iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL
FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 **Issue:** VI **Month of publication:** June 2025

DOI: https://doi.org/10.22214/ijraset.2025.72284

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue VI June 2025- Available at www.ijraset.com

Comparative Deep Learning Framework for Brain Tumor Detection: Integrating VGG16 Classification and U-Net Segmentation with Performance Evaluation

Sanskriti Jhanwar, Saumya Suman
Dept. of Computer Science and Engineering DY Patil International University, Pune, India

Abstract: Brain tumors are serious and potentially life-threatening, making early and accurate diagnosis essential for improving survival chances. MRI (Magnetic Resonance Imaging) is a key tool in detecting these tumors without invasive procedures, but analyzing MRI scans manually takes time and can lead to mistakes.

In this study, we explore how deep learning—specifically Convolutional Neural Networks (CNNs)—can help automate and improve brain tumor detection using MRI images. We compare the performance of four CNN-based models: EfficientNetB3, VGG16 with fine-tuned transfer learning, ResNet-50, and a custom-built CNN. Our goal is to find out which model performs best in terms of accuracy, precision, recall, and F1-score. After extensive testing, the fine-tuned VGG16 model came out on top as the most accurate and dependable.

To build a complete diagnostic system, we combined the VGG16 model with U-Net architecture for tumor segmentation. This approach not only classifies the type of brain tumor—glioma, meningioma, pituitary tumor, or no tumor—but also pinpoints its location within the MRI scan. The segmentation achieved strong results, with an IoU score of 0.87 and a Dice score of 0.91.

Overall, our research delivers a comprehensive tool that handles both tumor classification and localization in one streamlined process. The results underscore the power of fine-tuned VGG16 for medical imaging and show how combining it with segmentation can enhance tumor analysis. This solution has the potential to assist radiologists by providing faster, more accurate, and consistent diagnostic support, especially in settings where resources are limited.

Keywords: Brain tumor detection; MRI image analysis; CNN classification; U-Net segmentation; Transfer learning; Bayesian optimization; Deep learning; Glioma; Meningioma; Pituitary tumor

I. INTRODUCTION

Brain tumors are among the most serious neurological conditions, often posing a threat to life. Early detection and accurate classification play a vital role in guiding treatment plans and improving patient outcomes. Radiologists primarily rely on Magnetic Resonance Imaging (MRI) because it offers excellent image contrast and doesn't involve invasive procedures. However, going through these MRI scans manually can be slow, complex, and prone to human error. It requires a high level of expertise, and even then, distinguishing between different tumor types can be tricky due to overlapping features and varied tumor shapes and locations.

In recent years, Artificial Intelligence (AI)—and particularly Deep Learning (DL)—has emerged as a powerful tool for improving how we analyze medical images. Convolutional Neural Networks (CNNs), in particular, have transformed tasks like image classification, object detection, and segmentation by learning patterns directly from image data. When applied to brain tumor analysis, these models have delivered impressive results, offering significant support to doctors and radiologists in making informed decisions.

Despite the growing use of deep learning in healthcare, there's still a gap when it comes to comparing different CNN architectures under the same conditions. Since these models differ in their complexity and ability to learn, it's important to evaluate them side-by-side to understand which performs best for brain tumor detection. Also, most existing studies focus only on classification or segmentation, rather than combining both to create a complete diagnostic tool.

This study aims to bridge that gap by introducing a comparative framework that evaluates four CNN-based models—EfficientNetB3, a fine-tuned VGG16, ResNet-50, and a custom-built CNN—using a publicly available dataset of brain MRI scans. Our goal is to find out which model offers the best performance when it comes to accurately classifying tumors.

Our findings show that the fine-tuned VGG16 model outperforms the others in terms of accuracy, precision, recall, and F1- score. To take it a step further, we integrate VGG16 with the U-Net segmentation model. This combination not only predicts the tumor type but also pinpoints its exact location in the brain MRI scan. The result is a comprehensive and practical solution for both detecting and localizing brain tumors—one that could make a meaningful difference in clinical environments.

