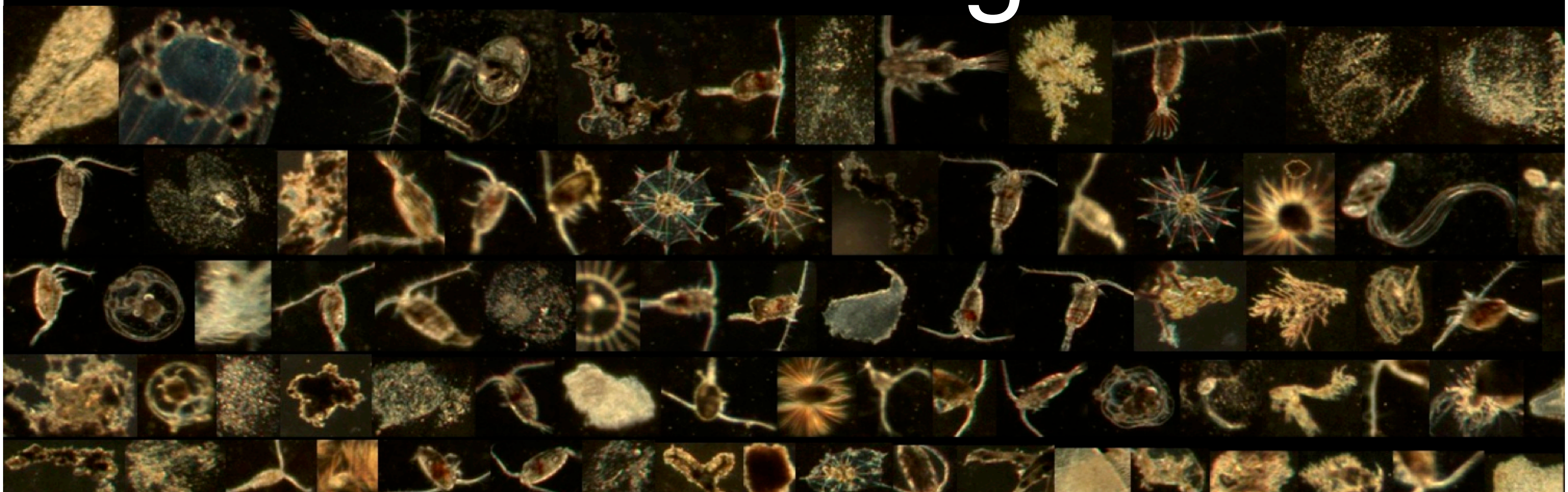
Automated Analysis of Planktonic Image Data



Eric Orenstein

Guest lecture – ECE228/SIO209

April 25th, 2018

Automated Analysis of Planktonic Image Data

i. Studying plankton
ii. Scripps Plankton Camera – system overview
iii. Transfer learning and deep feature extraction for planktonic image data sets
   - Orenstein and Beijbom, *Applications of Computer Vision (WACV),* 2017 *IEEE Winter Conference on*.
iv. High-temporal resolution *in situ* imaging reveals dynamics of copepod-parasite interaction

Studying plankton

1 mm

Studying plankton

- Basis of marine food web

Studying plankton

1 mm

- Basis of marine food web
- Link atmosphere and deep ocean

Studying plankton

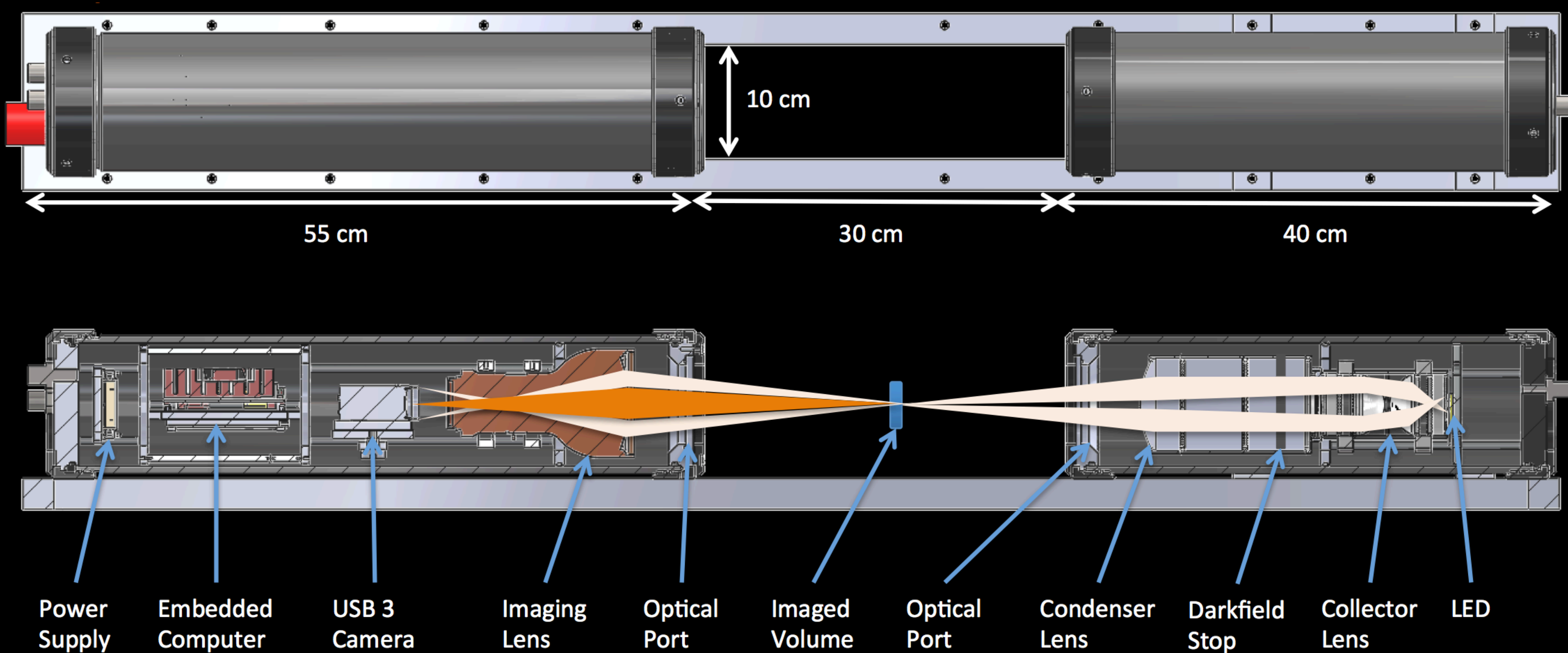
1 mm

- Basis of marine food web
- Link atmosphere and deep ocean
- Influence biogeochemical cycles

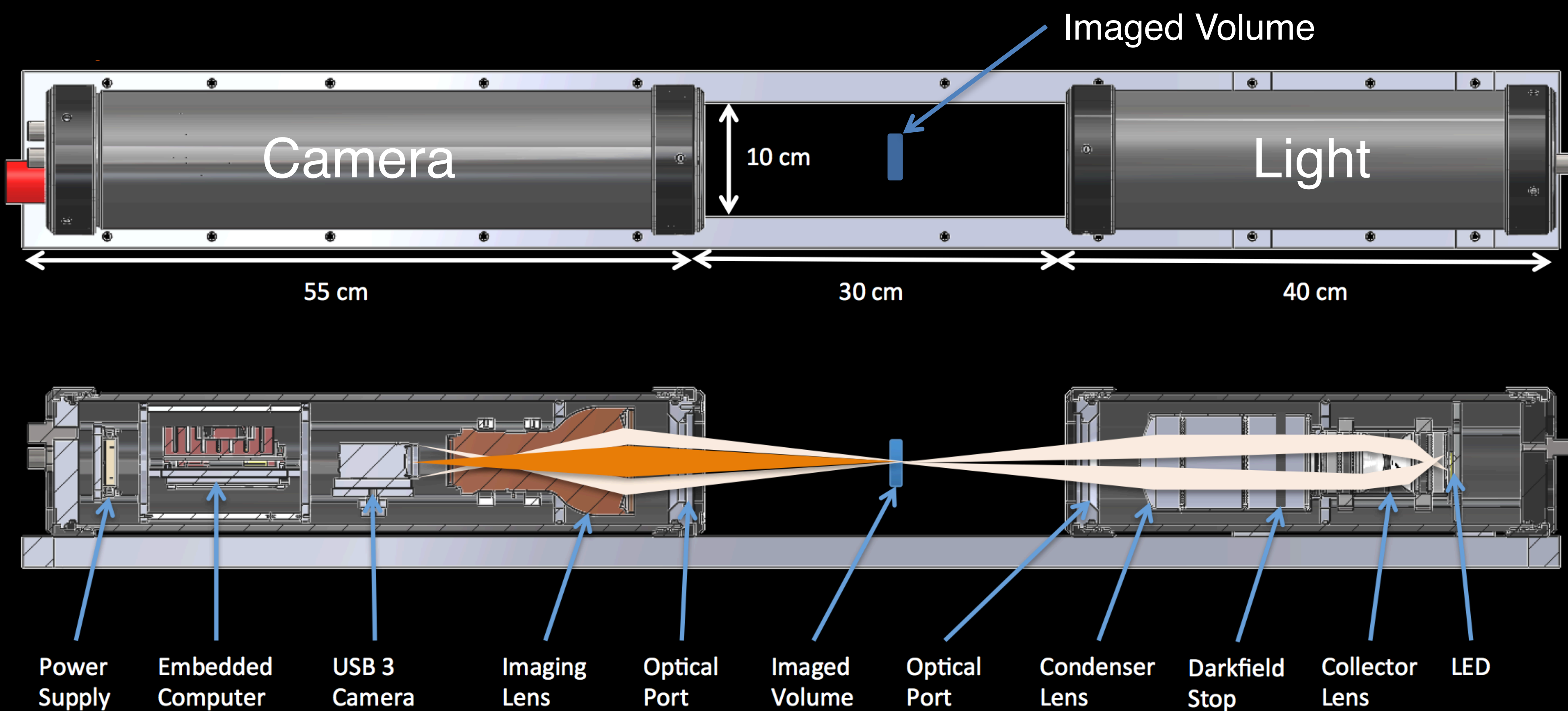
Studying plankton

- What organisms live in the ocean?
- How many of them are there?
- What are they doing?
  - Changes over time?
  - Interactions?
- Difficult to answer due to technological limitations

