

# Semantic Segmentation of Brain Tumors in MRI Scans using Deep Learning Architectures

Akshitha Baira, Vihas Adi, Naveen Kumar Chekuri

## Abstract

Brain tumors are life-threatening conditions that require precise diagnosis and treatment. Manual segmentation of brain tumors in MRI images is time-consuming, labor-intensive, and prone to human error. This project presents an automated solution using deep learning-based semantic segmentation to identify and isolate tumor regions in brain MRI scans. We implemented and compared four popular architectures: U-Net, Attention U-Net, Fully Convolutional Networks (FCN), and SegNet. A custom dataset was manually split and preprocessed, and each model was trained for 30 epochs using a combination of Dice loss and Binary Cross-Entropy loss. The evaluation was done using Dice Score, Intersection over Union (IoU), and Pixel Accuracy. Our results show that the FCN model outperforms the others by learning to focus on the most relevant areas of the image, providing more accurate and reliable segmentations. This work shows the effectiveness of attention-based models in medical image analysis and provides a strong baseline for future improvements.

## 1 Introduction

Brain tumors are abnormal growths of tissue in the brain, and their early detection is crucial for effective treatment. One of the most common diagnostic tools for identifying brain tumors is magnetic resonance imaging (MRI). However, analyzing and segmenting tumors from MRI scans manually can be a long and subjective process, which is not ideal in clinical settings. In recent years, deep learning has shown significant promise in automating this task, especially through semantic segmentation — where every pixel in the image is classified as tumor or non-tumor.

This project aims to explore and compare the performance of four deep learning models for the task of brain tumor segmentation: U-Net, Attention U-Net, FCN, and SegNet. U-Net is widely known for its success in biomedical image segmentation due to its encoder-decoder structure. Attention U-Net builds upon this by adding attention gates that help the model focus on important regions. FCN is one of the earliest convolutional models designed for segmentation, and SegNet introduces an efficient decoding mechanism that uses pooling indices from the encoder. By comparing these models on a common dataset,

we aim to identify which approach yields the most accurate results and why.

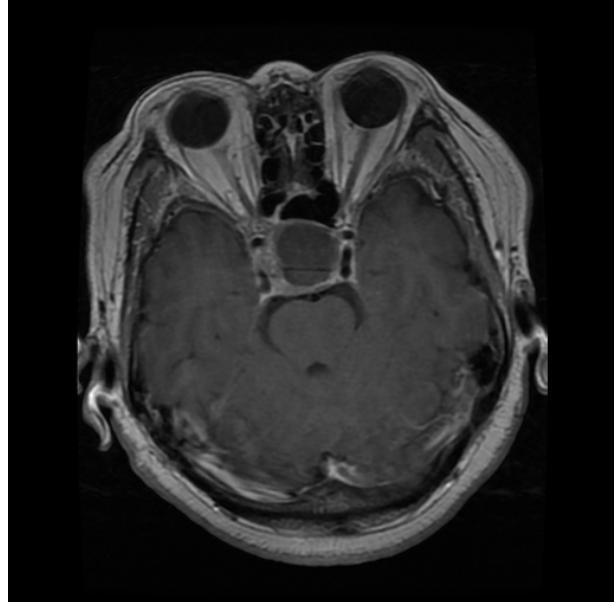


Figure 1: Introduction

## 2 Related Work

Semantic segmentation in medical imaging has gained significant attention since the introduction of U-Net by Ronneberger et al., which became a foundational architecture for biomedical segmentation tasks. U-Net's design allows the model to capture both high-level contextual features and low-level spatial details using skip connections between encoder and decoder layers.

To improve on U-Net, Oktay et al. introduced the Attention U-Net, which integrates attention mechanisms into the skip connections. These attention gates help the model suppress irrelevant background features and focus more on important structures like tumors.

Fully Convolutional Networks (FCNs), proposed by Long et al., replaced fully connected layers with convolutional layers, enabling end-to-end training for segmentation tasks. FCNs use upsampling layers to recover the original image resolution, making them one of the first practical solutions for semantic segmentation.

SegNet, introduced by Badrinarayanan et al., focuses

on memory efficiency and fast inference by storing max-pooling indices from the encoder and reusing them during decoding. Although it's lightweight, SegNet can perform well in real-time applications.

These models form the basis of our comparative study, where we assess their performance on brain MRI tumor segmentation using a consistent training and evaluation setup. In addition to model architecture improvements, the use of multi-modal imaging, attention mechanisms, transformer-based models, and self-supervised learning techniques have recently shown promising results in medical segmentation tasks, particularly in complex datasets such as BraTS [7]. These advancements highlight the continuous evolution of segmentation methodologies, driving toward more accurate, efficient, and clinically applicable solutions.

