Cornell University
MRI Research Lab

RSANet: Recurrent Slice-wise Attention Network for Multiple Sclerosis Lesion Segmentation

Hang Zhang, Jinwei Zhang, Qihao Zhang, Jeremy Kim, Shun Zhang, Susan A. Gauthier, Pascal Spincemaille, Thanh D. Nguyen, Mert Sabuncu, and Yi Wang

Objective

Since manual delineation of MS lesions is a time-consuming and highly operator-dependent task, which is influenced by lesion size, shape and conspicuity, it is of vital importance to develop efficient tool for multiple sclerosis (MS) lesion segmentation.

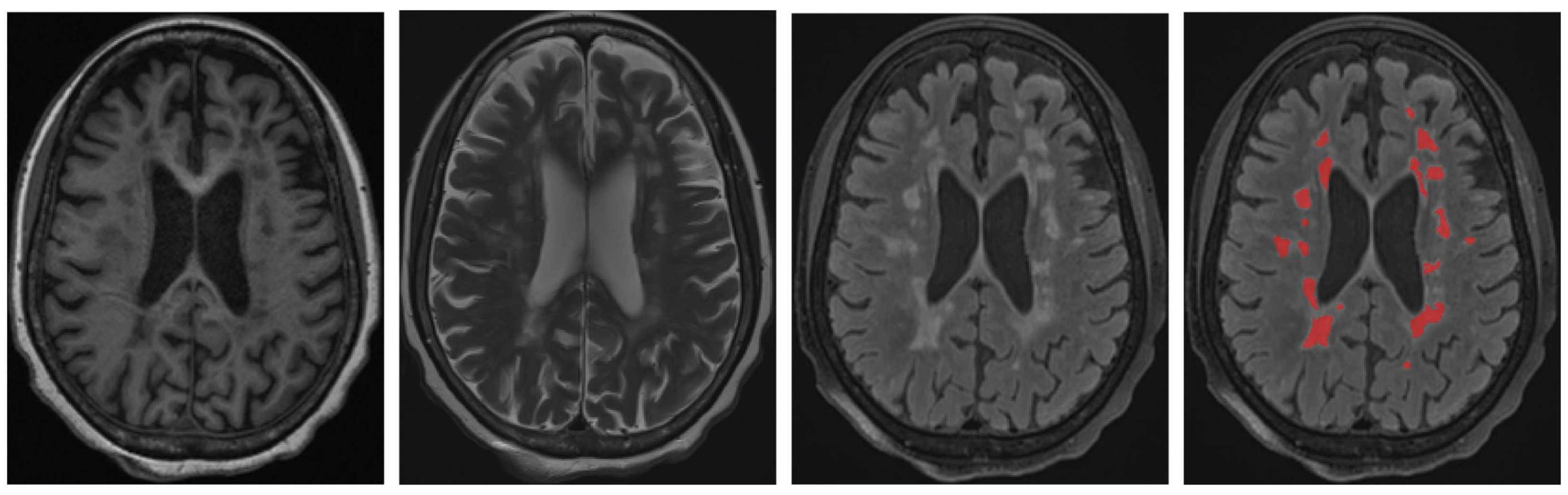

Fig.1. Example **T1, T2, T2-FLAIR** images and corresponding mask traced by a human expert and marked in red.

T2 weighted MRI images is a clinically important marker for MS lesion

Overview

Fully automated MS lesion segmentation tool is still a distant goal, as there are several inherent challenges in MS lesion data structure and limitations in previous methods:

- ❑ MS lesions vary a lot by lesion location, size, shape and conspicuity;
- ❑ Previous methods are not able to capture slice-wise correlations which are important for 3D MRI images;
- ❑ 3D MRI images are inherently large compared to 2D natural images, there lacks efficient methods to capture long-range dependencies.

Three main contributions of our method:

1. We propose a slice-wise attention (SA) block to get over the drawbacks of current methods in capturing slice-wise correlations;
2. We further propose a recurrent SA (RSA) block to capture long-range dependencies from all voxels;
3. Our RSA block reduces dramatically of GPU memory consumption and floating-point operations when computing the attention map.

