WINDOWZ: An Accessible Machine Learning Pipeline for Biomedical Image Segmentation and 3D Reconstruction

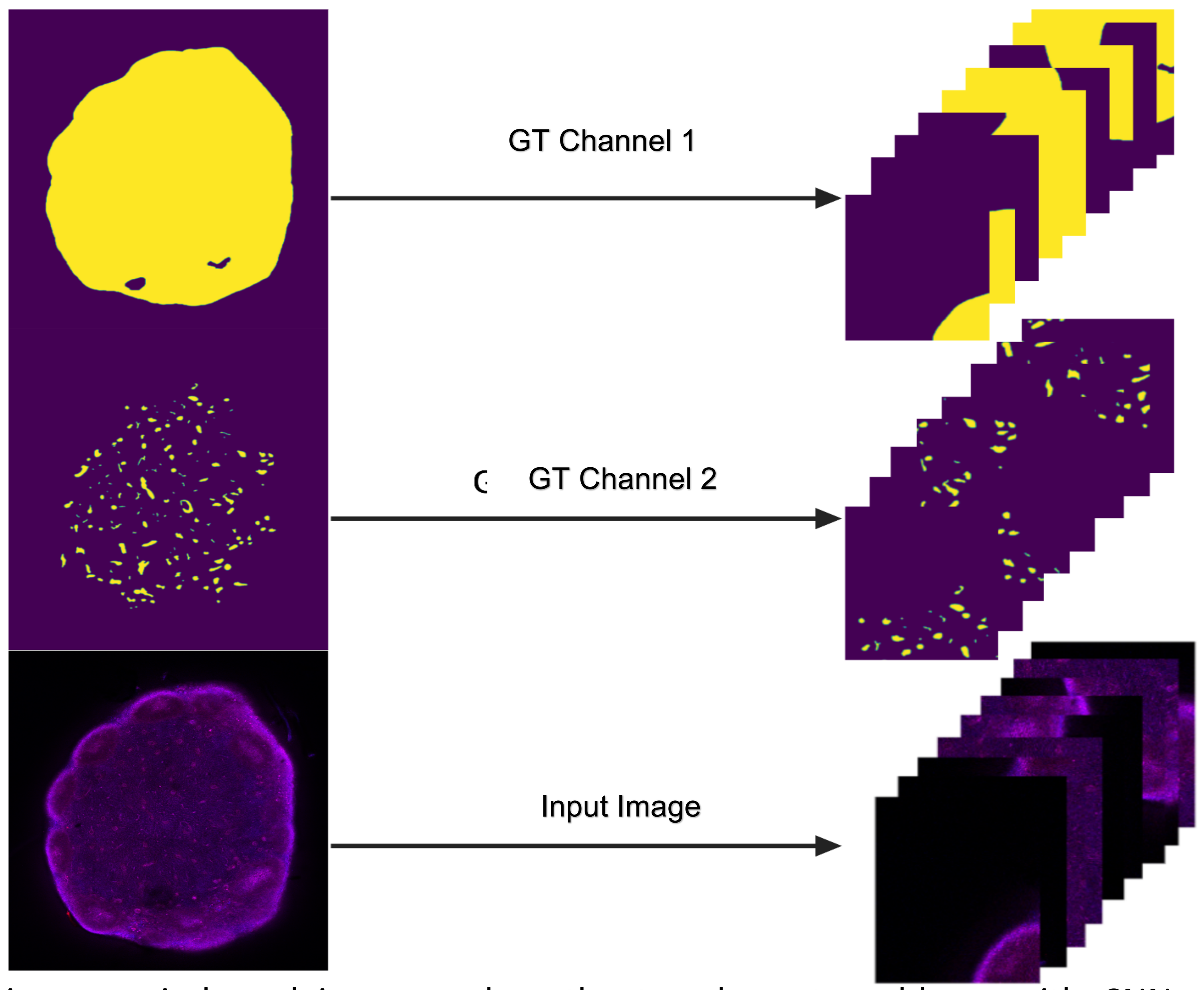
John Bean[1], Logan Hallee[2], Jason P. Gleghorn[2,3]

