

Incremental Machine Learning Based Segmentation with Expert-in-the-Loop Workflow Using FIB-SEM Imagery

Rengarajan Pelapur

Thermo Fisher Scientific, Houston, Texas, United States

In this work, we are demonstrating an integrated machine learning 3D platform for expert-in-the-loop based workflows. Interactive visualization, simple tools for marking ground-truth coupled with machine learning modules for handling large 3D imagery such as FIB-SEM (Focused Ion Beam Scanning Electron Microscopy) can result in more detailed image analysis. Incremental machine learning modules have a unique advantage in domains where manually marked data is hard to collect and share. With incremental machine learning, models can be updated as new data is collected without losing the classification knowledge from previous trainings. While there have been similar efforts to get trainable models in microscopy [1] but annotating and receiving expert time is challenging. Machine learning and deep learning have a lot of potential to become a regular part of image analysis workflows in electron microscopy (EM). By providing experts a fully integrated 3D platform with essential tools of large data visualization, annotation and incremental machine learning the goal is to make difficult image analysis, a few steps easier.

Following our work in [3], we have extended the code and integrated our work using the extensible platform of Avizo-Amira and derived applications. This platform gives us access to a rich set of image processing, interactive tools that will help us bring machine learning domain knowledge to experts in various domains.

To demonstrate such an integrated platform, we have chosen a publicly available dataset [4] which is a 5x5x5 micrometer sized section of a CA1 hippocampus region of a mouse brain. The resolution of acquisition of this section was done at 5x5x5 nanometers using a Focused Ion Beam Scanning Electron Microscope. The problem statement is focused on segmenting mitochondria which are typically 0.5 to 10 microns in diameter. Analyzing and annotating data at this resolution would take months of laborious effort including expert time [5]. Most of high-resolution EM data, due to such challenges, have the potential to go without proper image analysis. Detailed image analysis, specifically, segmentation of mitochondria are tough problems to solve even with state-of-the-art image processing and machine learning. Many of the algorithms that perform well on natural imagery do not perform well on EM data because there are many data specific challenges.

Our work relates directly to those done in [6],[7] and [5]. A fully 3D interactive, annotation oriented and simple invocation of powerful machine learning tools is a complete end-to-end workflow for experts and there is a lot of potential in literature to expand in this domain. Our incremental machine learning, classification and user feedback input back into the classifier in conjunction with a rich set of image processing filters is implemented using Avizo-Amira XPand extension [11]. Our machine learning classes are computational modules that use Avizo-Amira image processing application programming interface (API) calls. These calls are specific feature computation comprising of the pixel and region-level features. We compute multiple scales of these features in order to capture a range of varying degrees of subtle edges and varying patterns of an underlying texture. Adding normalized grayscale as an additional feature gives us a rich variation of the original grayscale values. Our scale space work follows classic scale space theory [12]. Random Forests implementation used is parallel C++ code modified from the original [13] using OpenMP [14].

Our expert-in-the-loop workflow begins by the user importing a 3D dataset into Avizo. Interactive volume rendering and ortho slice can be performed by visualizing the entire dataset. This is followed up by traversing to the segmentation workroom where tools can be used to mark up ground-truth. The user can choose to provide as sparse or dense annotations using 3D brush, blow, interpolation across slices and other semi-automated image processing tools. The user can provide any number of labels to train the classifier. Invoking the machine learning tool based on Random Forests is done by navigating to a processing workroom and providing Grayscale image and annotated labels as inputs to a module called 'Train'. Another input required is a file path where a trained model file will be written out. If the training module is invoked using an existing training file, then the module simply loads the random forest into memory and continues training using the new set of labels and information from the user.

'Predict' as another module, requires the training file and a target grayscale volume to classify. Classification is performed by computing the feature set and organizing the information per voxel. This provides a pixel-level classifier. The probability volume that is computed is then transferred to the project view by using an Avizo inherent data structure for generic handling of 3D data. As our modules conform to the Avizo-Amira engine, a side effect and advantage is that these modules can be used as part of the recipe interface. Recipes in Avizo-Amira are used for repeatable steps of image processing. The expert can then edit the missed predictions to retrain. This feedback loop will help improve the classification and help collect validation data on the entire dataset or multiple datasets. Our goal is also to make these modules available through the Xtra Library [15] online for users and researchers to download and use directly in related applications.