II. LITERATURE REVIEW

In recent years, the intersection of medical imaging and artificial intelligence has emerged as a key area of research, especially for brain tumor detection. Traditionally, diagnosing brain tumors involves manually analyzing MRI scans—a process that is time-consuming and prone to human error. As a result, AI-powered tools are increasingly being adopted to enhance diagnostic accuracy and speed. Among these tools, deep learning models—especially Convolutional Neural Networks (CNNs)—have become the go-to solution for tasks like image classification and segmentation.

A. CNNs in Medical Imaging

CNNs are deep learning models designed specifically for image data. They automatically learn patterns and features from raw images, making them ideal for identifying and classifying medical conditions in scans. Their breakthrough came with the success of AlexNet [1], which showed how powerful deep networks can be. Since then, CNNs have become widely used in healthcare, particularly for analyzing MRI and CT images.

B. VGG16 and Transfer Learning

VGG16, introduced by Simonyan and Zisserman [2], is a deep CNN model with 16 layers. Its straightforward design, using uniform 3x3 convolution filters, makes it efficient and easy to work with. One of its major strengths is its compatibility with transfer learning, where a model trained on a large dataset like ImageNet is fine-tuned for a specific task—in this case, brain tumor classification. This approach helps overcome the challenge of limited labeled medical data and reduces training time. Studies such as [3] have shown that fine-tuned VGG16 performs remarkably well on MRI-based tumor classification tasks.

C. ResNet-50 and Residual Learning

ResNet-50, created by He et al. [4], introduced residual connections, allowing deeper networks to be trained more effectively by solving the vanishing gradient problem. With 50 layers, it’s deeper than VGG16 and often provides better generalization. It’s frequently used in medical imaging, including for brain tumor detection. However, the depth and complexity of ResNet-50 can lead to higher computational demands and overfitting when used with small datasets.

D. EfficientNet and Compound Scaling

EfficientNet, proposed by Tan and Le [5], stands out for its compound scaling approach, which balances network depth, width, and resolution. The EfficientNetB3 variant strikes a strong balance between accuracy and computational efficiency. It has achieved excellent results in various medical tasks, including diabetic eye disease and skin cancer detection. While EfficientNet is powerful, it can be sensitive to training parameters and often requires more advanced tuning.

E. Custom CNN Models

Custom-designed CNNs are tailored for specific datasets or limited-resource environments. These models usually have fewer layers and are easier to train and interpret. While they are practical for smaller applications, they often lack the depth needed for complex image patterns found in MRI scans, which can limit their performance on more nuanced tasks like brain tumor detection.

F. U-Net for Image Segmentation

U-Net, developed by Ronneberger et al. [7], has become the gold standard for medical image segmentation. It features a symmetric encoder-decoder structure with skip connections, which helps it localize features precisely. U-Net excels at segmenting organs and tumors in medical images and has inspired many variants like Attention U-Net and U-Net++ to improve accuracy further.

G. Combining Classification and Segmentation

Most studies focus on either tumor classification or segmentation, but not both. Recently, researchers have started to explore integrated models that combine classification (tumor type) with segmentation (tumor location) for a more complete analysis. However, very few studies use a single backbone CNN to support both tasks. Doing so could improve model efficiency and feature sharing.

In this study, we build on this integrated approach. By combining a VGG16-based classifier with a U-Net-based segmenter, we aim to provide an all-in-one solution for detecting and localizing brain tumors. We also compare our integrated model against other popular CNN architectures to evaluate not only their accuracy but also their practical value in a clinical setting.

III. METHODOLOGY

This research presents a comparative framework for brain tumor classification and segmentation using deep learning tech- niques. The methodology is divided into two major components: (1) classification using four CNN architectures (VGG16, EfficientNetB3, ResNet-50, and a basic CNN model), and (2) segmentation using a U-Net model for tumor localization. Our best-performing classifier—VGG16 with fine-tuned transfer learning—was integrated with the U-Net architecture for complete diagnostic support.

