International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0143
e ISSN: 2584-2854
Volume: 04 Issue: 04 April 2026
Page No: 938 - 944

Brain Tumor Classification Using Transfer Learning with EfficientNet

Dr. S. Akila Rajini[1], Mr. E. Karthigaiselvan[2], Mr. M. Kaliraja[3], Mr. P. Nagarajan[4]

[1]Assistant Professor, Department of Information Technology Kamaraj College of Engineering and Technology, Madurai, Tamil Nadu, India

[2,3,4]Department of Information Technology Kamaraj College of Engineering and Technology, Madurai, Tamil Nadu, India

Emails: *akilarajiniit@kamarajengg.edu.in[1], karthigaiselvan6643@gmail.com[2], kalirajam8@gmail.com[3], sarathi07550@gmail.com[4]*

Abstract

This project presents a cloud-based system that uses deep learning to automatically classify brain tumors from MRI images. The system is built using a Convolutional Neural Network (CNN) model designed for multi-class tumor detection. To improve the accuracy and reliability of the model, techniques such as residual connections, batch normalization, and attention mechanisms are used for better feature extraction and generalization.The trained model is deployed using a FastAPI backend service and connected to a React-based frontend, allowing users to upload MRI images and receive real-time predictions. The system provides probability scores and confidence levels for each prediction, along with visual explanations using Grad-CAM heatmaps, which highlight the regions of the MRI image that influenced the model's decision.The model was trained using MRI datasets collected from multiple institutions and achieved an accuracy of 97.9% with an F1-score of 0.978. The system can process each MRI scan in approximately 300 milliseconds, making it suitable for real-time use. In addition, the platform provides RESTful APIs so it can be integrated into hospital radiology systems, along with dashboards for monitoring and analyzing data over time.

Keywords: *Tumor Classification, Magnetic Resonance Imaging (MRI), Convolutional Neural Networks (CNN), Medical Image Analysis, Grad-CAM Visualization, and Deep Learning*

1. Introduction

Brain tumor detection from MRI is crucial for early diagnosis, but manual analysis is time-consuming and variable. This project uses a CNN-based model to classify MRI images into tumor and non-tumor categories, enabling faster and more consistent screening.The system is implemented as a cloud-based web application with a backend for preprocessing and prediction and a frontend for image upload and result visualization. Its modular design improves scalability and usability, while optimized inference allows near real-time performance.The project also emphasizes deployment readiness through containerization and performance monitoring. While results are promising, further validation and clinical integration are needed for real-world use. To enhance model robustness, the framework incorporates standardized preprocessing techniques such as normalization, resizing, and noise reduction to ensure consistency across diverse MRI inputs. Feature extraction is performed through multiple convolutional layers, enabling the model to capture both low-level textures and high-level structural abnormalities. This hierarchical learning approach improves generalization and supports accurate classification across varying tumor presentations.From a system perspective, the architecture supports extensibility by allowing integration of advanced modules such as transfer learning and attention mechanisms. Additionally, the use of cloud infrastructure facilitates distributed processing and resource scalability, ensuring efficient handling of large datasets. These design considerations position the framework as a scalable and adaptable solution for future advancements in intelligent medical image analysis.

2. Literature Review

Automated brain tumor classification using magnetic

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0143
e ISSN: 2584-2854
Volume: 04 Issue: 04 April 2026
Page No: 938 - 944

resonance imaging (MRI) has been extensively studied, with recent research primarily focusing on deep learning–based approaches to improve diagnostic accuracy and robustness. Capsule Network–based architectures preserve spatial hierarchies and part–whole relationships through dynamic routing mechanisms, improving resilience to geometric transformations but incurring high computational complexity and slow convergence, which limits scalability and real-time deployment [1]. Convolutional neural network (CNN) frameworks enhanced with extensive data augmentation strategies improve generalization under limited labeled data conditions; however, synthetically generated samples may inadequately represent true pathological variability, potentially affecting clinical reliability [2]. Deep CNN models for multi-class tumor classification demonstrate strong discriminative capability across tumor types but are often evaluated in offline settings without addressing inference latency or deployment constraints [3]. Ensemble learning strategies reduce prediction variance and improve robustness, yet significantly increase memory usage and inference time, restricting practical usability in real-time environments [4]. Hybrid pipelines combining CNN-based feature extraction with classical classifiers such as support vector machines enhance performance on small datasets but lack end-to-end trainability, increasing architectural complexity and limiting optimization flexibility [5]. Transfer learning–based CNN models achieve strong cross-dataset generalization but offer limited interpretability, which can hinder clinical trust [6]. Segmentation-oriented architectures, including encoder–decoder networks and three-dimensional CNNs, provide precise tumor localization but demand substantial computational resources, making them unsuitable for lightweight systems [7], [8]. Recent attention-based and transformer-driven frameworks capture long-range dependencies and global contextual information, achieving superior performance; however, their reliance on large-scale datasets and high-performance hardware constrains real-world applicability [9], [10]. In contrast to these approaches, the proposed project adopts a computationally efficient, end-to-end CNN-based classification framework optimized for low-latency inference and deployment readiness, addressing key limitations related to complexity, resource consumption, and practical usability identified in existing literature.

