Semantic Segmentation of Adult Glioma

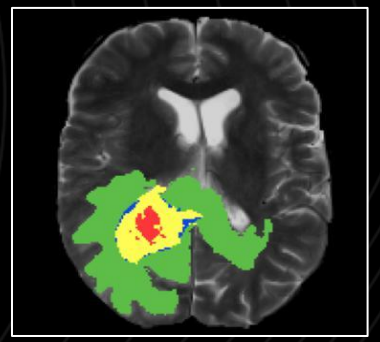
A STUDY ON BRAIN TUMOR SEGMENTATION

Contents

- Abstract
- Introduction
- Dataset
- Model
  - U-NET
  - Dilated Inception UNET
  - UNETR
  - Swin-Unet
  - UNETR + SEB
  - UNETR + CBAM
  - UNET + Attention
- Loss
- Results
- References

Introduction

- Gliomas are tumor that affect the Central Nervous System of any human and survival rate of patients diagnosed is low(<15 months)
- MRI is the preferred technique to analyze due to non-ionizing, non-invasive and good spatial and temporal resolution.
- Manual segmentation – Time consuming and introduces human error.
- Automatic and Semi-automatic segmentation techniques have allowed experts to work more efficiently.
- MR Images have 4 overlapping modalities – FLAIR,T1,T2,T1C.



Constituents of Gliomas

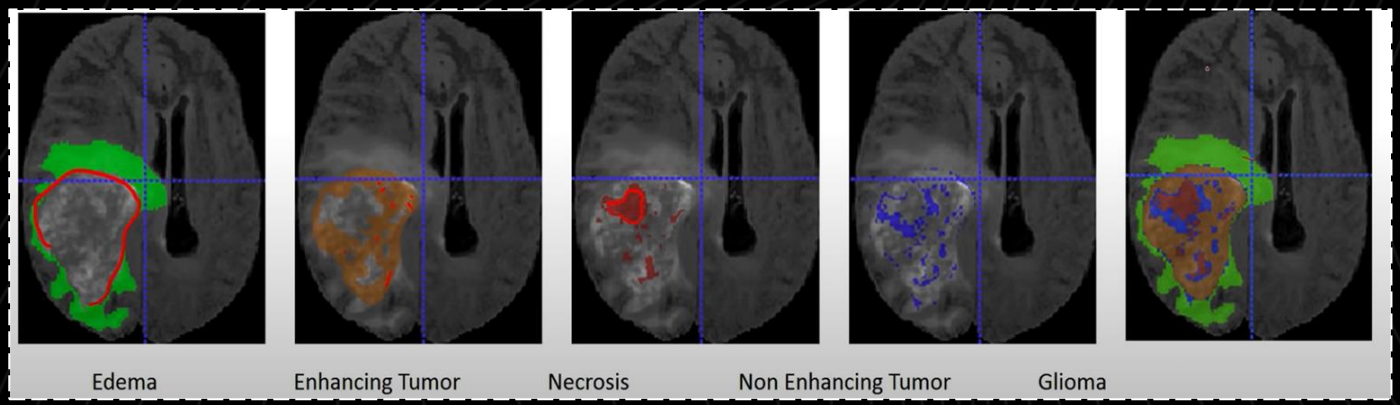
Constituents of gliomas are:

1. ***Edema*** : Collection of fluid or water. Best seen in FLAIR & T2. Fingerlike projection
2. ***Necrosis*** : Accumulation of dead cells. Best seen in T1.
3. ***Enhancing Tumor*** : Indicate breakdown of blood brain barrier. Best Seen in T1C.
4. ***Non Enhancing Region*** : Regions in region that are neither edema, necrosis or enhancing tumor.

Tumor core (TC) is the necrotic and non-enhancing tumor core

Whole tumor (WT) is the peritumoral edema

Enhancing tumor(ET) is the GD-enhancing tumor



Dataset – BraTS 2020

BraTS dataset link : https://www.kaggle.com/datasets/awsaf49/brats20-dataset-training-validation

- Composed of Low (75) and High (210) grade glioma.
- Contains modalities - T1, T1-Gd, T2, FLAIR.
- Annotations: Segmentation labels for enhancing tumor, peritumoral edema, and necrotic and non-enhancing tumor core.
- Dimension of dataset is 240 x 240 ×155. (155 slices)

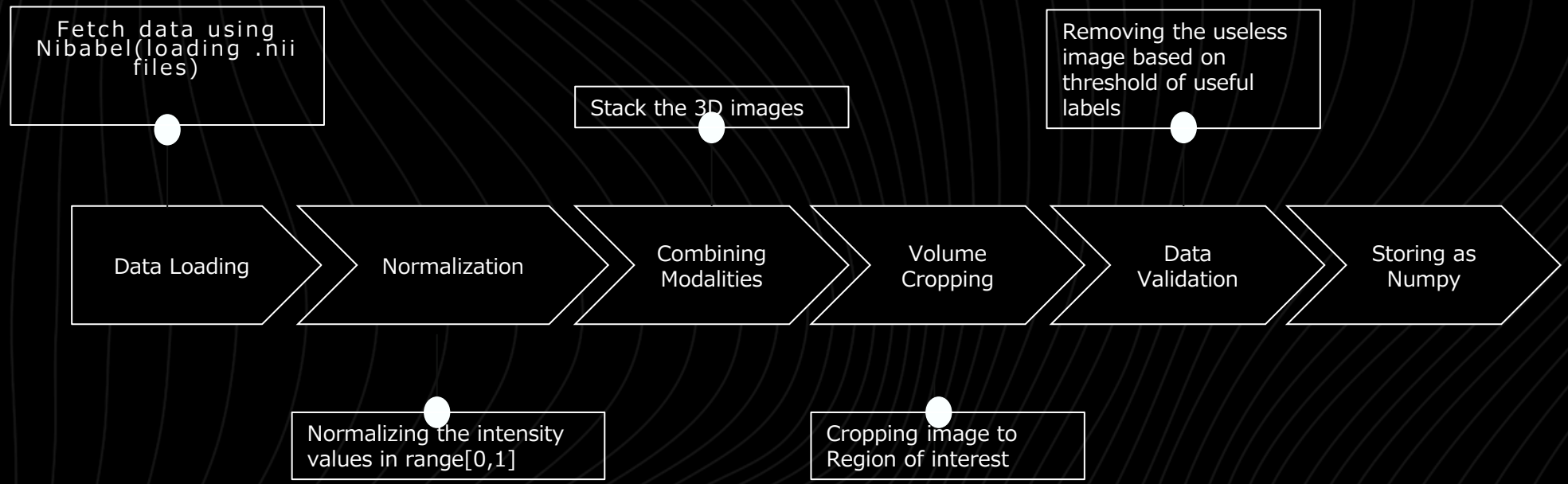
Pre-processing

This preprocessing pipeline prepares the medical image data for training machine learning models , enabling accurate and robust brain tumor segmentation.

