Land Cover Mapping

Nebojsa Jojic

People

CELA

- Lucas Joppa, Dan Morris and the rest of AI for Earth

Interns/contractors

- Kolya Malkin, Yale
- Caleb Robinson, Georgia Tech
- Le Hou, Stony Brook
- Anthony Ortiz, U of Texas El Paso/MILA
- Eli Cole, Caltech

AI Residents

- Andi Peng
- Blake Elias

(Some) external collaborators:

- Rachel Soobitsky, Jeff Allenby et al, Chesapeake Conservancy
- Jarlath O'Neil Dunne, U of Vermont
- Bistra Dilkina, USC
- Kai Kaiser, World Bank

Some of these people's faces

Caleb Robinson
Research Intern

Anthony Ortiz
Research Intern

Blake Elias
AI Resident

Kolya Malkin
Research Intern

Andi Peng
AI Resident

Nebojsa Jojic
Principal Researcher

Dan Morris
Principal Scientist and
Aspiring Rock Icon

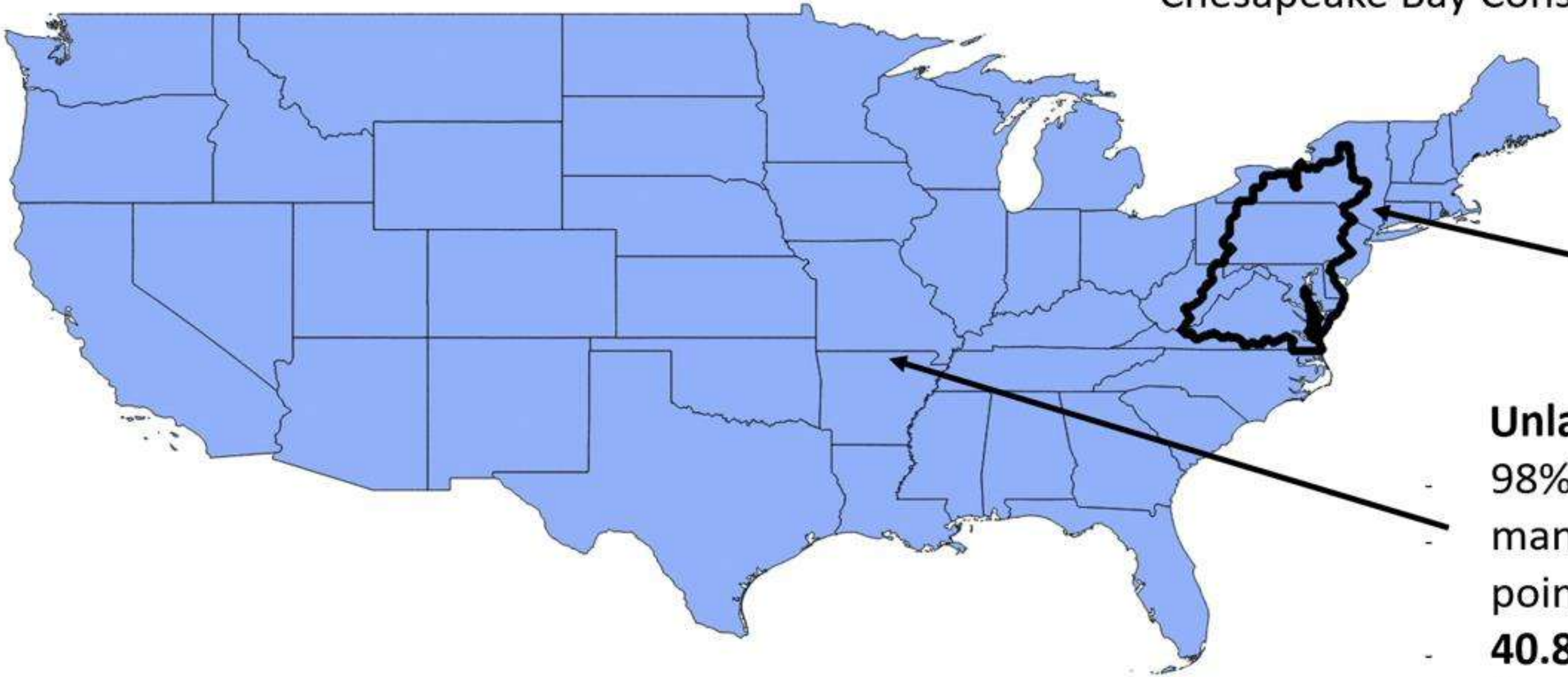
Bistra Dilkina
Assistant Professor of Computer Science
University of Southern California

https://www.microsoft.com/en-us/research/project/land-cover-mapping/

- Introduction
- Label super-resolution (ICLR 2019)
- First US-wide 1m land cover map (CVPR 2019)
- iNaturalist species observation (just starting on this)
- Hybrid intelligence approaches to land cover mapping (under review)
- Applications/collaborations
- Open problems

Mapping the US at 1m resolution (ICLR and CVPR 2019)

Chesapeake Bay Conservancy



Labeled

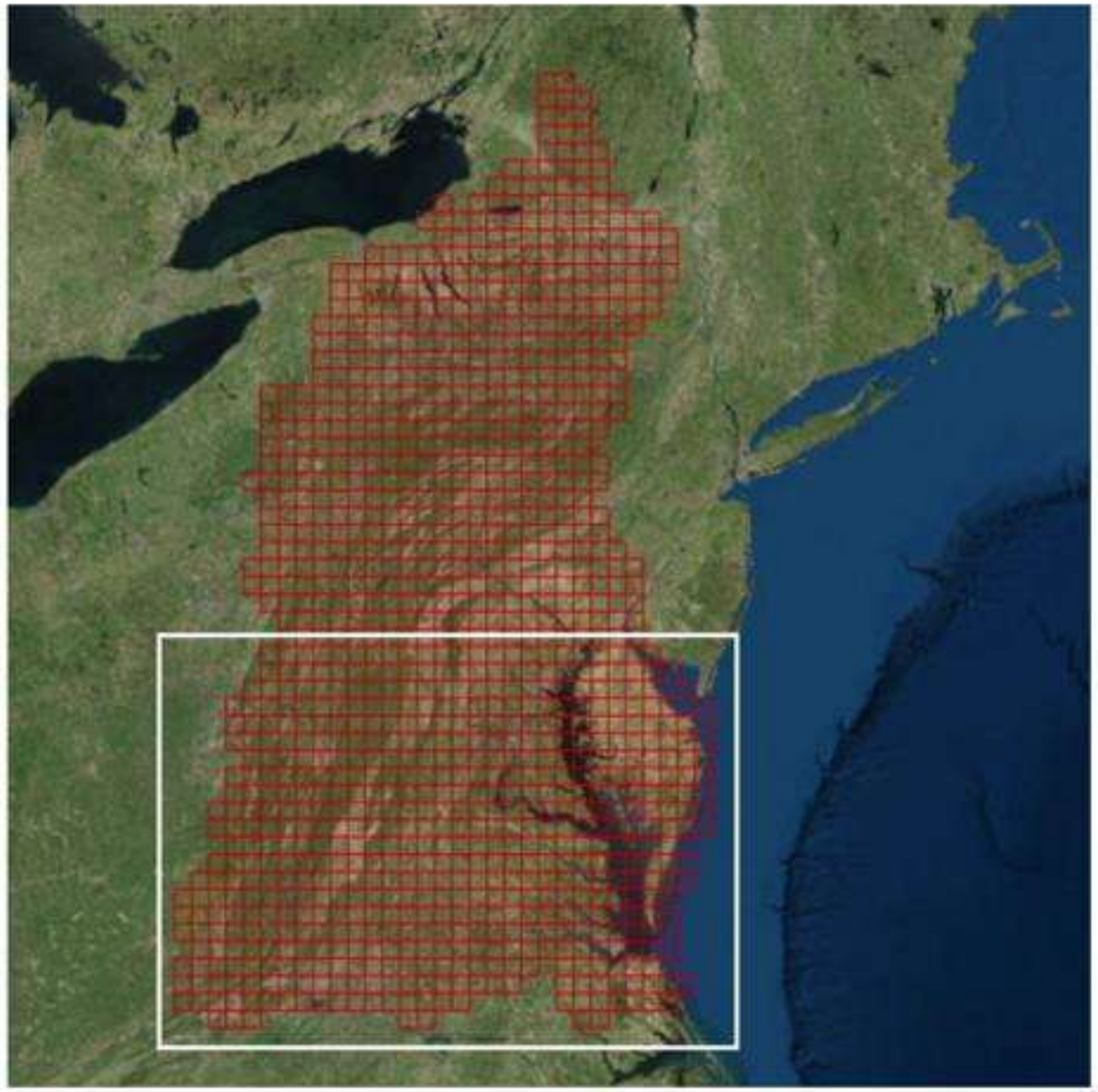
- 2% of area
- 1 time point
- 10 months
- $1.3 million

Unlabeled

- 98% of area
- many time points
- **40.8 years?**
- **$63.7 million?**

https://chesapeakeconservancy.org/conservation-innovation-center/high-resolution-data/

The Chesapeake Conservancy data



One of those 1000 tiles (16k X 16k pixels)

RGB image (there is also near IR)

Labels

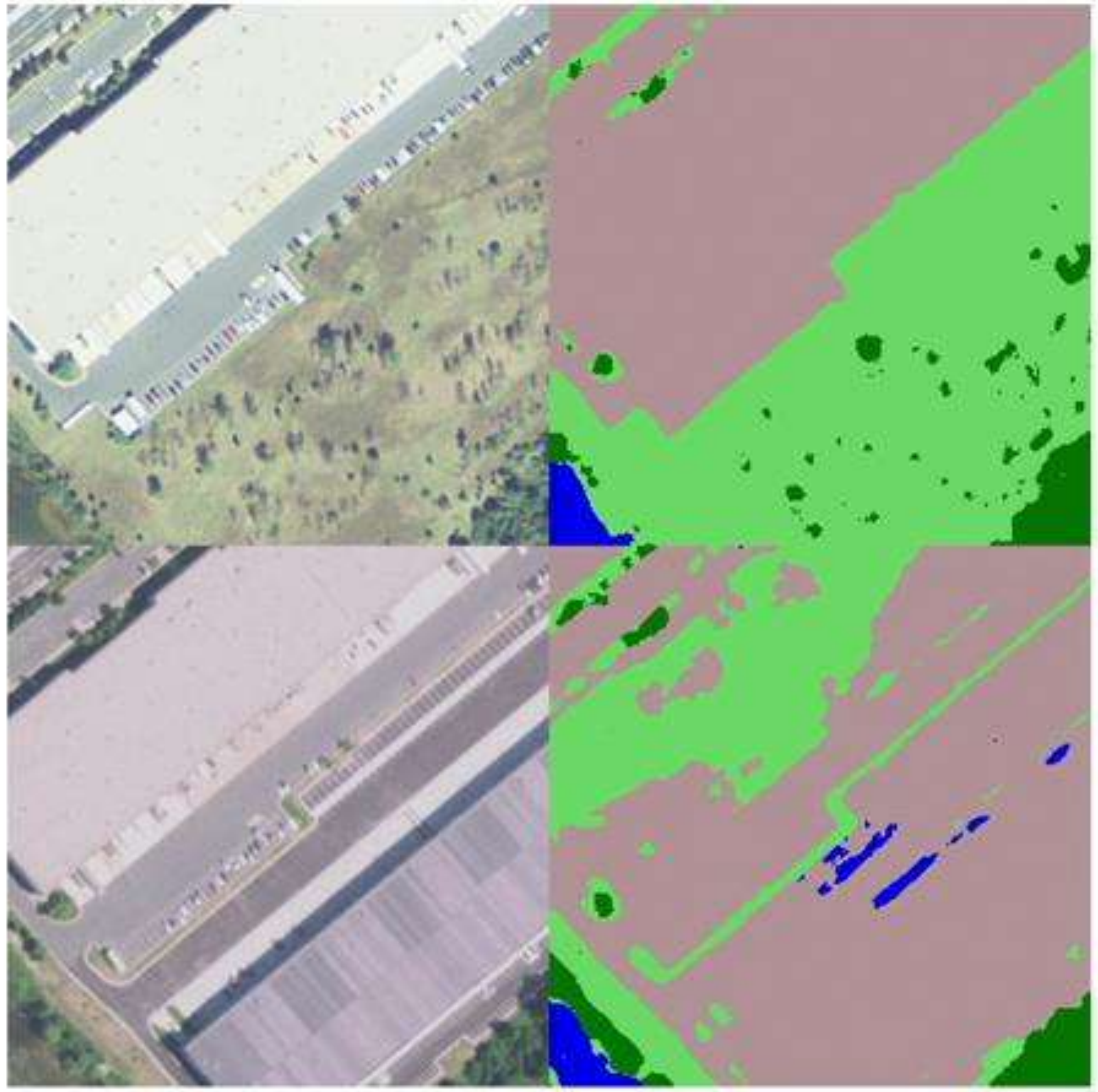
A small fraction of a tile: One pixel = $1m^2$

Why do we need high resolution land cover maps?

- **Change detection!**
    - Urban sprawl

Pre 2013
imagery/predicted
land cover

Post 2013
imagery/predicted
land cover

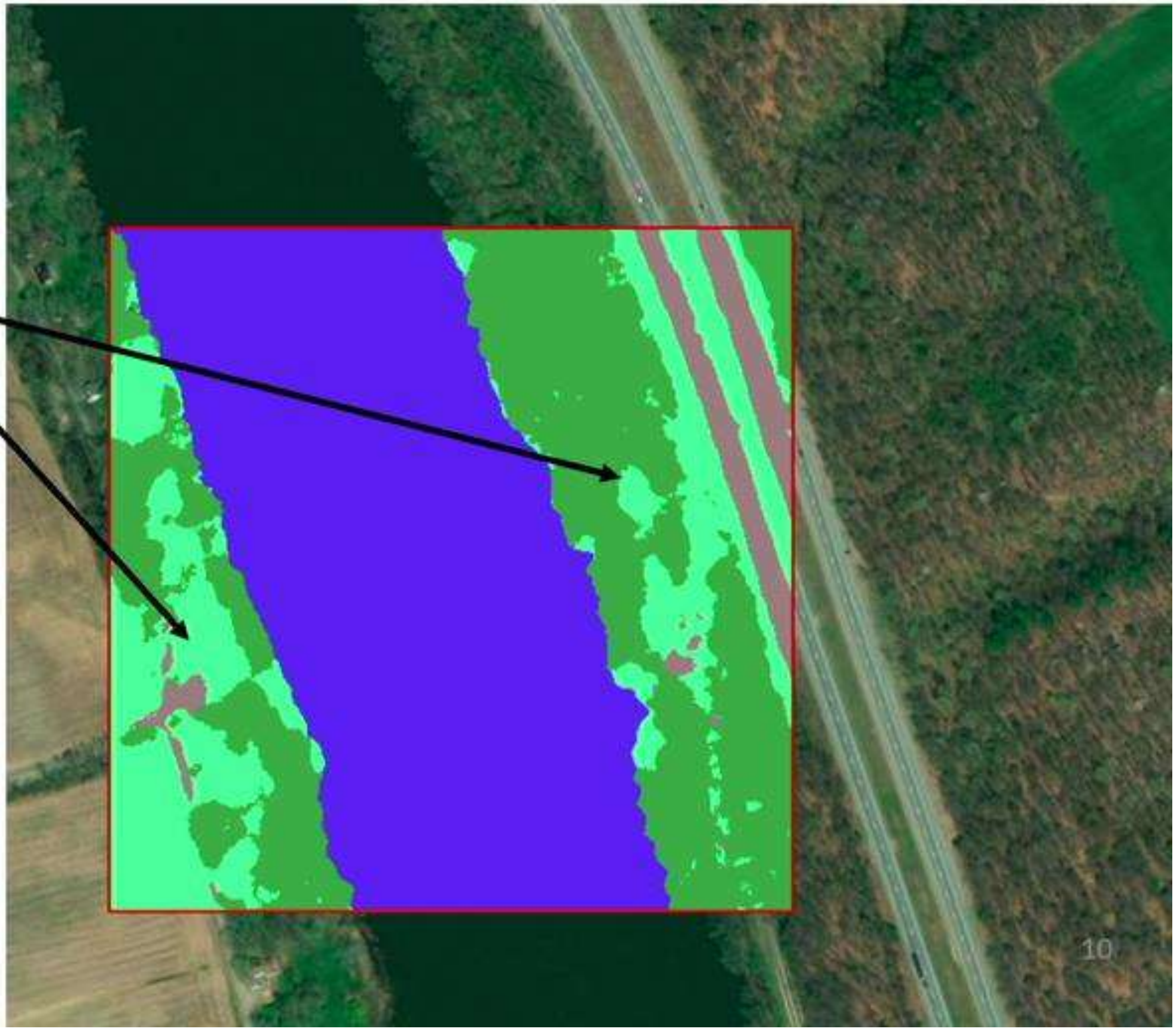


No actual land cover
change over years

Why do we need high resolution land cover maps?

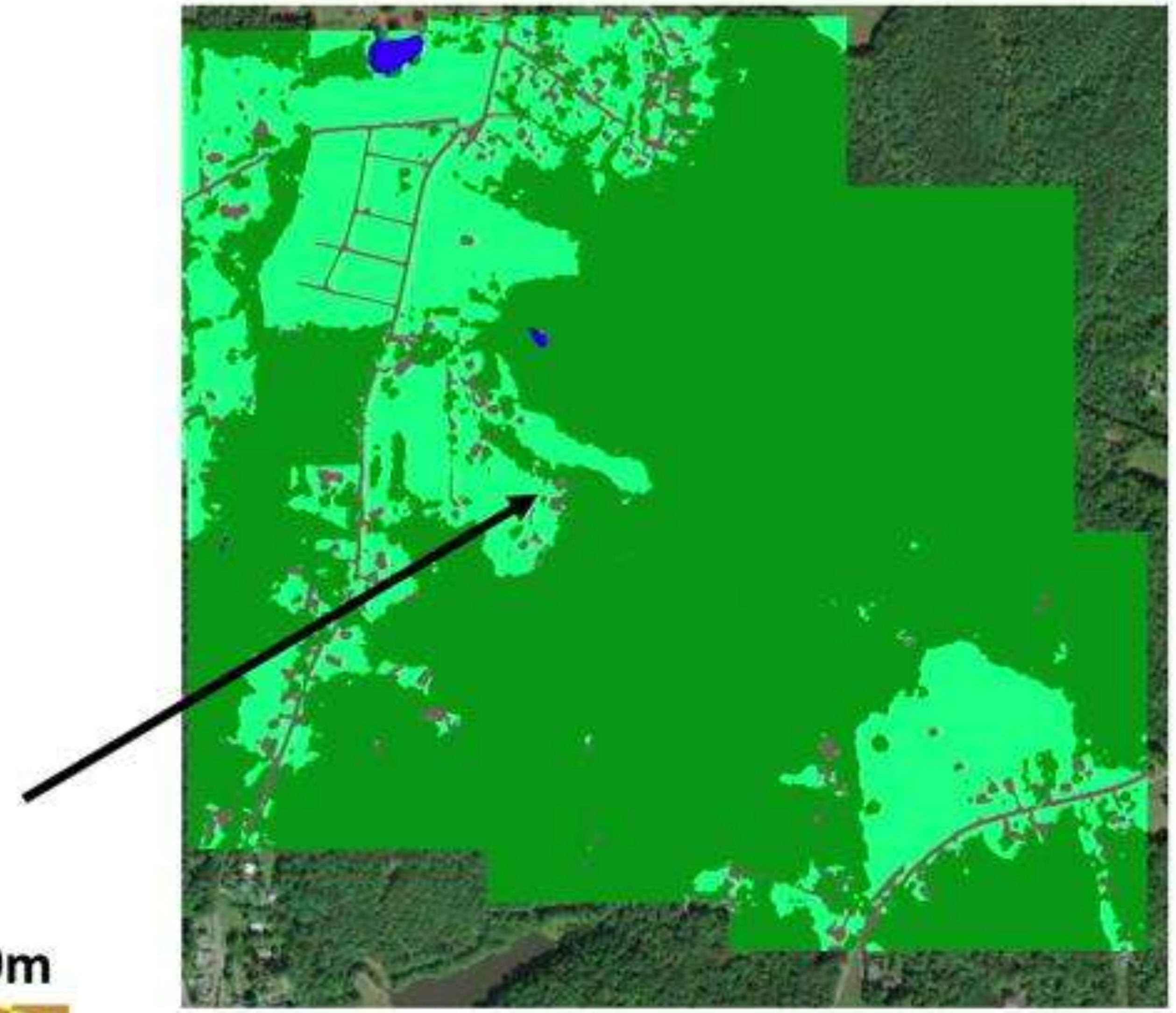
- Change detection
  - Urban sprawl
- **Conservation efforts!**
  - Where are riparian buffers?

(Chesapeake Conservancy measures water quality goals using land cover data)



Why do we need high resolution land cover maps?

- Change detection
  - Urban sprawl
- **Conservation efforts!**
  - Where are riparian buffers
  - Quantifying forest areas



At lower resolutions we will miss pockets of deforested areas at boundaries

High res at 1m

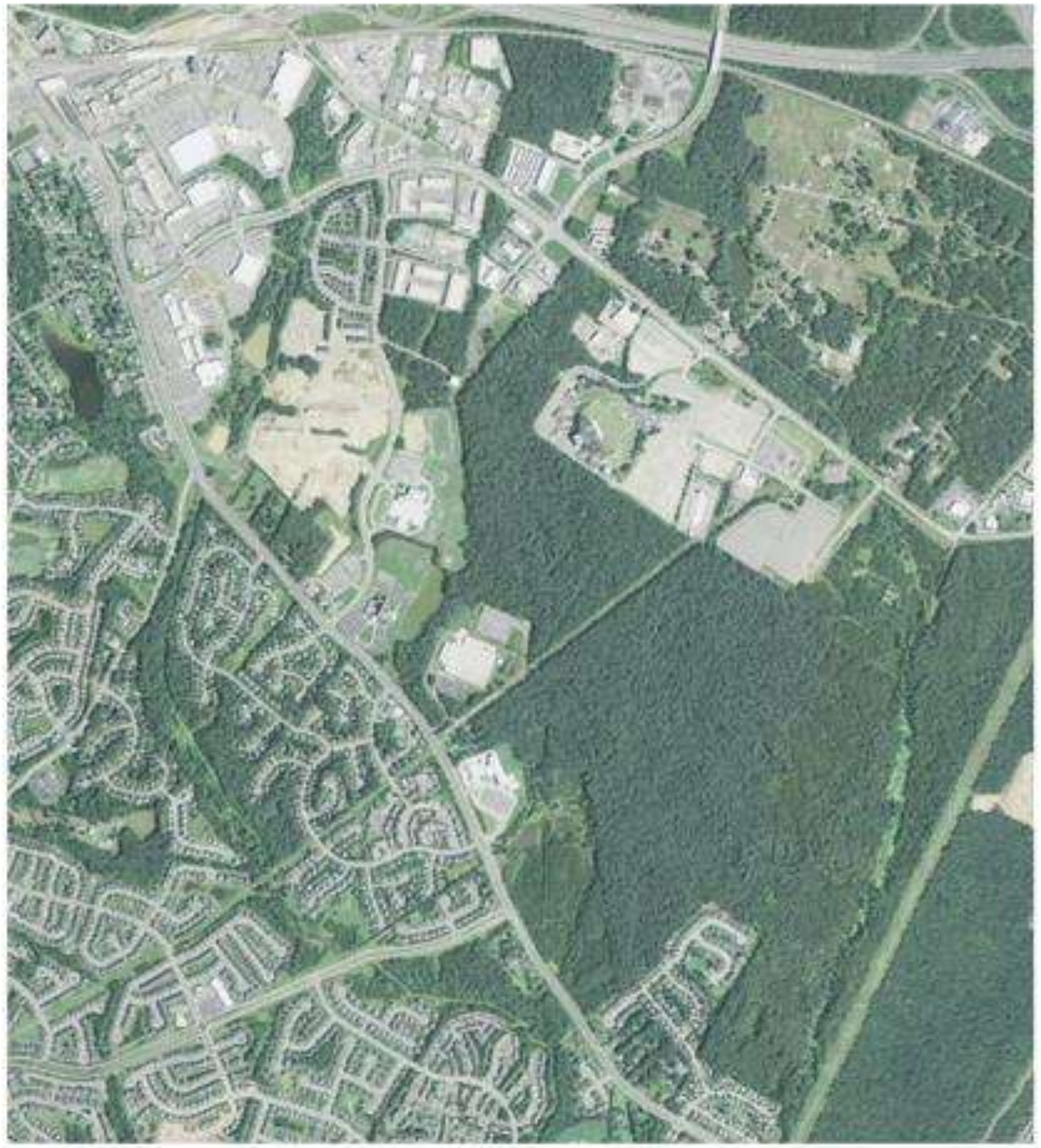
vs.