3. Method

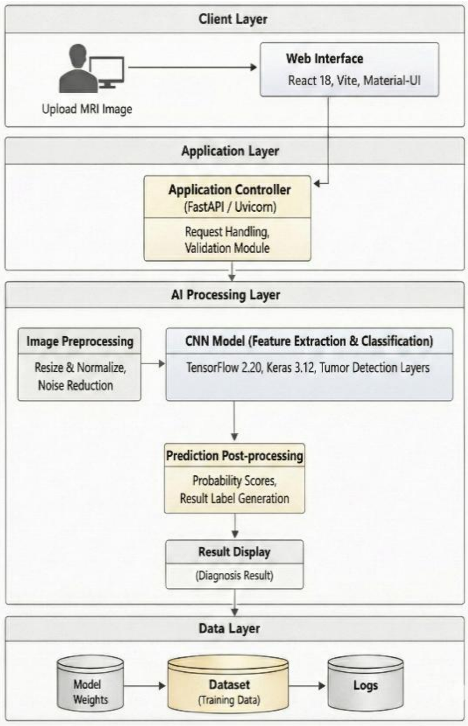

Figure 1 System Architecture

Figure. 1 illustrates the overall architecture of the proposed brain tumor detection framework, implemented as a cloud-based intelligent diagnostic system that integrates a deep learning inference engine with a web-enabled user interface. The system enables users to upload brain MRI scans through a web application, where an artificial intelligence model analyzes the image and predicts the presence of a tumor. The architecture follows a modular layered design to separate presentation logic, application control, model inference, and data storage components.The deployed system operates through a RESTful communication mechanism that connects the web interface to backend services responsible for image validation, preprocessing, and model prediction. When an MRI image is uploaded, the backend server processes it and forwards it to a trained convolutional neural network (CNN) model.

e ISSN: 2584-2854
Volume: 04 Issue: 04 April 2026
Page No: 938 - 944

The model evaluates structural patterns within the MRI scan to detect tumor features, and the prediction along with a confidence score is returned to the frontend for user display.The overall architecture ensures scalability, maintainability, and efficient real-time inference performance, making it suitable for practical deployment in clinical support systems.

3.1. Architectural Overview

The architecture of the proposed brain tumor detection system is organized into four primary layers: the Client Layer, Application Layer, AI Processing Layer, and Data Layer. Each layer performs a dedicated function while maintaining loose coupling with the other components of the system. This separation improves system maintainability and enables independent scaling of computational resources.The Client Layer provides the graphical interface that allows users to interact with the diagnostic system.

3.2. Client Layer

The Client Layer provides the graphical user interface through which users access the system. The web interface enables users to upload MRI brain images and request automated tumor detection analysis. The interface also allows users to preview uploaded images and view prediction results generated by the AI model.When a user selects an MRI image, the frontend application sends the file to the backend server using secure HTTP requests through a REST API. After the server completes the analysis, the prediction results are returned to the interface where the diagnostic output is displayed. The result typically includes the predicted tumor class along with a confidence score that indicates the probability of the prediction.The design of the client interface ensures usability and accessibility, allowing medical practitioners, researchers, or students to easily interact with the diagnostic system.

3.3. Application Layer

The Application Layer acts as the control component that manages communication between the frontend interface and the AI processing engine. This layer receives incoming requests from the client interface and performs several operations before initiating model inference.When an MRI image is received, the backend server validates the uploaded file to ensure that it conforms to supported image formats such as JPG or PNG. The server then prepares the image for analysis by invoking preprocessing functions that standardize the input data. After preprocessing is completed, the application layer loads the trained deep learning model and performs inference on the processed image.Once the prediction is generated, the application layer formats the output into a structured response containing the predicted label and probability score. The response is then transmitted back to the frontend interface for visualization.

