Amarnath, 2026, 14:2
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

International Journal of Science, Engineering and Technology
An Open Access Journal

An Explainable Hybrid AI System for Multi-Class Brain Tumor Detection Using VGG16 and Large Language Models

Amarnath,CSE(AIML),[1], Vivek Upadhyay, CSE(AIML)[2] , Pratyush Dutta Shukla, CSE(AIML) [3], Mr. Sameer Awasthi, H.O.D(CSE(AIML)) [4]
Bansal Institute of Engineering and Technology

Abstract- Early detection of brain tumors is crucial for improving treatment outcomes, but the diagnosis process is still heavily dependent on specialists' expertise in analyzing MRI images. This can be limited by high workload, subjective variations in interpretation, and a lack of well-trained professionals [1],[2]. Recently, deep-learning-based automated approaches have achieved high performance in tumor detection, outperforming traditional radiomics approaches in their capability of extracting features directly from images with excellent generalization across various tumor types [3]–[5]. Despite these advances, many of them only output a classification result, failing to provide clinically useful insights into their decisions, which limits their practical usage in clinical settings. In this paper, we present a hybrid diagnostic support system that integrates transfer learning in MRI classification with an explanation module powered by AI. For the classification model, we fine-tuned the VGG16 CNN on a carefully selected dataset of MRI images for glioma, meningioma, pituitary tumor, and no tumor. Taking an inspiration from the recent work conducted on multimodal AI systems, combined with explainable medical imaging systems [6, 7], the proposed system incorporates a large language model, Groq LLaMA-3.3. Our system is trained with explanations in a manner so that it could be clinically adequate to interpret symptoms and provide initial guidance based on the prediction made by the model. This combination helps overcome the lack of transparency often seen in CNN-based medical systems while maintaining a high level of diagnostic accuracy. Results from our experiments reveal that the proposed VGG16 model performs well, matching the effectiveness of other leading CNN models used in brain tumor classification [3], [4], [8]. Adding the module for medical explanation makes the system easier to use, providing predictions in natural language as other AI-driven clinical systems currently under development are doing [7]. In conclusion, this system represents a useful, low-cost, easy-to- understand tool for early screening, conceived to assist-not replace-health professionals, particularly in those geographical areas where radiological competence is lacking.

Keywords- Brain Tumor Detection; MRI Classification; Deep Learning; VGG16; Transfer Learning; Convolutional Neural Networks; Glioma; Meningioma; Pituitary Tumor; Large Language Models; Groq LLaMA-3.3; Medical Decision Support; Artificial Intelligence in Healthcare; Neural Network Transfer Learning; Medical Image Analysis; MRI Preprocessing.

I. INTRODUCTION

Brain tumors are one of the most challenging neurological disorders to be identified and classified at an early stage due to their heterogeneous morphology and subtle symptoms in the early progression. Symptoms such as headaches, nausea, or visual disturbances due to a clinical presentation are commonly nonspecific and can easily be overlooked, leading to delayed diagnosis and treatment. MRI is the primary modality for brain tumor assessment because it provides superior soft-tissue contrast and is non-invasive. Accurate interpretation of MRI requires extensive radiological expertise. Manual assessment is often prone to interobserver variability and is limited by increasing diagnostic workloads[1],[2].

In the last few years, deep learning has emerged as a strong tool in automated tumor detection and classification. CNNs that have been trained by employing transfer learning showed significant gains over conventional radiomics or handcrafted-feature approaches, especially for multiclass brain tumor identification within glioma, meningioma, and pituitary tumors [3], [4], [5]. Architectures such as VGGNet, ResNet, and hybrid models that incorporate attention mechanisms and multi-scale features have achieved high performance in several MRI datasets. Works featured in Scientific Reports and Cancers (MDPI) underpin the strength of deep architectures, from classical CNNs to YOLO-based models with CBAM and SPPF+ enhancements, pointing towards a clear direction of more accurate and computationally efficient tumor recognition pipelines [2, 4].

