

Deeply Supervised Multi-Scale Fully Convolutional Networks for Segmentation of White Matter Hyperintensities

Guodong Zeng and Guoyan Zheng

Institute for Surgical Technology and Biomechanics, University of Bern, Switzerland

Abstract. We present a method to address the challenging problem of segmentation of White Matter Hyperintensities (WMH) from multi-modality MR images (T1 and FLAIR). Our method is based on deeply supervised multi-scale fully convolutional networks (FCNs), that are executed in two sequential stages and can directly map a whole volumetric data to its volume-wise labels. In order to alleviate the potential gradient vanishing problem during training, we designed multi-scale deep supervision. Validated on the 60 training datasets of the MICCAI 2017 Grand Challenge on WMH segmentation, our method achieved an average Dice Similarity Coefficient of 74.6%, an average precision of 69.0% and an average recall of 82.5%, when we randomly chose 45 datasets for training and tested our method on the remaining 15 datasets.

1 Method

Figure 1 illustrates our two-stage method for the automatic WMH segmentation in multi-modality MR images. We first develop FCN-1 which is used at Stage one to extract the brain and to exclude other structures, such as skull, eyes, etc. The resulting mask is then applied to the multi-modality MR images. We further develop FCN-2 which is used at Stage two to get the final segmentation of WMH.

At both stages, multi-scale deep supervision with long and short skip connections is employed to improve the performance. More specifically, Both FCN-1 and FCN-2 consist of two parts, i.e., the encoder part (contracting path) and the decoder part (expansive path). Long and short skip connections are used to help recover spatial context lost in the contracting encoder. See Fig. 1 for details. We further inject two down-scaled branch classifiers into our networks in addition to the classifier of the main network. By doing this, segmentation is performed at multiple output layers. As a result, classifiers in different scales can take advantage of multi-scale context.

T1 weighted and FLAIR MR images are from different modality and provide complementary information for our task. Thus, in both FCN-1 and FCN-2 we construct separate encoder path for different modalities and then concatenate the output of the encoder path at the beginning of the expansive path such that

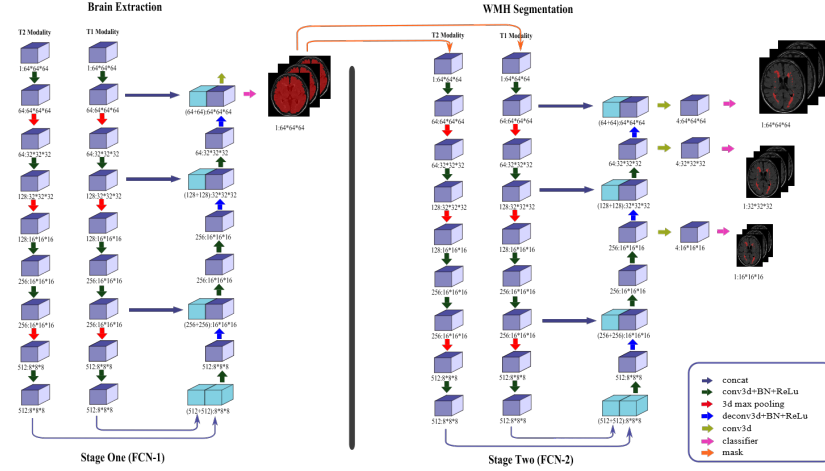


Fig. 1. A schematic illustration of our proposed network architecture. For each block, the numbers on the left side of the colon indicate the number of feature stack while the numbers on the right side of the colon represent the data size.

the decoder can fuse the complementary information to generate more accurate segmentation results.

The proposed network was implemented in python using TensorFlow framework and trained on a desktop with a 3.6GHz Intel(R) i7 CPU and a GTX 1080 Ti graphics card with 11GB GPU memory.

We conducted a validation study on the training data provided by MICCAI 2017 WMH segmentation challenge organizer, which contain T1 and FLAIR images of 60 patients from three difference clinical centers. More specifically, we randomly chose 45 datasets for training and tested our method on the remaining 15 datasets. In order to enlarge the training samples, data augmentation was utilized. Specifically, each training data was rotated (90, 180, 270) degrees around the y axis of the image and flipped horizontally (taking the z axis as the vertical direction).

Our trained FCN-1 can directly extract brain region from input images while the trained FCN-2 can estimate labels of lesion from the extracted brain region. Validated on 15 datasets, our method achieved an average Dice Similarity Coefficient of 74.6%, an average precision of 69.0% and an average recall of 82.5%.