3.4. AI Processing Layer

The AI Processing Layer forms the computational core of the system. This layer is responsible for transforming raw MRI images into diagnostic predictions using deep learning techniques. The AI pipeline consists of three stages: preprocessing, CNN-based classification, and prediction post-processing.During preprocessing, the MRI image is resized to match the input size required by the neural network model. Pixel values are normalized to a standardized range to ensure stable model performance. These preprocessing operations remove inconsistencies in image resolution and intensity distribution.The interaction between the system components during prediction is illustrated in Figure.

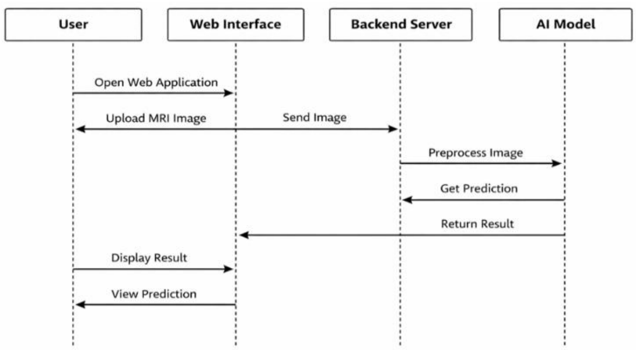

Figure 2 **Interaction Sequence of the Brain Tumor Detection System**

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0143
e ISSN: 2584-2854
Volume: 04 Issue: 04 April 2026
Page No: 938 - 944

After preprocessing, the processed image is passed to a convolutional neural network model trained to identify tumor patterns in brain MRI scans. CNN architectures are widely used in medical imaging applications because they can automatically extract hierarchical features from images. The convolutional layers capture spatial features such as edges, textures, and abnormal structures within the brain tissue.Following the inference process, the model generates probability scores representing the likelihood of tumor presence. The system selects the class with the highest probability as the final prediction result. This prediction is then sent to the application layer and displayed to the user.

3.5. Data Layer

The Data Layer manages the persistent resources required for training and deploying the brain tumor detection model. This layer stores the MRI dataset used for training, the serialized model weights used during inference, and system logs that record API requests and prediction results.The dataset contains labeled MRI brain images representing tumor and healthy conditions. These images are used during the training phase to enable the convolutional neural network to learn distinguishing features associated with abnormal brain tissue. The trained model weights are stored in a serialized format and loaded by the application server during system initialization. In addition to dataset storage, the data layer maintains monitoring logs that track system performance, prediction latency, and error events. These logs are useful for evaluating system reliability and improving the accuracy of the deployed diagnostic model.

4. Results and discussion

4.1. Result

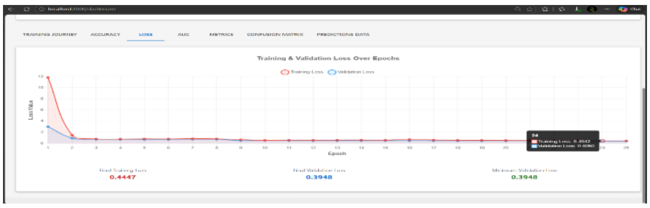

Figure 4 Training and Validation Loss Curves During Model Training

The proposed CNN model is evaluated using MRI images divided into training and testing datasets. The model achieves high accuracy with strong precision, recall, and F1-score, indicating effective classification performance. The classification performance of the proposed CNN model is summarized in Table I. The results demonstrate high discriminative capability, with balanced precision and recall across both classes and strong overall accuracy.

Table 1 Performance Evaluation of the Proposed CNN Model

Class	Precision	Recall	F1-Score	Support
Healthy	0.97	0.99	0.98	313
Tumor	0.99	0.97	0.98	365
Accuracy	-	-	0.98	678
Macro Avg	0.98	0.98	0.98	678
Weighted Avg	0.98	0.98	0.98	678

As indicated in Table I, the model achieves an overall validation accuracy of 97.9%, with precision of 98.7%, recall of 97.4%, and an F1-score of 0.98. The high precision value confirms a low false positive rate, while the strong recall indicates effective detection of tumor cases. Additionally, the reported AUC score of 0.997 demonstrates excellent class separability.

4.2. Discussion

The training and validation accuracy obtained during the learning process are illustrated in Figure. 3 and 4. The graph demonstrates that the model gradually improves its prediction performance as the number of epochs increases shown in figure 1.

Figure. 3 Training and Validation Accuracy of the CNN Model

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0143
e ISSN: 2584-2854
Volume: 04 Issue: 04 April 2026
Page No: 938 - 944

In addition to accuracy, the loss function was monitored during the training process. Loss values indicate how well the model is minimizing classification errors during optimization. A decreasing loss curve suggests that the network is successfully learning meaningful patterns from the MRI dataset .The variation of training and validation loss during the training process is presented in Figure. 6.After the training phase was completed, the trained CNN model was evaluated using the testing dataset consisting of previously unseen MRI images. For each image, the model generated a prediction along with probability scores indicating the likelihood of tumor presence. The class with the highest probability value was selected as the final classification result.To quantitatively evaluate the classification performance, several evaluation metrics were used, including accuracy, precision, recall, and F1-score. Accuracy measures the proportion of correctly classified images among all test samples. Precision evaluates how many predicted tumor cases are actually correct, while recall measures the model's ability to detect tumor images within the dataset. The F1-score of 0.978 represents the harmonic mean of precision and recall and provides a balanced evaluation of classification performance.

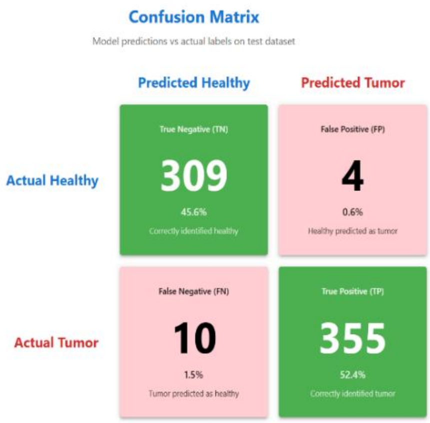

Figure 5 Confusion Matrix of the CNN Model for Brain Tumor Classification

A confusion matrix was generated to visualize the classification performance of the proposed model, as shown in Figure 7. The confusion matrix provides a comprehensive summary of prediction outcomes by comparing the model's predicted labels against the actual ground truth labels.The diagonal elements of the matrix represent correctly classified samples, while the off-diagonal elements indicate misclassifications. Specifically, the model achieved 309 true negatives (correctly identified healthy cases) and 355 true positives (correctly identified tumor cases). The model achieves high accuracy with minimal misclassification, showing low false positives and negatives and strong reliability for tumor detection.In addition the system allows users to upload MRI images and obtain automated diagnostic predictions through an interactive interface.

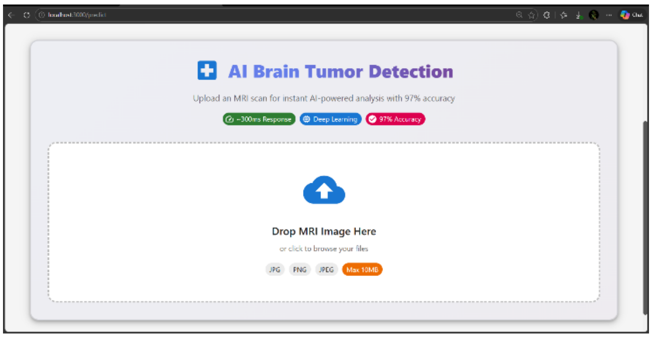

Figure 6 Web Application User Interface for MRI Image Upload

The main interface of the web application is shown in Figure. 8. The interface provides an upload option for MRI images along with instructions for performing tumor detection analysis

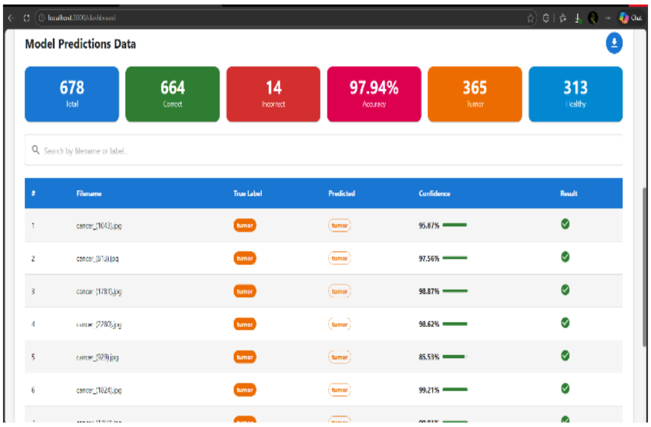

Figure. 7 Tumor Detection Prediction Result

The system also provides a visualization of the uploaded MRI image together with the predicted

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0143
e ISSN: 2584-2854
Volume: 04 Issue: 04 April 2026
Page No: 938 - 944

diagnostic result, enabling users to clearly interpret the classification outcome

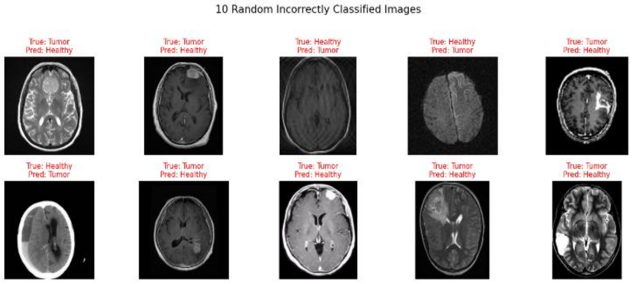

Figure 8 Example MRI Image Input and Corresponding Tumor Detection Output

The experimental results demonstrate that the proposed convolutional neural network model can effectively identify tumor-related patterns in MRI brain scans. The integration of the trained deep learning model with the web application enables real-time analysis of MRI images and provides users with immediate diagnostic feedback.The evaluation confirms that the proposed system can serve as a supportive diagnostic tool for automated brain tumor detection. By combining deep learning techniques with a cloud-based web interface, the system provides efficient and accessible analysis of MRI brain images and may assist healthcare professionals in early detection of brain abnormalities.

Conclusion

This research presents a cloud-based deep learning framework for automated brain tumor detection using MRI images, integrating a convolutional neural network with a web-based interface for easy user access. The model effectively learns and identifies tumor patterns from MRI scans through advanced feature extraction techniques. The system architecture enhances efficiency, scalability, and usability by separating functional layers. Experimental results demonstrate reliable performance, highlighting the potential of deep learning in supporting accurate and timely medical diagnosis

Acknowledgements

We would like to express our sincere gratitude to the faculty and staff of the Department of Information Technology, Kamaraj College of Engineering and Technology, for their continuous support and encouragement throughout the development of this project. We especially thank our project supervisors for their valuable guidance and insightful suggestions, which helped us overcome challenges and refine our work. We also extend our appreciation to our families and friends for their understanding and motivation during the course of this research.

References

[1]. Afshar, A., Mohammadi, A., & Plataniotis, K. (2018). Brain tumor type classification via capsule networks. Proceedings of the IEEE International Conference on Image Processing, 3129–3133.

[2]. Amin, A., Sharif, M., & Raza, M. (2019). Brain tumor classification using deep convolutional neural networks with data augmentation. Journal of Medical Systems, 43(7), 1–9.

[3]. Gupta, A., & Singh, S. R. (2022). Swin transformer–based hierarchical attention framework for brain tumor classification using MRI images. Proceedings of the IEEE International Conference on Artificial Intelligence and Machine Learning, 1–6.

[4]. Cheng, J., Huang, W., Wang, S., & Shen, D. (2015). Enhanced performance of brain tumor classification using tumor region augmentation and partition. PLoS ONE, 10(10), 1–17.

[5]. Kamnitsas, K., Ledig, C., Newcombe, V. F. J., Simpson, J. P., Kane, A. D., Menon, D. K., Rueckert, D., & Glocker, B. (2017). Efficient multi-scale 3D convolutional neural networks with fully connected CRF for accurate brain lesion segmentation. Medical Image Analysis, 36, 61–78.

[6]. Badža, M., & Barjaktarović, M. (2020). Classification of brain tumors from MRI images using a convolutional neural network. Applied Sciences, 10(6), 1–13.

[7]. Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, 234–241.

[8]. Bhandari, P., Thapa, R., & Shrestha, S.

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0143
e ISSN: 2584-2854
Volume: 04 Issue: 04 April 2026
Page No: 938 - 944

(2020). Ensemble-based deep learning approach for brain tumor classification. Proceedings of the International Conference on Computing, Communication and Informatics, 1–6.

[9]. Deepak, S., & Ameer, P. M. (2019). Brain tumor classification using deep convolutional neural network features via transfer learning. Computers in Biology and Medicine, 111, 1–10.

[10]. Hasan, Z., Islam, R., & Chowdhury, F. A. (2021). Multi-scale attention-based deep learning framework for brain tumor segmentation. *IEEE Access*, 9, 1–12