Brain Tumor Segmentation via Noised Multi-level Wavelet Feature Extraction

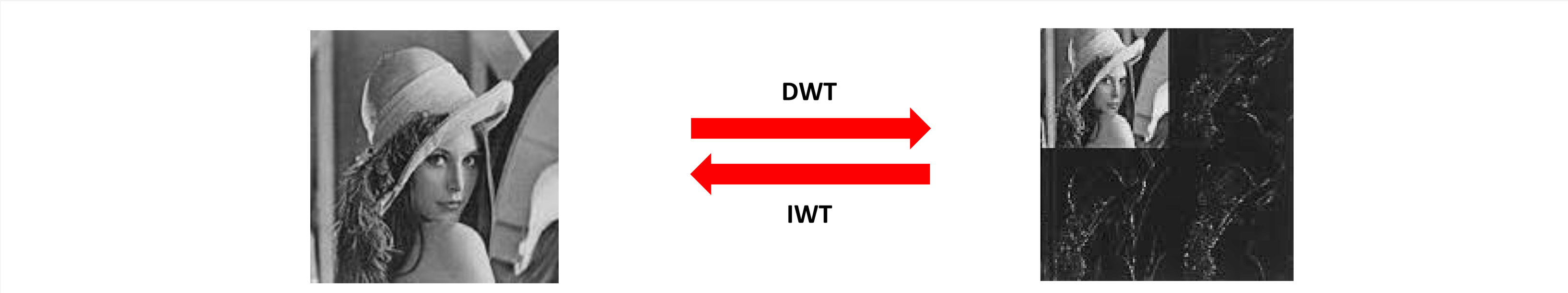
Medical Imaging & Vision Lab, @POSTECH
Daehyeon Choi, Hyuna Cho, Wonhwa Kim
{daehyeonchoi, hyunacho, wonhwa} @ postech.ac.kr

Our code is available

Summary

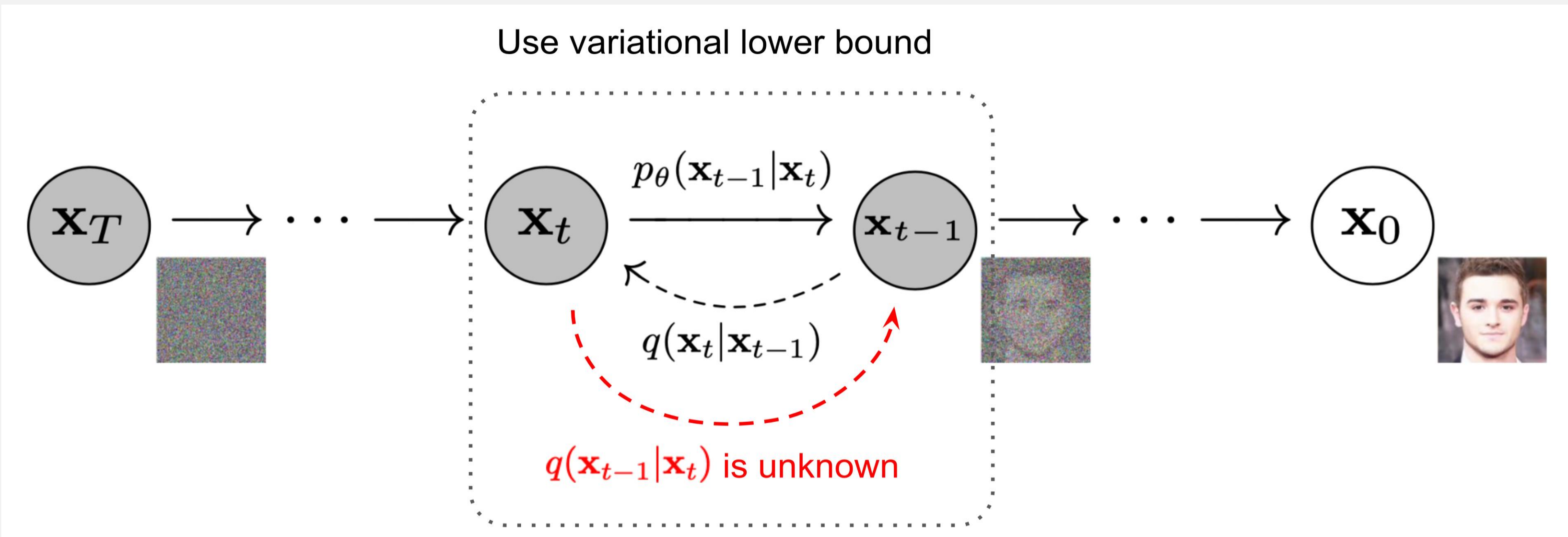
- Implemented '**WaveUNet**', that is able to extract useful features in image, with multi-level wavelet transform and various diffusion noise scales
- Proposed '**Symmetric Contrastive Loss**', simple but strong logic
- Based on AttentionUNet, added a symmetric contrastive loss and had a **better performance** for tumor cores (NCR, ET).
- Feature extractor's feature matching loss is hard to converge, need to modify the architecture or loss function in later

