DEEP LEARNING FOR MULTI-SITE MS LESIONS SEGMENTATION: TWO-STEP INTENSITY STANDARDIZATION AND GENERALIZED LOSS FUNCTION

Francesca Galassi, Solène Tarride, Emmanuel Vallée, Olivier Commowick, Christian Barillot

DEEP LEARNING FOR MULTI-SITE MS LESIONS SEGMENTATION: TWO-STEP INTENSITY STANDARDIZATION AND GENERALIZED LOSS FUNCTION

Francesca Galassi⋆*, Solène Tarride*⋆*, Emmanuel Vallée* †*, Olivier Commowick*⋆*, Christian Barillot*⋆

⋆ INRIA, CNRS, INSERM, IRISA, Empenn, Rennes, France
† Orange Labs, Lannion, France

ABSTRACT

We present an improved CNN framework for the segmentation of Multiple Sclerosis (MS) lesions from multimodal MRI. It uses a two-step intensity normalization and a cascaded network with cost sensitive learning. Performance was assessed on a public multi-site dataset.

Index Terms— MRI, MS lesions, Segmentation, CNN

1. INTRODUCTION

Performance of deep learning based classifiers in MS lesions segmentation is affected by intensity variability of MRI data as well as by high class imbalance. This work aims to mitigate the issue of intensity variability by using a two-step normalization and class imbalance by cost sensitive learning of a CNN cascaded network.

2. MATERIAL AND METHODS

We trained and evaluated the framework on the multi-site MICCAI 2016 MS lesion segmentation challenge dataset [1]. It includes 3D T1-w and 3D FLAIR images of 53 MS patients from four MR sites (15 train and 38 test datasets). One site was not included in the train dataset.

MR images were denoised, rigidly registered towards T1-w, skull-stripped and bias corrected. We acquired 20 MRI datasets of healthy subjects and generated an atlas representing the average intensity and shape of the controls per modality. We normalized each modality of each MS patient using the corresponding modality template image and k-means standardization. Finally, we normalized each image independently by subtracting its mean and dividing by its standard deviation.

We trained a two-CNN 3D patch-wise cascaded network [2] using a Generalized Dice Loss (GDL) [3], rather than the standard cross-entropy, where the contribution of a label was corrected by the inverse of its volume.

3. RESULTS

Lesion load varied from $\sim 0.5cm^3$ to $\sim 70cm^3$. Average performance scores are reported in Table 1, with/without the use of intensity normalization (IN) and/or GDL. The paired Wilcoxon-test for the F1-score ([2] + IN + GDL) vs [2] returned a p-value equal to 6.16e-07.

Table 1. Average scores on all test set and unseen site only

	Dice-score	F1-score	F1-score (unseen site)
[2] @ MICCAI16	0.541	0.490	0.280
[2] + IN	0.583	0.514	0.516
[2] + GDL	0.602	0.548	0.534
[2] + IN + GDL	0.624	0.604	0.614

4. DISCUSSION AND CONCLUSIONS

Network performance improved with the use of the two-step intensity normalization and Generalized Dice Loss. Top ranked methods at the MICCAI 2016 challenge obtained : (DSC=0.541, F1-score=0.490) and (DSC=0.591, F1-score=0.386) [1]. The framework outperformed all challengers, with higher precision and recall (DSC=0.624, F1-score=0.604). Scores were consistent across all sites. A marked improvement over [2] was achieved on the site that was not included in the train dataset.

5. REFERENCES

[1] O. Commowick., A. Istace., M. Kain et al., Objective Evaluation of Multiple Sclerosis Lesion Segmentation using a Data Management and Processing Infrastructure. Scientific reports, vol. 8, no. 1, Sept 2018.

[2] S. Valverde, M. Cabezas, E. Roura. et al., Improving automated multiple sclerosis lesion segmentation with a cascaded 3D convolutional neural network approach, Neuroimage, vol. 15, no. 155, pp. 159-168, Apr 2017.

[3] C.H. Sudre, W. Li, T. Vercauteren et al., Generalised Dice Overlap as a Deep Learning Loss Function for Highly Unbalanced Segmentations, DLMIA, ML-CDS, Lecture Notes in Computer Science, vol. 10553, Springer, 2017.