



# Deep Learning-Based Neuron Detection in Brain CLARITY Imaging

Checkpoint Presentation

Members: Prerna Singh

Mentor: Dr. Jeremias Sulam

# Background: CLARITY imaging

Clear

Lipid-exchanged

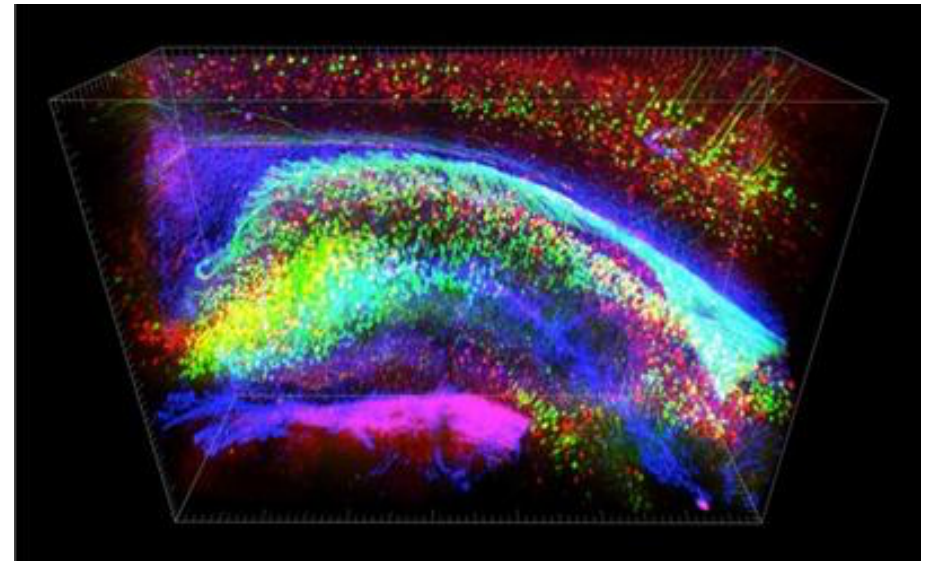
Acrylamide-hybridized

Rigid

In-situ hybridization-compatible

Tissue

Hydrogel



<https://en.wikipedia.org/wiki/CLARITY>

# Significance of CLARITY for Brain Imaging

- Connectome project
  - Local circuit wiring
  - Relationships between neural cells★
- Neurological diseases
  - 3D view of brain structures