Scripps Plankton Camera

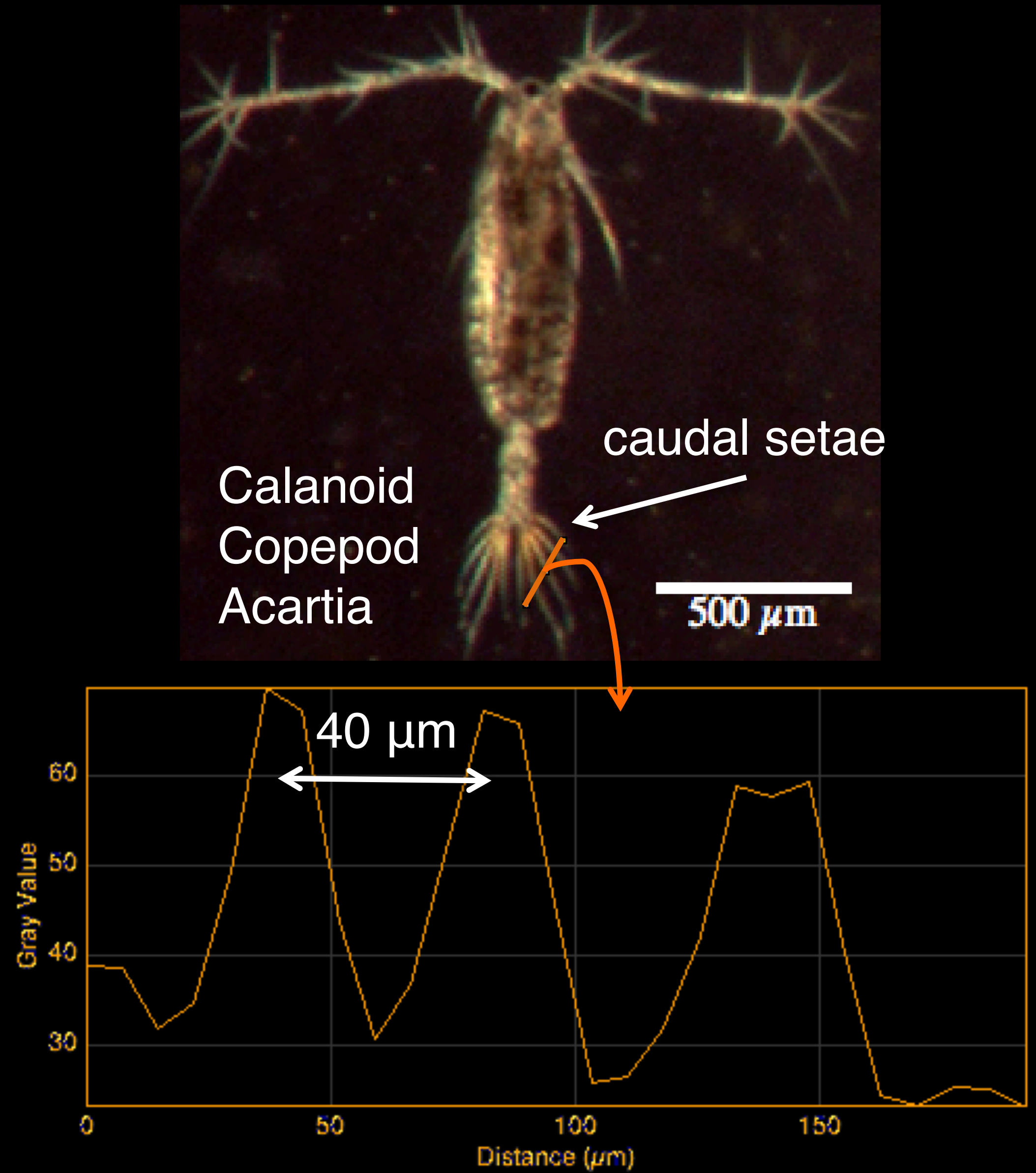


Scripps Plankton Camera

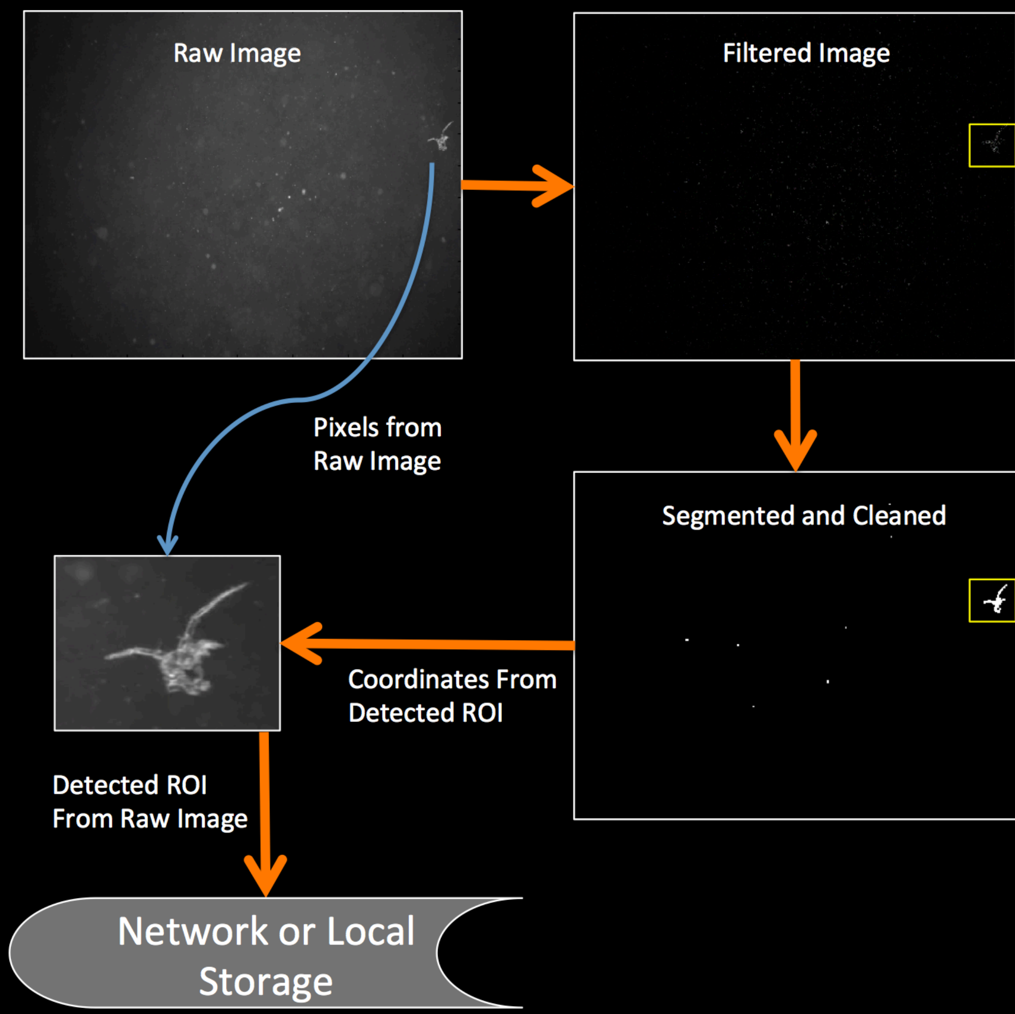


SPC - Specifications

Property	Specification
Field of View	25 mm x 20 mm
Resolution	7.4 µm pixels 35 lp/mm @ 40 % contrast
Depth of Field	400 µm @ 35 lp/mm @ 20 % contrast
Hi-Resolution Volume	0.2 mL per Frame
Blob-Detection Volume	10 mL Frame
Data Rate	Up to 8 fps with ROI processing