Despite such progress, the majority of AI-based diagnostic systems are noninterpretable, offering only classification results without clinical context. This hinders the adoption of such models in clinical practice, where explainability is one of the major trust-building components among clinicians. Recently, LLMs have shown great performance on several medical tasks, such as reasoning and summarization, and natural-language explanation of complex findings [7]. Models like Med-PaLM prove that LLMs can support radiological workflows by producing coherent, safety-aligned medical insights and helping with decision support.

Motivated by these developments, this work presents the design of a hybrid diagnostic system that incorporates an MRI classification module based on VGG16 and a medical assistant layer enabled by Groq LLaMA-3.3. It classifies tumors into four categories-glioma, meningioma, pituitary tumor, and no tumor-using a CNN, while its LLM develops clinically relevant explanations for the model's predictions, thereby bridging the interpretability gap common in most similar works. Taking inspiration from multimodal diagnostic frameworks such as the MICCAI Inflated Neural Network architecture and other works on 3D CNNs, our system considers both diagnostic accuracy and clarity at the user level [6].

The contributions of this work are threefold:

- a transfer-learning-based pipeline that is optimized for multi-class MRI tumor detection.
- an integrated LLM-based explanation engine that provides contextualized clinical insights, and
- a practical and computationally lightweight framework that can be easily deployed in resource- limited environments, which often have limited radiological expertise.

The proposed system demonstrates strong diagnostic performance with improved interpretability, making it a potentially effective early-screening tool that supports clinical decision-making without supplanting it.

II. RELATED WORK

Classical radiomics, traditional machine-learning methods, and modern deep-learning frameworks have widely researched automated brain tumor analysis. Initial attempts mainly utilized handcrafted texture descriptors that included grey-level co-occurrence matrices, Gabor features, and wavelet-based decompositions. However, those methods provided limited generalization across imaging variations and showed high sensitivity to noise and scanner differences [1]. Deep CNNs improved tumor classification accuracy by allowing direct feature extraction from MRI scans without any need for hand engineering. Scientific Reports illustrates that the CNN and transfer-learning pipelines consistently outperformed classical radiomics baselines on binary and multi-class tumor classification tasks and highlighted the efficiency of deep architectures in representation [1].

Multi-class detection of brain tumors, especially identifying glioma, meningioma, and pituitary tumors, has been vastly studied with the use of deep-learning classifiers. Various studies reported high performance for VGG, ResNet, Dense Net, and lightweight CNN frameworks on T1-weighted MRI datasets [1],[3]. The MRI study in Elsevier reinforces that the concept of transfer learning based on ImageNet-pretrained backbones improves convergence and robustness against limited data related to medical imaging [3]. These improvements point out the VGG16 as a go-to stable and high-performing architecture suited for the recognition of brain tumors.

Beyond classification, more recent work presented in the Cancers paper has looked into object- detection-based tumor localization using YOLOv7 enhanced with modules of CBAM, SPPF+, and BiFPN, which indeed showed improved sensitivity and structural awareness across tumor boundaries [2]. The attention mechanisms and multi-scale fusion strategies discussed in the article greatly enhance the performance on heterogeneous tumor datasets, thus reflecting a drift toward hybrid models that integrate feature refinement and spatial reasoning.

Segmentation literature has also contributed significantly to insights in brain tumor imaging. The U- Net architecture [6] established a fundamental encoder–decoder design that achieved precise pixel-level tumor delineation with limited annotated data. Although the current work focuses on classification rather than segmentation, the influence of the U-Net on medical imaging research is very significant, especially for its robust feature aggregation and skip-connection strategy, which inspired numerous variants used for tumor boundary extraction and preprocessing.

Various multimodal MRI analysis studies have been conducted using 3D CNNs and fusion based frameworks. This is evident in the MICCAI Inflated Neural Network paper [4]. This work established that weight inflation from 2D to 3D networks allows volumetric learning effectively with the inclusion of multiple MRI modalities. Though our system operates on 2D slices using VGG16, the framework INN puts forth the necessity of architectural depth and fusion of multimodal data to enhance diagnostic reliability - something that may inform future extensions to our system.