Low res at 30m

"High-resolution" matters

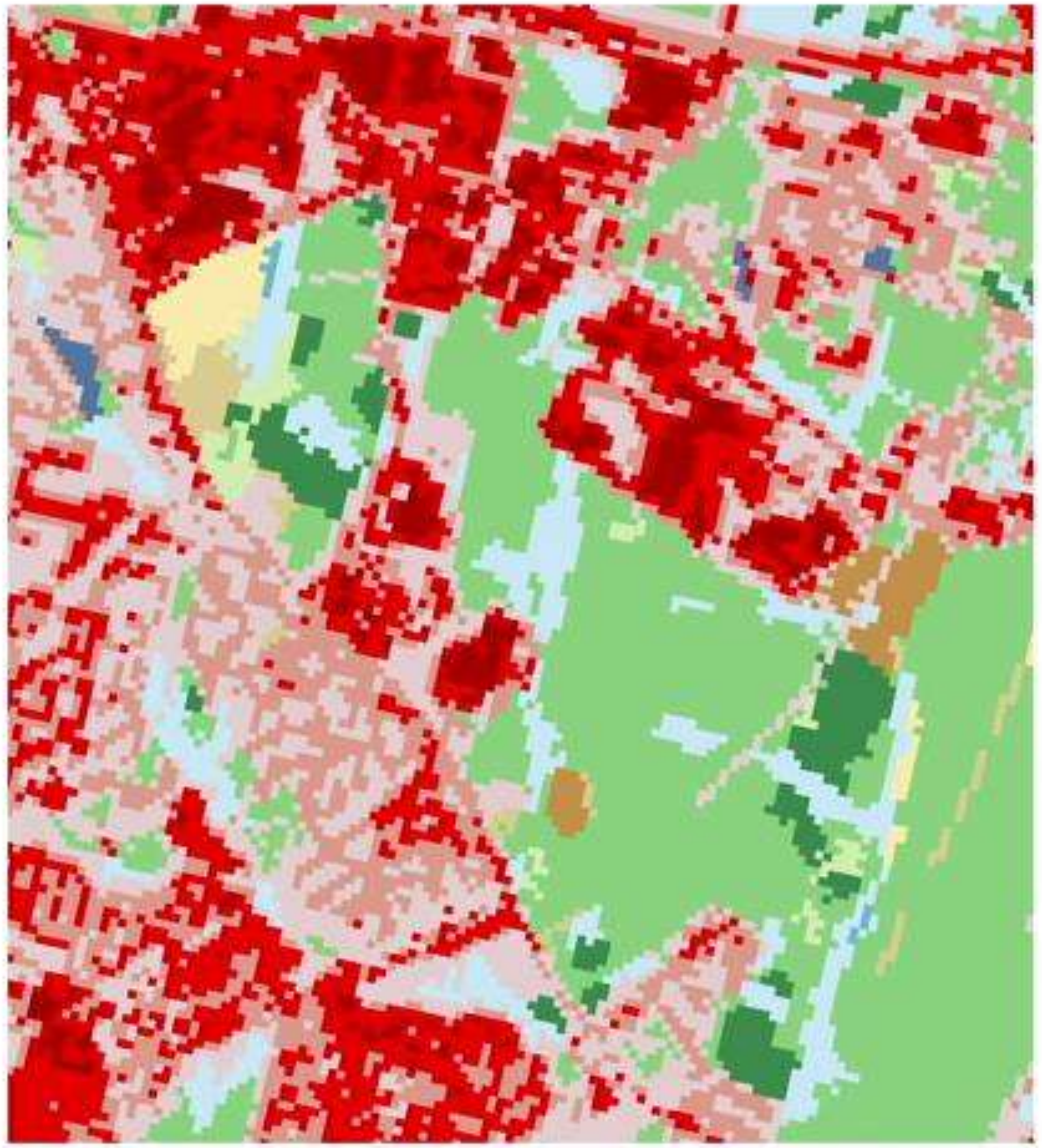
Chesapeake high res (HR) labels (1m)

National Aggriculture Imagery Program (NAIP) image (1m)



Chesapeake high res (HR) labels (1m)

National land cover database labels (30m)



Generally…

- Input images of various resolutions
  - NAIP at 1m is collected every couple of years in US
  - Landsat provides 30m images globally on a weekly bases
- Label data is sparse and usually at low resolution
  - Till this last Christmas, the latest NLCD year was 2011
- The tasks/applications are varied
  - Wetland destruction in the Gulf
  - Coffee farm mapping in South America
  - Resource mapping in the developing world
  - Population mapping in Africa
  - Disaster response

+6 TB of imagery per day...

= 300**M** km^2/day
Just from Planet's satellites [1]

Note: Planet Labs launched its first satellites in 2013; new startups plan to take images at video frame rate

[1] http://www.planet.com/products/

Projected market growth

	2018	2023
Geospatial analytics:	$41 billion	$86 billion
Cloud:	$272 billion	$623 billion
AI software:	$9.5 billion	$71 billion
Wine:	$108 billion	$450 billion

Breakup of GIS market (https://www.marketsandmarkets.com/)

By Type

- Surface & Field Analytics
- Network & Location Analytics
- Geovisualization
- Others

Breakup of GIS market

By Application

- Surveying
- Medicine & Public Safety
- Disaster Risk Reduction & Management
- Climate Change Adaptation
- Other

Breakup of GIS market

By Vertical

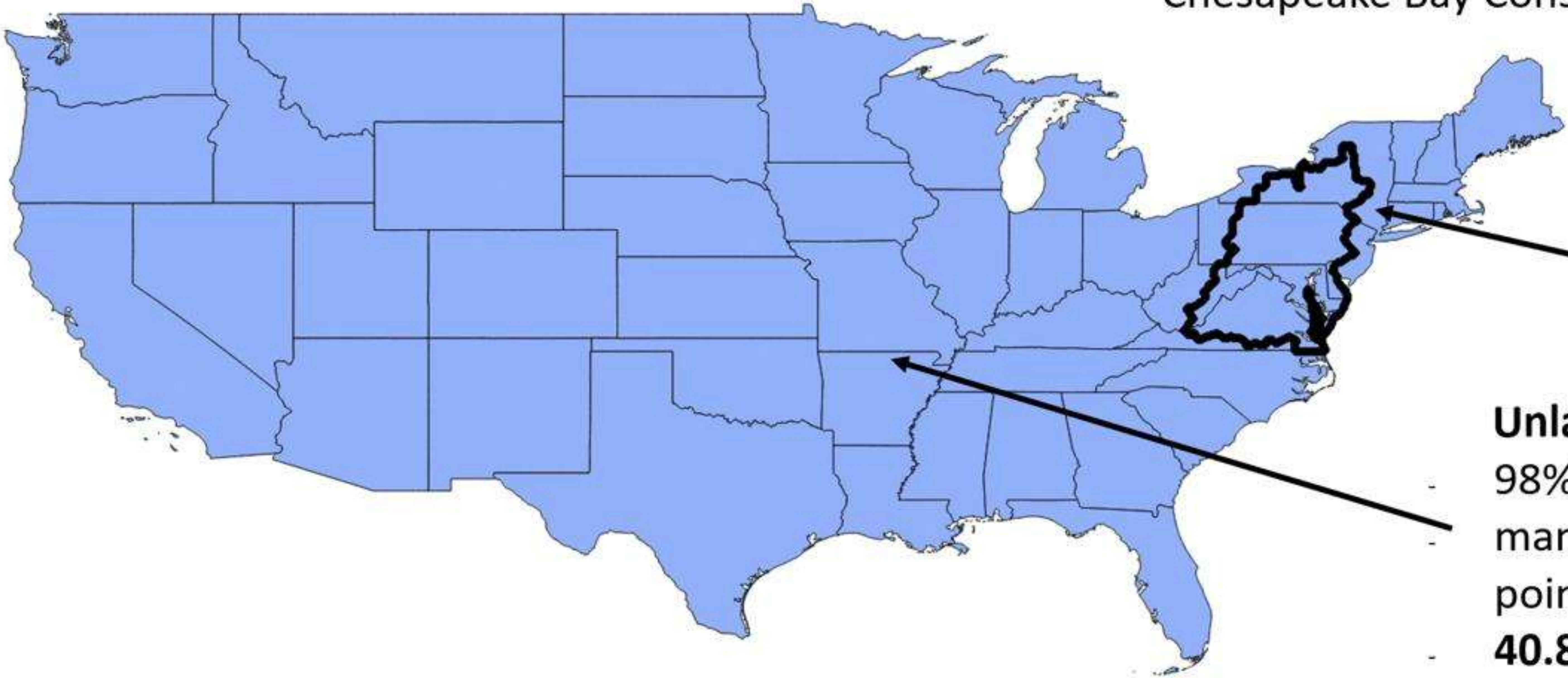
- Business
- Automotive
- Utility & Communication
- Government
- Defense & Intelligence
- Natural Resources
- Other

The bootstrap problem

- Land cover mapping is expensive
- Less well funded agencies do not always know if it is worth it

Mapping the US at 1m resolution (ICLR and CVPR 2019)

Chesapeake Bay Conservancy



Labeled
- 2% of area
- 1 time point
- 10 months
- $1.3 million

Unlabeled
- 98% of area
- many time points
- **40.8 years?**
- **$63.7 million?**

https://chesapeakeconservancy.org/conservation-innovation-center/high-resolution-data/

Domain adaptation issues…

- Rest of the US looks very different and there are no labels
- The desired labels themselves may change (e.g. wetlands)
- Can we adapt or retrain the models using some other form of guidance?

(or you can refer to this in terms of meta learning, domain transfer, learning to learn, semi-supervised learning, sample efficiency, …)

Domain adaptation issues…

- Rest of the US looks very different and there are no labels
- The desired labels themselves may change (e.g. wetlands)
- Can we adapt or retrain the models using some other form of guidance?

(or you can refer to this in terms of meta learning, domain transfer, learning to learn, semi-supervised learning, sample efficiency, …)

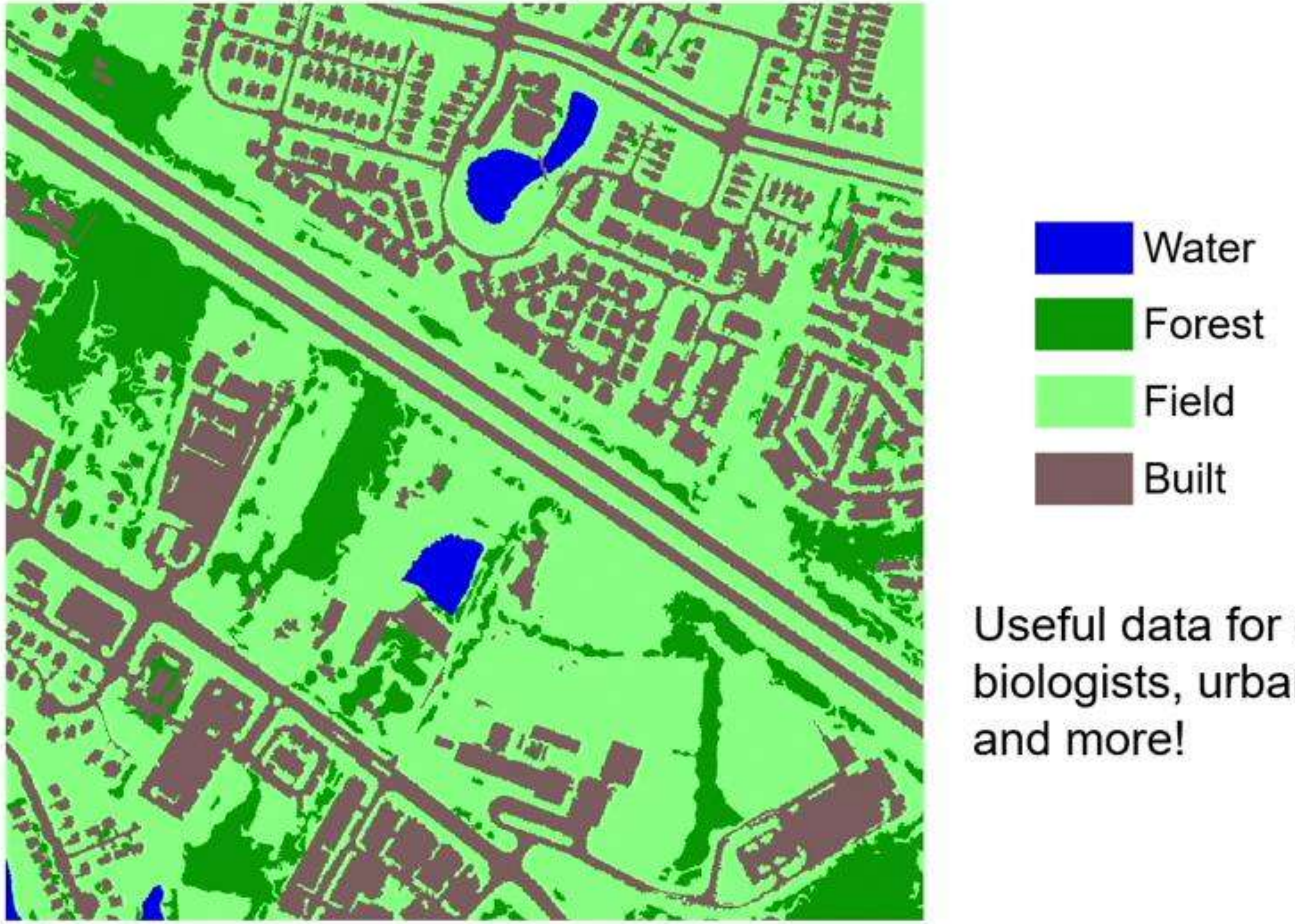
Thus, rapidly growing interest in the ML research community

- Introduction
- **Label super-resolution (ICLR 2019)**
- First US-wide 1m land cover map (CVPR 2019)
- iNaturalist species observation (just starting on this)
- Hybrid intelligence approaches to land cover mapping (under review)
- Applications/collaborations
- Open problems

1 pixel = 1 meter squared

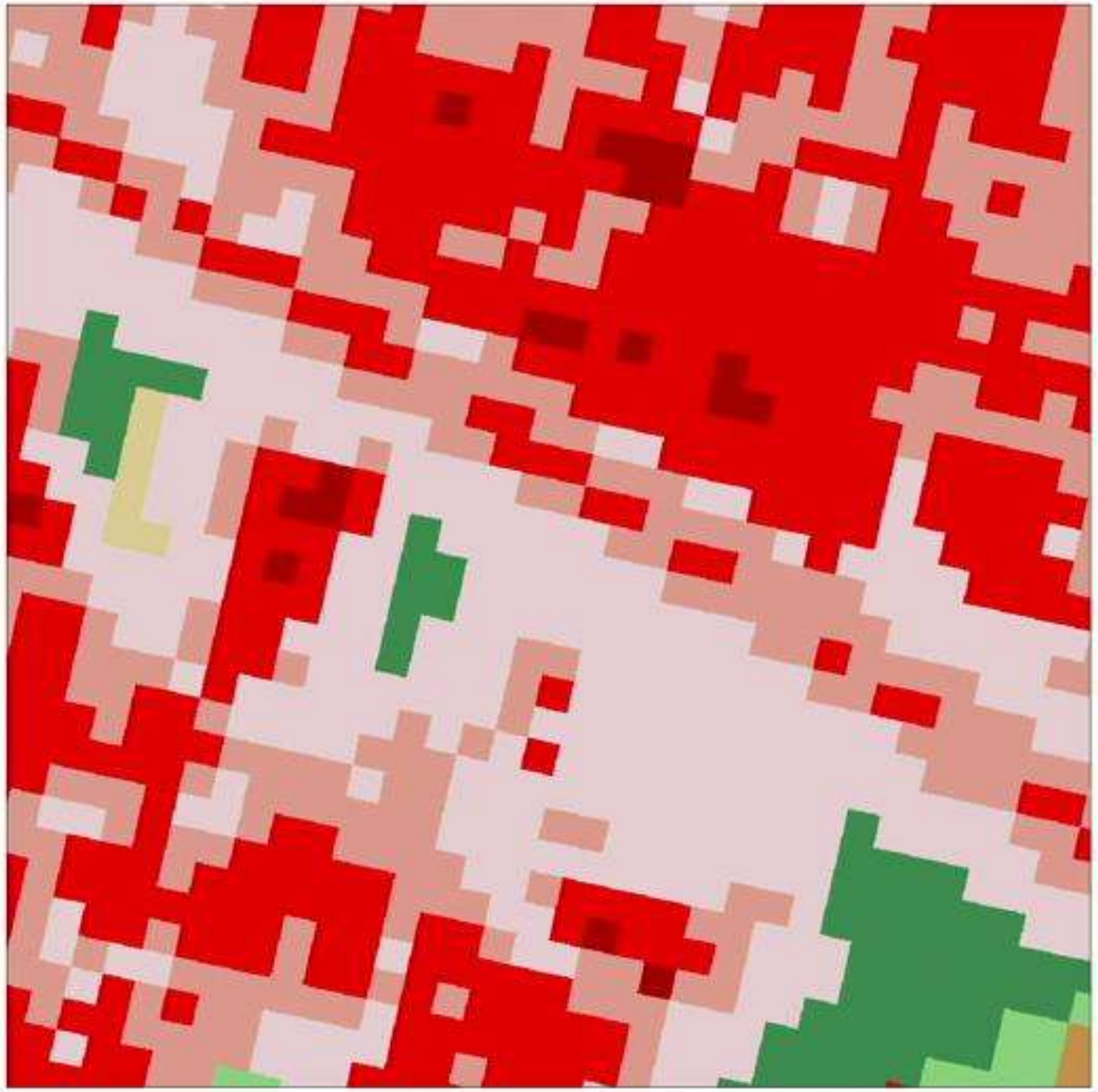
High-Resolution Satellite/Aerial Imagery

(NAIP Imagery)



Useful data for conservation biologists, urban planners, and more!

High-Resolution Land Cover Map



We have access to low-resolution labels in most of US!

Low-Resolution Land Cover

(NLCD Imagery)

Label super resolution (ICLR 2019)

(1m imagery) NAIP

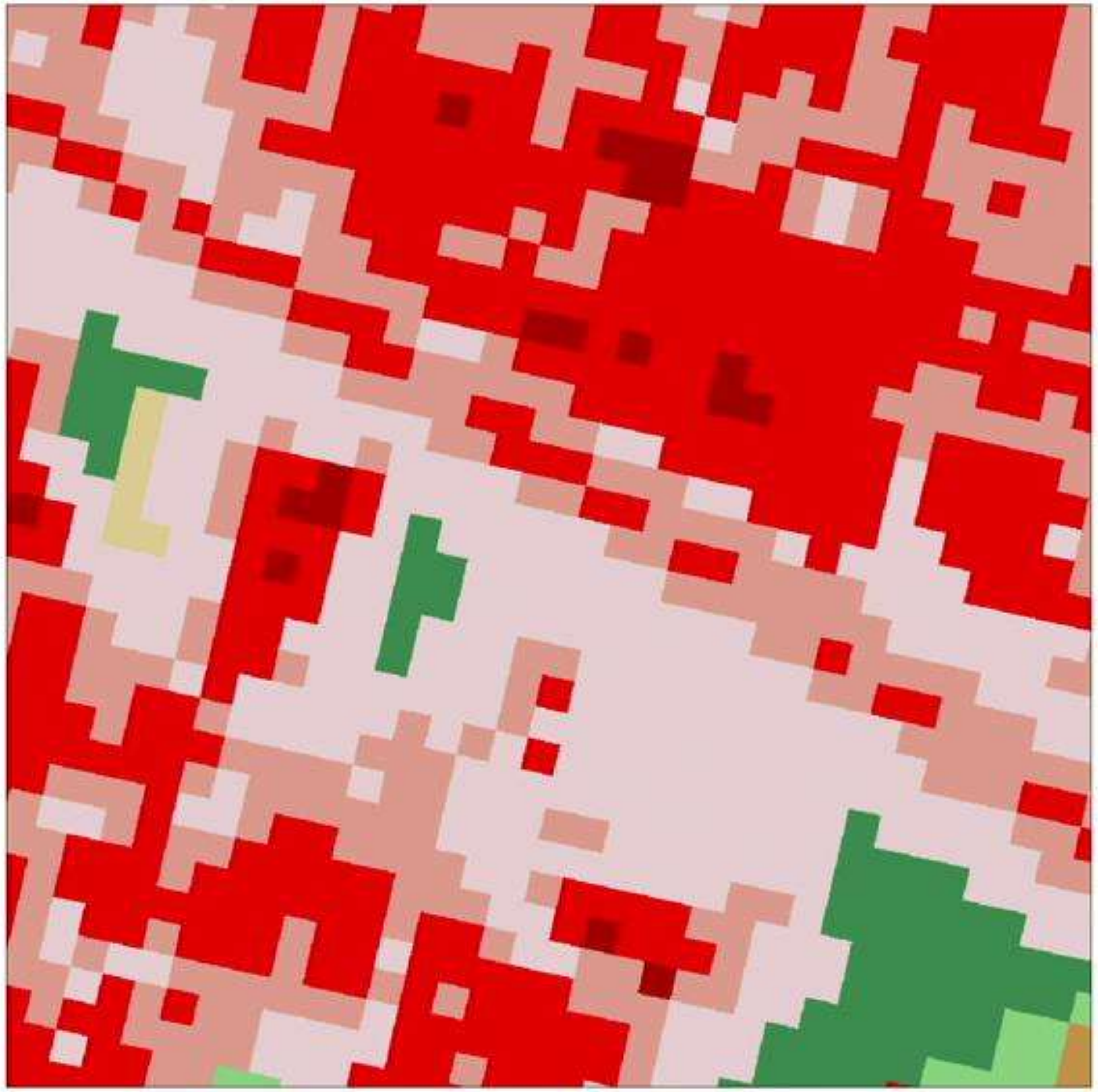
(30m land cover labels) NLCD

(1m land cover labels)

(1m imagery) NAIP

(30m land cover labels) NLCD

(1m land cover labels)



(1m imagery) NAIP

(30m land cover labels) NLCD

(1m land cover labels)

(1m imagery) NAIP

(30m land cover labels) NLCD

(1m land cover labels)

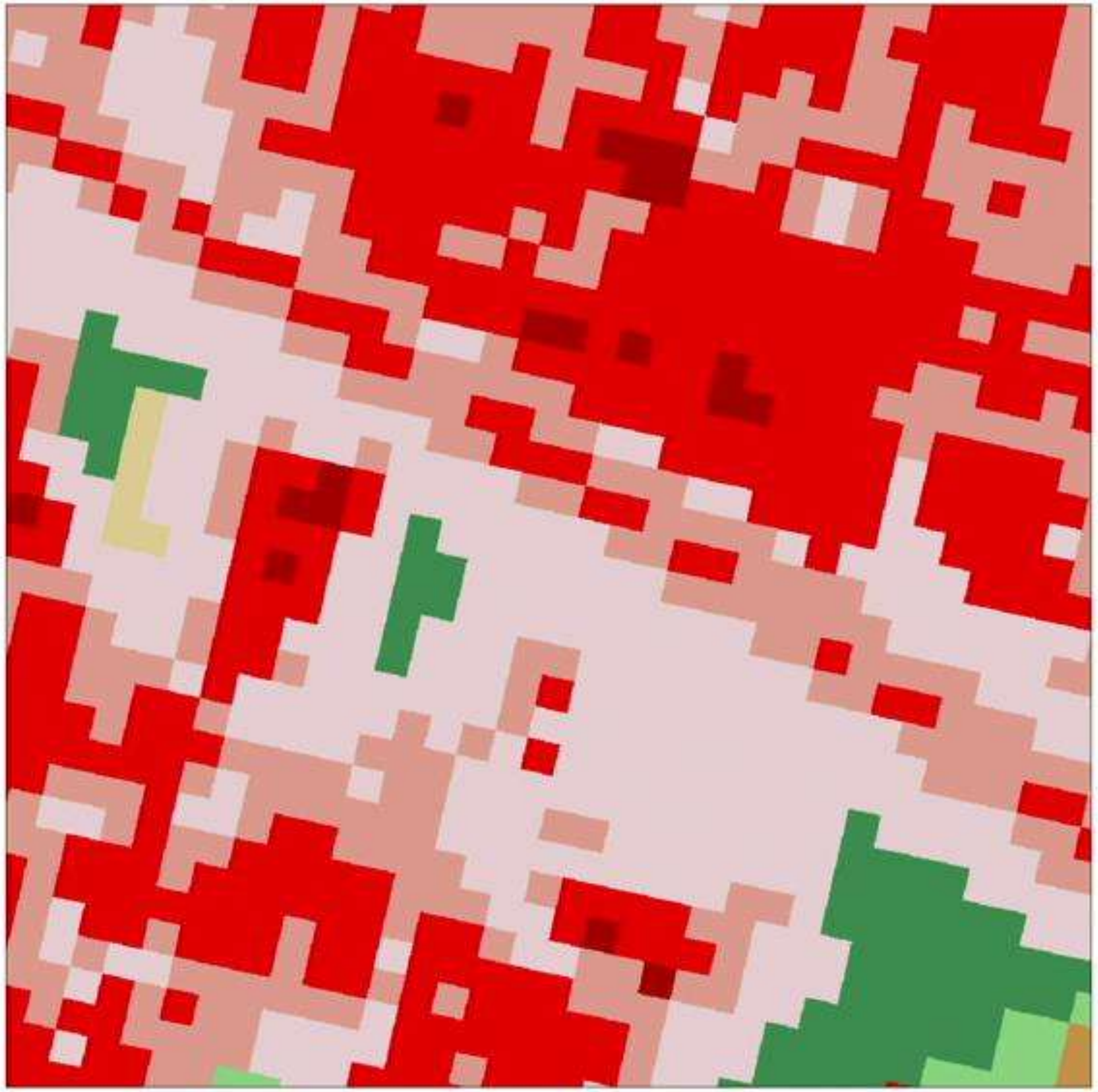
(1m imagery) NAIP

+

(30m land cover labels) NLCD

↓

(1m land cover labels)



