MA 598 MACHINE LEARNING SEMINAR
SUMMARY 3

ID 5107

3. U-Net: Convolutional Networks for Biomedical Image Segmentation

What is their primary result? This paper introduces a novel network architecture and training strategy that relies on data augmentation of a starting set of annotated training samples.

Why is this important? The paper shows that such a network is capable of outperforming the prior best method—a sliding-window CNN developed by Ciresan et al.—as well as significantly lowering the time required to train the network, and achieves better segmentation.

What are their key ideas? The architecture developed by the writers improves on the sliding-window CNN by extending a fully convolutional. The way in which it is extended is further explained in the architecture section. The network uses very little training data by augmenting the samples through random elastic deformations. This is important in biomedical segmentation since such deformations are the most common variation in tissue.

What's the architecture? The network architecture consists of a contracting path and expansive path. The contracting path consists of the repeated application of two 3-by-3 convolutions followed by a ReLU and a 2-by-2 max pooling operation with stride 2 for downsampling. At each downsampling path, the number of feature channels is doubled.

The expansive path consists of an upsampling of the feature map followed by a 2-by-2 upconvolution that halves the number of feature channels, a concatenation with the cropped feature map from the contracting path, and two 3-by-3 convolutions, each followed by a ReLU.

How did they train and evaluate it? The network is trained by feeding it the input images and their corresponding segmentation maps with a stochastic gradient descent implementation of Caffe.

More interesting is the data augmentation that is part of the training. The initial data is augmented by through random elastic deformations and gray value variations. This augmentation teaches the network the desired invariance and robustness required of it. The deformations are generated using random displacement vectors on a coarse 3-by-3 grid. The vectors themselves are sampled from a Gaussian with a 10 pixel standard deviation

Date: January 25, 2019.