

# COLLECTIONS OF CLASSIFIERS TUNED FOR CELL FINDING WITH AN APPLICATION TO BUILDING DIGITAL CELL ATLASES OF DROSOPHILA EMBRYOS

William Beaver<sup>1</sup>, Adam Paré<sup>2</sup>, David Kosman<sup>2</sup>, Ethan Bier<sup>2</sup>, William McGinnis<sup>2</sup>, and Yoav Freund<sup>1</sup> {wbeaver, yfreund}@ucsd.edu

I. Department of Computer Science and Engineering, UCSD, La Jolla, CA 2. Cell and Developmental Biology Department, UCSD, La Jolla, CA

#### Data

#### Goals

#### Biology:

Combine

rho @ 63x

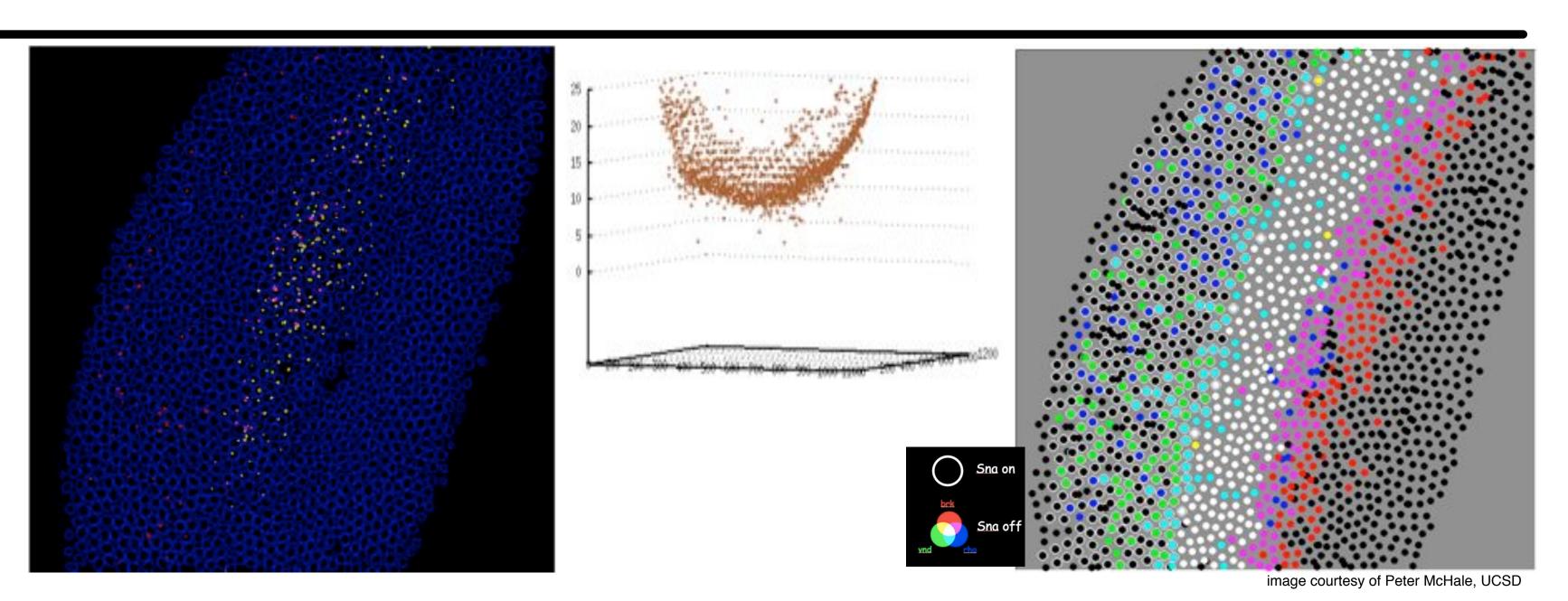
- quantitative analysis of image data
- Digital cell atlas

#### Computer Science:

- Classifier based algorithms (object detection)
- Classifier reuse (definition, storage, distribution, ...)