(1m imagery) NAIP

+

(30m land cover labels) NLCD

↓

(1m land cover labels)

(1m imagery) NAIP

(30m land cover labels) NLCD

(1m land cover labels)

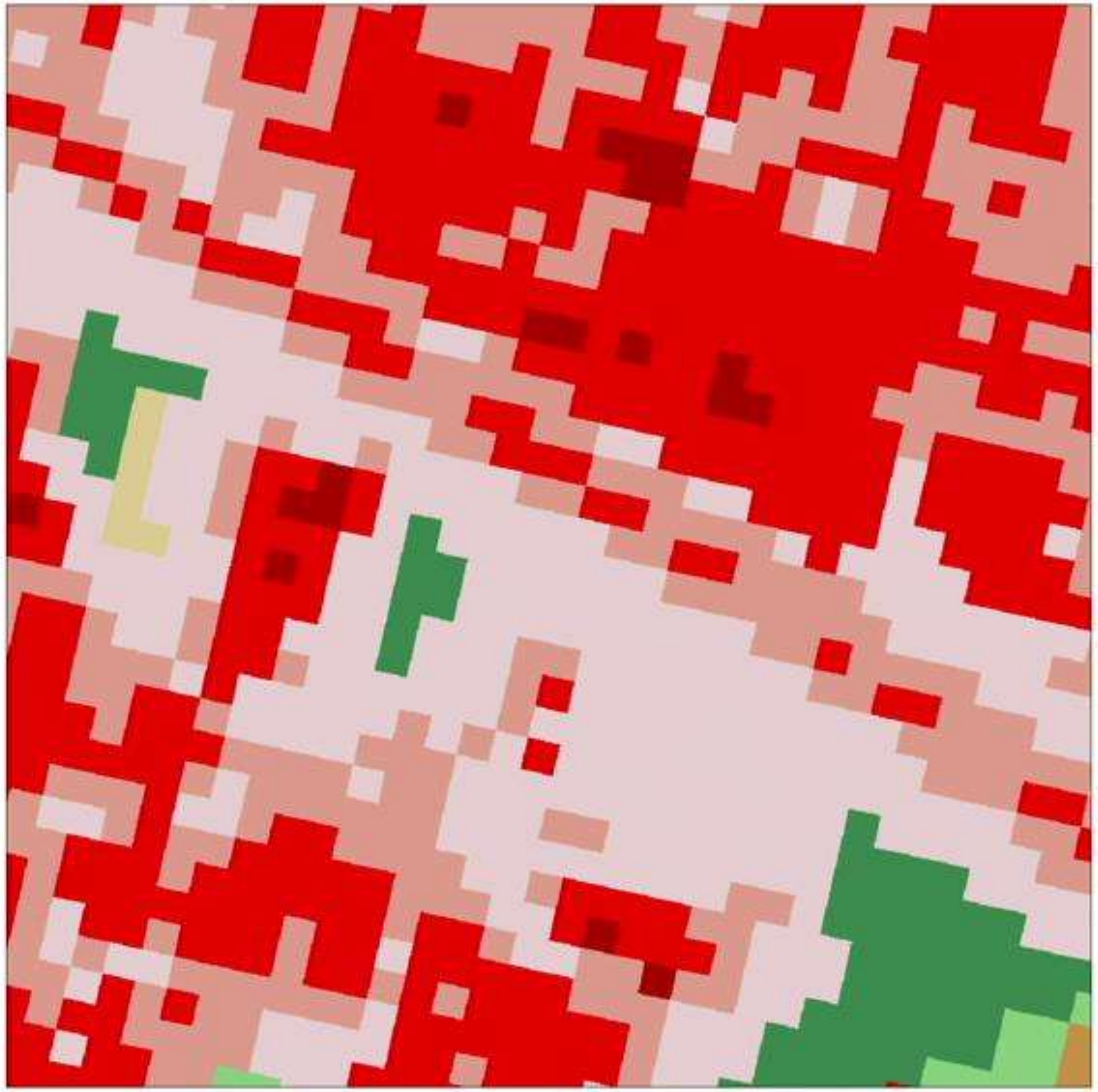
(1m imagery) NAIP

+

(30m land cover labels) NLCD

↓

(1m land cover labels)



(1m imagery) NAIP

+

(30m land cover labels) NLCD

↓

(1m land cover labels)

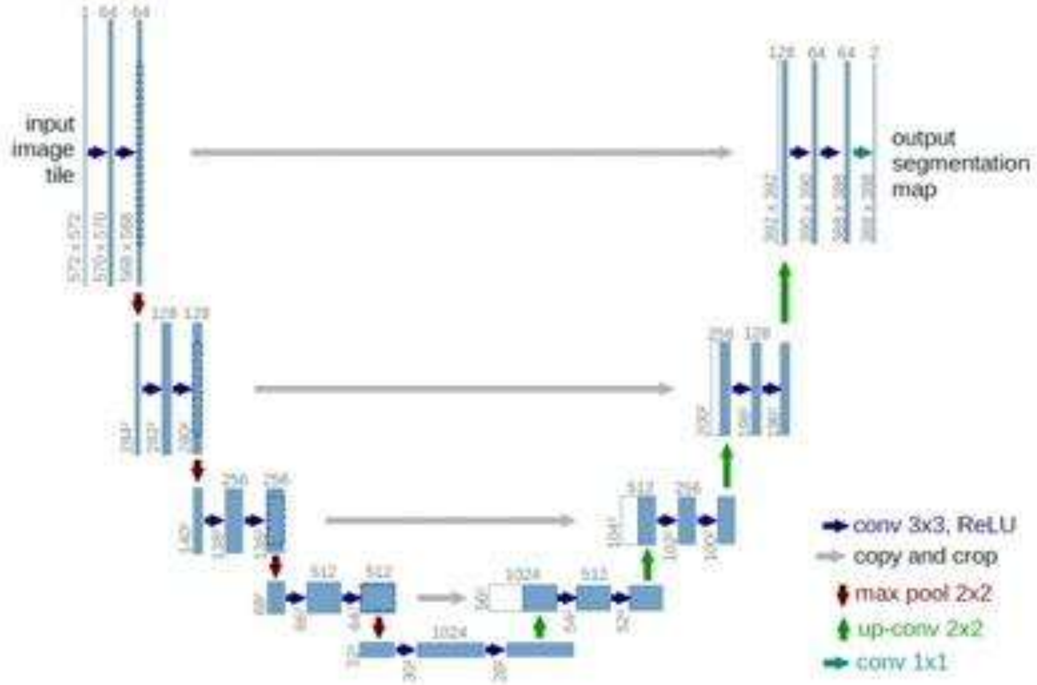
(1m imagery) NAIP

(30m land cover labels) NLCD

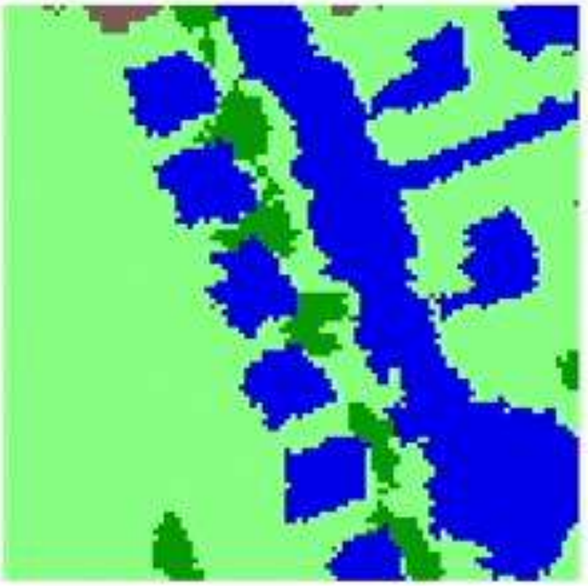
(1m land cover labels)

Setup

High-resolution input



High-resolution predictions

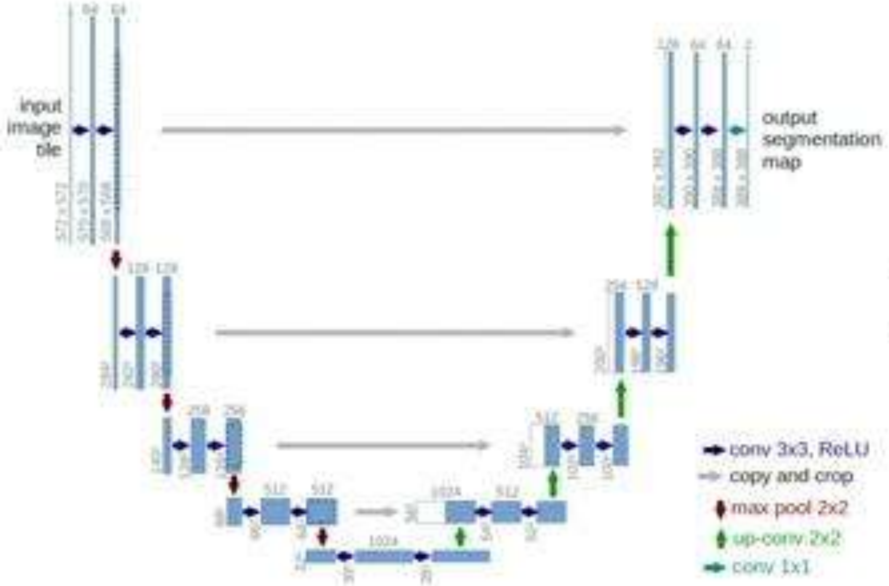


We want to train a CNN to make high-resolution (1 meter) land cover predictions using low-resolution (30 meter) labels

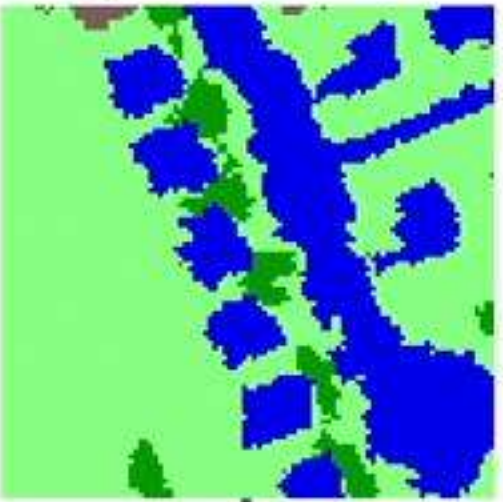
Image from: "U-net: Convolutional networks for biomedical image segmentation."

If we had high-resolution labels...

High-resolution input

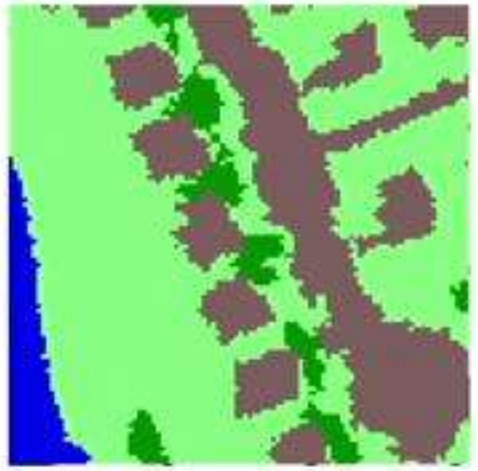


High-resolution predictions



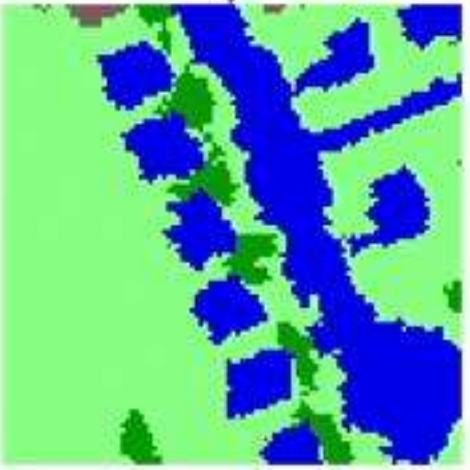
... then we could train the model with standard cross-entropy loss:

Cross Entropy
(pixel-wise comparison)



High-resolution ground truth

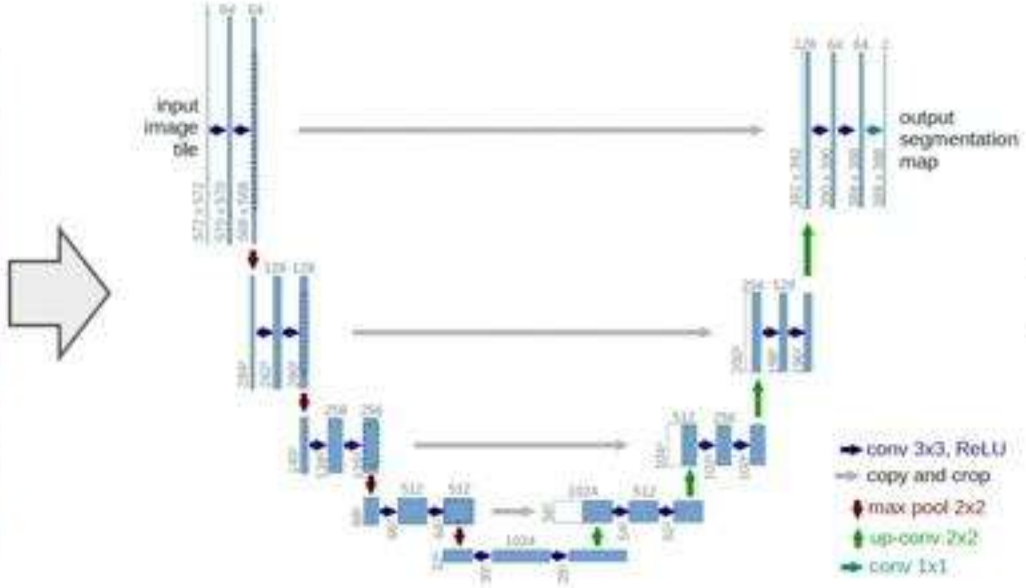
,



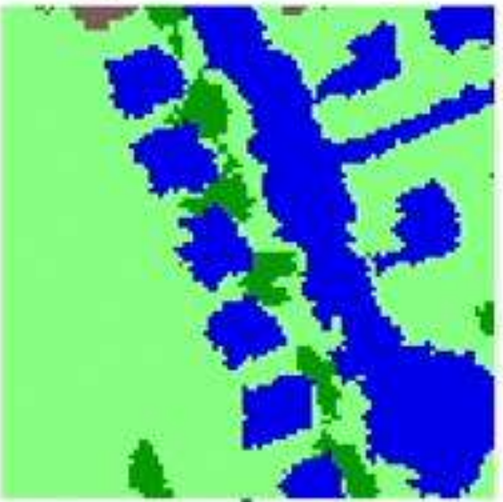
High-resolution predictions

Super-Resolution Loss

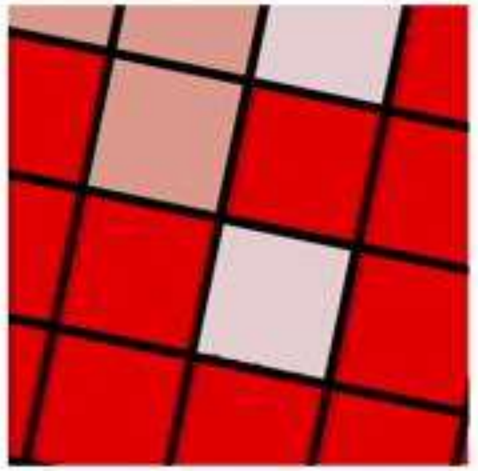
High-resolution input



High-resolution predictions

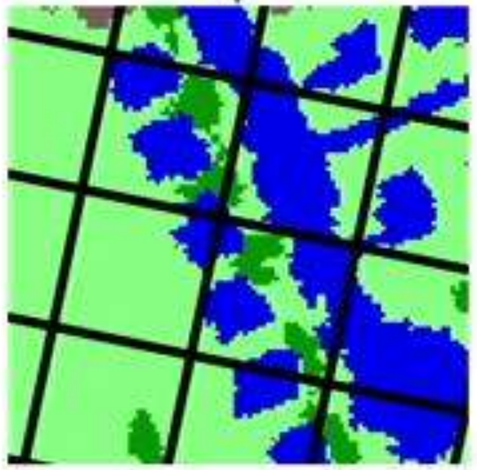


Super-Resolution Loss
(**block-wise** comparison)



Low-resolution ground truth

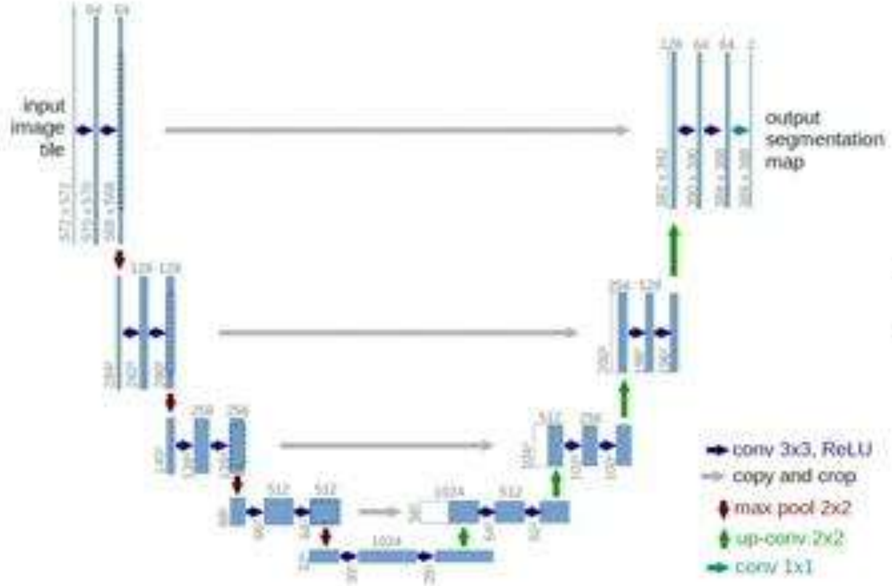
,



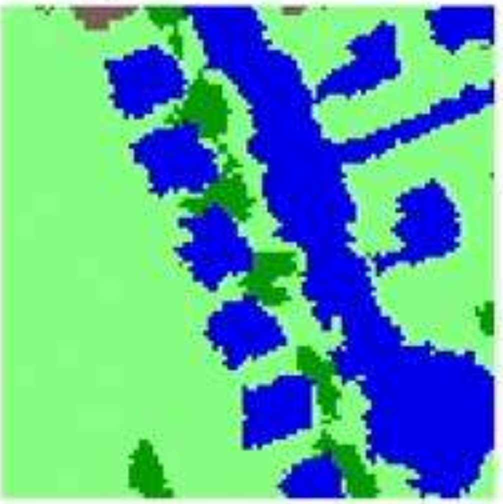
High-resolution predictions

Super-Resolution Loss

High-resolution input

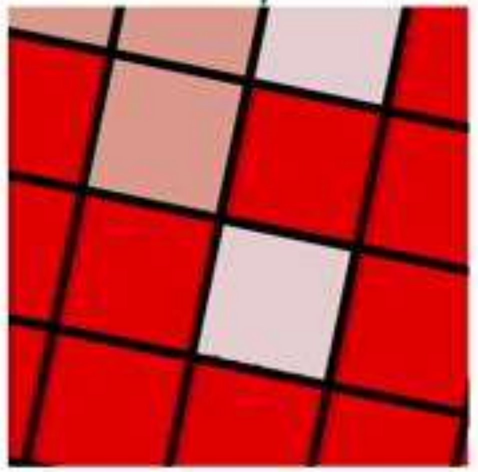


High-resolution predictions



Super-Resolution Loss
(**block-wise** comparison)

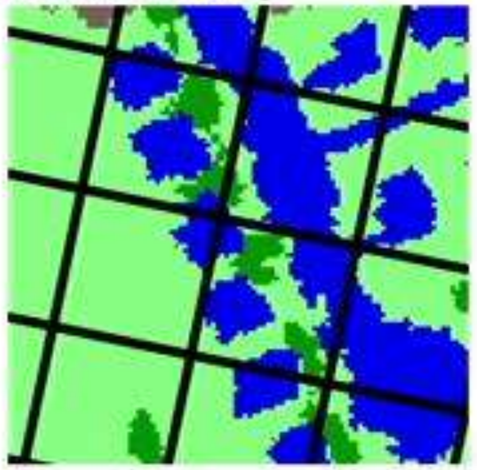
30m block - 1 label



Low-resolution ground truth

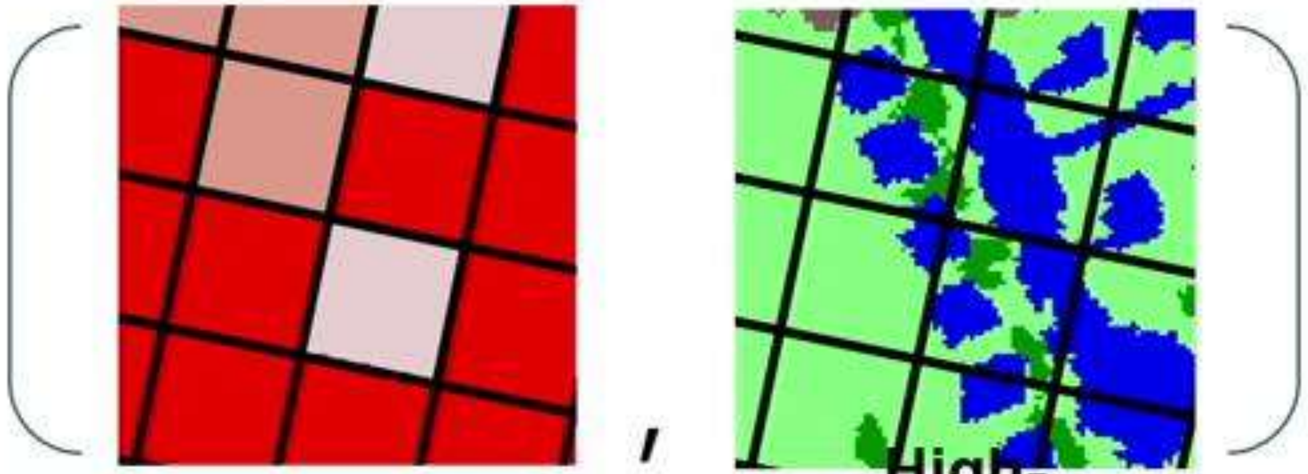
,

30m block - 900 predictions



High-resolution predictions

Super-Resolution Loss
(**block-wise** comparison)

