NEURAL INFORMATION PROCESSING SYSTEMS

3D Multi Modal Semantic Segmentation Using Adversarial Learning

Shivangi Aneja, Hongwei Li, Prof. Dr. Bjoern Menze

Image Based Biomedical Modelling Lab, Technical University of Munich, Germany

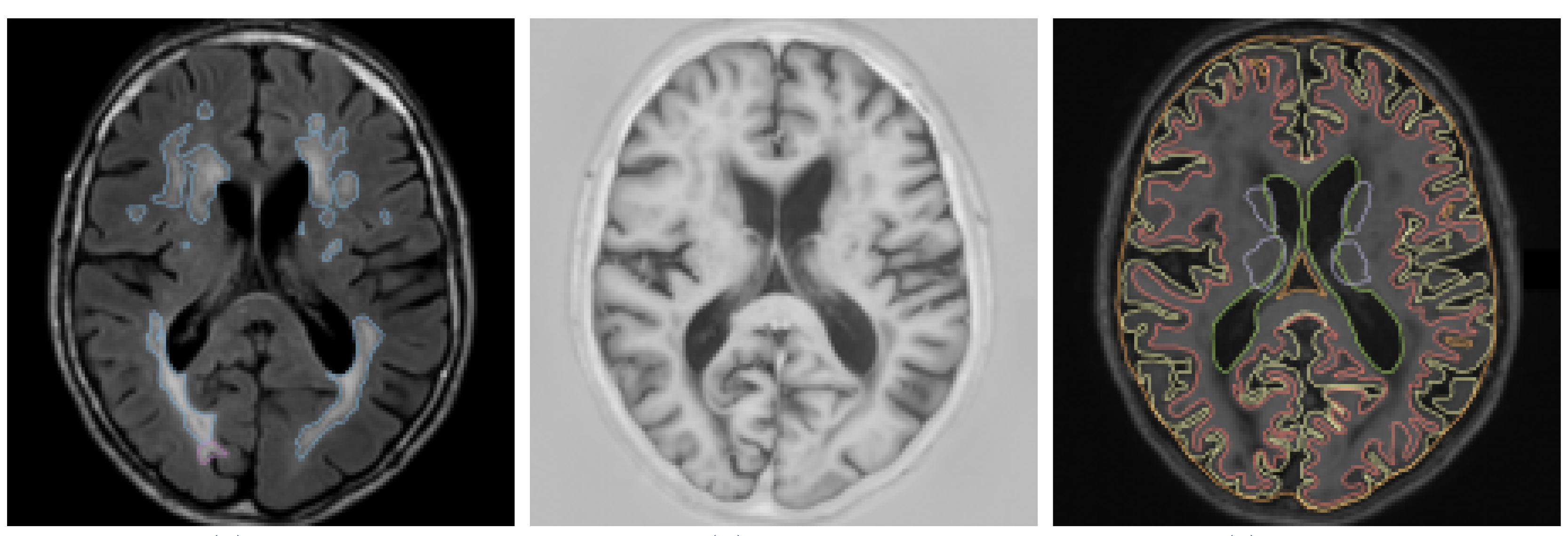
Women in Machine Learning

Introduction

Problem: Semantic segmentation of MRI scans is tedious and requires expertise.
Motivation: Inspired with success of GANs for semi-supervised segmentation in computer vision, both annotated and unannotated MRI scans are used for segmentation.
Contribution: Attempt to improve the segmentation results on 3D multi-modal medical images by making use of unannotated MRI scans.

Dataset

Our method is evaluated on problem of MRI brain segmentation task.
Annotated Dataset : Provided by [1]. Consists of 7 MRI scans (T1, T1-IR, and T2-FLAIR) with manual segmentation by experts. Include patients with diabetes, dementia and Alzheimers.



(a) T1 (b) T1-IR (c) FLAIR

Figure: Dataset Modalities

Unannotated Dataset : Provided by [2]. This dataset consists of brain MRI scans of (T1 and T2-FLAIR). So we used only two modalities (FLAIR and T1) for our task.

