# Brain Tumor Segmentation

Thesis Subtitle

**EPHREM Enzo**

A thesis presented for the degree of

Bachelor of Computer Science



Department of Computer Science

University of Claude Bernard Lyon 1

France

December 2022

# Abstract

In this paper, presented a brain tumor segmentation method using deep learning. A tumor can appear anywhere in the brain, it can take different shape, size and contrast. That is why the use of the different sequences in MRI techniques will be exploited to benefit the segmentation. By extracting the most important features from the MRIs inputs a corresponding segmentation will be generated. To do so, we can distinguish 2 path in the neural network the first one, called contracting path which will be dedicated to extract important and useful information. And the second one the expansive path that will, based on the information extracted, predict the segmentation of the tumor.

# Contents

# 1 Introduction

Segmentation of brain tumors is an important task in the field of medical image analysis. It is used to accurately identify and locate tumors within brain images, which can help doctors diagnose and treat patients. Segmentation, involves painting the exact location and boundaries of a tumor within an image. Involving specialized algorithms and software tools that are designed to analyze medical images and extract relevant information. By accurately segmenting brain tumors, doctors can make not only more informed decisions about a patient's treatment and care but also monitor the progress of the tumor over time. In the context of brain tumor segmentation, deep learning algorithms can be trained to recognize tumorous tissues.

# 2 Methodology

By focusing only on the Flair, T1ce and T2 sequences, and after going through the pre-processing, we are ready to give the brain tumor to the U-Net. After splits of seconds we get our prediction that should undergo post-process to make our result readable. Using multiple display functions provided, we can easly look at the accurate 3D mask generated.

# 3 Data

To train a CNN for brain tumor segmentation, a large dataset of medical images with corresponding ground truth labels is needed, labels are also called masks and they represent the segmentation of a tumor in a brain. Provided in the kaggle BraTS2020 Dataset (Training + Validation) [1] in the validation folder, for each of the 369 patients, 4 sequences of MRI (Flair, T1, T1ce and T2) and one corresponding mask.

## 3.1 Schema

```
dataset
├── BraTS20_Training_001/
│   ├── BraTS20_Training_001_flair.nii
│   ├── BraTS20_Training_001_seg.nii
│   ├── BraTS20_Training_001_t1.nii
│   ├── BraTS20_Training_001_t1ce.nii
│   └── BraTS20_Training_001_t2.nii
├── BraTS20_Training_{ID}/
│   ├── BraTS20_Training_{ID}_flair.nii
│   ├── BraTS20_Training_{ID}_seg.nii
│   ├── BraTS20_Training_{ID}_t1.nii
│   ├── BraTS20_Training_{ID}_t1ce.nii
│   └── BraTS20_Training_{ID}_t2.nii
└── ....
```

## 3.2 Sequences

There are various types of brain MRI sequences Fig. 1 that are used to obtain different types of information about the brain.

**T1-weighted**

This sequence produces images that highlight the different types of tissue in the brain, such as gray matter, white matter, and cerebrospinal fluid. It is commonly used to diagnose brain tumors and other abnormalities.

**T1ce (T1-weighted contrast-enhanced)**

A recovery of the T1-weighted sequence but with a contrast agent to make certain structures in the brain more visible on the resulting images.

**T2-weighted**

This sequence produces images that highlight areas of the brain where there is fluid accumulation, such as in the case of brain swelling or inflammation. It is commonly used to diagnose conditions such as hydrocephalus and multiple sclerosis.

**FLAIR (fluid-attenuated inversion recovery)**

This sequence is similar to T2-weighted imaging, but it produces more detailed images of the brain's fluid-filled spaces. It is commonly used to diagnose conditions such as brain tumors and cerebral edema.

