# Scalable Strategies for Image Analysis in Neuroscience

### Walter J. Scheirer

Computer Vision Research Laboratory, Department of Computer Science and Engineering



# Remarkable recent advances in visual recognition

5 o Clock Shadow --Arched Eyebrows --Attractive --Bags Under Eyes --Bald --Banas --Big Lips --Big Nose --Black Hair --Blond Hair --Blurry --Brown Hair --Bushy Eyebrows --Chubby --Double Chin --Eyeglasses --Goatee --Gray Hair --Heavy Makeup --High Cheekbones --Male --



Mouth Slightly Open --Mustache --Narrow Eyes --No Beard --Oval Face --Pale Skin --Pointy Nose --Receding Hairline --Rosy Cheeks --Sideburns --Smiling --Straight Hair --Wavy Hair --Wearing Earrings --Wearing Hat --Wearing Lipstick --Wearing Necklace --Wearing Necktie --Young --Dominance --Trustworthiness --

# Trouble...



I think it's a group of baseball players posing for a photo and they seem  $\textcircled{\begin{subarray}{c} {\begin{subarray}{c} {\begin{subaray}{c} {\bed} {\begin{subarray}{subarray} {\bed} {\begin{subaray}{\$ 



Nguyen et al. CVPR 2015

## REVERSE

### **Study** Natural System



## FORWARD

### **Build** Artificial System

Behavior, fMRI, 2-photon Imaging, Confocal Microscopy, Electron Microscopy, Electrode Recordings

Computer Vision Machine Learning



## Vision Systems that Work: Models







Neuroscience and Machine Learning Have a Common Problem:

Just because a system works doesn't mean we understand it.

Even when we have considerable access to the system, it's not always easy to answer "why" questions

Maybe we can build tools that help address both issues and learn something in the process

### New Experiments

Establishing causal links between cortical areas and function

Understanding decision boundaries between classes

Studying the mechanisms of sensory integration

Searching for mesoscale cortical computing circuits



Brain Science at the Notre Dame Computer Vision Research Lab

# IARPA MICrONS



# Machine Learning Algorithms from Wet Lab Experimentation



# Experimental workflow



# 2D Segmentation and 3D Reconstruction for EM



Elia Shabazi





EM: M. Joesch, Cox Lab @ Harvard

# 2D Segmentation and 3D Reconstruction for X-ray



Elia Shabazi



X-Ray: N. Kasthuri, Argonne National Laboratory / U. Chicago

## Psychophysics on the Model



# Strategies for Image Analysis

## Dense Segmentation and Reconstruction



#### Visual Computing Group @ Harvard

## Software tools for dense reconstruction

Package	Method	Link					
Randomer Forests	Decision Forests	http://ttomita.github.io/RandomerForest					
Gala	Active Learning	https://github.com/janelia-flyem/gala					
VESICLE	Deep Learning	http://openconnecto.me/vesicle					
Synapse Segmenter	Context Features + Adaboost	http://cvlab.epfl.ch/software/synapse					
ATMA	3D Pixel Features + Random Forests	https://github.com/RWalecki/ATMA					
ZNN	CNN	https://github.com/seung-lab/znn-release					
PRIM	CRF	http://github.com/funkey/prim					
ilastik	Random Forests	http://ilastik.org/					
Rhoana	CNN	https://github.com/Rhoana					

# Preeminent 2D segmentation method: CNN

Feature learning for strongly invariant membrane representations



Interlude: is deep learning as good as we think it is?

# An alternative to dataset testing



# The framework



### You don't have to use tricky manipulations

### GoogleNet Output

### Label: Hammerhead Shark



### Label: Syringe



### Label: Blow Dryer



### Label: Trimaran



### Label: Mosque



### Label: Missile



# Same transformation across 2D and 3D objects



# Various transformations across 3D objects



# Back to image analysis...

# Error rates (MICrONS targets)

### $100 \times 100 \times 100 \ \mu m^3$

1mm x 1mm x 0.1mm

### 1mm x 1mm x 1mm

Precision:  $\ge 70\%$ Recall:  $\ge 70\%$ NID:  $\le 0.95$  @ 50th pctl.  $\le 0.50$  @ 75th pctl.  $\le 0.20$  @ 95th pctl. VI < 1.75 nats Precision:  $\ge 85\%$ Recall:  $\ge 85\%$ NID:  $\le 0.95$  @ 15th pctl.  $\le 0.35$  @ 50th pctl.  $\le 0.15$  @ 85th pctl. VI < 1.0 nats Precision: ≥ 97.5% Recall: ≥ 97.5% NID: ≤ 0.80 @ 10th pctl. ≤ 0.15 @ 50th pctl. ≤ 0.05 @ 75th pctl. VI ≤ 0.25 nats

Think about error propagation for even the best of these numbers...

