STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **x**, Month 2025, pp 0–14.
Published online in International Academic Press (www.IAPress.org)

A Fine-Funed CNN for Multiclass Classification Of Brain tumors On Figshare CE-MRI and its Raspberry Pi Deployment

Alaeddine Hmidi[1,2,*], Neji Kouka [3], Lina Tekari [4]

[1] *Kairouan University, Higher Institute of Applied Sciences and Technology of Kasserine, 1200, Kasserine, Tunisia*
[2] *Monastir University,Faculty of Sciences, Laboratory of Electronics and Microelectronics, LR99ES30, 5000, Monastir, Tunisia*
[3] *MARS laboratory, ISITCom hammam sousse, Sousse University, Sousse, Tunisia*
[4] *Lucian Blaga University of Sibiu, Faculty of Medicine, Bulevardul Victoriei 10, Sibiu 550024, Romania*

Abstract This paper introduces a fine-tuned Convolutional Neural Network (CNN) for multiclass classification of brain tumors on contrast-enhanced T1-weighted MRI scans. The proposed model integrates batch normalization, dropout, and lightweight convolutional blocks to extract discriminative features while maintaining computational efficiency suitable for embedded deployment. Experiments were conducted on the Figshare dataset comprising 3,064 MRI slices from 233 patients with gliomas, meningiomas, or pituitary tumors. Images were preprocessed through resizing, normalization.
The model was trained using the Adam optimizer with a learning rate of 1e-4, a batch size of 32, and 100 epochs. Evaluation metrics included accuracy, precision, recall, and F1-score. The fine-tuned CNN achieved an overall accuracy of 94.08%, with class-specific performance indicating strong results for pituitary tumors (precision 95.65%, recall 95.96%) and meningiomas (precision 90.20%, recall 88.81%), while glioma classification showed high sensitivity (recall 96.85%) but lower precision (75.00%). To validate real-world applicability, the model was converted to TensorFlow Lite and deployed on a Raspberry Pi 4, achieving an inference time of approximately 60 ms per image.
These findings demonstrate that fine-tuned CNNs can offer a competitive and resource-efficient solution for computer-aided diagnosis of brain tumors, balancing accuracy and practicality in clinical environments with limited computational resources.

Keywords Deep learning, Fine-Tuned CNN, Brain tumor, MRI, Raspberry Pi, Transfer learning.

DOI: 10.19139/soic-2310-5070-2635

1. Introduction

Brain tumors represent one of the most complex and severe medical challenges due to their potential to impair vital functions and their significant heterogeneity in terms of type, location, and behavior. According to the World Health Organization (WHO), brain tumors are among the deadliest forms of cancer, with nearly 300,000 new cases diagnosed worldwide annually.

These tumors can develop in various areas of the brain, potentially leading to profound effects on patients' cognitive and motor skills. Brain tumors are generally categorized into two main types: benign and malignant. Malignant tumors are particularly hazardous due to their ability to invade surrounding tissues rapidly, while benign tumors, despite being non-cancerous, can still pose significant risks if they exert pressure on critical brain areas. These complexities make the diagnosis and treatment of brain tumors particularly challenging.

Classifying brain tumors is crucial for guiding treatment decisions and assessing the prognosis. However, achieving early and accurate diagnosis remains difficult due to the variability of symptoms, the intricate structure of the brain, and the diverse nature of tumor types.

*Correspondence to: Alaeddine Hmidi (Email: alaeddine.hmidi@issatkas.u-kairouan.tn). Kairouan University, Higher Institute of Applied Sciences and Technology of Kasserine, 1200, Kasserine, Tunisia.

ISSN 2310-5070 (online) ISSN 2311-004X (print)

Conventional diagnostic methods for brain tumors include biopsies, cerebrospinal fluid (CSF) analysis, and imaging techniques such as radiography. While biopsies allow for tissue examination, they carry risks such as infection and bleeding, with an accuracy rate of about 86.97% [1]. CSF analysis also involves risks like bleeding and allergic reactions [2]. Radiographic imaging, although useful, exposes patients to radiation, thereby increasing the risk of secondary cancers. Currently, magnetic resonance imaging (MRI) is widely used, providing detailed images of the brain to localize and characterize tumors. However, the interpretation of MRI scans relies heavily on human expertise, leading to potential diagnostic errors due to subjectivity, image complexity, and the large amount of data that needs to be analyzed. Furthermore, the diagnostic process is often time-consuming and expensive, highlighting the need for faster and more reliable methods.

To address these issues, computer-aided diagnosis (CAD) systems have emerged as a promising solution to automate tumor detection, reducing the need for human intervention and enabling quicker and more accurate analysis of MRI images. CAD systems can analyze MRI scans and generate precise diagnostic reports to assist radiologists [3]. The integration of machine learning (ML) and deep learning (DL) techniques within CAD systems has significantly improved the accuracy of tumor detection [4] [5] [6].

Traditional ML methods typically involve several steps, including preprocessing, feature extraction, and classification. However, these techniques often rely on manual feature extraction, which can result in the loss of important information [7]. Recent advances in AI, particularly in DL, offer an effective alternative. Convolutional Neural Networks (CNNs), a specialized DL architecture, have shown great promise in interpreting complex medical images, especially for tumor detection and classification. CNNs can automatically extract relevant features from images without the need for manual intervention during feature extraction [8]. These models enable faster and more accurate tumor classification, significantly reducing human errors, enhancing diagnostic precision, and speeding up analysis times, which is crucial in clinical settings where time is critical.

Despite the progress made, current models still face limitations, particularly in handling image variability, mitigating overfitting, and optimizing feature extraction. To overcome these challenges, it is necessary to develop more robust and accurate AI models capable of processing the complex and diverse data from brain tumor MRI images efficiently.

This paper proposes an innovative approach using an optimized Convolutional Neural Network, referred to as Fine-Tuned Convolutional Neural Network (CNN), for brain tumor classification. This method utilizes an advanced architecture that incorporates regularization techniques and convolutional kernel optimization to extract finer and more relevant features from MRI images. The primary goal of this study is to develop a model capable of classifying three major types of brain tumors: gliomas, meningiomas, and pituitary tumors, using MRI images from various modalities. The model was evaluated using the Figshare dataset, a public collection of 3,064 brain tumor MRI images [9], and demonstrated impressive results with an accuracy of 94.08%, outperforming traditional models. The main contributions of this study are:

- Development of a lightweight and optimized CNN for multiclass brain tumor classification on MRI (glioma, meningioma, pituitary), achieving high performance (94.08% accuracy), competitive with state-of-the-art models while remaining simple and efficient.
- Successful deployment on Raspberry Pi 4 with real-time inference (60 ms/image), demonstrating feasibility in resource-constrained clinical environments.

In conclusion, this study demonstrates how AI, particularly convolutional neural networks, can revolutionize brain tumor diagnosis by providing a faster, more accurate, and more reliable method compared to current approaches. This model could significantly enhance computer-aided diagnostic (CAD) systems, support radiologists in clinical decision-making, and improve the efficiency of patient treatment processes.