SPC – Data Logging



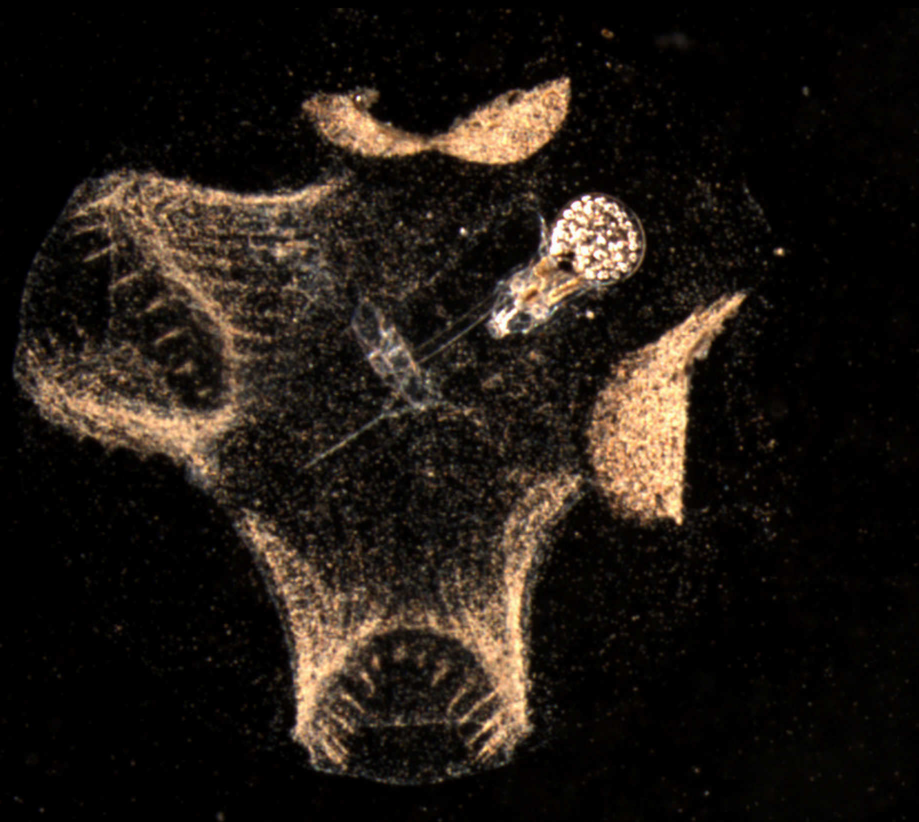
SPC - Deployment

SPC - Deployment

SPC - Deployment

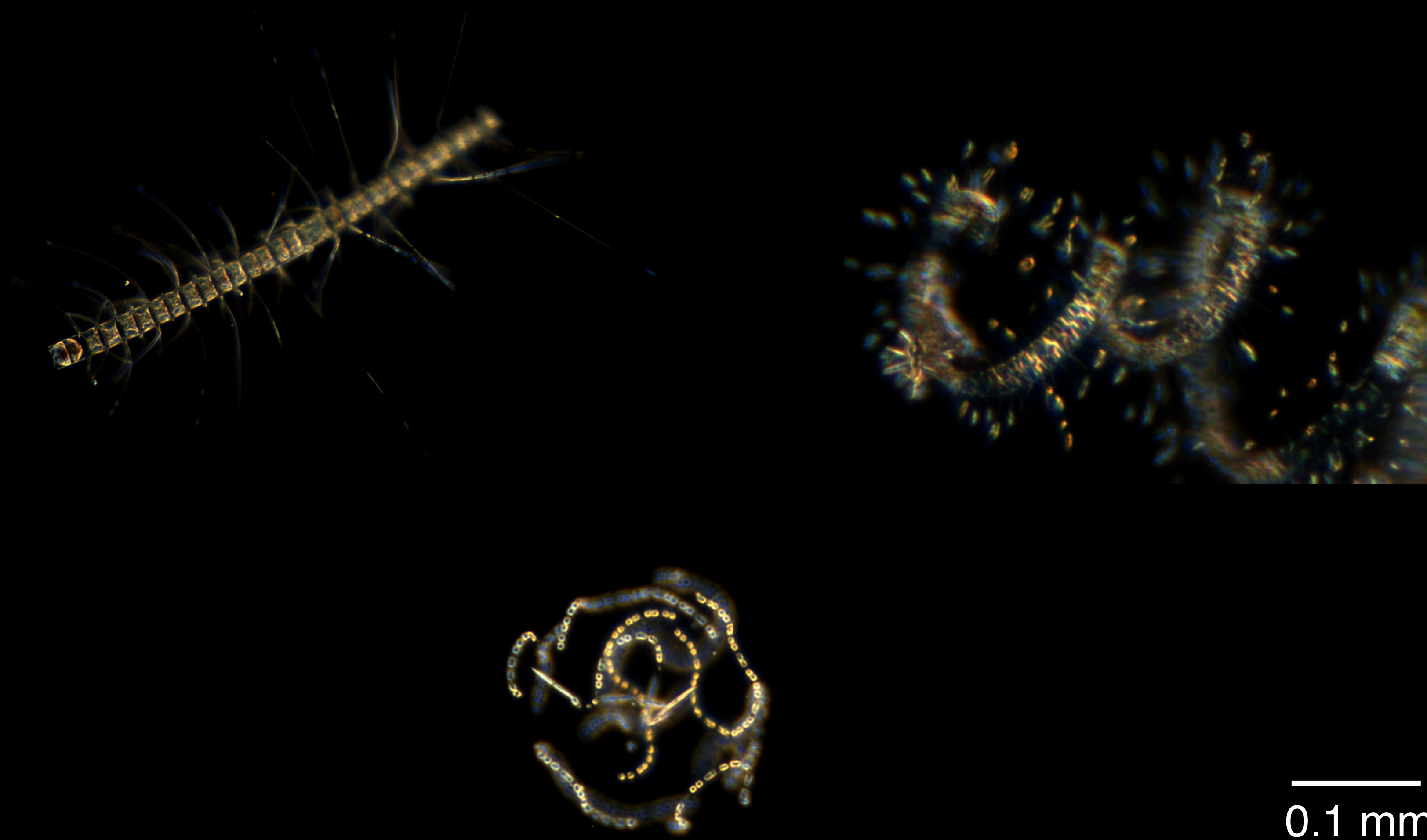
Photo: Emily Kelly

Fragile objects

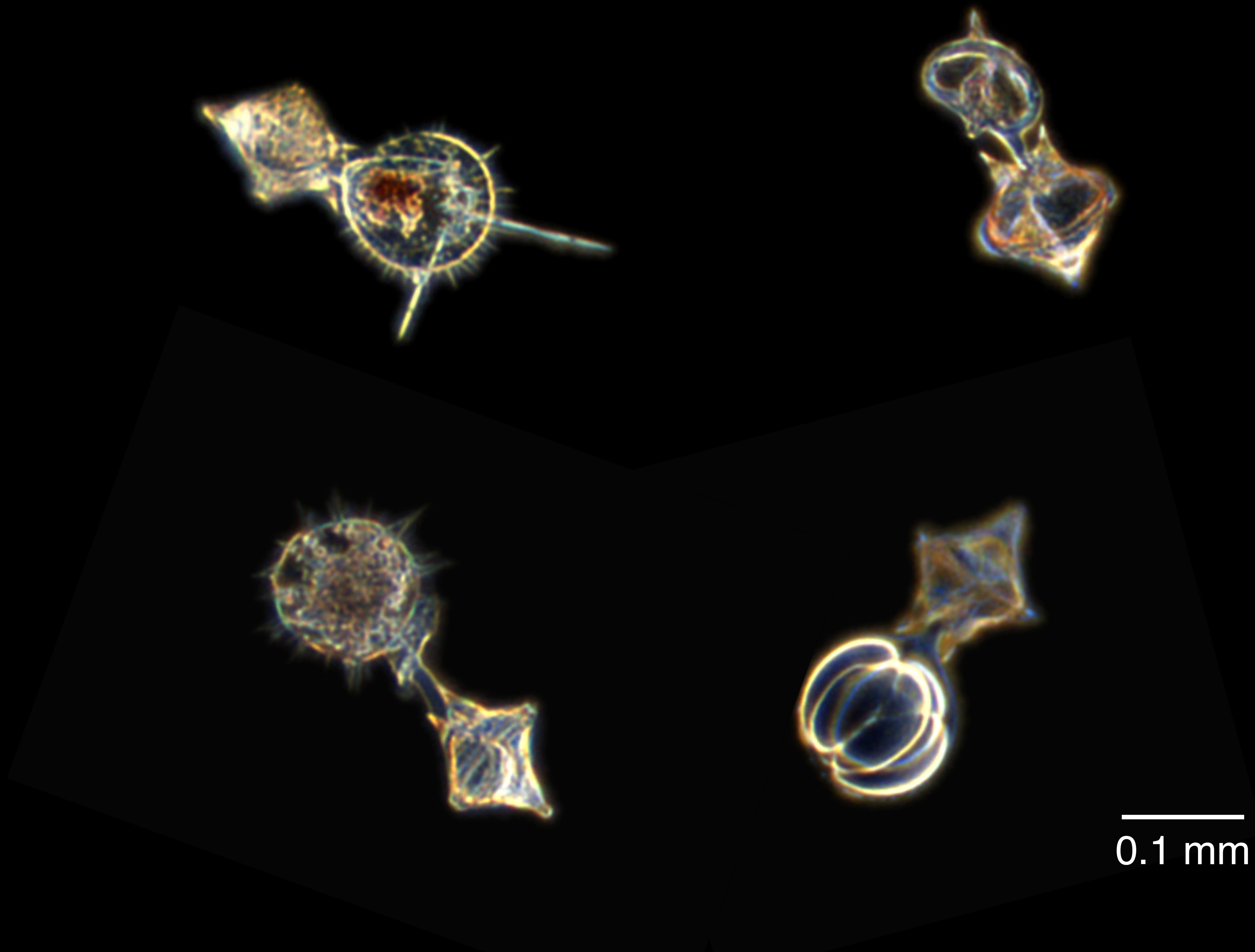


Background – SPC

Fragile objects

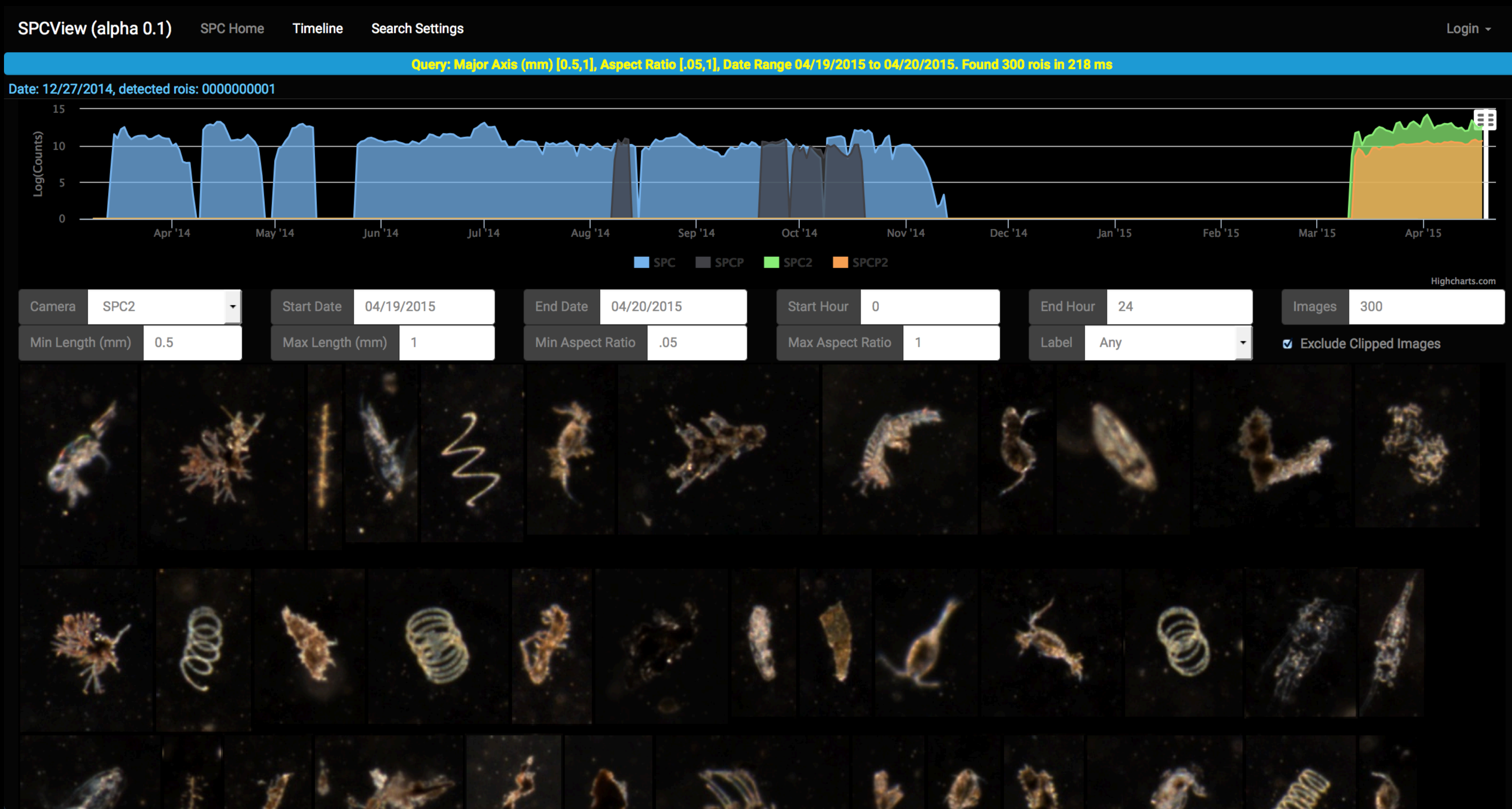


Short duration events



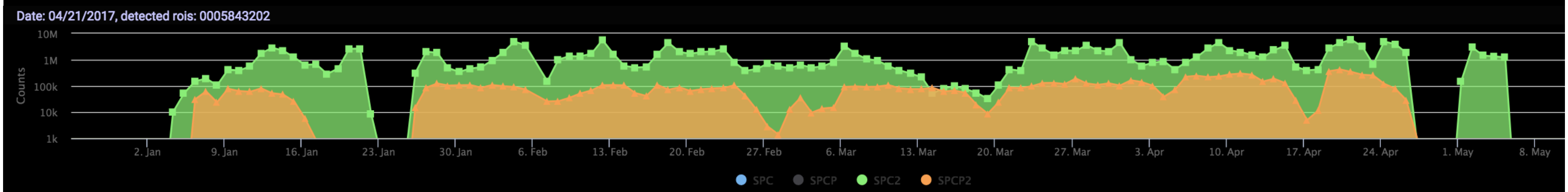
Background – SPC

spc.ucsd.edu

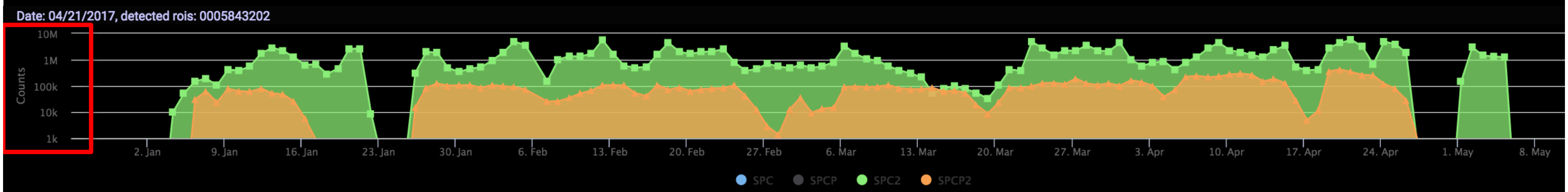


SPCView (alpha 0.1)
SPC Home
Timeline
Search Settings
Login
Query: Major Axis (mm) [0.5,1], Aspect Ratio [.05,1], Date Range 04/19/2015 to 04/20/2015. Found 300 rois in 218 ms
Date: 12/27/2014, detected rois: 0000000001
