Breast Volume and Fibroglandular Tissue Segmentation in MRI using a Deep Learning Unet

Gregory Kuling[1], Homa Fashandi[2], YingLi Lu[2], Hongbo Wu[1], and Anne L. Martel[1,2]
[1]Department of Medical Biophysics, University of Toronto, Toronto, Ontario M4N 3M5, Canada
[2]Physical Sciences, Sunnybrook Research Institute, Toronto, Ontario M4N 3M5, Canada

Purpose: The amount of fibroglandular tissue (FGT) relative to breast volume has been shown to have significant correlation with personalized risk of incidence for breast cancer [Dontchos]. Manual analysis of three-dimensional MR scans is a time consuming and subjective task making it difficult to retrieve consistent measurements. Since manual segmentation is often variable, we present here a fully automatic method for the segmentation of FGT and breast volume using a deep learning model called a Unet [Ronneberger]. This convolutional network architecture has been used for fast and precise image segmentation in many tasks, and with breast MRI it has been shown to efficiently segment anatomy in T1w images without fat-suppression (WOFS) [Dalmış]. Our goal is determining whether combining both fat suppressed, and non-fat suppressed images improves the accuracy of segmentation.

Method: We compared 2D Unets trained on T1w WOFS, T1w with fat suppression (FS), and a combination of both scan types. Our dataset has 98 patient scans acquired with a GE 1.5T scanner with an average resolution of [0.388, 0.388, 3.0] mm. T1w WOFS and T1w FS images were co-registered to correct for any motion and were resampled to an isotropic voxel size (2.0mm).

This task has 3 classes to label: background, fat tissue (FT) and fibroglandular tissue (FGT). To segment the breast, we manually segmented the breast tissue from the breast-air boundary and the chest wall-breast boundary using ITK-SNAP. Within the breast mask we used a combination of thresholding techniques, and k-Means clustering algorithms with the FS and WOFS data to determine the FGT ground truth. Specifically, taking the overlap of the Otsu threshold of WOFS and FS images independently, taking the Otsu threshold of the difference between MR signals of FS and WOFS, using the Niblack threshold to correct intensity inhomogeneity before taking the Otsu threshold of the corrected image, and a 2-dimensional k-means cluster of FS and WOFS images. The four FGT ground truth methods were combined using the STAPLE algorithm and then manually edited for accuracy.

We adopted a classical machine learning experiment protocol using independent train/validation/test sets, hyperparameter grid searching, and then evaluation of results on the test set (Table 1). The loss function was defined as the sum of Dice Similarity Coefficients (DSC) of the three classes [Zou].

Table 1: Details on the Unet parameters and experimental set-up.

Validation Scheme

Data Set	# of Scans
Train	68 (33% used for validation in training
HP Validation	15
Test	15

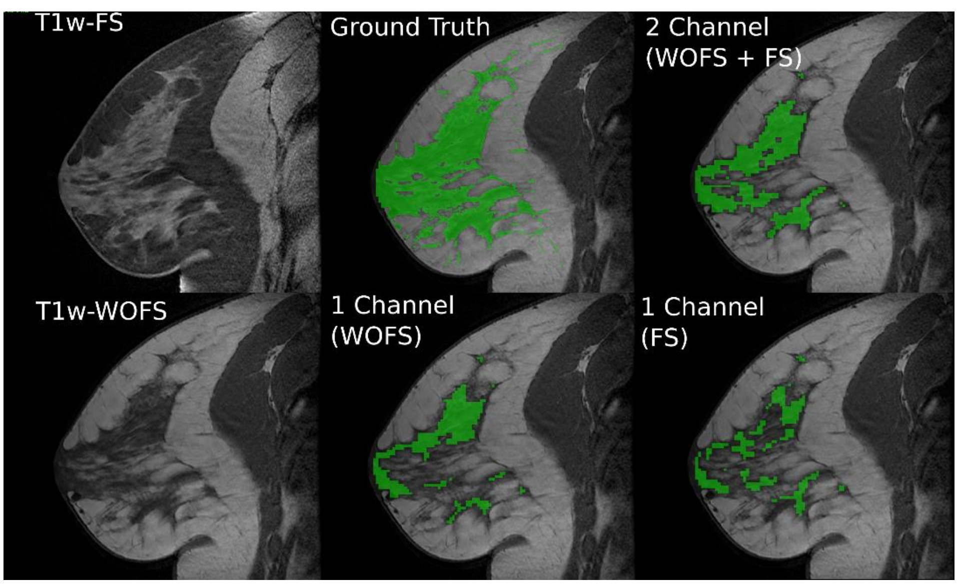
perparameter Grid Search

Hyperparameter	Values	Description
Unet Depth	[3,4,5]	Amount of Convolutional layers. Increased depth results in more complex segmentation.
Dilation rate	[1,2]	The spacing of convolution filter pixels. Changes the spatial recognition of the network.
Dropout rate	[0,0.2,0.5]	Randomly dropping input units at each update during training time. This helps the network generalize.
Optimizer	[sgd, adam]	Stochastic Gradient Descent (SGD) with Neserov momentum, adaptive moment (ADAM) estimation optimizer

Table 2: Results for the varying inputs on different tissue segmentations.

Input	Fat Tissue	Fibroglandular Tissue
2 Channel (WOFS + FS)	0.95±0.09	0.85±0.22
1 Channel WOFS	0.94±0.11	0.82±0.23
1 Channel FS	0.89±0.17	0.78±0.28

Results and Discussion: To evaluate the accuracy of our segmentation models, we used the DSC, which is a measure of segmentation. Our top performing models are shown in Table 2. The figure below shows thresholded predictions > 0.5 generated by our Unet models alongside FS and WOFS images that are co-registered. The DSC values and visual inspection of the results suggest that a combination of WOFS and FS images improves accuracy. Clearly down-sampling of image resolution from 512x512 voxel slices to 128x128 voxel slices significantly hinders the network's capability to segment smaller portions of FGT, making it difficult to distinguish between the finer segmentations that are present in the ground truth images. Future work will investigate applying this model on higher resolution images to capture more details in each scan. Furthermore, we will be working towards using these new FGT masks to segment out background parenchymal enhancement (BPE), another strong indicator for individual risk of developing breast cancer.



Conclusion: We have presented preliminary segmentation results using a deep Unet model to segment FGT in breast MRI. We have shown that the best results are achieved using a 2-channel combination of T1w MR images WOFS and with FS.

References:
Ronneberger, et al. (2015, October). In *International Conference on MICCAI* (pp. 234-241). Springer, Cham.
Dontchos, B. N., et al. (2015) *Radiology*, *276*(2), 371-380.
Dalmış, M. U., et al. (2017). *Medical physics*, *44*(2), 533-546.
Zou, K. H., et al. (2004). *Academic radiology*, *11*(2), 178-189.