

Efficient Brain Tumor Segmentation in MRI Images using U-Net

Nirdesh Patel, Shivam Patel, Sujal Patel, Asst. Prof. Ms. Monali Parikh

¹²³Research Scholar, Institute of Information Technology, Krishna School Of Emerging Technology & Applied Research, KPGU University, Varnama, Vadodara, Gujarat, India

⁴ Assistant Professor, Department of Information Technology and Engineering, Krishna School Of Emerging Technology & Applied Research, KPGU University. Varnama, Vadodara, Gujarat, India

Abstract - The Brain tumor segmentation from MRI images is a crucial task in medical imaging, aiding in diagnosis, treatment planning, and monitoring patient progress. This project presents an efficient solution for automatic brain tumor segmentation using the U-Net deep learning architecture. U-Net, with its encoder-decoder structure, is specifically designed for biomedical image segmentation tasks, enabling precise localization of tumor regions in MRI scans. The model was implemented using Python, leveraging TensorFlow and Keras for building and training the network, while NumPy was used for data manipulation. A user-friendly interface was created using Streamlit, allowing for real-time image uploads and visualization of segmentation results. The project employs MRI images from a publicly available dataset, where the images are preprocessed to standard dimensions and normalized for consistent performance. The model was trained using binary cross-entropy as the loss function and optimized using the Adam optimizer, achieving high accuracy in segmenting tumor regions.

1. INTRODUCTION

Deep learning has become of significant interest and utilization for medical image analysis in recent years by virtue of advancements in computer vision. Despite this growth, deep learning in medical imaging still faces challenges that need to be addressed and diminished. Medical image segmentation consists of anatomical structure labeling within templates, then image registration algorithms are used to map the templates to a particular subject domain through non-linear transformations. After that, a voting process (or maybe a more sophisticated method such as Bayesian weighting) is applied to select the correct label from the labels space and assign it to the segments.

Brain tumors are such abnormal growths found within the human cranium. Given the complex and sensitive nature of the brain, a noninvasive technology, i.e. Magnetic Resonance Imaging (MRI), is the most popular pick for brain tumor diagnosis. These images are 3-D scans of a patient's brain and can be visualized on either of its three respecting image planes (Coronal, Sagittal and Transversal). Each perspective plane displays its information regarding a potential abnormal growth within the cranium. This classification of MRI scans based on the perspective planes has been noted to improve the analytical results while detecting brain tumors.

Brain tumor segmentation aims to autonomously and accurately identify the size and location of a brain tumor from MRI scans. While traditional machine learning techniques require hand-crafted features to perform well, most of the current research is focused on using deep learning networks to segment a region of interest (ROI) from an input image. Although considerable success has been achieved using deep learning, they either require large amounts of annotated data or they depend on aggressive data augmentation techniques. Magnetic Resonance Imaging (MRI) is the most commonly used non-intrusive technique for medical image acquisition. Brain tumor segmentation is the process of algorithmically identifying tumors in brain MRI scans. While many approaches have been proposed in the literature for brain tumor segmentation, this paper proposes a lightweight implementation of U-Net. Apart from providing real-time segmentation of MRI scans, the proposed architecture does not need a large amount of data to train the proposed lightweight U-Net. Moreover, no additional data augmentation step is required. The lightweight U-Net shows very promising results on BITE dataset and it achieves a mean intersection-over-union (IoU) of 89% while outperforming the standard benchmark algorithms.

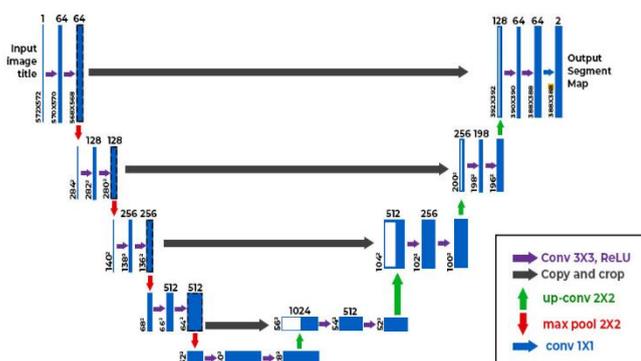
2. WORKING PRINCIPLE

The U-Net architecture operates based on a fully convolutional neural network (FCNN) structure specifically designed for efficient biomedical image segmentation. This network architecture is organized into a contracting path (encoder) and an expansive path (decoder), forming a symmetric U-shaped design. This structure enables the U-Net to capture contextual information effectively while preserving spatial accuracy.

