# 3D Convolutional Neural Network with Skip Connections for WMH Segmentation

Miguel Luna[1], Sang Hyun Park[1]

[1] Department of Robotics Engineering,
Daegu Gyeongbuk Institute of Science and Technology, Republic of Korea,
shpark13135@dgist.ac.kr

**Abstract.** To address the problem in white matter hyperintensities (WMH) segmentation challenge at MICCAI 2017, we propose a deep learning based method that can utilize features from multi-modality images, *i.e.*, T1 and FLAIR images in this challenge.

## 1 Method

Our proposed method consists of training and testing stages. In the training stage, we first normalize the voxel spacing of preprocessed T1 and FLAIR images to $1\times1\times3$ $mm$ and their intensity values from zero to one. From the normalized images, we extract pairs of T1 and FLAIR 3D image patches (see Sec. 1.1) and then construct a 3D deep convolutional neural network (CNN) consisting of 18 convolutional layers to predict the labels of voxels in the patch (see Sec. 1.2). The outputs of our CNN model are $n$ probability patches where $n$ is the number of labels and the size of each probability patch is the same with the size of 3D image patches. For this challenge, $n$ was set as two, *i.e.*, one is for WMH and another is for others, and the patch size was heuristically set as $27\times27\times9$. In the testing stage, we similarly normalize the voxel spacing and intensity values of the preprocessed T1 and FLAIR images, extract pairs of T1 and FLAIR 3D image patches, and then predict their labels. The image patches are extracted with a certain interval less than the patch size so that all voxels can be contained in multiple patches. The intervals on three dimensions were set as 9, 9, and 3, respectively, and thus most voxels were included in 27 or less than 27 patches. To make a final segmentation label, we average the probabilities estimated from the multiple patches for each label at each voxel and then choose the label with the highest average probability as the final label of that voxel.

### 1.1 Extraction of Training Patches

Due to the relatively small size of WMH, using all patches extracted from entire image is inefficient and requires huge memories. To extract informative training patches, we learn a simple training model only using the patches extracted near the WMH, and then apply this model to the training images. The segmentations generated by this simple training model are inaccurate and especially include

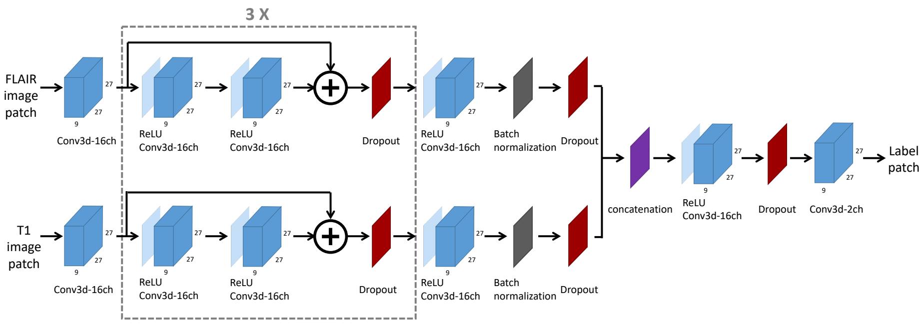


**Fig. 1.** Structure of our proposed convolutional neural network with skip connections

many false positives since many background patches are not included in the training set, but many certain background regions are effectively distinguished. We sample the training patches again near the true WMH and false positives to train a fine training model. Specifically, from the left top to the right bottom with 9, 9, 3 intervals on three dimensions, we extract the patches which contain more than five WMH or false positive voxels. For training of all sixty images, we used around 100,000 pairs of T1 and FLAIR patches.

### 1.2 Deep Convolutional Neural Network with Skip Connection

Figure 1 shows our proposed end to end 3D convolutional neural network. The network includes eight layers for T1 and FLAIR images, respectively, a concatenate operator for connecting layers in two threads, and two more layers to predict the final label patch. Specifically, for the eight layers in each threads, three skip connections are used from second to seven layers to make the deep structure using a relatively small number of training data [1]. Dropout is also used to prevent overfitting after the three skip connections. The outputs of the eighth layer in both threads are then merged into a concatenate operation [2], followed by ReLU, convolution, and dropout. Finally, the last convolution layer reduces the number of channels to $n$. The outputs of final layer are compared to the truth label patches and the errors are backpropagated in the training stage. Except for the final layer, we use 16 channels in the entire network to allow the skip connections and keep the network simple and fast.

## References


1. He, K., Zhang, X., Ren, S., Sun, J.: Deep Residual Learning for Image Recognition. In: Proceedings of Computer Vision and Pattern Recognition. (2015)
2. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: Proceedings of Medical Image Computing and Computer Assisted Intervention. (2015)