Methodology

Slice-wise attention (SA) block:

Unlike other attention-based methods, we develop SA block to capture the slice-wise correlations among different slices of 3D MRI images. The Slice-wise attention map A, as is shown in Fig. 2, is computed as follows:

$$A_{ij} = \frac{\exp(M_2[i,:]M1[:,j])}{\sum_{k=1}^{D}\exp(M_2[i,:]M1[:,j])}$$

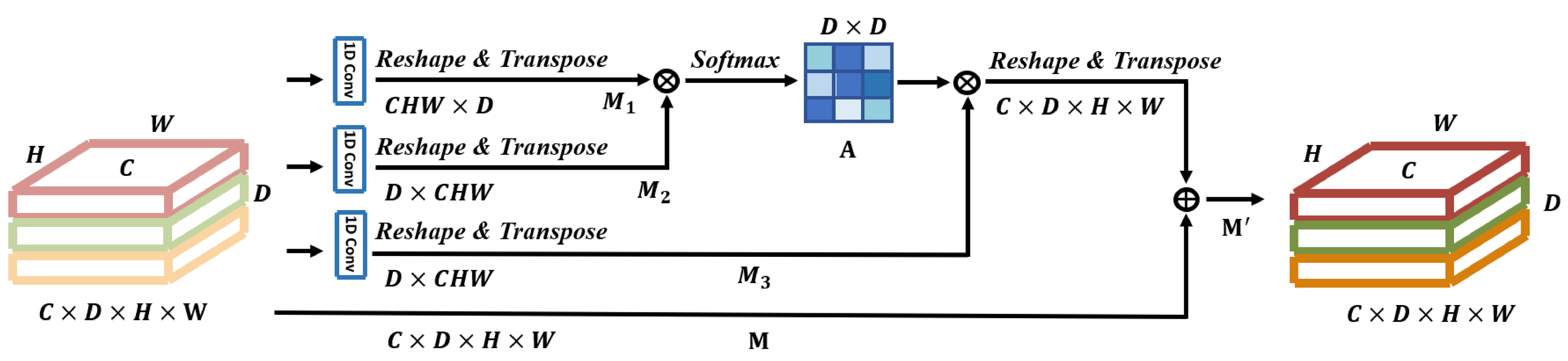

Fig.2. Details of the slice-wise attention block

Recurrent slice-wise attention (RSA) block:

By recurrently aggregating information from three SA blocks along sagittal, coronal and axial directions, our RSA block could capture global long-range dependencies.

$$M' = \alpha RT(AM_3) + M$$

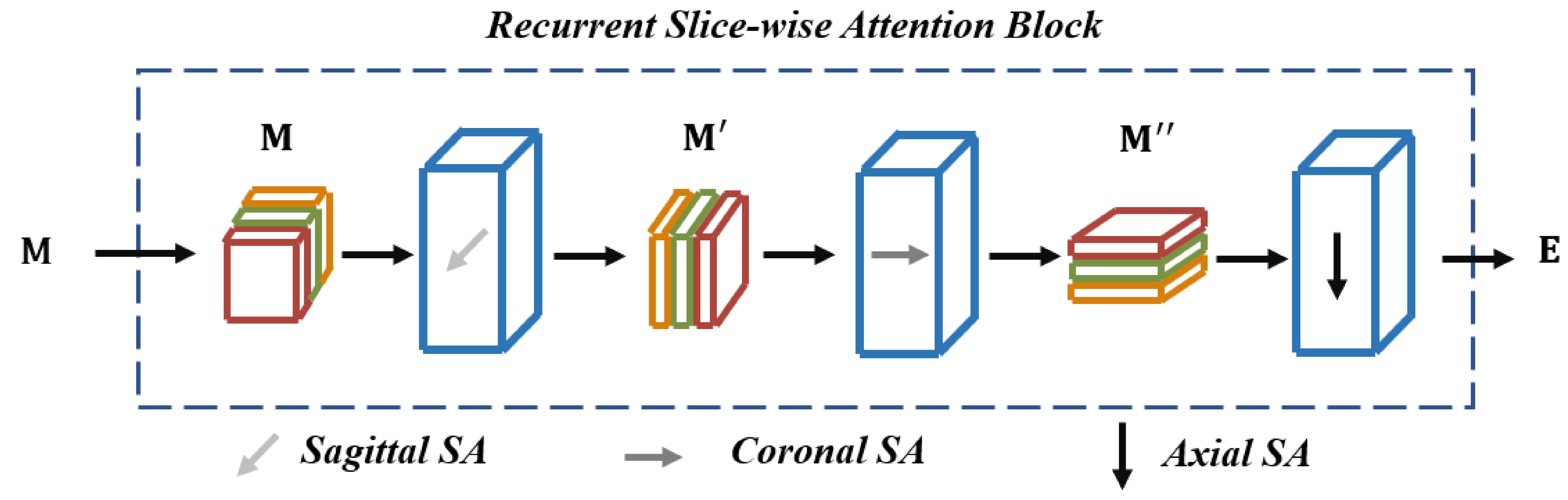

Fig.3. Details of the RSA block. The input is the feature map M. RSA block takes M and recurrently produces M', M'' as intermediate results. Finally, RSA block will output E, where each voxel is a weighted sum of all other voxels.