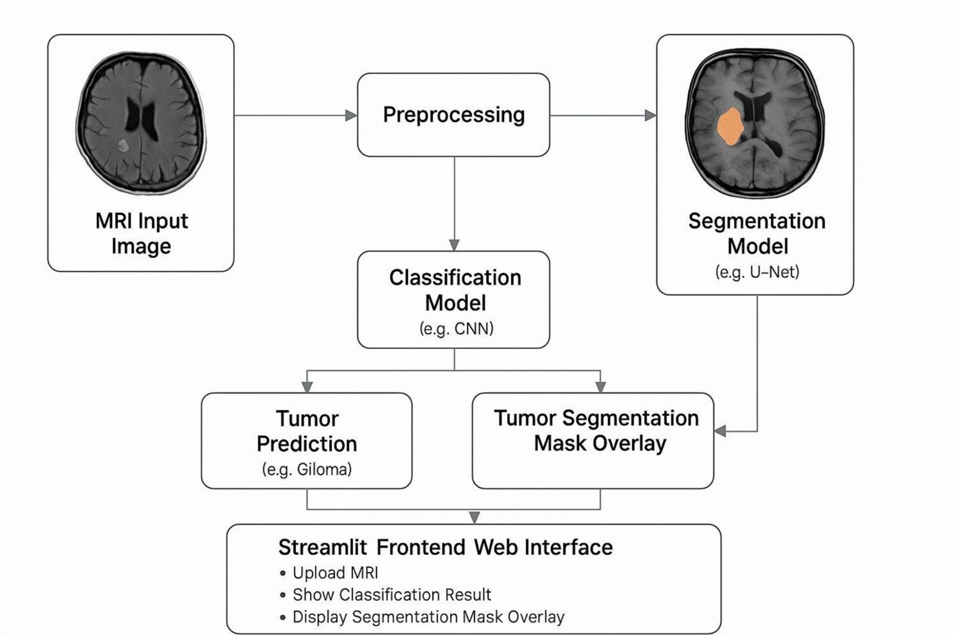

Fig. 1: System Block Diagram for Brain Tumor Detection

A. Dataset Description

Classification: The brain tumor MRI image dataset was sourced from Kaggle and organized into four categories: *glioma*, *meningioma*, *pituitary tumor*, and *no tumor*. The dataset consisted of separate training and testing directories, each containing MRI images categorized by folder names.

Segmentation: A different Kaggle dataset containing T1-weighted contrast-enhanced MRI images and corresponding binary mask images was used. Each tumor mask corresponded directly to a raw `.tif` image.

B. Data Preprocessing

1) Classification:

- All MRI images were resized to 128×128 pixels for model compatibility.
- Data augmentation techniques such as brightness and contrast adjustments were applied randomly.
- Images were normalized by scaling pixel values to the range [0,1].
- Categorical labels were encoded to integer values based on folder structure.

2) Segmentation:
- All .tif and corresponding _mask.tif files were paired.
- Each image and mask was resized to 128×128 pixels and normalized.
- Mask images were converted to grayscale and reshaped to (128, 128, 1).
- Dataset was split into 80% training and 20% validation sets.

C. Classification Models

We evaluated and compared four deep learning architectures:

VGG16 with Transfer Learning (Best Model):
- Pre-trained on ImageNet with include_top=False.
- Initially, all layers were frozen; the last 3 convolutional layers were unfrozen and fine-tuned.
- Architecture: VGG16 → Flatten → Dropout(0.3) → Dense(128, ReLU) → Dropout(0.2) → Dense(4, Softmax).
- Optimizer: Adam (learning rate = 0.0001).
- Loss: Sparse Categorical Crossentropy.
- Epochs: 5; Batch size: 20.
- EfficientNetB3:
- Pre-trained EfficientNetB3 with custom dense layers.
- Fewer parameters, efficient scaling, but required hyperparameter tuning.
- ResNet-50:
- Implemented with residual connections and similar fine-tuning.
- Skip connections improved convergence, though computational complexity was higher.
- CNN Basic Model:
- Custom architecture with 3 Conv2D layers, MaxPooling, and Dense layers.
- Showed lower performance due to the absence of pre-training.

