Multiple sclerosis lesion detection and segmentation using a convolutional network of 3D patches

Sergi Valverde, Mariano Cabezas, Eloy Roura, Sandra González-Villà, Arnau Oliver, Joaquim Salvi and Xavier Lladó

Vision and robotics research institute (VICOROB), University of Girona, Girona (Spain)

INTRODUCTION and PURPOSE:

Magnetic resonance imaging (MRI) of the brain has been widely used during the last years in clinical practice. This image modality presents a high contrast for soft tissues, including white matter lesions (WML). Expert tracing of these lesions may be a time-consuming task and prone to variability due to the large number of MRI slices and the different conditions between experts. In order to reduce theses limitations, here we present a novel supervised automated 3D voxelwise lesion detection and segmentation classifier that is based on a cascade of two identical Convolutional Neural Network (CNN) architectures with posterior refinement.

METHOD:

CNN architecture:

Isotropic 3D patches centred around the voxel of interest are selected to increase the number of training samples

3D patches reduce the number of false positives in areas where normal tissue reassembles the lesion intensity profile

CNN architecture (**Figure 1**):

- preprocessed FLAIR, T1, T2 and PD modalities
- 15 x 15 x 15 x 4 input
- 32 x 7 x 7 x 7 convolutional layer
- average pooling with size 2 and stride 2
- 64 x 3 x 3 x 3 convolutional layer
- average pooling with size 2 and stride 2
- dropout layer with (t = 0.5)
- 256 units fully connected layer
- 2 units soft-max layer

