

An Explainable DenseNet121-Based Brain Tumor Classification from MRI Images Using Grad-CAM

¹Indana Naga Satish, ²Chunduri Kiran kumar ³Cherukuri Usha

¹PG Student, Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram, AP, India, nagasatishindana@gmail.com

²Associate Professor, Department of MCA, Koneru Lakshmaiah Education Foundation, Vaddeswaram, AP, India, kirancv117@gmail.com

³Assistant Professor, Department of IT, NRI Institute of Technology, Guntur, AP, India, ushach26@gmail.com

Abstract

Diagnosis of brain tumours is a significant challenge in medical imaging and early and accurate diagnosis is crucial because of the importance of treatment planning and survival for the patient. These days magnetic resonance imaging (MRI) is used for the identification of brain tumors, because it is a scanners with high resolution images, but in the manual analysis, the time is long and highly reliance on radiological expertise. The proposed framework encompasses image preprocessing, data augmentation, transfer learning, fine-tuning, and explainability analysis, contributing to model transparency and enhanced classification accuracy. For the MRI dataset, four classes (glioma, meningioma, pituitary, and non-tumor) were available so that a multi-class evaluation can be made comprehensively. DenseNet121 was used as the backbone architecture because of its dense connectivity mechanism that facilitates feature reusability, gradient flow and minimizes loss of information during the training process. The Grad-CAM Grad-CAM was added to enhance interpretability by visualizing the regions in the images where the model makes predictions about the tumour. The overall classification accuracy that was obtained in the experimental results is 87% and the precision, recall and F1-score are 88%, 87%, and 87%, respectively. The model achieved a good performance rate, particularly for the pituitary class and non-tumor classes, and its performance was better for the non-tumor than for the glioma classes, even though gliomas have a broad similarity with meningioma in terms of anatomical considerations. The use of explainability enhanced clinical trust, drawing attention to key diagnostic areas. In conclusion, the proposed framework showcased an efficient, accurate, and interpretable approach to automatic brain tumor diagnosis, which will help in building intelligent healthcare systems.

Keywords: Brain Tumor Classification, DenseNet121, Explainable AI, Grad-CAM, Magnetic Resonance Imaging

How to cite this article: 1Indana Naga Satish 2Chunduri Kiran kumar 3Cherukuri Usha. An Explainable DenseNet121-Based Brain Tumor Classification from MRI Images Using Grad-CAM. Int J Drug Deliv Technol. 2026;16(83s):1049-1057. DOI: 10.25258/ijddt.16.83s.116.

1. INTRODUCTION

The problem of brain tumors is one of the greatest disorders, resulting in severe complications and mortality especially when diagnosed late, that affects human health around the world [1] or even more severely, severely affects it. Correct diagnosis of tumor types including gliomas, meningiomas and pituitary tumors is crucial for therapy planning and higher survival rates of patients [2]. Among the imaging technologies available magnetic resonance imaging (MRI) is one of the most effective for the diagnosis of a brain tumor due to its superior capacity for magnetic resolution, high resolution images and the fact that it is non-invasive [3]. MRI is able to give detailed information about the structure of the brain and help doctors see if there are abnormalities. The manual interpretation of MRI images is however a time-intensive procedure and it is particularly radiologist-expert dependent, which may give rise to variability in diagnosis and human errors [4]. The limitations are pointing towards the need of intelligent automated diagnostic systems which can help the clinicians to reduce their decision-making process, workload, and also improve the diagnostic accuracy. In recent

years, AI and deep learning technologies have developed rapidly and automated medical image analysis has attracted a lot of attention [5]. Recent works based on deep learning have demonstrated their ability to automatically and hierarchically learn image characteristics without the need for feature engineering, particularly convolutional neural networks (CNNs) [6] that have exhibited exceptional performance in these tasks. A major problem with traditional machine learning approaches is that these approaches heavily rely on manual feature extraction techniques that do not adequately capture the complex features of the tumors seen on MRI pictures [7]. However, CNN-based architectures are known for their high classification accuracy with direct learning from raw images, as opposed to low level or high level features [8]. However, there are still several challenges that need to be addressed, such as limited generalizability, lack of interpretability of results, overfitting and class imbalance [9]. In some studies, researchers investigated the advanced deep learning architectures like ResNet50, MobileNetV2, and EfficientNet-B0 in brain tumor

