Segmentation of Brain Structures with an Ensemble of Multi-Dimensional Convolutional Gated Recurrent Units

Robin Camarasa *[1, 2], Corentin Doué *[1, 2], Marleen de Bruijne[1, 3], and Florian Dubost[1]

[1]Biomedical Imaging Group Rotterdam, Department of Medical Informatics, and Department of Radiology, Erasmus MC, Rotterdam, The Netherlands
[2]Ecole des Mines de Saint-Etienne, Saint-Etienne, France
[3]Department of Computer Science, University of Copenhagen, Denmark

August 15, 2018

1 Method overview

Method. Our method is an ensemble of Multi-Dimensional Convolutional Gated Recurrent Units (MD-GRU) [1]. MD-GRUs have significantly less parameters than commonly used U-net methods, which can avoid overfitting on the training dataset. Furthermore, using MD-GRU mimics the slice by slice manual annotation process by considering a spatial dimension as a sequential dimension.

Material. The experiments were launched on four GPUs : one Nvidia GeForce GTX 1070, two Nvidia GeForce GTX 1070-Ti, and one Nvidia GeForce GTX 1080-Ti. The method uses Keras 2.2 [2] with Tensorflow 1.9 [3] as backend.

Dataset. The dataset of the challenge contains the T2-FLAIR, T1-weighted (T1-w) and T1-weighted inversion recovery (IR) 3D scans of 7 subjects. We did not use any other data to train our algorithms.

2 Preprocessing

Brain extraction. The skull is removed using a 3D U-Net with 16 convolutionnal layers trained on the FLAIR and the T1-w scans. The ground truth is a pixelwise binarization (0 background, 1 brain structure). During training, we apply on-the-fly random translations and rotations to the images. The loss function is computed as the number of false positives plus ten times the number of false negatives to favor masks larger than the brain over smaller ones. Adadelta [4] is used as optimizer. The computed mask is then applied to the 3 modalities (FLAIR, T1 and IR).

Tiling. The images are split in tiles of 70 * 70 * 22 voxels both at train and at test time, with a 50% overlap, to fit in the GPU memory and give a more robust (averaged) prediction.

*Contributed equally

Normalization. To be robust to outliers a 1%-99% percentile normalization is applied to brain-extracted data of each MRI sequence. Intensity values between 1% and 99% are consequently rescaled between 0 and 1.

3 Recurrent Neural Network

Model. The network takes a 3 channels array of tiles as an input : FLAIR, T1-w and IR ; and outputs a 9 channels array of tiles : cortical gray matter, basal ganglia, white matter, white matter lesions, cerebrospinal fluid in the extracerebral space, ventricles, cerebellum, brain stem and background, infarction and other. The architecture is composed of 3 MD-GRU layers linked with channelwise fully connected layers, as in [1]. Our MD-GRU applies a non padded Convolutional Gated Recurrent Unit [1], with a 2D convolution followed by batch-normalization in several directions before summing the outputs (6 parallel branches). The two first MD-GRU layers use the three spatial directions, both forward and backward, while the last MD-GRU layer only goes forward (3 parallel branches). The parameters are optimized with Adadelta [4] with Keras' default learning rate (1.0) and the activation function of the last layer is a softmax.

Training. Five models were trained with an averaged one-versus-all Dice loss on randomly selected train and validation splits. During training, on-the-fly random translations, rotations and flipping were used for the 500 first epochs, then random elastic deformations were added for the last 300 epochs. Training one model lasts 2 days on a single GPU.

4 Post-processing

Reconstruction. A reconstruction algorithm transforms the tiled output of a network into a full size output. Overlapping predictions are averaged.

Ensemble of five models. After reconstruction, the probabilistic outputs of the five models are averaged class-wise. For each voxel, the prediction is then the class with highest probability.

5 Acknowledgment

This research was funded by The Netherlands Organisation for Health Research and Development (ZonMw), Project 104003005.

References

[1] Andermatt, S., Pezold, S. and Cattin, P., 2016. Multi-dimensional gated recurrent units for the segmentation of biomedical 3D-data. In Deep Learning and Data Labeling for Medical Applications (pp. 142-151). Springer, Cham.

[2] Chollet, F., 2015. Keras.

[3] Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M. and Kudlur, M., 2016, November. Tensorflow: a system for large-scale machine learning. In OSDI (Vol. 16, pp. 265-283).

[4] Zeiler, M.D., 2012. ADADELTA: an adaptive learning rate method. arXiv preprint arXiv:1212.5701.