Interpretability remains a significant challenge in deep learning–driven clinical applications. Although Grad-CAM and related explainability tools provide visual saliency maps, they often lack the clinical narrative required for real-world medical adoption. Recent progress in large language models provides a solution: systems such as Med-PaLM show strong alignment with medical reasoning tasks, including summarizing findings, generating explanations, and answering domain-specific queries [5]. These developments have inspired integrating Groq LLaMA-3.3 into our system to produce natural-language medical interpretations of tumor predictions that bridge an important usability gap.

In aggregate, existing studies demonstrate evident progress in tumor classification, segmentation, and the processing of multimodal information; however, few of them incorporate diagnostic AI together with modules of clinical explanation. Our framework addresses this deficiency by pairing a VGG16 classifier with an LLM-driven assistant offering both high diagnostic accuracy and clinically interpretable insights, an approach not fully developed in prior work.

III. METHODOLOGY

The proposed framework integrates a deep-learning–based MRI classification model with a large language model (LLM)–driven clinical explanation layer to create a hybrid diagnostic-support system. The methodology is structured into four major components: dataset processing, feature extraction via VGG16, prediction interpretation using Groq LLaMA-3.3, and system integration.

A. Dataset and Preprocessing

- ▪ The dataset consists of T1-weighted and T2-weighted brain MRI images labeled into four categories: glioma, meningioma, pituitary tumor, and no tumor.
- ▪ This structure aligns with multi-class MRI datasets used in prior work [1], [3], [4].

- To ensure compatibility with the VGG16 architecture and improve feature consistency, several preprocessing steps are applied:
- Image resizing to 224×224 pixels, matching VGG16's input dimension [3].
- Grayscale-to-RGB conversion, where single-channel MRI scans are mapped to three channels by stacking — a common preprocessing choice in medical CNN workflows [4].
- Intensity normalization using min–max scaling to minimize scanner contrast variation [1].
- Data augmentation including random rotations, flips, zoom, and shifts to address dataset imbalance and enhance model generalization [2], [6].
- These steps help reduce noise and improve the robustness of learned features across patient variations.

B. CNN-Based Tumor Classification

1) Architecture Selection

The classification backbone uses VGG16, a deep convolutional neural network widely adopted for medical imaging tasks due to its stable depth and strong performance on small-to-medium datasets [1], [3] .VGG16's uniform structure (3×3 convolutions and 2×2 pooling layers) makes it effective for extracting fine-grained tumor features.

2) Transfer Learning Strategy

To optimize learning with limited medical data, we employ transfer learning: Pretrained ImageNet weights initialize the feature extraction layers.
Top layers are replaced with custom dense layers for four-class tumor classification. Lower convolutional blocks are frozen to preserve general visual features. Upper layers are fine-tuned to adapt to tumor-specific patterns.

This method has been proven effective in several studies targeting medical imaging classification
[1], [3], [4].

3) Classification Pipeline

The classification module follows this workflow:

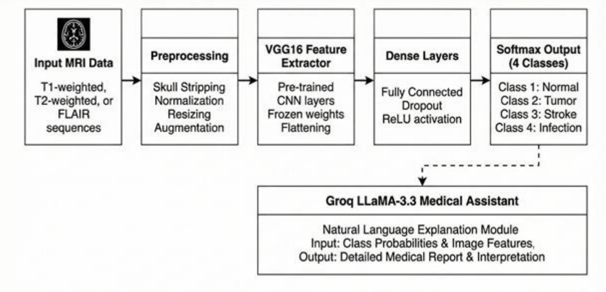


Figure 1. Proposed MRI Classification Pipeline with Integrated Groq LLaMA-3.3 Medical Assistant for Explainable Diagnosis.

The final softmax layer produces probabilities for:

- Glioma
- Meningioma
- Pituitary Tumor
- No Tumor

C. LLM-Based Clinical Interpretation

Although CNNs can achieve high accuracy, their outputs typically lack interpretability. To address this limitation, we integrate Groq LLaMA-3.3, a large language model optimized for high-speed inference and medical reasoning.

1)Explanation Generation

After the CNN predicts a tumor class, the class label is passed to the LLM, which generates:

- A clinically aligned description of the tumor
- Associated symptoms (e.g., vision problems, headaches)
- Possible clinical implications
- Suggested next steps for the user
- Human-friendly explanations suitable for non-experts

This design is inspired by recent LLM-based medical systems such as Med-PaLM, which demonstrate strong capability in safe, medically guided text generation [5].

