



# PROOFREADING OF AUTOMATIC SEGMENTATIONS IN CONNECTOMICS

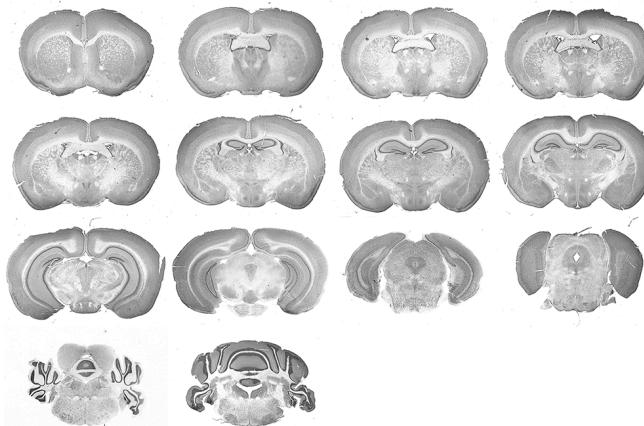
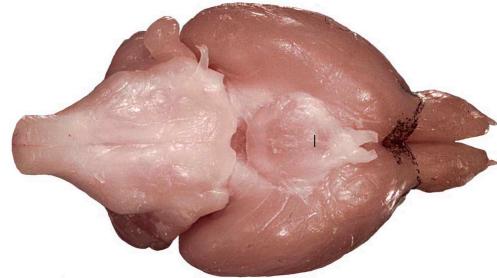
Daniel Haehn

Advisor: Hanspeter Pfister

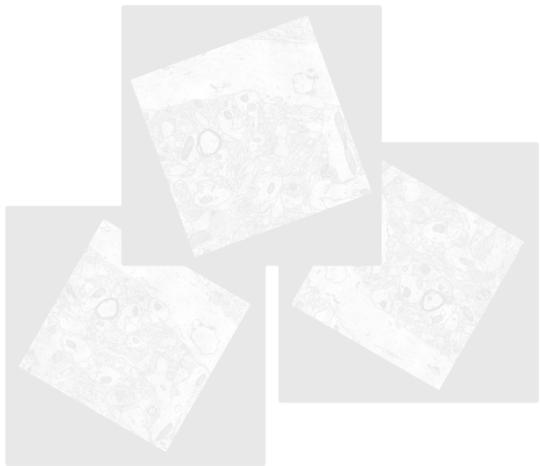
Image by Lichtman lab

# CONNECTOMICS

Goal: Fully understand the wiring diagram of the brain



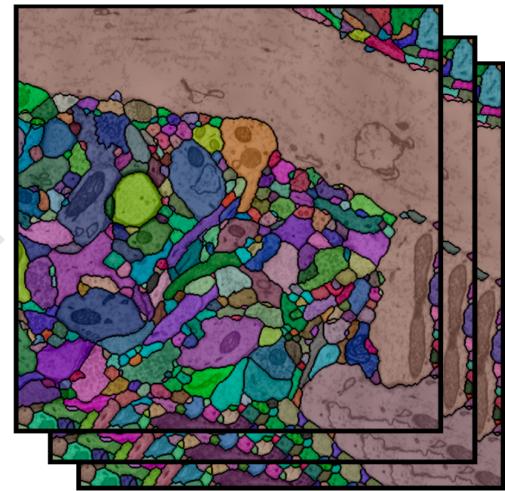
## Acquisition



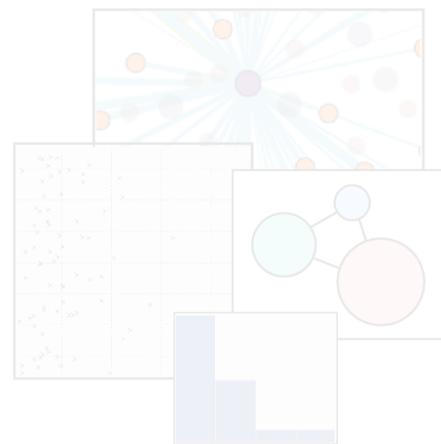
## Registration



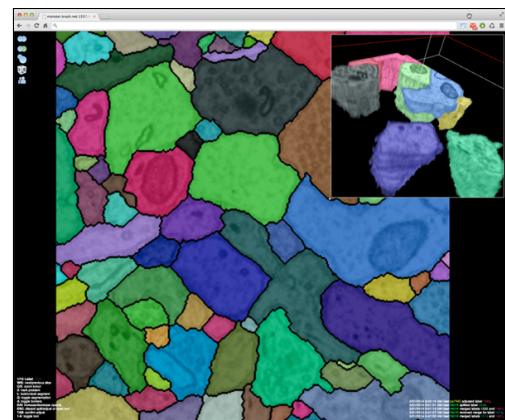
## Labeling



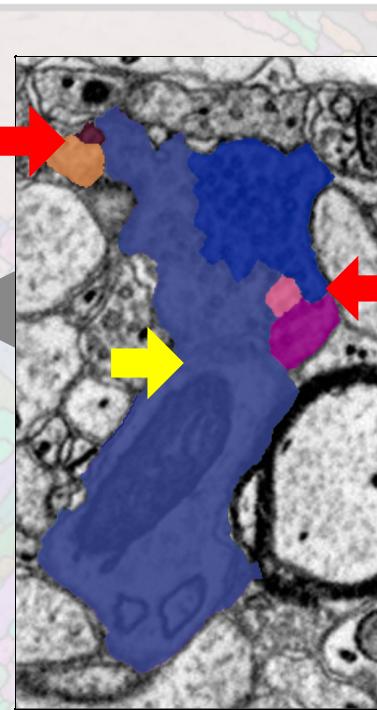
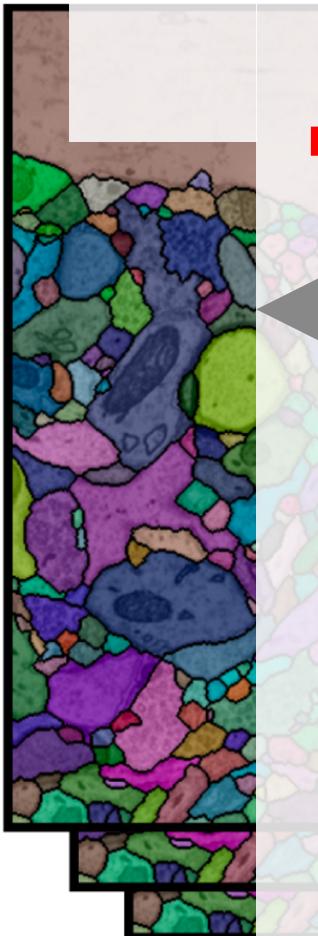
## Analysis



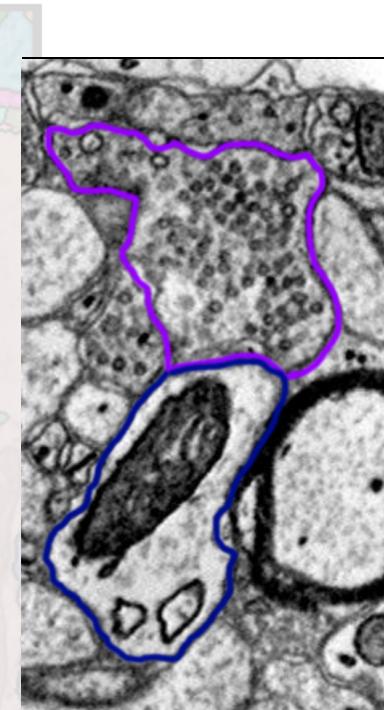
## Proofreading



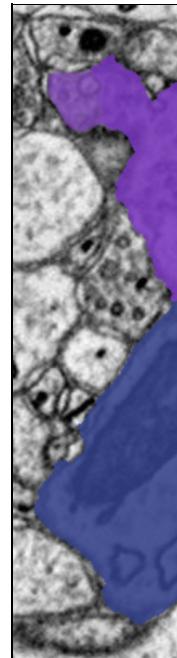
# LABELING ERRORS



Automatic Labeling



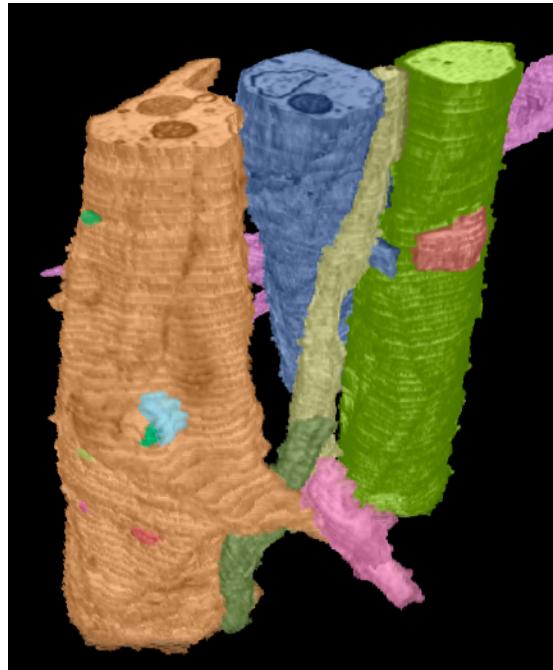
Original Image  
with cell boundaries



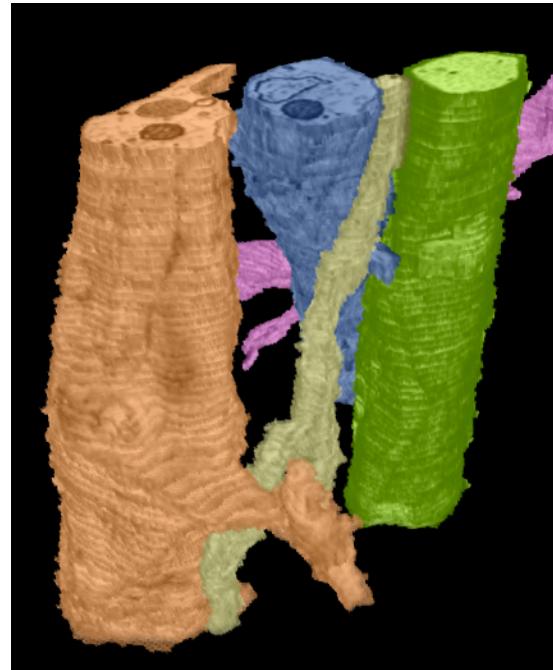
Corrected

# PROOFREADING

"Manual Correction of Automatic Labeling"



