Deep Convolutional Encoder Networks for Multiple Sclerosis Lesion Segmentation

Tom Brosch[1,4], Youngjin Yoo[1,4], Lisa Y.W. Tang[4], David K.B. Li[2,4], Anthony Traboulsee[3,4], and Roger Tam[2,4]

[1]Department of Electrical and Computer Engineering, UBC [2]Department of Radiology, UBC
[3]Division of Neurology, UBC [4]MS/MRI Research Group, University of British Columbia, Vancouver, Canada

Introduction

- Multiple sclerosis (MS) is an inflammatory and demyelinating disease of the central nervous system, and is characterized by the formation of lesions, primarily visible in the white matter on conventional magnetic resonance images (MRIs).
- Imaging biomarkers based on the delineation of lesions, such as lesion load and lesion count, have established their importance for assessing disease progression and treatment effect.
- We propose a new method for segmenting MS lesions that processes entire MRI volumes through a neural network to automatically learn features tuned for lesion segmentation.

Contributions

- Our network processes entire volumes instead of patches, which removes the need to select representative patches, eliminates redundant calculations where patches overlap, and therefore scales up more efficiently with image resolution.
- Our approach combines feature learning and segmentation in a single model, which allows for the automatic learning of features that are tuned towards lesion segmentation.
- We propose a new objective function based on a weighted combination of sensitivity and specificity, designed to deal with unbalanced classes, as is the case for lesions, which typically comprise less than $1\,\%$ of the image voxels.

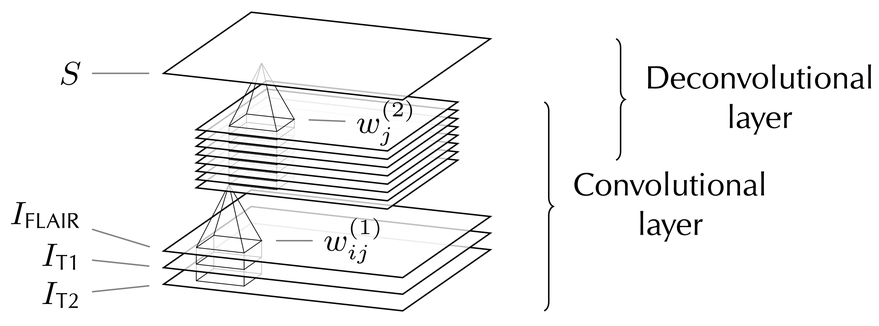
Method

- The task of segmenting MS lesions is defined as finding a function s that maps multi-modal images I, e.g., $I = (I_{\text{FLAIR}}, I_{\text{T1}}, I_{\text{T2}})$, to corresponding lesion masks S.
- Given a training set (I_n, S_n), finding s is modeled as an optimization problem of the following form

$$\hat{s} = \arg\min_{s \in \mathcal{S}} \sum_n E(S_n, s(I_n)) \tag{1}$$

where $\mathcal{S}$ is the set of possible segmentation functions, and E is an error measure.

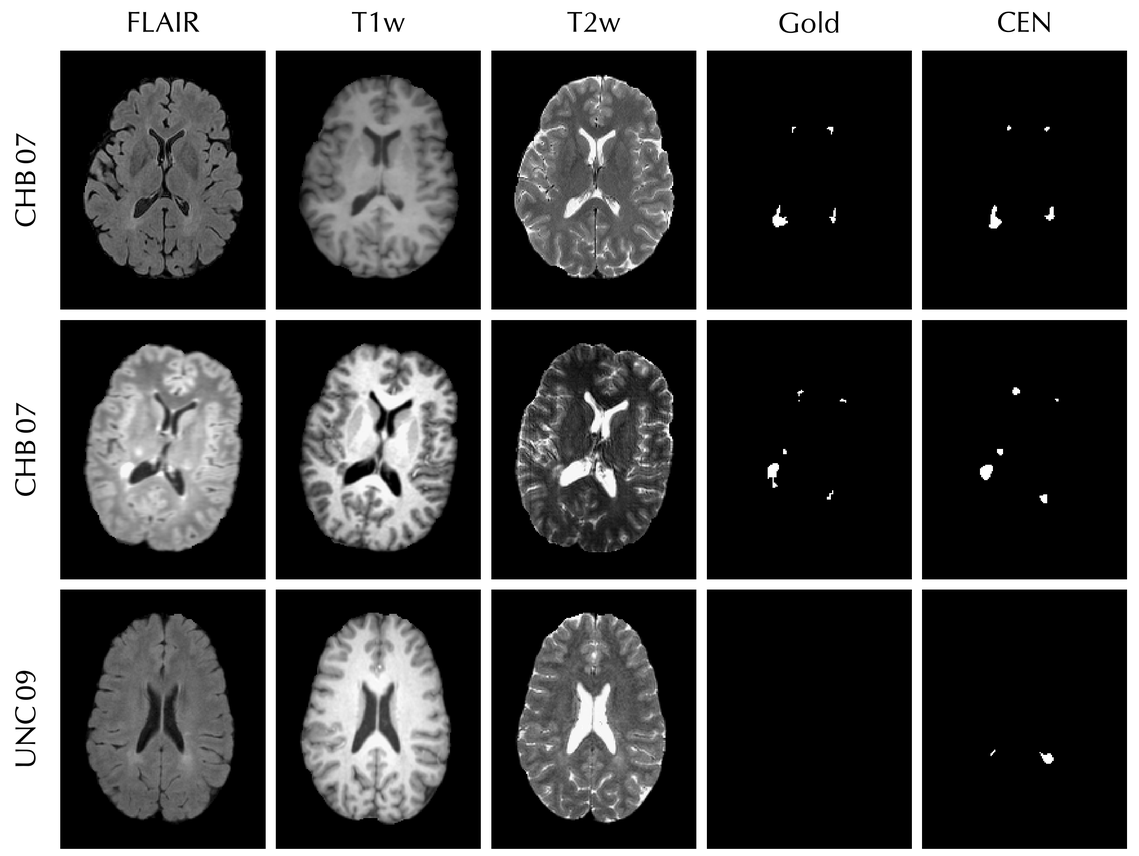
- The set of possible segmentation functions, $\mathcal{S}$, is modeled by the following 3-layer convolutional encoder network:



- The input layer is composed of the image voxels of different modalities.
- The convolutional layer extracts features from the input layer at each voxel location.
- The deconvolutional layer uses the extracted features to predict a lesion mask and thereby classify each voxel of the image in a single operation.
- The error measure, E, is a weighted sum of the mean squared difference of the lesion voxels (sensitivity) and non-lesion voxels (specificity), reformulated to be error terms.

$$E = r\frac{\sum_{\vec{p}} \left(S(\vec{p}) - y^{(2)}(\vec{p})\right)^2 S(\vec{p})}{\sum_{\vec{p}} S(\vec{p})} + (1-r)\frac{\sum_{\vec{p}} \left(S(\vec{p}) - y^{(2)}(\vec{p})\right)^2 \left(1 - S(\vec{p})\right)}{\sum_{\vec{p}} \left(1 - S(\vec{p})\right)} \tag{2}$$

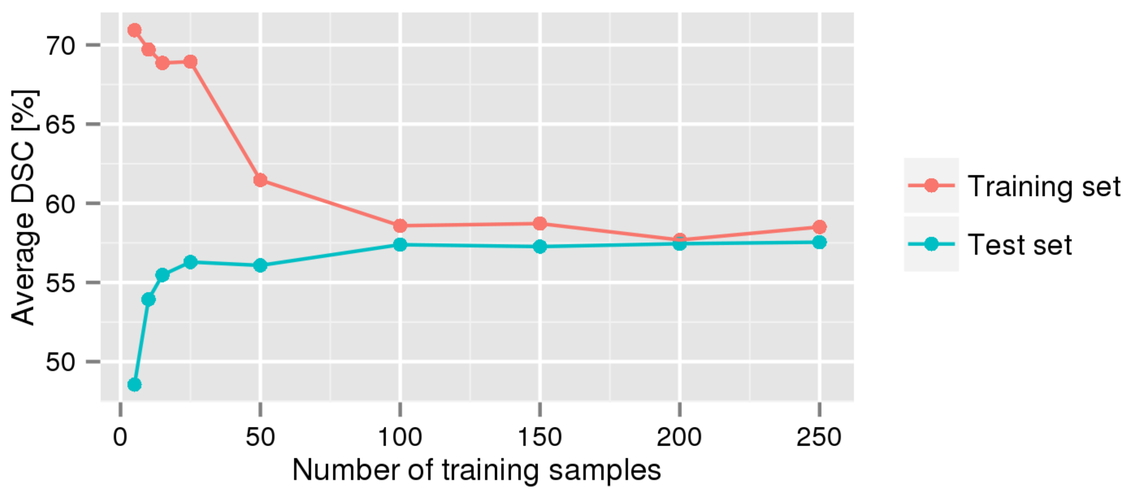
Evaluation



CHB 07 (DSC = $60.58\,\%$), CHB 07 (DSC = $60.58\,\%$), UNC 09 (DSC = $9.01\,\%$)

Method	TPR	PPV	DSC
Souplet et al. [?]	20.65	30.00	—
Weiss et al. [?]	33.00	36.85	29.05
Geremia et al. [?]	39.85	40.35	—
Our method	39.71	41.38	35.52

- BioMS results
- 100 samples need to prevent overfitting.



Conclusions

- Our model performs on par with the state-of-the-art on a small data set
- Performs can be much improved when a suitable training set is available

Acknowledgement

This work was supported by the Natural Sciences and Engineering Research Council of Canada and the Milan and Maureen Ilich Foundation.

Contact Information

✉ tombr@msmri.medicine.ubc.ca
http://tbrosch.blogspot.com/