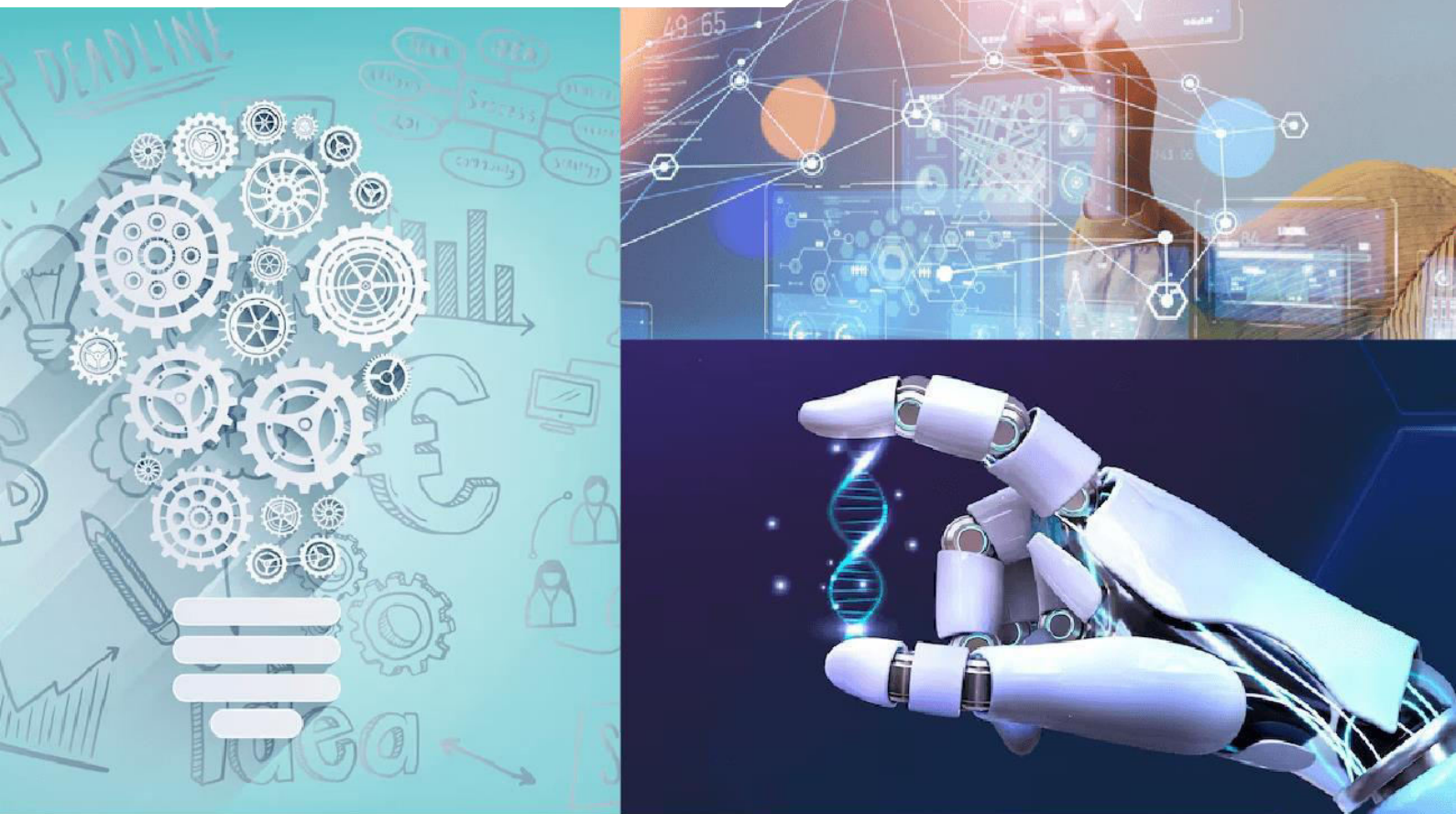




International Journal of Advanced Research in Education and Technology (IJARETY)

Volume 13, Issue 1, January - February 2026

Impact Factor: 8.152



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



Brain Tumor Classification Using MRI Images and Convolutional Neural Networks

Suhas A N

Master of Computer Applications, CMR Institute of Technology, Bengaluru, India

ABSTRACT-Brain tumors are among the most critical and life-threatening neurological disorders, requiring accurate and early diagnosis to improve patient survival rates. Conventional diagnostic approaches rely heavily on manual examination of Magnetic Resonance Imaging (MRI) scans by radiologists, which is time-consuming, subjective, and prone to inconsistencies. With the advancement of artificial intelligence, deep learning techniques have emerged as powerful tools for automating medical image analysis.

This research proposes an automated brain tumor classification system using Convolutional Neural Networks (CNNs) combined with transfer learning techniques. The system classifies MRI images into four categories: Glioma, Meningioma, Pituitary Tumor, and No Tumor. Pre-trained models such as VGG16, MobileNetV2, and ResNet50 are fine-tuned on MRI datasets to improve classification accuracy while reducing training time.

To address the challenge of model interpretability, Gradient-weighted Class Activation Mapping (Grad-CAM) is incorporated to highlight tumor-relevant regions in MRI scans. The system is deployed as a web-based application that enables real-time image classification and automated report generation. Experimental results demonstrate that the proposed system achieves an accuracy of 93.5%, proving its effectiveness as a reliable decision-support tool in medical diagnosis.

KEYWORDS: Brain Tumor Classification, MRI Imaging, Deep Learning, CNN, Transfer Learning, Grad-CAM, Medical Image Analysis

I. INTRODUCTION

Brain tumors are among the most critical neurological disorders, significantly impacting human health and survival rates. These tumors arise due to abnormal and uncontrolled cell growth within the brain and are commonly categorized into types such as glioma, meningioma, and pituitary tumors. Accurate classification of these tumors is essential for determining appropriate treatment strategies and improving patient outcomes. Magnetic Resonance Imaging (MRI) is widely regarded as the most effective imaging modality for brain tumor detection due to its high-resolution visualization of soft tissues [1][11].

However, manual analysis of MRI scans by radiologists is a complex and time-consuming task. It requires significant expertise and is often subject to inter-observer variability, which can lead to inconsistent diagnoses. In many healthcare settings, especially in resource-limited regions, the shortage of skilled radiologists further delays diagnosis and treatment. Therefore, the development of automated and reliable systems for brain tumor classification has become an important research area [12][13].

Recent advancements in artificial intelligence, particularly deep learning, have revolutionized medical image analysis. Convolutional Neural Networks (CNNs) have demonstrated remarkable success in image classification tasks by automatically learning hierarchical feature representations from raw data [2][14]. These models eliminate the need for manual feature engineering and have been successfully applied in various medical domains, including tumor detection and disease diagnosis [17].

Transfer learning has emerged as a powerful technique in situations where labeled medical datasets are limited. By leveraging pre-trained models such as VGG16, ResNet50, and MobileNetV2, which are trained on large-scale datasets like ImageNet, it is possible to achieve high classification accuracy with reduced training time [3][4][5]. These models provide a strong foundation for extracting meaningful features from MRI images.

DOI:10.15680/IJARETY.2026.1302049

Despite the high accuracy of deep learning models, their black-box nature remains a major challenge in clinical applications. Clinicians require transparency and interpretability to trust automated systems. To address this issue, explainability techniques such as Gradient-weighted Class Activation Mapping (Grad-CAM) have been introduced. Grad-CAM generates visual heatmaps that highlight the regions in an image that influence the model's predictions, thereby improving interpretability and clinical trust [6].

In this research, an automated brain tumor classification system is proposed using CNN-based transfer learning models combined with Grad-CAM visualization. The system is deployed as a web-based application to enable real-time classification and report generation. The primary objective is to develop an accurate, efficient, and interpretable system that can assist healthcare professionals in making faster and more reliable diagnostic decisions.

II. LITERATURE SURVEY

The application of machine learning and deep learning techniques in brain tumor classification has been extensively studied over the years. Early approaches primarily relied on traditional machine learning algorithms, where handcrafted features such as texture, shape, and intensity were extracted from MRI images and used with classifiers like Support Vector Machines (SVM) and k-Nearest Neighbors (k-NN). Although these methods provided baseline results, they were limited by their dependency on manual feature extraction and poor generalization across datasets [19].

The introduction of deep learning, particularly Convolutional Neural Networks (CNNs), significantly improved the performance of medical image classification systems. Krizhevsky et al. demonstrated the effectiveness of deep CNNs with the development of AlexNet, which achieved breakthrough results in image classification tasks [2]. This was followed by the development of deeper architectures such as VGGNet [3], ResNet [4], and MobileNetV2 [5], each contributing improvements in feature extraction, training stability, and computational efficiency.

Transfer learning has become a widely adopted approach in medical imaging applications due to the limited availability of annotated datasets. By utilizing pre-trained models trained on large datasets like ImageNet, researchers can fine-tune these models for specific tasks such as brain tumor classification. Studies have shown that transfer learning significantly improves accuracy while reducing training time [3][4][21].