The encoder path systematically down-samples the input MRI images through a series of convolutional and max-pooling layers, capturing increasingly abstracted, high-level features. Each convolutional layer utilizes small 3X3 filters with rectified linear unit (ReLU) activations, allowing for effective feature extraction. The down-sampling process reduces spatial dimensions, enabling the network to generalize over broader contexts within the image.

During training, the network minimizes a loss function that quantifies the discrepancy between the predicted segmentation map and the ground-truth mask, often using metrics like Dice similarity or cross-entropy loss to encourage accurate tumor boundary recognition. The U-Net's design, with its symmetrical structure and integration of skip connections, ensures a robust balance between spatial localization and contextual feature abstraction, making it highly effective for segmenting intricate structures like brain tumors in MRI images.

3. U-Net Architecture



Element of a system

In a system, inputs are the essential resources, data, or energy required to initiate operations. They form the starting point of any system by providing the raw materials or information that will undergo transformation. Inputs vary widely depending on the type of system; for instance, in a manufacturing system, they may be raw materials, while in an information system, they can be data streams or signals.

The processes within a system are the mechanisms or actions that transform inputs into desired outputs. Processes are typically structured and sequential, ensuring that each step contributes to achieving the system's objectives. They are at the core of the system's functionality and vary from computational algorithms in digital systems to physical changes in industrial systems. Processes define how efficiently and accurately the system can achieve its intended outcomes.

Outputs are the results of these processes—products, services, or information that represent the primary purpose of the system. Outputs illustrate the system's effectiveness and serve as a measure of its performance. In a computing system, for example, outputs might include processed data or reports, while in an engineering context, they may be finished products or assembled components.

To optimize performance and ensure stability, many systems incorporate feedback mechanisms. Feedback consists of information about the output or state of the system that is redirected back into the system to influence future operations. Positive feedback can amplify certain processes, while negative feedback helps correct deviations, enabling the system to maintain accuracy and adapt to changing conditions. Feedback loops are particularly valuable in complex systems where dynamic adjustments are essential.

Control mechanisms are the protocols and rules governing the system's operations. They ensure that the system follows set standards and meets specific objectives, functioning as a regulatory framework within the system. Controls may include quality checks, decision rules, or algorithms, depending on the system's requirements. Effective control mechanisms are critical for maintaining consistent and accurate outputs.

The boundaries of a system define what is inside the system versus what lies outside of it. Boundaries set the limits for the system's scope, interactions, and internal components, delineating the system from its external environment. Boundaries are essential for understanding the system's operational domain and its interaction points with external entities.

