New MS lesion Segmentation with Lesion-wise Metrics Learning

Reda Abdellah Kamraoui[1], Vinh-Thong Ta[1], José V Manjon[3], and Pierrick Coupé[1]

[1] Univ. Bordeaux, Bordeaux INP, CNRS, LaBRI, UMR5800, PICTURA, F-33400 Talence, France
[2] ITACA, Universitat Politècnica de València, 46022 Valencia, Spain

Abstract. Recently, the performance of automatic Multiple Sclerosis segmentation methods has been evaluated using lesion-wise detection metrics (*e.g.,* lesion-wise true positive rate). However, due to the complexity of implementing a differentiable loss function based on such metrics, most deep learning methods are trained with traditional voxel-wise metrics such as the Dice similarity coefficient. In this paper, to address this issue, we propose to use a convolutional neural network to estimate lesion-wise metrics from ground truth and predicted masks. Combined with the Dice Similarity Coefficient, this lesion-wise metric estimation model is used to optimize our new lesions segmentation method.

Keywords: Metric Learning · New lesion segmentation · Detection Metrics.

1 Introduction

The detection of new lesions is an important bio-marker in Multiple Sclerosis (MS) that allows clinicians to adapt the patient treatment and assess the evolution of the disease. Automating new lesion detection can alleviate the clinicians workflow. To assess such automatic methods and compare them to expert segmentation, the research community relies on several segmentation metrics. Recent works [3] question the relevance of voxel-wise metrics (such as Dice) compared to detection metrics, which are used for MS diagnostic and clinical evaluation of the patient evolution. Besides, other works [6] suggest that multiple complementary metrics are needed to provide a better understanding of the automatic method performance. Deep learning methods have shown encouraging results in the task of MS segmentation [1, 2]. Many segmentation methods are optimized by maximizing the Dice between the prediction and ground truth. To the best of our knowledge, no prior work has successfully implemented a deep learning loss function based on lesion-wise metrics. In this paper, we propose a novel metric learning framework designed for lesion-wise metric optimization.

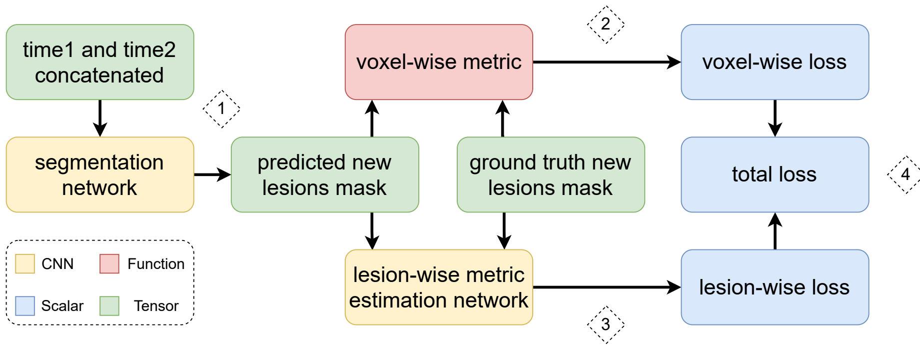

Fig. 1. The training framework

2 Method and Material

2.1 Method overview

The proposed method uses a Convolutional Neural Network (CNN) to estimate lesion-wise metrics. The CNN is used in addition to classic voxel-wise metrics to train a segmentation model. As shown in Fig. 1, the training of our new MS lesion segmentation network is performed with a 4 step cycle. First, the segmentation network is fed with a concatenation of 3D FLAIR patches relative to the two time points. As an output, the network predicts the new lesions mask. Second, a voxel-wise loss is computed from the predicted and the ground-truth new lesions masks. This segmentation error is estimated using the Dice similarity coefficient:

$$VoxelWise_{loss} = 1 - Dice(pred, true) \tag{1}$$

Third, the lesion-wise metric estimation network produce a lesion-wise loss term from the predicted and the ground-truth masks. The training of the metric estimation network is performed separately (see lesion-wise metric estimation network). Indeed, the weights of this network are frozen during the training of the segmentation network.

$$LesionWise_{loss} = 1 - F(pred, true) \tag{2}$$

where F(pred,true) is the metric estimation for the prediction and ground-truth masks.

Fourth, both voxel-wise and lesion-wise loss terms are combined into a total loss:

$$Total_{loss} = VoxelWise_{loss} + \alpha \times LesionWise_{loss} \tag{3}$$

This aggregated term is back propagated through the metric estimation network and back to the segmentation network. The gradient estimated at the level of the segmentation model can be used to update its weights using an optimization method (e.g. gradient descent).

