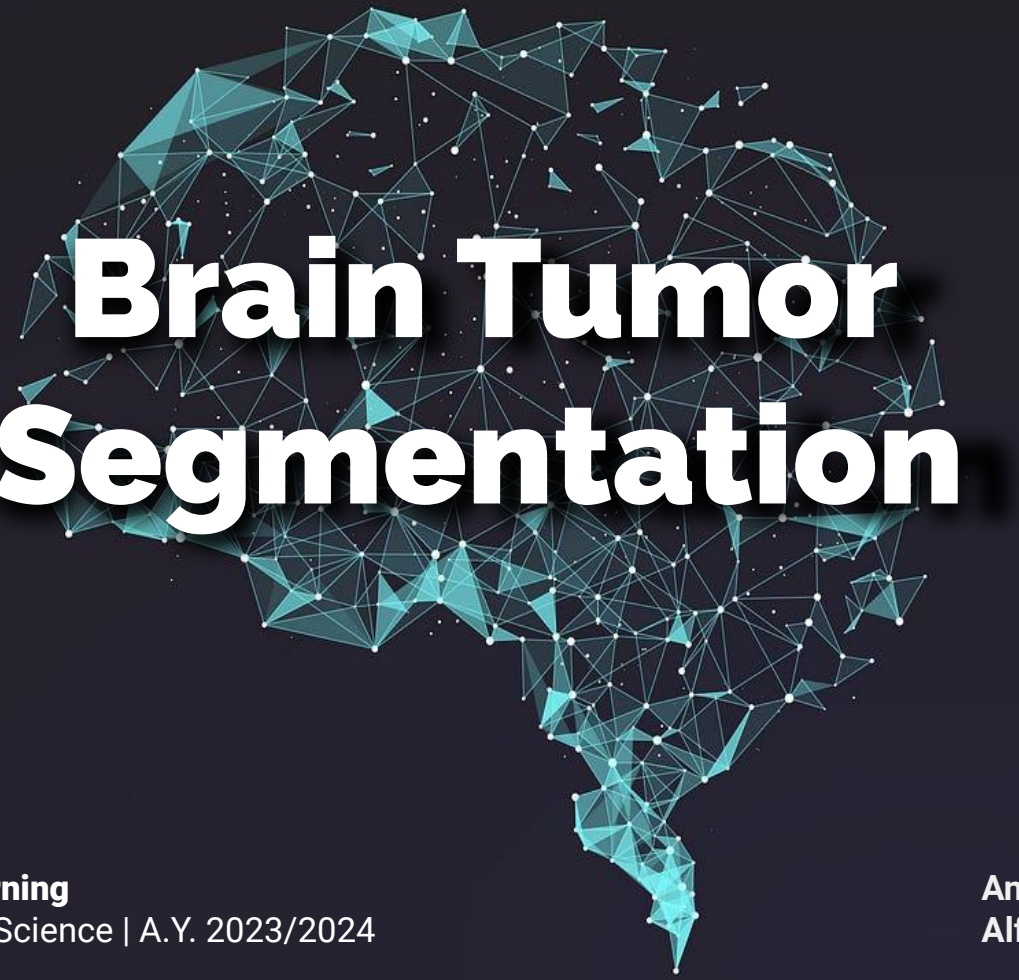


# Brain Tumor Segmentation

**Foundation of Deep Learning**
Master's Degree in Data Science | A.Y. 2023/2024


**Andrea Corsico** 874835
**Alfredo Sestito** 918555

# Index

# Objectives

- Our project aims to perform ***multiclass semantic segmentation*** of MRI scans of inherently heterogeneous (in appearance, shape, and histology) brain tumors, namely **gliomas**.

- Semantic segmentation is the task of assigning a ***class label*** to every single pixel of an input image. Semantic segmentation models create a ***segmentation map of an input image***. A segmentation map is, essentially, a reconstruction of the original image in which each pixel has been **color coded** by its semantic class to create segmentation masks.

# Dataset

The ***BraTS*** dataset consists of **484 images for the training** set and **266 for the test set**, however, we used only those from training because they came with **labels**.

