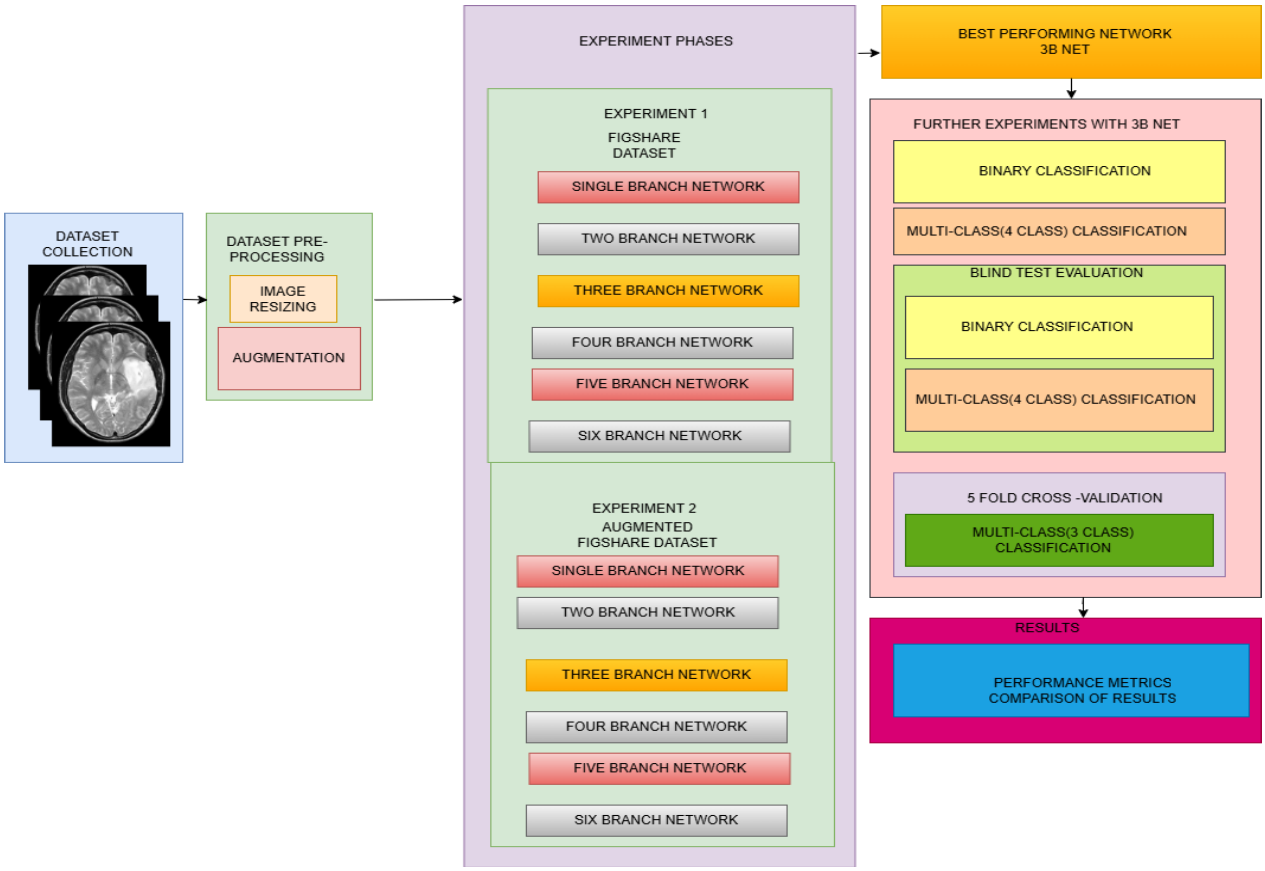

FIGURE 1. Workflow of the proposed method.

both binary and multi-class classification tasks. Initially, experiments were conducted using single to six branch networks on the original Figshare dataset. Subsequently, the same experiments were performed on the augmented Figshare dataset, designed to address class imbalance and enhance model robustness. The 3B Net, a three-branch network, emerged as the best-performing model from these experiments. Further experiments with the 3B Net included both binary classification(tumor/no tumor) and multi-class classification (glioma, meningioma, pituitary tumor, and no tumor). Additionally, blind test evaluations were conducted for binary and four-class classification tasks using separate datasets and five-fold cross-validation for the augmented three-class classification dataset to assess the model's generalizability and robustness. The results from these experiments were compiled and analyzed based on performance metrics such as precision, recall, accuracy, specificity, F1 score, and F2 score, demonstrating the effectiveness of the 3B Net and its potential applications in clinical settings. The subsequent sections provide a detailed explanation of each step in the proposed methodology.

A. PREPROCESSING

Initially, the input images are resized from their original dimensions of $512 \times 512 \times 3$ to $224 \times 224 \times 3$. This resizing reduces computational load while preserving the essential content and features of the images. MRI images are processed by isolating the outermost points of the brain's contour, allowing for the removal of extraneous areas and focusing on the region of interest. This process involves several steps in image processing: first, the image is converted to grayscale, followed by the application of a Gaussian blur filter. The blurred image undergoes thresholding, and a sequence of erosion and dilation operations are carried out to reduce noise. The most prominent contours are then detected, and the four extreme points—top, bottom, left, and right—are extracted. Using these points, the original image is cropped to remove any irrelevant areas that could introduce noise during the training.

To ensure the network does not learn only a narrow subset of the data, the samples are shuffled, allowing the system to train on a diverse set of images.After removing the irrelevant regions from the brain tumor MRIs, the dataset is divided into training and testing sets with an 80:20 split. Specifically, 80% of the samples are allocated for training the model, while the remaining 20% are set aside for testing and evaluating the model on unseen data. Figure 2 presents sample MRI images of brain tumors.

B. DATA AUGMENTATION

As the size of the training set is limited, data augmentation was employed to train the deep convolutional neural

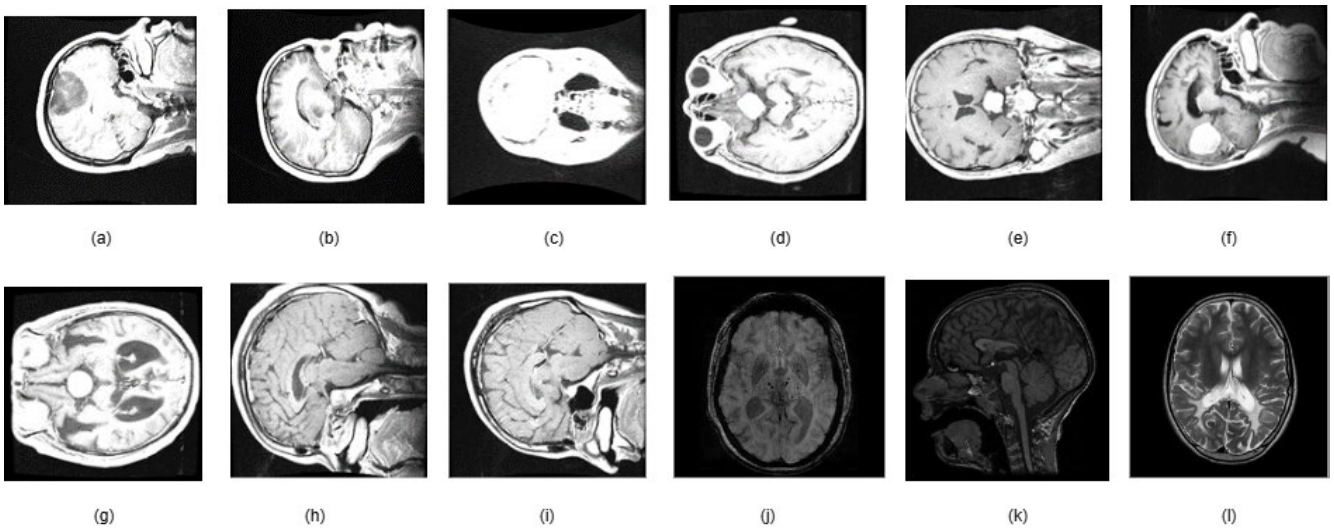

FIGURE 2. Sample images in dataset.

TABLE 1. Class-wise distribution of the dataset before and after data augmentation.

Tumor Type	Tumor slices before data augmentation	Tumor slices after data augmentation
Glioma	1,426	3,420
Meningioma	708	3,396
Pituitary Tumor	930	3,719
Total	**3,064**	**10,535**

network (DCNN) architecture effectively. This technique is essential for enhancing the model's robustness and preventing overfitting by synthetically increasing the number of training samples. Data augmentation involves applying various image transformations such as rotation, scaling, flipping, mirroring, and cropping to generate additional data samples without requiring new data collection. This study applied various data augmentation techniques, such as horizontal and vertical flipping, rotation, and height and width shifts. These methods were strategically employed to each image within the meningioma, glioma, and pituitary tumor categories. Augmentation was specifically performed by a factor of 2 for glioma, 4 for pituitary tumor, and 5 for meningioma images to ensure a more balanced representation across the different classes. Table 1 provides the distribution of instances per class before and after the augmentation process, highlighting the increased dataset size achieved through this method.

C. PROPOSED HYBRID 3B NET AND EFFICIENTNETB2 MODEL

This study introduces a novel hybrid DL model that integrates a custom 3B Net with EfficientNetB2 for multi-class classification of brain tumor images. The goal is to leverage the complementary strengths of a custom convolutional feature extraction module and a pre-trained model to improve classification performance. Figure 3 shows the proposed Hybrid 3 BNet and EfficientNetB2 model for brain tumor classification.The model architecture begins with an initial convolutional layer comprising 128 filters with a kernel size of 9, followed by ReLU activation and max pooling with a pool size of 5 and a stride of 2. This layer performs the initial feature extraction from the input images. The network is then divided into three branches, each employing different convolutional kernel sizes (5, 7, and 9) to capture features at various scales. Each branch consists of a convolutional layer with 96 filters, ReLU activation, and batch normalization. The outputs of these three branches are concatenated to form a combined feature map. Figure 4 shows the proposed 3B Net architecture.

Subsequently, a global max pooling layer reduces the dimensions of the feature map. A 1×1 convolutional layer with 576 filters and ReLU activation is then applied to match the number of channels with those expected by Efficient-NetB2. This is followed by max pooling and upsampling to align the feature map shape with the EfficientNetB2 output dimensions.EfficientNetB2, pre-trained on ImageNet, is incorporated into the model for feature extraction. To adapt the pre-trained features to the specific task of brain tumor classification, the last 20 layers of EfficientNetB2, excluding batch normalization layers, are unfrozen for fine-tuning. This fine-tuning enables the model to leverage the pre-trained knowledge while learning domain-specific features from the brain tumor dataset.