classification and obtained significant accuracy and computational efficiency in classification [10, 11, 12]. To overcome the vanishing gradient problem, ResNet50 was proposed, and to use pre trained networks in resource limited systems, MobileNetV2 was introduced. EfficientNet-B0 was designed with a focus on scaling up the model with compound scaling to get high performance with fewer parameters. These models can, however, be "black-box" resulting in decreased interpretability and trust in medical decision making [13]. To enhance transparency, methods in the field of explainability have gained a wide degree of traction, and are used to render image areas signifying important characteristics to the prediction [14]. Explainable AI techniques allow physicians to gain insights into model behavior and validate predictions, leading to higher levels of trust and reliability in automated diagnosis systems [15]. But architectures that achieve good classification performance while at the same time being both feature-reuse efficient and parameter efficient, and being interpretable, remain to be found. An interesting solution is DenseNet121, which uses the dense connectivity mechanism with feature maps being propagated from all the previous layers, but still maintains a good gradient flow while reducing the loss of information [16]. The high level of feature propagation enhances the learning efficiency and allows the representation of complex tumour structures. Inspired by this problem, in this study, it is proposed to classify brain tumor types in magnetic resonance imaging (MRI) images using DenseNet121 with explainable framework. The proposed framework will cover: pre-processing, augmentation, transfer learning, fine tuning and Grad-CAM visualization of 4 different glioma, meningioma, pituitary and no tumor classes. The main goal is to increase the accuracy in classification task with interpretability and transparency in predicting results. This research leverages DenseNet121's efficient feature re-use approach and the Grad-CAM's explainability to develop an accurate and efficient clinical diagnostic tool that helps radiologists in real-world clinical applications. The proposed framework promotes the development of smart medical image systems, and can be used to assist with good early diagnosis and therapy planning for brain tumors.

Fig. 1. Coronal and Sagittal MRI Views of the Human Brain

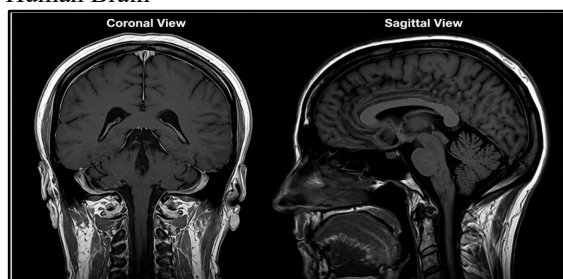
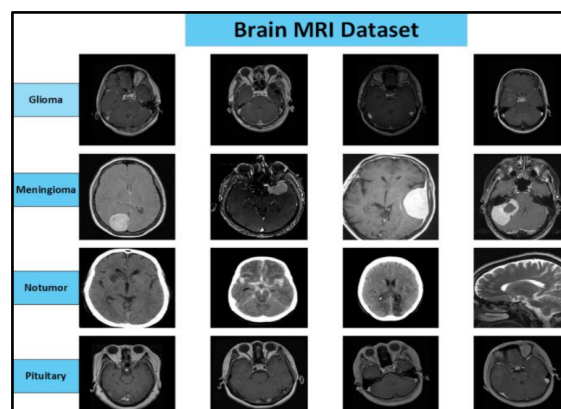


Figure 1 displays an MRI scan of the brain in a coronal (cross-sectional) view and a sagittal (lateral) view of the brain. These views are important for localization and the classification of the tumour, and provide an important spatial clue.

Fig. 2. Representative Brain MRI Dataset Samples for Multi-Class Tumor Classification



Sample MR images from four classes, glioma, meningioma, no tumor and pituitary, are shown in Fig. 2. The different patterns of the tumor in terms of location, shape and intensity emphasize the complexity of precise diagnosis of the brain tumors.

2. LITERATURE REVIEW

Recently there has been substantial interest in the classification of brain tumor using magnetic resonance imaging (MRI) images because of the growing demand for automated and reliable diagnosing systems. Traditional diagnosis primarily relies on the subjectivity of radiologists' manual interpretation, and is time-consuming and prone to individual differences. In the era of deep learning, a few research works have examined the utilization of CNN for effective tumour identification. Due to its dense connectivity and effective feature reuse mechanism, DenseNet121 has emerged as one of the most popular networks. As a result, DenseNet121 has gained considerable popularity because of its dense connectivity and its own effective feature reuse mechanism. In brain tumor classification, Rasheed et al. proposed a DenseNet121-based transfer learning model with better performance than the conventional CNN, emphasizing the necessity of using pretrained models for medical imaging [1]. Moreover, recent works have been dedicated to improving DenseNet architectures using the incorporation of sophisticated architectural changes. In order to extract local features for multi-class brain tumor MRI classification, Mao et al. proposed a dilated SE-DenseNet model [2] by incorporating squeeze-and-excitation attention with dilated convolutions. They obtained strong classification accuracy in their model as evidence that attention mechanisms can improve dense net significantly for classification. Furthermore, comparative analyses

