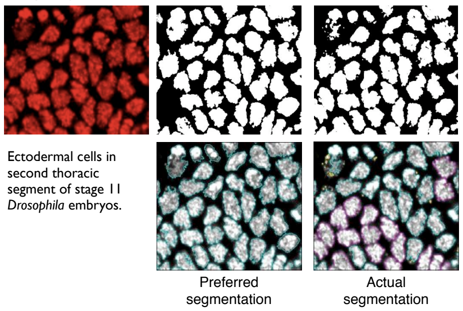
SEGMENTATION OF NUCLEI IN CONFOCAL IMAGE STACKS USING PERFORMANCE BASED THRESHOLDING

William Beaver[1], David Kosman[2], Gary Tedeschi[2], Ethan Bier[2], William McGinnis[2], and Yoav Freund[1]
{wbeaver, yfreund}@ucsd.edu
1. Department of Computer Science and Engineering, UCSD, La Jolla, CA   2. Cell and Developmental Biology Department, UCSD, La Jolla, CA

Problem



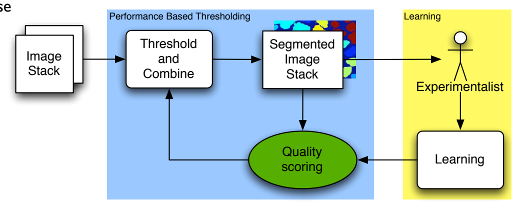
Ectodermal cells in second thoracic segment of stage 11 *Drosophila* embryos.

Challenge

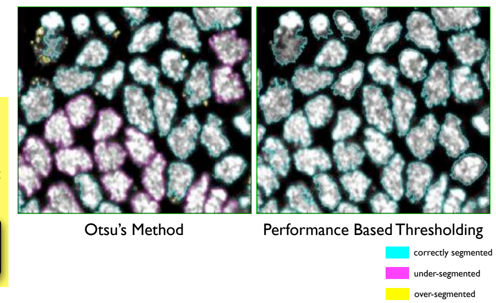
- Tuning required for computer vision algorithms (for both initial use and as experimental conditions change)
- Tuning requires computer vision expertise
- experimentalist ≠ computer vision expert (typically neither side wants =)

Our Approach

- use machine learning to *mimic* biologist
- intuitive interface engages biologist in training machine



Results

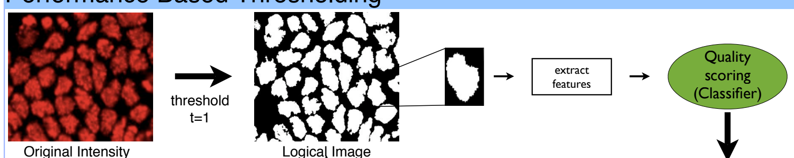


method	count	diff	% found	US	Conn
Mean	2535	-1126	69%	7572	347
Otsu	2741	-920	75%	4134	186
PBT	3505	-156	96%	20	173

Performance of various thresholding algorithms in detecting objects vs. ground truth segmentation from CT-FISH confocal image stack of DAPI. The actual number (ground truth) of 2D nuclei segments in tested 18 frame stack containing approximately 600 nuclei, each spanning about 9 frames, is 3661.

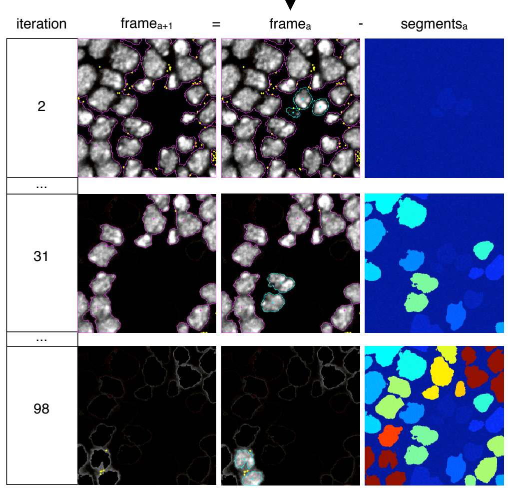
diff = difference from ground truth,
US = under-segmentation
Conn = Connected (nucleus not detected as a separate segment but combined with another nucleus.)

Performance Based Thresholding

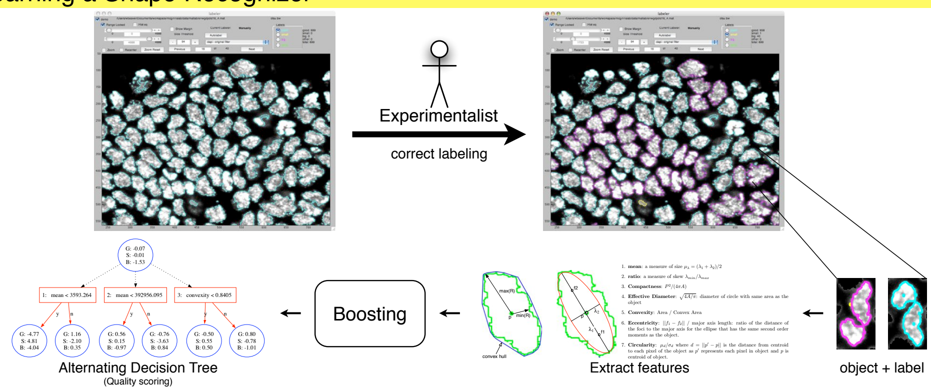


for all frame_a, in stack **do**
  $\text{segments}_a = \{\}$
  for all thresholds t from low to high **do**
    $B = \text{frame}_a > t$
    CE = 2D connected elements in B.
    for all ce_j in CE **do**
      p = F(ce_j)
      if p ids "good" **then**
        mask (remove) ce_j from frame_a.
        add ce_j to segments_a.
      end if
    end for
  end for
  Connect segments between segments_a and segments_{a-1}
end for

Pseudocode for Performance Based Thresholding (PBT)



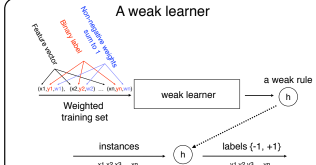
Learning a Shape Recognizer



Boosting Details

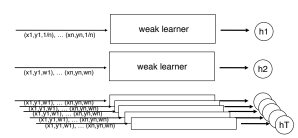
for more information see: www.cse.ucsd.edu/~yfreund/papers/IntroToBoosting.pdf

A weak learner



The weak requirement: $\left|\frac{\sum_{i=1}^{n} y_i \hat{y}_i w_i}{\sum_{i=1}^{n} w_i}\right| > \gamma > 0$

The Boosting Process



$$F_T(x) = \alpha_1 h_1(x) + \alpha_2 h_2(x) + \ldots + \alpha_T h_T(x)$$
$$f_T(x) = sign(F_T(x))$$

AdaBoost

$F_0(x) \equiv 0$

for $t = 1..T$

$w_i^t = \exp(-y_i F_{t-1}(x_i))$

Get h_t from $weak-learner$

$\alpha_t = \frac{1}{2}\ln\left(\sum_{i:h_t(x_i)=1, y_i=1} w_i^t \Big/ \sum_{i:h_t(x_i)=1, y_i=-1} w_i^t\right)$

$F_{t+1} = F_t + \alpha_t h_t$

[Freund, Schapire 1997]