Departments of Computer & Information Sciences[1], Bioinformatics[2], and Biomedical Engineering[3]

Introduction

In cell biology, image segmentation helps identify and track individual cells, enabling the study of cell behavior, growth, and interactions. It also aids in the detection of abnormalities, such as tumors or diseased cells, and the identification of critical cell regions—which are both essential for diagnosing and understanding diseases. However, segmenting cells by hand is difficult and time consuming. Convolutional Neural Networks (CNNs) can automate image segmentation tasks by extracting the features of an image, and reconstructing said features into a segmented map. CNNs produce high quality and accurate results, but traditionally they need to be trained on large datasets. We propose an alternative approach to CNN implementation that makes image segmentation more lightweight and accessible. In addition to 2D segmentation, our novel pipeline offers the capability to produce fully segmented 3D point clouds

Windowed approach to image segmentation

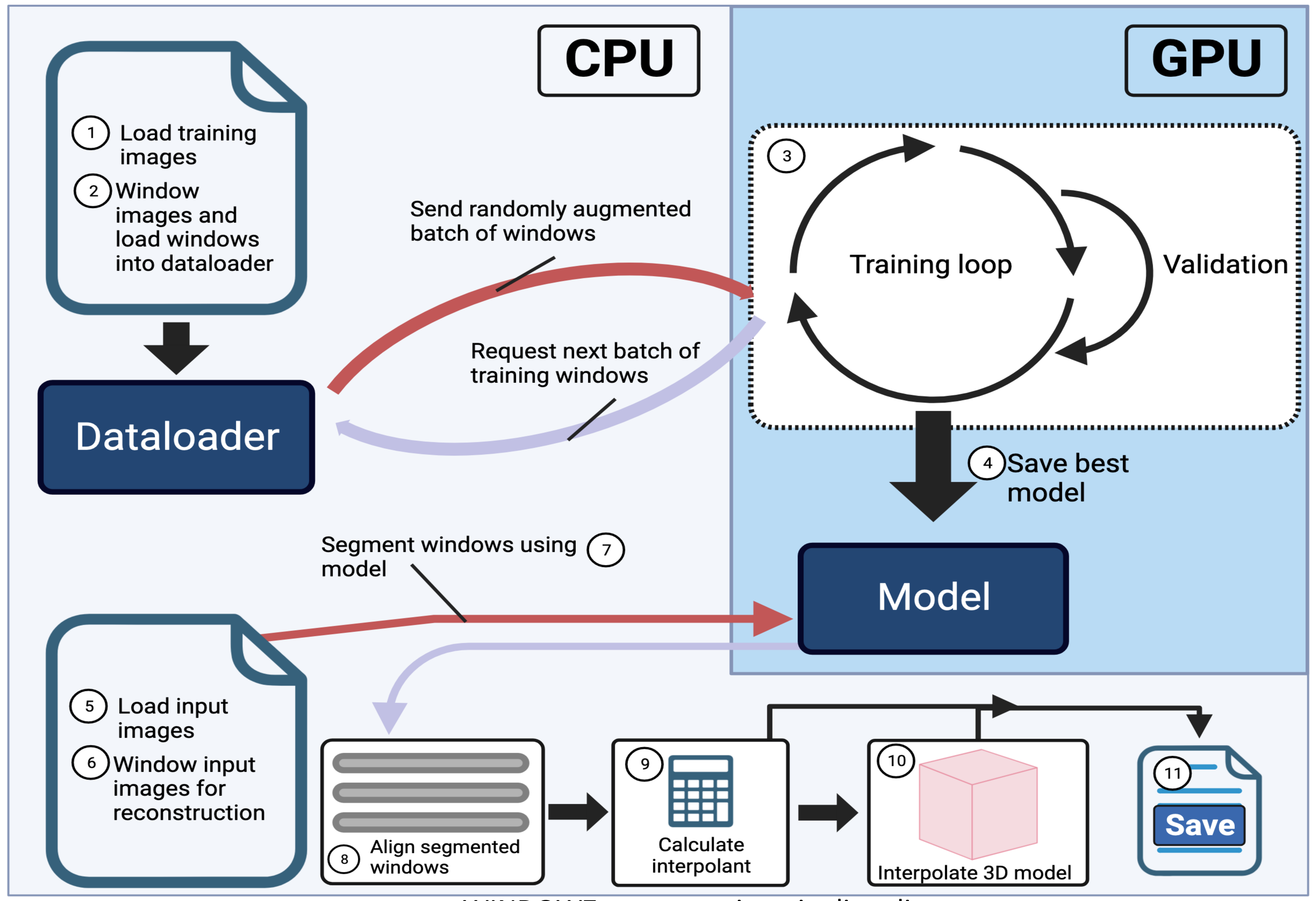


Training on windowed images solves the two largest problems with CNNs by increasing the total dataset size while decreasing the size of the individual samples—overall decreasing the amount of data and processing power needed to utilize CNNs.

Old approach

Our old approach to windowed image segmentation was comprised of three scripts, one for each key functionality. This methodology is not ideal as it statically augments the windows, meaning that the trainer is exposed to the same images every batch, which leads to overtraining. This approach also fails to utilize computational resources effectively.

