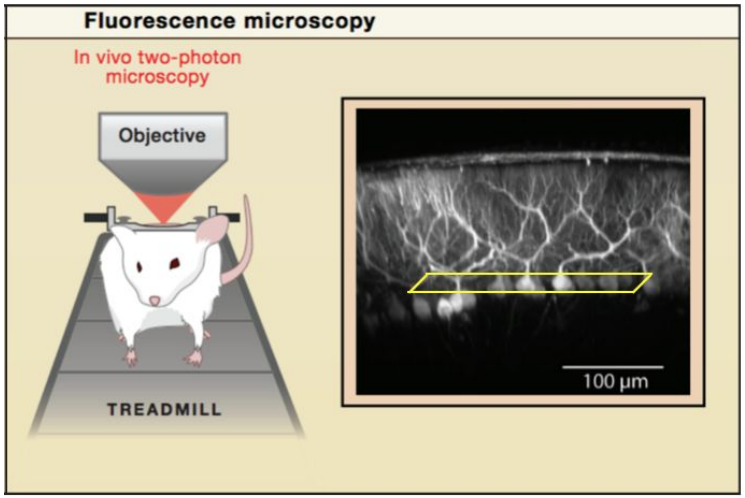
Large scale analysis of calcium imaging data

October 12, 2017

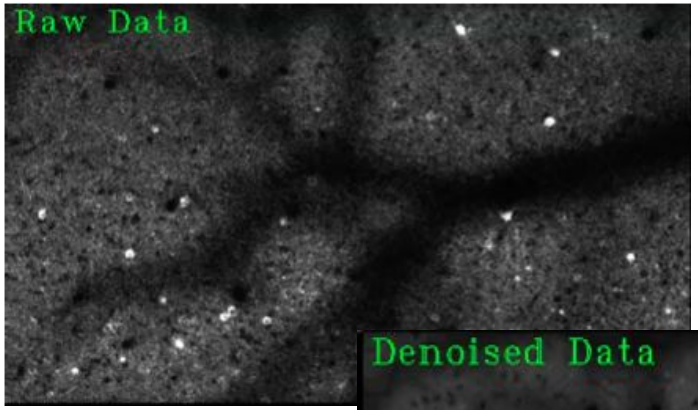
Data Minions(Yidi Zhang, Yang Sun, Shangying Jiang, Yiran Xu, Zhaopeng Liu) +
Andrea Giovannucci and Eftychios Pnevmatikakis (mentors from Simons Foundation, Flatiron Institute)

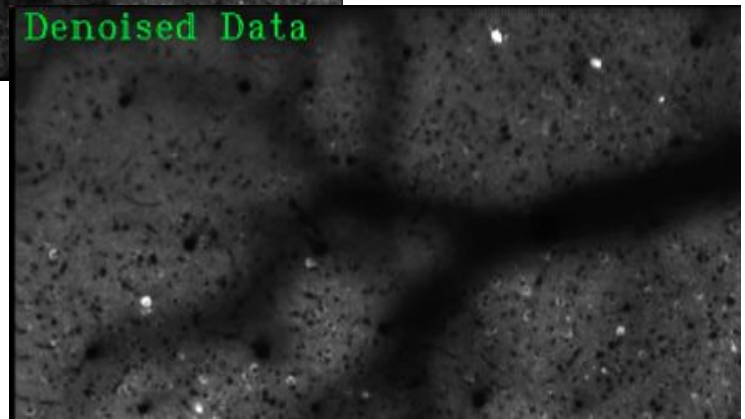
Data

- **Input data** is raw movies of various sizes (0.2 - 1 Megapixels per frame and about 10-100 kilo-frames) with binary masks representing neuronal locations.



X. Richard Sun, Andrea Giovannucci, Allyson E. Sgro, Samuel S.-H. Wang. Cell. 2014.

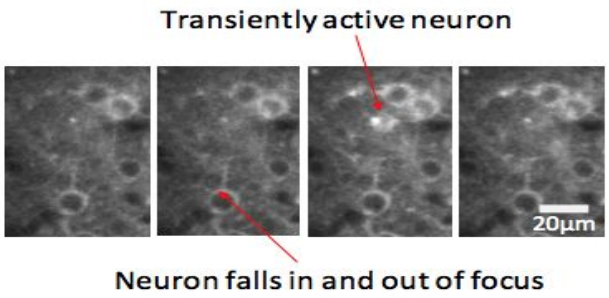




Understanding the problem(Questions)

1. Improve CNN model

- Classify spatial components extracted from the constrained non-negative matrix factorization (CNMF) algorithm into three classes: neurons, processes and noise artifacts.
- Improve existing CNN model (only based on spatial features)
- Further include temporal features (to detect transiently visible neurons).



2. Implement online algorithm

- Implement optimized version of the online algorithm (unsupervised learning, run in real-time) in Java or C++
- Incorporate the classifier developed in Q.1
- Implement as ImageJ Plugin in Java

Methodology

1. **Training data creation and preprocessing.**
    a. Only spatial
    b. Both spatial and temporal
2. **Train neural networks to perform classification.**
    a. Starting from CNN based approach
    b. Explore ResNet and VGG-like architectures
    c. Explore/Compare metaparameters, architectures and augmentation techniques
3. **Compare the proposed algorithm against test sets of manually annotated data.**
4. **Optimization and extension of online OnACID (Online analysis of Calcium Imaging data) algorithm.**

Rubric (Final Results: Report, Source Code, Illustrative movies)

Question A	• Generate images demonstrating proper components classification.
	• Illustrate performance on training sets and test sets of various architectures.
Question B	• Beat consistently vanilla algorithm on test set.
	• Implement plugin of ImageJ (Java) with online algorithm running on zebrafish data.