Modified U-Net for WMH Segmentation

Pablo Duque, Mariano Rincón, Jose Manuel Cuadra

Department of Artificial Intelligence, UNED

1 Pre-processing

The first step in the preprocessing is to generate the brain mask using HD-BET (Isensee et al., 2019) from the T1 image. In a previous work (Duque, Cuadra, Jiménez, & Rincón-Zamorano, 2019) we used the method explained in (Smith, 2002) but more recent work proposed in (Isensee et al., 2019) has proven better results basing their method en neural networks as well.

Then T1 images are co-registered to the FLAIR space to be able to generate a brain mask for FLAIR images as well. This was done using BRAINFit and BRAINResample modules from 3DSlicer (Fedorov et al., 2012).

A non linear contrast enhancement transformation is run on the FLAIR image to enhance areas where WMH is present, this has shown an improvement in DICE compared to not using this technique (Duque et al., 2019). Brain masks are then applied to both T1 and FLAIR images by simply multiplying both matrices to the mask.

2 Methods

2.1 Data augmentation

All slices that contain any level of WMH in the training set were augmented 6 times. The first five were done by applying random affine data transformations where values were picked from a normal distribution within the following ranges; for rotations with $[-30°, 30°]$ angles, shifts applied to both the x and y axis [-30%, 30%] of the total width and height, respectively, zoom on both axes with values in the ranges [0.9, 1.2] and shears in the range [-0.2, 0.2]. The sixth augmentation was done by applying all 5 transformations to each slice to further improve generalization.

2.2 Modified U-Net

The proposed solution uses a fully convolutional neural network based on the attention gated U-Net architecture (Schlemper et al., 2019).

Our proposed attention gated U-Net has three levels instead of four, which means it has only three pooling layers in the contracting path and three transposed strided convolutions (instead of upsampling layers) in the expanding path. This allowed us to reduce the number of parameters while keeping the same performance since the focus on small features in obtained thanks to the increased convolution kernel size and an initial stack of convolutions. All convolutional kernels are increased to size 11 to capture richer local data besides convolutions within the attention gates. RELU activations are used in all convolution layers. Within the attention gates convolution kernels are (5, 5) besides the obvious (1, 1) convolutions. Pooling layers are kept of (2, 2). There is an initial stack of 5 convolutions before the U-Net pattern of two convolutions and pooling starts then that pattern is applied three times in the contracting path. Initialization of all convolutional kernels is done by using the He normal (He, Zhang, Ren, & Sun, 2015) besides in the attention gates in which we use the Glorot or Xavier uniform (Glorot & Bengio, 2010).

In the reconstruction path of the U-Net is where attention gates are located. Reconstruction upsampling are always done by using (2, 2) strided transposed convolutions (sometimes wrongly referred as deconvolutions) to match the (2, 2) pooling on the downsampling path. Attention gates use batch normalization at the end of their process.

References

Duque, P., Cuadra, J. M., Jiménez, E., & Rincón-Zamorano, M. (2019). Data Preprocessing for Automatic WMH Segmentation with FCNNs. In *Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics)* (Vol. 11487 LNCS, pp. 452–460). Springer Verlag. doi: 10.1007/978-3-030-19651-$6_4$4

Fedorov, A., Beichel, R., Kalpathy-Cramer, J., Finet, J., Fillion-Robin, J. C., Pujol, S., ... Kikinis, R. (2012). 3D Slicer as an image computing platform for the Quantitative Imaging Network. *Magnetic Resonance Imaging*. doi: 10.1016/j.mri.2012.05.001

Glorot, X., & Bengio, Y. (2010). Understanding the difficulty of training deep feedforward neural networks. In *Journal of machine learning research.*

He, K., Zhang, X., Ren, S., & Sun, J. (2015). Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In *Proceedings of the ieee international conference on computer vision.* doi: 10.1109/ICCV.2015.123

Isensee, F., Schell, M., Pflueger, I., Brugnara, G., Bonekamp, D., Neuberger, U., ... Kickingereder, P. (2019). Automated brain extraction of multisequence MRI using artificial neural networks. *Human Brain Mapping*. doi: 10.1002/hbm.24750

Schlemper, J., Oktay, O., Schaap, M., Heinrich, M., Kainz, B., Glocker, B., & Rueckert, D. (2019). Attention gated networks : Learning to leverage salient regions in. *Medical Image Analysis*, *53*, 197–207. Retrieved from https://doi.org/10.1016/j.media.2019.01.012 doi: 10.1016/j.media.2019.01.012

Smith, S. M. (2002). Fast robust automated brain extraction. *Human Brain Mapping*. doi: 10.1002/hbm.10062