Multi-modality Aggregation Network with Self-attention and Deep Feature Reconstruction

The multi-modality aggregation network with self-attention and deep feature reconstruction is used to segment the brain MR images. The proposed algorithm is automatic and takes about 10 seconds to process each sample on a NVIDIA GTX 1080Ti GPU. In this network, multi-modality images including T1, T1-IR and T2-FLAIR are used for segmentation. Part of the encoding structure of the network is taken from VGG16, to extract the multi-scale features of multi-modality image. The network uses self-attention mechanism to better capture spatial context information, and uses deep features to reconstruct the features at all levels. The network is two-dimensional and no other training data is added to help with the training. The preprocessing phase uses the classic data enhancement strategy.