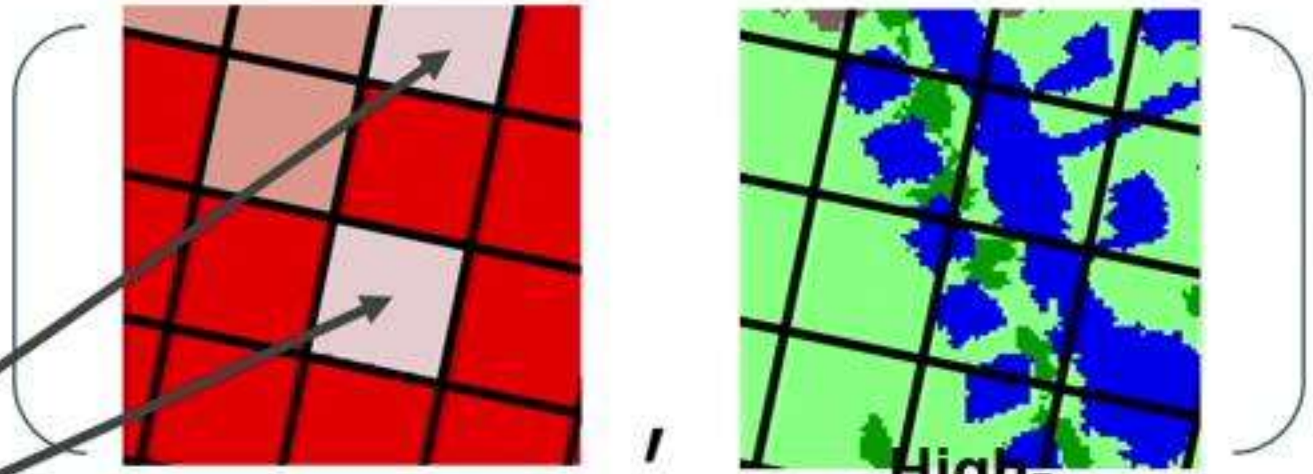


Low-resolution ground truth , High-resolution predictions

We exploit the fact that we know the **joint distribution** between *low-resolution* and *high-resolution* labels

	High resolution classes			
Low resolution classes	**Water**	**Forest**	**Field**	**Built**
Developed, Open Space	**0%**	**42%**	**46%**	**12%**
Developed, Low Intensity	**1%**	**30%**	**34%**	**35%**
Developed, Medium Intensity	**1%**	**14%**	**21%**	**64%**

Super-Resolution Loss
(**block-wise** comparison)

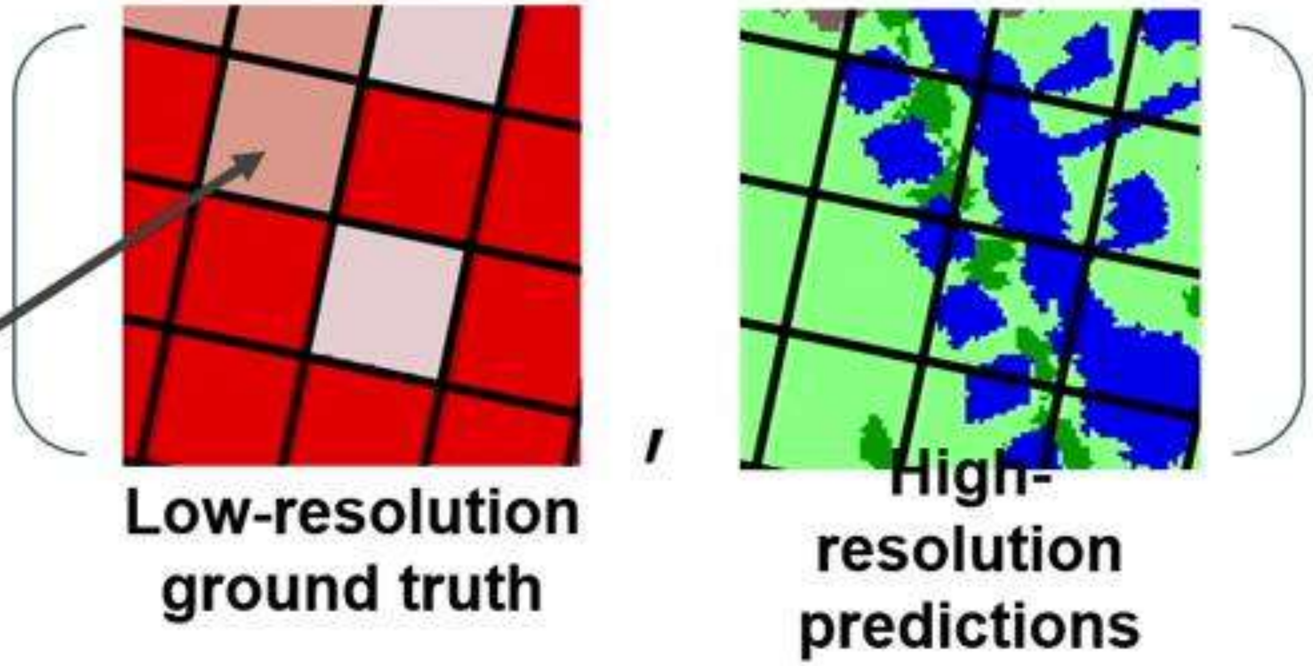


Low-resolution ground truth , **High-resolution predictions**

Developed, Open Space:
On Average - Should contain 0% water labels, 42% forest labels, … (+/- something)

	High resolution classes			
Low resolution classes	**Water**	**Forest**	**Field**	**Built**
Developed, Open Space	**0%**	**42%**	**46%**	**12%**
Developed, Low Intensity	**1%**	**30%**	**34%**	**35%**
Developed, Medium Intensity	**1%**	**14%**	**21%**	**64%**

Super-Resolution Loss
(**block-wise** comparison)

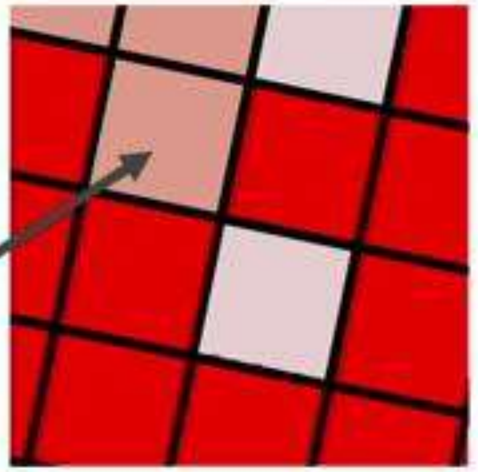


Developed, Low Intensity: On average - Should contain 1% water labels, 30% forest labels, ...

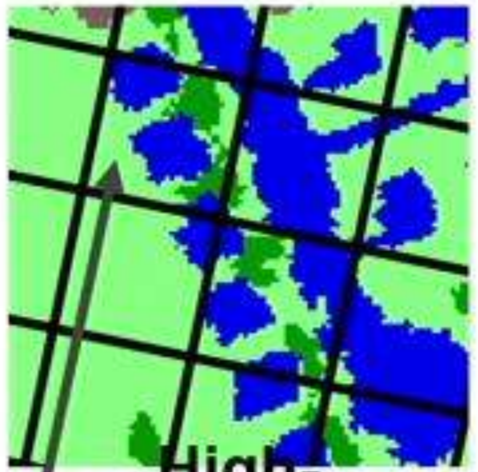
	High resolution classes			
Low resolution classes	**Water**	**Forest**	**Field**	**Built**
Developed, Open Space	**0%**	**42%**	**46%**	**12%**
Developed, Low Intensity	**1%**	**30%**	**34%**	**35%**
Developed, Medium Intensity	**1%**	**14%**	**21%**	**64%**

Super-Resolution Loss

(**block-wise** comparison)



Low-resolution ground truth



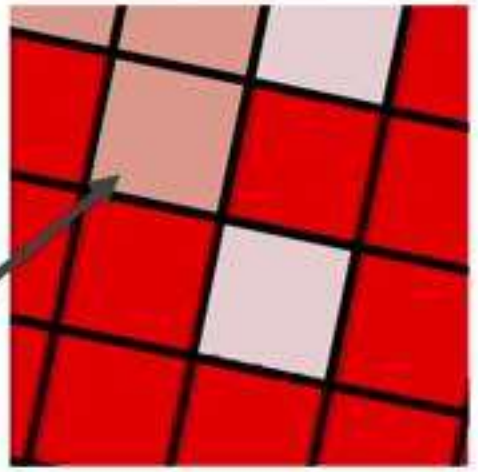
High-resolution predictions

Developed, Low Intensity: On average - Should contain 1% water labels, 30% forest labels, ...

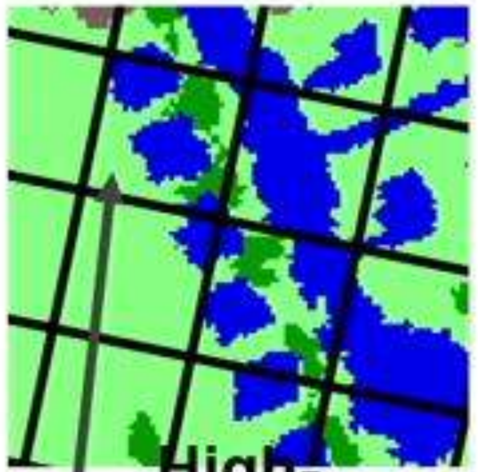
Label counting: Count predicted class labels to get a similar distribution.

E.g. but we predicted 50% water labels, 20% forest labels, ...

Super-Resolution Loss
(**block-wise** comparison)



Low-resolution ground truth



High-resolution predictions

1% water, 30% forest, 34% field, 35% built

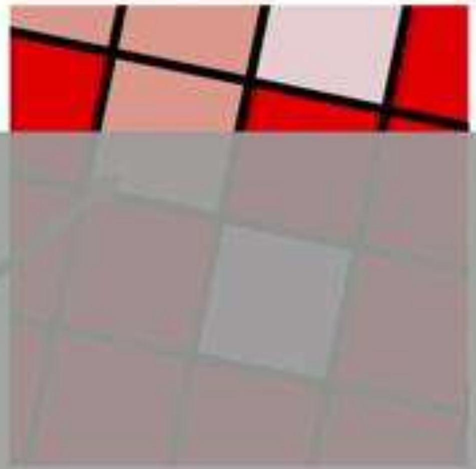
Ground Truth:
<u>Expected distribution</u>

50% water, 10% forest, 40% field, 0% built

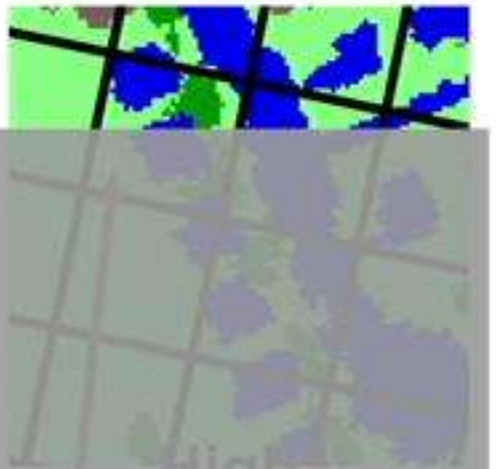
<u>Predicted distribution</u>

Compare with a differentiable distribution based measure

Super-Resolution Loss
(**block-wise** comparison)



Low-resolution



High-resolution predictions

- **Predictions at 1m**
- **Summarization at 30m**
- **Compare to ground truth at 30m**

1% water, 30% forest, 34% field, 35% built

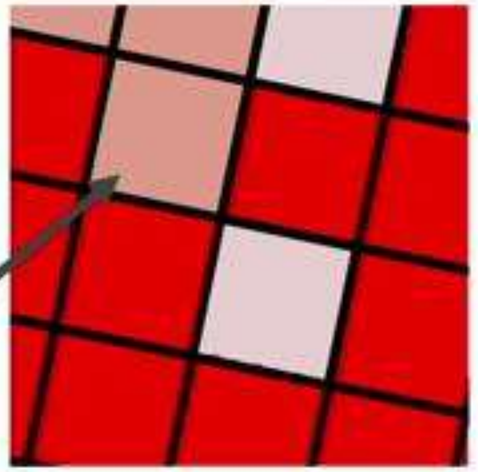
Ground Truth: Expected distribution

50% water, 10% forest, 40% field, 0% built

Predicted distribution

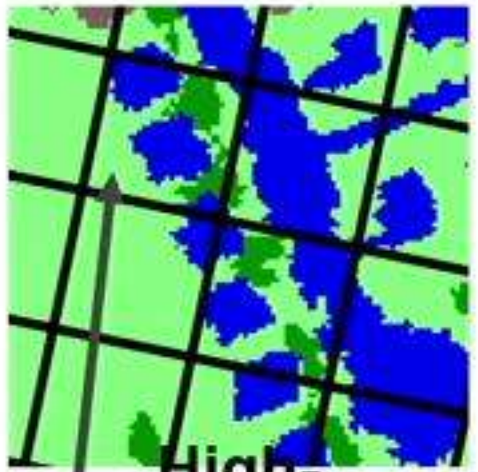
Compare with a differentiable distribution based measure

Super-Resolution Loss
(**block-wise** comparison)



Low-resolution ground truth

,



High-resolution predictions

1% water, 30% forest, 34% field, 35% built

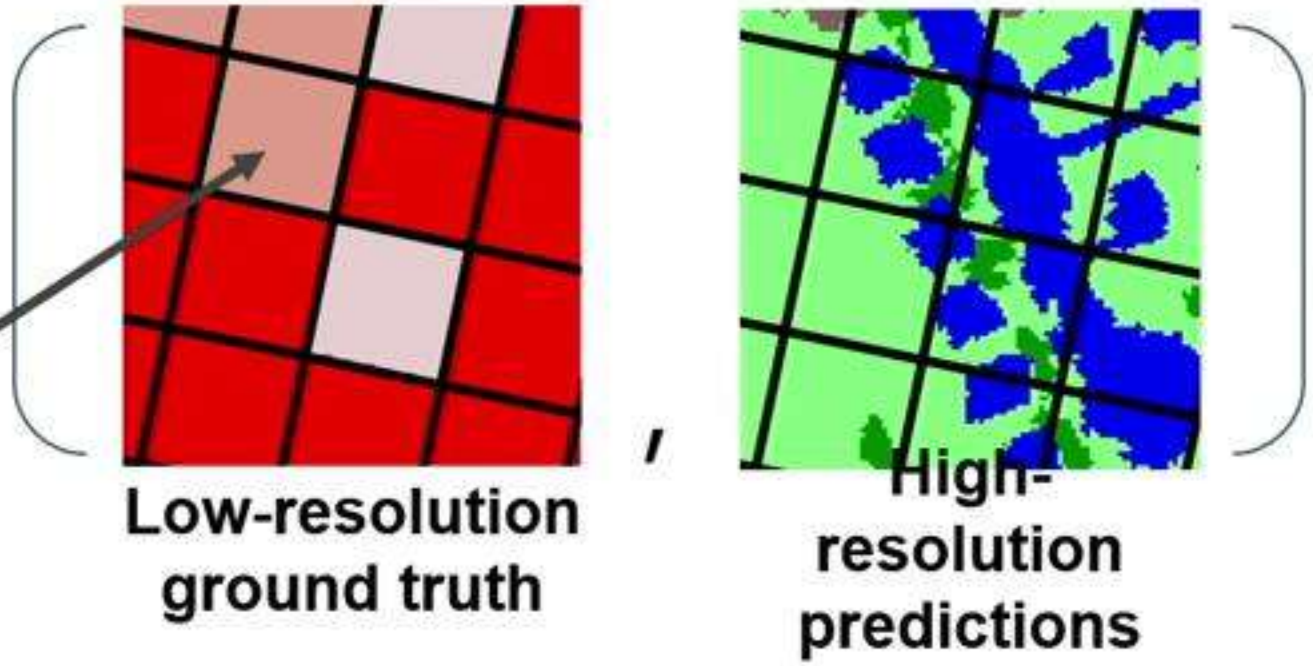
Ground Truth:
Expected distribution

50% water, 10% forest, 40% field, 0% built

Predicted distribution

Compare with a differentiable distribution based measure

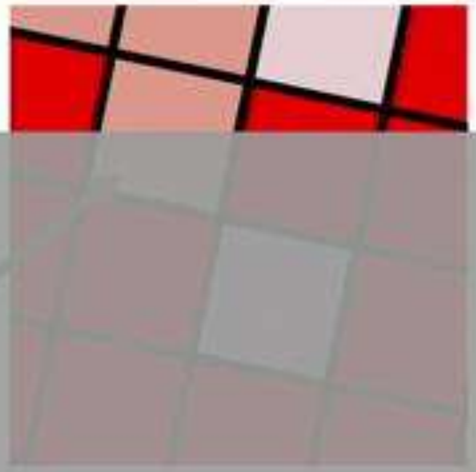
Super-Resolution Loss
(**block-wise** comparison)



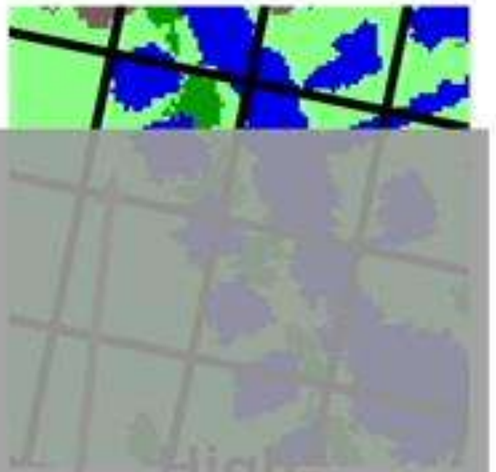
Developed, Low Intensity: On average - Should contain 1% water labels, 30% forest labels, …

	High resolution classes			
Low resolution classes	**Water**	**Forest**	**Field**	**Built**
Developed, Open Space	**0%**	**42%**	**46%**	**12%**
Developed, Low Intensity	**1%**	**30%**	**34%**	**35%**
Developed, Medium Intensity	**1%**	**14%**	**21%**	**64%**

Super-Resolution Loss
(**block-wise** comparison)



Low-resolution



High-resolution predictions

- **Predictions at 1m**
- **Summarization at 30m**
- **Compare to ground truth at 30m**

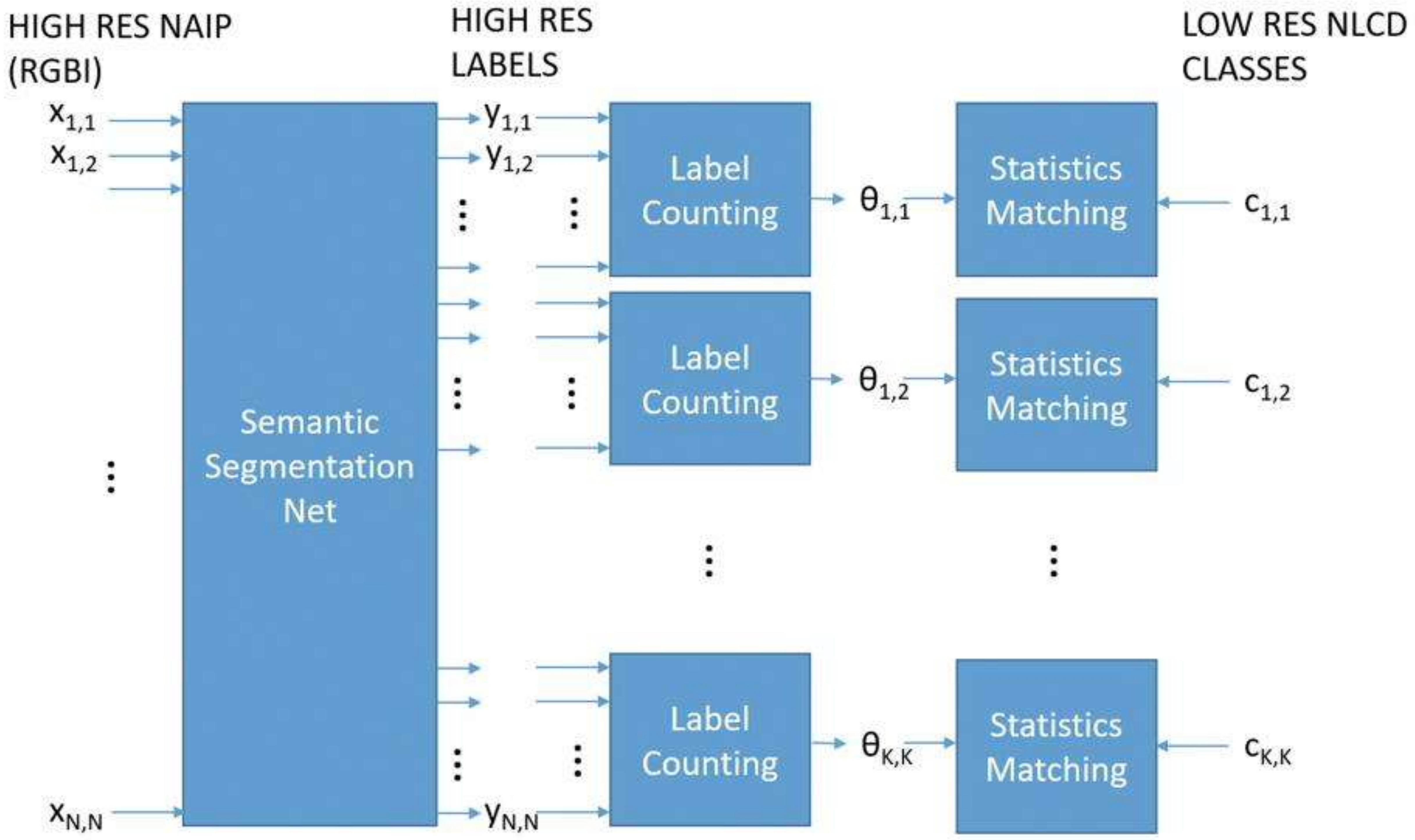
1% water, 30% forest, 34% field, 35% built

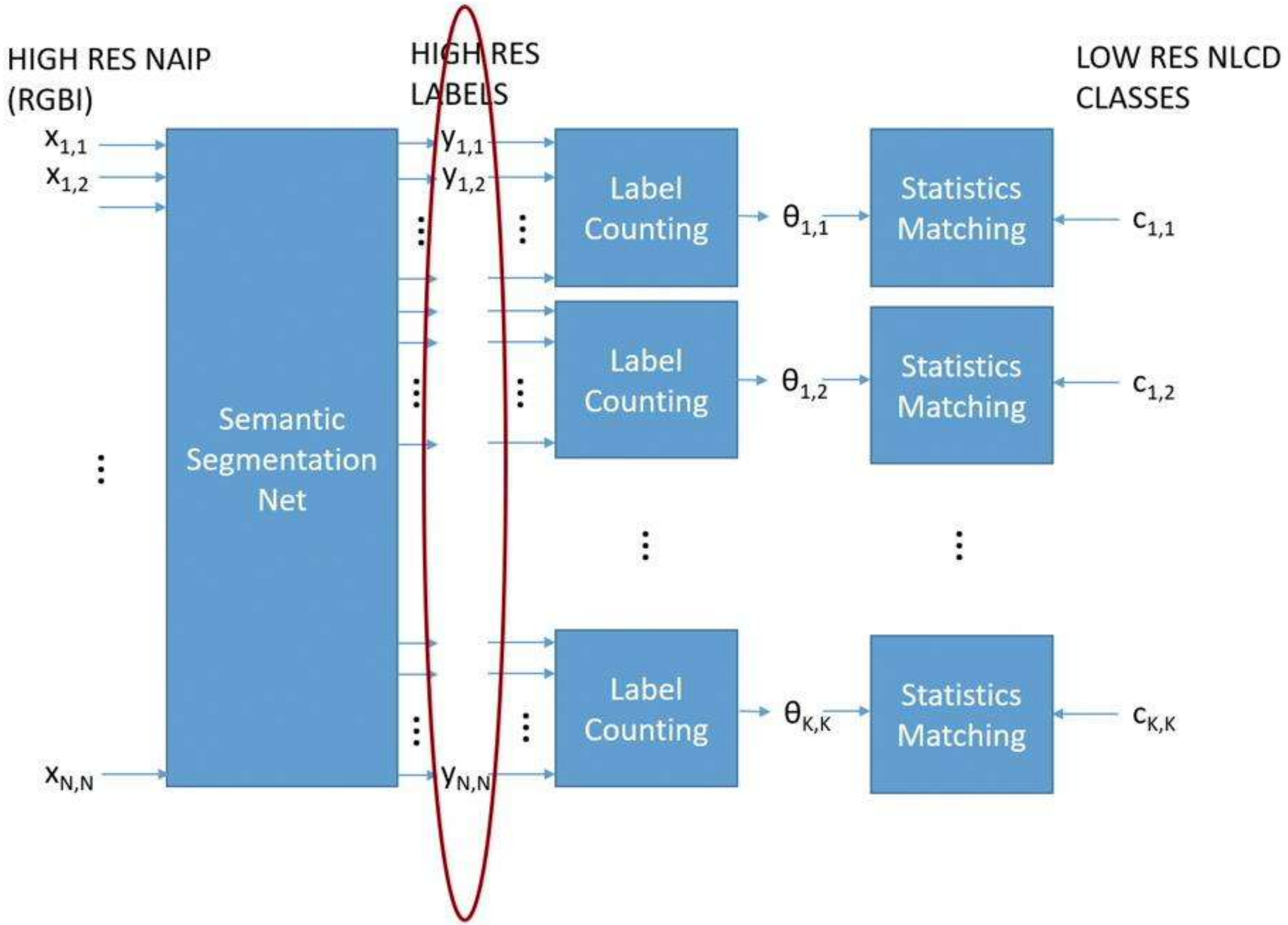
Ground Truth:
Expected distribution

50% water, 10% forest, 40% field, 0% built

Predicted distribution

Compare with a differentiable distribution based measure


HIGH RES NAIP (RGBI)
$x_{1,1}$
$x_{1,2}$
$x_{N,N}$
Semantic Segmentation Net
HIGH RES LABELS
$y_{1,1}$
$y_{1,2}$
$y_{N,N}$
Label Counting
Label Counting
Label Counting
$\theta_{1,1}$
$\theta_{1,2}$
$\theta_{K,K}$
Statistics Matching
Statistics Matching
Statistics Matching
LOW RES NLCD CLASSES
$c_{1,1}$
$c_{1,2}$
$c_{K,K}$

