Stochastic Modelling and Computational Sciences

ENHANCED MRI-BASED BRAIN TUMOR DETECTION VIA TUNED ALEXNET- CNN

S. M. Zakariya[1,*] and Mohammad Sarosh Umar[2]
[1]Electrical Engineering Section, University Polytechnic, Aligarh Muslim University, Aligarh, India
[2]Department of Computer Engineering, ZHCET, Aligarh Muslim University, Aligarh, India
[1]smzakariya.ubp@amu.ac.in

ABSTRACT
Early and accurate detection of brain tumors is crucial for better patient outcomes, but traditional methods that rely on manual image analysis are often time-consuming and prone to errors. This study presents a deep learning-based approach using a hyperparameter-optimized AlexNet Convolutional Neural Network (CNN) model to classify brain tumors from MRI scans. A diverse dataset comprising 7023 MRI images across four classes—glioma, meningioma, pituitary, and no tumor—was employed. The model was trained with and without data augmentation using varied epochs (10, 15, 20), a batch size of 32, and the Adamax optimizer with a learning rate of 0.001. The AlexNet architecture incorporated ReLU and softmax activation functions for effective feature extraction and classification. Experimental results demonstrated high classification performance, achieving a peak training accuracy of 99.97%, test accuracy of 96.15%, and AUC values exceeding 99.7% across all classes. Evaluation metrics including precision, recall, F1-score, and ROC analysis confirm the robustness of the model. These findings highlight the potential of a fine-tuned AlexNet model in supporting automated, accurate, and efficient brain tumor diagnosis using MRI data.

Keywords: *Brain Tumor, AlexNet-CNN, MRI Classification, Hyperparameter Tuning, Deep Learning*

1. INTRODUCTION
Accurate and timely detection of brain tumors remains crucial in combating neurological diseases. Magnetic Resonance Imaging (MRI) is a leading technique for obtaining detailed anatomical insights [1]. Deep learning has significantly advanced medical image analysis, enhancing tumor detection and classification precision [2]. Gliomas, the most common primary brain tumors, and pituitary tumors, due to their critical location, require accurate identification for effective treatment [3] and [4].

Deep learning, a subset of AI, excels at detecting complex patterns in MRI data, offering improved diagnostic accuracy [5]. This study focuses on applying deep learning, particularly convolutional neural networks (CNNs), to detect and classify gliomas, pituitary tumors, and healthy brain tissue using MRI scans [6].

Neurological conditions such as stroke, hemorrhage, multiple sclerosis, and tumors pose major health challenges. Early diagnosis is key to effective treatment. MRI remains the gold standard for brain tumor detection [7] and [8].

The paper is organized as follows: Section 2 provides an overview of deep learning methods used for brain tumor analysis. Section 3 describes the dataset, preprocessing steps, and model development process. Section 4 highlights the experimental results, while Section 5 wraps up the study with key insights, limitations, and suggestions for future research.

2. RELATED WORKS
Early and accurate detection of brain tumors is crucial for improving survival and quality of life. Traditional diagnostic methods, relying on manual image interpretation, are often slow and error-prone, prompting the use of machine learning—especially CNNs—to improve detection accuracy [9]. AlexNet's breakthrough performance in the ILSVRC demonstrated deep learning's potential in image analysis [10], leading to its application in medical imaging. Studies show CNNs can automatically learn hierarchical features from MRI scans, improving tumor detection and segmentation [11] and [12]. CNNs have shown high accuracy in glioma segmentation [13] and in differentiating various tumor types [14]. AlexNet has been adapted for brain tumor detection through fine-tuning, transfer learning, and specialized preprocessing [15]. Ensemble models combining AlexNet with VGGNet,

ResNet, and InceptionNet further boost detection performance [16]. However, challenges such as class imbalance, limited labeled data, and medical image variability remain [27]. Addressing these requires data augmentation, regularization, and multimodal imaging [18]. Overall, CNNs like AlexNet mark a major step forward in brain tumor detection [19], and this paper explores their application, strengths, and limitations.

3. METHODOLOGY IMPLIED

Figure 1 shows the major steps involved in this work.

Figure 1: Methodology of the Work

3.1 Brain Tumor MRI Dataset: The MRI dataset used in this study was obtained from Kaggle and combines three different sources: Figshare, the SARTAJ dataset, and Br35H. In total, the dataset includes 7,023 human brain MRI images categorized into four classes: glioma, meningioma, pituitary tumor, and no tumor. Notably, all 'no tumor' images were taken specifically from the Br35H dataset [20]. The dataset is organized into two main folders—Training and Testing—each containing four subfolders named after the respective classes. Figure 2 displays sample MRI images from the dataset, while Table 1 provides detailed descriptions for each class.

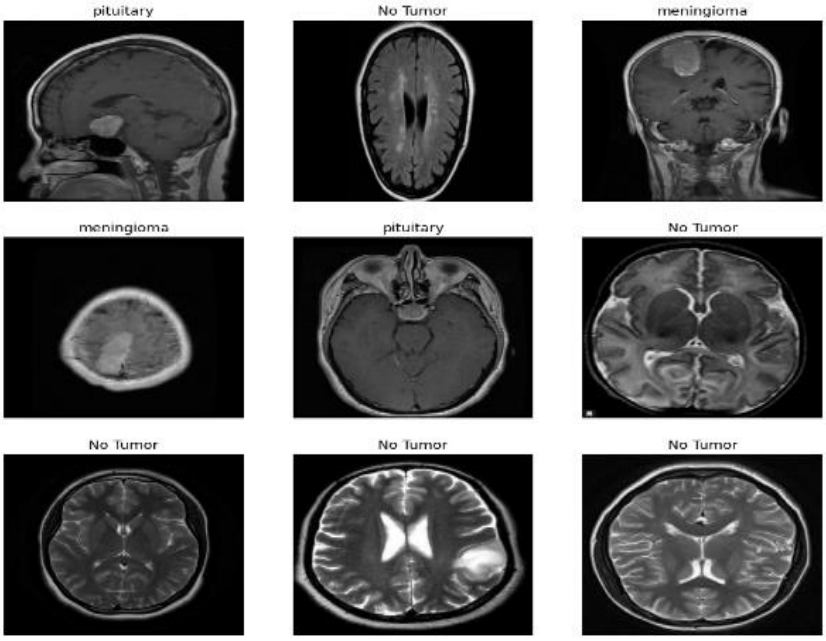


Figure 2: Brain Tumor MRI Dataset Samples

Table 1: Description overview of Brain Tumor MRI dataset

S. No.	Sub-Class	No. of Samples	Description
1	Glioma	1,621	Affected
2	Meningioma	1,757	Affected
3	Pituitary	1,6451	Affected
4	No Tumor	2,000	Normal (Unaffected)

3.2 Data Preparation and Model Training: To ensure consistency and improve processing efficiency, all images were resized to 240×240×3 pixels and converted to grayscale. The dataset was then divided into training (80%), validation (10%), and testing (10%) sets. AlexNet, a popular CNN architecture, was trained across multiple epochs with early stopping applied to prevent overfitting. Throughout training, the model's performance was monitored on the validation set using key metrics like accuracy, precision, recall, and AUC-ROC to evaluate its generalization ability and guide further optimization.

3.3 Performance Evaluation: After completing the training and validation phases, the final performance of the AlexNet model was evaluated using a separate testing dataset. To thoroughly assess its effectiveness in classifying different brain tumor subtypes from MRI images, key metrics such as loss, accuracy, precision, recall, and AUC-ROC were calculated. The description flowgraph of the work is given in Figure 3.

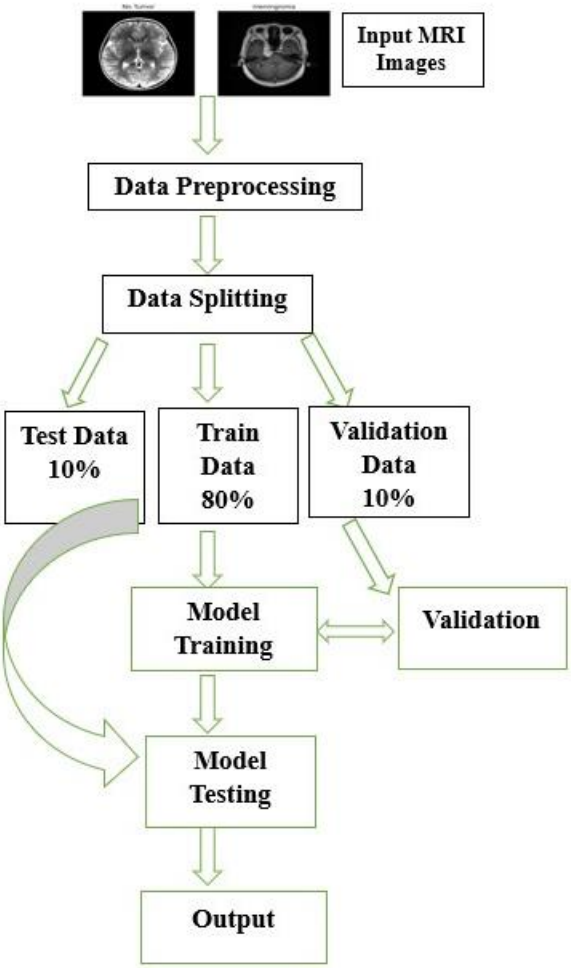


Figure 3: Description flowgraph of the work

3.4 Training Configuration: AlexNet was trained over 10, 15, and 20 epochs using a batch size of 32. The Adamax optimizer, with a learning rate of 0.001, was used to dynamically update the model's weights during training. ReLU activation functions were applied to the input and hidden layers to introduce non-linearity, while the softmax function in the output layer provided class probability predictions. The model relied on 2D convolution operations, a core component in CNN-based image processing tasks [21]. Given an input image I and a filter/kernel K, the 2D convolution operation can be represented mathematically as follows:

$$C(i,j) = (I * K)(i,j) = \sum_m \sum_n I(m,n).K(i-m, j-n) \quad (1)$$

Where: C(i, j) is the value at position (i, j) in the output feature map. (m, n) is the pixel value of the input image at position (m, n). ($i - m$, $j - n$) is the value of the filter/kernel at position ($i - m$, $j - n$). The double summation is performed over all valid positions of the filter/kernel in the input image. '$*$' denotes the convolution operation.

ReLU (Rectified Linear Unit) and Softmax are commonly used activation functions in neural networks, including convolutional neural networks (CNNs) for various tasks like classification.