Several studies have also explored advanced architectures and hybrid approaches. Capsule Networks were introduced to preserve spatial relationships between features, providing improved classification performance in certain cases [8]. Similarly, generative models such as Generative Adversarial Networks (GANs) have been used for pre-training and data augmentation to enhance model robustness [10].

Recent advancements include the use of transformer-based models such as Vision Transformers (ViT), which utilize self-attention mechanisms to capture global contextual information from images. Although these models show promising results, they require large datasets and high computational resources, limiting their practical applicability in many medical scenarios [18][15].

A significant challenge in deploying deep learning models in healthcare is the lack of interpretability. To address this issue, explainability techniques such as Grad-CAM have been widely adopted. Grad-CAM provides visual explanations by highlighting important regions in the input image, enabling clinicians to understand and validate model predictions [6]. This has improved trust and usability of AI-based systems in clinical environments.

Several researchers have also emphasized the importance of preprocessing techniques such as normalization, skull stripping, and data augmentation to improve model performance and generalization [7][9]. These preprocessing steps help reduce noise and variability in MRI images, leading to more robust classification results.

Despite these advancements, many existing systems lack real-time deployment capabilities and user-friendly interfaces. Additionally, some models fail to provide interpretable outputs, limiting their adoption in clinical practice. The proposed system addresses these gaps by integrating transfer learning, explainability, and a web-based deployment framework for real-time brain tumor classification and analysis.

III. PROPOSED METHODOLOGY

The proposed system follows a structured pipeline consisting of data collection, preprocessing, model training, classification, and deployment. The methodology is designed to ensure high accuracy, efficiency, and interpretability. The complete workflow integrates deep learning models with explainability techniques to support reliable medical diagnosis.

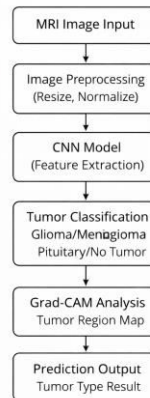


Fig 1: Overall System Architecture of Brain Tumor Classification System

3.1 Dataset

The dataset used in this research consists of labeled MRI images categorized into four classes: Glioma, Meningioma, Pituitary Tumor, and No Tumor. The dataset is obtained from publicly available medical imaging sources and is divided into training and testing sets using an 80:20 ratio to ensure proper evaluation of model performance.

To improve model generalization and address class imbalance, data augmentation techniques are applied to the training dataset. These include rotation, horizontal and vertical flipping, zooming, and brightness adjustments. These transformations increase dataset diversity and help the model learn robust features under varying imaging conditions.

3.2 Data Preprocessing

Preprocessing plays a crucial role in improving model performance by ensuring uniformity in input data. The following steps are performed:

- Resizing all images to 224×224 pixels to match model input requirements
- Normalizing pixel values to the range $[0,1]$ for stable training
- Applying data augmentation techniques to enhance generalization
- Encoding labels into categorical format for multi-class classification

These preprocessing steps reduce noise, improve feature consistency, and minimize the risk of overfitting during training.

3.3 Transfer Learning Models

Three pre-trained Convolutional Neural Network (CNN) architectures are used in this study:

- VGG16: A deep architecture with 16 layers that uses small convolution filters. It is known for its simplicity and effectiveness in feature extraction.
- MobileNetV2: A lightweight architecture that uses depthwise separable convolutions, making it suitable for real-time and resource-constrained applications.
- ResNet50: A deep residual network that uses skip connections to overcome vanishing gradient problems and improve training stability.

These models are pre-trained on large datasets and fine-tuned for MRI image classification, enabling efficient learning even with limited medical data.

3.4 Model Training

The models are trained using the Adam optimizer and categorical cross-entropy loss function, which are suitable for multi-class classification tasks.

Training is performed in two phases:

1. Feature Extraction Phase: Pre-trained layers are frozen to retain learned features

2. Fine-Tuning Phase: Top layers are unfrozen to adapt the model to MRI data

Early stopping is applied based on validation loss to prevent overfitting. A learning rate scheduler is also used to improve convergence and training stability.

3.5 Grad-CAM Explainability

To enhance interpretability, Grad-CAM (Gradient-weighted Class Activation Mapping) is implemented. Grad-CAM generates heatmaps that highlight the regions of the MRI image that influence the model's predictions.

This helps in:

- Understanding model decision-making
- Verifying that the model focuses on tumor regions
- Increasing trust among medical professionals

The overall data flow of the proposed system is illustrated in Fig. 2.

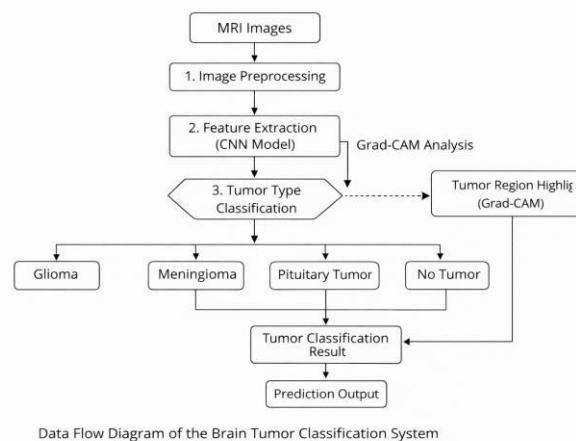


Fig.2: Data Flow Diagram of the Brain Tumor Classification System

The data flow diagram represents the complete workflow of the proposed system, including preprocessing, feature extraction using CNN models, tumor classification, and Grad-CAM-based visualization. It illustrates how the system processes MRI images to generate accurate and interpretable predictions.

3.6 System Deployment

The system is deployed as a web application using the Streamlit framework. The application provides a user-friendly interface for real-time interaction.

Key features include:

- Uploading MRI images
- Performing tumor classification
- Displaying Grad-CAM visualizations
- Generating downloadable PDF reports

This deployment ensures accessibility and practical usability in clinical environments.

IV. RESULTS AND DISCUSSION

The performance of the proposed brain tumor classification system is evaluated using standard metrics such as accuracy, precision, recall, and F1-score. These metrics provide a comprehensive assessment of the model's ability to correctly classify MRI images into different tumor categories.

Among the three transfer learning models used in this study, VGG16 achieved the highest classification accuracy of 93.5%, demonstrating superior performance in feature extraction and classification. The ResNet50 model achieved an accuracy of 92.8%, benefiting from residual connections that improve training stability. The MobileNetV2 model

achieved an accuracy of 91.2%, offering a good balance between accuracy and computational efficiency, making it suitable for real-time applications.

The classification performance of the best-performing model (VGG16) is further analyzed using a confusion matrix, as shown in Fig. 3.

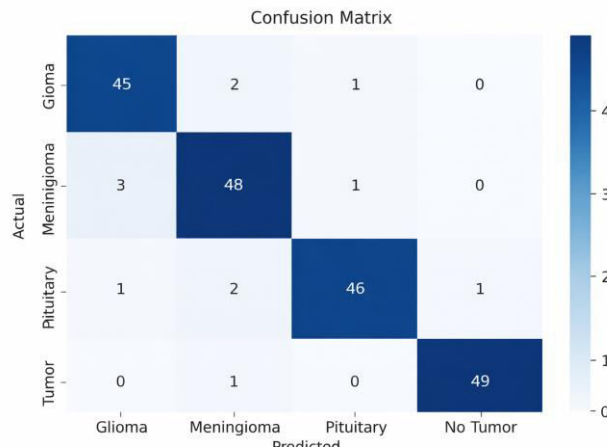


Fig. 3: Confusion Matrix of Brain Tumor Classification

The confusion matrix illustrates the model's ability to correctly classify MRI images across all four categories: Glioma, Meningioma, Pituitary Tumor, and No Tumor. The results indicate that the model achieves high classification accuracy with minimal misclassification between classes. The "No Tumor" class shows particularly high precision, which is important in avoiding false-positive diagnoses.

To enhance model interpretability, Grad-CAM visualization is applied to highlight the regions of the MRI image that influence the model's predictions. A sample Grad-CAM output is shown in Fig. 4.

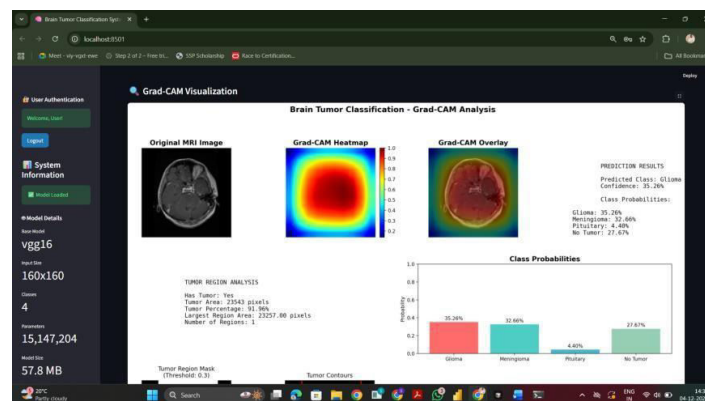


Fig. 4: Grad-CAM Visualization of Tumor Region

The Grad-CAM visualization highlights the tumor regions within the MRI image, confirming that the model focuses on clinically relevant areas while making predictions. This improves the transparency of the system and allows medical professionals to verify the model's decision-making process. In addition to classification accuracy, the system demonstrates a fast response time of approximately **3 seconds per prediction**, making it suitable for real-time applications.

The integration of classification, visualization, and report generation ensures that the system can be effectively used as a decision-support tool in clinical and research environments. Overall, the results validate that the proposed system achieves high accuracy, strong generalization, and reliable interpretability, making it a practical solution for automated brain tumor classification.

V. CONCLUSION

This research presents an efficient and accurate brain tumor classification system using deep learning and transfer learning techniques. The system utilizes pre-trained CNN models such as VGG16, ResNet50, and MobileNetV2 to classify MRI images into four categories: Glioma, Meningioma, Pituitary Tumor, and No Tumor. Among the evaluated models, VGG16 achieved the highest classification accuracy of 93.5%, demonstrating its effectiveness in medical image analysis. The integration of Grad-CAM significantly enhances the interpretability of the system by providing visual explanations of model predictions, thereby increasing trust among medical professionals.

The system is implemented as a web-based application using Streamlit, enabling real-time tumor classification, visualization, and automated report generation. Its user-friendly interface and fast response time make it suitable for practical deployment in clinical and academic environments. The proposed system can serve as a reliable decision-support tool for radiologists, reducing workload and improving diagnostic efficiency. For future work, the system can be further enhanced by integrating advanced architectures such as EfficientNet and Vision Transformers, incorporating multi-modal medical data, and performing large-scale clinical validation to improve robustness and generalization.

REFERENCES

- [1] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, 2015.
- [2] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems*, 2012.
- [3] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *Proc. ICLR*, 2015.
- [4] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE CVPR*, 2016, pp. 770–778.
- [5] M. Sandler *et al.*, "MobileNetV2: Inverted residuals and linear bottlenecks," in *Proc. IEEE CVPR*, 2018.
- [6] R. R. Selvaraju *et al.*, "Grad-CAM: Visual explanations from deep networks via gradient-based localization," in *Proc. IEEE ICCV*, 2017, pp. 618–626.
- [7] J. Cheng *et al.*, "Enhanced performance of brain tumor classification via tumor region augmentation," *PLOS ONE*, vol. 10, no. 10, 2015.
- [8] P. Afshar, A. Mohammadi, and K. N. Plataniotis, "Brain tumor type classification via capsule networks," in *Proc. IEEE ICIP*, 2018.
- [9] H. H. Sultan, N. M. Salem, and W. Al-Atabany, "Multi-classification of brain tumor images using deep neural network," *IEEE Access*, vol. 7, pp. 69545–69559, 2019.
- [10] N. Ghassemi, A. Shoeibi, and M. Rouhani, "Deep neural network with generative adversarial networks pre-training for brain tumor classification," *Biomedical Signal Processing and Control*, vol. 57, 2020.
- [11] A. Esteva *et al.*, "A guide to deep learning in healthcare," *Nature Medicine*, vol. 25, pp. 24–29, 2019.
- [12] G. Litjens *et al.*, "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, pp. 60–88, 2017.
- [13] J. Ker, L. Wang, J. Rao, and T. Lim, "Deep learning applications in medical image analysis," *IEEE Access*, vol. 6, pp. 9375–9389, 2018.
- [14] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.
- [15] A. Vaswani *et al.*, "Attention is all you need," in *Advances in Neural Information Processing Systems*, 2017.
- [16] J. Devlin *et al.*, "BERT: Pre-training of deep bidirectional transformers for language understanding," in *NAACL-HLT*, 2019.
- [17] D. Shen, G. Wu, and H. Suk, "Deep learning in medical image analysis," *Annual Review of Biomedical Engineering*, vol. 19, pp. 221–248, 2017.
- [18] A. Dosovitskiy *et al.*, "An image is worth 16×16 words: Transformers for image recognition at scale," in *Proc. ICLR*, 2021.
- [19] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*. Springer, 2009.
- [20] S. Zagoruyko and N. Komodakis, "Wide residual networks," in *Proc. BMVC*, 2016.
- [21] M. Tan and Q. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in *Proc. ICML*, 2019.
- [22] C. Szegedy *et al.*, "Going deeper with convolutions," in *Proc. IEEE CVPR*, 2015.

International Journal of Advanced Research in Education and Technology

ISSN: 2394-2975

Impact Factor: 8.152