# Training Time

### AlexNet:



# Rethinking the problem

- Make samples better with high-contrast tissue prep.
- Sparse reconstruction vs. dense reconstruction
- Avoid overfitting with unsupervised methods
- Cell-specific reconstruction strategies
- Solve this problem more like people do

### Assisted Reconstruction Technique for Electron Microscopic Interrogation of Structure (ARTEMIS)

(a) tagging a genetically identified cell with an electron-dense tracer
(b) enhancing the electron-dense staining of these tracers
(c) imaging the cell rapidly at relatively low resolution
(d) re-imaging small volumes at higher resolution to map connectivity



M. Joesch, D. Mankus, M. Yamagata, A. Shahbazi, R. Schalek, A. Peleg, M. Meister, J. W. Lichtman, W. J. Scheirer, J. R. Sanes, "Reconstruction of Genetically Identified Neurons Imaged by Serial-Section Electron Microscopy," eLife, Vol. 5, e15015 2016

# **ARTEMIS Staining Strategy**



## Improved contrast to noise ratio

standard

intensity [uint8]

reduced



intensity [uint8]

# Cytosolic Apex (enhanced tissue)



# 4nm/pixel, detail from $680\mu \times 680\mu$ section



# Learning-free 2D segmentation and 3D reconstruction



# Starburst Amacrine Cell







Keeley et al. J. Comp. Neurol. 2007

# Step 1: Pre-processing

Original EM images are of very high resolution (100000 x 50000 pixels)

Process local neighborhoods in 2048 x 2048 tiles

Each tile can be processed separately for 2D segmentation

# Step 2: Adaptive Thresholding on Local Intensity

### Key step enabled by ARTEMIS

Apply a Wiener filter, modulating via known pixel ranges of ARTEMIS markers:





# Step 3: Cluster-based Image Segmentation

Choose threshold to minimize the intraclass variance of the black and white pixels

Prune non-ARTEMIS marked segments: 3 x n matrix to calculate weights based on local pixel neighborhoods

**Remove Artifacts** 



# Step 4: Assess morphology

Catalog the properties of each segment:

- Centroid
- Convex Area
- Area
- Bounding Box
- Extent
- Solidity

- •Extrema
- Major Axis Length
- Minor Axis Length
- •Equiv. Diameter
- Eccentricity



