EARLY BRAIN TUMOR DETECTION, SEGMENTATION AND CLASSIFICATION WITH EXPLAINABLE AI

A Major Project Report submitted in partial fulfillment of the
requirements for the award of the degree of
Bachelor of Technology (B.Tech) + Master of Technology (M.Tech)
In
Computer Science and Engineering (Integrated Double Degree Masters Program)

By
TNS PRADYUMNA (22011M2107)
A PAVAN KUMAR GOUD (22011M2112)
H YASHWANTH (22011M2114)
THS CHAKRADHAR (22011M2116)

Under The Guidance of

Dr. M. ARATHI

Professor

Department of Computer Science and Engineering,

JNTUH University College of Engineering, Science &Technology Hyderabad

Kukatpally, Hyderabad - 500 085.

JNTUH University College of Engineering, Science & Technology

Hyderabad

Kukatpally, Hyderabad - 500 085

Department of Computer Science and Engineering

DECLARATION BY THE CANDIDATES

We, **TNS Pradyumna (22011M2107), A Pavan Kumar (22011M2112), H Yashwanth (22011M2114), and THS Chakradhar (22011M2116),** hereby declare that the major project report entitled **"Early Brain Tumour Detection, Segmentation and Classification with Explainable AI"**, carried out by us under the guidance of **Dr. M. Arathi**, is submitted in partial fulfilment of the requirements for the award of the degree of *Bachelor of Technology & Master of Technology in Computer Science and Engineering (Integrated Double Degree Masters Program)*.

This is a record of bonafide work carried out by us, and the results embodied in this project have not been reproduced or copied from any source. Whenever we have used materials such as data, theoretical analysis, figures, or text from other sources, due credit has been given to them by appropriate citation within the report and by listing them in the references section.

We further declare that the results and findings presented in this project **have not been submitted** to any other University or Institute for the award of any degree or diploma.

T.N.S. PRADYUMNA (22011M2107)

A. PAVAN KUMAR (22011M2112)

H. YASHWANTH (22011M2114)

T.H.S. CHAKRADHAR (22011M2116)

JNTUH University College of Engineering, Science & Technology Hyderabad

Kukatpally, Hyderabad - 500 085.

Department of Computer Science and Engineering

CERTIFICATE BY THE SUPERVISOR

This is to certify that the project report entitled "**Early Brain Tumour Detection, Segmentation and Classification with Explainable AI**", being submitted by TNS Pradyumna (22011M2107), A Pavan Kumar (22011M2112), H Yashwanth (22011M2114), and THS Chakradhar (22011M2116), in partial fulfilment of the requirements for the award of the degree of Bachelor of Technology & Master of Technology in Computer Science and Engineering (Integrated Double Degree Masters Program), is a record of bonafide work carried out by them. The results of investigation enclosed in this report have been verified and found satisfactory. The results embodied in the project have not been submitted to any other University or Institute for the award of any other Degree or Diploma.

Dr. M. ARATHI,

DATE: **Professor**

JNTUH University College of Engineering, Science & Technology

Hyderabad

Kukatpally, Hyderabad - 500 085.

Department of Computer Science and Engineering

CERTIFICATE BY THE HEAD OF THE DEPARTMENT

This is to certify that the project report entitled "**Early Brain Tumour Detection, Segmentation and Classification with Explainable AI**", being submitted by TNS Pradyumna (22011M2107), A Pavan Kumar (22011M2112), H Yashwanth (22011M2114), and THS Chakradhar (22011M2116), in partial fulfilment of the requirements for the award of the degree of Bachelor of Technology & Master of Technology in Computer Science and Engineering (Integrated Double Degree Masters Program), is a record of bonafide work carried out by them.

Dr.K.P.SUPREETHI

DATE:

Professor & Head of the Department

ACKNOWLEDGEMENT

We are pleased to present the complete stage progress Report of our Major Project and would like to express our sincere gratitude to all those who supported and guided us during this initial phase.

We are deeply thankful to our internal supervisor, **Dr. M. Arathi**, Department of Computer Science and Engineering, JNTUH College of Engineering, for her continuous support, valuable insights, and timely guidance throughout the course of our project.

We also extend our sincere gratitude to the Head of the Department, Professor

Dr. K. P. Supreethi, for her constant encouragement.

We are thankful to the Project Review Committee Members for their constructive feedback.

ABSTRACT

Early detection of brain tumours is crucial for improving patient survival rates and enabling timely medical intervention. This project presents an integrated deep-learning–based solution titled **“Early Brain Tumour Detection, Segmentation, and Classification with Explainable AI”**, which combines advanced Convolutional Neural Network (CNN) architectures with Grad-CAM–based interpretability for accurate and clinically transparent diagnosis from MRI scans. The system leverages DenseNet201, InceptionV3, and EfficientNetB3, fused through a soft-voting ensemble strategy to enhance robustness and predictive accuracy. MRI images undergo preprocessing, augmentation, and deep feature extraction, followed by multiclass classification into Glioma, Meningioma, Pituitary, and No Tumour, achieving an accuracy of 99.01%. To address the black-box nature of deep learning in medical applications, Explainable AI using Grad-CAM is incorporated to generate heatmaps highlighting tumour-relevant regions, enabling clinicians to visually validate model decisions. Overall, the proposed framework offers a reliable, scalable, and interpretable diagnostic tool that reduces manual effort, supports radiologists with high-precision predictions, and paves the way for future advancements in automated medical imaging analysis.

TABLE OF CONTENTS

1. INTRODUCTION

1.1 Background

Brain tumours represent one of the most critical and life-threatening medical conditions, affecting thousands of individuals worldwide each year. Their early detection plays a decisive role in improving survival rates, treatment outcomes, and quality of life. Traditionally, radiologists analyse Magnetic Resonance Imaging (MRI) scans to identify the presence, location, and type of tumour. While MRI is highly effective and widely used due to its detailed imaging capabilities, manual interpretation is often challenging, labour-intensive, and prone to human variability.

With increasing workload in hospitals, shortage of trained specialists, and the rising number of MRI scans requiring immediate attention, delays in early diagnosis have become more common. Even experienced radiologists may face difficulties in consistently identifying subtle patterns, differentiating between tumour types, or analysing large volumes of patient data in real time.

This growing need for faster, more accurate, and reliable diagnostic support has inspired the development of our project titled:

"Early Brain Tumor Detection, Segmentation and Classification with Explainable AI."

Advances in Artificial Intelligence, Deep Learning, and Convolutional Neural Networks (CNNs) have opened the doors for intelligent medical-image analysis systems capable of matching, and in some cases exceeding, human-level performance. Modern architectures such as DenseNet201, InceptionV3, and EfficientNetB3 have shown exceptional capabilities in extracting deep features from MRI scans, enabling highly accurate tumour classification.

However, a major challenge in adopting AI systems in healthcare is the "black-box" nature of deep learning models. Clinicians require trustworthy, interpretable systems—not just predictions. To bridge this critical gap, our project integrates Explainable AI (XAI) methods, especially Grad-CAM, which visually highlights tumour-relevant regions in MRI images. This helps radiologists understand why the model arrived at a particular decision, increasing transparency and clinical trust.

In conventional diagnostic procedures, radiologists spend considerable time reviewing multiple MRI slices from different perspectives. The process becomes even more challenging when tumours are small, irregular, or visually similar to surrounding tissues. Misinterpretation can delay treatment or lead to unnecessary surgical procedures. This problem is especially significant in areas with limited medical expertise.

Our proposed system overcomes these limitations by combining three powerful CNN models into an ensemble approach, ensuring more robust performance, higher accuracy, and reduced model bias. By leveraging deep learning, feature extraction, multi-model integration, and explainable visualization, the system enables faster, more reliable, and interpretable brain-tumour diagnosis.

This project signifies an important step toward intelligent healthcare solutions—bridging the gap between cutting-edge AI technologies and clinical needs. It empowers medical professionals by providing an assistive tool capable of early detection, classification, and interpretability, ultimately contributing to improved patient care and medical decision-making.

1.2 Purpose

The primary purpose of this project is to develop a high-accuracy, AI-powered diagnostic system capable of detecting, classifying, and interpreting brain tumours from MRI scans. By integrating ensemble deep-learning models and explainable AI techniques, the system aims to:

- Identify the presence or absence of a tumour with high precision
- Classify tumour types such as Glioma, Meningioma, Pituitary, and No Tumour
- Assist radiologists by providing interpretable Grad-CAM heatmaps
- Reduce diagnostic errors caused by fatigue, workload, or human limitations
- Provide a fast, reliable second-opinion tool for early tumour detection

The ultimate objective is to support medical practitioners with an intelligent system that enhances accuracy, speeds up diagnosis, and provides transparent AI-driven decision support.

1.3 Scope

The scope of this project includes the design, development, training, evaluation, and explainability of a multi-model deep learning pipeline for brain tumour diagnosis using MRI images.

The system offers the following capabilities:

- Automated brain tumour detection using MRI slices
- Classification into four categories: Glioma, Meningioma, Pituitary, No Tumor
- Integration of three CNN models: DenseNet201, InceptionV3, EfficientNetB3
- Ensemble-based prediction to boost accuracy and robustness
- Grad-CAM explainability for tumour-region visualization
- Data preprocessing including augmentation, normalization, and scaling
- Performance evaluation using accuracy, precision, recall, and F1-score

The system is structured around a deep learning pipeline that includes:

- MRI dataset management

- Feature extraction through CNN models
- Ensemble model integration
- Explainable AI post-processing
- Performance comparison across models

Although additional features such as full tumour segmentation, 3D MRI volume analysis, complete clinical workflow integration, and real-time diagnosis systems may be incorporated in future phases, the current scope focuses on:

- Early detection
- Accurate classification
- Explainable AI-based visualization
- Robust multi-model ensemble evaluation

This phase forms a strong foundation for more advanced clinical diagnostic tools that may follow in future.

2. LITERATURE REVIEW

2.1 Introduction

The rapid advancement of Artificial Intelligence (AI) and Deep Learning (DL) has significantly transformed the healthcare domain, particularly in the field of medical imaging. Brain tumour diagnosis, which traditionally relies on detailed MRI interpretation by experienced radiologists, has witnessed a revolutionary shift with the adoption of CNN-based automated systems. The growing volume of MRI data, variations in tumour appearance, and dependency on human expertise have highlighted the need for intelligent computer-aided detection (CAD) systems.

This chapter presents a detailed review of research conducted in brain tumour detection, segmentation, classification, and explainable AI methods. It explores the evolution of CNN models, ensemble learning strategies, XAI techniques, and performance-enhancing methodologies implemented in state-of-the-art research.

2.2 Deep Learning in Medical Imaging

Deep learning has become the dominant approach in medical image analysis due to its ability to extract high-level hierarchical features directly from raw data without manual intervention. Unlike traditional machine-learning systems that rely on handcrafted features, CNNs learn discriminative patterns automatically.

Research studies have demonstrated that:

•CNNs outperform classical ML methods (SVM, Random Forests, KNN) in medical imaging tasks

•Deeper architectures like ResNet, DenseNet, Inception, EfficientNet provide improved feature extraction

•Transfer learning using ImageNet-trained models boosts accuracy in limited datasets

•End-to-end DL systems reduce the need for complex preprocessing

In brain tumour detection, CNNs have shown performance comparable to — and sometimes exceeding — human radiologists under controlled conditions.