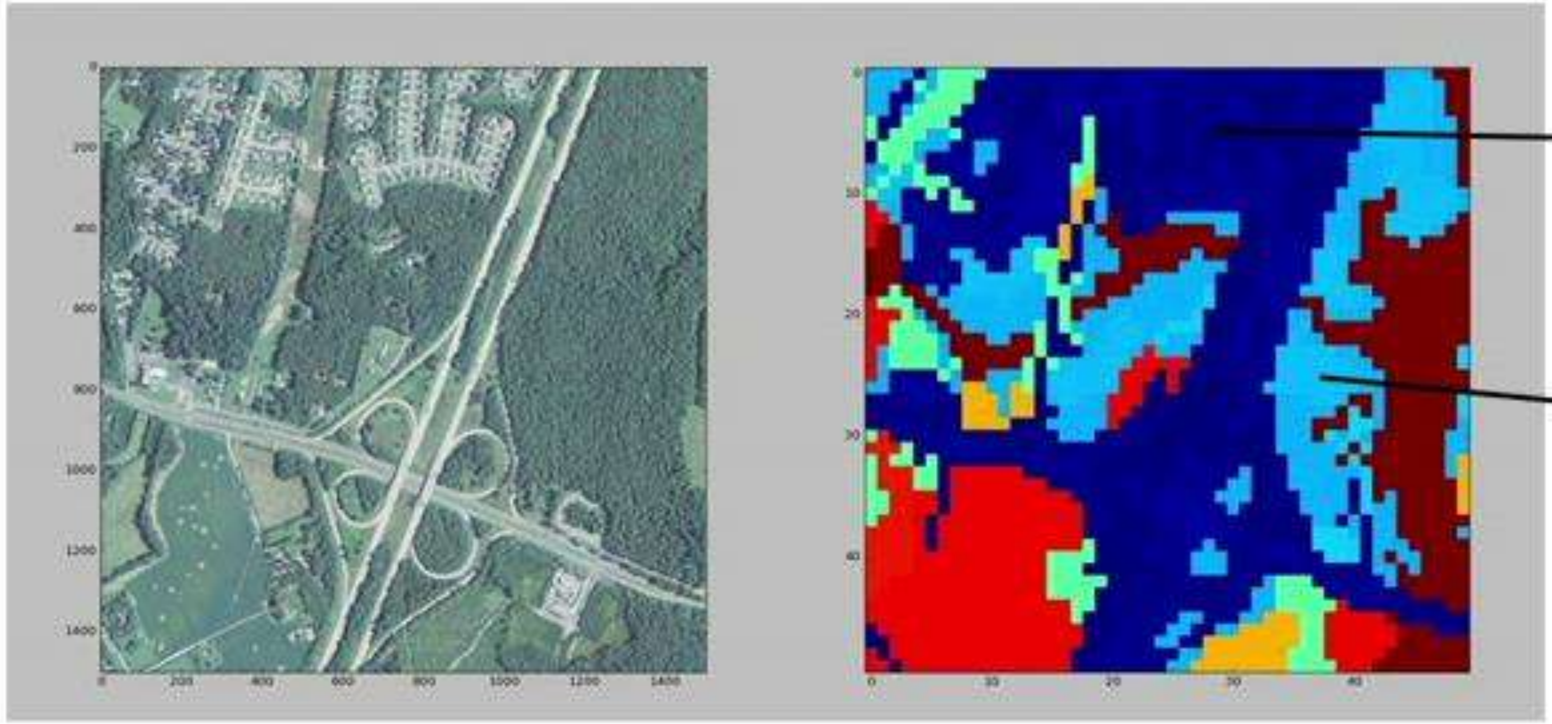


HIGH RES NAIP
(RGBI)
HIGH RES
LABELS
LOW RES NLCD
CLASSES
$x_{1,1}$
$x_{1,2}$
$x_{N,N}$
Semantic
Segmentation
Net
$y_{1,1}$
$y_{1,2}$
$y_{N,N}$
Label
Counting
Label
Counting
Label
Counting
$\theta_{1,1}$
$\theta_{1,2}$
$\theta_{K,K}$
Statistics
Matching
Statistics
Matching
Statistics
Matching
$c_{1,1}$
$c_{1,2}$
$c_{K,K}$

Label super resolution

- Given ONLY pairs

x c



National Land Cover Database labels:

22 Developed, low intensity:

Impervious: 20% - 49%

Fields, trees

41 Deciduous forest:

Trees: >20%

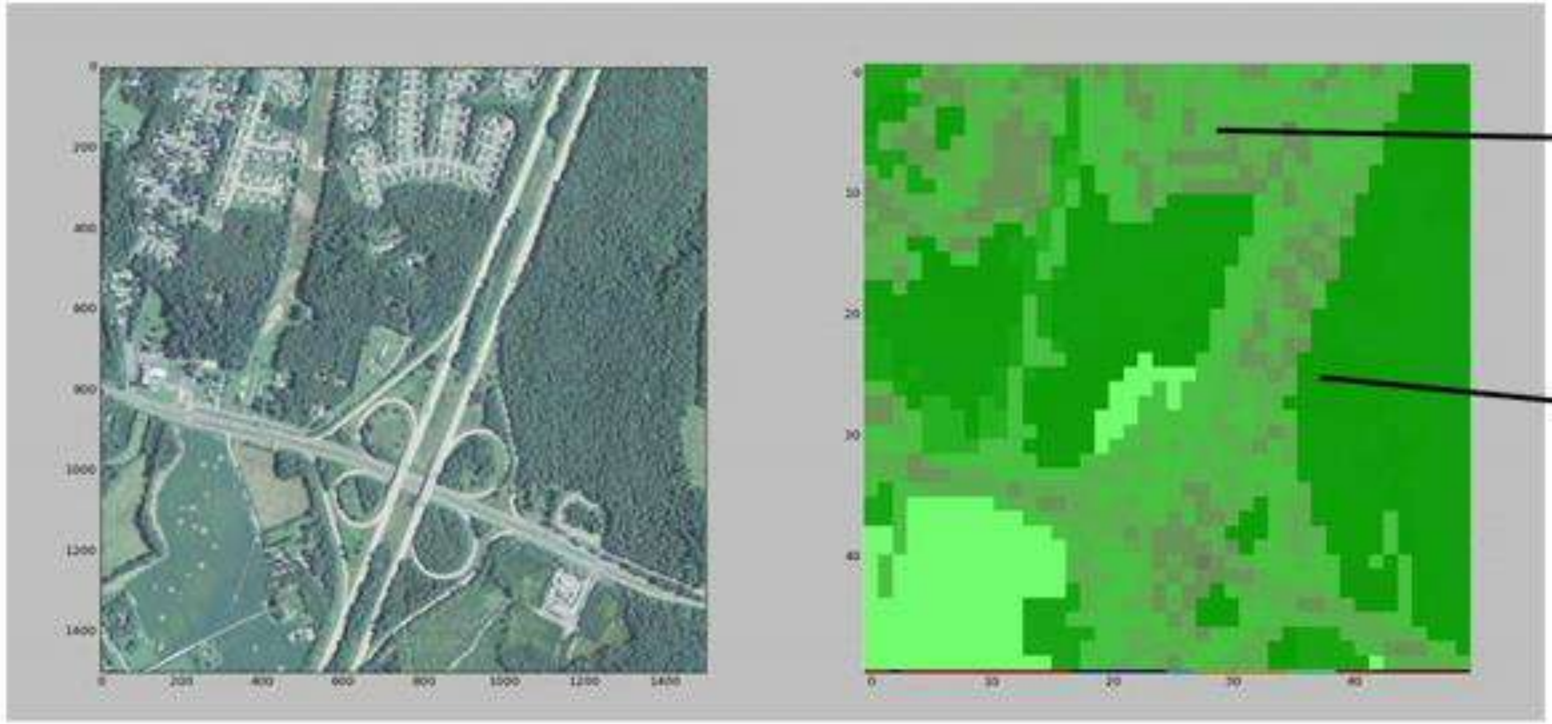
Fields make the rest

etc.

Label super resolution

- Given ONLY pairs

x c



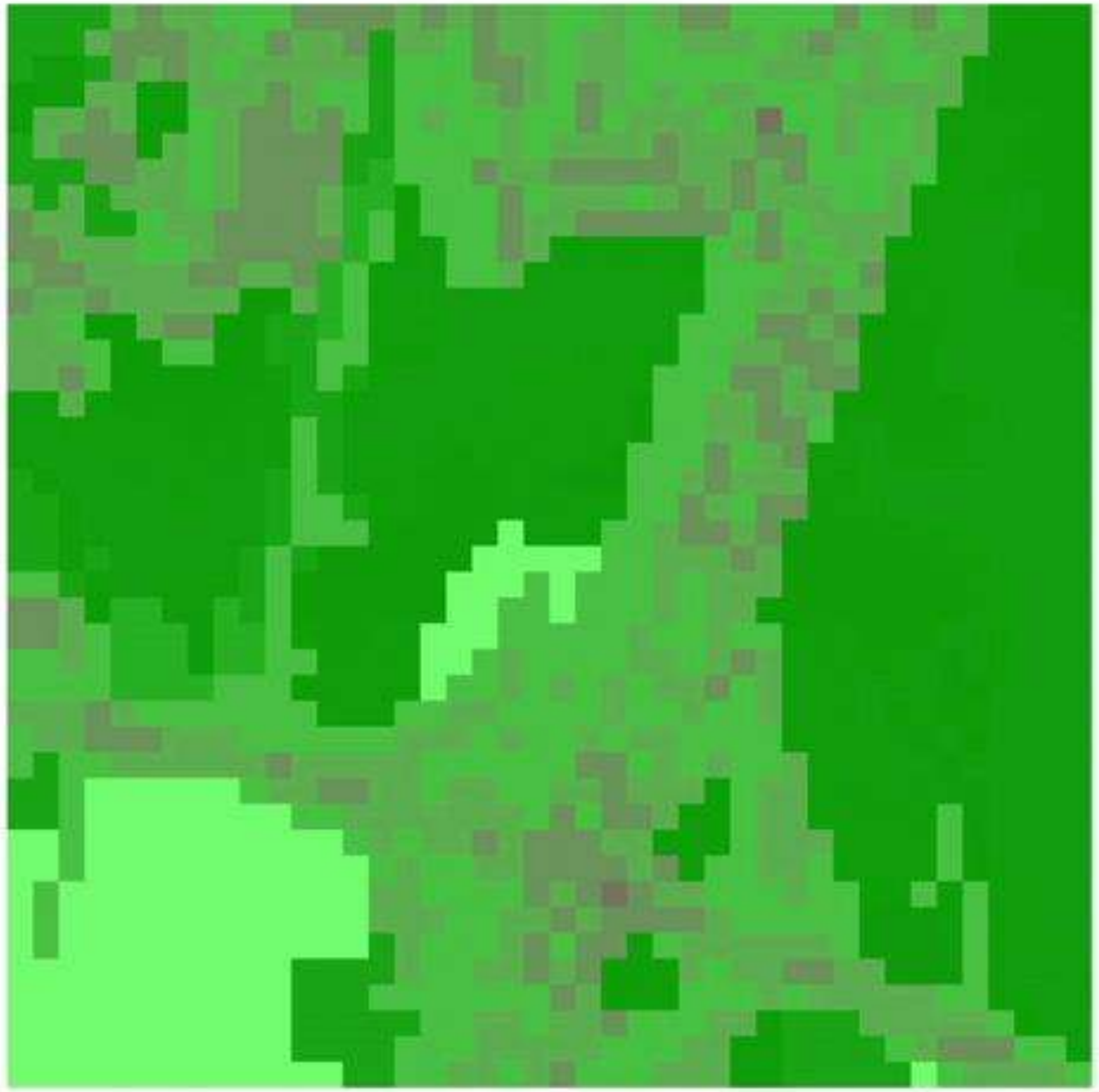
National Land Cover Database labels:
22 Developed, low intensity:
  Impervious: 20% - 49%
  Fields, trees

41 Deciduous forest:
  Trees: >20%
  Fields make the rest

etc.

Label super resolution: Infer 1m structure!

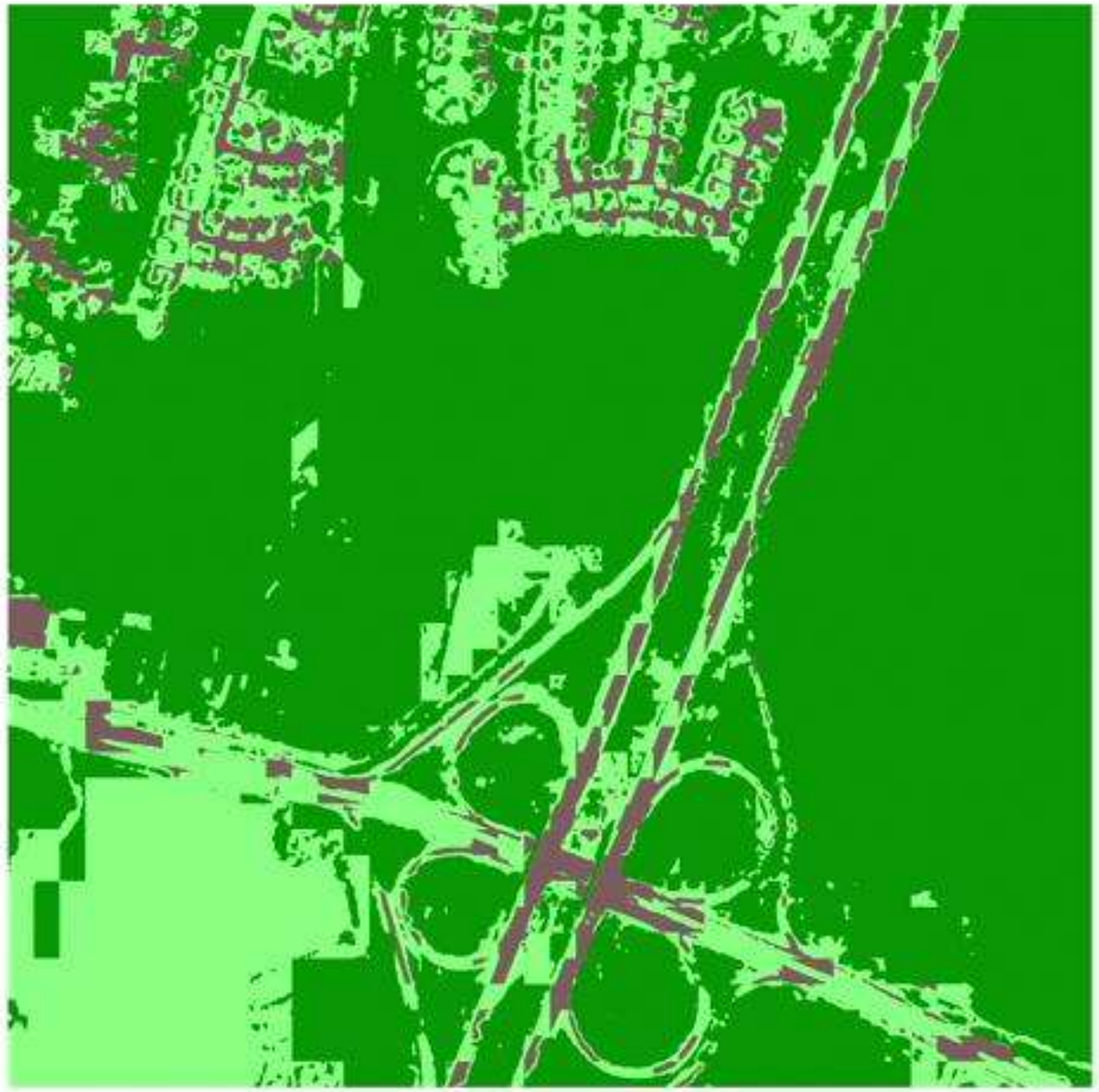
Low res NLCD labels



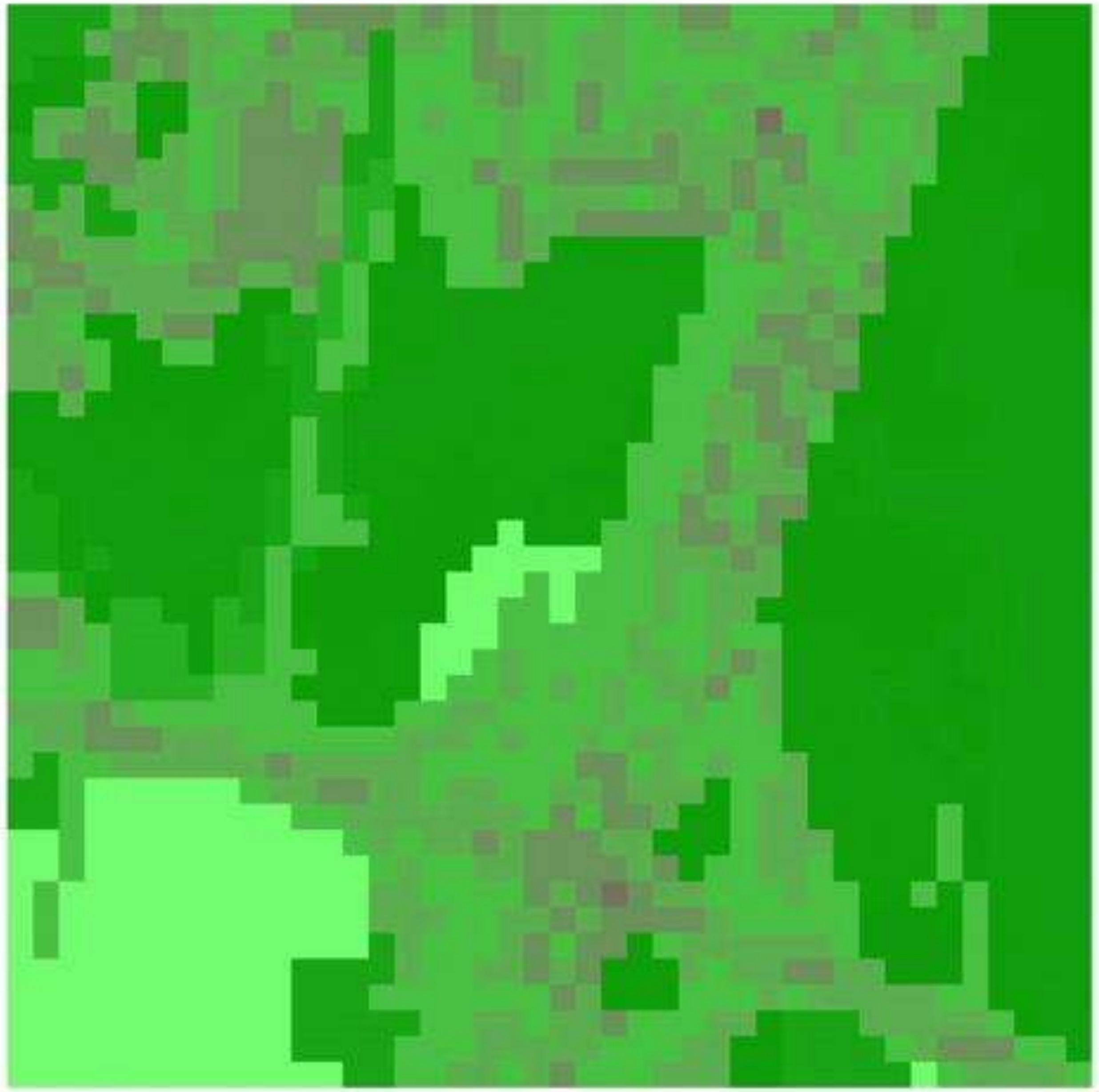
Iterations of gradient descent: Soft map

Iterations of gradient descent: Hard map

Iterations of gradient descent

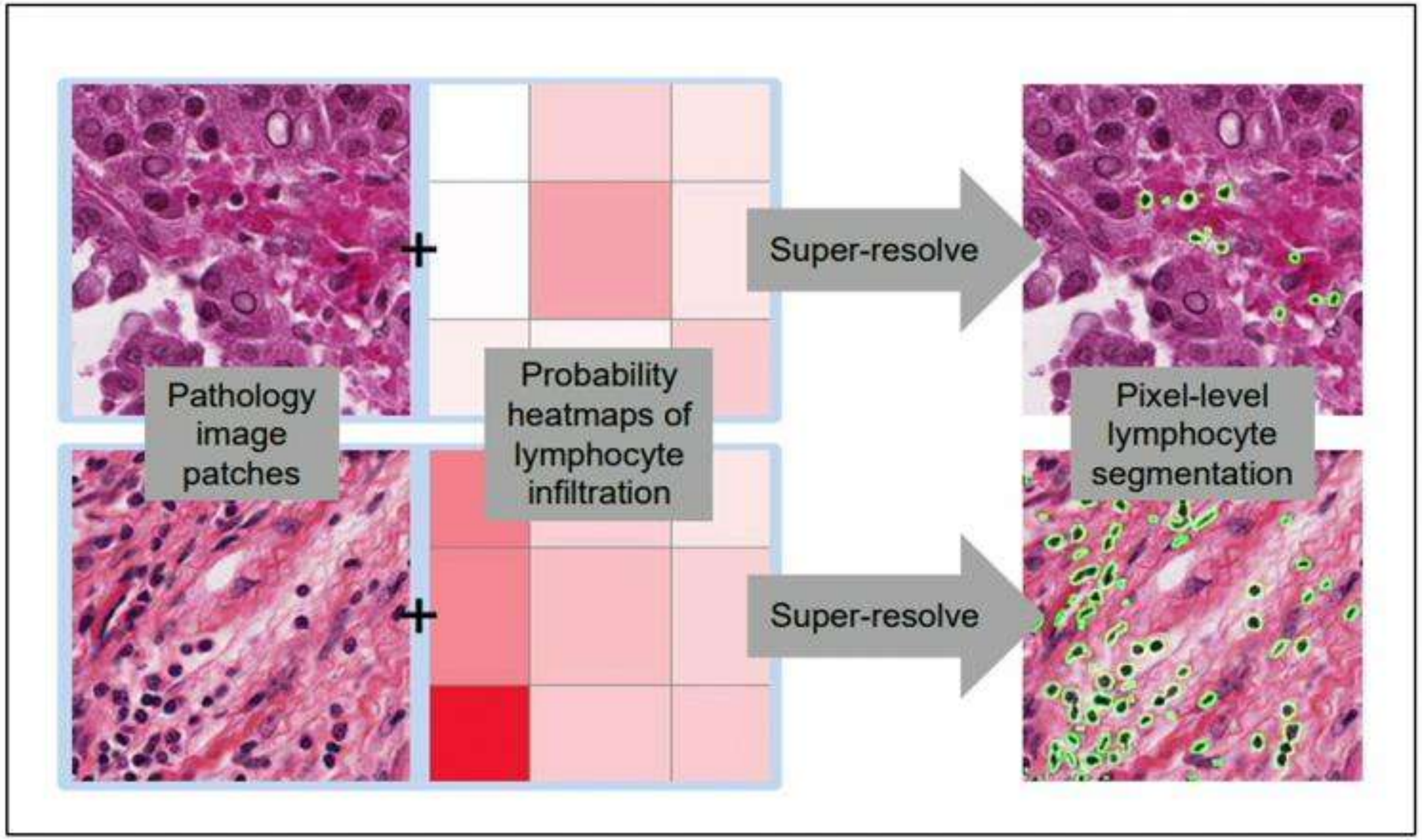


Low res NLCD labels



Iterations of gradient descent

Also works on pathology images



- Introduction
- Label super-resolution (ICLR 2019)
- **First US-wide 1m land cover map (CVPR 2019)**
- iNaturalist species observation (just starting on this)
- Hybrid intelligence approaches to land cover mapping (under review)
- Applications/collaborations
- Open problems

