SPOTLIGHT

Pumping up the volume

Catherine G. Galbraith[1,2]

The time and cost of annotating ground-truth images and network training are major challenges to utilizing machine learning to automate the mining of volume electron microscopy data. In this issue, Gallusser et al. (2023. *J. Cell Biol.* https://doi.org/10.1083/jcb.202208005) present a less computationally intense pipeline to detect a single type of organelle using a limited number of loosely annotated images.

Ever since Robert Hooke peered into his microscope and identified the first cell over 350 yr ago, scientists have been trying to see the cell's contents with better resolution in the hope that if every protein and every interaction can be ascertained, then the mysteries of cell biology will be revealed. After having spent centuries improving resolution and cataloging the newly uncovered cellular contents, microscopy is now at the point where focused ion-beam scanning electron microscopy (FIB-SEM) can reveal the entire cell with a 3D isotropic resolution better than 5 nm, approaching protein resolution (1). However, the density of FIB-SEM data is so great that manual cataloging is intractable—it could take up to 60 person-yr to identify and annotate the boundaries of all the organelles within a cell (2). Hence, the need to develop machine-learning (ML) based organelle segmentation tools to identify organelles automatically.

ML approaches rely on manually annotating organelle boundaries of only a small subset of the data. These "ground truth" data are then used to train deep convoluted networks to automatically recognize the structures in the remaining unannotated data. A recent pipeline created by the COSEM project automates the segmentation of up to 35 organelles/intracellular features from relatively few, but precisely annotated data sets (2). To achieve this level of detail, annotating and labeling every voxel in a 1 μm³ area

block took one person 2 wk. Large multichannel 3D-U-Nets were then trained to predict many (14) or few (4) of the annotated entities, with the time for inference of the 14 organelle network >500,000 iterations (2).

Gallusser et al. (3) present an alternative approach for identifying a single type of organelle within a data set, called automated segmentation of intracellular substructures in electron microscopy (ASEM). ASEM allows for looser annotation by relaxing the requirement of voxel accuracy. The authors found that simple structures like mitochondria, nuclear pores, and clathrin-coated structures could be annotated with established tools, such as Ilastic (4) and Volume Annotation and Segmentation Tool (5). For annotating more complicated structures, such as Golgi, the authors developed a new graph-cut method for annotating (available at https://github.com/kirchhausenlab/incasem). The graph-cut tool operates on the whole 3D volume (rather than 2D slices), using several sparse 2D brush strokes (seeds) in a few 2D arbitrarily spaced planes to annotate the volume. The choice of tool for annotation of each type of organelle was made empirically, driven by the intent to decrease pipeline cost by annotating as accurately as possible with the least amount of time. On average, the ASEM project produced data annotated for one type of organelle at roughly 0.8 h per 1 μm (3).

ML learning on the annotated images was done with 3D U-Nets based on the architecture used in Funke et al. (6) with three down-sampling layers with a factor of two and two convolutional layers on each down-sampling level. Although ASEM and COSEM used similarly defined and implemented U-Nets, there is a significant difference. COSEM was implemented to identify a minimum of four organelles or intracellular features simultaneously, while ASEM uses a separate U-Net to identify each organelle. Obtaining network predictions of three organelles (mitochondria, ER, and Golgi apparatus) on the same FIB-SEM volume required training three networks and superimposing the findings (Fig. 1). While this can be significantly faster (100,000–150,000 iterations/organelle), the predictions for each organelle are independent. Users may need to use some corrections to ensure that the same voxel is not assigned to multiple organelles in the same data set.

It is important to note that ASEM performed well with specimens obtained by both chemical fixation (CF) and high-pressure freezing with freeze substitution (HPFS). While HPFS produces FIB-SEM specimens with better ultrastructural preservation, CF is more likely to be employed clinically (biopsy specimens). Moreover, since models trained on mitochondria or ER annotations prepared by CF performed

....
[1]Oregon Health and Science University, Portland, OR, USA; [2]Quantitative and Systems Biology Program in Biomedical Engineering and The Knight Cancer Institute, Portland, OR, USA.