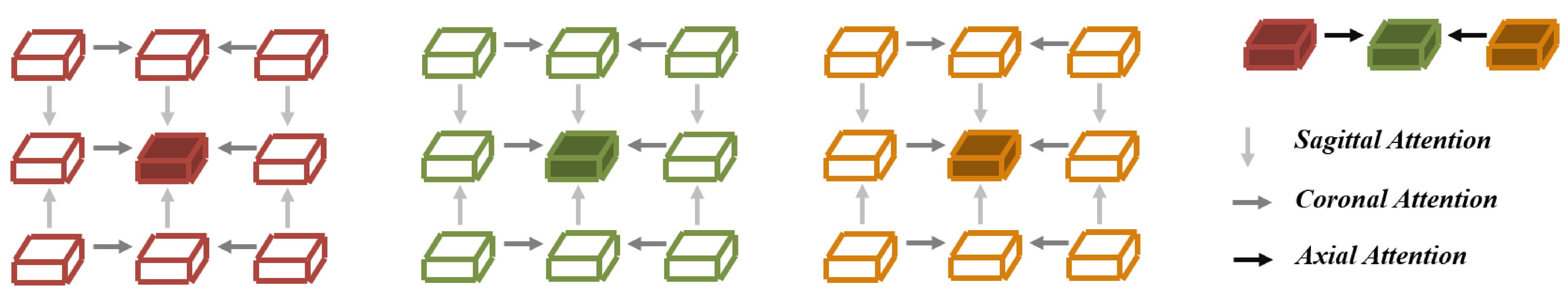

Fig.4. An example of information propagation in RSA block.

Experimental Results

Data set statistics:

- Multi-modalities T1, T2, T2FLAIR;
- 43 patients scanned on a GE 3T scanner;
- Size varies from 230 × 320 × 44 to 260 × 320 × 60;
- Voxel size 0.7 × 0.7 × 3.0 mm.

Method	Sample avg. dice	Voxel avg. dice	Sample avg. IoU	Voxel avg. IoU
3D U-Net	63.984%	69.640%	48.754%	53.506%
NCL-010	64.346%	70.473%	49.231%	54.500%
NCL-101	64.069%	70.121%	49.103%	54.090%
NCL-111	64.074%	70.185%	48.833%	54.170%
RSA-010	65.300%	70.207%	50.248%	54.200%
RSA-101	65.949%	**71.589%**	50.896%	**55.847%**
RSA-111	**66.011%**	71.054%	**50.917%**	55.201%

Table 1. Quantitative comparison of MS lesion segmentation.

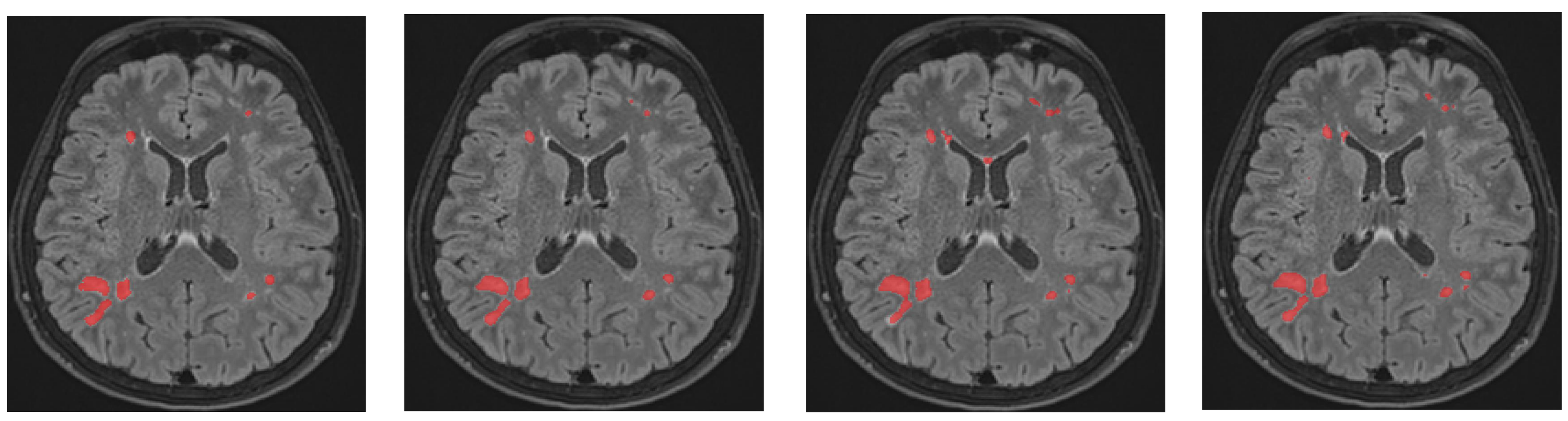

Fig. 5. Example segmentation result. From left to right are ground truth label, results of RSA-111, NCL-010 and 3D U-Net.

References

1. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) MICCAI 2015. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4 28
2. Wang, X., Girshick, R., Gupta, A., He, K.: Non-local neural networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 7794–7803 (2018)