The features extracted from both the 3B Net and EfficientNetB2 are concatenated to form fused features.To mitigate overfitting, a dropout layer with a 50% drop rate is incorporated. These fused features are then flattened using a global max pooling layer. The flattened features are passed through dense layers with 1024 and 512 units, each followed

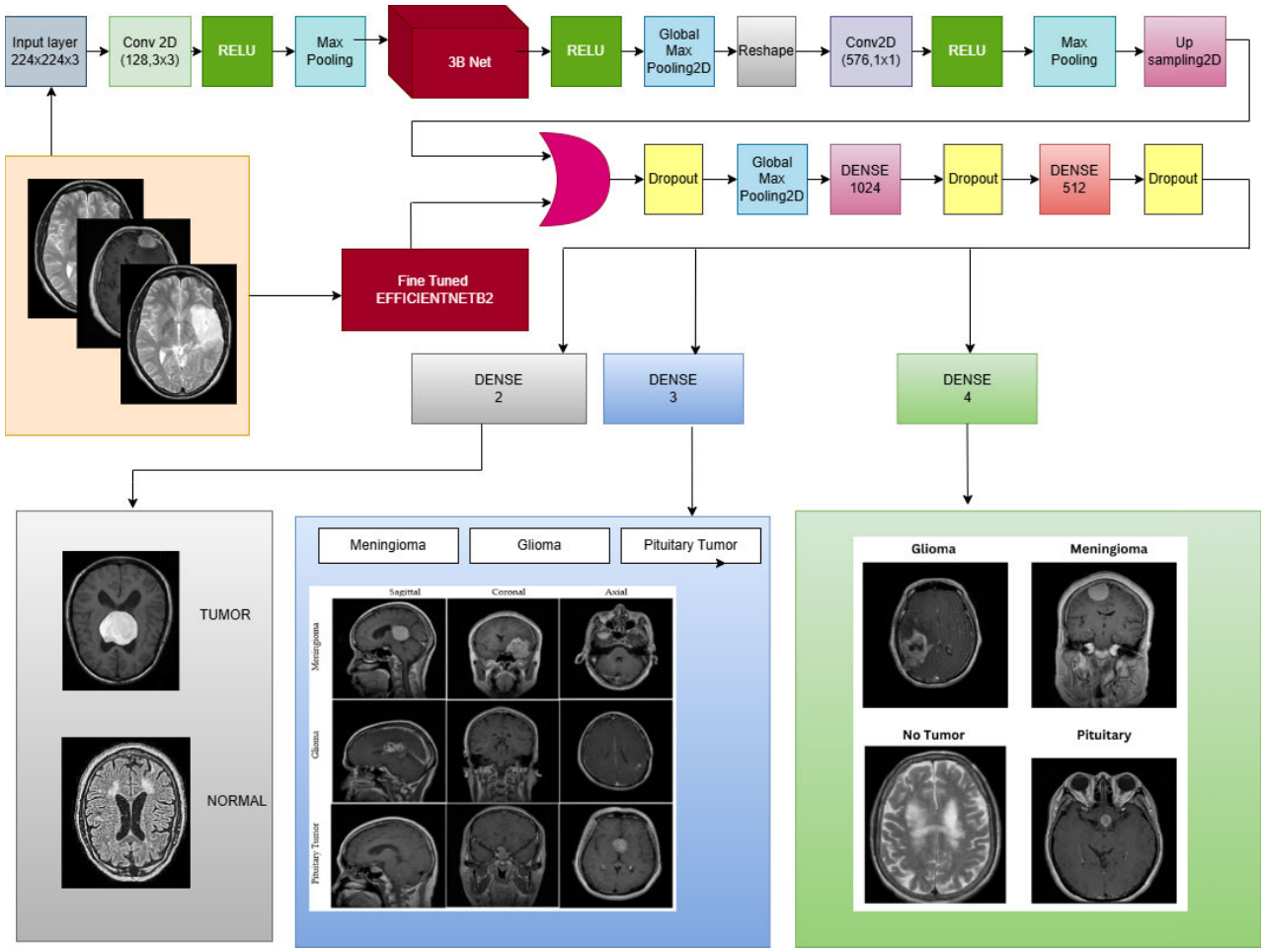

FIGURE 3. Proposed Hybrid 3B Net and EfficientNetB2 model.

by ReLU activation and dropout layers, to learn high-level representations. The final output layer uses a softmax activation function with dense layers containing 2 units for binary classification (tumor vs. no tumor), 3 units for three-class classification (meningioma, glioma, and pituitary tumors), and 4 units for four-class classification (no tumor, meningioma, glioma, and pituitary tumors). The model is configured with the Adam optimizer, set to a learning rate of 1e-4. It employs categorical cross-entropy as the loss function and uses accuracy as the primary evaluation metric. To enhance training efficiency, the ReduceLROnPlateau callback is implemented, which reduces the learning rate by 20% if the validation loss does not improve after three consecutive epochs, with a minimum set at 1e-6 for the learning rate.

The training procedure involves splitting the dataset into training and test sets. The model is trained on the training set, with its performance monitored to prevent overfitting. After training, the model's performance is evaluated on the test set using accuracy and other relevant metrics such as recall, precision, F1-score and F2-score. In summary, this hybrid model combines the strengths of a custom 3B Net and a pre-trained EfficientNetB2 model to effectively classify brain tumor images. By integrating a structured feature extraction approach with deep, pre-trained features and fine-tuning the last 20 layers of EfficientNetB2, the model aims to achieve high accuracy and robustness in the multi-class classification of brain tumors.

D. COMPARISON WITH EXISTING MULTI-BRANCH CNN ARCHITECTURES

The proposed 3B Net introduces a novel approach to multi-branch convolutional networks by integrating hierarchical feature fusion and depth-wise EfficientNetB2 integration to enhance classification performance while maintaining computational efficiency. Unlike conventional multi-branch CNNs, which primarily concatenate features from independent branches at a fixed depth, our model leverages feature fusion across multiple depth levels to improve spatial and contextual understanding of brain tumor images.

1) KEY DIFFERENCES FROM EXISTING MULTI-BRANCH CNNs

1) **Hierarchical Feature Fusion**: Unlike conventional multi-branch CNNs that concatenate features at a single depth, our model integrates features across

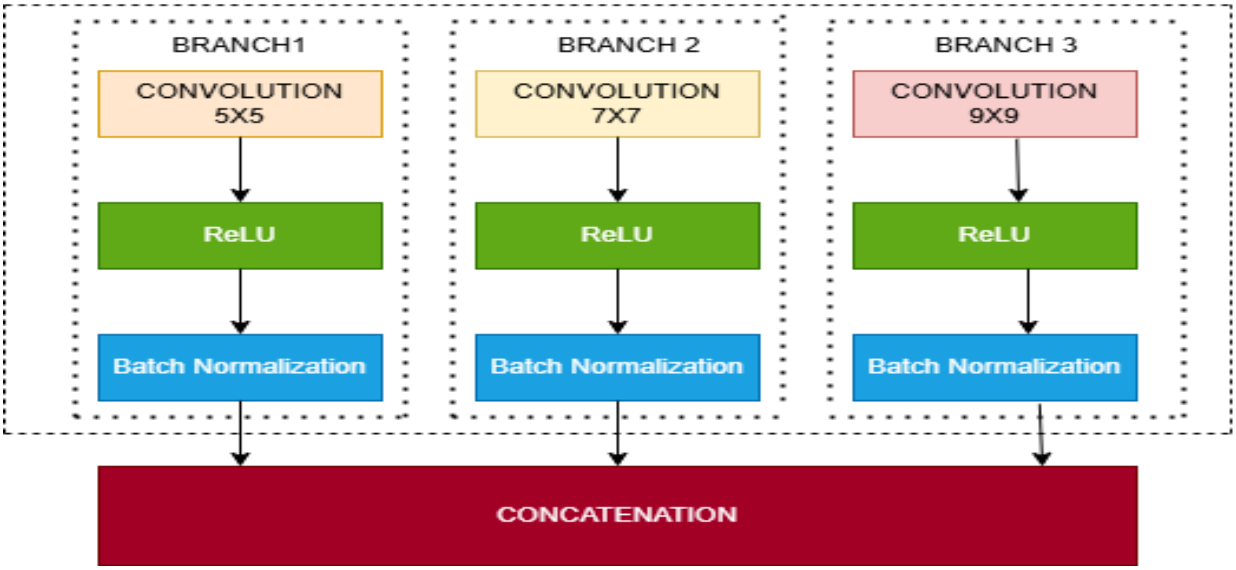

FIGURE 4. Architecture of 3B Net.

multiple depth levels, enhancing spatial and contextual understanding.

2) **Multi-Scale Feature Extraction with EfficientNetB2**: We employ three parallel branches with different kernel sizes ($5\times5, 7\times7, 9\times9$) to capture multi-scale features, which are then aligned with EfficientNetB2 using a 1×1 convolutional bottleneck layer.
3) **Depth-Wise Parameter Sharing**: Our model optimizes computational efficiency by reducing redundant filters through depth-wise parameter sharing and selective fine-tuning of EfficientNetB2 layers. Instead of freezing most layers of EfficientNetB2, our approach fine-tunes the last 20 layers (excluding batch normalization layers), allowing the model to leverage pre-trained knowledge while learning domain-specific features.

E. MODEL INTERPRETABILITY USING GRAD-CAM

To enhance the interpretability of our proposed 3B Net integrated with EfficientNetB2, we employed Gradient-weighted Class Activation Mapping (Grad-CAM) to visualize the important regions in MRI images that contributed to the model's classification decisions. Grad-CAM generates class-discriminative heatmaps by highlighting the spatial regions most influential in determining the output prediction. This approach provides valuable insights into how the model identifies tumor and non-tumor regions, ensuring its transparency and clinical applicability.

III. EXPERIMENTAL SETUP

The experimental setup for this study is meticulously designed to ensure comprehensive evaluation and robustness of the proposed DL model. It encompasses multiple critical components, beginning with the dataset selection, which form the foundation for training and validating the models. The system and software requirements are carefully outlined to provide a reproducible environment, detailing the hardware and software specifications used. Additionally, hyperparameter settings are meticulously optimized to enhance the model's performance, balancing between underfitting and overfitting. Finally, the performance evaluation metrics are clearly defined, ensuring a thorough assessment of the model's effectiveness across various classification tasks. Each of these aspects is discussed in detail in the following sections.

A. DATASET

In this study, we utilized multiple publicly available datasets to perform a comprehensive evaluation of our DL models for brain tumor classification. These datasets encompass a variety of brain tumor types and imaging modalities, providing a robust foundation for training and testing our models. The diversity and volume of these datasets enable us to conduct extensive experiments, including binary and multiclass classification tasks. By augmenting the dataset, we aimed to address class imbalance and enhance the model's ability to generalize. Below, we provide a detailed description of each dataset used in our experiments. The class-wise distribution of the dataset is presented in Table 2.

1) DATASET 1:FIGSHARE BRAIN TUMOR DATASET

The primary dataset utilized in this study is the publicly available Figshare brain tumor dataset [27], (https://figshare.com/articles/dataset/brain_tumor_dataset/1512427/8). This dataset includes 3064 T1-weighted Contrast-Enhanced Magnetic Resonance Images (CE-MRI) obtained from 233 patients. The data was collected between 2005 and 2010 from Nanfang Hospital, General Hospital, and Tianjin Medical University in China. The dataset is organized into three categories of brain tumors: meningioma (708 slices

TABLE 2. Summary of brain tumor datasets.

SL. NO.	DATASET	CLASSIFICATION TYPE	CLASSES	TOTAL NO. OF IMAGES
1	FIGSHARE	3 CLASS	GLIOMA	1426
			MENINGIOMA	708
			PITUITARY TUMOR	930
2	AUGMENTED FIGSHARE	3 CLASS	GLIOMA	3420
			MENINGIOMA	3396
			PITUITARY TUMOR	3719
3	Br35H: Brain Tumor Detection 2020	2 CLASS	NO TUMOR	1500
			TUMOR	1500
4	Brain Tumor Classification (MRI) Dataset	4 CLASS	GLIOMA	926
			MENINGIOMA	937
			PITUITARY TUMOR	901
			NO TUMOR	500
5	Brain-Tumor-Dataset	4 CLASS	GLIOMA	1038
			MENINGIOMA	1318
			PITUITARY TUMOR	681
			NO TUMOR	1255
6	Brain MRI Images for Brain Tumor Detection Dataset	2 CLASS	NO TUMOR	98
			TUMOR	155

from 82 patients), glioma (1426 slices from 89 patients), and pituitary tumor (930 slices from 62 patients). The original MRI images have a resolution of 512 × 512 pixels and are stored in MATLAB (.mat) format, which includes class labels, patient identifiers, image data, tumor boundaries, and tumor masks. To prevent data leakage, we ensured that no patient contributed images to both the training and test datasets. This was achieved by assigning all slices from a given patient exclusively to either the training or test set, thereby maintaining the independence of these datasets.