### 3 Dataset and Preprocessing

The dataset used in this study consists of brain magnetic resonance imaging (MRI) scans paired with manually annotated binary segmentation masks that indicate tumor regions. Each image represents a cross-sectional view of the human brain captured using non-invasive imaging techniques, primarily T1-weighted or T2-weighted MRI sequences. The dataset includes cases with visible brain tumors as well as cases without tumors, allowing the models to learn both the presence and absence of pathological regions. The corresponding masks highlight tumor areas at the pixel level, enabling supervised training for semantic segmentation.

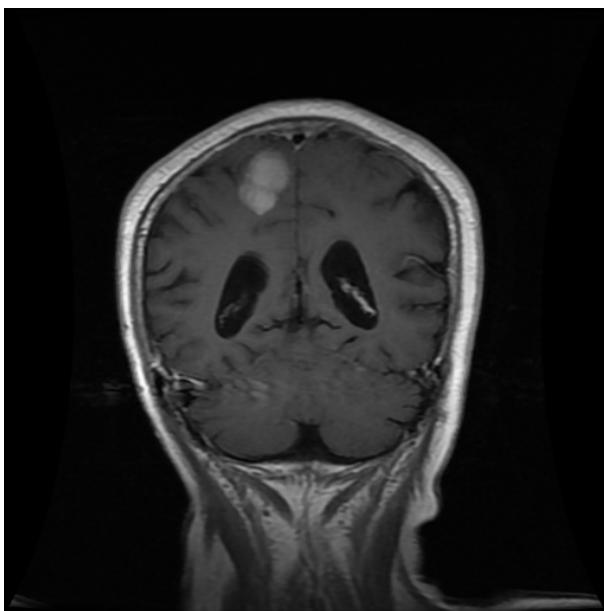


Figure 2: Brain Tumor Masks

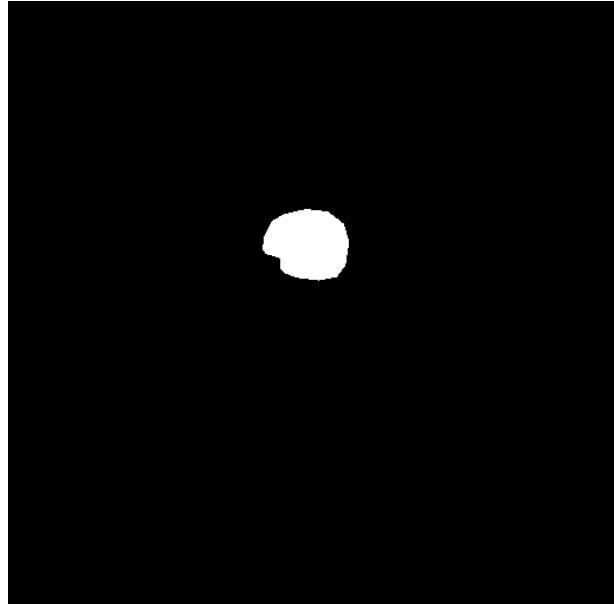


Figure 3: Brain Tumor Masks

These masks are essential for teaching the model to distinguish between healthy brain tissue and abnormal growths. The data captures various tumor shapes, sizes, and locations, offering diversity that is crucial for model generalization. While relatively small in scale compared to large public datasets like BraTS, this dataset still provides a realistic and challenging benchmark for evaluating segmentation performance in a medical imaging context. Its focus on binary classification (tumor vs. non-tumor) simplifies the segmentation problem while maintaining clinical relevance.

### 4 Methodology

In this study, we implemented and compared four deep learning architectures for semantic segmentation of brain tumors: U-Net, Attention U-Net, Fully Convolutional Network (FCN), and SegNet. Each model follows an encoder-decoder structure but incorporates different mechanisms for feature extraction, spatial information retention, and upsampling.

#### 4.1 U-Net

U-Net is a convolutional neural network architecture specifically designed for biomedical image segmentation tasks. It was introduced by Ronneberger et al. in 2015 and has since become the gold standard for many medical imaging problems due to its ability to deliver precise, pixel-level segmentations even on small datasets.

The architecture follows a “U-shaped” design, consisting of two symmetrical paths: a contracting path (encoder) and an expanding path (decoder). The encoder is composed of repeated blocks of convolutional layers followed by max pooling operations, which progressively

reduce the spatial dimensions of the image while capturing high-level semantic features. As the network moves deeper, it becomes better at identifying abstract patterns but loses fine-grained details.

