



EMORY
UNIVERSITY

Department of
Biomedical Informatics
Emory University School of Medicine

Biomedical Image Analysis for Tumor Characterizations

Jun Kong, PhD

Assistant Professor
Department of Biomedical Informatics
Department of Mathematics and Computer Science
Winship Cancer Institute
Emory University

Nov 11, 2016



Research Areas

- Biomedical Image Analysis
 - 2D Whole-Slide Microscopy Image Analysis
 - 3D Pathology Image Analysis
- Computer-Aided Diagnosis
- Objection Representation and Tracking
- Machine Learning and Pattern Recognition
- Integrative Study with Imaging Data, Genomics, and Clinical Outcome for Oncology Translational Research

Digital Pathology



- Traditional Pathology: **qualitative** in nature, **time-consuming**, and prone to severe observer **variability**



- Digital Pathology: easy to be archived, queried, transferred, annotated, and shared

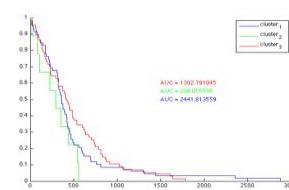
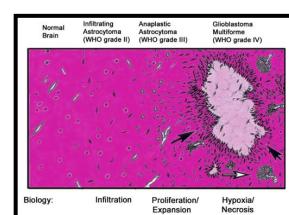
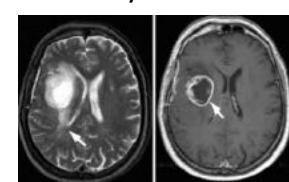


- Computer-based Microscopy Image Analysis : quantitative, reproducible, scalable, and definitive

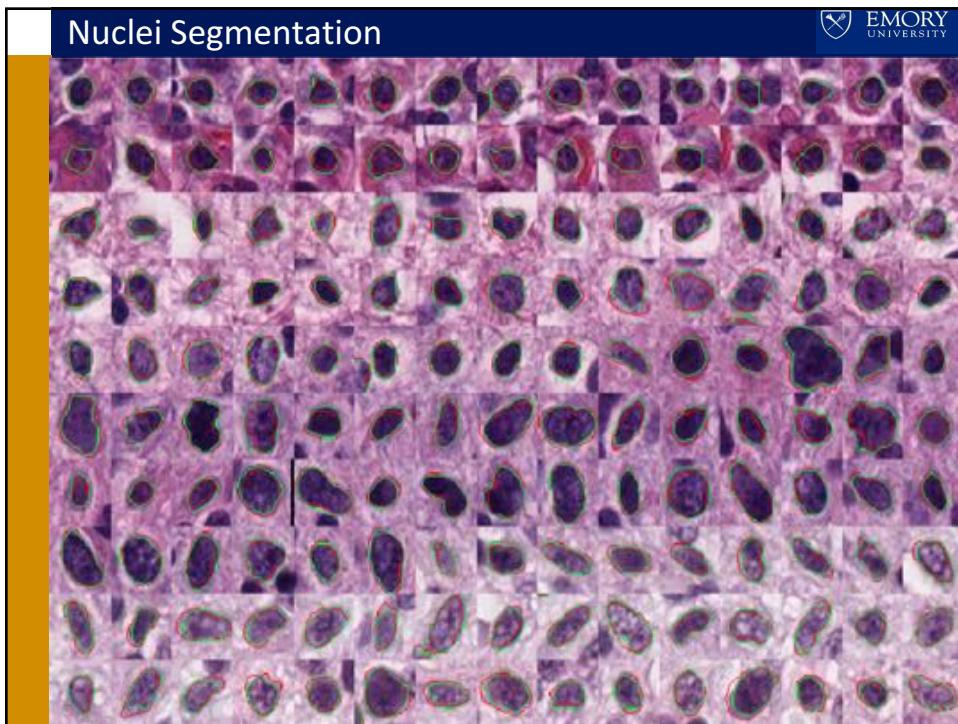
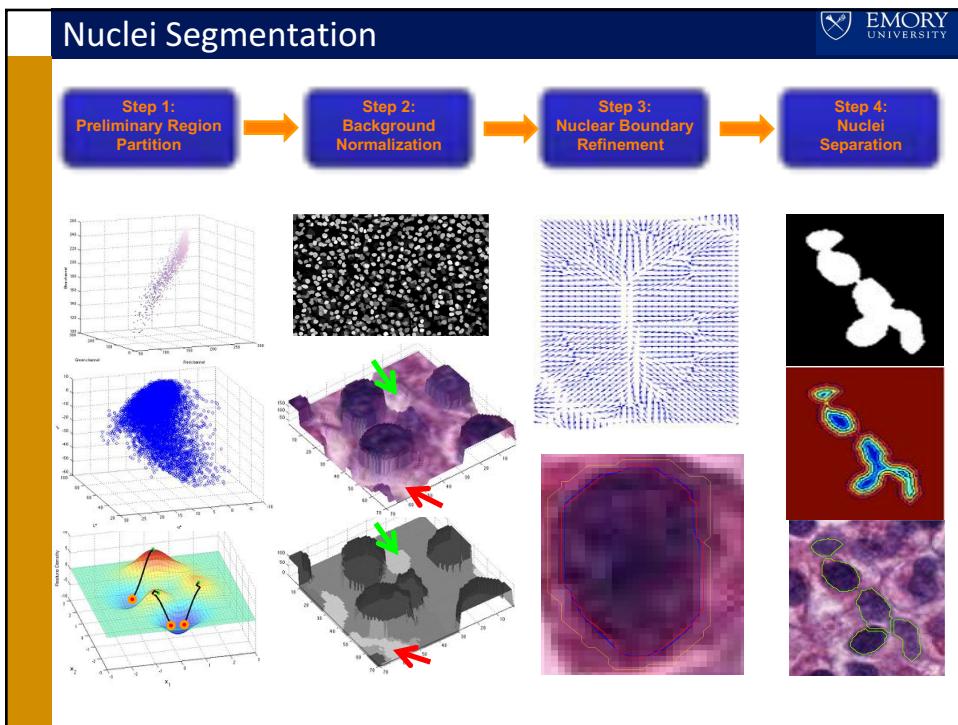
Introduction of Glioblastoma (GBM)

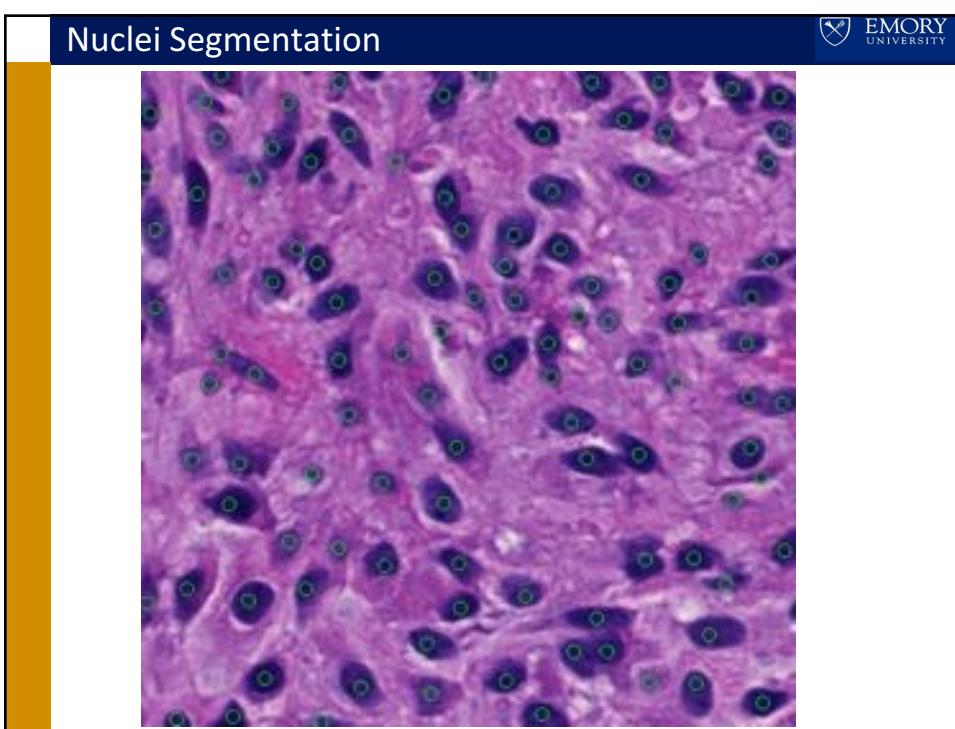
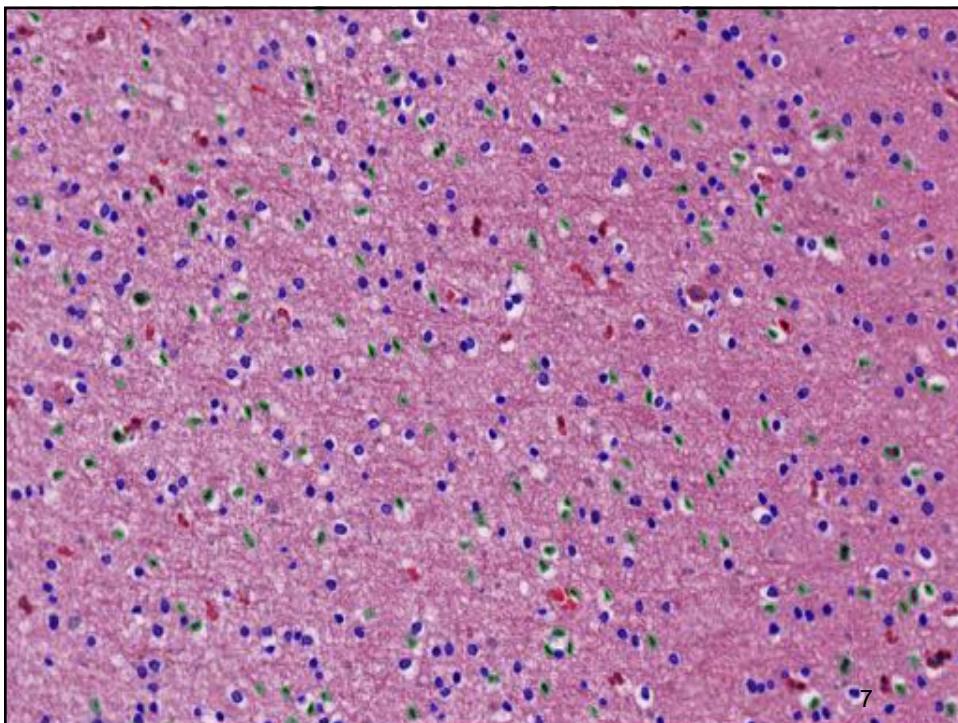


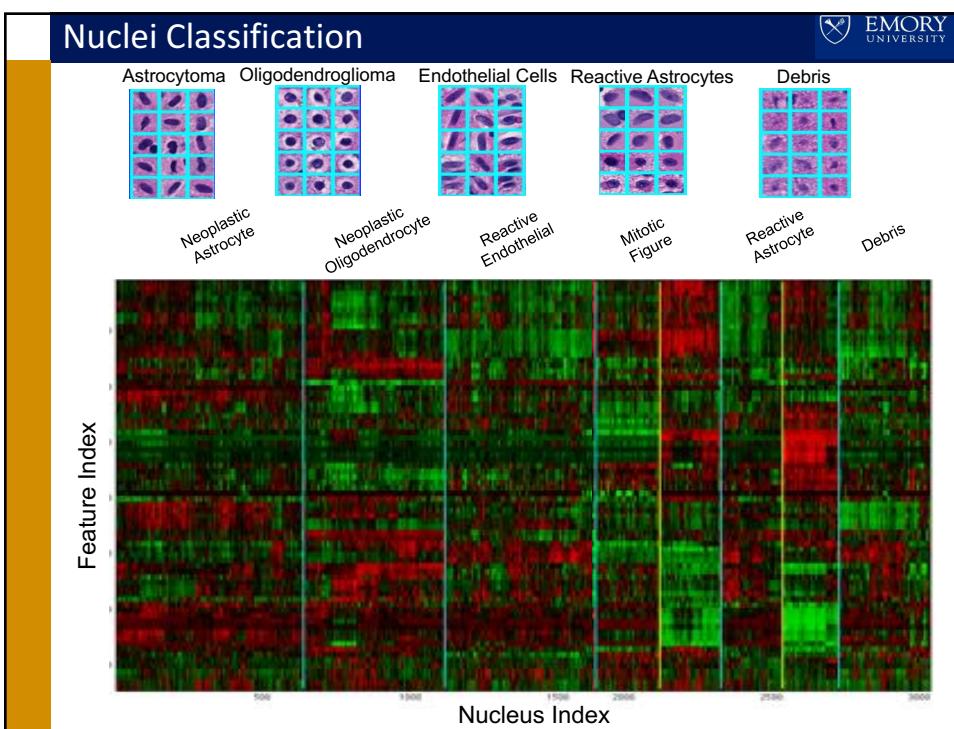
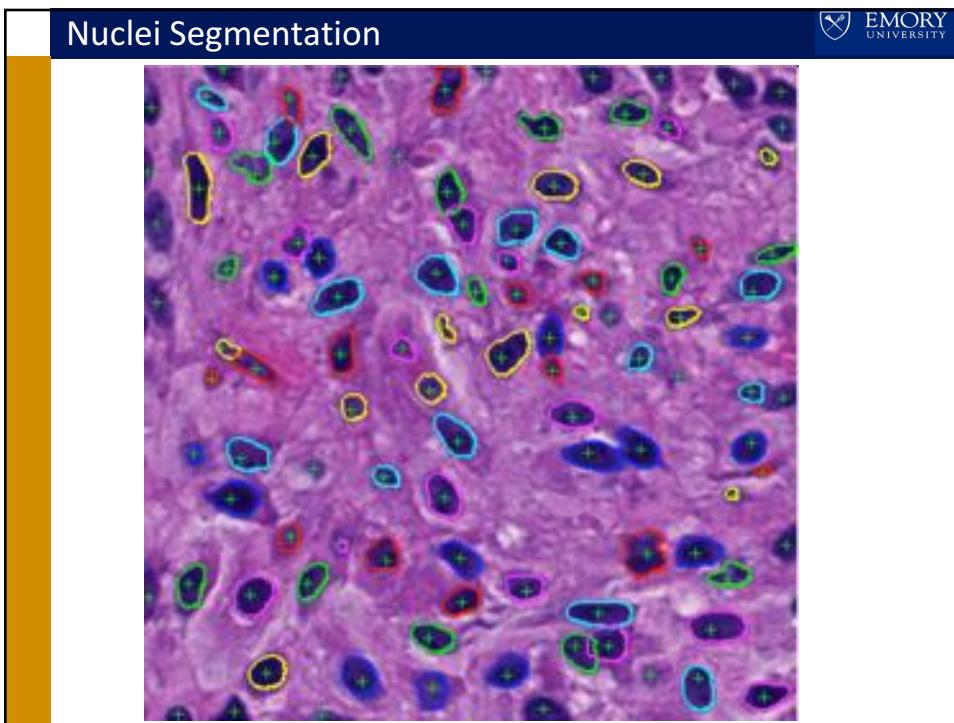
- Most frequent** brain cancers in the central nervous system
 - Uniformly fatal
 - Infiltrative and inclined to progression
 - Resistance to conventional therapies
 - Challenge on histological review
 - Large inter-reader variability
- Heterogeneity**
 - Grade IV Astrocytoma
 - Oligodendrogloma in abundance
 - Mixed Oligoastrocytomas
- Dismal** survival
 - 14 weeks** after resections on average
 - De Novo** for the majority



GBM = Devastation + Mystery!









Confusion Matrix Evaluation

The mean confusion matrices without and with SFFS (in %) are presented when QDA+MAP classification method is applied to 100 five-fold cross validation runs.

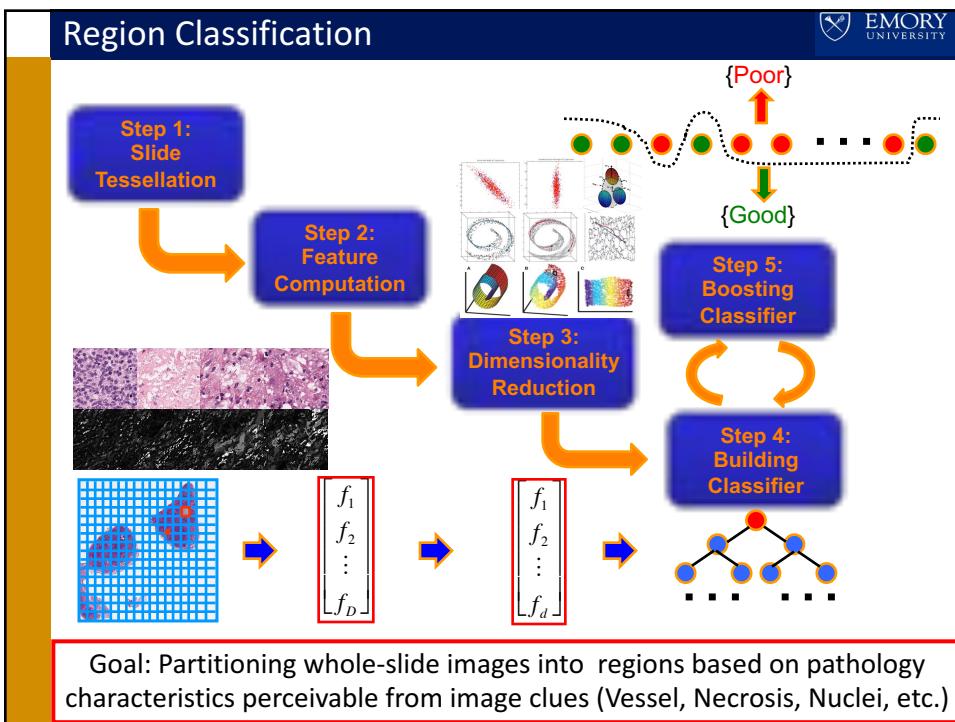
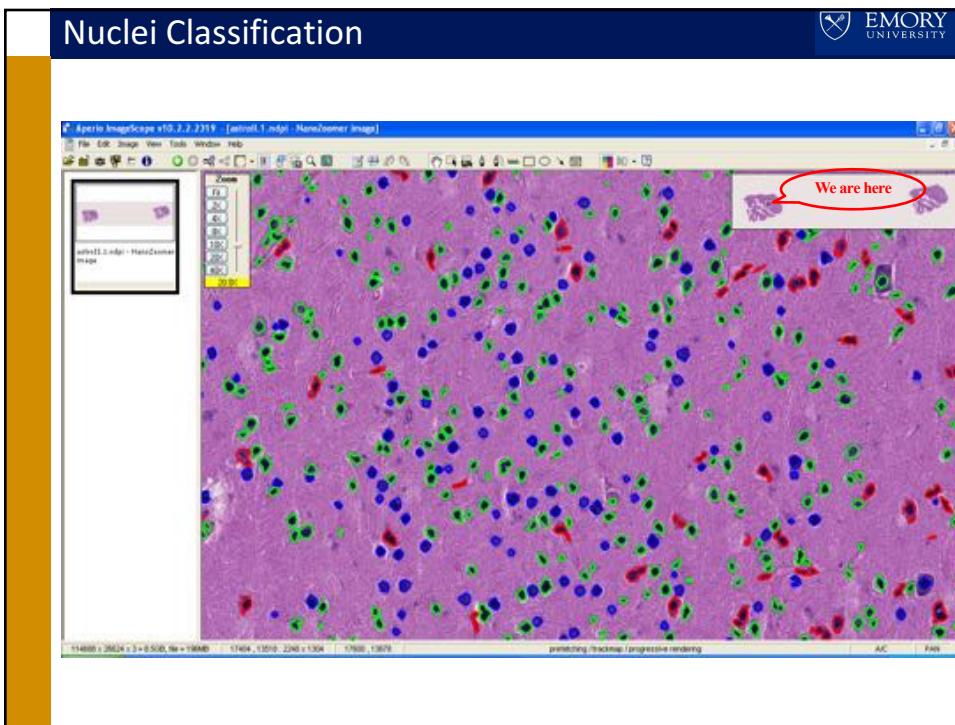
	Neoplastic Astrocyte	Neoplastic Oligodendrocyte	Reactive Endothelial	Mitotic Figure	Reactive Astrocyte	Debris
Neoplastic Astrocyte	89.84, 89.72	2.06, 1.75	5.78, 5.15	0.50, 0.79	0.59, 0.72	1.23, 1.87
Neoplastic Oligodendrocyte	4.02, 2.18	90.42, 92.39	3.12, 1.99	0.94, 1.34	0.15, 0.44	1.35, 1.66
Reactive Endothelial	14.33, 8.08	0.91, 0.71	80.26, 85.26	0.23, 0.19	0.41, 2.07	3.86, 3.69
Mitotic Figure	4.53, 2.14	1.05, 0.79	2.27, 1.34	86.42, 92.93	2.62, 1.25	3.11, 1.55
Reactive Astrocyte	11.13, 7.97	2.77, 2.89	19.95, 8.73	1.48, 1.78	63.24, 77.43	1.42, 1.20
Debris	10.00, 3.76	1.40, 1.35	11.65, 7.33	3.18, 2.24	0.12, 1.24	73.65, 84.08

Nuclei Classification



- The mean confusion matrices without and with SFFS (in %) are presented when QDA+MAP classification method is applied to 100 five-fold cross validation runs.