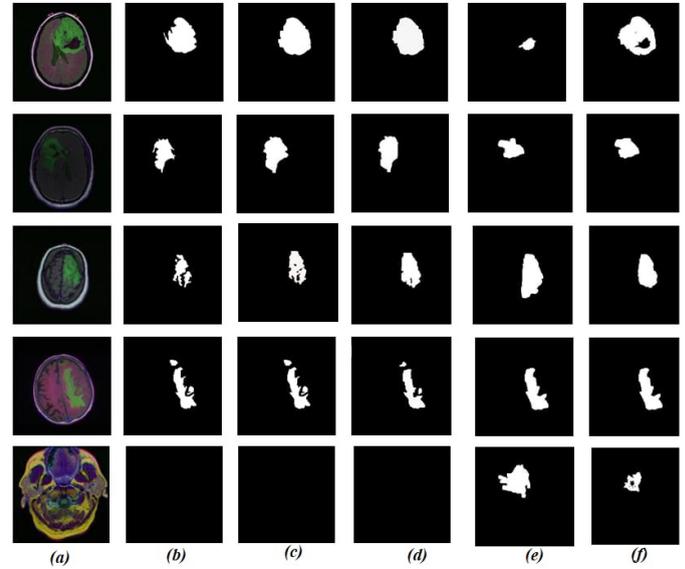
3.1 Image Preprocessing

Image segmentation is usually ready with the image processing step to enhance the image. Image enhancement is the first step in understanding the image. In the experimental result section, we used three datasets namely TCIA, FIGSHARE and Brats 2019. Therefore, in this section, we will discuss them preprocessing step for each dataset as follows. In TCIA dataset, we applied two preliminary processing techniques for image enhancement; Noise removal and contrast enhancement. Preliminary processing techniques attempt to reduce the artifacts introduced by the imaging method. Also, noise reduction is an effective step to improve results before processing (e.g., edge detection on image). Note that, the used images are loaded into gray. For noise removal, we used two filters which are median filter and soft filter. Next, we will illustrate the two filters. The median filter is a non-linear digital filtering technique. It used to remove signals or noise from an image. The soft weighted median filter (SWMF) is a new method to remove noise in image processing. This filter is used for two noise images. The first type is the constant value noise (FVN), a type of noise whose value does not change, such as salt and pepper noise. The second type is the random value noise (RVN), a type of random value noise that reaches a variable value such as Gaussian noise and Speckle noise.

3.2 U-net Segmentation

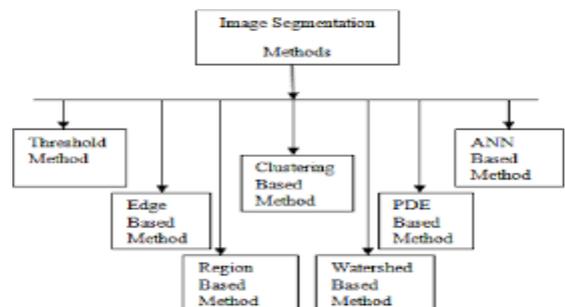
Usually, medical images contain clear details of the human body shown (for example, cerebral tumor), so that the natural shape of the body edge is variable. A structure has been proposed to deal with the segmentation of organisms using elaborate models that combine a high-level representation of deep decoding layers with a representation of the appearance of shallow encoding layers to produce the segmentation of details. The U-net method has shown very high results in natural images. It also applies to biomedical images. Encoding path (sampling down) which contains 5 layers of convolution. As each block has two 3×3 filter. This led to an increase in the number of feature maps from 1 to 1024. For encoding, the max pooling with a stride 2 is performed at the end of all blocks except the last block, so feature maps are reduced from 256×256 to 16×16 .

3.3 MRI Images



Tumor image segmentation. (a) original image, (b) ground truth, (c) Modified U-net segmentation, (d) U-net segmentation, (e) TH segmentation, and (f) K-means segmentation.

Manual segmentation of a brain tumor requires a group of clinical experts to locate the exact location of the tumor and its type. In other words, it needs more skills, experiences, and time. Thus, it is inefficient for segmentation of a brain tumor. Instead of the manual segmentation method, computer-based automated segmentation method offers the optimal solution to save surgeon time and provide reliable and accurate results and at the same time it reduce the efforts by experienced clinicians to finish diagnostic or evaluation procedures for each patient. In past, many algorithms of machine learning were proposed to segment normal and abnormal brain tissue using MRI images. These algorithms have many disadvantages such as the challenge of choosing the features that enable the process to be fully automated and the requirement of computer engineering with medical experience.



4. Software and implementation

4.1 Training

Stochastic gradient descent (SGD) is used for training the network, and to evaluate model in the last layer, the energy function is calculated using the SoftMax pixel-wise over the final feature map (after the 1×1 convolution layer) where the loss function used is the cross-entropy loss function. The SoftMax pixel-wise function is given by:

$$p_k(x) = \frac{\exp(a_k(x))}{\sum_{k'=1}^K \exp(a_{k'}(x))}$$

where $a_k(x)$ is the function of activation corresponding to channel (k) and a pixel position at (x). K is the number of classes (labels within the segmented image).