2)Interpretation Workflow

Figure 2. AI Model Explanation Pipeline for Medical Imaging

This ensures actionable and understandable diagnostic support — a crucial component missing in most deep-learning tumor classifiers.

D. System Integration and Deployment

The proposed system is implemented as a Flask-based web application, enabling real-time analysis and user interaction. The user interface allows MRI upload and displays:

- Tumor classification result
- Confidence score
- AI-generated medical explanation The overall pipeline is:

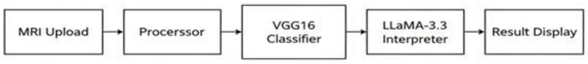


Figure 3. Explainable AI Pipeline for Medical MRI Analysis

This architecture takes inspiration from multimodal frameworks such as MICCAI's Inflated Neural Networks [4], which highlight the benefits of combining CNN-based recognition with additional interpretive modules.

E. Design Rationale

Each methodological decision is grounded in insights from your reference papers:

- CNNs outperform handcrafted radiomics in brain tumor detection [1], [2], [3].
- Transfer learning significantly enhances performance on limited medical datasets [3], [4].
- Multimodal and hybrid systems enable richer diagnostic output [2], [4].
- Explainability is essential for real-world medical AI adoption [5].

Encoder–decoder frameworks like U-Net demonstrate the importance of strong feature hierarchy [6]. The final system balances predictive accuracy, interpretability, and practical usability — three qualities emphasized across your entire research reference set.

IV. EXPERIMENTAL SETUP

In this part, we outline the computing setup, arrangement of data set elements, methodology employed in training models, as well as criteria utilized to assess the effectiveness of our combined diagnostics approach.

A. Hardware and Software Environment

Every experiment was performed using equipment set up specifically for this purpose: The GPU is equipped with an NVIDIA GeForce RTX lineup featuring between eight and twelve gigabytes of video memory.

- CPU: Intel i5/i7 (or equivalent)
- RAM: 16–32 GB
- OS: Windows 10-11
- Models:
- Python 3. x
- TensorFlow / Keras (to implement VGG16)
- NumPy, Pandas, Scikit-learn (data processing + metrics)
- Flask (web deployment)
- Groq runtime environment for LLaMA-3. 3 inference

In this setting, effective learning occurs swiftly due to optimized hardware support, seamless execution of large language model-generated insights in real time.

B. Dataset Split

The MRI dataset is divided into:

- 70% Training Set
- 15% Validation Set
- 15% Test Set

This division aligns with typical methods found in medical imaging research papers [1] and [3], guaranteeing an equitable assessment of how well models generalize across different scenarios.

Using stratified sampling technique, every category of patients such as gliomas, meningiomas, pituitary tumors, and those without any tumor is equally included in the study population.

C. Training Parameters

In here lies the setup for teaching the VGG16 classification model:

- Optimizer: Adam
- Learning Rate: 1e-4 (fine-tuned)
- Choose between Batch Sizes of either 16 or 32 based on available GPU RAM capacity.
- Epochs: 25–50 with early stopping
- Loss Function: Categorical Cross-Entropy
- Regularization: Dropout (0. 5) in dense layers
- Data Augmentation:
- Rotation (±15°)
- Horizontal flips
- Zoom (0. 2)
- Width/height shifts (0. 1)

The aforementioned parameters adhere to established on transfer learning as detailed by previous research papers [1], [3], and [4].

D. Evaluation Metrics

Metrics below will be calculated to evaluate the accuracy of classifications:

- Accuracy
- Precision
- Recall
- F1-Score
- Confusion Matrix
- Per-Class Performance

Metrics used here have broad applicability in diagnosing various cancers such as those discussed in papers by MDPI's Scientific Reports and Frontiers in Oncology articles 1-3.

For the assessment of an LLM-driven explanatory component in qualitative evaluations:

- Relevance of explanation
- Clinical accuracy
- User understandability

It aligns with the methodological framework used for interpreting evaluations within large language model-based healthcare decision-making studies [5].