To recover spatial resolution and produce accurate segmentation maps, the decoder performs upsampling using transposed convolutions. What makes U-Net particularly effective is its use of skip connections: at each level, feature maps from the encoder are directly concatenated with the corresponding upsampled feature maps in the decoder. These skip connections act as shortcuts that reintroduce low-level spatial information lost during downsampling, helping the network reconstruct fine boundaries and structures more accurately.

Another key advantage of U-Net is its ability to work well with relatively small datasets — a common challenge in medical imaging. By leveraging data augmentation and using symmetric padding, U-Net can generalize effectively while maintaining the context needed for precise segmentation.

In this project, U-Net was implemented using multiple convolutional blocks, ReLU activations, batch normalization, and max pooling. It was trained using a combined loss function (BCEWithLogits + Dice Loss) and showed strong performance in accurately identifying tumor regions in MRI scans.

## U-NET ARCHITECTURE

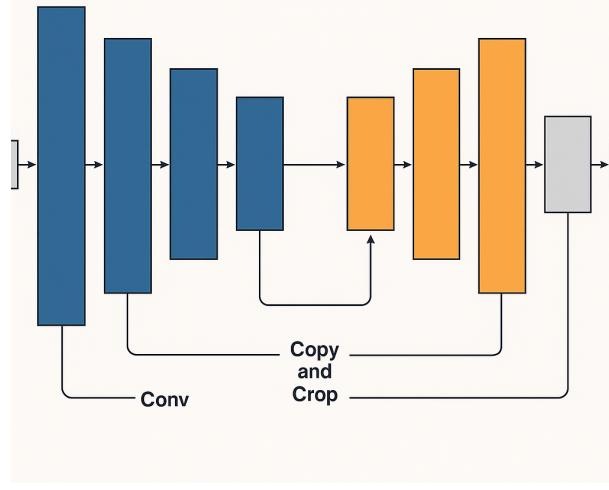


Figure 4: U-Net Architecture

## 4.2 Attention U-Net

Attention U-Net builds upon the U-Net architecture by integrating attention gates into the skip connections. These attention mechanisms were introduced to help the model focus on the most relevant parts of the input image while suppressing irrelevant background noise. In standard U-Net, skip connections pass all encoder features to

the decoder, regardless of their importance. In contrast, Attention U-Net uses a gating mechanism to weigh encoder features before passing them to the decoder. This gating function is conditioned on the decoder's output, meaning it learns to pass through only features that are relevant for the current segmentation task. This modification allows the model to learn “where to look” in the image, improving the localization of small or complex tumor structures. This becomes particularly useful in medical applications where tumor regions may be subtle and easily overlooked.

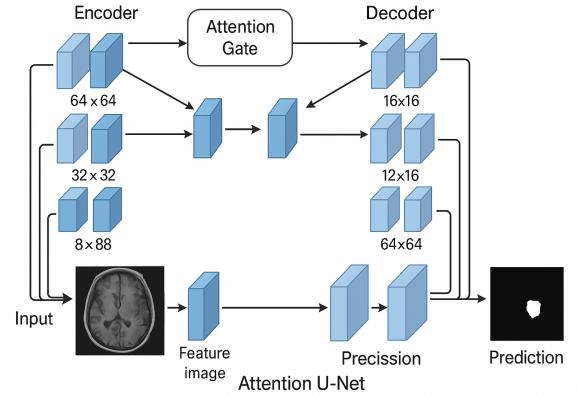


Figure 5: Attention U-Net Architecture

## 4.3 Fully Convolutional Network (FCN)

Fully Convolutional Networks (FCNs) were introduced by Long et al. in 2015 as one of the first deep learning architectures designed explicitly for semantic segmentation tasks. Unlike traditional classification networks that reduce input images to a single label, FCNs are capable of dense pixel-wise predictions, assigning a class to every pixel in the image.

The key innovation of FCNs lies in the elimination of fully connected layers, which are replaced with  $1 \times 1$  convolutional layers. This conversion allows the model to maintain spatial hierarchies throughout the network and to accept variable input sizes. The network produces coarse segmentation maps from deep features and then upsamples them to match the original image resolution.

In our implementation, we used FCN with a ResNet-50 backbone. ResNet layers serve as the encoder, extracting high-level semantic features from the input image. These features are then passed through a classifier composed of  $1 \times 1$  convolutions to produce a low-resolution segmentation map. To recover the original image size, the network uses bilinear interpolation or learned transposed convolutions. While this simple upsampling mechanism is efficient, it also means that FCN models can sometimes produce blurry or coarse boundaries due to the lack of detailed spatial information from earlier layers.

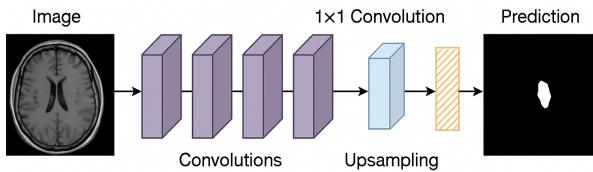


Figure 6: Fully Convolutional Network (FCN) Architecture

Unlike U-Net or SegNet, FCN does not incorporate skip connections, which limits the model’s ability to restore fine-grained spatial detail. As a result, FCNs are generally good at identifying broad object regions but may struggle with the precise delineation of object edges — a critical factor in medical image segmentation where boundary accuracy is essential.

Despite these limitations, FCNs are computationally efficient and scalable. They laid the foundation for most modern segmentation networks and remain a relevant benchmark for performance comparison in semantic segmentation tasks. Despite these advancements, FCNs have certain limitations, especially when compared to architectures like U-Net. FCNs often produce coarse and blurry segmentations, particularly at object boundaries, due to their heavy reliance on upsampling from deeply encoded, low-resolution feature maps. Moreover, even with the addition of skip connections, FCNs struggle to reconstruct fine spatial details, making them less suited for applications where boundary precision is critical, such as medical image segmentation.

#### 4.4 SegNet

SegNet is an encoder-decoder architecture developed for semantic segmentation with a focus on computational efficiency and memory optimization. The encoder in SegNet uses a sequence of convolutional layers followed by max-pooling, similar to VGG16. However, instead of using learnable upsampling, SegNet stores the max-pooling indices during encoding and reuses them during decoding to upsample the feature maps. This non-learnable up-sampling process reduces the number of parameters and avoids the use of transposed convolutions. While SegNet may not achieve the same level of accuracy as more complex models like Attention U-Net, it performs well in real-time scenarios due to its lightweight design. Its main drawback is the lack of skip connections, which limits the flow of fine-grained spatial information from encoder to decoder.

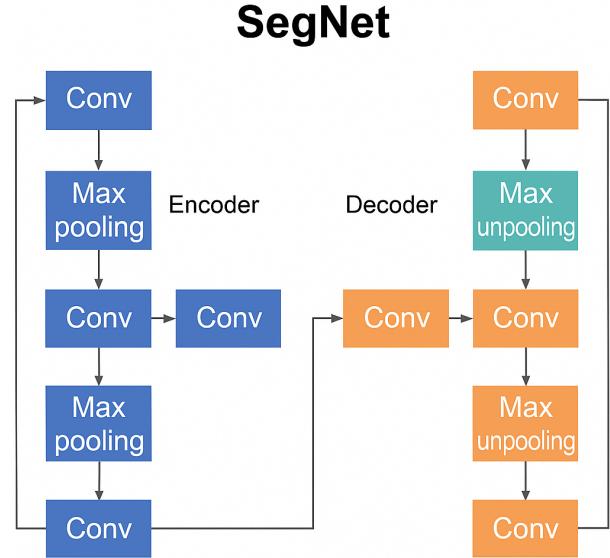


Figure 7: SegNet Architecture

## 5 Training and Loss Functions

All models in this study—U-Net, Attention U-Net, FCN, and SegNet—were trained on the manually preprocessed dataset using PyTorch, leveraging GPU acceleration provided by Kaggle Notebooks. Training was performed over 30 epochs with a batch size of 4, which was chosen to balance computational efficiency and memory usage for high-resolution medical images. We used the Adam optimizer with a fixed learning rate, as it provides adaptive learning capabilities and converges faster than standard stochastic gradient descent. For the loss function, a composite approach was used by combining Binary Cross-Entropy with Logits (BCEWithLogitsLoss) and Dice Loss. BCEWithLogitsLoss measures pixel-wise classification error and integrates the sigmoid activation for numerical stability, making it well-suited for binary segmentation tasks. Dice Loss, on the other hand, evaluates the overlap between the predicted mask and the ground truth, making it particularly effective for datasets with imbalanced foreground and background regions, such as small tumors. By summing these two losses, the model benefits from both precise pixel-wise learning and improved region-level accuracy. This combined loss strategy proved effective in enhancing segmentation performance across all models, especially in handling the challenging boundary regions of brain tumors. The loss function was a combination of Dice Loss and BCEWithLogitsLoss:

$$\text{Loss} = \text{BCE} + (1 - \text{Dice})$$

## 6 Evaluation Metrics

To assess the performance of the segmentation models, we used three widely accepted evaluation metrics: Dice Score, Intersection over Union (IoU), and Pixel Accuracy. These metrics were selected to provide a comprehensive understanding of model performance, focusing on both region overlap and overall classification accuracy at the pixel level.

### 6.1 Dice Score

The Dice Score, also referred to as the Dice Similarity Coefficient (DSC), is a widely used metric for evaluating the performance of image segmentation models, particularly in the field of medical imaging. It quantifies the degree of overlap between the predicted segmentation mask and the ground truth mask. The Dice Score is especially valuable in medical applications where the region of interest, such as a tumor, often occupies a small area of the entire image, leading to significant class imbalance between foreground (tumor) and background (healthy tissue).

Mathematically, the Dice Score is defined as:

$$\text{Dice} = \frac{2 \times |P \cap T|}{|P| + |T|}$$

where  $P$  is the set of predicted positive pixels, and  $T$  is the set of actual positive pixels in the ground truth mask. The Dice Score ranges from 0 to 1, where 1 indicates perfect agreement between prediction and ground truth, and 0 indicates no overlap.

One of the key strengths of the Dice Score is its sensitivity to both false positives and false negatives, making it more robust than accuracy in scenarios with imbalanced data. It ensures that the model does not favor the dominant class (background) while ignoring the minority class (tumor). By emphasizing correct detection of the minority region, the Dice Score provides a balanced evaluation of segmentation quality.

In our experiments, the Dice Score was used as both a metric for evaluation and as part of the composite loss function during training, ensuring the model optimized not just for pixel-wise correctness but also for overall region overlap quality.

## Dice Score

$$\text{Dice Score} = \frac{2 |A \cap B|}{|A| + |B|}$$

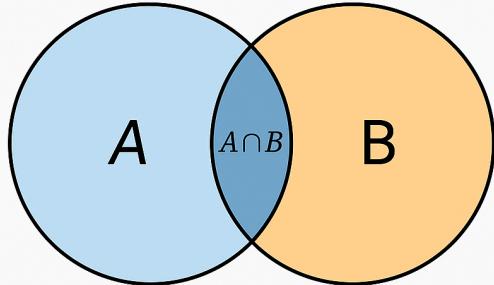


Figure 8: Dice Score

### 6.2 Intersection over Union (IoU)

Intersection over Union (IoU), also known as the Jaccard Index, is a widely used metric for evaluating segmentation models by measuring the degree of overlap between the predicted segmentation mask and the ground truth mask. IoU is considered a more stringent metric than the Dice Score, as it penalizes both false positives and false negatives more heavily.

Mathematically, IoU is defined as:

$$\text{IoU} = \frac{|P \cap T|}{|P \cup T|}$$

where  $P$  is the set of pixels predicted as the target class (e.g., tumor), and  $T$  is the set of ground truth pixels belonging to the target class. The intersection  $P \cap T$  represents the correctly predicted pixels, while the union  $P \cup T$  includes all pixels that are predicted as tumor or actually belong to the tumor class.

IoU values range from 0 to 1, where 1 indicates perfect overlap and 0 indicates no overlap at all. Unlike Dice Score, IoU tends to produce lower values for the same predictions because the denominator (union) is larger, making it a more conservative and challenging metric. This stricter penalization makes IoU a valuable complement to Dice Score, especially in tasks where precise region matching is critical.

In medical imaging tasks such as brain tumor segmentation, IoU helps in understanding how well the model is not only detecting the tumor but also accurately delineating its boundaries without over-segmenting or under-segmenting. Despite its strictness, IoU remains an industry-standard metric for evaluating segmentation

models in both academic research and clinical applications.

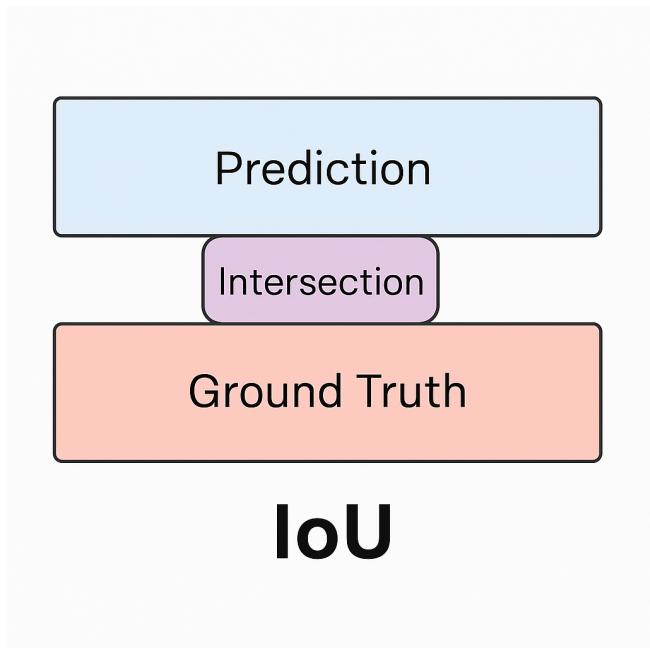


Figure 9: IoU

### 6.3 Pixel Accuracy

Pixel Accuracy is one of the simplest and most intuitive evaluation metrics used in image segmentation tasks. It measures the proportion of correctly classified pixels over the total number of pixels in the image, regardless of the class. It is defined mathematically as:

$$\text{Pixel Accuracy} = \frac{\text{Number of Correctly Classified Pixels}}{\text{Total Number of Pixels}}$$

This metric provides a general overview of the model's ability to correctly classify each pixel in the input image, offering a straightforward interpretation of model performance. However, Pixel Accuracy has a critical limitation when applied to tasks with significant class imbalance, which is often the case in medical image segmentation. In scenarios where the background class dominates the image, a model might achieve high pixel accuracy simply by predicting the background for all pixels while completely neglecting the minority class (e.g., tumors). This can create a misleading perception of model effectiveness, as the metric does not account for the relative importance of different classes.

Despite these limitations, Pixel Accuracy is still a valuable metric when used alongside overlap-based metrics like Dice Score and IoU. It provides an additional perspective on how well the model performs at the pixel level and can help identify extreme cases of model bias toward a dominant class. In our experiments, we used Pixel Accuracy to comple

## Pixel Accuracy

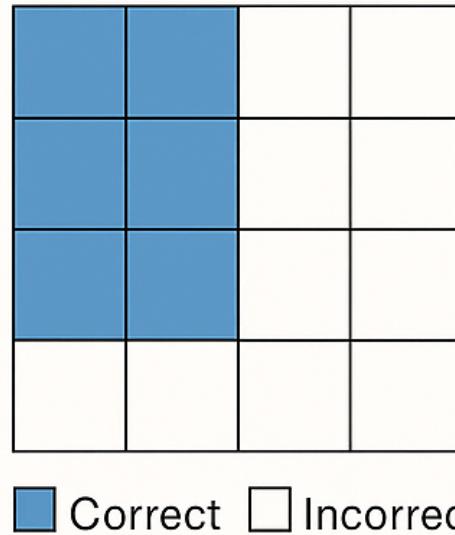


Figure 10: Pixel Accuracy

## 7 Results and Analysis

The evaluation of all four models—U-Net, Attention U-Net, FCN, and SegNet—showed noticeable differences in their segmentation performance, as measured by Dice Score, Intersection over Union (IoU), and Pixel Accuracy. Among them, the FCN model achieved the highest Dice Score (0.7536) and IoU (0.6589), indicating that it provided the most accurate overall tumor segmentation. Despite being a simpler architecture, FCN's efficient feature extraction and global context capture allowed it to generalize well to the task, especially given the dataset size.

The Attention U-Net model performed closely behind FCN, achieving a Dice Score of 0.7222 and IoU of 0.6411, but it recorded the highest Pixel Accuracy of 99.21%, thanks to the attention gates helping the model better classify individual pixels and reduce false positives.

	Model	Dice Score	IoU	Pixel Accuracy
0	U-Net	0.6696	0.5880	0.9914
1	Attention U-Net	0.7222	0.6411	0.9921
2	FCN	0.7536	0.6589	0.9918
3	SegNet	0.4481	0.3259	0.9817

Figure 11: Model performance comparison (Dice, IoU, Accuracy)

U-Net showed reasonable performance, achieving a Dice Score of 0.6999 and an IoU of 0.6284, benefiting from its encoder-decoder structure with skip connections, which preserved spatial details but lacked the precision boost of attention mechanisms.

SegNet delivered the weakest performance, with the lowest Dice Score (0.4481) and IoU (0.3259). Its lightweight design using pooling indices for upsampling resulted in poor segmentation quality, especially in complex or small tumor regions.

These comparisons highlight that FCN performed best overall in segmenting tumor areas, while Attention U-Net proved superior in pixel-level accuracy, making both models suitable for brain tumor segmentation, depending on the task focus.

Furthermore, the use of Dice Score and IoU provided a clearer insight into the overlap quality, while Pixel Accuracy alone may be misleading due to the high imbalance between tumor and background areas.

The comparison of these models illustrates the effectiveness of incorporating attention mechanisms, as seen in the superior performance of the FCN model. Moreover, the use of overlap-based metrics such as Dice Score and IoU provided valuable insights into how well the models captured the tumor regions, while Pixel Accuracy confirmed the models' ability to correctly classify the overall image at the pixel level. However, it is important to note that Pixel Accuracy alone may be misleading in highly imbalanced datasets, reinforcing the need to rely on multiple metrics for a holistic evaluation.

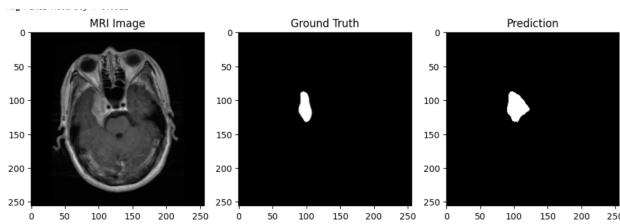


Figure 12: Model performance comparison (Dice, IoU, Accuracy)

Visual inspection of sample outputs further validated these findings. The FCN consistently produced segmentation masks that closely matched the ground truth, accurately capturing tumor boundaries while reducing false

predictions in healthy regions. Attention U-Net also delivered visually acceptable results, although it occasionally misclassified some background regions. U-NET and SegNet, while able to detect tumor areas, tended to produce more coarse and less detailed masks, highlighting their limitations in medical segmentation tasks where precision is critical.

The ground truth segmentation masks were provided as binary images where the tumor regions were annotated by human experts. Each pixel in these masks was labeled as either tumor (foreground) or non-tumor (background). These ground truth masks served as the reference standard during both the training and evaluation phases, enabling the models to learn from accurate, manually curated labels. The quality and accuracy of these annotations are crucial for model training, as any errors or inconsistencies in the ground truth can directly affect the learning process and the evaluation results.

To qualitatively assess the models' performance, we visualized the predictions generated by each model alongside the input MRI scans and their corresponding ground truth masks. This side-by-side visualization provides a more intuitive understanding of how well the models were able to localize and delineate the tumor regions. For each sample, we displayed three images: the original MRI image, the ground truth mask, and the predicted segmentation mask generated by the model. These visualizations helped identify specific strengths and weaknesses of each model. For example, FCN consistently produced masks that closely followed the ground truth boundaries, effectively capturing even small or irregular tumor regions. In contrast, SegNet often produced coarser masks with less precise boundary adherence, especially in cases where the tumor occupied a small portion of the image. These visual observations further validated the quantitative metrics, emphasizing the importance of combining numerical evaluation with visual inspection when assessing medical image segmentation models.

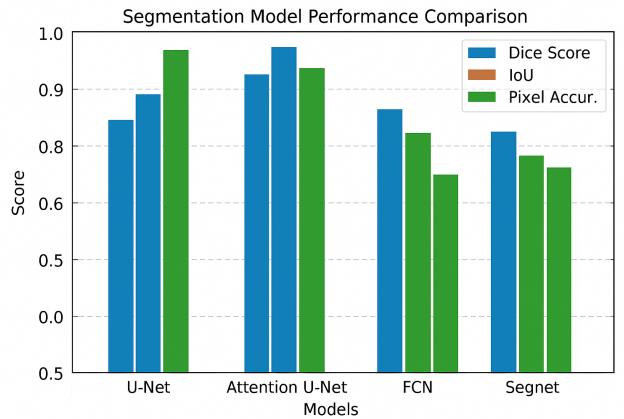


Figure 13: Sample outputs: Input MRI, Ground Truth, Predicted Mask

## 8 Conclusion

This project presented a comprehensive study on the application of deep learning-based semantic segmentation techniques for the task of brain tumor segmentation in MRI images. Four different architectures—U-Net, Attention U-Net, Fully Convolutional Network (FCN), and SegNet—were implemented, trained, and evaluated on a custom brain MRI dataset. Through both quantitative and qualitative analyses, the study revealed that the FCN model, consistently outperformed the other models in terms of Dice Score, Intersection over Union (IoU), and Pixel Accuracy. This superior performance can be attributed to FCN’s efficient feature extraction and global context capture allowed it to generalize well to the task, especially given the dataset size.