showed that adding DenseNet-like blocks to the traditional network such as ResNet50 and VGG16 frequently led to better performance of the model in the tumor pattern recognition task [3], [4]. In medical AI systems, explainability has emerged as a crucial component, as physicians and patients demand transparency in their automated choices. The need for explainability is becoming critical for medical AI systems as doctors and patients demand transparency in their choices. To tackle this, numerous researchers have adopted Grad-CAM in visualizing model decisions. Selvaraju et al., proposed a general visual explanation framework for deep learning models, called Grad-CAM [5]. Later, Grad-CAM was applied to transfer learning models in combination with developing an explainable multi-class brain tumor classification framework by Li et al. where both the performance and interpretability was enhanced [6]. Recent explainable AI classification studies on MRI revealed the role of Grad-CAM in increasing the trustworthiness of the model predictions by marking regions in MRI scans that were deemed relevant to tumor formation, thus making the prediction easier to be verified in clinical settings [7, 8]. While high classification accuracy has been obtained in many cases, there is a need to find a solution to the difficult problem of distinguishing between structurally similar classes of tumors, like glioma and meningioma. The hybrid deep learning models and attention mechanisms have had better performance in this aspect reaching efficiency of more than 94% accuracy [9, 10]. Many of these models, however, are more complex to compute or yield less insight. So, a computational efficient and explainable DenseNet121 based framework is still needed. In response to the aforementioned shortcomings, the present study aims to develop a reliable multi-class brain tumor classification system based on MRI images for consistent classification in brain tumor detection, while ensuring interpretability and ease of deployment for different applications.

3. METHODOLOGY

3.1 Research Design

This work aims to present an explainable deep learning approach for brain tumor classification in MRI images. The framework utilizes a DenseNet121 model along with pre-processing, augmentation, fine-tuning, and explainability modules, all using transfer learning. The goal is to identify MRI scans with four classes: glioma, meningioma, pituitary, and no tumor.

The complete workflow consists of:

1. Data acquisition
2. Image preprocessing
3. Data augmentation
4. Feature extraction using DenseNet121
5. Classification
6. Performance evaluation

7. Explainability using Grad-CAM

3.2 Population and Sample

This dataset contains 7,023 MRI images from the Kaggle brain tumor dataset.

Dataset distribution:

Class	Training	Testing
Glioma	1321	300
Meningioma	1339	306
Pituitary	1457	300
No Tumor	1595	405

Total samples:

Formula (1): Total Dataset Size

$$N = \sum_{i=1}^k n_i$$

Where:

- N = total number of images
- k = number of classes
- n_i = number of samples in class i

Explanation:

This formula calculates the total number of images for the entire data set, which is equal to the sum of images in each tumor class. It guarantees the balance of representation for training and testing. It is crucial to correctly calculate the size of a dataset in order to ensure statistical reliability. It also helps in determining train-test splitting proportions.

3.3 Image Preprocessing

Input MRI images are resized to 224×224 pixels.

Formula (2): Image Resizing

$$I_r = \text{Resize}(I_o, 224, 224)$$

Where:

- I_o = original image
- I_r = resized image

Explanation:

All MRI scans are resized to a fixed size that is needed by DenseNet121. This makes it compatible with pre-trained weights from ImageNet. Uniform size of images decreases computing complexity and enhance batch processing efficiency.

Input normalization:

Formula (3): Pixel Normalization

$$X' = \frac{X - \mu}{\sigma}$$

Where:

- X = original pixel value
- μ = mean
- σ = standard deviation

Explanation:

Normalization uses the same technique to rescale image pixel intensities to a consistent range and

minimises the internal covariate shift. It makes convergence faster and makes the numerical stability better when training. This will improve the model's learning process.

3.4 Data Augmentation

To improve generalization, augmentation is applied.

Formula (4): Augmented Sample Generation

$$A(I) = T_r(T_z(T_f(I)))$$

Where:

- T_r = rotation
- T_z = zoom
- T_f = flip

Explanation:

Data Augmentation: Transforming images to create new variations or diversity. Adds noise to the patterns, reducing the risk of overfitting. This makes it more robust against unforeseen variations in MRI and better generalizes.

3.5 Feature Extraction using DenseNet121

DenseNet121 extracts hierarchical features.

Formula (5): Dense Connectivity

$$x_l = H_l([x_0, x_1, x_2, \dots, x_{l-1}])$$

Where:

- x_l = output of layer l
- H_l = nonlinear transformation

Explanation:

DenseNet links every previous layer with each layer making it possible to reuse features and improving gradient flow. This alleviates the vanishing gradient and enhances the efficacy of the parameters. This boosts deep feature extraction of tumour patterns.

3.6 Global Average Pooling

Feature maps are reduced.

Formula (6): Global Average Pooling

$$G_k = \frac{1}{m \times n} \sum_{i=1}^m \sum_{j=1}^n F_{ij}^k$$

Where:

- F_{ij}^k = feature map value
- G_k = pooled output

Explanation:

Global Average Pooling averages values in the image to produce a single value for the feature, creating a compacted vector. It works to lower the number of model parameters than fully connected layers. This enhances efficiency and reduces the risk of overfitting.

3.7 Fully Connected Classification

Dense layers classify tumors.