WINDOWZ pipeline

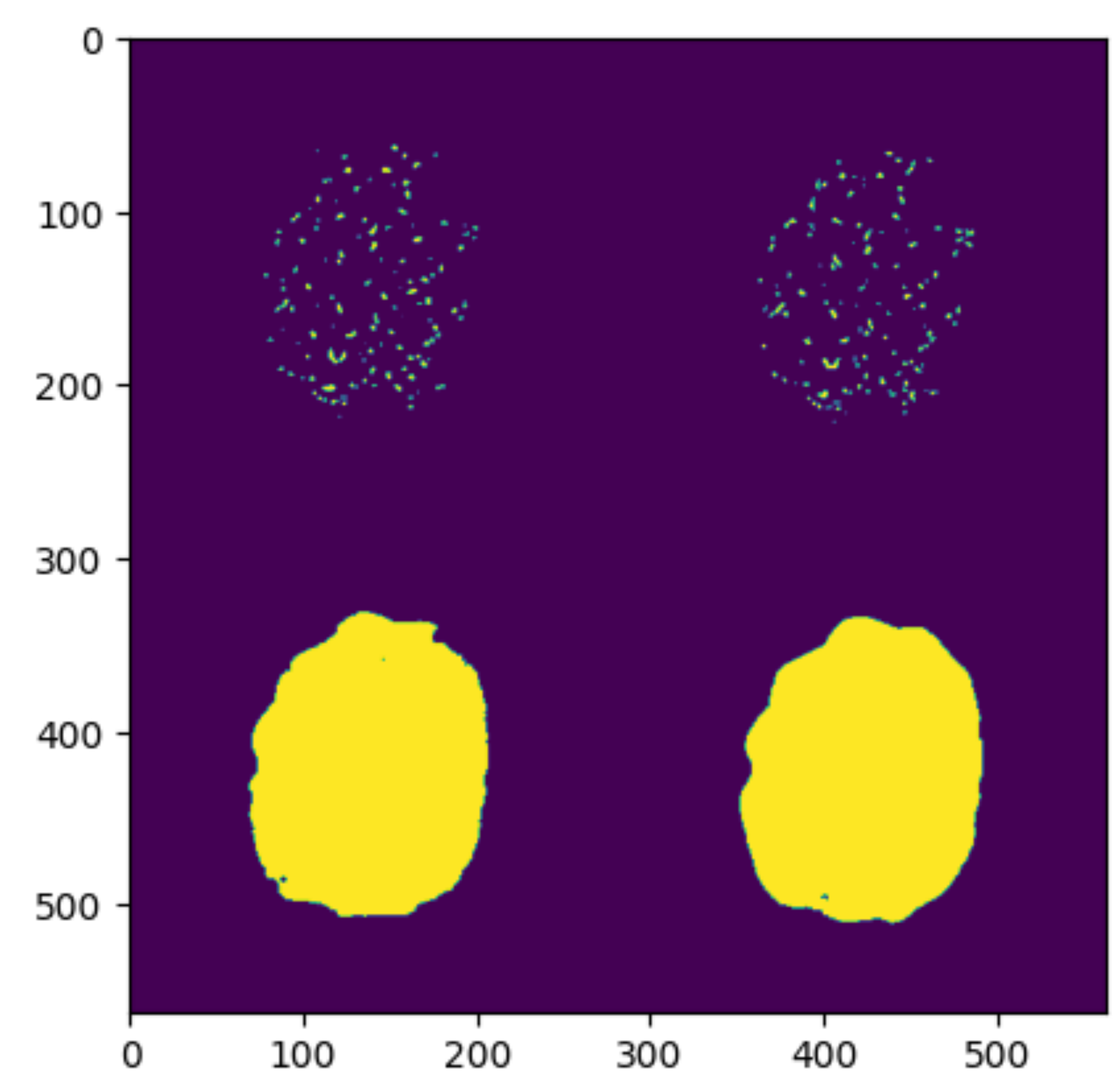


WINDOWZ segmentation pipeline diagram

The user must configure WINDOWZ before use. We have created a robust manual to guide the user through configuration. Once configured, WINDOWZ requires no further user interaction.

1. Load training images from user configured file paths
2. Window training images and load windows into dataloader
3. Begin training model, requesting augmented batches from the dataloader for next epoch while training
4. Save the best iteration of the segmentation model
5. Load input images to be segmented
6. Window the input images
7. Pass the windows through the model to segment
8. Align the segmented windows so they can be neatly vertically stacked, then reconstruct each image in 2D
9. Calculate the interpolant
10. Interpolate points between each segmented image, and stack the image in 3D to create a high-resolution point cloud
11. Save interpolant and high-resolution point cloud

WINDOWZ performance improvements



WINDOWZ 2D segmentation example | Left: prediction, right: ground truth

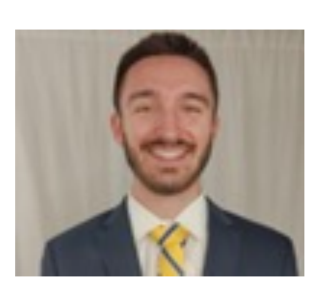
By augmenting the images as they are needed, WINDOWZ allows for extreme variation between training batches. WINDOWZ also utilizes parallel CPU and GPU operations to increase overall efficiency and reduce runtime. Furthermore, WINDOWZ's approach to configuration fosters an accessible user experience and streamlined hyperparameter finetuning.

Discussion

WINDOWZ is an accessible image segmentation and reconstruction pipeline that reduces the barrier to entry of convolution neural networks. The tools it provides enable 3D anatomical models and will have significant impacts across numerous applications including tissue morphometry, morphodynamics, and spatial pharmacology. Furthermore, our pipeline makes CNNs accessible to areas where datasets are limited, allowing more users to participate in similar studies. WINDOWZ is also much faster than conventional convolutional neural network architectures because of its windowed approach. This encourages rapid experimentation and precise finetuning. WINDOWZ makes convolutional neural networks for image segmentation more accessible, more efficient, and more accurate.

Acknowledgments

Funding for this research was provided in part by the Delaware INBRE program, the University of Delaware Graduate College through the Unidel Distinguished Graduate Scholar Award, and the National Institutes of Health R21AI157889.


jbeans@udel.edu
lhallee@udel.edu
gleghorn@udel.edu