U-Net: Convolutional Networks for Biomedical Image segmentation

Summary written by Vedant Jain
April 14, 2021

Summary: This paper describes a novel approach to training deep networks that requires very few training images while still providing high model accuracy. Essentially the architecture involves a convolution path which contracts the images to capture specific context based features and then an expansion path that helps localize these selected regions. Furthermore, they show that in the ISBI challenge for segmentation of neuronal structures their model was not only very fast requiring close to less than a minute on a GPU but they also won the contest.

Related work: CNN's have been used for visual recognition tasks for quite some time, Yet they lack in their use for localization of specific features. Furthermore, they require large amounts of training and testing data. [2] But to actually determine the class of a specific pixel required a different approach. Ciresan et al, showed a simple pioneering approach in which they took an image and then broke it into many small patches which were fed through a CNN, and then each patch was classified.[1] Drawbacks to this method include redundancy with overlapping pixels during training and testing as well as the increased amount of parameters to include in the network training.

Approach: The major contribution of Ronneberger et al. Involves a completely new style of network that includes both a contracting path and an expansive path. Each step of the contracting involves shrinking the image by a 3X3 convolution, 2X2 max pool with stride 2. The expansion part involves up-sampling feature output and then a 2X2 convolution and then two 3X3 convolutions followed by ReLU and then a final 1X1 convolution to map the features to the selected class values. The model is trained using a batch size of 1 and a cross entropy loss function, that is applied at the level of the individual pixel. Furthermore, they performed data variation methods like shifting and rotations to make the network robust to deformations. They also made use of the dropout function.

Datasets, Experiments and Results: The authors of this paper attempted their segmentation network on three different segmentation tasks. The first one was on neuronal structures in electron microscopic recordings. This they got from the ISBI competition. This includes 30 training images, which came with annotated ground truth. The network had a wrapping and pixel error of .000353 and .0611 respectively. They also attempted the algorithm on the ISBI cell tracking challenge where they had the highest success rate of .9203 on the PhC-U373 dataset and .7756 on the DIC-HeLa dataset.

Strengths: This paper describes a very straightforward approach to image segmentation. The author does a good job of directly providing the evidence for why their network works as well as the exact mechanisms for the method. It a minimalist paper that provides a couple simple figures that directly describe the network, how they tested it, and some example features extracted.

Weaknesses: This paper lacks elaboration on exactly why they chose the convolution steps that they did. For example, I am curious why they used only 3X3 convolutions or max pool with a stride of 2. It would be good for them to have experimented with different methods and described why or why they didn't choose them. Furthermore, the figures could have described some of the performance differences based on different intermediate steps in their convolutions. I would've also liked to see them experiment with other activation functions.

Reflections: Olaf Ronneberger et al, describe a simple yet powerful network to perform image segmentation. They describe how a convolution (shrinking) arm and then a expanding (upsampling) arm of a neural network can be used to highly accurately segment images. The authors test this method on three different datasets describing the highest accuracy on the ISIB cell tracking challenge. This method makes it possible to train very high accuracy networks that require very little training data.

References

[1] G. L. G. A. S. J. Ciresan, D.C. Deep neural net- works segment neuronal membranes in electron microscopy images. 2012.

[2] S. I. H. G. Krizhevsky, A. Imagenet classification with deep con- volutional neural networks. 2012.