MRI Brain Tumor Segmentation

Saied Salem, Omar Mansour, Mohamed Ibrahim, Nouran Mohamed

***Abstract*—Automatic MRI brain tumor segmentation is of vital importance for the disease diagnosis, monitoring, and treatment planning. We present a method for classification and segmentation of brain tumors based on deep learning analysis of brain contrast T1ce (t1ce) MR images. To achieve this goal, three different deep learning networks are investigated, i.e., VAE and Unet with attention, Z-net, and DeepLabv3+ models. Experiments are performed on the Brain Tumor Segmentation (BraTS 2020) Challenge, composed of 369 brain MRI volumes, each with 155 cross-sections. Using the available computational resources, the Unet with VAE network achieves the highest Dice Similarity Coefficients (DSC) of the three investigated CNNs.**

I. Introduction

Brain tumor is an epidemic cause of cancer death. In the USA, 700.000 people are diagnosed with brain tumors (80% benign and 20% malignant). In 2020, the American Cancer Society (ACS) for brain tumors estimated about 23,890 malignant tumors of the brain and around 18,020 deaths from malignant brain tumors. Multi-modal MRI scans are commonly used to grade brain tumors based on size and imaging appearance. As a result, imaging plays an important role in the diagnosis and treatment administered to patients. Deep learning-based approaches in general, and convolutional neural networks (CNNs) in particular have been utilized to achieve superior performance in the fields of object detection and image segmentation. In this paper, we propose to utilize 3 different networks for segmentation. The network's inputs are the MRI slices, and the output is the segmented mask of the tumor.

II. Related Work

As the segmentation problem of a brain tumor with 3D data is hard to solve with regular machine learning techniques, deep learning networks are the most common solution for such a problem. In Choudhury et al. [2], various networks are investigated as U-net, VGG16-Segnet, and Deep Lab V3+, where they sliced the 3D brain images along 3 orientations: sagittal, coronal, and axial, and used different combinations of them along with 18 models, which improve accuracy but lead to a weakness of heavy and complex computations and training time. They didn't change the resolution of the image, and neither did we. In Nassar et al. [2], they used a 3D narrow band filter on the input images, which improved accuracy, but increased training time and needed more memory. They used ResNet as a backbone, just as we did. In Myronenko, et al. [8], they added a variational auto-encoder branch to reconstruct the input image itself in order to regularize the shared decoder and impose additional constraints on its layers due to the limited size of the training dataset. In Oktay, et al. [7], they added an attention gate to the U-Net architecture, which will improve its prediction performance across different datasets.

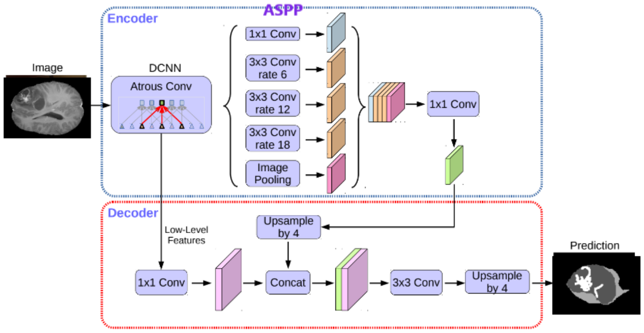

Figure 1. Deeplabv3 network archticture

III. Dataset and Features

Our dataset is the BraTs2020 challenge MRI brain images. BraTs 2020 utilizes multi-institutional pre-operative MRI scans and primarily focuses on the segmentation of intrinsically heterogeneous (in appearance, shape, and histology) brain tumors, namely gliomas. Scans describe T1, T2, T1ce and Flair volumes. Annotations comprise the GD-enhancing tumor (ET), the peritumoral edema (ED), and the necrotic and non-enhancing tumor core (NCR/NET).

All BraTS multimodal scans are available as NIFTI files (.nii.gz), and imaging datasets have been segmented manually, by one to four raters, following the same annotation protocol, and their annotations were approved by experienced neuroradiologists. The provided data are distributed after their preprocessing, i.e., co-registered to the same anatomical template, interpolated to the same resolution (1 mm^3) and skull-stripped. It consists of 369 volumes, and each volume is composed of 155 slices with dimensions of (240,240) and a resolution of 1 mm.

IV. Methods

A. Deeplabv3

In the preprocessing needed for the deeplab method, we selected only the slices that contained the tumor from every volume (83 out of 155) while neglecting the other slices, and then the slices were normalized and every slice was converted to RGB. The segmentation true masks were one-hot encoded. An example is shown in Fig. 2.