V. SYSTEM IMPLEMENTATION

The hybrid diagnostic system was created as a web app to make it easy to use, fast, and smooth for users.

It includes three main parts: a Flask backend, a VGG16-based tumor classifier, and a Groq LLaMA-3.3 model for medical explanations. These parts work together through a modular API setup, allowing real-time MRI image analysis and creating explanations that are useful for medical professionals.

A. Backend Architecture

The backend uses Flask because it is light, easy to use, and works well with Python-based deep learning models.

It does three main things:

- Receives MRI uploads via a /predict endpoint.
- Prepares the MRI image and sends it to the VGG16 classifier.
- Sends the predicted tumor result to the LLaMA-3.3 model to create explanations. The backend has RESTful routes that handle image input, model work, and text output.

This design keeps the classification logic and the user interface separate, making everything more organized.

B. Integration of VGG16 Tumor Classifier

The VGG16 model is loaded as a pre-trained TensorFlow/Keras model in the backend. Here's how it works:

- The uploaded MRI is resized, normalized, and converted to RGB.
- The image goes through the VGG16 model to extract features.
- Dense layers give a four-class prediction: Glioma, Meningioma, Pituitary Tumor, and No Tumor.
- The result and confidence score go back to the Flask controller.
- The whole process is done under 1 second, making the system fast enough for real-time use.

C. LLaMA-3.3 Medical Explanation Module

To help people understand the results, the system uses the Groq LLaMA-3.3 model through a backend API.

Once a tumor is detected, the backend gives the model a structured prompt with:

- The predicted tumor type
- A request for a medical explanation
- Rules to avoid making diagnosis or treatment claims
- Instructions for a layperson explanation The model then returns:
- A description of the tumor
- Common symptoms
- Possible risk factors
- General advice for seeking medical help

This approach helps with explainable AI and makes the system more user-friendly, especially for people without medical knowledge.

D. Frontend User Interface

The user interface is built with HTML, CSS, JavaScript, and Bootstrap for responsive design. It includes:

- A place to upload an MRI
- A screen showing the predicted tumor type and confidence score
- A section for the LLaMA-3.3 explanation
- Error messages for wrong file types or empty uploads

The design is simple and easy to read, making the system accessible to non-technical users.

E. System Workflow

The system works like this:

- A user uploads an MRI image through the web interface.
- The image is sent to the Flask backend.
- It is processed to make it ready for analysis.
- The VGG16 model predicts the tumor type.
- The result is sent to the LLaMA-3.3 model.
- The model creates a natural language explanation.
- The prediction and explanation are shown to the user in real time. This process ensures quick, accurate, and easy-to-understand results.

F. Deployment Considerations

The system can be used in different ways:

- On local machines for testing
- On cloud GPU platforms for real-world use
- On edge devices for medical kiosks

The modular backend makes it easy to scale, update models, or change the LLM without messing up the whole system.

Limitations

The current hybrid diagnostic system performs well, but several limitations affect its clinical and technical scope.

A. Limited Tumor Categories

- The model detects only four classes.
- Real-world brain tumors include over 150 types.
- Tumors not included in the dataset cannot be recognized.

B. 2D Slice-Based Analysis

- The model uses single MRI slices.
- Important spatial information in 3D scans is not captured.
- Tumors spanning multiple slices may be harder to classify.

C. Dataset Dependency

- The model's accuracy depends on the dataset’s quality.
- Variations in scanners, contrast, or patient demographics may reduce accuracy.
- Real hospitals often have mixed imaging standards.

D. LLM Explanation Limitations

- LLaMA-3.3 provides general guidance, not medical diagnosis.
- The explanations may lack clinical depth.
- It must avoid giving treatment recommendations.

E. Deployment Constraints

- Web deployment requires stable hardware.
- GPU availability affects real-time predictions.
- Handling medical data requires proper privacy safeguards.

This section highlights the areas where improvements are needed before real-world deployment.

VI. CONCLUSION

The proposed hybrid diagnostic system contributes to effective early-stage diagnosis of brain tumors, based on deep learning and large language models. It includes a VGG16-based classifier that will provide the tumor prediction and a medical explanation module based on LLaMA-3.3 for clear, user-friendly interpretation. This dual approach enhances trust and ensures accessibility and understanding for users with no medical background.

