Developing an Optimize Deep Learning Framework for Brain Tumors Detection and Classification

Sayani Ghosh[1], Souvik Mazumder[2], Prasenjit Kundu[3], Arpita Dey[4]*

[1]Research Scholar, Institute of Engineering and Management, University of Engineering and Management, Kolkata, India

[2]Research Scholar, Swami Vivekananda University, Kolkata, India
[3]Associate Professor, MBA Dept., Institute of Engineering and Management, University of Engineering and Management, Kolkata, India
[4]Final year MBA Student, Institute of Engineering and Management, Kolkata, India

Abstract: Brain tumors are considered one of the present health challenges. Perhaps the early spotting could be an important factor in their effective treatment and in the good outcome of the patients. But, due to the asymptomatic nature and location of brain tumor, early detection through tissue diagnosis is difficult. This research design an empirical deep learning- based framework to detect and classify brain tumors from Magnetic Resonance Imaging (MRI) images datasets. The research work aims at improving the quality and reliability of radiological images through preprocessing techniques such as resizing, noise reduction, normalization, and segmentation to classify tumors more precisely. The aim of this study to develop hybrid deep learning based model to ensure that benign and malignant brain tumors are appropriately differentiated and also attempt to provide a framework for predicting the type of malignant tumors. This article evaluated three deep learning frameworks on radiological tumor images dataset to train the model and based on accuracy, validation and performance, the most optimize model is selected to apply on test dataset to automatically classified brain tumors. Authors of our study are hopeful for adding value to diagnostic precision. Automation in the classification process optimizes efficiency for faster and more reliable diagnostic support. Future research are expected to incorporate multi-modal data to enhance the diagnostic accuracy and smoothness of its clinical application. This may most likely improve the speed and precision of the diagnosis made by health professionals using deep learning (DL) in the diagnosis of brain tumors and help to improve patient care.

1 Introduction

The brain is primarily divided into three regions, namely the brain stem, the cerebrum, and the cerebellum. Each region is specialized to perform some unique functions that are

*Corresponding author: prasenjit.kundu@iem.edu.in

obligatory in enhancing the over-all neural activity of the body. The brain stem links the brain to the spinal cord and is primarily accountable for managing necessary autonomic activities such as regulating oxygen intake and pulse. The cerebrum is said to be the significant functional part and governs higher order functions like cognitive function, emotional involvement, and sensory perception, as well as regulation of voluntary movements of the body. The core thought process area and the area of decision making are also attributed to it. Cerebellum is located behind the cerebrum and it governs the fine motor skills and maintains balance and coordination of movement [1]. The weight of a typical human intracranial is about 1.2 to 1.4 kilograms. In volume, the average size of male and female brains is about 1260 cm^3 and 1130 cm^3 respectively [2]. WHO (World Health Organization) has identified the grades of brain tumors into four categories which range from grade I to IV. Grades I and II grow gradually and act as less aggressive whereas grades III and IV are the swiftly increasing ones showing aggressive behaviour thus, in most cases they have a poor prognosis [3]. Brain tumors are classified based on their grades, which describe the rate of growth and potential dissemination.
Grade I(Well differentiated): These are moderate growth of tumors that do not spread fast and have a higher likelihood of being completely removed by surgery, offering a favourable long-term prognosis (e.g., pilocytic astrocytoma).

Grade II (Moderately differentiated): These tumors develop gradually and may spread to nearby tissues. There is a potential for progression to a higher-grade tumor. Despite apparent surgical removal, recurrence is possible (e.g., oligodendroglioma).
Grade III (Poorly differentiated): These tumors grow more rapidly than moderately differentiated tumors and may invade surrounding tissues. Surgery alone is often insufficient, and radiation or chemotherapy is usually required in addition to surgery (e.g., anaplastic astrocytoma).
Grade IV (Undifferentiated) gliomas spread quickly in the brain. They can also use blood vessels to grow faster, like glioblastoma multiforme [4].

Spotting and segmentation of brain tumors have been significant applications in medical imaging. Most of them are cases of distinguishing between benign and malignant using non-invasive methods. Recently ML has become more and more significant in diagnosis tasks of medicine by partly automating the diagnostic process and improving the accuracy of the results. In this respect, image pre-processing is indispensable because it enhances the quality of the images, extracts meaningful features, and makes them ready to be analysed. The current research is based on different ML algorithms with much emphasis on pre-processing and feature extraction for the best outcome in classifying brain tumors. Pre-processing is a vital step involving multiple serial processes to be applied to raw images, cleaning the noisy data from the images, enhancing those features that are important, standardizing format and size, and increasing the quality and usability of the image. Generally, pre-processing involves:

Image Resizing: Images are resized to a uniform size for consistency across the entire dataset.
Noise Reduction: Processes like Gaussian blurring deletes unwanted noise and details unessential for better clarity of images.
Normalization: Normalization adjusts pixel intensity values to make raw image data suitable for model training.
Segmentation: Segmentation separates the tumor area from the background to focus on important features.