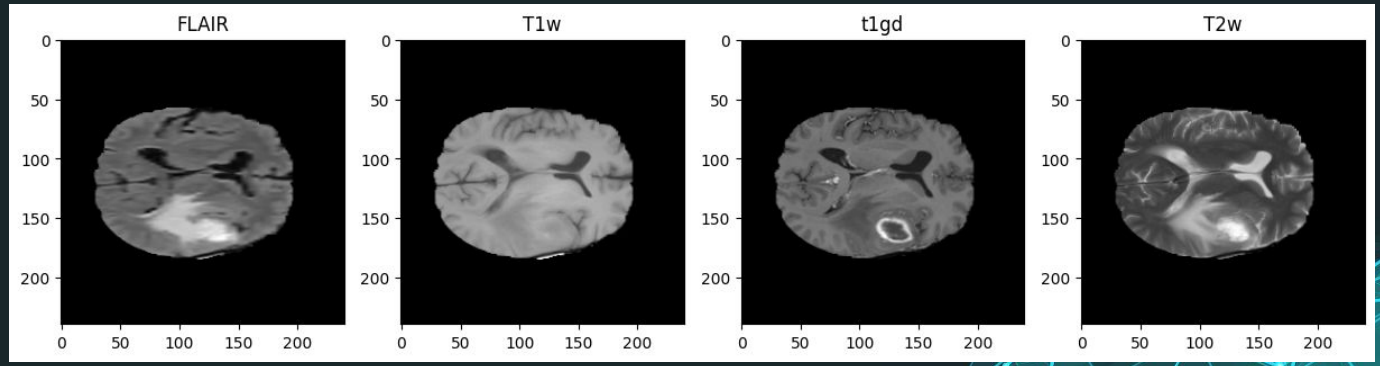
Modalities are different types of procedures or scanning techniques in the context of **Medical Imaging**. Images and labels have a 240x240x155 shape.

**Images have 4 channels representing modalities:**

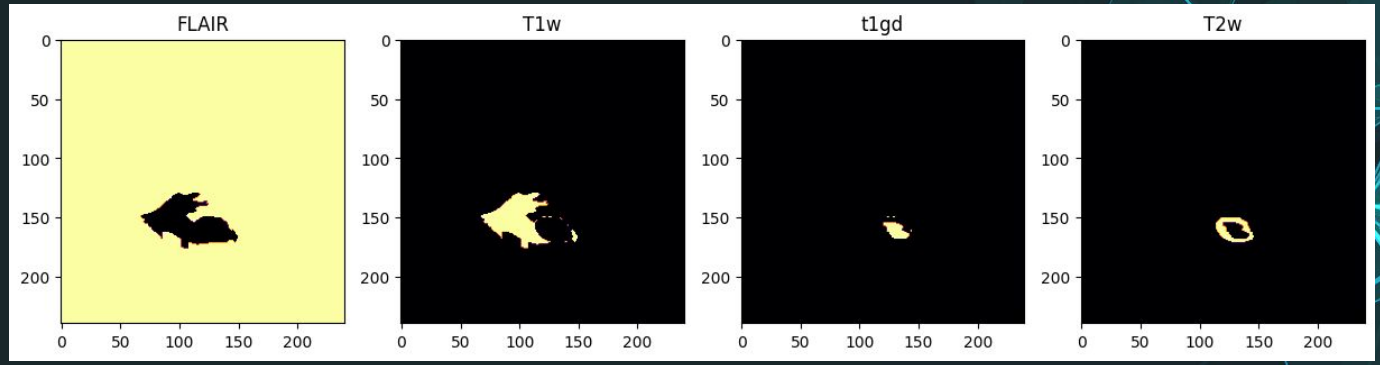
- *FLAIR*
- *T1w*
- *t1gd*
- *T2w*

**Labels have 4 classes:**

- *background (≈ 0.95)*
- *edema (≈ 0.03)*
- *non-enhancing tumor (≈ 0.01)*
- *enhancing tumour (≈ 0.01)*




The 4 channels representing the images modalities.




The 4 channels of the labels.