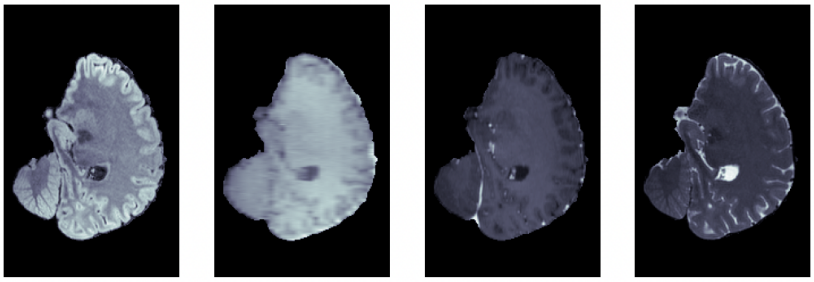


Figure 1: Different types of sequences in brain MRI : (a) Flair; (b) T1; (c) T1ce; (d) T2

## 3.3 Masks

A mask is an image that is used to identify the location and boundaries of a tumor in a medical image of the brain. In our dataset the mask images are labeled to identify different tissue types within the brain tumor. Each voxel (the contraction of volume and pixel), has a value of 0, 1, 2 or 4. Label 0 represents a healthy tissue.

**Label 1: Necrosis**

When a brain tumor is present, the abnormal growth of cells can lead to necrosis in the surrounding tissue. This can happen if the tumor grows too large and begins to press on nearby brain tissue, restricting the blood flow and oxygen supply to the affected area. The lack of oxygen and nutrients can cause the cells in the surrounding tissue to die, leading to necrosis. In some cases, the tumor itself can undergo necrosis, which can cause the tumor to shrink in size. However, necrosis can also cause the release of harmful substances into the surrounding tissue, leading to further damage and potentially causing complications.

**Label 2: Enhancing tumor**

An enhancing tumor in a brain tumor is a type of tumor that appears to be growing or becoming more active on imaging tests, such as MRI scans. When a brain tumor is present, the abnormal growth of cells can cause changes in the surrounding tissue, including the formation of new blood vessels. These new blood vessels can appear as a ”halo” or ”ring” of bright spots on an MRI scan, indicating that the tumor is growing or becoming more active. This is known as enhancement, and it can be a sign that the tumor is growing or that treatment is not working effectively. It is important to monitor for signs of enhancement in a brain tumor, as it can help doctors determine the best course of treatment.

**Label 4: Edema**

Edema is the medical term for swelling caused by an accumulation of fluid in the body's tissues. In the case of a brain tumor, edema can occur if the tumor grows and begins to press on nearby brain tissue. This can restrict the flow of cerebrospinal fluid, the clear liquid that surrounds and protects the brain and spinal cord. The buildup of fluid in the brain can cause the tissue to swell, leading to edema. Edema in the brain can be a serious condition and can cause a range of symptoms, including headache, nausea, and difficulty with balance and coordination. It is important to monitor for signs of edema in a brain tumor and to seek medical treatment if it occurs.


Figure 2: Mask Labels : (a) Whole Tumor; (b) Necrosis; (c) Enhancing tumor; (d) Edema

# 4 Data processing

Pre-processing and preparing brain MRI images and masks before training the model will benefit in numerous ways. Scans in the dataset (brain and mask) have a shape of 240x240x155. and have a .nii extension (.nii for a file type NIfTI-1).

**NiBabel**

NiBabel is a library that support common medical and neuroimaging file formats including .nii files. It gives full or selective access to header (meta) information and provides a numpy interface NumPy of the scans.

## 4.1 Brain MRI pre-processing

First of all it is important to try and reduce the neural network input data size given that it is considerably high (240x240x155 = 8 928 00 floats, amounting to approximately 34 MB per channel if using float32) especially with low computational resources at hand. The T1 channel can be discarded as the T1ce channel contains the same information in an enhanced way. Furthermore the borders are mostly 0-valued and can be cropped reducing the shape to 128x128x128 per channel. The second step is scaling each MRI image using the min-max scaling. Min-max scaling Eq. 1 is a method used to normalize data. It works by transforming the data so that all the values lie between 0 and 1. This is often useful when working with neural networks, as it can help to improve the convergence of the algorithm and can also make it easier to compare different features of the data. It also reduces bias triggered by relatively large range of values, and can help with computing time.

$$x_{scaled} = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{1}$$

## 4.2 Mask MRI pre-processing

It is important to remember that there are few values that masks can take for given coordinates Section 3.3. Since the label 3 is empty, replacing the label 4 by the label 3 before categorizing the masks would help eliminating an unnecessary class. Post-processing (processing the prediction) will handle switching all label 3 to 4 afterwards. Before handing out our labels to our model, to match the shape of the brain images, we also crop it from 240x240x155 to 128x128x128.

# 5 Model architecture

In deep learning, a model architecture refers to the specific design of a neural network, including the number of layers, the number of neurons in each layer, and the connections between them. The architecture of a model determines its ability to learn from data and make predictions, so choosing an appropriate architecture is an important part of building a successful deep learning model. A U-Net [6] is a type of convolutional neural network (CNN) that is often used for image segmentation tasks. It is called a U-Net because it has a ”U” shaped architecture that is made up of two parts: an encoder that processes the input image and extracts features from it, and a decoder that expands the feature map and

produces the segmented output. The U-Net is typically used for tasks such as medical image segmentation, where it can be trained to identify and segment specific structures in medical images. This is why it will be our chosen architecture for the segmentation of the tumor. The following diagram was made with PlotNeuralNet

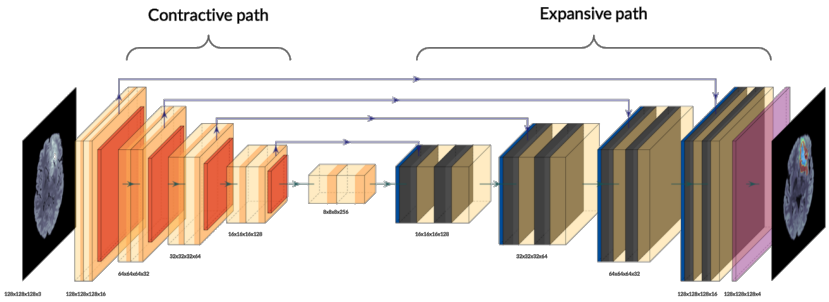


Figure 3: U-Net model architecture for segmentation

The encoding part, is made of 4 convolutional blocks. Knowing that a convolutional block is made of 2 Conv3D with a ReLu activation and a MaxPooling3D. And the decoding part 4 blocks of, 1 Conv3DTranspose and a combination of 2 Conv3D with a ReLu activation as well. Connecting the encoding and decoding parts, are 2 Conv3D called the bottleneck. Note that dropout layers are present throughout the model, and that the model ends with a Conv3D layers with a soft max activation. Finally arrows in Fig 3 represents concatenations.

**Conv3D**

As the name suggests, a Conv3D layer performs 3D convolutions over the input data. The purpose of a Conv3D layer is to learn spatial relationships in the input data by applying filters (also referred as kernels). The filters are typically small 3D volumes, and they slide across the input data, performing dot multiplications and summing the results to produce a new 3D volume. By applying multiple filters to the input data, a Conv3D layer can learn to detect a variety of spatial patterns in the data. In the suggested neural network, all Conv3D have kernels equal to (3, 3, 3) except the last one (1, 1, 1). The number of filters applied depends on if we are in the encoding or decoding part and how deep.

**MaxPooling3D**

MaxPooling3D is a technique used to reduce the size of the 3D input tensor. The input tensor is divided into a set of non-overlapping 3D patches, and for each patch, the maximum value is taken and used as the output of the maxpooling layer for that patch. The goal is to down-sample the feature maps, which can help to extract more robust features from the input data and improve the performance of the model. The output tensor has a smaller size than the input tensor, as it has fewer patches. In our neural network, having a pool_size of (2, 2, 2) the input tensor will be halved.

**Conv3DTranspose**

Conv3DTranspose is used to upsample the input tensor, which means it increases the spatial resolution of the feature maps produced by the layer. It is typically used to generate high-resolution output from low-resolution input. Like a Conv3D, the Conv3DTranpose takes a number of filters and a 3D kernel. The chosen kernel is (2, 2, 2) and the number of filters matches the number of filters in the Conv3D of the same block.

# 6 Segmentation loss function and metrics

A loss function is a function that is used to measure the performance of a model on a given dataset. The goal of training a deep learning model is to find the set of model parameters that minimizes the loss function, so that the model can make accurate predictions on new data. Metrics, on the other hand, are used to evaluate the performance of a model on a given dataset. Unlike the loss function, which is used during training to update the model's parameters, metrics are used to evaluate the model's performance after training is complete. There are many different types of metrics that can be used in deep learning. In general, the loss function and metrics are used together to evaluate the performance of a deep learning model and to guide the training process.

## 6.1 Loss function

The loss function used for the model is a combination of two well known loss functions. The Dice Loss Eq. 3, and the Categorical Focal Loss Eq. 4. Both provided by the segmentation model 3D library [7].

$$LossFunc = DiceLoss + (1 * FocalLoss) \tag{2}$$

**Dice Loss**

The dice loss is commonly used while training any model for segmentation. It follows a simple equation:

$$\begin{gathered} DiceLoss(gt, pr) = 1 - DiceCoef(gt, pr) \\ \text{where } gt = \text{ground truth,} \\ pr = \text{prediction} \end{gathered} \tag{3}$$

The next section will get deeper into what the Dice coefficient is, Section 6.2.

**Categorical Focal Loss**

Categorical focal loss is a loss function that can be used in the training of deep learning models for classification tasks. It is designed to address the issue of class imbalance, which occurs when the number of examples for some classes is much larger than the number of examples for other classes. This can lead to a model that performs poorly on the underrepresented classes, as it is more likely to predict the more frequently occurring classes. Having a multi-class segmentation and an imbalance in the types of tissue in the tumor (Necrosis, Enhancing tumor, Edema), justifies the contribution of the Categorical Focal Loss in our model's loss function.

$$\begin{gathered} FocalLoss(gt, pr) = -gt \cdot \alpha \cdot (1 - pr)^{\gamma} \cdot \log(pr) \\ \text{where } gt = \text{ground truth,} \\ pr = \text{prediction,} \\ \alpha = 0.25, \gamma = 2.0 \end{gathered} \tag{4}$$

## 6.2 Dice coefficient

Used as the main metric, the Dice Coefficient has proven to be a very good indicator of segmentation correctness. The Dice coefficient is a measure of the similarity between two sets. It is defined as the ratio of the size of the intersection of the sets to the size of the union of the sets. The Dice coefficient is often used to evaluate the performance of image segmentation, where the sets being compared are the sets of pixels that are predicted to belong to a particular class and the set of pixels that are actually in that class. The Dice coefficient can range from 0, indicating no overlap between the sets, to 1, indicating complete overlap. It is calculated as follows:

$$\begin{gathered} DiceCoeff(gt, pr) = 2 * \frac{|gt \cap pr|}{|gt + pr|} \\ \text{where } gt = \text{ground truth,} \\ pr = \text{prediction,} \\ |gt \cap pr| = \text{set of elements to both gt and pr} \end{gathered} \tag{5}$$

# 7 Results and Analysis

## 7.1 Training

The model was trained on 295 of the 369 patients, in batches of 2. The other 74 were used as validation. Figure 4 (a) and Figure 5 (b) respectively shows the variation of the loss function and the FScore (Dice Coefficient) during the 50 epochs.

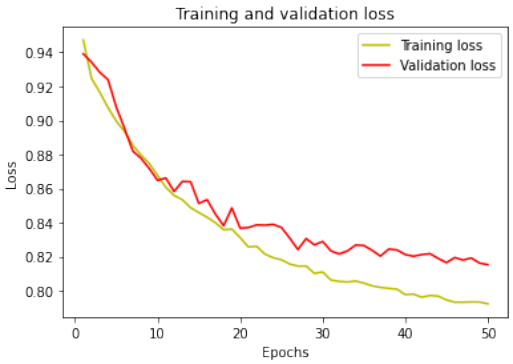


Figure 4: Model's loss

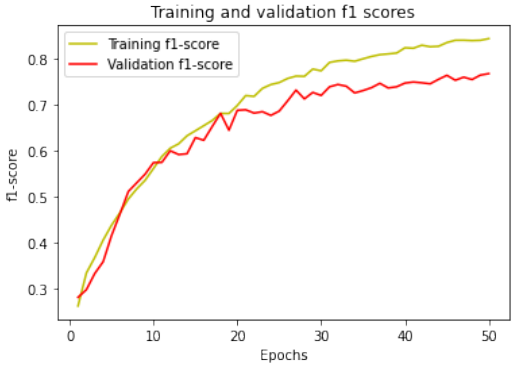


Figure 5: Model's F1 Score

## 7.2 BraTS Challenge validation results

The Brain Tumor Segmentation (BraTS) challenge is an annual competition organized by the Multimodal Brain Tumor Segmentation (BRATS) consortium. It is designed to evaluate and compare the performance of different algorithms for segmenting brain tumors in magnetic resonance imaging (MRI) scans. In the BraTS challenge, participants are given a dataset of MRI scans with brain tumors and are asked to develop algorithms to automatically segment the tumors in the scans. The algorithms are then evaluated on their ability to accurately identify and outline the tumor regions in the MRI scans. In the scope of this report, a validation zip file of the predicted masks to the validation dataset has been submitted as a validation evaluation to the model. The table that follows shows results of the submission.

| | Dice_ET | Dice_TC | Dice_WT |
|---|---|---|---|
| mean | 0.605 | 0.686 | 0.805 |
| median | 0.726 | 0.815 | 0.860 |
| 25quantile | 0.450 | 0.561 | 0.773 |
| 75quantile | 0.829 | 0.890 | 0.908 |

Table 1: Dice scores of ET-enhancing tumor, TC-tumor core (necrosis), WT-whole tumor

Based on Table 1 the model is capable of accurately predicting the shape of the whole tumor, but have a hard time accurately distinguishing between the core tumor and the enhancing tumor.

# 8 Conclusion

With a whole tumor segmentation dice score that competes with the state-of-the-art methods, this basic U-Net with more class focused loss functions can achieve higher greater and better results.

# References


[1] Kaggle, "BraTS2020 Dataset (Training + Validation)", https://www.kaggle.com/datasets/awsaf49/brats20-dataset-training-validation

[2] B. H. Menze, A. Jakab, S. Bauer, J. Kalpathy-Cramer, K. Farahani, J. Kirby, et al. "The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)", IEEE Transactions on Medical Imaging 34(10), 1993-2024 (2015) DOI: https://ieeexplore.ieee.org/document/6975210

[3] S. Bakas, H. Akbari, A. Sotiras, M. Bilello, M. Rozycki, J.S. Kirby, et al., "Advancing The Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features", Nature Scientific Data, 4:170117 (2017) DOI: https://www.nature.com/articles/sdata2017117

[4] S. Bakas, M. Reyes, A. Jakab, S. Bauer, M. Rempfler, A. Crimi, et al., "Identifying the Best Machine Learning Algorithms for Brain Tumor Segmentation, Progression Assessment, and Overall Survival Prediction in the BRATS Challenge", arXiv preprint https://arxiv.org/abs/1811.02629

[5] U.Baid, et al., The RSNA-ASNR-MICCAI BraTS 2021 Benchmark on Brain Tumor Segmentation and Radiogenomic Classification, arXiv:2107.02314, 2021. arXiv preprint, https://arxiv.org/abs/2107.02314

[6] Ronneberger, Olaf and Fischer, Philipp and Brox, Thomas, "U-Net: Convolutional Networks for Biomedical Image Segmentation" (2015), arXiv preprint https://arxiv.org/abs/1505.04597

[7] Solovyev, Roman and Kalinin, Alexandr A and Gabruseva, Tatiana, "3D convolutional neural networks for stalled brain capillary detection", Computers in Biology and Medicine 2022, doi: 10.1016/j.compbiomed.2021.105089 https://github.com/ZFTurbo/segmentation_models_3D