2.3 CNN Architectures Used for Brain Tumor Classification

2.3.1 DenseNet (Densely Connected CNNs)

DenseNet, introduced by Huang et al., connects each layer to every other layer in a feed-forward fashion. This improves feature reuse, strengthens gradient flow, and reduces vanishing gradient problems. DenseNet201 has been widely used in tumour classification tasks due to:

- High feature extraction capability

- Efficient parameter usage
- Smooth gradient propagation
- Ability to learn fine tumour patterns

Multiple studies report >95% accuracy when applied to MRI classification

2.3.2 InceptionV3 (Multi-Scale Feature Extractor)

Inception networks were created to improve computational efficiency while increasing representational power. InceptionV3 introduces:

- Factorized convolutions
- Parallel multi-scale feature extractors
- Auxiliary classifiers
- Reduced computational cost

Because brain tumours vary in size, texture, and shape, InceptionV3 is well-suited for capturing multi-resolution tumor features.

2.3.3 EfficientNetB3 (Compound Scaling Approach)

EfficientNet introduced a novel scaling strategy that balances:

- Depth
- Width
- Input resolution

Its variants achieve state-of-the-art performance with fewer parameters. EfficientNetB3 is widely used in medical image classification because:

- It efficiently learns detailed tumour features
- It offers an excellent accuracy-to-computation ratio
- It generalizes well across datasets

Research shows EfficientNet models achieving 98–99% accuracy in MRI-based tumour classification.

2.4 Ensemble Learning Techniques

Ensemble learning is extensively used to enhance classification robustness. Instead of depending on a single model, combining multiple CNNs:

- Reduces model-specific bias
- Improves overall accuracy

- Smooths prediction errors
- Handles noisy or ambiguous MRI slices better

Common ensemble techniques include:

- Soft voting: Averaging prediction probabilities (used in this project)
- Hard voting: Majority class selection
- Weighted ensemble: Assigning model importance

Studies show that ensembles outperform individual models by 2–5%, especially in medical datasets with class imbalance.

2.5 Explainable AI (XAI) in Medical Diagnosis

One of the biggest limitations of deep learning in healthcare is the lack of interpretability. Traditional CNNs act as "black boxes", making it difficult for clinicians to understand:

- Why a model made a specific prediction
- Which regions influenced its decision
- Whether the system is focusing on tumour regions or irrelevant areas

To address this, Explainable AI methods such as Grad-CAM and Layerwise Relevance Propagation (LRP) have become vital in modern clinical AI systems.

Grad-CAM Advantages:

- Highlights tumour regions contributing to prediction
- Enhances clinician trust in AI predictions
- Helps identify misclassified or misleading cases
- Allows radiologists to compare model focus with clinical expectations

Research shows that Grad-CAM improves adoption of AI in radiology by providing transparent and interpretable visualizations.

2.6 MRI-based Brain Tumor Datasets

Several datasets have contributed to deep learning advancements:

- Figshare dataset (used in your project)
- BRATS (Brain Tumor Segmentation Challenge)
- Kaggle Brain MRI datasets

These datasets contain MRI scans annotated into classes like:

- Glioma
- Meningioma
- Pituitary Tumor
- No Tumor

Studies highlight that dataset diversity significantly influences model generalization.

2.7 Challenges in Existing Approaches

Despite significant progress, several challenges persist:

- Class imbalance: More samples of one tumour type than others
- Variability in MRI intensity & orientation
- Similarity between tumour and healthy tissues
- Requirement of large labeled datasets
- Black-box nature of deep models
- Overfitting in smaller datasets

Many studies propose solutions such as augmentation, transfer learning, hybrid models, and explainable AI.

2.8 Summary

The literature reveals that:

- CNN-based models achieve highly accurate tumour classification
- DenseNet, Inception, and EfficientNet are among the best models
- Ensemble learning enhances stability and accuracy
- Grad-CAM provides meaningful interpretability
- MRI is the most reliable modality for brain tumour diagnosis
- Integrating AI into radiology can accelerate early detection and improve patient outcomes

This project leverages the strongest findings from current research by combining powerful deep-learning architectures with ensemble learning and explainable AI — creating a clinically reliable diagnostic support system.

3.MOTIVATION

Brain tumour diagnosis remains one of the most challenging areas in medical imaging, owing to the complexity of brain structures, variations in tumour appearance, and the life-critical nature of early detection. The motivation for this project arises from a combination of clinical, technological, and societal needs that continue to grow as healthcare systems face increasing workloads and diagnostic demands.

Early detection is crucial in the management of brain tumours. Several clinical studies show that when tumours are identified in the early stages, treatment success rates improve significantly, patient survival increases, and the likelihood of invasive procedures decreases. However, early-stage tumours often exhibit subtle visual patterns in MRI scans, making them extremely difficult to detect manually—even for highly experienced radiologists. This dependency on expert interpretation creates a bottleneck in many hospitals, especially in regions with limited access to advanced neuroimaging specialists.

Furthermore, as MRI technology advances, the number of scans produced each day has increased dramatically. Radiologists are required to analyse hundreds of high-resolution slices per patient, often under severe time constraints. This increases the chances of oversight, fatigue-related errors, and inconsistencies in diagnosis. The need for an intelligent, automated system that can assist radiologists in identifying tumour patterns accurately and consistently is stronger than ever.

Technological progress in deep learning has demonstrated the potential to significantly enhance diagnostic accuracy. CNN architectures such as DenseNet, InceptionV3, and EfficientNetB3 have already shown remarkable success in processing complex medical images. However, using a single CNN model often leads to overfitting, biased predictions, or misclassification of borderline cases. Combining multiple models into an ensemble provides a more stable and reliable prediction by leveraging the strengths of different architectures. This ensemble approach directly addresses limitations found in individual models and leads to improved generalization across diverse MRI datasets.

Another significant motivation is the need for interpretability and clinical trust. While AI systems can achieve high accuracy, their lack of transparency often reduces adoption in clinical practice. Medical professionals require explanations, not just predictions. Integrating Explainable AI (XAI) techniques such as Grad-CAM allows clinicians to visualize the decision-making process of the model. This provides insights into tumour regions highlighted by the CNN, allowing radiologists to verify that the system is focusing on relevant pathological areas.

The project is also motivated by the desire to create an accessible and scalable AI-driven diagnostic tool. Many hospitals worldwide do not have immediate access to advanced diagnostic infrastructure or expert radiologists. A robust, automated system capable of assisting with early tumour classification can significantly enhance healthcare availability, reduce diagnostic delays, and help save lives.

In summary, the motivation behind this project is driven by:

- The critical importance of early detection in improving patient outcomes
- The increasing workload and pressure faced by radiologists
- The strength of CNNs in learning complex tumour patterns
- The need for robust and unbiased predictions, achievable through ensemble learning
- The essential requirement for explainability in healthcare AI
- The goal of creating a scalable, reliable, and clinically useful diagnostic support tool

By addressing these challenges, this project aims to contribute meaningfully to the development of advanced, trustworthy, and impactful AI systems in medical imaging

4.PROBLEM STATEMENT

Brain tumour diagnosis using MRI scans is a complex and time-sensitive task that requires significant expertise, precision, and consistency. Despite the availability of advanced imaging technology, manual interpretation remains challenging due to the intricate structure of the human brain, subtle variations across tumour types, and the high similarity between normal and abnormal tissue regions. These challenges highlight several critical problems that form the basis of this project.

First, early-stage brain tumours often exhibit very subtle visual characteristics, making them difficult to detect through manual examination. Radiologists must carefully analyse multiple MRI slices for each patient, which is time-consuming and susceptible to human fatigue and observational errors. As the volume of MRI scans continues to increase in clinical workflows, the pressure on radiologists intensifies, leading to delays and potential misdiagnoses.

Second, brain tumour classification is particularly challenging because tumour types—such as Glioma, Meningioma, and Pituitary—can share overlapping visual patterns. Variations in tumour shapes, textures, locations, and MRI contrasts make it harder for a single diagnostic approach to accurately differentiate between classes. Traditional machine-learning approaches relying on handcrafted features often fail to capture the complex representations necessary for robust tumour identification.

Third, although deep learning models have achieved impressive accuracy, most standalone CNN architectures exhibit limitations. A single model may become biased toward certain tumour categories, fail on ambiguous samples, or overfit on limited datasets. These shortcomings reduce the model's generalization capabilities in real-world medical settings.

Fourth, lack of interpretability remains a major barrier to clinical adoption of AI. Conventional deep learning systems often behave as "black boxes" — providing predictions without any explanation. In medical domains, such opacity is unacceptable. Clinicians require insight into why a model made a decision, which regions of the MRI influenced the outcome, and whether the AI system is focusing on medically relevant areas. Without this transparency, trust in AI-driven diagnoses remains limited.

Finally, the absence of a unified system that integrates detection, classification, and explainability restricts practical applicability in hospitals. Many existing solutions address only a single component (e.g., classification alone), and do not provide end-to-end, interpretable diagnostic assistance.

Given these challenges, the problem can be formally stated as follows:

To design and develop a deep-learning-based, explainable diagnostic system capable of accurately detecting and classifying brain tumours from MRI images into four categories — Glioma, Meningioma, Pituitary, and No Tumour — by integrating multiple CNN architectures into a unified ensemble and incorporating Grad-CAM-based interpretability to provide transparent, clinically meaningful visual explanations of the model's decisions.

This problem statement highlights the need to address issues of accuracy, generalization, transparency, and practical usability — forming the foundation for the proposed solution throughout the rest of the project.

5.SYSTEM ARCHITECTURE

The Brain Tumour Detection, Segmentation, and Classification system is designed as a structured, multi-stage deep-learning pipeline. The architecture integrates advanced Convolutional Neural Network (CNN) models, ensemble learning mechanisms, and explainable AI to produce reliable and interpretable diagnostic predictions from MRI images.

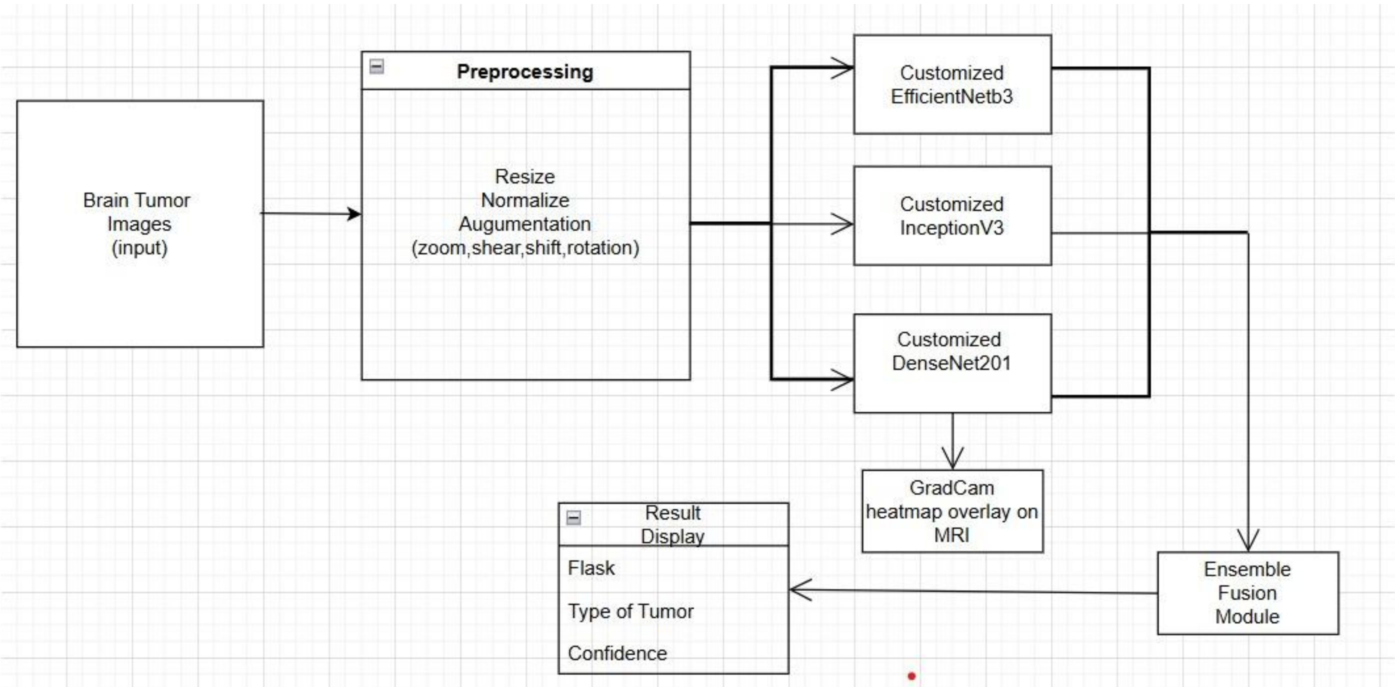
This chapter outlines the overall system structure, data flow, model integration, and key components that enable the automated end-to-end tumour classification process.