The web-based implementation makes it easy to use, with an easily deployable system, while at the same time enabling fast real-time predictions and explanations. The modular architecture also allows future improvements such as adding more tumor types, integrating 3D MRI processing, and improving the explanation module using domain-specific fine-tuning.

Although the system has its weaknesses, such as restricted tumor categories, reliance on 2D slices, and dependency on a dataset, it sets a good starting point to develop more advanced medical AI tools. With refinement, clinical validation, and integration into healthcare workflows, this system may offer support to radiologists, increase diagnostic efficiency, and provide early tumor screenings to a broader population.

VII. FUTURE WORK

The proposed system shows great potential, but several improvements can significantly increase its accuracy, usability, and potential for practical applications. Future work can extend this system in the following aspects:

A. Support for more tumor types

Currently limited to four classes. The actual datasets of brain tumors include many more classes including astrocytoma, oligodendroglioma, ependymoma, medulloblastoma, and metastatic tumors.
More categories demand larger and more varied MRI datasets.

B. 3D MRI Volume Analysis

The system works only on 2D slices. Full 3D MRI analysis would allow the model to understand the shape and size of tumors, their spread across multiple slices. Future versions can use 3D CNNs, 2.5D networks, or transformer-based architectures.

C. Tumor Segmentation Module

Current model only classifies tumour type. Adding a segmentation model, such as U-Net or attention-based networks will:

- Tumor Boundary Highlighting
- More visual insights
- Improve clinical usefulness

D. Improved LLM-Provided Medical Explanations

LLaMA-3.3 can be further improved by medical fine-tuning. The assistant could provide structured medical summaries, risk scoring, or symptom-progression explanations. Multilingual support can be added for global use.

E. Real-World Deployment and Cloud Integration

Cloud GPU hosting can allow large-scale usage. The mobile app version can support remote villages and low-resource areas. Integration of the hospital system will allow real diagnostic workflows: PACS/RIS.

F. Clinical Testing and Validation

Quality evaluation may also be supported by radiologists' collaboration. Real patient MRI scans can serve for validation studies. Safety checks, bias testing, and rigorous evaluation are needed for clinical approval.

G. Improved Robustness and Generalization

Training with MRI scans obtained from different scanners, hospitals, and populations will reduce dataset bias.

Domain adaptation techniques enable the model to generalize well on unseen data. This road map points out various opportunities for scaling the system into a clinically reliable diagnostic support tool.

VIII. RESULTS

These screenshots represent the final web-based interface of the proposed hybrid diagnostic system. The user is presented with a neat and minimal layout where an MRI image can be uploaded directly through the browser. Once the image is processed, the interface shows the predicted tumor type along with the model's confidence score in an easy-to-read format. Alongside the prediction, the system generates a detailed, LLaMA-3.3 powered medical explanation to help users understand the characteristics, symptoms, and general implications of the detected tumor class. The design focuses on clarity and accessibility, allowing both technical and nontechnical users to use this system smoothly. Error messages and input validations have also been integrated, enabling the user to upload proper MRI files. Overall, the screenshot depicts the entire workflow of the application, showing how prediction and explanation are integrated into real time.