Image

Stack



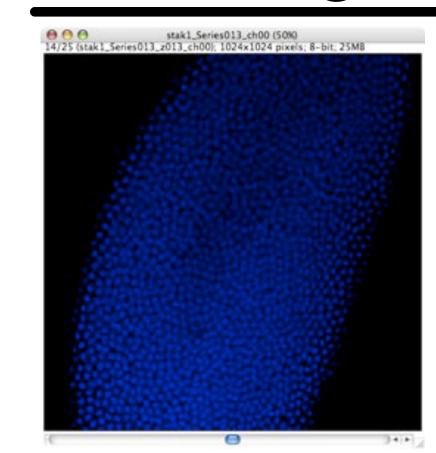
### Challenges

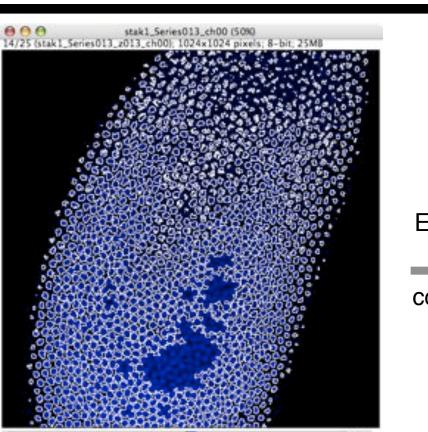
- Tuning is required for computer vision algorithms (for both initial use and as experimental conditions change)
- Tuning requires computer vision expertise
- experimentalist ≠ computer vision expert (typically neither side wants =)
- Building classifiers not as easy as it should be
- Sharing classifiers not as easy as it should be

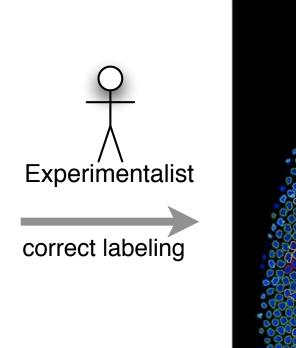
### Approach

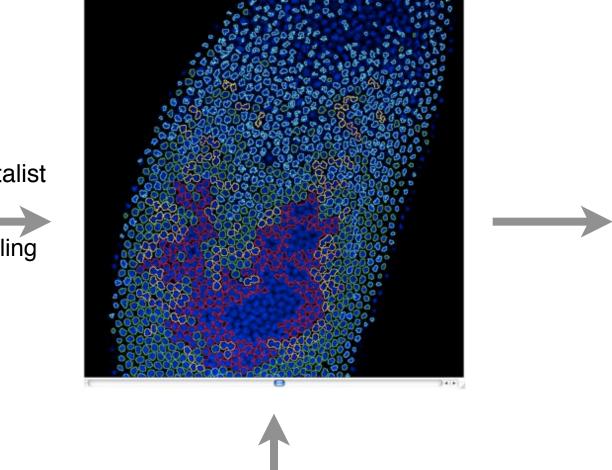
- use machine learning to mimic biologist Image
- intuitive software interface engages biologist in training machine
- learn many classifiers as needed

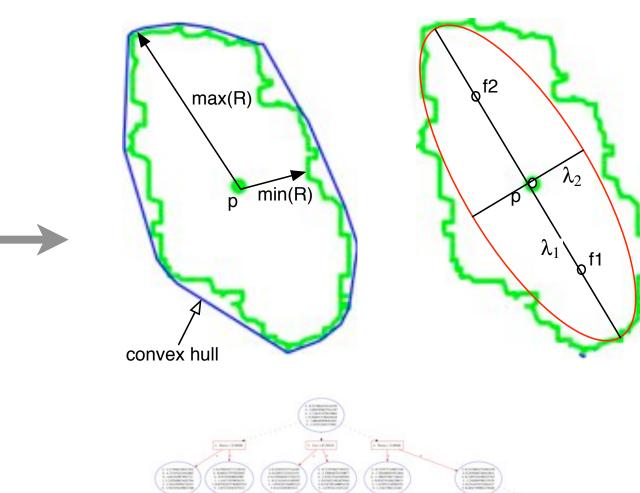
## Learning Process

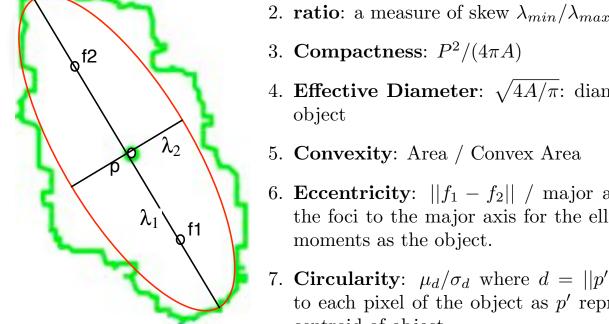












3. Compactness:  $P^2/(4\pi A)$ 4. Effective Diameter:  $\sqrt{4A/\pi}$ : diameter of circle with same area as the 5. Convexity: Area / Convex Area