	GMM		QDA ("Nuc.", "Cyto.", "Nuc.+Cyto.")	
	Nuc	Nuc + Cyto	No SFFS	SFFS (20, 22, 34)
Neoplastic Astrocyte	72.90	89.62	74.47, 88.46, 89.84	73.82, 85.46, 89.72
Neoplastic Oligodendrocyte	85.02	90.53	84.94, 77.23, 90.42	86.14, 79.48, 92.39
Reactive Endothelial	64.19	80.36	64.64, 66.57, 80.26	66.13, 72.84, 85.26
Mitotic Figure	79.07	86.45	78.44, 77.07, 86.42	78.77, 84.63, 92.93
Reactive Astrocyte	52.90	63.68	51.17, 54.40, 63.25	51.87, 66.24, 77.43
Debris	64.99	74.14	63.31, 60.53, 73.65	64.42, 66.07, 84.08
Overall	70.53	82.08	70.44, 72.63, 81.97	71.04, 77.00, 87.43





Method -- Feature Computation

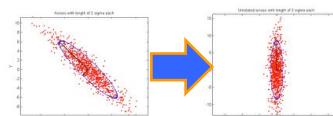
- Co-occurrence Matrix Features
 - the second-order statistics are demonstrated playing crucial roles in the human texture perception
- Tamura Features
 - perceptual relevant features proved to be highly correlated with human perceptions;
- Wavelet Features
 - multi-scale features from decomposed image channels;
- Neighborhood Difference Matrix Features
 - texture features particularly sensitive to the spatial intensity change;
- Gray Level Run Length Matrix Features
 - higher-order statistical texture features;



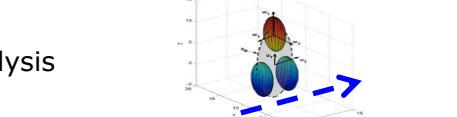
Method – Dimensionality Reduction

- For comparison purposes, we use and compare the following methods with our data:

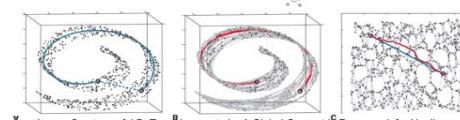
– Principal Component Analysis



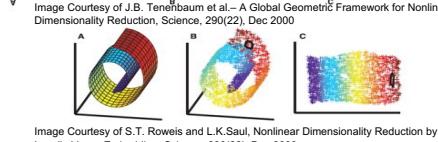
– Linear Discriminant Analysis

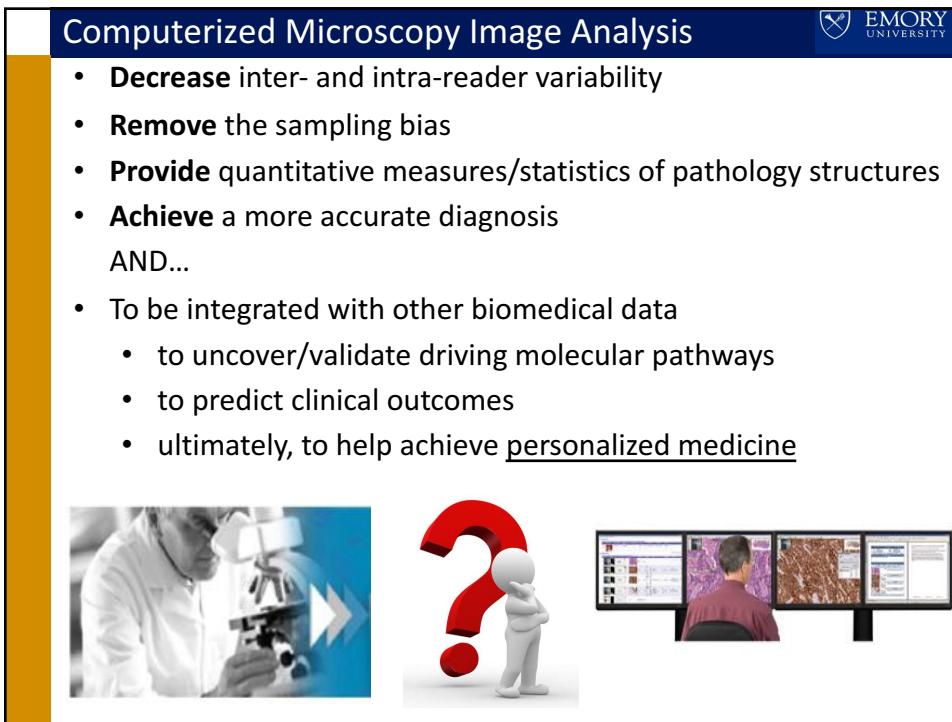
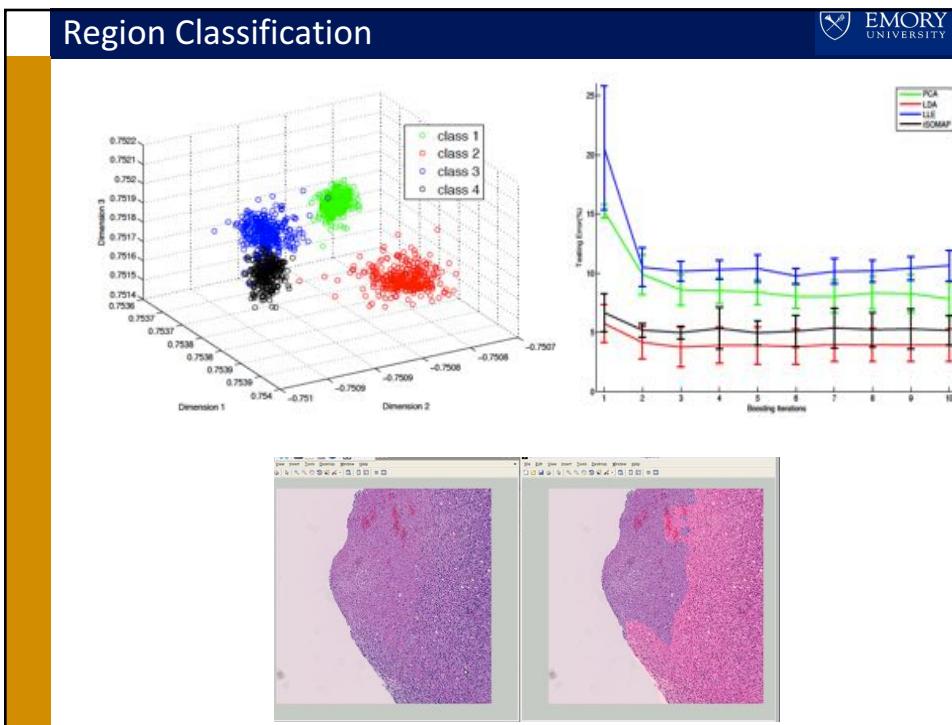


– Iso-Metric Mapping



– Local Linear Embedding



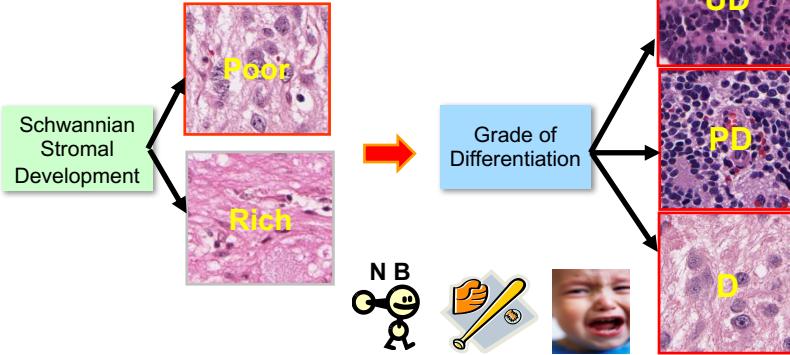


Computer Aided Diagnosis



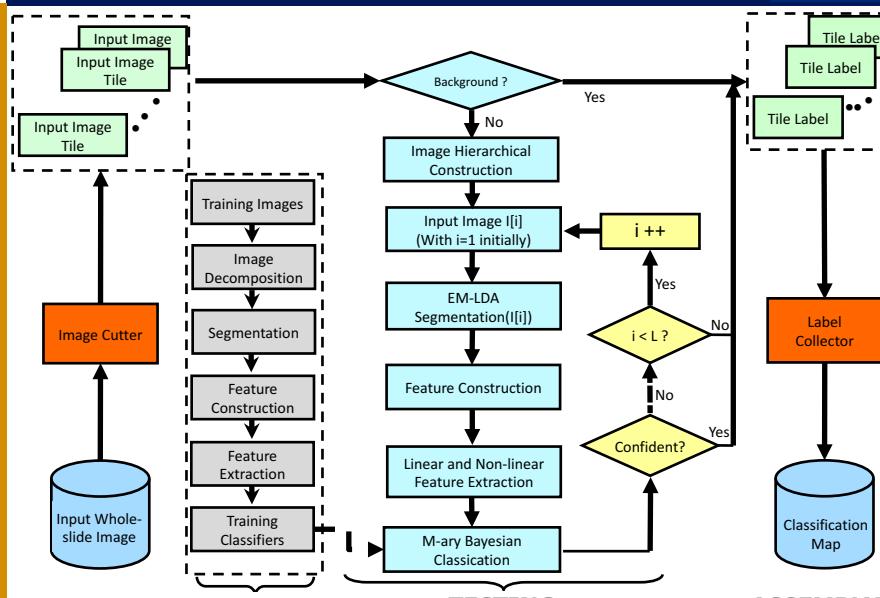
- Neuroblastoma is a cancer of sympathetic nervous system
- *Neuro* indicates "nerves", and *Blastoma* refers to a cancer that affects immature or developing cells.

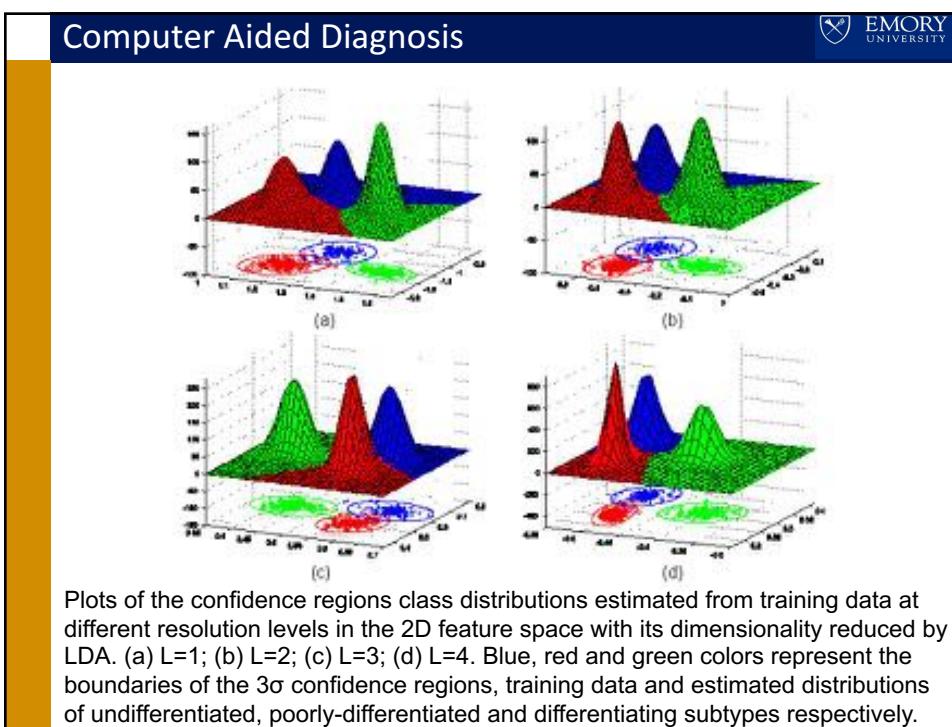
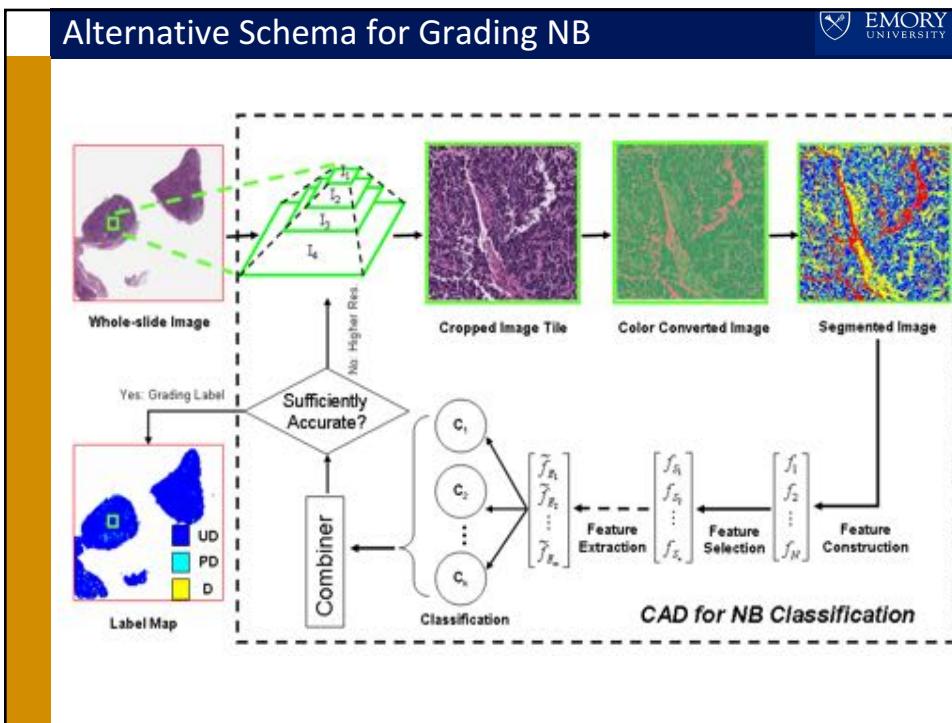
International NB Pathology Classification

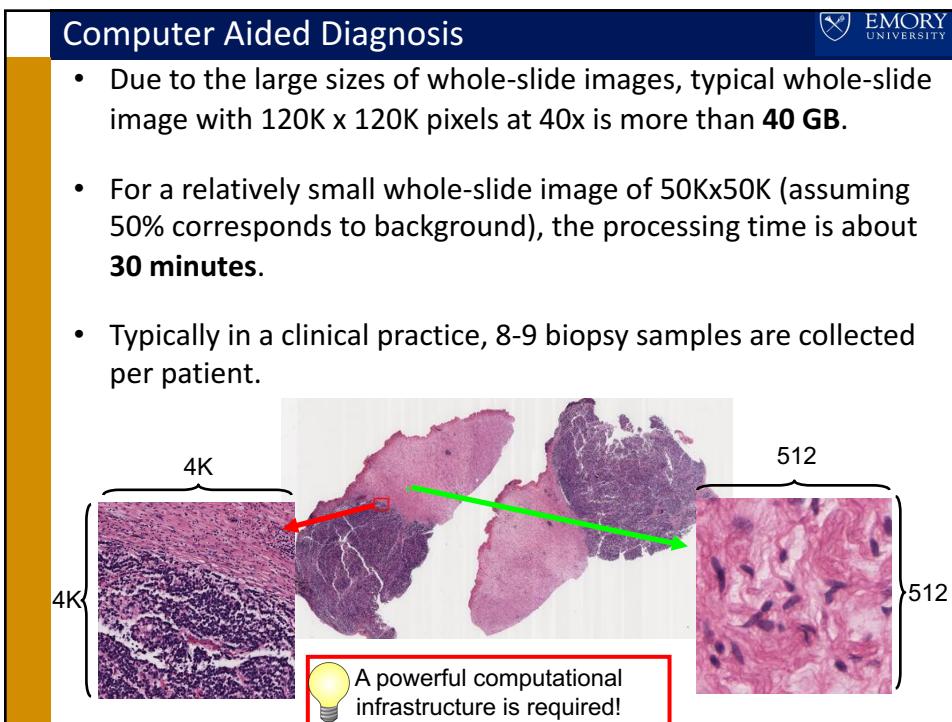
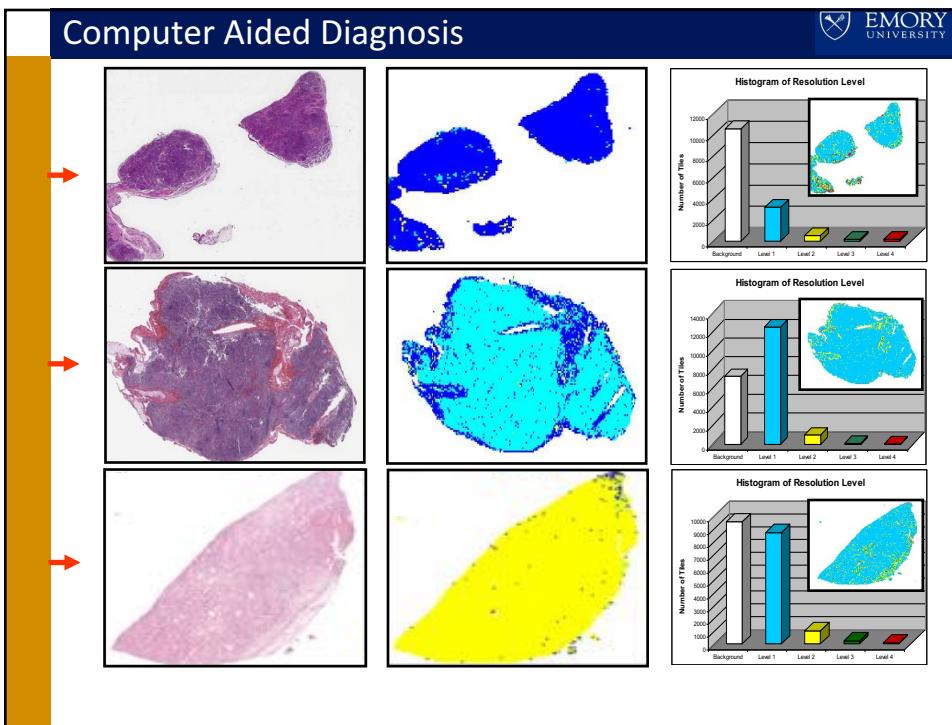


Abstract/Genetic Question: X-class classification given tissue content

Computer Aided Diagnosis



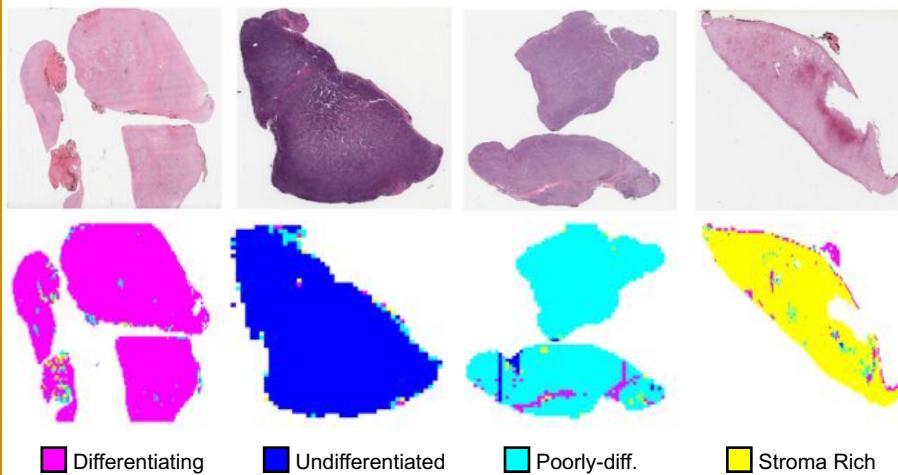




Computer Aided Diagnosis

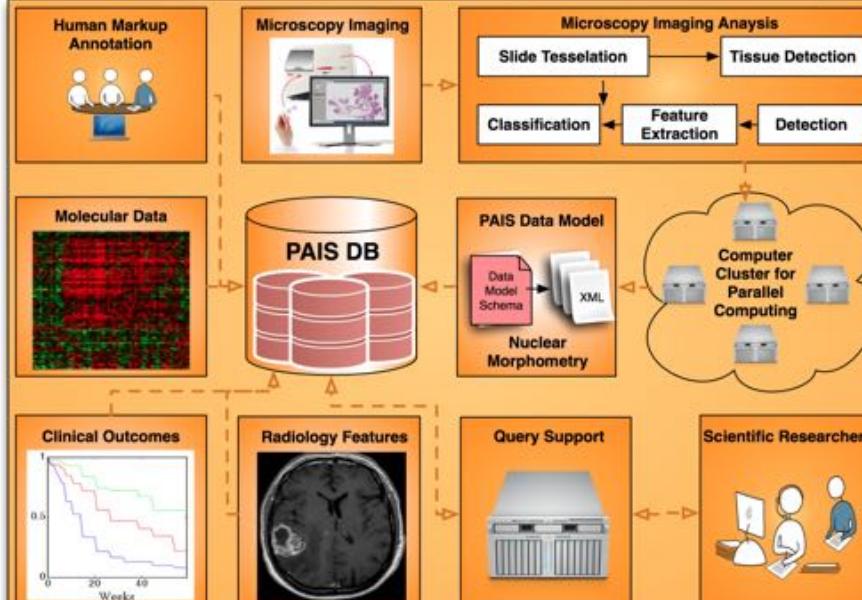


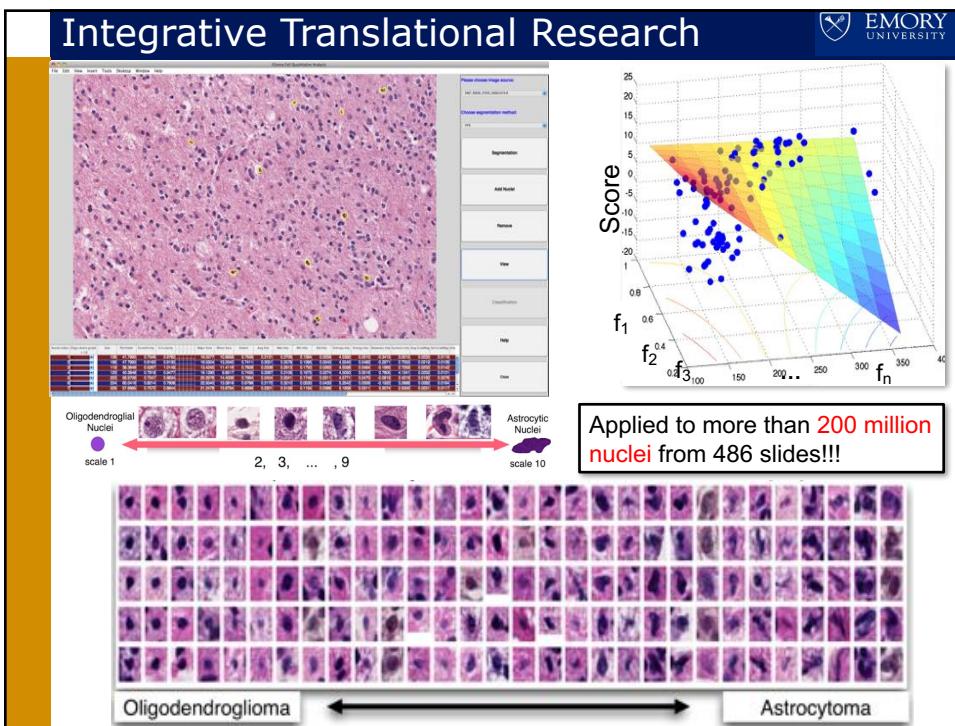
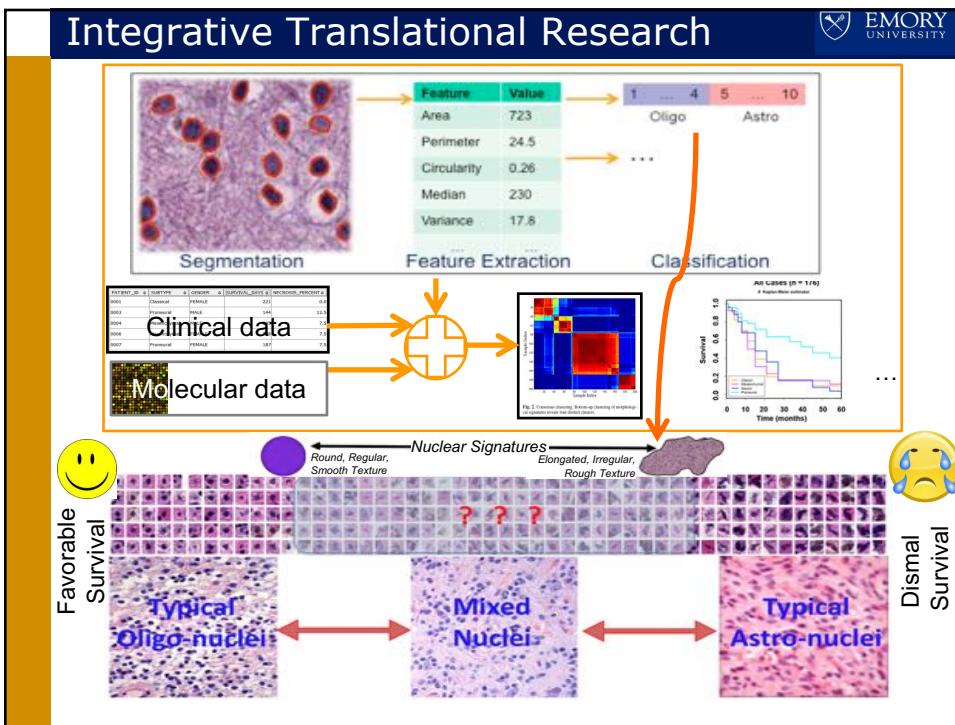
- Problems associated with digital pathology is that scalable image analysis algorithms need to be developed to remedy image blurring, artifacts, variation, and large scale ($\sim 100\text{k} \times 100\text{k}$)

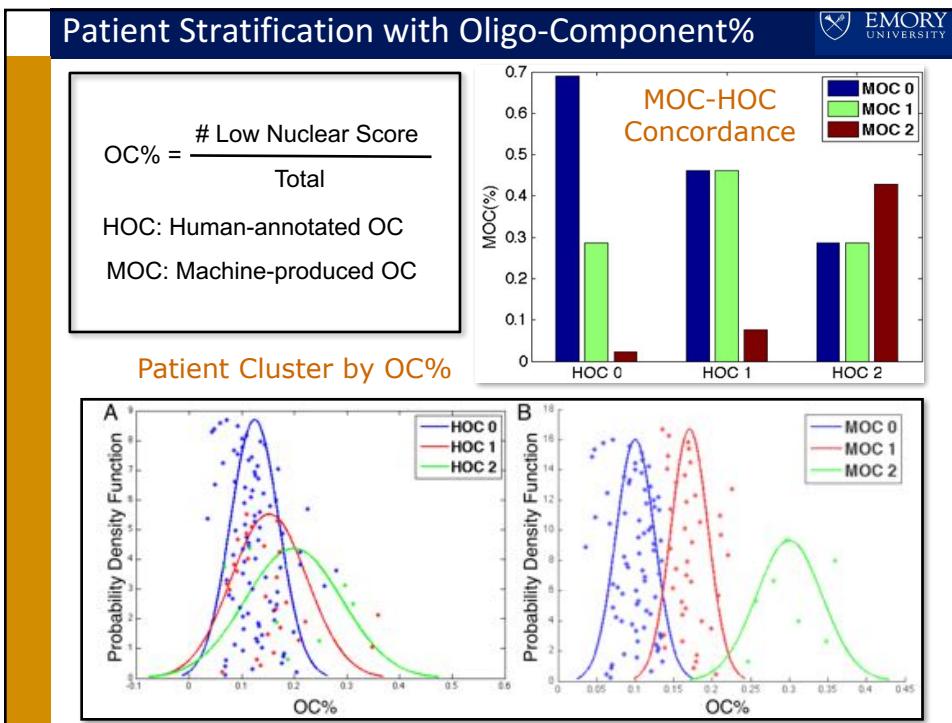


Each pixel in the classification map is associated with a 512×512 image region.

Integrative Translational Research



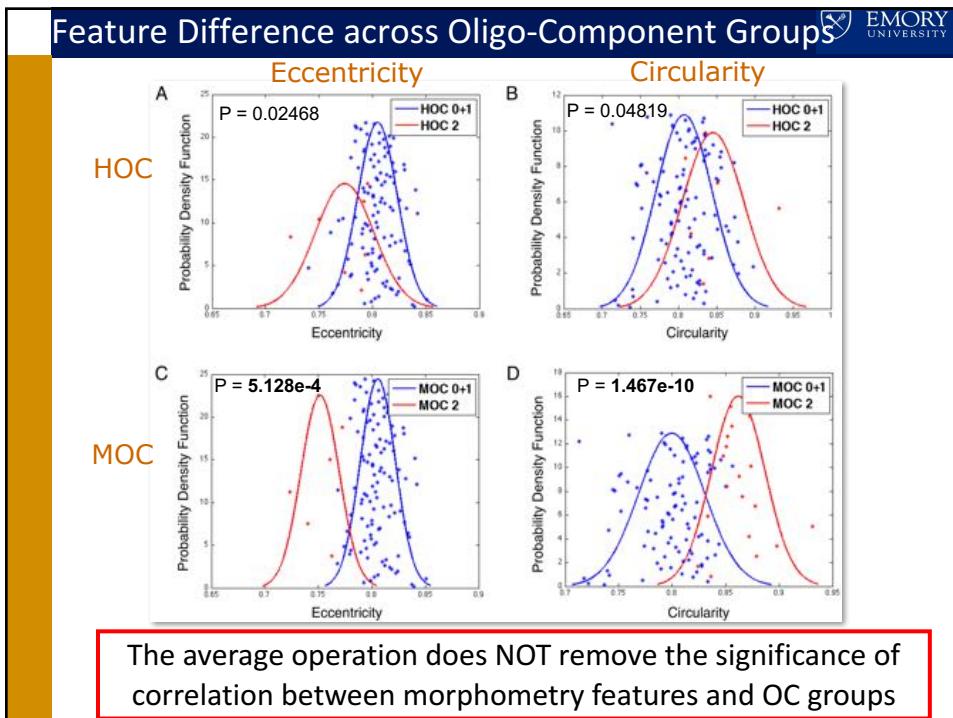
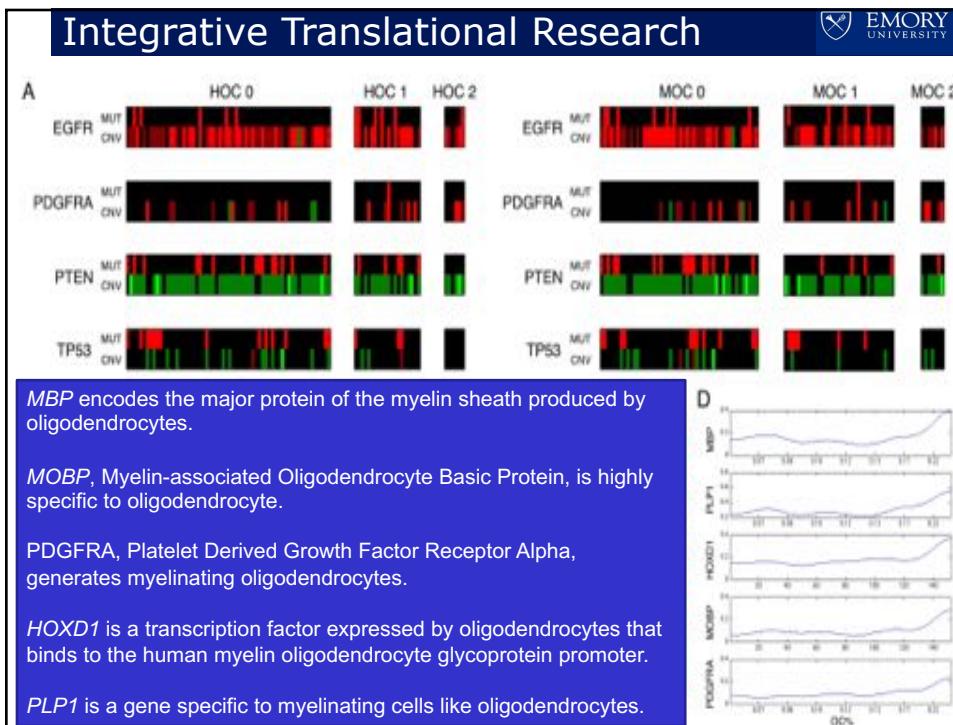


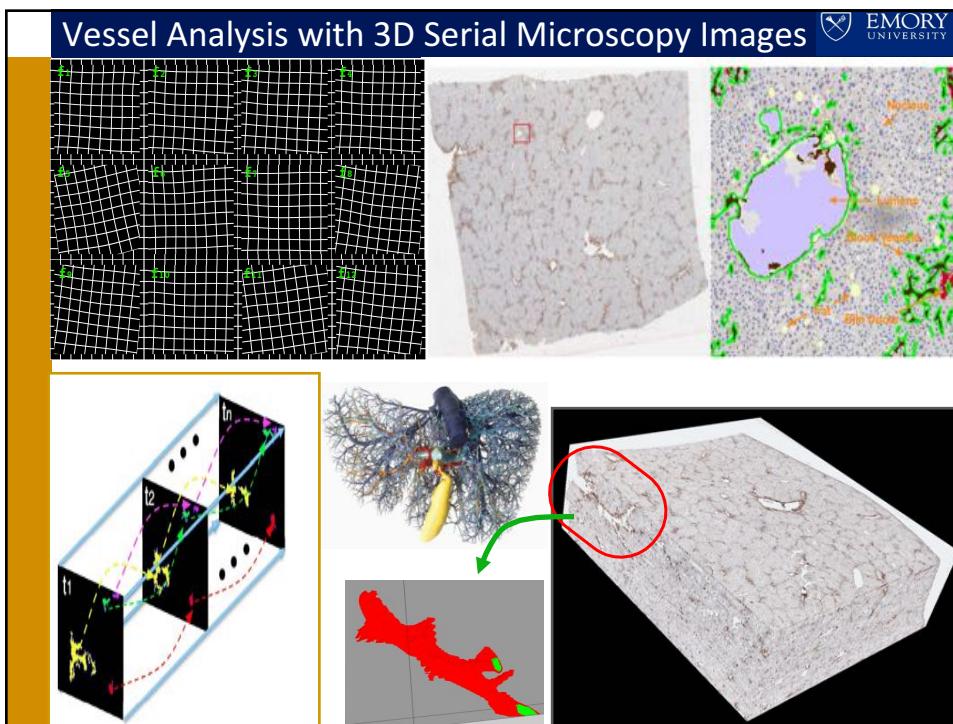
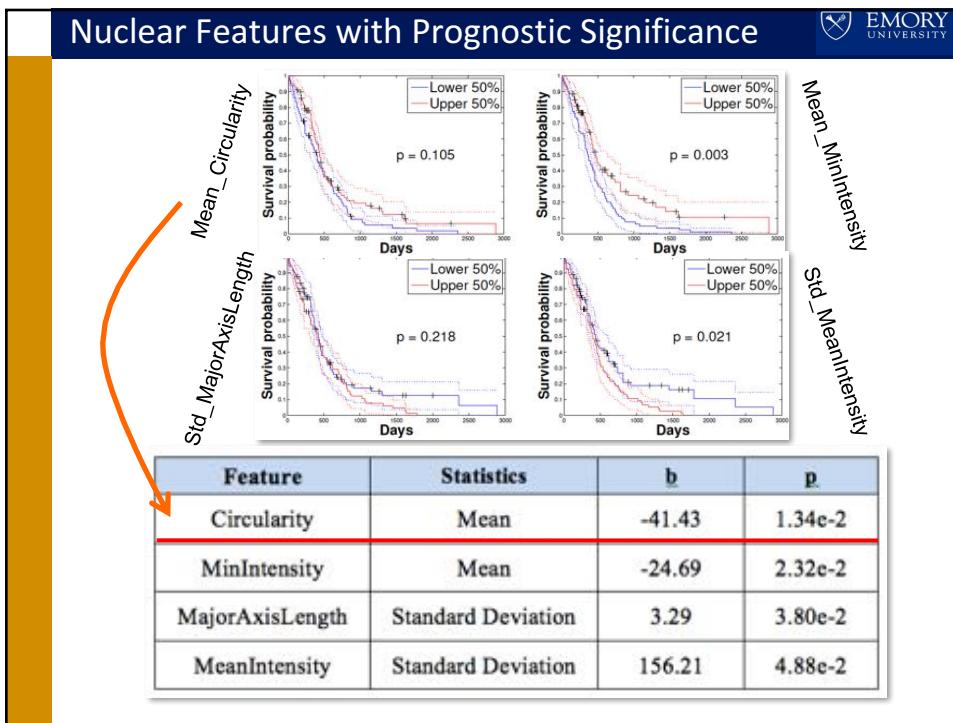


Integrative Translational Research

EMORY UNIVERSITY

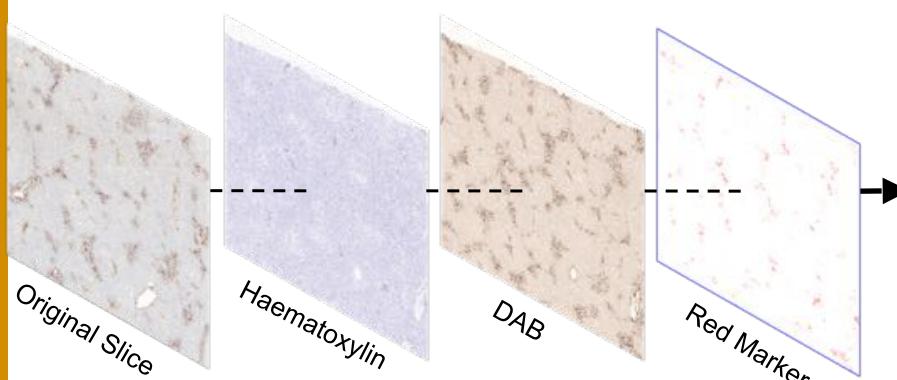
Association	MOC-0	MOC-1	MOC-2
Concordance	Enriched in HOC-0	Similar to HOC-1	Enriched in HOC-2
Gene Expression Class Association	Enriched by Classical	None	Enriched by Proneural
Somatic Mutation Association	Enriched by PTEN and RB1 mutant	Depleted by PTEN mutatnt	None
Copy Number Association	None	None	Enriched by PDGFRA amplification
Gene Expression Association	25 over expressed; 242 under expressed	2 under expressed	MOPB, MBP over expressed (4 total)
Genes Correlated with OC%	MBP, HOXD1, PLP1, MOBP, and PDGFRA (194 total)		





Vessel Analysis with 3D Serial Microscopy Images  EMORY UNIVERSITY

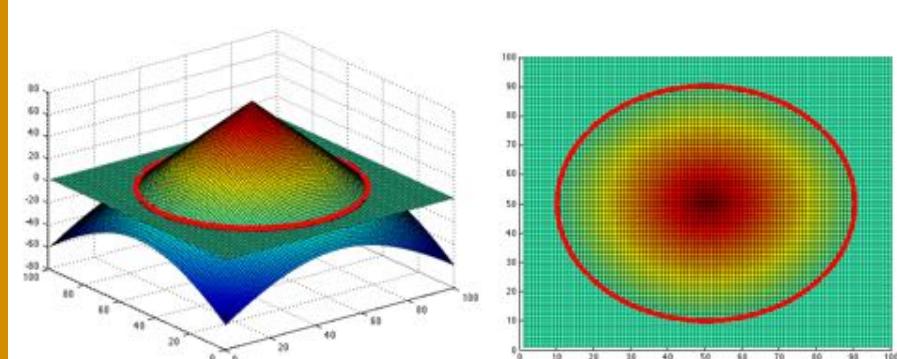
Color Deconvolution



- Find each stain channel
 - Vessels are dyed by DAB stain

Vessel Analysis with 3D Serial Microscopy Images  EMORY UNIVERSITY

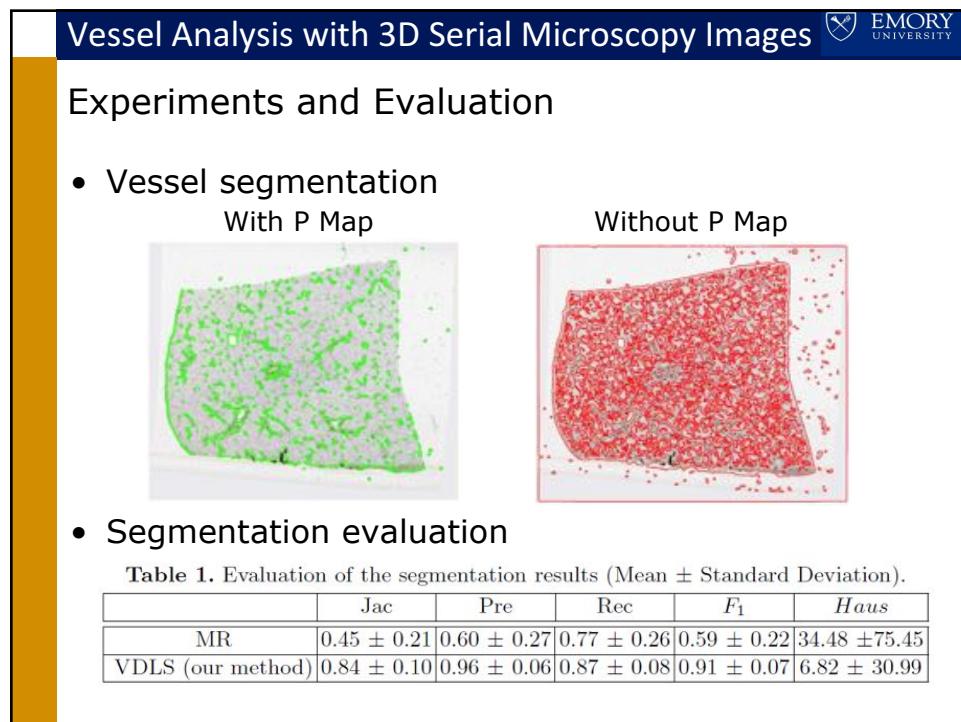
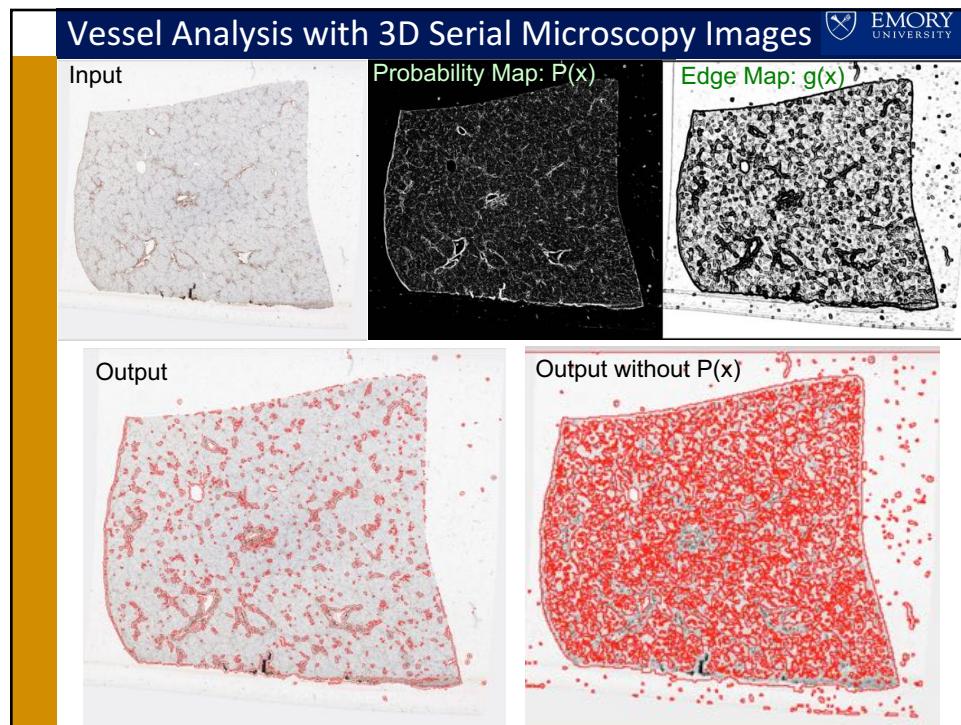
A Variational Level Set Method

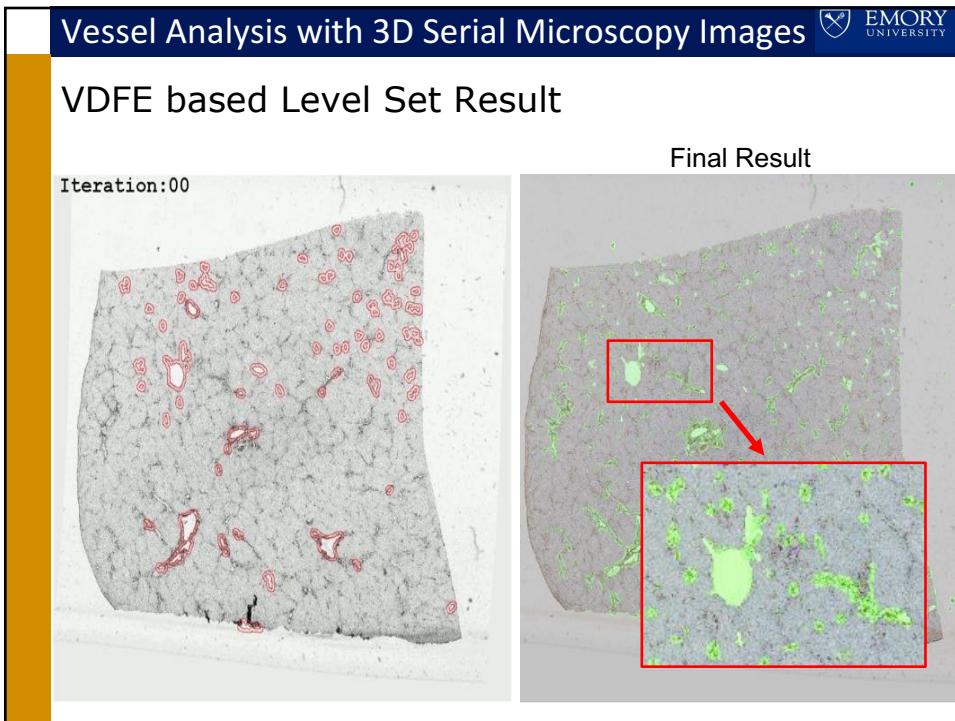


Curve C is represented via a Lipschitz function Φ by $C = \{(x, y) | \phi(x, y) = 0\}$

$$\frac{\partial \phi}{\partial t} = |\nabla \phi| F, \quad \phi(0, x, y) = \phi_0(x, y)$$

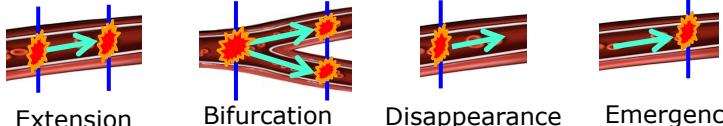
$$\begin{cases} C = \partial \omega = \{(x, y) \in \Omega : \phi(x, y) = 0\}, \\ inside(C) = \omega = \{(x, y) \in \Omega : \phi(x, y) > 0\} \\ outside(C) = \Omega \setminus \bar{\omega} = \{(x, y) \in \Omega : \phi(x, y) < 0\} \end{cases}$$





