TANGENCE
Number 137, 2025

ISSN: 1710-0305
PISSN: 1189-4563

Transfer Learning-Driven MRI-Based Classification Pipelines for Brain Tumor Diagnosis: Glioma, Meningioma, and Pituitary Tumor Discrimination

Vivek Verma[1] ,Vaibhav Singh[1] ,Diksha Sharma[2], Varsha Sharma[2] ,Aditi Sharma[2]

[1]Independent Researcher

Computer Science Engineering

[1]Centre for Advanced Studies, Dr. A.P.J. Abdul Kalam Technical University, Lucknow, India

Email: vivekverma9920@gmail.com

[1]Independent Researcher

Computer Science Engineering

[1]Centre for Advanced Studies, Dr. A.P.J. Abdul Kalam Technical University, Lucknow, India

Email: singhvbhv0302@gmail.com

[2]Assistant Professor

Computer Science Engineering Assistant Pr

[2]Institute of Engineering and Technology, Lucknow, India

Email: dsharma.csed.cf@ietlucknow.ac.in

[2]Assistant Professor

Computer Science Engineering
Assistant Profes

[2]Institute of Engineering and Technology, Lucknow, India

Email:var6197@gmail.com

[2]Assistant Professor

Computer Science Engineering
Assistant Profes

[2]Institute of Engineering and Technology, Lucknow, India

Email: asharma.csed.cf@ietlucknow.ac.in

*Corresponding author(s). E-mail(s): vivekverma9920@gmail.com

Abstract: MRI-based classification of brain tumors is an important step in delivering timely and effective treatment. In this study, we tested a method that uses CNNs trained via transfer learning to classify the three main types of brain tumors (gliomas, meningiomas, and pituitary tumors). Mendeley dataset is taken into consideration, containing 6,056 MRI images (2004 brain glioma, 2004 brain meningioma, 2048 brain pituitary). Two types of CNN architectures were tested, AlexNet (trained from scratch) and InceptionV3 (using the weights from ImageNet). All of the images were preprocessed before being fed to the models

ISSN: 1710-0305
PISSN: 1189-4563

using image resizing, normalization, and extensive augmentation to ensure accuracy and minimize class imbalance between the tumor categories. Effectively stratified train-test splits of the data allowed for fair performance evaluation of both models. The AlexNet model consistently achieved 94% accuracy, with a precision, recall, and F1-score of 94%, indicating that it could provide reliable performance when classifying brain tumors based on MRI. In contrast, the InceptionV3 model using transfer learning and fine-tuning performed even better than AlexNet, achieving 98% accuracy with a precision, recall, and F1- score of 98%. These results indicate that pre-trained convolutional neural network architectures provide increased classification reliability, significantly reduce training time, and are applicable to medical datasets that contain limited numbers of instances. The findings of this research study illustrate the potential for developing highly accurate and efficient automated deep learning technology to accurately diagnose neuro-oncology diseases using transfer learning. This type of technology will provide a strong basis for Clinical Decision Support Systems (CDSS) that aid radiologists with the interpretation of diagnostic medical images**.**

Keywords: Brain Tumor, Glioma, Meningioma, Pituitary Tumor, Transfer Learning, Image Classification, Brain Tumor Classification, Convolutional Neural Netrwok, Magnetic Resonance Imaging

1. Introduction

1.1 Background on Brain Tumors and MRI Diagnosis

The abnormal intracranial growths of the cells of brain or nearby areas are the brain tumors whose localization, size and heterogeneity make diagnosis challenging. Standard MRI with T1, T2 and FLAIR sequences provides high contrast, non-invasive visualization of such lesions and is the imaging modality of choice for initial detection and characterization. [1][2]

1.2 Challenges in Manual MRI Interpretation

Manually interpreting MRI scans for brain tumor diagnosis takes a lot of time and it also heavily relies on the radiologist's expertise. It experiences significant inconsistencies both between different readers and within the same reader, for example, segmentation variability for brain tumors showed average Dice scores approximately 0.75 (95% CI 0.701-0.808) across various methods.[3][4] The variation in tumor characteristics (such as size, shape, location, and contrast) along with the imaging artifacts and differing acquisition protocols, makes it more difficult to accurately delineate and classify lesions, increasing the chance of misdiagnosis or postponement of treatment. [5][6]

1.3 Research Gap and Motivation

Although there has been significant advancement in the automation of MRI based brain tumor classification machine learning and deep learning methods, there are still several gaps exist in the literature. Many studies are based on single deep learning architectures and fail to compare multiple pretrained or transfer learning models to determine which would provide the greatest level of multiclass classification (for example, gliomas, meningiomas, pituitary tumors). [7][8] Class imbalances, heterogeneous datasets (including different protocols and MRI machines), and inefficient computational processes are rarely accounted for in the development of machine learning models. [9][10] Hence, there is a need to create an efficient, accurate, and robust classification system that uses transfer learning to diagnose multi-class brain tumors across multiple types of imaging equipment.

1.4 Contributions of This Work

In this study we proposed a classification pipeline for MRI using transfer learning to enable accurate

identification of glioma, meningioma, and pituitary tumors. Key contributions of this study include comparative analysis of pretrained CNN architectures, the use of effective preprocessing techniques and augmentation techniques and the enhanced accuracy in diagnosing brain tumors, thus providing a practical approach for automation of the diagnostic process for brain tumors.

1.5 Paper Organization

This paper is structured as follows: Section 2 of this paper contains related work on MRI-based brain tumor classification and transfer learning. Section 3 of this paper contains details about the dataset, preprocessing, and proposed transfer learning pipeline. Section 4 of this paper describes experiments and evaluation metrics. Sections 5 and 6 present results, discussion, and clinical implications, followed by conclusions in Section 7.

2. Related Work

2.1 Related Research works and Studies for Brain Tumor Classification

Table 1. Comparative analysis of existing techniques and mythologies for Brain Tumor Classification

S.No.	Author	Year	Technology Used	Performance	Key Insight
1	Zhang [11]	2011	Wavelet transform, principle component analysis and back propagation (BP) NN	ccuracies on both training and test images are 100%	Applied this method on 66 images (18 normal, 48 abnormal) and the computation time per image is only 0.0451 s.
2	W. H. rahim [12]	2013	Principle Component Analysis (PCA), and Back-Propagation Neural Network	Classification accuracy of 96.33%	3×58 datasets of MRI Brain segital images have been used for tainting and testing
3	N. Abdullah [13]	2011	Support vector machine (SVM)	Accuracy of 65%	Determination of normal and abnormal brain image is based on symmetry which is exhibited in the axial and coronal images
4	Kumar [14]	2017	Genetic algorithm and support vector machine (SVM)	curacy between 80% and 90%	Parameters used for analyzing the images are given as: entropy, smoothness, root mean square error (RMS), kurtosis and correlation
5	Z. Jia [15]	2025	Fully Automatic Heterogeneous Segmentation using Support Vector Machine (FAHS- SVM)	98.51% accuracy in detecting abnormal and normal tissue	Proposes the separation of the whole cerebral venous system into MRI imaging with the addition of a new, fully automatic algorithm based on structural, morphological, and relaxometry details

