PET-Disentangler: PET Lesion Segmentation via Disentangled Healthy and Disease Feature Representations

Tanya Gatsak, Kumar Abhishek, Hanene Ben Yedder, Saeid Asgari Taghanaki, Ghassan Hamarneh

School of Computing Science, Simon Fraser University, Burnaby, Canada

tgatsak@sfu.ca

Introduction

PET imaging is an invaluable tool in clinical settings and the development of automatic lesion detection methods for PET is crucial. Many previous deep learning segmentation approaches aim to train models to better learn disease features.
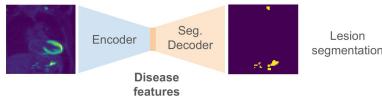
We aim to learn more robust disease features by disentangling a PET image into disease and normal healthy anatomical features.
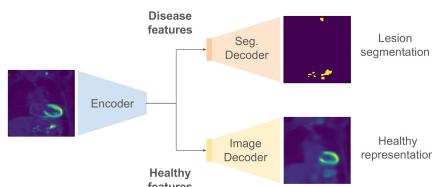

Method

PET-Disentangler for disentangling PET volumes into disease and healthy latent space features. The encoder and decoders follow a revised 3D UNet architecture, wherein the encoder is modified to split the image into healthy and disease features.

Method

The disease features are passed to a segmentation decoder to produce a segmentation prediction. An image decoder uses the segmentation prediction and healthy features to reconstruct the input image.
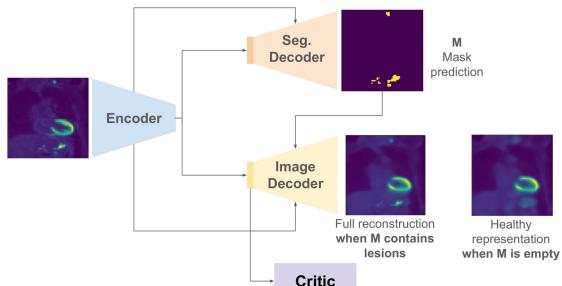

The image decoder uses SPADE blocks to combine the disentangled components and reconstruct the input volume. A critic network learns the healthy anatomy distribution by optimizing the Wasserstein distance with gradient penalty. ComboLoss is used to optimize segmentation, L_1 and L_2 losses are used to optimize reconstruction.

Quantitative Results

Segmentation accuracy on the test set reported using the Dice similarity coefficient.

	UNet	PET-Disentangler[full-skip]	PET-Disentangler[no-skip]	PET-Disentangler
Healthy	0.6005 ± 0.4067	0.4458 ± 0.4722	0.4812 ± 0.5028	0.7617 ± 0.4112
Disease	0.4828 ± 0.3196	0.6497 ± 0.2837	0.5235 ± 0.3039	0.6422 ± 0.2623

Qualitative Results
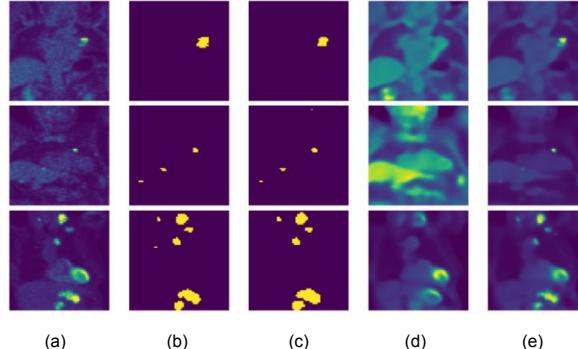
(a) Coronal slices, (b, c) Ground truth and predicted segmentation masks, respectively. (d) Healthy component. (e) Reconstruction using healthy and disease components.

Conclusion

We presented a novel PET disentanglement method that improves upon 3D UNet segmentation baseline results. Our findings highlight the effectiveness of learning disentangled representations and incorporating healthy anatomy modeling.