Vessel Association  EMORY UNIVERSITY

Two-stage Vessel Association

- Local bi-slide mapping and global vessel structure association
 - Four vessel association cases
 
 - Similarity functions
 - Shape descriptor, spatial relationship.
 - .One-to-one: $s(v_i^t, v_j^{t+1}) = \mu_1 g(v_i^t, v_j^{t+1}) + \mu_2 d(v_i^t, v_j^{t+1})$
 - One-to-two: $s(v_i^t, v_{j_1}^{t+1}, v_{j_2}^{t+1}) = \mu_1 g(v_i^t, v_{j_1}^{t+1} \cup v_{j_2}^{t+1}) + \mu_2 d(v_i^t, v_{j_1}^{t+1} \cup v_{j_2}^{t+1})$
 - One-to-none: $s(v_i^t, v_\emptyset^{t+1}) = d(v_i^t, \Omega_t)$
 - None-to-one: $s(v_\emptyset^{t-1}, v_i^t) = d(v_i^t, O_t)$

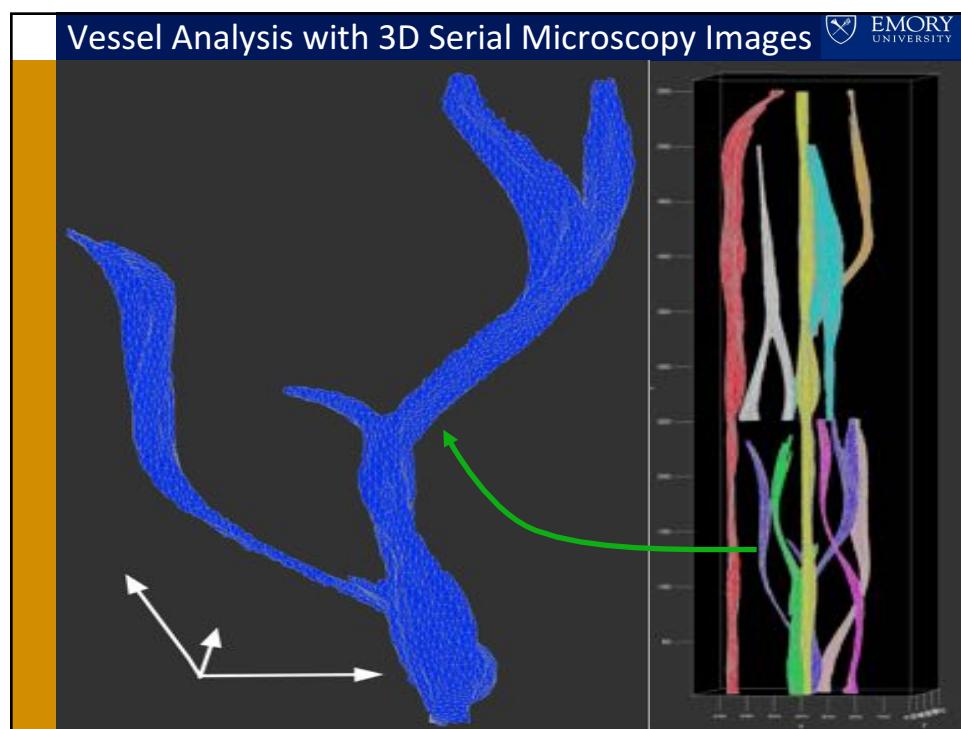
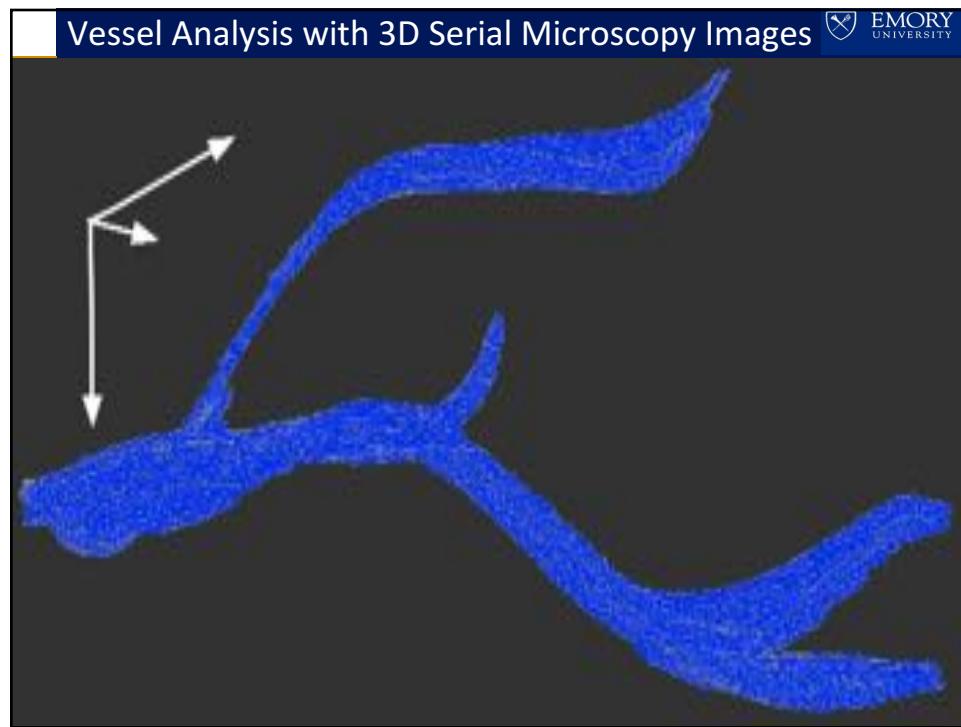
Vessel Association

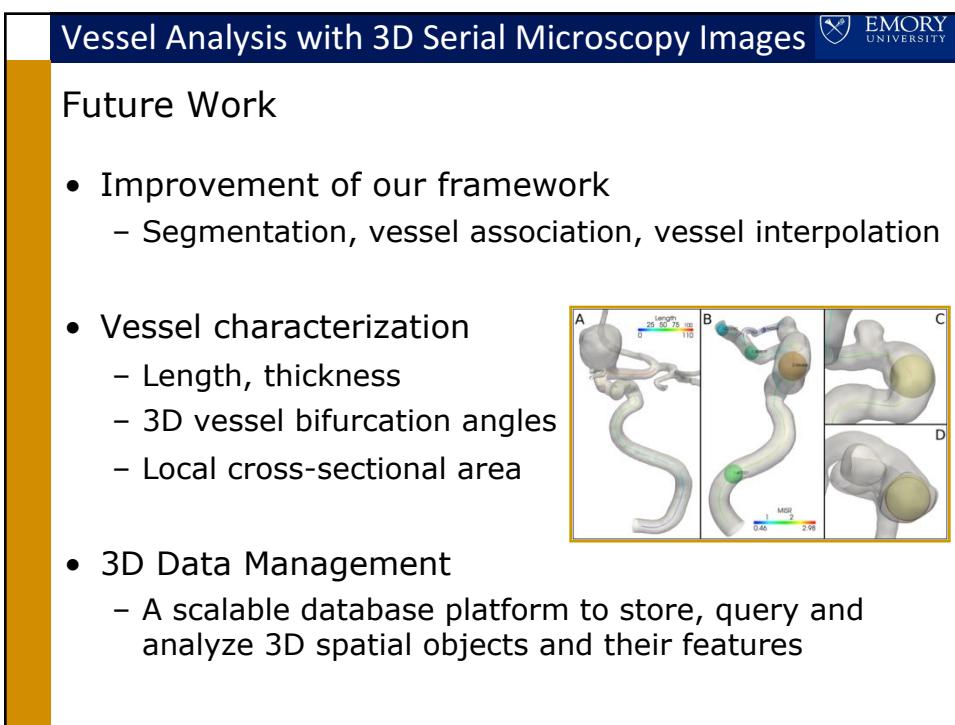
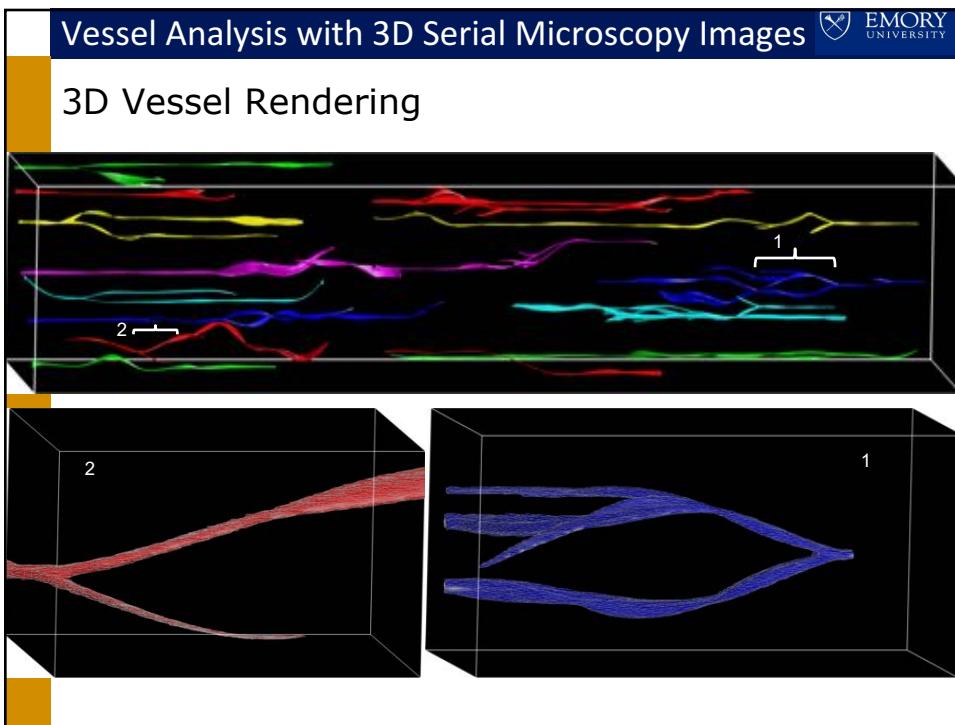
Two-stage Vessel Association

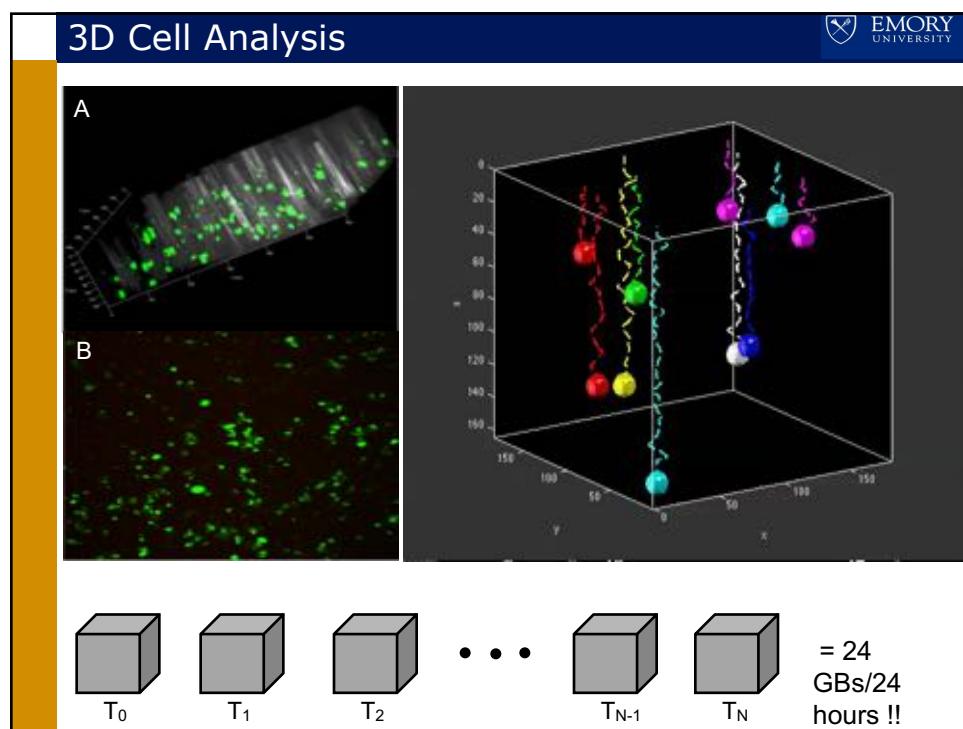
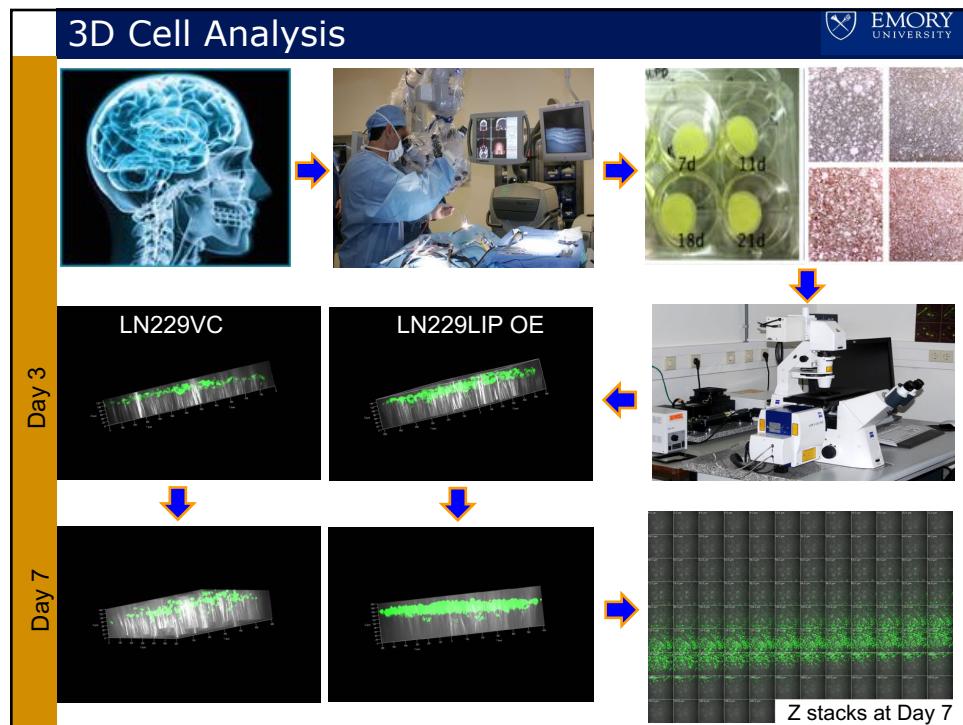
- Global vessel structure association
 - Bayesian Maximum A Posteriori (MAP) framework

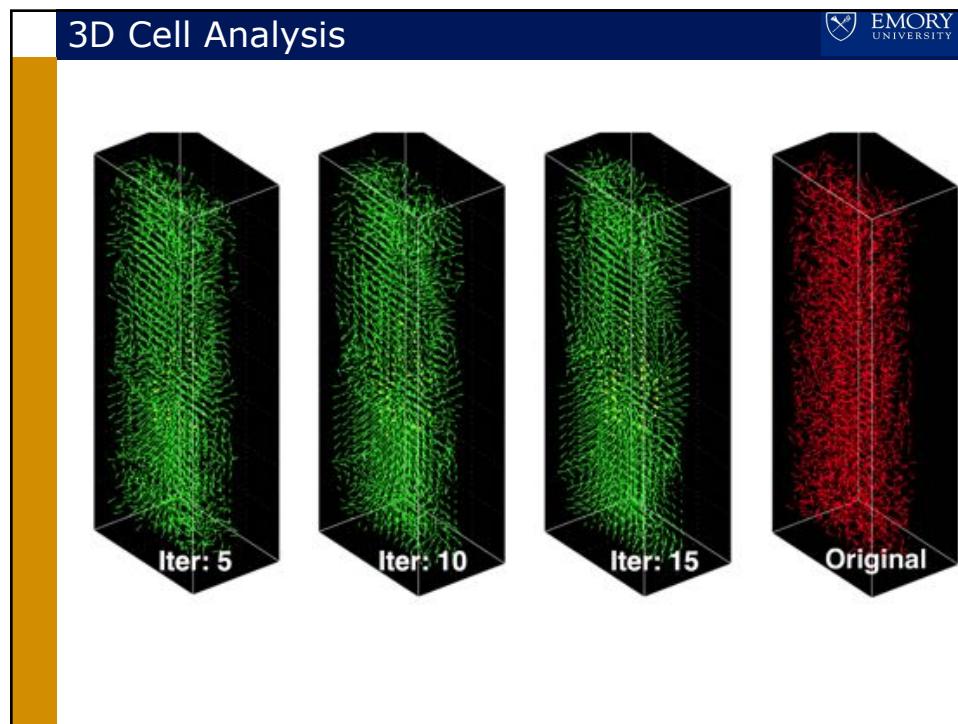
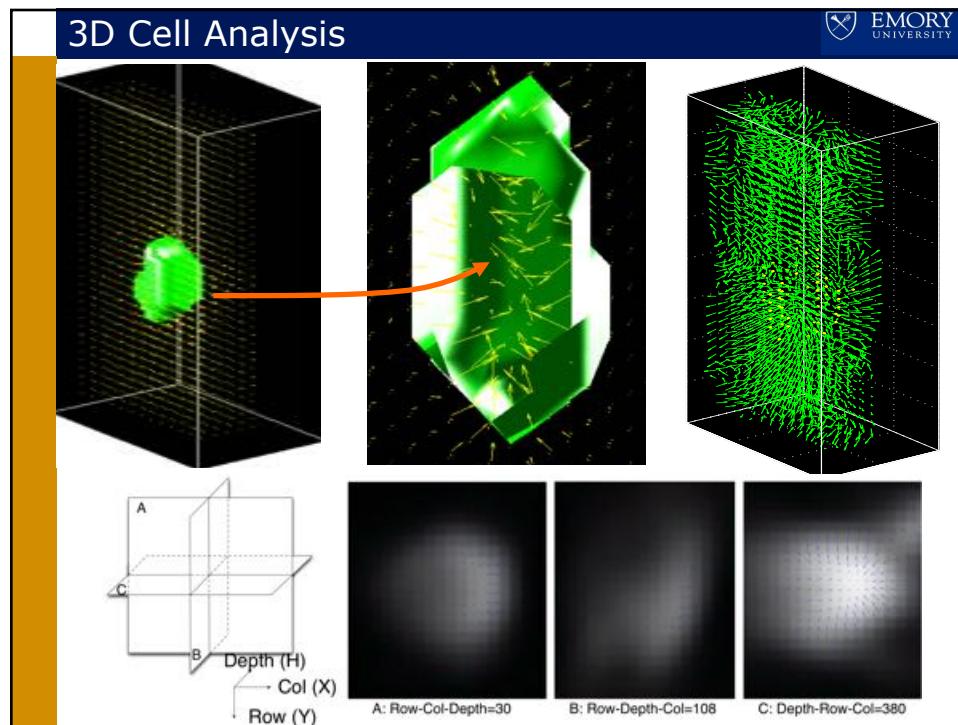
$$\mathbf{V}^* = \arg \max_{\mathbf{V}} P(\mathbf{V}|\mathbf{B}) = \arg \max_{\mathbf{V}} \prod_{V_k \in \mathbf{V}} P(V_k|\mathbf{B}) = P_{\emptyset \rightarrow 1}(B_s^k|B_\emptyset)$$

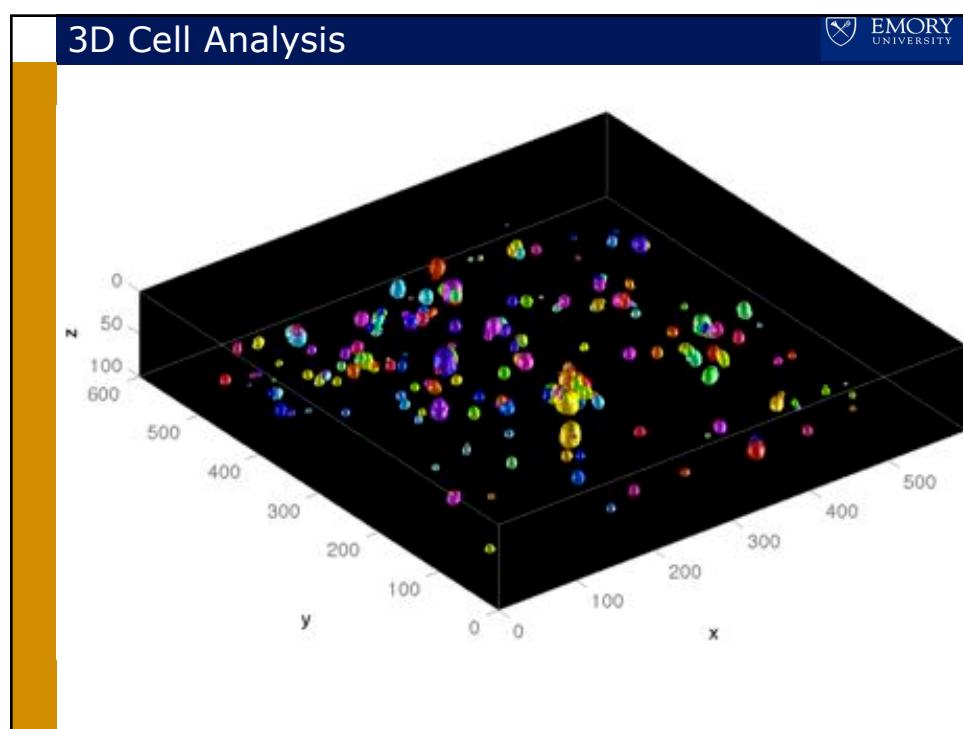
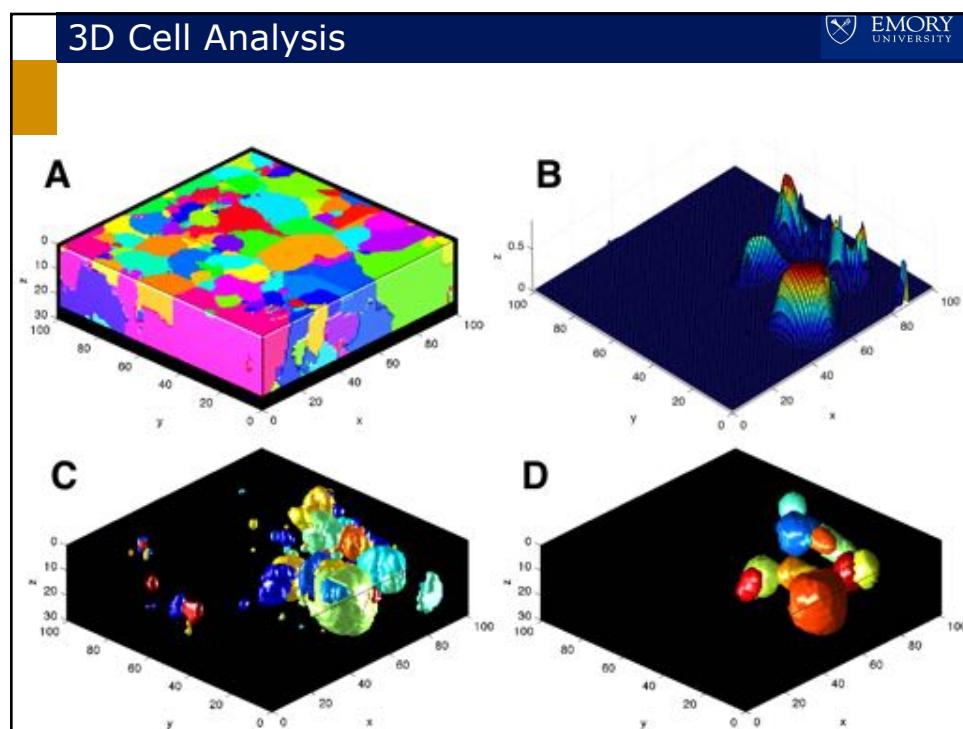
$$\prod_{B_i^k, B_j^k \in V_k} P_{1 \rightarrow 1}(B_j^k|B_i^k) \quad \prod_{B_m^k, B_{n_1}^k, B_{n_2}^k \in V_k} P_{1 \rightarrow 2}(B_{n_1}^k, B_{n_2}^k|B_m^k) \quad \prod_{B_e^k \in V_k} P_{1 \rightarrow \emptyset}(B_\emptyset|B_e^k)$$