Log(Counts)
15
10
5
0
Apr '14
May '14
Jun '14
Jul '14
Aug '14
Sep '14
Oct '14
Nov '14
Dec '14
Jan '15
Feb '15
Mar '15
Apr '15
SPC
SPCP
SPC2
SPCP2
Highcharts.com
Camera
SPC2
Start Date
04/19/2015
End Date
04/20/2015
Start Hour
0
End Hour
24
Images
300
Min Length (mm)
0.5
Max Length (mm)
1
Min Aspect Ratio
.05
Max Aspect Ratio
1
Label
Any
Exclude Clipped Images

What now?

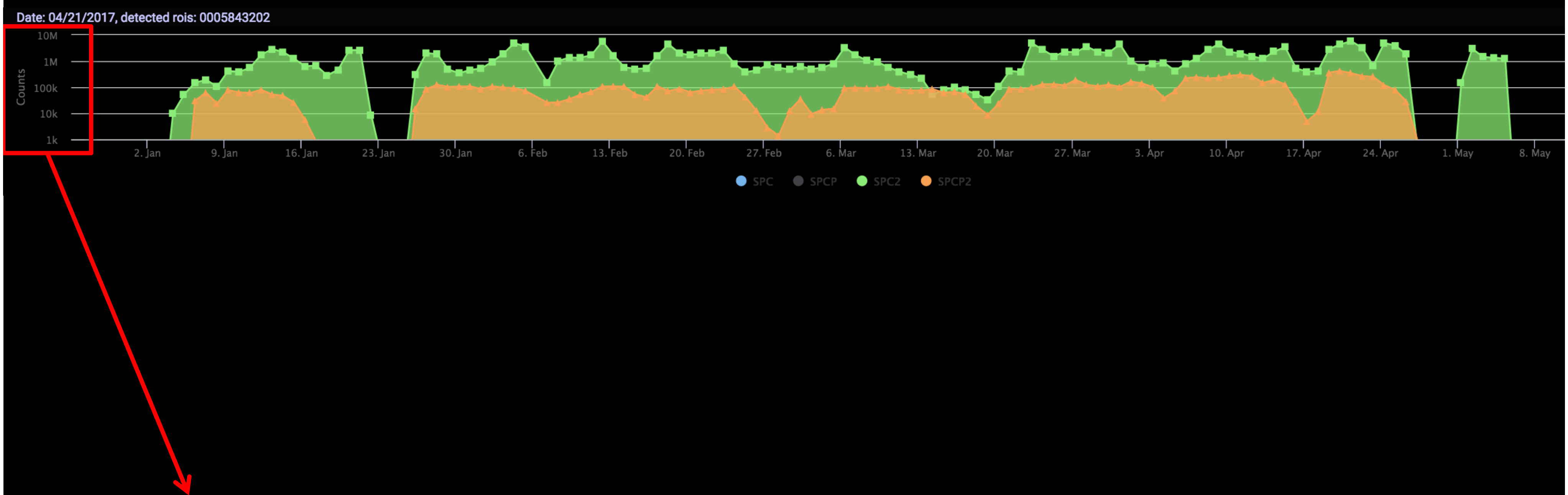

Date: 04/21/2017, detected rois: 0005843202
10M
1M
100k
10k
1k
Counts
2. Jan
9. Jan
16. Jan
23. Jan
30. Jan
6. Feb
13. Feb
20. Feb
27. Feb
6. Mar
13. Mar
20. Mar
27. Mar
3. Apr
10. Apr
17. Apr
24. Apr
1. May
8. May
SPC
SPCP
SPC2
SPCP2

What now?