Before

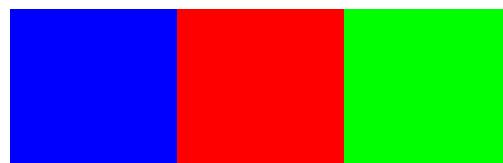


After

# PROOFREADING OPERATORS

MERGE

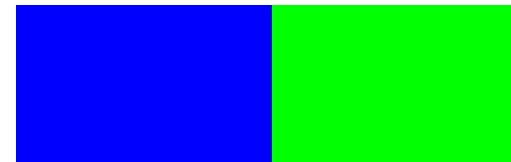
Before



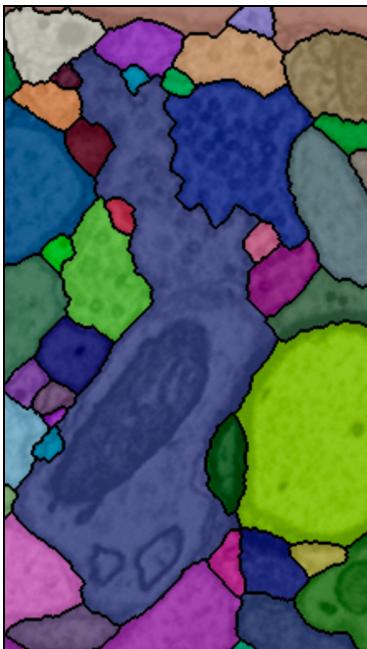
After



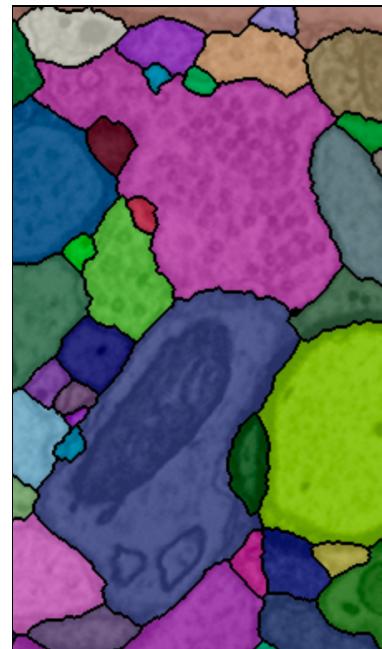
SPLIT



# PROOFREADING RESULT



Before Proofreading

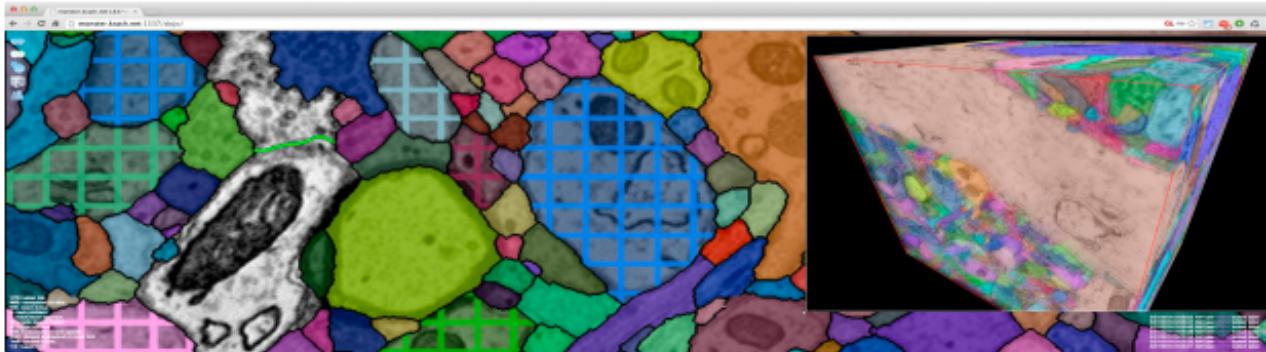


After Proofreading



# Design and Evaluation of Interactive Proofreading Tools for Connectomics

Daniel Haehn, Seymour Knowles-Barley, Mike Roberts,  
Johanna Beyer, Narayanan Kasthuri, Jeff W. Lichtman, and Hanspeter Pfister



**Fig. 1: Proofreading with Dojo.** We present a web-based application for interactive proofreading of automatic segmentations of connectome data acquired via electron microscopy. Split, merge and adjust functionality enables multiple users to correct the labeling of neurons in a collaborative fashion. Color-coded structures can be explored in 2D and 3D.

**Abstract**—Proofreading refers to the manual correction of automatic segmentations of image data. In connectomics, electron microscopy data is acquired at nanometer-scale resolution and results in very large image volumes of brain tissue that require fully automatic segmentation algorithms to identify cell boundaries. However, these algorithms require hundreds of corrections per cubic micron of tissue. Even though this task is time consuming, it is fairly easy for humans to perform corrections through splitting, merging, and adjusting segments during proofreading. In this paper we present the design and implementation of *Mojo*, a fully-featured single-user desktop application for proofreading, and *Dojo*, a multi-user web-based application for collaborative proofreading. We evaluate the accuracy and speed of *Mojo*, *Dojo*, and *Raveler*, a proofreading tool from Janelia Farm, through a quantitative user study. We designed a between-subjects experiment and asked non-experts to proofread neurons in a publicly available connectomics dataset. Our results show a significant improvement of corrections using web-based *Dojo*, when given the same amount of time. In addition, all participants using *Dojo* reported better usability. We discuss our findings and provide an analysis of requirements for designing visual proofreading software.

D. Haehn, et al., “Design and Evaluation of Interactive Proofreading Tools for Connectomics,” IEEE Transactions on Visualization and Computer Graphics, vol. 20, no. 12, pp. 2466-2475, 2014.

## NeuroBlocks - Visual Tracking of Segmentation and Proofreading for Large Connectomics Projects

A. K. Al-Awami, J. Beyer, **D. Haehn**, N. Kasthuri, J. W. Lichtman, H. Pfister, M. Hadwiger

IEEE Transactions on Visualization and Computer Graphics, vol. 22, no. 1, pp. 738-746, 2016.

## Automatic Neural Reconstruction from Petavoxel of Electron Microscopy Data

A. Suissa-Peleg, **D. Haehn**, S. Knowles-Barley, V. Kaynig, T. R. Jones, A. Wilson, R. Schalek, J. W. Lichtman, H. Pfister

Microscopy and Microanalysis, vol. 22, S3, pp. 536-537, 2016.

## Imaging a 1 mm<sup>3</sup> Volume of Rat Cortex Using a MultiBeam SEM

R. Schalek, D. Lee, N. Kasthuri, A. Peleg, T. R. Jones, V. Kaynig, **D. Haehn**, H. Pfister, D. Cox, J. W. Lichtman

Microscopy and Microanalysis, vol. 22, S3, pp. 582-583, 2016.

# Guided Proofreading of Automatic Segmentations in Connectomics

Daniel Haehn<sup>1</sup> Verena Kaynig<sup>1,2</sup> James Tompkin<sup>1</sup> Jeff W. Lichtman<sup>2</sup> Hanspeter Pfister<sup>1</sup>

<sup>1</sup>Harvard Paulson School of Engineering and Applied Sciences

<sup>2</sup>Harvard Center for Brain Science  
Cambridge, MA 02138, USA

haehn@seas.harvard.edu

## Abstract

Automatic cell image segmentation methods in connectomics can lead to split and merge errors, which require correction through proofreading. To aid error correction, we develop two classifiers that are able to recommend candidate errors and their corrections to the user. These classifiers are informed by training a convolutional neural network with known errors in automatic segmentations by comparison to expert-labeled ground truth. Our network architecture is able to detect potentially erroneous regions by considering a large context region around a segmentation boundary. With recommendations, proofreading of mouse cortex electron microscopy image segmentations can reduce variation of information scores on two different datasets from 0.4764 to 0.3996 and from 0.4847 to 0.3946, which we find is an improvement on pure manual and pure automatic cases.

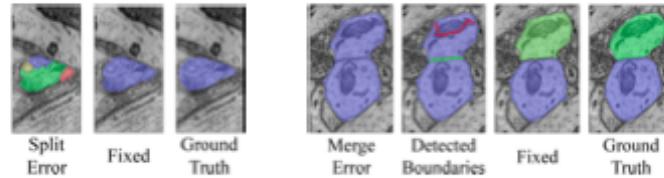
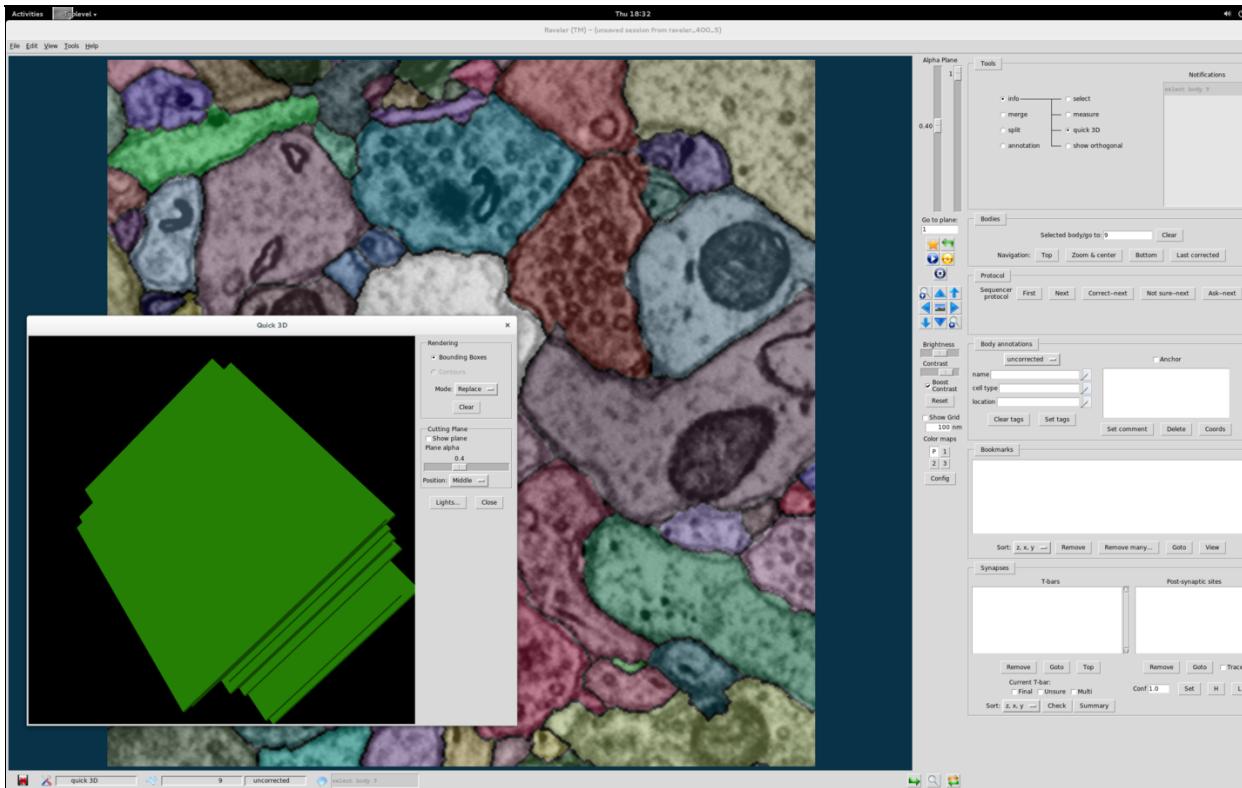


Figure 1: Split and merge error examples, their corrections, and their ground truths.

tion to be grouped into geometrically-consistent cells across registered sections, or cells are segmented across registered sections in 3D directly. Using dynamic programming techniques [22] and a GPU cluster, these classifiers can segment  $\approx 1$  terabyte of data per hour [16], which is sufficient to keep up with the 2D data capture process on state-of-the-art electron microscopes (though 3D registration is still an expensive offline operation).

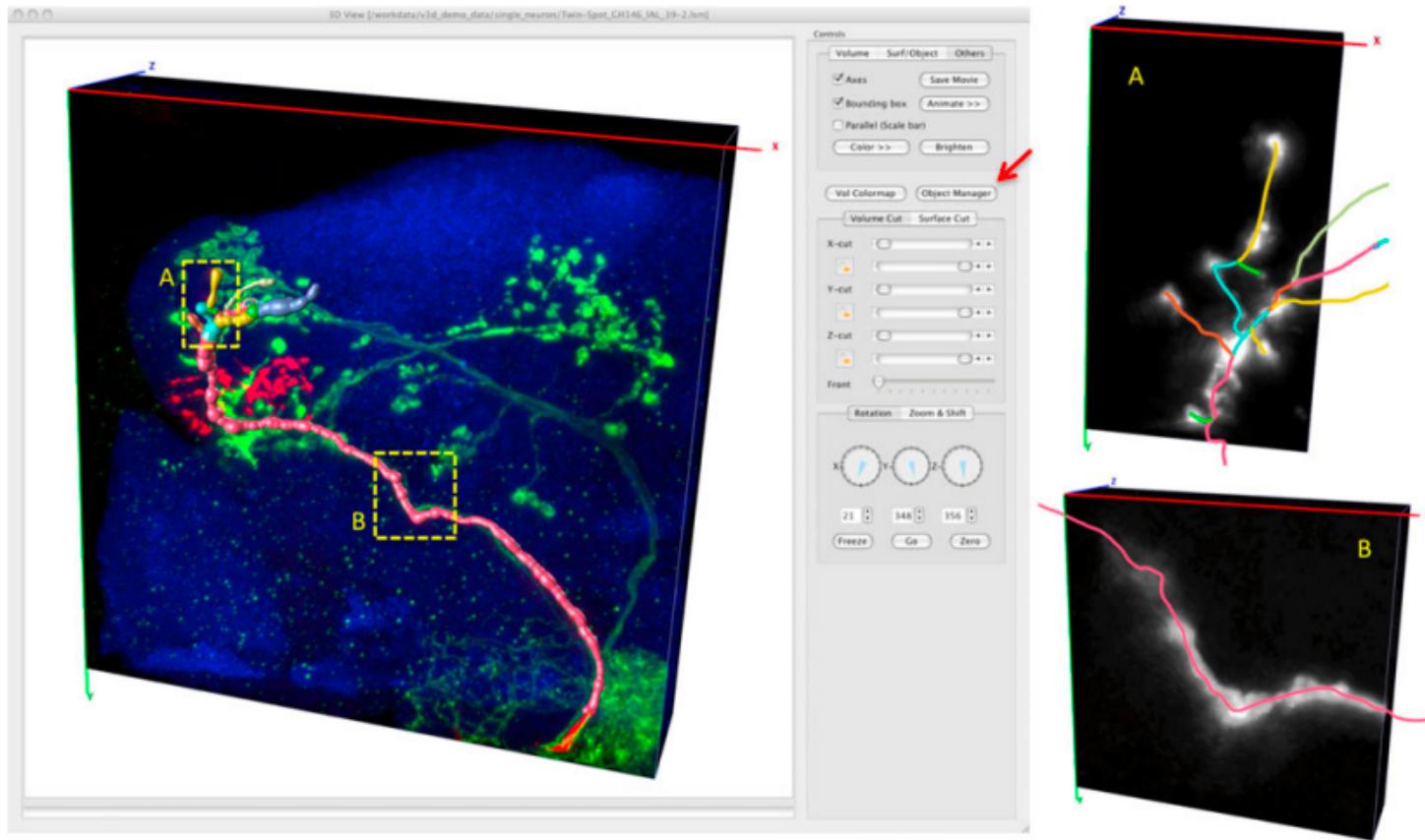
All automatic methods make errors, and we are left with

# RELATED WORK



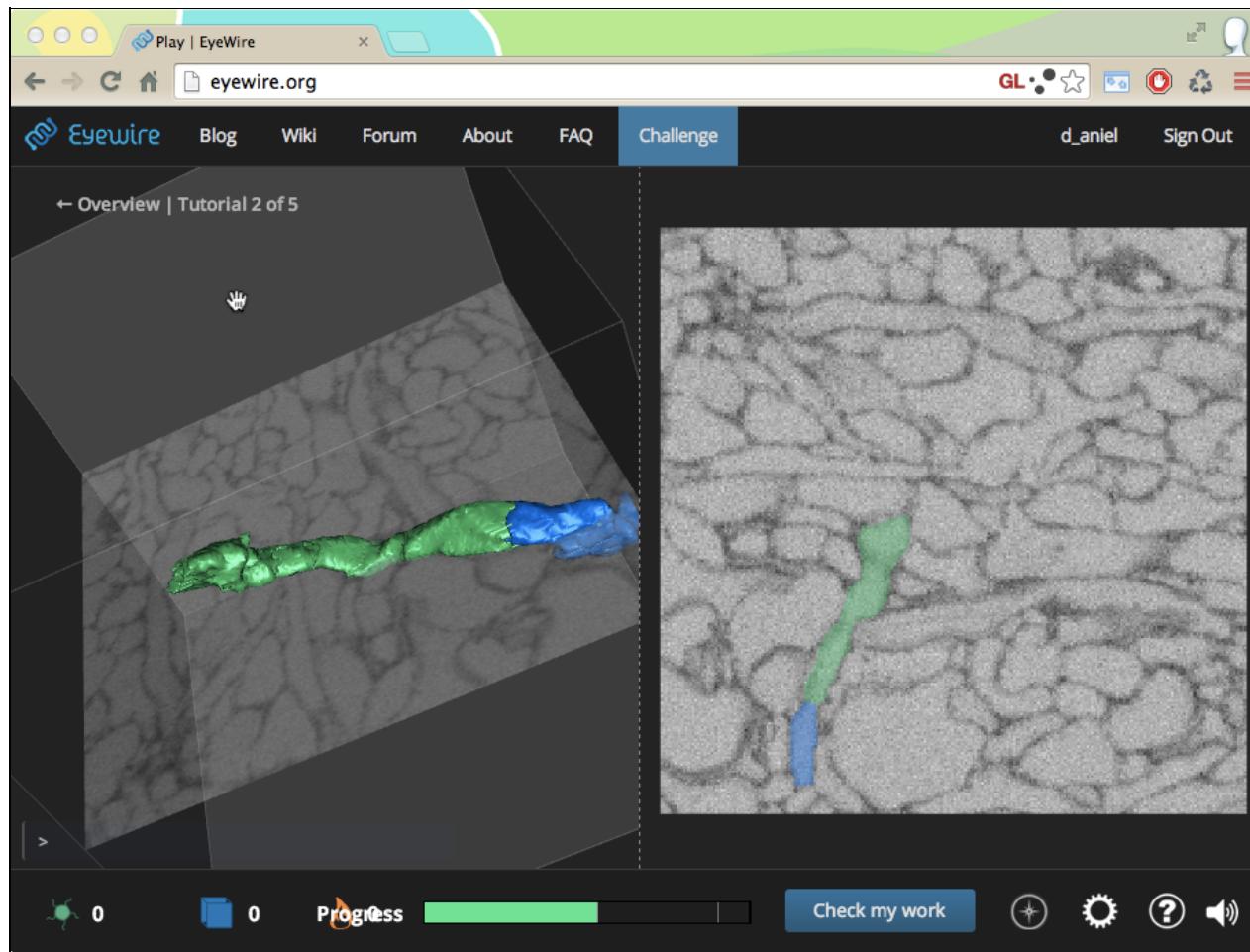
Chklovskii et al.: Semi-automated reconstruction of neural circuits using electron microscopy,  
Current Opinion in Neurobiology, 2010.

# RELATED WORK (2)



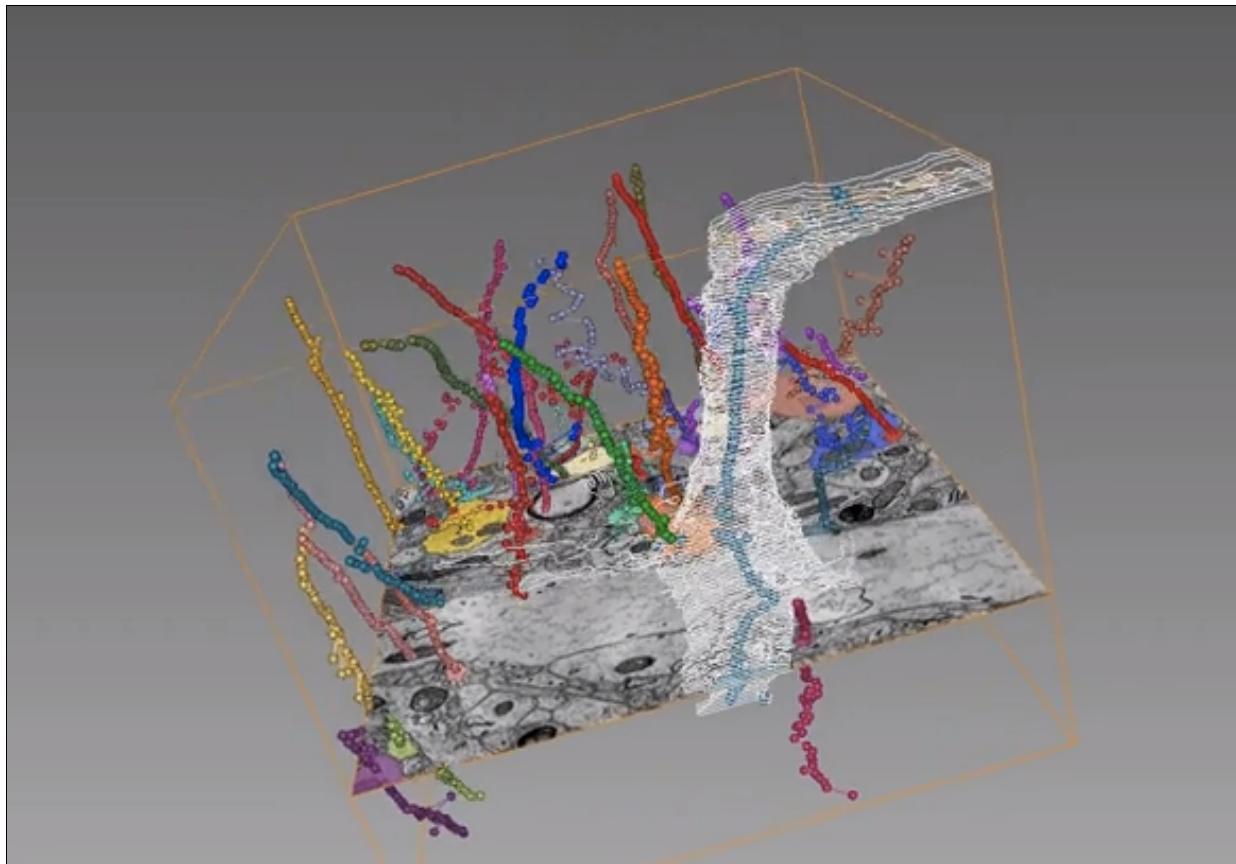
Peng et al.: Proof-editing is the Bottleneck Of 3D Neuron Reconstruction: The Problem and Solutions,  
Neuroinformatics, 2011.

# RELATED WORK (3)



Seung et al.: EyeWire, <http://eyewire.org>, 2012.

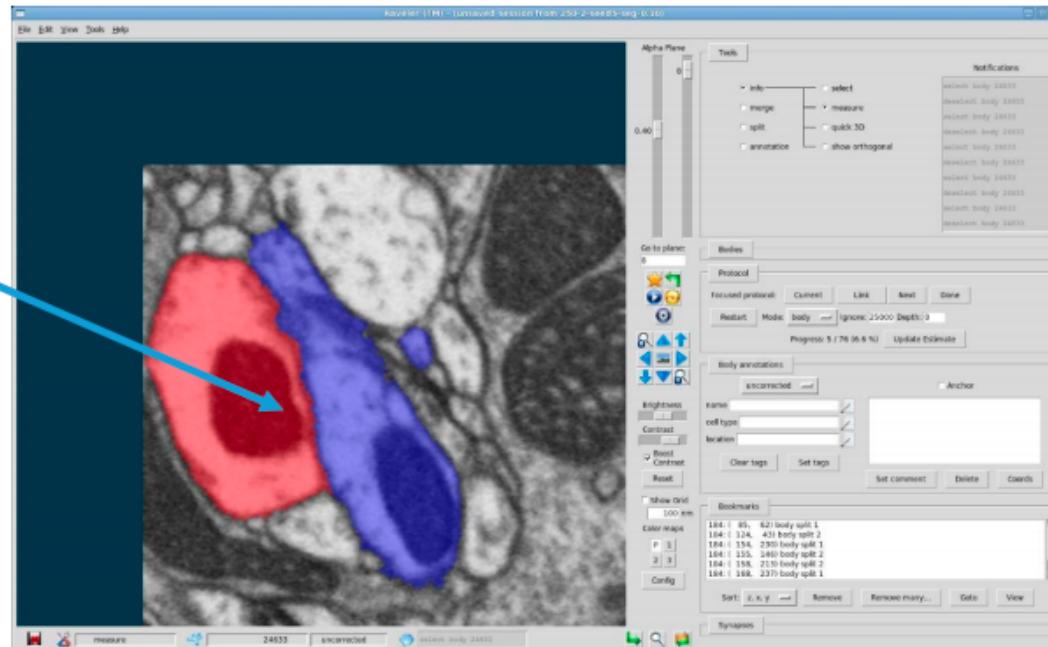
# RELATED WORK (4)



Sicat et al.: Graph abstraction for simplified proofreading of slice-based volume segmentation,  
EUROGRAPHICS, 2013.

# RELATED WORK (5)

focused protocol  
(simple yes/no  
decision)



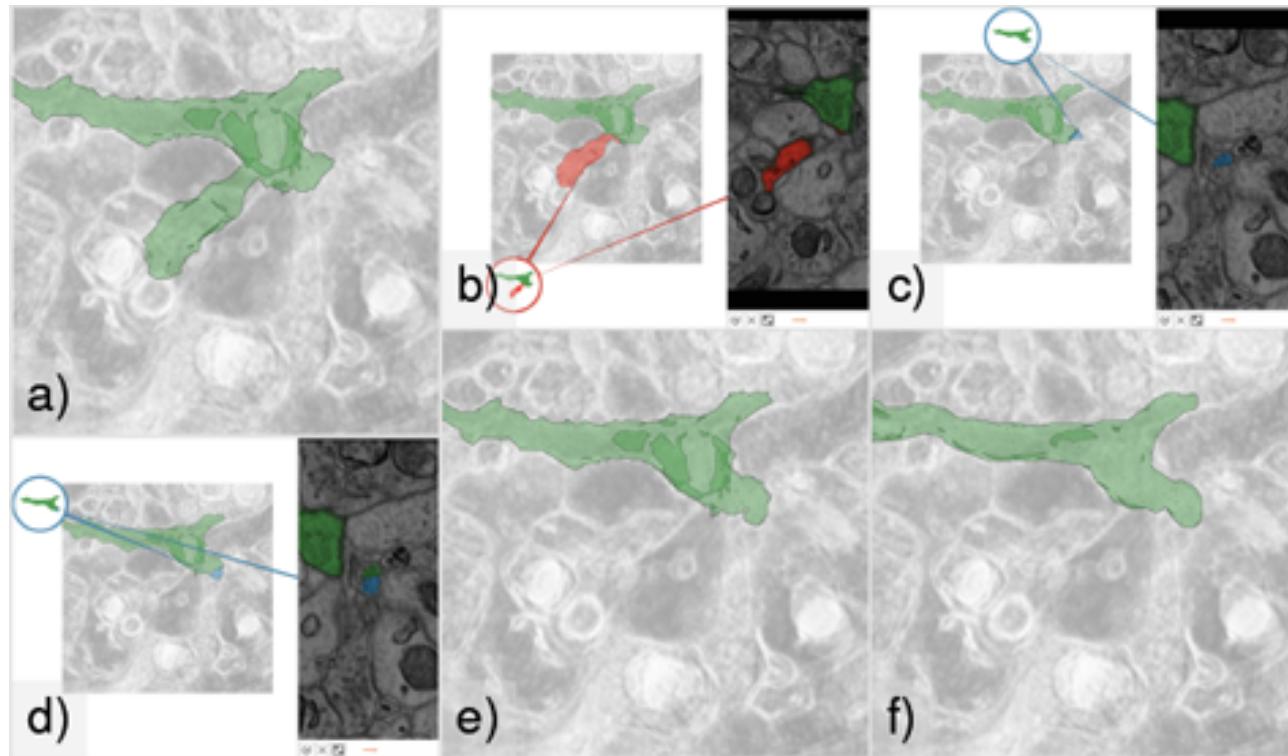
Plaza et al.: Focused Proofreading: Efficiently Extracting Connectomes from Segmented EM Images,  
CoRR, 2014.

# RELATED WORK (6)



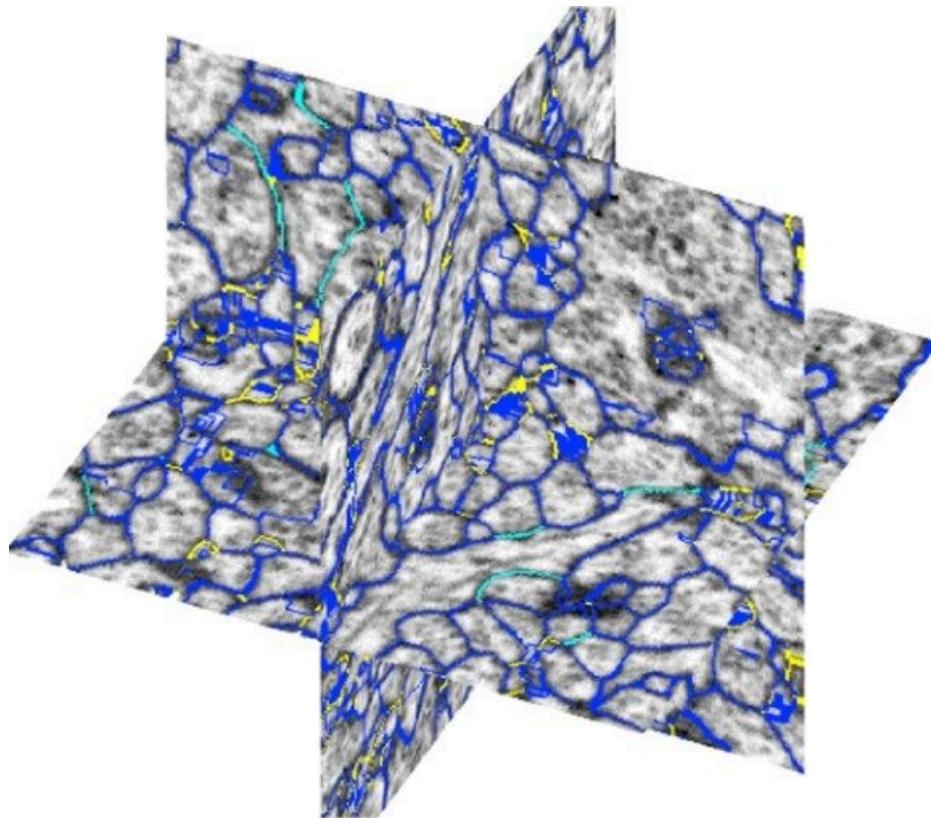
Al-Awami et al.: NeuroBlocks - Visual Tracking of Segmentation and Proofreading for Large Connectomics Projects, IEEE SciVis, 2015.

# RELATED WORK (7)



Karimov et al.: Guided Volume Editing based on Histogram Dissimilarity,  
Computer Graphics Forum, 2015.

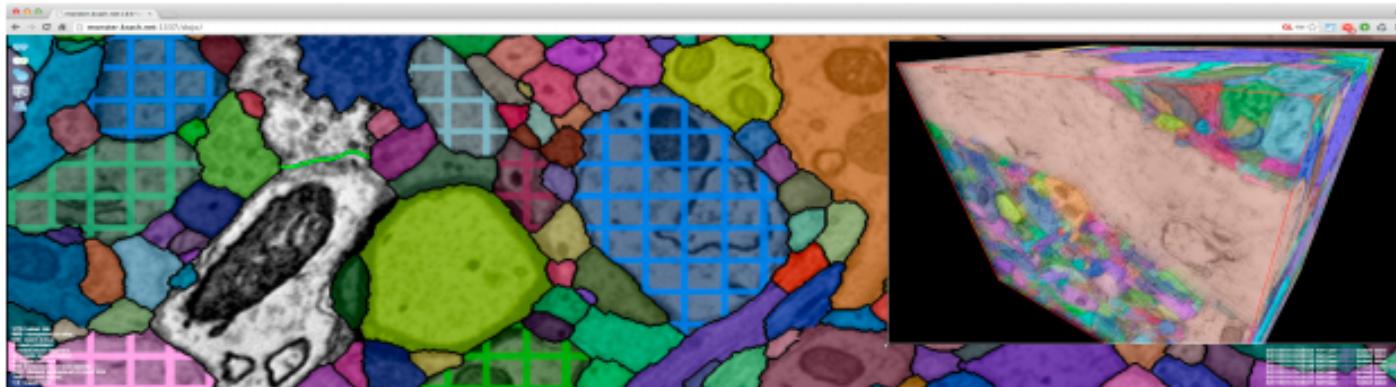
# RELATED WORK (8)



Uzunbas et al.: An efficient conditional random field approach for automatic and interactive neuron segmentation, Medical Image Analysis, 2016.

# Design and Evaluation of Interactive Proofreading Tools for Connectomics

Daniel Haehn, Seymour Knowles-Barley, Mike Roberts,  
Johanna Beyer, Narayanan Kasthuri, Jeff W. Lichtman, and Hanspeter Pfister



**Fig. 1: Proofreading with Dojo.** We present a web-based application for interactive proofreading of automatic segmentations of connectome data acquired via electron microscopy. Split, merge and adjust functionality enables multiple users to correct the labeling of neurons in a collaborative fashion. Color-coded structures can be explored in 2D and 3D.

**Abstract**—Proofreading refers to the manual correction of automatic segmentations of image data. In connectomics, electron microscopy data is acquired at nanometer-scale resolution and results in very large image volumes of brain tissue that require fully automatic segmentation algorithms to identify cell boundaries. However, these algorithms require hundreds of corrections per cubic micron of tissue. Even though this task is time consuming, it is fairly easy for humans to perform corrections through splitting, merging, and adjusting segments during proofreading. In this paper we present the design and implementation of *Mojo*, a fully-featured single-user desktop application for proofreading, and *Dojo*, a multi-user web-based application for collaborative proofreading. We evaluate the accuracy and speed of *Mojo*, *Dojo*, and *Raveler*, a proofreading tool from Janelia Farm, through a quantitative user study. We designed a between-subjects experiment and asked non-experts to proofread neurons in a publicly available connectomics dataset. Our results show a significant improvement of corrections using web-based *Dojo*, when given the same amount of time. In addition, all participants using *Dojo* reported better usability. We discuss our findings and provide an analysis of requirements for designing visual proofreading software.

# MOJO

File Edit View

Adjust Tool Merge Tool Split Tool Join Splits 3D Commit Change Cancel Change Save Segmentation

Segmentation Visibility Border Lines Merge Method Draw Global Replace Split Method Draw Split Line

Mojo v2.0

Location 414,259,0

Selection

Lock Size 0 Id 0

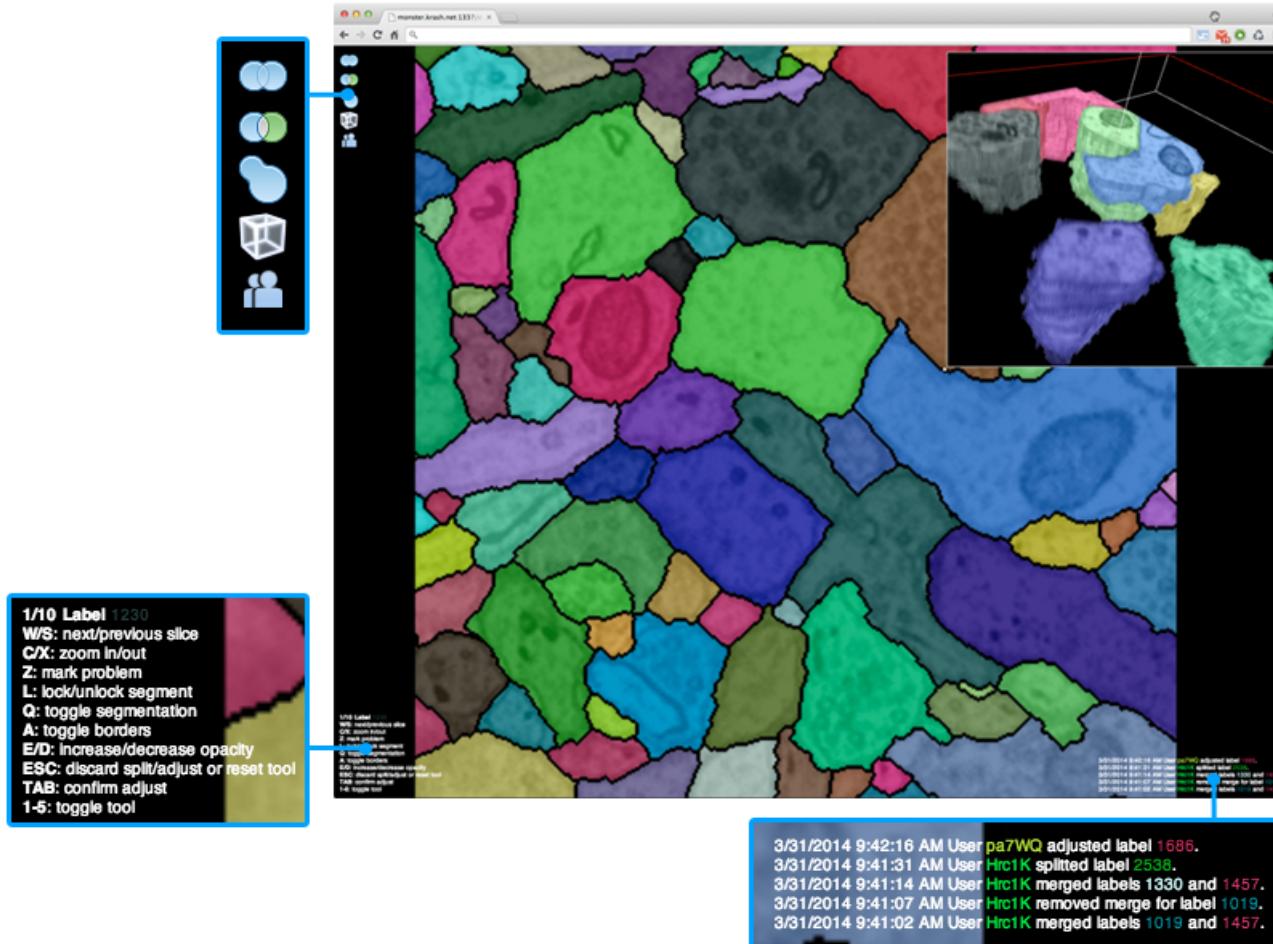
<< < > >> Page 1 of 1

Lock	Name	Size	Id
<input type="checkbox"/>	segment1980	90715	1980
<input type="checkbox"/>	segment273	89656	273
<input type="checkbox"/>	segment379	82790	379
<input type="checkbox"/>	segment583	78337	583
<input type="checkbox"/>	segment34	77362	34
<input type="checkbox"/>	segment1313	74037	1313
<input type="checkbox"/>	segment1230	73050	1230
<input type="checkbox"/>	segment209	71968	209
<input type="checkbox"/>	segment2547	59066	2547
<input type="checkbox"/>	segment1295	55780	1295
<input type="checkbox"/>	segment1795	53644	1795
<input type="checkbox"/>	segment975	49469	975
<input type="checkbox"/>	segment111	48623	111
<input type="checkbox"/>	segment181	36233	181
<input type="checkbox"/>	segment608	36047	608
<input type="checkbox"/>	segment1431	31168	1431
<input type="checkbox"/>	segment1402	28824	1402
<input type="checkbox"/>	segment2111	27994	2111
<input type="checkbox"/>	segment1581	24095	1581
<input type="checkbox"/>	segment2538	24298	2538
<input type="checkbox"/>	segment1488	23252	1488
<input type="checkbox"/>	segment1229	21142	1229
<input type="checkbox"/>	segment972	21024	1273
<input type="checkbox"/>	segment180	20030	180
<input type="checkbox"/>	segment1801	16041	807
<input type="checkbox"/>	segment424	17012	424
<input type="checkbox"/>	segment1457	16058	1457
<input type="checkbox"/>	segment337	16051	337
<input type="checkbox"/>	segment1240	15981	2140

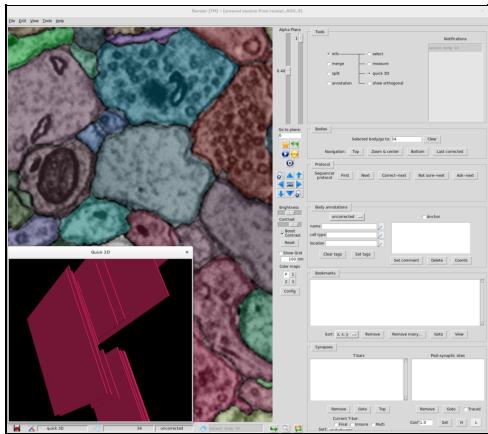
Image 1 of 10 Selection: Target: segment1795 [size=53k]

# **DOJO - LIVE DEMO**

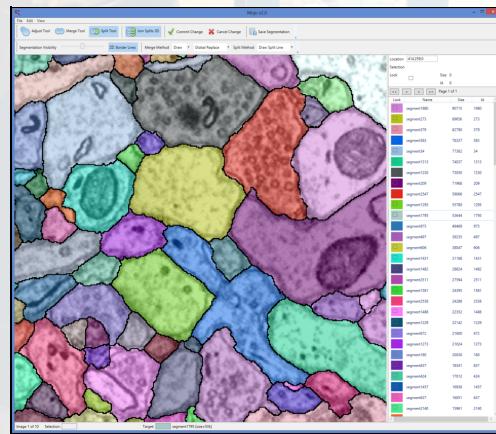
# DOJO (2)



# EXPERIMENT

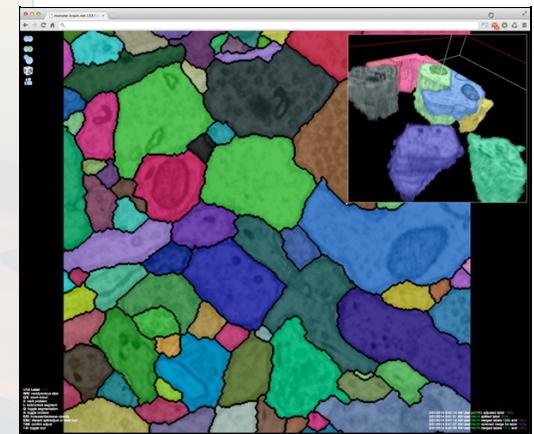


Raveler



VS

Mojo



Dojo

# **EXPERIMENT (2)**

between-subjects experiment with untrained  
participants (N=30)

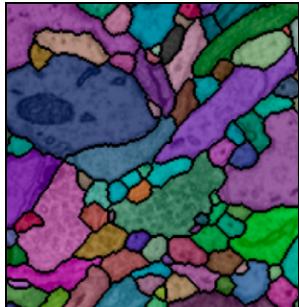
proofread a small dataset

in 30 minutes

also asked 2 experts to segment the dataset  
from scratch

# **EXPERIMENT (3)**

## **DATASET**



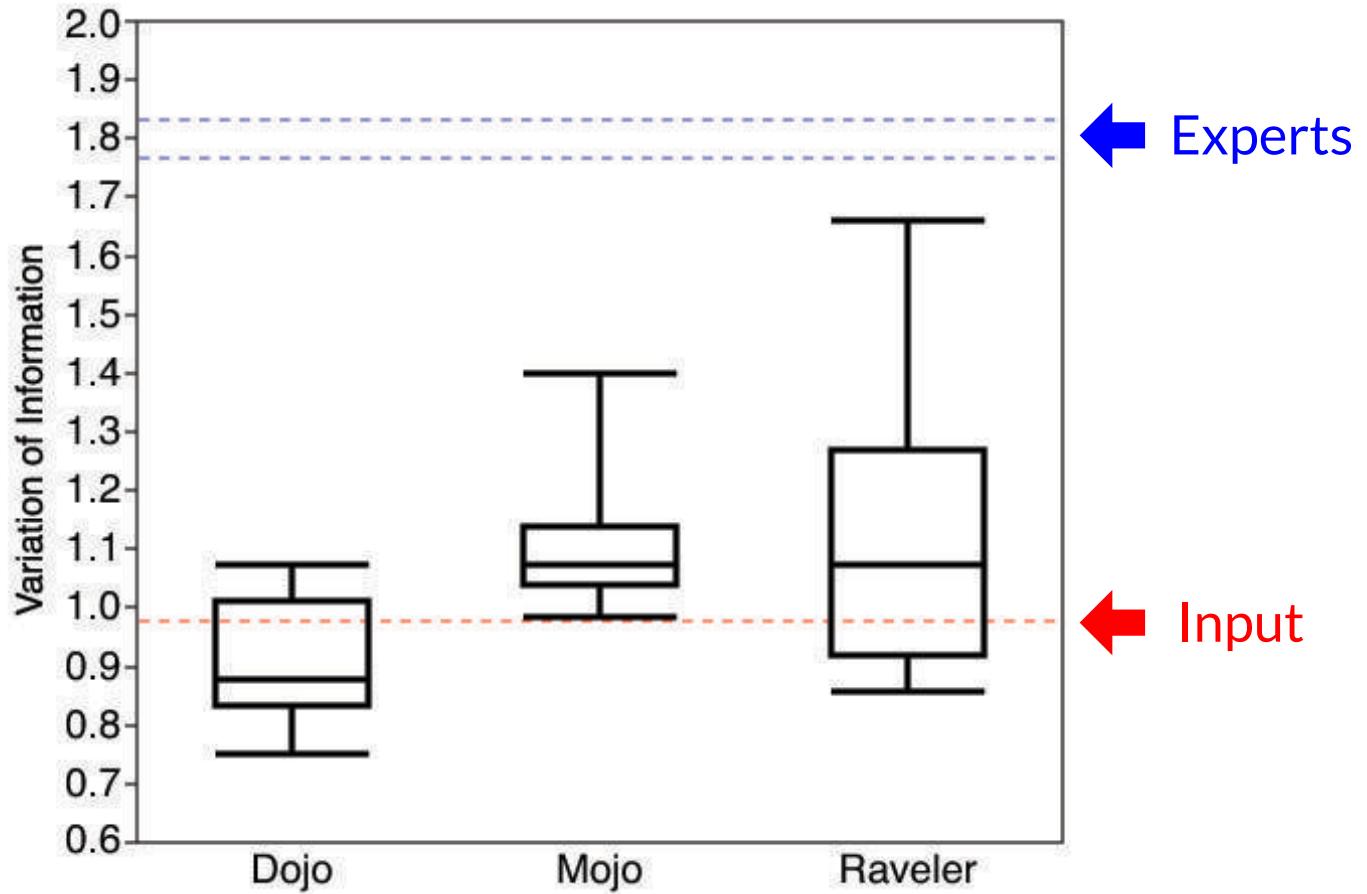
400x400x10 voxels, 6x6x30 nm/voxel

cut from publicly available dataset  
(ISBI 2013 challenge)

with ground-truth available  
automatically segmented

# RESULTS

The lower, the better



The majority of time is spent finding errors.

# Guided Proofreading of Automatic Segmentations in Connectomics

Daniel Haehn<sup>1</sup> Verena Kaynig<sup>1,2</sup> James Tompkin<sup>1</sup> Jeff W. Lichtman<sup>2</sup> Hanspeter Pfister<sup>1</sup>

<sup>1</sup>Harvard Paulson School of Engineering and Applied Sciences

<sup>2</sup>Harvard Center for Brain Science  
Cambridge, MA 02138, USA

haehn@seas.harvard.edu

## Abstract

Automatic cell image segmentation methods in connectomics can lead to split and merge errors, which require correction through proofreading. To aid error correction, we develop two classifiers that are able to recommend candidate errors and their corrections to the user. These classifiers are informed by training a convolutional neural network with known errors in automatic segmentations by comparison to expert-labeled ground truth. Our network architecture is able to detect potentially erroneous regions by considering a large context region around a segmentation boundary. With recommendations, proofreading of mouse cortex electron microscopy image segmentations can reduce variation of information scores on two different datasets from 0.4764 to 0.3996 and from 0.4847 to 0.3946, which we find is an improvement on pure manual and pure automatic cases.

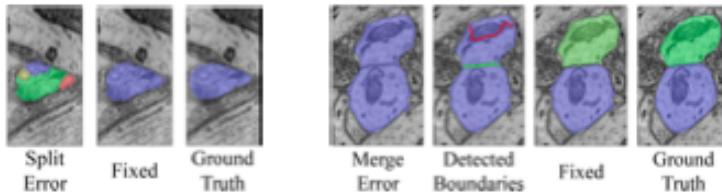
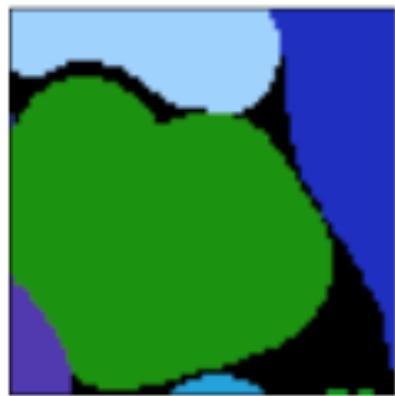


Figure 1: Split and merge error examples, their corrections, and their ground truths.

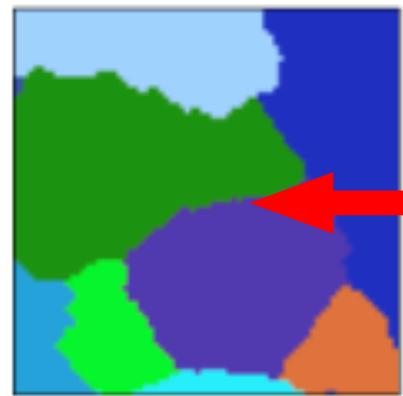
tion to be grouped into geometrically-consistent cells across registered sections, or cells are segmented across registered sections in 3D directly. Using dynamic programming techniques [22] and a GPU cluster, these classifiers can segment  $\approx 1$  terabyte of data per hour [16], which is sufficient to keep up with the 2D data capture process on state-of-the-art electron microscopes (though 3D registration is still an expensive offline operation).

# **GUIDED PROOFREADING - LIVE DEMO**

# SPLIT ERROR DETECTION

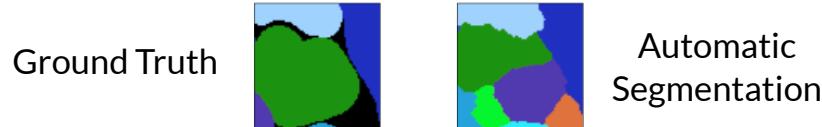


Ground Truth

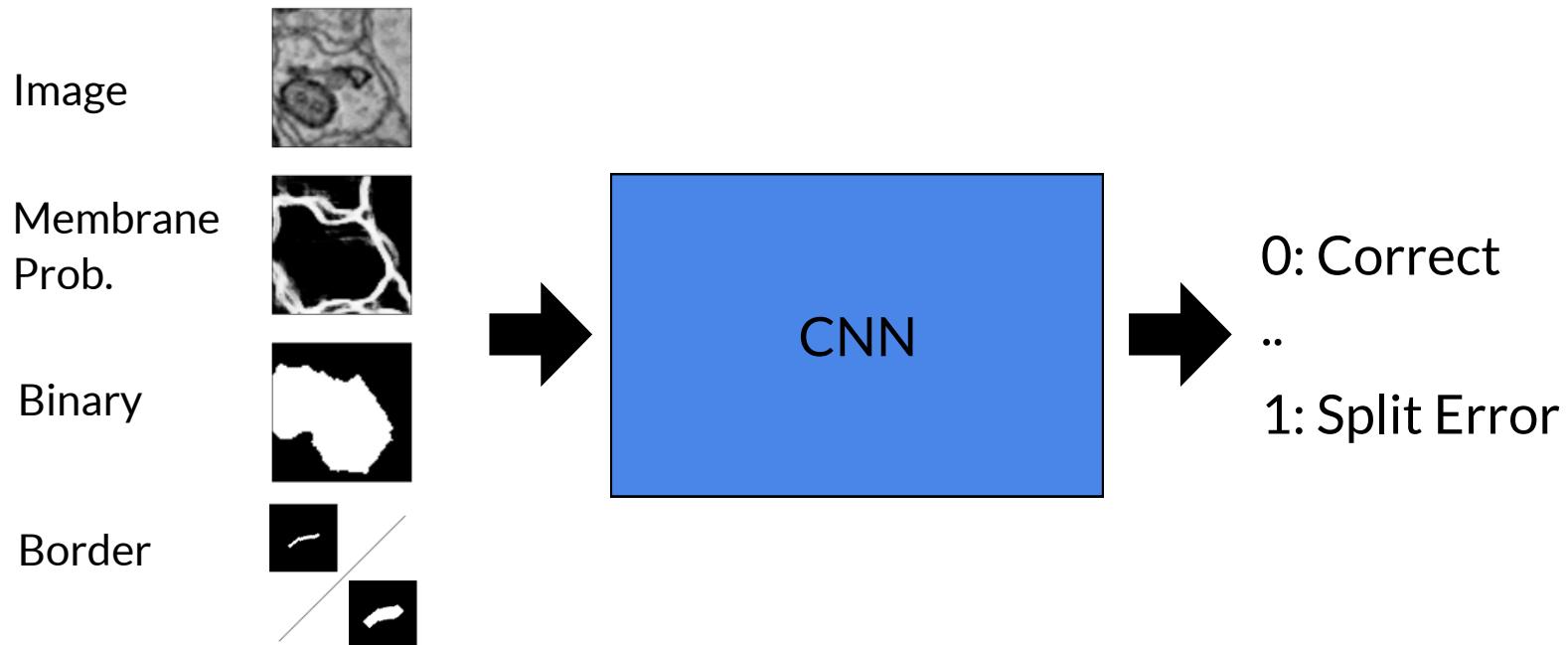


Automatic  
Segmentation

# SPLIT ERROR DETECTION (2)



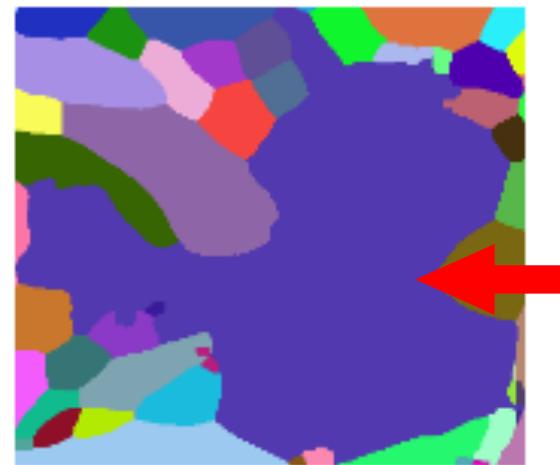
Inspired by Bogovic et al.: **Learned versus Hand-Designed Feature Representations for 3d Agglomeration**, CoRR, 2013.



