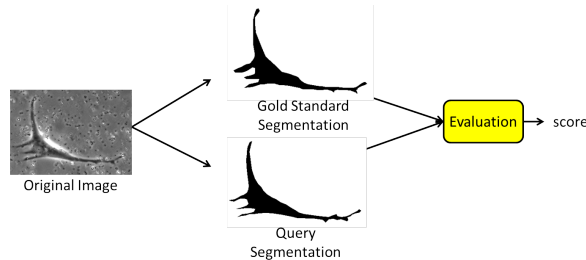# SAGE: An Approach and Implementation Empowering Quick and Reliable Quantitative Analysis of Segmentation Quality

Danna Gurari[1], Seule Ki Kim[1], Eugene Yang[1], Brett Isenberg[2], Tuan A. Pham[2], Alberto Purwada[2], Patricia Solski[2], Matthew Walker[3], Joyce Y. Wong[2], Margrit Betke[1]

[1] Department of Computer Science; [2] Department of Biomedical Engineering; [3] Department of Biology

Scope: Segmentation Analysis

Select *measure* to produce score indicating how similar a query segmentation is to a gold standard segmentation.

Y. J. Zhang. A Survey on Evaluation Methods for Image Segmentation. *Pattern Recognition*, 29(8): 1335-1346, 1996.
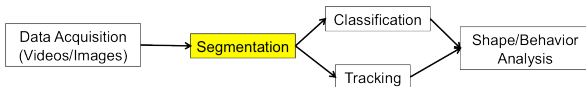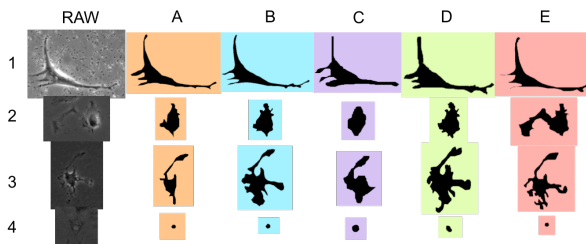
Motivation

Computer Vision

- *Demonstrate effective algorithm design* with good scores...
  Which evaluation measure?

Biology/Biomedical Engineering

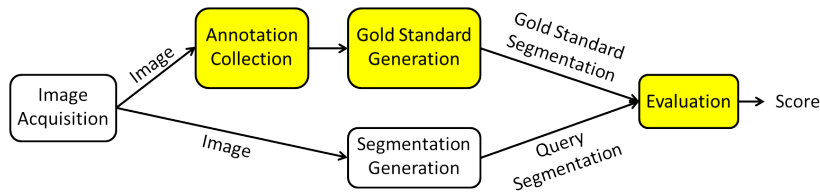- *Trust analyses* derived from segmentations...

Which gold standard?

Acknowledgments

Key Contribution #1: Segmentation Analysis Model

Observing that the performance score depends on the gold standard segmentation, we propose an analysis approach that introduces the consideration of how to establish the gold standard. It links annotation collection approaches with gold standard generation methods and evaluation algorithms into a unified framework we call SAGE (**S**egmentation **A**nnotation Collection, **G**old Standard Generation, and **E**valuation).

Overview of SAGE (yellow) within the context of analyzing a query segmentation.

Key Contribution #3: How to Model Gold Std Segmentation

Case studies showing the impact of annotation collection and gold standard generation on establishing trusted (i.e., high-consensus) gold standard segmentations in practice.

	ID	# of Images	Imaging Modality	Object
Image Library:	1	35	Phase Contrast	Neonatal rat smooth muscle cells
	2	48	Phase Contrast	Fibroblast cells of the Balb/c 3T3 mouse strain
	3	36	Phase Contrast	Vascular smooth muscle cells from rabbit aortas
	4	35	MRI	Left renal artery and the iliac bifurcation of a New Zealand White Rabbit

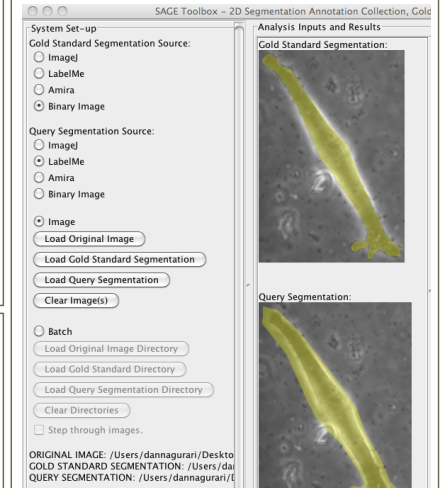|	ID	Education Level	Worked with cell images	Worked with MRI images	Used ImageJ	Used Amira
Annotators:	A	Undergrad	3 mths	None	Yes	No
	B	Post-doc	14 yrs	3 mths	Yes	No
	C	PhD student	10 yrs	1 yr	Yes	No
	D	Post-doc	2 mths	None	Yes	No
	E	PhD student	3 wks	1 yr	Yes	No

1) **Which annotation tool?** *Preference for Amira over ImageJ*
2) **Who annotates?** *Education and experience not important*
3) **Should fusion methods be used?** *Preference for original over fused annotations*

Key Contribution #2: Toolbox

Freely available implementation of SAGE:

http://www.cs.bu.edu/~betke/SAGE

Annotation Collection

Gold Standard Generation

User can select an original or fused annotation to represent the gold standard. Supported fusion algorithms are STAPLE [Warfield 2004] and TPM [Meyer 2006]

Evaluation