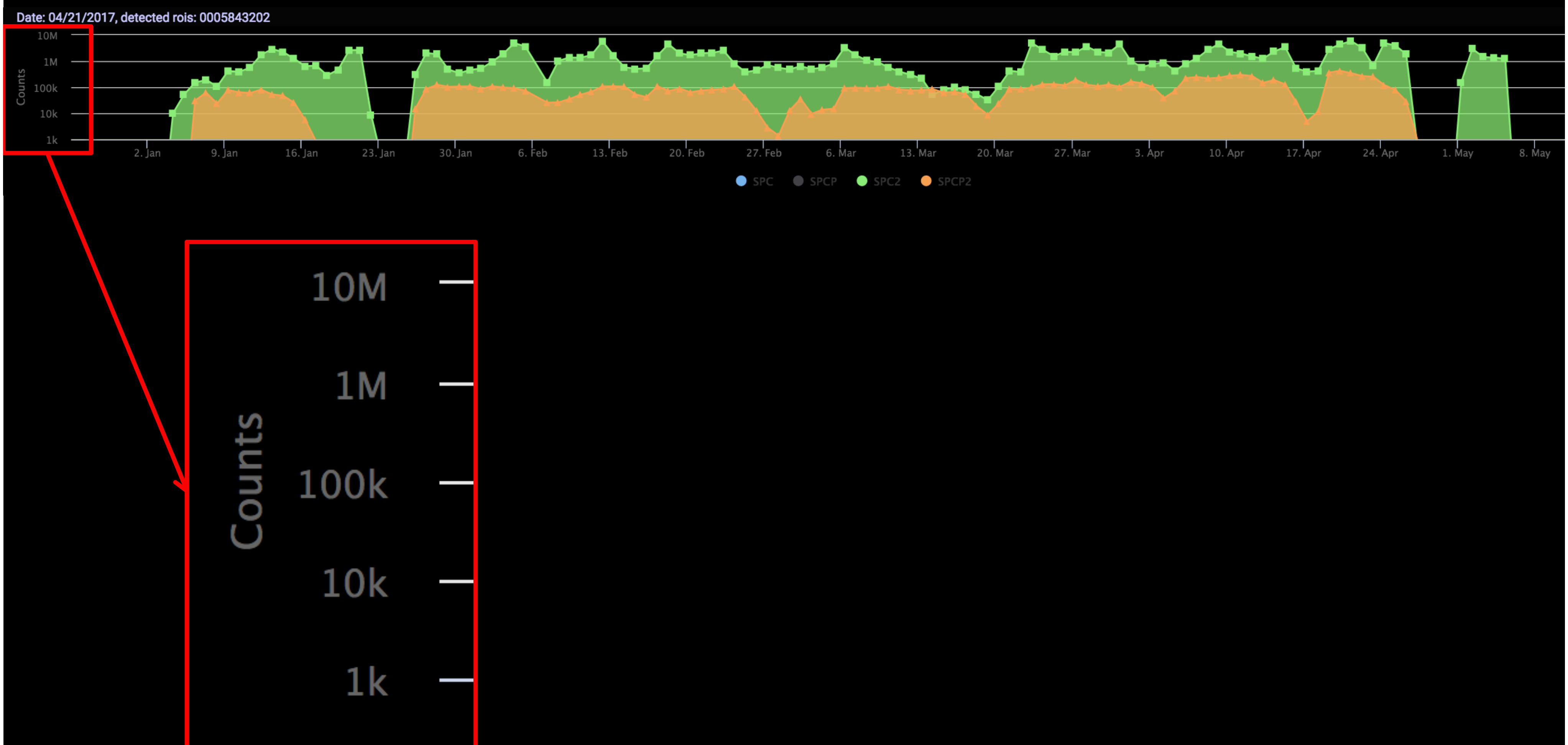


Date: 04/21/2017, detected rois: 0005843202
10M
1M
100k
10k
1k
Counts
2. Jan
9. Jan
16. Jan
23. Jan
30. Jan
6. Feb
13. Feb
20. Feb
27. Feb
6. Mar
13. Mar
20. Mar
27. Mar
3. Apr
10. Apr
17. Apr
24. Apr
1. May
8. May
SPC
SPCP
SPC2
SPCP2

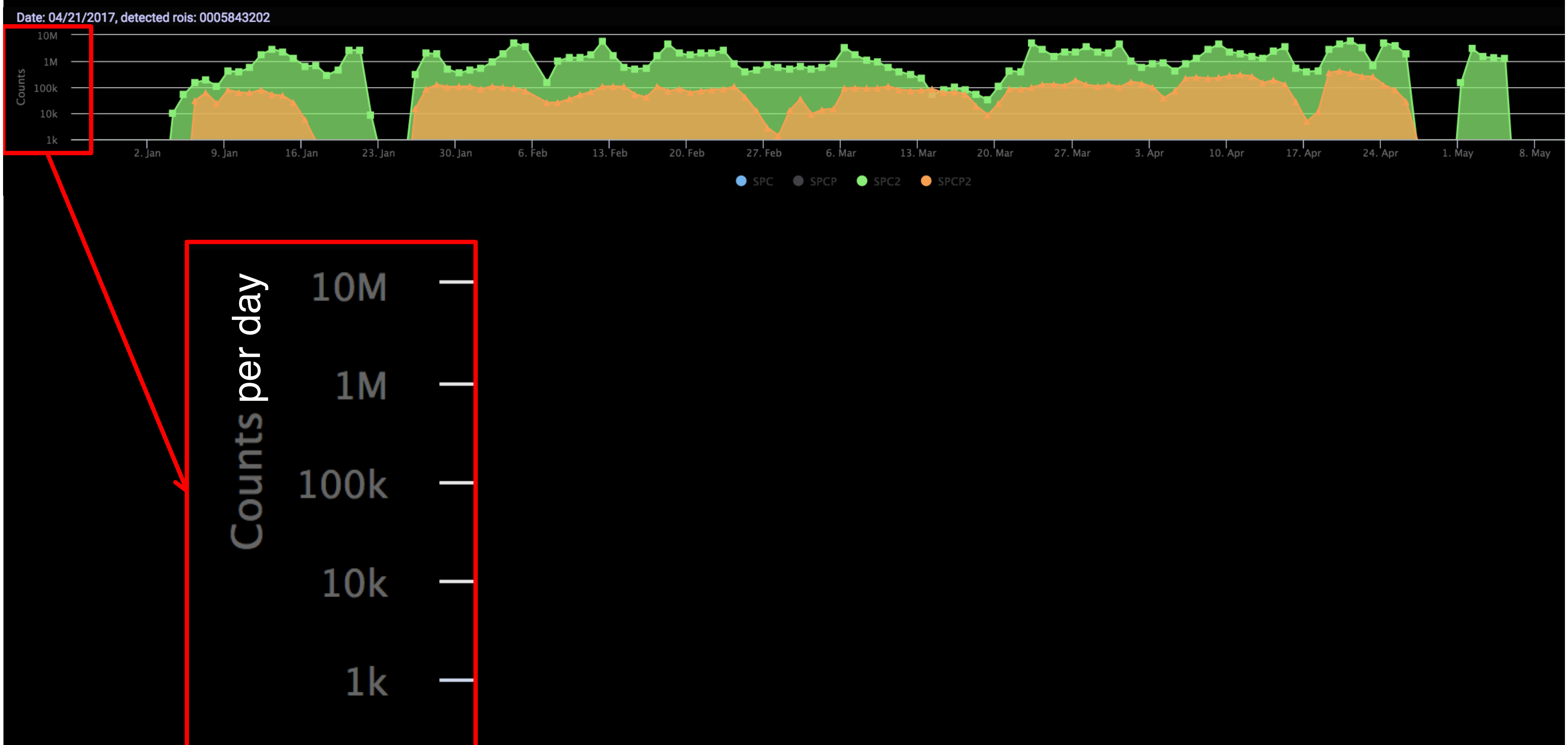
What now?


Date: 04/21/2017, detected rois: 0005843202
Counts
10M
1M
100k
10k
1k
2. Jan
9. Jan
16. Jan
23. Jan
30. Jan
6. Feb
13. Feb
20. Feb
27. Feb
6. Mar
13. Mar
20. Mar
27. Mar
3. Apr
10. Apr
17. Apr
24. Apr
1. May
8. May
SPC
SPCP
SPC2
SPCP2

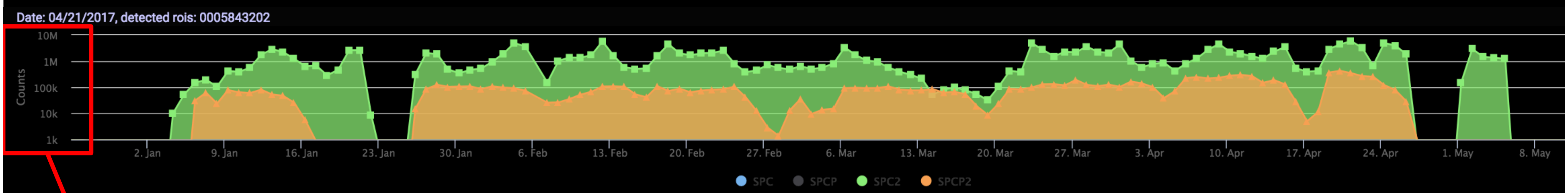
What now?


Date: 04/21/2017, detected rois: 0005843202
Counts
10M
1M
100k
10k
1k
2. Jan
9. Jan
16. Jan
23. Jan
30. Jan
6. Feb
13. Feb
20. Feb
27. Feb
6. Mar
13. Mar
20. Mar
27. Mar
3. Apr
10. Apr
17. Apr
24. Apr
1. May
8. May
SPC
SPCP
SPC2
SPCP2
Counts
10M
1M
100k
10k
1k

What now?


Date: 04/21/2017, detected rois: 0005843202
Counts
10M
1M
100k
10k
1k
2. Jan
9. Jan
16. Jan
23. Jan
30. Jan
6. Feb
13. Feb
20. Feb
27. Feb
6. Mar
13. Mar
20. Mar
27. Mar
3. Apr
10. Apr
17. Apr
24. Apr
1. May
8. May
SPC
SPCP
SPC2
SPCP2
Counts per day
10M
1M
100k
10k
1k

What now?



- Data set contains over 800 million ROIs
- ~100k have been labeled
- <<1% of the images have been labeled

What now?