# MERGE ERROR DETECTION



Ground Truth



Automatic  
Segmentation

# MERGE ERROR DETECTION (2)

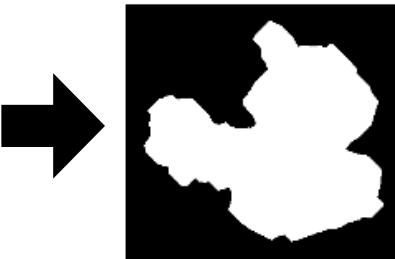
Ground Truth



Automatic Segmentation



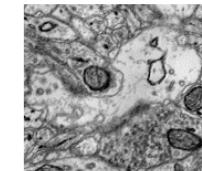
Binary



Dilated



Watershed

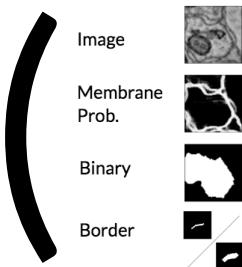


Image



Gradient

1 -



CNN

0: Correct  
..  
1: Split Error



Merge Error

Invert Split Error CNN probability  
of each generated border

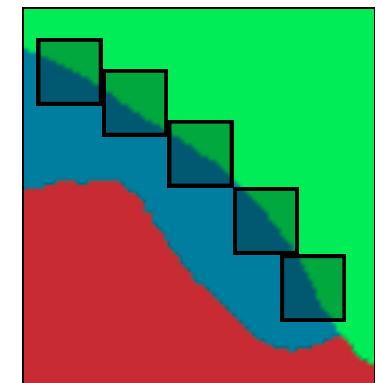
# CNN INPUTS

Patch size 75x75 pixels, center pixel on membrane

	Ground Truth	Segmentation	Image	Probability	Binary	Border	Large Border
Correct	A multi-colored patch showing several distinct regions.	A multi-colored patch showing a different segmentation from the ground truth.	A grayscale patch showing a textured surface.	A grayscale patch showing a network-like structure.	A binary mask showing a single large white blob.	A binary mask showing a small white line segment.	A binary mask showing a small white blob.
Split Error	A multi-colored patch where a boundary between two regions is split into two separate segments.	A multi-colored patch where a boundary between two regions is split into two separate segments.	A grayscale patch showing a textured surface.	A grayscale patch showing a network-like structure.	A binary mask showing a boundary split into two separate segments.	A binary mask showing a boundary split into two separate segments.	A binary mask showing a boundary split into two separate segments.

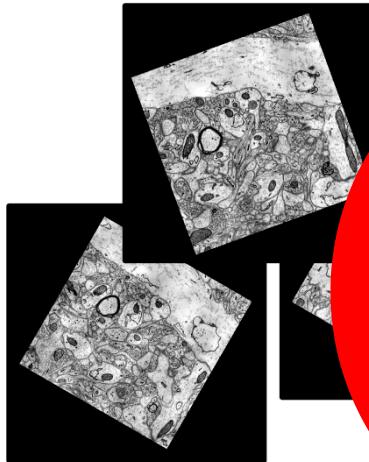
Patch size covers ~80% of all segment boundaries

For larger boundaries: sample multiple times (non-overlapping) and combine (weighted arithmetic mean)



# WHY 2D?

Acquisition



Registration



3D Alignment is very expensive

Fusion step after proofreading required (NeuroProof)

# PATCH GENERATION

Left Cylinder Volume from Kasthuri et al.: **Saturated Reconstruction of a Volume of Neocortex**, Cell, 2015.

2048x2048x300 voxels, 3x3x30nm/voxel

ground truth 180k labels, automatic segmentation 500k labels

## Training/Validation

2048x2048x250 voxels

266,080 correct patches

266,080 split error patches

normalized -0.5 .. 0.5

data augmentation: rotation

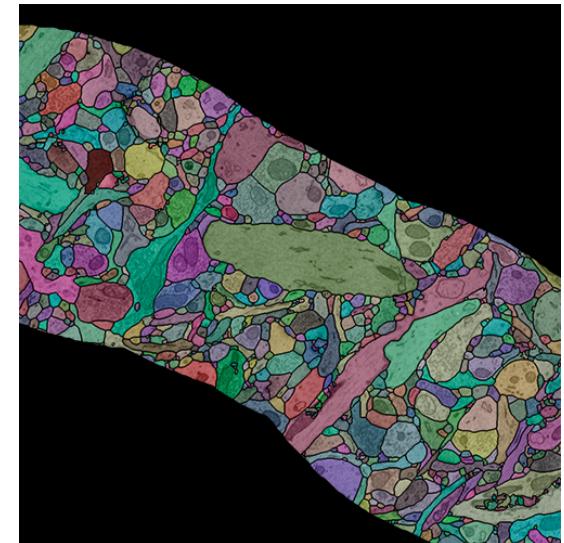
## Test

2048x2048x50 voxels

26,816 correct patches

26,816 split error patches

normalized -0.5 .. 0.5



Patches from Segmentation rather than Ground Truth



Ground Truth



Automatic Segmentation



Maximum Overlap



Wrong Boundaries

# **NETWORK ARCHITECTURE**

# **PARAMETERS**

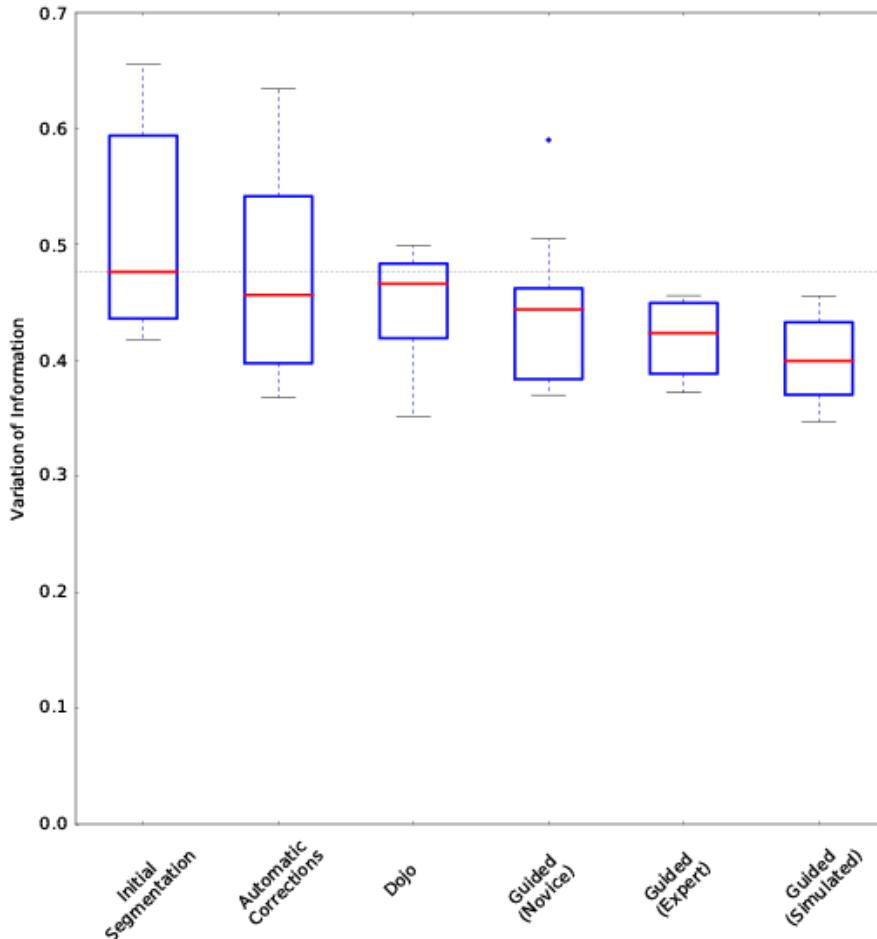
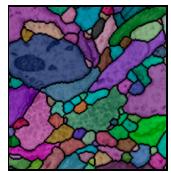
# **RESULTS**

# AUTOMATIC CORRECTION

# **USER IN THE LOOP**

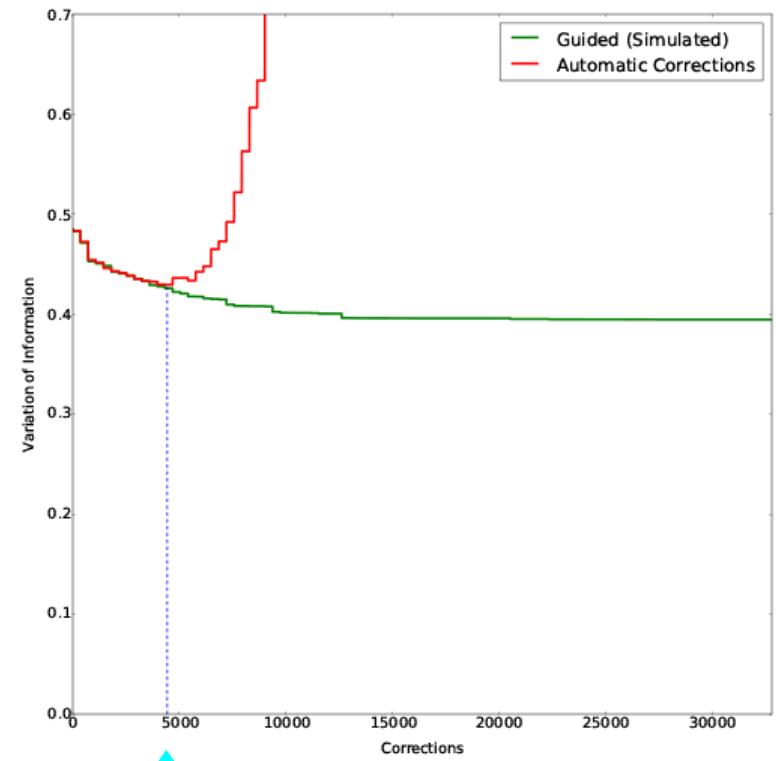
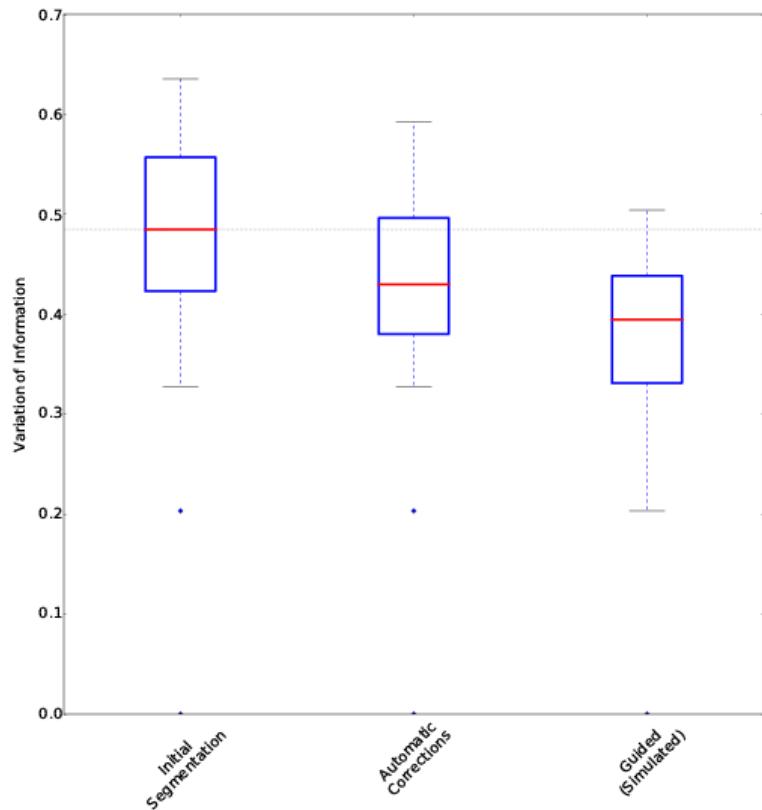
# EXPERIMENTS

Guided Proofreading vs. best Dojo user (30 minutes, same data)