Setup

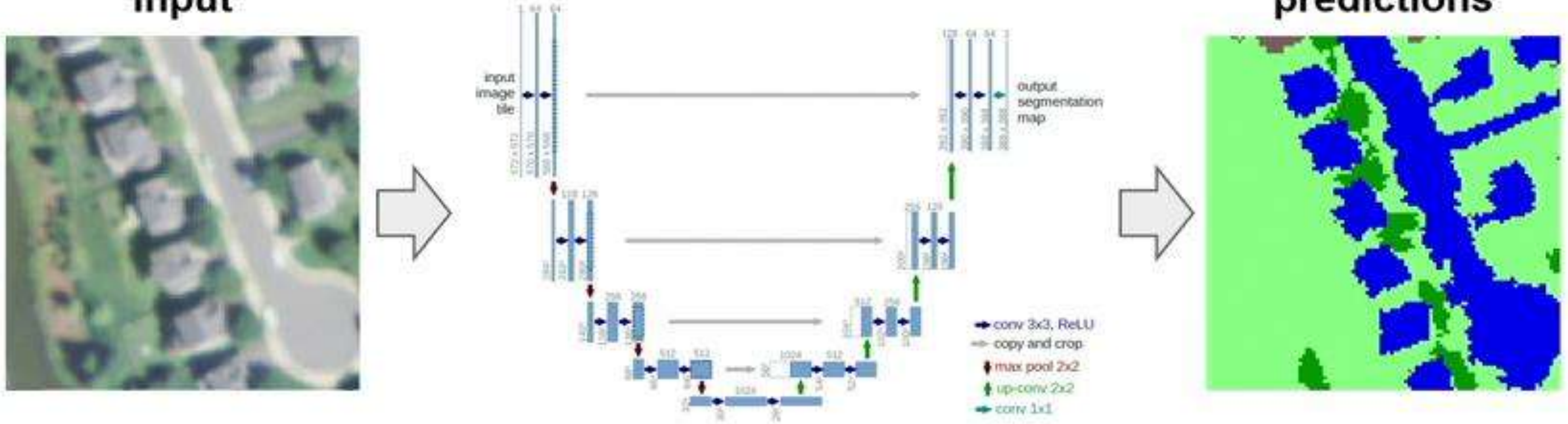
We have **limited** high-resolution (1m) land cover labels, only in the **Chesapeake Bay area**

We want to train a model that works over the **entire US**

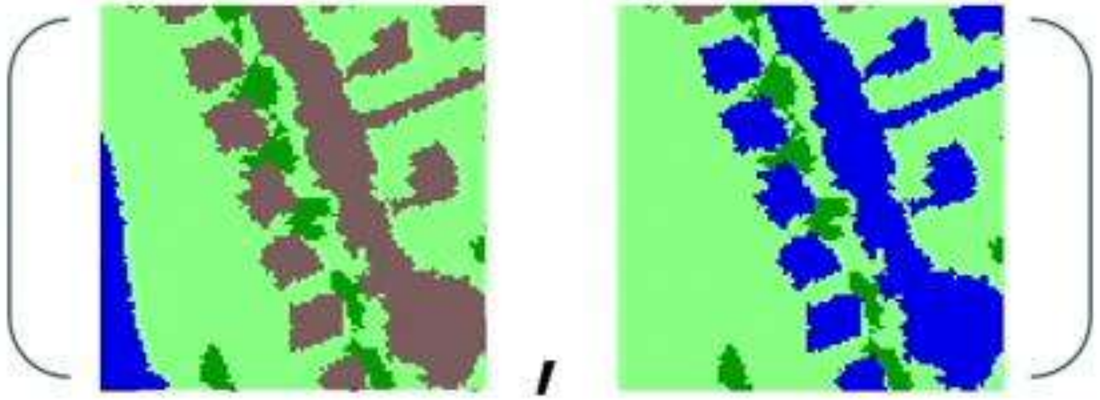
Over 8 trillion pixels and ~55TB of data

High-resolution input

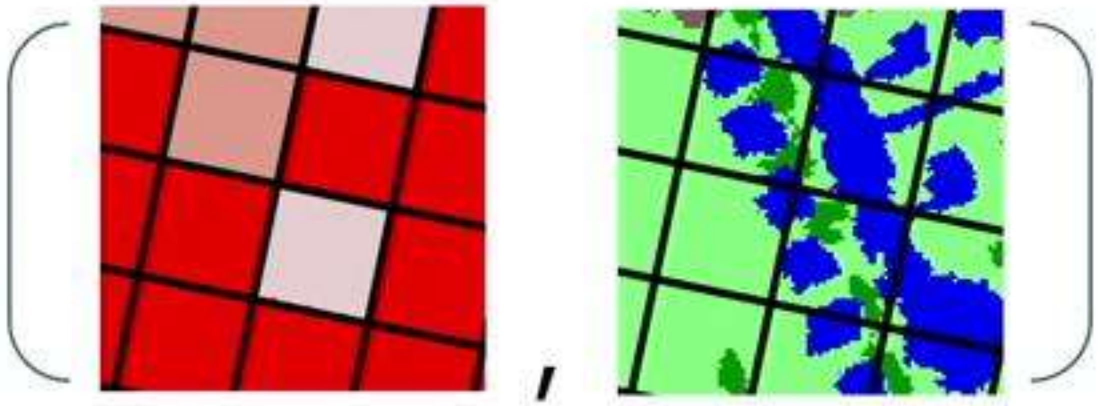
High-resolution predictions

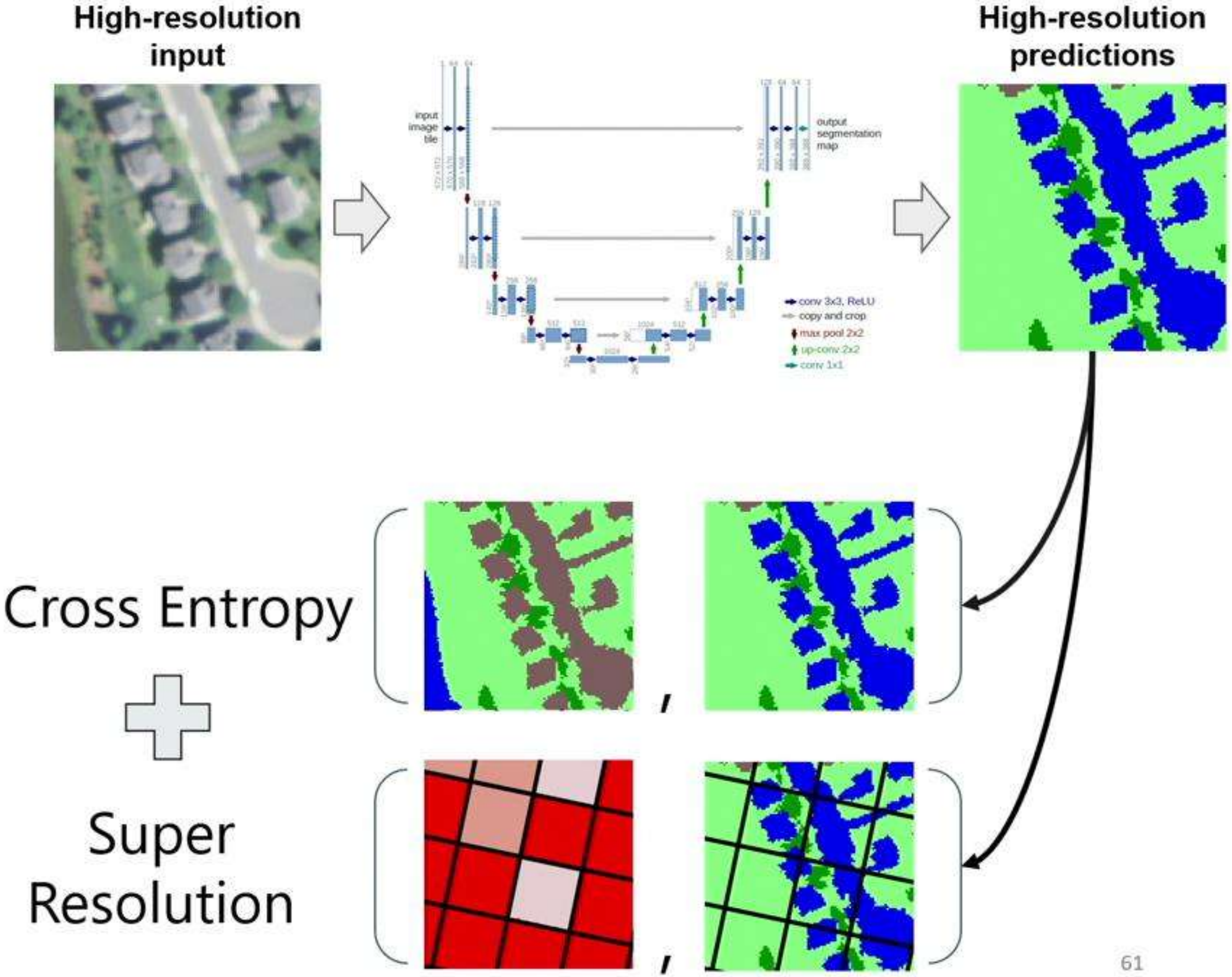


Cross Entropy



Super Resolution





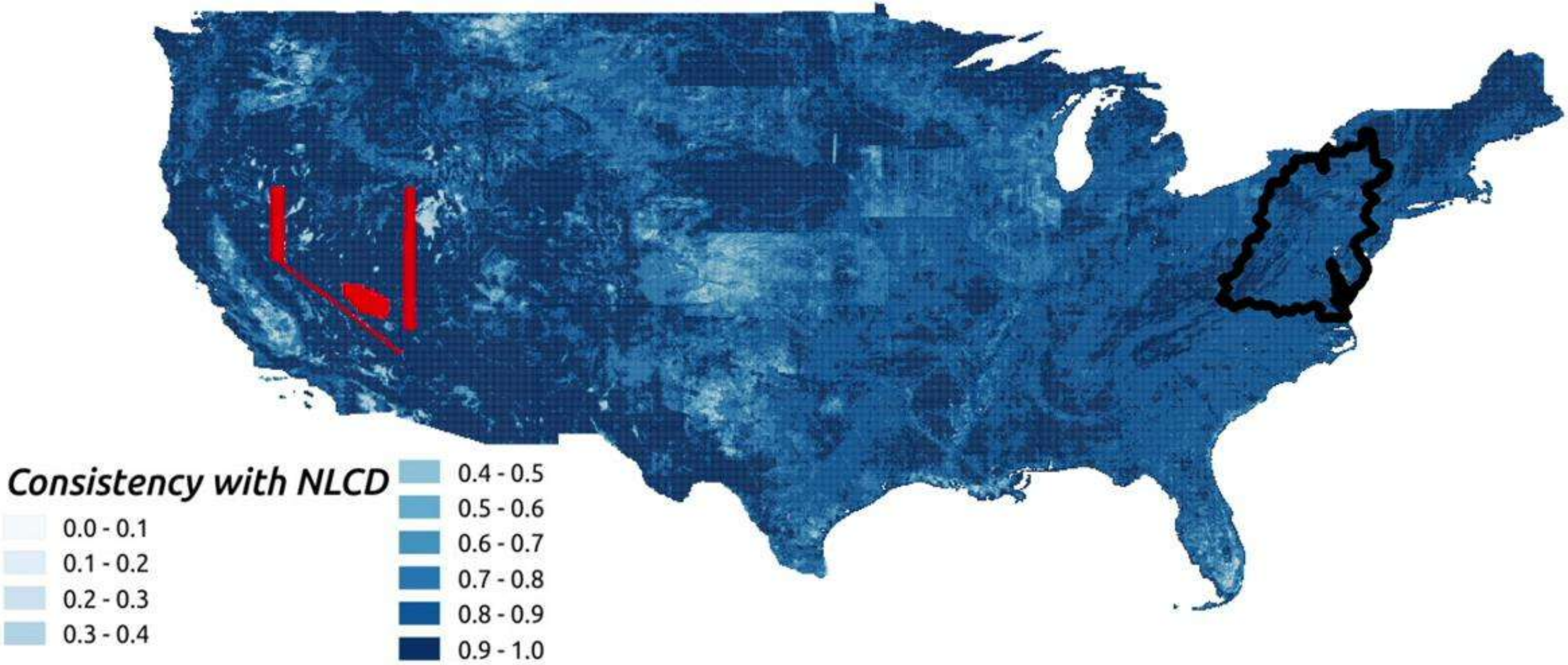
High-resolution input
High-resolution predictions
input image tile
output segmentation map
conv 3x3, ReLU
copy and crop
max pool 2x2
up-conv 2x2
conv 1x1
Cross Entropy
Super Resolution

Single Unet for the entire US

- Training
  - Sampled patches of land all over US and used super-res loss
  - Sampled patches of land from Chesapeake and used high-res loss
  - Color augmentation
  - Additional inputs (multiple time point Landsat)

  (the model is undertrained)
- Inference took 10 days on 40 GPUs

Land cover quality (a bound)

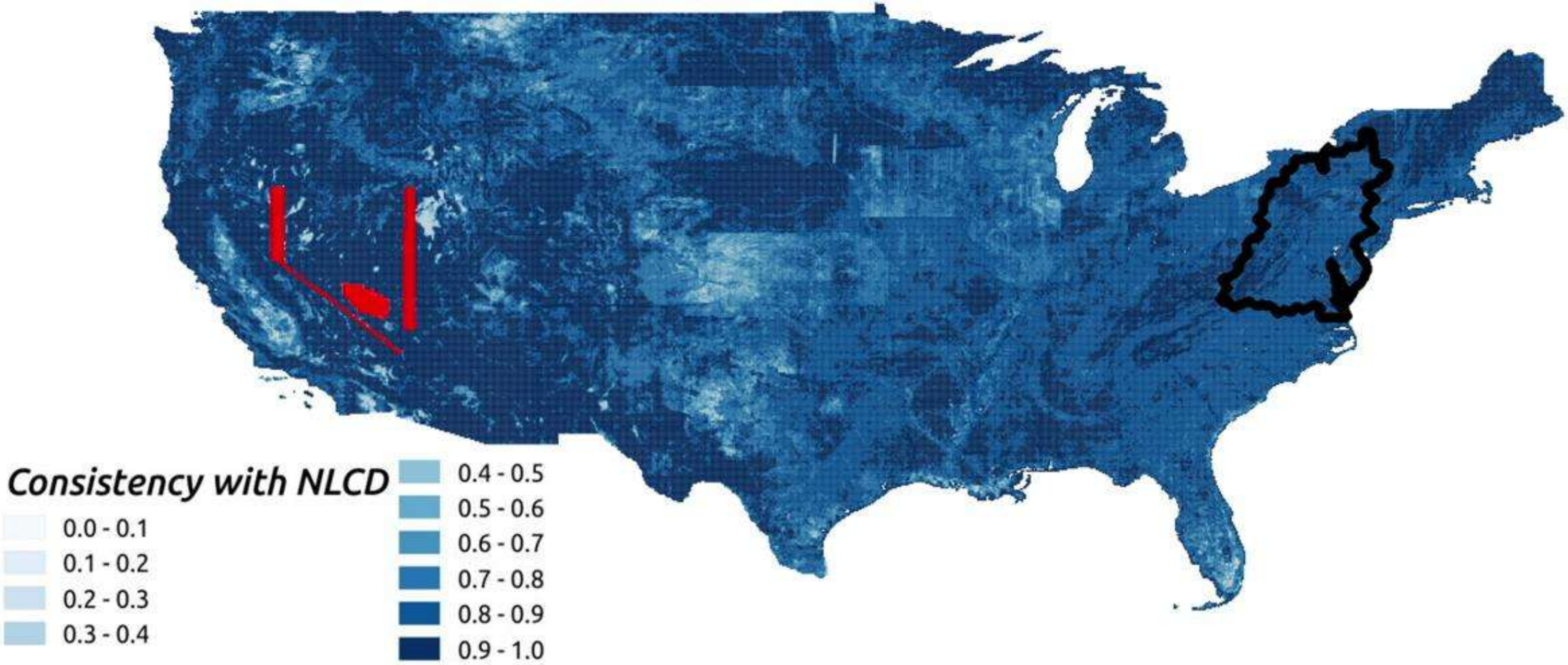


Beyond coarse labels

- NLCD is rarely updated and available only in the US
- Summary statistics for large areas are difficult to define consistently

- Point labels
  - Location precise at 1m (or close)
  - Possibly indirect label description
  - Very sparse

Land cover quality (a bound)



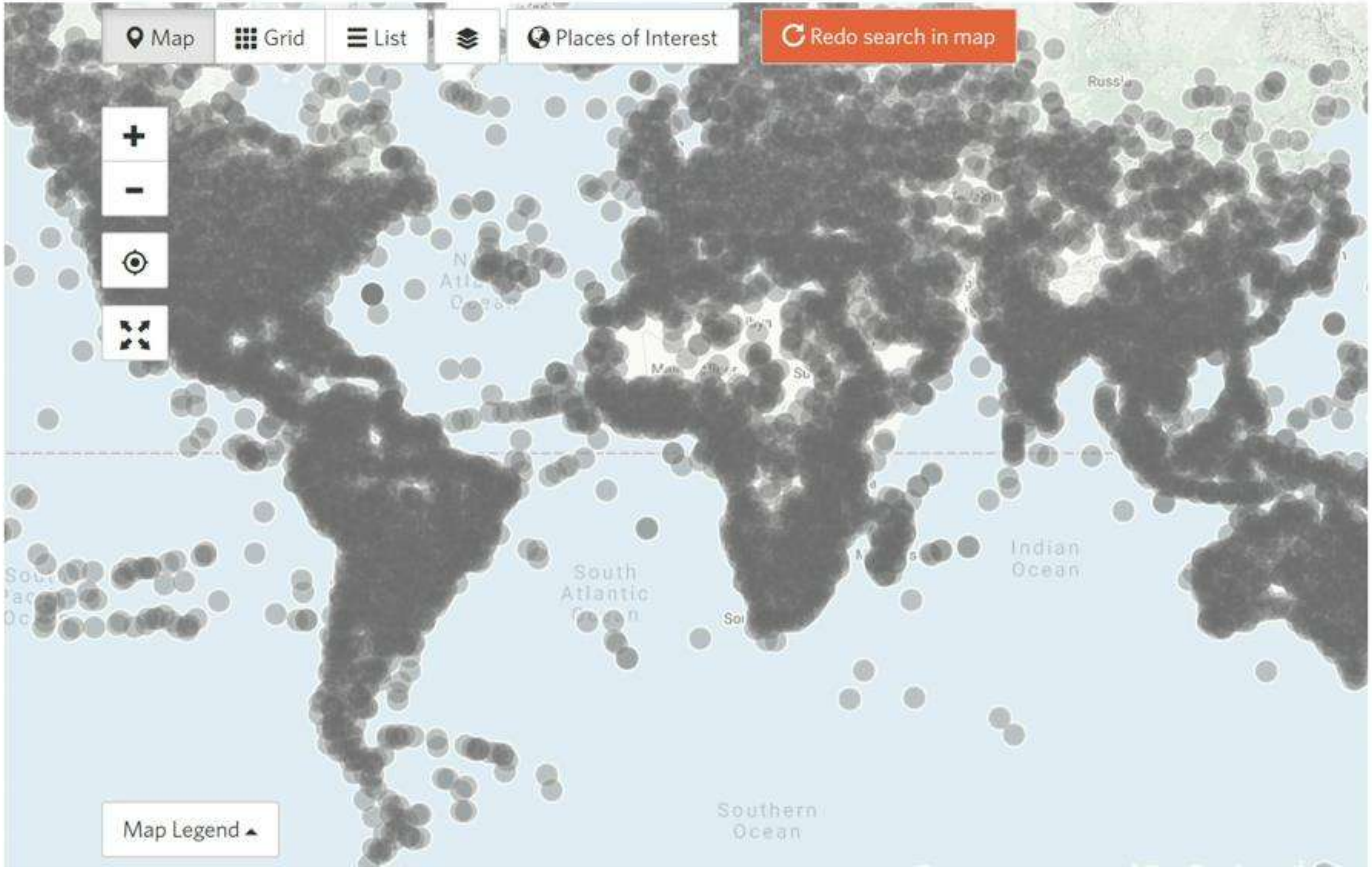
Beyond coarse labels

- NLCD is rarely updated and available only in the US
- Summary statistics for large areas are difficult to define consistently

- Point labels
  - Location precise at 1m (or close)
  - Possibly indirect label description
  - Very sparse

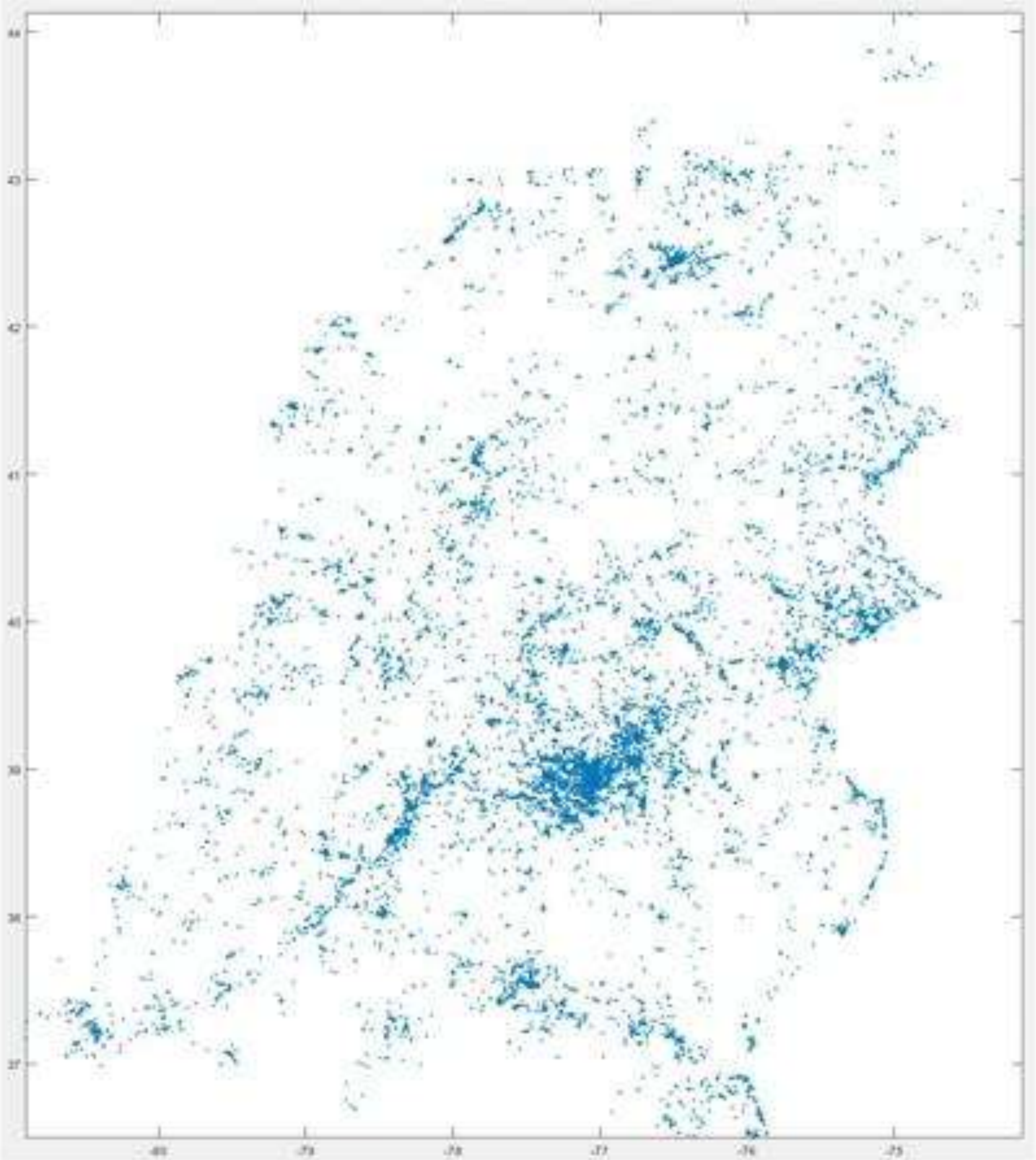
- Introduction
- Label super-resolution (ICLR 2019)
- First US-wide 1m land cover map (CVPR 2019)
- **iNaturalist species observation (just starting on this)**
- Hybrid intelligence approaches to land cover mapping (under review)
- Applications/collaborations
- Open problems

iNaturalist



Indirect point guidance

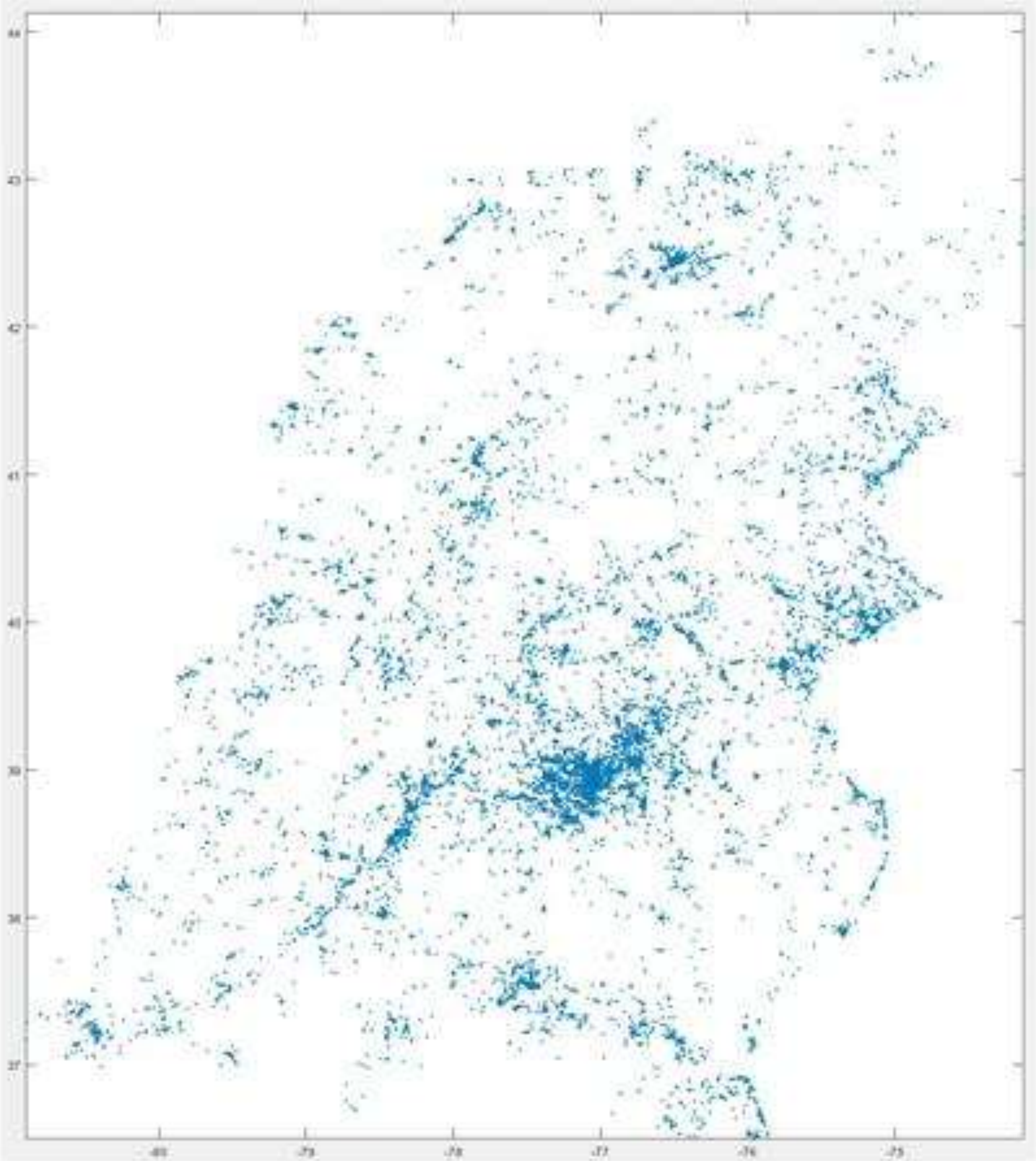
- Landmarks, e.g. OpenStreet, geotagged photos, **iNaturalist**

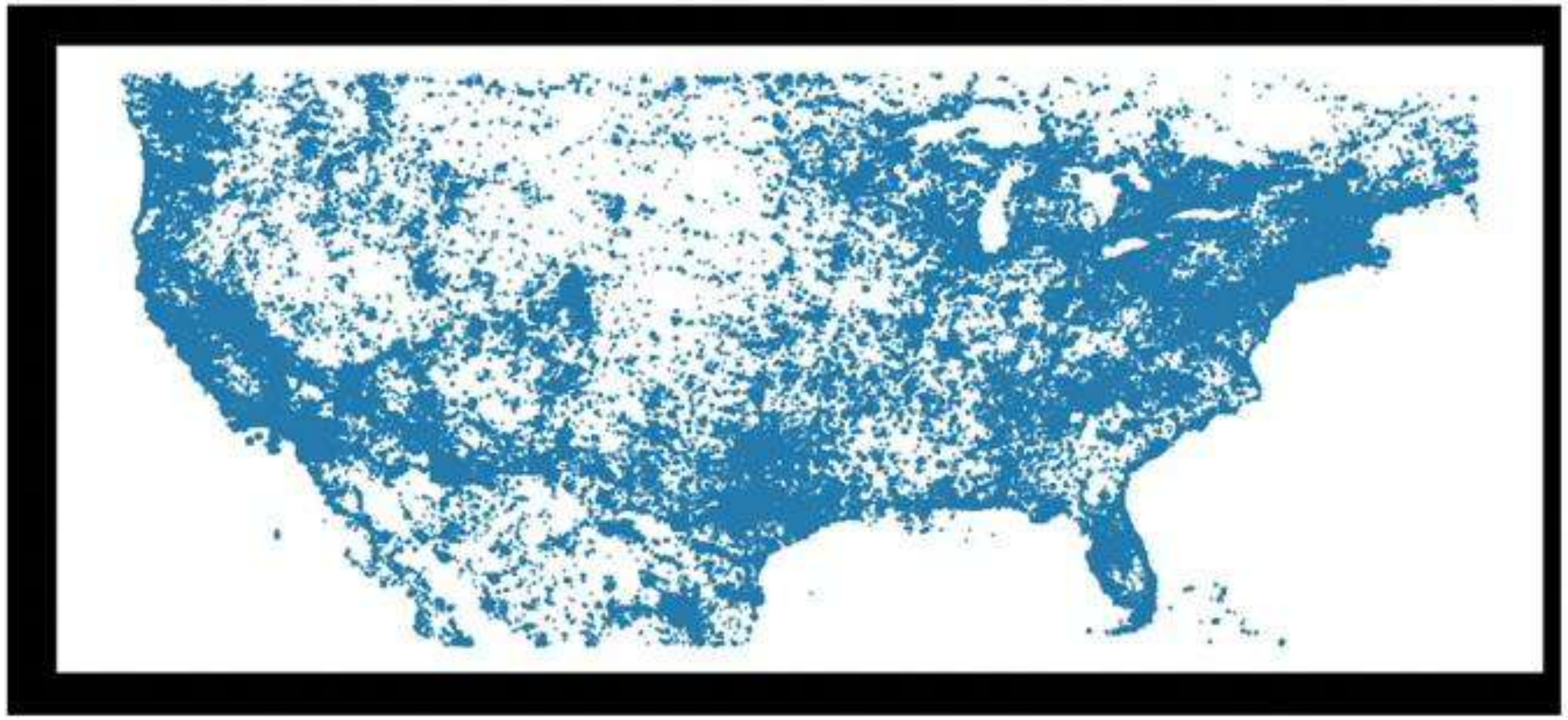


63k observations in Chesapeake

Indirect point guidance

- Landmarks, e.g. OpenStreet, geotagged photos, **iNaturalist**

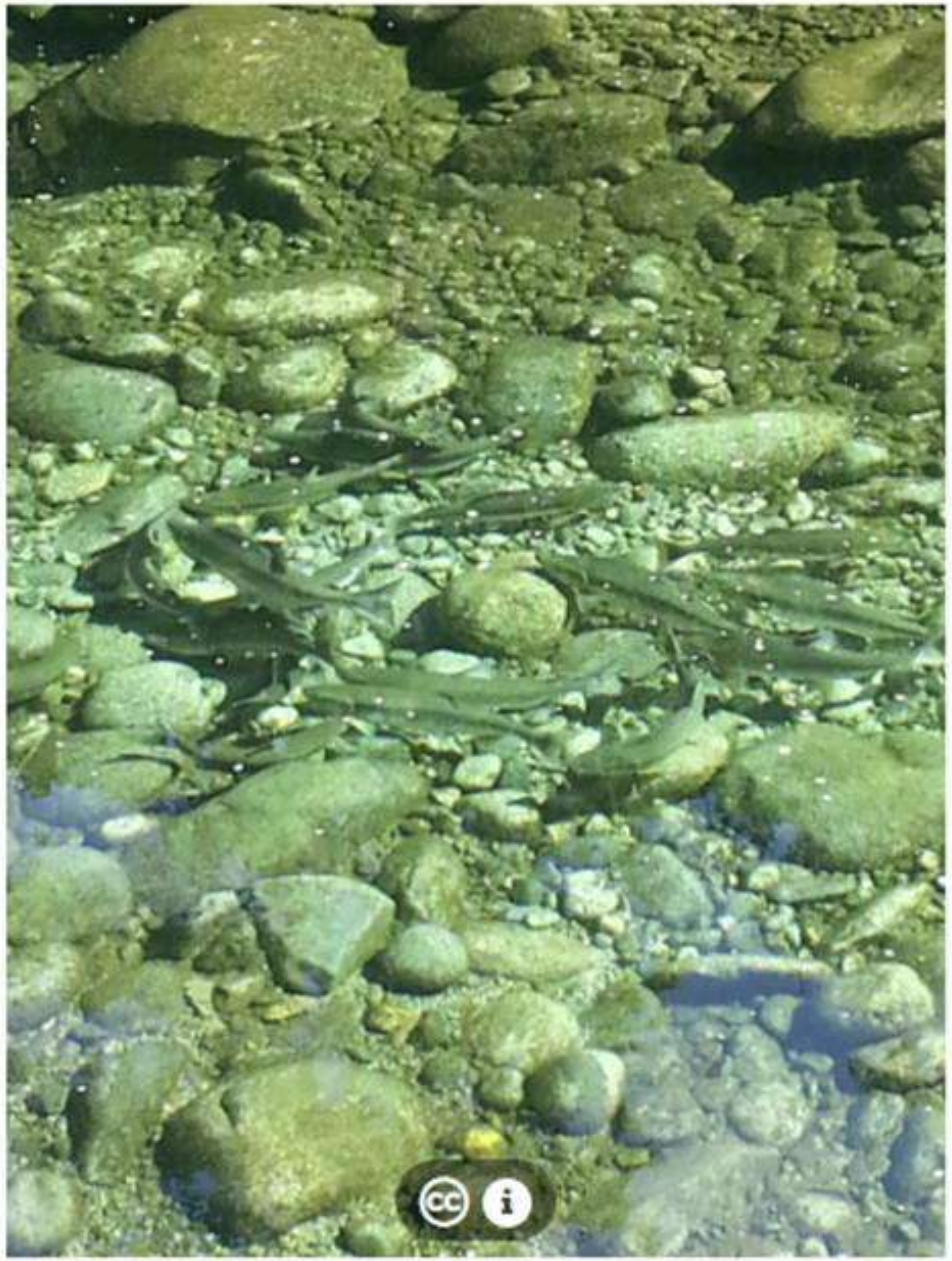




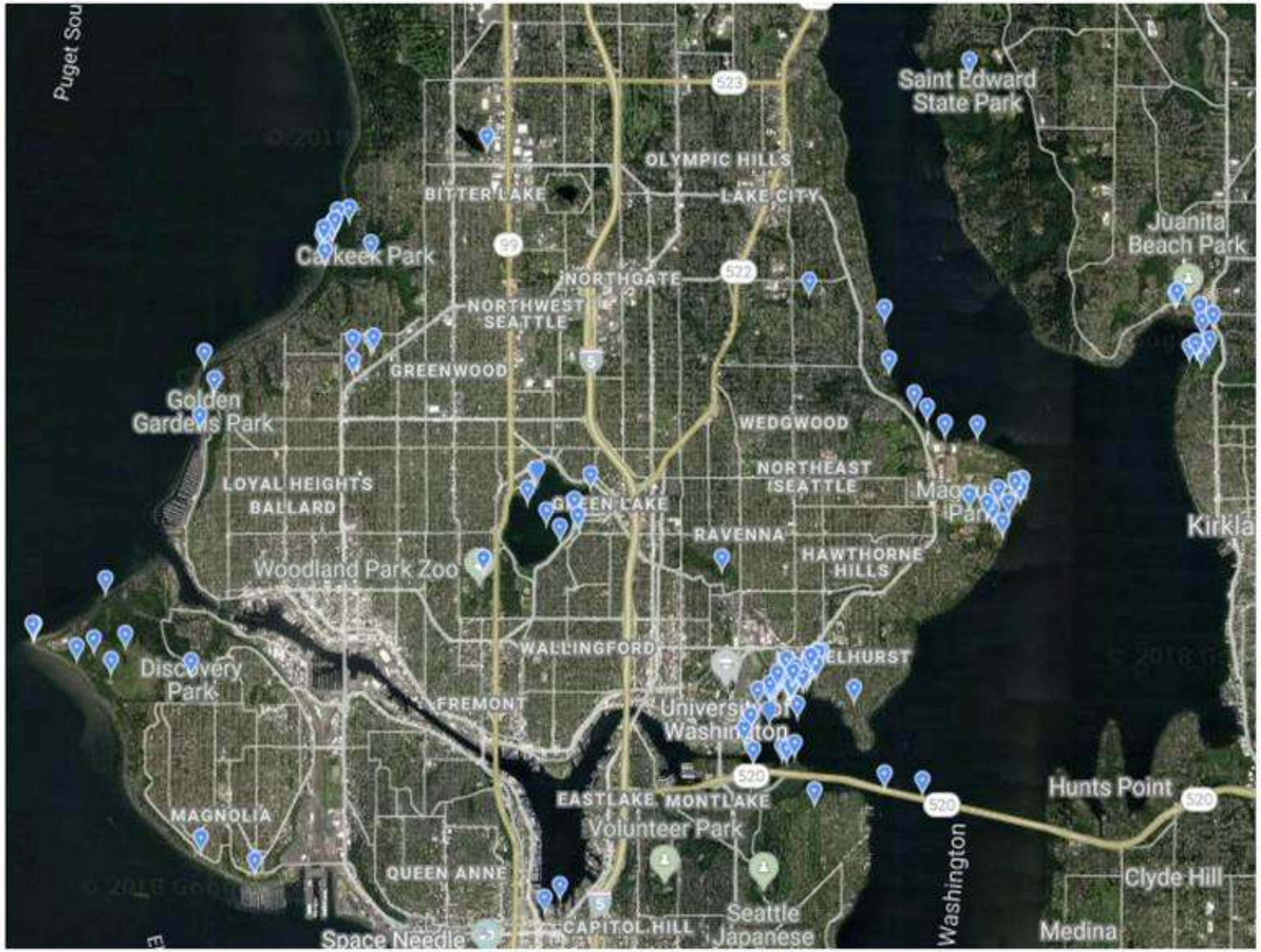
63k observations in Chesapeake


Pacific Salmons and Trouts
Observer: kriscu
Date: September 23, 2017
3 IDs | Needs ID View »

Found in a small group of a dozen salmon; staying in deeper pools of the creek

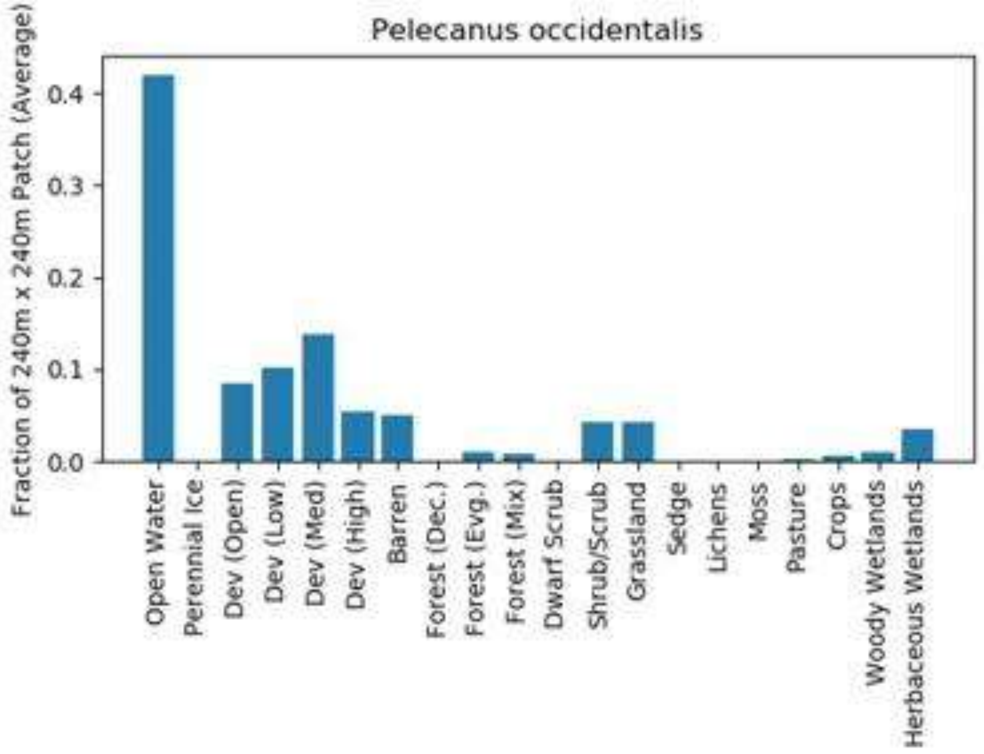


Sightings of Haliaeetus (Sea egles and fish-eating eagles)

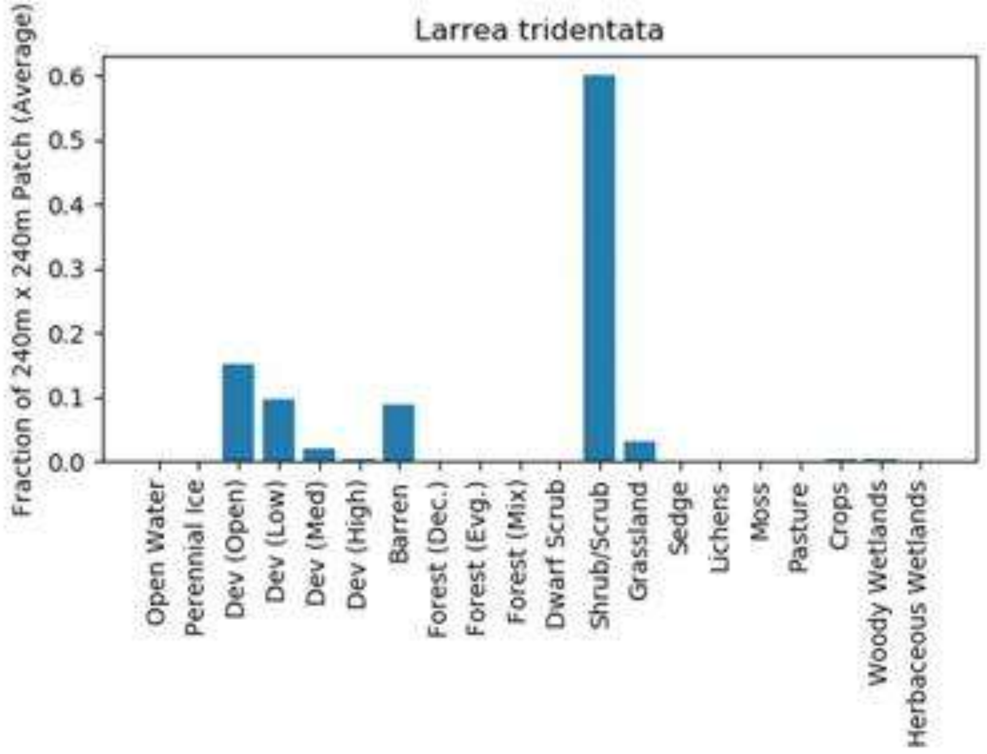


Species and Land Cover

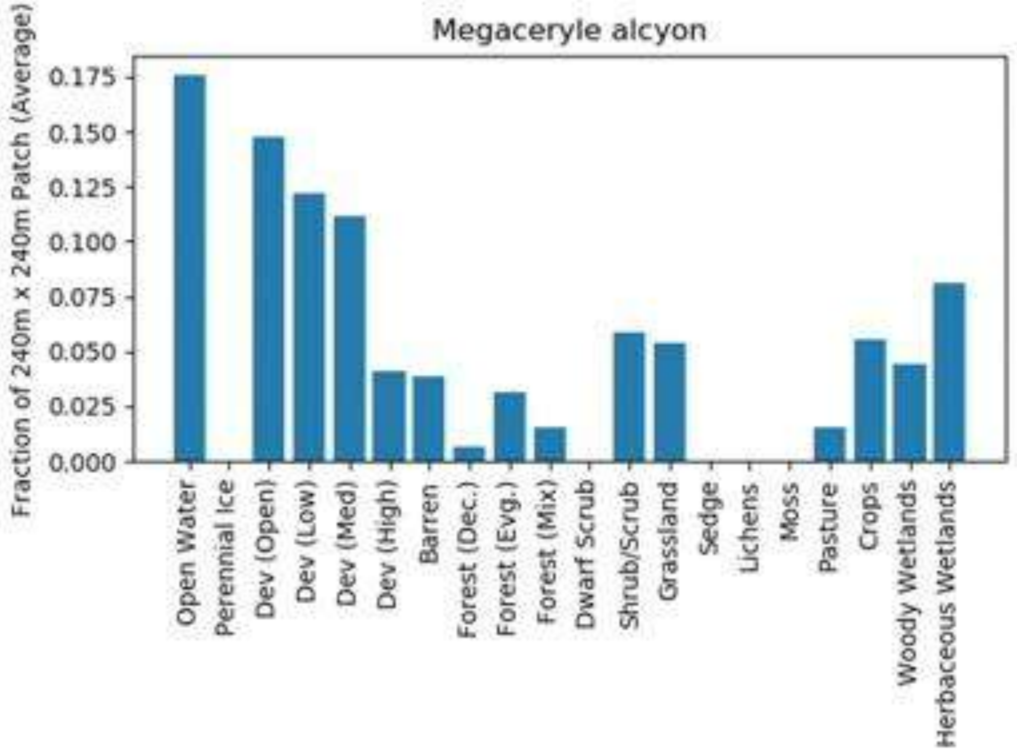
- Different species tend to occur in areas with different types of land cover.
- For a given species, can compute average local land cover distribution over all observations of that species.