## Step 5: Database assessment

Analysis moves to abstract representations stored in a MySQL database

Rapid search for large and small segments repeated across layers

en warea	1,658,974+	rows total (ap	proximately), limited to 1	1,000													Not	NK 9	le wor			ws 🕆 Fa
🤌 id	Layer	pArea	Cent_1	Cent_2	00×	00y	88	860-	MajA.	MnA	Eccent	Convièrea	ex3	ex2	ex3	ex4	ed	ex6	ex?	ext	ed	ex 10
- 1	207	61	4.21311475409836	4.21311475409836	0.5	0.5	12	12	24.3500347479635	7.62295669545576	0.847236922378132	78	0.5	12.5	12.5	12.5	1.5	0.5	0.5	0.5	0.5	0.5
- 2	207	24	1	299.5	0.5	287.5		25	27.712812921102	1.15479053837927	0.999131567256817	24	0.5	1.5	1.5	1.5	1.5	0.5	0.5	0.5	287.5	287.5
3	207	348	3.12837837637637638	443.155405405405	0.5	426.5	- 6	35	30.4733727106805	6.79314873093564	0.974836558382174	155	0.5	1.5	6.5	6.5	1.5	0.5	0.5	0.5	426.5	426.5
	267	750	6.9666666666667	\$26.796	0.5	467.5	20	86	67.7522638196955	20.1262129289624	0.95486007562018	960	0.5	1.5	20.5	20.5	1.5	0.5	0.5	0.5	467.5	467.5
5	207	73	1.87671232876712	594.698630136986	0.5	580.5	3	29	29.0695743837198	3.43810136986661	0.992981296250498	27	0.5	1.5	3.5	3.5	1.5	0.5	0.5	0.5	580.5	580.5
	207	681	2.5345080763583	707.607929515429	0.5	620.5	- 6	201	299.913059532927	5.54597605139982	0.999615130620117	901	0.5	1.5	6.5	6.5	1.5	0.5	0.5	0.5	620.5	620.5
,	207	346	2.90839636408036	845.575342465753	0.5	826.5	7	39	33.6680/990917687	7.07620879253889	0.977663625091435	176	0.5	1.5	7.5	2.5	1.5	0.5	0.5	0.5	826.5	826.5
•	267	30	1	806.5	0.5	871.5		30	34.6410161513775	1.15470053837922	0.999444290037663	30	0.5	1.5	1.5	1.5	1.5	0.5	0.5	0.5	871.5	871.5
•	207	30	1	1,004.5	0.5	999.5		30	34.6410362513775	1.15409053807922	0.999444290037663	30	0.5	1.5	1.5	1.5	1.5	0.5	0.5	0.5	999.5	999.5
10	207	409	2.38630806845966	1,101.63080684987	0.5	1,007.5	7	140	361.087900229449	5.6290526120781	0.999393441227636	747	0.5	1.5	7.5	7.5	1.5	0.5	0.5	0.5	1,037.5	1,007.5
- 11	207	230	3.67226890796303	1,209.74789915966	0.5	1,186.5	30	47	41.5458962101644	9.2247853890946	0.97503791590431	277	0.5	1.5	30.5	30.5	1.5	0.5	0.5	0.5	1,186.5	1,186.5
13	267	129	2.7984496124031	1,277.3488372093	0.5	1,258.5	- 4	35	30.6475531902746	6.30762982156757	0.978391297141347	545	0.5	1.5	6.5	6.5	1.5	0.5	0.5	0.5	1,259.5	1,259.5
13	207	90	2.0666666666666	1,318.544444444	0.5	1,301.5	- 1	24	31.5706403362757	4.22588620818123	0.993000940051246	20.3	0.5	1.5	4.5	4.5	1.5	0.5	0.5	0.5	1,301.5	1,301.5
14	207	113	1.36283183840758	1,386.54867256637	0.5	1,382.5	- 2	72	79.1600554909125	2.22914505363414	0.999603429336087	134	0.5	1.5	2.5	2.5	1.5	0.5	0.5	0.5	1,352.5	1,152.5
15	207	368	2.30357142857140	1,461.74404761905	0.5	1,405.5		60	53.1086342893938	5.14529860185817	0.995295823513619	215	0.5	1.5	6.5	6.5	1.5	0.5	0.5	0.5	1,405.5	1,405.5
16	207	25	-	1,596	0.5	1,543.5		25	28.8675134394813	1.15479053837927	0.999199679743744	25	0.5	1.5	1.5	1.5	1.5	0.5	0.5	0.5	1,543.5	1,90.5
17	207	97	1.23711340206386	1,671.07236494945	0.5	1,628-5		24	85.6522850089977	1.89640405473261	0.99975404011561	122	0.5	1.5	2.5	2.5	1.5	0.5	0.5	0.5	1,628.5	1,628.5
18	207	267	3.28089887640449	1,753.13857677903	0.5	1.721.5		78	61.1484183207111	8.47245993829945	0.990356934908996	386	0.5	1.5	8.5	8.5	1.5	0.5	0.5	0.5	6,721.5	1,721.3
19	207	20		1,905.5	0.5	1,895.5	-	20	23.094010767585	1.15470053837927	0.998748217771909	20	0.5	1.5	1.5	1.5	1.5	0.5	0.5	0.5	1,895.5	1,895.5
20	207	240	2.4075	1,961.908333333333	0.5	1,924.5		72	68.9779478125766	5.21335825705465	0.99713973580821	28.5	0.5	1.5	6.5	6.5	1.5	0.5	0.5	0.5	1,924.5	1,924.5
21	207	1,128	3.53540099290.78	2,140.85726950355	0.3	2,823.5	12	227	233.564986984756	8.95701304904005	0.999364403676523	1,912	0.3	1.5	12.5	12.5	1.5	0.5	0.5	0.5	2,023.5	2,023.5