These steps have become very important pre-processing techniques that allow ML algorithms to improve their accuracy and reliability in performing several tasks related to medical image classification.

2 Literature Review

Various regions of the brain perform specialized functions. The frontal lobe assists with thinking, reasoning also controlling movements. The parietal lobe enables us to feel touch and understand space and direction. The temporal lobe manages memory and hearing. The occipital lobe helps us see and understand what is in front of us. Each part works together to help us function. The cerebral cortex also plays an important role. Which is essentially the brains layer. Boasts an area peppered with nerve cells known as cortical neurons that work together to process and handle incoming information effectively [5]. The brain stem measures about 7–10 cm in length and links the brain to the spinal cord. It contains important nerves that control eye movements. It also helps with keeping our balance. The brain stem plays a role in controlling our breathing. It is vital for many basic body functions. The transmission of nerve signals originating from the thalamus, in the brain passes through the brain stem before reaching the cord and branching out into parts of the body. The brain stem comprises three segments: Midbrain: Responsible for motor control, hearing, vision and eye movements. Pons: Aids in breathing, communication between different brain areas, and sensory information, such as touch. Medulla Oblongata: Regulates automatic functions such as blood pressure, swallowing, and sneezing [6].Deep learning techniques especially Convolutional Neural Networks (CNNs) have demonstrated significant success in accurately classifying brain tumors from MRI scans [7]. Hybrid deep learning frameworks further improve performance by integrating multiple network architectures to manage tumor variability effectively [8]. Advanced feature extraction methods, particularly those utilizing deep features, contribute substantially to increasing classification precision and improving model interpretability [9]. Moreover, studies focusing on CNN-based models emphasize the importance of optimization strategies and data augmentation to develop robust and reliable brain tumor diagnostic systems [10]. Cancerous tumors can grow fast and spread in the brain. This makes the condition worse. The brain usually replaces old or damaged cells with new and healthy cells. If mature cells are not cleared then they can turn into tumors. Tumors can be very harmful. MRI and CT scans can help find brain tumors early which is improving the chances of recovery [11].

3 Research Gap

• Most of the existing research focuses solely on MRI images, missing the opportunity to combine data from other sources like CT scans, genetic profiles or patient medical histories. Integrating these different types of data could lead to more accurate and reliable models for brain tumor classification but this area remains largely unexplored.

• Many studies have not addressed how ML models can fit into the daily routines of healthcare professionals. Without user-friendly platforms or interfaces, these models are difficult to adopt in real-world clinical settings. This is a significant gap that needs to be addressed to make these technologies truly impactful.

• Most research has tested models on specific MRI datasets. These datasets do not show the variety found in real life. Models need testing on data from different MRI machines. They should also work well for different patient groups. Testing in varied imaging conditions is important too. This is a key area for future research.

3.1 Research Objective

The aim of this study is to create an optimized and precise framework that will detect and classify brain tumors. It will work with MRI images to identify tumors. The proposed framework will also determine whether the tumor is cancerous or non-cancerous using a hybrid deep learning technique. It uses machine learning (ML) to do this. Advanced image processing helps study the MRI scans. This can help doctors diagnose faster and more accurately.

4 RESEARCH METHODOLOGY

The proposed framework suggests developing a ML based system that could identify and classify brain tumors based on a dataset of MRI images (Dataset Link 1, 2024; Dataset Link 2, 2024). This system uses more powerful deep models combined with various pre-processing techniques in order to achieve enhanced classification performance. The overall methodology includes the following stages: data collection, data preparation, model optimization, training process, performance testing and output generation.

4.1 Data Collection

The MRI images of the brain tumors are used for this purpose, fetched from publicly available sources. The total images comprising the dataset, amount to 2363. They have been divided into 1654 training images and 709 testing images which constitute 30% of the data. All of these images are described as either benign, malignant or no tumor, thus providing information in the process of supervised learning concerning the classification of tumors.

4.2 Data Pre-processing

Pre-processing steps were done to improve and make the MRI dataset more consistent:
Resize Images: Resizing each one of the MRI images into the same size in pixels, 256 x 256 pixels, in order to ensure consistency within the dataset and allow comparability with the input size of the deep learning models.
Noise Reduction: The unwanted noise in the MRI images is reduced using Gaussian blurring that enhances the features' clarity in the images.
Normalization: The pixel values of the images are adjusted to range from 0 to 1 for better consistency and faster training.
Segmentation: The tumor regions are segmented to isolate the region of interest concerning the background, making the focus of the model on the relevant features.