- Boost human annotation efforts with machine learning
- Test a variety of ML approaches
  - Random Forests (Blaschko et al., 2005)
  - Support Vector Machines (Sosik et al., 2007)
  - Convolutional Neural Networks (Orenstein et al., 2015)
- Goal: Test the effectiveness of out-of-domain data for plankton classification

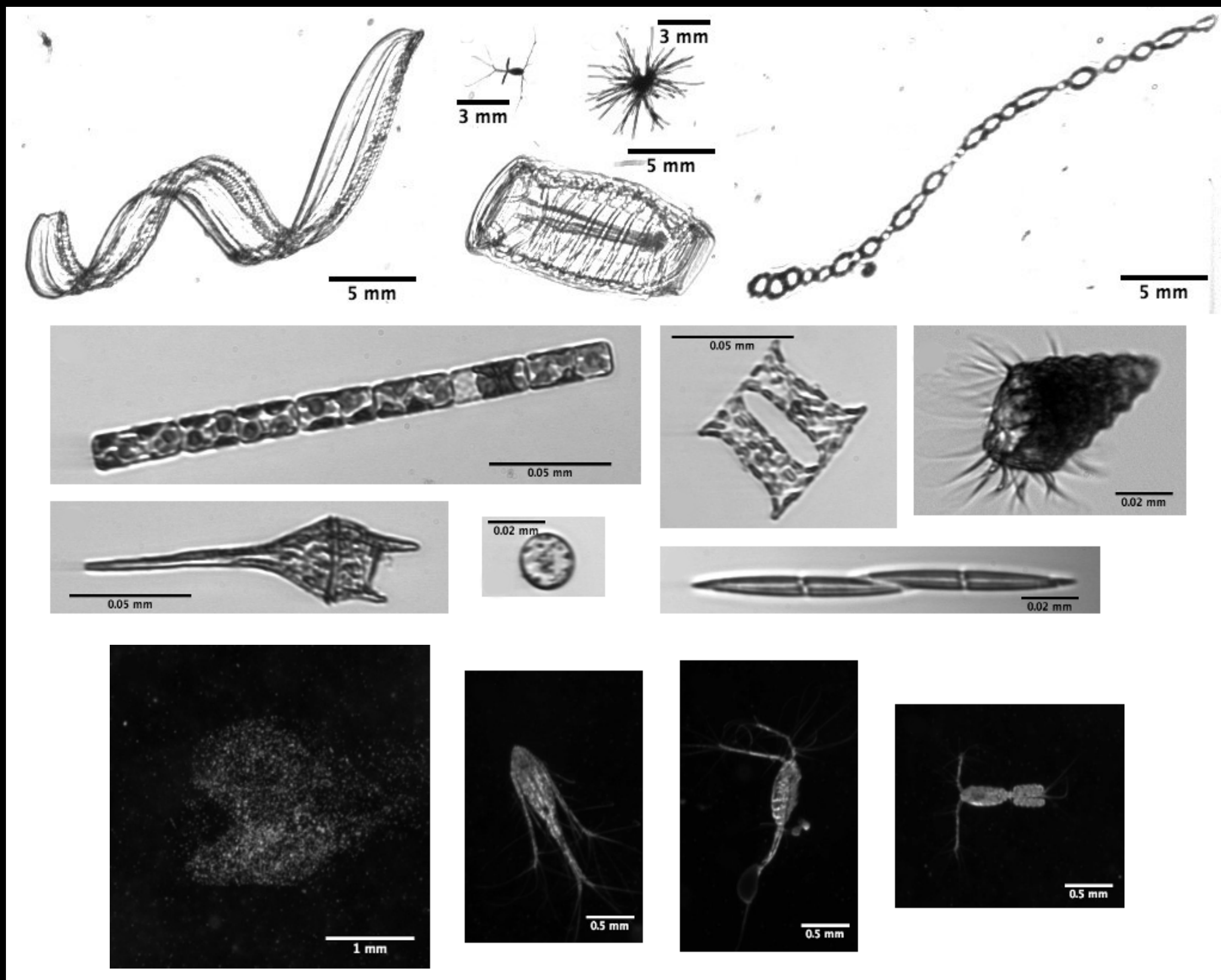
In situ imaging systems

In Situ Ichthyoplankton Imaging System (ISIIS)

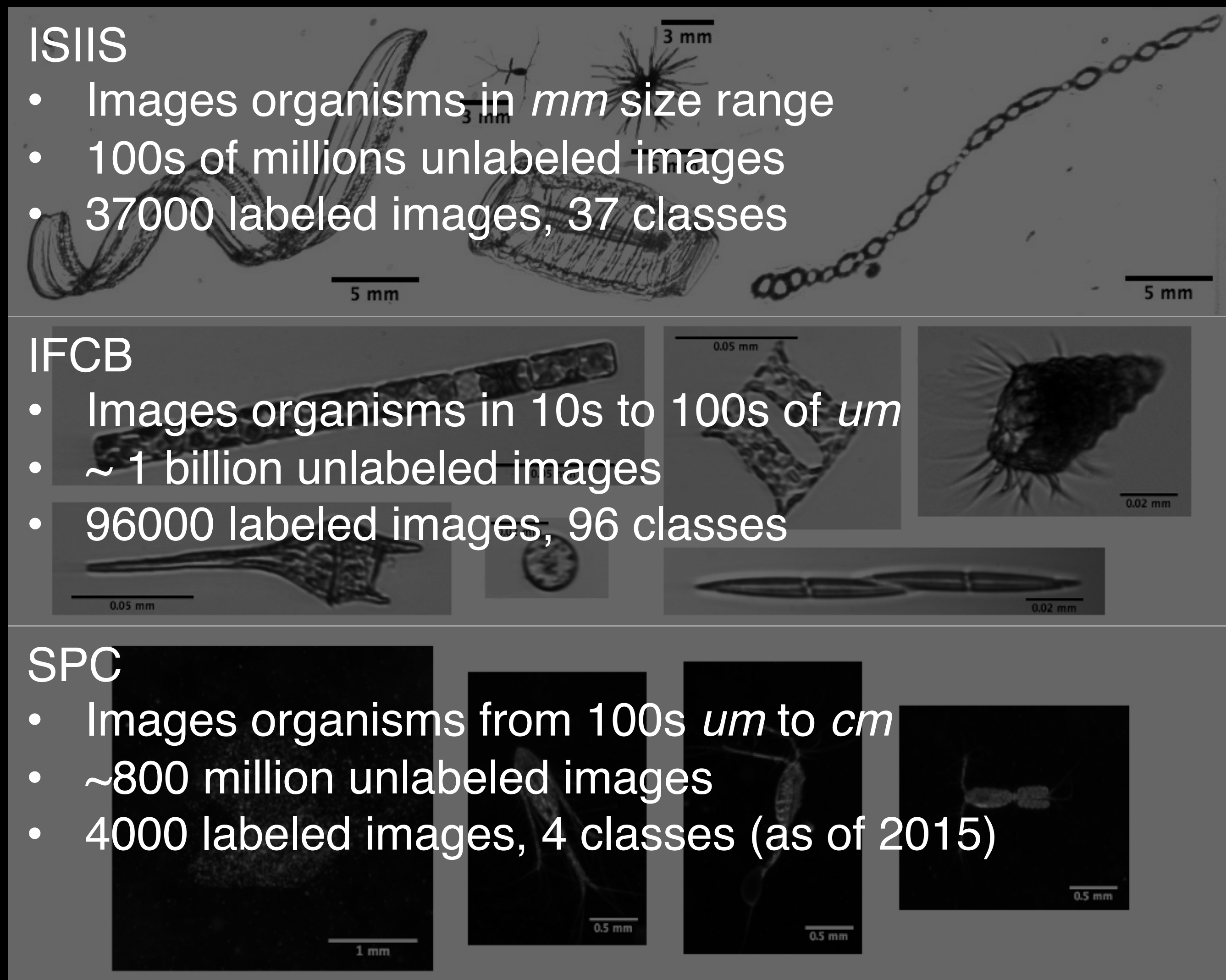
Imaging FlowCytobot (IFCB)

Scripps Plankton Camera (SPC)