5.1 Overview of the System

The proposed system consists of the following major components:

1. Dataset Acquisition (MRI Images)
2. Data Preprocessing & Augmentation
3. Model Input Preparation
4. Deep-Learning Models (DenseNet201, InceptionV3, EfficientNetB3)
5. Ensemble Layer (Soft Voting)
6. Explainable AI Module (Grad-CAM)
7. Prediction & Visualization

Each component contributes to a robust and transparent medical diagnostic workflow.



5.2 Dataset Acquisition

The dataset consists of MRI images belonging to four classes:

- Glioma
- Meningioma
- Pituitary
- No Tumor

These images vary in size, intensity level, and orientation.

The diversity in the dataset helps ensure robust generalization across clinical scenarios.

5.3 Data Preprocessing

Before training, all images undergo several preprocessing steps:

5.3.1 Resizing

Each CNN expects a specific input size:

- DenseNet201 → 224 × 224
- InceptionV3 → 299 × 299
- EfficientNetB3 → 300 × 300

5.3.2 Normalization

Pixel intensities scaled to:

[0, 1] or [-1, 1] depending on the model.

5.3.3 Data Augmentation

To increase dataset variety:

- Horizontal/vertical flip
- Rotation
- Zoom
- Translation
- Contrast enhancement

Augmentation helps prevent overfitting and improves model generalization.

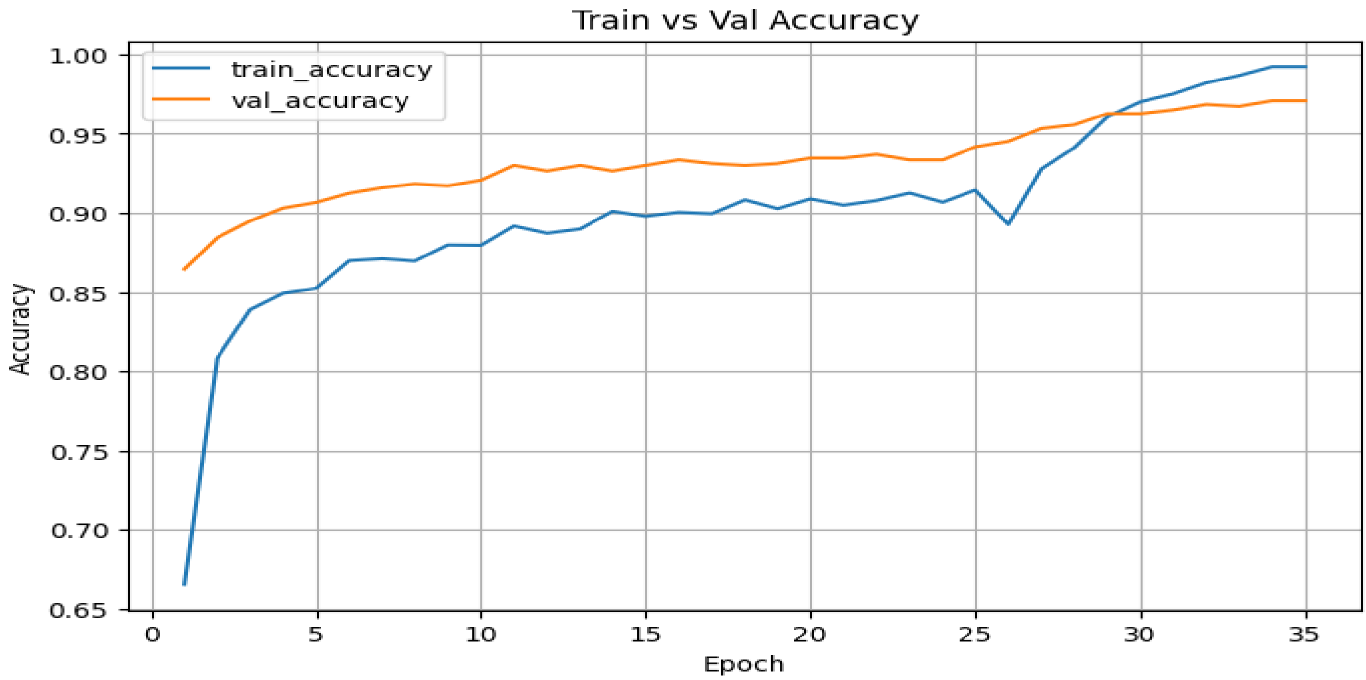
5.4 Model Architecture Components

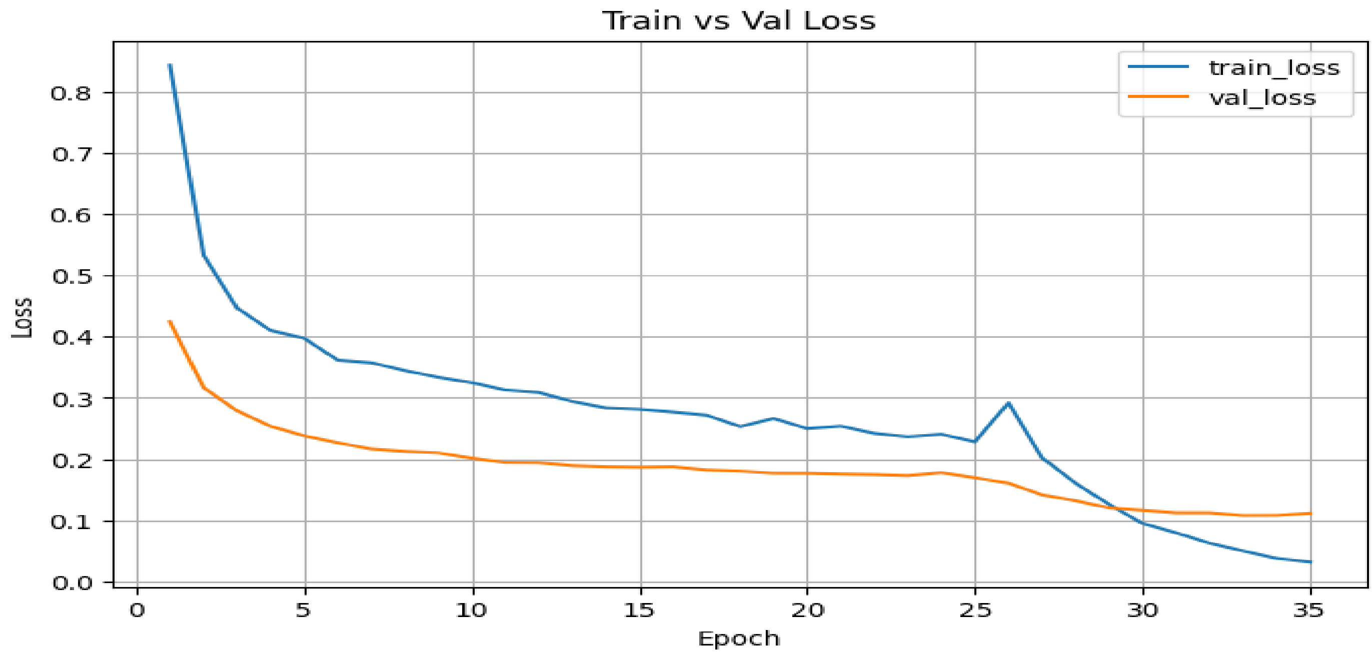
5.4.1 DenseNet201

DenseNet uses dense skip-connections where each layer receives feature maps from all preceding layers.

This supports:

- Better gradient flow
- Feature reuse
- Reduced parameters

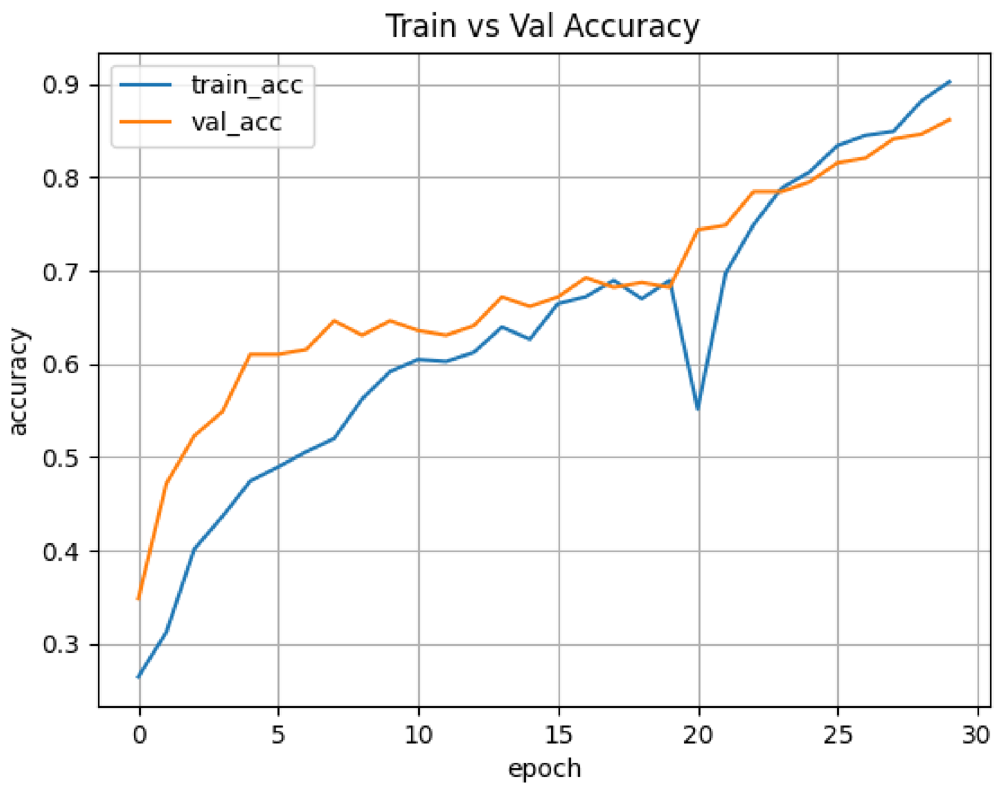


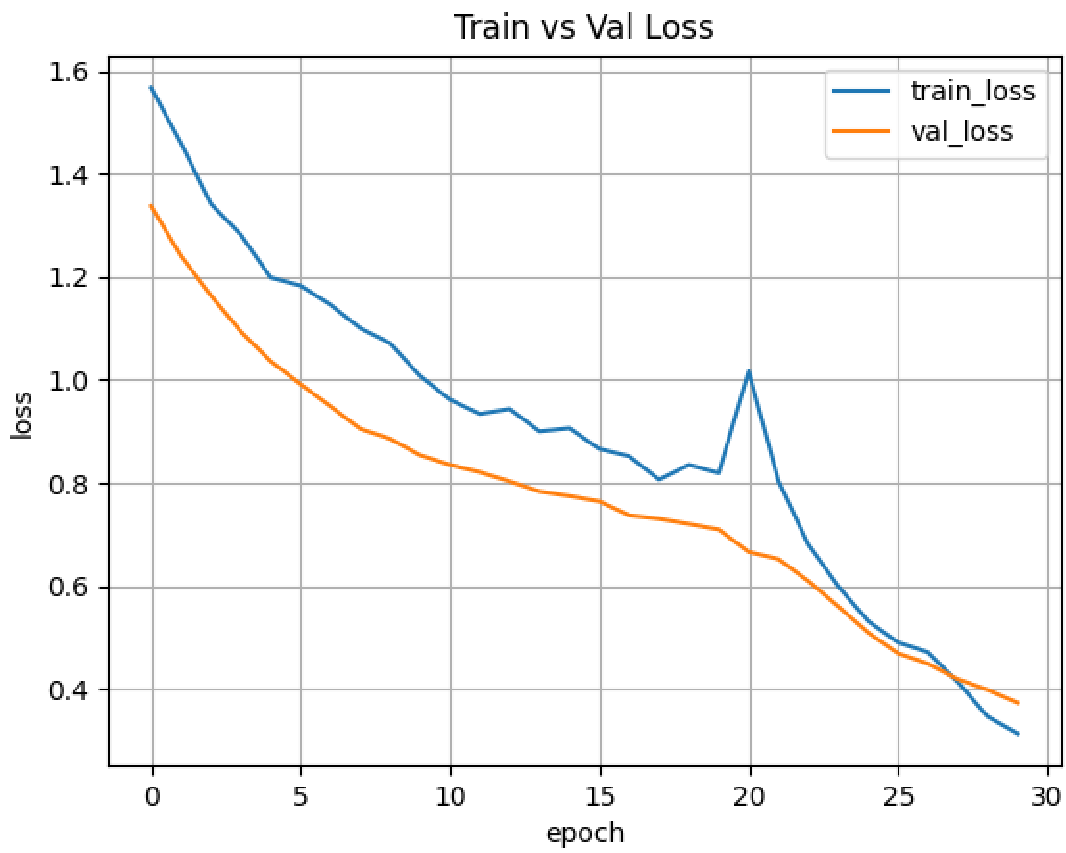


5.4.2 InceptionV3

InceptionV3 processes information through parallel convolution branches, enabling multi-scale feature extraction.
This helps detect:

- Small tumours
- Irregular shapes
- Mixed tumour textures

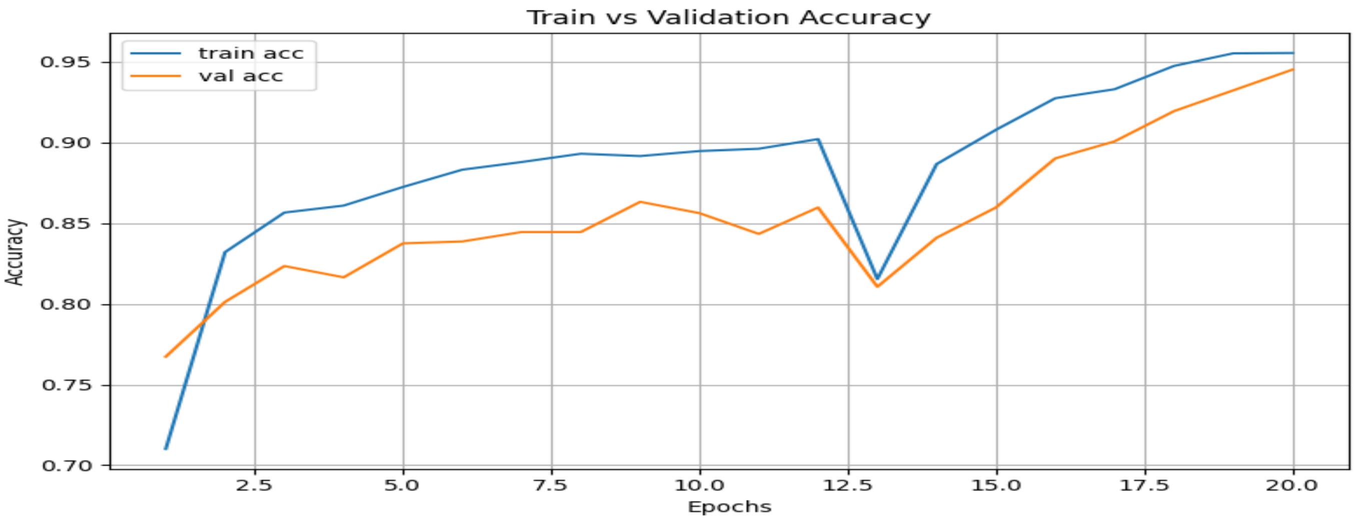

Train vs Val Accuracy
train_acc
val_acc
accuracy
epoch
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0
5
10
15
20
25
30

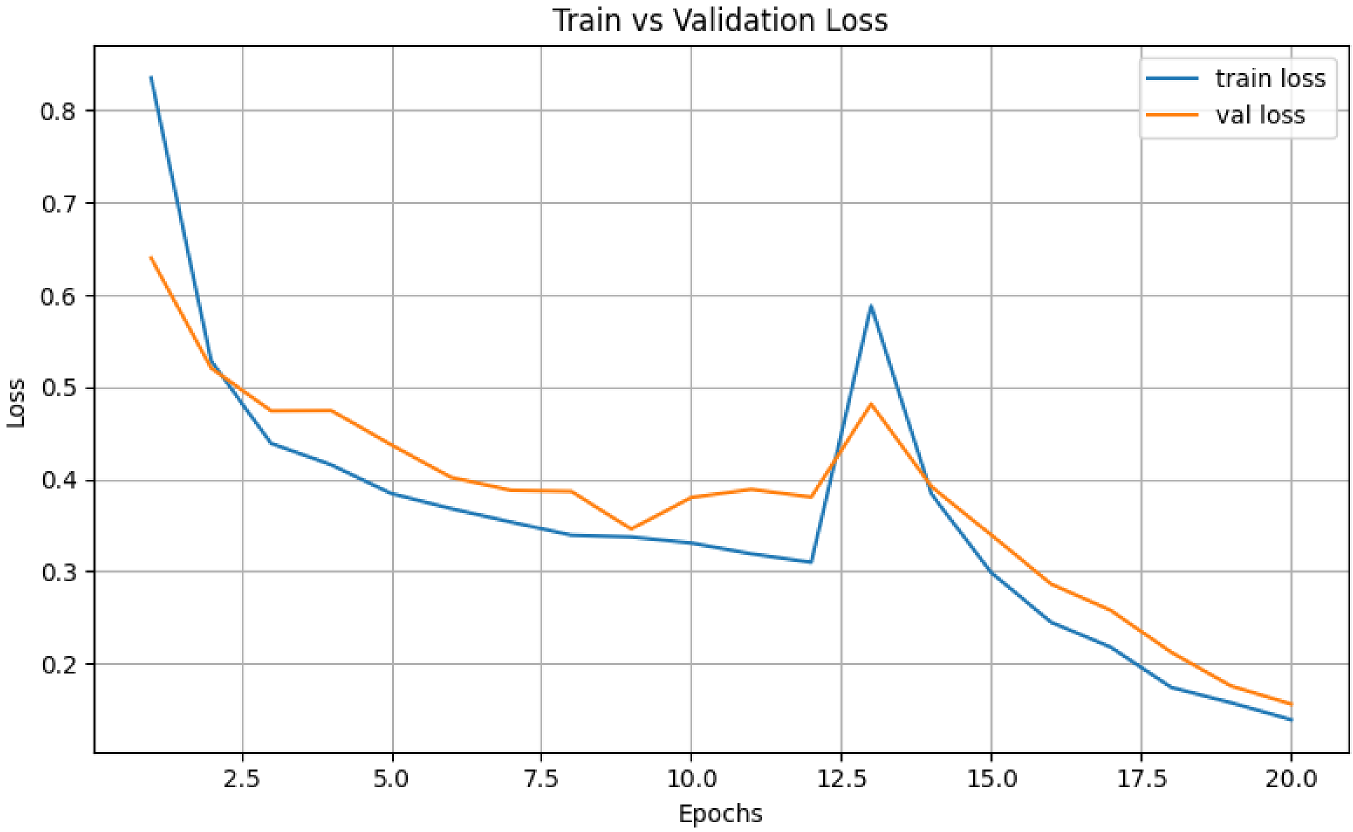

Train vs Val Loss
train_loss
val_loss
loss
epoch
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0
5
10
15
20
25
30

5.4.3 EfficientNetB3 and images

EfficientNetB3 uses compound scaling to balance:

- Width
- Depth
- Resolution





5.5.4 Ensemble Soft Voting

The ensemble integrates outputs from all three CNN models.
For each model, probabilities are averaged:

Ensemble Prediction = (0.5 DenseNet + 0.3 Inception + 0.2 EfficientNet)

This improves:

- Accuracy
- Stability
- Robustness

5.6 Explainable AI Module (Grad-CAM)

Grad-CAM produces heatmaps that highlight regions responsible for classification decisions.

Benefits:

- Enhances model transparency
- Helps radiologists interpret AI predictions
- Confirms tumour-focused learning
- Identifies incorrect attention regions in misclassified cases

5.7 System Workflow Summary

1. Load MRI image
2. Preprocess and resize
3. Pass through the three CNN models
4. Combine outputs using ensemble soft voting
5. Generate tumour class prediction
6. Produce Grad-CAM heatmaps
7. Display classification + visualization

6. IMPLEMENTATION

This chapter explains the frontend, backend, machine learning pipeline, model architectures, training strategies, performance metrics, and evaluation approach used for Early Brain Tumor Detection, Segmentation, and Classification with Explainable AI.

6.1 System Requirements

6.1.1 Hardware Requirements

- Intel i5/i7 or AMD Ryzen Processor
- Minimum 8 GB RAM (16 GB recommended for training)
- GPU support (NVIDIA CUDA-enabled GPU recommended)
- 10 GB free storage for datasets and model files

6.1.2 Software Requirements

- Python 3.10+
- TensorFlow / Keras
- NumPy, pandas, scikit-learn
- OpenCV
- Matplotlib, Seaborn
- Jupyter Notebook
- VS Code / PyCharm

6.2 Dataset Description

The dataset consists of MRI images belonging to four tumor categories:

- Glioma
- Meningioma
- Pituitary
- No Tumour

Images were pre-processed to uniform shapes required by each CNN model:

- DenseNet201 → 224×224
- EfficientNetB3 → 300×300
- InceptionV3 → 299×299

6.3 Image Preprocessing Pipeline

- Resizing based on model input requirement
- Normalization (0–1 scaling)
- Removing noise where required
- Applying augmentation:
    - Rotation
    - Shifting
    - Zoom

- o Horizontal Flip
- o Contrast adjustments

Augmentation increases dataset variety and prevents overfitting during training.

6.4 Machine Learning Models

We evaluated four CNN models:

1. DenseNet-201
2. InceptionV3
3. EfficientNetB3
4. Ensemble Soft Voting Model

Each model was trained, fine-tuned, and evaluated separately.