Our extended tests of using the entire training data of 165 slices in the sub-volume [4] resulted in increased probability of mitochondria isolation. We have extended our previous work [3] to custom modules extending the Avizo engine to provide a complete 3D incremental machine learning platform combining powerful volume rendering and annotation tools for data collection and training models for future accurate segmentation of EM data.

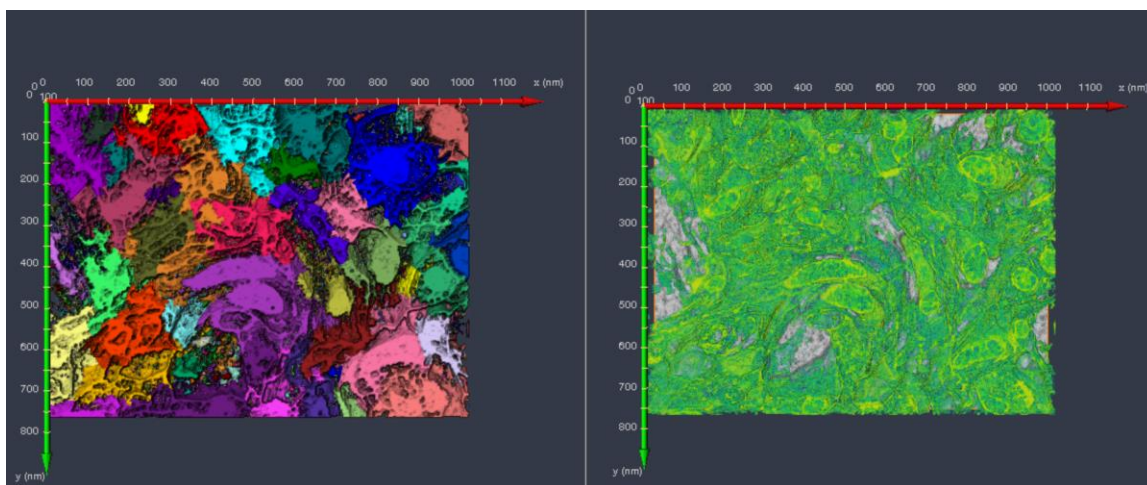


Figure 1. (from L-R) : Figure (L) showing separated mitochondria and associated membranes with multiple labels and (R) showing probability volume with Random Forests based prediction with opaque segmented mitochondria