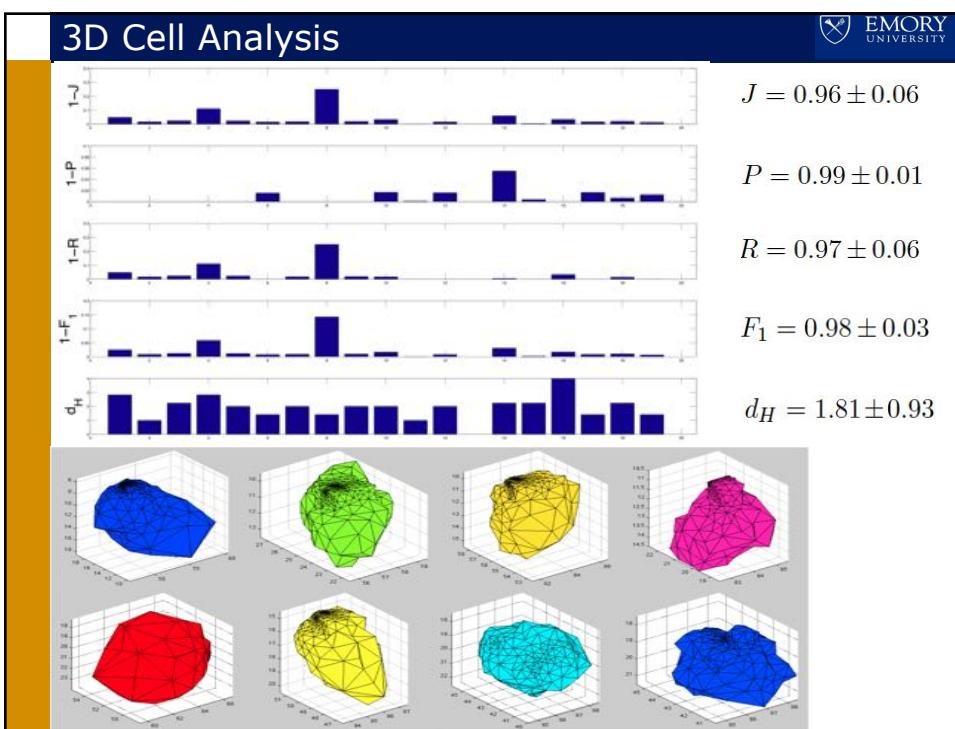
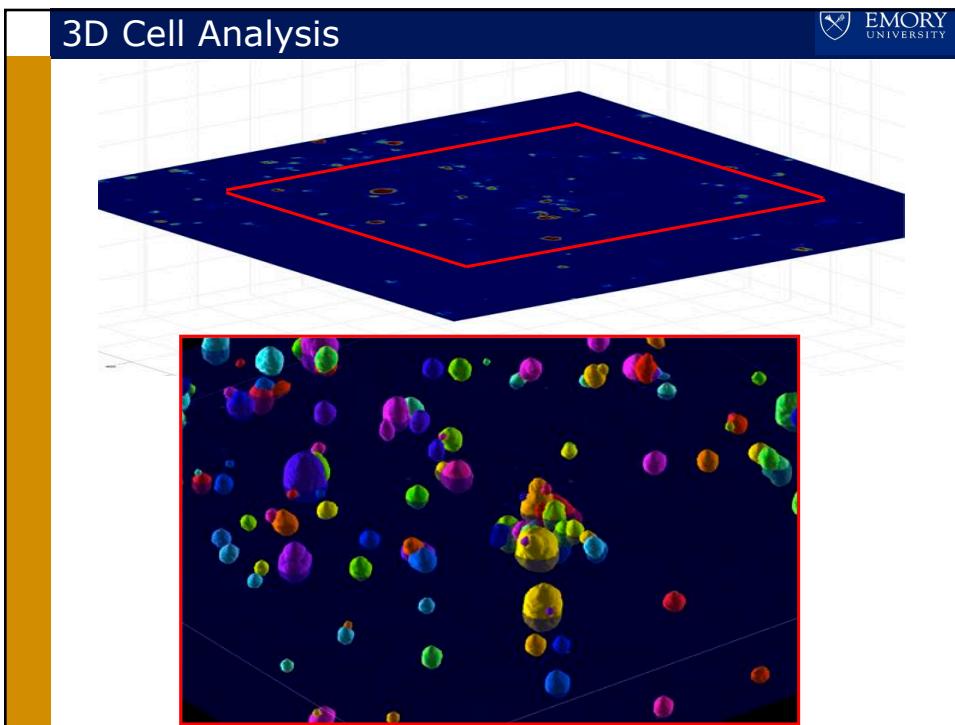





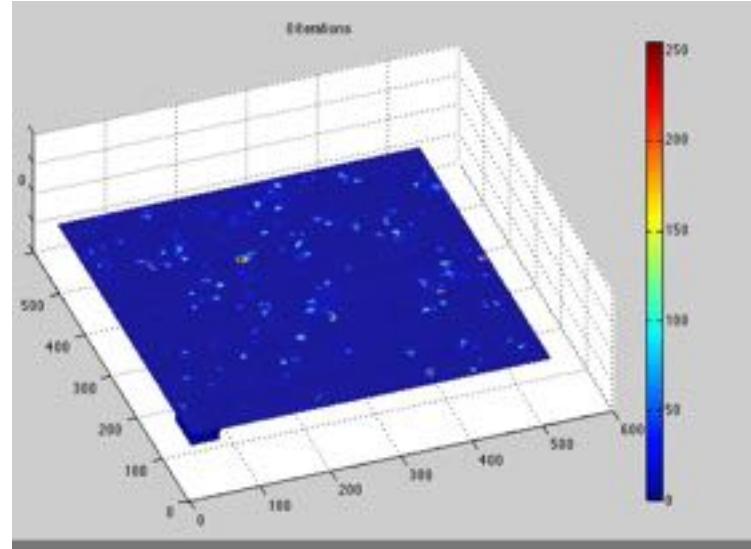




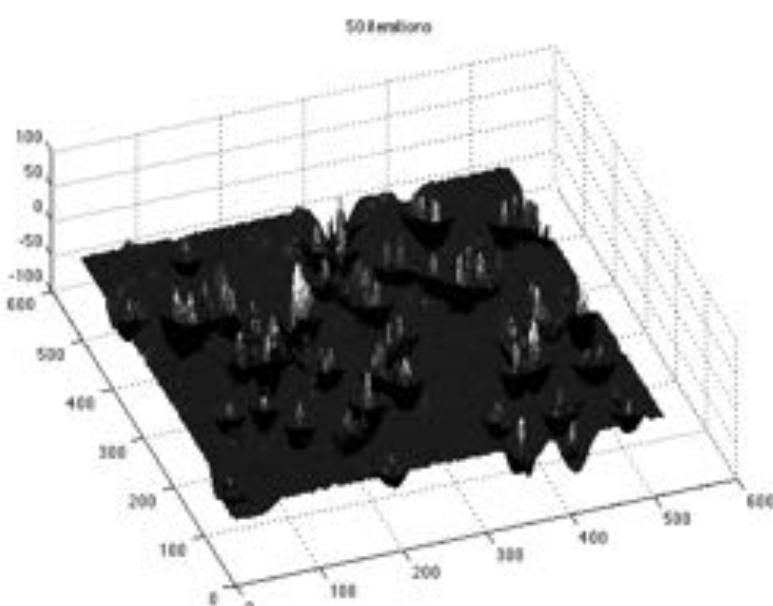


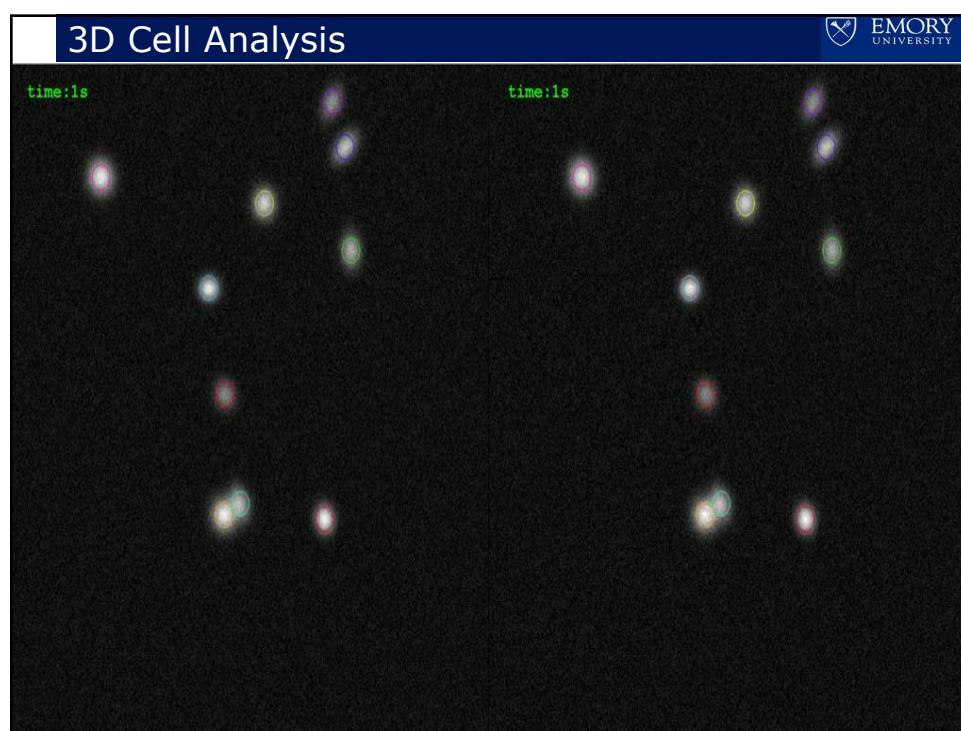
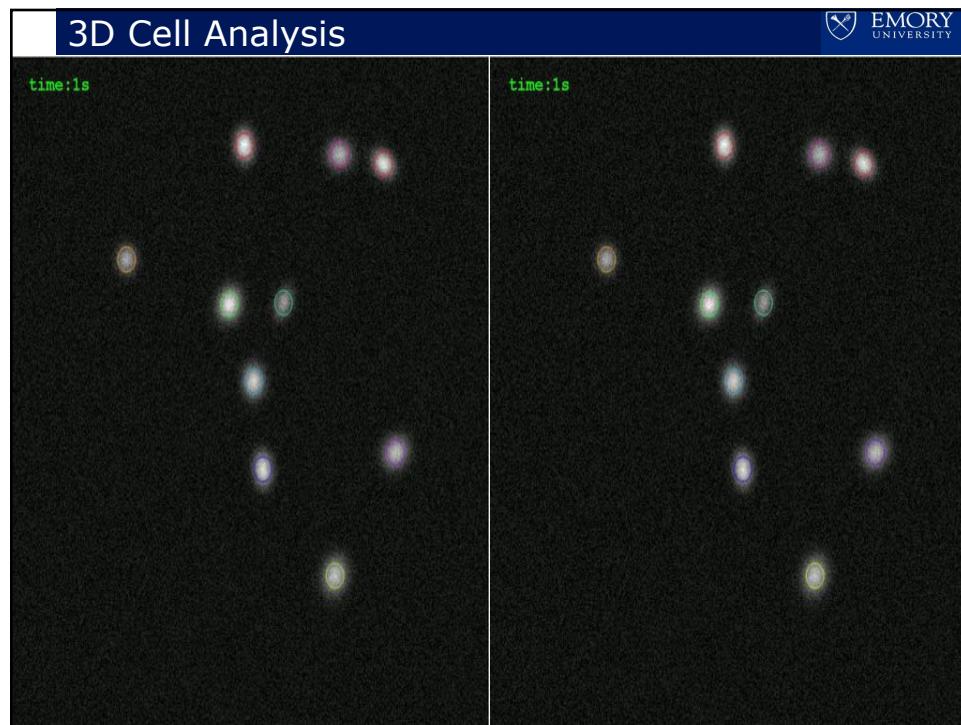


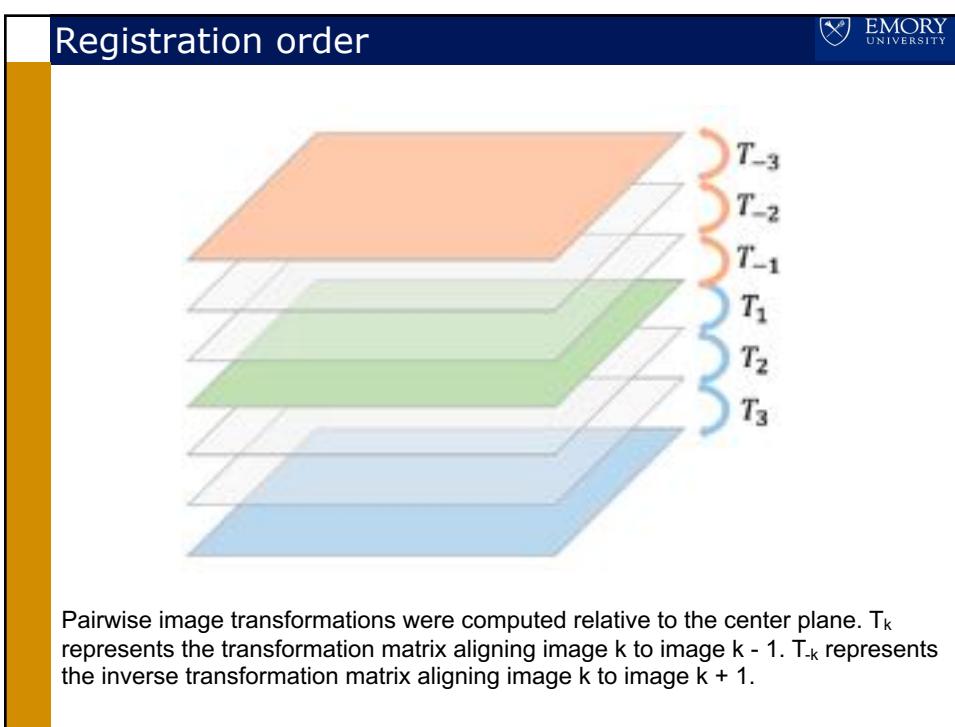
This Method is Generic

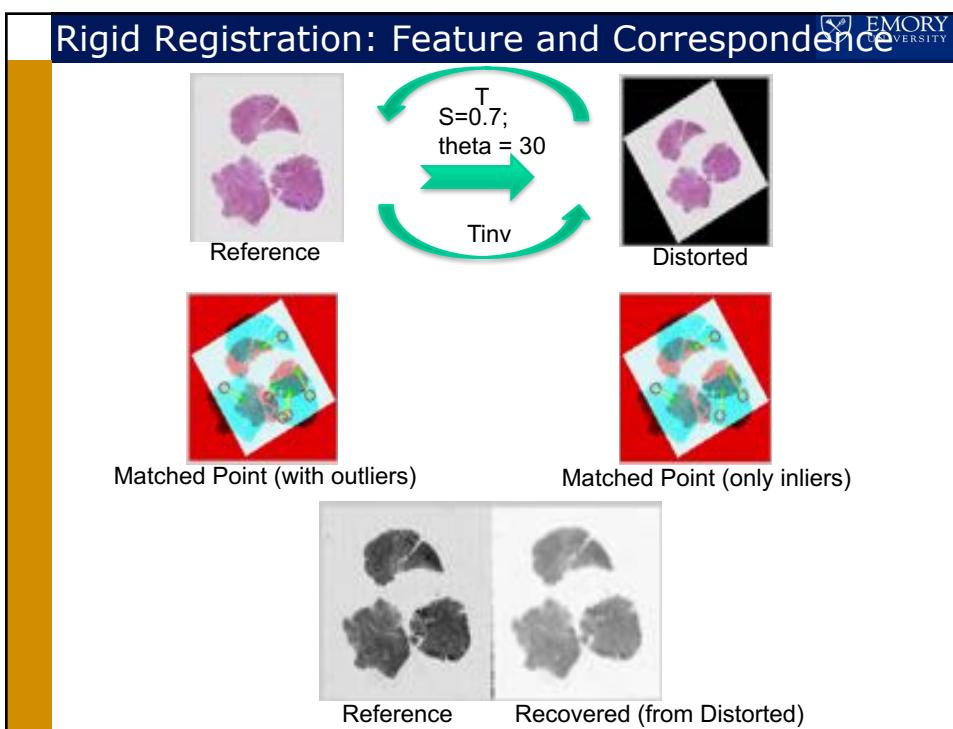
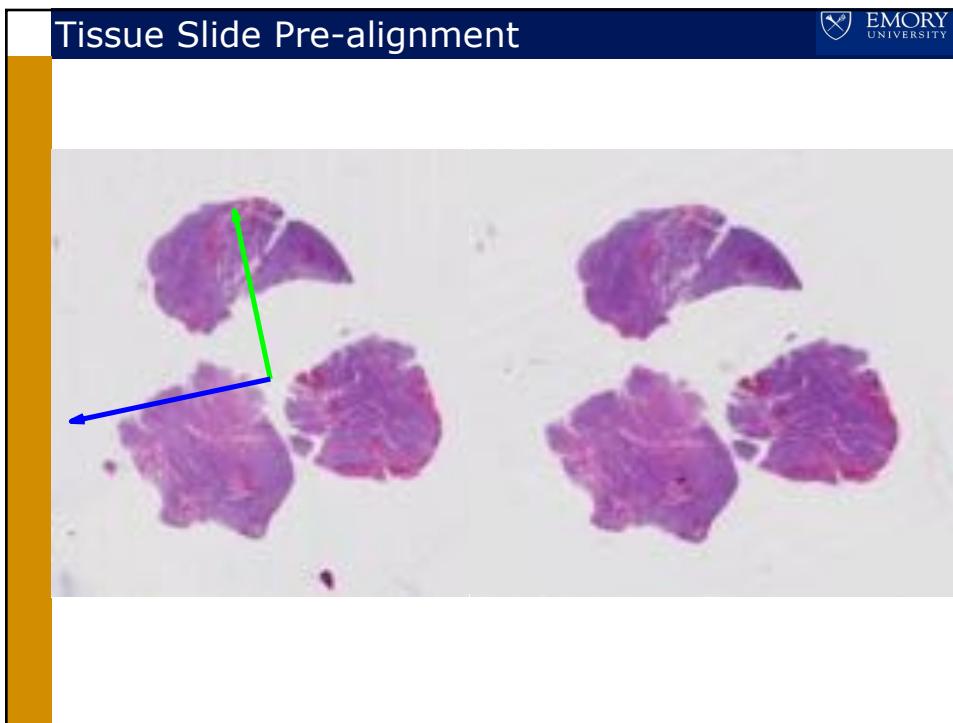


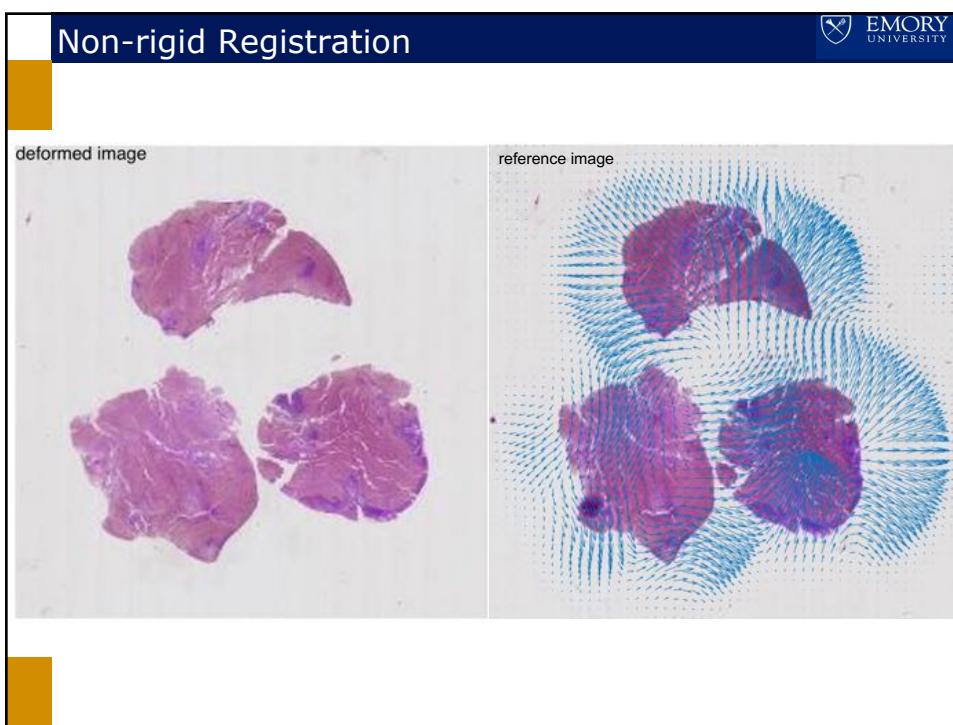
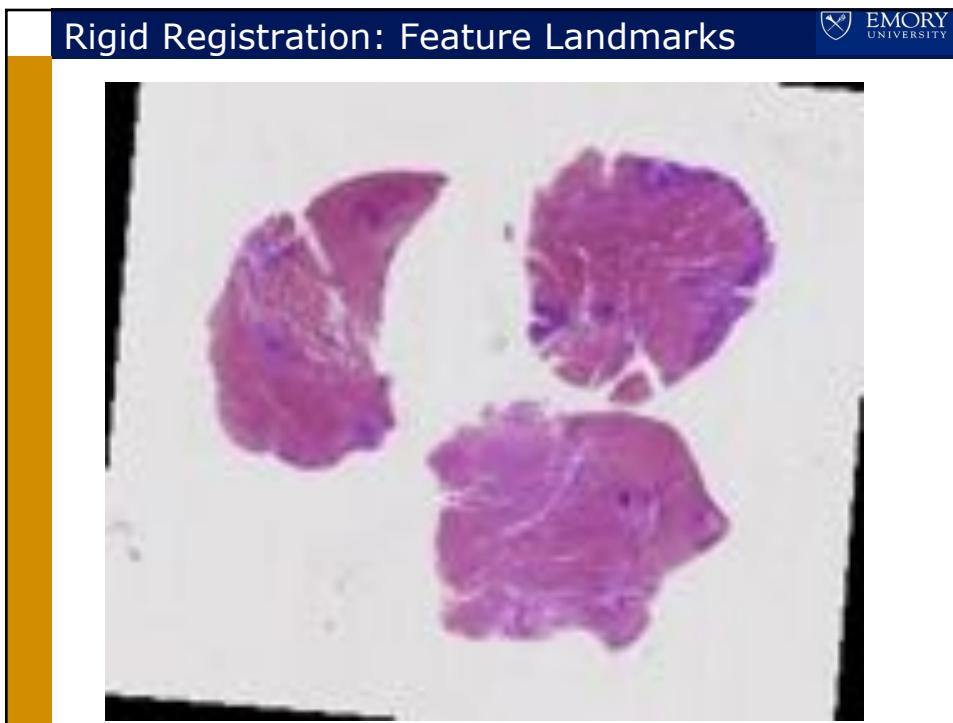
Another Possible Cell Segmentation Method