2) DATASET 2: AUGMENTED FIGSHARE BRAIN TUMOR DATASET

To address the class imbalance in the dataset and enhance the model's robustness, data augmentation techniques were employed. Specifically, images of meningioma, pituitary and glioma tumors were augmented by factors of 5, 4, and 2, respectively. This augmentation resulted in a more balanced dataset with the following distribution: 3420 images of glioma, 3396 images of meningioma, and 3719 images of pituitary tumors. Data augmentation techniques were applied solely to the training set to further enhance model generalization.

3) DATASET 3: BRAIN TUMOR DETECTION 2020 DATASET

For binary classification, the Brain Tumor Detection 2020 dataset from Kaggle [28] was employed. This dataset includes a total of 3000 images, equally divided between images containing tumors (1500) and normal brain scans (1500). The dataset is publicly available at https://www.kaggle.com/datasets/ahmedhamada0/brain-tumor-detectionKaggle

4) DATASET 4: BRAIN TUMOR CLASSIFICATION (MRI) DATASET

For four-class classification, we employed the Brain Tumor Classification (MRI) dataset [29]. The training directory consists of 395 images without tumors, 826 images with gliomas, 822 with meningiomas, and 827 with pituitary tumors. The testing directory consists of 105 images without tumors, 115 with meningiomas, 100 images with gliomas, and 74 with pituitary tumors. The dataset is publicly available at https://www.kaggle.com/datasets/sartajbhuvaji/brain-tumor-classification-mriKaggle

5) DATASET 5: BRAIN-TUMOR-DATASET

For another four-class classification task, we employed the Brain-Tumor-Dataset [30]. The training directory comprises 445 images without tumors, 981 with meningiomas, 891 images with gliomas, and 959 with pituitary tumors. The testing directory includes 236 images without tumors, 337 with meningiomas, 147 images with gliomas, and 296 with pituitary tumors. The dataset is publicly available at https://www.kaggle.com/datasets/mohamedmetwalysherif/braintumordataset Kaggle.

6) DATASET 6: BRAIN MRI IMAGES FOR BRAIN TUMOR DETECTION DATASET

This dataset consists of a total of 253 images, of which 98 are non-tumor images and 155 are tumor images [31].

By utilizing these diverse datasets, we aimed to create robust models capable of performing accurate and generalizable brain tumor classification across various scenarios. The dataset is publicly available at https://www.kaggle.com/datasets/navoneel/brain-mri-images-for-brain-tumor-detectionKaggle.

7) DATA SPLITTING AND LEAKAGE PREVENTION

To ensure the integrity of our model evaluation and prevent data leakage, we implemented the following strategies:

1) **Patient-Wise Data Splitting**: We assigned all slices from a single patient exclusively to either the training or testing set. This approach prevents the model from encountering slices from the same patient in both sets, thereby avoiding overestimation of performance due to data leakage.
2) **Training Set Augmentation**: Data augmentation techniques were applied solely to the training set. This practice ensures that augmented variations of the training data do not appear in the testing set, maintaining the independence of the evaluation process.

While the Figshare dataset provides subject-level identifiers, allowing for precise patient-wise data splitting, other datasets used in our study lacked such metadata. In these cases, each image was treated as an independent sample, as subject-wise grouping information was unavailable. This limitation may introduce potential biases, such as inadvertently including multiple slices from the same patient in both training and test sets. To mitigate this risk, we relied on the dataset's predefined structure and ensured that any data augmentation was confined to the training set. By implementing these measures, we aimed to uphold the robustness and generalizability of our model, ensuring that performance metrics accurately reflect its capability to handle unseen data.

B. SYSTEM AND SOFTWARE REQUIREMENTS

The proposed model was implemented using Python, with TensorFlow and Keras libraries. The model was trained on a computer equipped with an Intel Core i5-12450H CPU running at 2.00 GHz, a 64-bit operating system, and 16 GB of RAM. The system also featured a 512 GB SSD for storage. Additionally, an NVIDIA GeForce RTX 3050 GPU was employed to accelerate the training process. A summary of the system specifications is presented in Table 3.

TABLE 3. System and software requirements.

Name	Specifications
Operating System	Windows 11, 64-bit
CPU	Intel Core i5-12450H at 2.00 GHz
Memory	16 GB
GPU	NVIDIA GeForce RTX 3050
Development Environment	Python 3.10
Software Library	TensorFlow

C. HYPER-PARAMETER TUNING

Hyperparameter tuning [32] plays a crucial role in optimizing the performance of DL models. In this study, we employed a systematic approach to fine-tune the hyperparameters of the proposed Hybrid 3B Net integrated with EfficientNet B2, to achieve optimal performance in brain tumor classification.The hyperparameter tuning [33] was performed using different methods, like Grid Search [34], Random Search [35], and Bayesian Optimization [36] on the validation dataset to identify the optimal combination of these hyperparameters.

TABLE 4. Hyper parameter Search space and best value identified.

Hyper-parameter	Search space	Best value
Conv 1 filters	[64,128,192,256]	128
Conv 1 kernel size	[3,5,7,9]	9
Branch 1 filters	[32, 64, 96, 128]	96
Branch 2 filters	[32, 64, 96, 128]	96
Branch 3 filters	[32, 64, 96, 128]	96
Conv 2 filters	[256, 320, 384, 448, 512, 576]	576
Dropout rate	[0.3, 0.4, 0.5, 0.6, 0.7]	0.5
Dense 1 units	[512, 640, 768, 896, 1024]	1024
Dense 2 units	[256, 384, 512]	512
Learning rate	[1e-6 to 1e-3] (log scale)	1e-6
Optimizer	[Adam, RMSprop, SGD]	Adam

1) SEARCH SPACE SELECTION AND BAYESIAN OPTIMIZATION

To ensure a well-defined search space, we selected hyperparameter ranges based on prior studies and empirical observations. The search space included a broad set of values for key hyperparameters, including the number of convolutional filters, kernel sizes, dropout rates, and learning rates. Bayesian Optimization was employed to efficiently explore this space, iteratively selecting the best hyperparameter combinations to maximize validation accuracy while minimizing overfitting. Table 4 summarizes the hyperparameter search space and the best values identified during the tuning process.

2) COMPUTATIONAL EXPENSE OF HYPERPARAMETER TUNING

The Bayesian Optimization process was performed on a system with an NVIDIA GeForce RTX 3050 GPU, Intel Core i5-12450H CPU, and 16GB RAM. The tuning process required 50 iterations, where each iteration trained the model for 10 epochs before evaluating its performance. The total computational time for the optimization process was approximately 8 hours, ensuring an optimal balance between accuracy and efficiency. This systematic tuning significantly

enhanced the model's classification performance, as detailed in Section IV.

D. PERFORMANCE EVALUATION METRICS

The confusion matrix (CM) is a widely-used tool for assessing the predictive accuracy of a model on test data. It consists of equal numbers of rows and columns, representing the true class labels and the predicted labels. The matrix includes values that show the count of correct and incorrect predictions for each class. True Positive (TP) refers to cases where positive instances are correctly identified, while True Negative (TN) denotes cases where negative instances are correctly classified. False Positive (FP), also known as Type-1 errors, arise when negative instances are mistakenly classified as positive, and False Negative (FN), or Type-2 errors, occur when positive instances are incorrectly labeled as negative. The performance of AI models is assessed using metrics such as recall (sensitivity), precision, accuracy, specificity, F1 score, and F2 score. These metrics are detailed in the following subsections.

1) PRECISION

Precision measures the accuracy of positive predictions made by the model and is calculated as:

$$Precision = \frac{TP}{TP + FP} \tag{1}$$

In brain tumor classification, precision is crucial as it reflects the model's ability to accurately identify the presence of tumors without mistakenly classifying healthy tissue as diseased. High precision ensures that the predictions made by the model are reliable, thereby supporting medical professionals in making well-informed decisions regarding patient diagnosis and treatment.

2) RECALL/SENSITIVITY

Sensitivity, also known as recall or the true positive rate (TPR), measures the proportion of actual positives that are correctly identified by the model. It is calculated as:

$$Recall/Sensitivity/TPR = \frac{TP}{TP + FN} \tag{2}$$

This metric is crucial in medical diagnostics, such as brain tumor classification, as it directly impacts the accurate identification of tumors. High sensitivity ensures that the model can effectively detect true positive cases, minimizing the risk of missing actual instances of tumors, which is essential for timely intervention and treatment.

3) SPECIFICITY

Specificity, or the true negative rate (TNR), measures the proportion of actual negatives that are correctly identified by the model. It is calculated as:

$$Specificity/TNR = \frac{TN}{TN + FP} \tag{3}$$

In brain tumor classification, specificity is vital as it measures the model's ability to correctly identify non-tumor cases, ensuring that healthy individuals are accurately distinguished from those with tumors. High specificity reduces the risk of false positives, which is essential for preventing unnecessary treatments and providing accurate diagnoses.

4) F-1 SCORE

The F1-score provides a balanced evaluation of model performance by calculating the harmonic mean of precision and recall. It is given by:

$$F1\text{-}score = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{4}$$

A high F1-score in medical diagnosis tasks, such as brain tumor classification, reflects the model's effectiveness in accurately identifying both tumor and non-tumor cases while minimizing the occurrence of false positives and false negatives. This balance between precision and recall is crucial for ensuring reliable and comprehensive diagnostic performance.

5) F-2 SCORE

The F2-score is used to evaluate the effectiveness of a model, particularly when recall is prioritized over precision. It introduces a configuration parameter, β, which can be adjusted to give more or less weight to precision and recall. The formula is:

$$F2\text{-}score = \frac{5 \times Precision \times Recall}{4 \times Precision + Recall} \tag{5}$$

In brain tumor classification, a high F2-score is crucial because it prioritizes recall, ensuring that most tumors are correctly identified, which is vital for accurate diagnosis and timely treatment.

6) ACCURACY

Accuracy measures the proportion of correctly predicted labels among all labels and is calculated as:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{6}$$

This metric provides an overall measure of the model's predictive performance on the test data, indicating the percentage of correct predictions made by the model.

IV. RESULTS AND DISCUSSION

This section presents the results of various experiments conducted to evaluate the performance of different DL architectures for the classification of brain tumors using MRI. The experiments were performed on multiple datasets, including the original and augmented Figshare brain tumor dataset, the Brain Tumor Detection 2020 dataset, Brain Tumor Classification(MRI) dataset, Brain -Tumor-dataset and Brain MRI images for Brain Tumor Detection dataset. The primary objective was to compare the effectiveness of

TABLE 5. Performance of different network architectures on original figshare dataset.

Model	Class	A	P	R	F1	S	F2
1B	0	96.85	91.72	96.85	94.22	95.78	92.58
	1	75.35	91.45	75.35	82.63	98.03	78.10
	2	98.39	93.85	98.39	96.06	97.22	97.44
	Average	92.35	92.30	92.35	92.10	95.82	92.34
2B	0	96.85	92.95	96.85	94.86	93.77	96.05
	1	80.28	91.94	80.28	85.71	98.00	82.37
	2	97.85	94.79	97.85	96.30	97.69	97.22
	Average	93.32	93.27	93.32	93.18	96.38	93.31
3B	0	96.85	94.86	96.85	95.85	96.45	95.55
	1	86.62	92.48	86.62	89.45	97.96	87.73
	2	98.39	96.83	98.39	97.60	98.61	98.07
	Average	**94.95**	**94.91**	**94.95**	**94.90**	**97.30**	**94.94**
4B	0	96.85	92.95	96.85	94.86	93.77	96.05
	1	78.87	91.06	78.87	84.53	97.81	81.04
	2	98.39	94.82	98.39	96.57	97.68	97.65
	Average	93.16	93.08	93.16	92.99	96.31	93.14
5B	0	95.10	92.83	95.10	93.96	94.64	93.86
	1	77.46	90.16	77.46	83.33	97.62	79.71
	2	99.46	92.96	99.46	96.10	96.74	98.09
	Average	92.35	92.26	92.35	92.15	95.93	92.33
6B	0	98.60	91.56	98.60	94.95	97.11	92.17
	1	76.76	97.32	76.76	85.83	99.41	80.15
	2	99.46	95.36	99.46	97.37	97.90	98.61
	Average	93.81	94.04	93.81	93.57	96.44	93.86

*R-Recall,*A-Accuracy,*F1-F1-score,*P-Precision,*F2-F2-Score,*S-Specificity,*0-Glioma,*1-Meningioma,*2-Pituitary tumor

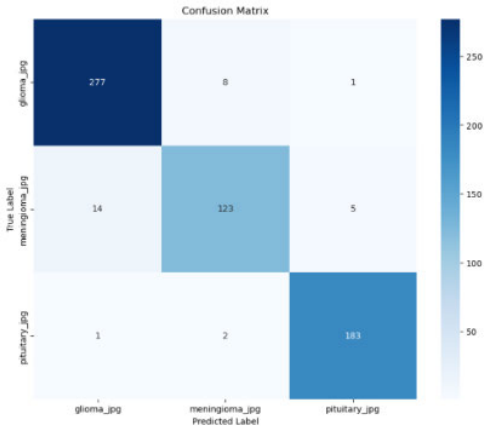

FIGURE 5. Confusion matrix of EXPERIMENT 1.

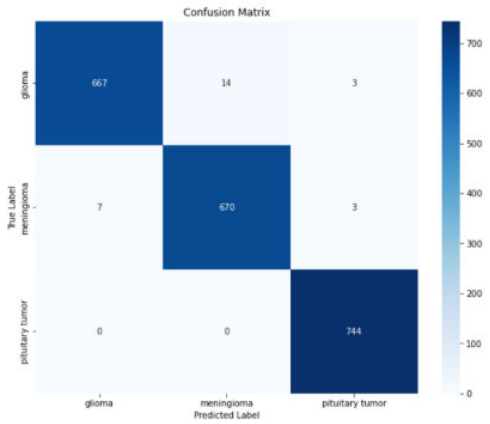

FIGURE 6. Confusion matrix of EXPERIMENT 2.

the proposed 3B Net architecture against other state-of-the-art networks in binary and multiclass classification tasks. The results are discussed in terms of key performance metrics such as precision, recall, accuracy, F1 score, specificity, and F2 score.

A. EXPERIMENT 1: CLASSIFICATION ON ORIGINAL FIGSHARE BRAIN TUMOR DATASET

In the first experiment, we evaluated the performance of six different network architectures (1B Net, 2B Net, 3B Net, 4B Net, 5B Net, and 6B Net) on the original Figshare brain tumor dataset, consisting of 3064 images. The results demonstrated that the 3B Net architecture achieved superior performance across multiple metrics. Detailed results for each network are summarized in Table 5. Accuracy and loss plots are shown in Figure 11 and the confusion matrix in Figure 5.

B. EXPERIMENT 2: CLASSIFICATION ON AUGMENTED FIGSHARE BRAIN TUMOR DATASET

To further enhance model performance, we augmented the Figshare dataset to balance the class distribution and

TABLE 6. Performance of different network architectures on augmented figshare dataset.

Model	Class	A	P	R	F1	S	F2
1B	0	97.66	98.09	97.66	97.88	99.10	97.75
	1	97.65	98.08	97.65	97.86	99.10	97.73
	2	100	99.20	100	99.60	99.56	99.84
	Average	98.48	98.46	98.44	98.45	99.25	98.48
2B	0	96.49	97.63	96.49	97.06	98.90	96.72
	1	97.06	97.35	97.06	97.20	98.76	97.12
	2	100	98.67	100	99.33	99.27	99.73
	Average	97.91	97.88	97.85	97.86	98.96	97.91
3B	0	97.51	98.96	97.51	98.23	99.51	97.80
	1	98.53	97.95	98.53	98.24	99.03	98.41
	2	100	99.20	100	99.60	99.56	99.84
	Average	**98.72**	**98.70**	**98.68**	**98.69**	**99.36**	**98.72**
4B	0	97.22	98.52	97.22	97.87	99.31	97.48
	1	98.09	97.80	98.09	97.94	98.96	98.03
	2	100	99.07	100	99.53	99.49	99.81
	Average	98.48	98.46	98.44	98.45	99.24	98.48
5B	0	97.81	98.09	97.81	97.95	99.10	97.86
	1	97.94	98.23	97.94	98.09	99.17	98.00
	2	100	99.47	100	99.73	99.71	99.89
	Average	98.62	98.60	98.60	98.60	99.32	98.62
6B	0	97.37	98.38	97.37	97.87	99.24	97.57
	1	97.94	97.65	97.94	97.80	98.89	97.88
	2	100	99.33	100	99.67	99.63	99.87
	Average	98.48	98.45	98.44	98.44	99.25	98.48

*R-Recall,*A-Accuracy,*F1-F1-score,*P-Precision,*F2-F2-Score,*S-Specificity,*0-Glioma,*1-Meningioma,*2-Pituitary tumor

increase the number of training samples. This augmentation included rotation, flipping, shifting, and scaling. We then reevaluated the six network architectures on this augmented dataset (totaling 10535 images). The results are presented in Table 6. Accuracy and loss plots are shown in Figure 12 and the confusion matrix in Figure 6. These experiments showed that the 3B Net consistently outperformed other networks on both the original and augmented datasets, achieving higher accuracy, precision, recall, F1 score, and F2 scores.

C. EXPERIMENT 3: BINARY CLASSIFICATION ON BRAIN TUMOR DETECTION 2020 DATASET USING 3B NET

Encouraged by the results from the Figshare dataset, we employed the 3B Net architecture for binary classification on the Brain Tumor Detection 2020 dataset. This dataset presented a distinct binary classification challenge. The 3B Net achieved exceptional performance, as outlined in Table 7. The results demonstrate the effectiveness of the 3B Net in binary classification tasks, with an accuracy and F1 score of 99.50%. Accuracy and loss plots are shown in Figure 13 and the confusion matrix in Figure 7.

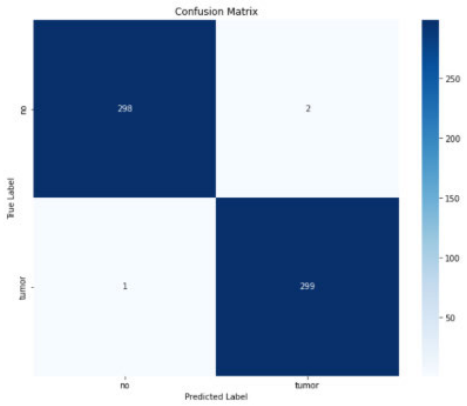

FIGURE 7. Confusion matrix of EXPERIMENT 3.

D. EXPERIMENT 4: FOUR-CLASS CLASSIFICATION ON BRAIN TUMOR CLASSIFICATION (MRI) DATASET USING 3B NET

Experiment 4 involved the application of the 3B Net for four-class classification on the Brain Tumor Classification (MRI) dataset, which consists of images categorized into four types of brain tumors. The performance metrics are summarized in Table 8. Accuracy and loss plots are shown in Figure 14 and the confusion matrix in Figure 8. The 3B Net demonstrated robust performance in the four-class classification task, maintaining high accuracy and precision across all classes.

E. EXPERIMENT 5: BLIND TEST EVALUATION FOR BINARY CLASSIFICATION

To evaluate the generalization ability of the 3B Net, we conducted a blind test evaluation. The model was trained

TABLE 7. Performance of 3B Net on Brain Tumor Detection 2020 Dataset (Binary Classification).

Model	Class	A	P	R	F1	S	F2
3B	0	99.33	99.67	99.33	99.50	99.67	99.40
	1	99.67	99.34	99.67	99.50	99.34	99.60
	Average	99.50	99.50	99.50	99.50	99.50	99.50

*R-Recall,*A-Accuracy,*F1-F1-score,*P-Precision,*F2-F2-Score,*S-Specificity,*0-No Tumor,*1-Tumor

TABLE 8. Performance of 3B Net on brain tumor classification (MRI) dataset (Four-Class Classification).

Model	Class	A	P	R	F1	S	F2
3B	0	96.77	97.83	96.77	97.30	99.12	96.98
	1	97.34	97.34	97.34	97.34	98.89	97.34
	2	96.25	97.47	96.25	96.86	99.64	96.49
	3	100	98.37	100	99.18	99.34	99.67
	Average	97.80	97.79	97.80	97.79	99.24	97.79

*R-Recall,*A-Accuracy,*F1-F1-score,*P-Precision,*F2-F2-Score,*S-Specificity,*0-Glioma,*1-Meningioma,*2-No tumor,*3-Pituitary tumor

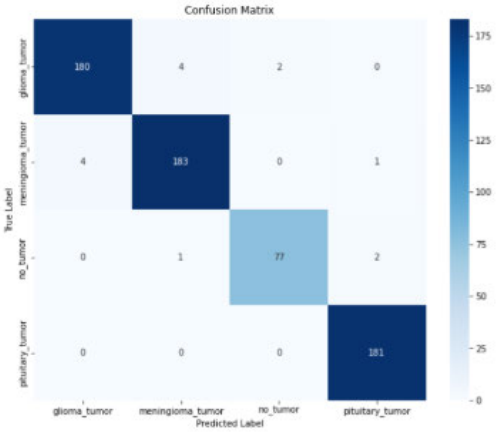

FIGURE 8. Confusion matrix of EXPERIMENT 4.

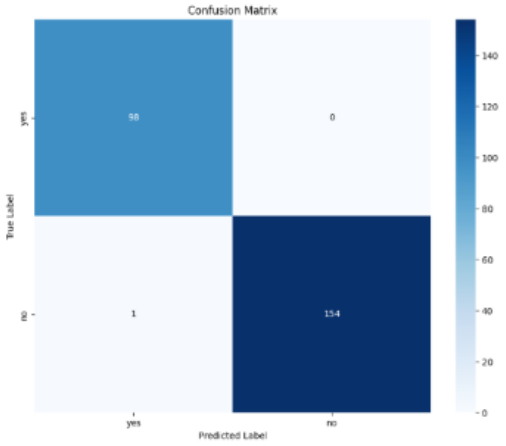

FIGURE 9. Confusion matrix of EXPERIMENT 5.

with 3000 images collected from the Br35H:Brain Tumor Detection 2020 dataset and tested on an unseen set of 253 images collected from Brain MRI Images for Brain Tumor Detection Dataset. The results are summarized in Table 9. Accuracy and loss plots are shown in Figure 15 and the confusion matrix in Figure 9.

F. EXPERIMENT 6: BLIND TEST EVALUATION FOR FOUR CLASS CLASSIFICATION

In this experiment, we conducted a blind test evaluation for four-class classification using the 3B Net model. The training was performed on Brain Tumor Classification(MRI) Dataset comprising 3264 images, while testing was carried out on a separate set of 703 images collected from the testing directory of Brain-Tumor-Dataset. The classes included glioma, meningioma, no tumor, and pituitary tumors. The results obtained from this evaluation are summarized in Table 10. The results indicate exceptional performance, with the model achieving nearly perfect metrics across all classes, demonstrating the robustness and accuracy of the 3B Net

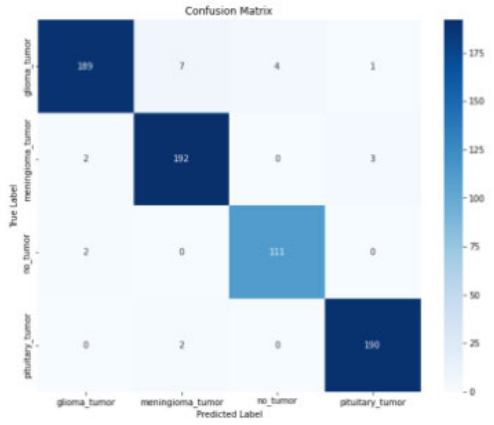

FIGURE 10. Confusion matrix of EXPERIMENT 6.

in a four-class classification setting. Accuracy and loss plots are shown in Figure 16 and the confusion matrix in Figure 10.