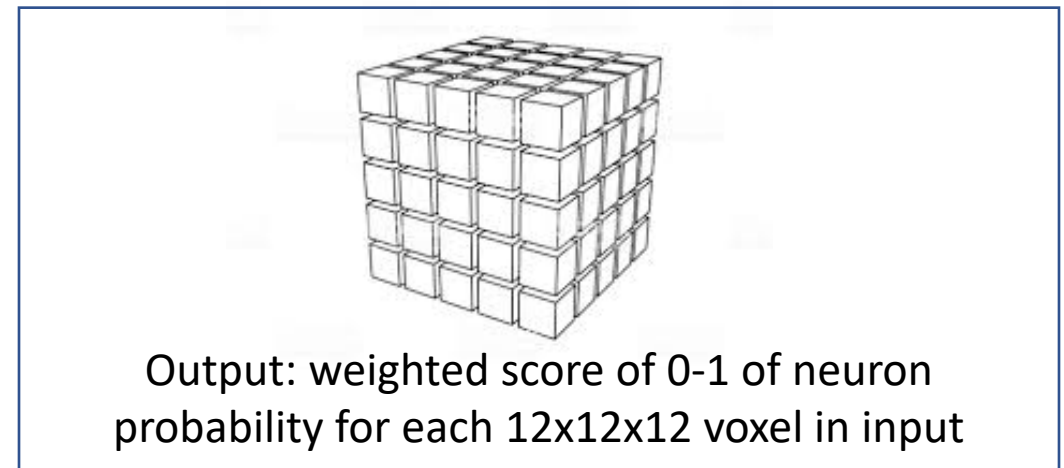
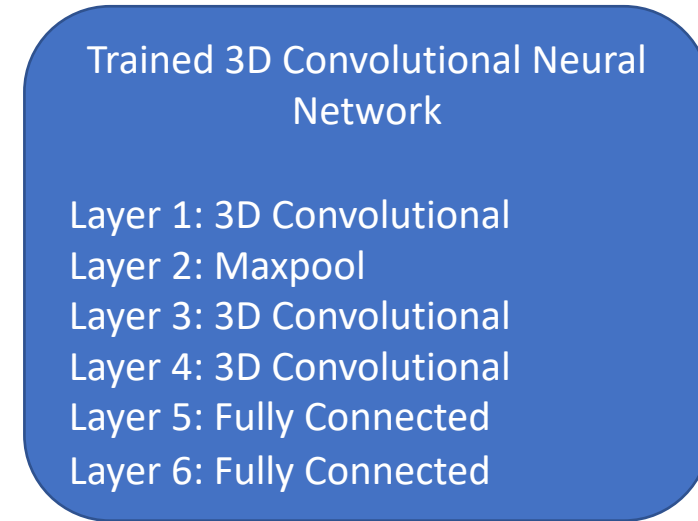
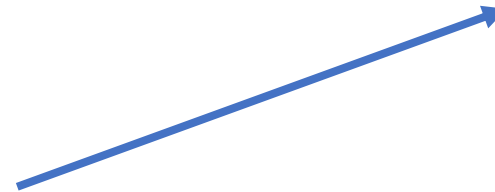
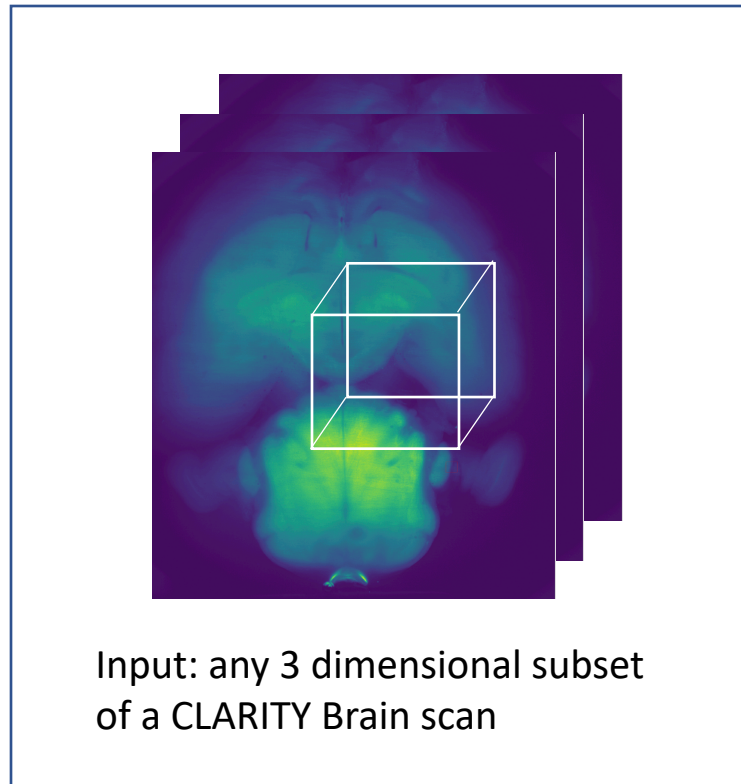


<http://www.humanconnectomeproject.org/>

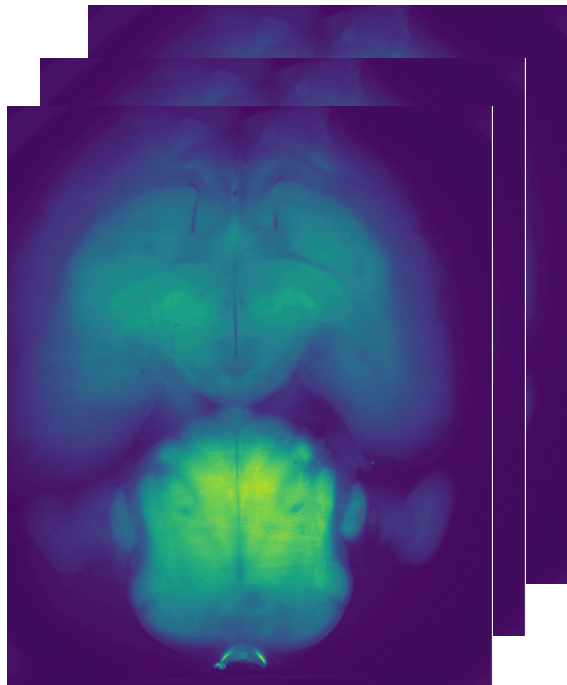
# Goal

- Develop a **robust 3D CNN** that can predict, with improved precision and accuracy when compared to other models, how many fluorescent neurons are present within a section of a brain imaged with CLARITY
  - Previous models: template matching, blob detection
  - Maximum accuracy achieved was  $\sim 59\%$  on training data with these techniques