ReLU is a simple non-linear activation function defined as:

$$ReLU\,(x) = max(0, x) \quad (2)$$

In multi-class classification tasks, the softmax function is used in the output layer of a neural network to convert raw output scores (logits) into probabilities. This helps the model express its confidence in each class prediction. Given an input vector z with K elements (where K is the number of classes), the Softmax function is defined as:

$$Softmax\,(z)_i = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}} \quad (3)$$

Where: The ith element of the output probability vector is Softmax(z)$_i$. Euler's number is e (approximately 2.71828). The ith element of the input vector is z_i. The denominator in the softmax function adds up the exponentials of all the input values, ensuring that the final output values represent probabilities that sum to 1. This makes softmax ideal for multi-class classification problems, where the model needs to assign a probability to each class.

3.5 AlexNet Architecture Layers: The AlexNet model consists of one input layer, five convolutional layers, three max-pooling layers, two fully connected hidden layers, and a final output layer [22]. The overall architecture of AlexNet is illustrated in Figure 4.

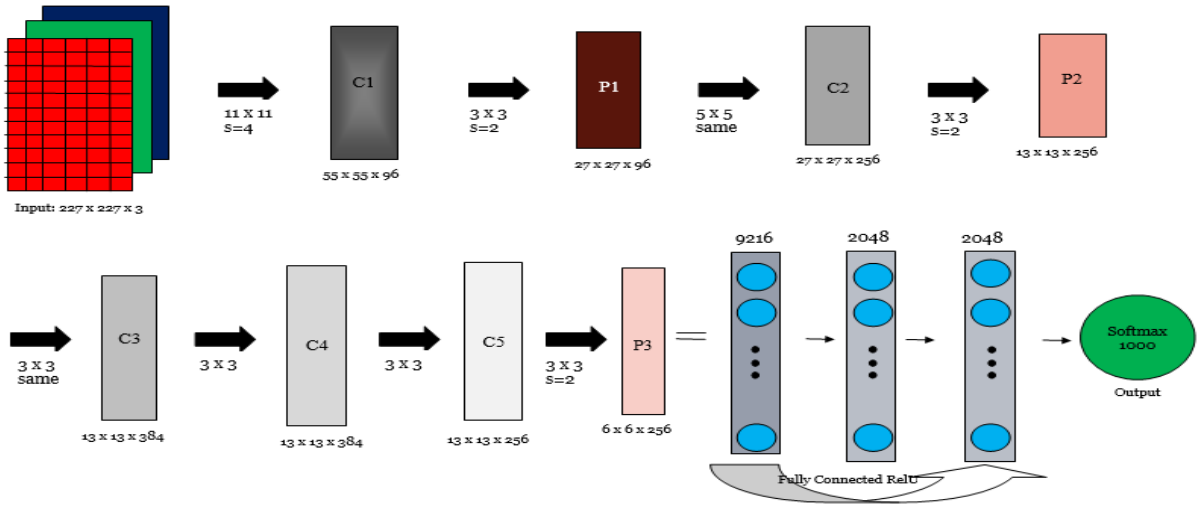


Figure 4: Architecture of AlexNet Model: C1, C2, C3, C4 and C5 are the Convolutional Layers. P1, P2, and P3 are the Maximum Pooling. 's' indicates strides

This architecture follows a pattern of alternating convolutional layers—which extract important features—and max-pooling layers that reduce spatial dimensions. These are followed by fully connected layers that handle the classification task. The final softmax layer outputs the probabilities for each class.

3.6 Data augmentation: Data augmentation enhances model robustness and generalization, especially when training data is limited. Table 2 outlines the applied augmentation techniques.

Table 2: Data Augmentation Parameters Used During Model Training

S. No.	Parameter	Value
1	Horizontal flip	True
2	Vertical flip	True
3	Height shift range	0.1
4	Width shift range	0.1
5	Rotation	90^0
6	Shear range	0.2
7	Zoom range	0.2

3.7 Hyperparameter (Fine-Tuning) used in AlexNet Model: Fine-tuning AlexNet for classifying brain tumors involves the following steps. The process generally includes preparing dataset, modifying the AlexNet architecture to suit specific problem, and then training the model.

Figure 5 shows AlexNet model with fine tuning and hyperparameters. The hyperparameter optimization strategy used in training is summarized in Table 3.

Table 3: Optimization of fine-tuned (hyper-parameters) used in this work

Model	Hyperparameters	Value
CNN-AlexNet	Optimizer function	Adamax
	Loss function	categorical_crossentropy
	Metrics	Train and Test Loss, AUC, Accuracy
	Epochs	10, 15 and 20
	Batch size	32
	Learning rate	0.001

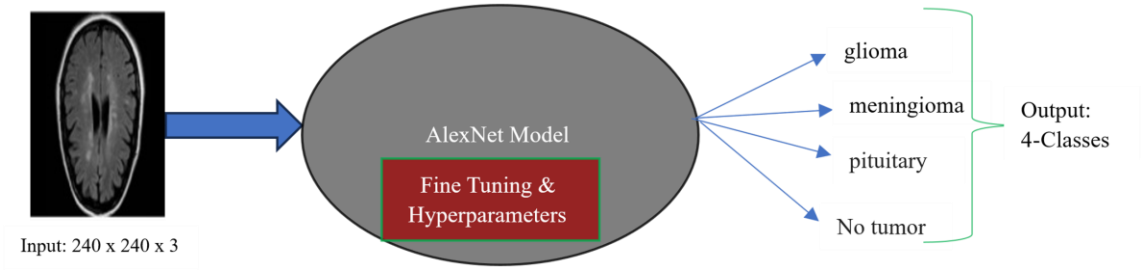


Figure 5: AlexNet model for the classification of brain tumor disease

3.8 Early Stopping and Learning Rate Reduction: To prevent overfitting and improve model convergence, two strategies were used: early stopping and learning rate reduction. Early stopping halted training if validation accuracy didn't improve after 3 epochs, while the learning rate was reduced by a factor of 0.1 if performance plateaued for 2 epochs. These two callbacks—EarlyStopping and ReduceLROnPlateau—helped optimize training by stopping at the right time and fine-tuning the learning pace.

3.9 Performance Metrics: After training and validating the model, a full set of performance metrics was calculated to assess how well AlexNet performed and how effectively it could generalize to new data [33]. One of the key metrics was accuracy, which measures the percentage of correctly classified samples out of the total.

$$Accuracy = \frac{TruePositive + TrueNegative}{TruePositive + TrueNegative + FalsePositive + FalseNegative} \quad (4)$$

Precision: This metric reflects how many of the model's positive predictions were actually correct. It helps measure the model's ability to minimize false positives.

$$Precision = \frac{TruePositive}{TruePositive + FalsePositive} \quad (5)$$

Recall: Recall shows how well the model captures actual positive cases. It's calculated as the proportion of true positives out of all actual positive examples, highlighting the model's ability to detect relevant instances.

$$Recall = \frac{TruePositive}{TruePositive + FalseNegative} \quad (6)$$

F1 – Score: The F1-score is the harmonic mean of precision and recall. It provides a balanced measure that considers both false positives and false negatives, offering a single, comprehensive performance score.

$$F1 - Score = 2x \frac{PrecisionxRecall}{Precision + Recall} \quad (7)$$

Loss: The measure of discrepancy between predicted and actual class labels, computed using appropriate loss functions such as categorical cross-entropy loss L. In a Convolutional Neural Network (CNN), the loss function is a measure of how well the model's predictions match the actual labels in the training data [23]. L is calculated as follows:

$$Loss = -\sum_{i=1}^{c} y_i * log[\hat{y}_i] \quad (8)$$

Where: the number of classes is c, the actual probability that belongs to class i is y_i, and the predicted probability that belongs to class i is $\hat{y}_i$.

AUC-ROC: This metric summarizes how well the model distinguishes between classes. It's based on the ROC curve, which plots the trade-off between sensitivity (true positive rate) and the false positive rate. A high AUC indicates strong performance [24].

True Positive Rate (TPR): The proportion of actual positives correctly identified—e.g., how many tumors the model detects accurately.

False Positive Rate (FPR): The proportion of negatives wrongly classified as positives—e.g., non-tumor cases labeled as tumors.

ROC Curve: Visualizes the balance between TPR and FPR. The closer the curve is to the top-left corner, the better the model's performance.

$$TPR = \frac{TruePositive}{(TruePositive + FalseNegative)} \quad (9)$$

$$FPR = \frac{FalsePositive}{(FalsePositive + TrueNegative)} \quad (10)$$

4. EXPERIMENTAL RESULTS AND PERFORMANCE EVALUATION:

The experimental results clearly show AlexNet's strong ability to classify brain tumors from MRI images. It achieved a training accuracy of 99.97% and a training AUC of 100%, indicating excellent learning of complex patterns. Tables 4 and 5 summarize the average training and testing loss, accuracy, and AUC across different epochs. The model also performed well on unseen data, with a test accuracy of 96.15% and a test AUC of 99.75%.

Tables 6 and 7 report precision, recall, and F1 scores at epochs 10, 20, and 30, showing consistent performance in identifying tumors. Figures 6 and 7 visually illustrate the model's results across four tumor types, both with and without data augmentation. These metrics and visualizations reflect the model's ability to generalize and detect subtle differences in tumor features over time.