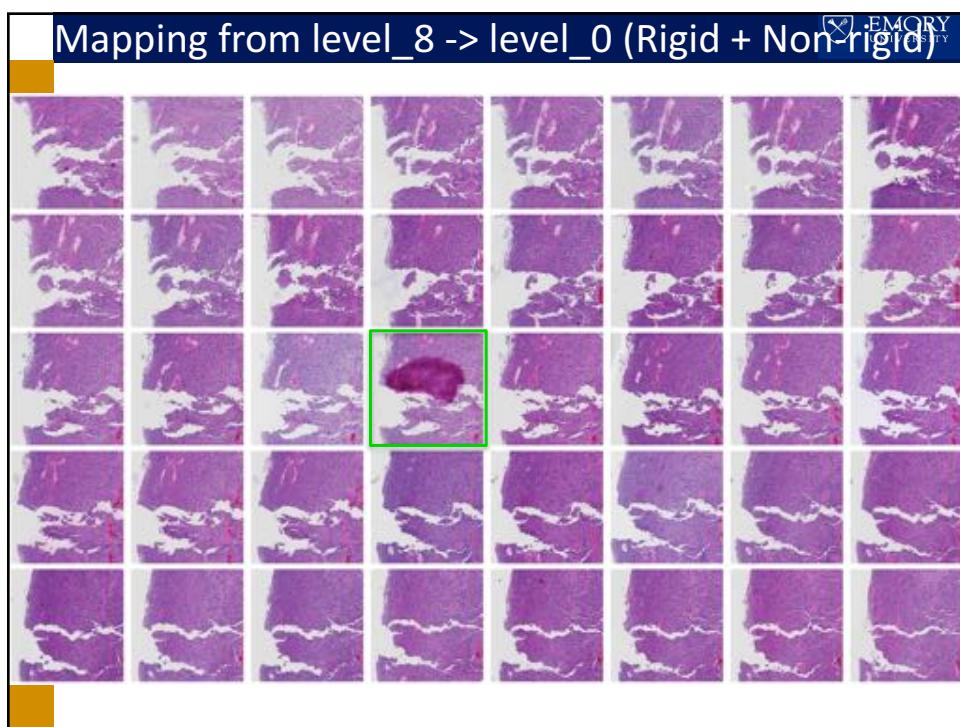
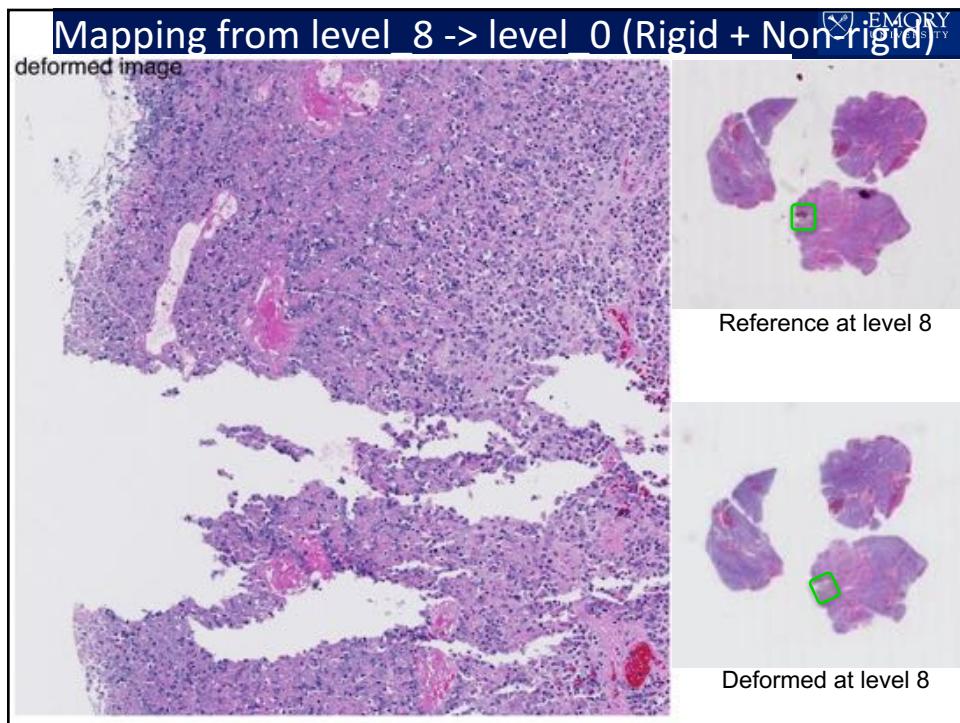


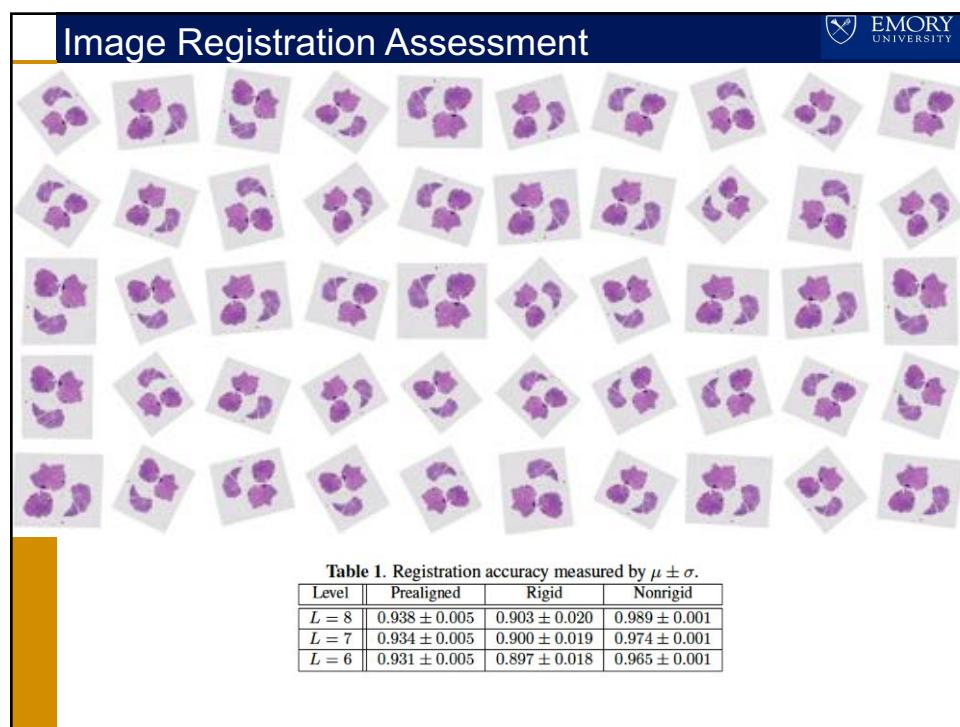
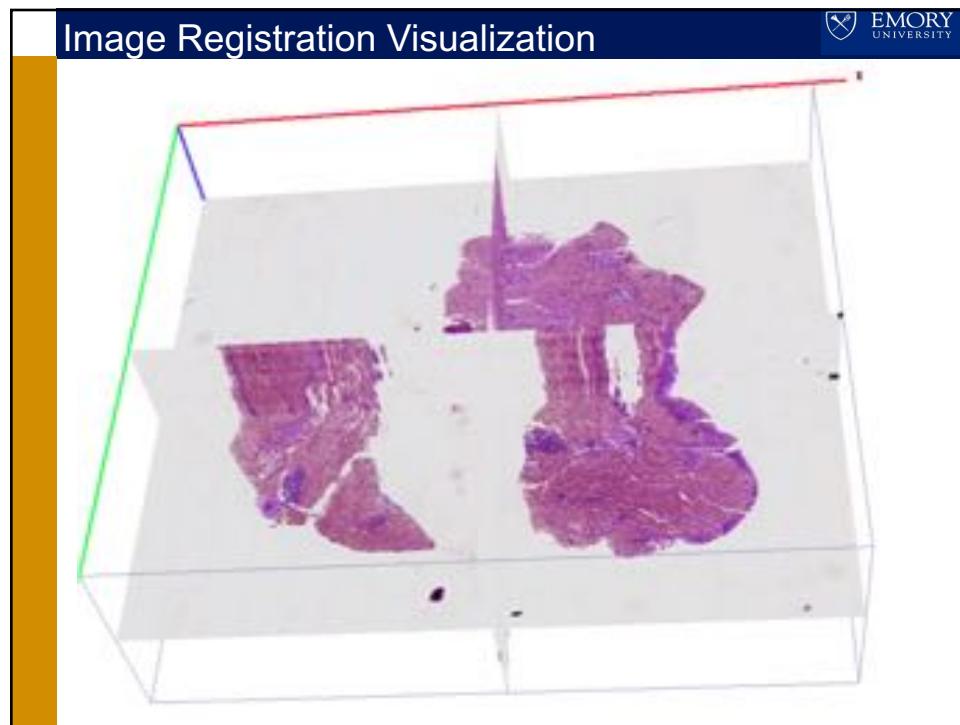


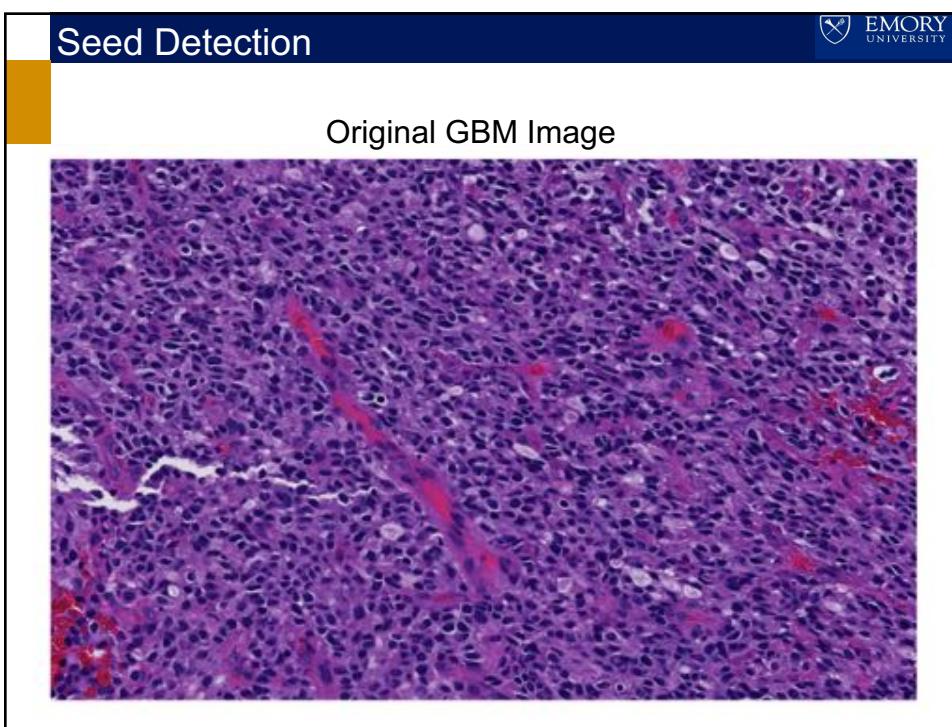
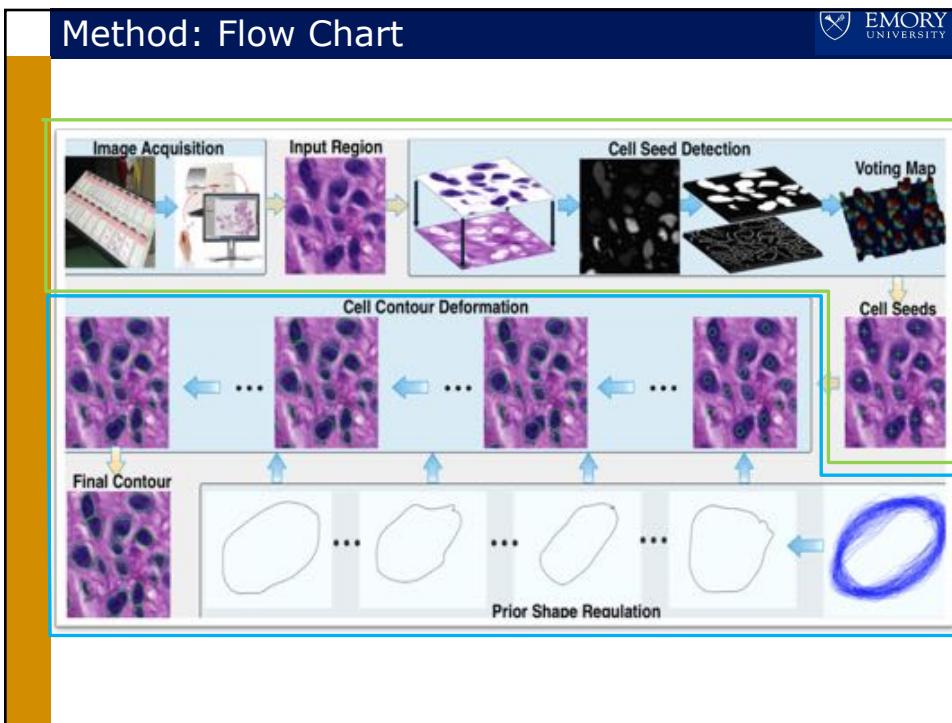


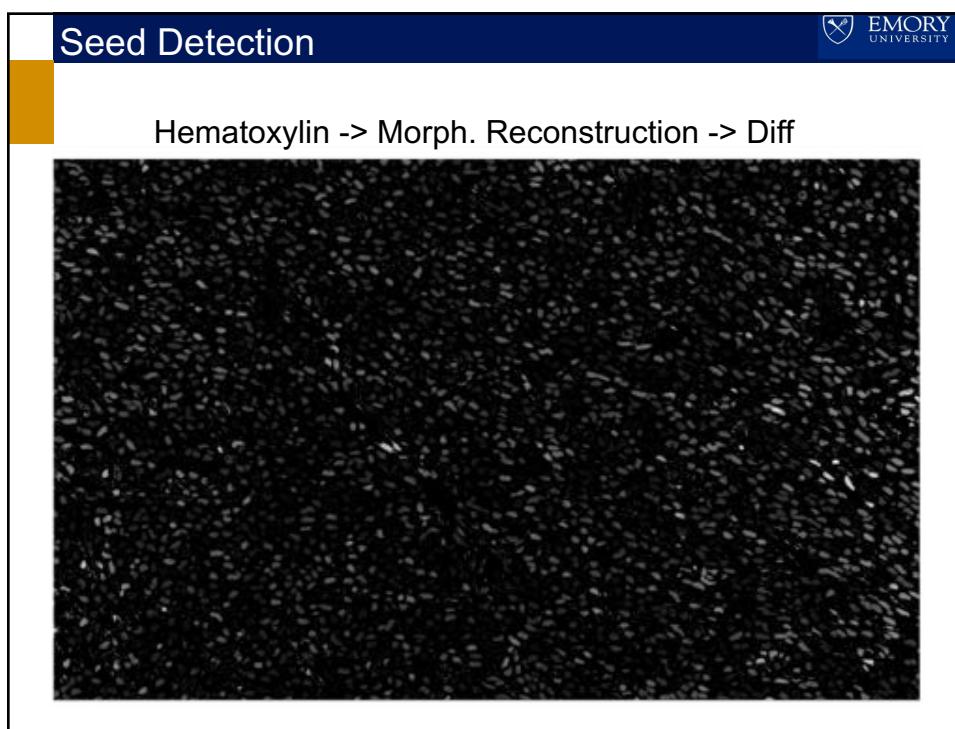
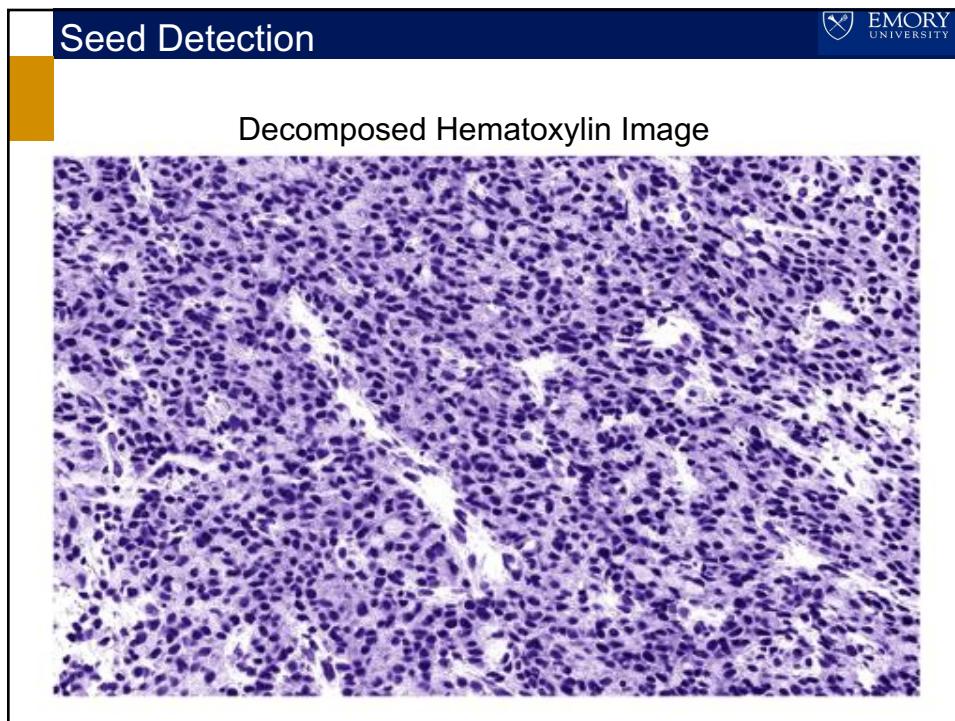


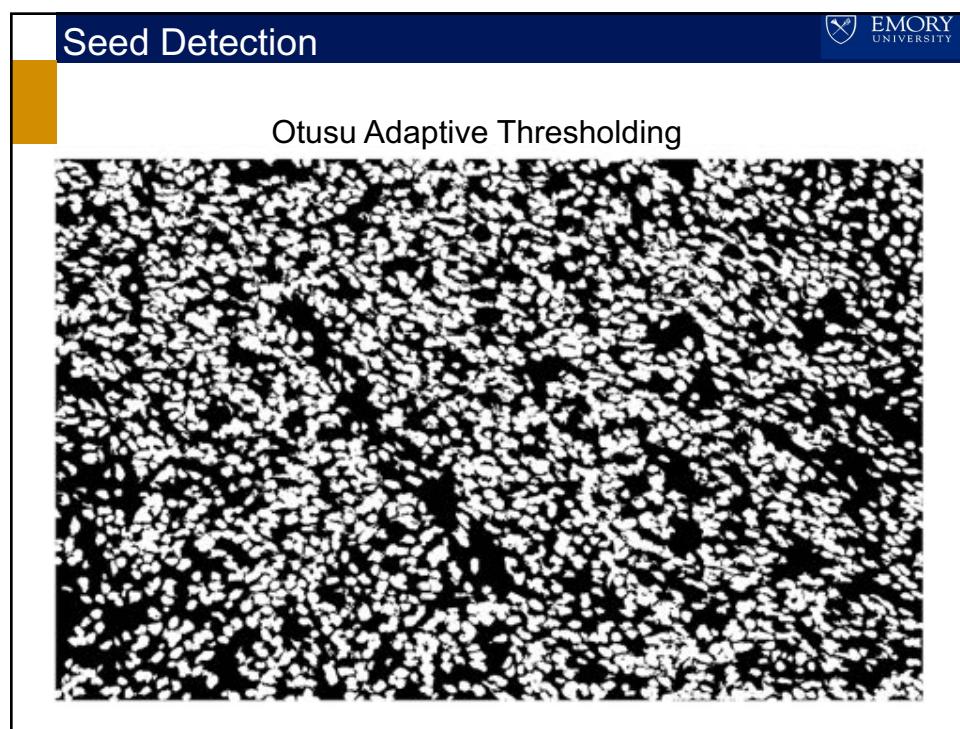
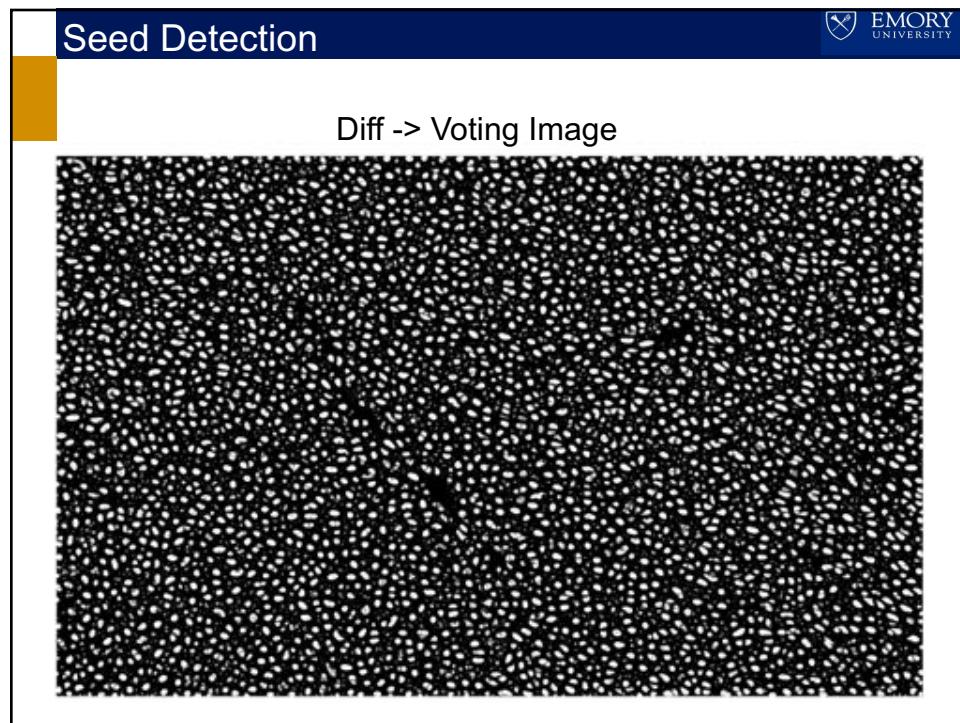


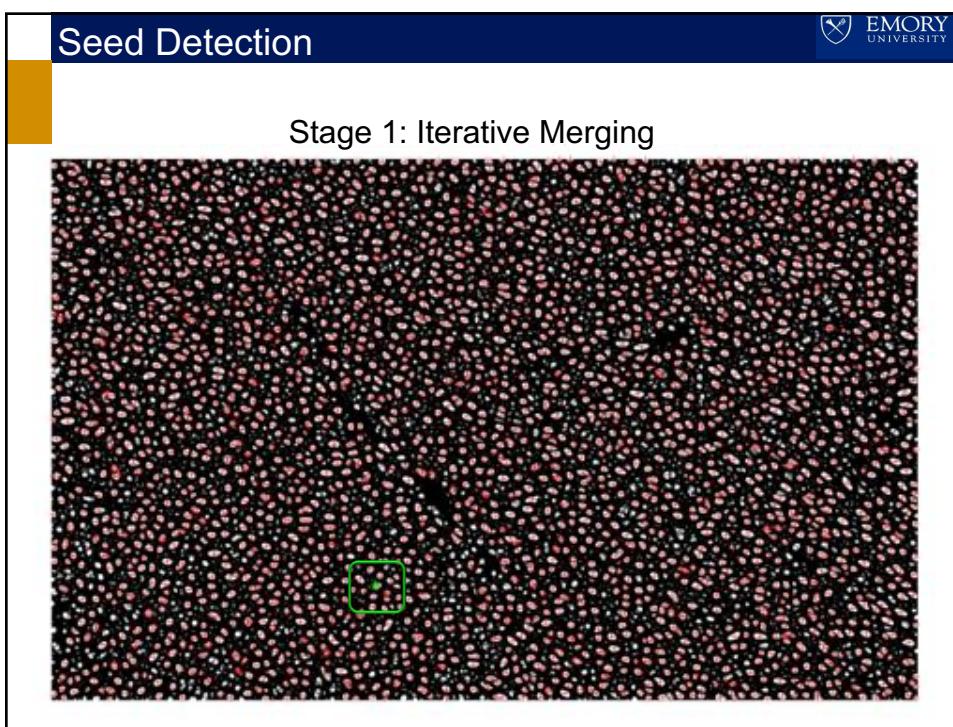
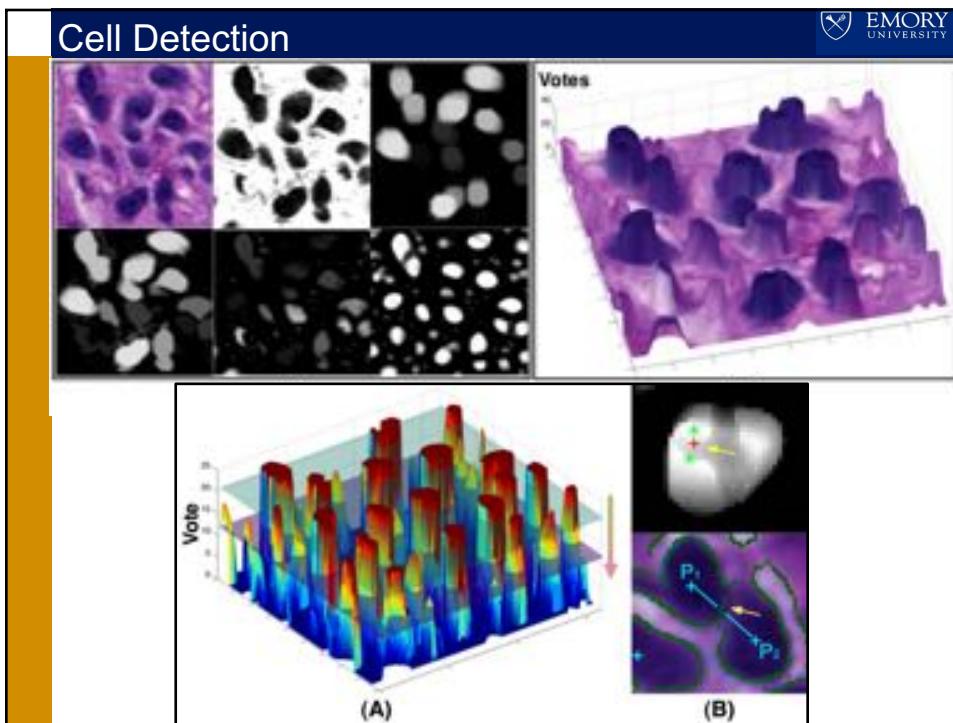


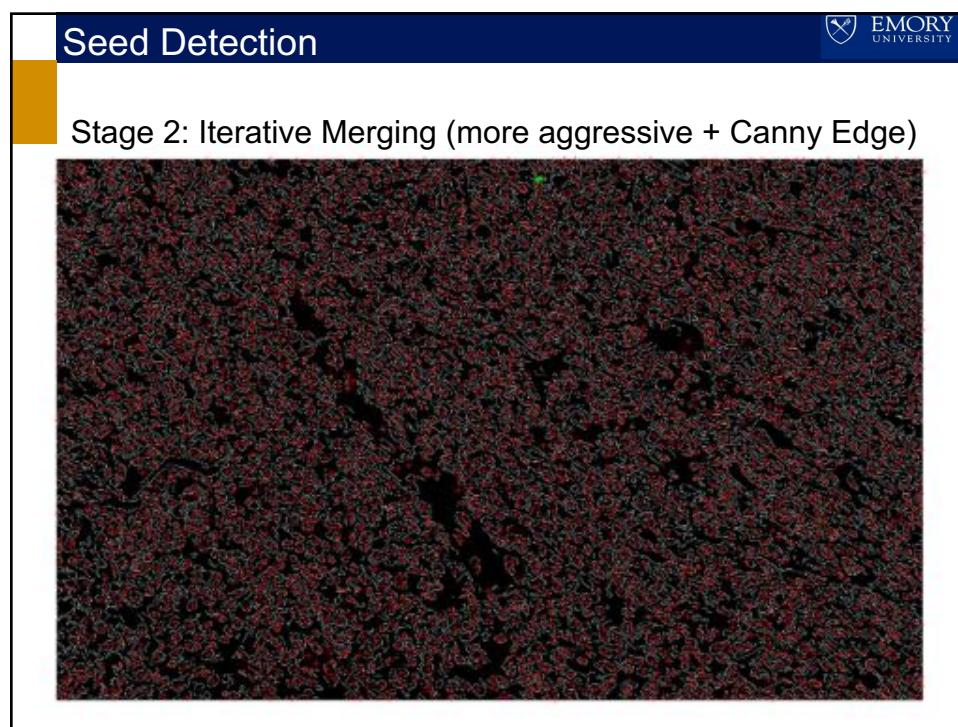
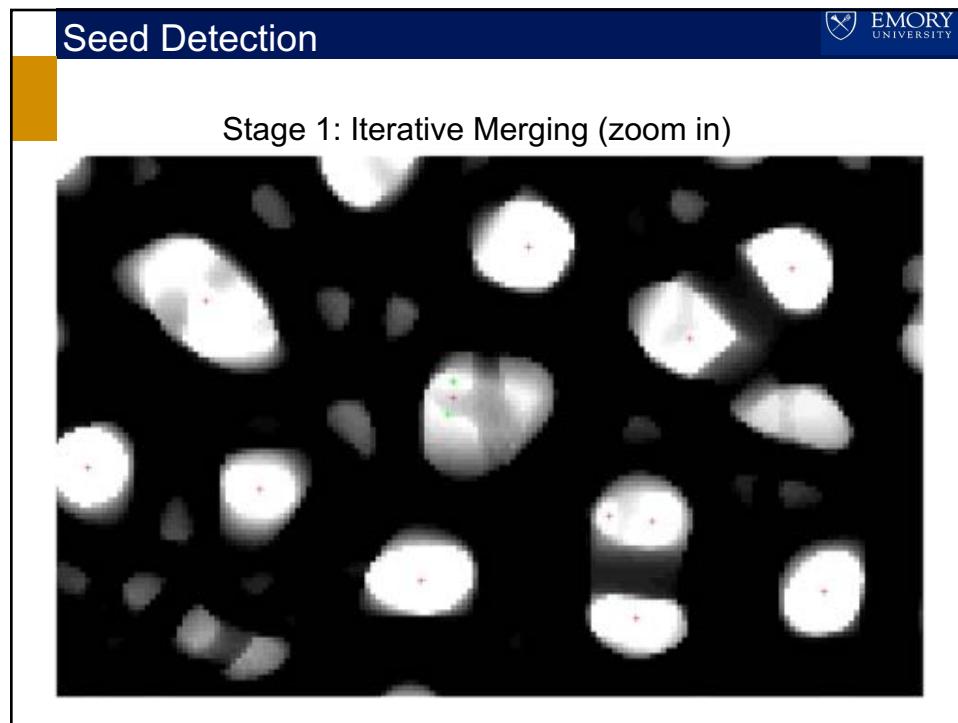


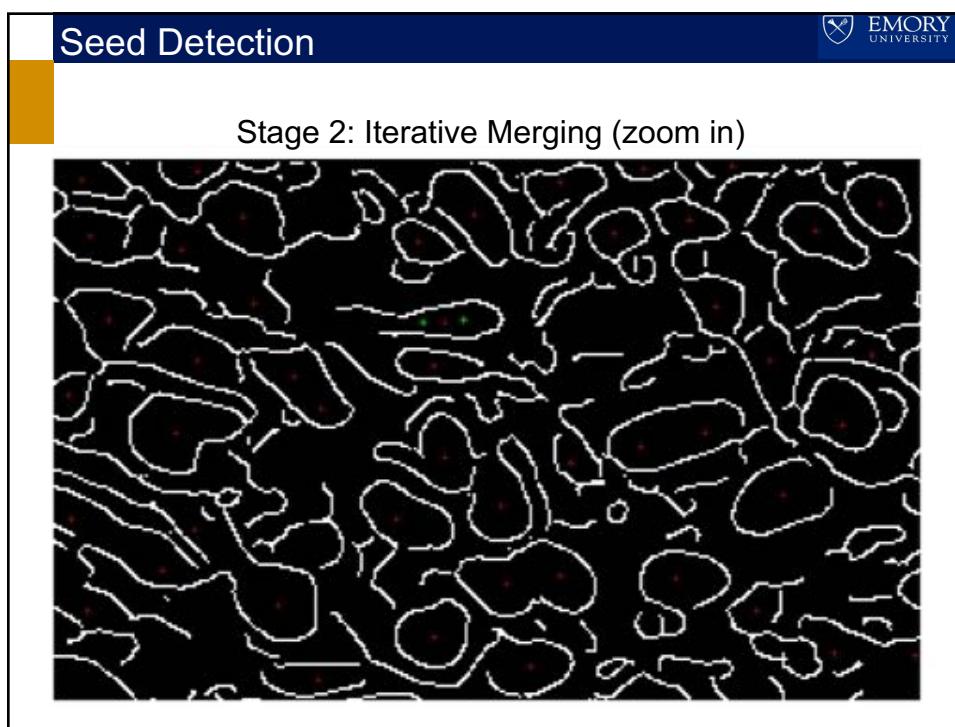
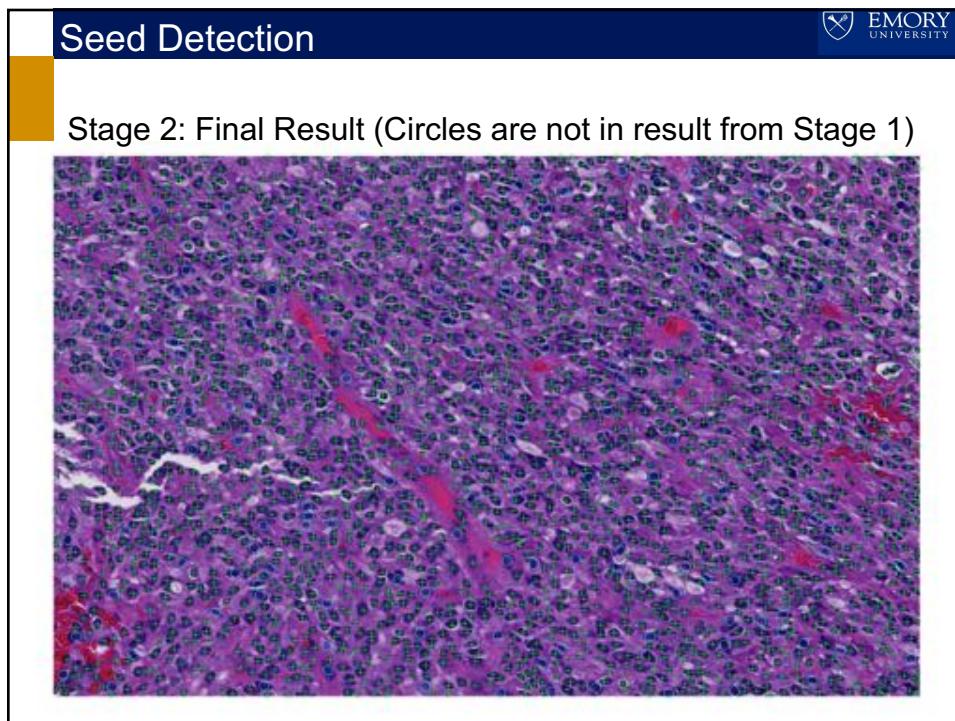






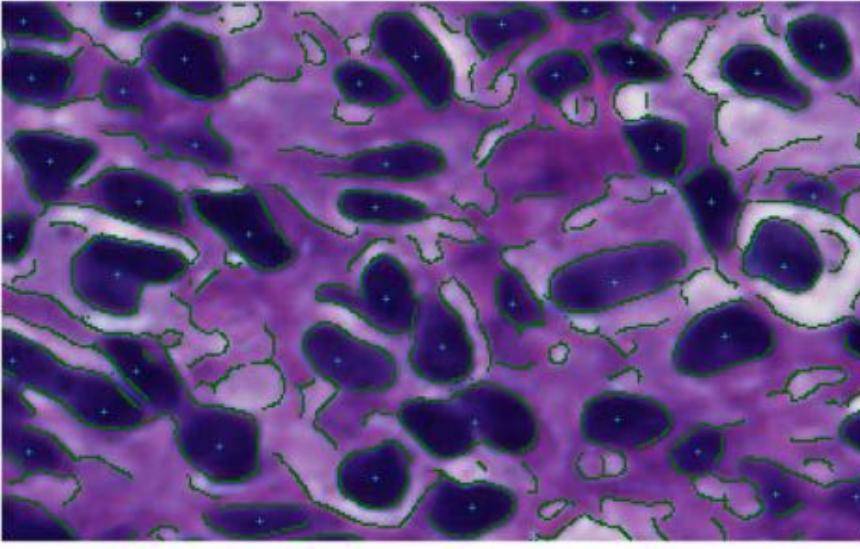






Seed Detection

Stage 2: Final Result (zoom in)



Seed Detection

Algorithm 1 New Seed Detection Algorithm

```

Input:  $I_{color}$ : original image;  $L \leftarrow \emptyset$ ;  $V(x) \leftarrow 0, \forall x$ ; parameters  

 $(\rho(r), \{s_i\}, \alpha, \beta, D_1, D_2)$   

Output:  $L$ : a list of seeds for cells

1: {Initialization Phase}
2: Decouple  $I_{color}$  into two stain channels: hematoxylin  $I_H$  and eosin  $I_E$ 
3: Construct mask image:  $\Psi \leftarrow \text{complement}(I_H)$  and marker image:  $\Phi \leftarrow I_H \oplus \rho$ 
4: Compute reconstructed image  $R_{\Phi}^{X,\rho}(\Psi) \leftarrow \text{morphRecon}(\Phi, \Psi, \rho)$ 
5: Compute difference image  $\Delta \leftarrow \Psi - R_{\Phi}^{X,\rho}(\Psi)$ 
6:  $I_{gray}(x) \leftarrow \text{rgb2Gray}(I_{color})$ ;  $M(x) \leftarrow \text{otsuThresholding}(I_{gray}(x))$ 
7: {Voting Map Construction}
8: for all  $i \in (1, 2, \dots, N)$  do
9:   for all  $x \in \Omega$  do
10:     $\lambda_2(x) \leftarrow \text{eigHessian}(\Delta(x), s_i)$ 
11:    if  $\text{sign}(\lambda_2(x)) < 0$  then
12:       $V(x) \leftarrow V(x) + 1$ 
13: {Detecting and Merging Seeds}
14:  $v_{un} \leftarrow \text{descendSort}(V(x))$ ;  $v_{i-1} \leftarrow \text{flip}(\text{normalize}(v_{un}))$ 
15: Let  $v_c$  and  $v_p$  be the current and previous voting value in  $v_{un}$ 
16: for all  $v_c$  and  $v_p \in v_{un}$ , where  $v_c \leftarrow v_{un}(i)$  do
17:   find objects  $O(v_c) \leftarrow \text{label}(V(x) \geq v_c)$ 
18:   for all  $o \in O(v_c)$  do
19:     if  $\text{size}(o) \geq A(o, \beta, v_i) = \beta + \exp(\alpha v_{i-1}(i)) \&&$ 
         $M(\text{centroid}(o)) = 1 \&& o \cap L = \emptyset$  then
           $L \leftarrow L \cup \{\text{centroid}(o)\}$ 
20: while anyL.pairwiseDistance()  $\leq D_1$  = true do
21:   find  $p \in L$  s.t. anyL.pairwiseDistance(p)  $\leq D_1$  = true  $\&\&$ 
    sumL.pairwiseDistance(p)  $\leq$  sumL.pairwiseDistance(p')  $\forall p' \in L$ 
22:   find  $Q = \{q \mid (\text{L.pairwiseDistance}(p, q) \leq D_1) = \text{true}\}$ 
23:    $L \leftarrow L \cup \text{Mean}(\{p\} \cup Q)$ 
24:    $L \leftarrow L \setminus \{\{p\} \cup Q\}$ 
25: {Merging Seeds with Edge Information}
26: while anyL.pairwiseDistance()  $\leq D_2 \&& \exists \text{.isBlockedByEdge}() = \text{true}$  do
27:   for all  $p \in L$  do
28:     find  $Q(p) = \{q \mid (\text{L.pairwiseDistance}(p, q) \leq D_2 \&&$ 
       $\text{.isBlockedByEdge}(p, q)) = \text{true}\}$ 
29:     find  $p \in L$  s.t.  $Q(p) \neq \emptyset \&& \text{sumL.pairwiseDistance}(p, Q(p)) \leq$ 
       $\text{sumL.pairwiseDistance}(p', Q(p')) \forall p' \in L$ 
30:      $L \leftarrow L \cup \text{Mean}(\{p\} \cup Q(p))$ 
31:    $L \leftarrow L \setminus \{\{p\} \cup Q(p)\}$ 
32:    $L \leftarrow L \setminus \{\{p\} \cup Q(p)\}$ 

```

Cell Detection

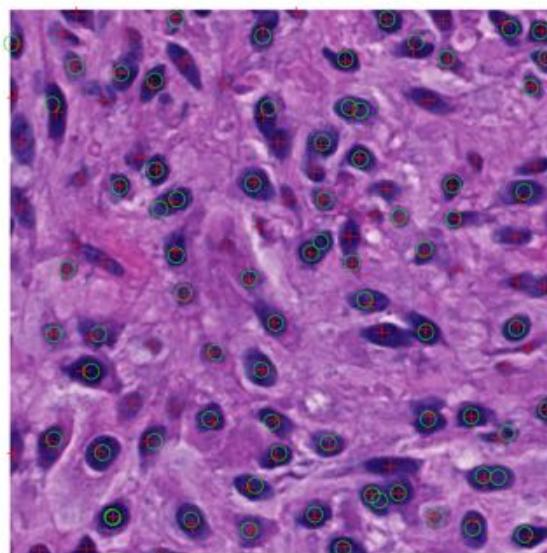


Table 1. Image-wise seed detection performance. The aggregated rates of these metrics in reference to the true number of cells are also shown in the last column.