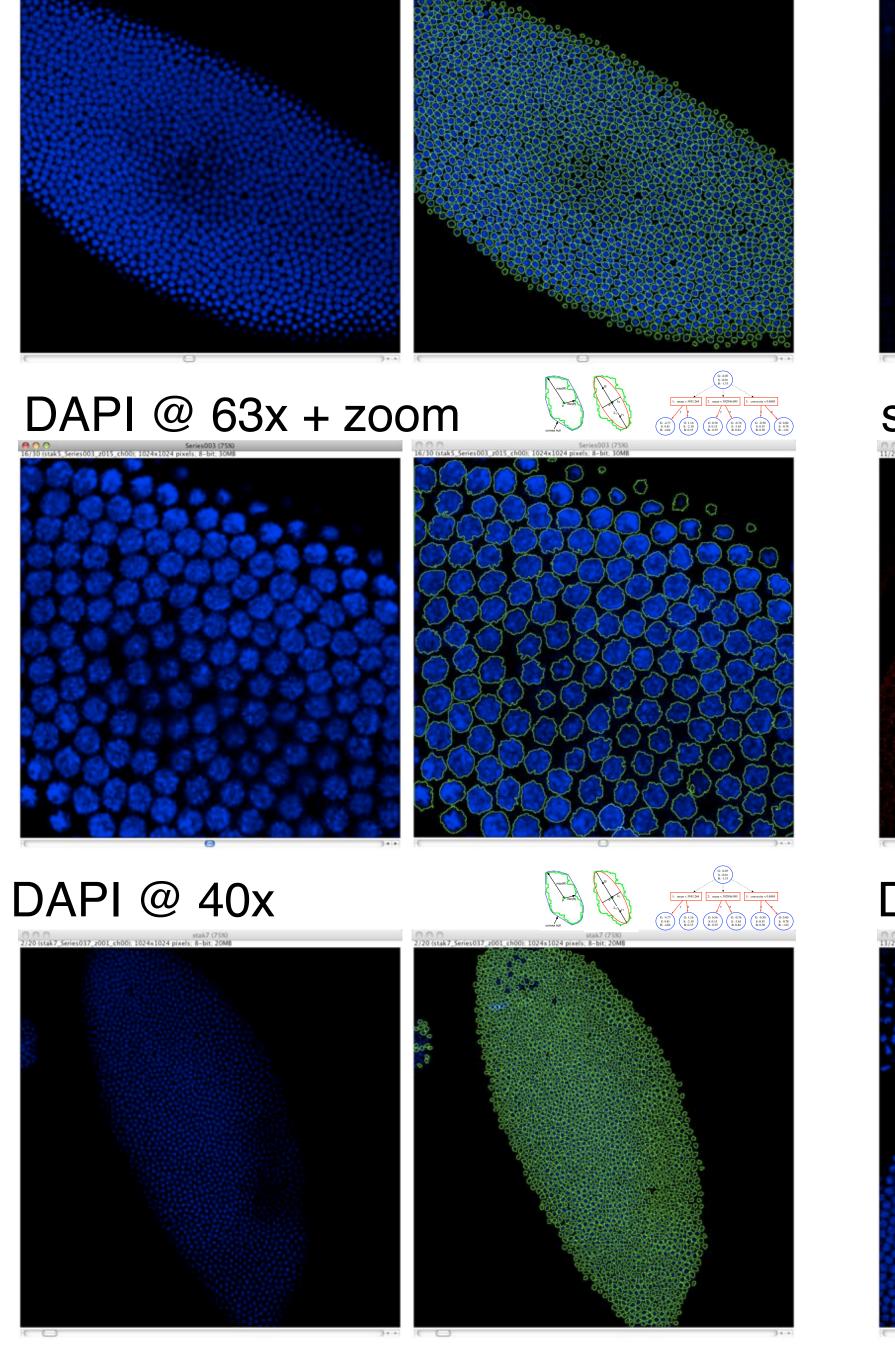
1. **mean**: a measure of size  $\mu_{\lambda} = (\lambda_1 + \lambda_2)/2$ 