2D Haar Wavelet Transform



- Simple variation of wavelet transform, involving Discrete Wavelet Transform (DWT) & Discrete Inverse Wavelet Transform (IWT)
- L = $\frac{1}{\sqrt{2}}[1\ \ 1]$ & H = $\frac{1}{\sqrt{2}}[-1\ \ 1]$ represent low-pass & high-pass filters, construct 4 kernels $(\boldsymbol{LL^T}, \boldsymbol{LH^T}, \boldsymbol{HL^T}, \boldsymbol{HH^T})$
- Decompose the input $X \in R^{H\times W}$ into 4 subbands (X_{ll}, X_{lh}, X_{hl}, and X_{hh}) with dimensions $\frac{H}{2} \times \frac{W}{2}$
- Accurate reconstruction of the original signals X from frequency components through IWT

Diffusion Model

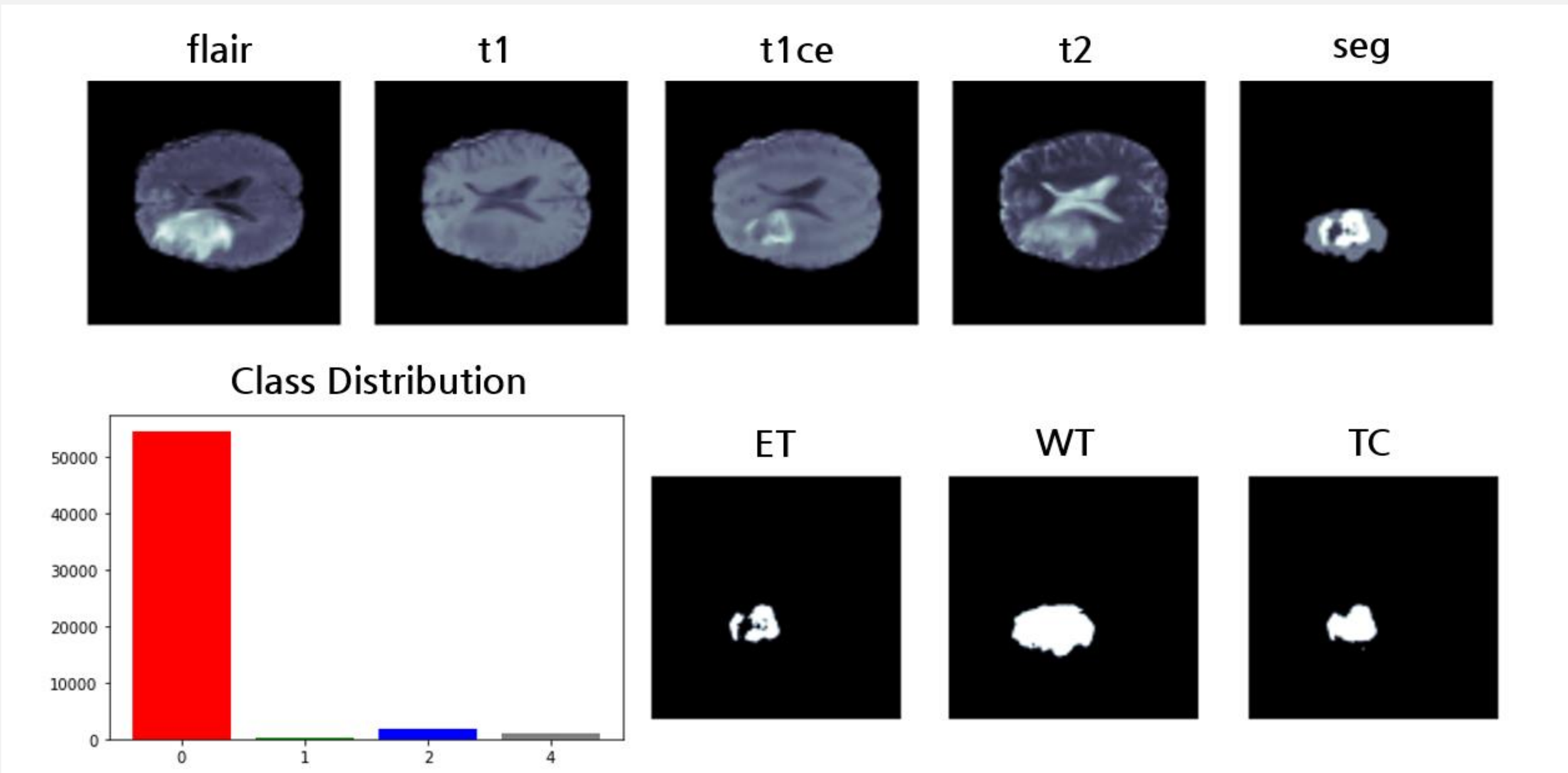


- Generative model, goes through a forward process and a reverse process of **'denoising'** to generate data with a distribution similar to the original.
