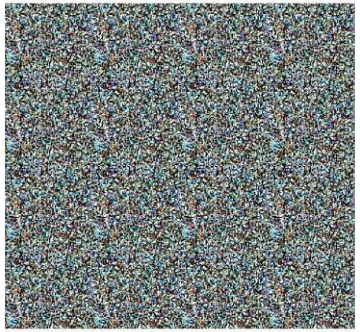
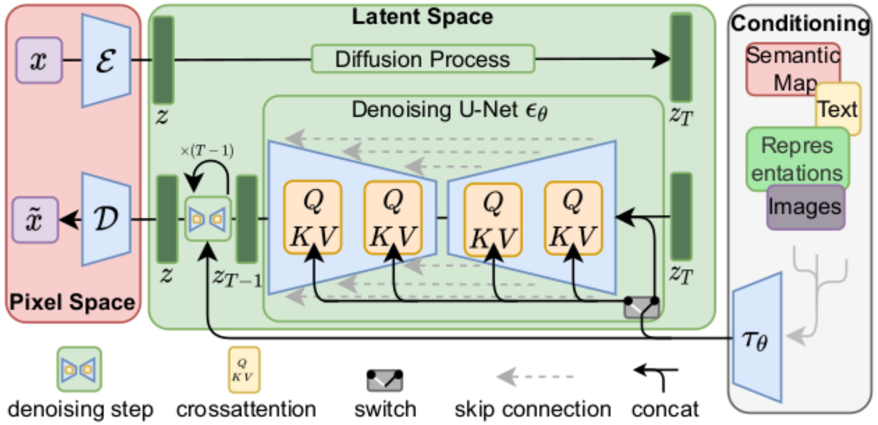
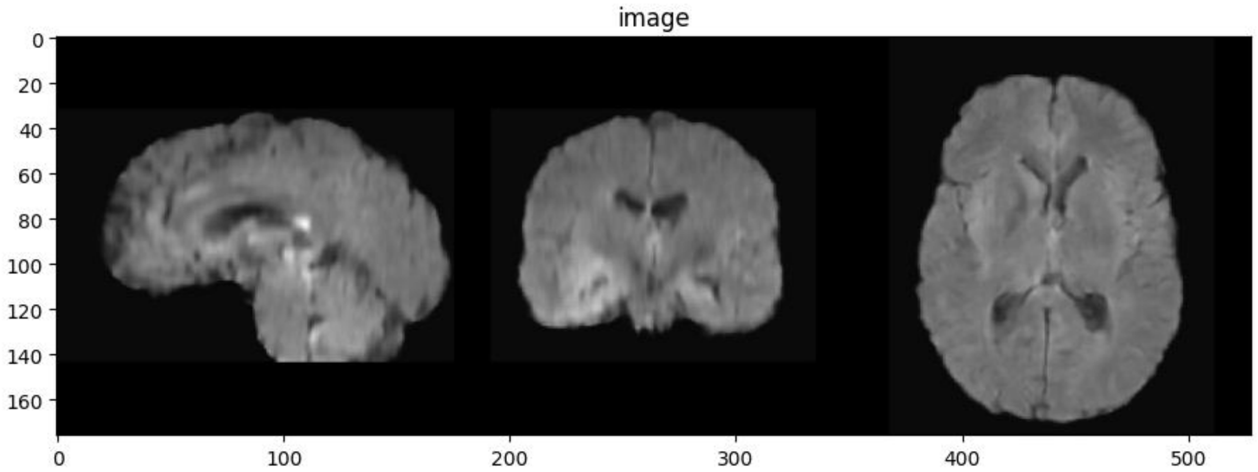


Random Noise input





3D MRI image Generation from Random Noise data

New 3D MRI image samples will be generated using a Latent Diffusion Models (LDM) trained on BraTS 2016 and 2017 data from the Medical Segmentation Decathalon.

The **Latent Diffusion Model (LDM)** is an approach to high-resolution image synthesis that reduces computational demands by separating the **compression** and **generation** phases. Instead of directly applying diffusion in pixel space (which is costly), LDMs first encode images into a **lower-dimensional latent space** using an **autoencoder**. The diffusion process then operates in this compressed space, making sampling more efficient while retaining perceptual quality.