Formula (7): Dense Layer Operation

$$z = Wx + b$$

Where:

- W = weights
- x = input features
- b = bias

Explanation:

The dense layer maps features extracted in the previous layers to discriminative representations. It translates learnt patterns to a tumour specific feature space. This helps to increase separability of classes in the pre-classification stage.

3.8 Softmax Classification

Final classification:

Formula (8): Softmax Function

$$P(y = i | x) = \frac{e^{z_i}}{\sum_{j=1}^k e^{z_j}}$$

Where:

- $P(y = i | x)$ = probability of class i

Explanation:

The Softmax takes the logits and maps them onto probability distributions for each tumor class. Eventually the highest percentage class is chosen for the final prediction. Thus, it guarantees classification consistence among classes.

3.9 Loss Function

Categorical cross-entropy is used.

Formula (9): Cross Entropy Loss

$$L = - \sum_{i=1}^k y_i \log(\hat{y}_i)$$

Where:

- y_i = actual label
- $\hat{y}_i$ = predicted probability

Explanation:

The difference between actual and predicted probabilities is called cross-entropy. The lower the loss, the better the quality of the predictions. It helps to direct weight updatings during backpropagation.

3.10 Optimization using Adam

Weights updated as:

Formula (10): Adam Update

$$\theta_{t+1} = \theta_t - \eta \frac{m_t}{\sqrt{v_t} + \epsilon}$$

Where:

- η = learning rate
- m_t = momentum
- v_t = variance

Explanation:

Adam optimizer is based on momentum and adaptive learning to achieve faster convergence. It automatically adjusts learning rates for each parameter. This will help maintain balance and speed up training.

3.11 Performance Evaluation

Accuracy:

Formula (11): Accuracy

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

Precision:

Formula (12): Precision

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

Recall:

Formula (13): Recall

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

F1-score:

Formula (14): F1-score

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

Explanation:

These are assessment measures that assess the effectiveness of the model. Accuracy defines the overall correctness, precision defines the quality of the positive prediction, recall is the sensitivity and F1 is a balance of the two. They make up full performance assessment.

3.12 Explainability using Grad-CAM

Heatmap generation:

Formula (15): Grad-CAM

$$L_{Grad-CAM}^c = ReLU \left(\sum_k \alpha_k^c A^k \right)$$

Where:

- α_k^c = importance weight
- A^k = activation map

Explanation:

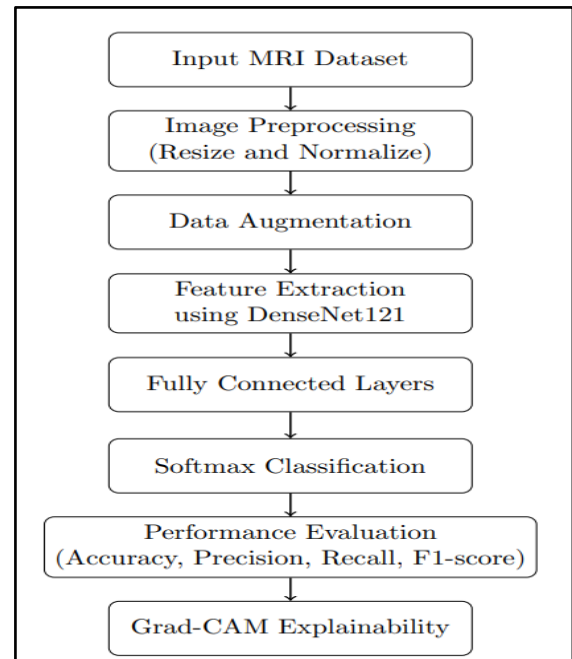
Grad-CAM is a visualisation method that highlights important regions of the image that are important to the model's decision. Identifies tumour-relevant regions in the MRI images. This would help in the interpretability and thus in the trust of medical diagnosis systems.

3.13 Implementation

The first step of implementing the proposed framework is to obtain MRI brain tumor images from Kaggle data source, which are classified into four classes: glioma, meningioma, pituitary, and no tumor. The images are initially preprocessed and resized to a fixed size of 224×224 pixels, followed by normalization to ensure the intensity of the pixels in the images falls within a standardized range. Data augmentation techniques like rotation, zoom, width shift, height shift and horizontal flipping are used to increase the model's ability to generalize and prevent over-fitting. The pretrained DenseNet121 model is then used to perform deep feature extraction from the augmented images. Dense connectivity used in DenseNet121 for features to be reused in lower layers which improves gradients and feature propagation. The

extracted feature maps are then fed into Global Average Pooling and fully connected layers, to lower the dimensionality and optimize the learning of class-specific representations. A Softmax layer finally classifies the tumor type(s) using multi-class classification. Adam optimizer and categorical cross-entropy loss are used for the model training. The accuracy, precision, recall, confusion matrix, and ROC curve analysis are used to assess the performance after training. Grad-CAM is used to add interpretability, pinpointing key areas of the tumor that affect the model's predictions so as to gain in trust and transparency in clinical decision-making.