Table 4: Train and Test Loss, Accuracy, AUC at different Epochs, batch size = 32 Without Augmentation

Epoch = 10	Epoch = 15	Epoch = 20
Train Loss: 0.0338	Train Loss: 0.0070	Train Loss: 0.0067
Train Accuracy: 99.27%	Train Accuracy: 99.96%	Train Accuracy: 99.89%
Train AUC: 99.97%	Train AUC: 100.00%	Train AUC: 100.00%
----------	----------	----------
Test Loss: 0.1586	Test Loss: 0.1116	Test Loss: 0.1051
Test Accuracy: 94.72%	Test Accuracy: 97.72%	Test Accuracy: 97.72%
Test AUC: 99.34%	Test AUC: 99.47%	Test AUC: 99.65%%

Table 5: Train and Test Loss, Accuracy, AUC at different Epochs, batch size = 32 With Augmentation

Epoch = 10	Epoch = 15	Epoch = 20
Train Loss: 0.1266	Train Loss: 0.0215	Train Loss: 0.0022
Train Accuracy: 95.03%	Train Accuracy: 99.50%	Train Accuracy: 99.97%
Train AUC: 99.67%	Train AUC: 99.99%	Train AUC: 100.00%
----------	----------	----------
Test Loss: 0.2772	Test Loss: 0.1622	Test Loss: 0.0113
Test Accuracy: 89.63%	Test Accuracy: 94.63%	Test Accuracy: 96.15%
Test AUC: 98.70%	Test AUC: 99.35%	Test AUC: 99.75%

Table 6: Class-wise Precision, Recall, F1-Score, and Average Accuracy of AlexNet Model at Different Training Epochs without Data Augmentation

Class	Precision	Recall	F1-Score	Avg. Accuracy
	Epoch = 10			
glioma	0.94	0.94	0.94	0.9472
meningioma	0.92	0.88	0.90	
No Tumor	0.96	0.97	0.97	
pituitary	0.96	0.98	0.97	
	Epoch = 15			
glioma	0.97	0.98	0.97	0.9772
meningioma	0.97	0.95	0.96	
No Tumor	0.99	0.99	0.99	
pituitary	0.98	0.99	0.98	
	Epoch = 20			
glioma	0.98	0.97	0.98	0.9772
meningioma	0.98	0.96	0.97	
No Tumor	0.99	0.99	0.99	
pituitary	0.97	0.98	0.97	

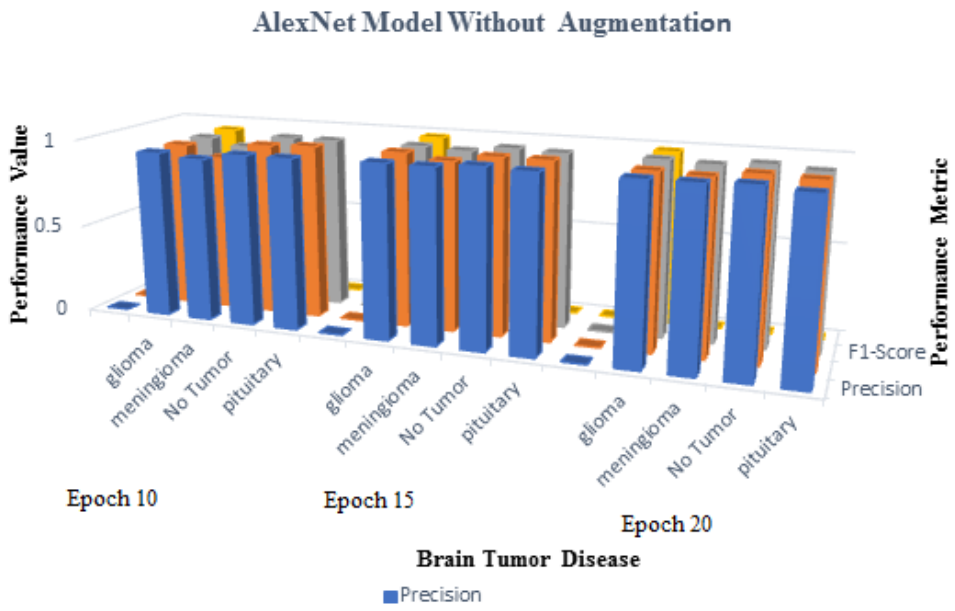


Figure 6: Performance graph of Brain Tumor Disease using Hype-tuned AlexNet Model without Augmentation

Table 7: Class-wise Precision, Recall, F1-Score, and Average Accuracy of AlexNet Model at Different Training Epochs with Augmentation

Class	Precision	Recall	F1-Score	Avg. Accuracy
	Epoch = 10			
glioma	0.93	0.79	0.86	0.8963
meningioma	0.81	0.89	0.84	
No Tumor	0.97	0.96	0.96	
pituitary	0.89	0.95	0.92	
	Epoch = 15			
glioma	0.91	0.94	0.92	0.9463
meningioma	0.92	0.91	0.91	
No Tumor	0.98	0.99	0.99	
pituitary	0.97	0.95	0.96	
	Epoch = 20			
glioma	0.90	0.92	0.91	0.9615
meningioma	0.92	0.89	0.90	
No Tumor	0.98	0.99	0.99	
pituitary	0.96	0.95	0.96	

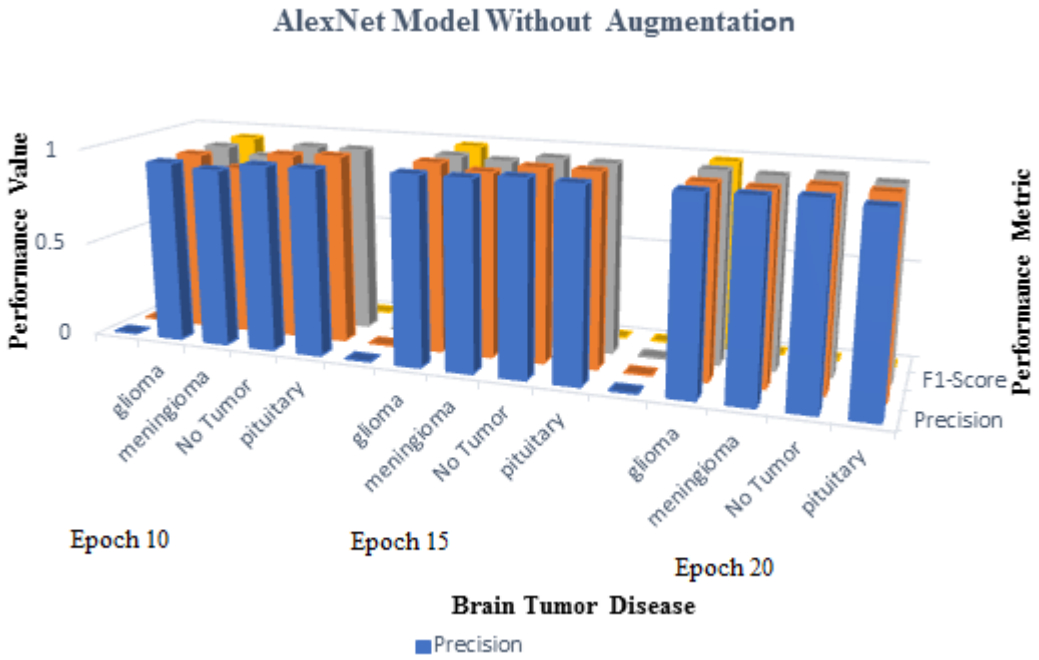


Figure 7: Performance graph of Brain Tumor Disease using Hype-tuned AlexNet Model with Augmentation

Figure 8 shows training and validation metrics of the hyper-tuned AlexNet model over 10 epochs without data augmentation, presented in four plots.

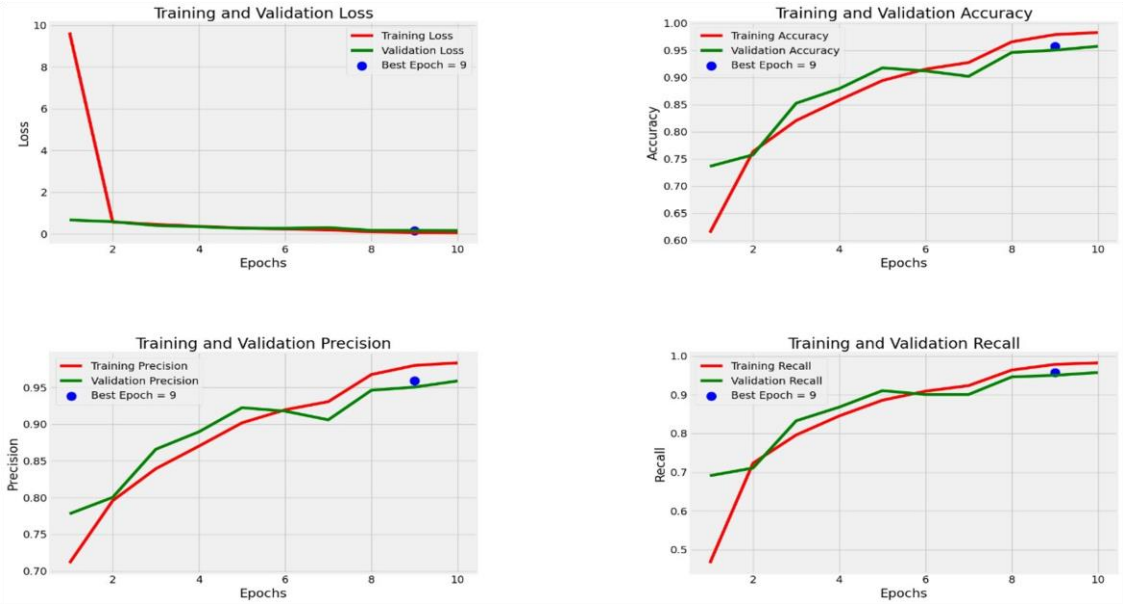


Figure 8: Training progress for the AlexNet model: loss value, accuracy value, precision value and recall value during training and validation process, Epoch = 10, batch size = 32 without Augmentation

The model shows strong performance across all metrics (loss, accuracy, precision, recall), with improvements continuing until epoch 9, after which performance stabilizes. The best epoch for the model across all metrics appears to be epoch 9. At this point, both training and validation results show minimal discrepancy, indicating that the model generalizes well and is not overfitting.

The gap between training and validation metrics is relatively small, which suggests that the model is well-trained without significant overfitting.

Similarly, Figures 9–13 represent four plots showing the training and validation metrics for a machine learning model over 10, 15, and 20 epochs, with and without augmentation, highlighting the best epoch at different points.