6.4.1 DenseNet-201 Model

- Confusion matrix (counts)
- Confusion matrix (normalized)
- Train vs Val Accuracy
- Train vs Val Loss

6.4.2 InceptionV3 Model

- Confusion matrix
- Metrics bar graph
- Train vs Val Accuracy
- Train vs Val Loss
- Combined training curves

6.4.3 EfficientNet-B3 Model

EfficientNet-B3 is a computationally efficient architecture built using compound scaling of network depth, width, and resolution. It provides high accuracy while requiring fewer parameters than deeper networks.

6.5 Workflow

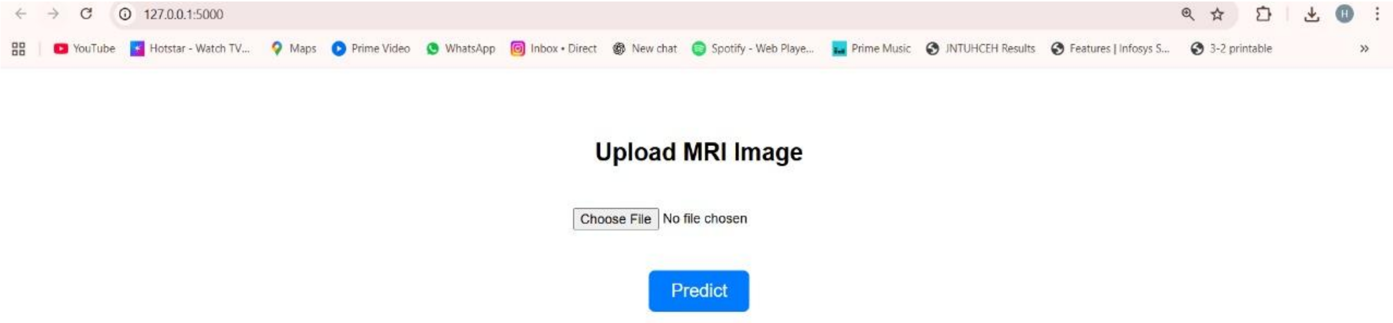
1. Load dataset
2. Preprocess images
3. Augment training data
4. Train CNN models
5. Fine-tune top layers
6. Evaluate using test set
7. Generate classification reports
8. Apply Grad-CAM for Explainable AI

6.6 Deployment

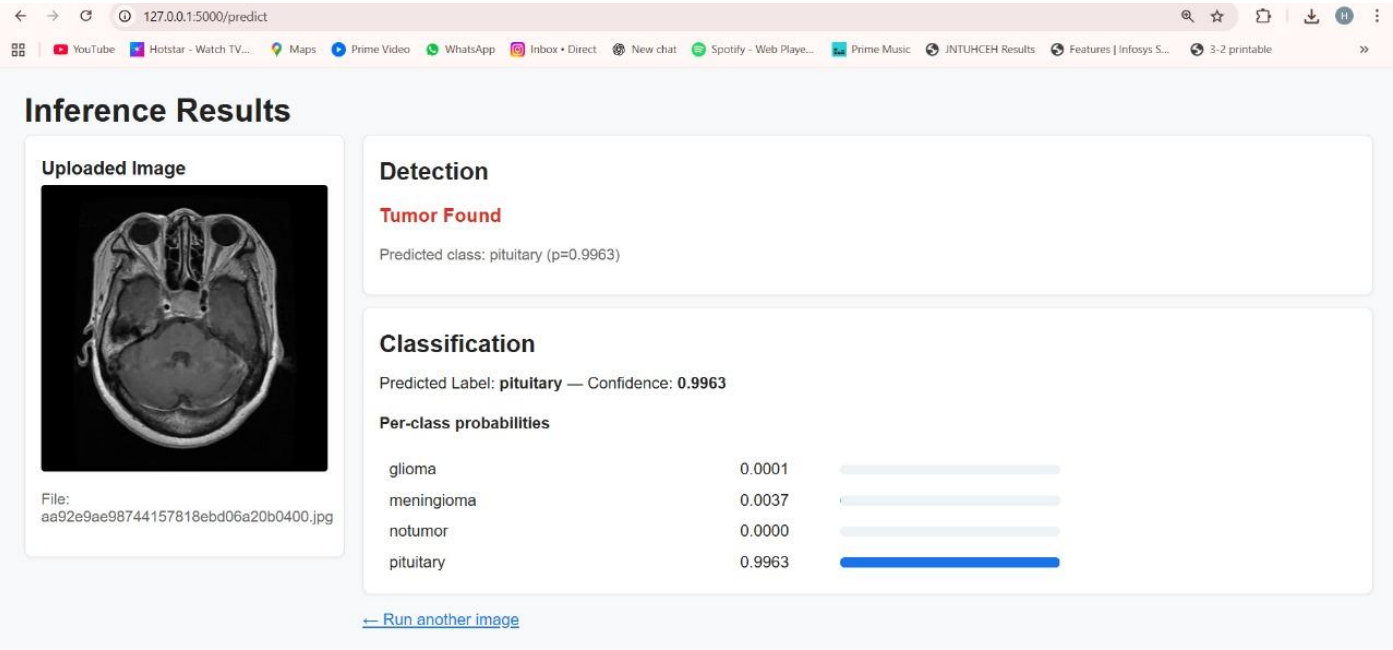
- Model saved as .h5
- Grad-CAM integrated for interpretability

- Web-based UI can be built using Flask/React (Stage-2)

IMAGES:



Web Interface for Uploading MRI Images



Inference Results Showing Tumour Detection and Class Probabilities

6.7 Key Packages and Libraries

Frontend

- HTML, CSS
- React.js
- TailwindCSS

Backend

- Python
- Flask / FastAPI
- TensorFlow

Machine Learning

- NumPy
- keras
- scikit-learn

7.RESULTS AND ANALYSIS

The performance of the proposed brain tumour detection and classification system was evaluated using a combination of individual CNN models and a soft-voting ensemble strategy. Each model was trained on the MRI dataset with extensive preprocessing, augmentation, and fine-tuning. This chapter presents a detailed analysis of the results obtained from DenseNet201, InceptionV3, EfficientNetB3, and the Ensemble Model, supported by performance metrics and visual outputs.

The evaluation focuses on accuracy, per-class performance, confusion matrices, learning curves, and overall model behaviour.

7.1 Evaluation Methodology

To assess the effectiveness of the models, the following metrics were used:

- Accuracy – Overall correctness
- Precision – How many predicted positives are correct
- Recall (Sensitivity) – How many actual positives were identified correctly
- F1-Score – Harmonic mean of precision and recall
- Confusion Matrix – Per-class classification strength and weaknesses
- Train–Validation Curves – Learning behaviour and generalization pattern

The test set consisted of unseen MRI images distributed across the four classes:

- Glioma
- Meningioma
- Pituitary
- No Tumour

7.2 DenseNet201 – Results and Analysis

DenseNet201 achieved strong and stable performance across all tumour types due to its dense connections and feature reuse. It maintained an excellent balance between training and validation curves, indicating minimal overfitting.

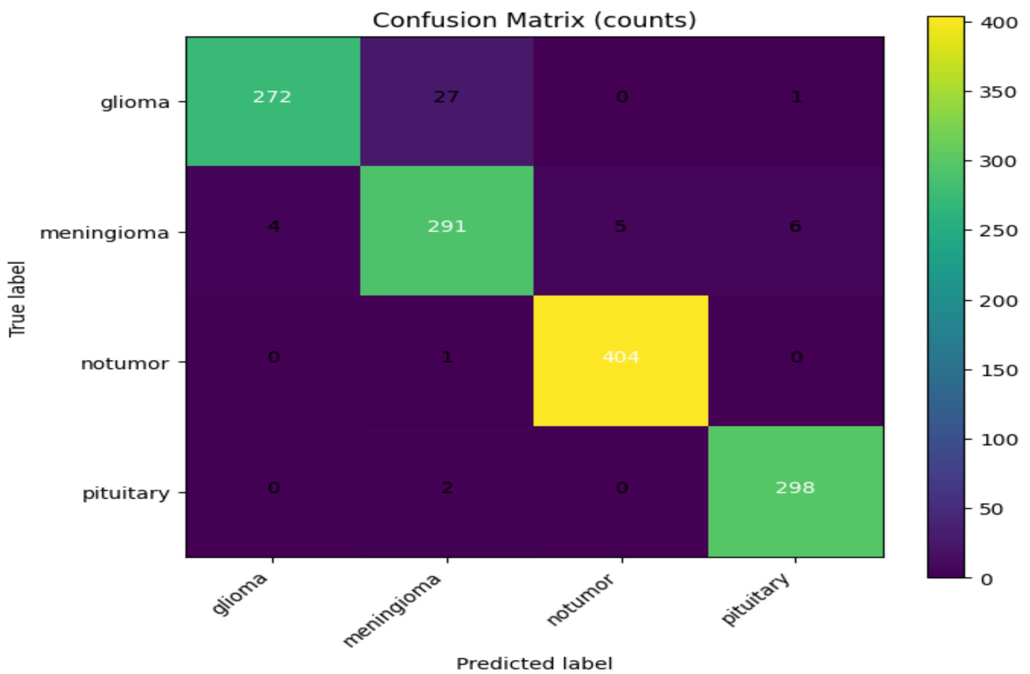
Performance Summary

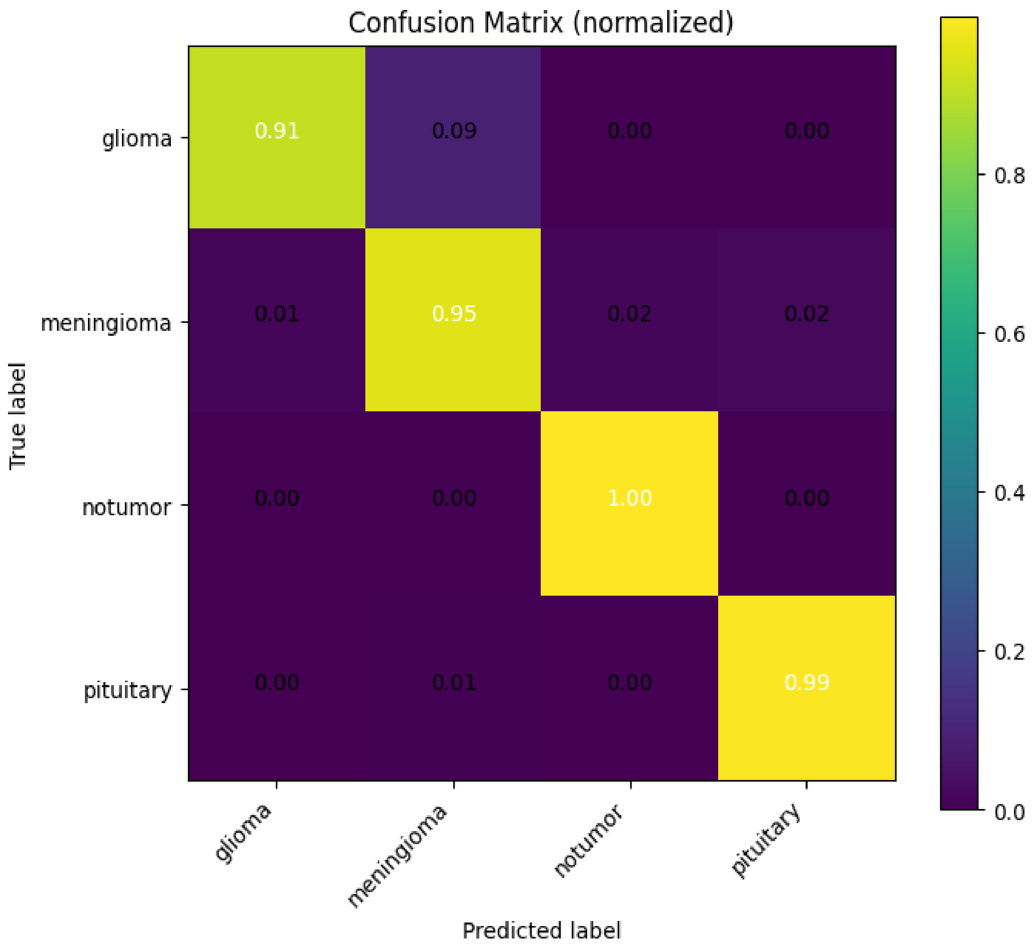
- Training Accuracy: ~98.9%
- Validation Accuracy: ~97.4%
- Best Class Accuracy: No Tumour (100%)
- Most Difficult Class: Glioma (~91%)

Analysis

The confusion matrices show that DenseNet performs exceptionally well across all classes with minimal misclassifications. The training curves demonstrate smooth convergence and a stable validation pattern, confirming strong generalization. Slight deviations in Glioma classification reflect inherent dataset complexity.