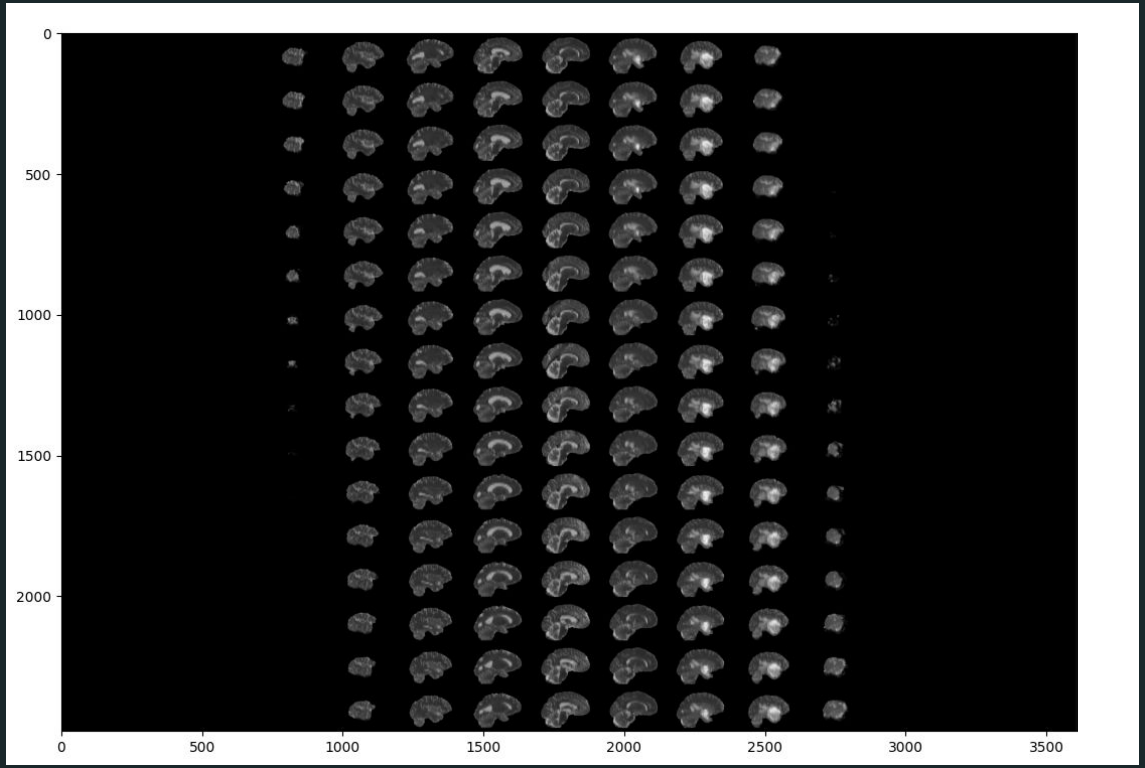
# Preprocessing

Firstly we split our data to **train** (70%), **validation** (15%) and **test set** (15%), resulting respectively in 338, 73 and 73 observations.

**Train and validation** were used to train the model from **scratch** while test set was kept aside for **final evaluation**.

All slices of a 240x240x155 volume are shown in the figure.

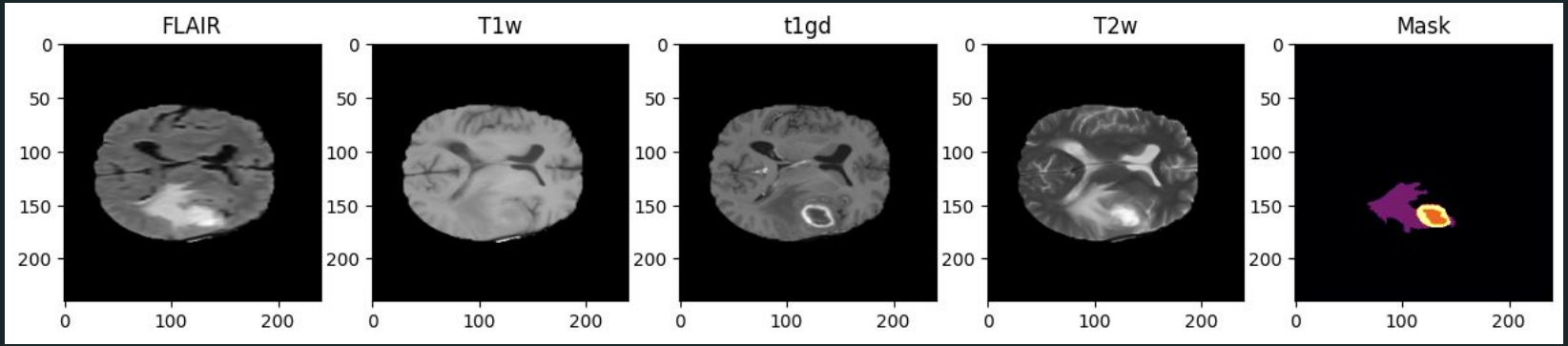
Many of the slices along the third axis are composed entirely of class 0 labels.



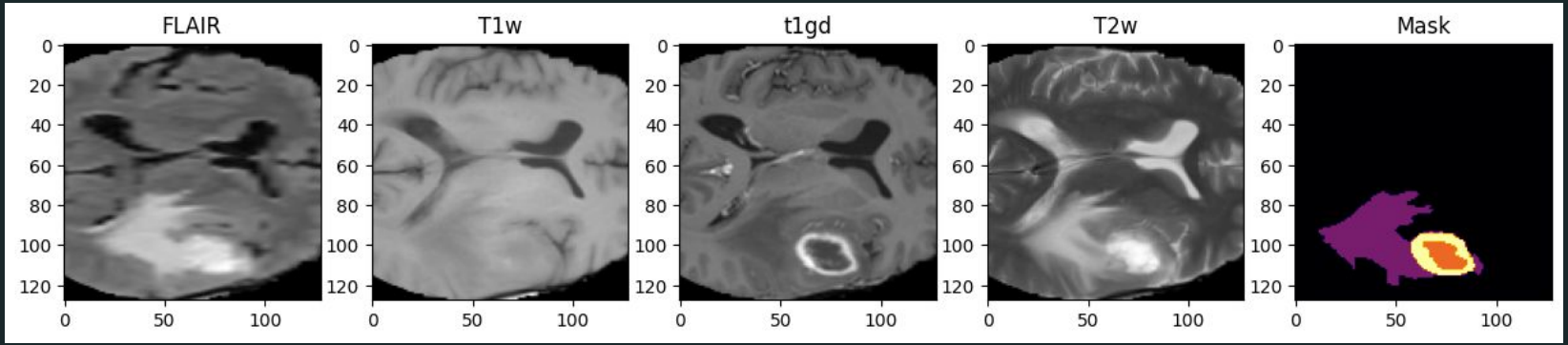
# Preprocessing

Then we **conducted** our preprocessing according to different works on this topic (Isensee F. et al., 2018; medical segmentation decathlon, 2018)

- **normalize** data between 0 and 1 for each channel within each subject
- volumes are **downsampled** from the middle to 128x128x128x4 in order to reduce computational complexity and enhance the representation of minority classes by eliminating background noise
- apply **One Hot Encoding** to the categorical target variables



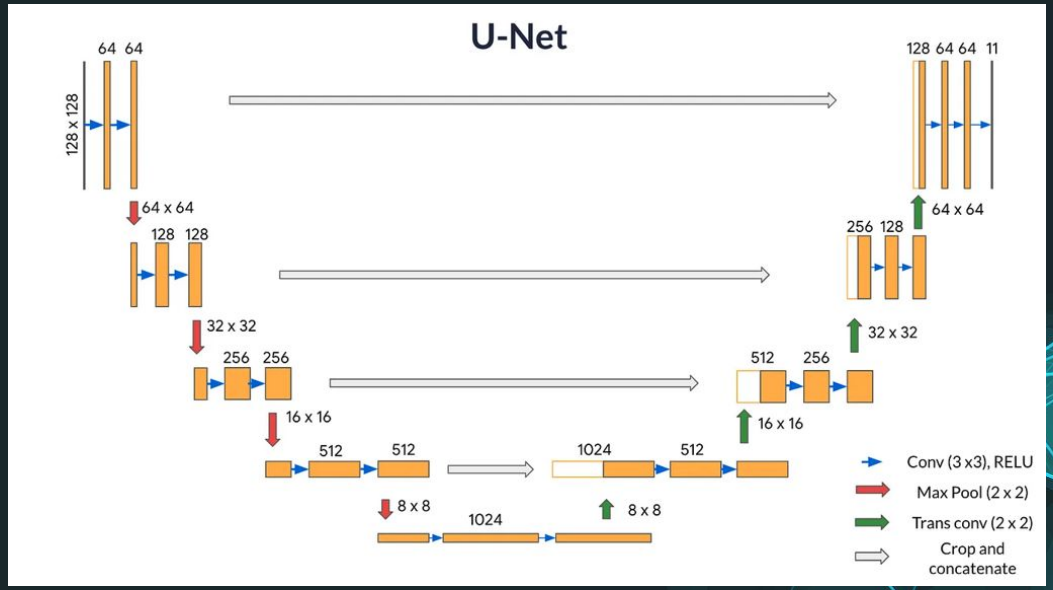

Uncutted volume and relative mask




Cutted volume and relative mask

# U-Net

U-Net is a **convolutional neural network** (CNN) initially developed for **biomedical image segmentation** but applicable to various **task** (Ronneberger O., 2015).



# U-Net

## Key features

- **encoder - decoder architecture:**
  Consists of a **contracting path** and an **expansive path**, where the main encoding and decoding structures are repeated with **different levels** of complexity

- **skip connections:**
  Directly link **corresponding layers** in the encoder and decoder, preserving **spatial information**, crucial to **reassemble** extracted feature

## POSSIBLE VARIATIONS

- input dimensionality
- Initial number of filters
- number of layers
- use of techniques like batch normalization or spatial dropout

# Fixed Architecture

- **Encoding Block (Contracting Path)**
  - Convolutional Layers: Extract features with two 3x3 convolutional layer each followed by ReLU activation, typically the number of filters  follows powers of two, increasing in contracting path and decreasing  in extensive path.
  - Max Pooling: 2x2 pooling reduces spatial dimensions while retaining key features.
- **Decoding Block (Expansive Path)**
  - Up-Convolution: 2x2 up-convolutions to increase spatial dimensions.
  - Concatenation: Combines up-sampled features with corresponding encoder features.
  - Convolutional Layers: Two 3x3 convolutions refine features.
- **Final Output**
  - 1x1 Convolution: Reduces feature channels to the number of classes for the final segmented output.
  - Activation Function: Sigmoid or Softmax for binary and multiclass segmentation respectively.

# Fixed Parameters

## Metrics

Dice-Sørensen coefficient:

$$DSC = \frac{2\,TP}{2\,TP + FP + FN}$$

Intersection over Union:

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|}.$$

## Optimizer

***Adam*** with learning rate schedule:

*Initial rate*: 5e-4
*Exponential decay rate*: 0.985 per epoch

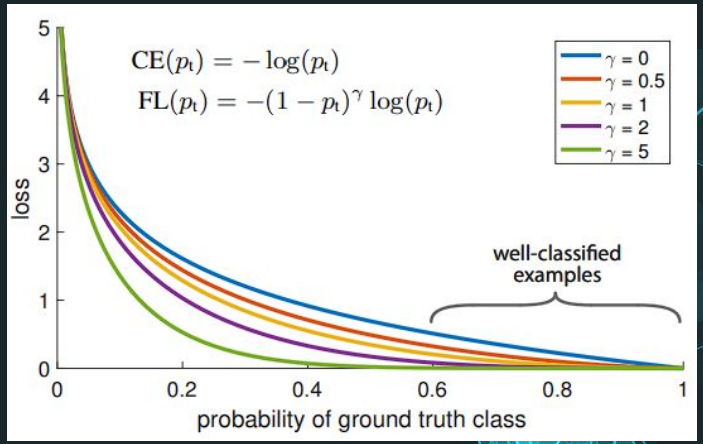
## Loss Function

Categorical Focal Cross-entropy

FL(p□)= - $\alpha$ ( 1 - p□ )$^{\gamma}$ CategoricalCE(y□ , y□)

where Categorical Cross-Entropy is:

- $\Sigma$ y□ *log*(y□)

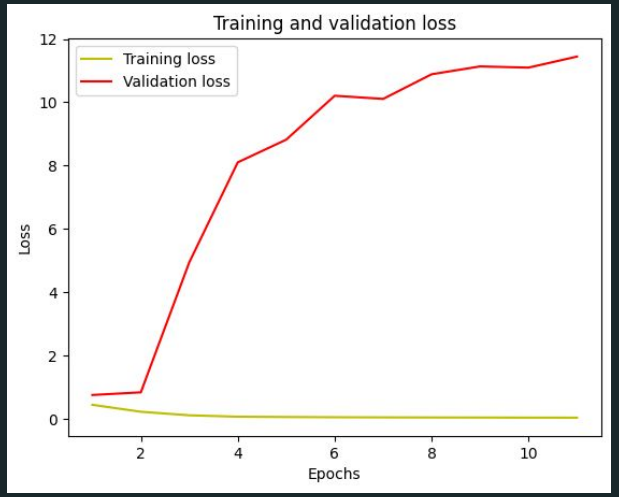


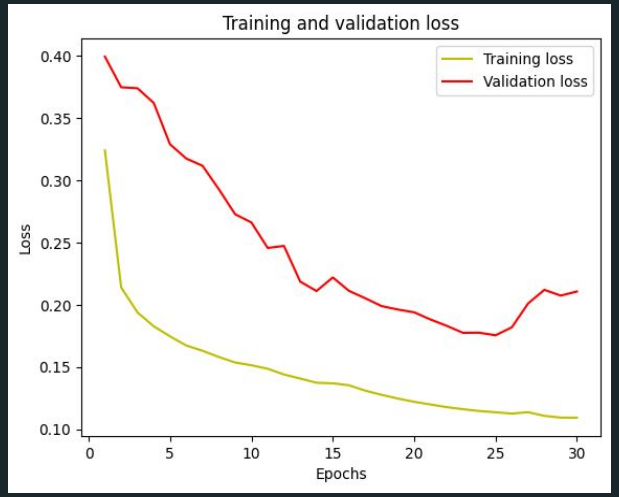
# 3D Unet

We initially tested the 3D version of the UNet but the model exhibited heavy overfitting from the early epochs (image on the left).

We tried to address this problem implementing Data Augmentation inside the training phase but this approach resulted too computationally expensive, so we decided to apply data augmentation as preprocessing, doubling the size of the train set and increasing its variability. This approach improved the situation but overfitting persisted (image on the right).

Given these observations, we decided to focus our work on 2D networks.

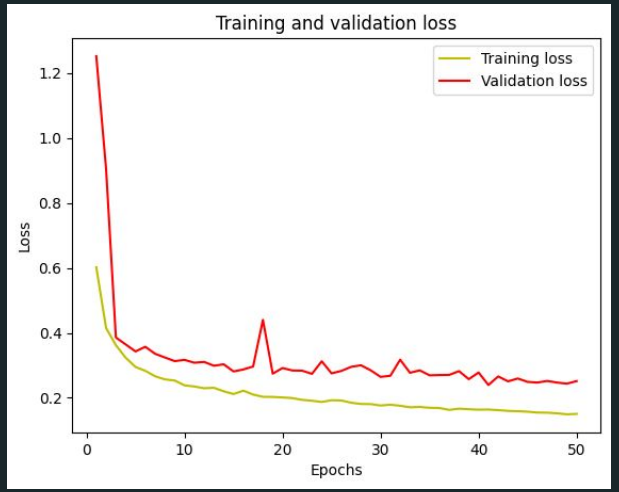


# 2D Unet

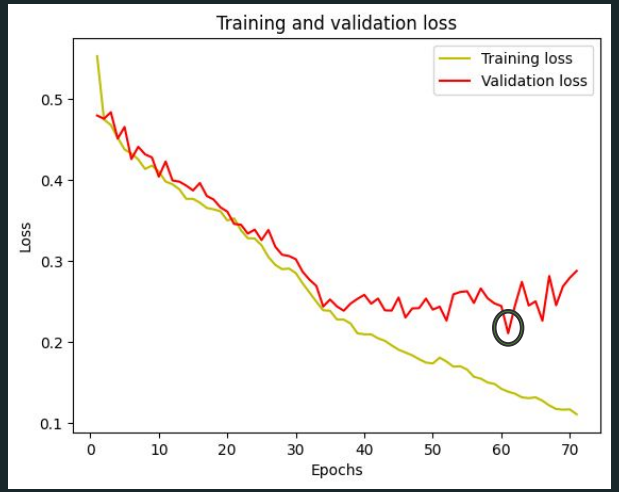
## *Model 1*

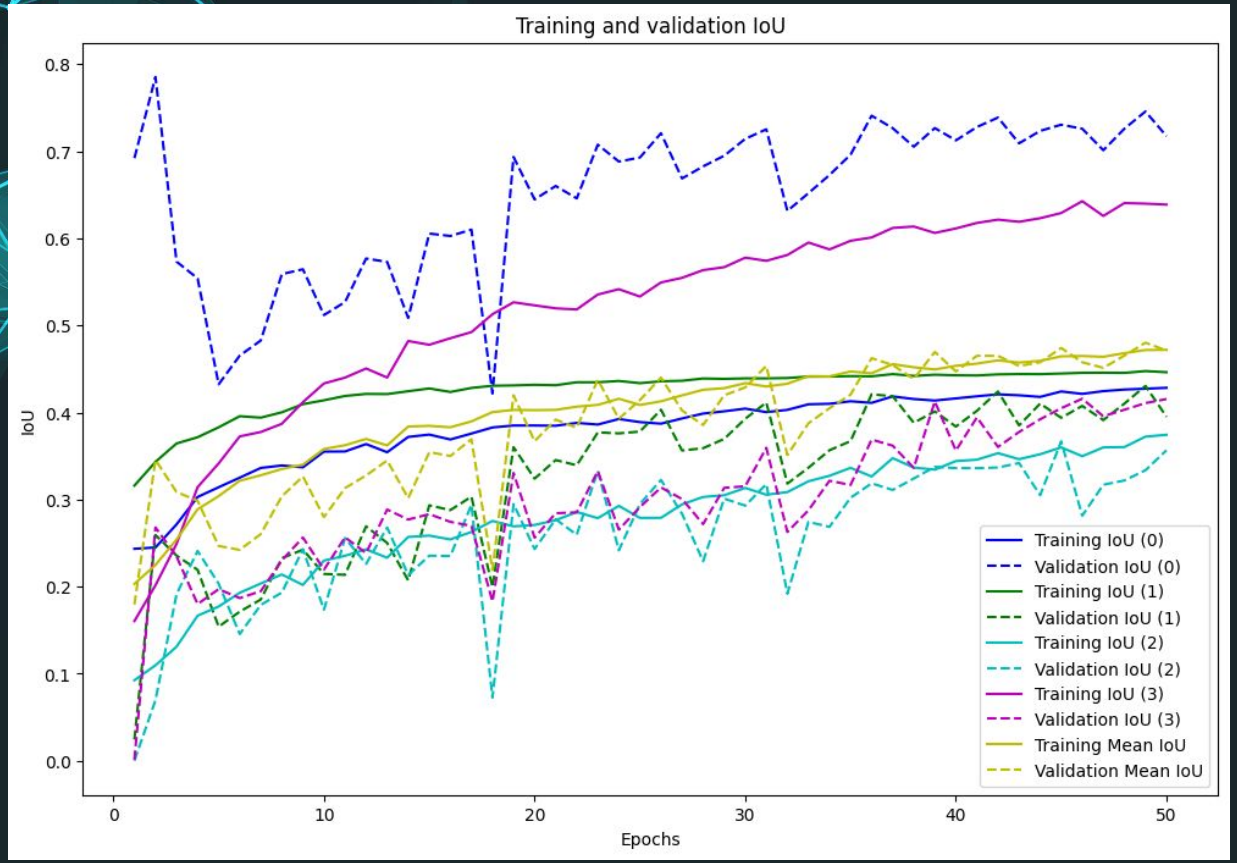
- 3 Convolutions
- 32 Starting filters
- Spatial dropout (0.2)
- Batch Normalization



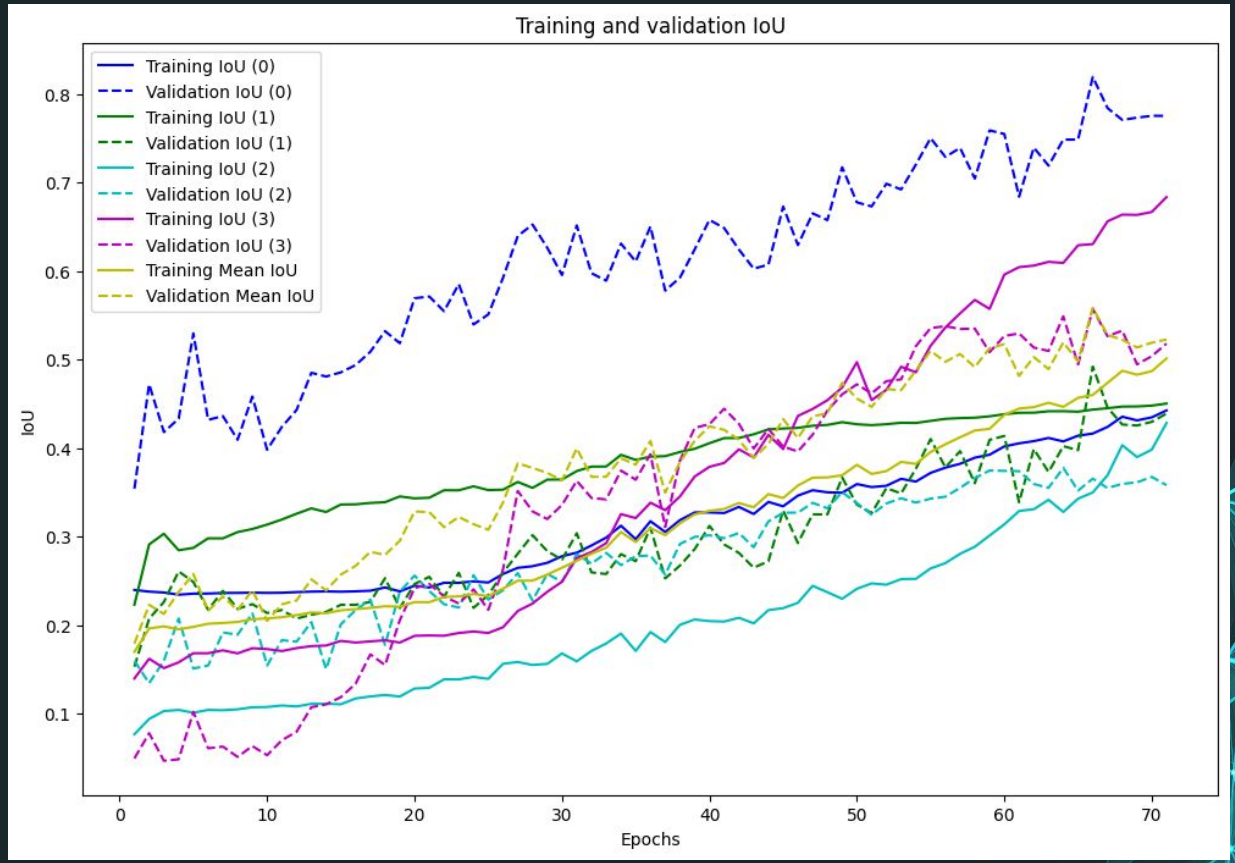

## *Model 2*

- 4 Convolutions
- 64 Starting filters
- NO dropout
- NO Batch Normalization






Average and “single class” IoU of Model 1




Average and “single class” IoU of Model 2

# Evaluation

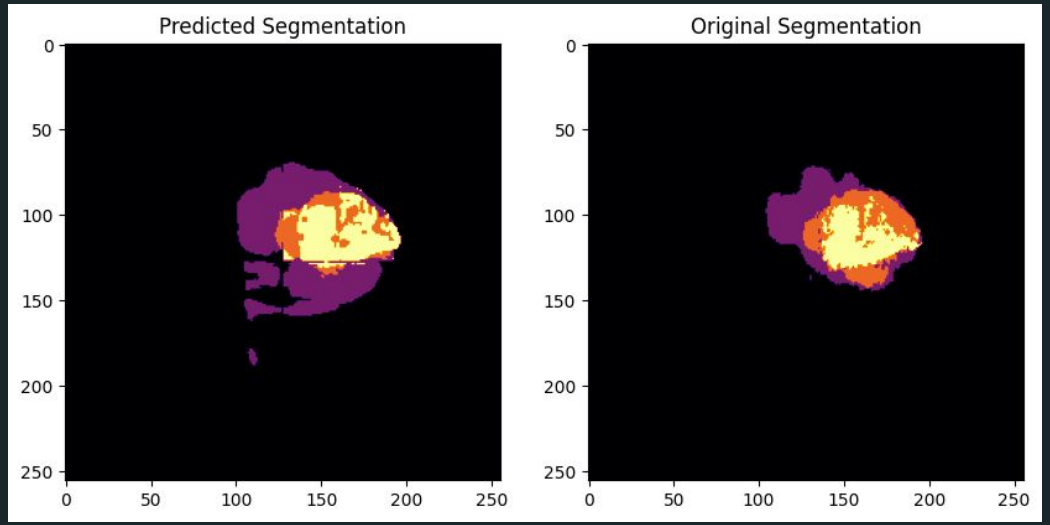
The evaluation was done on the test set, previously pre-processed like the training and test set.

| | Model 1 | Model 2 |
|---|---|---|
| *Loss* | 0.303 | 0.214 |
| *Accuracy* | 0.923 | 0.941 |
| *Dice Coefficient* | 0.421 | 0.446 |
| *Mean IoU* | 0.461 | 0.489 |
| *IoU class 0* | 0.936 | 0.681 |
| *IoU class 1* | 0.327 | 0.321 |
| *IoU class 2* | 0.245 | 0.368 |
| *IoU class 3* | 0.336 | 0.581 |

# Prediction

Prediction was made using model 2.

The sample image was upscaled from 240x240x155 to 256x256x155 to fit the input shape without overlapping.



# Limitations & future developments

- **COMPUTATIONAL CHALLENGE**

  One of the primary challenges faced is the computational complexity of applying data augmentation, essential to achieve good results, especially in 3D networks. This process is resource and time consuming, wich limited our ability to fully explore its benefits.

  Despite that, Data Augmentation (especially more complex, like the one used in 3D networks) remains the main avenue for improving our results.

# Limitations & future developments

- **HYPER-PARAMETERS TUNING**

  **FIXED**: Some parameters such as the optimizer and the loss function were kept constant to maintain consistency while comparing different configurations. However exploring different possibilities could potentially enhance the model's performance.

  **NON-FIXED**: Techniques like batch/group normalization, dropout and L2 regularization were explored during the project, however detailed experimentations are crucial for achieving better generalization and prevent overfitting.

# References


- Antonelli, M., Reinke, A., Bakas, S. *et al.* The Medical Segmentation Decathlon. *Nat Commun* 13, 4128 (2022). https://doi.org/10.1038/s41467-022-30695-9.
- Isensee F., Maier-Hein K.H., Kickingereder P., Bendszus M., Wick W., and 3rd International Workshop on Brainlesion, BrainLes 2017 Held in Conjunction with Medical Image Computing for Computer Assisted Intervention , MICCAI 2017 2017 09 14 - 2017 09 14. 2018. "Brain Tumor Segmentation and Radiomics Survival Prediction: Contribution to the BRATS 2017 Challenge." *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 10670 LNCS: 287–97. https://doi.org/10.1007/978-3-319-75238-9_25.
- Ronneberger O., Fischer P., Brox T., and 18th International Conference on Medical Image Computing and Computer-Assisted Intervention, MICCAI 2015 2015 10 05 - 2015 10 09. 2015. "U-Net: Convolutional Networks for Biomedical Image Segmentation." Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 9351: 234–41. https://doi.org/10.1007/978-3-319-24574-4_28.
- Zeineldin, R.A., Karar, M.E., Burgert, O., Mathis-Ullrich, F. (2023). Multimodal CNN Networks for Brain Tumor Segmentation in MRI: A BraTS 2022 Challenge Solution. In: Bakas, S., *et al.* Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2022. Lecture Notes in Computer Science, vol 13769. Springer, Cham. https://doi.org/10.1007/978-3-031-33842-7_11


THANK YOU