- Not only synthesize high-quality images, but also be able to extract **useful feature representations**
- Hierarchical features can be extracted from various noise levels through UNet, a denoise function.

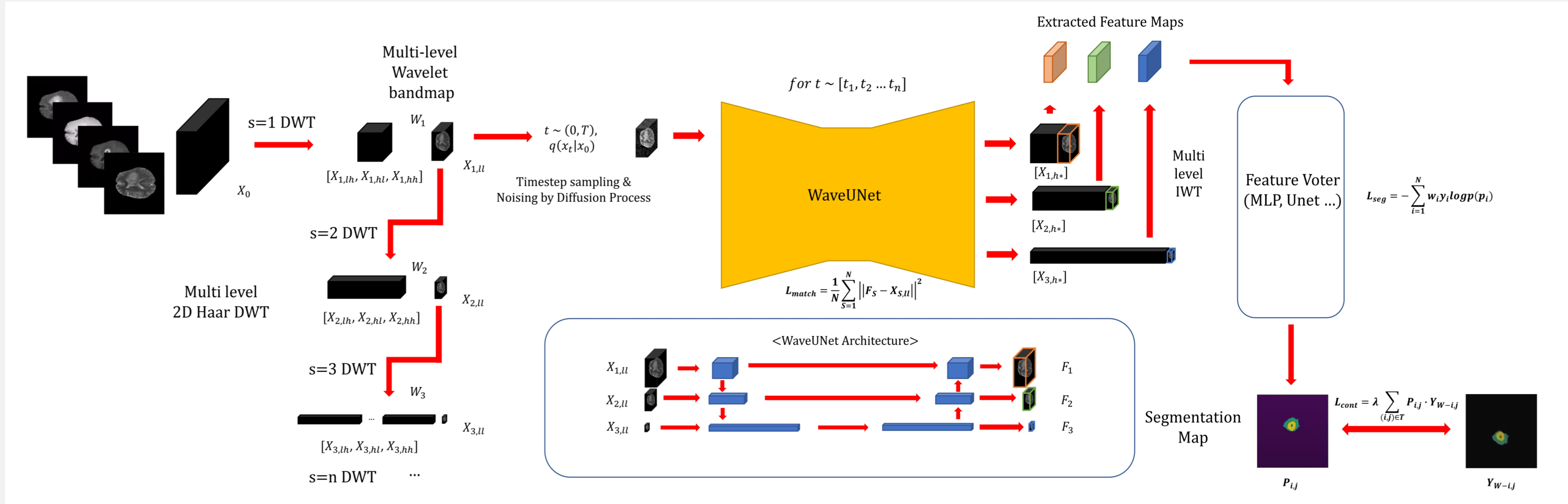
Dataset: BraTS 2021

BraTS Challenge
- Challenge in **MICCAI**
- Evaluate state-of-the-art methods for the tumor segmentation in mpMRI scans

Preprocessing
- Sliced 10 timesteps(70-79) to solve **2D Brain Tumor Segmentation task**
- Sliced data have 240×240 resolution with **4 modalities** (t1, t1ce, t2, flair), depending on whether a contrast agent is administered or not
- Each pixels are labeled one of 4 classes
  (label 0: Background, label 1: NCR, label 2: Edematous, label 4: ET)
- Preprocessed with min-max normalization to range each pixel values 0-1
- No data augmentation



Method 1: Multi-level Wavelet Feature Extractor



WaveUNet
1. Input image is decomposed by Multi-level DWT, first low-frequency subbands $X_{1,ll}$ gets noise and goes into WaveUNet
2. Higher level's low-frequency subbands are concatenated residually into WaveUNet's layers
3. WaveUNet is trained to mimics each level's low frequency subbands, returns to image domain by IWT

Feature Voter
- Can apply any general segmentation models (MLP, CNN, …)
- WaveUNet gives extracted multi-level feature map to Feature Voter, makes a per class score map

Loss Function

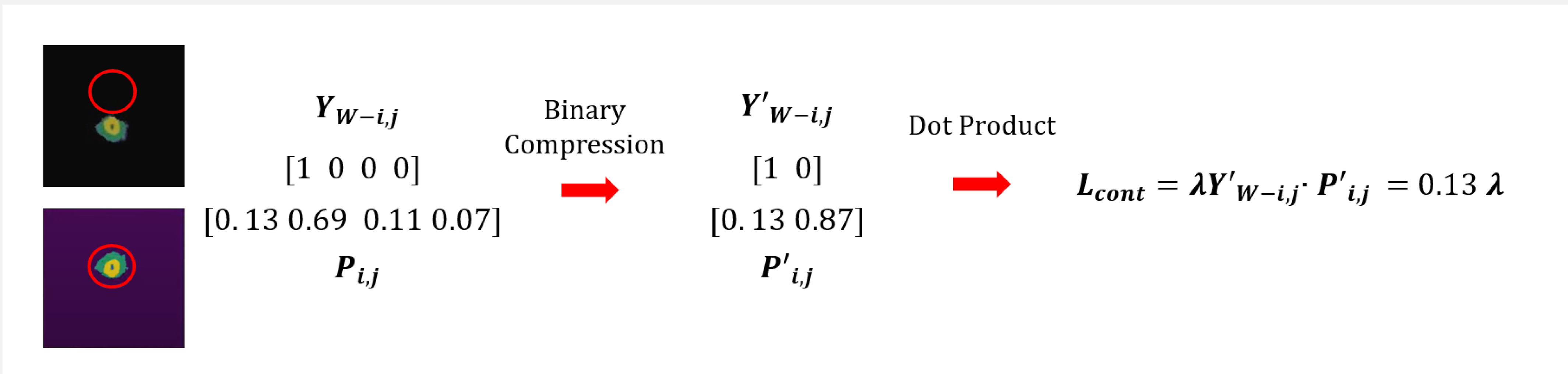
1. **Feature matching Loss:** $\boldsymbol{L_{match}} = \frac{1}{N}\sum_{S=1}^{N}\left|\left|\boldsymbol{F_S - W_{S,ll}}\right|\right|_2^2$

- Feature extractor is learned to reduce the **L2 norm** of multilevel wavelet inputs & feature maps it extracts

2. **Segmentation Loss:** $\boldsymbol{L_{seg}} = -\sum_{i=1}^{N}\boldsymbol{w_i y_i logp(y_i)}$

- Weighted multi-class cross-entropy loss
- Class distribution was highly imbalanced, weights as reciprocal of the class distribution for stable learning

Method 2: Symmetric Contrastive Loss



- Tumor regions are almost **asymmetric**
  If there is a tumor in one region based on the x-axis, there is very likely to no tumor in the opposite.
- Penalize not to be similar model's prediction to opposite region's label (Just tumor region)

Experiments & Results

Experiment 1.
- Used baseline model (AttentionUNet) learns to compare the effects by adding symmetric contrastive loss
- Each model trained for 100 epochs, hyperparameter $\boldsymbol{\lambda}$ **= 0, 0.1, 0.3, 0.5**
- Trained on 12,510 train datasets, and measure **Dice Score** for each classes with 400 validation images.

Experiment 2.
- Feature extraction with WaveUNet, give feature maps to AttentionUNet, λ=0.1 -> Not going well…

Model | Class	NCR	Edema	ET
AttentionUNet, λ=0	93.79	87.44	93.86
AttentionUNet, λ=0.1	94.29	86.51	94.27
AttentionUNet, λ=0.3	92.84	78.04	94.08
AttentionUNet, λ=0.5	-	-	-