4.2.1 Model Selection and Architecture

ResNet152V2 is a deep neural network with 152 layers. It uses skip connections to solve the vanishing gradient problem. These connections help the model learn better by linking certain layers. This makes training deep networks easier. ResNet is very useful for image classification. It detects important patterns in images. Because of this, it works well for medical image analysis. The "V2" version improves training speed and accuracy with batch normalization.

InceptionResNetV2 is a hybrid architecture that combines the "Inception" model's multi-scale feature extraction with the "ResNet" architecture's residual connections. It is designed to capture a wider range of image features at different scales while maintaining efficiency. This architecture excels in capturing complex patterns, which is useful in distinguishing between benign and malignant brain tumors in MRI images.

Xception (Extreme Inception) is an advanced architecture that enhances the Inception model by substituting standard convolutional layers with depthwise separable convolutions. This change allows the model to learn more efficient feature representations with fewer parameters, resulting in improved performance with less computational cost. Xception is particularly useful in image classification tasks where efficiency and accuracy are both critical.

Three different deep learning models are used for comparison and segmentation:
ResNet152V2: A pre-trained model with frozen layers, is used to extract features. Extra dense and dropout layers are added to customize it for the classification task.
InceptionResNetV2: Another high-performing architecture, InceptionResNetV2, is employed for its ability to capture more complex patterns. Similar to the ResNet152V2, it is fine-tuned with additional dense layers.
Xception: A third architecture, Xception, is used to evaluate its performance in comparison to the other models. It is known for its depth-wise separable convolutions, which can enhance feature extraction.

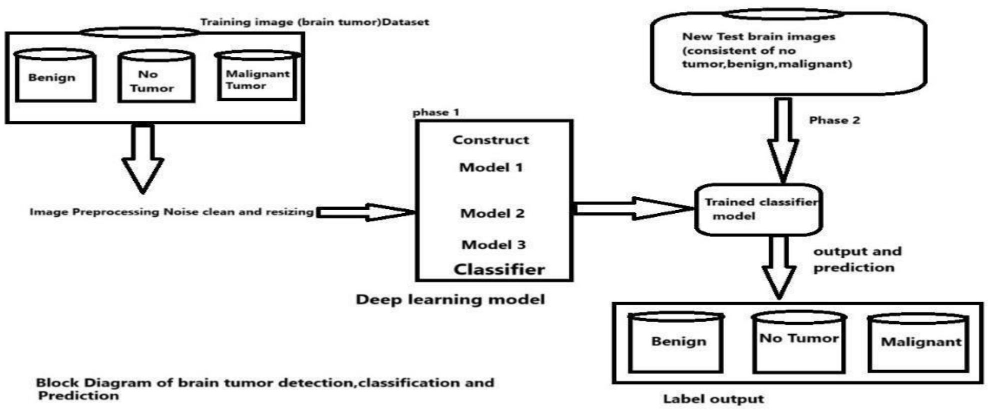


Figure 1: Block Diagram of detection, classification and prediction

The study is depicted in Figure 1 is showcasing the block diagram of our work for brain tumor detection, classification, and prediction process using a deep learning model that includes multiple stages: image pre-processing, model construction, training and label prediction.

4.2.2 Model Training

The models are trained on MRI images using categorical cross-entropy as the loss function and Adam as the optimizer. They are trained for several epochs to ensure accurate learning. The training is carried out using a 70-30 train-validation split to ensure the models are not overfitting. During training, ResNet152V2 required 60 epochs, InceptionResNetV2 completed in 20 epochs, and Xception trained for 60 epochs.

4.2.3 Model Evaluation

After training the models are tested on a validation set to measure accuracy and loss:

- ResNet152V2 achieved 90% accuracy on test set.
- Xception reached 89% accuracy.
- InceptionResNetV2 an accuracy of 86% had.

Confusion matrices also reports were used to check precision, recall and F1-scores. These scores were calculated for benign, malignant along with no tumor categories.

4.2.4 Prediction and Deployment

ResNet152V2, identified as the best model from validation results, is used to classify new MRI scans. A custom function is developed to load new MRI images, preprocess them, and predict their class using the trained model. The predicted results are categorized into folders representing different tumor types (e.g., Glioma Tumor, Meningioma Tumor, No Tumor) for easy interpretation.

