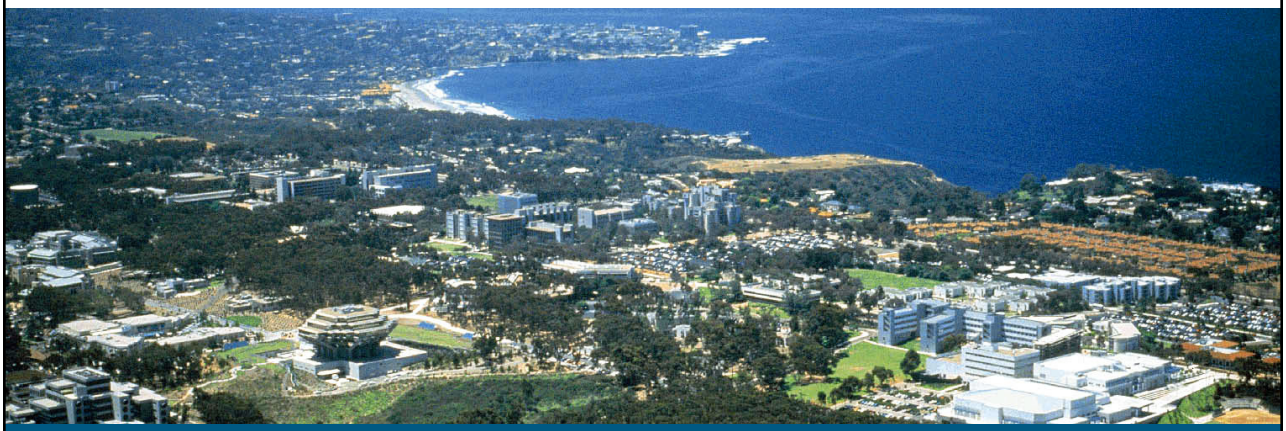
Segmenting Large Electron Microscopic Image Volumes

An Introduction to NBCR image analysis and segmentation tools



National Biomedical Computation Resource Summer Training Program @ UC San Diego

Chris Churas & Matthias Haberl (original instructions created by Alex Perez P.H.D)
August 7th, 2017
churas@ncmir.ucsd.edu & mhaberl@ucsd.edu

Hands-On Session 1

The Manual and Automated Segmentation of Organelles in 3D EM Data

Goals

By the end of this session, you will be able to:

1. Use IMOD program to get information about SBEM image stack
2. Generate a training dataset from training data with CHMutil/IMOD that can be run by CHM

1. Download and Unzip the Dataset

A. Download the ZIP file entitled *download_dataset_1.zip* from: https://github.com/CRBS/nbcrtrainingvm/blob/master/download_dataset_1.zip

B. Unzip the contents to an easily accessible location:

i. For Windows 7/8: Unzip the file into a new folder in your home directory. If you used the IMOD cygwin install, the home directory will be C:\Users\<username> (for windows 10 it may be C:\cygwin\home\<username>)

ii. For Mac: Unzip the file into a new folder (e.g. /Users/<username>/download_dataset_1).

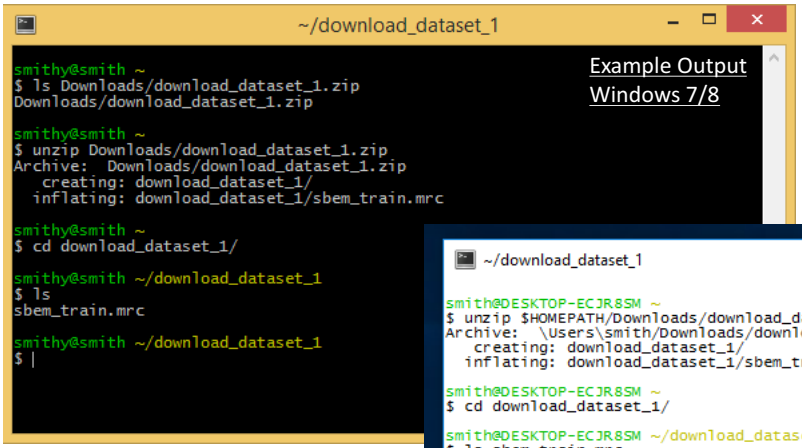
iii. For Linux: Unzip the file into a new folder (e.g. /home/<username>/download_dataset_1).

