Triplanar U-Net Ensemble Network (TrUE-Net) for segmentation of WMHs on brain MR images

Vaanathi Sundaresan, Mark Jenkinson, Ludovica Griffanti

Wellcome Centre for Integrative Neuroimaging, FMRIB, Nuffield Department of Clinical Neurosciences, University of Oxford, UK

September 4, 2020

Team Name: FMRIB_TRUENET_2

1 Data preparation

We used both FLAIR and T1-weighted images as inputs for the model. We considered the unprocessed (original) FLAIR and T1 images, reoriented the images to match the orientation of the standard MNI space, performed bias-field correction using FSL FAST [Zhang et al., 2001] and performed skull-stripping FSL BET [Smith, 2002]. We then registered the T1 image to FLAIR space using linear rigid-body registration (FSL FLIRT) [Jenkinson and Smith, 2001] and cropped the FOV for both images close to the brain. We then extracted 2D slices from the volumes from all three planes.

2 Data normalisation and slice extraction

For each subject, we applied Gaussian normalisation on FLAIR and T1 by subtracting the mean intensity and dividing by the standard deviation intensity value to normalise the intensity values to range between 0 and 1. For the axial plane, we cropped the slices to a dimension of 128 $\times$ 192 voxels. For sagittal and coronal slices, we cropped and resized the extracted slices to 192 $\times$ 120 and 128 $\times$ 80 voxels respectively, using bilinear interpolation.

3 Data augmentation

Data augmentation was applied using translation(x/y-offset = [-10, 10]), rotation (θ = [-10, 10]) and random noise injection (Gaussian, $\mu = 0$, σ^2 = [0.01, 0.09]), increasing the dataset by a factor of 5 and 3 for axial and sagittal/coronal planes respectively.

4 Model

In TrUE-Net, we combined three 2D U-Nets in parallel within an ensemble model, each one detecting WMHs from a different plane. In the ensemble architecture, variation in the individual probability maps (due to noise or spurious structure) reduces when they are combined with equal weights in the ensemble network.

Fig. 1 shows the architecture of the proposed TrUE-Net. For each plane, the 2D model takes FLAIR and T1-weighted slices as input channels and provides the probability map in the corresponding plane. In each plane, we trimmed the depth of the classic U-Net [Ronneberger et al., 2015] to obtain the 3-layer deep U-Net model to reduce the computational load. Our model mostly uses 3 $\times$ 3 convolutional layers, except for the initial 5 $\times$ 5 convolutional layers in the cases of sagittal and coronal U-Nets, since larger receptive fields could adapt well to variations in lesion characteristics and lower resolution along the z-direction. Each convolution layer is

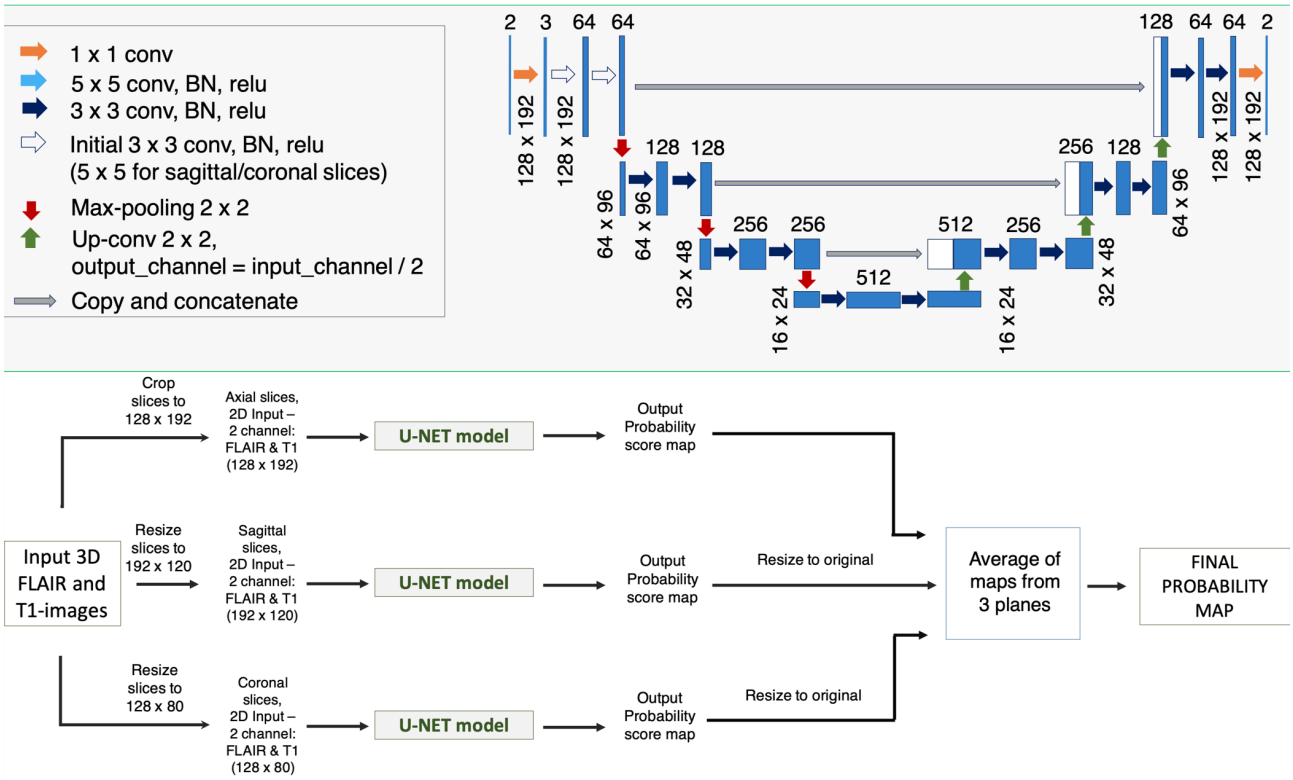

Figure 1: Triplanar U-Net ensemble network (TrUE-Net) architecture.

followed by batch normalisation and *ReLU* (rectified linear unit) layers. We added a 1×1 convolutional layer finally before the softmax layer for predicting the probability maps. In the ensemble model, training of U-Nets in the individual planes occurs independently using the slices extracted from the corresponding planes from the resized training images. During testing, we assembled the probability maps of slices from each of the three networks into a 3D probability map. We then obtained the final probability map by averaging the individual 3D probability maps from the 3 planar orientations, after resizing them back to their original dimensions.

4.1 Loss function

We used the sum of the voxel-wise cross-entropy (CE) loss function and the Dice loss (DcL) as the total cost function. We chose equal weights empirically for the Dice loss and the weighted CE loss while adding them to obtain the total loss. We weighted the CE loss function using a spatial weight map to up-weight the areas that are more likely to contain the less represented class (i.e. WMH). The weight value at each voxel is the sum of its distance from the ventricles and the distance from the gray matter.

4.2 Post-processing

We padded the probability maps with zeros to get them back to their original dimensions. We later masked the probability map with the white matter mask obtained from a dilated and inverted cortical CSF tissue segmentation combined with other deep grey exclusion masks (using FSL FAST [Zhang et al., 2001] and *make_bianca_mask* command in FSL BIANCA [Griffanti et al., 2016]). Finally, we thresholded the probability maps at 0.5.

References

[Griffanti et al., 2016] Griffanti, L., G. Zamboni, A. Khan, L. Li, G. Bonifacio, V. Sundaresan, U. G. Schulz, W. Kuker, M. Battaglini, P. M. Rothwell, et al., 2016, BIANCA (Brain Intensity AbNormality Classification

Algorithm): a new tool for automated segmentation of white matter hyperintensities: NeuroImage, **141**, 191–205.

[Jenkinson and Smith, 2001] Jenkinson, M., and S. Smith, 2001, A global optimisation method for robust affine registration of brain images: Medical image analysis, **5**, 143–156.

[Ronneberger et al., 2015] Ronneberger, O., P. Fischer, and T. Brox, 2015, U-net: Convolutional networks for biomedical image segmentation: International Conference on Medical image computing and computer-assisted intervention, Springer, 234–241.

[Smith, 2002] Smith, S. M., 2002, Fast robust automated brain extraction: Human brain mapping, **17**, 143–155.

[Zhang et al., 2001] Zhang, Y., M. Brady, and S. Smith, 2001, Segmentation of brain MR images through a hidden markov random field model and the expectation-maximization algorithm: IEEE transactions on medical imaging, **20**, 45–57.