In situ imaging systems



In situ imaging systems

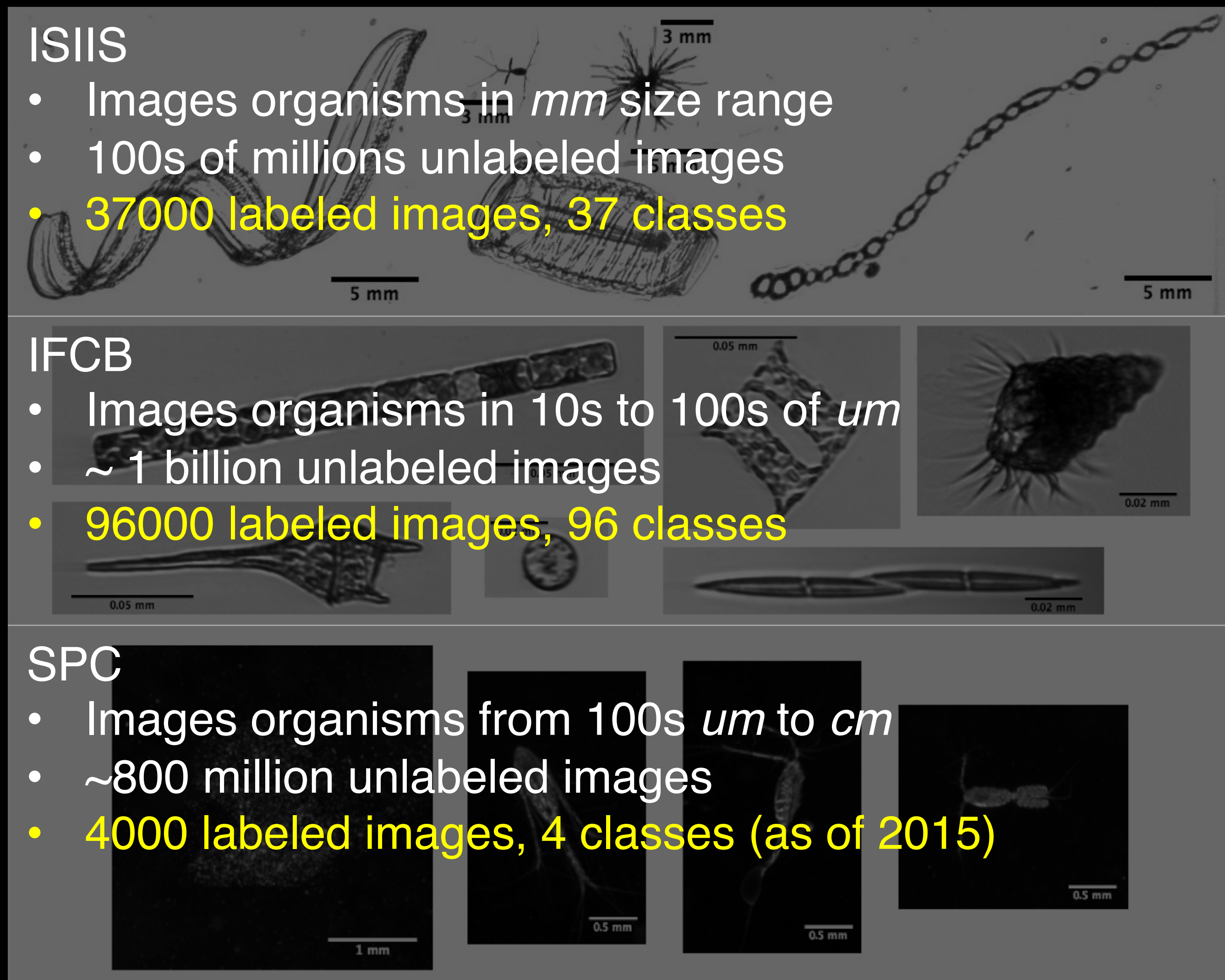


ISIIS
- Images organisms in *mm* size range
- 100s of millions unlabeled images
- 37000 labeled images, 37 classes

IFCB
- Images organisms in 10s to 100s of *um*
- ~ 1 billion unlabeled images
- 96000 labeled images, 96 classes

SPC
- Images organisms from 100s *um* to *cm*
- ~800 million unlabeled images
- 4000 labeled images, 4 classes (as of 2015)

In situ imaging systems



ISIIS
- Images organisms in *mm* size range
- 100s of millions unlabeled images
- 37000 labeled images, 37 classes

IFCB
- Images organisms in 10s to 100s of *um*
- ~ 1 billion unlabeled images
- 96000 labeled images, 96 classes

SPC
- Images organisms from 100s *um* to *cm*
- ~800 million unlabeled images
- 4000 labeled images, 4 classes (as of 2015)

Machine learning

- Random forests
  - Ensemble classifier
  - Operates on ‘hand-engineered’ features
- Convolutional Neural Networks (CNNs)
  - Representation learning
  - Operates directly on raw data
- Both types require labeled data for training and evaluation

Feature extraction

Feature extraction

$$= [x_1]$$

Feature extraction

$$= [x_1, x_2]$$

Feature extraction

$$= [x_1, x_2, x_3]$$

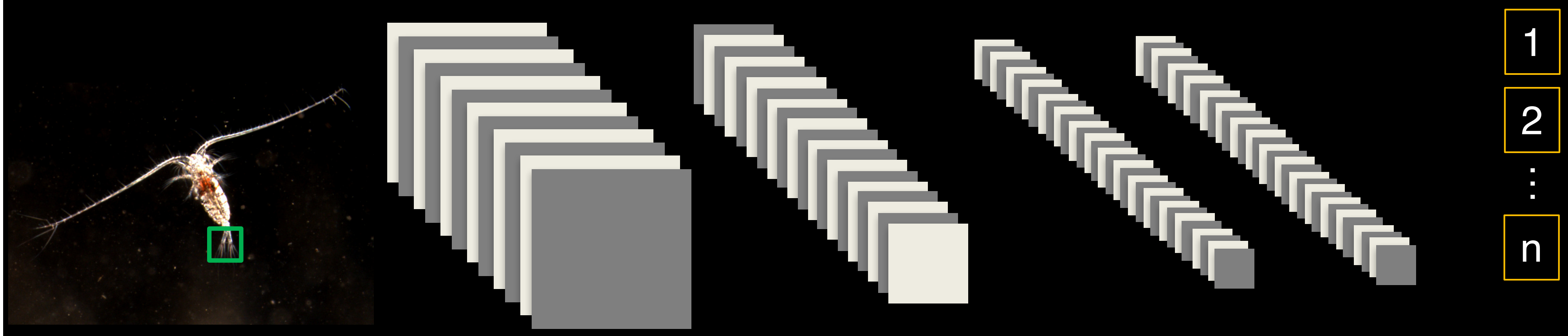
Feature extraction

$$= [x_1, x_2, x_3, \ldots, x_n]$$

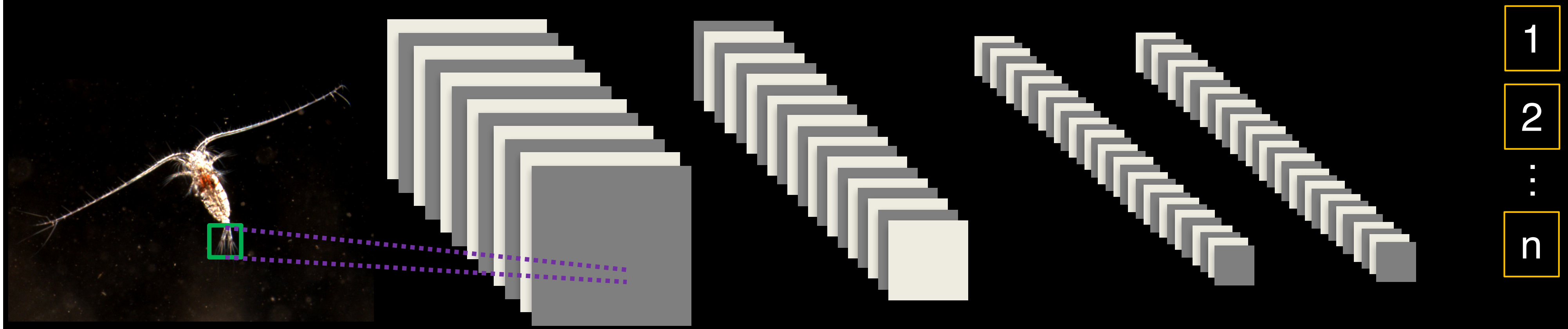
Feature extraction

$$= \begin{bmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{j1} & \cdots & x_{jn} \end{bmatrix}$$