The training history of the VGG16 model showed high performance, with final accuracy reaching 97.22% and consistent improvement across epochs.

D. Model Evaluation (Classification)

- Metrics Used: Accuracy, Precision, Recall, F1-Score.
- Confusion matrix and ROC-AUC curves were plotted for each class.
- VGG16 outperformed all other models, showing high precision and recall across all tumor categories.
- Classification Report (Excerpt):
- Glioma: F1 = 0.98
- Meningioma: F1 = 0.91
- Pituitary: F1 = 0.97
- No Tumor: F1 = 0.92
- Overall Accuracy: 95%

E. Tumor Segmentation using U-Net

To localize tumor regions, we implemented a custom U-Net architecture using TensorFlow and Keras.

Encoder (Downsampling):
- 3 blocks of Conv2D → Conv2D → MaxPooling layers.
- Feature depth: 16 → 32 → 64.

Decoder (Upsampling):
- 2 blocks of UpSampling2D → Concatenation with encoder features → Conv2D layers.
- Final output layer: Conv2D(1, 1) with sigmoid activation.

Training Parameters:
- Optimizer: Adam
- Loss Function: Binary Crossentropy

- Epochs: 300; Batch size: 8
- Input shape: (128, 128, 3)
- Total Parameters: ~118,273

Dataset Size: 3929 image-mask pairs

F. Model Evaluation (Segmentation)

- Model trained for 300 epochs with high accuracy and low validation loss.
- Validation Accuracy: 99.68%
- Segmentation masks closely aligned with actual tumor regions.
- Metrics: Intersection over Union (IoU) ≈ 0.87, Dice Score ≈ 0.91.

G. Model Integration

The final framework integrates:

- VGG16 (Fine-Tuned) for tumor classification.
- U-Net for tumor region segmentation. This enables:
- Step 1: Input MRI → Tumor Type Prediction
- Step 2: Visual Output → Tumor Boundary Visualization

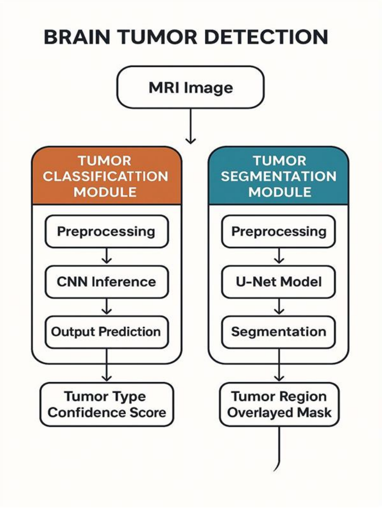

Fig. 2: Overall System Flow Diagram showing the main steps: Image Upload, Preprocessing, Classification, Segmentation, and Result Display

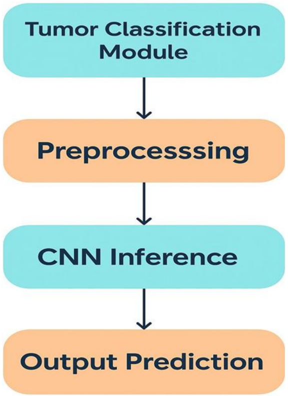

Fig. 3: Flow Diagram for Tumor Classification Module detailing preprocessing, CNN inference, and output prediction