- . **Eccentricity**:  $||f_1 f_2||$  / major axis length: ratio of the distance of the foci to the major axis for the ellipse that has the same second order moments as the object. to each pixel of the object as p' represents each pixel in object and p is
- Boosting

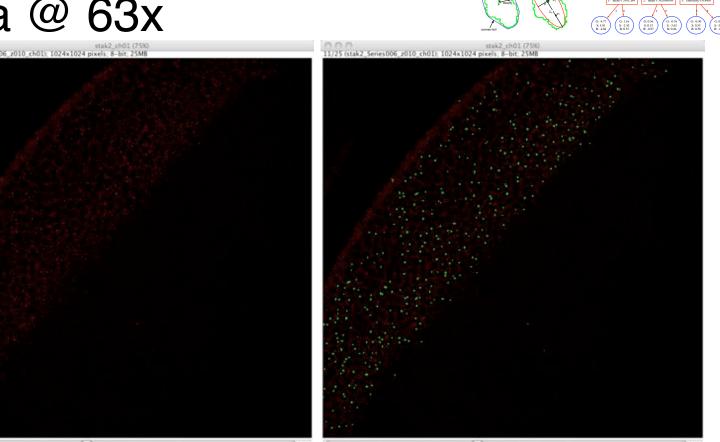
Classifier (Alternating Decision Tree)