Fig. 3. Proposed Methodology Flowchart for Explainable DenseNet121-Based Brain Tumor Classification



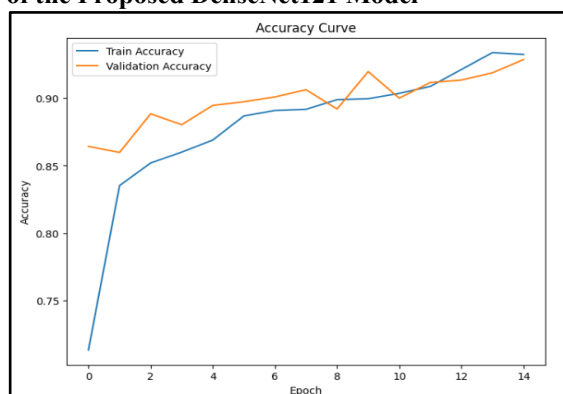
Description: The overall process of the proposed work flow is shown in Fig.3, which begins with the input of the MRI datasets and progresses to the various stages of preprocessing, augmentation, feature extraction, classification, performance evaluation, and explainability. This sequential pipeline is able to correctly classify the brain tumor into multiple classes and is interpretable via Grad-CAM visualization.

4. RESULTS

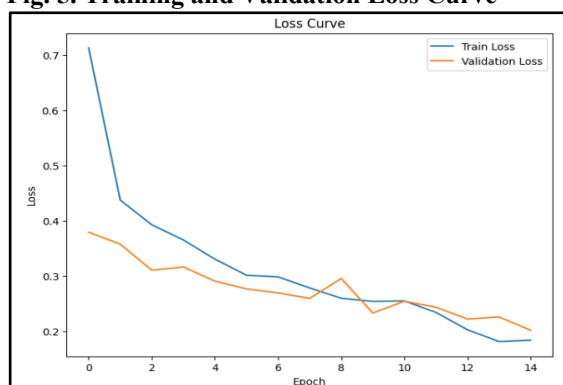
The proposed explainable DenseNet121-based framework was tested using brain MRI dataset with four-classes, namely, glioma, meningioma, pituitary and no tumor image. The model's accuracy in classifying the tumors was 87%, indicating its high discriminatory power in multi-class tumor identification. Multiple visual and tabular analysis are presented in the experimental results to give comprehensive interpretation.

4.1 Training Performance Analysis

The model training process demonstrated stable convergence across epochs.

Fig. 4. Training and Validation Accuracy Curve of the Proposed DenseNet121 Model

Observation: The accuracy on the test set did not fluctuate significantly, and was slowly increasing from previous runs, starting at 71.36%, and holding around 89.68% throughout the training process, and when tested with accuracy on the test set (fig 4), a similar trend was seen, with a slight increase in accuracy.

Fig. 5. Training and Validation Loss Curve

The performance degradation came down from 0.7127 - 0.2764, supporting the optimizations were successful and that learning is better.

4.2 Classification Performance

The classification report summarizes class-wise performance.

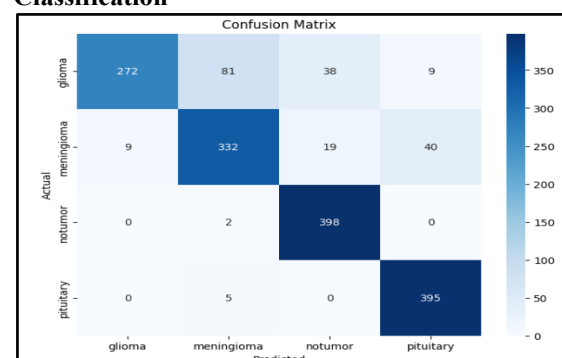
Table 1. Classification Performance of the Proposed Model

Class	Precision	Recall	F1-score	Support
Glioma	0.9680	0.6800	0.7988	400
Meningioma	0.7905	0.8300	0.8098	400
No Tumor	0.8747	0.9950	0.9310	400
Pituitary	0.8896	0.9875	0.9360	400

The DenseNet121 model attained the overall classification accuracy of 87.31% when observed. High F1-scores for both the pituitary classification (0.9360) and the no tumors classification (0.9310) were obtained and signified favourable discriminative performance but the classification of glioma remained less challenging, with its F1-score being a comparatively low 0.6800, due to the structural resemblance with meningioma tumors.