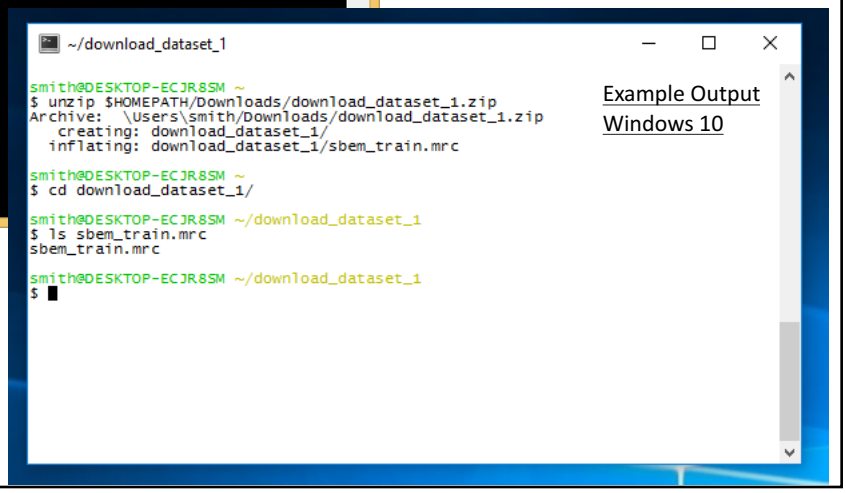
C. Open your terminal program

D. Enter the directory to which you unzipped the dataset by using the *ls* command in the terminal:

cd ~/download_dataset_1

E. List the directory's contents using the *ls* command. You should see one file, *sbem_train.mrc*





These screen shots show use of *unzip* command from terminal to unzip the dataset file on Windows 7/8/10

2. Determine the Properties of the Dataset

A. At the command line, type *header sbem_train.mrc*

- The program *header* is part of the IMOD software suite*, and reads pertinent details from MRC files. The output should contain the following:

```
Number of columns, rows, sections .....   500   500   2
Map mode ..............................   0  (bytes - signed in file)
Start cols, rows, sects, grid x,y,z ...   0   0   0   500   500
2
Pixel spacing (Angstroms).............   77.98    77.98    300.0
Cell angles ...........................   90.000  90.000  90.000
Fast, medium, slow axes ...............   X  Y  Z
Origin on x,y,z .......................   0.000    0.000    0.000
Minimum density .......................   119.00   ( -11.000   in file)
Maximum density .......................   180.00   ( 54.000   in file)
Mean density ..........................   152.41   ( 24.446   in file)
tilt angles (original,current) ........   0.0  0.0  0.0  0.0  0.0  0.0
Space group,# extra bytes,idtype,lens .   0    0    0    0
```

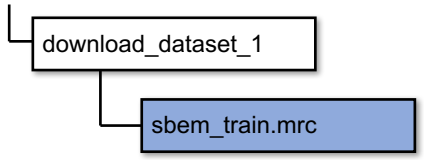
Voxel dimensions of the file, in (X Y Z).

0 = 8-bit, unsigned integers
1 = 16-bit, signed integers
2 = 32-bit, floating point
6 = 16-bit, unsigned integers

Voxel sizes, in Angstroms per voxel, in (X Y Z). This dataset has anisotropic voxels, with 7.798 nm/voxel in X and Y, and 30 nm/voxel in Z. Anisotropic voxels are common in SBEM datasets.