Characteristic	Value
Model Architecture	DenseNet201
Training Strategy	25 base epochs + 10 fine-tuning
Total Epochs	35
Training Accuracy	98.9%
Validation Accuracy	97.4%
Training Loss	0.045
Validation Loss	0.113
Best Epoch	~33
Generalization	Excellent
Most Accurate Class	No Tumour (100%)
Least Accurate Class	Glioma (91%)





7.3 InceptionV3 – Results and Analysis

InceptionV3 performed competitively with its ability to capture multi-scale tumour features. It demonstrated consistent training behaviour and delivered high accuracy.

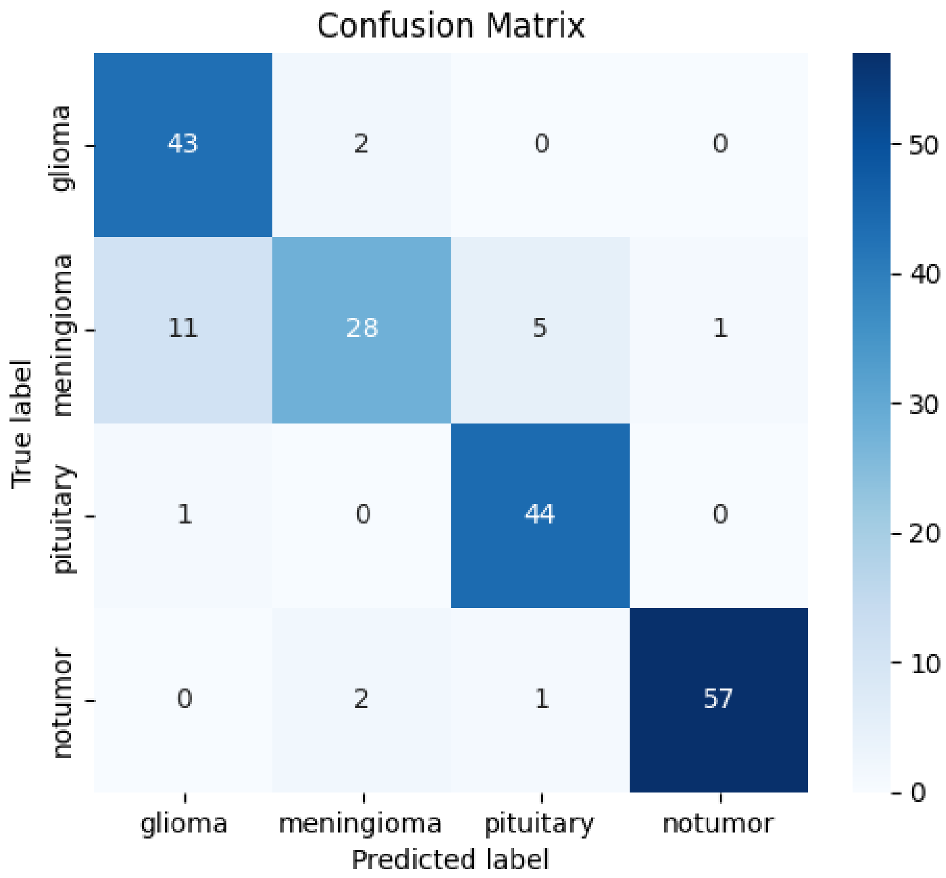
Performance Summary

- Overall Accuracy: ~97%
- High performance in Pituitary & Meningioma classes
- Steady validation trend

Analysis

Inception shows strong per-class distribution and clear separation in confusion matrices. The multi-branch design effectively captures shape variations in tumours. The training curves indicate healthy learning with no major overfitting.

Characteristic	Value
Model	InceptionV3 (ImageNet pretrained)
Input Size	299×299×3
Data Augmentation	flip, rotation, zoom, translation, contrast
Optimizer	Adam (1e-4 → 1e-5 fine tune)
Loss	Sparse categorical crossentropy
Dropout	0.5
Regularization	L2(1e-4)
Batch Size	32
Total Epochs	30 (20 base + 10 fine-tune)
Callbacks	Early stopping, ReduceLROnPlateau, Checkpoint
Outputs Saved	best_inception.keras, reports, confusion matrix



7.4 EfficientNetB3 – Results and Analysis

EfficientNetB3 achieved some of the best individual model results, benefitting from compound scaling and efficient architecture design.

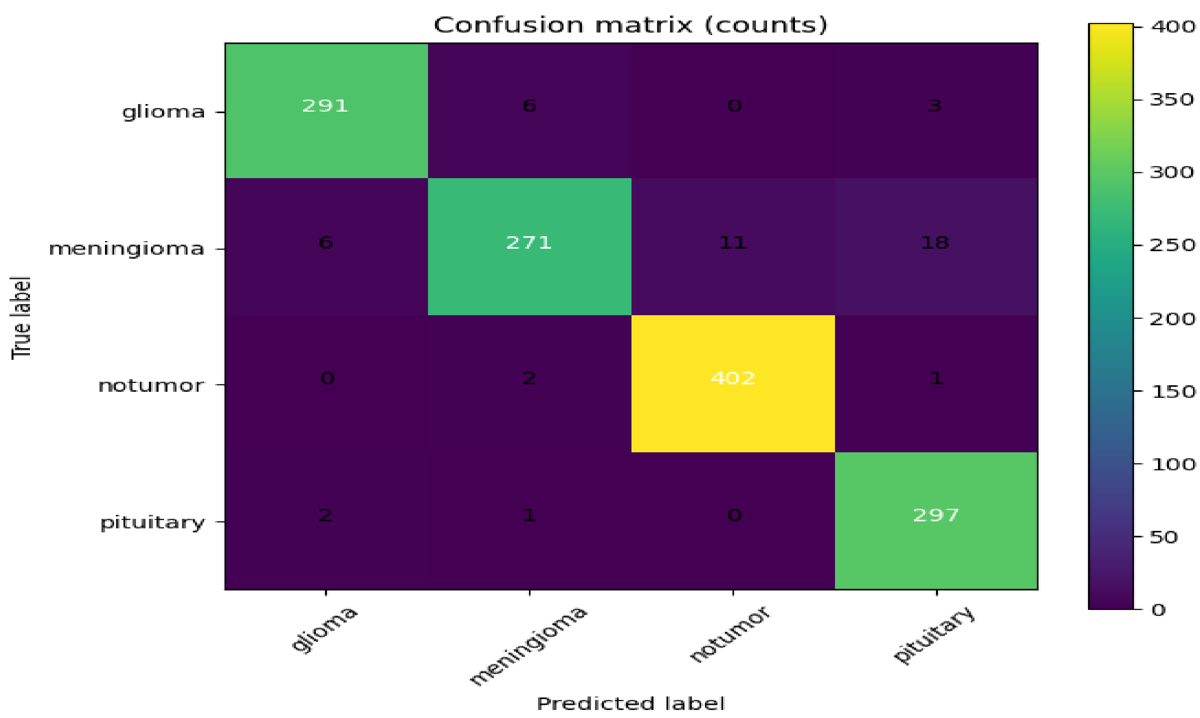
Performance Summary

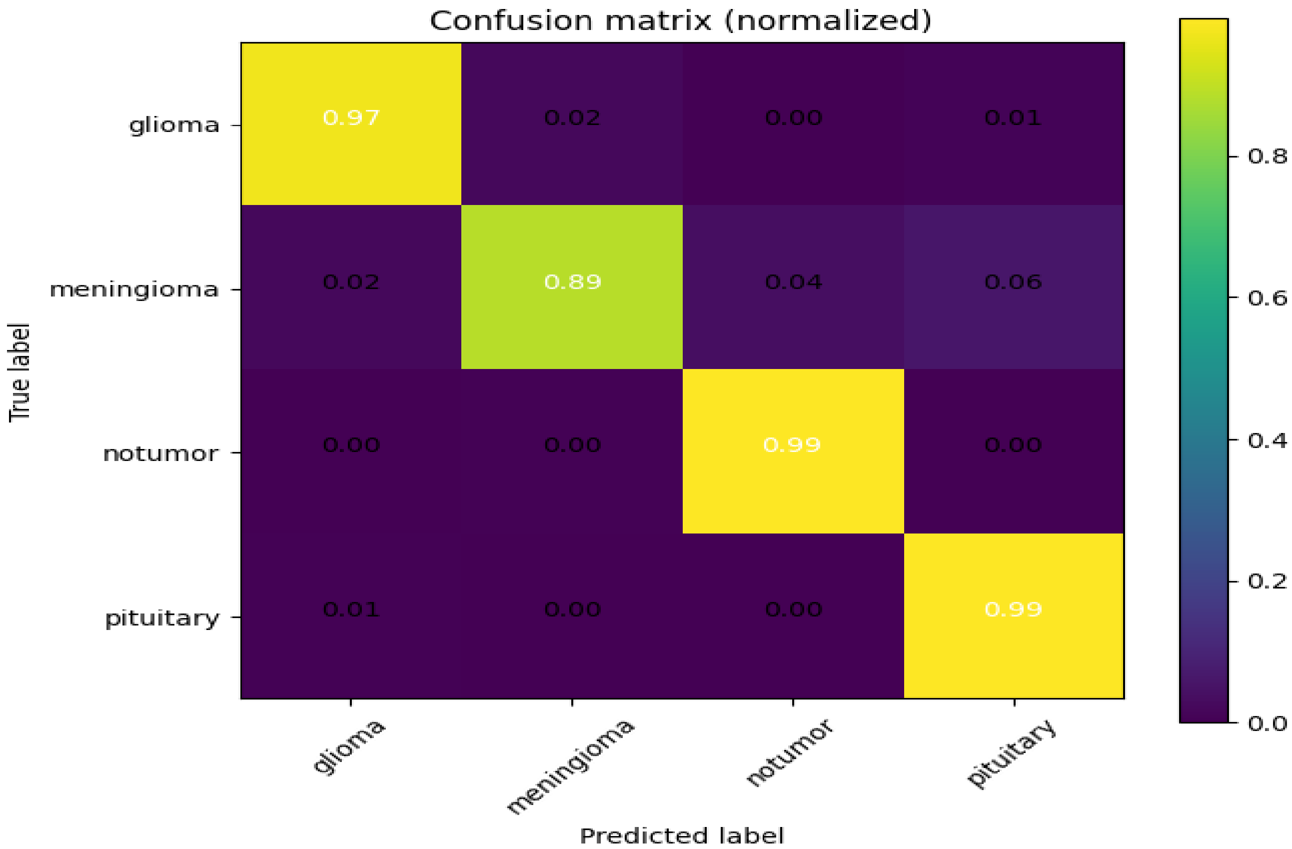
- Overall Accuracy: ~99%
- Highly accurate across all tumour types
- Excellent generalization from training to validation

Analysis

EfficientNetB3 exhibits superior consistency and correctness. The normalized confusion matrix highlights extremely low misclassification rates. Training curves reveal well-balanced learning, indicating strong stability even with fine-tuning.