Training Configuration of Autoencoder:

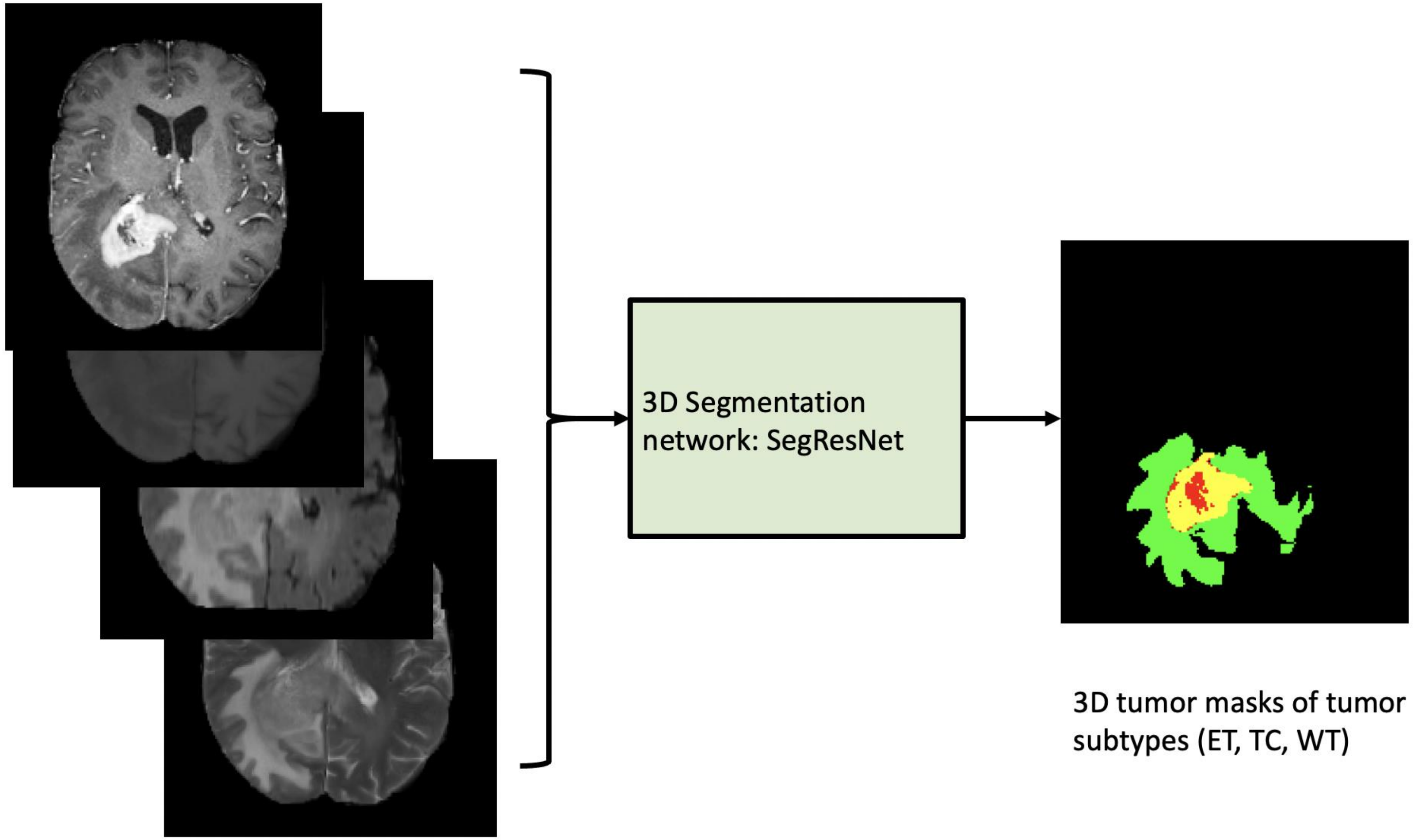
The autoencoder was trained using the following configuration:
GPU: at least 32GB GPU memory
Actual Model Input: 112 x 128 x 80
AMP: False
Optimizer: Adam
Learning Rate: 1e-5
Loss: L1 loss, perceptual loss, KL divergence loss, adversarial loss, GAN BCE loss

Input
1 channel 3D MRI Flair patches
Output
1 channel 3D MRI reconstructed patches
8 channel mean of latent features
8 channel standard deviation of latent features

Training Configuration of Diffusion Model:

The latent diffusion model was trained using the following configuration:
GPU: at least 32GB GPU memory
Actual Model Input: 36 x 44 x 28
AMP: False
Optimizer: Adam
Learning Rate: 1e-5
Loss: MSE loss

Training Input
8 channel noisy latent features
a long int that indicates the time step
Training Output
8 channel predicted added noise



4x 3D brain MRIs modalities
(T1c, T1, T2, Flair)

3D tumor masks of tumor
subtypes (ET, TC, WT)

volumetric (3D) segmentation of brain tumor

A pre-trained model for volumetric (3D) segmentation of brain tumor subregions from multimodal MRIs based on BraTS 2018 data.

The model is trained to segment 3 nested subregions of primary brain tumors (gliomas): the "enhancing tumor" (ET), the "tumor core" (TC), the "whole tumor" (WT)

Training configuration

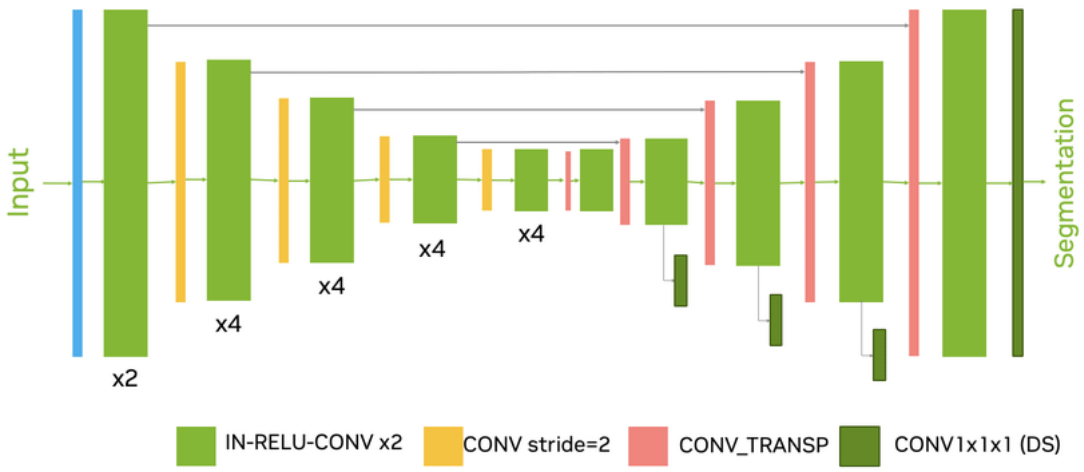
The training was performed with the following:
- GPU: At least 16GB of GPU memory.
- Actual Model Input: 224 x 224 x 144
- AMP: True
- Optimizer: Adam
- Learning Rate: 1e-4
- Loss: DiceLoss

Input

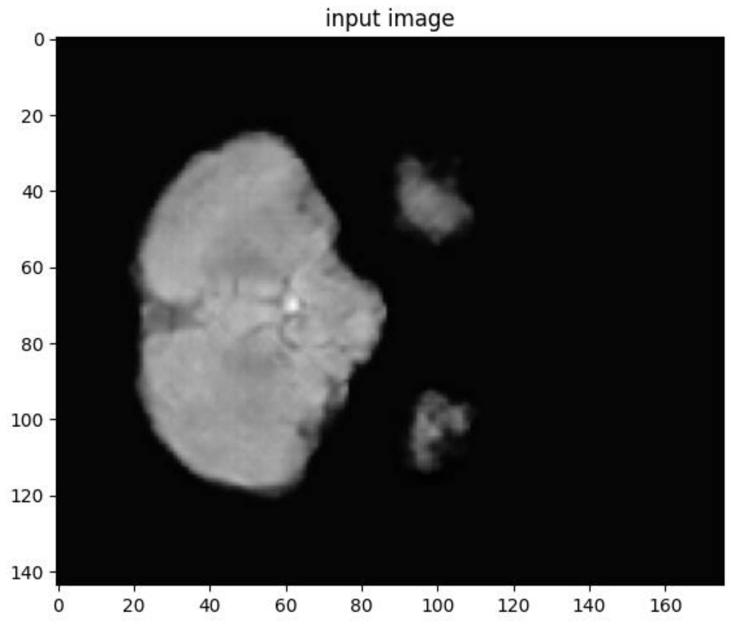
4 channel aligned MRIs at 1 x 1 x 1 mm - T1c - T1 - T2 - FLAIR

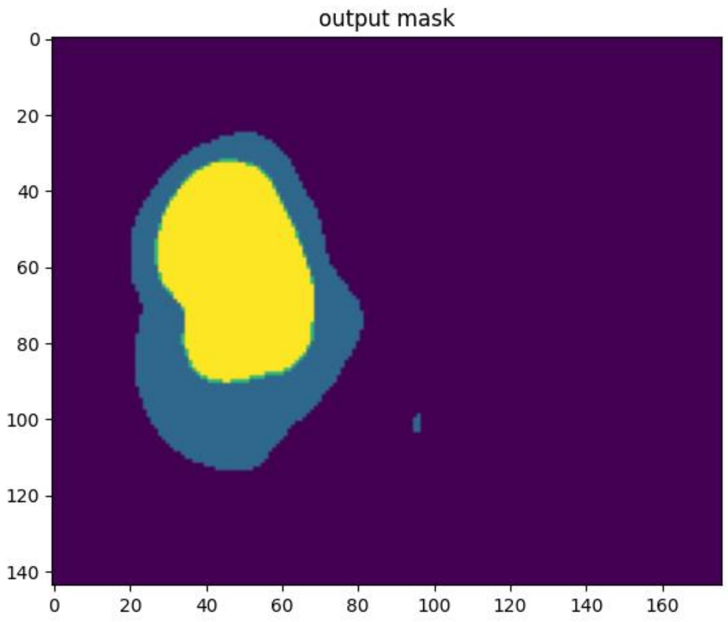
Output

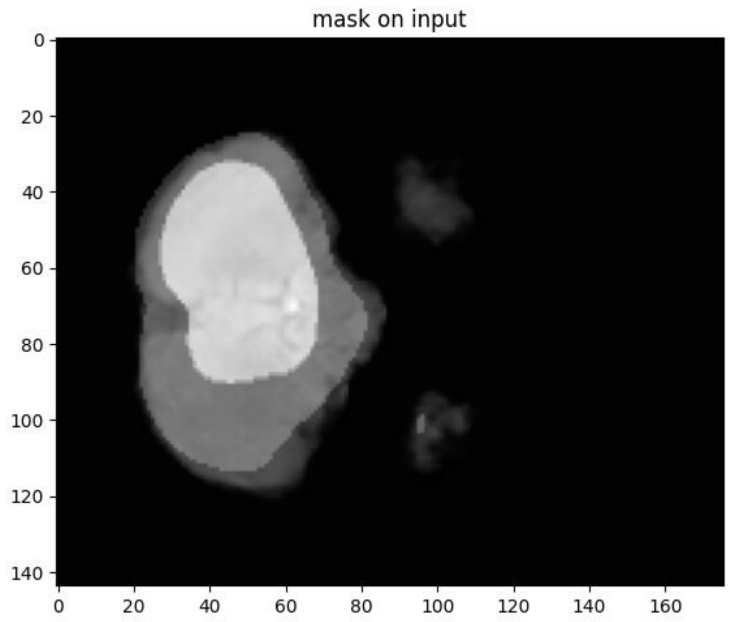
3 channels - Label 0: TC tumor subregion - Label 1: WT tumor subregion - Label 2: ET tumor subregion



This segmentation approach follows encoder-decoder based CNN architecture with an asymmetrically larger encoder to extract image features and a smaller decoder to reconstruct the segmentation mask .Also add an additional branch to the encoder endpoint to reconstruct the original image, similar to autoencoder architecture. The motivation for using the auto-encoder branch is to add additional guidance and regularization to the encoder part, since the training dataset size is limited. We follow the variational auto-encoder (VAE) approach to better cluster/group the features of the encoder endpoint.







Final Segmented Results