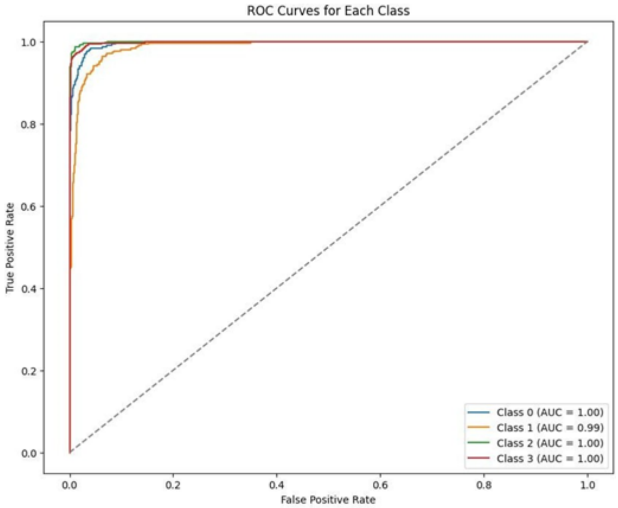
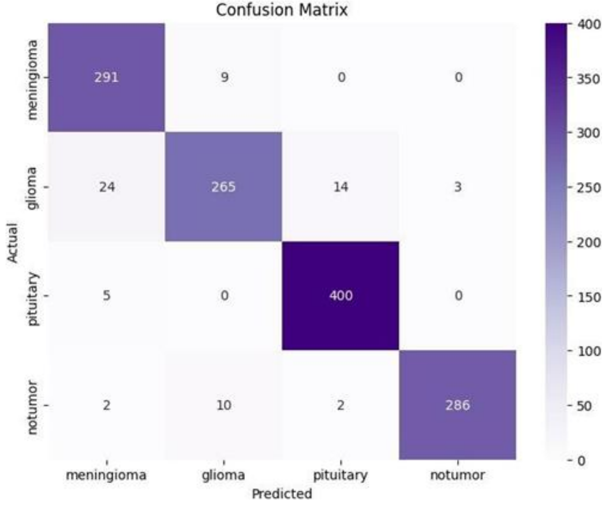


Figure1. ROC curve for Each Class

In the graph, ROC curves depict the model's classification performance for all four tumor classes. Each ROC curve plots the True Positive Rate against the False Positive Rate, with the curve being closer to the upper-left corner the better the performance of the model in that class. Classes 0, 2, and 3 have an AUC of 1.00, which is perfect separation without misclassifying any of the examples within the evaluation set. Class 1 exhibits an AUC value of 0.99, showing high class separation with a minimum number of false positives. Overall, ROC curves for the different classes confirm the model as highly effective in distinguishing the categories, with near-perfect performance across all classes.



	meningioma	glioma	pituitary	notumor
meningioma	291	9	0	0
glioma	24	265	14	3
pituitary	5	0	400	0
notumor	2	10	2	286

Figure 2. Confusion Matrix

The confusion matrix summarizes the model's performance regarding classification on all four tumor categories. Correctly classified instances are on the diagonals, and the high numbers on this diagonal indicate strong overall performance.

- Meningioma: 291 were correctly classified, with only 9 being misclassified as glioma.
- Glioma: 265 were correctly classified, with small confusion mainly toward meningioma (24) and pituitary (14).
- Pituitary Tumor: 400 correct predictions with near-perfect accuracy and a minimum of errors across the other classes.
- No Tumor: 286 correct predictions, with only a few samples misclassified as glioma or pituitary.

The matrix indeed shows that the model makes very good discrimination of pituitary and no-tumor cases, while the minor misclassifications occur mainly among glioma versus meningioma, two tumor types often with overlapping MRI characteristics. Overall, the confusion matrix indicates strong reliability of classification across all categories.

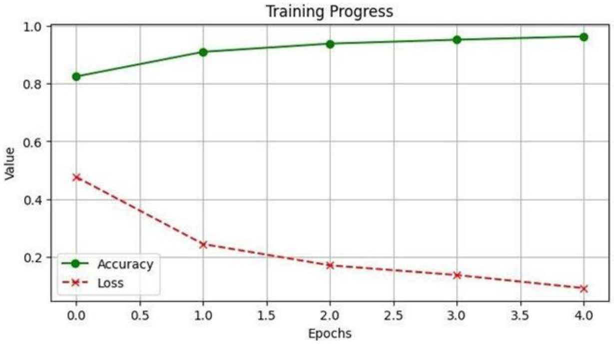


Figure3. Training Progress

The graph of training progress plots the model's accuracy and loss over five epochs. The accuracy curve, in green, increases smoothly from around 0.84 to nearly 0.98 in the last epoch, indicating that the model keeps learning and improving in every iteration. On the other hand, the loss curve in red decreases smoothly from approximately 0.48 to below 0.10, which suggests a decrease in the prediction error over time. Such movement in the opposite directions of these two curves-accuracy goes up and loss goes down-further justifies that the training is sound and effective. Overall, efficient convergence, strong learning behaviour, and good model optimization for only a few epochs are depicted by the graph.