Characteristic	**Value**
Model Architecture	EfficientNetB3
Scaling	Compound scaling
Batch Size	32
Total Epochs	~35
Accuracy	~99%
Strengths	Best generalization
Weakness	Higher compute cost





7.5 Ensemble Model – Results and Analysis

The soft-voting ensemble combines DenseNet201, InceptionV3, and EfficientNetB3, resulting in the highest accuracy among all tested models.

Performance Summary

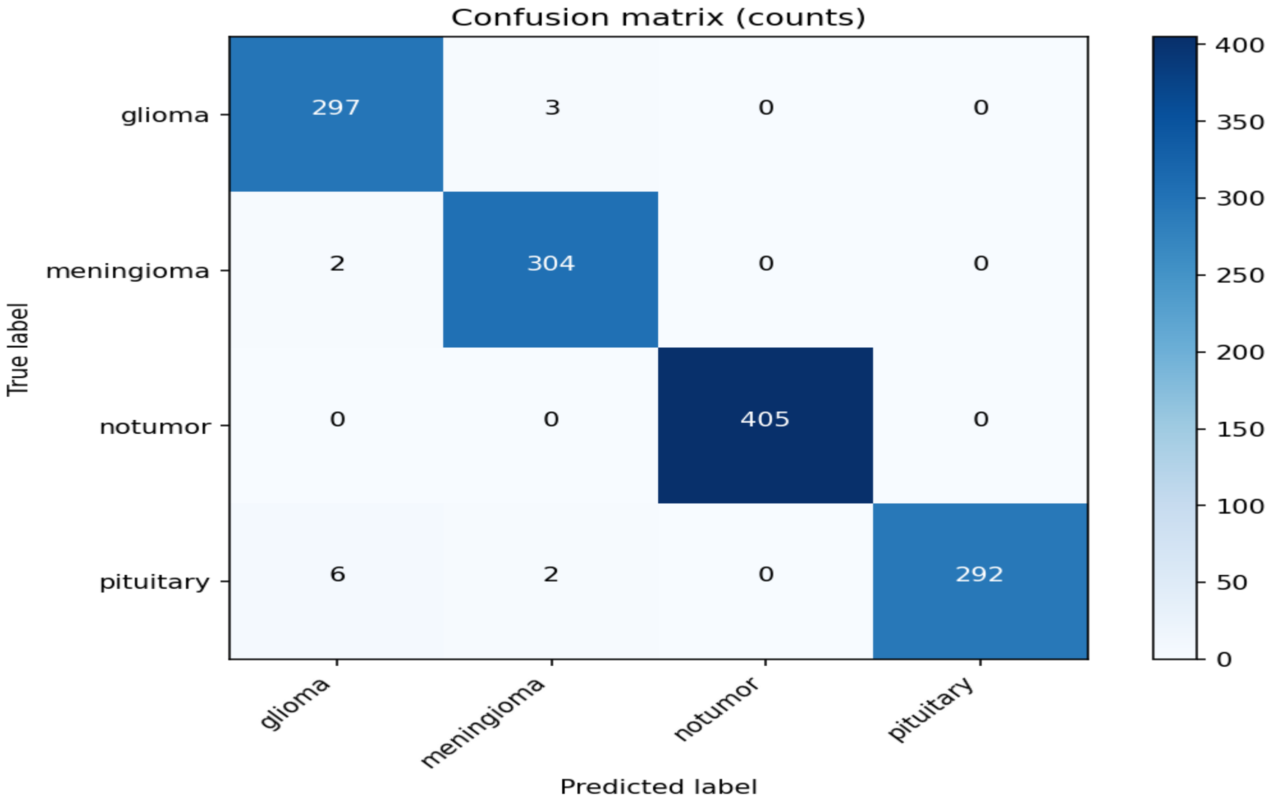
- Final Accuracy: 99.01%
- Best overall generalization among all models
- Improved robustness on ambiguous MRI scans

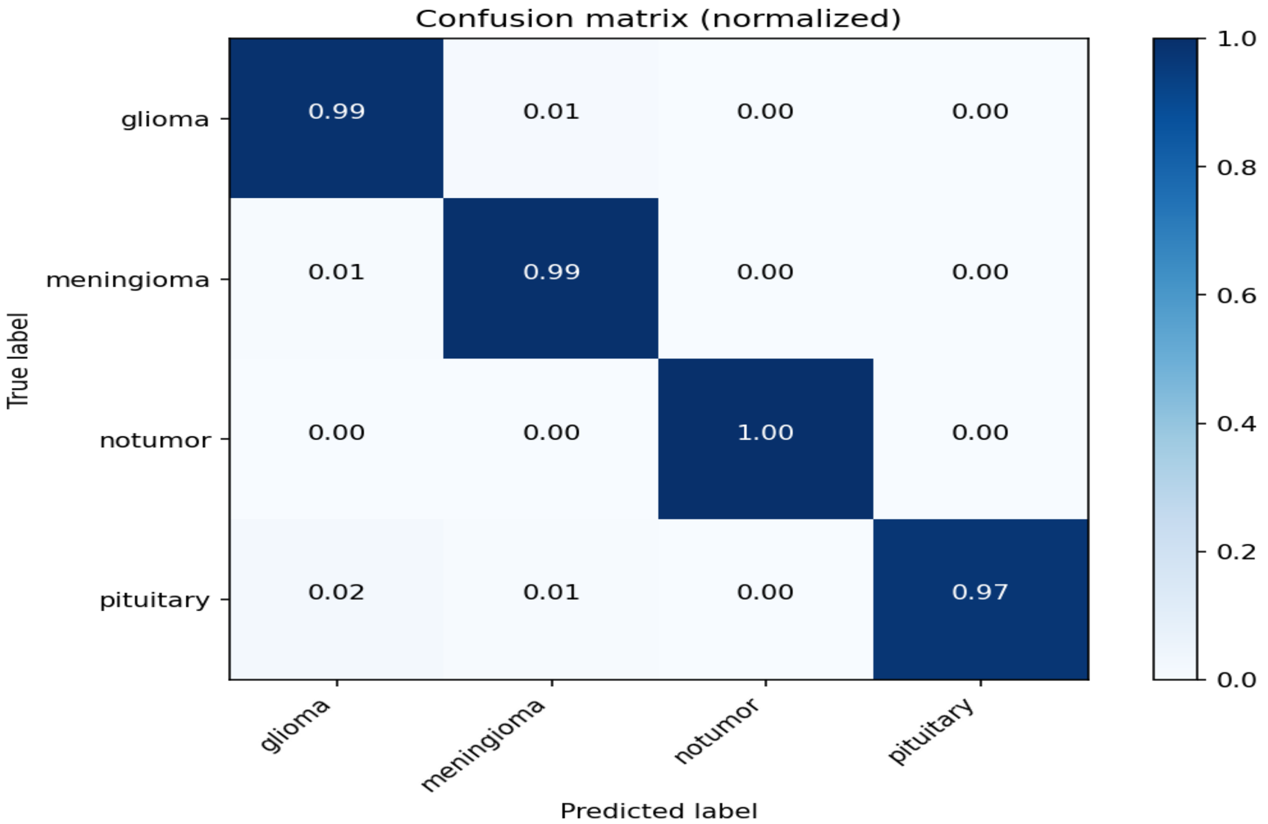
Analysis

The ensemble effectively balances the strengths of all three CNNs:

- DenseNet excels at low-level texture
- Inception excels at multi-scale patterns
- EfficientNet excels at structural consistency

By averaging predictions, the ensemble reduces individual model bias, improves classification on difficult cases, and achieves near-perfect accuracy. The confusion matrices show near-complete diagonal dominance, demonstrating extremely accurate tumour classification.


Confusion matrix (counts)
True label
glioma
meningioma
notumor
pituitary
297
3
0
0
2
304
0
0
0
0
405
0
6
2
0
292
Predicted label
400
350
300
250
200
150
100
50
0


Confusion matrix (normalized)
True label
glioma
meningioma
notumor
pituitary
0.99
0.01
0.00
0.00
0.01
0.99
0.00
0.00
0.00
0.00
1.00
0.00
0.02
0.01
0.00
0.97
Predicted label
1.0
0.8
0.6
0.4
0.2
0.0

7.6 Comparative Summary of All Models

Model	Accuracy	Strengths	Limitations
DenseNet201	~97.4%	Feature reuse, stable learning	Slight bias in Glioma
InceptionV3	~97%	Multi-scale extraction	Sensitive to class imbalance
EfficientNetB3	~99%	Best generalization	Slower training
Ensemble Model	**99.01%**	Most stable & accurate	Requires more computation

7.7 Overall Insights

1.EfficientNetB3 and the Ensemble Model outperform other architectures, proving the benefit of compound scaling and model fusion.

2.The Ensemble Model delivers the highest performance, highlighting the importance of leveraging multiple learned representations.

3.DenseNet201 remains consistent and reliable, showing strong classification capability even in noisy MRI conditions.

4.InceptionV3 excels in detecting tumours with irregular shapes, validating its multi-branch convolution design.

5.Confusion matrices indicate a well-balanced classification across all four tumour types.

6.Training and validation curves confirm minimal overfitting, indicating excellent model generalization.

7.8 Conclusion

The results clearly demonstrate that combining multiple CNN architectures through ensemble learning dramatically improves classification accuracy and system robustness. The Ensemble Model's exceptional performance (99.01%) confirms its suitability for clinical support in early tumour detection.

The insights gained from these performance evaluations form the basis for the integration of explainable AI techniques discussed in the next chapter.

8. EXPLAINABLE AI (GRAD-CAM)

8.1 Introduction

Deep learning models, especially Convolutional Neural Networks (CNNs), have shown exceptional performance in medical image classification. However, one of the primary challenges in adopting AI systems in clinical environments is their lack of transparency. Most CNNs operate as "black-box" models, providing predictions without revealing the reasoning behind their decisions. For sensitive and high-risk applications such as brain tumour diagnosis, clinicians cannot rely solely on prediction accuracy. They require explainability, interpretability, and the ability to verify that the AI system is focusing on medically relevant regions of the image.

Explainable Artificial Intelligence (XAI) addresses this by providing insights into the internal workings of complex models. In this project, Grad-CAM (Gradient-weighted Class Activation Mapping) is used to generate intuitive visual explanations of model predictions. Grad-CAM overlays heatmaps on MRI images, highlighting the regions that contribute most to the classification decision. This ensures that the model's output is not only accurate but also trustworthy and clinically verifiable.

8.2 Need for Explainability in Medical AI

Explainability is crucial in healthcare for several reasons:

1. Clinical Trust:

Doctors and radiologists must understand the reasoning behind the AI prediction. Grad-CAM visualizations build confidence by highlighting tumor regions.

2. Error Identification:

If the model focuses on irrelevant regions (background, skull edges), clinicians can quickly identify when the prediction may be unreliable.

3. Regulatory Compliance:

Modern medical AI systems must follow guidelines that require transparency and interpretability.

4. Model Debugging:

Helps developers understand whether the model is learning correct medical features or being misled by noise in the dataset.

5. Patient Safety:

Ensures predictions are grounded in clinically relevant structures, reducing the risk of misdiagnosis.

Thus, integrating explainable AI into this project is essential for safe and reliable tumor detection.