* For an exhaustive list of IMOD programs and help documents, visit: http://bio3d.colorado.edu/imod/doc/program_listing.html

Directory Structure:



~/download_dataset_1

Example Output

```
smithy@smith ~/download_dataset_1
$ header sbem_train.mrc

 RO image file on unit   1 : sbem_train.mrc     Size=        489 K

 Number of columns, rows, sections .....     500     500       2
 Map mode ..............................     0    (byte)
 Start cols, rows, sects, grid x,y,z ...     0       0       0     500     500       2
 Pixel spacing (Angstroms)..............    77.98       77.98       300.0
 Cell angles ...........................    90.000    90.000    90.000
 Fast, medium, slow axes ...............     X     Y     Z
 Origin on x,y,z .......................    -999.0        0.000       -14.00
 Minimum density .......................    119.00
 Maximum density .......................    180.00
 Mean density ..........................    152.41
 tilt angles (original,current) ........    0.0   0.0   0.0   0.0   0.0   0.0
 Space group,# extra bytes,idtype,lens .        0         0         0         0

    10 Titles :
raw2mrc: Converted to mrc format.                          19-Feb-12  12:22:31
NEWSTACK: Images copied                                    21-Feb-12  11:37:20
NEWSTACK: Images copied                                    21-Feb-12  12:19:51
NEWSTACK: Images copied              , floated to range    22-Feb-12  08:54:59
NEWSTACK: Images copied                                    24-Feb-12  12:04:59
NEWSTACK: Images copied, transformed, floated to means     28-Feb-12  22:16:28
BOXSTARTEND:  10 starts clipped out and averaged            3-Aug-15  13:50:47
NEWSTACK: Images copied                                     3-Aug-15  13:52:31
NEWSTACK: Images copied                                     3-Aug-15  13:52:50
NEWSTACK: Images copied                                    14-Jul-17  08:47:41


smithy@smith ~/download_dataset_1
$
```

3. Create IMOD Model File for Training Labels

A. Open the training data stack in 3dmod: *3dmod sbem_train.mrc*

B. In this course, we will be generating training data for mitochondria. Using the techniques described in the preceding lecture, manually segment all instances of mitochondria in the two tiles of your training data stack.

As an additional reference, you can view the file:

https://github.com/CRBS/nbcrtrainingvm/blob/master/Generating_CHM_training_data_with_IMOD.pdf

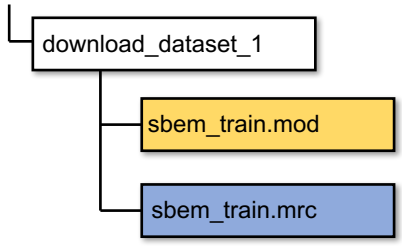
Or visit the following URLs:

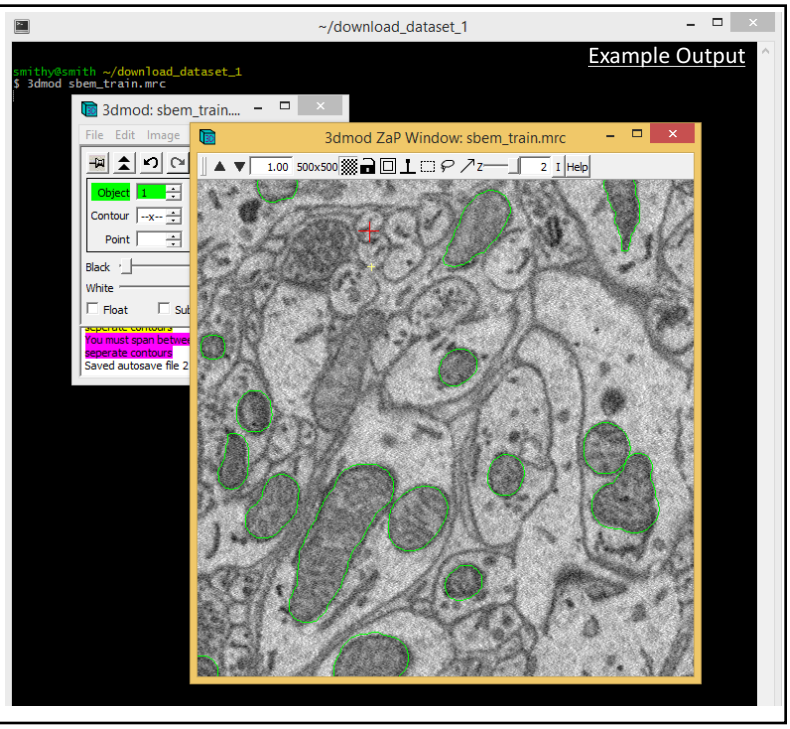
http://bio3d.colorado.edu/imod/doc/3dmodguide.html – Detailed information about manual segmentation and model file structure in IMOD.