4.2.5 Implementation and Testing:

A simple interface is made using Python. Users can load their MRI images into it. The system gives predictions using a trained model. It is tested with new MRI images. These images were not used in training the model. This helps check if the system is reliable.

5 Discussion, Result & Analysis

This study used modern deep learning to identify and classify brain tumors. Three models were tested: Xception, InceptionResNetV2 and ResNet152V2. Each model was fine-tuned using specific layers. This includes the findings and what they mean. The results, performance of the models during testing, validation, and training phases were analysed and discussed below.

5.1 Model Performance Analysis

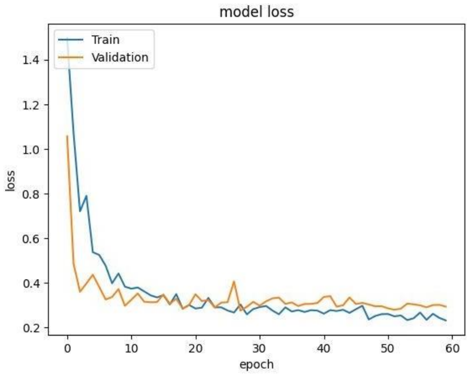
MRI images were grouped into glioma, meningioma and no tumor to test and compare the performance of three models. The results showed that the models had varying accuracy, precision, recall and F1-scores. The results are summarized below:

5.1.1 ResNet152V2 Model:

The ResNet152V2 model reached 90% validation accuracy. It showed strong performance overall. There was some confusion between benign tumors and healthy brain scans. This happened because their MRI features looked similar. However the model was very good at telling benign and malignant tumors apart. It performed well during the evaluation phase. The validation loss plot showed a steady drop. This means the model was learning well. However, there was a small gap between training and validation losses. This suggests mild overfitting. Adding more diverse data could help fix this issue. Using extra regularization techniques could also improve the model.

The model performance analysis is shown in Figure 2 and Figure 3 which illustrate the

ResNet152V2 model's effectiveness in the training phase by displaying its accuracy (Figure 3) and training and validation loss (Figure 2).



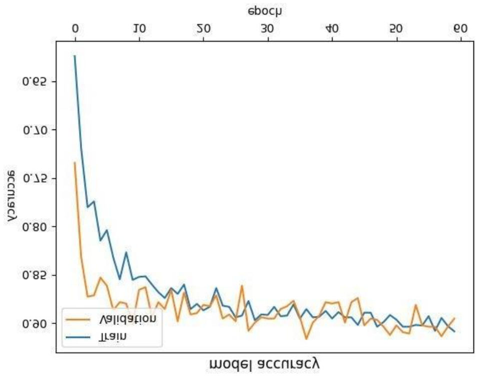
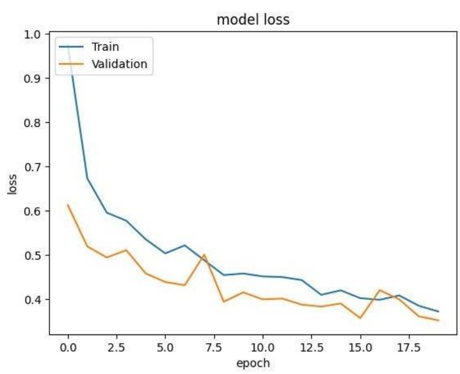


Figure 2: Model Loss InceptionResNetV2 Model **Figure 3:** Model AccuracyInceptionResNetV2 Model

5.1.2 InceptionResNetV2 Model:

The model's validation accuracy is 86%. This is slightly lower than expected. But it was able to reduce false negatives. This means the model correctly identified more tumors that were missed before. It did this by learning to classify tumors better. This is very important in clinical settings. Failing to detect a tumor can lead to severe outcomes. The model also performed well with malignant tumors. It had better precision for these cases. This means fewer healthy tissues were wrongly classified as tumors. It also had better recall for malignant tumors. This means it caught more aggressive cancer cases. Altogether the model is more reliable in finding critical tumors.

How accurate the model is (Figure 5) and the training and validation loss (Figure 4) of InceptionResNetV2 are displayed in Figures 4 and 5 which demonstrate the model's performance in segmentation of tasks.