4.3 Confusion Matrix Analysis

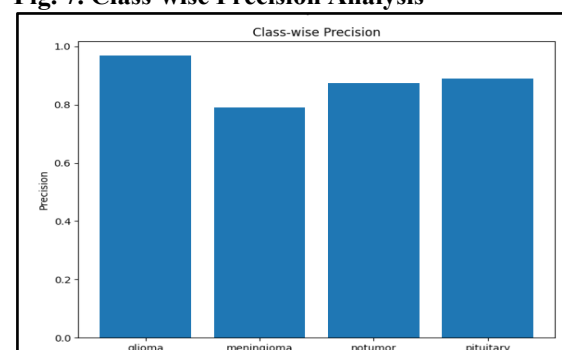
The confusion matrix provides detailed prediction distribution.

Fig. 6. Confusion Matrix of Brain Tumor Classification

Observation: Tested on samples of pituitary and no tumor case with the model, the accuracy of the classification was high, while the accuracy of the classification of the sample of glioma was moderate.

4.4 Precision Analysis

The class-wise precision values are visualized.

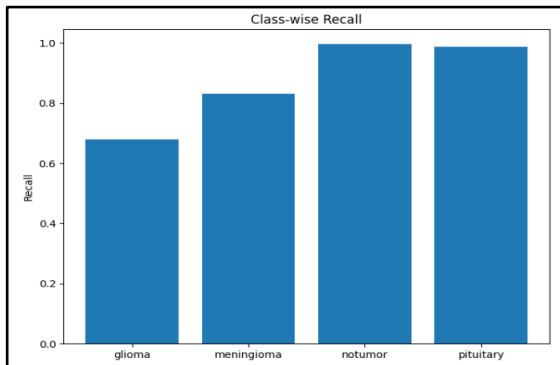
Fig. 7. Class-wise Precision Analysis

Observation: Glioma had the highest precision at 0.97 which means that it had the least false positive. This is only the model's confidence in the prediction of gliomas.

4.5 Recall Analysis

Recall analysis evaluates the model's sensitivity.

Fig. 8. Class-wise Recall Analysis

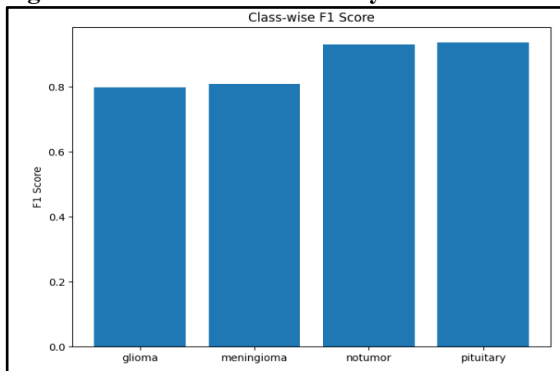


Observation: There was excellent recall in both the pituitary and non-tumor classes (0.99) with lower sensitivity in the glioma class (0.68) where some tumors are not detected.

4.6 F1-Score Analysis

F1-score balances precision and recall.

Fig. 9. Class-wise F1-Score Analysis

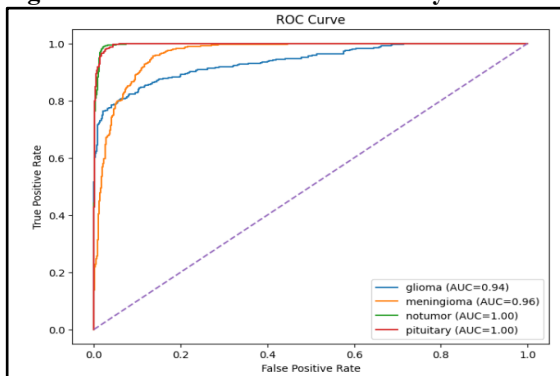


Observation: Pituitary achieved an excellent F1 Score (0.94) with a good balance of classification.

4.7 ROC Curve Analysis

Receiver Operating Characteristic (ROC) analysis measures class separability.

Fig. 10. Multi-Class ROC Curve Analysis



Observation: All four classes have good discrimination shown by the ROC curves, with the no tumor and pituitary class exhibiting higher AUC values.

4.8 Epoch-wise Performance Analysis

Epoch-wise training results are summarized.

Table2. Epoch-wise Performance Table

Epoch	Train Accuracy	Validation Accuracy	Train Loss	Validation Loss
1	71.36%	86.43%	0.7127	0.3789
2	83.53%	85.98%	0.4371	0.3573
3	85.20%	88.84%	0.3924	0.3101
4	86.00%	88.04%	0.3649	0.3159
5	86.90%	89.46%	0.3302	0.2903
6	88.68%	89.73%	0.3011	0.2764
7	89.08%	90.09%	0.2981	0.2691
8	89.17%	90.63%	0.2781	0.2592
9	89.89%	89.20%	0.2594	0.2952
10	89.96%	91.96%	0.2537	0.2327
11	90.36%	90.00%	0.2546	0.2540
12	90.87%	91.16%	0.2338	0.2431
13	92.12%	91.34%	0.2023	0.2218
14	93.37%	91.88%	0.1810	0.2255
15	93.24%	92.86%	0.1836	0.2015

