We propose to segment the White Matter Hyperintensities (WMH) of presumed vascular origin by the following two steps: (1) train a fully convolutional neural network (FCN) with short-cut architecture to perform efficient end-to-end learning and inference suitable for WMH characteristics using FLAIR images; (2) train another FCN to segment the white matters using T1 images; (3) apply the segmented T1 mask to remove false positives from the output of WMH FCN. Some details about FCN architectures are described below.

We adapt and modify the U-Net architecture [1]. To help preserve very small structures of WMH, we trim the previously deeper U-Net architecture to contain only three pooling layers, resulting in four different spatial resolutions. The number of channels in each convolutional layer is the same as in U-Net architecture. The overall network contains a fine-to-coarse downsampling path and a coarse-to-fine upsampling path. For each convolutional layer in the downsampling path, we use 3×3×3 kernels with stride of 2 and padding of 1 followed by a batch normalization layer and a ReLu layer. The pooling layer uses max pooling with stride of 2 in each dimension. In the upsampling path, the convolutional layer uses the same parameters with those in the downsampling path except that the last convolutional layer uses a 1×1×1 kernel. In addition, the pooling layer in the upsampling path is replaced by deconvolution layer with 2×2×2 kernels and strides of 2 in each dimension. It should be noted that an additional upsampling path with short-cuts from the corresponding pooling layers to the upsampling layers helps better capture small scale WMH, since the finer scale information before pooling layers are added to help predict the output.

In our end-to-end training, the loss function is computed over all voxels in a training image. Due to the sparsity of WMH existing in each training data, the distribution of WMH and background voxels is highly biased. Therefore, we use a global class-balancing weight in the loss function as follow:

$$loss = -\beta \sum_{j \in Y_+} \log \hat{y}_j \; - (1-\beta) \sum_{j \in Y_-} \log(1 - \hat{y}_j \;),$$

where $\hat{y}_j$ is the computed value after the final convolutional layer, Y_+ and Y_- represent the set of the foreground and background labels, Y is the sum of Y_+ and Y_-, and β = mean(Y_+/Y) is global weight precomputed over the total training data. Our network was implemented using the *Caffe* library [2], and trained on an NVidia K40 GPU.

Another FCN similar to the WHM is also trained to segment the white matter. Applying the segmented white matter to WHM helps to remove the false positives outside of white matter region, such as skull, nose, or necks. The ground truth mask for the white matter segmentation is obtained using the FSL software developed by Oxford [3], since the white matter mask is not available from the challenge organizer.

Intensive data augmentation strategy is used to increasing the variations of training data set. Each axial slice serves as one original training data. And we further rotate it in 8 angles and applied the left-to-right flip operation. This results in totally ~50000 training data and 5000 validation data.

[1] Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." MICCAI, 2015.
[2] Jia, Yangqing, et al. "Caffe: Convolutional architecture for fast feature embedding." Proceedings of the 22nd ACM international conference on Multimedia. ACM, 2014.
[3] M. Jenkinson, C.F. Beckmann, T.E. Behrens, M.W. Woolrich, S.M. Smith. FSL. NeuroImage, 62:782-90, 2012