Mode	No. Of Subjects	Subject List
Training (Annotated)	4 subjects	[4, 5, 7, 70]
Training (Unannotated)	4 subjects	[0, 2, 4, 6]
Validaion	1 subject1	[148]
Test	2 subjects	[1, 14]

Table: Data Splitting

Segmentation Classes

Cortical gray matter, Basal ganglia,White matter, White matter lesions, Cerebrospinal fluid, Ventricles, Cerebellum, Brain stem.

Pre-processing: Removed noise and reduced variation across subjects. Each scan is of size 240 × 240 × 48. Every scan is bias field corrected using the N4ITK algorithm. Then scans are cropped such that 10 pixels from each side is removed to get rid of black background borders and Z-normalized. After pre-processing, the final dimension of each scan is 240 × 240 × 48.

Network Architecture

For baseline, 3D-UNet network was trained with cross-entropy loss. The architectural pipeline for our proposed method is shown below.

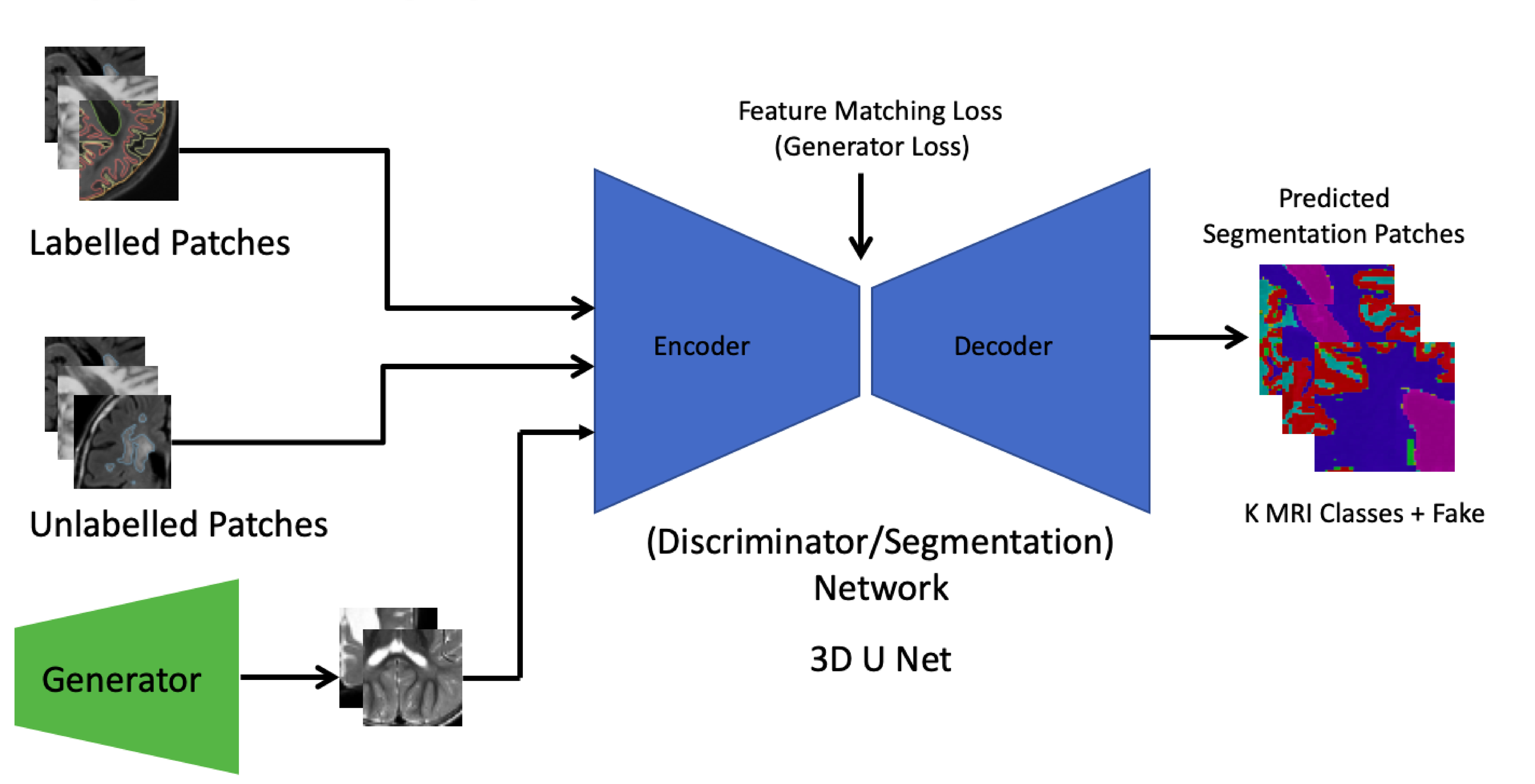


Figure: Network Architecture

Loss Functions & Metrics

To include unlabelled images during training, GAN is used. The task of the discriminator(segmentation) network is to determine whether the image patch is real (labelled or unlabelled) or fake (from generator). So, the discriminator predicts (K+1) classes, where K is the number of class labels for the real samples and the additional class corresponds to fake samples from the generator. Given a 3D image patch $x_{H \times W \times D}$, the network predicts segmentation mask $y_{H \times W \times D \times (K+1)}$.

$$L_{disc} = L_{labelled} + L_{unlabelled} + L_{fake} \tag{1}$$

$$L_{labelled} = -E_{x,y \sim p_{data}(x,y)} \sum_{i=1}^{H \times W \times D} log[p_{model}(y_i|x, y_i < K+1)] \tag{2}$$

$$L_{unlabelled} = -E_{x \sim p_{data}(x)} \sum_{i=1}^{H \times W \times D} log[1 - p_{model}(y_i = K+1|x)] \tag{3}$$

$$L_{fake} = -E_{z \sim Noise} \sum_{i=1}^{H \times W \times D} log[p_{model}(y_i = K+1|G_{\theta_G}(z))] \tag{4}$$

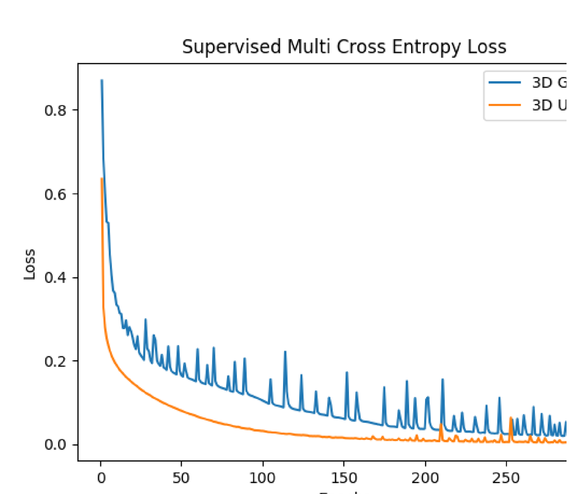
$$L_{generator} = ||E_{x \sim p_{data}(x)} Enc(x) - E_{z \sim Noise} Enc(G_{\theta_G}(z))||_2 \tag{5}$$

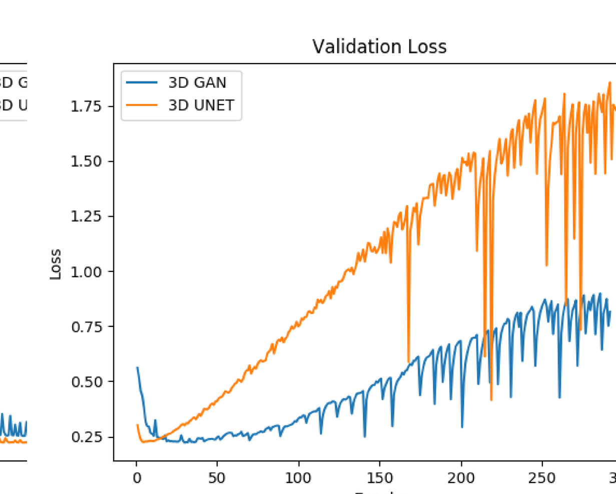
For training generator network, we just minimize fake loss L_{fake}. Given a ground-truth segmentation map G and a segmentation map P generated by an algorithm.

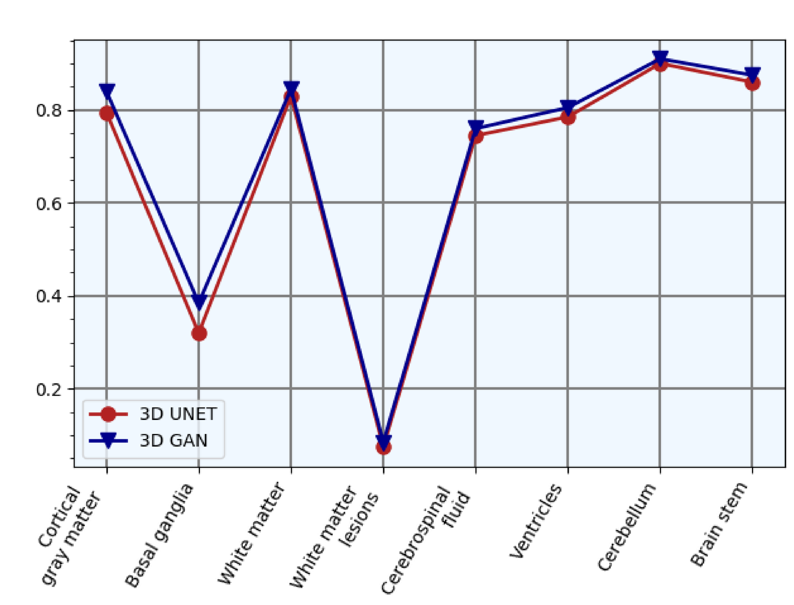
$$Dice = \frac{|G| \cap |P|}{|G| + |P|}, V_S = \frac{|V_G - V_P|}{V_G}$$