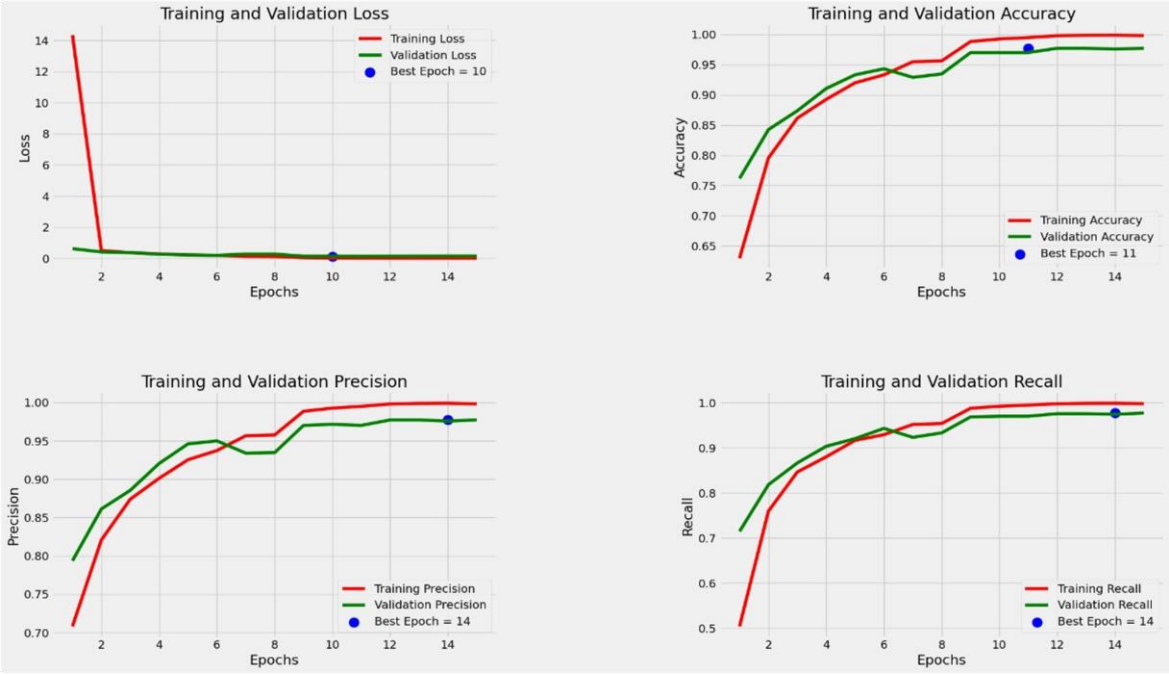


Figure 9: Training progress of the AlexNet model: loss value, accuracy value, precision value and recall value during training and validation process at Epoch = 15, batch size = 32 without Augmentation

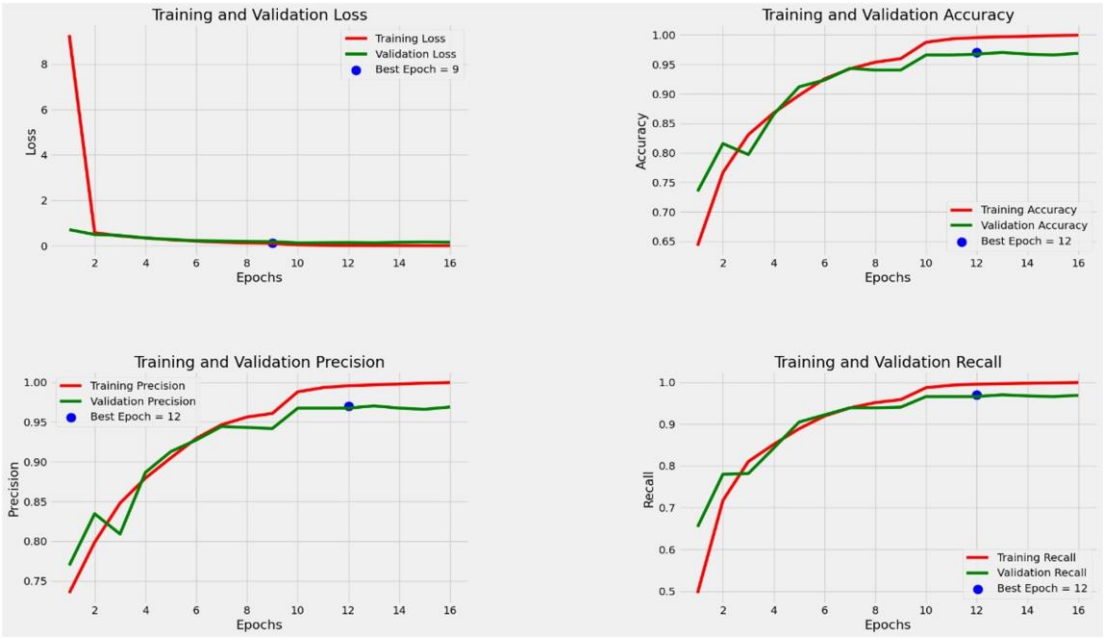


Figure 10: Training progress of the AlexNet model: loss value, accuracy value, precision value and recall value during training and validation process at Epoch = 20, batch size = 32 without Augmentation, Early stopping

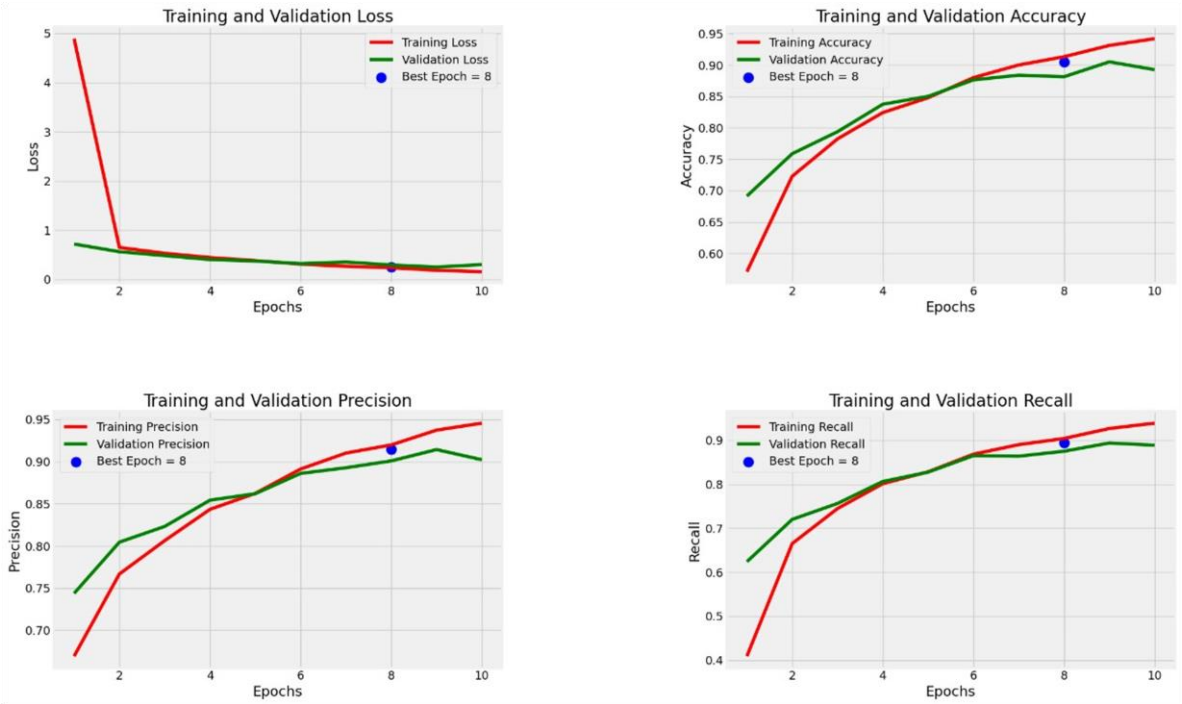


Figure 11: Training progress of the AlexNet model: loss value, accuracy value, precision value and recall value during training and validation process at Epoch = 10, batch size = 32 with Augmentation

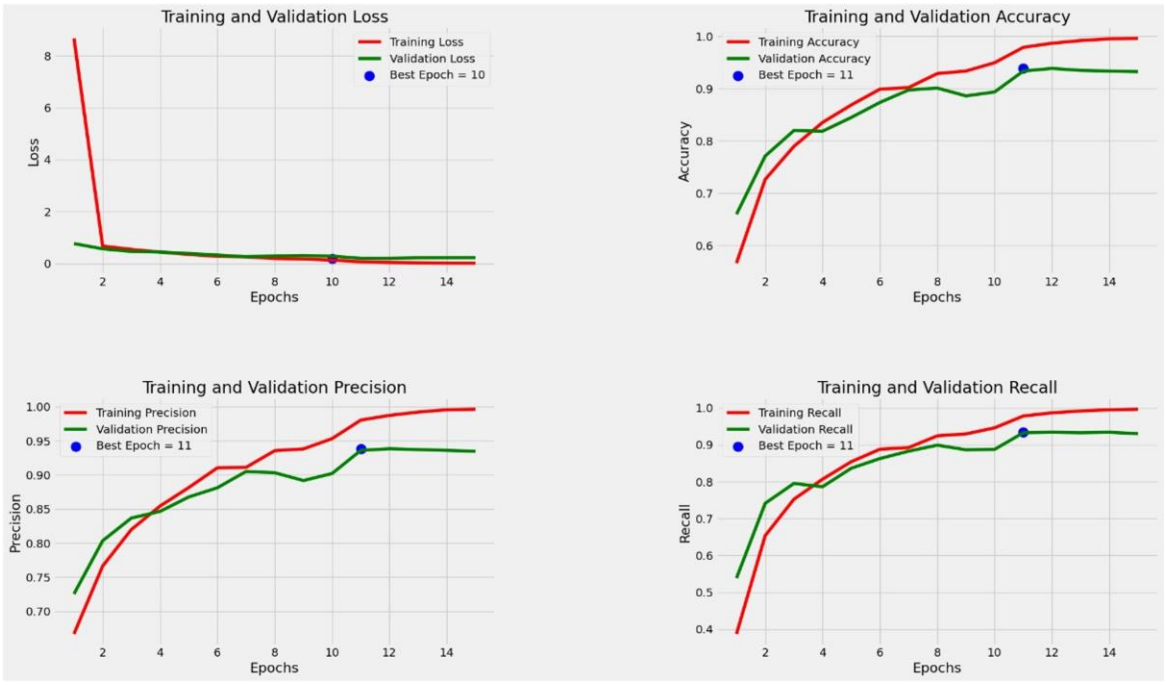


Figure 12: Training progress of the AlexNet model: loss value, accuracy value, precision value and recall value during training and validation process at Epoch = 15, batch size = 32 with Augmentation