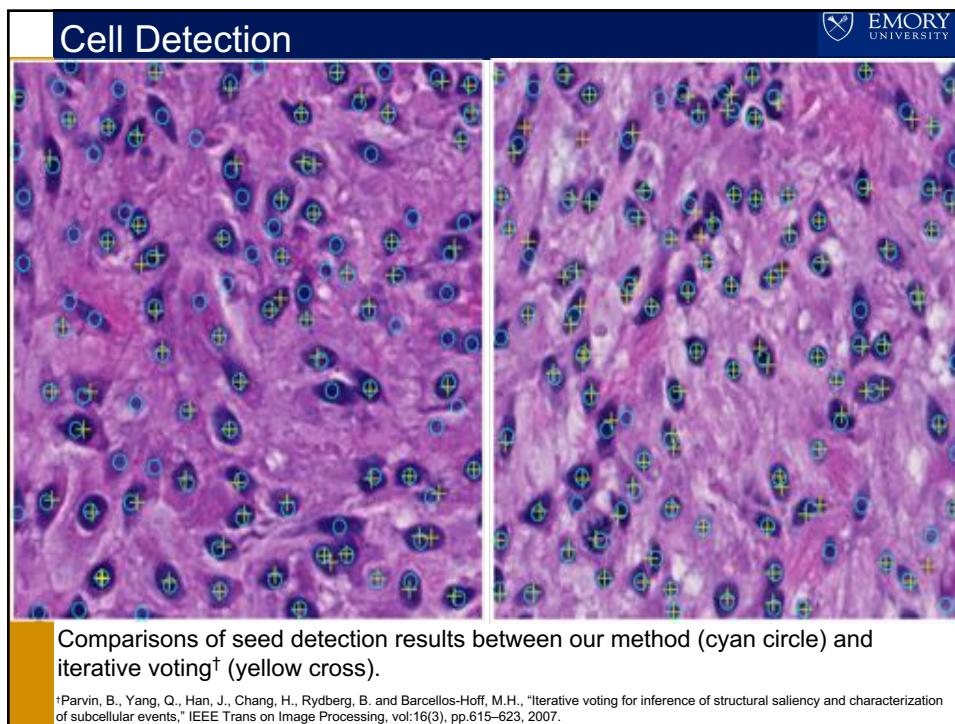
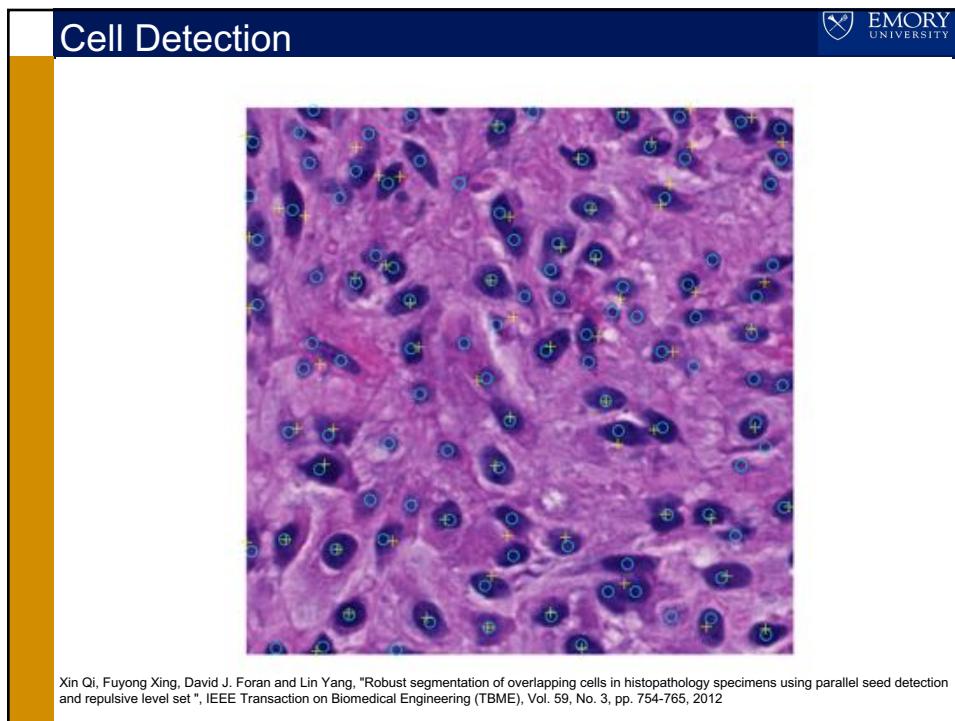
Metric	Mean	Std	Med	Min	Max	Overall Rate
C	2.15	1.51	2.00	0.00	5.00	1.59%
M	2.85	1.99	2.00	0.00	7.00	2.11%
F	1.30	1.18	1.00	0.00	5.00	0.96%
O	0.33	0.66	0.00	0.00	3.00	0.24%
U	0.50	0.72	0.00	0.00	2.00	0.37%

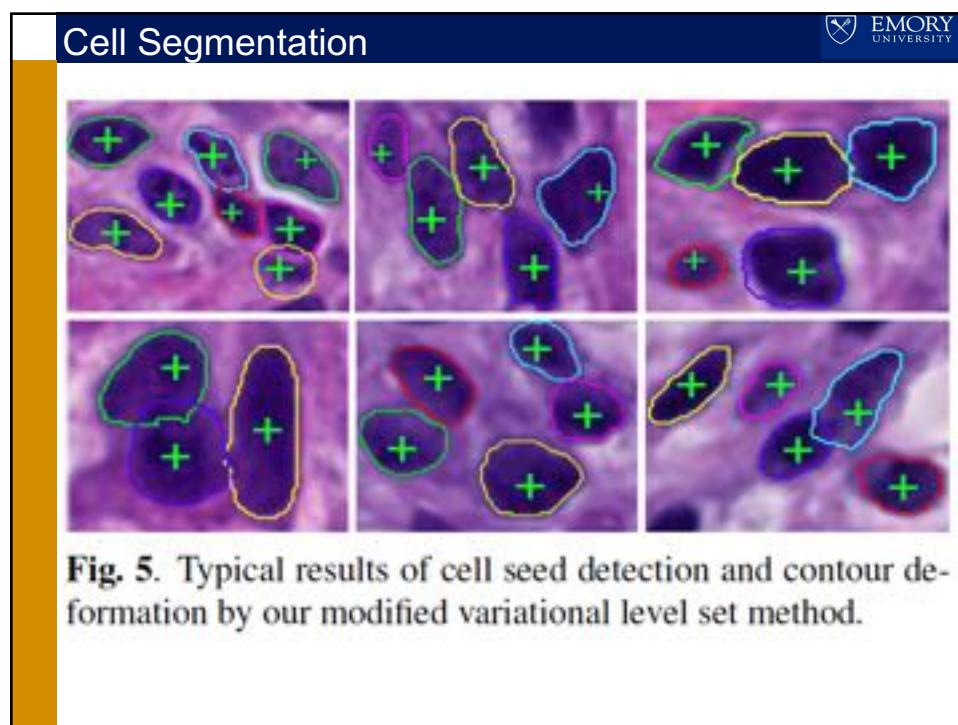
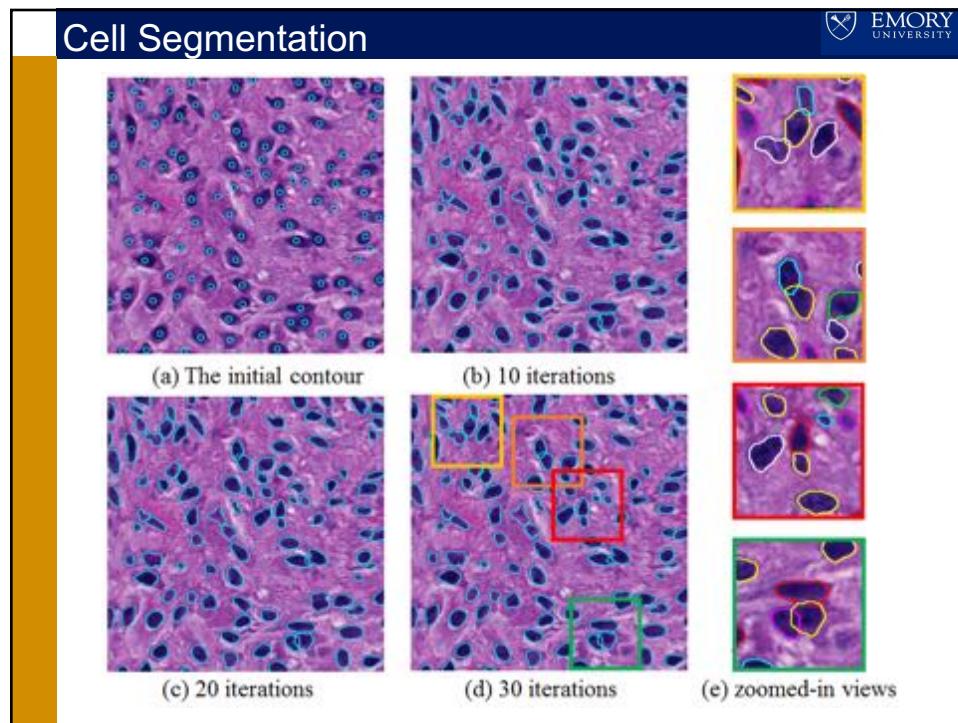
- The total number of human annotated cells for seed detection is 5396.
- Note that we evaluate our approach with non-touching and occluded cells in each image separately.
- Four metrics are computed from each image to show seed detection performance: (1)Cell Number Error; (2)Miss Detection (M); (3)False Recognition (F); (4)Over- (O); (5)Under- Segmentation (U)

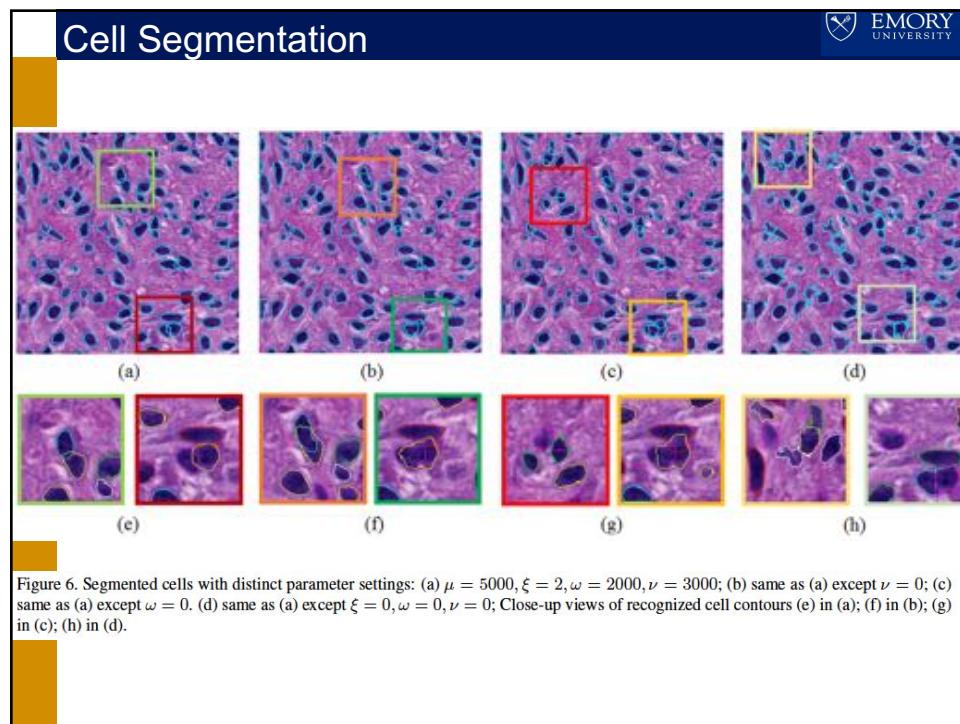
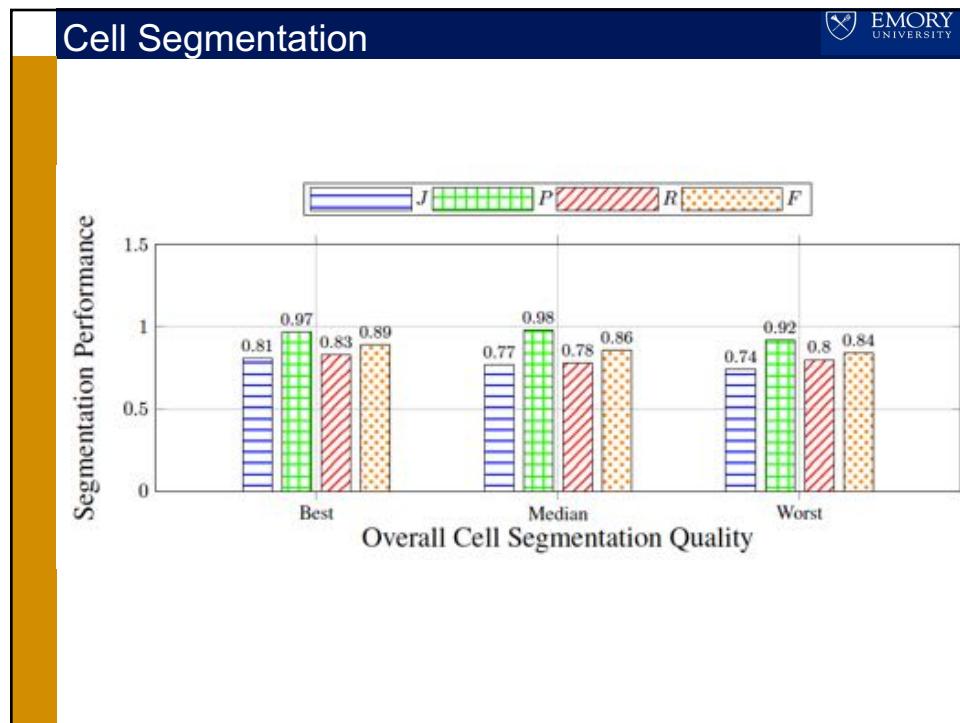
Cell Detection

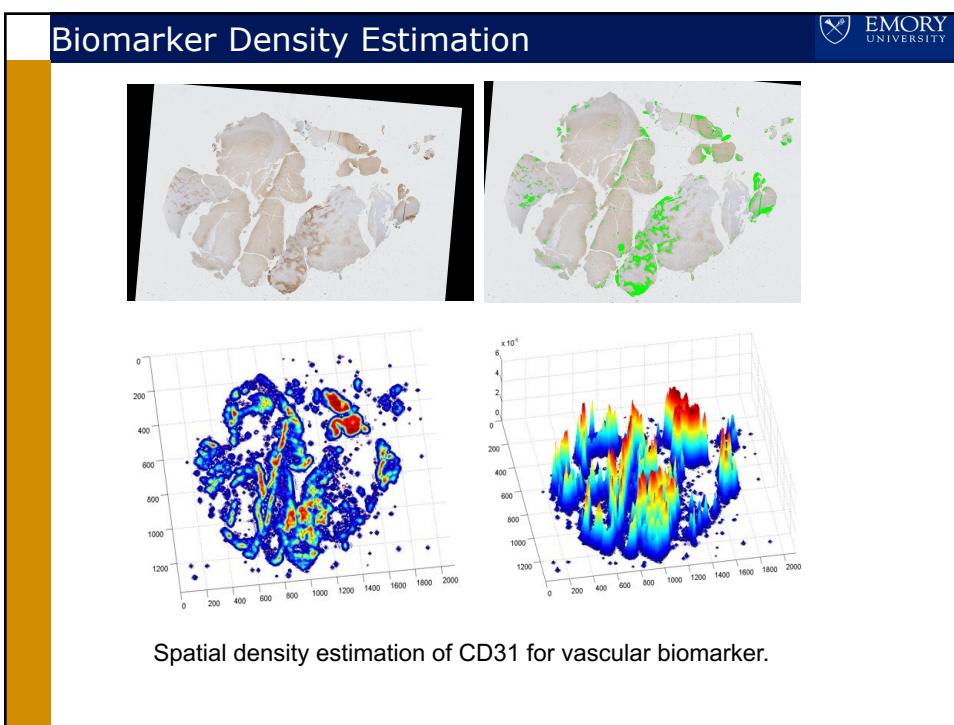
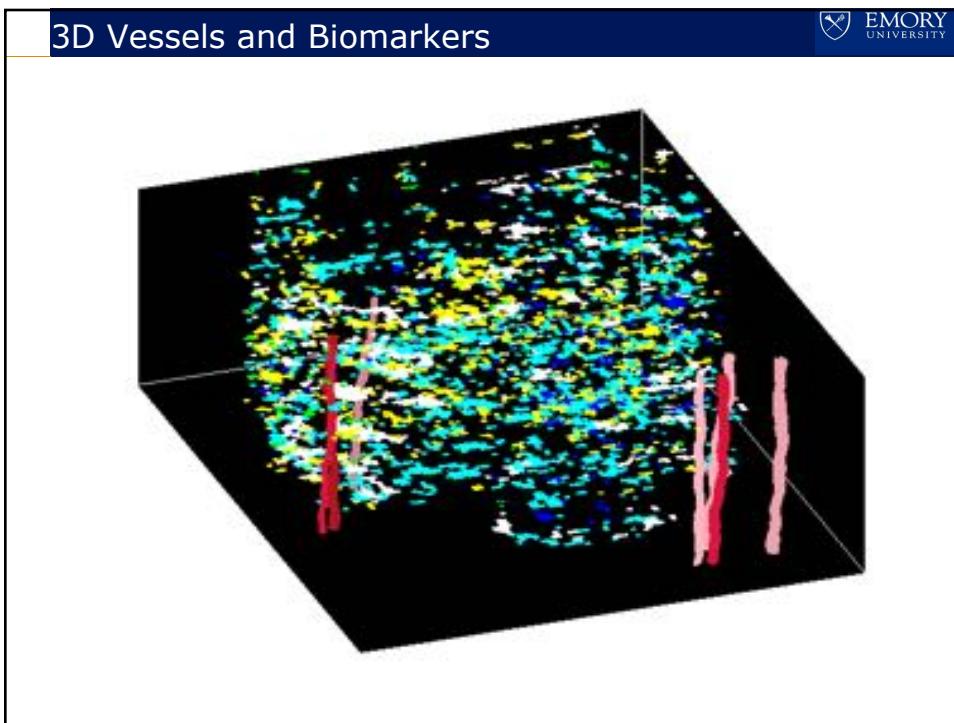


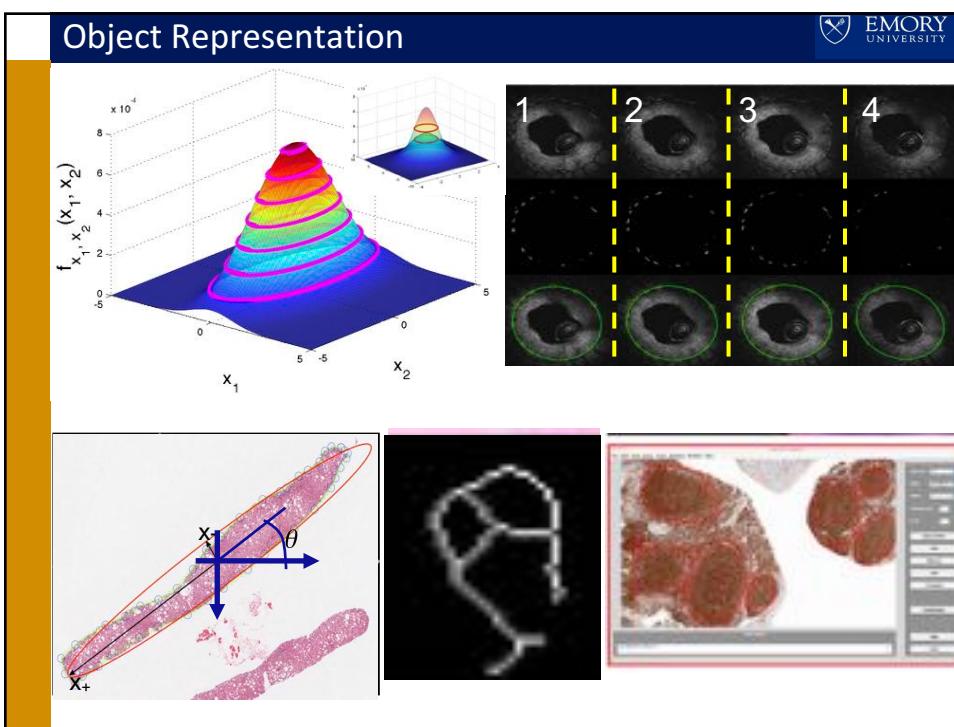
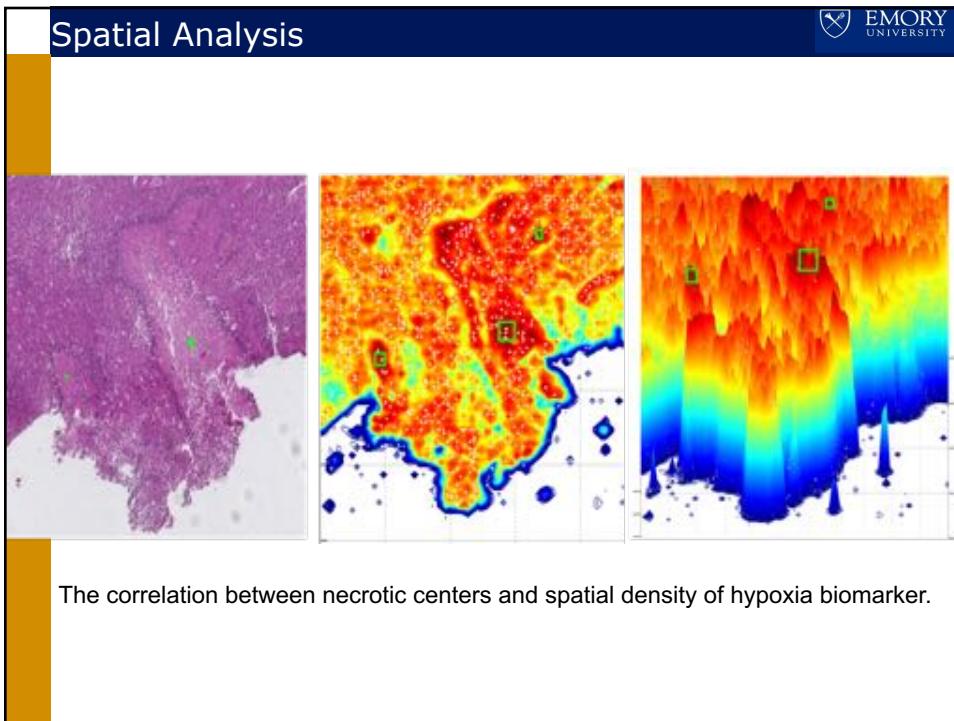
Barcellos-Hoff, M.H. et al., "Iterative voting for inference of structural saliency and characterization of subcellular events," IEEE Trans on Image Processing , vol:16(3), pp.615–623, 2007.





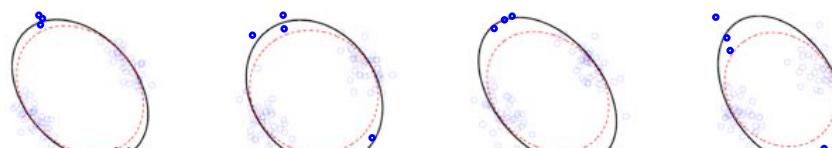
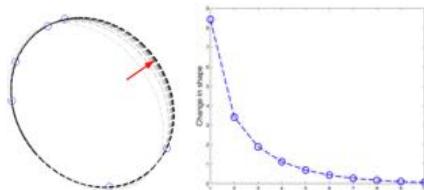




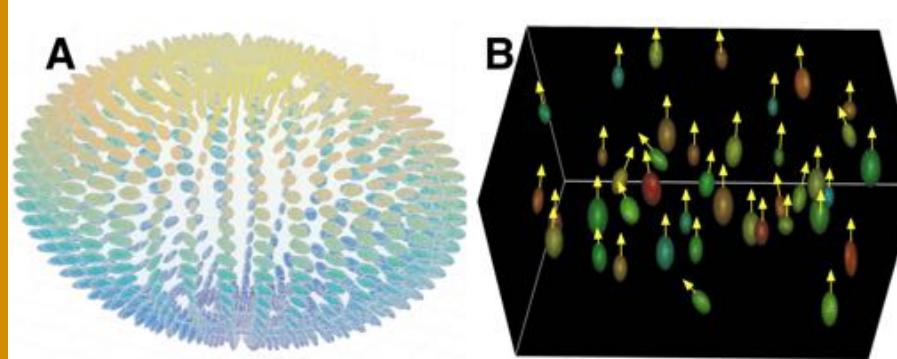


Superiority

- Robustness
- Convergence rate
- Resistance to “High Curvature Bias”
- Straightforward geometric interpretation



3D Cell Representation and Directionality





Thanks!

Students interested in this field, please contact:

Jun Kong, PhD

jun.kong@emory.edu; 404-712-0055

<http://insilico.cci.emory.edu/~jkong>

<http://www.bmi.emory.edu/junkong>

571, Psychology Building

Dept of Biomedical Informatics

Dept of Math and Computer Science

Winship Cancer Institute