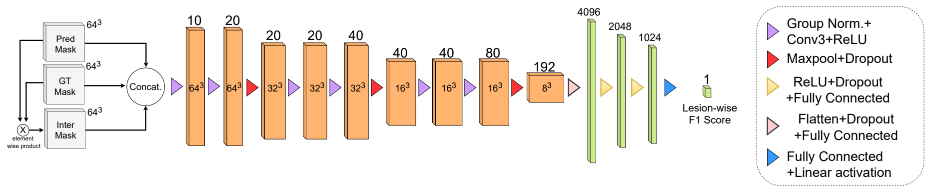

Fig. 2. The lesion-wise metric estimation block

2.2 Lesion-wise metric estimation network

The proposed network is a CNN that predicts an approximation of lesion-wise metrics from probability masks. As shown in the Fig. 2, the network is composed of convolutional layers and fully connected layers. This model is trained with mask pairs (*i.e.* representing the predicted mask and the ground-truth mask) and their respective lesion-wise metric. For better network optimization and metric estimation, the mask pairs should be selected to produce a balanced distribution of output values in the metric defined range (*i.e.* in our case, a balanced distribution of values in the range [0,1]). It is worth mentioning that the masks are augmented to simulate continuous probability maps. Specifically, masks are altered with gaussian blur and small random intensity changes. For this work, the estimation network predicts the lesion-wise F1 score (defined by the challenge organizers [9]) as the following:

$$LF_1 = \frac{2 \times LTPR \times LPPV}{LTPR + LPPV} \tag{4}$$

where LTPR (lesion-wise true positive rate) is the rate of ground truth lesions that intersect with predicted lesions and LPPV (lesion-wise positive predictive value) is the rate of predicted lesions that intersect with ground truth lesions. The mean squared error is used to train this estimation network.

2.3 Data

The dataset provided by the MICCAI 2021 - Longitudinal Multiple Sclerosis Lesion Segmentation Challenge [4] was used to train our method. For preprocessing, our strategy used the docker [5] built with the Anima scripts[3] proposed by the challenge organizers. It includes bias correction, denoising and skull stripping.

3 Implementation details

For the new lesion segmentation, a 3D U-Net architecture similar to the one proposed by [7] has been selected. As input, the network receives a concatenation of FLAIR patches of size 64^3 from the two times points. The model output

[3] anima.irisa.fr

predicts the new lesion mask. Image quality data augmentation [7] is used when training the new lesion segmentation model. The models are optimized with Adam [8] using a learning rate of 0.0001 and a momentum of 0.9. We found empirically that $\alpha = 0.2$ is a good tradeoff between Dice and the estimation of LF_1.

References

1. Carass, A., Roy, S., Jog, A., Cuzzocreo, J.L., Magrath, E., Gherman, A., Button, J., Nguyen, J., Prados, F., Sudre, C.H., et al.: Longitudinal multiple sclerosis lesion segmentation: resource and challenge. NeuroImage **148**, 77–102 (2017)
2. Commowick, O., Cervenansky, F., Ameli, R.: Msseg challenge proceedings: Multiple sclerosis lesions segmentation challenge using a data management and processing infrastructure (2016)
3. Commowick, O., Istace, A., Kain, M., Laurent, B., Leray, F., Simon, M., Pop, S.C., Girard, P., Ameli, R., Ferré, J.C., et al.: Objective evaluation of multiple sclerosis lesion segmentation using a data management and processing infrastructure. Scientific reports **8**(1), 1–17 (2018)
4. the challenge dataset: https://portal.fli-iam.irisa.fr/msseg-2/data/
5. the pre-processing docker: https://github.com/Inria-Empenn/lesion-segmentation-challenge-miccai21/
6. García-Lorenzo, D., Francis, S., Narayanan, S., Arnold, D.L., Collins, D.L.: Review of automatic segmentation methods of multiple sclerosis white matter lesions on conventional magnetic resonance imaging. Medical image analysis **17**(1), 1–18 (2013)
7. Kamraoui, R.A., Ta, V.T., Tourdias, T., Mansencal, B., Manjon, J.V., Coupé, P.: Towards broader generalization of deep learning methods for multiple sclerosis lesion segmentation. arXiv preprint arXiv:2012.07950 (2020)
8. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014)
9. the evaluation criteria for the MSSEG-2 challenge: https://portal.fli-iam.irisa.fr/files/2021/06/MS_Challenge_Evaluation_Challengers.pdf