For the network architecture, the DCNN backbone used is ResNet 50. The encoder module processes multiscale contextual information by applying atrous convolution at multiple scales, while the decoder module refines the segmentation results along object boundaries [5]. Atrous Convolution allows us to enlarge the field of view of filters to incorporate more context. When r = 1, it is the standard convolution we usually

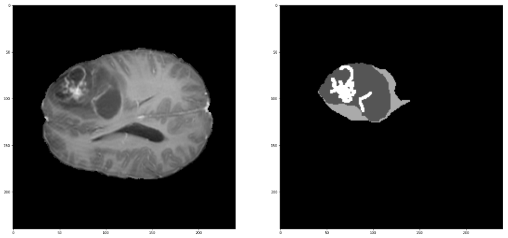

Figure 2. Example shows a slice along with its true mask

use. When r > 1, it is the atrous convolution, which is the stride to sample the input sample during convolution, and it is represented mathematically in equation 1. In addition, in the input feature map, a parallel atrous convolution with different rates is applied and fused together.

$$y[i] = \sum_{k} x[i + r.k]w[k] \tag{1}$$

For the loss function we used the dice loss as in equation 3

$$Dice = \frac{2\,|A \cap \mathrm{B}|}{|A| + |B|} \tag{2}$$

$$Dice\ Loss = 1 - Dice \tag{3}$$

B. 3D Unet with attention and VAE

In the preprocessing needed for 3D Unet with VAE and attention network, the height and width were cropped (from 240 to 128), and the slices were discarded (from 155 to 128) since no useful information was included in those parts. The volume was then normalized. The segmentation true masks were one hot encoded in order to define four classes (GD-enhancing tumour (ET), peritumoral edoema (ED), and necrotic and non-enhancing tumour core (NCR/NET). We use Group Normalization (GN) , which outperforms BatchNorm when the batch size is small (in our case, one batch).

The network architecture is shown in Fig. 3 where each green block is a ResNet-like block with the GroupNorm normalization. The output of the segmentation decoder has three channels (with the same spatial size as the input), followed by a sigmoid for segmentation maps of the four tumor subregions. The VAE branch reconstructs the input image into itself and is used only during training to regularize the shared encoder.

For the loss function, we used the dice loss as in equation 4, where L2 is an L2 loss and $L_{KL}$is a standard VAE penalty term.

$$L = L_{dice} + 0.1 * L_{L2} + 0.1 * L_{KL} \tag{4}$$

We tried another variation of the current network by adding the attention block taken from the architecture shown in Fig. 3. But due to lack of computational resources we had to change the architecture to be less complex by removing some of the conv layers to be able to add that attetion blocks.

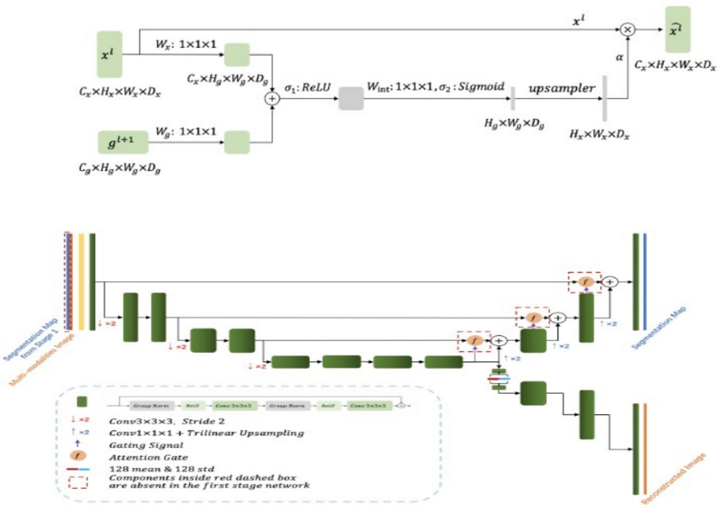
x^l
W_x: 1×1×1
$C_x \times H_g \times W_g \times D_g$
$C_x \times H_x \times W_x \times D_x$
σ_1: ReLU
W_{int}: 1×1×1, σ_2 : Sigmoid
upsampler
$H_g \times W_g \times D_g$
$H_x \times W_x \times D_x$
α
$\widehat{x^l}$
$C_x \times H_x \times W_x \times D_x$
g^{l+1}
W_g: 1×1×1
$C_g \times H_g \times W_g \times D_g$
Segmentation Map from Stage 1
Multi-modalities Image
↓ ×2
↑ ×2
Segmentation Map
Reconstructed Image
Group Norm
ReLU
Conv 3×3×3
↓ ×2 Conv3×3×3, Stride 2
↑ ×2 Conv1×1×1 + Trilinear Upsampling
Gating Signal
f Attention Gate
128 mean & 128 std
Components inside red dashed box are absent in the first stage network

Figure 3. Attention block architecture

C. Znet

As an experiment, we implemented the znet architecture as shown in Fig. 4, which is based on the ideas of skip-connection, encoder-decoder architectures, and data amplification. The architecture consists of an encoding part (analytic downsampling) and a decoder part (synthetic upsampling).