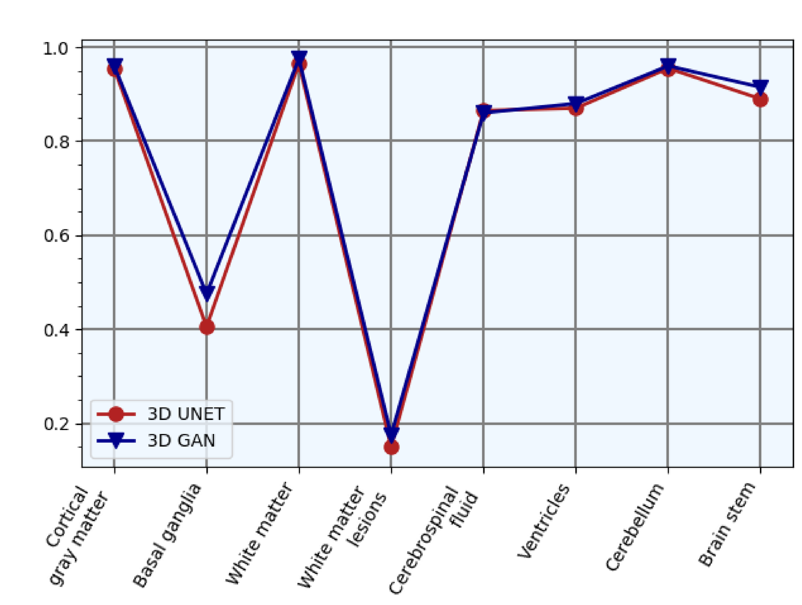
Results

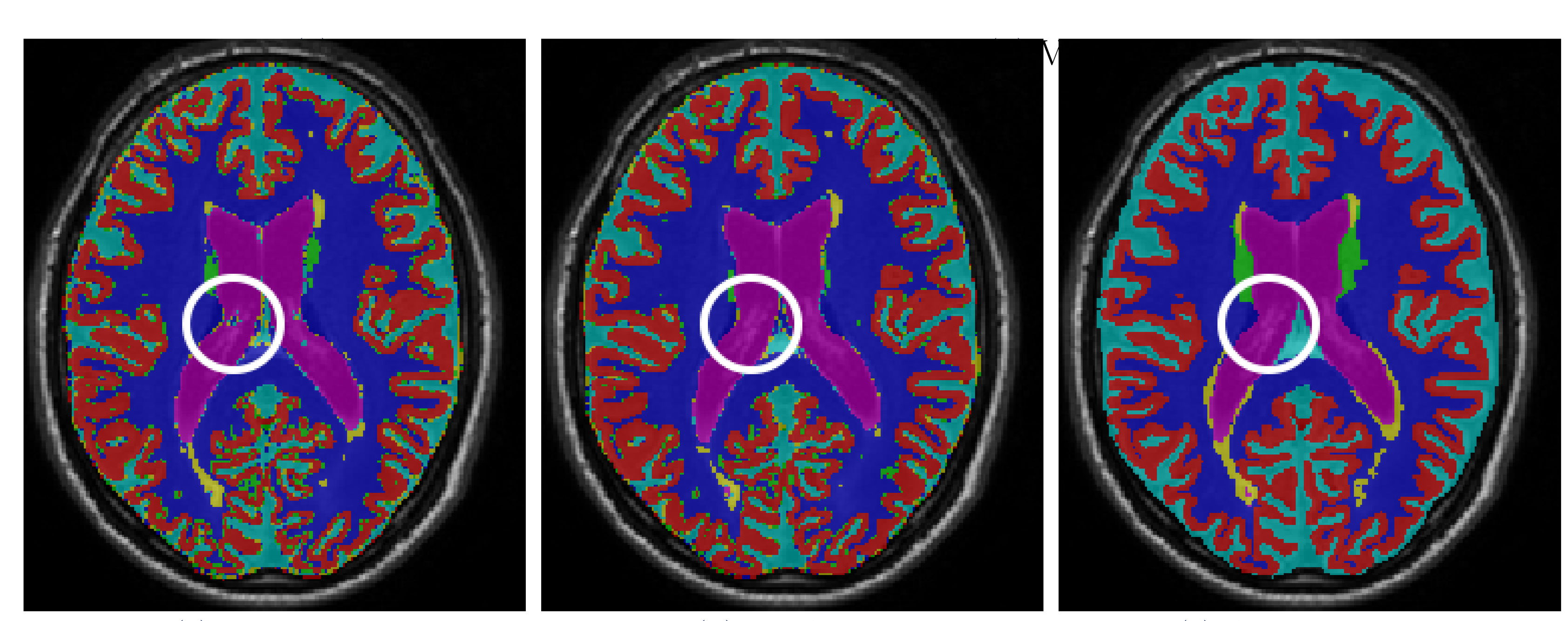
In case of 3D-Unet, due to limited training data (only 4 MRI scans), the network overfits within first few epochs. But for 3D-GAN, even though the training curve has some fluctuations but validation curve is relatively more stable.











(c) 3D U-Net (d) 3D GAN (e) Ground Truth

Figure: Qualitative Results

Conclusion

Slight improvements are observed when the model is trained with both annotated and unannotated scans in comparison to when it is only trained with annotated scans. This is due to the fact that model does not overfits and generalizes better

References

[1] Mr brains18 dataset.
Grand Challenge on MR Brain Segmentation at MICCAI 2018, 2018.

[2] White matter hyperintensities (wmh) dataset.
WMH Segmentation Challenge, 2017, 2017.