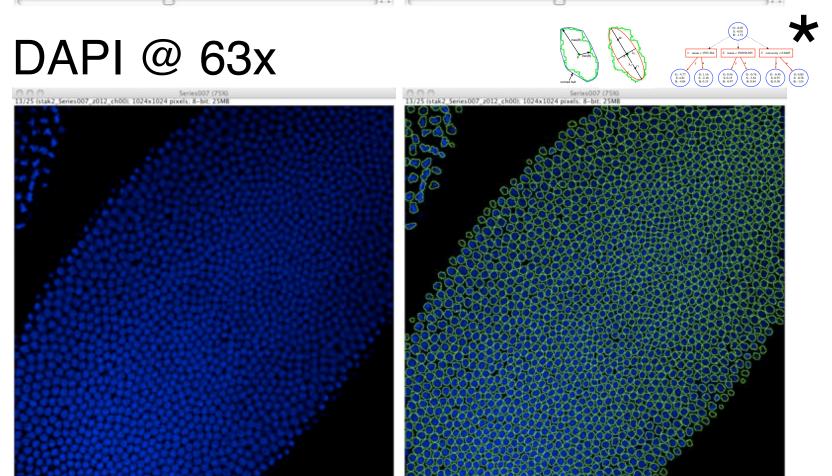
#### Results

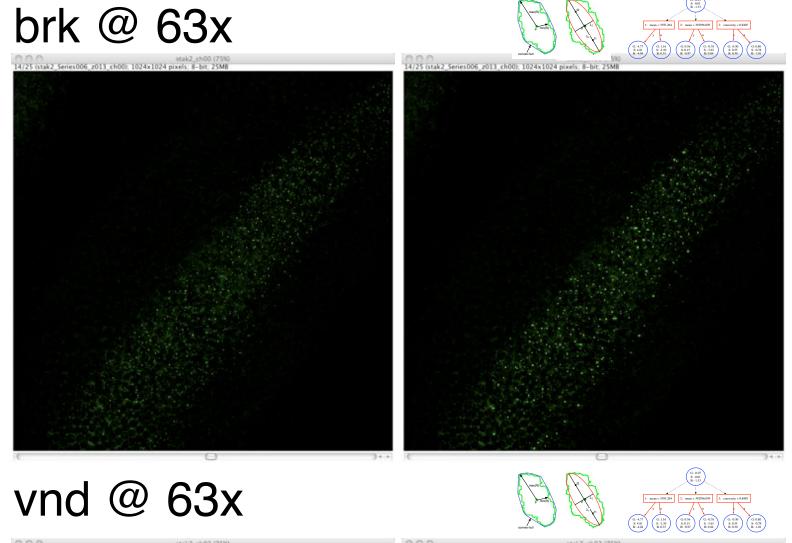
**DAPI** @ 63x

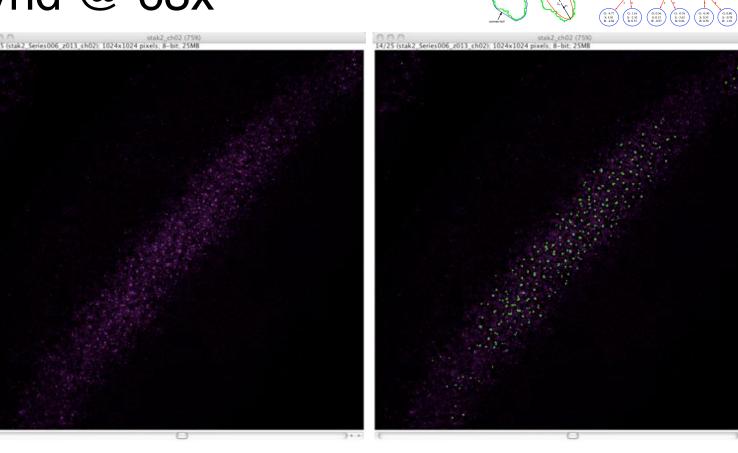


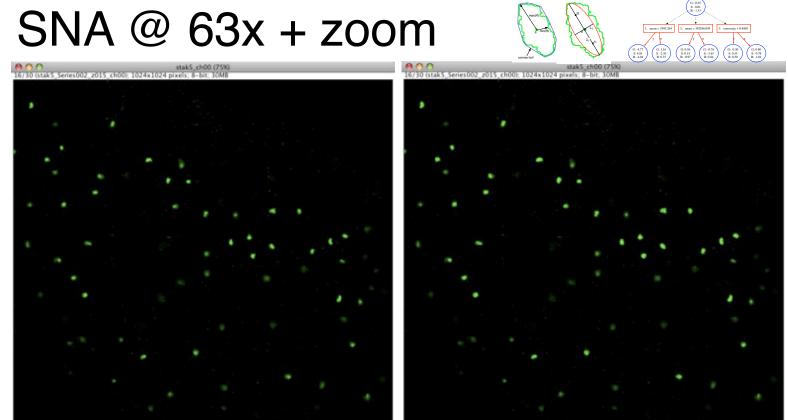
# sna @ 63x





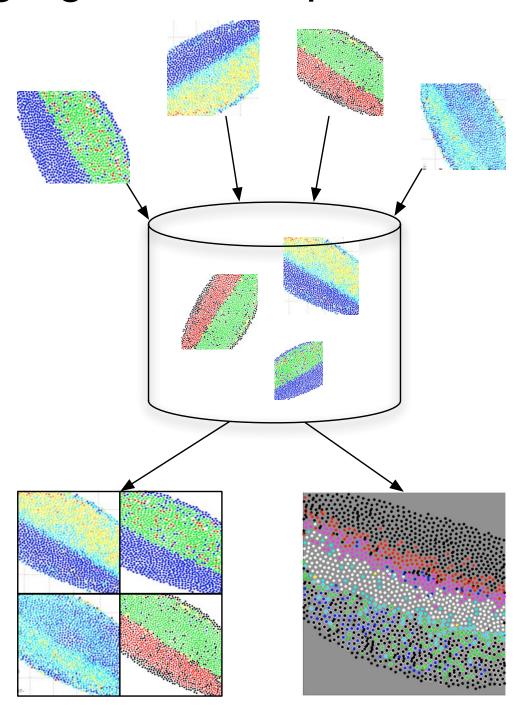






### Our Vision

Individual spatially & temporally aligned models allow quantitative comparison and aggregation of experiments



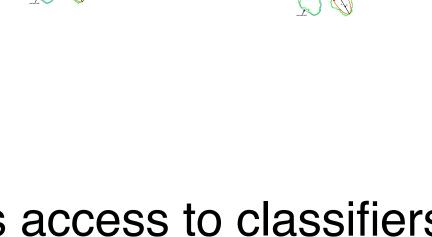
Construct models of embryonic development that can be queried

Provide tools to drive interactive inquiry and experimentation

Store classifiers and training data in a central repository. Standardizes interfaces & metadata

Collections of classifiers used, on demand, for:

- reuse in existing algorithms,
- development of new algorithms,
- repurposing of classifiers,
- adaption of existing classifiers
- many other uses....



Allow external entities access to classifiers, algorithms and tools

Other experimentalists/labs contribute classifiers back to classifier repository