Brain Tumour Segmentation Using U-Net

Safwan Hasan
Department of Electrical and Computer Engineering
Toronto Metropolitan University

***Abstract*—This paper focuses on the application of a U-Net convolutional neural network (CNN) model for segmenting brain tumors in medical images. Brain tumor segmentation is a critical task in medical imaging, as accurate identification of tumor regions within brain scans facilitates improved diagnosis, treatment planning, and patient monitoring. Using a Brain Tumor Image Semantic Segmentation Dataset from Kaggle that includes 2146 annotated images [12], this project aims to develop a robust and reliable model capable of distinguishing between tumor and non-tumor regions.**

The U-Net model, known for its encoder-decoder structure with skip connections, has been widely adopted for biomedical image segmentation due to its ability to preserve spatial information while performing pixel-wise classification. This paper leverages this architecture to tackle the complexities of brain tumor segmentation. This study will not only highlight the effectiveness of the U-Net model in medical image segmentation but also explores various enhancements and modifications that can further improve its performance. In this paper, a standard U-Net model is applied to the dataset, providing successful segmentations as a result.

***Index Terms*—Convolutional Neural Network(CNN),U-Net,tumor**

I. Introduction

Brain tumors pose significant challenges in the field of medical diagnosis and treatment due to their complex structure and varying appearance in MRI scans. These tumors can be either malignant or benign [2], each requiring distinct approaches for treatment. Accurate segmentation of brain tumors is crucial for several reasons: it aids in precise diagnosis, guides surgical planning, facilitates targeted therapy, and allows for effective monitoring of treatment response and disease progression. However, achieving accurate segmentation of brain tumors is often a very difficult process due to the complexity of brain tumors [1].

Image segmentation, the process of partitioning an image into meaningful regions, is a fundamental task in medical imaging. In the context of brain tumor segmentation, it involves distinguishing the tumor region from the surrounding healthy brain tissue, resulting in a binary map where each pixel is classified as either part of the tumor or the background [3].Traditional methods of manual segmentation by radiologists are time-consuming and impractical for large-scale clinical applications.

Deep learning, particularly convolutional neural networks (CNNs), has revolutionized the field of image segmentation. Models such as U-Net, introduced by Olaf Ronneberger et al. [3] have become the standard of biomedical image segmentation tasks. A U-Net model can be applied to a labeled segmented dataset in order to learn features and output segmentations [1]. The U-Net architecture consists of a contracting path (encoder) to capture context and a symmetric expanding path (decoder) to enable precise localization. The inclusion of skip connections between corresponding layers of the encoder and decoder helps retain important spatial information, making U-Net particularly effective for tasks requiring fine-grained segmentation [1].

This project leveraged a standard U-Net model, known for its effectiveness in biomedical image segmentation, to segment brain tumors in MRI images [1,2]. This project leveraged the U-Net model for the task of brain tumor segmentation using a publicly available dataset from Kaggle [12]. The dataset comprises 2146 MRI images annotated in COCO format, encompassing two classes: tumor and non-tumor. The preprocessing steps include auto-orientation of pixel data and resizing images to 640 x 640 pixels to ensure uniformity.

By developing and fine-tuning a U-Net model for brain tumor segmentation, it is the aim to provide a tool that can assist radiologists and oncologists in making more accurate and efficient diagnoses. The research contributes to the ongoing advancements in medical image analysis, showcasing the potential of deep learning models in transforming clinical practice.

II. Literature Review

1) This study reviews eight U-Net model variants, including traditional 2D, 3D, U-Net++, Inception, Attention, Residual, Recurrent, and Dense, focusing on their application to brain tumor segmentation [1]. The Attention U-Net was found to provide the highest accuracy, leveraging a mechanism that highlights informative regions of feature maps for improved segmentation. Each variant offers unique architectural benefits, such as improved gradient propagation, feature extraction, and temporal information integration.

2) The SAResU-Net model [2] integrates shuffle attention blocks and residual modules within a 3D U-Net framework for enhanced brain tumor segmentation. Testing on the BraTS 2019 and 2020 datasets, the model achieved high DSC values, indicating significant improvements in segmentation accuracy. The self-ensemble module and attention mechanisms capture local spatial and channel information, contributing to its superior performance.

3) This study employs modified U-Net and M-Net models for segmenting the brachial plexus nerve region in ultrasound images [3], using a Kaggle dataset [4], utilizing pre-processing techniques to enhance image quality. Despite achieving lower

DSC scores with U-Net due to dataset inconsistencies, the modified M-Net demonstrated more promising results. The models were optimized through average pooling, dropout layers, and specific convolutional operations to improve segmentation performance.

4) A proposed 3D U-Net model, enhanced with channel attention mechanisms, showed significant improvements over traditional 2D and 3D U-Net models in brain tumor segmentation [5]. Utilizing the BraTS 2018 dataset, the model incorporated global pooling and activation functions to refine segmentation accuracy. Results indicate a notable advancement in effectively segmenting whole tumors, enhanced tumors, and tumor cores.

5) This study presents a U-Net-based model for glioma segmentation, optimized with 1x1 convolution layers and evaluated on the BraTS 2015 dataset [6]. The approach addresses data imbalance and utilizes intensity normalization to standardize MRI images. The model outperformed others in various metrics, demonstrating competitive Dice Similarity Coefficients despite slight over-segmentation tendencies.

6) Pre-trained SegResNet and nnU-Net models were evaluated on pediatric MRI data for segmenting diffuse midline glioma, leveraging adult glioblastoma pre-training to enhance performance [7]. The nnU-Net with pre-training showed the best results, achieving high Dice scores and improved training speed. Pre-processing steps and cross-validation were crucial in achieving superior segmentation compared to previous pediatric efforts.

7) Reswave-Net, a U-Net variant incorporating residual connections and wavelet decomposition, was developed to enhance segmentation accuracy in brain tumor MRIs [8]. Using the BraTS 2020 dataset, the model achieved notable Dice scores and Hausdorff distances across tumor types. Residual connections and wavelet decomposition addressed gradient issues and added spectral details, significantly improving performance over standard methods.

8) Mathews and Mohamed review deep learning's role in automating brain tumor segmentation, emphasizing U-Net's effectiveness in learning complex patterns [9]. U-Net adaptations like spatiotemporal-separable 3D U-Net and ensemble approaches enhance segmentation by processing various MRI modalities. The review highlights U-Net's growing application in medical imaging, achieving high accuracy compared to traditional methods.

9) This paper introduces the Attention-Sharp-U-net, enhancing traditional U-net architecture with a grid-based attention block and sharp block for improved segmentation [10]. Tested on the BraTS2020 dataset, it outperformed baseline models in Dice and Jaccard scores. The model effectively segments brain tumor tissues with high accuracy and efficiency, addressing limitations of existing methods.

10) Kong and Zhang propose a multi-modal approach using a cascade network of 3D U-Nets, refining segmentation accuracy for tumor substructures [11]. Deep residual modules and a Tversky loss function with category weights improve performance, particularly for small tumors. Evaluated on the BraTS2018 dataset, the method shows robust results, enhancing clinical decision-making in brain tumor identification.

III. Dataset

The goal is to segment brain tumor images using a U-Net machine learning model, using a publicly available Kaggle dataset [12]. This dataset has 2146 images and are annotated in COCO format. In addition, the following preprocessing was applied to each image: auto-orientation of pixel data and resize to 640 x 640. This dataset has two classes: class 0 (non-tumor) and class (tumor).

IV. Problem Statement

The objective is to develop a deep learning model for accurately segmenting brain tumor regions in MRI images using a publicly available Kaggle dataset [12]. This model aims to improve the precision of diagnosis and treatment planning by providing accurate and detailed tumor segmentation, thereby contributing to advancements in medical image analysis and patient outcomes.

A. Additional Data Processing

The true masks for the input images are not directly provided within the dataset. Instead, they are derived from the annotations available in the COCO (Common Objects in Context) format files. The COCO format is widely used in computer vision for object detection, segmentation, and captioning tasks, and it includes detailed annotations for each image in the dataset.

1. Extraction of True Masks

The process of obtaining the true masks involves parsing the COCO annotation files, which include information about the segmentation polygons or masks for each object within the images. These annotations are typically stored in JSON format and contain detailed information such as the image IDs, category IDs, and segmentation data for each object. This information is extracted to construct binary masks where each pixel represents whether it belongs to a particular object or not.

2. Resizing and Normalization

Once the true masks are extracted, the next step is to preprocess both the images and the masks to ensure they are in a consistent format suitable for training the model.