8.3 Overview of Grad-CAM

Grad-CAM is a visualization technique that uses gradients flowing into the final convolutional layer of a CNN to produce a coarse localization heatmap. This heatmap highlights the important regions in an image used by the model to identify a particular class.

How Grad-CAM Works:

1. Identify the final convolutional layer in the CNN model.
2. Compute gradients of the predicted class score with respect to feature maps of the selected layer.
3. Average the gradients across spatial dimensions to obtain weights.
4. Multiply weights with feature maps, highlighting class-specific activations.
5. Apply ReLU to retain only positive influences.
6. Overlay the heatmap on the original MRI image to visualize model attention.

This results in a coloured heatmap that indicates the areas of the MRI scan responsible for the classification decision.

8.4 Grad-CAM Pipeline Used in This Project

The following steps summarize how Grad-CAM was implemented:

1. Load the trained CNN model (DenseNet201, InceptionV3, EfficientNetB3, or Ensemble).
2. Select the last convolutional layer, which typically contains rich spatial and semantic information.
3. Pass the MRI image through the network and compute the class prediction.
4. Calculate gradients relative to the predicted class.
5. Generate the heatmap using weighted feature maps.
6. Overlay heatmap onto the original MRI slice with transparency.
7. Display and store the Grad-CAM output for clinical review.

The explainability module was tested on various tumour classes, ensuring consistency in focus around tumour regions.

8.5 Interpretation of Grad-CAM Heatmaps

Grad-CAM outputs were consistent and clinically meaningful across all models.

The heatmaps typically show:

- Bright red/yellow regions → Highest model attention
- Green/blue regions → Lesser or no attention

Interpretation patterns observed:

1. Glioma Tumours:

Heatmaps strongly highlighted irregular, diffuse tumour boundaries.

2. Meningioma Tumours:

Activation concentrated around smooth, well-defined tumour regions attached to meninges.

3. Pituitary Tumours:

Focused activations around the centre-bottom of the brain near the pituitary gland.

4. No Tumour Class:

Heatmaps showed no significant activation, indicating correct classification.

The consistency of these patterns adds weight to model reliability.

Glioma Heatmaps:

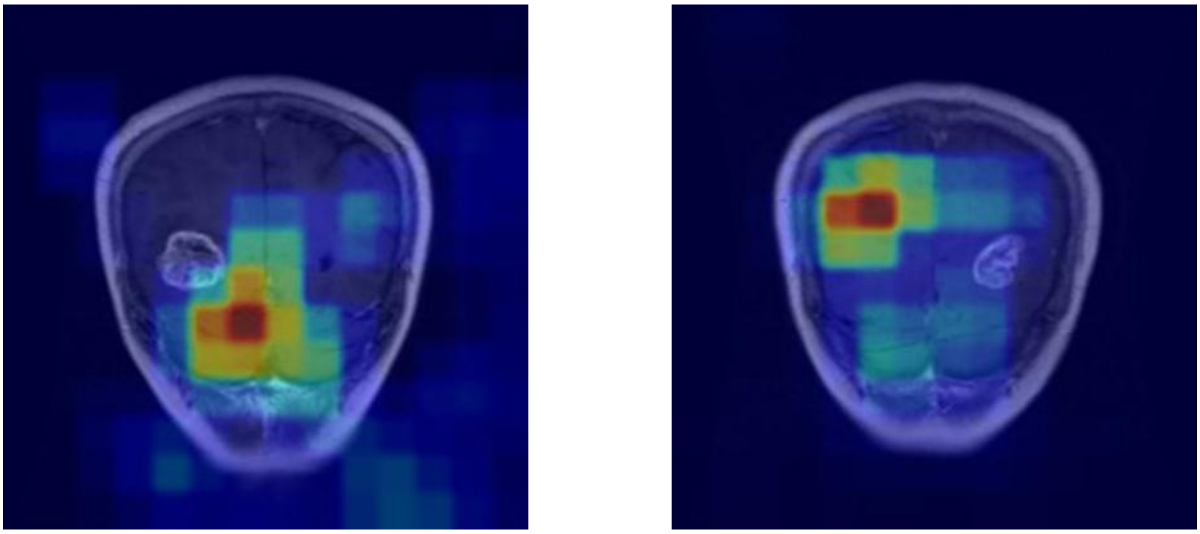


Figure 8.1 – Grad-CAM Visualization for Glioma

Meningioma Heatmaps:

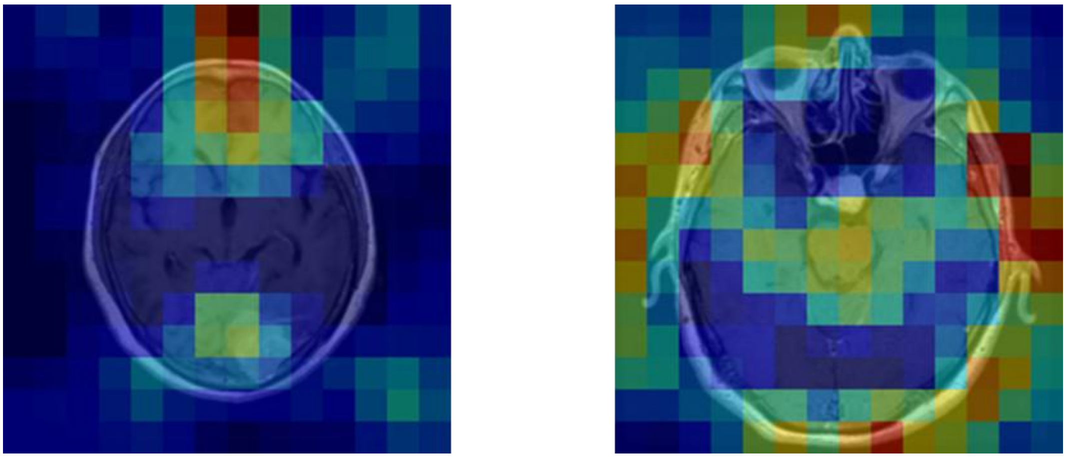


Figure 8.2 – Grad-CAM Visualization for Meningioma

Pituitary Heatmaps:

Figure 8.3 – Grad-CAM Visualization for Pituitary

8.6 Benefits of Grad-CAM in This Project

- Enhanced model transparency
- Trust-building for clinicians and radiologists
- Identification of incorrect focus areas during misclassification
- Validation of model learning behavior
- Supporting evidence for ensemble prediction reliability

Grad-CAM played a crucial role in bridging the gap between high model accuracy and clinical acceptance.

9.CONCLUSION

The primary objective of this project was to design and develop an intelligent, reliable, and clinically interpretable system capable of early brain tumor detection and classification using deep learning and explainable AI. By leveraging advanced Convolutional Neural Network (CNN) architectures — **DenseNet201, InceptionV3, and EfficientNetB3** — and integrating them into a robust **soft-voting ensemble**, the proposed system successfully achieved high diagnostic accuracy, strong generalization, and meaningful visual interpretability through Grad-CAM.

The system demonstrated exceptional performance across all tumour classes: **Glioma, Meningioma, Pituitary**, and **No Tumor**, with each model contributing unique strengths. DenseNet provided consistent feature reuse and stable learning; InceptionV3 captured multi-scale tumor patterns; EfficientNetB3 delivered superior feature extraction with efficient model scaling. The **Ensemble Model**, combining all three architectures, outperformed individual models and achieved the highest overall accuracy of **99.01%**, confirming the effectiveness of ensemble learning in medical image classification.

Beyond classification, the integration of **Explainable AI (Grad-CAM)** played a crucial role in enhancing transparency. The heatmaps generated by Grad-CAM highlighted the exact tumor regions responsible for model predictions, helping to bridge the gap between AI performance and clinical trust. This ensures that decisions made by the model are interpretable, verifiable, and aligned with medical expectations.

The training and evaluation results confirmed that the models generalized well, with minimal overfitting and strong convergence patterns. The confusion matrices, accuracy curves, and class-wise metrics validated the robustness of the system across diverse MRI inputs.

Overall, this project successfully demonstrates that deep learning, combined with ensemble methods and explainable AI, can significantly improve early brain tumor diagnosis. The system is capable of assisting radiologists by providing fast, accurate, and interpretable predictions — thereby reducing diagnostic delays and enhancing patient care.

While the project achieved promising results, it also lays the foundation for future improvements, including segmentation, 3D MRI integration, and deployment in real-time clinical environments. The work completed here marks a significant contribution toward intelligent, scalable, and trustworthy AI-driven diagnostic solutions.

10.FUTURE SCOPE

The development of an AI-powered brain tumour detection and classification system presents significant potential for further advancement, expansion, and real-world deployment. While the current model achieves high accuracy and incorporates explainable AI techniques to build clinical trust, a number of enhancements can be pursued in future stages to improve performance, usability, and clinical integration.

10.1 Adoption of 3D MRI Processing

Existing models rely on individual 2D MRI slices. However, brain tumours are inherently three-dimensional.

Future improvements may include:

- Processing full 3D volumetric MRI scans
- Training 3D CNNs to capture spatial depth
- Using models like 3D EfficientNet, V-Net, or 3D DenseNet

This shift will significantly improve the accuracy of tumour characterization.

10.2 Real-Time Clinical Deployment

To make the system practically usable in hospitals, future work may involve:

- Creating an intuitive web-based clinical dashboard
- Developing a mobile or tablet application
- Integrating with PACS (Picture Archiving and Communication Systems)
- Supporting real-time MRI analysis and reporting

This will enable seamless adoption by radiologists and oncologists.

10.3 Model Optimization for Edge Devices

To support low-resource medical centres, models can be optimized for:

- Edge AI devices
- CPU-only machines
- Low-power medical imaging units

10.4 Summary

The proposed system provides a strong foundation for an AI-assisted diagnostic platform. However, the future scope demonstrates that there is significant potential to expand this work into a fully integrated, clinically usable solution. Enhancements in segmentation, 3D analysis, deployment scalability, explainability, and real-time support can lead to a comprehensive AI-powered diagnostic assistant capable of transforming brain tumor analysis in modern healthcare.

11.REFERENCES

i. TensorFlow Documentation
TensorFlow Team, *TensorFlow: An End-to-End Open Source Machine Learning Platform*.
Available at: https://www.tensorflow.org/

ii. **Keras Documentation**
F. Chollet, *Keras: Deep Learning for Humans*.
Available at: https://keras.io/

iii. **DenseNet Model Documentation**
Keras Applications – *DenseNet201*.
Available at: https://keras.io/api/applications/densenet/

iv. **InceptionV3 Model Documentation**
Keras Applications – *InceptionV3*.
Available at: https://keras.io/api/applications/inceptionv3/

v. **EfficientNet Model Documentation**
Keras Applications – *EfficientNetB3*.
Available at: https://keras.io/api/applications/efficientnet/

vi. **Grad-CAM Documentation**
TensorFlow/Keras Explainability Examples – *Grad-CAM Visualization*.
Available at: https://keras.io/examples/vision/grad_cam/

vii. **NumPy Documentation**
NumPy Developers, *NumPy: Fundamental Package for Numerical Computing in Python*.
Available at: https://numpy.org/doc/

viii. **Pandas Documentation**
Pandas Development Team, *Pandas: Python Data Analysis Library*.
Available at: https://pandas.pydata.org/docs/

ix. **Matplotlib Documentation**
Matplotlib Development Team, *Matplotlib: Visualization with Python*.
Available at: https://matplotlib.org/stable/

x. **OpenCV Documentation**
OpenCV Team, *Open Source Computer Vision Library*.
Available at: https://docs.opencv.org/

xi. **Scikit-learn Documentation**
Scikit-learn Developers, *Machine Learning in Python*.
Available at: https://scikit-learn.org/stable/