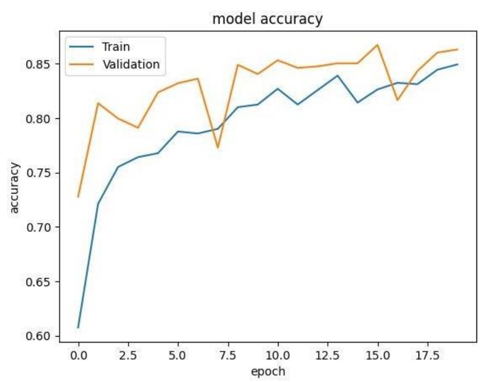
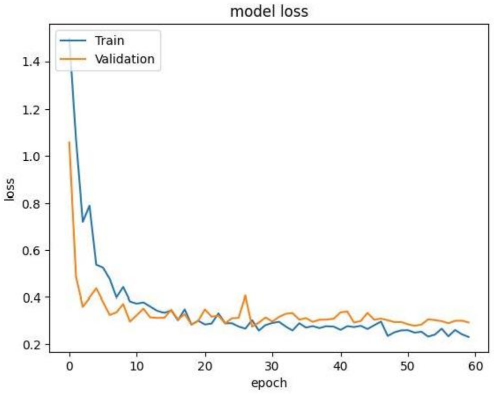


Figure 4: Model Loss InceptionResNetV2 Model **Figure 5**: Model AccuracyInceptionResNetV2 Model

5.1.3 Xception Model:

The Xception model achieved a validation accuracy of 89%. This means it performed well on new data. It shows the model can generalize effectively. The confusion matrix gave more insights. It showed high precision for all three classes. This means the model made fewer wrong predictions. The recall values were also high. This means the model correctly detected most cases. Both precision and recall were balanced. The model handled all three classes equally well. There were fewer mistakes overall. This makes the model reliable for segmentation tasks.

The performance of the Xception model is represented in Figure 6 and Figure 7 which illustrate the model's training and validation loss (Figure 6) and accuracy (Figure 7) demonstrating its generalization capability and effective classification performance.



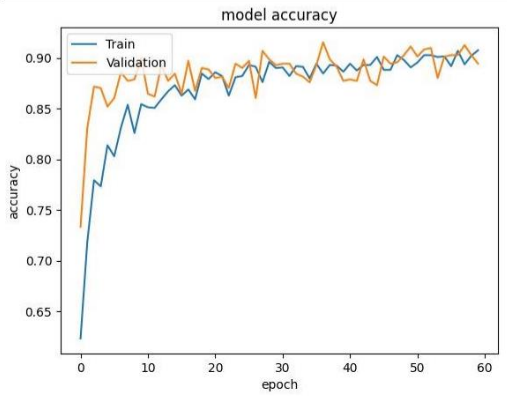


Figure 6: Model Loss InceptionResNetV2 Model **Figure 7:** Model Accuracy InceptionResNetV2 Model

5.2 Performance Comparison of Different Models for Brain Tumor segmentation

Table 1 The study compares the performance of ResNet152V2, InceptionResNetV2, and Xception models for brain tumor detection. These models were trained using the Adam optimizer, which helps adjust the learning rate for better learning. The categorical cross-entropy loss function was used to measure classification errors. The performance of the models was evaluated using key metrics. These include accuracy (correct predictions), precision (correctly identified tumors among predicted ones), recall (correctly detected actual tumors), and F1-score (a balance between precision and recall). These metrics help in understanding how well the models classify MRI images.

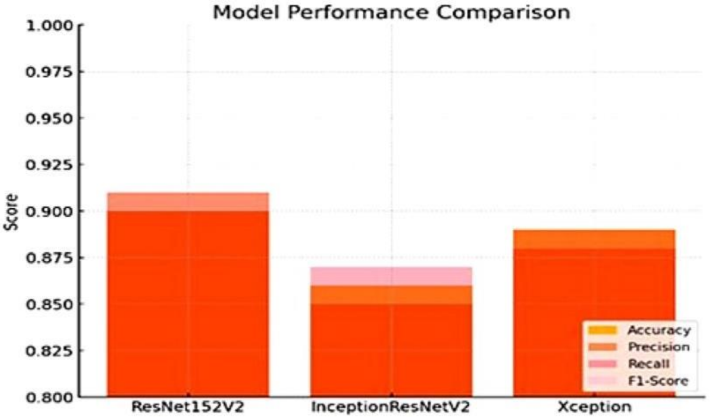


Figure 8: Model Performance Comparison

This is the comparison of performances of three deep learning models shown in (Figure 8) : ResNet152V2, InceptionResNetV2 and Xception in terms of correctness, exactness, Recall and F1-Score. y-axis performance scores range between 0.8 and 1.0, and there are listings of three models on the x-axis.

ResNet152V2: Provides the best overall performance, where all the metrics— correctness rate, exactness, detection rate and F1-score—score above 0.9.

InceptionResNetV2: Performs weakest out of the three models, with scores for most metrics closer to 0.85.