- **Resizing**: The images and their corresponding masks are resized to a uniform resolution of 128 x 128 pixels. This resizing is crucial because it standardizes the input dimensions for the model and ensures that the network can process the images efficiently. Resizing also helps in reducing computational complexity and memory usage, which is important for training deep learning models.
- **Normalization**: To further preprocess the images, each pixel value is normalized by dividing by 255. This normalization step scales the pixel values to a range between 0 and 1. Normalization is essential because it helps in stabilizing and speeding up the training process. By scaling the input values, the model can learn more

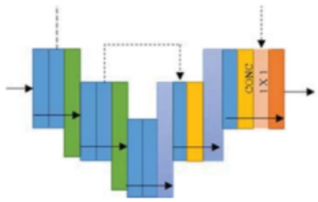

Fig. 1. Standard U-Net [1]

effectively and converge faster. It also prevents potential issues related to numerical stability and helps the model achieve better performance.

3. Final Dataset Preparation

After resizing and normalizing, the images and masks are organized into training, validation, and test sets. This structured organization ensures that the model is evaluated on separate data during training and testing phases, providing a clear assessment of its performance and generalization capabilities.

B. Model

A standard U-Net model is used for training, testing, and validating the dataset, as shown in Figure 1. The U-Net architecture, known for its efficacy in biomedical image segmentation, consists of a symmetric encoder-decoder structure with skip connections that enable precise localization and context assimilation. A standard U-Net model consists of:

1) **Encoder**: The encoder compresses the input image into a lower-dimensional representation through successive convolutional and max-pooling layers.
2) **Bottleneck**: The bottleneck serves as the transition between the encoder and decoder, capturing high-level features of the image.
3) **Decoder**: The decoder up samples the encoded representation to the original image size, utilizing transposed convolutions and concatenating features from the encoder through skip connections to enhance spatial resolution.

In this paper, the model has 4 convolutional layers and each layer is followed by a 2x2 max pooling operation that reduces the spatial dimensions of the feature maps. . The decoder section mirrors the encoder with 4 up-sampling layers. Each decoder block consists of a 2x2 up-sampling (using transposed convolutions) followed by concatenation with the corresponding skip connection from the encoder. This is followed by convolutional layers to refine the feature map. After the encoder blocks, there is a bottleneck (a convolution block) with the largest number of filters (1024 in this case) which processes the deepest feature maps. The final output is obtained using a 1x1 convolution with a sigmoid activation function. This produces the segmentation mask where each pixel's value is between 0 and 1, indicating the probability of that pixel belonging to the target class.

C. Loss Function

The model uses a Dice Loss (DL) function, which is very common in binary segmentation tasks [3]. In the implementation, a small smoothing term is added to avoid division by zero, especially when dealing with empty masks. This smooth term (`1e-15` in this case) ensures numerical stability by preventing division by zero.

$$\frac{2 * |S_{gt} \cap S_{pred}|}{|S_{gt}| + |S_{pred}|} = \frac{2TP}{2TP + FP + FN} \quad (1)$$

Fig. 2. Dice Loss Function Equation [3]

V. RESULTS

In this section, the results of the model training, including the chosen hyper parameters and performance metrics are presented.

The table below summarizes the model parameters used during training. After experimenting with various settings, a learning rate of 1e-4 and 50 epochs were found to produce the best results. The Adam optimizer was selected for its effective adaptive learning rate capabilities.

Hyper Parameter	Model
Epoch	50
Optimizer	Adam
Learning Rate	1e-4

TABLE I
MODEL PARAMETERS

The following table presents the performance metrics of the model on the test dataset. These metrics show that the model performs well according to the evaluation criteria.

- **Test Loss**: The model achieved a test loss of 0.249, which reflects the overall prediction error.
- **Test Dice Coefficient**: The Dice coefficient of 0.744 indicates a good level of overlap between the predicted and actual masks, suggesting effective segmentation.
- **F1 Score**: An F1 score of 0.758 shows a balanced performance between precision and recall.
- **Precision**: With a precision of 0.832, the model has a high rate of correct positive predictions.
- **Recall**: The recall of 0.696 indicates that the model identifies a good proportion of relevant positive instances, though there is room for improvement.
- **Jaccard Score**: The Jaccard score of 0.610 further supports the model's effectiveness, showing a satisfactory agreement between the predicted and true masks.

Test Loss	0.249
Test Dice Coefficient	0.744
F1 Score	0.758
Precision	0.832
Recall	0.696
Jaccard Score	0.610

TABLE II
PERFORMANCE METRICS

VI. Conclusion

In this study, we successfully developed and evaluated a U-Net model for binary segmentation tasks. The model was trained using a learning rate of 1e-4 and for 50 epochs, with the Adam optimizer chosen for its adaptive learning capabilities. The results demonstrate that the model performs well, with key metrics meeting the standards set for this project.

The performance metrics show that the model achieved a test loss of 0.249, indicating a reasonable level of prediction accuracy. The Dice coefficient of 0.744 highlights a strong overlap between the predicted and true segmentation masks, while the F1 score of 0.758 confirms a good balance between precision and recall. Additionally, the precision of 0.832 and recall of 0.696 reflect the model's ability to make accurate positive predictions and identify relevant instances, respectively. The Jaccard score of 0.610 further supports the effectiveness of the segmentation.

These results suggest that the U-Net model is capable of providing effective segmentations for the task at hand. However, there is still room for improvement. Future work could involve experimenting with different hyperparameters, enhancing data augmentation techniques, or exploring alternative model architectures to further boost performance. Additionally, fine-tuning the model or incorporating more advanced methods could help address any remaining challenges and improve the overall segmentation quality.

Overall, the model demonstrates promising performance and provides a solid foundation for further research and development in this area.

References

[1] S. Saluja, M. C. Trivedi, and A. K. Dubey, "U-Net variants for Brain tumor segmentation: Performance and limitations," *2023 International Conference on Computational Intelligence, Communication Technology and Networking (CICTN)*, Apr. 2023. doi:10.1109/cictn57981.2023.10141477

[2] Y. Zhang, Y. Han, D. Liu, and J. Zhang, "SARESU-net: Shuffle attention residual U-Net for brain tumor segmentation," *2022 15th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, Nov. 2022. doi:10.1109/cisp-bmei56279.2022.9979978

[3] N. Abraham, K. Illanko, N. Khan, and D. Androutsos, "Deep learning for semantic segmentation of brachial plexus nervesin ultrasound images using U-Net and M-Net," *2019 3rd International Conference on Imaging, Signal Processing and Communication (ICISPC)*, Jul. 2019. doi:10.1109/icispc.2019.8935668

[4] "Ultrasound nerve segmentation," Kaggle, https://www.kaggle.com/c/ultrasound-nerve-segmentation (accessed Jun. 23, 2024).

[5] S. Chen, G. Hu, and J. Sun, "Medical image segmentation based on 3D U-net," *2020 19th International Symposium on Distributed Computing and Applications for Business Engineering and Science (DCABES)*, Oct. 2020. doi:10.1109/dcabes50732.2020.00042

[6] T. Yang and J. Song, "An automatic brain tumor image segmentation method based on the U-net," *2018 IEEE 4th International Conference on Computer and Communications (ICCC)*, Dec. 2018. doi:10.1109/compcomm.2018.8780595

[7] X. Liu *et al.*, "Automatic segmentation of rare pediatric brain tumors using knowledge transfer from adult data," *2023 IEEE 20th International Symposium on Biomedical Imaging (ISBI)*, Apr. 2023. doi:10.1109/isbi53787.2023.10230757

[8] S. E., A. K. Vinod, J. R. S., A. P. R., and S. K. G., "Reswavenet: A wavelet based residual U-Net for brain tumour segmentation and overall survival prediction," *2023 International Conference on Control, Communication and Computing (ICCC)*, May 2023. doi:10.1109/iccc57789.2023.10165135

[9] C. Mathews and A. Mohamed, "Review of automatic segmentation of MRI based brain tumour using U-Net Architecture," *2020 Fourth International Conference on Inventive Systems and Control (ICISC)*, Jan. 2020. doi:10.1109/icisc47916.2020.9171057

[10] M. R. Goni and N. I. Ruhaiyem, "Salient feature extraction using attention for Brain Tumor segmentation," *2022 IEEE International Conference on Computing (ICOCO)*, Nov. 2022. doi:10.1109/icoco56118.2022.10031677

[11] L. Kong and Y. Zhang, "Multi-modal Brain tumor segmentation using cascaded 3D U-net," *2021 6th International Conference on Image, Vision and Computing (ICIVC)*, Jul. 2021. doi:10.1109/icivc52351.2021.9526932

[12] P. K. Darabi, "Brain tumor image dataset: Semantic segmentation," Kaggle, https://www.kaggle.com/datasets/pkdarabi/brain-tumor-image-dataset-semantic-segmentation?resource=download (accessed Jun. 23, 2024).