The rest of this paper is as follows: Section 2 provides an overview of related studies. Section 3 introduces the database along with the preprocessing steps. Section 4 describes the proposed model, while Section 5 focuses on the training protocol and implementation aspects. Section 6 explains the evaluation metrics. Section 7 discusses the tumor classification results and compares them with existing approaches. Section 8 addresses inference and deployment on embedded devices, with a case study on the Raspberry Pi. Finally, Section 9 concludes the work and highlights potential directions for future research.

2. A brief review of existing methods

Recent advances in machine learning (ML) and deep learning (DL) have significantly improved the accuracy of brain tumor detection [10–30]. Numerous studies have employed a variety of techniques ranging from traditional feature extraction methods to sophisticated models like Convolutional Neural Networks (CNNs) [13] and Capsule Networks (CapsNet) [15]. These approaches have demonstrated substantial improvements in classification performance, particularly when combined with preprocessing techniques, transfer learning, and innovative architectures. Below, we summarize key works in this field, highlighting their methods, classifiers, and performance results. Table 1 provides a summary of relevant work in brain tumor detection.

Reference	Method/Approach	Techniques Used	Classifier/Model	Accuracy (%)
Cheng et al. [10]	Manual tumor contour extraction	GLCM, Density histograms, BoW	SVM	High accuracy
Ismail et al. [11]	Statistical feature extraction	Gabor filters, DWT	MLP	91.9
Tahir et al. [12]	Image preprocessing	Noise reduction, edge detection, contrast enhancement	SVM	86
Paul et al. [14]	CNN	CNN	CNN	90.26
Afshar et al. [15]	Capsule Networks	Capsule Networks	Capsule Networks	Not specified
Sultan et al. [16]	16-layer CNN	CNN	CNN	96.1 - 98.7
Hossain et al. [17]	Fuzzy C-Means clustering with CNN	CNN + Fuzzy C-Means	CNN	97.9
Anaraki et al. [20]	CNN + Genetic Algorithms	CNN, GA	CNN	90.9
Talo et al. [21]	ResNet34 for transfer learning	ResNet34	CNN	Not specified
Swati et al. [22]	Fine-tuning of VGG19	VGG19	CNN	Not specified
Lu et al. [23]	Enhanced AlexNet	AlexNet	CNN	90.7
Sajjad et al. [24]	Augmented VGG19	VGG19	CNN	90.7
PDCNN [25]	Parallel Dilated CNN	Dilated convolutions, ensemble models	PDCNN	98.35 - 98.67
[26]	Hybrid methods	PLS, pre-trained DL models	Hybrid (SVM, KNN, etc.)	98.95
[27]	EfficientNet with transfer learning	EfficientNetB0-B4, fine-tuning	EfficientNet	99.06

Table 1. Summary of Related Works in Brain Tumor Detection

The studies reviewed here highlight several important developments in brain tumor classification. Traditional feature extraction methods, such as Gray Level Co-occurrence Matrix (GLCM), Density Histograms, and Discrete Wavelet Transform (DWT), have been used to extract relevant tumor features and provide robust classification results. In particular, the use of Support Vector Machine (SVM) classifiers and Multilayer Perceptron (MLP) classifiers has been common in earlier works, showing respectable accuracy rates. In more recent years, deep learning models, particularly Convolutional Neural Networks (CNNs), have achieved state-of-the-art results in brain tumor detection. These models benefit from the ability to automatically learn relevant features from raw image data, eliminating the need for manual feature extraction. Innovative architectures, such as Capsule Networks (CapsNet) [15] and Parallel Dilated Convolutional Neural Networks (PDCNN), have shown promise in overcoming the limitations of traditional CNNs, capturing both global and local features more effectively.

The integration of transfer learning, using pre-trained models such as VGG19 [31], ResNet34 [32], and EfficientNet [33], has further boosted the performance of tumor detection systems, especially when coupled with data augmentation techniques. These hybrid approaches, combining deep learning with traditional machine learning techniques like fuzzy clustering and genetic algorithms, have yielded even better results, as seen in the high accuracies achieved by Hossain et al. [17] and Anaraki et al. [20].

Overall, the increasing use of deep learning techniques, especially CNNs and their variants, along with advanced preprocessing and transfer learning strategies, marks a significant improvement in the field of brain tumor detection. Despite these advancements, challenges remain, such as improving model generalization across diverse populations and ensuring the interpretability of these models in clinical settings. Our work aims to address these challenges by developing more robust and interpretable models, ultimately enhancing the accuracy and reliability of brain tumor diagnostics.

3. Dataset and Preprocessing:

In order to assess the performance of our proposed model, we used the CE-MRI dataset available on Figshare [9]. This comprehensive dataset includes 3,064 contrast-enhanced T1-weighted MRI slices from 233 patients diagnosed

with meningiomas, gliomas, or pituitary tumors. The data was collected between 2005 and 2010 from hospitals across China. Initially, the images had a resolution of 512×512 pixels but were downsampled to 256×256 pixels to optimize computational efficiency. The dataset is organized into three anatomical views—axial, coronal, and sagittal—and is distributed as follows: 708 slices for meningiomas, 1,426 slices for gliomas, and 930 slices for pituitary tumors. Figure 1 shows an example of a Figshare dataset. (a) Glioma. (b) Meningioma. (c) Pituitary tumor.

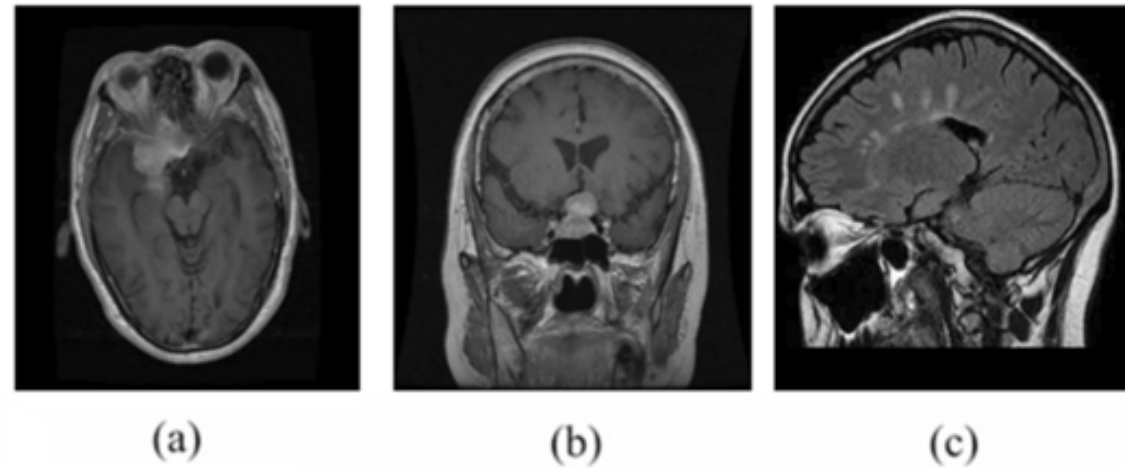


Figure 1. Sample of Figshare dataset. a Glioma. b Meningioma. c Pituitary tumor.