Xception: They achieve mid-range performance where their metrics are floating just below ResNet152V2 but above InceptionResNetV2. The chart highlights ResNet152V2 as the best- performing model for all metrics, which may be an indication that it is more reliable when related to the task being evaluated. Otherwise, InceptionResNetV2 falls behind when comparing the two models. This allows us to draw a clear view of each model's strengths and weaknesses for decision-making and adaptation.

Model	Class	Precision	Recall	F1-Score	Accuracy
ResNet152V2	Glioma	0.87	0.89	0.88	0.9
ResNet152V2	Meningioma	0.87	0.85	0.86	
ResNet152V2	No Tumor	0.96	0.94	0.95	
InceptionResNetV2	Glioma	0.85	0.84	0.84	0.86
InceptionResNetV2	Meningioma	0.86	0.84	0.85	
InceptionResNetV2	No Tumor	0.88	0.95	0.91	
Xception	Glioma	0.88	0.91	0.89	0.89
Xception	Meningioma	0.88	0.88	0.88	
Xception	No Tumor	0.98	0.93	0.96	

Table 1: Performance Comparison of models

5.3 Comparative Analysis

A comparative assessment of the three models showed ResNet152V2 outperforming both InceptionResNetV2 and Xception in terms of segmentation accuracy, exactness and recall. Because it was able to capture intricate spatial patterns within the MRI images, false positives and false negatives were reduced, thus making this architecture most suitable for this type of classification task.

The confusion matrix shows that the ResNet152V2 model performs well overall in classifying MRI images into Glioma Tumor, Meningioma Tumor and No Tumor categories with the highest accuracy in detecting the No Tumor class. It correctly classifies most images (260 Glioma, 240 Meningioma, and 140 No Tumor), but there are some misclassifications, particularly between Glioma and Meningioma (33 Meningioma images classified as Glioma and 25 Glioma images as Meningioma). This overlap suggests shared features between these tumor types, making them harder to distinguish.

Test	Test Statistic	P- Value
Paired t-test (ResNet vs Inception)	0.142857143	0.909665529
Paired t-test (ResNet vs Xception)	-0.333333333	0.795167235
ANOVA	0.055555556	0.981559988

Table 2: Statistical Test Results and Model Comparisons

The following Table 2 gives the results from statistical tests comparing ML models between ResNet, Inception, and Xception.

t-test, paired: ResNet vs. Inception

Test Statistic: 0.1429

P-Value: 0.9097

The high p-value indicates no statistically significant difference between the performances of ResNet and Inception.

Paired t-test (ResNet vs. Xception):

Test Statistic: -0.3333

P-Value: 0.7952

Again, the high p-value indicating no notable difference in performance between ResNet and Xception.

ANOVA:
Test Statistic: 0.0556

P-Value: 0.9816 This test compares the performances across all three models—ResNet, Inception, Xception—at once. The very high p-value implies no statistically meaningful differences between the models' results.

Post-hoc Analysis: Following the ANOVA test, Tukey's HSD (Honestly Significant Difference) test was applied to perform pairwise comparisons between the three models—ResNet152V2, InceptionResNetV2, and Xception. This post-hoc test was chosen to control for Type I error across multiple comparisons and to determine which specific models had statistically significant differences in performance.

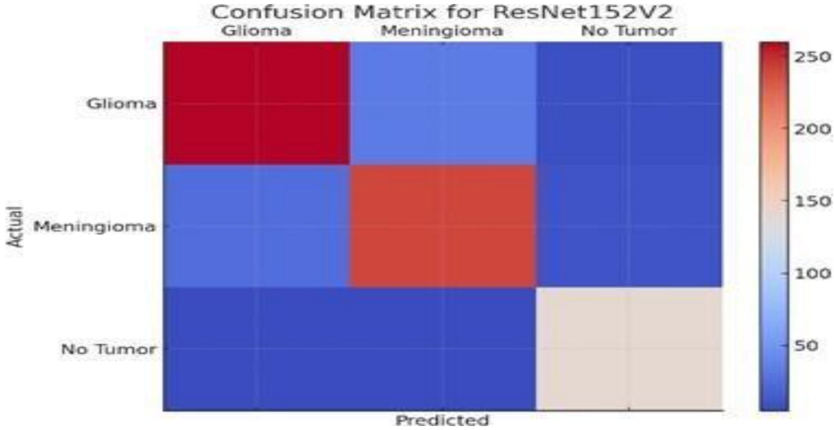


Figure 9: Confusion matrix

The performance of the ResNet152V2 model is illustrated in Figure 9, which presents the confusion matrix, showing the classification distribution among Glioma, Meningioma and No Tumor categories.

Visualization of The Results: The accuracy and loss curves for each model showed a consistent rise over the epochs, indicating that learning was successful.