ISSN: 1710-0305
PISSN: 1189-4563

6	Badža, M. M. [16]	2020	nvolutional Neural Network (CNN)	curacy of 96.56%	Tested on T1-weighted contrast-enhanced magnetic resonance images
7	an HA [17]	2020	VGG-16, ResNet-50, and Inception-v3 models	VGG-16 achieved 96%, ResNet-50 achieved 89% and Inception-V3 achieved 75% accuracy	Experiment is tested on a very small dataset but the experimental result shows that our model accuracy result is very effective and have very low complexity rate
8	Ullah [18]	2020	anced Deep Neural Network (DNN)	ained 95.8% accuracy	Extracted features from an enhanced MR brain image using a discrete wavelet transform and these feature are further reduced by color moments i.e. mean, standard deviation, and skewness
9	Assam [19]	2021	Feed Forward - ANN (FF-ANN), a hybrid classifiers called: Random Subspace with Random Forest (RSwithRF) and Random Subspace with Bayesian Network (RSwithBN)	FF-ANN gives 95.83% accuracy, RSwithRF gives 97.14% and RSwithBN gives 95.71% accuracy	Used images with brain tumor, acute stroke and alzheimer, besides normal images, from the public dataset developed by harvard medical school, for evaluation purposes
10	S. Chetana [20]	2022	Transfer learning-based CNN-pretrained VGG-16, ResNet-50, and Inception-v3 models	VGG-16 gives accuracy of 96.0%, Inception-v3 gives 78% and ResNet50 gives 95.0%	I brain tumor images dataset consisting of 233 images

2.2 Identified Research Gaps

The results of investigation into the current MRI-based detection of brain abnormalities indicates the need for further investigation into a significant number of important limitations. A major limitation of early studies using wavelet transformations, PCA's, and BP-NNs is that although they were able to achieve significant levels of accuracy, their data sets typically consisted of very small numbers of images, e.g. 66-image dataset used in [11] by Y. Zhang and the 3×58-image dataset evaluated in [12] by W. H. Ibrahim, which severely limited the generalizability of the model. Analogously, models that rely on symmetry analysis and handcrafted feature extraction techniques, such as those described in [13] by N. Abdullah and [14] by S. Kumar, possess little to no capacity to adequately represent the full spectrum of tumor features due to a lack of representation of features from multiple tumor types, leading to decreased performance across many tumor types and MRI modalities.

ISSN: 1710-0305
PISSN: 1189-4563

Second, while recent methods utilizing deep learning (i.e., CNN and transfer learning based model) have demonstrated enhanced performance [16-18, 20], there remain dataset limitations that were recognized in many studies regarding limited sample sizes as well as missing many of the multimodal MRIs which precluded these networks from being able to identify heterogeneous tumor features. Additionally, most of the available literature centres predominantly on classification rather than development and evaluation of complete or end-to-end solutions for segmentation and diagnosis although there has been development of the FAHS-SVM segmentation algorithm for venous anatomical structures [15], highlighting an important gap regarding how to adequately create joint segmentation or classification models. Furthermore, as noted in numerous papers utilizing public datasets [19], studies vary in how they assess a dataset. The protocols utilized in assessing the datasets used vary significantly due to the use of different MRIs, imbalanced classes and a lack of cross validation between datasets, therefore making it virtually impossible to establish a standardized benchmark for reliably comparing models. Additionally, although the VGG-16 and RSwithRF methods are very high accuracy models [17-20], they do not appear to have been studied for their computational efficiency and scalability or whether they are applicable for clinical use in a real time setting, thus creating an opportunity for developing computationally optimized and clinically applicable solutions.

3. Materials and Methods

3.1 Dataset Description

3.1.1 Source and Characteristics of MRI Dataset

The Bangladesh Brain Cancer MRI Dataset [21] contains 6056 images that were collected in multiple hospitals across Bangladesh under the guidance of trained professionals. The MRI images were resized to 512 by 512 pixels and represent three different types of tumors: 2004 Brain Glioma, 2004 Brain Menin, and 2048 Brain Tumor. It provides machine learning and deep-learning experts with a valuable, comprehensive resource for developing and assessing algorithms for the automatic diagnosis of brain tumors.

3.1.2 Tumor Classes: Glioma, Meningioma, Pituitary

The dataset separates the different brain tumor types into 3 classifications: glioma, a malignant tumor that arises from glial cells, meningioma, which is generally a benign tumor that is located in the meninges, and pituitary tumor, which is found within the pituitary gland. This allows MRI multi-class classification for diagnosing & treating the brain tumor types accurately. [12]

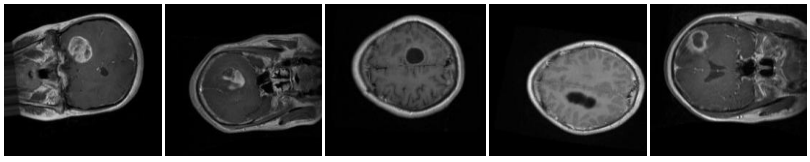

Fig.1. Representative MRI slices illustrating glioma-affected brain regions from the dataset

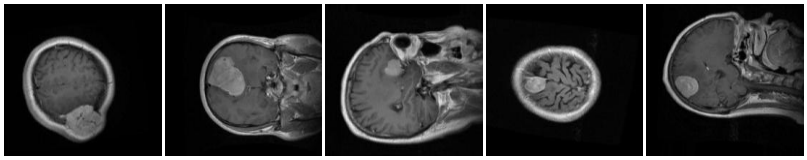

Fig.2. Representative MRI slices illustrating meningioma-affected brain regions from the dataset

ISSN: 1710-0305
PISSN: 1189-4563

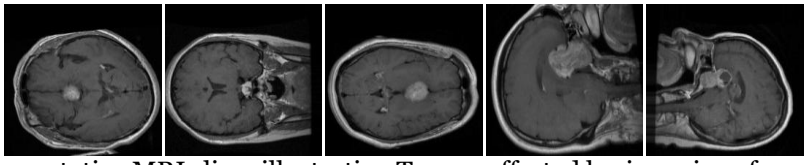

Fig.3. Representative MRI slices illustrating Tumor-affected brain regions from the dataset

3.2 Preprocessing Pipeline

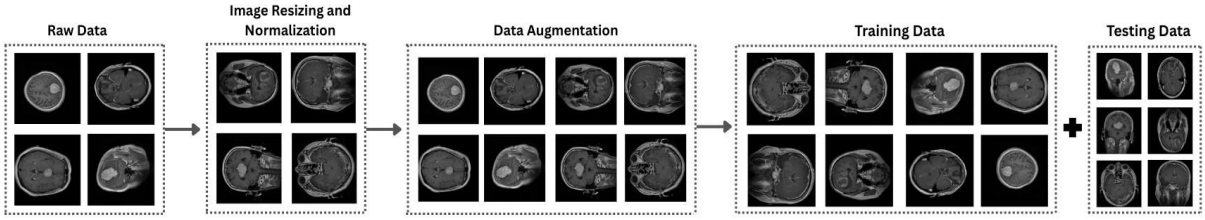

Fig.4. Preprocessing pipeline illustrating the sequential operations applied to the brain tumor dataset, including image resizing and intensity normalization, data augmentation techniques, and the train–validation split strategy

3.2.1 Image Resizing and Normalization

The deep learning architecture AlexNet along with InceptionV3 requires input images to be in standard format of 224 x 224 pixels, so we resized the input MRI images into the standard format to match the input requirements of the networks. Standardizing all images to the same size increases efficiency because it permits one GPU processor to work through a complete set of images. By standardizing the size of every image, the order of all pixel array positions remains intact when using them for convolution. InceptionV3 and AlexNet models also require that all pixel array values within an original input image be scaled back to a numeric range of [0, 1]. The normalization of pixel values increases the stability of the training process through gradient updates, thereby enhancing the learning rate. In addition to the standard image resizing and normalization processes, InceptionV3 requires additional specific image preprocessing when using the pretrained weights on the dataset so that the correct weights can be utilized when performing image classification. All of these preprocessing techniques assist in providing a uniform dataset input to the model, reduce computation resources needed to train the model, and enhance the ability of the model to rapidly learn and extract critical characteristics of images from MRI scans for the purpose of predicting their classification.

3.2.2 Data Augmentation Techniques

In order to mitigate overfitting and promote generalization of the models, we used augmented data for both AlexNet and InceptionV3. We used different augmentation techniques such as random rotations, zooming in, shifting the image both vertically and horizontally, flipping horizontally, changing brightness levels, and performing shear transformations; these techniques were intended to simulate realistic variations in MRI imaging. The increased effective number of images in the training dataset due to augmentations exposes the model to a greater variety of imaging environments which improves its ability to extract features effectively. While augmentations of AlexNet trained from scratch provided a tremendous benefit by way of additional examples due to limited amount of training data, augmentations for InceptionV3 provided a benefit as well from pretraining on trained imaging features and fine-tuning using high-level imaging features. These three approaches combined allowed both networks greater ability to generalize to

ISSN: 1710-0305
PISSN: 1189-4563

previously unseen images, and reduced the likelihood of biasing towards specific images due to orientation or brightness variation.

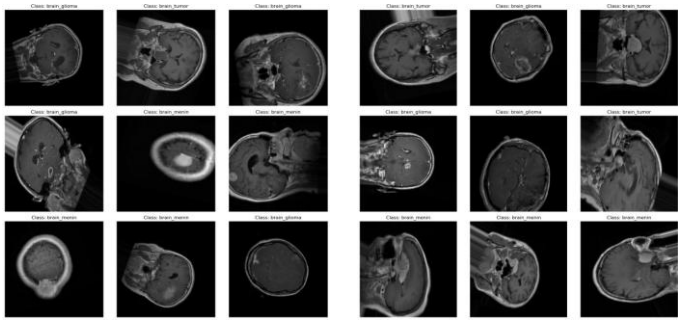

Fig.5. Augmented MRI images generated from the brain tumor dataset. The figure illustrates a variety of applied augmentation transformations

3.2.3 Train–Test Split Strategy

To maintain an even representation of all classes within the training and testing datasets for both AlexNet and InceptionV3, we used stratified splitting. We have taken five different train-to-test ratios that are 10:90, 15:85, 20:80, 25:75, and 30:70, to achieve the best balance between training and testing. The larger amounts of training data provided enough information for the models to learn the necessary features from the data, while the corresponding testing datasets provided a way of monitoring possible overfitting and also allowing for tuning of hyperparameters. The fine-tuning of InceptionV3 was enhanced by this method, by preserving previously learnt features from the pre-training stage. By using this method, both models were trained on diverse samples and thus reduced the chances of overfitting, as well as providing accurate metrics for comparative purposes for classification of brain MRIs.

3.3 Transfer Learning Framework

3.3.1 Overview of Transfer Learning Paradigm

Transfer learning is a machine learning approach that takes advantage of what has already been learned by using large data sets to help with the performance of a machine learning task being trained with a limited data set (target task). In this research project, AlexNet was developed from scratch due to its simpler architecture than InceptionV3 and InceptionV3, on the other hand, utilized the pre-trained weights from ImageNet to jumpstart its transfer learning capabilities. By retaining the learned features from previous data sets within each model (and using them as inputs), it allows these models to be trained to identify relevant features in brain MRI scans much more quickly than if they were trained on the MRI images alone. Thus, using transfer learning allows both types of models to perform better when classifying the different types of brain tumors (increased accuracy) and to not be overfit (avoid classification inaccuracies due to overfitting).

3.3.2 Selection of Pre-Trained CNN Architectures

In this study we selected AlexNet and InceptionV3 models because they possess features that complement each other. AlexNet has a shallower architecture and fewer layers so that we can easily demonstrate how well it performs after being trained from scratch with the MRI images. On the other hand, the more advanced and deeper architecture of InceptionV3 has Inception modules that allow effective capturing of features at multiple scales with the use of pretrained ImageNet weights to develop superior classifiers from

ISSN: 1710-0305
PISSN: 1189-4563

features generated by millions of previously seen images. Therefore, we are capitalising on the advantages that both training methods have to offer that is the simplicity of AlexNet and the transfer learning capabilities of InceptionV3. Consequently, we are able to make a direct comparison of the feature extraction, rate of training convergence and classification accuracy for brain tumors between the two models.

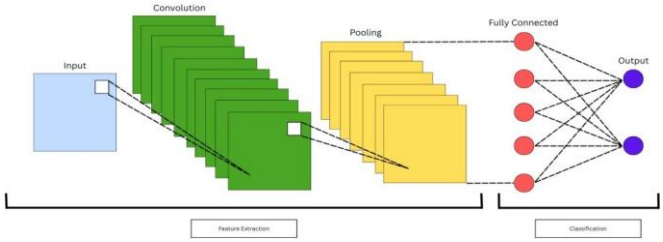

Fig.6. Basic Architecture of Convolution Neural Network

3.3.3 Fine-Tuning Strategies and Layer Freezing

In this study, for the fine-tuning of the pretrained InceptionV3 model we selectively unfreeze layers of the model to adapt it to the MRI image dataset retaining useful learned features that were acquired during original training. For the training of the model, we initially kept the lower layers of InceptionV3 frozen while we retrained the upper layers on MRI images to capture domain specific features. Then we took a gradual approach to reach our goal of fully fine-tuning InceptionV3 using the lower learning rate so as to avoid catastrophic forgetting. AlexNet on the other hand was trained from scratch and hence did not need to keep any of the layers frozen; however, various techniques of regularization and dropout were employed to mitigate the potential for overfitting.

3.4 Proposed Classification Pipeline

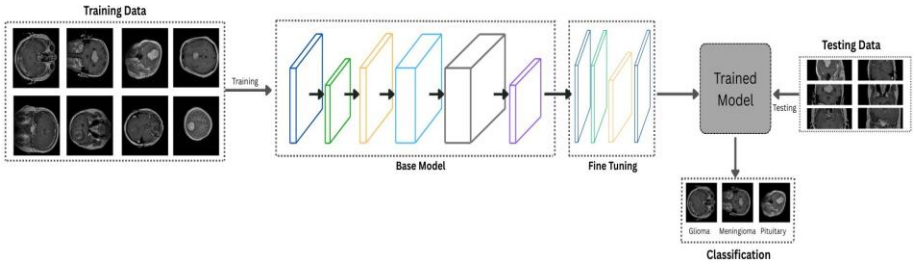

Fig.7. Proposed classification pipeline for brain tumor detection, illustrating the sequential stages of feature extraction, model architecture with custom layers, and optimization strategy

3.4.1 Feature Extraction Process

The feature extraction process changes the input image into an understandable representation for the classification to occur based on that representation. An example of this is AlexNet where the features were learned directly from MRI images by processing them with a series of convolutional and pooling layers that learned hierarchical feature representations of the model (edge detection, textures). With InceptionV3, the model used the pretrained weights from ImageNet as a good set of general purposes for learning features and fine-tuned these features onto the brain MRI data in order to learn the specifics of this domain. Both

ISSN: 1710-0305
PISSN: 1189-4563

AlexNet and InceptionV3 utilized convolutional operations to capture spatially embedded information within the image efficiently, allowing InceptionV3 to utilize its multiple-scale inception modules to learn the ability to capture fine and coarse features simultaneously, allowing the model to have a better ability to differentiate between tumor groupings and improve the overall classification.

3.4.2 Model Architecture and Custom Layers

The AlexNet architecture has five convolutional layers separated by max pooling layers, followed by fully connected layers for classification using dropout and softmax layers. The InceptionV3 architecture adds the ability to learn multi-scale features through the use of Inception modules along with deep convolutional layers and then performs global average pooling, followed by Densely Connected Layers and dropout layers for regularization. To perform brain MRI classification, we added custom layers onto both architectures. Dropout layers are used to control overfitting and Dense Layers are used as inputs for the determination of final classifications and decisions. Each of the models contains an optimal amount of depth, feature richness and the computational power needed for superior performance.

3.4.3 Optimization Algorithms and Hyperparameters

We used the Adam Optimizer for training both models, automatically adjusting learning rates based on each parameter to increase the speed of convergence. The initial learning rate 0.0001 is used for AlexNet during the training phase, where the model is trained from scratch.InceptionV3's initial learning rate of 0.0001 during transfer learning and 0.00001 during fine-tuning. A loss function of categorical cross-entropy with label smoothing (0.1) is used to improve generalization. The batch size of 32 is used for training, along with early stopping and reduction of the learning rate on plateaus to prevent overfitting. Dropout rates of 0.5 in dense layers are used as a regularization technique. The hyperparameters that we set for training both models are selected to maximize efficiency and stability while also allowing for good generalization on the MRI dataset.

3.5 Evaluation Metrics

3.5.1 Accuracy, Precision, Recall, F1-Score

Accuracy – Proportion of correct predictions over all samples.

$$\text{Accuracy} = \frac{True\ Positive + True\ Negative}{True\ Positive + True\ Negative + False\ Positive + False\ Negative}$$

Precision – Proportion of correctly predicted positives among predicted positives.

$$\text{Precision} = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

Recall (Sensitivity) – Proportion of correctly predicted positives among actual positives.

$$\text{Recall} = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

F1-Score – Harmonic mean of precision and recall.

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

ISSN: 1710-0305
PISSN: 1189-4563

3.5.2 Confusion Matrix Analysis

The confusion matrix shows how the classification is performing by showing the number of True Positives, True Negatives, False Positives, and False Negatives from a classification model. It provides insight to see how the errors are being made on each individual class and provides a better way to see where misclassification is occurring. It will allow you to calculate other evaluation metrics such as accuracy, precision, recall, and F1-score.

4. Experimental Setup

4.1 Hardware and Software Specifications

The selection of computer hardware (an Intel Core i7 processor, 32 GB of RAM, and an NVIDIA GeForce RTX 3080 Graphics card) allowed researchers to perform their deep-learning experiments as efficiently as possible in the shortest amount of time, compared to traditional computer systems. The use of high-speed HDDs to store/transfer MRI data provided researchers direct access for write/read operations in comparison to traditional methods. Windows 11 was chosen for use as the Operating System and Python 3.10 was selected as the Programming Language. Data generated by AlexNet and InceptionV3 was used to facilitate research and testing on various combinations of hyperparameters, augmentations, and tuning techniques within the earliest timeframe possible.

4.2 Implementation Environment

The Keras API in TensorFlow 2.x made it possible to use this framework for the implementation of both AlexNet and Inception V3, allowing the design of models flexibly and providing for pretraining the models and setting up a pre-processing pipeline. A number of the features offered by TensorFlow provided the means for creating unique/custom layer designs with dropout and the global pooling layer types. Augmenting the training set using the ImageDataGenerator class allowed for developing an ensemble of possible input data for training and allows for loading of batches in real-time. By enabling the use of the Graphics Processing Unit to perform the forward and backward passes of the model more quickly through the use of CUDA and cuDNN and to visualize the training and validation metrics through TensorBoard, this development environment provided a clear path to support both the training from scratch of AlexNet and the fine-tuning of pre-trained models using Inception V3.

4.3 Training Parameters and Batch Settings

To optimize GPU memory and gradient stability, each model had a batch size of 32 that was utilized throughout the training process on our two models. The Adam optimizer was used at the beginning of training with an initial learning rate of 1e−4, which was then decreased to 1e−5 during its successful fine-tuning on the InceptionV3 model. Early stopping with patience set for five epochs also assisted with limiting overfitting; along with using the ReduceLROnPlateau method of the Adam Optimizer to help fine-tune the model, reducing the learning rate relative to the validation loss. For the two hundred twenty epochs that were run and ultimately the ten total epochs that were used to fine-tune InceptionV3, we froze all the non-fine-tuned layers at the end of the training phase. Regularization due to dropout was introduced to one layer in the case of both Models where all fully connected layers (the input and output connections) utilized dropout rates of 50 percent. As such parameters ensure that convergence occurs in all instances of our fine-tuning process even while also providing for effective training of MRIs utilizing the InceptionV3 architecture.

ISSN: 1710-0305
PISSN: 1189-4563

4.4 Cross-Validation Protocol

The use of stratified k-fold cross-validation is a method used to help evaluate how the models we train perform robustly and produce splits with less bias in regard to the imbalanced data. The training-testing ratio of the dataset was selected multiple times (10:90, 15:85, 20:80, 25:75, and 30:70) to ensure that the splits produced during stratified k-fold cross-validation had the same distributions of all classes across the various folds. Since each fold used to train and validate the model was separate and independent, we can use the average performance metrics (accuracy, loss, and F1 score) to provide a more reliable estimate of how our final model will perform when encountering previously unseen images from MRI scans, whilst minimizing the chance of overfitting. We were able to see how both AlexNet and InceptionV3 models generalized to new images. Cross-validation has provided valuable data to identify how hyperparameter tuning will be affected by the increasing volume of training data for both models. With cross-validation results, end-users may gain further insight into how they might select which model best suits their purposes in the future.

5. Results

5.1 Quantitative Results for Each Model

5.1.1 Performance Comparison of CNN Architectures

Table 2. Performance Report of Inception V3 model on Brain Tumor Dataset for different Train-Test ratios

Train Validation Ratio	Model Used	Precision (%)	Recall (%)	F1-Score (%)	Accuracy (%)
90:10	Inception V3	98.00	98.00	98.00	98.34
85:15		98.00	98.00	98.00	97.91
80:20		98.00	98.00	98.00	98.10
75:25		98.00	98.00	98.00	98.08
70:30		97.00	97.00	97.00	97.14
Average		97.80	97.80	97.80	97.91

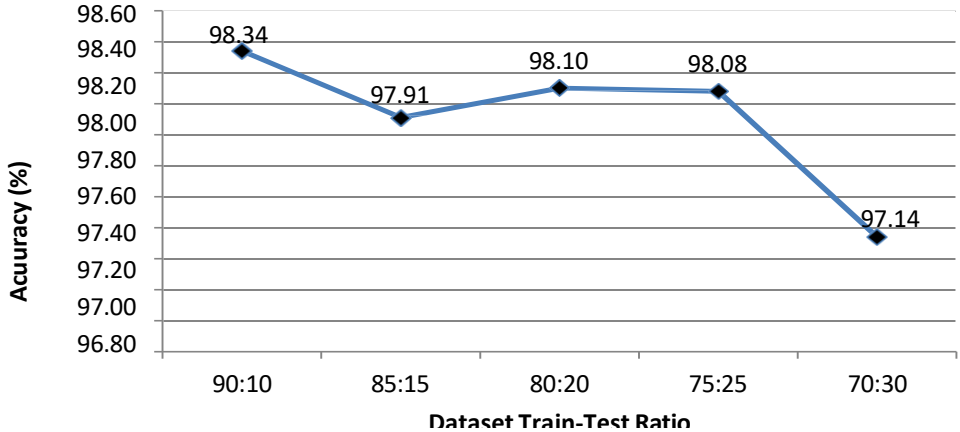

Fig.8. Performance comparison of the Inception-V3 model on the brain tumor dataset across varying train–test split ratios

ISSN: 1710-0305
PISSN: 1189-4563

Table 3. Performance Report of AlexNet model on Brain Tumor Dataset for different Train-Test ratios

Train Validation Ratio	Model Used	Precision (%)	Recall (%)	F1-Score (%)	Accuracy (%)
90:10	AlexNet	94.00	94.00	94.00	93.54
85:15		92.00	92.00	92.00	93.83
80:20		94.00	94.00	94.00	93.30
75:25		93.00	93.00	93.00	93.79
70:30		92.00	92.00	92.00	90.42
Average		93.00	93.00	93.00	92.97

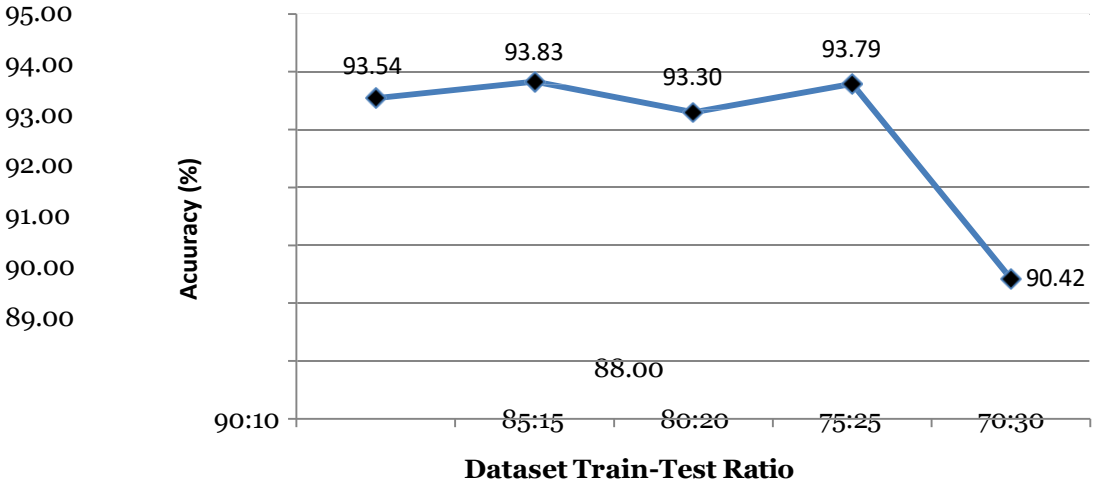

Fig.9. Performance comparison of the AlexNet model on the brain tumor dataset across varying train–test split ratios

5.1.2 Class-Wise Metrics for Glioma, Meningioma, and Pituitary

Fig.10. Confusion matrix of the Inception-V3 model trained and tested on a 90:10 train–test split of the brain tumor dataset

ISSN: 1710-0305
PISSN: 1189-4563

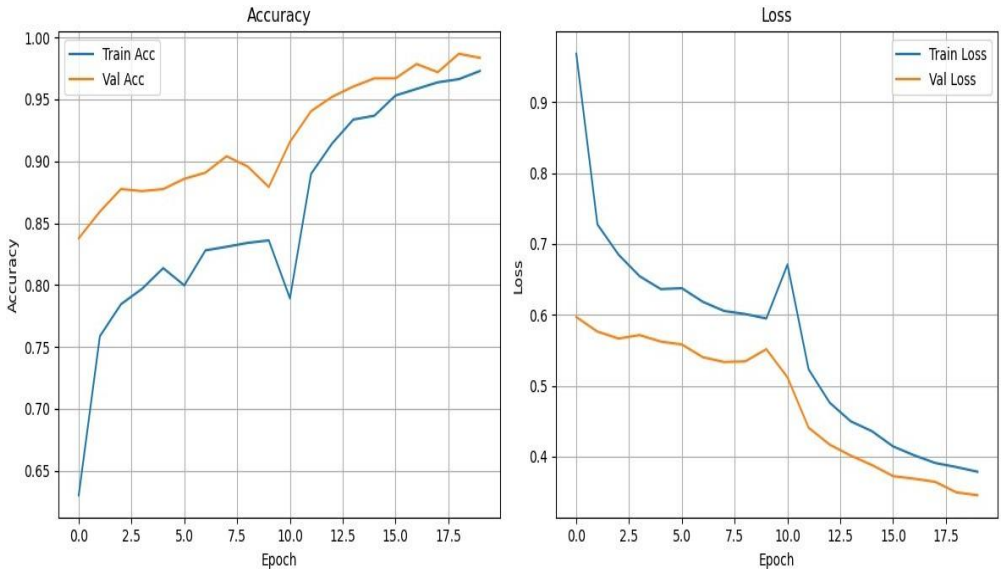

Fig.11. Training and validation performance of the Inception-V3 model on the brain tumor dataset with a 90:10 train–test split

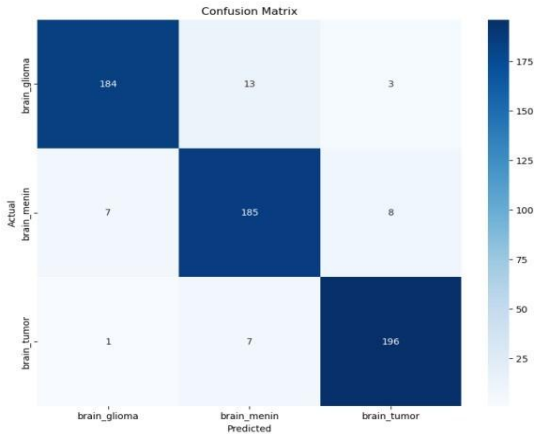

Fig.12. Confusion matrix of the AlexNet model trained and tested on a 90:10 train–test split of the brain tumor dataset

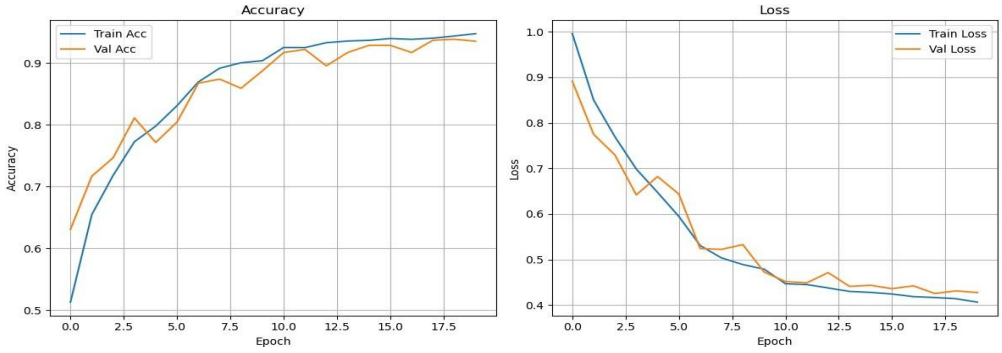

Fig.13. Training and validation performance of the AlexNet model on the brain tumor dataset with a 90:10 train–test split

ISSN: 1710-0305
PISSN: 1189-4563

5.2 Comparison with State-of-the-Art Methods

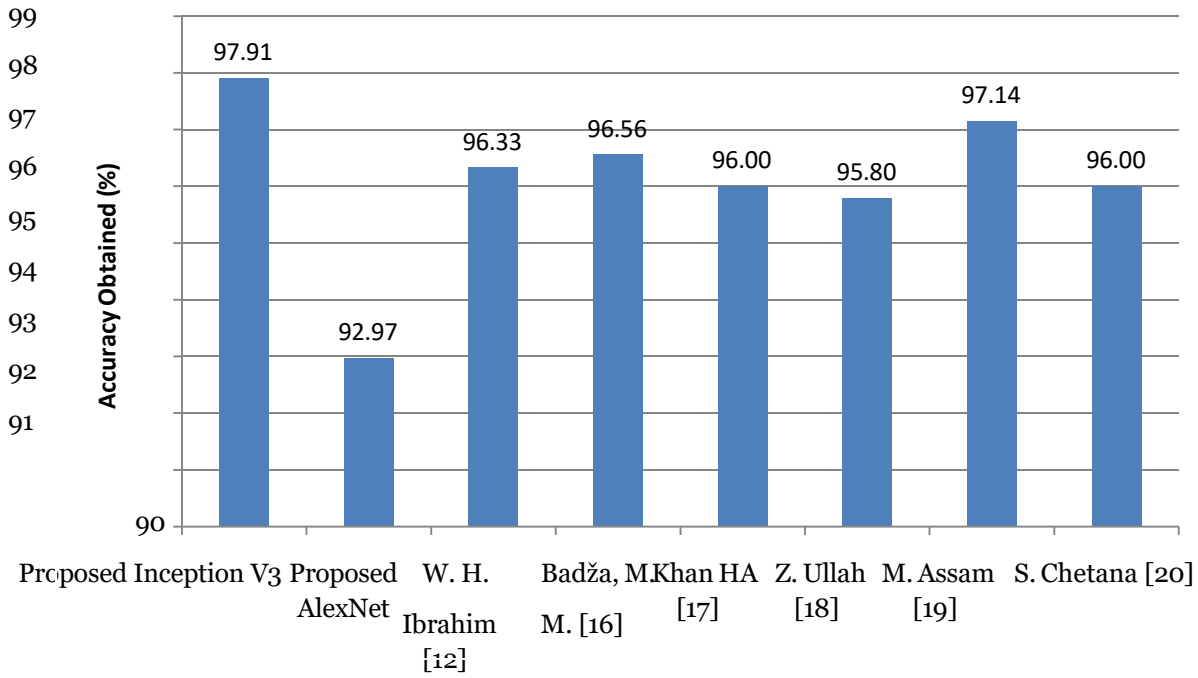

Fig.14. Comparative analysis of the proposed classification models against existing methodologies for brain tumor classification

6. Discussion

6.1 Analysis of Transfer Learning Impact

Pretrained weights from ImageNet provided the foundation for InceptionV3 to rapidly learn advanced feature vectors that can be fine-tuned for brain MRI classification; whereas AlexNet's architecture does not allow the transfer of learned weights as effectively as InceptionV3. A large portion of AlexNet's success was due to the extensive training performed on multiple epochs and applying regularization techniques during its training. Transfer Learning also decreased the training time, provided greater convergence stability and increased generalization capability with respect to the active training dataset. In summary, this study demonstrates that using pretrained deep learning architectures to develop classifiers for medical images is an efficient way to obtain salient feature representation from the underlying data, thus allowing for optimal knowledge reuse and adoption to be implemented in the development of deep learning classifiers in settings where the number of labelled MRI training examples is very limited.

6.2 Interpretation of Classification Trends

The models used for this study are able to accurately identify different types of brain tumors, and the results showed that InceptionV3 outperformed AlexNet by consistently identifying tumors more accurately. The majority of misclassified samples exhibited some type of subtle boundary or low contrast within the image due to the characteristics of these images. The major differences between the networks are that AlexNet is more sensitive to noise and differed in terms of image orientation while InceptionV3 is able to characterize finer characteristics using multi-scale inception modules. The results also indicate that the deeper the

ISSN: 1710-0305
PISSN: 1189-4563

model and the greater the level of pre-training, the better its capacity to manage the variability between multiple tumor classes and the complexity of MRI images. High levels of precision and recall are found for both networks, which demonstrated excellent balanced detection of all three classes across all datasets. The classification results affirm the role and applicability of CNN-based methods in the analysis of MRI for assisting in the automated detection of tumors and indicate that using transfer learning can improve the performance of models used to detect various types of tumor presentation.

6.3 Challenges Encountered

Deep Convolutional Neural Networks (CNNs) represent a powerful tool for processing brain MRI data. In training, AlexNet struggled against overfitting from its high number of parameters compared with the limited training dataset. Therefore, to combat overfitting, extensive data augmentation and dropout methods were employed. Compared with AlexNet, InceptionV3's lower propensity for overfitting is largely due to the use of transfer learning; however, careful fine-tuning was still necessary to prevent catastrophic forgetting. The problem with class imbalance was dealt with by using stratified splitting so that each tumor type was represented in the training sample. Other challenges included inconsistencies in the quality and standardization of MRI and image orientation. However, with careful image augmentation, preprocessing and learning rate scheduling, the two model types trained effectively with the potential for greater accuracy; however, the effect of outliers (extreme variations in data) on the robustness of both models was still considerable.

6.4 Clinical Relevance and Diagnostic Implications

Both deep learning models have demonstrated excellent accuracy and stability suggesting that both neural networks have excellent potential for clinical use in diagnosing all types of brain tumors. The benefits of AlexNet include being a low-complexity and rapid assessment option for initial diagnoses, while the strength of InceptionV3 is its ability to identify tumors with much more precise and accurate results, as well as being ideal for providing diagnostics on injuries requiring immediate attention. Along with providing radiologists with reduced workloads through the automation of classifying tumors from MRIs, the automated MRI classification will yield rapid preliminary assessments of patients and will greatly improve the early detection of tumors. Furthermore, both models are capable of monitoring the progression response of tumors to treatment by means of repeated imaging (MRI) provided to the patient. The use of both CNN models will help improve the consistency and speed with which we interpret diagnostic images. Overall, these examples of the potential impact that deep learning models, particularly through the process of transfer learning can have in augmenting clinical decision-making in the area of neuro-oncology.

6.5 Limitations of the Current Study

The study did show Promise. However, it does have some limitations. For example, Even though the use of augmentations did help with the size of the data set, the final dataset size is still relatively small compared to other datasets making the generalizability of these results to other populations of patients that are not represented in this dataset difficult to achieve. AlexNet took considerably longer to train and had a much higher likelihood of overfitting than the InceptionV3 Model. Therefore, all features learned by the InceptionV3 model from the pre-training of images in the natural useable world, may not capture every nuance of the particular characteristic found in the medical picture domain. In addition to only using MRI Scans, this study did not include multimodal imaging and thus may limit the applicability in the clinical setting. The variations of real-world scanning protocols, noise and artifacts were also not fully investigated. Future work should include using larger multi-institutional datasets, training different architectures of CNNs and including multimodal imaging to continue to validate and improve the clinical deployment of the models.

ISSN: 1710-0305
PISSN: 1189-4563

7. Conclusion and Future Work

7.1 Summary of Key Findings

This study examined how well both Convolutional Neural Networks (CNNs), specifically AlexNet and InceptionV3, could classify MRI scans of brain tumors and proved that AlexNet is able to achieve a higher rate of accurate classifications (94%) than those that will be produced by any other models developed from this data; therefore, this study shows that it can act as a baseline against which to measure the success of all future CNNs trained on MRI brain tumor data. Although InceptionV3's transfer learning provided the best accuracy of 98%, it was further supported by precision, recall, and F1 score measures, which also showed InceptionV3 had superior ability to generalize over the traditional CNN model of AlexNet. Also, adding augmentation of samples during training reduced overfitting and class imbalance to lower degrees. Thus, CNNs should be regarded as useful tools for classifying brain tumors, with pretrained models like InceptionV3 being much more trustworthy than traditional models like AlexNet.

7.2 Contributions to the Field

This research advances the fields of medical imaging and deep learning through demonstrating the effectiveness of CNNs for automatic brain tumor classification. The comparison made between classic CNNs (namely AlexNet) and modern pretrained CNN models (like InceptionV3) demonstrates how the application of transfer learning can increase both accuracy and speed of training. Additionally, the protocols established within the work for preprocessing, augmentation and fine-tuning the MRI data creates solutions to problems associated with small data sets and an unfair distribution of class labels. These contributions will provide a resource for future studies in neuro-oncology and encourage researchers to implement deep learning models into their practice as a reliable, reproducible and clinically applicable tool for diagnosing brain tumors on the basis of MRI data.

7.3 Recommendations for Future Research

Future research endeavors will seek to utilize larger multi-site datasets to enhance generalizability and decrease possible bias when developing models. By utilizing multimodality imaging (e.g., fusing MRI scans with CT or PET) to create better-defined tumors, new and more robust methods of feature extraction can be developed using enhanced CNN construction and attention mechanisms like EfficientNet. In addition, developing optimal transfer learning models through methods such as selective frozen layer models and domain-specific pre-training will create opportunities to further enhance the accuracy of results produced using deep learning methods. Moreover, by employing methods based on Explainable Artificial Intelligence, clinicians will be able to understand how models produce results in more detail than would otherwise be possible. Conducting longitudinal studies to monitor tumor growth and deploy models to the clinic in real-time will provide evidence of the practical application of these methods. All combined, creating more accurate diagnostic approaches using deep learning will firmly establish them as valuable clinical tools.

7.4 Potential Integration into Clinical Workflows

The integration of both AlexNet and InceptionV3 can improve the radiology workflow the most. The con for AlexNet is that it will not provide the accurate diagnosis of advanced preliminary results (high-quality data) but will provide speed. Conversely, the advantage of InceptionV3 is that it can render the greatest accuracy when determining the appropriate and essential decisions concerning the diagnosis. A hospital using both AlexNet and InceptionV3 will likely reduce workload for radiologists and will expedite early diagnosis (tumors) and continual assessment of how to treat patients. These systems, when integrated with a PACS or EHR system, may also help radiologists by providing additional support in reaching an educated diagnosis at the time of patient presentation combined with an imaging report. For example, if a user-

friendly interface with visualization tools is built into the systems, this may create an enhanced capability that will assist radiologists to interpret the results. Therefore, CNNs are an effective and valid complement to neuro-oncology MRI research articles.

References

[1] Kamepalli, HariKishore & Kalaparti, Viswanadh & Kesavadas, Chandrasekharan. (2023). Imaging Recommendations for the Diagnosis, Staging, and Management of Adult Brain Tumors. Indian Journal of Medical and Paediatric Oncology. 44. 026-038. 10.1055/s-0042-1759712.

[2] Abd-Ellah MK, Awad AI, Khalaf AAM, Hamed HFA. A review on brain tumor diagnosis from MRI images: Practical implications, key achievements, and lessons learned. Magn Reson Imaging. 2019 Sep;61:300-318. doi: 10.1016/j.mri.2019.05.028. Epub 2019 Jun 5. PMID: 31173851.

[3] Rios Piedra EA, Taira RK, El-Saden S, Ellingson BM, Bui AAT, Hsu W. Assessing Variability in Brain Tumor Segmentation to Improve Volumetric Accuracy and Characterization of Change. IEEE EMBS Int Conf Biomed Health Inform. 2016 Feb;2016:380-383. doi: 10.1109/BHI.2016.7455914. Epub 2016 Apr 21. PMID: 28670648; PMCID: PMC5489257.

[4] Dorfner FJ, Patel JB, Kalpathy-Cramer J, Gerstner ER, Bridge CP. A review of deep learning for brain tumor analysis in MRI. NPJ Precis Oncol. 2025 Jan 3;9(1):2. doi: 10.1038/s41698-024-00789-2. PMID: 39753730; PMCID: PMC11698745.

[5] Li,Q. (2025). Brain Tumor Detection Based on MRI Images and Artificial Intelligence. Applied and Computational Engineering,139,67-75.

[6] Alshuhail A, Thakur A, Chandramma R, Mahesh TR, Almusharraf A, Vinoth Kumar V, Khan SB. Refining neural network algorithms for accurate brain tumor classification in MRI imagery. BMC Med Imaging. 2024 May 21;24(1):118. doi: 10.1186/s12880-024-01285-6. PMID: 38773391; PMCID: PMC11110259.

[7] Khan MA, Hussain MZ, Mehmood S, Khan MF, Ahmad M, Mazhar T, Shahzad T, Saeed MM. Transfer learning for accurate brain tumor classification in MRI: a step forward in medical diagnostics. Discov Oncol. 2025 Jun 9;16(1):1040. doi: 10.1007/s12672-025-02671-4. PMID: 40490586; PMCID: PMC12149374.

[8] Jain, J., Kubadia, M., Mangla, M., & Tawde, P. (2023). Comparison of Transfer Learning Techniques to Classify Brain Tumours Using MRI Images. Engineering Proceedings, 59(1), 144. https://doi.org/10.3390/engproc2023059144

[9] Gorenshtein A, Liba T, Goren A. Lightweight Transfer Learning Models for Multi-Class Brain Tumor Classification: Glioma, Meningioma, Pituitary Tumors, and No Tumor MRI Screening. J Imaging Inform Med. 2025 Sep 19. doi: 10.1007/s10278-025-01686-1. Epub ahead of print. PMID: 40973910.

[10] Rasheed, M., Jaffar, M.A., Akram, A. et al. Improved brain tumor classification through DenseNet121 based transfer learning. Discov Onc 16, 1645 (2025). https://doi.org/10.1007/s12672-025-03501-3

[11] Y. Zhang, Z. Dong, L. Wu, and S. Wang, “A hybrid method for MRI brain image classification,” Expert Systems with Applications, vol. 38, no. 8, pp. 10049–10053, 2011, doi: 10.1016/j.eswa.2011.02.012.

[12] W. H. Ibrahim, A. A. A. Osman and Y. I. Mohamed, "MRI brain image classification using neural networks," 2013 INTERNATIONAL CONFERENCE ON COMPUTING, ELECTRICAL AND ELECTRONIC

ENGINEERING (ICCEEE), Khartoum, Sudan, 2013, pp. 253-258, doi: 10.1109/ICCEEE.2013.6633943.

[13] N. Abdullah, U. K. Ngah and S. A. Aziz, "Image classification of brain MRI using support vector machine," 2011 IEEE International Conference on Imaging Systems and Techniques, Batu Ferringhi, Malaysia, 2011, pp. 242-247, doi: 10.1109/IST.2011.5962185.

[14] S. Kumar, C. Dabas, and S. Godara, “Classification of Brain MRI Tumor Images: A Hybrid Approach,” Procedia Computer Science, vol. 122, pp. 510–517, 2017, doi: 10.1016/j.procs.2017.11.400.

[15] Z. Jia and D. Chen, "Brain Tumor Identification and Classification of MRI Images Using Deep Learning Techniques," in IEEE Access, vol. 13, pp. 123783-123792, 2025, doi: 10.1109/ACCESS.2020.3016319.

[16] Badža, M. M., & Barjaktarović, M. Č. (2020). Classification of Brain Tumors from MRI Images Using a Convolutional Neural Network. Applied Sciences, 10(6), 1999. https://doi.org/10.3390/app10061999

[17] Khan HA, Jue W, Mushtaq M, Mushtaq MU. Brain tumor classification in MRI image using convolutional neural network. Math Biosci Eng. 2020 Sep 15;17(5):6203-6216. doi: 10.3934/mbe.2020328. PMID: 33120595.

[18] Z. Ullah, M. U. Farooq, S.-H. Lee, and D. An, “A hybrid image enhancement based brain MRI images classification technique,” Medical Hypotheses, vol. 143, art. no. 109922, 2020, doi: 10.1016/j.mehy.2020.109922.

[19] M. Assam, H. Kanwal, U. Farooq, S. K. Shah, A. Mehmood and G. S. Choi, "An Efficient Classification of MRI Brain Images," in IEEE Access, vol. 9, pp. 33313-33322, 2021, doi: 10.1109/ACCESS.2021.3061487.

[20] Srinivas, Chetana, K. S., Nandini Prasad, Zakariah, Mohammed, Alothaibi, Yousef Ajmi, Shaukat, Kamran, Partibane, B., Awal, Halifa, Deep Transfer Learning Approaches in Performance Analysis of Brain Tumor Classification Using MRI Images, Journal of Healthcare Engineering, 2022, 3264367, 17 pages, 2022. https://doi.org/10.1155/2022/3264367

[21] Rahman, Md Mizanur (2024), “Brain Cancer - MRI dataset”, Mendeley Data, V1, doi: 10.17632/mk56jw9rns.1