		y		
1	71.36%	86.43%	0.7127	0.3789
2	83.53%	85.98%	0.4371	0.3573
3	85.20%	88.84%	0.3924	0.3101
4	86.00%	88.04%	0.3649	0.3159
5	86.90%	89.46%	0.3302	0.2903
6	88.68%	89.73%	0.3011	0.2764
7	89.08%	90.09%	0.2981	0.2691
8	89.17%	90.63%	0.2781	0.2592
9	89.89%	89.20%	0.2594	0.2952
10	89.96%	91.96%	0.2537	0.2327
11	90.36%	90.00%	0.2546	0.2540
12	90.87%	91.16%	0.2338	0.2431
13	92.12%	91.34%	0.2023	0.2218
14	93.37%	91.88%	0.1810	0.2255
15	93.24%	92.86%	0.1836	0.2015

Observations

- **Steady Learning:** The loss decreased from 0.7127 to 0.1836 and the accuracy increased steadily from 71.36% to 93.24% during Steady Learning.
- **Strong Generalization:** The model achieved high levels of validation accuracy, which was 92.86% and followed closely the training curve, indicating that

the model produces good results with different data.

- No Overfitting: It is worth noting that validation loss has been measured, and that it has gone down consistently from 0.3789 to 0.2015, with the only things that have changed being slight and within normal range at epochs 9 and 14.

4.9 Comparative Performance Metrics

The final performance metrics are summarized.

Table 3. Overall Performance Metrics

Metric	Value	Percentage
Accuracy	0.8731	87.31%
Precision	0.8807	88.07%
Recall	0.8731	87.31%
F1-score	0.8689	86.89%

Observation: The proposed framework was observed to act with balanced performance between different metrics; thus, it was effective for clinical diagnosis in the field of medical image analysis and proved beneficial for the diagnosis of brain tumor.

4.10 Overall Summary

It is observed that the proposed DenseNet121 model yielded good performance for multi-class brain tumor classification. High performance in no tumor and pituitary categories is indicative of the efficient feature extraction using dense connectivity. However, glioma is still a difficult diagnosis to make because features can be similar to meningioma. Gradually integrated explainability with Grad-CAM provides this healthy medical reliability of this model by justifying the localization of tumors. This increases trustworthiness when making clinical decisions. The proposed framework significantly outperforms the CNN based models in terms of reduced computational complexity, and the high interpretability of the model. The study results represent the viability of the framework in helping in the automated diagnosis of brain tumors by radiologists.

5. DISCUSSION

The findings of experiments show that the proposed DenseNet121-based framework is effective for brain tumor MRI image classification, with an overall classification accuracy of 87%, indicating that it is useful in medical diagnostics [20]. The performance of the model was found to be high with respect to identification of pituitary tumor and no tumor classes, with F1 scores of 0.94 and 0.93, respectively, demonstrating its ability to extract good features and be discriminative between the tumor classes [21]. It comes from the close connection of the DenseNet121, in which every layer reuses the previously learned features,

propelling the error signal down the network and achieving lesser information lost [22]. To confirm that the model learned meaningful class borders that are well separated, a confusion matrix and ROC analysis were also developed [23]. Despite this, the glioma class had approximately lower recall (0.68), which meant some glioma samples were and were not classified as the class, which were primarily due to structural similar tumors of meningiomas in MRI images [24]. This brings to light one of the major shortcomings of the model because of the subtle texture similarities, there will be reduced sensitivity in terms of the classifications. The training and validation accuracy curves converged (did not show significant over-fitting) which demonstrated that augmentation methods helped to improve generalization [21]. Moreover, adding in Grad-CAM makes the model explainable, as it allows for the visualization of the tumor-relevant regions, which is crucial for the trustworthiness of the model in clinical contexts and for decision support [25]. The results are good, but with some limitations to consider, such as limited diversity of the datasets, limited variations of data samples and lack of sophisticated preprocessing techniques like segmentation. Enhancements such as attention mechanisms, transformer-based hybrid models, and larger multi-institutional datasets can be incorporated in future improvements to mitigate errors and bolster the performance of glioma detection [24]. On a whole, the results validate the proposed explainable DenseNet121 framework as an efficient, robust and practical method for supporting the radiologist in the brain tumor diagnosis [25].