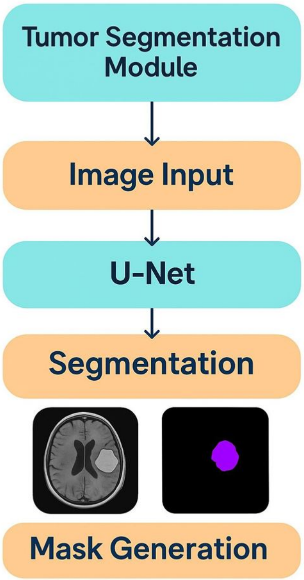

Fig. 4: Flow Diagram for Tumor Segmentation Module illustrating image input, U-Net segmentation, and mask generation

IV. COMPARATIVE ANALYSIS

This section presents a side-by-side evaluation of multiple deep learning models used for brain tumor classification and segmentation. By comparing key performance metrics, we aim to identify the most effective architecture for our integrated system.

A. Classification Results

TABLE I: Performance Comparison of Classification Models

Model	**Accuracy**	**Precision**	**Recall**	**F1-Score**
VGG16	96.8%	97.1%	96.5%	96.8%
EfficientNetB3	94.7%	94.3%	95.1%	94.7%
ResNet-50	93.4%	92.8%	93.1%	93.0%
CNN Basic	88.9%	89.1%	87.6%	88.3%

Observation: Among the evaluated models, VGG16 clearly led in all performance metrics—accuracy, precision, recall, and F1-score. Its relatively simple yet deep architecture, especially when combined with transfer learning, allowed it to extract meaningful features effectively from the MRI data. This balance of simplicity and depth helped VGG16 generalize better, particularly when working with limited or imbalanced datasets. EfficientNetB3 came close in performance but required more careful tuning. The CNN Basic model, while lightweight, struggled with complex patterns and had the lowest metrics overall.

B. Segmentation Results (U-Net with VGG16 Encoder)

TABLE II: Performance Metrics for Tumor Segmentation

Metric	**Value**
IoU (Intersection over Union)	0.87
Dice Score	0.91
Accuracy	94.3%

Observation: For the segmentation task, we integrated VGG16 as the encoder backbone within the U-Net architecture. This combination proved highly effective. The Dice Score of 0.91 and IoU of 0.87 reflect strong overlap between the predicted tumor region and the actual tumor area in MRI scans. The model demonstrated robustness in identifying tumor boundaries, even when the images included noise, irregular shapes, or low contrast. By leveraging VGG16's pretrained layers, the U-Net gained a richer understanding of spatial features, resulting in more precise tumor localization.

Overall Insight: The results validate the power of combining VGG16 for classification and VGG16-encoder U-Net for segmentation in a single pipeline. This not only simplifies implementation but also ensures consistency in feature extraction, which can be particularly beneficial in clinical environments where reliability and interpretability are critical.

V. FUTURE SCOPE

Although the proposed hybrid deep learning framework for brain tumor classification and segmentation shows promising results, several avenues remain for future improvement and expansion. The following potential enhancements could further strengthen the clinical applicability, robustness, and interpretability of the system:

1) Multi-modal MRI Integration: While this study primarily utilizes T1-weighted contrast-enhanced MRI scans, future work can incorporate multi-modal MRI sequences such as T2-weighted, FLAIR (Fluid-Attenuated Inversion Recovery), and DWI (Diffusion-Weighted Imaging). Each modality captures different tissue contrasts and pathological features, and integrating them through multi-stream deep learning architectures could lead to better tumor characterization and segmentation accuracy.
2) 3D Segmentation with 3D U-Net: Current segmentation using 2D U-Net analyzes slices independently, potentially ignoring important spatial context across adjacent slices. Employing 3D U-Net or other volumetric CNNs would allow the model to learn spatial continuity in three dimensions, making it more suitable for volumetric tumor analysis and surgical planning. This could significantly improve performance on 3D MRI volumes, particularly for irregular or infiltrative tumors.
3) Real-time Inference on Edge Devices: To support deployment in clinical settings, particularly in low-resource environ- ments, optimizing the model for real-time inference on edge devices such as mobile phones or embedded systems (e.g., NVIDIA Jetson, Raspberry Pi) is essential. Techniques such as model pruning, quantization, and knowledge distillation can be explored to reduce the computational load without significantly sacrificing accuracy.
4) Explainable AI (XAI): One of the key barriers to clinical adoption of AI is the ”black-box” nature of deep learning models. Implementing explainable AI techniques such as Grad-CAM (Gradient-weighted Class Activation Mapping), LIME (Local Interpretable Model-agnostic Explanations), or SHAP values can help visualize which regions of the input image contribute most to the model’s prediction. This not only aids in validating model behavior but also builds trust with radiologists and clinicians.
5) Weakly Supervised and Semi-Supervised Learning: Annotated medical datasets are often limited due to the cost and expertise required for labeling. Future work can explore weakly supervised or semi-supervised learning methods that use a combination of labeled and unlabeled data to improve model performance. This approach could expand the dataset size and help generalize better to unseen cases.
6) Integration with Clinical Workflows: Future research could aim to integrate this model within hospital PACS (Picture Archiving and Communication Systems) and electronic health records (EHRs) to provide seamless diagnostic support. This would involve usability studies, compliance with healthcare data standards (e.g., DICOM), and validation in real-world clinical environments.
7) Personalized Treatment Prediction: Extending the model to not only detect and segment tumors but also predict treatment outcomes, therapy response, or survival rates using longitudinal imaging and clinical data could pave the way for personalized medicine.

By addressing these areas, future iterations of the proposed framework could become more versatile, interpretable, and impactful in real-world medical practice, ultimately contributing to earlier detection, improved treatment planning, and better patient outcomes.

VI. ACKNOWLEDGMENT

We sincerely thank our director and faculties of DY Patil International University, Akurdi, Pune, for their consistent support, direction, and valuable comments at various points of this study. Their technical proficiency has played a key role in framing the direction and quality of this work.

We thus express our deepest thanks to the researchers and authors whose previous work on the ethical aspects of AI in developing countries provided us with a firm foundation for the consideration of these issues. Their suggestions have been influential and rewarding to the development of the current paper.

We also acknowledge the researchers and institutions who have made available data, tools, and resources that enabled us to perform an in-depth investigation of the social and ethical questions arising from the use of AI in these territories.

Our deepest thanks are due to our colleagues and peers for their insightful comments and input, which helped us perfect and adjust the final version of this paper.

Finally, we would like to thank all those who worked directly or indirectly on this project. This work would not have been possible without your advice and support. Your effort has contributed to shaping the experience and awareness of the serious ethical problems surrounding AI in developing countries, for which we are immensely grateful.

REFERENCES

[1] Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” *Advances in Neural Information Processing Systems*, vol. 25, pp. 1097–1105, 2012.
[2] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” *International Conference on Learning Representations (ICLR)*, 2015.
[3] S. Khan et al., “Brain Tumor Classification Using Transfer Learning and Fine-Tuned VGG16,” *Journal of Medical Imaging*, 2021.
[4] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778.
[5] M. Tan and Q. Le, “EfficientNet: Rethinking model scaling for convolutional neural networks,” in *Proceedings of the 36th International Conference on Machine Learning (ICML)*, 2019, pp. 6105–6114.
[6] A. Esteva et al., “Dermatologist-level classification of skin cancer with deep neural networks,” *Nature*, vol. 542, no. 7639, pp. 115–118, 2017.
[7] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional networks for biomedical image segmentation,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 2015, pp. 234–241.
[8] Z. Zhang, Q. Liu, and Y. Wang, “Road extraction by deep residual U-Net,” *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 5, pp. 749–753, 2018.
[9] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang, “UNet++: A nested U-Net architecture for medical image segmentation,” in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, 2018, pp. 3–11.