# EXPERIMENTS (2)

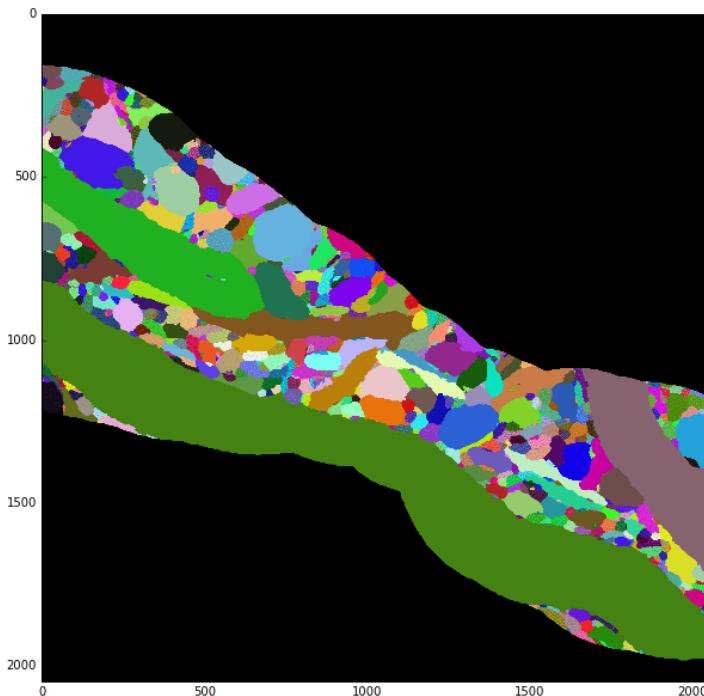
2k x 2k x 50 voxels, Automatic Corrections vs. Simulated User



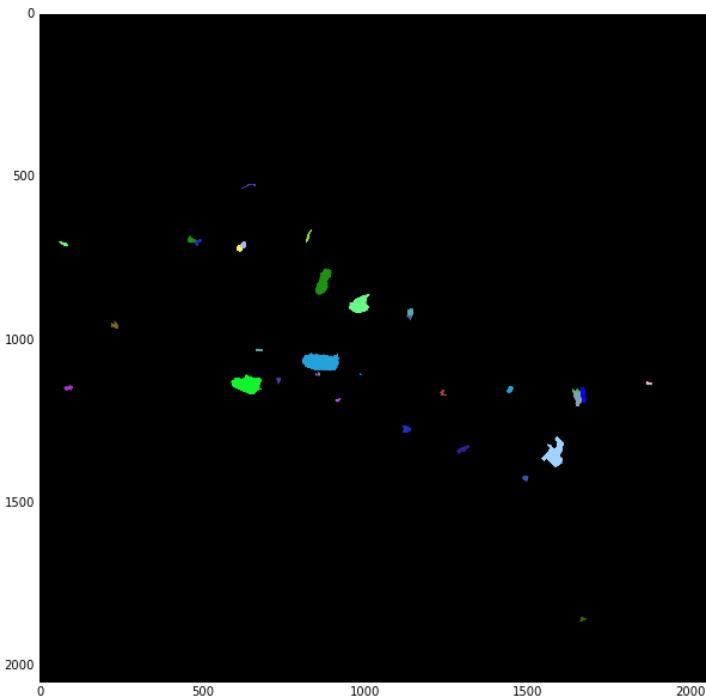
$\uparrow p_t = 0.95$

# EXPERIMENTS (3)

2k x 2k x 50 voxels, Automatic Corrections ( $p_t=.95$ )



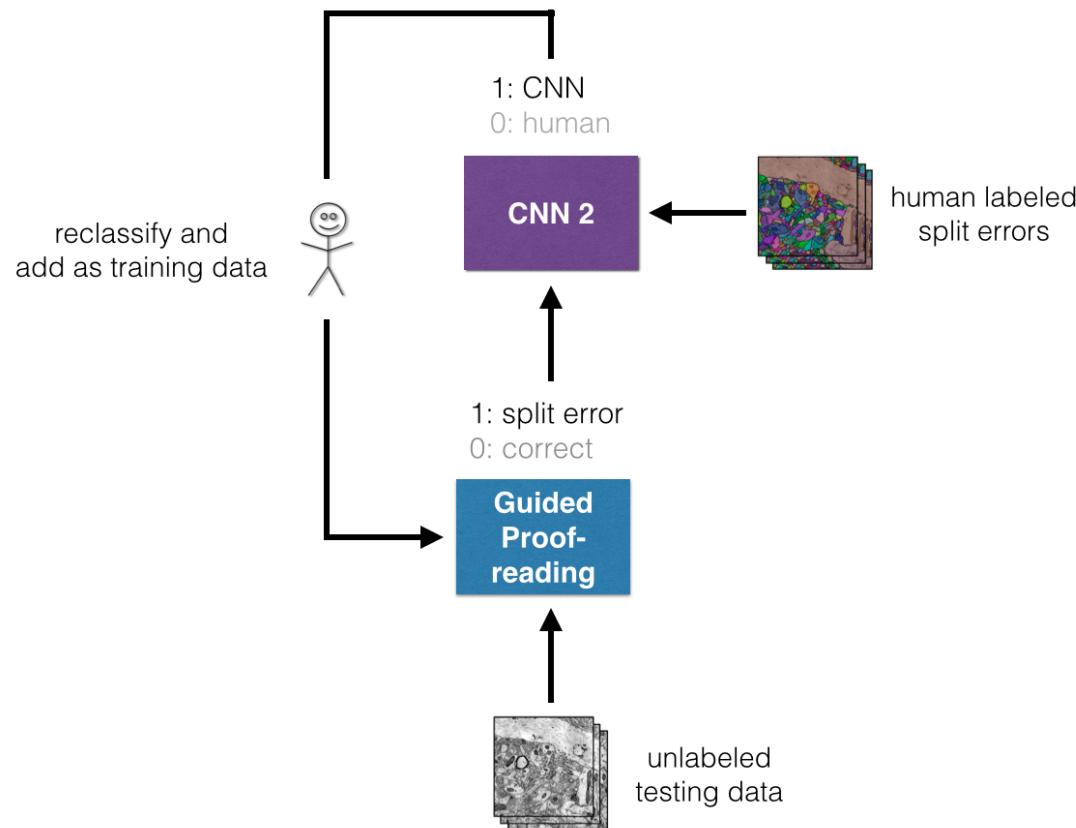
Snapshot every 30 corrections



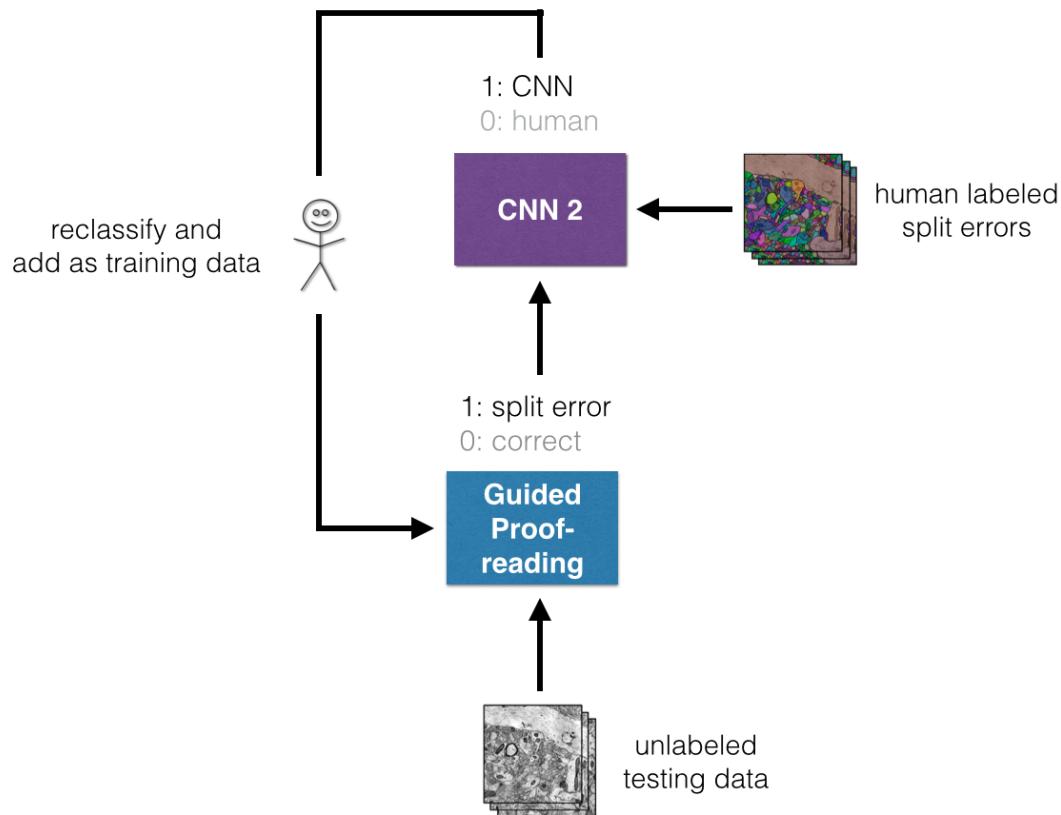
Difference to input segmentation

# FUTURE WORK

Active Learning, inspired by  
Goodfellow et al.: **Generative Adversarial Nets**, NIPS 2014



# FUTURE WORK (2)



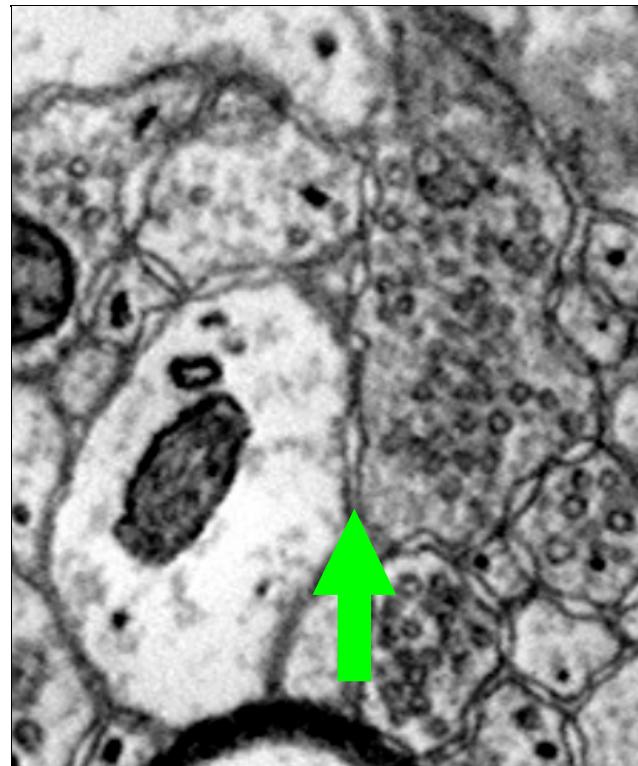
Find pattern of wrong classifications in Guided Proofreading

Ideally, CNN2 is not able to distinguish input from human or our classifier

Fine-tune our classifier

# FUTURE WORK (3)

Proofreading of Synapses



# **CONCLUSION**

**THANK YOU!**