



Lesion synthesis for extending MRI training datasets and improving automatic multiple sclerosis lesion segmentation

, Sergi Valverde, Mariano Cabezas, Deborah Pareto, Arnau Oliver, Joaquim Salvi, Àlex Rovira, Xavier Lladó, Mostafa Salem

Abstract:

Background: Image synthesis is gaining attention in many domains, including medical imaging. For instance, the generation of synthetic lesions can be used as a solution to the lack of large datasets with multiple sclerosis (MS) lesions manually annotated, which is one of the main limitations to train robust and generalisable supervised machine learning algorithms. **Objectives:** To propose a fully convolutional neural network (CNN) model for MS lesion synthesis in magnetic resonance images. **Materials and methods:** T2-FLAIR and T1-w images from a dataset of 65 patients with a clinically isolated syndrome or early relapsing-remitting MS were used to train a CNN able to synthesise new lesions. The inputs of the CNN were processed images without lesions, while the original images with lesions were the outputs of the CNN. To obtain the input images, we computed a white matter hyperintensity (WMH) mask along with several intensity level masks that encoded the intensity profiles of the WMH voxels. Then, the WMH mask was filled with intensities resembling white matter. The CNN architecture to perform the image synthesis consisted of two encoders (one per each modality) that learned the latent representation for the input modalities, and two decoders, that allowed to generate new lesions in both modalities. To evaluate the synthesis, we tested a state-of-the-art MS lesion segmentation approach (Valverde et al. 2017) on an in-house dataset and the public ISBI2015 challenge dataset, showing the performance in different scenarios such as using synthetic images for data augmentation. **Results:** For the in-house dataset, when adding to a single original image several synthetic ones, the performance of the lesion segmentation increased the sensitivity from 41% to 50% and the positive predictive value (PPV) from 53% to 65%. Repeating the experiment on the ISBI2015 dataset, the sensitivity increased from 44% to 51% and the PPV from 76% to 78%. With the inclusion of few original images and the synthetic data, we were able to increase the detection performance to that of the segmentation algorithm fully trained using the entire available training set, yielding a comparable human expert rater performance. **Conclusions:** The proposed CNN was able to generate T1-w and T2-FLAIR images with synthetic MS lesions. The combination of original images with synthetic ones of the same domain increased the lesion segmentation accuracy, reducing also the number of manually annotated images of the database. **Disclosure:** M. Salem: nothing to disclose. S. Valverde: nothing to disclose. M. Cabezas: nothing to disclose. D. Pareto: has received speaking honoraria from Novartis and Biogen. A. Oliver: nothing to disclose. J. Salvi: nothing to disclose. A. Rovira serves on scientific advisory boards for Novartis, Sanofi-Genzyme, Icometrix, SyntheticMR, Bayer, Biogen and OLEA Medical, and has received speaker honoraria from Bayer, Sanofi-Genzyme, Bracco, Merck-Serono, Teva Pharmaceutical Industries Ltd, Novartis, Roche and Biogen. X. Lladó: nothing to disclose.

Published In:

Multiple Sclerosis Journal -ECTRIMS (JCR CN IF:5.649 Q1(23/199)), Stockholm. Sweden , Vol. 25 , pp. 463-463