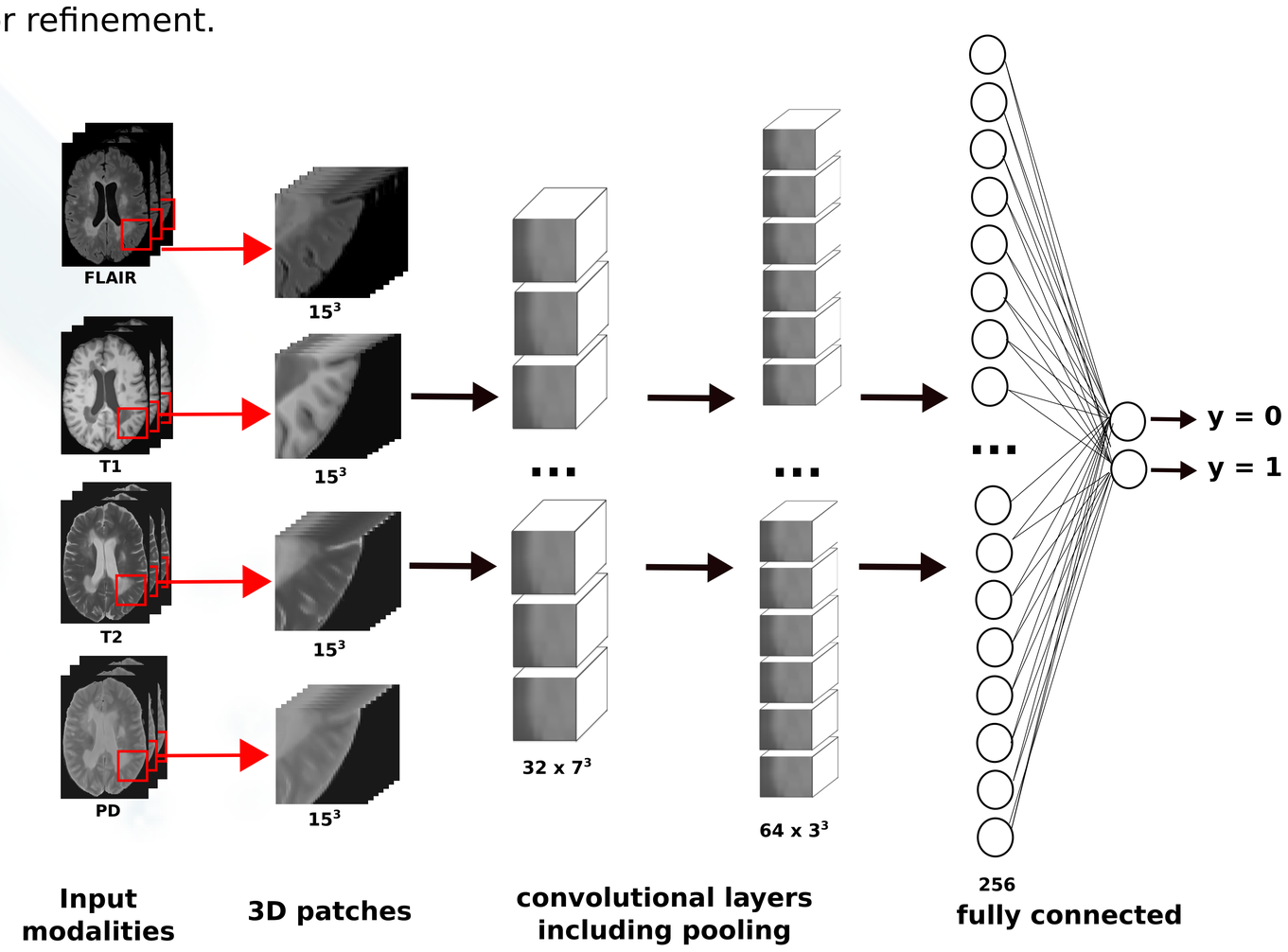

***Figure 1**: Architecture of the convolutional network taking candidate patches of size 15 × 15 × 15 as inputs*

Training:

A cascade of two identical classifiers is implemented in order to reduce the effect of false positives (**Figure 2**):

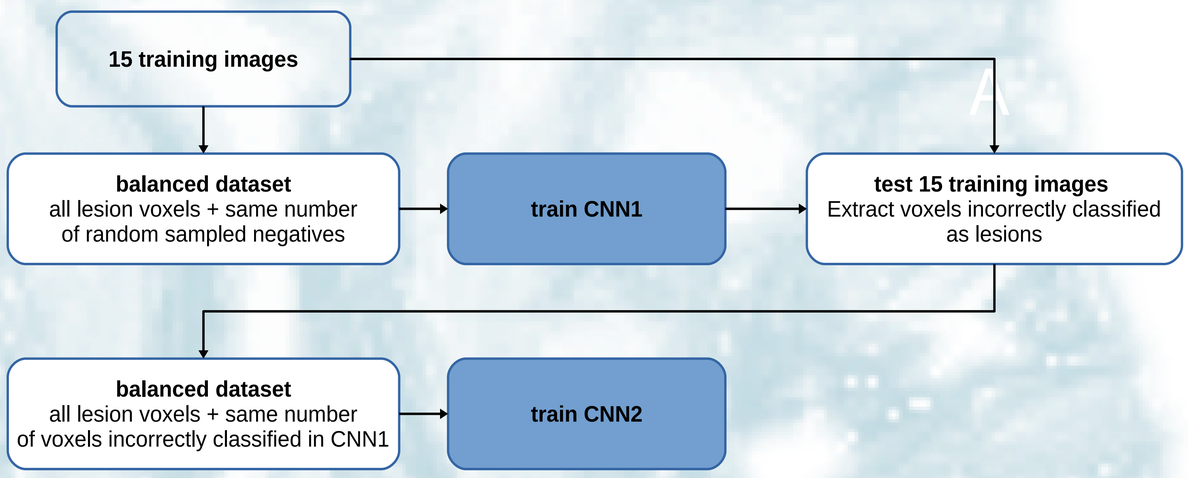

***Figure 2:** The proposed architecture is trained in two steps: First, we use all the voxels defined as lesion in the consensus and a random sampling of the same size of all the candidates that are not lesion. After training the CNN1, we test all the training images with the CNN1, obtaining a probabilistic map for each image voxel. These maps are used to select the most challenging false positives. Then, we randomly select a sample of this false positive voxels of the same size as the lesion voxels and we train a second classifier CNN2 using this dataset.*

Testing:

For each new image to test, all its voxels are evaluated using both nets CNN1 and CNN2. The probabilistic maps obtained in each case are multiplied and obtained lesion regions are filtered (**Figure 3**):

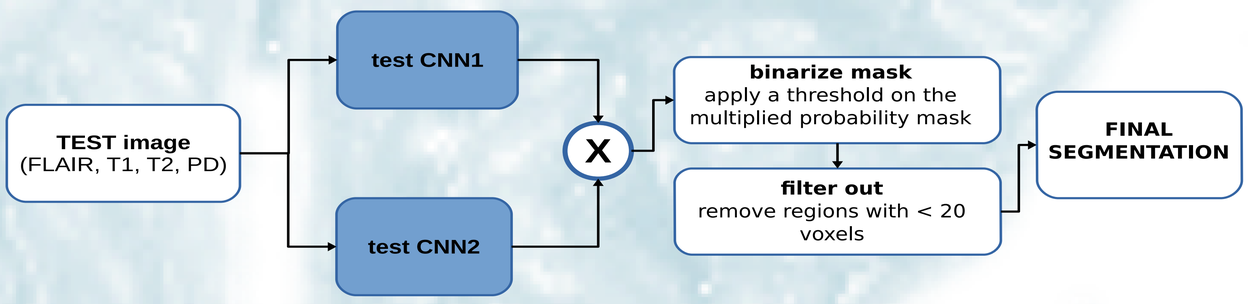

Figure 3: *Since both networks use the same positive voxels, when testing all lesions are considered in both nets. However, since we used different negative voxels, the probabilistic maps present different false positive detections that do not overlap after ouput multiplication, reducing the number of false positives. Finally, the probabilistic map is binarized and small regions are removed.*

EVALUATION:

A quantitative evaluation of the proposed pipeline is shown in (**Figure 4**):

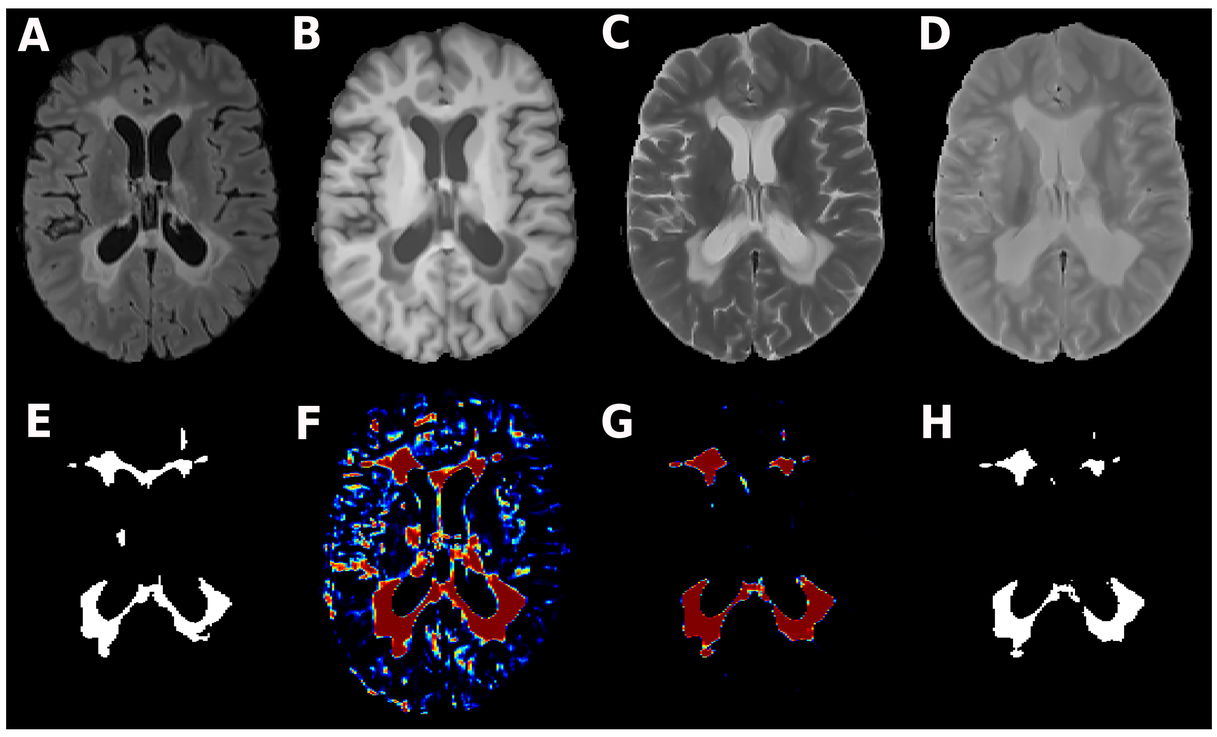

***Figure 4**: Quantitative evaluation of the training case 01016SACH after training it using leave-one-out. Pre processed input images A) FLAIR , B) T1, C) T2 and D) PD. E) Consensus ground-truth annotation. F) Output map after testing the input image with the F) CNN1 net and G) CNN2 net. H) Final resulting output segmentation mask after multiplying the CNN1 and CNN2 outputs, binaryzing and removing regions with small area.*

CONCLUSION:

The presented pipeline is our first attempt to segment MS lesion using a simple deep learning approach. Although qualitative evaluations are still not available, evaluations on training data suggest that these techniques are very promising. In this aspect, we are currently studying more sophisticated pipelines with the aim of increasing the accurary of automated MS lesion segmentation.

ACKNOWLEDGMENTS:

This work has been partially supported by "La Fundació la Marató de TV3", by Retos de Investigación TIN2014-55710-R, and by MPC UdG 2016/022 grant. The authors gratefully acknowledge the support of theNVIDIA Corporation with their

CONTACT INFORMATION:

Sergi Valverde, PhD.
University of Girona (UdG). Av. Lluís Santaló s/n
(P4-building). 17071, Girona (Spain)
Phone: (+34) 972419812 E-mail: svalverde@eia.udg.edu