5.4 Confusion Matrices

All three models had some misclassified cases in their confusion matrices. The

InceptionResNetV2 model often confused benign tumors with no tumors. This was a key issue in its performance. ResNet152V2 had the highest accuracy. It performed the best at classifying brain tumors. The results show it is the most reliable model for MRI images.

Practical Challenges in Clinical Deployment

Data Variety and Integration **-** Right now, the model mainly works with MRI scans, which may limit how well it performs in different situations. To improve this, it should also be trained on other medical data like CT scans, patient history, and genetic profiles. Using more types of data can make the model more reliable for diagnosing brain tumors. Also, differences in MRI machines and patient demographics can make training more complex.

Approval from Medical Authorities **-**Before AI models can be used in real hospitals, they must go through thorough testing and approval from organizations like the FDA (Food and Drug Administration) in the U.S. or CE certification in Europe. These approvals ensure that the model is safe and accurate for medical use.

Real-Time Performance and Efficiency**-** Hospitals need AI models that work quickly and efficiently. The system should be able to analyze large amounts of data in real time so doctors can get instant and accurate results. The model should also be optimized to work on different types of hospital computers and systems.

Training for Medical Staff- To make AI useful in real-world hospitals, it needs to have easy-to-use interfaces. Doctors and nurses must be trained to use AI-generated results in their decision-making process. The model should assist medical professionals rather than replace them.

Data Privacy and Security **-** Since medical data is highly sensitive and strict rules must be followed to protect patient privacy. The AI system should have strong security measures to prevent unauthorized access and keep patient information safe.

6 Conclusion

This research employs deep learning techniques to identify and categorize brain tumors in MRI scans. It studies three models: Xception, InceptionResNetV2 and ResNet152V2. These models are compared to see how well they work. The goal is to identify three types of cases: glioma, meningioma and no tumor. The study looks at how accurately each model can tell these apart. This helps in understanding which model is the best for this task. All the models performed well in classifying the cases. But their accuracy, precision, recall and F1-scores were quite different. ResNet152V2 was the best model for this task. It reached the best accuracy of 90% with high precision and recall for all categories. This means it made fewer mistakes and detected more correct cases. It could also recognize complex patterns in MRI images. This made it the best choice for the task. The Xception model came second. It had a balanced accuracy of 89%. InceptionResNetV2 performed slightly worse. Its accuracy was 86% which was lower than the other two models. Despite these small differences, all three models showed promise for clinical use, provided more study and development were conducted.

This study focuses on using deep learning models to detect and classify brain tumors in

MRI scans. The study uses three models—ResNet152V2, InceptionResNetV2 and Xception—to compare their ability to classify tumors as Glioma, Meningioma or No Tumor. All models could distinguish between the classes but their accuracy, precision, recall and F1-scores differed significantly.

7 Limitation and Future Scope

This study shows that ML can help detect brain tumors in MRI scans, but it has some limitations. The model's accuracy is moderate and its reliance solely on MRI limits its generalizability. Class imbalance and tumor variability also affected classification performance. Importantly this research is not a fully confirmatory test and requires further validation through clinical trials and broader assessments.

Data Limitations:

The study uses just one dataset of MRI images, which might not represent all types of brain tumors. Different machines, patient backgrounds, and scan conditions weren't included, making it hard to apply the model in different settings.

The dataset has an uneven number of tumor types (glioma, meningioma, no tumor), which could cause the model to be biased toward the more common tumor types.

Lack of Multi-modal Data: The model only uses MRI scans, but adding other types of data like CT scans, genetic information, and patient history could make the model more accurate and flexible for different situations.

Overfitting: Some models, like ResNet152V2, seem to "memorize" the data rather than learning to generalize. This could lead to poor performance on new data. Using more varied data and techniques to prevent overfitting could help.

Model Interpretability: Deep learning models are often seen as "black-boxes," meaning we don’t fully understand how they make decisions. This is a problem for doctors who need to trust the model’s results. Using tools to visualize how the model makes predictions could make it more understandable and reliable for clinical use.

Real-time Application: The model hasn't been tested to work in real-time, which is important for doctors in emergency situations. More testing in real-life settings is needed.

Regulatory and Ethical Concerns: The model hasn't been checked for compliance with medical regulations, like FDA (Food and Drug Administration) approval, or for privacy and security concerns related to patient data. It would need to meet strict rules before being used in hospitals.

Future work should incorporate multi-modal data (e.g., CT scans, patient history) and explore alternative diagnostic methods to enhance accuracy. Data augmentation can help balance the dataset, and clinical trials are needed to confirm the model's effectiveness in real-world settings. Developing user-friendly systems for healthcare integration will further support its adoption in clinical workflows.