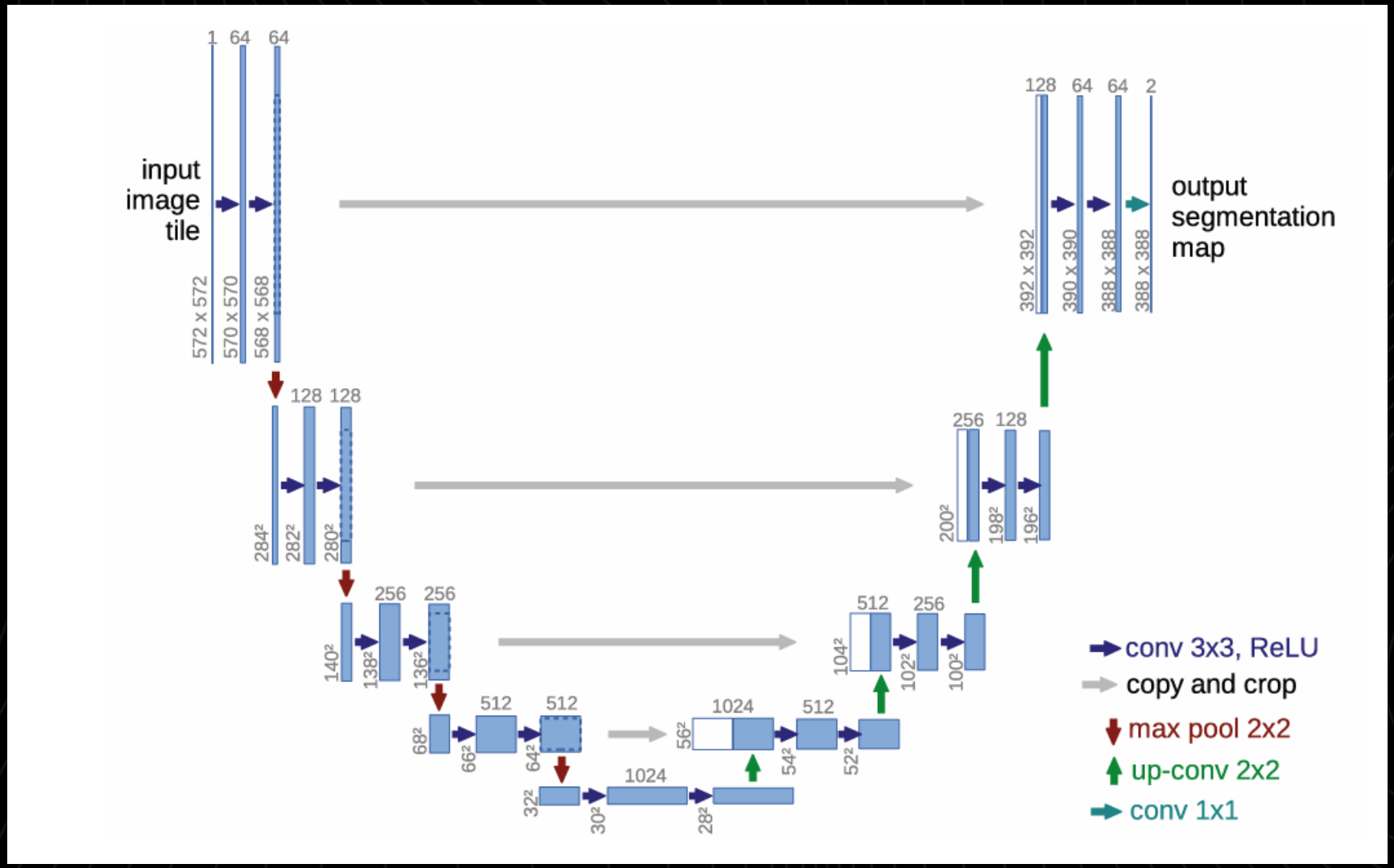
The main steps involved are:

- Loading Images & Masks:
  - Iterate through list of image file
  - Load each image and mask using nibable library
- Data Scaling:
  - Reshape the 3D img to 2D matrix to apply scaling using minmaxscaler
  - Reshape back to 3D Shape
- Combining Modalities:
  - Stack the preprocessed images (FLAIR,T1,T1C,T2) and create .npy array file.
  - Crop them into 128x128x128x4
  - Convert mask into a npy. File aswell and perform same crop operation (128x128x128)

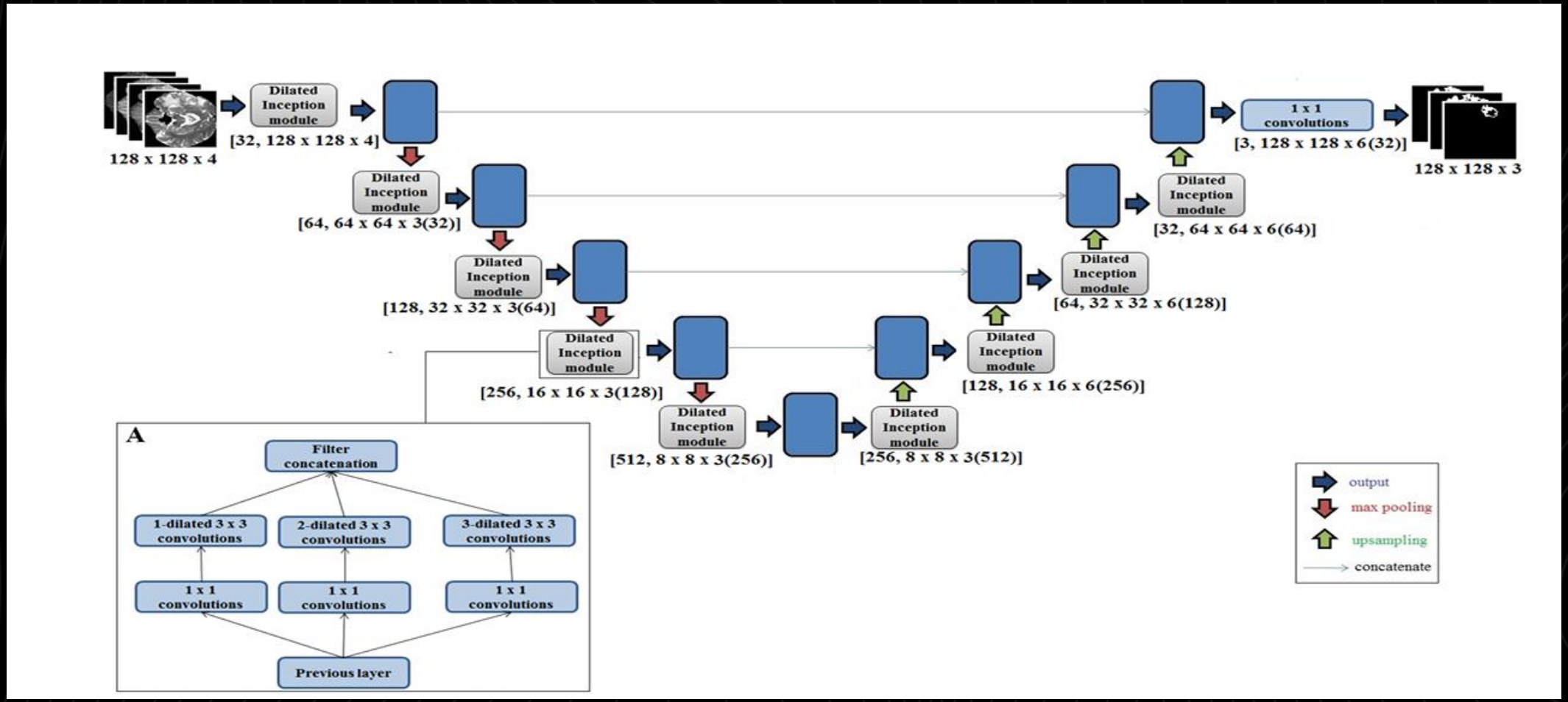
Contd.


Fetch data using Nibabel(loading .nii files)
Stack the 3D images
Removing the useless image based on threshold of useful labels
Data Loading
Normalization
Combining Modalities
Volume Cropping
Data Validation
Storing as Numpy
Normalizing the intensity values in range[0,1]
Cropping image to Region of interest

