P1143
A quantitative analysis of deep learning methods for multiple sclerosis white matter lesion segmentation

A. Clèrigues[1], S. Valverde[1], J. Bernal[1], D. Pareto[2], J.C. Vilanova[3], L. Ramió-Torrentà[4], À. Rovira[2], A. Oliver[1], X. Llado[1]
[1]University of Girona. Research institute of Computer Vision and Robotics, Girona, [2]Magnetic Resonance Unit, Dept. of Radiology, Vall d'Hebron University Hospital, Barcelona, [3]Girona Magnetic Resonance Center, [4]Neuroimmunology and Multiple Sclerosis Unit, Neurology Department, Dr. Josep Trueta University Hospital, Girona; Neurodegenerative and Neuroinflammation Research Group, IDIBGI, Girona; University of Girona, Girona, Spain

Background: Several multiple sclerosis (MS) lesion segmentation methods based on convolutional neural networks (CNN) have been proposed recently. Contrary to previous supervised learning methods, CNNs do not require manual feature engineering or prior guidance. With the increase of computing facilities, these methods are nowadays suitable for automated lesion segmentation tasks.
Aim: To perform a quantitative analysis and evaluation of six different brain lesion segmentation CNN models, studying the impact of their characteristics for accurate MS lesion segmentation.
Methods: The set of methods comprises (i) two U-NET models with either 2D (UNET2D) and 3D (UNET3D) input patches, (ii) a 3D U-NET with residual connections (UNET3D-R), (iii) two fully convolutional multi-resolution architectures using either multi-branching (FCB) or skip-connections (FCS) and (iv) a cascaded 3D CNN (CASC). The evaluation of the different CNN approaches was carried out on a clinical dataset (containing T1w and FLAIR images from 30 patients) and the international ISBI2015 challenge dataset (composed of 20 cases with T1w and FLAIR images). The Dice overlap (DSC) coefficients between the resulting segmentation and manual lesion annotations were calculated and compared to state-of-the-art methods.
Results: On the clinical dataset, UNET3D-R reported the highest overlap (DSC 0.66), followed by the UNET3D (DSC 0.63) and CASC (DSC 0.62). UNET2D yielded the lowest performance among the CNN methods (DSC 0.54), being still significantly superior to other state-of-the-art methods like LST (DSC 0.51). On ISBI2015, the top performance was shown by UNET3D (DSC 0.71), followed by CASC (DSC 0.68) and FCS (DSC 0.67). Similar to the previous case, UNET2D yielded the lowest performance (DSC 0.57).
Conclusions: In general, the CNNs approaches based on 3D input patches were found to be more sensitive to capture the inherent lesion anatomy when compared to 2D strategies. Concerning architecture, in both experiments, the popular u-shaped approaches were more accurate than both fully convolutional and the cascaded approaches. We plan to release all these automated CNN lesion segmentation methods to the MS research community, aiming to benefit from the improvements introduced by these new machine learning strategies.

Disclosure

A. Clérigues: nothing to disclose.
S. Valverde: nothing to disclose.
J. Bernal: nothing to disclose.
D. Pareto: has received speakung honoraria from Novartis.
J.C. Vilanova: nothing to disclose.
Lluís Ramió-Torrentà: has received compensation for consulting services and speaking honoraria from from Biogen, Novartis, Bayer, Merck, Sanofi, Genzyme, Teva Pharmaceutical Industries Ltd, Almirall, Mylan.
A. Rovira: serves on scientific advisory boards for Biogen Idec, Novartis, Genzyme, and OLEA Medical, and has received speaker honoraria from Bayer, Genzyme, Sanofi-Aventis, Bracco, Merck-Serono, Teva Pharmaceutical Industries Ltd, OLEA Medical, Stendhal, Novartis and Biogen Idec. A. Oliver: nothing to disclose.
X. Lladó: nothing to disclose.