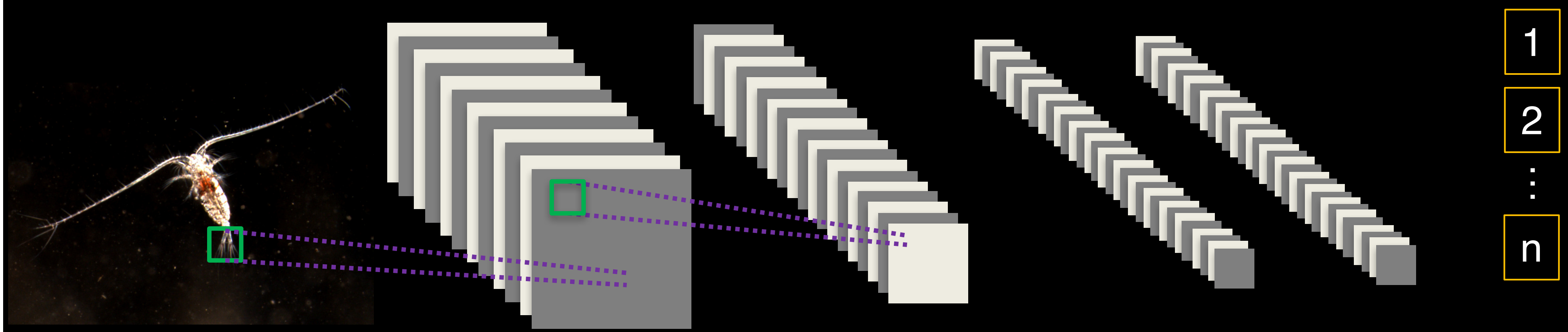
Convolutional Neural Networks



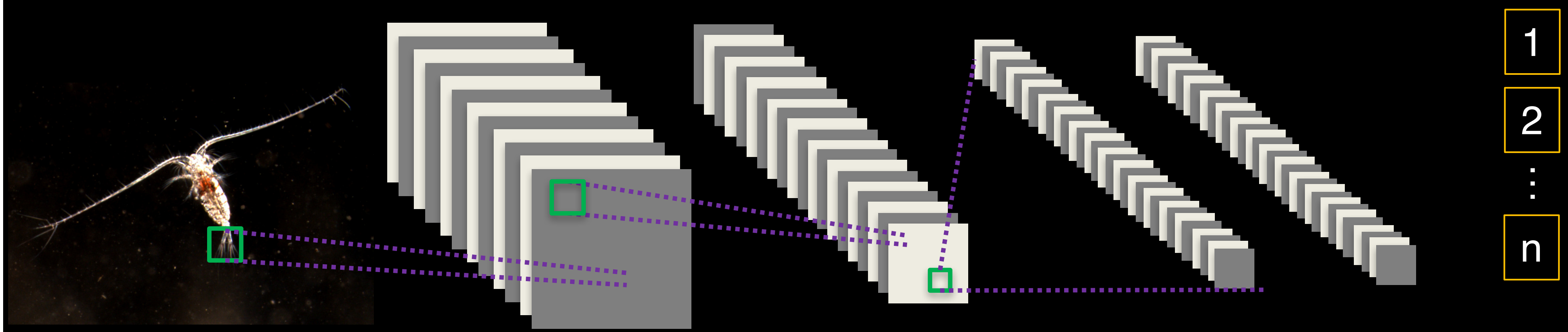
Convolutional Neural Networks



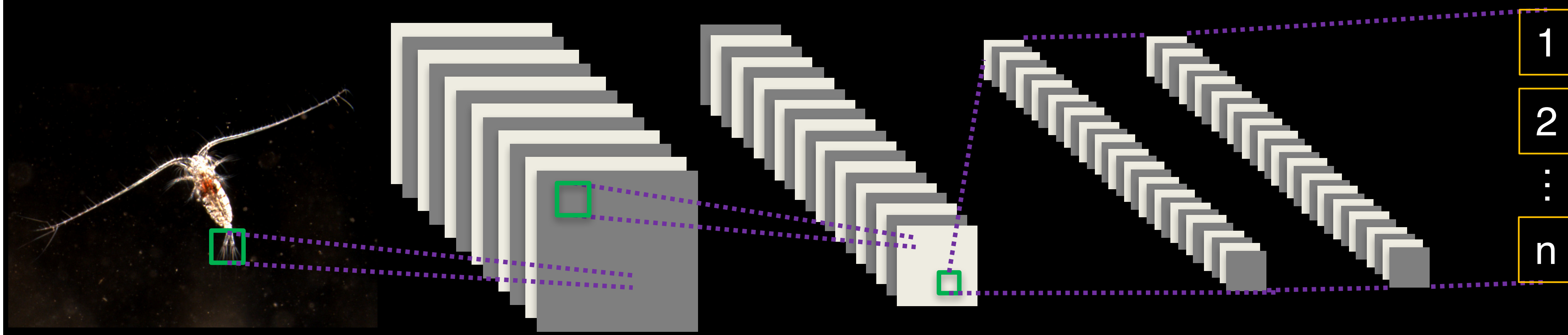
Convolutional Neural Networks



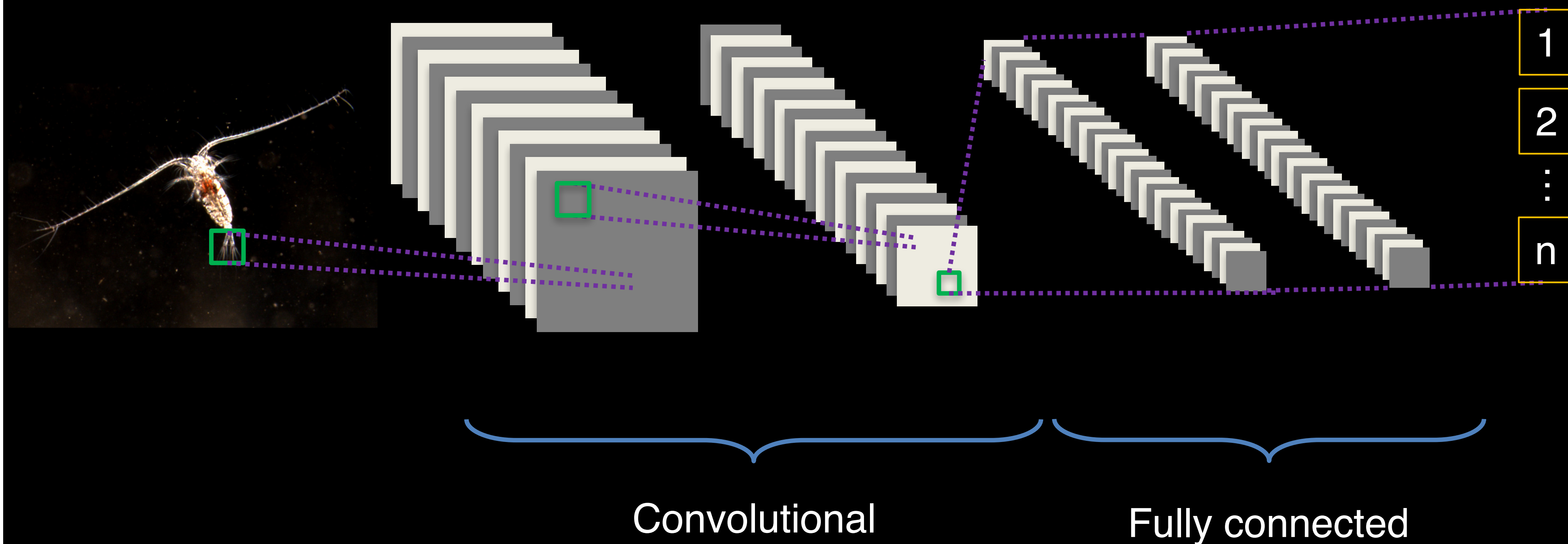
Convolutional Neural Networks



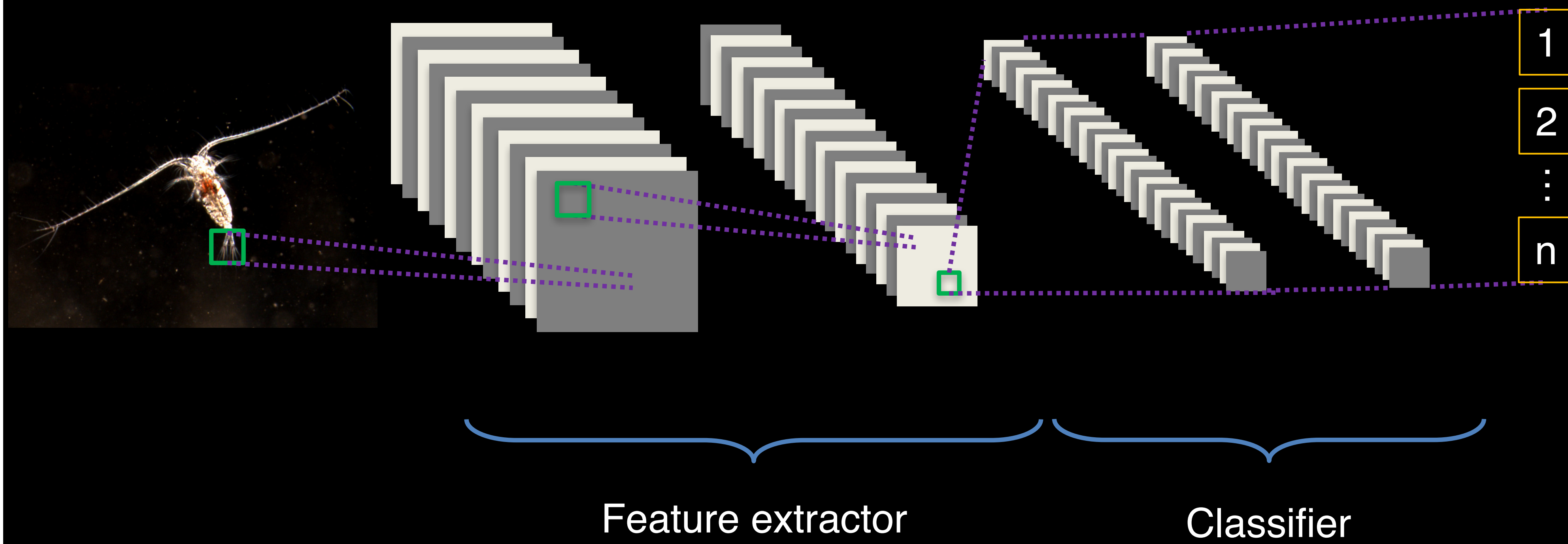
Convolutional Neural Networks



Convolutional Neural Networks



Convolutional Neural Networks



Convolutional Neural Networks

- Fine tuning
  - Cut off top layers of a network and retrain
  - Decreases training time, increases accuracy with small training set (Yosinski et al., 2014)
- Deep features
  - CNN weights used for margin or ensemble classifiers (Donahue et al., 2013)

Fine-tuning and feature extraction

- Deep nets based on AlexNet:
  - 2 from scratch with ISIIS and IFCB
  - 2 fine-tuned from ImageNet with ISIIS and IFCB
  - 2 double fine-tuned with both plankton datasets
  - ImageNet weights with no plankton data

Fine-tuning and feature extraction

- Random Forests:
  - Hand-engineered features
  - Deep features
  - Deep + hand-engineered features

Results

CNN		*ifcb*	*ImageNet ➔ifcb*	*ImageNet ➔isiis ➔ifcb*	*isiis*	*ImageNet ➔isiis*	*ImageNet ➔isiis ➔ifcb*	*ImageNet*
	IFCB	0.78	**0.86**	**0.86**	-	-	-	-
	ISIIS	-	-	-	0.71	**0.83**	**0.83**	-

- Double fine-tuning had a slight positive effect on classifier accuracy
- Suggests that a machine classifier treats IFCB and ISIIS data sets as very similar

Results

		ifcb	*ImageNet* → *ifcb*	*ImageNet* → *isiis* → *ifcb*	*isiis*	*ImageNet* → *isiis*	*ImageNet* → *isiis* → *ifcb*	*ImageNet*
Deep feat	IFCB	-	-	-	0.65	0.77	-	**0.81**
	ISIIS	0.56	**0.65**	-	-	-	-	0.63
	SPC	0.57	0.69	0.67	0.52	0.65	0.65	**0.71**
Deep + Hand	IFCB	-	-	-	0.68	0.78	-	**0.81**
	ISIIS	0.65	**0.70**	-	-	-	-	0.66
	SPC	0.67	0.76	0.74	0.69	0.74	0.74	**0.77**

- Addition of hand-engineered features had slight positive effect

Results

- Demonstrated value of out-of-domain data
  - Weights from standard object detectors useful
  - Potential to combine ocean image data for more powerful models
- Further development needed for deployment on real data
  - Data–set shift: prior probability of classes changing with time (Moreno-Torres et al., 2012)

Thank you

Eric Orenstein
e1orenst@ucsd.edu