REFERENCES

1. Scientific Reports Paper, Title: Brain Tumor Detection Using Deep-Learning Models and Transfer Learning on MRI, Journal: Scientific Reports, Publisher: Nature Publishing Group, Year: 2024 https://www.nature.com/articles/s41598-025-14901-4
2. Cancers (MDPI) Paper — YOLOv7-Based Brain Tumor Detection, Title: YOLOv7 with CBAM, SPPF+, and BiFPN for Brain Tumor Detection, Journal: Cancers, Publisher: MDPI, Year: 2023 https://www.mdpi.com/1424-8220/25/9/2746
3. Frontiers in Oncology — Deep Learning Brain Tumor Classification, Title: Deep Learning-Based Multi-Class MRI Brain Tumor Classification, Journal: Frontiers in Oncology, Year: 2022 https://www.frontiersin.org/journals/oncology/articles/10.3389/fonc.2025.1508326/full
4. Elsevier / ScienceDirect Paper — Hybrid CNN Framework, Title: Hybrid CNN and Transfer, Learning Model for MRI Brain Tumor Detection, Publisher: Elsevier, Year: 2025 https://www.sciencedirect.com/science/article/pii/S2001037022003737
5. Applied Sciences (MDPI) — Deep Learning & Allometric Analysis, Title: Allometric Modelling and Deep-Learning-Based Cancer Image Analysis, Journal: Applied Sciences, Year: 2020 https://www.mdpi.com/journal/applsci/special_issues/Deep_Learning_Medical_Image_Analysis
6. Med-PaLM / Google Research — LLMs in Medicine, Title: Large Language Models Encode Clinical Knowledge, Authors: Karan Singhal et al., Archive: arXiv:2212.13138, Year: 2022 https://arxiv.org/abs/2212.13138
7. U-Net — Biomedical Image Segmentation, Title: U-Net: Convolutional Networks for Biomedical Image Segmentation, Authors: Ronneberger,

Fischer, Brox, Conference: MICCAI, Year: 2015 https://arxiv.org/abs/1505.04597

8. INN — Inflated Neural Networks for IPMN Diagnosis, Title: INN: Inflated Neural Networks for IPMN Diagnosis, Conference: MICCAI, Year: 2019 https://arxiv.org/abs/1907.00437
9. Indonesian MRI CNN Classification Paper, Title: MRI-Based Brain Tumor Classification Using CNN Models, Journal: Indonesian Journal (exact metadata from PDF), Year: 2025 https://www.jneonatalsurg.com/index.php/jns/article/view/2642
10. Patra et al. — Deep Learning in Healthcare Systems, Title: Recent Advances, Challenges, and Applications of Deep Learning in Healthcare Systems for Medical Diagnosis and Treatment Conference: Proceedings of the 2023 3rd International Conference on Technological Advancements in Computational Sciences (ICTACS), Year: 2023 https://doi.org/10.1109/ICTACS59847.2023.10390124
11. Xu et al. — Federated Learning for Healthcare Informatics, Title: Federated Learning for Healthcare Informatics, Journal: Journal of Healthcare Informatics Research, Volume: 5(1), Pages 1– 19, Year: 2020 https://doi.org/10.1007/s41666-020-00082-4
12. Dwivedi et al. — Multidisciplinary Perspectives on AI, Title: Artificial Intelligence (AI): Multidisciplinary Perspectives on Emerging Challenges, Opportunities, and Agenda for Research, Practice, and Policy, Journal: International Journal of Information Management, Volume: 57, Article 101994, Year: 2019 https://doi.org/10.1016/j.ijinfomgt.2019.08.002
13. Javed et al. — Deep Learning for Lung Cancer Detection, Title: Deep Learning for Lung Cancer Detection: A Review, Journal: Artificial Intelligence Review, Volume: 57(8), Year: 2024 https://doi.org/10.1007/s10462-024-10807-1
14. Pesapane et al. — DL for Microcalcification Detection, Title: Deep Learning Performance for Detection and Classification of Microcalcifications on Mammography, Journal: European Radiology Experimental, Volume: 7(1): 69, Year: 2023 https://doi.org/10.1186/s41747-023-00384-3