WMH segmentation challenge MICCAI 2017
Team name - Achilles

1 Method

This section describes the method used for the 2017 MICCAI WMH segmentation challenge.

1.1 Preprocessing

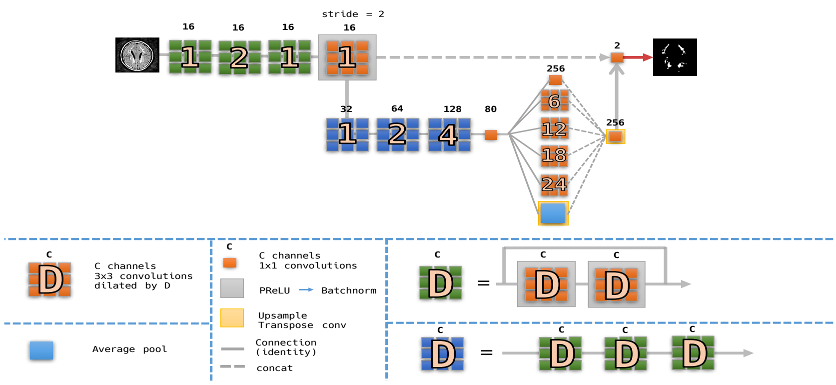
Initially, all images were resized to 200x200x100 and underwent two preprocessing steps at the image level. First, histogram normalization [1] was applied with minimum-maximum percentile values as $[1, 99]\%$. The threshold is chosen as the mean intensity per-subject image. Once this normalization is applied, we further apply in subject variance normalization.

1.2 Network architecture

The network architecture described below contains similar elements to HighResNet [2] and DeepLab v3 [3].

The architecture is shown below and contains the following elements:

- Atrous (dilated) convolutions
  Receptive field is larger using the same number of parameters, allowing us to easily capture large scale context in the image.
- Atrous spatial pyramid pooling
  At the final classification layer we apply atrous convolutions of different rates to capture multi-scale context information at the final layer.
- Residual connections
  Improve flow of information in the network through identity connections



All convolutional layers were initialized with the Kaiming Uniform initialization procedure [4]. The PyTorch framework was used for the implementation.

1.3 Training

During training, 90% of the dataset was used for training and 10% for validation. No additional data outside the competition was used and only the FLAIR.nii.gz images were considered. The training procedure corresponds to taking random 71^3 patches from the training set (batch size 1) and applying the following train time augmentations:

- $[0.8, 1.1]$ scaling
- $0 - 10°$ rotation on every axis

Optimizer: Adam optimization with ($lr = 0.0001$ and $\beta_0 = 0.9$, $\beta_1 = 0.999$).
Loss function: Binary dice loss
Notes:

- The network was trained until validation loss stabilized. The network was trained in AWS on a p2.xlarge instance (12GB GPU memory).
- Any voxels labelled as 2 (other disease) were set to the background class 0.

1.4 Evaluation

During evaluation, the image is first resized and normalized as done to the training images. After that, a sliding window approach is taken. The stride is set to 50 and patch size 70^3. For each patch:

1. Include 15 more voxels around the patch (if they exist) to form a 100^3 patch at maximum
2. Make 3 test time predictions (2 test time augmented patches as during training and 1 non-augmented) and average the results.
3. Remove the extra 15 voxels to obtain final prediction of the 70^3 patch

The complete prediction is then resized to the original input size.

References

[1] L. Nyul, J. Udupa, and X. Zhang, "New variants of a method of MRI scale standardization," *IEEE Transactions on Medical Imaging*, vol. 19, no. 2, pp. 143–150, 2000.

[2] W. Li, G. Wang, L. Fidon, S. Ourselin, M. J. Cardoso, and T. Vercauteren, *On the Compactness, Efficiency, and Representation of 3D Convolutional Networks: Brain Parcellation as a Pretext Task*, pp. 348–360. Cham: Springer International Publishing, 2017.

[3] L.-C. Chen, G. Papandreou, F. Schroff, and H. Adam, "Rethinking atrous convolution for semantic image segmentation," 2017.

[4] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," *CoRR*, vol. abs/1502.01852, 2015.