The original images, stored in the MATLAB-compatible .mat format, were converted into NumPy arrays to facilitate handling in Python. Finally, the dataset was split into training, validation, and testing sets, with 80% allocated for training and 20% reserved for validation and testing. All images were normalized by scaling pixel values and resized to the model's input dimensions of 256 × 256 pixels, ensuring consistency across the dataset.

4. Proposed Model

In this work, we introduce a refined Convolutional Neural Network architecture for multiclass classification of brain tumors in MRI images. Unlike traditional binary classification, which only distinguishes the presence or absence of a tumor, our model provides a more detailed analysis by categorizing various types of brain tumors, such as glioblastoma, meningioma, and others, thereby supporting more accurate and personalized clinical decision-making.

The proposed Fine-Tuned CNN architecture, as shown in Figure 2, incorporates several enhancements over standard models to improve performance on complex image analysis tasks, including classification and detection. Key modifications include the use of Batch Normalization (BN) and Dropout techniques to regularize the model and prevent overfitting, while ensuring faster convergence and high accuracy. These improvements allow the model to handle variations in input data, such as differences in lighting, orientation, or perspective, by leveraging data augmentation strategies.

This architecture is specifically designed for deployment in real-world clinical settings, with optimization for environments with limited computational resources, such as mobile devices or embedded systems. The model's compatibility with TensorFlow Lite ensures its practicality for real-time applications, enabling fast and efficient tumor detection in clinical practice. The training pipeline for the model follows a structured process consisting of five main steps:

1. **Input Step**: MRI images are collected and split into training, validation, and testing datasets for the training process.

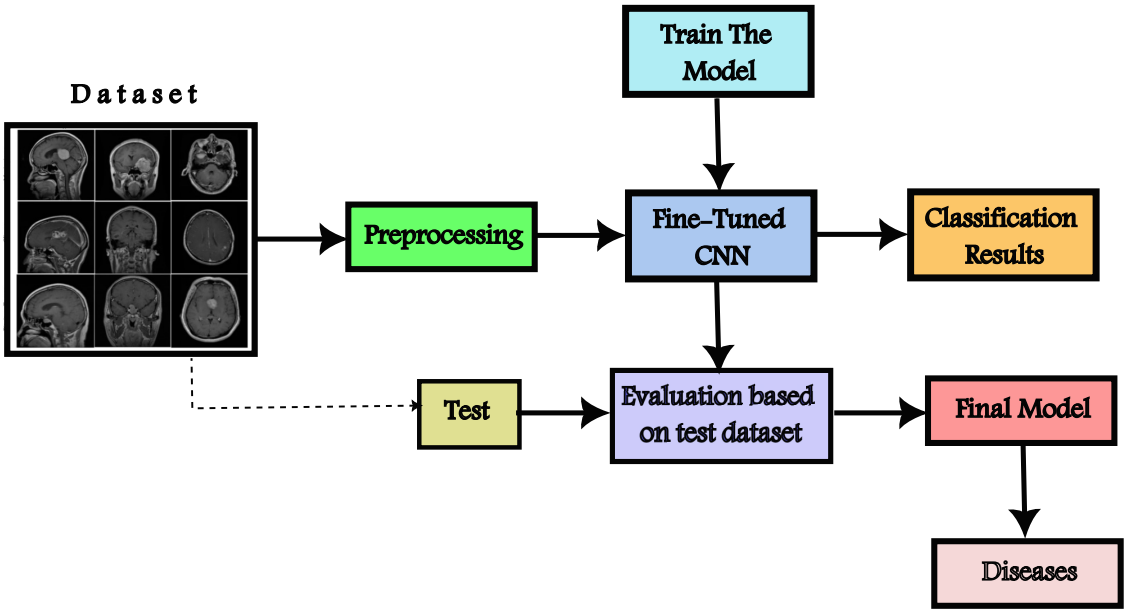


Figure 2. Illustration of proposed Model.

2. **Image Preprocessing**: Images are normalized by scaling pixel values from the range [0, 255] to [0, 1], and resized to a standard size (256 × 256 pixels) to match the model's input dimensions.
3. **Model Training**:
   - **Input**: Augmented images are fed into the network.
   - **Feature Extraction**: The model extracts essential features such as edges, textures, and regions of interest indicative of possible tumor locations.
   - **Tumor Detection**: The extracted features are analyzed to locate and classify potential tumor areas in the MRI images.
4. **Classification and Evaluation**: The model is evaluated on previously unseen data to determine its accuracy in classifying different tumor types.
5. **Final Model**: After evaluation, the model is fine-tuned and validated to ensure it provides robust tumor classifications in real-world applications.

The structure of the proposed model is outlined in **Table 2**, which details the network layers and their operations.

4.1. Model Architecture

The network begins with several Convolutional Layers designed to capture low-level features such as edges and textures. These layers progressively increase in complexity, with the network learning to identify more intricate features like tumor shapes and boundaries as it deepens.

Each convolutional layer is followed by a MaxPooling Layer , which reduces the spatial resolution of the feature maps, helping the network focus on the most significant aspects of the images. This helps in minimizing computational cost while maintaining key information necessary for classification.

To prevent overfitting, a Dropout Layer is introduced after the first fully connected Dense Layer , with a dropout rate of 50%. This improves generalization and ensures the model does not rely too heavily on any single feature, enhancing its ability to perform well on unseen data.

Finally, a Softmax Layer at the output produces a classification score across three categories, corresponding to different tumor types, ensuring that the model can effectively classify images into one of the three tumor classes.

Table 2. A Fine-Tuned CNN Architecture for Tumor Classification.

Layer	Operation	Input Size	Output Size	Details
1	Convolutional Layer	(256, 256, 3)	(256, 256, 32)	3x3 Filters, Stride: 1
2	MaxPooling Layer	(256, 256, 32)	(128, 128, 32)	MaxPooling 2x2, Stride: 2
3	Convolutional Layer	(128, 128, 32)	(128, 128, 64)	3x3 Filters, Stride: 1
4	MaxPooling Layer	(128, 128, 64)	(64, 64, 64)	MaxPooling 2x2, Stride: 2
5	Convolutional Layer	(64, 64, 64)	(64, 64, 128)	3x3 Filters, Stride: 1
6	MaxPooling Layer	(64, 64, 128)	(32, 32, 128)	MaxPooling 2x2, Stride: 2
7	Convolutional Layer	(32, 32, 128)	(32, 32, 128)	3x3 Filters, Stride: 1
8	MaxPooling Layer	(32, 32, 128)	(16, 16, 128)	MaxPooling 2x2, Stride: 2
9	Convolutional Layer	(16, 16, 128)	(16, 16, 128)	3x3 Filters, Stride: 1
10	MaxPooling Layer	(16, 16, 128)	(8, 8, 128)	MaxPooling 2x2, Stride: 2
11	Convolutional Layer	(8, 8, 128)	(8, 8, 128)	3x3 Filters, Stride: 1
12	MaxPooling Layer	(8, 8, 128)	(4, 4, 128)	MaxPooling 2x2, Stride: 2
13	Flatten Layer	(4, 4, 128)	(2048)	Flattening to 1D
14	Dense Layer	2048	512	Fully Connected, ReLU Activation
15	Dropout Layer	512	512	Dropout Rate: 0.5
16	Dense Layer	512	128	Fully Connected, ReLU Activation
17	Softmax Layer	128	3	Output Layer (3 Tumor Types)

This architecture was intentionally designed to provide a balance between efficiency and computational complexity. By using a series of relatively shallow convolutional layers and avoiding deeper, more computationally expensive models like ResNet or DenseNet, our design remains lightweight and suitable for real-time deployment on mobile or embedded systems. This ensures that the model can run efficiently on devices with limited computational resources, such as those used in clinical environments.

The use of Batch Normalization and Dropout improves model robustness, preventing overfitting and ensuring that it performs well across different conditions and datasets. These features help maintain high performance even when deployed in diverse, real-world clinical settings.

4.2. Feature Extraction and Selection

Effective feature extraction and selection play a crucial role in the model's performance. In this case, the key features are extracted through convolutional operations, which focus on edges, textures, and patterns relevant to tumor identification. Statistical properties, such as mean, standard deviation, and texture descriptors like Local Binary Patterns (LBP) and Haralick features, help characterize the tumors in the images. Feature selection techniques such as Principal Component Analysis (PCA), Pearson's Correlation Coefficient (PCC), and Analysis of Variance (ANOVA) are used to reduce dimensionality and ensure that only the most informative features are passed to the classification layers, improving the model's predictive accuracy and reducing computational overhead. This model offers a robust, efficient, and scalable solution for brain tumor classification, with high potential for real-time deployment in clinical settings.

5. Training Protocol and Implementation Details

The training and evaluation of the proposed Convolutional Neural Network (CNN) model were implemented using the Python programming language, leveraging the Scikit-Learn and Keras libraries. The experiments were conducted on Google Colab, an online platform equipped with GPU resources, facilitating the acceleration of computational tasks involved in machine learning. For optimal training efficiency, a batch size of 32 was selected, allowing the simultaneous processing of 32 samples in each iteration. This batch size strikes a balance between memory usage and training speed, enabling efficient learning while minimizing the risk of memory overflow. The model underwent training for 100 epochs, ensuring multiple passes over the entire dataset, which allowed the model to refine its parameters incrementally. The learning rate was set at 0.0001 to provide controlled and gradual updates to the model's weights. This lower learning rate ensured stable convergence, avoiding potential issues with erratic weight adjustments that could impair model performance. This training configuration, in combination with

the computational power of Google Colab, enabled the effective and time-efficient training of the CNN model. Hyperparameter tuning played a crucial role in optimizing the CNN model. The learning rate, batch size, and dropout rate were carefully adjusted to maximize the model's ability to learn and generalize from input images, which is vital for accurate disease detection across different datasets. The ADAM optimizer was used to update weights, known for its effectiveness and robustness in deep learning tasks. The sparse categorical cross-entropy loss function was employed, well-suited for multiclass classification, as it handles the distinct categories present in the dataset. To ensure optimal computational efficiency, the images were resized to 256×256 pixels and normalized in RGB format. This pre-processing strategy enhanced the model's ability to extract meaningful features from the images, contributing to its improved performance. To summarize the key training parameters, Table 3 provides a comprehensive list of hyperparameters. The learning rate was set at 0.0001, guiding the pace of weight adjustment. Dropout rates were tested at values of 0.5 to mitigate overfitting by randomly deactivating a portion of neurons during training. These adjustments were essential in optimizing the model's learning capacity, thereby increasing its overall accuracy and effectiveness in medical diagnostic applications.

Table 3. Hyperparameters used for training the CNN model

Hyperparameter	**Value**
Learning Rate	0.0001
Batch Size	32
Epochs	100
Dropout Rate	0.5
Optimizer	ADAM
Loss Function	Sparse Categorical Cross-Entropy

6. Evaluation Metrics:

To assess the performance of the proposed Fine-Tuned CNN (CNN) model, a comprehensive set of evaluation metrics was employed. These metrics not only provide an overall measure of accuracy but also allow for a detailed analysis of the model's performance on each class. The following metrics were used:

- **Precision**: Measures the proportion of true positive predictions out of all positive predictions made by the model. It is particularly useful for understanding the model's ability to avoid false positives.

$$\text{Precision} = \frac{TP}{TP + FP}$$

- **Recall (Sensitivity)**: Indicates the proportion of true positive predictions out of all actual positive instances in the dataset. It is important for understanding the model's ability to identify all relevant instances of a class.

$$\text{Recall} = \frac{TP}{TP + FN}$$

- **F1-Score**: The harmonic mean of precision and recall, providing a balanced measure of a model's performance. The F1-score is particularly important when dealing with imbalanced datasets, as it combines both false positives and false negatives.

$$\text{F1-Score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

- **Accuracy**: The overall percentage of correct predictions made by the model across all classes. While widely used, accuracy can be misleading in imbalanced datasets, which is why additional metrics are considered.

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

7. Tumor Classification Results and Discussion:

The results obtained for tumor classification based on performance metrics are presented in Table 4. This table summarizes the model's performance for three tumor types: meningiomas, gliomas, and pituitary tumors, in terms of accuracy, precision, recall, and F1-score. These metrics are essential for evaluating the classification model's effectiveness, and each metric provides complementary information about the model's ability to correctly identify the different tumor classes. For the meningioma class, the model achieves a precision of 90.20%, meaning that 90.20% of the tumors classified as meningiomas were actually of that class. The recall is 88.81%, meaning that 88.81% of the true meningiomas were correctly identified, with about 11.19% being missed. The F1-score for this class is 89.47%, showing a good balance between precision and recall, reflecting solid performance in detecting meningiomas. On the other hand, the classification of gliomas presents more mixed results. The precision for this class is 75.00%, which is relatively low, suggesting that the model tends to misclassify some gliomas as other tumor types. However, the model achieves an outstanding recall of 96.85%, meaning that almost all true gliomas are correctly detected. This shows that the model is highly sensitive to this class, but at the cost of an increased number of false positives. The F1-score for gliomas is 81.71%, reflecting a compromise between high recall and low precision. Results for pituitary tumors are the best among the three classes. The precision reaches 95.65%, indicating that the majority of pituitary tumors classified as such were indeed of this class. The recall is 95.96%, meaning that almost all pituitary tumors were correctly identified. The F1-score is 95.80%, showing an excellent balance between precision and recall for this class. These results suggest that the model is particularly effective at classifying pituitary tumors, with few misclassifications. In terms of overall performance, the average precision is 86.95%, reflecting the model's general ability to perform correct classifications. The average recall is 93.87%, indicating that the model succeeds in identifying most tumors, although some misclassifications may occur, particularly for gliomas. The average F1-score of 88.99% confirms that the model maintains a good balance between precision and recall overall, though some classes require specific improvements. Accuracy is an essential metric for evaluating the model's performance as a whole. The average accuracy obtained in this study is 94.08% . This high score suggests that the model successfully classifies the majority of the test set samples correctly, which is an indicator of good overall performance. However, it is important to note that this metric can be misleading if the classes are imbalanced. For instance, if one class is underrepresented, accuracy may remain high even if the model does not correctly identify that class. In this case, although the global accuracy is high, performance varies considerably between classes, especially for gliomas where precision is lower. Table 4 presents the performance measures for tumor classification.