https://www.youtube.com/watch?v=BsNSVLIQ-cE – Useful video illustrating the use of IMOD's Drawing Tools

C. Save the model file as *sbem_train.mod*

Directory Structure:





4. Generate a Training Label Stack

The goal of this step is to generate a new MRC stack with the same dimensions as the training images we previously segmented. However, in this stack, all pixels inside of the traced contours will have values of one, and all pixels outside of the traced contours will have values of zero. In this way, a binary label stack will be created that will serve to tell the CHM training algorithm where mitochondria are. We will use the IMOD program *imodmop* to generate this stack.

A. Create the binary label stack using the following command:

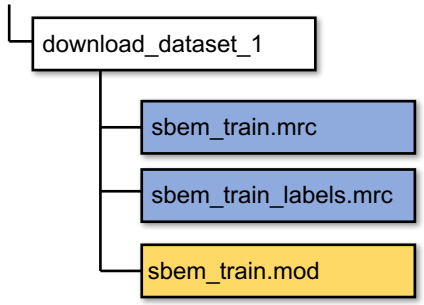
imodmop –mask 1 sbem_train.mod sbem_train.mrc sbem_train_labels.mrc

To understand the meaning of these arguments and the *imodmop* syntax, view the imodmop man page in the terminal by typing *man imodmop* (to exit the man page type '*q*' without the quotes)

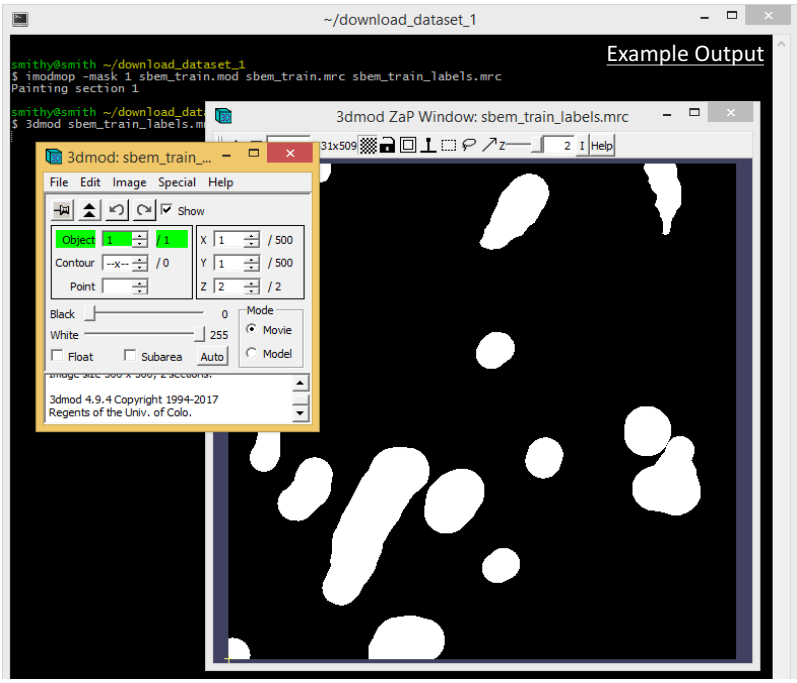
B. Visualize the label stack using 3dmod:

3dmod sbem_train_labels.mrc

Directory Structure:



Example Output



5. Generate Training Image and Label PNGs

The CHM algorithm requires inputs in the form of PNG files. Thus we must first convert from MRC to PNG.

A. First, we need to make sub-directories under the *train* folder for images and labels (The '*-p*' argument for *mkdir* will automatically create any missing parent directories):

mkdir –p train/images
mkdir train/labels

B. Next, convert the MRC stack of raw training images to individually numbered PNGs using the IMOD program *mrc2tif*. The '-p' argument forces conversion to PNG, rather than TIF and the 'x' at end is needed by *mrc2tif* as a filename prefix:

mrc2tif –p sbem_train.mrc train/images/x

C. Convert the MRC stack of training labels to individually numbered PNGs using the IMOD program *mrc2tif:*

mrc2tif –p sbem_train_labels.mrc train/labels/x

Directory Structure: