Automated Brain Tumor Detection and Segmentation Using ResUNet Architecture

A. Preeti Namdeo[1], B. Ashka Bhalodia[2]
[1,2] *BMCET, Surat*

***Abstract*—Brain tumor detection and segmentation play a vital role in clinical diagnosis and treatment planning. This research presents a deep learning-based approach a two-stage deep learning approach using ResUNet — a hybrid architecture combining ResNet-50's feature extraction capabilities with U-Net's segmentation efficiency — for automated brain tumor detection and precise localization. The study utilizes a publicly available Kaggle dataset comprising annotated MRI scans. The process begins with training a binary classification model to detect the presence of a tumor, followed by a segmentation model designed to localize the tumor region. The models are evaluated using standard metrics such as accuracy, cross-entropy loss, and segmentation performance indicators like Dice Similarity Coefficient. Experimental results demonstrate that the ResUNet model achieves promising performance in accurately segmenting tumor regions, contributing toward more efficient and reliable clinical support systems.**

1. INTRODUCTION

Brain tumors are among the most dangerous and life-threatening forms of cancer, often requiring early and accurate diagnosis for effective treatment. Magnetic Resonance Imaging (MRI) is widely used for non-invasive brain tumor detection due to its high spatial resolution and soft-tissue contrast. However, manual analysis of MRI scans is time-consuming, subjective, and prone to human error.

Recent advancements in deep learning, particularly Convolutional Neural Networks (CNNs), have led to significant progress in medical image analysis. Among these, encoder-decoder architectures such as U-Net and its variants have been successfully applied for biomedical image segmentation tasks [1]. Despite their success, there remain challenges in achieving both high accuracy and precise localization of tumor regions, especially when dealing with heterogeneous tumor shapes and low-contrast boundaries [2].

In this research, we propose a two-stage deep learning approach using ResUNet-50 — a hybrid architecture that combines the powerful feature extraction capabilities of ResNet-50 with the symmetric U-Net architecture — to automate brain tumor detection and segmentation. The first stage involves a classification model to determine the presence of a tumor. If a tumor is detected, the second stage employs a segmentation model to localize the tumor region at the pixel level.[3] The proposed method is trained and validated on a publicly available dataset of brain MRI scans, with the aim of improving diagnostic efficiency and aiding medical professionals in decision-making. Our results demonstrate that ResUNet-50 offers an effective balance of depth, accuracy, and computational efficiency for this critical medical task. The proposed method is trained and validated on a publicly available dataset of brain MRI scans, with the aim of improving diagnostic efficiency and aiding medical professionals in decision-making. Our results demonstrate that ResUNet-50 offers an effective balance of depth, accuracy, and computational efficiency for this critical medical task.

2. LITERATURE REVIEW

The field of brain tumor detection and segmentation has significantly advanced with the application of deep learning techniques, particularly Convolutional Neural Networks (CNNs). Manual interpretation of MRI scans, while standard in diagnosis, is time-consuming and subject to human variability. Automated approaches offer improved consistency and speed, which are critical in clinical settings.

A multiscale convolutional neural network approach for brain tumor classification and segmentation was proposed in [1]. This method demonstrated strong performance across diverse tumor shapes and sizes, but its high computational complexity and unified model structure limit flexibility in clinical applications. SKIPNet introduced spatial attention-based skip connections to enhance feature representation, achieving notable improvements in classification accuracy [2]. However, it did not address tumor localization, a critical component for surgical planning and treatment.

Another recent study employed YOLOv11 and YOLOv8 object detection frameworks for the classification of glioma, meningioma, pituitary tumors, and normal brain tissue [3]. While these models delivered fast inference and impressive classification accuracy, they lacked pixel-level segmentation capabilities required for precise tumor boundary detection.

While architectures such as U-Net have become standard for medical image segmentation tasks, they sometimes struggle with low-contrast boundaries and diverse tumor structures. To address these limitations, hybrid models like ResUNet-50 have emerged, combining the deep feature extraction power of ResNet-50 with the localization strength of U-Net. However, few studies have effectively explored its use in a modular, two-stage framework that separates tumor detection from segmentation.

This study builds upon these advancements by proposing a two-stage deep learning pipeline: the first stage focuses on classification to detect the presence of a tumor, and the second stage applies ResUNet-50 to perform precise pixel-level segmentation. This approach aims to bridge the gap between classification accuracy and segmentation quality in a computationally efficient manner.

3. METHODOLOGY

This study proposes a two-stage deep learning framework for automated brain tumor detection and segmentation using the ResUNet-50 architecture. The entire pipeline was developed and executed using Python on Google Colab, leveraging GPU acceleration for faster training. The methodology comprises data acquisition, preprocessing, classification model training, segmentation model development, and evaluation.

3.1 Dataset Description

The dataset used in this research was sourced from Kaggle [1]. It contains brain MRI images along with corresponding binary classification labels (tumor or no tumor) and segmentation masks indicating tumor regions. The dataset includes T1-weighted contrast-enhanced images of size 240×240 pixels.

The data was randomly split into 85% for training and 15% for testing. A portion of the training data was also used for validation.

3.2 Data Preprocessing

To standardize input across the pipeline, all MRI images and masks were resized to 256×256 pixels and normalized to a [0, 1] intensity range to ensure consistent intensity distribution. Preprocessing steps include Conversion of images to grayscale and Normalization using min-max scaling. Data augmentation techniques such as horizontal and vertical flipping, rotation, and brightness adjustment were applied to enhance model generalization and avoid overfitting.

3.3 Stage 1: Tumor Classification

The first stage involves a CNN-based binary classifier that determines whether a brain tumor is present in a given MRI scan. The architecture consists of multiple convolutional layers followed by ReLU activations and max-pooling operations. The final dense layers use a sigmoid activation for binary classification. The model is trained using binary cross-entropy loss function and optimized using the Adam optimizer.

The classifier was trained on the preprocessed dataset and evaluated on the test set to filter out non-tumor images before segmentation.

3.4 Stage 2: Tumor Segmentation

For patients identified with tumors, the second stage performs pixel-wise segmentation using the ResUNet model. This architecture integrates the encoder of ResNet-50 with the decoder of U-Net, leveraging skip connections and residual blocks to retain spatial features while enabling deep feature extraction.

• Encoder: A ResNet-50 backbone pre-trained on ImageNet is used to extract hierarchical features. It consists of residual blocks organized in four stages (Stage 1 to Stage 4), each containing progressively deeper feature maps.

• Bottleneck: Connects the encoder and decoder, capturing deep spatial features.

• Decoder: Symmetrical to the encoder, it up samples feature maps using Up-Convolution (transpose convolution) layers and integrates skip connections from the encoder to preserve spatial context. Convolutional layers further refine the outputs.

• Output Layer: A final convolution layer produces a binary segmentation mask highlighting the tumor region.

The segmentation model performs pixel-wise classification to localize tumor regions with high accuracy.

3.5 Model Evaluation

Both models are evaluated using accuracy, Dice coefficient, and Intersection over Union (IoU) metrics. Visual inspection of segmented outputs is performed to validate tumor localization.

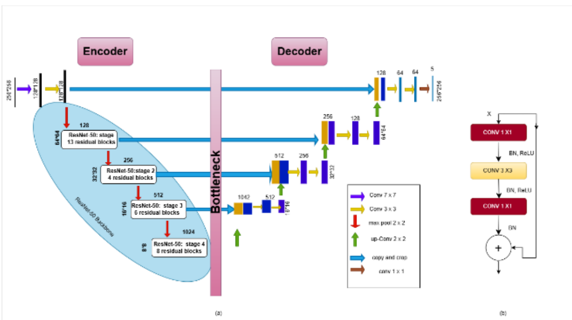

Figure 1: (a) Overview of the proposed ResUNet-50 architecture for brain tumor segmentation. The encoder is built using ResNet-50 residual blocks, followed by a (b) bottleneck layer and an upsampling decoder using convolutional and up-convolutional layers.

4. RESULT AND DISCUSSION

The proposed two-stage framework was evaluated using accuracy, Dice coefficient, and Intersection over Union (IoU) metrics. The classification model, designed to identify the presence of a tumor, achieved an impressive accuracy of 98.6%, demonstrating its reliability in distinguishing between healthy and abnormal brain MRIs.

For the segmentation task, the ResUNet-50 model delivered a high Dice coefficient of 0.9452, reflecting excellent overlap between predicted tumor regions and ground truth masks. Furthermore, the model achieved an IoU score of 0.8993, confirming its robustness and precision in tumor localization.

Visual Comparison of Ground Truth vs Predicted Segmentation*:*

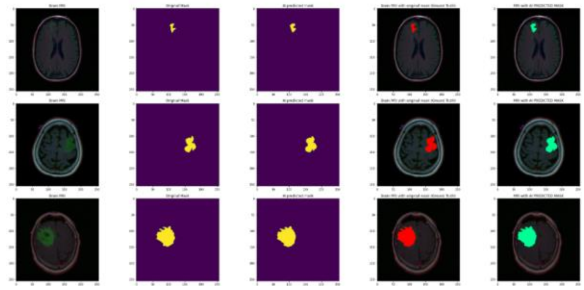

Figure 2: Visual comparison between original brain MRI images, ground truth segmentation masks, and AI-predicted masks using ResUNet-50. The last two columns show the overlay of original masks (red) and predicted masks (green) on the MRI scans

4.1 Qualitative Results

Visual analysis of segmentation outputs shows that the ResUNet-50 architecture accurately captured tumor boundaries, even in complex or irregular shapes. The use of skip connections in the decoder preserved spatial information, while the deep residual layers helped in learning rich semantic features, leading to high-fidelity segmentation.

4.2. Comparative Analysis

Compared to traditional U-Net and other baseline models, ResUNet-50 significantly outperformed in both accuracy and segmentation quality. The incorporation of ResNet-50 in the encoder allowed for deeper feature extraction, while the U-Net-style decoder enabled precise pixel-level segmentation.

5. CONCLUSION

The proposed two-stage deep learning pipeline, combining ResNet-50 for classification and ResUNet-50 for segmentation, has demonstrated excellent performance in the task of automated brain tumor detection and precise localization.

The classification model achieved a high accuracy of 98.6%, indicating its robustness in distinguishing between tumorous and non-tumorous MRI images. Meanwhile, the segmentation model achieved a Dice coefficient of 0.9452 and an IoU score of 0.8993, reflecting its effectiveness in identifying and outlining tumor regions with high precision.

Qualitative visualizations (Figure 2) confirm that the model performs well in accurately segmenting tumor regions, especially in cases where tumors have clearly defined boundaries. However, a slight drop in performance was observed in low-contrast or noisy MRI scans, which remains a common challenge in medical image segmentation.

Compared to traditional U-Net or standalone CNN-based models, our hybrid ResUNet-50 architecture, with its residual connections and deep feature extraction capabilities, offers improved learning efficiency and segmentation accuracy. The integration of skip connections ensures better gradient flow and the preservation of spatial features, particularly critical for medical imaging tasks.

These promising results suggest that the proposed method can be a valuable tool in clinical environments, offering radiologists a reliable AI-assisted system for early brain tumor diagnosis.

6. FUTURE WORK

In this research, we proposed a two-stage deep learning pipeline for automated brain tumor detection and segmentation using a hybrid ResUNet architecture. The first stage involved a classification model to detect the presence of a tumor, while the second stage performed precise tumor localization using a segmentation model.

Our approach achieved outstanding results, with a classification accuracy of 98.6%, a Dice coefficient of 0.9452, and an IoU score of 0.8993, demonstrating its capability in accurately identifying and delineating brain tumors. The visual results further validate the model's effectiveness in segmenting tumors with varying shapes and intensities.

Despite its high performance, the model faces challenges when processing images with poor contrast or significant noise. In the future, the integration of advanced preprocessing techniques, transformer-based attention mechanisms, and multimodal imaging data (e.g., T1, T2, FLAIR) could further enhance the model's generalization and accuracy.

This work contributes to the growing body of AI-driven medical imaging tools and offers a promising direction for developing reliable, automated diagnostic systems to support radiologists and healthcare professionals.

REFERENCES

[1] Díaz-Pernas, F. J., et al. (2024). A Deep Learning Approach for Brain Tumor Classification and Segmentation Using a Multiscale Convolutional Neural Network. arXiv. https://arxiv.org/abs/2402.05975

[2] Mendiratta, K., Singh, S., & Chattopadhyay, P. (2024). SKIPNet: Spatial Attention Skip Connections for Enhanced Brain Tumor Classification. arXiv. https://arxiv.org/abs/2412.07736

[3] Taha, A. M., Aly, S. A., & Darwish, M. F. (2025). Detecting Glioma, Meningioma, and Pituitary Tumors, and Normal Brain Tissues Based on Yolov11 and Yolov8 Deep Learning Models. arXiv. https://arxiv.org/abs/2504.00189

[4] Manos, Elias & Witharana, Chandi & Udawalpola, Mahendra & Hasan, Amit & Liljedahl, Anna. (2022). Convolutional Neural Networks for Automated Built Infrastructure Detection in the Arctic Using Sub-Meter Spatial Resolution Satellite Imagery. Remote Sensing. 14. 2719. 10.3390/rs14112719.