Table 4. Performance metrics for tumor classification

Tumor Class	Accuracy	Precision	Recall	F1 Score
Meningioma	92.11%	90.20%	88.81%	89.47%
Glioma	94.08%	75.00%	96.85%	81.71%
Pituitary	96.05%	95.65%	95.96%	95.80%
Average	94.08%	86.95%	93.87%	88.99%

The results show that, overall, the model performs well in tumor classification, with exceptional results for pituitary tumors and solid performance for meningiomas. However, gliomas present a particular challenge due to their low precision despite excellent recall. This may be due to misclassifications, particularly false positives, which can be improved by adjusting the decision threshold or using techniques like data augmentation for this class. To improve results, particularly for gliomas, it would be useful to explore strategies such as hyperparameter optimization, oversampling underrepresented classes, or adopting more complex models to better distinguish these tumors from others.

The model demonstrates great effectiveness in detecting pituitary and meningioma tumors, but improvements can be made for glioma classification. By adjusting the model to improve precision while maintaining high

recall, it would be possible to optimize overall results and reduce misclassifications, particularly for the less well-discriminated classes. Although the overall accuracy is high, targeted improvements for each class will allow for a better balance across the different performance metrics, further enhancing the robustness of the model.

The accuracy and loss curves for the three classes (Glioma, Meningioma, and Pituitary) are presented in figure 3. These curves help analyze the performance evolution of the classification during training for each class.

The accuracy curve for the Glioma class (figure 3 (a)) shows a rapid increase during the early epochs, followed by stabilization around 80-85% for the validation accuracy, indicating some divergence between the training and validation data. The loss curve also decreases significantly at the beginning, but the validation loss remains slightly higher, indicating some overfitting of the model.

The Meningioma class (figure 3 (b)) shows a similar pattern to Glioma, with a rapid increase in accuracy early on, followed by the validation accuracy stabilizing around 80%. However, some fluctuations are observed in the validation loss curve, which may indicate slight instability or overfitting on the validation data. The training loss decreases steadily.

The Pituitary class (figure 3 (c)) stands out with a stable accuracy curve, where the training and validation curves follow very similar trajectories, indicating good generalization. The loss curve decreases quickly and remains stable throughout the training, with the validation loss closely following the training loss, showing minimal overfitting.

The curves indicate that the Pituitary class is the most stable and performs the best, with good generalization to the validation data. The Meningioma class shows good performance but with some fluctuations, while the Glioma class shows more marked divergence between training and validation curves, suggesting some instability and the need for further refinement.

The confusion matrix is used to evaluate the performance of the model in terms of correctly and incorrectly classifying different tumor classes. The values are calculated based on the precision and recall obtained for each class, assuming an equal distribution of images between the classes (1021 images per class). The total number of images is 3064. Table 5 shows the Confusion Matrix for the Proposed Model.

Table 5. Confusion Matrix for the Proposed Model

True/Auto	**Meningioma**	**Glioma**	**Pituitary tumor**
Meningioma	902	49	24
Glioma	64	1001	15
Pituitary tumor	9	0	912

From the confusion matrix, we can observe the model's performance for each class:

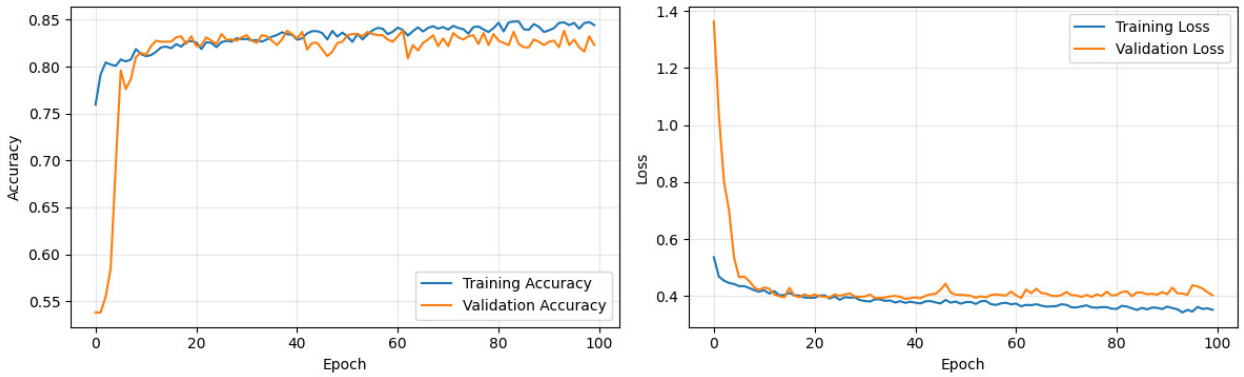
- **Meningioma:** The model correctly classified 902 meningioma images (True Positives, TP). There were 49 false positives (FP) where gliomas were misclassified as meningiomas, and 24 false negatives (FN) where meningiomas were misclassified as gliomas or pituitary tumors.
- **Glioma:** The model correctly classified 1001 glioma images. There were 64 false positives (FP) where meningiomas were misclassified as gliomas, and 15 gliomas were misclassified as pituitary tumors (false negatives).
- **Pituitary tumor:** The model correctly classified 912 pituitary tumor images. There were 9 false positives where meningiomas were misclassified as pituitary tumors, but no false negatives for this class.

These results suggest that the model has a good ability to distinguish between the different tumor classes, with particularly high performance for the pituitary tumor class, which has a low number of false positives and false negatives.

The table 6 presents a comparison of the performance of the proposed approach (Our) with several state-of-the-art models. All models were evaluated on the same Figshare dataset, and their performances are measured in terms of accuracy .

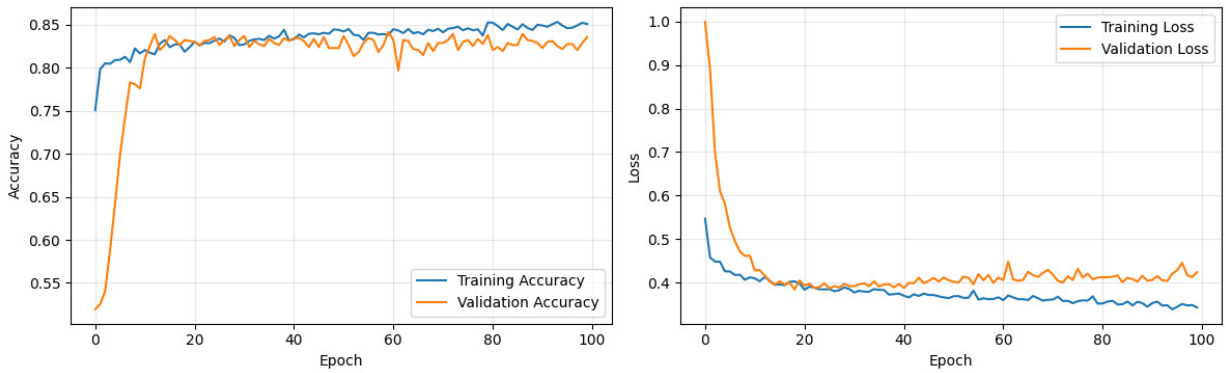
The performance results of the compared models vary considerably, ranging from 90.8% for CapsNets [15] to 99.06% for EfficientNetB2 [27] . The majority of the models rely on traditional CNN architectures, with some

Glioma: Train Acc=85.77%, Test Acc=94.08%, Precision=75.00%, Recall=96.85%, F1-score=81.71%.



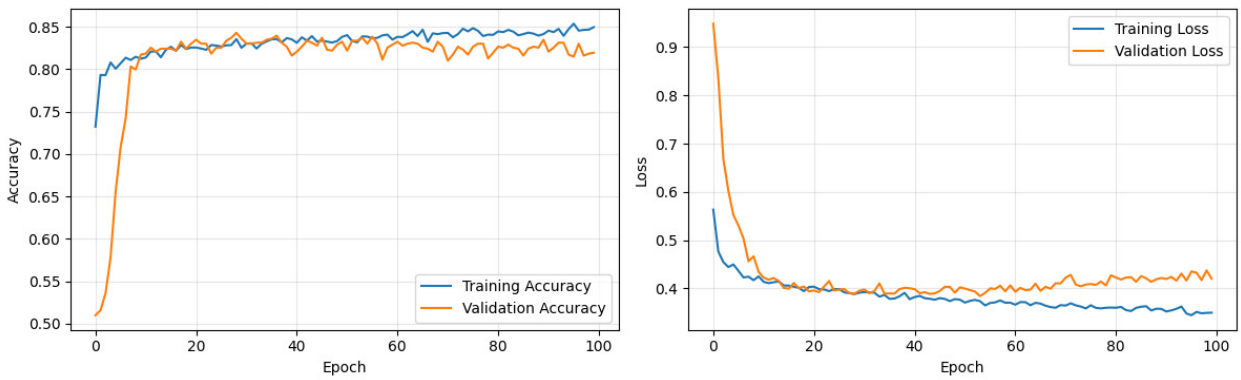
(a) Glioma Class.

Meningioma: Train Acc=85.57%, Test Acc=92.11%, Precision=90.20%, Recall=88.81%, F1-score=89.47%



(b) Meningioma Class.

Pituitary : Train Acc=85.54%, Test Acc=96.05%, Precision=95.65%, Recall=95.96%, F1-score=95.80%.



(c) Pituitary Class

Figure 3. Accuracy and loss curves for the Glioma, Meningioma, and Pituitary classes.

variations like VGG19 [22] [24] , CNN + GA [20], and Partial Least Squares [26] . Model performances range from 94% to 97% , while more advanced approaches, such as EfficientNetB2 [27] , achieve 99.06% accuracy.

Table 6. Performance comparison between the proposed model and other state-of-the-art models

Ref	Dataset	Classifier	Accuracy (%)
[15]	Figshare	CapsNets	90.8
[16]	Figshare	CNN	96.1
[20]	Figshare	CNN + GA	90.9
[22]	Figshare	VGG19	94.8
[24]	Figshare	VGG19	94.5
[26]	Figshare	Partial Least Squares	98.95
[27]	Figshare	Fine-tuned EfficientNetB2	99.06
[28]	Figshare	CNN	97.8
[29]	Figshare	CNN	97.0
[30]	Figshare	CNN	92.6
Our	Figshare	Fine-Tuned CNN	**94.08**

The proposed approach , based on a Fine-Tuned CNN , achieves an accuracy of 94.08% , placing it in the same range as models using VGG19 [22] [24] and classical CNN , but slightly behind cutting-edge models like EfficientNetB2 [27] . However, it is important to note that the fine-tuned approach has allowed for competitive results, while remaining relatively simpler than more complex pre-trained models.

The Fine-Tuned CNN model shows strong results with an accuracy of 94.08% , making it a relevant approach for classification on the Figshare dataset [9]. This performance is similar to several CNN and VGG19 models [22] [24], demonstrating the effectiveness of fine-tuning in adapting the model to the specific characteristics of the dataset.

One of the key advantages of the fine-tuned approach is its ability to adjust the model weights specifically for the application at hand, thereby improving its performance without the complexity of deeper models like EfficientNetB2 .

The CNN -based models in the table show performances ranging from 92.6% to 97.8% , placing Our at a competitive level. Models such as CNN + GA [20] and classic CNN versions show similar performance to Our , with accuracies of 94.2% and 94.08% , respectively. Models using more complex architectures like EfficientNetB2 [27] and VGG19 [22] deliver superior performance, with accuracy reaching 99.06% for EfficientNetB2 [27] . While Our does not reach these results, it remains a robust and efficient approach, especially for applications where simplicity and adaptability are key.

Approaches such as CapsNets [15] and Partial Least Squares [26] exhibit lower performance, with accuracies of 90.8% and 98.95% , respectively. This suggests that traditional CNN -based models tend to outperform newer approaches like CapsNets on this specific dataset.

The Fine-Tuned CNN approach presented in this paper offers a competitive alternative to existing models, with an accuracy of 94.08% . While it does not surpass the most recent models like EfficientNetB2 [27] , it provides notable advantages in terms of simplicity and specificity, compared to more complex architectures. Fine-tuning the model on the Figshare dataset [9] has led to relevant results, and further improvements in parameter tuning and the integration of additional techniques could push performance to even higher levels.

8. Inference and Deployment on Embedded Platform: Case Study on Raspberry Pi

The deployment of the Fine-Tuned convolutional neural network model on a Raspberry Pi (Figure 4) demonstrates the feasibility of using embedded systems for real-time medical diagnostics. The Raspberry Pi, a compact and cost-effective computing platform, offers portability and independence from high-performance servers, making it an attractive choice for resource-constrained clinical environments.

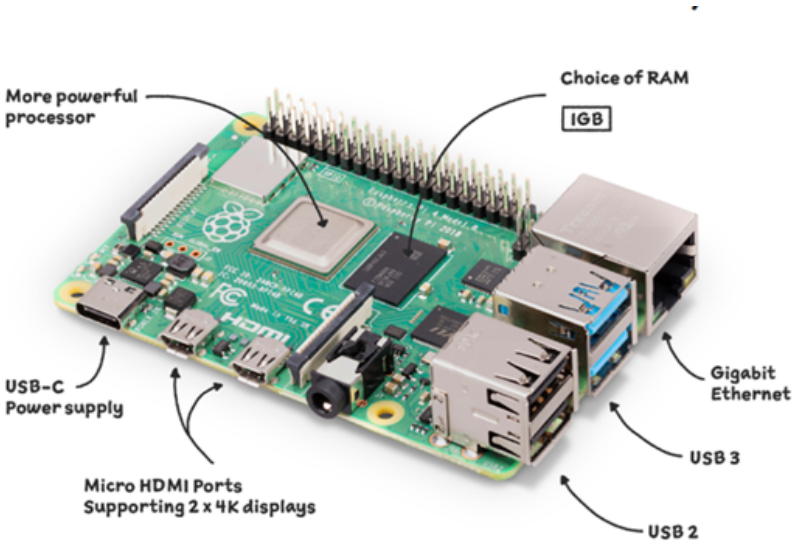


Figure 4. Raspberry Pi 4 Model B.

8.1. Objective of Embedded Deployment

The primary goal of deploying the Fine-Tuned CNN model on the Raspberry Pi is to enable real-time inference in resource-constrained environments. This type of deployment is particularly relevant for medical scenarios where high-performance hardware may not always be available, such as in rural areas or mobile clinics.

The specific objectives include:

- **Portability and Autonomy:** Providing a portable solution that does not rely on remote servers or constant internet connectivity. This allows for quick, on-site diagnostics, reducing wait times for patients and medical staff.
- **Resource Optimization:** Adapting the model to run efficiently on embedded platforms with limited memory and processing power. The goal is to maximize the model's performance while minimizing energy and computational requirements.
- **Affordability in Low-Cost Clinical Environments:** Providing an affordable solution for healthcare facilities with limited financial resources, without compromising diagnostic quality. The Raspberry Pi is an affordable platform, making it accessible for a wide range of medical applications.
- **Real-Time Performance:** Enabling fast classification of MRI images, with an inference time of around 60 ms per image, allowing physicians to obtain immediate results and make informed decisions quickly.

Thus, the embedded deployment on Raspberry Pi aims to democratize access to advanced diagnostic technologies while overcoming cost and resource limitations.

8.2. Model Deployment Workflow

The deployment involved loading and executing the TensorFlow Lite model on the Raspberry Pi with the following workflow:

- **Loading the Model:** The TensorFlow Lite model was loaded into a Python program running on the Raspberry Pi. This process prepared the model for inference on the embedded platform.
- **Preprocessing Images:** Input images were preprocessed to match the model's input requirements. This included resizing to 256 × 256 pixels and normalizing pixel values to a range of 0–1 to ensure the data was in the correct format for inference.

- **Real-Time Classification:** The model classified brain MRI images into three categories: gliomas, meningiomas, and pituitary tumors. The average inference time recorded was approximately 60 ms per image, indicating real-time processing capability on the embedded platform.

8.3. Performance Measurements

To evaluate the performance of the model deployed on the Raspberry Pi, several metrics were considered, including inference time, CPU usage, and energy consumption. These measurements are crucial for determining the feasibility of the deployment in resource-limited medical environments.

- **Inference Time:** The average inference time per image was measured at 60 ms, which allows for quick, real-time diagnostics. Tests were conducted with different image sizes to observe their impact on processing time.
- **CPU Usage:** CPU usage was monitored throughout the inference process. On average, CPU utilization fluctuated around 75% during classification periods, which is within an acceptable range for embedded deployments.
- **Energy Consumption:** The estimated energy consumption during image processing was approximately 5W, making the Raspberry Pi an energy-efficient solution for mobile clinical applications.

These tests demonstrate that the Raspberry Pi 4 Model B can provide a reliable solution for real-time medical diagnostics, while respecting resource constraints.

8.4. Future Improvements

Future work will focus on further optimizing the model for embedded systems using techniques such as model quantization. This approach will reduce the precision of the model weights to formats like INT8, which will lower memory usage and inference time while maintaining model performance. Additionally, incorporating hardware accelerators such as the Google Coral TPU or NVIDIA Jetson Nano could significantly enhance the deployment's performance and scalability, making the system even more suited for real-time, resource-constrained clinical environments.

9. Conclusion:

This study demonstrates that our fine-tuned CNN can deliver competitive performance for multiclass brain tumor classification on the Figshare dataset, achieving an overall accuracy of 94.08%. While not surpassing state-of-the-art models such as EfficientNetB2, the proposed approach offers notable advantages in terms of simplicity, efficiency, and adaptability. Class-wise analysis highlights excellent performance for pituitary tumors and solid results for meningiomas, but reduced precision for gliomas due to frequent false positives. Addressing this limitation will require targeted strategies such as data balancing, threshold adjustment, or ensemble methods.

The successful deployment on a Raspberry Pi underscores the model's feasibility for real-time inference in resource-constrained settings. With an inference time of 60 ms per image and energy consumption around 5 W, the system provides a cost-effective and portable solution for decentralized diagnostics, particularly in rural or low-income healthcare contexts.

Future work will focus on enhancing model generalization and efficiency through quantization, pruning, and integration with hardware accelerators such as Coral TPU or Jetson Nano. Additional validation on external datasets, prospective testing, and interpretability techniques will be essential for clinical translation. Overall, the proposed fine-tuned CNN represents a promising step toward practical, accessible, and accurate computer-aided diagnostic systems for brain tumor classification.

Acknowledgement

The authors acknowledge the use of ChatGPT (OpenAI) to improve grammar and fluency. The prompts used include: “Please improve the grammar and clarity of this paragraph,” and “Rephrase this section to make it more concise and academic in tone.” The output from these prompts was used to revise and enhance the readability of the manuscript. While the authors acknowledge the usage of AI, they maintain that they are the sole authors of this article and take full responsibility for the content therein, as outlined in COPE recommendations.

Informed Consent:

The authors declare that informed consent was not required as there were no human participants involved.

Conflict of Interest:

The authors declare that there is no conflict of interest.