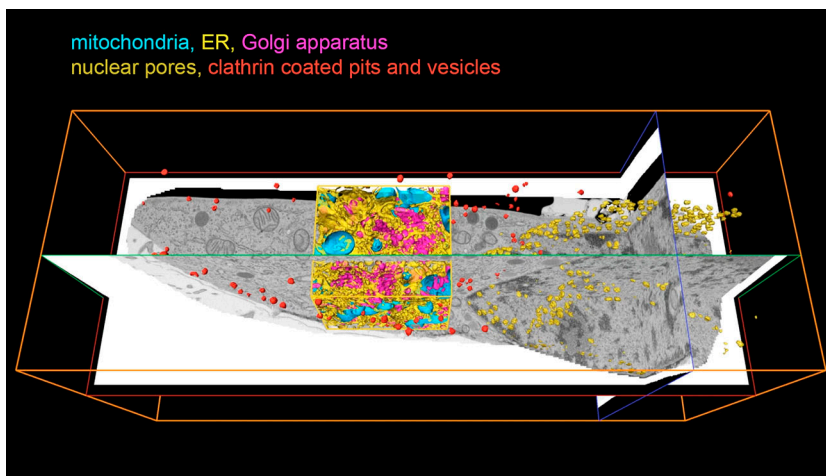Correspondence to Catherine G. Galbraith: galbrcat@ohsu.edu.

Figure 1. **Predictions of organelles on naïve interphase HPFS cell 15 SVG-A at 5 nm resolution.** The image is modified from a single frame from Video 3 (3). Organelles are color coded to match labels. The model was trained on interphase HPFS cell 19 HeLa and cell 20 HeLa at 5 nm resolution.

poorly on cells prepared by HPFS and vice versa, the authors showed that combining training data from both protocols allowed them to create generalist models that performed nearly as well on naïve data as models trained on data from the matched fixation method.

Gallusser et al. (3) also demonstrate the use of fine-tuning for improving segmentation. They retrain a model trained on one cell using a simplified transfer learning approach with ground truth annotations from a naïve cell. Examples indicate that only 5,000–10,000 iterations are needed to increase the prediction accuracy throughout the rest of the naïve cell. Fine-tuning worked well except when the pre-trained model already produced good segmentation (high F1 score) on naïve cells, where F1 = TP/(TP+[FP+FN]/2), TP is a true positive, FP is a false positive, and FN is a false negative.

The authors used ASEM to provide two biological demonstrations of new findings (Fig. 1). In the first, 10 nuclear pores were annotated to provide the ground truth to train a model that performed well after 100,000 training iterations to document a non-normal baseline distribution. In the second test of ASEM with relatively small structures, the model was trained with 15 endocytic plasma-membrane-coated pits representing different stages of clathrin coat assembly. After 80,000–100,000 training iterations, the model accurately recognized all endocytic coated pits in the trans-Golgi network, including caveolae. Rather than retrain, the caveolae were filtered out by size and appearance. Analysis of the resultant large, segmented data yielded eccentricities of the assembly pit consistent with a budding mechanism in which the growth of a clathrin coat drives membrane invagination, ultimately creating constriction and closure.

The ASEM pipeline was implemented on an NVIDIA DGX Station A100, a hardware system built for artificial intelligence and analytics. Despite using this beefy workstation, Gallusser et al. (3) posit that a standard workstation with a 12 GB GPU and 500 GB CPU would work and presume 64 GB CPU memory would also suffice since training

processes out-of-memory datasets. The latter processing power is within reach of many labs.

The spectacularly detailed online data made available by the COSEM project in the dedicated web resource OpenOrganelle (https://openorganelle.janelia.org/) allowed a direct comparison of F1 segmentation performance using the same data. Although ASEM is constructed to segment only a single organelle simultaneously, it yields F1 scores similar to or better than COSEM by segmenting the same data with fewer annotations and iterations.

This pipeline demonstrates that while the ultimate goal of cell biology may be to map the entire cell's contents, resource limitations will continue to demand that scientists restrict mining these incredibly rich data sets by concentrating on their organelles of interest. New tools such as the graph segmentation tool and the ASEM pipeline make these goals more attainable. Moreover, the ability to apply them to many different types of organelles holds the promise of improving workflows for a wide range of questions.

References

1. Xu, C.S., et al. 2021. *Nature*. https://doi.org/10.1038/s41586-021-03992-4
2. Heinrich, L., et al. 2021. *Nature*. https://doi.org/10.1038/s41586-021-03977-3
3. Gallusser, B., et al. 2023. *J. Cell Biol*. https://doi.org/10.1083/jcb.202208005
4. Berg, S., et al. 2019. *Nat. Methods*. https://doi.org/10.1038/s41592-019-0582-9
5. Berger, D.R., et al. 2018. *Front. Neural Circuits*. https://doi.org/10.3389/fncir.2018.00088
6. Funke, J., et al. 2019. *IEEE Trans. Pattern Anal. Mach. Intell.* https://doi.org/10.1109/TPAMI.2018.2835450