22	207	40	1.4	2,108.4	0.3	2,325-5		27	27/6529140468614	2.26487336479747	0.996640256402919	49	0.3	1.5	2.5	2.5	1.5	9.5	9.5	9.5	2,325.5	2,123-3
20	207			2,173.3	0.5	2,360.5	-	- 26	30.0222139978605	1.15409053837927	0.999260081289737	26	0.5	1.5	1.5	1.5	1.5	0.5	0.5	0.5	2,360.5	2,360.5
24	207		1.72368423052632	2,493.10526315789	0.5	2,475.5		36	34.0880943626201	3.2937129/2909775	0.99512099.7611787	8.7	0.5	1.5	3.5	3.5	1.5	0.5	0.5	0.5	2,475.5	2,475.5
- 20	207	235	2.90967742935464	2,353.73812903226	0.3	2,501.5		90	29.88.3087330.3693	7.52494963396458	0.982039544733666	230	0.3	1.5	7.5	7.5	1.5	0.5	0.5	0.5	2,531.5	2,531.5
26	207	279	3.66187050099712	2,407.37760764173	0.3	2,585-5	20		64.842284860746	9.46427670779401	0.98929072721969	449	0.3	1.5	20.5	39.5	1.5	9.5	0.5	9.5	2,585.5	2,985.5
	207	21		2,737	0.5	2,728.5	-	- 21	24.2487113059640	1.15409053837927	0.990005569685858	21	0.5	1.5	1.5	1.5	1.5	0.5	0.5	0.5	2,726.5	2,726.5
	207			2,825	0.5	2,770.5	-	- 2	27.712812921302	1.154/005383/927	0.999131567256817	24	0.5	1.5	1.5	1.5	1.5	0.5	0.5	0.5	2,770.5	2,170.5
- 2-				2,863.5		2,040.5		- 2-	32.333813074619	1.154/001383/922	0.999362041462373		0.5	1.5	1.5	1.5	1.5	0.5	0.5	0.5	2,049.5	2,040.5
	207			2,497	0.3	2,894.5			28.86.75134394613	1.19409050807927	0.9993399679740244		0.5		1.5	1.5	1.5	9.5	9.5	9.5	2,004.5	2,004.5
	207	62	F 596 022 0806 40 16	2,908.00094538129	0.5	2,915.5			44.2007040571532	2.14637584129901	0.998830147919769		0.5	1.5	2.5	2.5	1.5	9.5	0.5	0.5	2,915.5	2,913-3
			1 000000000000000	2,778.5	0.5	2,471.5		- 21	20.000013273095	1.134/003383/924	0.997445717412067		0.5		1.5	1.5	1.5	0.5	0.5	0.5	2,971.5	2,971.5
		-	1.85206362978723	2,041.9040301702		2,003.5			81.4089902324083	2.00300990.7000.29	0.990900907211/90		0.5						0.0	0.5	2,005.5	2,005.5
- 2 -	-		1.2010/12/2010/0	2, 101, 910, 00990, 022		3,153-5		- 22	51.00000000000	10. Photo 2010 1000	0.077673600746073	-			11.5	11.5				4.5	0.130.5	1,100.0
- 2-	2017	100	2.28.0757505079	3,272,933,999,921		2,240.5	- 1		24.400061793626	3.901/08/07/19/5	0.9999/3005/779579	64.9	0.5	1.2	6.5		1.5		0.5	0.5	2,047.5	1,000
			2.90909090909090	3,331,125	0.5	2,10.5	- 1		21.3331/22041/36	7.09021290027222	0.940153942894904	115	0.5		0.5	6.5	1.5	0.5	0.5	0.5	3,317.5	3,207.3
- 2	-		1.1530-01330-013	2,306.30700230709		2,170.5			35.7525722873962	1.82930273212048	0.990009021290030		0.5		4.5			0.5	0.5	0.5	2,279.5	1,000.5
- 2 -	20.0		2.000000000000000	1.444.996.000000		1.441.5	- 1	- 2	20.00000000000000	1 114700 10170	0.0000000000000000000000000000000000000									0.5	1.481.5	1,401.5
- 2-	207	12	2 0000245 107745	3,494.5		3,463.5		42	41.01726/0015	5.19474/53837927	0.000000000000000	22		1.5	- 13	1.5		0.5	0.5	0.5	1.445.5	3,403.5
	207	1//	2.90900401977401	3,538.30508474576		4,913.5	- 1		42.012296703001	6.80060/#3621081	0.90300901122090	212	0.5		- 13			0.5	0.5	0.5	2,212.5	2,313.5
		130	2.00401200401230	2,418.81330401338		2,798.5	- 1	- 2	20.0221200800908	6.20400970121073	0.982191257202909	104	0.5		6.3			0.5	0.5	0.5	2,298.3	2,798.5
				3,968		4,410-3		- 2	28.8679 (24)94613	1.19474/10807927	0.000100670740744			- 12						4.5	2,403-3	1,453-3
	207	30	1.35	3,795		1,493.5		12	10.0010253000426	2.2404/612987146	0.3973366-8004045	56		1.5				9.5	4.5	9.5	2,000.5	2,493.5
	207	244	2.30000000000000	2,794.51.300000009	0.5	2,172.5			42,4409404147904	3.39456564131528	0.992208913004147	182	0.5	1.5	5.5	3.5	1.5	0.5	0.5	0.5	2,771.3	3,171.5
	207	20		2,870.5		2,000.5		20	23.094030767585	1.15470053837927	0.990740217771909	20	0.5	1.5	1.5	1.5	1.5	0.5	0.5	0.5	2,000.3	2,000.5
-	207			3,922	- 23	2,930.5		20	26-3081123827238	1.19470050807927	0.9990094073020963	20	0.5		- 10	1.5	1.5	0.5	0.5	0.5	0,939.3	1,930.3
	202			4,021.5	9.5	4,008.5		26	30.0222139978605	1.154.005383.9927	0.999250081289737		9.5	6.5	6.5	1.5	8.5	9.5	9.5	9.5	4,008.5	4,008.3