Multi-modal Data Integration: Future research should combine different types of data, like CT scans, genetic information, and patient medical history, to make models more accurate and reliable for detecting and predicting tumors.
Data Augmentation and Balancing: To fix problems with unbalanced data and overfitting, techniques like rotating, flipping, and resizing images can be used. Also, methods like

SMOTE can balance the data so the model doesn't favor more common tumor types.

Explainable AI (XAI): Future research should emphasize improving the interpretability of deep learning models. Tools like Grad-CAM can be utilized to visualize which regions of MRI scans influence the model's decisions, thereby enhancing transparency and building clinicians' trust in AI outputs.

Cross-Dataset Validation: Models should undergo validation using data collected from diverse hospitals and imaging devices to ensure their generalizability across varied clinical environments. This approach can involve leveraging publicly available datasets or collaborating with multiple medical institutions.

Real-Time and Scalable Implementation: Future efforts should aim to make the model faster and able to handle large amounts of data in real-time. Solutions should also work in different healthcare systems, whether centralized or decentralized.

Regulatory Compliance and Clinical Trials: Research should address compliance with healthcare regulations, such as obtaining FDA approvals and safeguarding patient data privacy under HIPAA guidelines. Conducting clinical trials will be crucial to validate the model's effectiveness and safety in real-world healthcare applications.

Integration into Clinical Workflows: Future studies should aim to create user-friendly systems that allow seamless incorporation of AI-driven diagnostic support into existing healthcare operations, thereby assisting doctors in interpreting results and improving clinical decision-making.

8 References:

1. C. Watson, M. Kirkcaldie, G. Paxinos, The Brain: An Introduction to Functional Neuroanatomy (Academic Press, New York, 2010).

2. J. Nolte, The Human Brain: An Introduction to Its Functional Anatomy (Mosby Elsevier, Philadelphia, 2009).

3. D.R. Johnson, J.B. Guerin, C. Giannini, J.M. Morris, L.J. Eckel, T.J. Kaufmann, 2016 updates to the WHO brain tumor classification system: What the radiologist needs to know, Radiographics 37, 2164 (2017). https://doi.org/10.1148/rg.2017170033

4. E. Wright, E.K. Amankwah, S.P. Winesett, G.F. Tuite, G. Jallo, C. Carey, et al., Incidentally found brain tumors in the pediatric population: A case series and proposed treatment algorithm, J. Neurooncol. 141, 355 (2019). https://doi.org/10.1007/s11060-018-03021-5

5. M.W. Dubin, How the Brain Works (Wiley, New York, 2013).

6. M.A. Nuñez, J.C.F. Miranda, E. de Oliveira, P.A. Rubino, S. Voscoboinik, R. Recalde, Brain stem anatomy and surgical approaches, in Comprehensive Overview of Modern Surgical Approaches to Intrinsic Brain Tumors (Elsevier, Amsterdam, 2019) p. 53. https://doi.org/10.1016/B978-0-323-58126-1.00006-6

7. C. Kurian, J. Joseph, Automated brain tumor detection and classification using deep

learning techniques, Eur. Phys. J. Plus 137, 350 (2022). https://doi.org/10.1140/epjp/s13360-022-02588-3

8. M. Hasan, S. Rahman, A hybrid deep learning model for brain tumor classification, Eur. Phys. J. Spec. Top. 232, 1189 (2023). https://doi.org/10.1140/epjs/s11734-023-00792-5

9. P. Singh, R. Sharma, Feature extraction and selection methods for brain MRI analysis: a survey, Eur. Phys. J. Plus 137, 700 (2022). https://doi.org/10.1140/epjp/s13360-022-03021-6

10. S. Li, Y. Huang, Deep convolutional neural networks for brain tumor detection: a comprehensive study, Eur. Phys. J. E 45, 63 (2022). https://doi.org/10.1140/epje/s10189-022-00132-7

11. N.V. Shree, T.N.R. Kumar, Identification and classification of brain tumor MRI images with feature extraction using DWT and probabilistic neural network, Brain Inform. 5, 23 (2018). https://doi.org/10.1007/s40708-017-0075-5

12. Research Dataset 1, Google Drive (2024). Available at: https://drive.google.com/drive/folders/1-CY4dWPOnPVVPHmJPQNTYHz4J9Dn0vY_?usp=drive_link
13. Research Dataset 2, Google Drive (2024). Available at: https://drive.google.com/drive/folders/1D9GRT5WDTltfPj1RDKeJlkWsA6exRP?usp=drive_link