1) The encoder part consists of five blocks, where each encoder block contains double convolutions combined by batch normalization and the rectified activation function ReLU and followed by max-pooling. The output of the encoder block is concatenated with the encoder block input. Note, the encoder block input is interpolated to match the feature map of the encoder block output.
2) The decoder part consists of five blocks, the same as the encoder block except for the use of ConvTranspose2d instead of max-pooling.

V. Experiments/Results/Discussion:

At first, we tried to implement the deep lab network from scratch using PyTorch , and the implementation was inspired by [3], but it didn't go very well; so, we tried a pre-trained, ready-implemented version from the segmentation_models_pytorch library [6]. The dice coefficient metric was used as it is one of the best metrics to use in segmentation problems. The Adam optimizer with a learning rate of 0.001 was used. Due to the limited memory offered by Kaggle, a batch size of 10 slices is the maximum it can hold. In Fig, 5 below, we show the true and predicted masks for five slices and the dice score.

In the Znet, we used a similar approach to the deeplab by building it from scratch with PyTorch, but it still needs more investigation due to computational resource limitations.

For the Unet architecture, the two variations we discussed (with and without attention) gave the best results, with a dice coefficient of 0.7102 without attention and 0.698 when

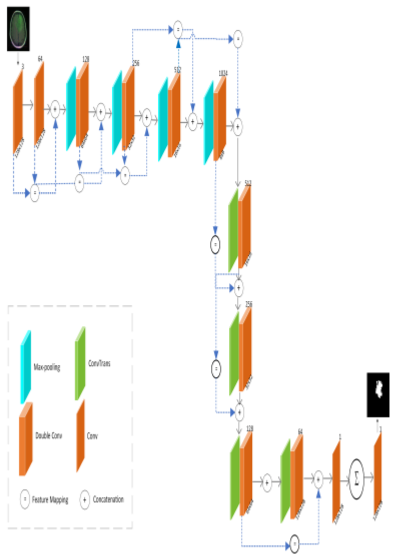

Figure 4. The znet architecture

attention was added. Examples of the output are shown below for both models. As a final step, we did inference on test data, and we got the following results, shown in Figs. 6 and 7.

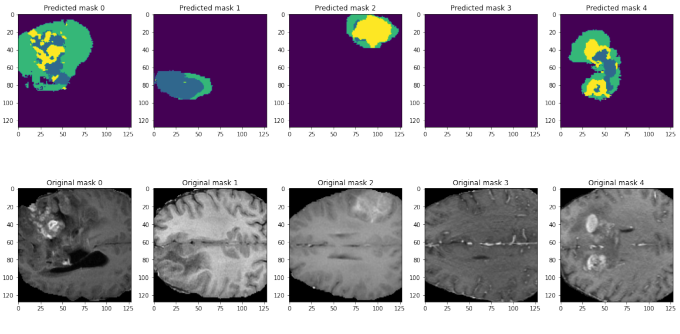

Figure 6. Results of the Unet with VAE without attention

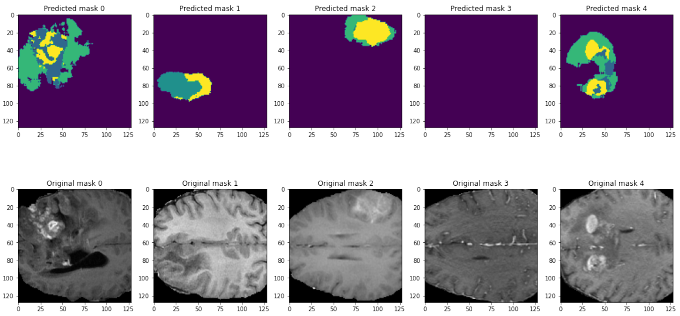

Figure 7. Results of the Unet with VAE with attention

VI. Conclusion

In this paper, the proposed system for brain tumor segmentation is based on deep learning experiments. The performance of BraTs 2020 is evaluated using the T1CE modality of MR images. According to the current results, the Unet with VAE network has the highest Dice Similarity Coefficients (DSC) of the three investigated CNNs. In the future, with more computational resources and a deeper understanding of PyTorch, Deeplabv3+ and Z-net models will be reevaluated as they have shown promising results.

Contribution

- Nouran Mohamed, Omar Mansour: Deeplab and Znet
- Saied Salem, Mohamed Ibrahim : Attention Unet with VAE

Source code

- https://github.com/Omarmansour557/MRI-Tumor-Segmentation
- https://drive.google.com/drive/folders/1i7CV-fZ_ULMLbnDNxxMXfahQHXW9QhFt

References

[1] Roy Choudhury, A. et al. (2019) "Segmentation of brain tumors using deeplabv3+," Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries, pp. 154–167. Available at: https://doi.org/10.1007/978-3-030-11726-9_14.

[2] Nassar, S., Mohamed, M. and Elnakib, A. (2021) "MRI brain tumor segmentation using deep learning. (dept. E)," MEJ. Mansoura Engineering Journal, 45(4), pp. 45–54. Available at: https://doi.org/10.21608/bfemu.2021.139470.

[3] Team, K. Keras documentation: Multiclass semantic segmentation using deeplabv3+, Keras. Available at: https://keras.io/examples/vision/deeplabv3_plus/#building-the-deeplabv3-model (Accessed: January 8, 2023).

[4] Papers with code - rethinking atrous convolution for Semantic Image segmentation Rethinking Atrous Convolution for Semantic Image Segmentation | Papers With Code. Available at: https://paperswithcode.com/paper/rethinking-atrous-convolution-for-semantic (Accessed: January 8, 2023).

[5] Papers with code - encoder-decoder with atrous separable convolution for Semantic Image segmentation Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation | Papers With Code. Available at: https://paperswithcode.com/paper/encoder-decoder-with-atrous-separable (Accessed: January 8, 2023).

[6] Welcome to segmentation models's documentation! Welcome to Segmentation Models's documentation! - Segmentation Models documentation. Available at: https://smp.readthedocs.io/en/latest/index.html (Accessed: January 8, 2023).

[7] Oktay, O. et al. (2018) Attention U-net: Learning where to look for the pancreas, arXiv.org. Available at: https://arxiv.org/abs/1804.03999 (Accessed: January 8, 2023).

[8] Myronenko, A. (2019) "3D MRI brain tumor segmentation using Autoencoder regularization," Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries, pp. 311–320. Available at: https://doi.org/10.1007/978-3-030-11726-9_28.

[9] Papers with code - encoder-decoder with atrous separable convolution for Semantic Image segmentation (no date) Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation | Papers With Code. Available at: https://paperswithcode.com/paper/encoder-decoder-with-atrous-separable (Accessed: January 8, 2023).

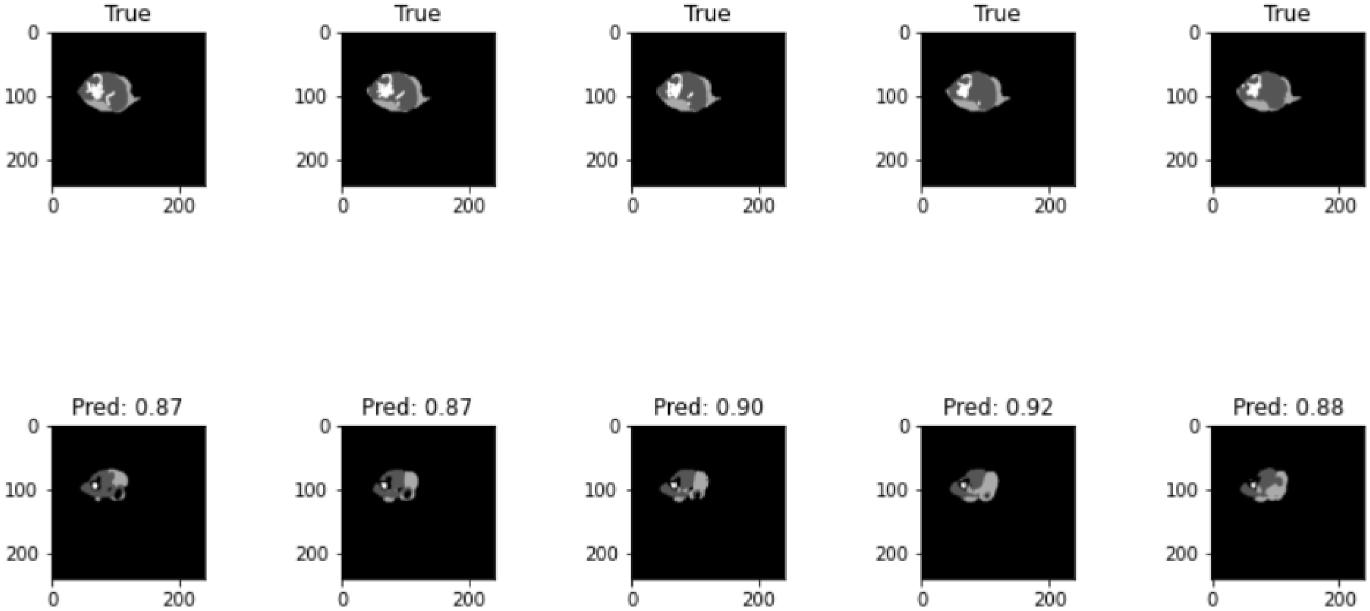

Figure 5. Results of deep lab model