TABLE 9. Performance of 3B Net on blind test evaluation (Binary Classification).

Model	Class	R	A	F1	P	S	F2
3B	0	98.67	98.67	98.67	98.67	98.67	98.67
	1	98.67	98.67	98.67	98.67	98.67	98.67
	Average	98.67	98.67	98.67	98.67	98.67	98.67

*R-Recall,*A-Accuracy,*F1-F1-score,*P-Precision,*F2-F2-Score,*S-Specificity,*0-No Tumor,*1-Tumor

TABLE 10. Performance of 3B Net on blind test evaluation for four class classification.

Model	Class	A	P	R	F1	S	F2
3B	0	94.03	97.93	94.03	95.94	99.22	94.78
	1	97.46	95.52	97.46	96.48	98.24	97.07
	2	98.23	96.52	98.23	97.37	99.32	97.88
	3	98.96	97.94	98.96	98.45	99.22	98.75
	Average	97.01	97.03	97.01	97.01	98.99	97.02

*R-Recall,*A-Accuracy,*F1-F1-score,*P-Precision,*F2-F2-Score,*S-Specificity,*0-Glioma,*1-Meningioma,*2-No tumor,*3-Pituitary tumor

TABLE 11. Performance of 3B Net on five fold cross-validation for three class classification.

Fold	Accuracy (%)	Loss
1	98.20	0.1057
2	97.91	0.0932
3	98.10	0.0944
4	97.86	0.0981
5	98.29	0.0699
Average	98.07	0.0923

G. EXPERIMENT 7: FIVE FOLD CROSS-VALIDATION FOR THREE CLASS AUGMENTED FIGSHARE DATASET

In this experiment, we performed a five-fold cross-validation using the augmented Figshare brain tumor dataset for three-class classification (meningioma, glioma, and pituitary tumors). This method involves partitioning the dataset into five segments. The model is trained on four of these segments while the fifth segment is used for validation. This procedure is repeated five times, ensuring each segment is used as the validation set exactly once. The results for each fold and the average performance are presented in Table 11. The average accuracy across the five folds is 98.07%, with an average loss of 0.0923. These results highlight the consistency and reliability of the 3B Net model when applied to the three-class augmented Figshare dataset, further supporting its efficacy for brain tumor classification tasks. Accuracy and loss plots are shown in Figure 17.

The experiments conducted in this study underscore the superior performance of the 3B Net architecture for brain tumor classification across various datasets and tasks. Initially, experiments on the original and augmented Figshare datasets demonstrated the 3B Net's consistent outperformance over other network architectures. Data augmentation played a crucial role in enhancing model performance, especially for the 3B Net, by addressing class imbalance and improving generalization. The model excelled in both binary and multiclass classification tasks, achieving high accuracy, recall, precision, and F1 scores. The proposed model achieved a binary classification accuracy of 99.50%, a three-class classification accuracy of 94.95% without augmentation and 98.72% with augmentation, and a four-class classification accuracy of 97.80%.In a blind test evaluation for binary classification, the 3B Net achieved an accuracy of 98.67% and for four-class classification 97.01%, further validating its robustness. Additionally, a five-fold cross-validation on the augmented Figshare dataset for three-class classification yielded an average accuracy of 98.07%, highlighting the model's reliability. The 3B Net's exceptional performance across these experiments confirms its effectiveness as a tool for brain tumor classification using MRI images, making it a valuable asset for medical diagnostics and research. Future work could explore further optimization and application of the 3B Net to other medical imaging tasks.

H. MODEL ROBUSTNESS AND DATASET CONSIDERATIONS

Our study utilized publicly available datasets from Figshare and Kaggle (as detailed in Table 2), which offer a diverse range of brain tumor MRI images. However, we acknowledge that these datasets often lack comprehensive metadata on imaging protocols, scanner models, and patient demographics, limiting our ability to thoroughly analyze potential biases. To address potential biases and enhance model robustness, we implemented several data augmentation techniques. These include adjustments to image contrast, brightness, and the introduction of noise, which aim to simulate variability in imaging conditions that may arise from different protocols or scanner characteristics [37]. While not a perfect substitute for true protocol and scanner diversity, these augmentations help improve the model's ability to generalize across varying imaging conditions. We recognize that the absence of detailed metadata restricts our capacity to fully assess the model's performance across diverse imaging conditions and demographic groups.

To address this limitation in future work, we plan to incorporate datasets with more comprehensive metadata,

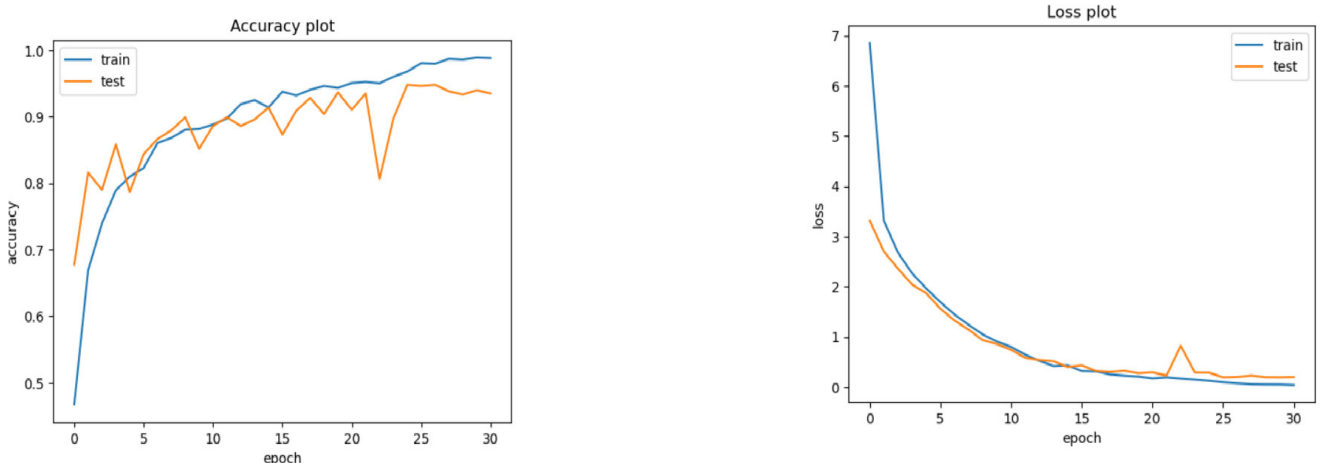

FIGURE 11. Accuracy and loss plots of 3BNet on the figshare dataset.

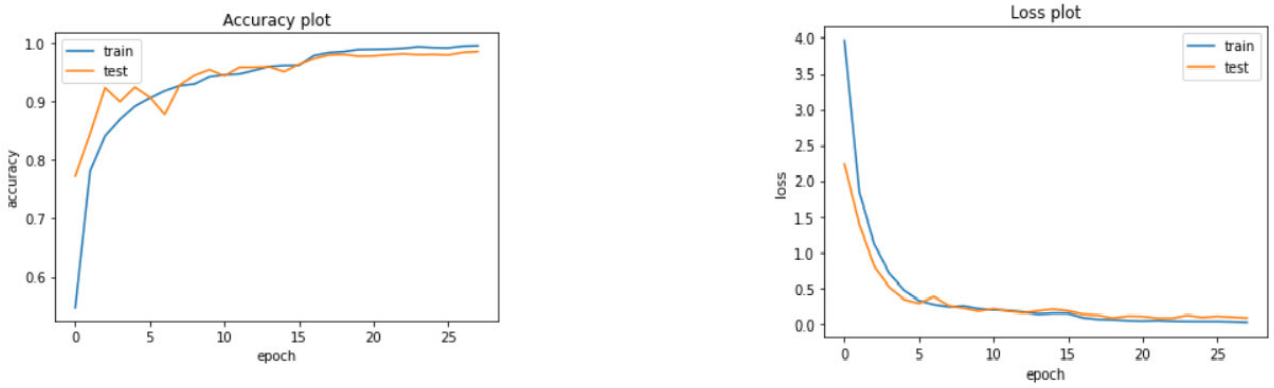

FIGURE 12. Accuracy and loss plots of 3BNet on augmented figshare dataset.

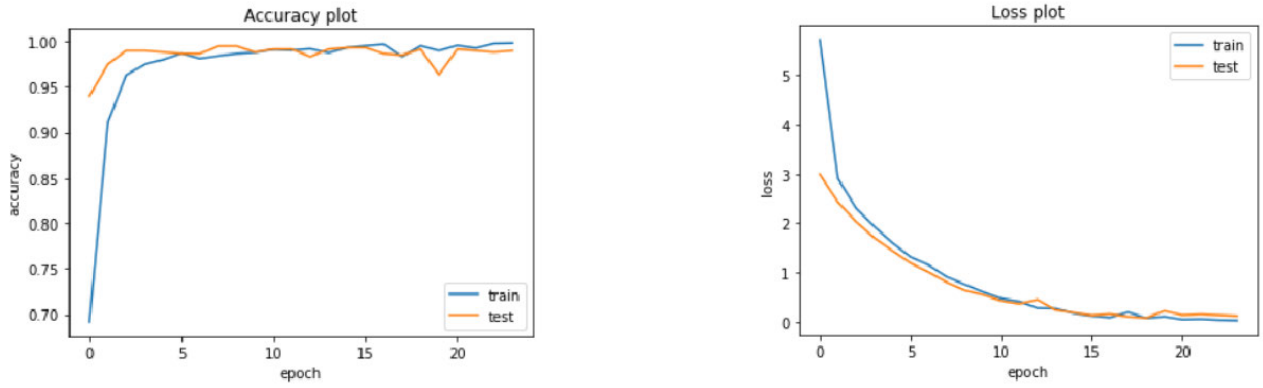

FIGURE 13. Accuracy and loss plots of 3BNet on brain tumor detection 2020 dataset.

such as the UCSF-PDGM dataset [38]. This will enable a more thorough evaluation of our model's robustness across different imaging protocols, scanner models, and demographic distributions. Additionally, we are exploring advanced bias mitigation techniques as highlighted in recent literature [39], [40]. These include pre-processing methods like data harmonization, in-processing techniques such as fairness-aware learning algorithms, and post-processing approaches like model calibration and pruning. Implementing these techniques will further enhance our model's fairness

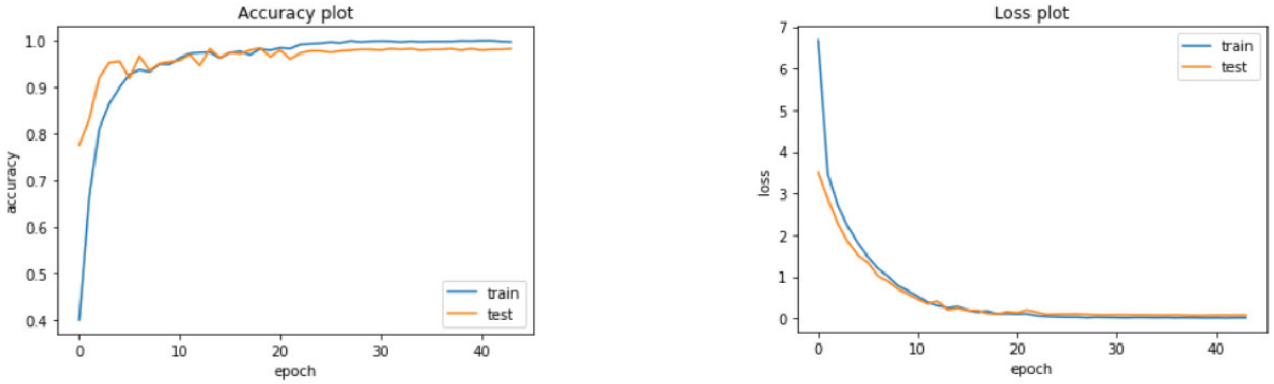

FIGURE 14. Accuracy and loss plots of 3BNet on brain tumor classification (MRI) dataset.

