A Recurrent Residual CNN for Segmentation of Brain Structures in Multimodal MR Images

Hua Ma[1], Shuai Chen[1], Chengjia Wang[2]

[1]Biomedical Imaging Group Rotterdam, Erasmus MC, Netherlands
[2]BHF Centre for Cardiovascular Science, University of Edinburgh, The United Kingdom

Network
The proposed network takes a series of MR slices as input. It follows an encoder / decoder architecture (similar to U-net). Each encoder / decoder processes only one slice each time; the relation between slices was modeled with a recurrent subnetwork using two convolutional LSTM layer. The complete network follows the architecture: encoder-RNN-decoder with skip connection from the encoder to the decoder to introduce information of high resolution. The basic building block of the encoder and decoder network is a residual network with two convolutions and a skip connection. Transposed convolution was used for upsampling in the decoder. The decoder also uses dropout of 0.2 to reduce overfitting. Softmax was used after the last residual layer in the decoder to output the probability for all 9 classes (including the background).

Input
As some brain structures are only visible in certain modalities, all three modalities (Flair, T1, IR) were used as the three input channels of the network. The network learns joint features from all three modalities.

Loss
A DICE loss between the predicted segmentation and the ground truth label was used as the loss function to train the network.

Data augmentation
The following transformations were used as data augmentation: random translation, rotation, flipping left / right, flipping up / down, elastic deformation.

Training
The model was trained using 5 dataset (out of 7). Dataset 5 and 7 were left out from training and used as validation set to select hyperparameters.