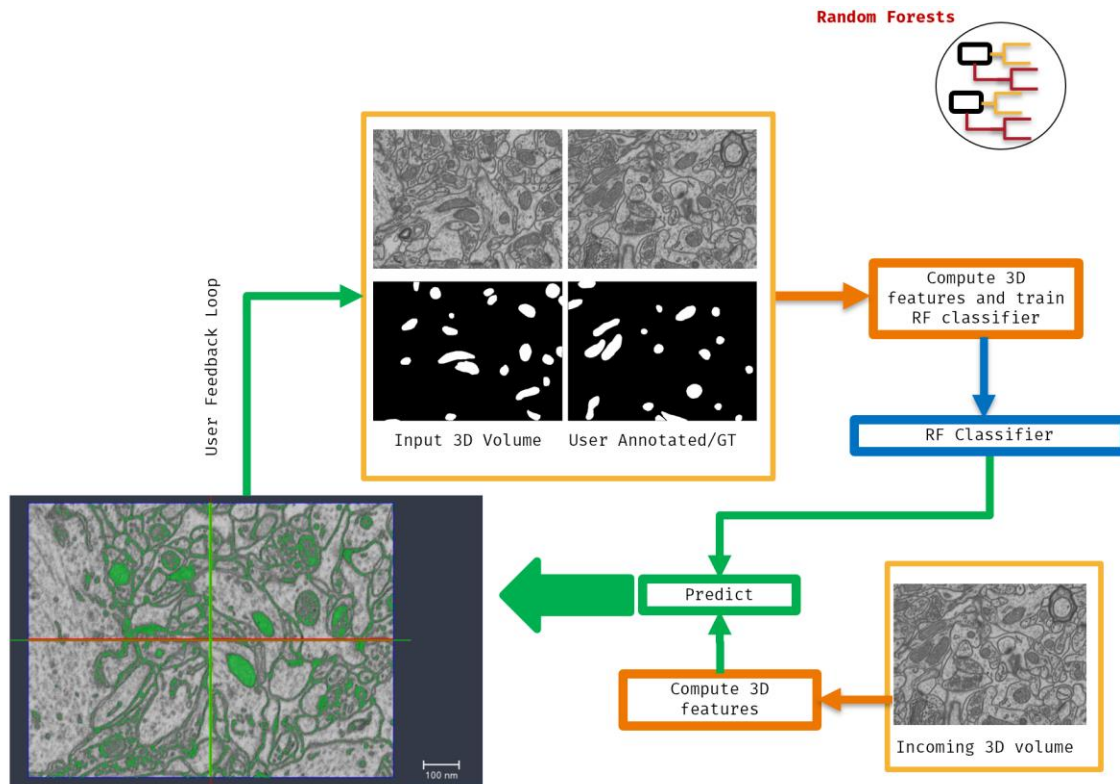


Figure 2. Figure showing user based annotations and feedback loop for incremental machine learning using random forests.

References

- [1] I. Arganda, et al.. "Trainable Weka Segmentation: a machine learning tool for microscopy pixel classification." *Bioinformatics*, 2017: 2424-2426.
- [2] L. Waller, and T. Lei "Computational imaging: Machine learning for 3D microscopy." *Nature*, 2015: 416-417.
- [3] R. Pelapur, and M. M. Heck. "3D Platform for Machine Learning Based Segmentation and Visualization Using FIB-SEM Imagery." *M&M*, 2019: 214-215.
- [4] Mitochondria Detection in EM Stacks. cvlab.epfl.ch/data/data-em/
- [5] A. Lucchi; Y. Li; P. Fua : Learning for Structured Prediction Using Approximate Subgradient Descent with Working Sets. 2013. Conference on Computer Vision and Pattern Recognition (CVPR), Portland, Oregon, USA, June 23-28, 2013. DOI : 10.1109/Cvpr.2013.259.
- [6] A. Kreshuk, C. Straehle, et al. "Automated segmentation of synapses in 3D EM data." In *2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pp. 220-223. IEEE, 2011.
- [7] C. Sommer, et al.. "Interactive learning and segmentation tool kit." *Systems Biology of Human Disease* (2010): 230-33.
- [8] H. Nguyen and Q. Ji. Shape-Driven Three-Dimensional Watersnake Segmentation of Biological Membranes in Electron Tomography. *IEEE TMI*, 616–628, 2008.
- [9] E., Paiva, et al., 2010. Detection of neuron membranes in electron microscopy images using a serial neural network architecture. *Medical image analysis*, 14(6), pp.770-783.
- [10] D. Cirosan et. al, 2012. Deep neural networks segment neuronal membranes in electron microscopy images. *ANIPS*, pp. 2843-2851.
- [11] Avizo XPand. (Feb 24, 2019)

- [12] T. Lindeberg. *Scale-space theory in computer vision*. Springer , 2013
- [13] RFLib, <http://vision.mas.ecp.fr/Personnel/teboul/RFLib.php>
- [14] L. Dagum et al.. OpenMP: an industry standard API for shared-memory programming. *IEEE CSE* pp.46-55, 1998
- [15] The Xtra Library, <https://xtras.amira-avizo.com/>.
- [16] A.A. Taha et al. Metrics for evaluating 3D medical image segmentation: *BMC MI*, p.29, 2015
- [17] Open Inventor, <http://developer.openinventor.com/>