Problems and Progress in Automating Electron Microscopy Segmentation

Matthew Guay[1], Zeyad Emam[2], Richard Leapman[1]

[1.] National Institute of Biomedical Imaging and Bioengineering, National Institutes of Health, Bethesda, MD, USA
[2.] AMSC Department, University of Maryland, College Park, MD, USA

Biologists have long used electron microscopy (EM) to examine the nanoscale structure of biological matter. *Image segmentation* groups image pixels together into labeled regions corresponding to image content, and it is a fundamental tool in the quantitative analysis of EM data. Historically, segmentation was done by hand; it is a tedious procedure, but could be completed within acceptable time frames for the quantities of data produced by electron microscopes. However, modern EM hardware is capable of generating gigapixel 2D images and teravoxel 3D image volumes thanks. One technology in particular which contributes to this trend is serial block-face scanning electron microscopy (SBF-SEM) [1], which generates large 3D image volumes by scanning the face of a large sample block, shaving off a few nanometers with a diamond ultramicrotome, and repeating this process until the sample is consumed. This technique offers to radically improve our understanding of the organization of biological systems, but it is infeasible to segment large data regions by hand.

Researchers have recognized the potential for automating microscope image segmentation for decades. Segmentation tasks can be formulated computationally, and a sufficiently-advanced computer vision algorithm ought to be able to parse the content from a wide variety of images as humans do. This is made difficult by at least two factors. First, the difficulty of a segmentation problem is data-dependent. For many EM analysis problems of interest to biologists, the image segmentation algorithms produced by the computer vision community have historically not been accurate enough to be useful. A segmentation algorithm becomes useful when the time required to run the algorithm and manually correct any errors becomes less than the time required to perform the same segmentation by hand. Second, EM segmentation problems currently lack a black-box software solution. Designing an algorithm to solve a particular segmentation problem still generally requires the work of a machine learning expert, a resource unavailable to many microscopy labs.

In the past few years it has become possible to build and train neural networks to solve segmentation problems, and research in this direction has greatly improved the accuracy of the segmentation algorithms available to the EM community. In 2015, Ronneberger et al. introduced u-nets [2], a now-prototypical example of the encoder-decoder neural networks being used in image-to-image translation problems such as segmentation. While a step forward, the original u-net alone is not enough to create a useful segmentation algorithm for problems of interest to our lab. Multiple research groups have worked to extend this initial work, adding new computational elements for more elaborate networks. One aspect that has undergone intense development has been the use of 3D data windows instead of 2D [3]. It is clear even in human experience that a single 2D image is often insufficient to resolve all features within that image; humans use 3D context by looking at adjacent regions along the z axis, and computer benefit from 3D context as well. We have also found in our own work that ensemble strategies significantly increase the performance of encoder-decoder segmentation algorithms [4]. For any classification problem, a statistical ensemble averages the predictions of multiple individual classifiers. These and other strategies for improving segmentation accuracy are a crucial part of the drive for practical automated EM segmentation.

Meanwhile, we have been working to solve the problem of machine learning expertise requirements by developing a software package, *genenet*, which allows the user to build an ensemble of algorithmically-designed encoder-decoder networks without requiring special training in machine learning. While this is only a stopgap solution while the research continues on segmentation algorithms which work across multiple tasks, it provides a powerful workflow for solving the segmentation challenges facing EM labs today.

[1] Denk, Winfried, and Heinz Horstmann. "Serial block-face scanning electron microscopy to reconstruct three-dimensional tissue nanostructure." PLoS biology 2.11 (2004): e329.
[2] Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." International Conference on Medical image computing and computer-assisted intervention. Springer, Cham, 2015.
[3] Çiçek, Özgün, et al. "3D U-Net: learning dense volumetric segmentation from sparse annotation." International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Cham, 2016.
[4] Guay, Matthew D., et al. "Neural Network Ensembles Will Enable Teravoxel Image Segmentation for Electron Microscopy" Biophysical Journal 114.3 (2018): 343a.
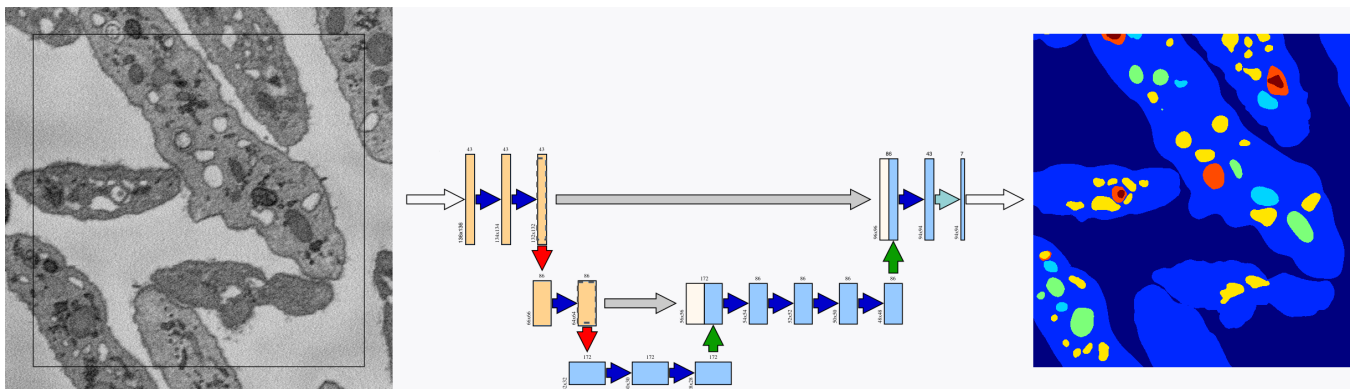

Figure 1. Demonstration of the segmentation process with encoder-decoder networks. An input image is decomposed into a multiscale collection of features by an encoder module (orange), then an encoder module (blue) synthesizes those features into a new output image.
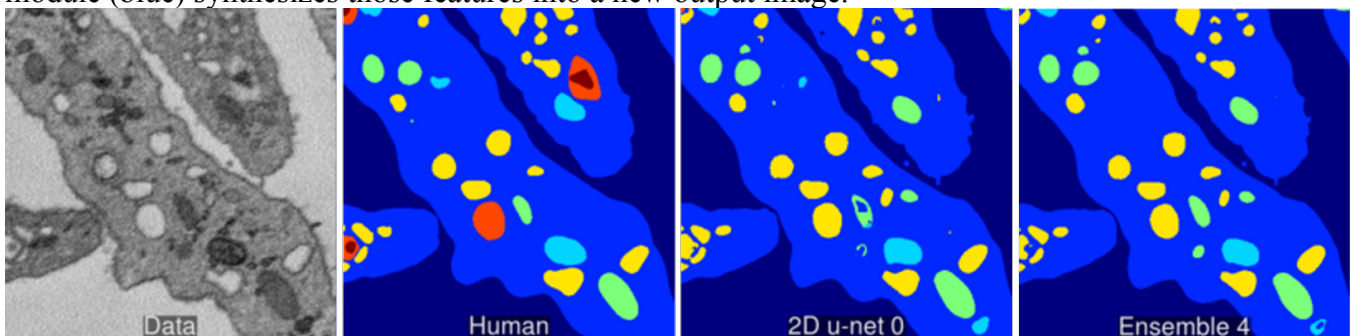

Figure 2. An example of the encoder-decoder ensemble approach to a segmentation problem. A data window (*left*) and human ground truth segmentation (*middle-left*) are compared with the output of the single best-performing network from a set of 80 randomly-generated networks (*middle-right*), as well as the output of an ensemble of the best four networks (*right*).