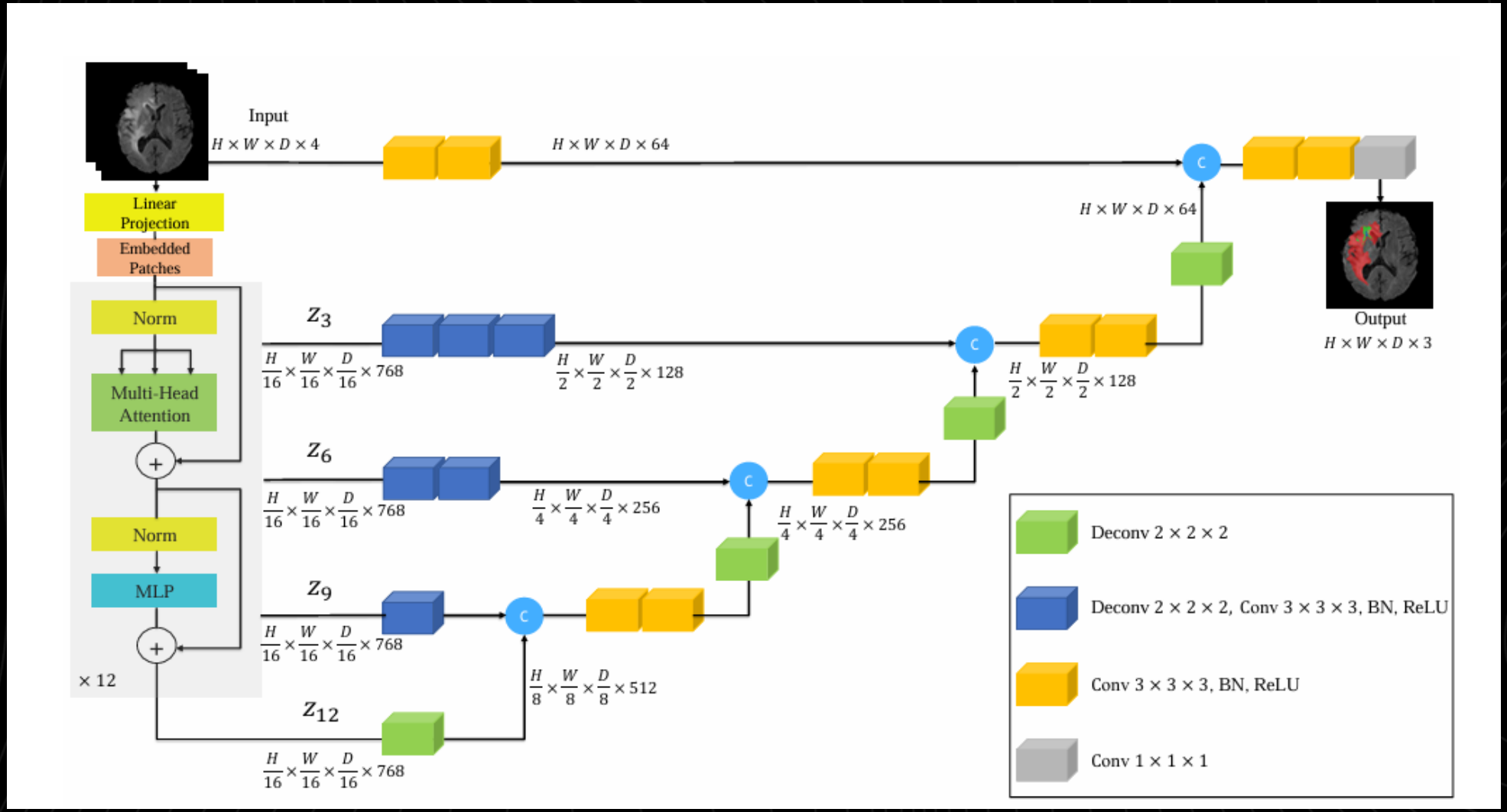
U-NET


input image tile
572 x 572
570 x 570
568 x 568
1 64 64
128 128
284²
282²
280²
256 256
140²
138²
136²
512 512
68²
66²
64²
1024
32²
30²
28²
1024
512
56²
54²
512
256
104²
102²
100²
256 128
200²
198²
196²
128 64 64 2
392 x 392
390 x 390
388 x 388
388 x 388
output segmentation map
conv 3x3, ReLU
copy and crop
max pool 2x2
up-conv 2x2
conv 1x1

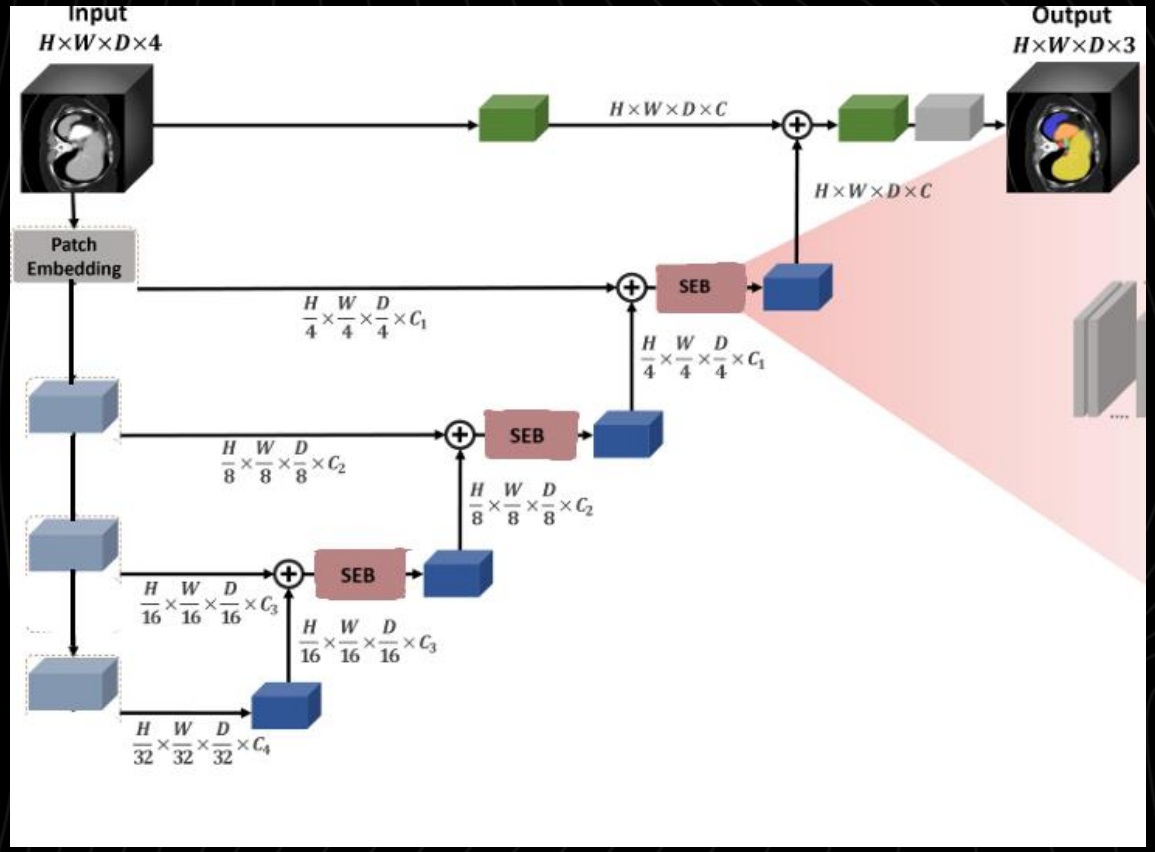
Dilated Inception UNET



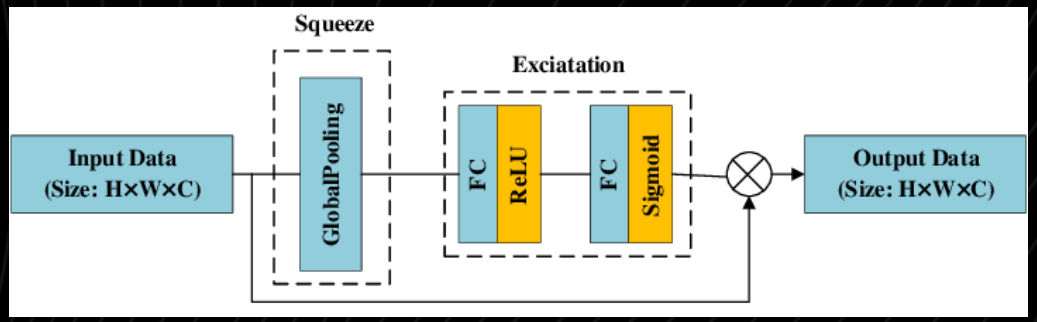
UNETR


Input
$H \times W \times D \times 4$
Linear Projection
Embedded Patches
Norm
Multi-Head Attention
Norm
MLP
× 12
z_3
z_6
z_9
z_{12}
$\frac{H}{16} \times \frac{W}{16} \times \frac{D}{16} \times 768$
$H \times W \times D \times 64$
$\frac{H}{2} \times \frac{W}{2} \times \frac{D}{2} \times 128$
$\frac{H}{4} \times \frac{W}{4} \times \frac{D}{4} \times 256$
$\frac{H}{8} \times \frac{W}{8} \times \frac{D}{8} \times 512$
C
Output
$H \times W \times D \times 3$
Deconv 2 × 2 × 2
Deconv 2 × 2 × 2, Conv 3 × 3 × 3, BN, ReLU
Conv 3 × 3 × 3, BN, ReLU
Conv 1 × 1 × 1

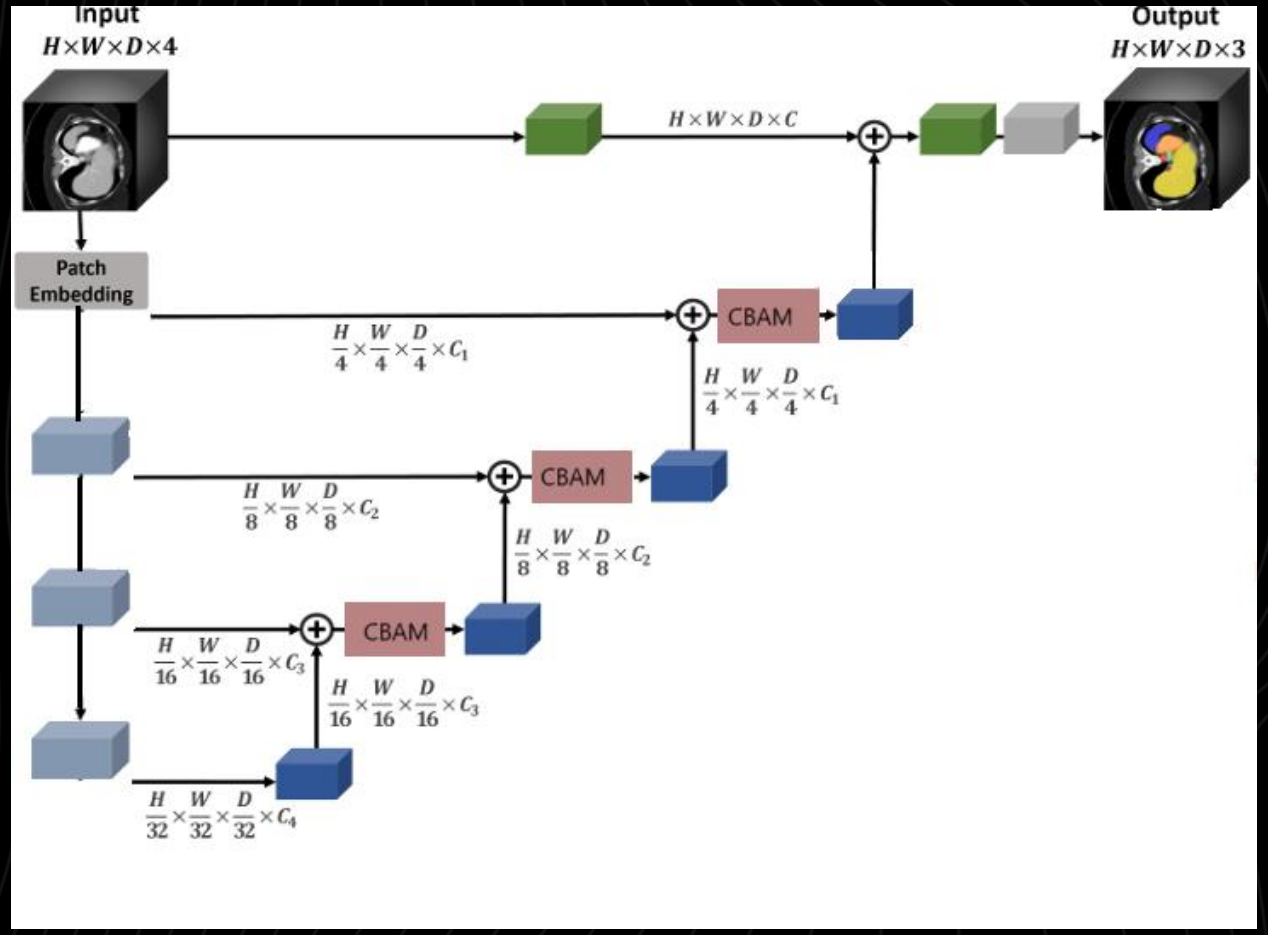
UNETR + SEB(Squeeze & Excitation Block)



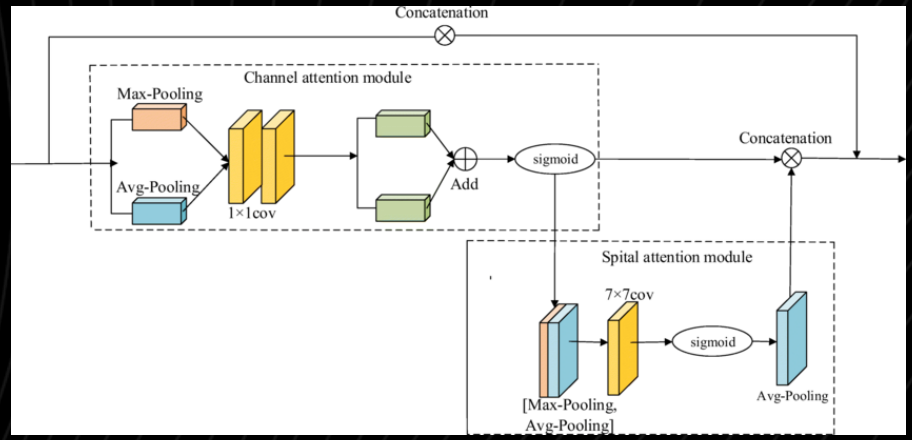
Squeeze & Excitation Block



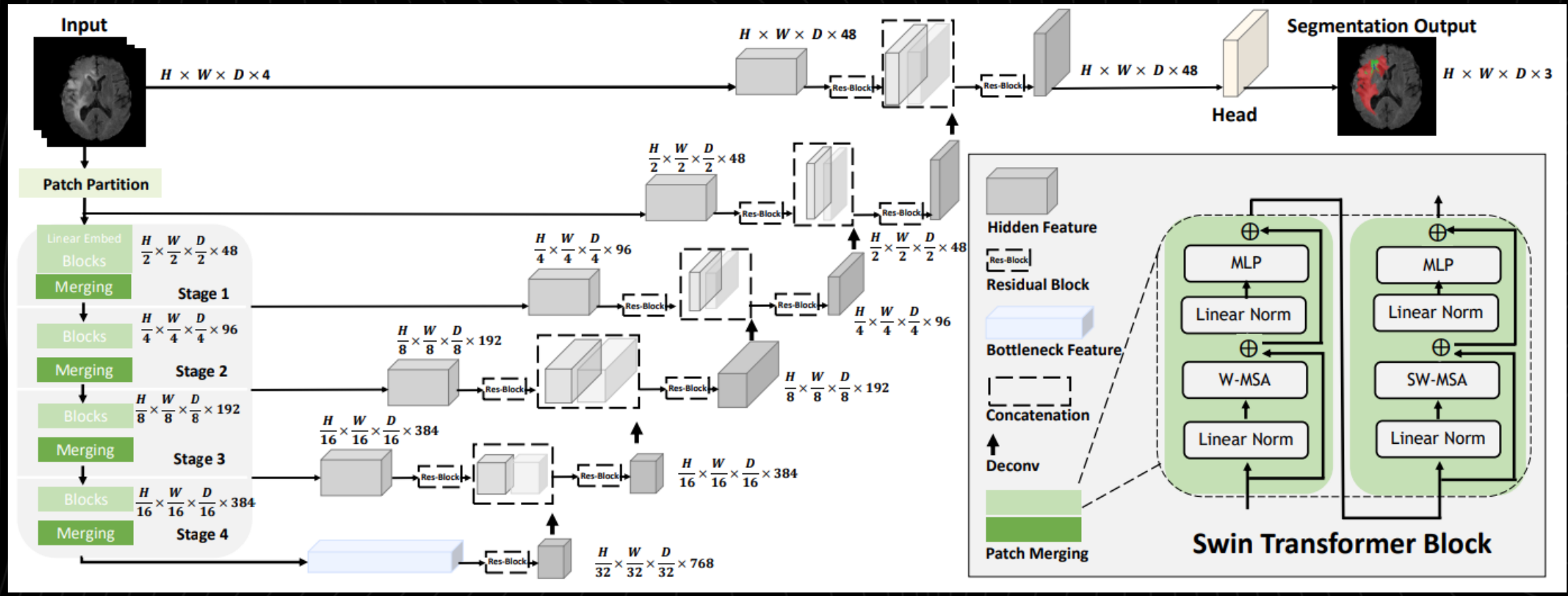
UNETR + CBAM (Convolutional Block Attention Module)



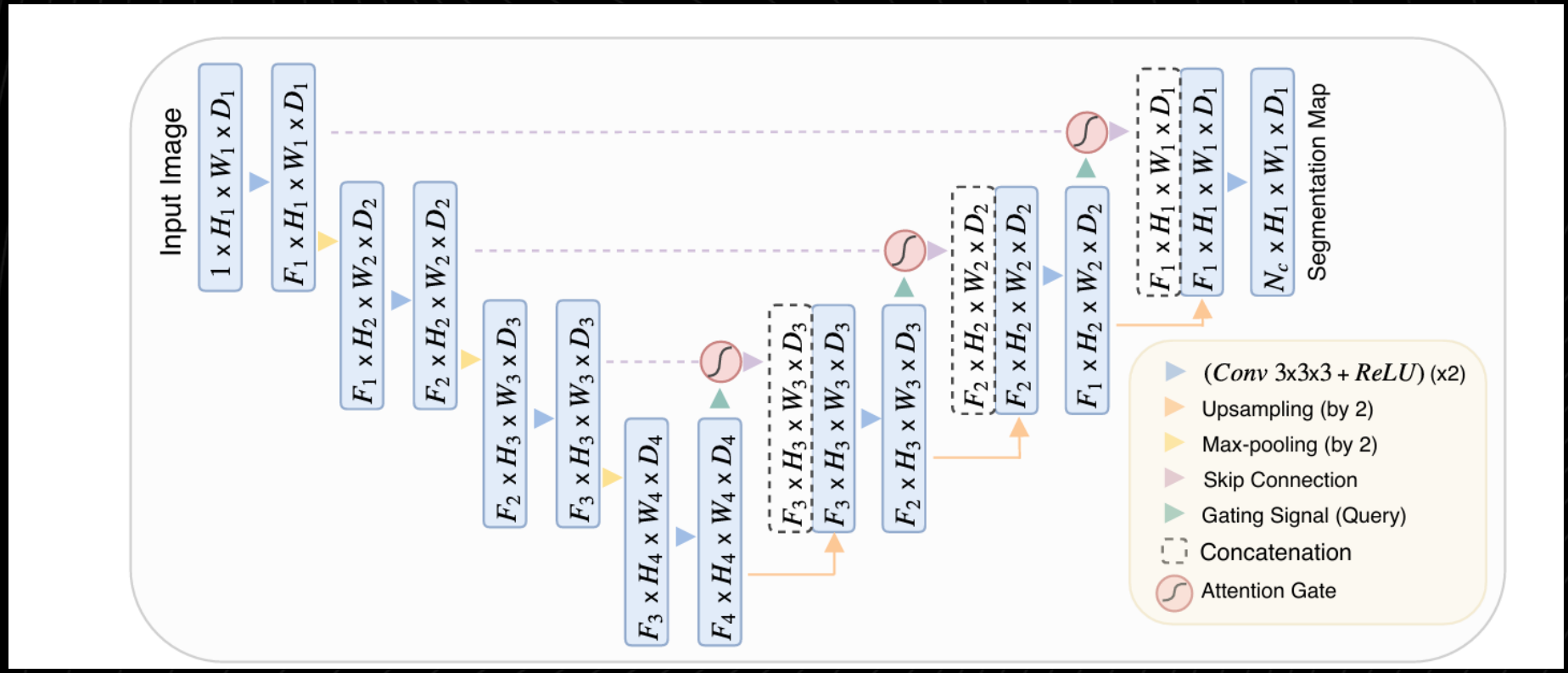
CBAM Block Module



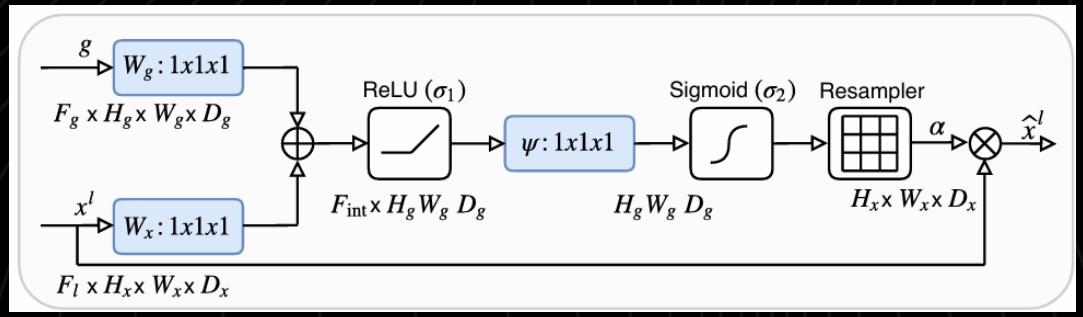
Swin-UNETR


Input
H × W × D × 4
Patch Partition
Linear Embed
Blocks
Merging
Stage 1
Stage 2
Stage 3
Stage 4
Res-Block
H × W × D × 48
Head
Segmentation Output
H × W × D × 3
Hidden Feature
Residual Block
Bottleneck Feature
Concatenation
Deconv
Patch Merging
MLP
Linear Norm
W-MSA
SW-MSA
Swin Transformer Block

UNET + Attention



Attention Gate



Dice Loss

- Dice loss, also known as the Sorsen-Dice Coefficient, measures the similarity between predicted segmentation masks and the ground truth masks. It is particularly useful for imbalanced datasets where the class of interest (I,e tumor region) might be small compared to background class
- Dice Coefficient for a single class is computed as

$$D = \frac{2\sum_i^N p_i g_i}{\sum_i^N p_i^2 + \sum_i^N g_i^2}$$

- Dice loss = (1 – D)

Tversky Loss

- The Tversky Index is a asymmetric similarity measure that is a generalisation of the dice coefficient and the Jaccard index.

$$TI = \frac{TP}{TP + \alpha FN + \beta FP}$$

- Tversky loss = 1 – TI
- The tversky index adds two parameters, α and β, where α + β = 1. In the case where α = β = 0.5, it simplifies into the dice coefficient. It simplifies to the Jaccard index if α = β = 1.
- By setting the value of α > β, you can penalize false negatives more. This becomes useful in highly imbalanced datasets where the additional level of control over the loss function yields better small scale segmentations than the normal dice coefficient.

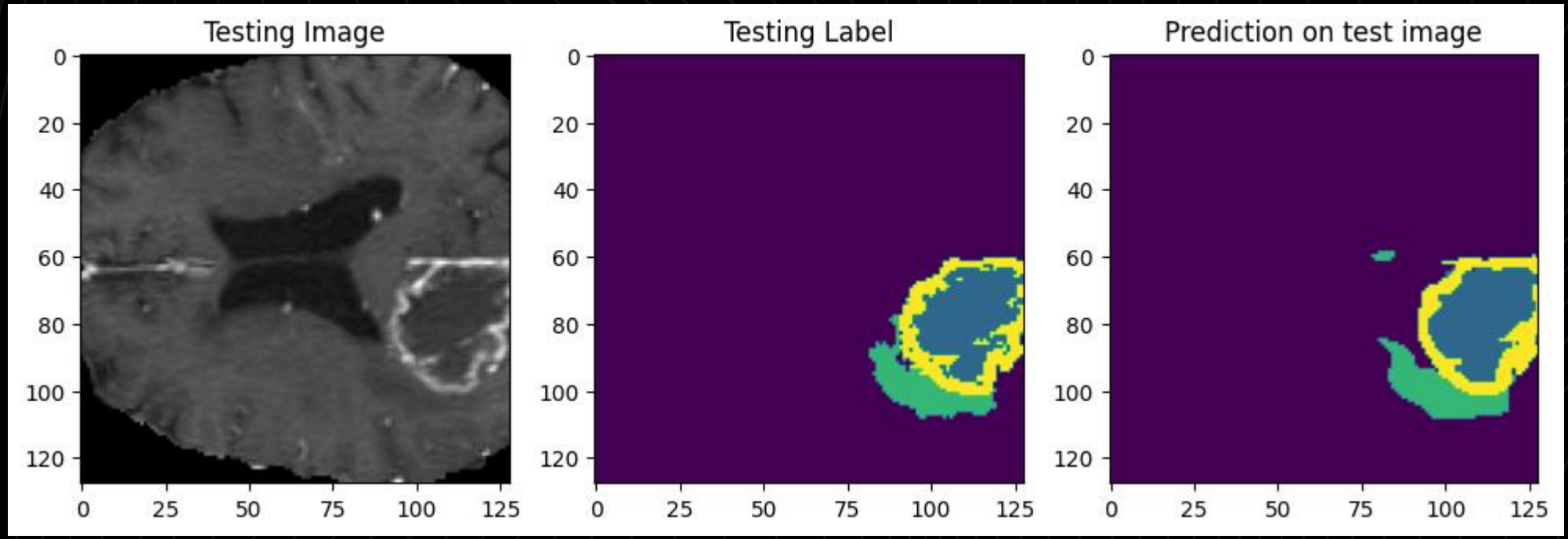
Losses

Model	Loss	Optimizer	Accuracy	Loss	TC	WT	ET
UNET	Dice	BCE+ Tversky	97.92	0.1433	0.76	0.8632	0.763
UNETR	Tversky	Adam	78.14	0.1366	0.6928	0.8561	0.7324
UNETR	DiceCE	Adam	79.81	0.1486	0.7562	0.8661	0.7759
UNETR+SEB	Tversky	Adam	79.63	0.1120	0.7477	0.8682	0.7823
UNETR+CBAM	DiceCE	Adam	79.77	0.1662	0.7481	0.8681	0.7854
Dilated Inception	Dice	Adam	67.97	0.3106	0.6366	0.8139	0.5888
UNET + Attention	Dice	Adam	98.80	0.1479	0.8999	0.8586	0.8123
Swin-UNETR *	Dice	Adam	55.57	0.5143	0.5052	0.8043	0.5557

* Final Results pending

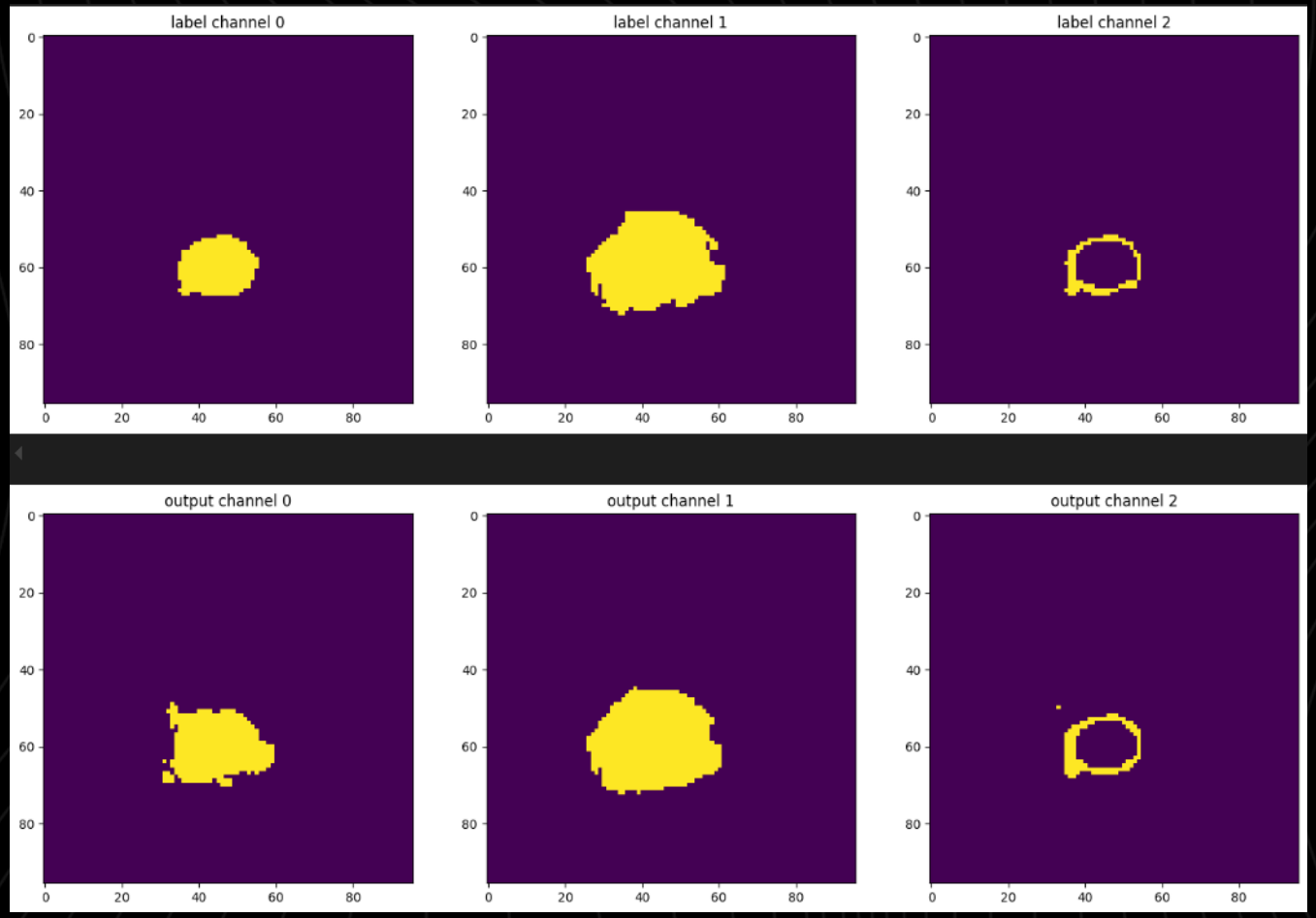
Results

- UNET



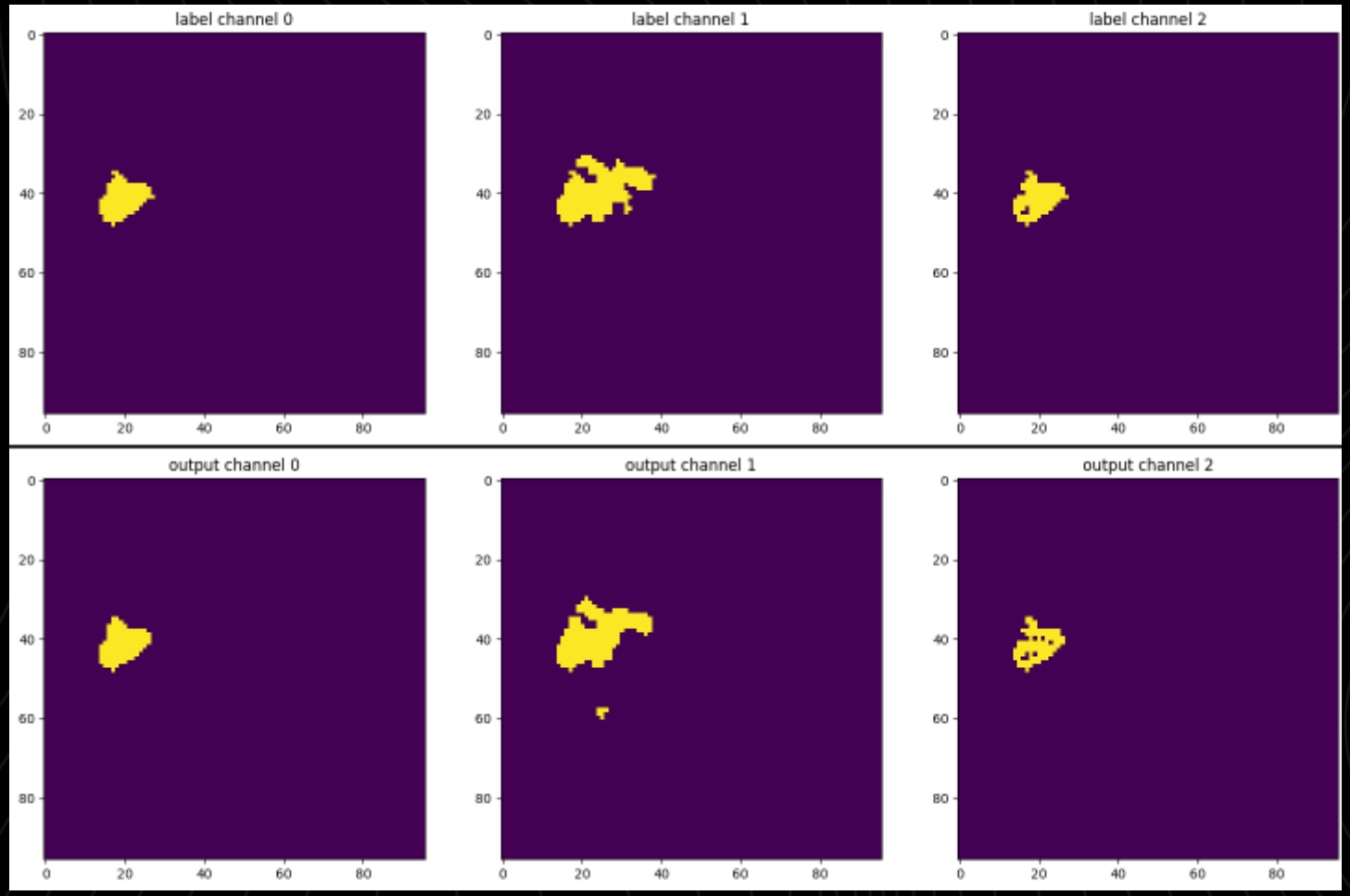
Results

- UNETR (Tversky loss)



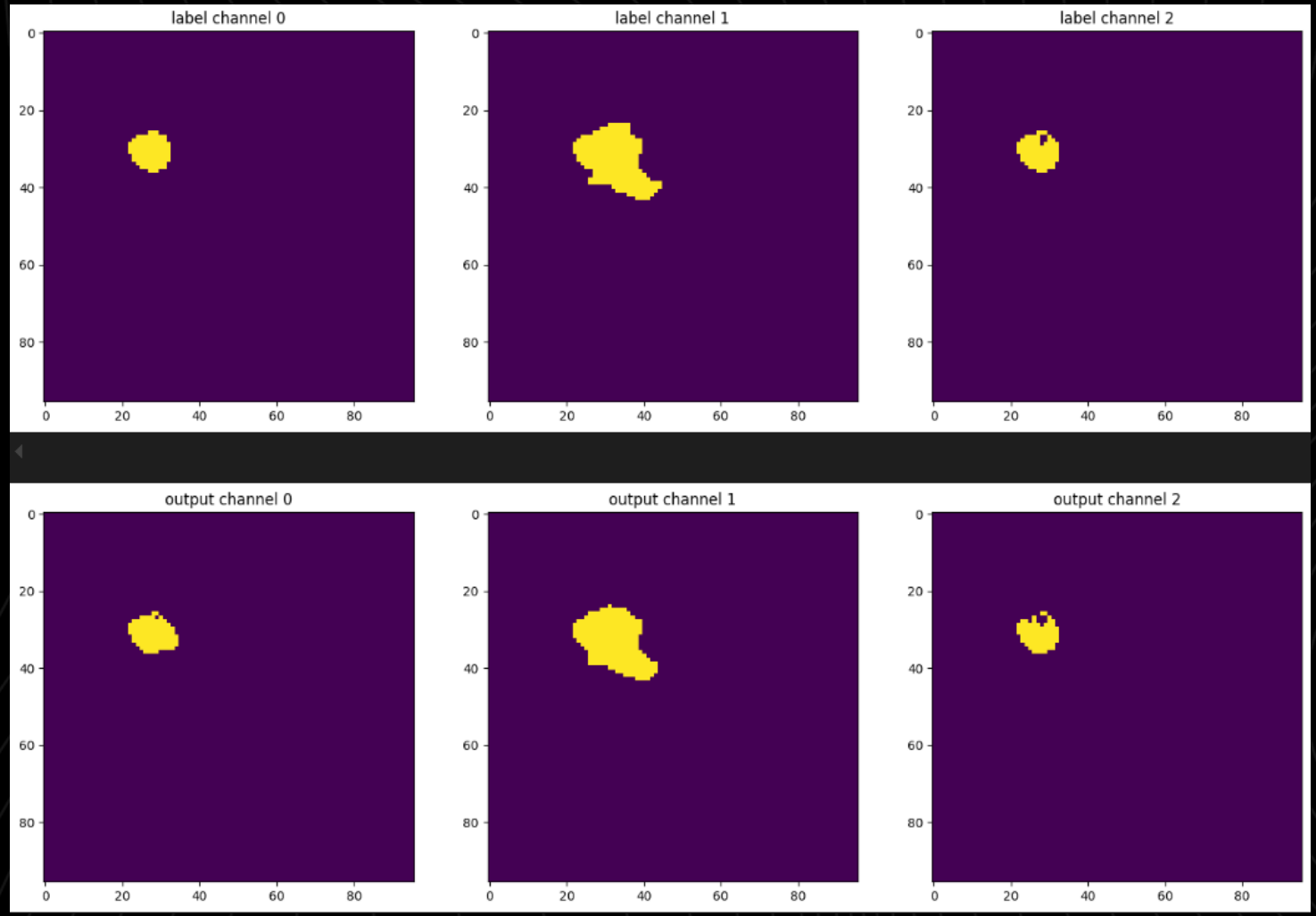
Results

- UNETR (DiceCE loss)



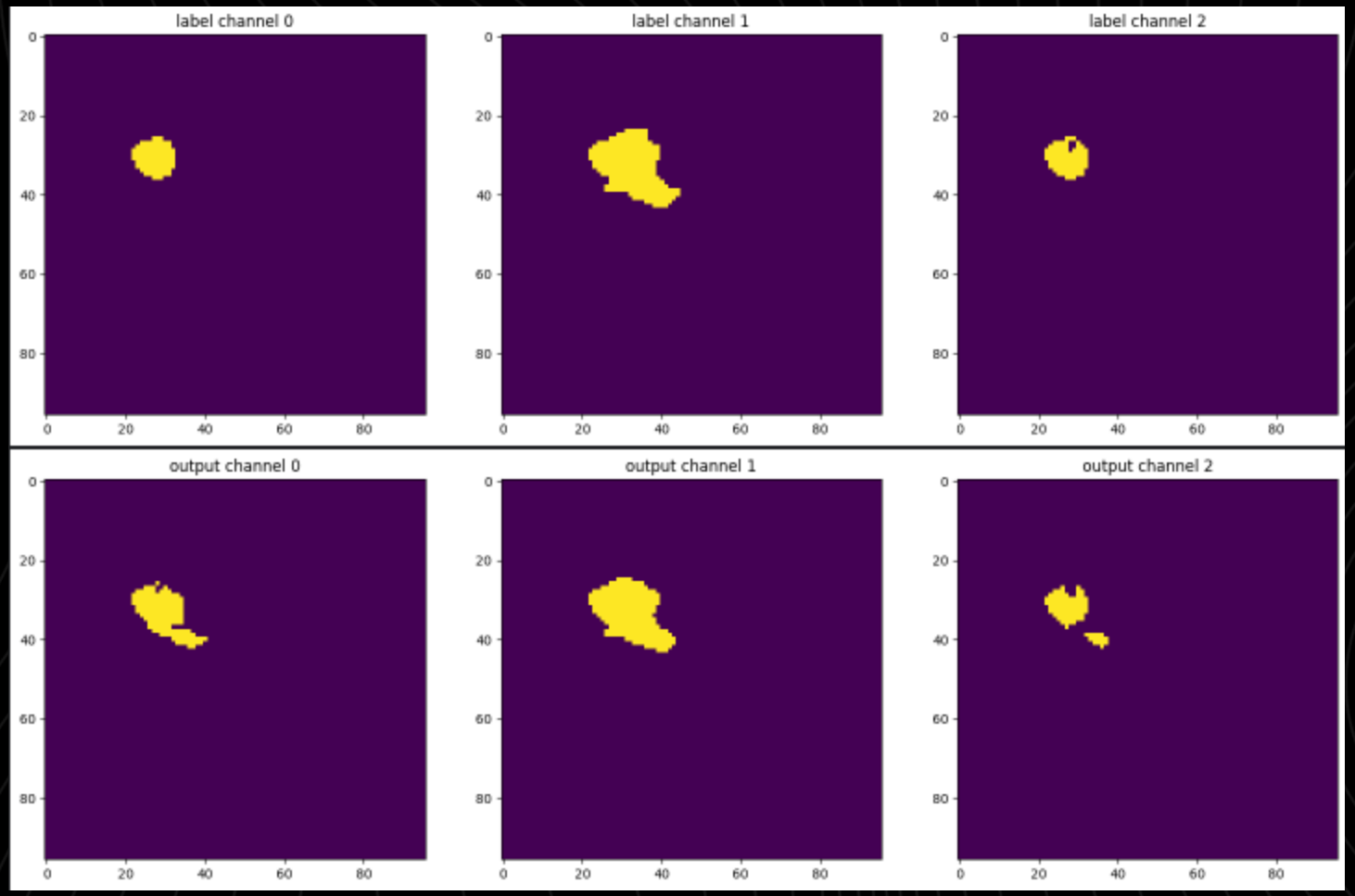
Results

- UNETR + SEB



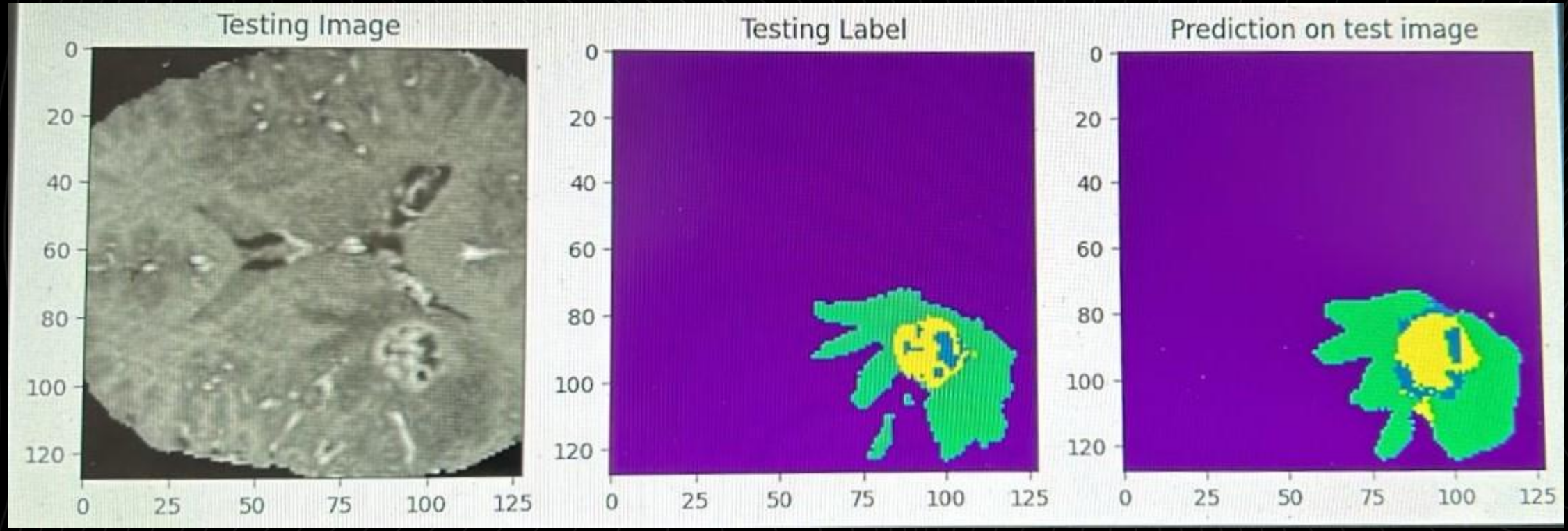
Results

- UNETR + CBAM



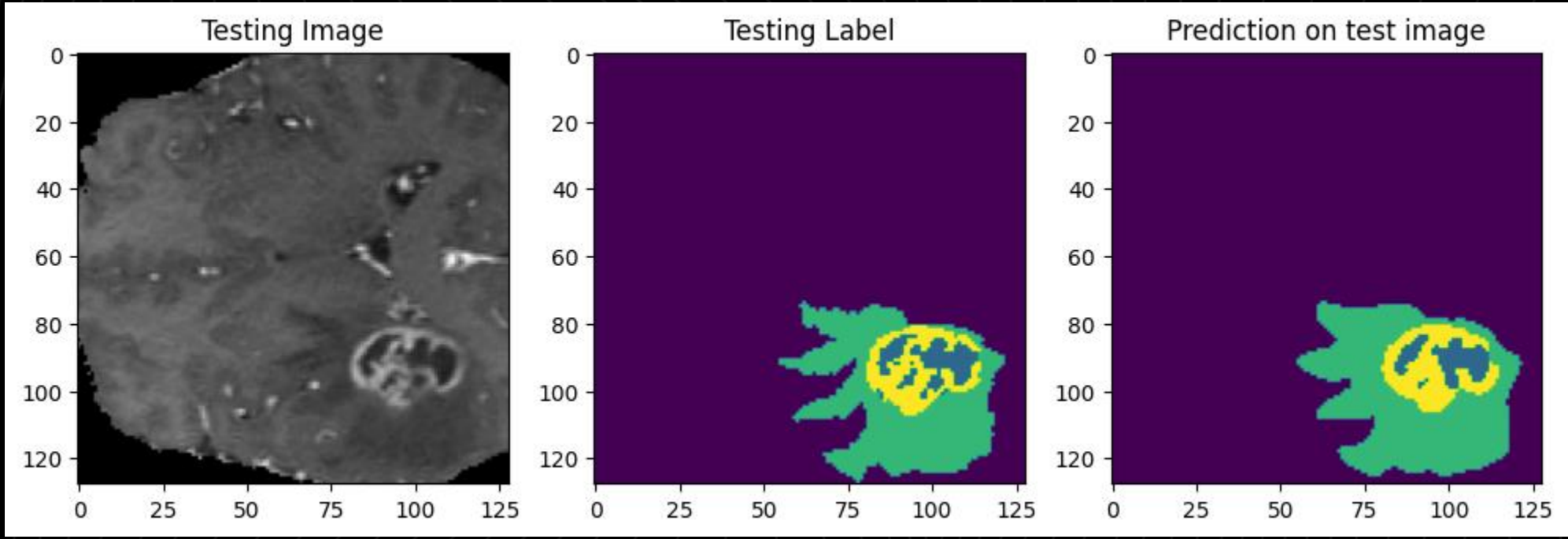
Results

- DIU-NET



Results

- UNET + Attention



References

- Attention Is All You Need
- U-Net: Convolutional Networks for Biomedical Image Segmentation (Olaf Ronneberger)
- MultiResUNet: Rethinking the U-Net architecture for multi modal biomedical image segmentation
- Inception-UDet: An Improved U-Net Architecture for Brain Tumor Segmentation
- DRI-UNet: dense residual-inception UNet for nuclei identification in microscopy cell image
- SwinBTS: A Method for 3DMultimodal Brain Tumor Segmentation Using Swin Transformer
- W-Net: A Deep Model for Fully Unsupervised Image Segmentation
- UNETR++: Delving into Efficient and Accurate 3D Medical Image Segmentation
- Depth-wise Squeeze and Excitation Block-based Efficient-Unet model for surface defect detection
- Dilated Inception U-Net (DIU-Net) for Brain Tumor Segmentation
- Attention U-Net: Learning Where to Look for the Pancreas