# Step 6: Graph-based segment search

Greedy strategy:

Starting from larger segment, exhaustively reconstruct within a bounding box for a limited number of layers

Register all layers into one 2D candidate

Identify direction of the process, and then expand bounding box according to it.

Repeat steps until the connectivity path reaches another large segment or another smaller path



## **Reconstruction Results**



# **Reconstruction Results**



# Quantitative Performance

Ground-truth: 444 APEX positive segments

Recall statistic: 91.8% for the 2D segmentation portion of the algorithm

Two orders of magnitude faster compared to Random Forests\* (supervised machine learning)

### Not perfectly accurate, but **much** faster

\*Kaynig et al. Medical Image Analysis 2015

# Quantitative Performance

	3D U-Net	Ours (Cluster Based Thresh.)	Ours (Adaptive Thresh.)				
Manual Annotation	40 Hr.	-	-				
Training	926 Min.	-	-				
Segmentation	233 Min.	8 Min.	8 Min.				
Reconstruction	_	128.8 Min.	16.82 Min.				

## Also works for X-ray





### Reconstruction

# **Reconstruction Results**





Cells

### Vasculature + Cells

# Apex Cells Imaged with X-ray



# Comparison to deep learning



# Bringing learning back in

# Open Set Machine Learning

Can we bring supervised machine learning back into the picture to handle unknown data?



Kasthuri et al. Cell 2015

# "There are known knowns..."

**known classes:** the classes with distinctly labeled positive training examples (also serving as negative examples for other known classes)

**known unknown classes**: labeled negative examples, not necessarily grouped into meaningful categories

unknown unknown classes: classes unseen in training



# Learning Objective



## MICrONS use case: "synaptomics"



# Model: Compact Abating Probability



# Binary RBF SVM incorporating a CAP model: W-SVM

Combine probabilities computed for both 1-class and binary RBF SVMs

1-class SVM CAP model is a conditioner



## Step 1: Train a 1-class SVM $f^{o}$



# Step 2: Fit Weibull over tail of scores from $f^{o}$



# Step 3: Train a binary SVM f



Class Label = '3' Known Negative Classes = '0', '1', '2'

# Step 4: Fit EVT distributions over tails of scores from f



# W-SVM Object Recognition



# Incremental learning



# Extreme Value Machine (EVM)





Addition of a new class

# EVM Learning Objective:



## Incremental learning: ImageNet



Bendale and Boult CVPR 2015

## Basic Machine Learning Benchmark: LETTER



# Thank you!

(web) www.wjscheirer.com (code coming soon) https://github.com/CVRL