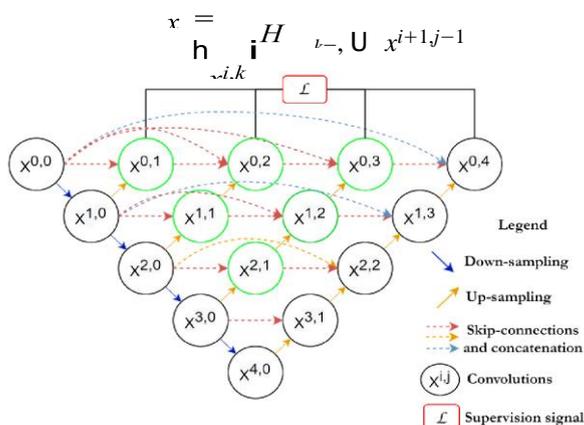
However, the energy function, which is the cross entropy that penalizes at each spatial location, is defined as:

$$E = \sum_{x \in \Omega} \omega(x) \log(p_{l(x)}(x))$$

where (l) is the exact label of each pixel. $w(x)$ is the weight map, which is defined as:

$$\omega(x) = \omega_c(X) + \omega_0 \cdot \exp\left(-\frac{(d_1(X) + d_2(X))^2}{2\sigma^2}\right)$$

where ω_c is the weight map used for class frequencies balancing. d_1 is the distance between the nearest cell and the border, while d_2 is the distance between the border and the second nearest cell.

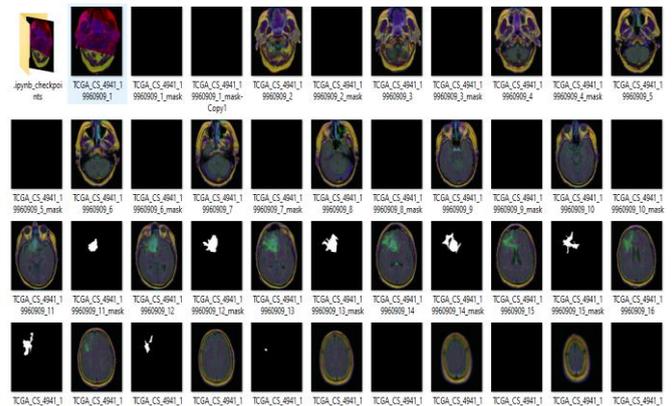


4.2 Dataset Taken from Kaggle for Training

For this study, a publicly available dataset from Kaggle was utilized to train and evaluate the modified U-Net model for brain tumor segmentation in MRI images. This dataset includes MRI scans with labeled segmentation masks, which provide the necessary ground truth for training deep learning models. The images are formatted consistently, enabling preprocessing techniques like noise reduction and normalization, which are critical for achieving high segmentation accuracy. The dataset's high-quality segmentation masks help in training the model to identify tumor regions with precision, supporting both visual inspection and quantitative performance evaluations.

(<https://www.kaggle.com/code/monkira/brain-mri-segmentation-using-unet-keras>)

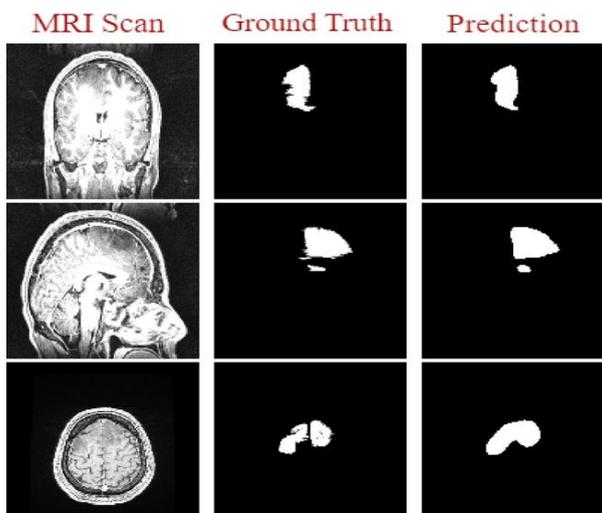
4.3 Labelling of 750 Dataset for MRI Segmentation



To effectively train our model, we labeled a total of 750 images from the dataset specifically for MRI Segmentation. Labeling involves annotating images with relevant information that indicates the Tumor. This step is critical in supervised machine learning, as the model learns from these labeled examples to recognize patterns associated with human figures. We ensured that the labels accurately reflect the content of the images, considering variations in Tumor size. This meticulous labeling process allows the model to generalize better and improves its accuracy when deployed in real-world applications.

4.4 Subjective evaluation

Training the network began in small iterations where we monitored the segmentation performance based on iterations alone. Early results using 10 epochs produced poor results across two of the image planes notably the sagittal and coronal planes. Poor segmentation results were expected based on the small epoch range used during training, however, adequate results were recorded on the transversal plane. This is likely due to the size of each dataset as the transversal dataset had the most images as the patient’s tumor is most prominent from this particular perspective. To further improve these results we increased the networks epoch range to 50 and monitored the results to see if the segmentation performance had improved. Using 50 epochs significantly improved the networks segmentation performance across all three perspective planes. Figure shows the results across all three perspectives using only 50 epochs and the proposed network architecture from Section Here, model 1 (or first model) is the standard U-Net architecture, whereas the model 2 (or the final model) is the proposed U-Net with optimized filter values. To further improve the networks segmentation performance we began undertaking small experiments using the entire dataset of images. This dataset contains all images extracted from all three perspective planes. The purpose of this study is to experiment on U-Nets ability to extract features from images from different perspectives. Running the network using this new dataset for 50 epochs only produced very promising results. Increasing the number of epochs to an extreme size does increase the networks overall segmentation accuracy but only by a very slight amount.



4.5 Accuracy and Output of Model

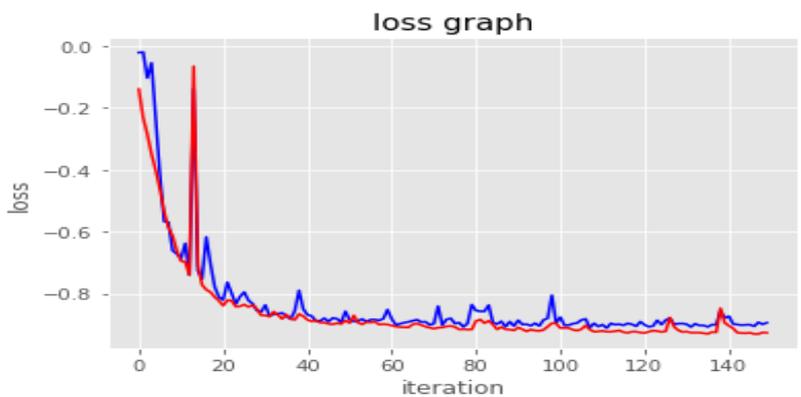
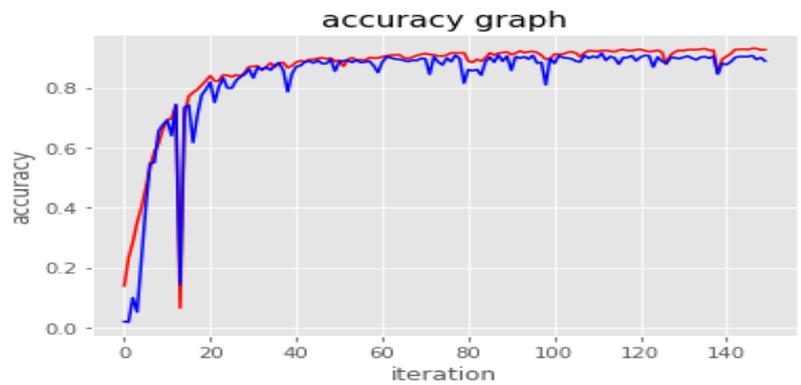
$$\text{Accuracy}(\%) = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \times 100$$

TP (True Positive): Tumor exists and is correctly detected.

TN (True Negative): No tumor exists and is not detected.

FP (False Positive): No tumor exists but is detected.

FN (False Negative): Tumor exists but is not detected



5. Result and outcome

The research demonstrated that the modified U-Net architecture substantially improves brain tumor segmentation accuracy in MRI images, achieving robust results across multiple datasets. This model was evaluated on three primary datasets—TCIA, Figshare, and BraTS 2019—where it outperformed traditional segmentation methods such as k-means and thresholding, as well as the original U-Net model. Key metrics revealed that the modified U-Net achieved up to a 2.2% improvement in the Dice Similarity Coefficient (DSC) over the basic U-Net, with accuracy rates reaching 99.84% on the Figshare dataset and 99.5% on the BraTS 2019 dataset. These results highlight the model's reliability in distinguishing tumor regions from healthy brain tissue, which is crucial for accurate diagnosis.

The study also assessed image quality using Peak Signal-to-Noise Ratio (PSNR) and Root Mean Square Error (RMSE), which confirmed the high fidelity of segmented images compared to ground truth data. The modified U-Net consistently produced images with improved PSNR values and reduced RMSE, indicating that its segmented outputs closely matched the original MRI images in quality. Additionally, preprocessing steps, including noise reduction with median and soft weighted median filters, were essential for enhancing segmentation accuracy by reducing artifacts and allowing the model to detect tumor boundaries more precisely.

The modified U-Net also demonstrated strong adaptability and generalization across different MRI datasets, effectively segmenting various tumor types, such as high-grade and low-grade gliomas. This adaptability underscores the model's potential for real-world clinical applications, where imaging conditions and datasets can vary widely. Overall, the modified U-Net emerges as a powerful tool for automated brain tumor segmentation, offering high accuracy, superior image quality, and robust generalization, making it a promising candidate for integration into clinical workflows for tumor diagnosis and surgical planning.

6. CONCLUSIONS

The results affirm that the modified U-Net is a highly effective tool for automated brain tumor segmentation, yielding high accuracy, improved image quality, and reliable generalization across datasets. These advancements position the modified U-Net as a viable candidate for clinical integration, aiding in accurate tumor diagnosis and planning in neurosurgery. These findings underscore the potential of deep learning approaches, particularly U-Net variations, to address the complex challenges associated with medical image segmentation.

REFERENCES

- 1 Integrating Edges into U-Net Models with Explainable Activation Maps for Brain Tumor Segmentation using MRI Images.
- 2 Optimized U-Net for Brain Tumor Segmentation. [NVIDIA, Santa Clara CA 95051, USA]
- 3 Brain Tumor Segmentation Using Modified U-net. [2022]
- 4 CU-Net: U-Net architecture for efficient brain-tumor segmentation on BraTS 2019 dataset [2019]
- 5 Using U-Net network for efficient brain tumor segmentation in MRI images.
- 6 Deep Learning-Based Brain Image Segmentation for Automated Tumour Detection.
- 7 J. Hu, X. Gu, X. Gu, Mutual ensemble learning for brain tumor segmentation, Neurocomputing 504 (2022) 68–81.

Websites

- <https://www.geeksforgeeks.org/u-net-architecture-explained/>
- https://pytorch.org/hub/mateuszbuda_brain-segmentation-pytorch_unet/#:~:text=This%20U%2DNet%20model%20comprises,64%2C%20128%2C%20and%20256.
- <https://www.sciencedirect.com/science/article/pii/S2772442522000429>
- <https://www.kaggle.com/datasets/mateuszbuda/lgg-mri-segmentation>
- <https://www.kaggle.com/code/monkira/brain-mri-segmentation-using-unet-keras>