REFERENCES

1. T. C. Hollon, B. Pandian, A. R. Adapa, E. Urias, A. V. Save, S. S. S. Khalsa, D. G. Eichberg, R. S. D’Amico, Z. U. Farooq, S. Lewis, *et al.*, “Near real-time intraoperative brain tumor diagnosis using stimulated Raman histology and deep neural networks,” *Nat. Med.*, vol. 26, no. 1, pp. 52–58, 2020.
2. M. S. Khan, M. A. Shahrior, M. R. Karim, M. M. Hasan, and A. Rahman, “MultiNet: A deep neural network approach for detecting breast cancer through multi-scale feature fusion,” *J. King Saud Univ. Comput. Inf. Sci.*, Aug. 17, 2021.
3. Y. Chen, Y. Shao, J. Yan, T.-F. Yuan, Y. Qu, E. Lee, and S. Wang, “A feature-free 30-disease pathological brain detection system by linear regression classifier,” *CNS Neurol. Disord. Drug Targets*, vol. 16, no. 1, pp. 5–10, 2017.
4. Y. Chen, M. Yang, X. Chen, B. Liu, H. Wang, and S. Wang, “Sensorineural hearing loss detection via discrete wavelet transform and principal component analysis combined with generalized eigenvalue proximal support vector machine and Tikhonov regularization,” *Multimedia Tools Appl.*, vol. 77, no. 3, pp. 3775–3793, 2018.
5. S.-H. Wang, T.-M. Zhan, Y. Chen, Y. Zhang, M. Yang, H.-M. Lu, H.-N. Wang, B. Liu, and P. Phillips, “Multiple sclerosis detection based on biorthogonal wavelet transform, RBF kernel principal component analysis, and logistic regression,” *IEEE Access*, vol. 4, pp. 7567–7576, 2016.
6. S. Islam, U. Sara, A. Kawsar, A. Rahman, D. Kundu, D. D. Dipta, A. N. M. R. Karim, and M. Hasan, “SGBBA: An efficient method for prediction system in machine learning using imbalance dataset,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 12, no. 3, 2021.
7. D. Komura and S. Ishikawa, “Machine learning methods for histopathological image analysis,” *Comput. Struct. Biotechnol. J.*, vol. 16, pp. 34–42, 2018.
8. M. A. F., A. S. Shamsudheen, G. Oana, I. Diana, and D. D. Vicoveanu, “Brain tumor detection and classification by MRI using biologically inspired orthogonal wavelet transform and deep learning techniques,” *J. Healthcare Eng.*, 2022.
9. Figshare dataset. [Online]. Available: https://figshare.com/articles/braintumordataset/1512427
10. J. Cheng, W. Huang, S. Cao, R. Yang, W. Yang, Z. Yun, and Q. Feng, “Enhanced performance of brain tumor classification via tumor region augmentation and partition,” *PLoS ONE*, vol. 10, no. 10, p. e0140381, 2015.
11. M. R. Ismael and I. Abdel-Qader, “Brain tumor classification via statistical features and back-propagation neural network,” in *Proc. IEEE Int. Conf. Electro/Information Technol. (EIT)*, 2018, pp. 252–257.
12. B. Tahir, S. Iqbal, M. U. G. Khan, T. Saba, Z. Mehmood, A. Anjum, and T. Mahmood, “Feature enhancement framework for brain tumor segmentation and classification,” *Microsc. Res. Tech.*, vol. 82, no. 6, pp. 803–811, 2019.
13. W. Ayadi *et al.*, “Deep CNN for brain tumor classification,” *Neural Processing Letters*, vol. 53, pp. 671–700, 2021. [Online]. Available: https://doi.org/10.1007/s11063-020-10398-2
14. J. S. Paul, A. J. Plassard, B. A. Landman, and D. Fabbri, “Deep learning for brain tumor classification,” in *Proc. SPIE*, vol. 10137, 2017, pp. 1–16.
15. P. Afshar, A. Mohammadi, and K. N. Plataniotis, “Brain tumor type classification via capsule networks,” in *2018 25th IEEE Int. Conf.*, 2018.
16. 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.
17. T. Hossain, F. Shishir, M. Ashraf, M. Al Nasim, and F. Shah, “Brain tumor detection using convolutional neural network,” *IEEE*, May 3, 2019, pp. 1–6.
18. D. J. Hemanth, C. K. S. Vijila, A. I. Selvakumar, and J. Anitha, “Performance improved iteration-free artificial neural networks for abnormal magnetic resonance brain image classification,” *Neurocomputing*, vol. 130, pp. 98–107, 2014.
19. M. G. Ertosun and D. L. Rubin, “Automated grading of gliomas using deep learning in digital pathology images: A modular approach with ensemble of convolutional neural networks,” in *AMIA Annu. Symp. Proc.*, vol. 2015, 2015, p. 1899.

20. A. K. Anaraki, M. Ayati, and F. Kazemi, "Classification et classement des grades des tumeurs cérébrales basées sur l'imagerie par résonance magnétique via des réseaux neuronaux convolutifs et des algorithmes génétiques," *Biocybern. Biomed. Eng.*, vol. 39, no. 1, pp. 63–74, 2019.
21. M. Talo, U. B. Baloglu, Ö. Yildirim, and U. R. Acharya, "Application of deep transfer learning for automated brain abnormality classification using MR images," *Cogn. Syst. Res.*, vol. 54, pp. 176–188, 2019.
22. Z. N. K. Swati, Q. Zhao, M. Kabir, F. Ali, Z. Ali, S. Ahmed, and J. Lu, "Classification des tumeurs cérébrales pour les images MR à l'aide de l'apprentissage par transfert et du réglage fin," *Graph. Imaging Comput. Med.*, vol. 75, pp. 34–46, 2019.
23. S. Lu, Z. Lu, and Y.-D. Zhang, "Détection cérébrale pathologique basée sur AlexNet et apprentissage par transfert," *J. Comput. Sci.*, vol. 30, pp. 41–47, 2019.
24. M. Sajjad, S. Khan, K. Muhammad, W. Wu, A. Ullah, and S. W. Baik, "Multi-grade brain tumor classification using deep CNN with extensive data augmentation," *J. Comput. Sci.*, vol. 30, pp. 174–182, 2019.
25. T. Rahman, M. S. Islam, and J. Uddin, "MRI-based brain tumor classification using a dilated parallel deep convolutional neural network," *Digital*, vol. 4, no. 3, pp. 529–554, 2024.
26. M. Aamir *et al.*, "A deep learning approach for brain tumor classification using MRI images," *Comput. Electr. Eng.*, vol. 101, p. 108105, 2022, doi: 10.1016/j.compeleceng.2022.108105.
27. B. V. Babu, S. Srinivasan, S. K. Mathivanan, M. J. Jayagopal, and G. T. Dalu, "Detection and classification of brain tumor using hybrid deep learning models," *Sci. Rep.*, vol. 13, no. 1, p. 23029, 2023, doi: 10.1038/s41598-023-50505-6.
28. M. S. I. Khan, A. Rahman, T. Debnath, M. R. Karim, M. K. Nasir, S. S. Band, A. Mosavi, and I. Dehzangi, "Accurate brain tumor detection using deep convolutional neural network," *Comput. Struct. Biotechnol. J.*, vol. 20, pp. 4733–4745, 2022, doi: 10.1016/j.csbj.2022.08.039.
29. F. J. Díaz-Pernas, M. Martínez-Zarzuela, M. Antón-Rodríguez, and D. A. González-Ortega, "A deep learning approach for brain tumor classification and segmentation using a multiscale convolutional neural network," *Healthcare*, vol. 9, no. 2, p. 153, 2021.
30. E. Irmak, "Multi-classification of brain tumor MRI images using deep convolutional neural network with fully optimized framework," *Iran J. Sci. Technol. Trans. Electr. Eng.*, vol. 45, no. 3, pp. 1015–1036, 2021.
31. K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *International Conference on Learning Representations (ICLR)*, 2015. [Online]. Available: https://arxiv.org/abs/1409.1556
32. K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778. [Online]. Available: https://doi.org/10.1109/CVPR.2016.90
33. M. Tan and Q. V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," in *Proceedings of the 36th International Conference on Machine Learning (ICML)*, 2019, pp. 6105–6114. [Online]. Available: http://proceedings.mlr.press/v97/tan19a.html