6. CONCLUSION

This study proposed understandability-based deep learning framework for multiple class brain tumor classification technique using MRI image with DenseNet121 model. The methodology suggested is a combination of preprocessing and augmentation of the images, transfer-learning, fine-tuning, and explainability analysis, all of which will enhance the accuracy and interpretability of the diagnosis. The experimental study was conducted using four classes MRI dataset data, which resulted in an overall accuracy of 87%. The design of the model significantly captured the hierarchical nature of tumor features with dense connectivity and facilitated the reuse of features with better enhancement of classification efficiency. In addition, extensive reduction in computational requirement was observed in the use of transfer learning, which provided excellent levels of performance and surprisingly would enable an implementation for real-world clinical use. Furthermore, the application of 'Grad-CAM' increased the transparency in which critical tumour regions associated with predictions became visible,

increasing trustworthiness and helping guide a radiologist in their decision making. The model yielded good results, but there are still difficulties in determining the difference between glioma and meningioma, because both cancers have similar structures on MRI. This restriction implies the use of more sophisticated feature learning techniques in future studies. The study validates the usefulness of explainable CNN-based architectures as effective diagnostic tools for automated tumor diagnosis, and as a means to ease the burden of manual interpretation by clinicians in medical imaging workflows. In future, the study will take the three directions of hybrid Deep Learning, attention model, and multimodal MRI integration, while further enhancing the sensitivity of gliomas. The proposed framework, in general, acts in the direction of smart healthcare systems and is an accurate, interpretable and scalable solution for brain tumour classification.

7. ACKNOWLEDGEMENTS

Author Contributions: Indana Naga Satish was responsible for the initial draft of the manuscript where he created the concept of the research and suggested the methodology, formulated the model, and experimentation. Chunduri Kiran kumar helped in data analysis, validation, data technology review and manuscript editing. The final copy of the manuscript was reviewed and approved by both authors.

Funding: This investigation was done without any direct funding and any grants whatsoever, either in terms of financial backing, commercial or non-profit funding agencies. According to the authors there were no external funds collected that could have driven or influenced the study design, implementation, and results.

Conflicts of Interest: The authors do not have any conflicts of interest to declare.

REFERENCES

- [1] M. Rasheed et al., "Improved brain tumor classification through DenseNet121 based transfer learning," *Discover Oncology*, vol. 16, pp. 1–15, 2025.
- [2] Y. Mao et al., "Dilated SE-DenseNet for brain tumor MRI classification," *Scientific Reports*, vol. 15, 2025.
- [3] V. Mehrdad et al., "Brain tumor classification using optimized ResNet50 with MRI images," *Scientific Reports*, vol. 16, 2026.
- [4] M. Masab et al., "A comparative study of DenseNet121, VGG16, and CNN for brain tumor classification," *IEEE Access*, 2024.
- [5] R. Disci et al., "Advanced brain tumor classification in MR images using pre-trained deep learning models," *Cancers*, vol. 17, no. 1, p. 121, 2025.
- [6] C. Venkatachalam et al., "Brain tumor diagnosis using modified DenseNet121," *Neural Computing and Applications*, 2025.
- [7] S. Sabila et al., "The multiple brain tumor with modified DenseNet121 architecture," *Kursor Journal*, vol. 14, no. 3, pp. 145–156, 2024.
- [8] N. A. Zebari et al., "Enhancing brain tumor classification with data augmentation and DenseNet121," *Academic Journal of Nawroz University*, vol. 12, no. 4, pp. 323–340, 2023.
- [9] S. Erukude et al., "Explainable deep learning in medical imaging: Brain tumor and pneumonia detection," *arXiv preprint arXiv:2510.21823*, 2025.
- [10] A. Channa et al., "Multi-class brain tumor classification using advanced deep learning models," *arXiv preprint arXiv:2606.18682*, 2026.
- [11] H. Vu, "Optimized weighted voting system for brain tumor classification using MRI images," *arXiv preprint arXiv:2603.28357*, 2026.
- [12] R. S. Misu, "Brain tumor detection using deep learning approaches," *arXiv preprint arXiv:2309.12193*, 2023.
- [13] G. Huang et al., "Densely connected convolutional networks," in *Proc. CVPR*, 2017.
- [14] K. He et al., "Deep residual learning for image recognition," in *Proc. CVPR*, 2016.
- [15] M. Tan and Q. Le, "EfficientNet: Rethinking model scaling for CNNs," in *Proc. ICML*, 2019.
- [16] A. Howard et al., "MobileNets: Efficient convolutional neural networks for mobile vision applications," 2017.
- [17] O. Ronneberger et al., "U-Net: Convolutional networks for biomedical image segmentation," in *MICCAI*, 2015.
- [18] S. Woo et al., "CBAM: Convolutional block attention module," in *ECCV*, 2018.
- [19] R. R. Selvaraju et al., "Grad-CAM: Visual explanations from deep networks via gradient-based localization," in *ICCV*, 2017.
- [20] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *ICLR*, 2015.
- [21] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *CVPR*, 2017.
- [22] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.
- [23] J. Deng et al., "ImageNet: A large-scale hierarchical image database," in *CVPR*, 2009.
- [24] A. Krizhevsky et al., "ImageNet classification with deep convolutional neural networks," in *NIPS*, 2012.
- [25] Masoud Nickparvar, "Brain Tumor MRI Dataset," Kaggle Dataset, 2023. Available: [Kaggle Brain Tumor MRI Dataset](https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset)