AutoDenseSeg (pxln@hotmail.com)

Abstract

This document briefly describes techniques we used in automatic segmentation of the prostate in transversal T2-weighted MR images for the PROMISE12 challenge. We use a densely connected CNN with 3D convolutions for semantic segmentation to tackle this problem. Details of each step are described next in each section.

Data processing

Each volume is normalized to has a zero mean and a unit variance. The training data is augmented by rotations (90, 180, 270 degrees) and flipping in axial plane. For training, we first crop a ROI region that encompasses the prostate for each volume. Then we randomly cropped 16x64x64 sub-volumes from this ROI region and background parts for training. The batch size is 8.

Network Structure

The proposed network consists of a down-sampling path and an up-sampling path, followed by a softmax layer for prediction. The down-sampling path follows a DenseNet-like network structure, including several dense blocks and transition blocks with 3D pooling. The up-sampling path uses transpose convolutions with stride 2 to restore the high-resolution signal. 3D convolutions are used to exploit spatial correlations in a 3D volume.

For better regularization, a dropout layer with a ratio 0.2 (80% probability to keep) is added after each convolution.

Deep supervision

Except the final main prediction, we take the smaller volumes with ½ of the original resolution produced during the down-sampling path to get an auxiliary prediction. Again, transpose convolutions are used for the up-sampling. Both predictions are used in loss metric.

Loss metric

We use cross entropy for the loss metric. For deep supervision, the weighted loss of the two predictions are used.

Evaluation

Implementation

The algorithm is implemented in a modified caffe, supporting 3D convolution and 3D pooling. We use a GeForce GTX 1080 Ti GPU for training. The input size is 16x64x64 with a batch size of 8 for training.

Inference process

During inference, we use overlapped sub-volumes for prediction. Only the main prediction is utilized. For overlapped regions, the prediction probabilities are averaged to get the final prediction. As a post-processing, we only retain the largest connected set with a label "foreground" as the prostate. No other post-processing is used.

Results

We split the 50 training cases into a training set with 45 cases and a validation set with 5 cases, i.e. 0, 10, 20, 30 and 40. The proposed scheme achieves a dice score of 0.879 on the validation set and 0.883 on the training set.