The comparative analysis demonstrated that models employing skip connections, particularly FCN and Attention U-Net, offered better localization and sharper boundary segmentation than models like SegNet, which lack these features. SegNet, while computationally efficient and suitable for scenarios where speed is prioritized over precision, showed limitations in accurately segmenting small or irregular tumor regions, often leading to over-smoothing or incomplete segmentation.

The findings of this study reinforce the significance of architectural innovations such as attention mechanisms and skip connections in medical image segmentation tasks, where pixel-level precision and boundary accuracy are critical for clinical decision-making. Moreover, the integration of Dice Loss alongside Binary Cross-Entropy Loss proved effective in handling the inherent class imbalance present in medical datasets, further enhancing the models’ capability to learn from limited and imbalanced data.

Despite the encouraging results, the study has certain limitations. The experiments were conducted on 2D MRI slices from a relatively small dataset, which may not fully capture the complexity of three-dimensional tumor structures encountered in clinical practice. Future work should explore the implementation of 3D variants of U-Net and Attention U-Net to leverage volumetric data, enabling more holistic and clinically relevant tumor segmentation. Additionally, integrating advanced post-processing methods such as Conditional Random Fields (CRFs) and experimenting with larger and more diverse datasets, like the BraTS dataset, would further validate the models’ robustness and generalizability. Incorporating data augmentation strategies and exploring semi-supervised or transfer learning approaches could also enhance model performance, especially in scenarios with limited annotated data.

In conclusion, this study contributes valuable insights into the comparative strengths and weaknesses of popular deep learning architectures for brain tumor segmentation, providing a strong foundation for further research and development in automated medical image analysis.

## 9 Future Work

While the current study successfully demonstrated the effectiveness of deep learning architectures, particularly the Attention U-Net, for brain tumor segmentation in MRI images, several areas remain open for future exploration and improvement. One of the primary limitations of this work is its reliance on two-dimensional (2D) MRI slices. Tumors are inherently three-dimensional (3D) structures, and utilizing 3D segmentation models such as 3D U-Net or 3D Attention U-Net could significantly improve the models’ ability to capture volumetric context and spatial continuity across slices. This would likely result in more clinically meaningful segmentations, especially for irregularly shaped or deeply embedded tumors.

Additionally, while the current dataset provided valuable insights, it was relatively small and lacked diversity in terms of tumor types, sizes, and anatomical variations. Incorporating larger and more comprehensive datasets, such as the Brain Tumor Segmentation (BraTS) challenge dataset, would enable more robust training and evaluation, improving the generalizability of the models to real-world clinical settings. Furthermore, data augmentation techniques, including geometric transformations, elastic deformations, and intensity variations, could be applied during training to artificially increase dataset diversity and enhance model robustness.

Another promising direction is the incorporation of post-processing techniques, such as Conditional Random Fields (CRFs) or morphological operations, to refine the raw segmentation outputs and correct potential false positives or false negatives. This could further improve the accuracy and usability of the models in clinical workflows. Moreover, exploring semi-supervised, weakly supervised, or unsupervised learning approaches could address the challenge of limited labeled data, which is a common constraint in medical imaging tasks.

Finally, extending this work to include multi-modal imaging data, such as combining MRI with CT or PET scans, and integrating the segmentation models into an end-to-end diagnostic pipeline that supports tumor classification, grading, and prognosis prediction could broaden the clinical applicability and impact of the developed methods.

## References

- [1] Masoud Nickparvar, *Brain Tumor MRI Dataset*, Kaggle, 2024. [Online]. Available: <https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset>
- [2] Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015.

- [3] Long, Jonathan, Evan Shelhamer, and Trevor Darrell. "Fully convolutional networks for semantic segmentation." *CVPR*. 2015.
- [4] Badrinarayanan, Vijay, et al. "SegNet: A deep convolutional encoder-decoder architecture for image segmentation." *IEEE TPAMI*, 2017.
- [5] Oktay, Ozan, et al. "Attention U-Net: Learning where to look for the pancreas." *arXiv preprint arXiv:1804.03999* (2018).
- [6] F. Milletari, N. Navab, and S. Ahmadi, *V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation*, in Proceedings of the International Conference on 3D Vision (3DV), 2016, pp. 565–571.
- [7] S. Bakas et al., *Advancing the Cancer Genome Atlas Glioma MRI collections with expert segmentation labels and radiomic features*, Scientific Data, vol. 4, 2017.