# Technical Approach- Design and Implementation of 3-dimensional CNN Model



# Technical Approach – First training and validation of 3-dimensional CNN training on reduced dataset



One 3-D CLARITY Brain Imaging Data annotated for Neurons

Neuron 12X12X12 voxels  
and non-neuron 12x12x12

→

voxels from these images  
are used for training

3D Convolutional Neural  
Network

Model applied to training data  
resulted in 95% accuracy: **limited  
need for data augmentation**

# Work to date

- 300x300x1 slice
  - 7/7 known cells were identified
  - 25211 more candidates
- 500x500x1
  - 12/12 known cells were identified
  - 100165 more candidates
- 2150x2650x1
  - Running now

# Technical Approach- Large Scale Training on Complete Dataset & CNN Validation

- Data includes 9 whole brains
- **Coarse validation**: report metrics (e.g. AUC) over a few random splits of data (6 train, 3 test)
- **Full Leave-one-out (LOO) Cross validation**: report cross-validated metrics (e.g. AUC) averaged over 9 folds. In each fold, all but 1 brain will be used for training, and the remaining one for testing

# Dependencies

| Dependency       | Solution                               | Date Expected | Date Required | Alternative                            |
|------------------|--|---------------|---------------|--|
| Computing power  | MARCC Access granted through Dr. Sulam | 2/10          | 2/25          | Very slow training                     |
| Computer         | Personal laptop                        | 2/10          | 2/10          | Desktop computer provided in Sulam Lab |
| Backups          | Private github code storage,           | 2/10          | 2/25          | code & data also stored on MARCC       |
| pytorch          | Install pytorch for use on MARCC       | 2/10          | 2/25          | Tensorflow                             |
| Data- Images     | Downloaded off of BOSS Neurodata       | 2/10          | 2/16          | --                                     |
| Data-Annotations | Provided by Dr. Sulam                  | 2/10          | 2/16          | --                                     |



NEW ISSUE



ALL RESOLVED

# Dependency issues

Matthew via RT <marcc-hel  
p@marcc.jhu.edu>



Wed 4/1/2020 3:22 PM  
Prerna Singh

Hello Prerna,

Excellent. Thank you for the information. I got an update from systems and they've narrowed down where in the system commands where it's occurring but it's still an ongoing process for fixing it. I'm very sorry for the inconvenience. Below is a response we send out sometimes.

"MARCC is effectively free to most users and we are severely limited in staff. We encourage dissatisfied users to consult with the University leadership to find ways to adequately fund and manage this resource rather than relying on diminished staff."

Matthew via RT <marcc-hel  
p@marcc.jhu.edu>



Fri 3/27/2020 3:17 PM  
Prerna Singh

Hello Prerna,

Thank you for letting me know. This time I caught it. The data directory is in fact unmounted from gpu032 but for your account and not mine. This is very unusual and I will bring it up to the systems group who are the administrators of the file system. I hope to update you soon with a fix. Sorry for the inconvenience.

Sincerely,  
Matthew Wang

Matthew via RT <marcc-hel  
p@marcc.jhu.edu>



Wed 4/1/2020 2:41 PM  
Prerna Singh

Hello Prerna,

Yes, I'll check with systems on an update on data. If the lrgmem also threw that cpu allocate issue then there's no point in continuing that line of investigation. I'll search around for more information on this error.

Sincerely,  
Matthew Wang

...

Matt via RT <marcc-help@  
marcc.jhu.edu>





Mon 3/23/2020 10:31 AM  
Prerna Singh

We are investigating an intermittent issue with the /data directories that may report failed permissions when trying to access files or directories within /data. There is no risk to data loss, but the issue may interrupt some jobs that read or write from /data.

# Deliverables

## Minimum:

- Trained 3D CNN model (pytorch) for neuron detection 
- Report in Jupiter notebook – explains training of CNN and predictions 

## Expected:

- Robust and validated trained model (pytorch) for neuron detection
- Packaged model with documentation –made available on the internet with equal or better performance than competing simpler alternatives

## Maximum:

- Academic paper describing packaged model, training, validation
- Surpass baseline accuracy of previous models by a significant margin



# Milestones

| Date | Milestone                              |
|------|--|
| 3/1  | CNN input script completed             |
| 3/15 | Preliminary model trained and designed |
| 4/5  | Preliminary model validated            |
| 4/18 | Complete model trained and validated   |
| 5/10 | Software package completed             |