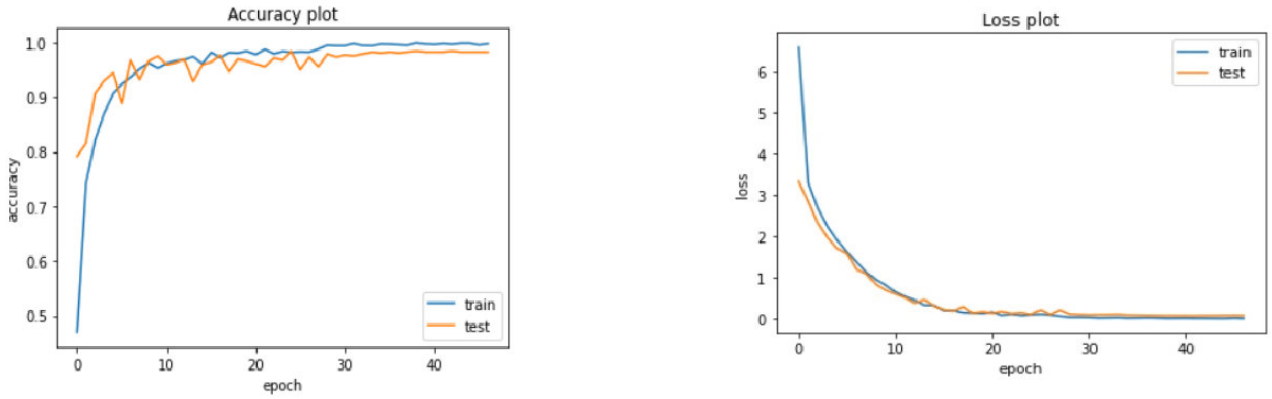

FIGURE 15. Accuracy and loss plots for binary class blind test evaluation.

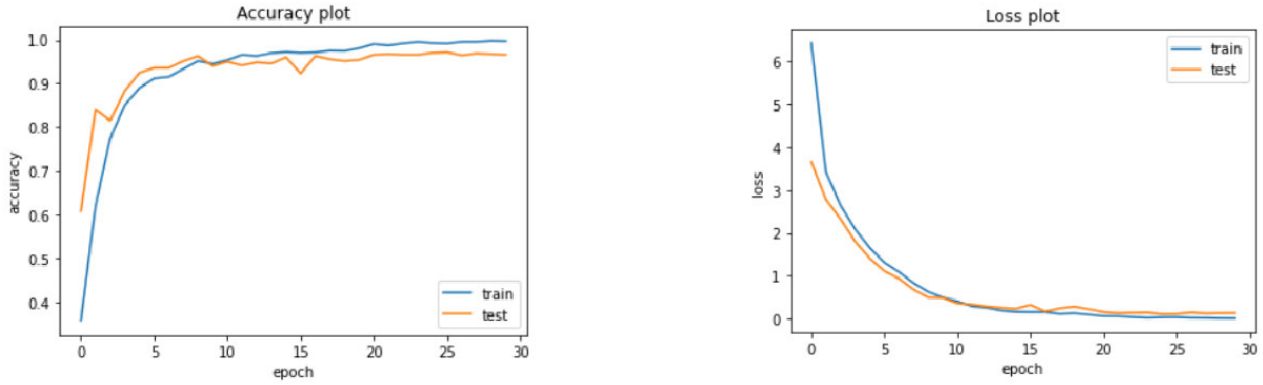

FIGURE 16. Accuracy and loss plots for four class blind test evaluation.

and generalizability across diverse patient populations and imaging conditions.

In conclusion, while our current study is constrained by the limited metadata in the utilized datasets, we have taken initial steps to enhance model robustness through data augmentation. We acknowledge the importance of assessing model performance across diverse imaging conditions and demographics, and we are committed to addressing these aspects more comprehensively in our future research.

I. GRAD-CAM VISUALIZATION OF TUMOR REGIONS

To further evaluate the interpretability of our 3B Net integrated with EfficientNetB2, we applied Grad-CAM to visualize the most relevant regions influencing the model's classification decisions. As shown in Figure 18, the generated heatmaps highlight key tumor-affected regions in MRI images, aiding in model transparency. The Grad-CAM visualizations confirm that our proposed model accurately highlights tumor regions, reinforcing its potential

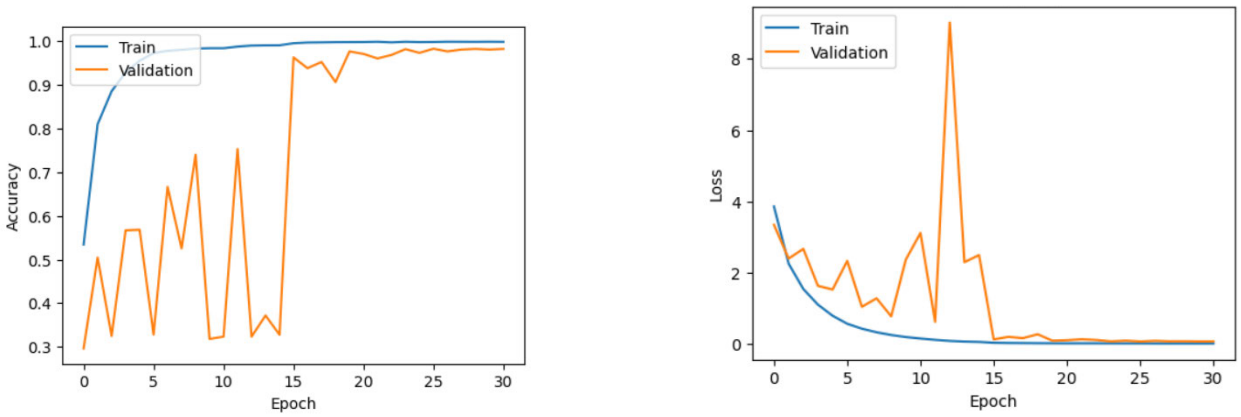

FIGURE 17. Accuracy and loss plots for three class five-fold cross-validation.

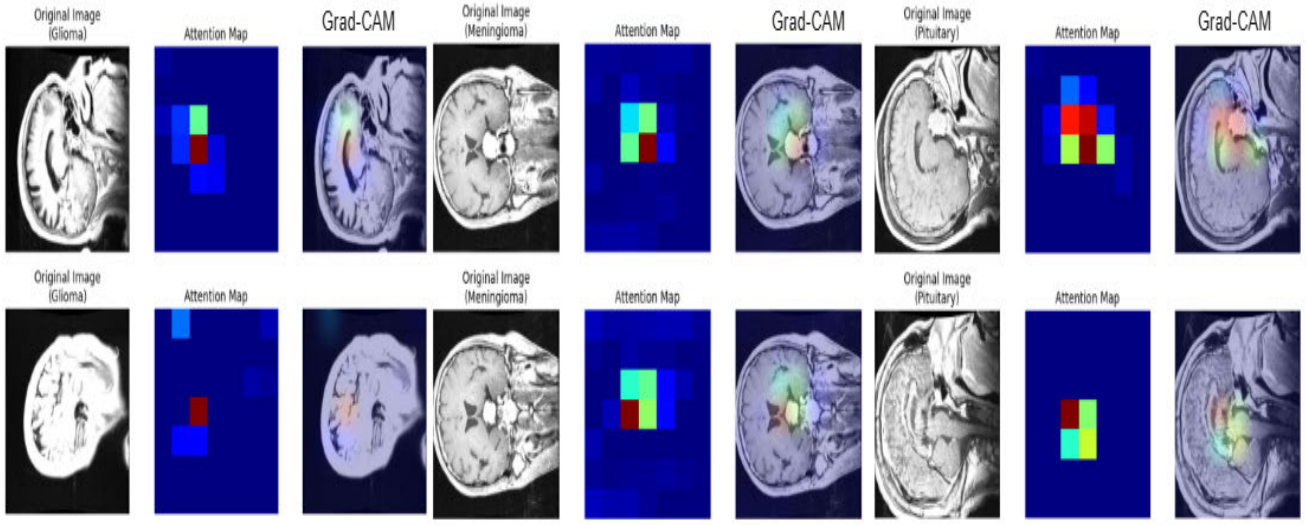

FIGURE 18. Grad-CAM Visualization.

clinical applicability. By ensuring model interpretability, this approach supports reliable decision-making in medical diagnostics.

J. COMPUTATIONAL EFFICIENCY ANALYSIS

In real-time clinical applications, inference time, computational expense, and memory requirements play a crucial role in determining the feasibility of deploying deep learning models. To ensure the practicality of our proposed Hybrid 3B Net with EfficientNetB2, we performed an evaluation of its inference speed, parameter count, and model size, comparing it with existing methods.

Table 12 presents a comparative analysis of our model against other EfficientNet-based approaches. The results indicate that our proposed method achieves a significantly lower inference time (69 ms) compared to the fine-tuned EfficientNetB2 models proposed by Zulfiqar et al. [16] and Babu Vimala et al. [41], which require 2 seconds for inference. Although our model has a higher parameter count (12.31M) than the compared methods, it achieves superior performance while maintaining a significantly reduced model size of 47.72 MB, making it more efficient for deployment in real-time clinical settings. These findings highlight the efficiency of our model, demonstrating a favorable trade-off between computational cost and classification accuracy. The reduced inference time makes it highly suitable for integration into real-time clinical workflows, where fast and accurate predictions are essential.

K. COMPARISON WITH STATE-OF-THE-ART METHODS

In this study, we compared the performance of our proposed 3B Net integrated with EfficientNetB2 against several state-of-the-art methods in brain tumor classification. The methods evaluated include traditional machine learning approaches like SVM and KNN, as used by Cheng et al. [25], and more advanced techniques such as Genetic Algorithm-CNN by Anaraki et al. [47] and CNN-based models by Afshar et al. [48]. Additionally, recent studies

TABLE 12. Comparison of inference time, parameters, and model size.

Reference	Model	Inference Time	Parameters (M)	Model Size (MB)
Zulfiqar et al. [16]	Fine-tuned EfficientNetB2	2 sec	7.72	90.0
Babu Vimala et al. [41]	Fine-tuned EfficientNetB2	2 sec	7.81	89.9
Proposed Method	**Hybrid 3B Net with EfficientNetB2**	**69 ms**	**12.31**	**47.72**

TABLE 13. Comparison with state-of-the-art methods.

Reference	Method	Classification Type	Accuracy (%)
Preetha et al. [42]	Fine-tuned EfficientNetB4	2 Class	99.33
Shah et al. [43]	Fine-tuned EfficientNetB0	2 Class	98.80
Seetha et al. [44]	CNN	2 Class	97.50
Khushi et al. [45]	Pre-trained customized EfficientNetB4	2 Class	98.61
Suryawanshi et al. [46]	Hybrid CNN-SVM	2 Class	98.01
Proposed Method	**3B Net integrated with EfficientNetB2**	**2 class**	**99.50**
Cheng et al. [25]	SVM & KNN	3 Class	91.28
Anaraki et al. [47]	Genetic Algorithm-CNN	3 Class	94.20
Afshar et al. [48]	CNN	3 Class	90.89
Suryawanshi et al. [46]	Hybrid CNN-SVM	3 Class	95.16
Ferdous et al. [49]	LCDEiT (Linear Complexity Data-Efficient image Transformer)	3 Class	98.11
Proposed Method	**3B Net integrated with EfficientNetB2**	**3 Class**	**98.72**
Mohsen et al. [50]	CNN	4 Class	96.97
Khan et al. [51]	Hierarchical Deep Learning-Based CNN	4 Class	92.13
Rajput et al. [24]	Fine-tuned VGG19, Inception-v3, and ResNet50	4 Class	90.00
Uysal et al. [52]	Vision Transformer (ViT)	4 Class	95.57
Ferdous et al. [49]	LCDEiT (Linear Complexity Data-Efficient image Transformer)	4 Class	93.69
Proposed Method	**3B Net integrated with EfficientNetB2**	**4 Class**	**97.80**

TABLE 14. Comparison of classification accuracy with 95% confidence intervals.

Method	Accuracy (%)	95% CI
Preetha et al. (Fine-tuned EfficientNetB4)	99.33	[98.82, 99.84]
Shah et al. (Fine-tuned EfficientNetB0)	98.8	[98.13, 99.47]
Seetha et al. (CNN)	97.5	[96.53, 98.47]
Khushi et al. (Pre-trained EfficientNetB4)	98.61	[97.88, 99.34]
Suryawanshi et al. (Hybrid CNN-SVM) - 2 Class	98.01	[97.14, 98.88]
Proposed Method (3B Net + EfficientNetB2) - 2 Class	**99.50**	**[99.06, 99.94]**
Cheng et al. (SVM & KNN)	91.28	[89.53, 93.03]
Anaraki et al. (GA-CNN)	94.2	[92.75, 95.65]
Afshar et al. (CNN)	90.89	[89.11, 92.67]
Suryawanshi et al. (Hybrid CNN-SVM) - 3 Class	95.16	[93.83, 96.49]
Ferdous et al. (LCDEiT)	98.11	[97.27, 98.95]
Proposed Method (3B Net + EfficientNetB2) - 3 Class	**98.72**	**[98.02, 99.42]**
Mohsen et al. (CNN)	96.97	[95.91, 98.03]
Khan et al. (Hierarchical CNN)	92.13	[90.46, 93.80]
Rajput et al. (Fine-tuned VGG19/Inception-v3/ResNet50)	90.0	[88.14, 91.86]
Uysal et al. (ViT)	95.57	[94.29, 96.85]
Ferdous et al. (LCDEiT) - 4 Class	93.69	[92.18, 95.20]
Proposed Method (3B Net + EfficientNetB2) - 4 Class	**97.80**	**[96.89, 98.71]**

employing Vision Transformers (ViT) [52] and Linear Complexity Data-Efficient image Transformers (LCDEiT) [49] were considered. These methods primarily focused on three-class classification tasks, similar to our approach. Our proposed method demonstrated significant improvements in accuracy, particularly after applying data augmentation. Before augmentation, the 3B Net achieved a classification accuracy of 94.95% in the three-class task, surpassing

several traditional DL methods. After augmentation, the accuracy increased substantially to 98.72%, showcasing the effectiveness of our data enhancement strategy.

Additionally, we evaluated our model in binary classification, where it achieved superior results compared to methods like the fine-tuned EfficientNetB0 used by Shah et al. [43], fine-tuned EfficientNetB4 used by Preetha et al. [42], hybrid CNN-SVM by Suryawanshi and Patil [46] and CNN used by Seetha and Raja [44]. Furthermore, in the four-class classification task, our method also exhibited strong performance, compared to methods proposed by Rajput et al. [24], Mohsen et al. [50],LCDEiT by Ferdous et al. [49], and Khan et al. [51], indicating its robustness across different classification scenarios. Notably, our method also surpasses ViT by Uysal and Erkan [52], demonstrating its superiority in handling complex classification tasks. This comprehensive comparison as shown in Table 13 highlights the efficacy of our 3B Net integrated with EfficientNetB2 in outperforming existing state-of-the-art methods, making it a highly competitive tool for brain tumor classification.

Since the methods listed in Table 13 were evaluated on different datasets, direct statistical hypothesis testing (e.g., t-tests) is not applicable for performance comparison. Instead, we report 95% confidence intervals (CIs) for accuracy to quantify the statistical reliability of each method. These confidence intervals provide a more robust comparison and ensure meaningful interpretation of performance differences.

To compute the CIs, we employed the Wilson score interval, which is suitable for proportion-based metrics such as classification accuracy. The 95% confidence interval for each method was calculated as:

$$CI = \hat{p} \pm Z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \tag{7}$$

where $\hat{p}$ represents the observed accuracy, $Z = 1.96$ corresponds to the 95% confidence level, and n is the total number of test samples used in each study. The reported CIs indicate the range within which the true accuracy is expected to fall with 95% confidence. A narrower CI reflects a more reliable accuracy estimate, while a wider CI suggests greater variability.

Our proposed 3B Net integrated with EfficientNetB2 demonstrates superior performance across binary, three-class, and four-class classifications, achieving 99.50%, 98.72%, and 97.80% accuracy, respectively, with consistently narrow confidence intervals. These results indicate the robustness and effectiveness of our method compared to existing approaches.

V. CONCLUSION

In this study, we explored various deep learning architectures for brain tumor classification using MRI images, ultimately demonstrating the superior performance of the 3B Net. The comprehensive comparison of single-branch to six-branch networks highlighted the 3B Net's ability to capture multi-scale features, significantly enhancing classification accuracy. Integrating EfficientNetB2's feature extraction capabilities further improved the model's performance, particularly in multi-class classification tasks.

Our experiments showed that data augmentation effectively addressed class imbalances, leading to better model generalization and robustness. The 3B Net achieved high accuracy, recall, precision, F1, and F2 scores in binary and multi-class classification tasks. Blind test evaluations and five-fold cross-validation confirmed the model's ability to generalize to unseen data, with strong performance across different tumor types. These results underscore the potential of the 3B Net as a valuable tool for automated brain tumor diagnosis, contributing to more accurate and efficient clinical decision-making.

VI. LIMITATIONS

Despite its promising performance, this study has some limitations. First, the evaluation was conducted on publicly available datasets, which may not fully represent the diversity of real-world clinical data. The absence of multi-institutional datasets limits the model's generalizability across different imaging protocols, scanner types, and patient demographics. Second, the dataset used lacks subject-level metadata, as it consists of pre-extracted 2D slices without information on the number of subjects or slices per subject. This limitation may introduce biases, such as multiple slices from the same patient appearing in both training and validation sets, potentially affecting model generalization. Third, while data augmentation addressed class imbalance, the model's performance may still be impacted by underrepresented tumor subtypes. Further validation on larger, more diverse datasets, including rare tumor types, is needed to enhance robustness.

Additionally, the computational complexity of the 3B Net model poses challenges for real-time deployment, particularly in hardware-constrained environments. Techniques such as model quantization and pruning should be explored to reduce computational overhead. Finally, this study focused solely on classification, without incorporating tumor segmentation or localization, which are crucial for comprehensive tumor analysis. External validation using independent datasets from different institutions is necessary to further assess the model's real-world applicability.

VII. FUTURE WORK

Future work could focus on further optimizing the 3B Net architecture and exploring its application to other medical imaging tasks. Additionally, the incorporation of advanced techniques such as attention mechanisms, transfer learning from larger datasets, and hybrid models combining CNNs with other types of neural networks could further enhance the model's performance.

Exploring the use of 3B Net in real-time diagnostic systems and its integration into clinical workflows will also be important steps towards practical implementation. Finally, expanding the dataset with more diverse and larger

samples, including multi-institutional and multi-modal data, could improve the model's robustness and generalizability, facilitating its broader adoption in medical imaging research and clinical practice.

VIII. DATA AVAILABILITY AND ETHICAL CONSIDERATIONS

This study utilizes publicly available medical imaging datasets obtained from Kaggle and Figshare (https://figshare.com/articles/dataset/brain_tumor_dataset/1512427/8, https://www.kaggle.com/datasets/ahmedhamada0/brain-tumor-detection, https://www.kaggle.com/datasets/sartajbhuvaji/brain-tumor-classification-mri, https://www.kaggle.com/datasets/mohamedmetwalysherif/braintumordataset, https://www.kaggle.com/datasets/navoneel/brain-mri-images-for-brain-tumor-detection). The dataset is licensed under Creative Commons Attribution (CC BY 4.0) and was used in accordance with the terms and conditions outlined by the dataset providers.

Ethical considerations were strictly followed to ensure compliance with regulations governing the use of medical images for research. Since the datasets used in this study are anonymized and publicly accessible, Institutional Review Board (IRB) approval was not required. No personally identifiable information or patient-sensitive data was processed in this research.

The study adheres to the principles outlined in the Declaration of Helsinki and follows ethical guidelines for medical research involving human subjects. Any reuse or redistribution of the dataset should comply with the original licensing agreements and ethical guidelines.

REFERENCES

[1] Q. T. Ostrom, M. Price, C. Neff, G. Cioffi, K. A. Waite, C. Kruchko, and J. S. Barnholtz-Sloan, "Cbtrus statistical report: Primary brain and other central nervous system tumors diagnosed in the United States in 2016–2020," *Neuro-Oncology*, vol. 25, pp. 1–100, Oct. 2023.

[2] A. C. Soc. *Key Statistics About Brain and Spinal Cord Tumors in Adults*. Accessed: Mar. 19, 2025. [Online]. Available: https://www.cancer.org/cancer/types/brain-spinal-cord-tumors-adults/about/key-statistics.html

[3] V. Health. (2023). *Brain Tumor Facts and Statistics*. Accessed: Aug. 11, 2024. [Online]. Available: https://www.verywellhealth.com/brain-tumor-facts-and-statistics-what-you-need-to-know-5524668

[4] Weill Cornell Med. (2023). *Early Detection Can be Key to Surviving a Brain Tumor*. Accessed: Aug. 11, 2024. [Online]. Available: https://neurosurgery.weillcornell.org/about-us/blog/early-detection-can-be-key-surviving-brain-tumor

[5] J. Amin, M. Sharif, A. Haldorai, M. Yasmin, and R. S. Nayak, "Brain tumor detection and classification using machine learning: A comprehensive survey," *Complex Intell. Syst.*, vol. 8, no. 4, pp. 3161–3183, Aug. 2022.

[6] R. Kaifi, "A review of recent advances in brain tumor diagnosis based on AI-based classification," *Diagnostics*, vol. 13, no. 18, p. 3007, Sep. 2023.

[7] S. Saeedi, S. Rezayi, H. Keshavarz, and S. R. Niakan Kalhori, "MRI-based brain tumor detection using convolutional deep learning methods and chosen machine learning techniques," *BMC Med. Informat. Decis. Making*, vol. 23, no. 1, p. 16, Jan. 2023.

[8] R. Preetha, M. J. P. Priyadarsini, and J. S. Nisha, "Comparative study on architecture of deep neural networks for segmentation of brain tumor using magnetic resonance images," *IEEE Access*, vol. 11, pp. 138549–138567, 2023.

[9] D. Rastogi, P. Johri, V. Tiwari, and A. A. Elngar, "Multi-class classification of brain tumour magnetic resonance images using multi-branch network with inception block and five-fold cross validation deep learning framework," *Biomed. Signal Process. Control*, vol. 88, Feb. 2024, Art. no. 105602.

[10] T. Agrawal, P. Choudhary, V. Kumar, P. Singh, M. Diwakar, and S. Kumar, "A comparative study of brain tumor classification on unbalanced dataset using deep neural networks," *Biomed. Signal Process. Control*, vol. 94, Aug. 2024, Art. no. 106256.

[11] B. Sandhiya and S. Kanaga Suba Raja, "Deep learning and optimized learning machine for brain tumor classification," *Biomed. Signal Process. Control*, vol. 89, Mar. 2024, Art. no. 105778.

[12] S. Deepak and P. M. Ameer, "Brain tumor classification using deep CNN features via transfer learning," *Comput. Biol. Med.*, vol. 111, Aug. 2019, Art. no. 103345.

[13] E. Şahin, D. Özdemir, and H. Temurtaş, "Multi-objective optimization of ViT architecture for efficient brain tumor classification," *Biomed. Signal Process. Control*, vol. 91, May 2024, Art. no. 105938.

[14] S. Khoramipour, M. Gandomkar, and M. Shakiba, "Enhancement of brain tumor classification from MRI images using multi-path convolutional neural network with SVM classifier," *Biomed. Signal Process. Control*, vol. 93, Jul. 2024, Art. no. 106117.

[15] 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 Appl.*, vol. 82, no. 20, pp. 31709–31736, Aug. 2023.

[16] F. Zulfiqar, U. Ijaz Bajwa, and Y. Mehmood, "Multi-class classification of brain tumor types from MR images using EfficientNets," *Biomed. Signal Process. Control*, vol. 84, Jul. 2023, Art. no. 104777.

[17] R. L. Kumar, J. Kakarla, B. V. Isunuri, and M. Singh, "Multi-class brain tumor classification using residual network and global average pooling," *Multimedia Tools Appl.*, vol. 80, no. 9, pp. 13429–13438, Apr. 2021.

[18] N. F. Aurna, M. A. Yousuf, K. A. Taher, A. K. M. Azad, and M. A. Moni, "A classification of MRI brain tumor based on two stage feature level ensemble of deep CNN models," *Comput. Biol. Med.*, vol. 146, Jul. 2022, Art. no. 105539.

[19] A. Kazemi, M. E. Shiri, A. Sheikhahmadi, and M. Khodamoradi, "Classifying tumor brain images using parallel deep learning algorithms," *Comput. Biol. Med.*, vol. 148, Sep. 2022, Art. no. 105775.

[20] A. Vidyarthi, R. Agarwal, D. Gupta, R. Sharma, D. Draheim, and P. Tiwari, "Machine learning assisted methodology for multiclass classification of malignant brain tumors," *IEEE Access*, vol. 10, pp. 50624–50640, 2022.

[21] B. S. Abd El-Wahab, M. E. Nasr, S. Khamis, and A. S. Ashour, "BTC-fCNN: Fast convolution neural network for multi-class brain tumor classification," *Health Inf. Sci. Syst.*, vol. 11, no. 1, p. 3, Jan. 2023.

[22] S. Asif, M. Zhao, X. Chen, and Y. Zhu, "BMRI-NET: A deep stacked ensemble model for multi-class brain tumor classification from MRI images," *Interdiscipl. Sci., Comput. Life Sci.*, vol. 15, no. 3, pp. 499–514, Sep. 2023.

[23] B. V. Isunuri and J. Kakarla, "EfficientNet and multi-path convolution with multi-head attention network for brain tumor grade classification," *Comput. Electr. Eng.*, vol. 108, May 2023, Art. no. 108700.

[24] I. S. Rajput, A. Gupta, V. Jain, and S. Tyagi, "A transfer learning-based brain tumor classification using magnetic resonance images," *Multimedia Tools Appl.*, vol. 83, no. 7, pp. 20487–20506, Aug. 2023.

[25] J. Cheng, W. Huang, S. Cao, R. Yang, W. Yang, Z. Yun, Z. Wang, and Q. Feng, "Enhanced performance of brain tumor classification via tumor region augmentation and partition," *PLoS ONE*, vol. 10, no. 10, Oct. 2015, Art. no. e0140381.

[26] H. H. Sultan, N. M. Salem, and W. Al-Atabany, "Multi-classification of brain tumor images using deep neural network," *IEEE Access*, vol. 7, pp. 69215–69225, 2019.

[27] J. Cheng, "Brain tumor dataset," 2017, Figshare, doi: 10.6084/m9.figshare.1512427.v8.

[28] A. Hamada. (2023). *Brain Tumor Detection Dataset*. Kaggle. Accessed: Mar. 6, 2025. [Online]. Available: https://www.kaggle.com/datasets/ahmedhamada0/brain-tumor-detection

[29] S. Bhuvaji. (2021). *Brain Tumor Classification MRI Dataset*. Kaggle. Accessed: Mar. 6, 2025. [Online]. Available: https://www.kaggle.com/datasets/sartajbhuvaji/brain-tumor-classification-mri

[30] M. M. Sherif. (2023). *Brain Tumor Dataset*. Kaggle. Accessed: Mar. 6, 2025. [Online]. Available: https://www.kaggle.com/datasets/mohamedmetwalysherif/braintumordataset

[31] N. Chakrabarty. (2021). *Brain MRI Images for Brain Tumor Detection*. Kaggle. Accessed: Mar. 6, 2025. [Online]. Available: https://www.kaggle.com/datasets/navoneel/brain-mri-images-for-brain-tumor-detection
[32] B. Bischl, M. Binder, M. Lang, T. Pielok, J. Richter, S. Coors, J. Thomas, T. Ullmann, M. Becker, A.-L. Boulesteix, D. Deng, and M. Lindauer, "Hyperparameter optimization: Foundations, algorithms, best practices, and open challenges," *WIREs Data Mining Knowl. Discovery*, vol. 13, no. 2, p. 1484, Mar. 2023.
[33] H. Li, P. Chaudhari, H. Yang, M. Lam, A. Ravichandran, R. Bhotika, and S. Soatto, "Rethinking the hyperparameters for fine-tuning," 2020, *arXiv:2002.11770*.
[34] D. M. Belete and M. D. Huchaiah, "Grid search in hyperparameter optimization of machine learning models for prediction of HIV/AIDS test results," *Int. J. Comput. Appl.*, vol. 44, no. 9, pp. 875–886, Sep. 2022.
[35] J. Bergstra and Y. Bengio, "Random search for hyper-parameter optimization," *J. Mach. Learn. Res.*, vol. 13, no. 1, pp. 281–305, Mar. 2012.
[36] V. Nguyen, "Bayesian optimization for accelerating hyper-parameter tuning," in *Proc. IEEE 2nd Int. Conf. Artif. Intell. Knowl. Eng. (AIKE)*, Jun. 2019, pp. 302–305.
[37] Z. Hussain, F. J. G. Fuentes-Guerra, D. Yi, and D. L. Rubin, "Differential data augmentation techniques for medical imaging classification tasks," in *AMIA Annu. Symp. Proc.*, vol. 2017, Jan. 2017, pp. 979–984.
[38] E. Calabrese, J. E. Villanueva-Meyer, J. D. Rudie, A. M. Rauschecker, U. Baid, S. Bakas, S. Cha, J. T. Mongan, and C. P. Hess, "The university of California San Francisco preoperative diffuse glioma MRI dataset," *Radiol., Artif. Intell.*, vol. 4, no. 6, Nov. 2022, Art. no. 220058.
[39] Z. Xu, J. Li, Q. Yao, H. Li, M. Zhao, and S. K. Zhou, "Addressing fairness issues in deep learning-based medical image analysis: A systematic review," *NPJ Digit. Med.*, vol. 7, no. 1, p. 286, Oct. 2024.
[40] A. S. Tejani, Y. S. Ng, Y. Xi, and J. C. Rayan, "Understanding and mitigating bias in imaging artificial intelligence," *RadioGraphics*, vol. 44, no. 5, May 2024, Art. no. 230067.
[41] B. Babu Vimala, S. Srinivasan, S. K. Mathivanan, Mahalakshmi, P. Jayagopal, and G. T. Dalu, "Detection and classification of brain tumor using hybrid deep learning models," *Sci. Rep.*, vol. 13, no. 1, p. 23029, Dec. 2023.
[42] R. Preetha, M. J. P. Priyadarsini, and J. S. Nisha, "Automated brain tumor detection from magnetic resonance images using fine-tuned EfficientNet-B4 convolutional neural network," *IEEE Access*, vol. 12, pp. 112181–112195, 2024.
[43] H. A. Shah, F. Saeed, S. Yun, J.-H. Park, A. Paul, and J.-M. Kang, "A robust approach for brain tumor detection in magnetic resonance images using finetuned EfficientNet," *IEEE Access*, vol. 10, pp. 65426–65438, 2022.
[44] J. Seetha and S. S. Raja, "Brain tumor classification using convolutional neural networks," *Biomed. Pharmacol. J.*, vol. 11, no. 3, pp. 1457–1461, Sep. 2018.
[45] H. M. T. Khushi, T. Masood, A. Jaffar, and S. Akram, "A novel approach to classify brain tumor with an effective transfer learning based deep learning model," *Brazilian Arch. Biol. Technol.*, vol. 67, Apr. 2024, Art. no. e24231137.
[46] S. Suryawanshi and S. B. Patil, "Efficient brain tumor classification with a hybrid CNN-SVM approach in MRI," *J. Adv. Inf. Technol.*, vol. 15, no. 3, pp. 340–354, 2024.
[47] A. Kabir Anaraki, M. Ayati, and F. Kazemi, "Magnetic resonance imaging-based brain tumor grades classification and grading via convolutional neural networks and genetic algorithms," *Biocybern. Biomed. Eng.*, vol. 39, no. 1, pp. 63–74, Jan. 2019.
[48] P. Afshar, K. N. Plataniotis, and A. Mohammadi, "Capsule networks for brain tumor classification based on MRI images and coarse tumor boundaries," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2019, pp. 1368–1372.
[49] G. J. Ferdous, K. A. Sathi, Md. A. Hossain, M. M. Hoque, andM. A. A. Dewan, "LCDEiT: A linear complexity data-efficient image transformer for MRI brain tumor classification," *IEEE Access*, vol. 11, pp. 20337–20350, 2023.
[50] H. Mohsen, E.-S.-A. El-Dahshan, E.-S.-M. El-Horbaty, and A.-B.-M. Salem, "Classification using deep learning neural networks for brain tumors," *Future Comput. Informat. J.*, vol. 3, no. 1, pp. 68–71, Jun. 2018.
[51] A. H. Khan, S. Abbas, M. A. Khan, U. Farooq, W. A. Khan, S. Y. Siddiqui, and A. Ahmad, "Intelligent model for brain tumor identification using deep learning," *Appl. Comput. Intell. Soft Comput.*, vol. 2022, pp. 1–10, Jan. 2022.
[52] F. Uysal and M. Erkan, "Multiclass classification of brain tumors with various deep learning models," *Eng. Proc.*, vol. 27, no. 1, p. 30, 2022.

R. PREETHA received the B.Tech. degree in electronics and communication engineering from M. G. University, Kottayam, Kerala, India, in 2009, and the M.E. degree in applied electronics from Anna University, Chennai, Tamil Nadu, India, in 2013. She is currently pursuing the Ph.D. degree with Vellore Institute of Technology. Her research interests include medical image processing, machine learning, and deep learning.

M. JASMINE PEMEENA PRIYADARSINI received the B.E. degree in electronics and communication engineering from the University of Madras, Chennai, Tamil Nadu, India, in 1992, the M.E. degree in microwave and optical communication from Madurai Kamaraj University, Madurai, Tamil Nadu, in 1995, and the Ph.D. degree from Vellore Institute of Technology, Vellore, Tamil Nadu, in 2016. She is currently a Professor and the Associate Dean with the School of Electronics Engineering, Vellore Institute of Technology. Her area of research interests include digital signal and image processing, wireless communication, applied electronics, and optical image processing.

J. S. NISHA received the B.Tech. degree in electronics and communication engineering from the University College of Engineering, Kariyavattom, the M.Tech. degree in signal processing from the College of Engineering Trivandrum, India, and the Ph.D. degree from the National Institute of Technology, Tiruchirappalli, India. She is currently an Assistant Professor (Grade 1) with the Department of Sensor and Biomedical Technology, School of Electronics Engineering, Vellore Institute of Technology, Vellore. Her research interests include medical image processing, machine learning, and deep learning. More specifically, cancer detection from colonoscopy images.