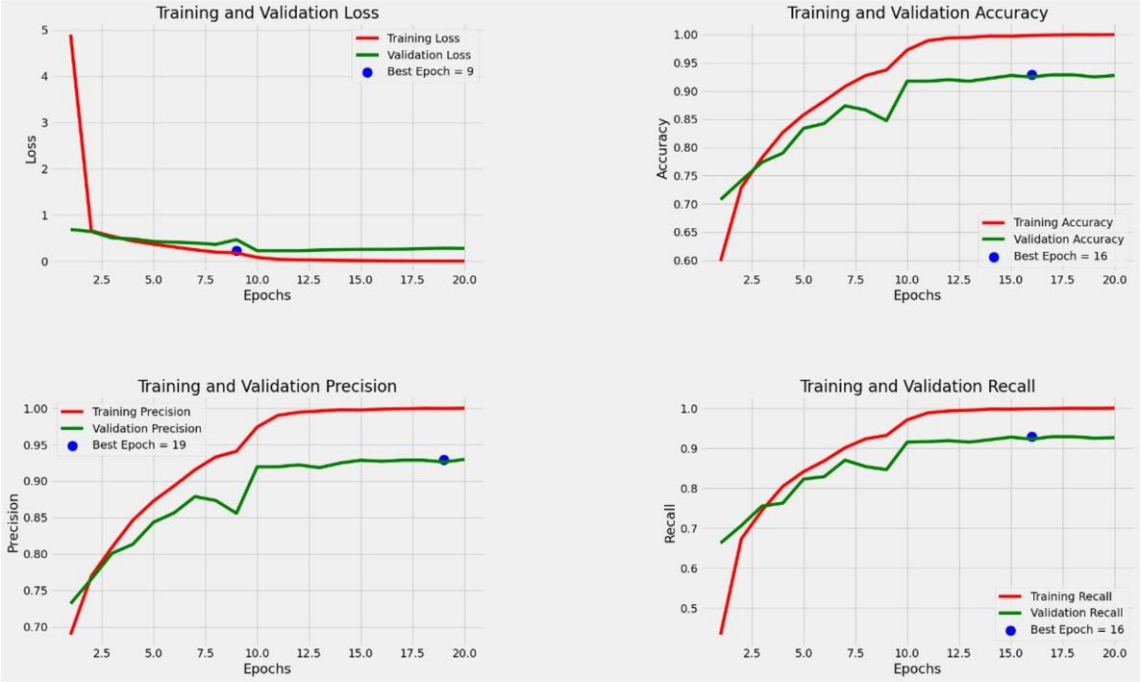


Figure 13: Training progress of the AlexNet model: loss value, accuracy value, precision value and recall value during training and validation process at Epoch = 20, batch size = 32 with Augmentation

A confusion matrix helps assess how accurately the model classifies brain tumors into four categories: glioma, meningioma, pituitary tumor, and no tumor. Figure 14 shows that most predictions align with the true labels, indicating strong performance. The "No Tumor" class was predicted perfectly with 195 correct predictions and no errors, while overall misclassifications were minimal.

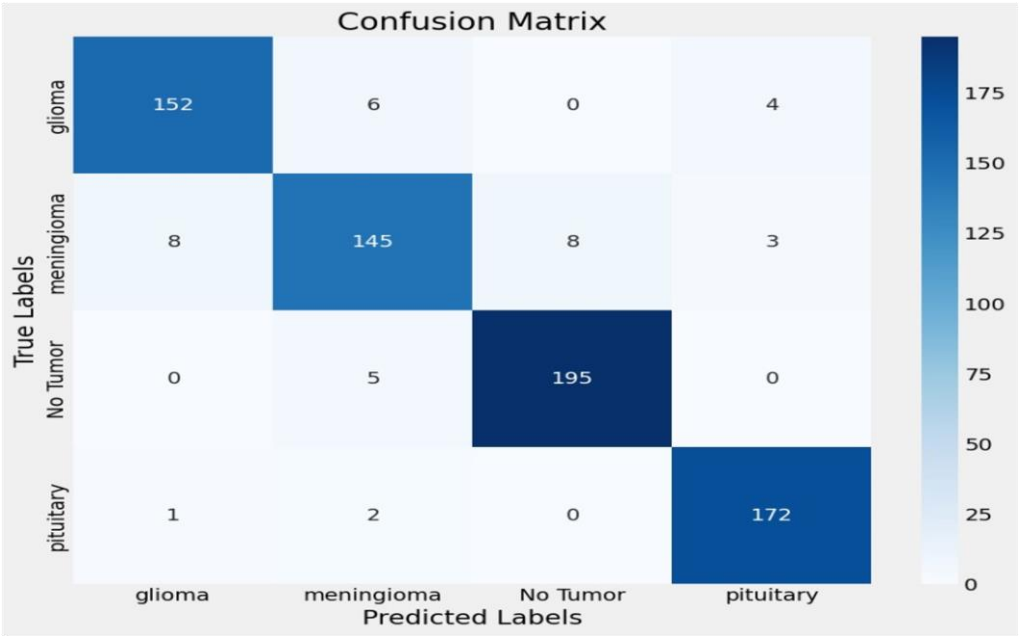


Figure 14: Confusion matrix for AlexNet model Performance at Epoch = 10, Batch Size = 32, without augmentation

Similarly, Figures 15 – 19 represent confusion matrices that evaluate the performance of a machine learning classification model over 10, 15, and 20 epochs, with and without augmentation.

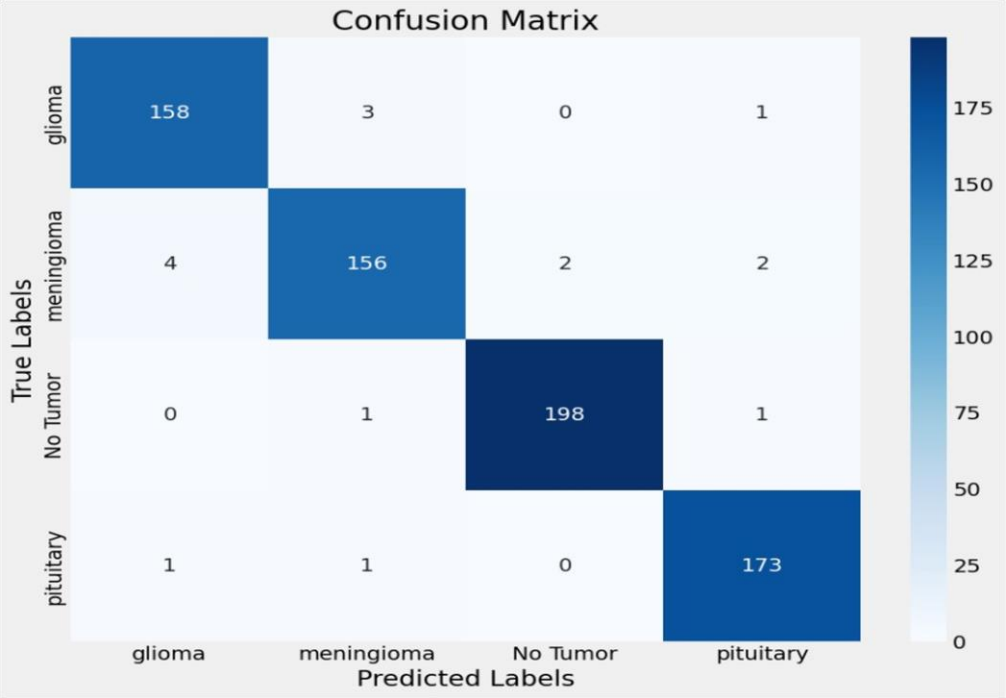


Figure 15: Confusion matrix for AlexNet model Performance at Epoch = 15, Batch Size = 32, without augmentation

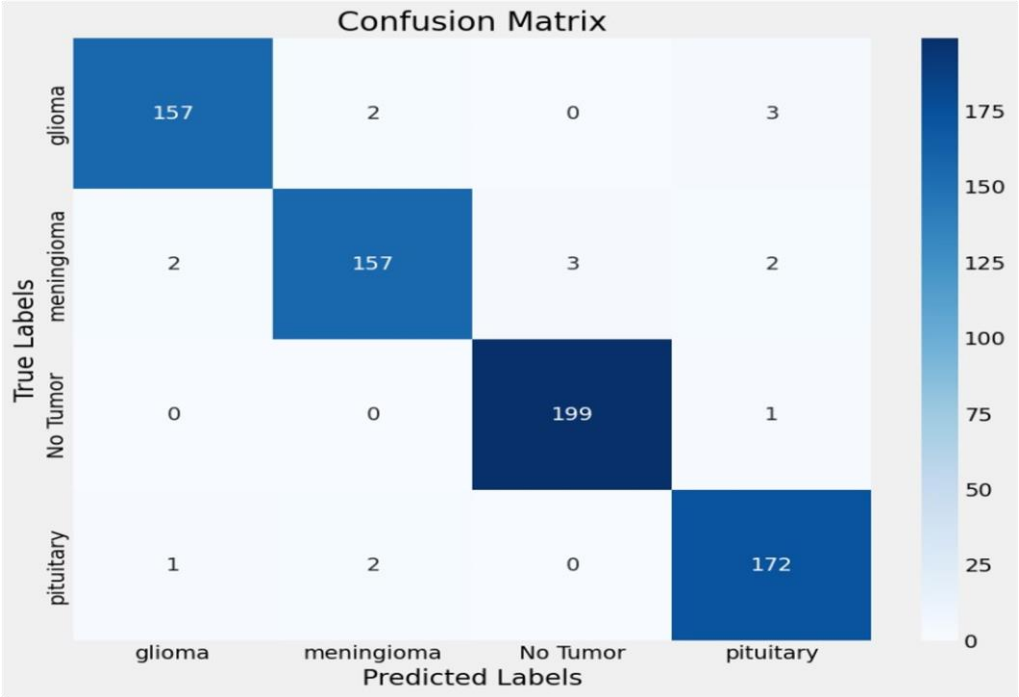


Figure 16: Confusion matrix for AlexNet model Performance at Epoch = 20, Batch Size = 32, without augmentation

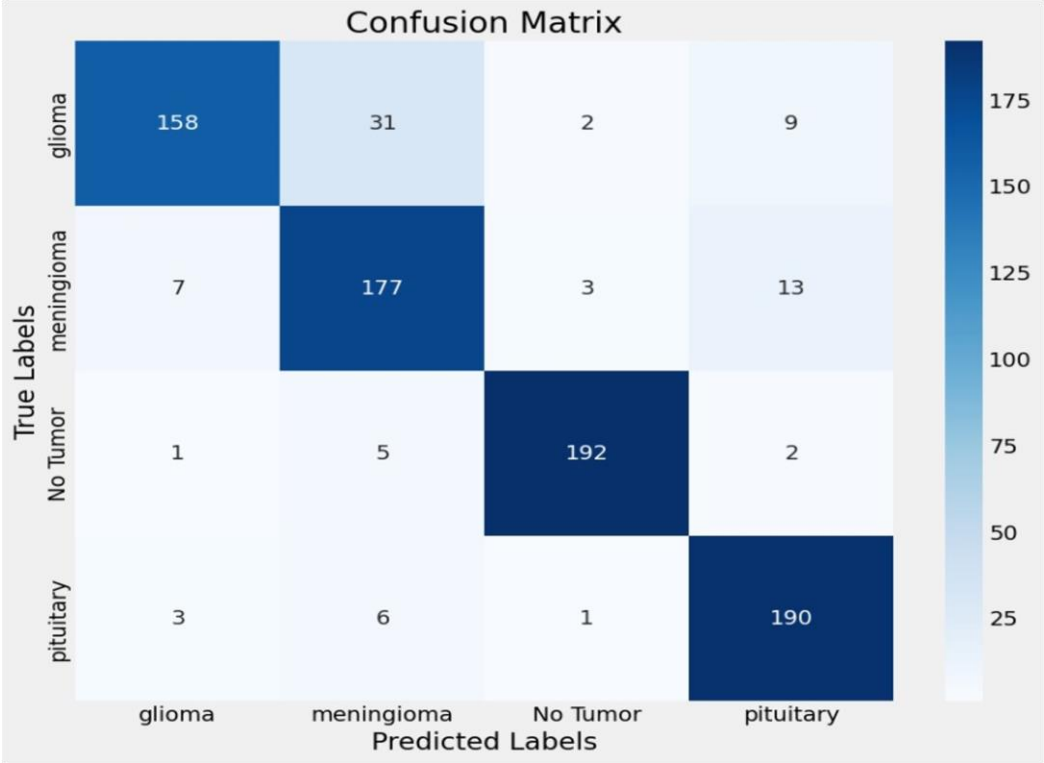


Figure 17: Confusion matrix for AlexNet model Performance at Epoch = 10, Batch Size = 32 with augmentation

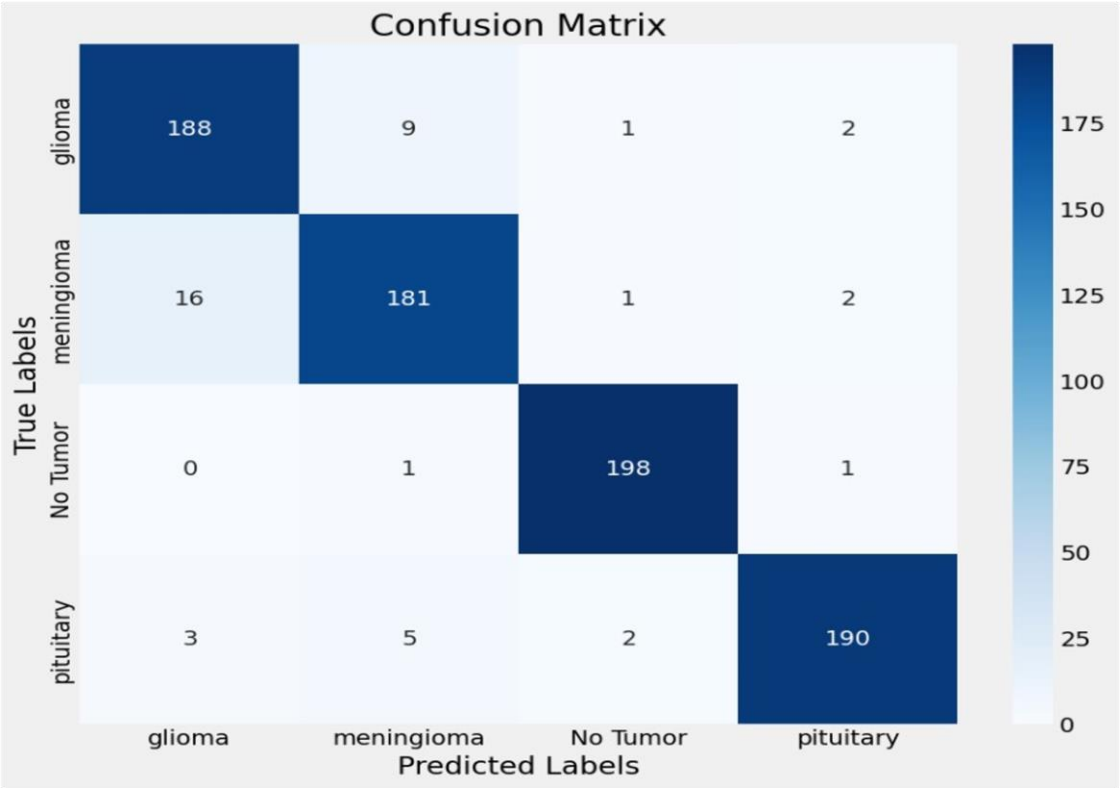


Figure 18: Confusion matrix for AlexNet model Performance at Epoch = 15, Batch Size = 32 with augmentation

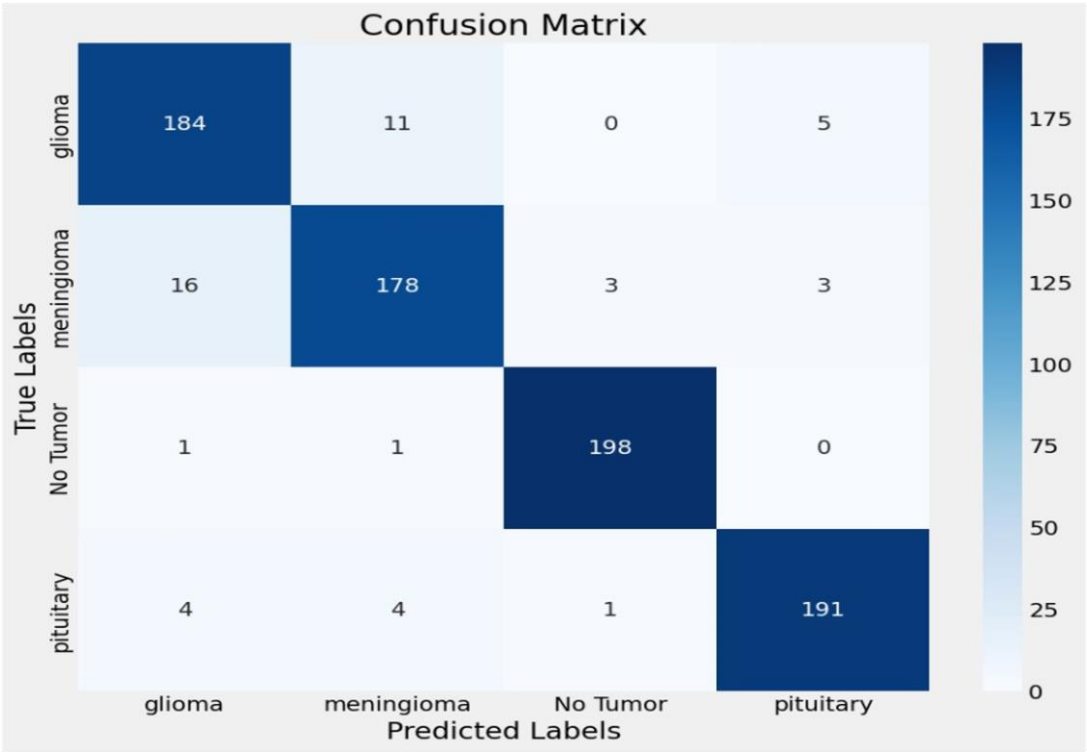


Figure 19: Confusion matrix for AlexNet model Performance at Epoch = 20, Batch Size = 32 with augmentation

Figure 20 shows the ROC curves for the AlexNet model across four classes: glioma, meningioma, pituitary tumor, and no tumor. The model performs exceptionally well, with AUC scores of 1.00 for glioma, no tumor, and pituitary, and 0.99 for meningioma. The curves closely hug the top-left corner, indicating high accuracy and minimal overlap between true and false positives.

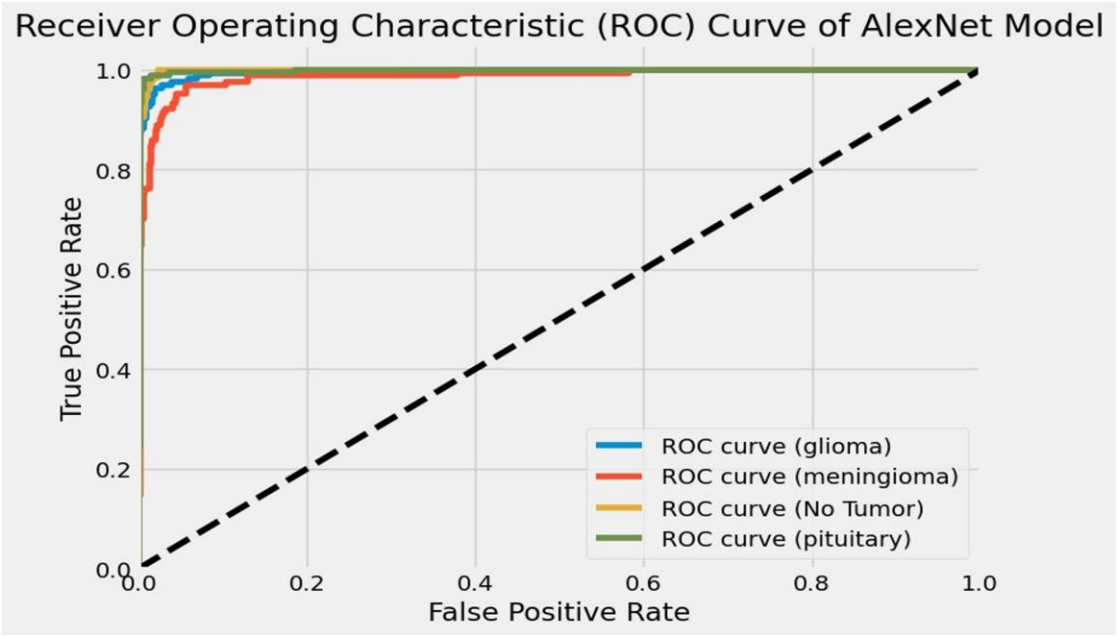


Figure 20: ROC curve for AlexNet model Performance at Epoch = 10, Batch Size = 32, without augmentation

Similarly, Figure 21 – 25 shows the Receiver Operating Characteristic (ROC) curve for the AlexNet model, representing its classification performance across four different categories over 10, 15, and 20 epochs, with and without augmentation.

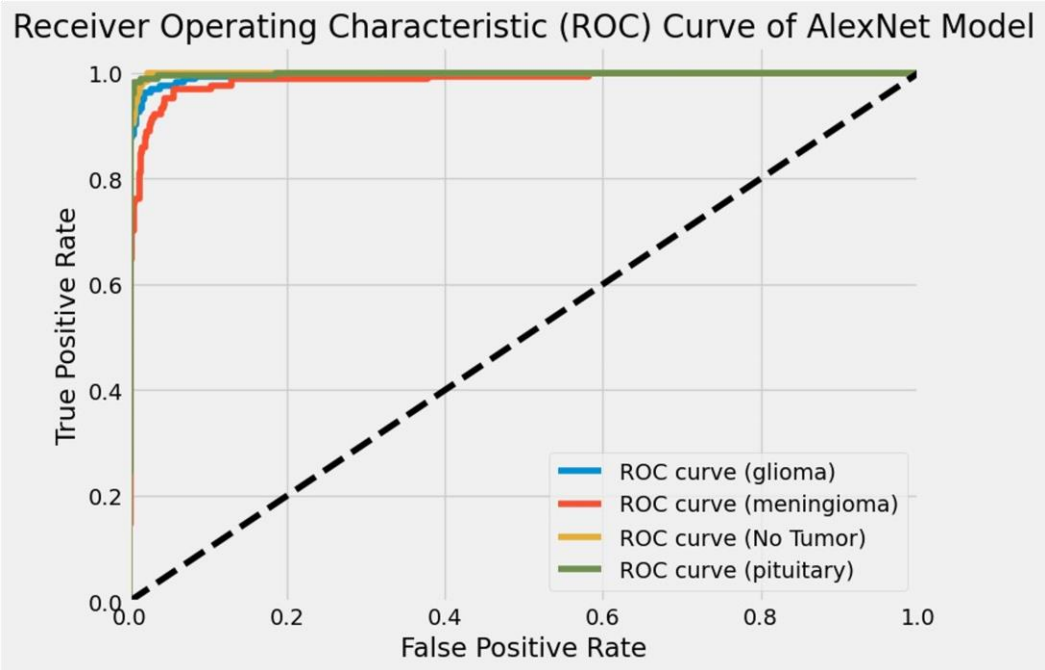


Figure 21: ROC curve for AlexNet model Performance at Epoch = 15, Batch Size = 32, without augmentation

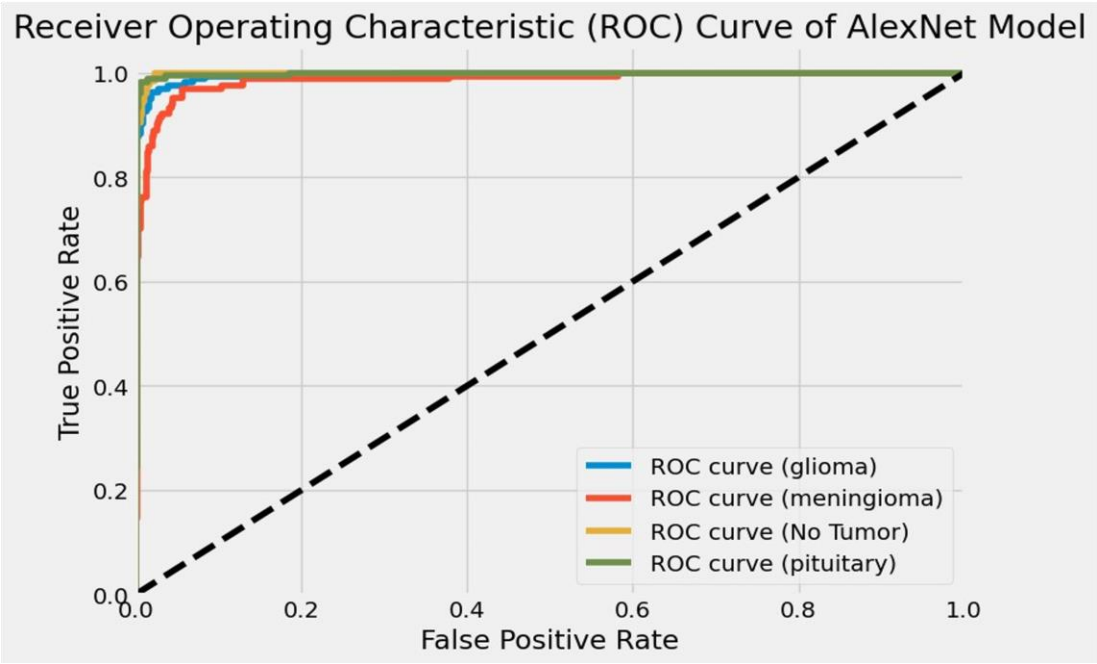


Figure 22: ROC curve for AlexNet model Performance at Epoch = 20, Batch Size = 32, without augmentation

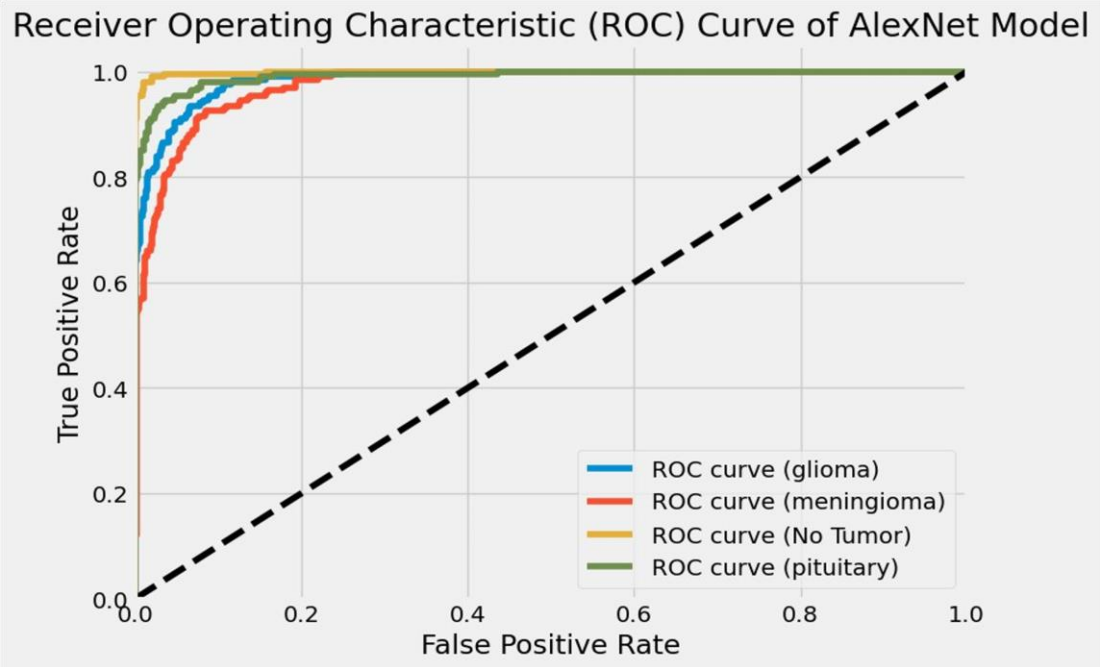


Figure 23: ROC curve for AlexNet model Performance at Epoch = 10, Batch Size = 32 with augmentation

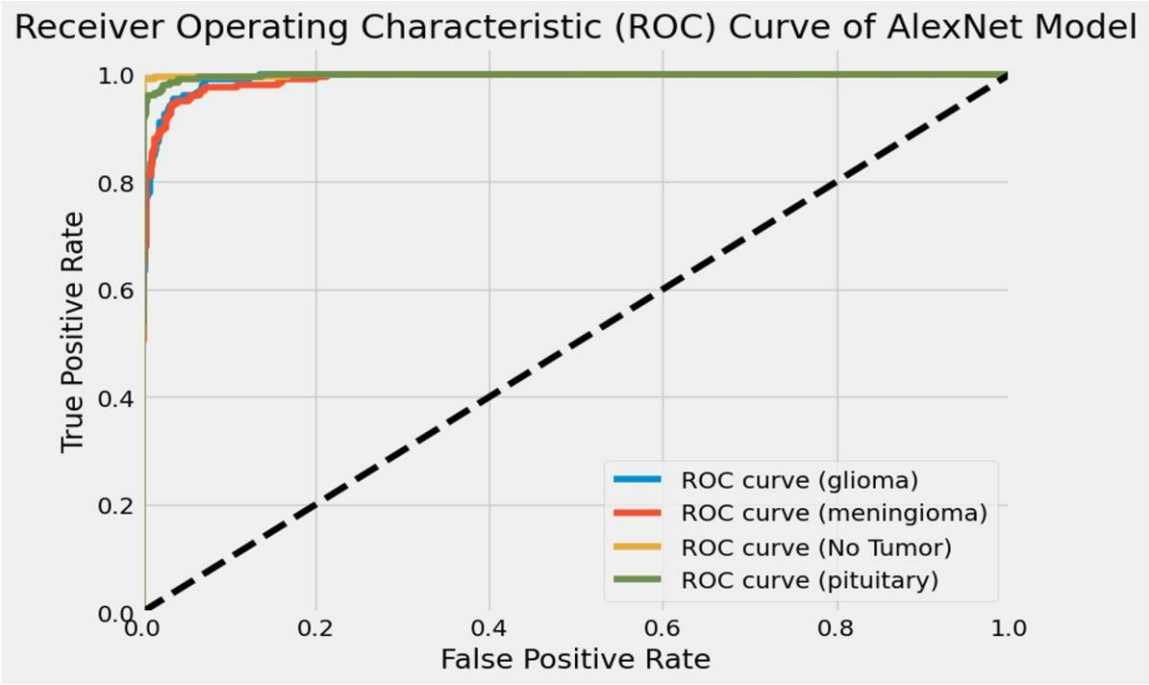


Figure 24: ROC curve for AlexNet model Performance at Epoch = 15, Batch Size = 32 with augmentation

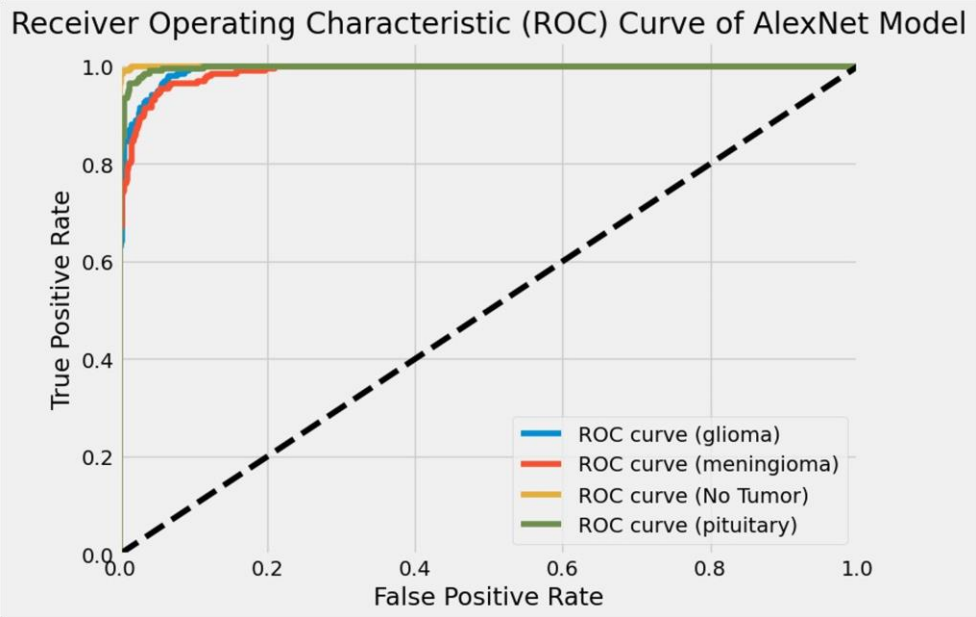
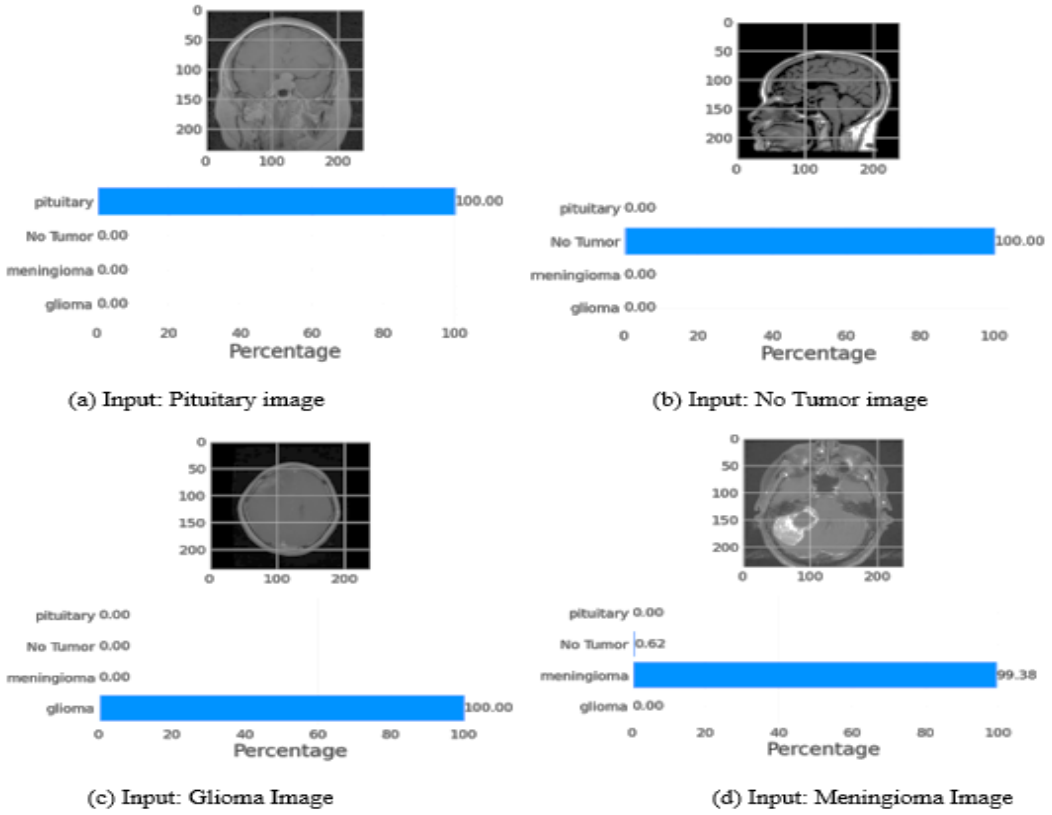


Figure 25: ROC curve for AlexNet model Performance at Epoch = 20, Batch Size = 32 with augmentation

Figure 26 demonstrates a sample of the prediction outcomes from the hyper-tuned AlexNet model. Four input images, one from each category, show an almost 100% probability of the predictions.



(a) Input: Pituitary image

(b) Input: No Tumor image

(c) Input: Glioma Image

(d) Input: Meningioma Image

Figure 26: Prediction results of the AlexNet model at epochs = 20, with augmentation, shown almost 100% probability.

CONCLUSION

This study demonstrates the efficacy of a hyperparameter-tuned AlexNet CNN model in accurately classifying brain tumor types using MRI data. The model was rigorously evaluated with and without data augmentation across multiple epochs, achieving a high training accuracy of 99.97% and a test accuracy of 96.15%, with AUC values approaching 100%. These results underscore the model's robustness, generalization ability, and its capacity to detect subtle imaging variations across tumor classes. The consistent improvement in performance with data augmentation further validates the significance of dataset diversity in medical image analysis. Future research may explore enhancing this framework through transfer learning, model ensembling, and evaluation on larger, multimodal MRI datasets to develop even more generalizable and clinically viable diagnostic tools.

REFERENCES

[1] A. Chattopadhyay and M. Maitra, "MRI-based brain tumour image detection using CNN based deep learning method," *Neuroscience Informatics*, vol. 2, no. 4, p. 100060, Dec. 2022, doi: 10.1016/j.neuri.2022.100060.

[2] Md. S. I. Khan *et al.*, "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.

[3] Md. S. I. Khan *et al.*, "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.

[4] M. Toğaçar, B. Ergen, and Z. Cömert, "BrainMRNet: Brain tumor detection using magnetic resonance images with a novel convolutional neural network model," *Med Hypotheses*, vol. 134, p. 109531, Jan. 2020, doi: 10.1016/j.mehy.2019.109531.

[5] S. Maqsood, R. Damaševičius, and R. Maskeliūnas, "Multi-Modal Brain Tumor Detection Using Deep Neural Network and Multiclass SVM," *Medicina (B Aires)*, vol. 58, no. 8, p. 1090, Aug. 2022, doi: 10.3390/medicina58081090.

[6] 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, doi: 10.1109/ACCESS.2019.2919122.

[7] I. Abd El Kader, G. Xu, Z. Shuai, S. Saminu, I. Javaid, and I. Salim Ahmad, "Differential Deep Convolutional Neural Network Model for Brain Tumor Classification," *Brain Sci*, vol. 11, no. 3, p. 352, Mar. 2021, doi: 10.3390/brainsci11030352.

[8] J. Seetha and and S. Selvakumar Raja, "Brain Tumor Classification Using Convolutional Neural Networks," *Biomedical & Pharmacology Journal*, vol. 11, no. 3, pp. 1457–1461, Sep. 2018.

[9] J. Amin, M. Sharif, A. Haldorai, M. Yasmin, and R. S. Nayak, "Brain tumor detection and classification using machine learning: a comprehensive survey," *Complex & Intelligent Systems*, vol. 8, no. 4, pp. 3161–3183, Aug. 2022, doi: 10.1007/s40747-021-00563-y.

[10] I. Singh, G. Goyal, and A. Chandel, "AlexNet architecture based convolutional neural network for toxic comments classification," *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 9, pp. 7547–7558, Oct. 2022, doi: 10.1016/j.jksuci.2022.06.007.

[11] G. Litjens *et al.*, "A survey on deep learning in medical image analysis," *Med Image Anal*, vol. 42, pp. 60–88, Dec. 2017, doi: 10.1016/j.media.2017.07.005.

[12] R. Ranjbarzadeh, A. Bagherian Kasgari, S. Jafarzadeh Ghoushchi, S. Anari, M. Naseri, and M. Bendechache, "Brain tumor segmentation based on deep learning and an attention mechanism using MRI multi-modalities brain images," *Sci Rep*, vol. 11, no. 1, p. 10930, May 2021, doi: 10.1038/s41598-021-90428-8.

[13] S. Deepak and P. M. Ameer, "Brain tumor classification using deep CNN features via transfer learning," *Comput Biol Med*, vol. 111, p. 103345, Aug. 2019, doi: 10.1016/j.compbiomed.2019.103345.

[14] A. M. Alqudah, "Brain Tumor Classification Using Deep Learning Technique - A Comparison between Cropped, Uncropped, and Segmented Lesion Images with Different Sizes," *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 8, no. 6, pp. 3684–3691, Dec. 2019, doi: 10.30534/ijatcse/2019/155862019.

[15] B. H. Menze *et al.*, "The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)," *IEEE Trans Med Imaging*, vol. 34, no. 10, pp. 1993–2024, Oct. 2015, doi: 10.1109/TMI.2014.2377694.

[16] M. Havaei *et al.*, "Brain tumor segmentation with Deep Neural Networks," *Med Image Anal*, vol. 35, pp. 18–31, Jan. 2017, doi: 10.1016/j.media.2016.05.004.

[17] Z. Liu *et al.*, "Deep learning based brain tumor segmentation: a survey," *Complex & Intelligent Systems*, vol. 9, no. 1, pp. 1001–1026, Feb. 2023, doi: 10.1007/s40747-022-00815-5.

[18] W. Chen, B. Liu, S. Peng, J. Sun, and X. Qiao, "S3D-UNet: Separable 3D U-Net for Brain Tumor Segmentation," 2019, pp. 358–368. doi: 10.1007/978-3-030-11726-9_32.

[19] C. Swarup, K. Udham Singh, A. Kumar, S. Kumar Pandey, N. varshney, and T. Singh, "Brain tumor detection using CNN, AlexNet & GoogLeNet ensembling learning approaches," *Electronic Research Archive*, vol. 31, no. 5, pp. 2900–2924, 2023, doi: 10.3934/era.2023146.

[20] Nickparvar and Msoud, "MRI dataset.," *Kaggle:https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset*, Sep. 2021.

[21] Driss Lamrani, Bouchaib Cherradi, Oussama El Gannour, Mohammed Amine Bouqentar, and Lhoussain Bahatti, "Brain Tumor Detection using MRI Images and Convolutional Neural Network," *(IJACSA) International Journal of Advanced Computer Science and Applications*, vol. 13, no. 7, pp. 452–460, 2022.

[22] Vaibhav Jayaswal, "Performance Metrics: Confusion matrix, Precision, Recall, and F1 Score," *https://towardsdatascience.com/performance-metrics-confusion-matrix-precision-recall-and-f1-score-a8fe076a2262*, Sep. 2020.

[23] Artem Oppermann, "How Loss Functions Work in Neural Networks and Deep learning," *https://builtin.com/machine-learning/loss-functions*, Dec. 2022.

[24] Sarang Narkhede, "Understanding AUC - ROC Curve," *https://towardsdatascience.com/understanding-auc-roc-curve-68b2303cc9c5*, Jun. 2018.