Brown Pelican



Chaparral

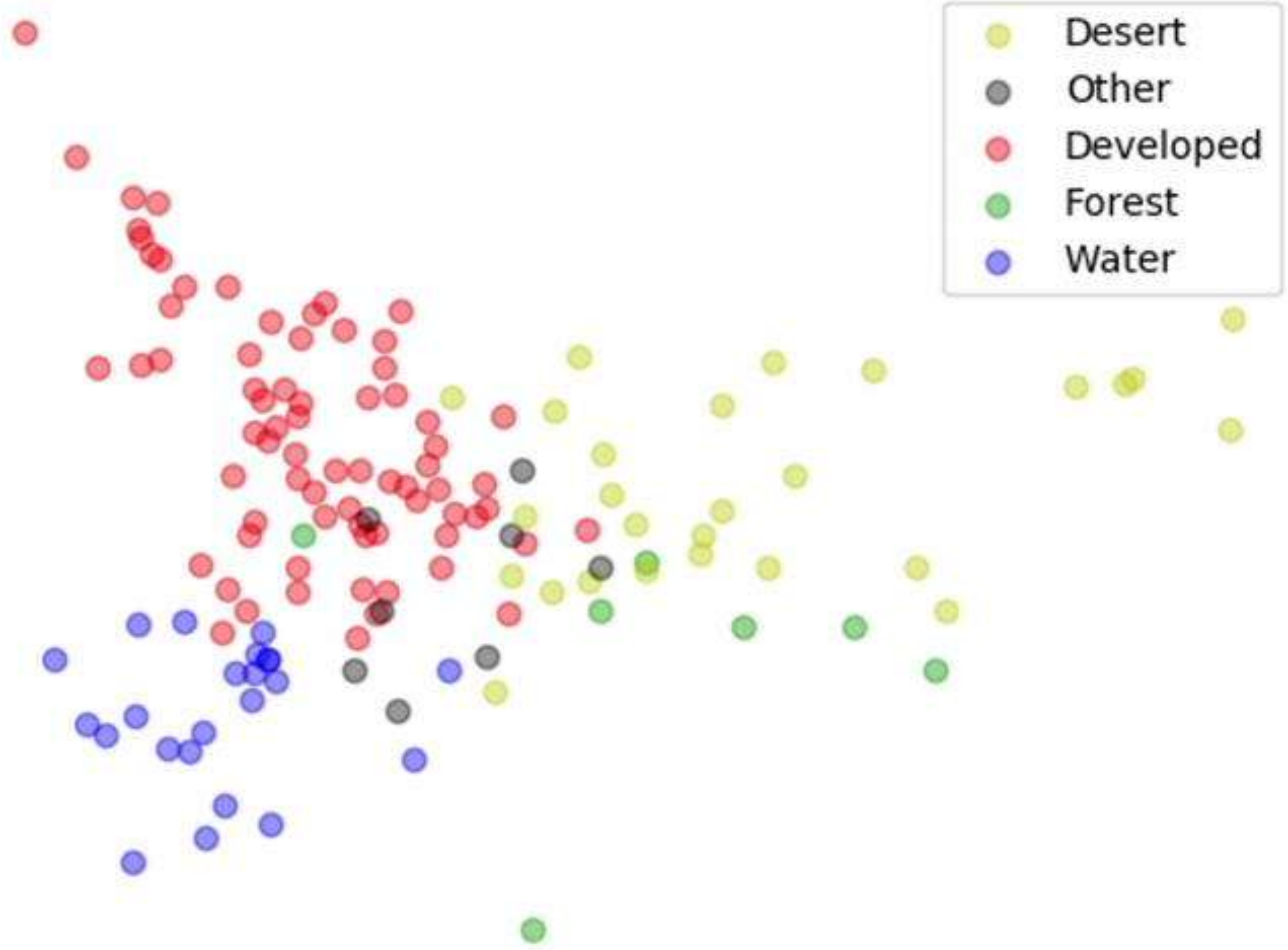


Belted Kingfisher

Species and Land Cover

- Applying linear dimensionality reduction like PCA to the average land cover distributions reveals some structure

- Each marker is one species, colored according to dominant land cover type.



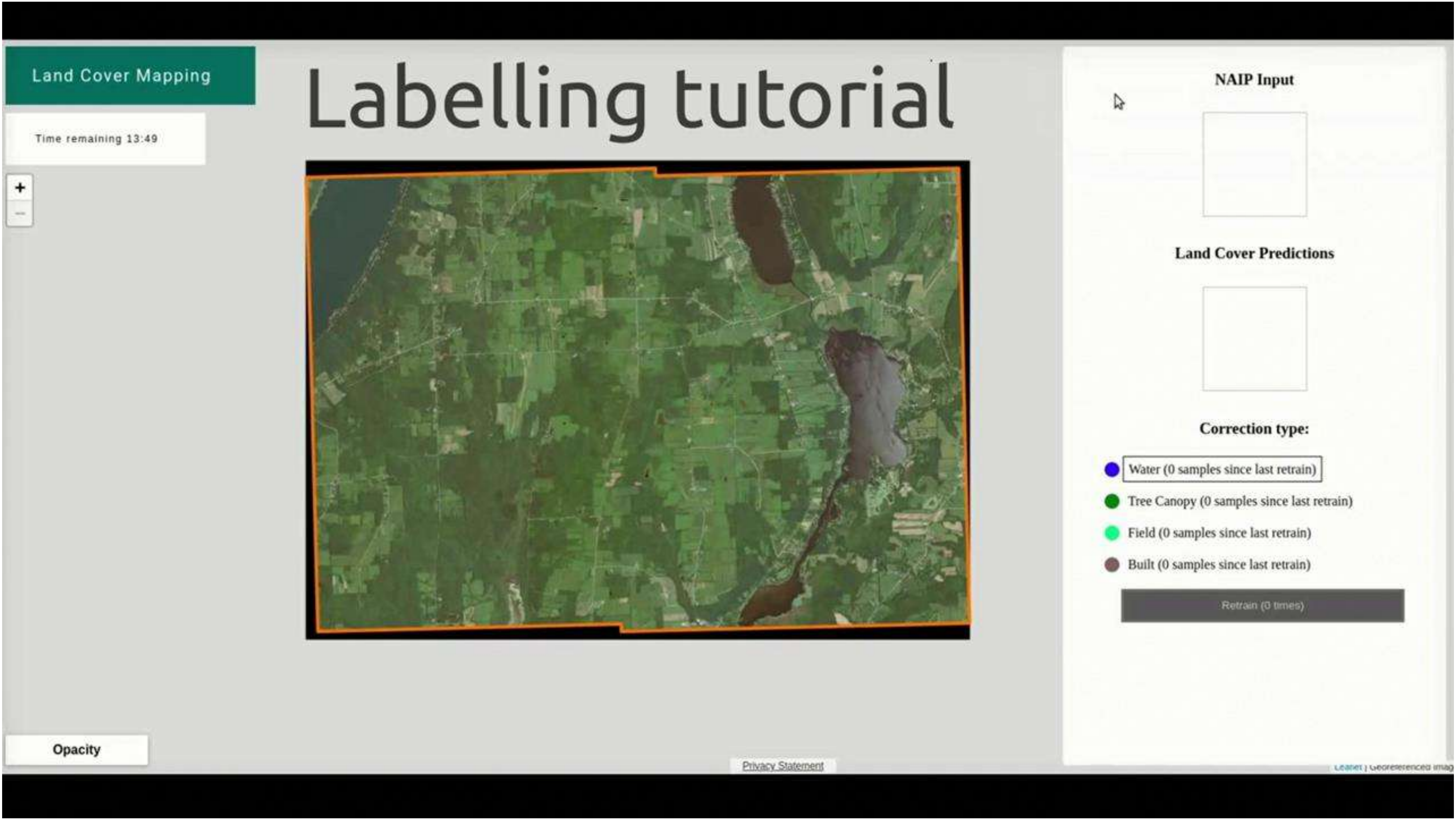

Map
Grid
List
Places of Interest
Redo search in map
Map Legend
Unknown
Laurel Creek Conse... • Jun 11, 2019
17m
Unknown
Laurel Creek Conse... • Jun 11, 2019
18m
Earthworms
(Subclass Oligochaeta)
Laurel Creek Conse... • Jun 11, 2019
18m
Unknown
Laurel Creek Conse... • Jun 11, 2019
18m
Unknown
Kent Island
Royal Oak
Oxford
Trappe
Secretary
Hurlock
Tilghman Island
Deale

- Introduction
- Label super-resolution (ICLR 2019)
- First US-wide 1m land cover map (CVPR 2019)
- iNaturalist species observation (just starting on this)
- **Hybrid intelligence approaches to land cover mapping (under review)**
- Applications/collaborations
- Open problems

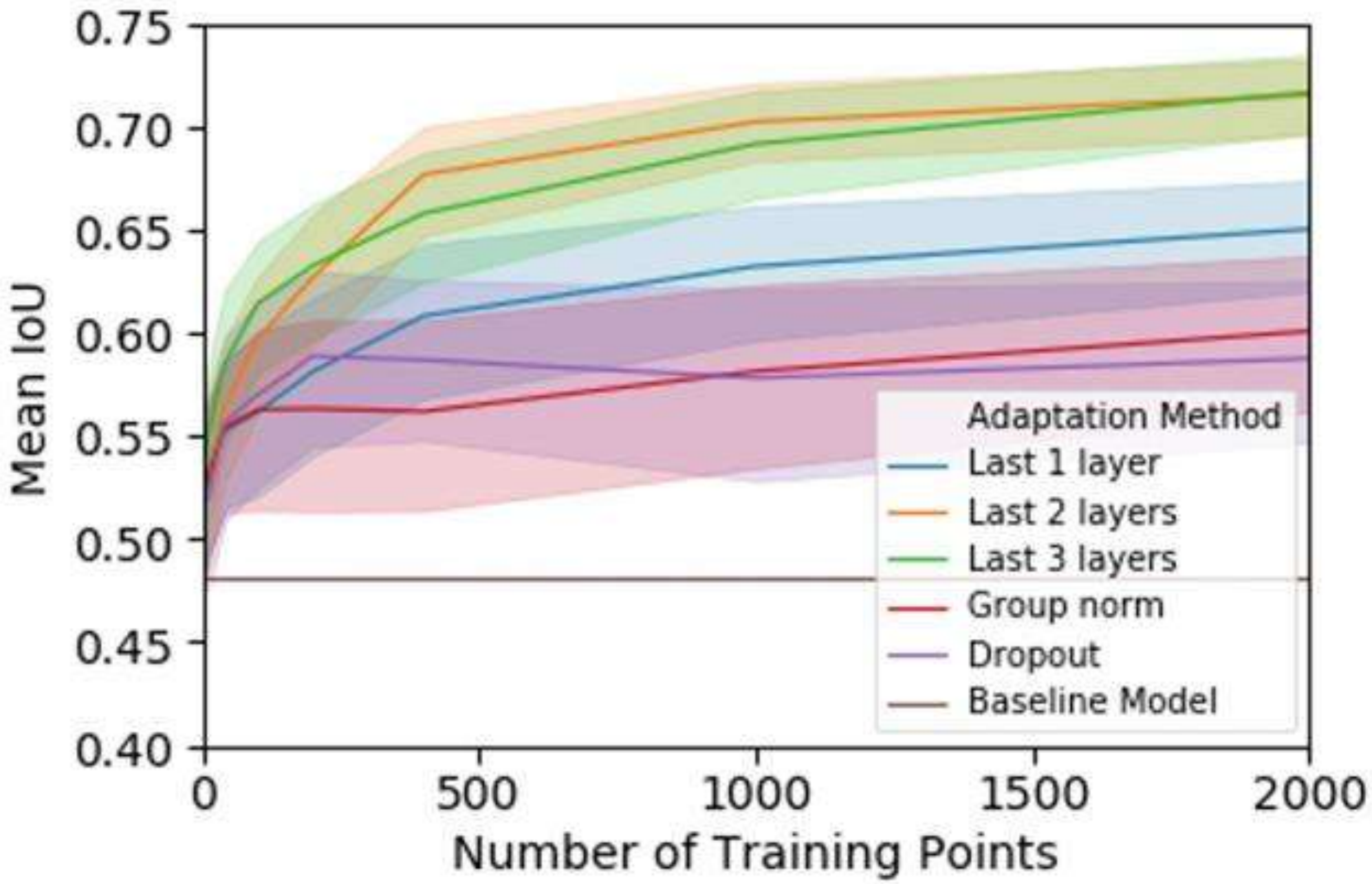
How many new point labels do we need?

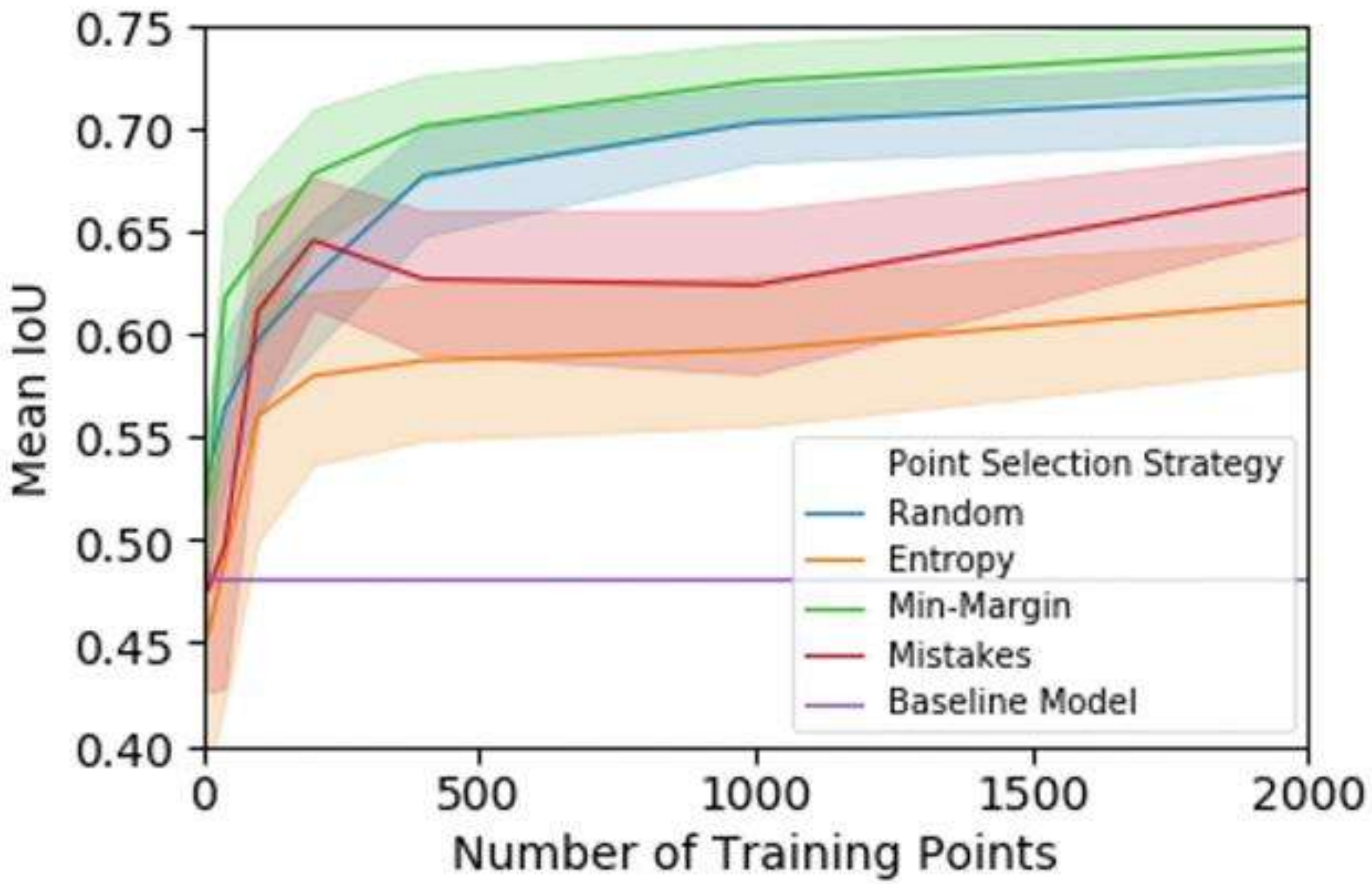
- We can ask humans to label 100s or 1000s of points and adapt the model by:
  - Retraining last k layers
  - Dropout configuration search
  - Group norm parameter adaptation
  - …
- If we also retrain the model on the fly as the labelers are using it, then this may further increase sample efficiency as humans find better points to label


Land Cover Mapping
NAIP Input
Land Cover Predictions
Correction type:
Opacity
Remember to check how your fine-tuned model is performing locally


Land Cover Mapping
Time remaining 13:49
Labelling tutorial
+
Opacity
Privacy Statement
NAIP Input
Land Cover Predictions
Correction type:
Water (0 samples since last retrain)
Tree Canopy (0 samples since last retrain)
Field (0 samples since last retrain)
Built (0 samples since last retrain)
Retrain (0 times)

Offline analysis of Fine-Tuning and Query Methods





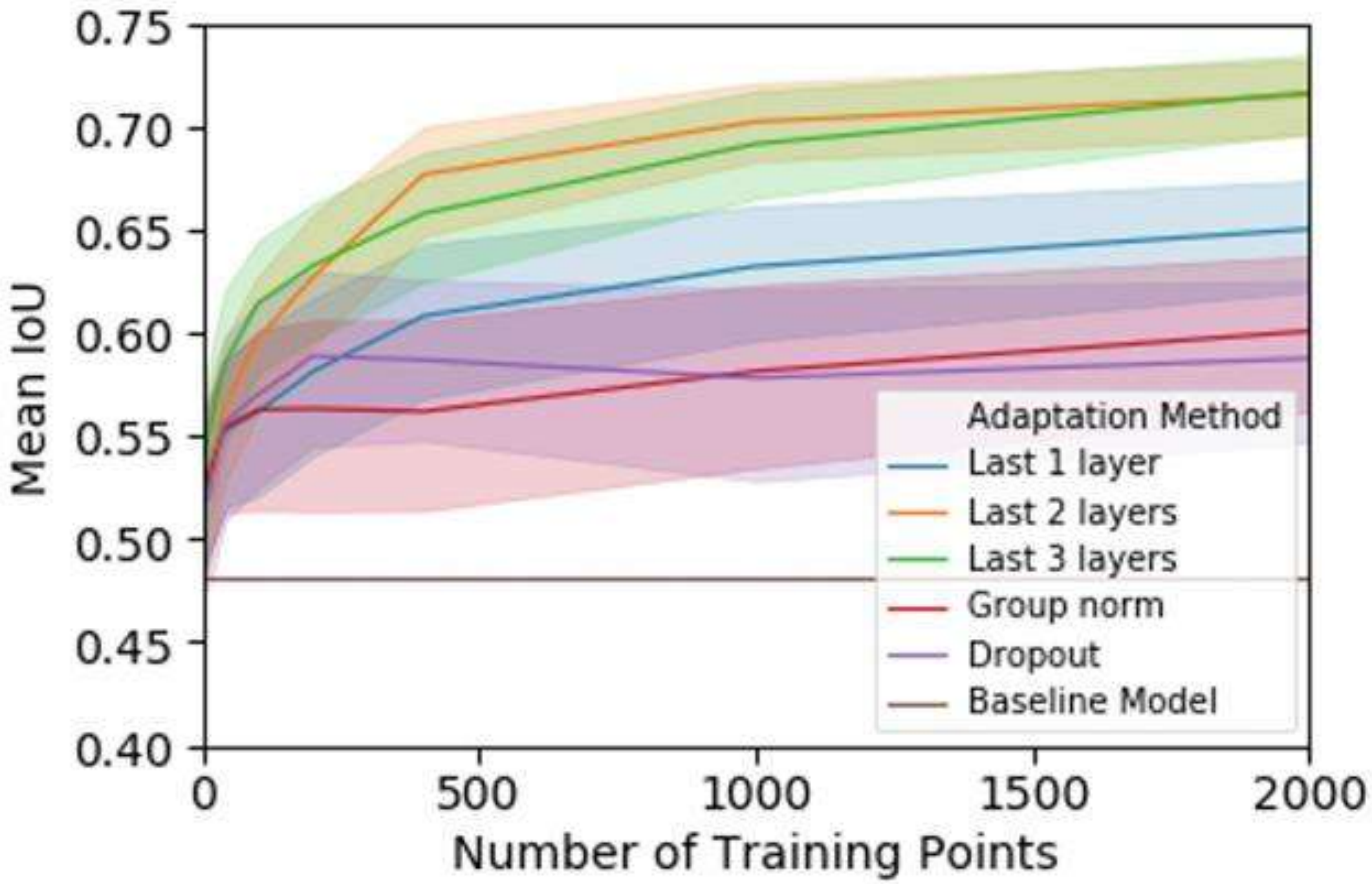
Different fine-tuning methods
Random query method

Last 2 Layers fine-tuning method
Different query methods

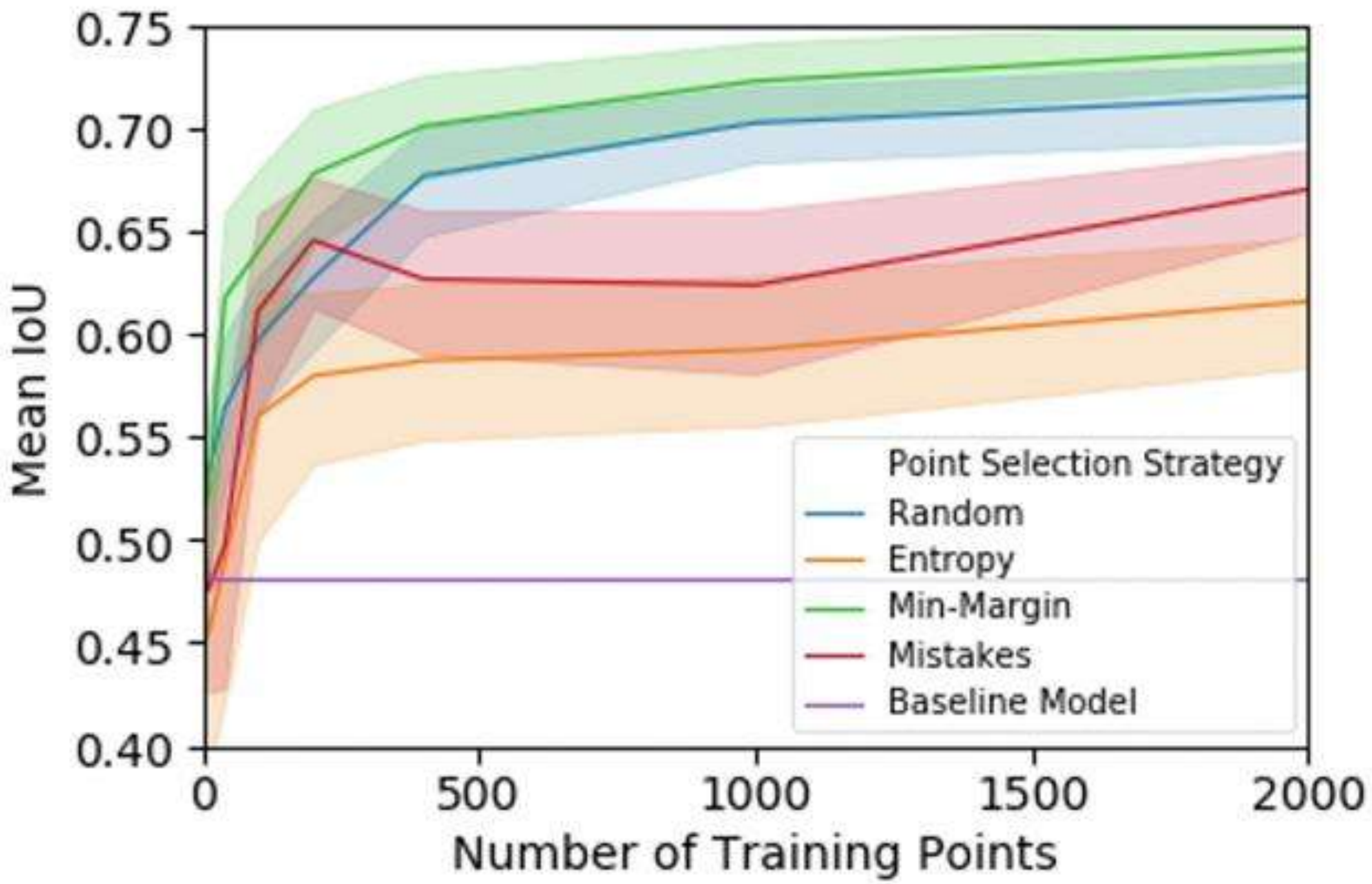
Our human-in-the-loop study

- 50 mechanical turkers with >95% approval
- Each working 15 minutes in each of the four different areas in NY state
- Always starting with the same base model pretrained in Maryland
- Two adaptation methods, last layer and last two layers fine tuning
    - first three tasks with one, and the fourth with the other
- Randomized orders

Offline analysis of Fine-Tuning and Query Methods



Different fine-tuning methods
Random query method

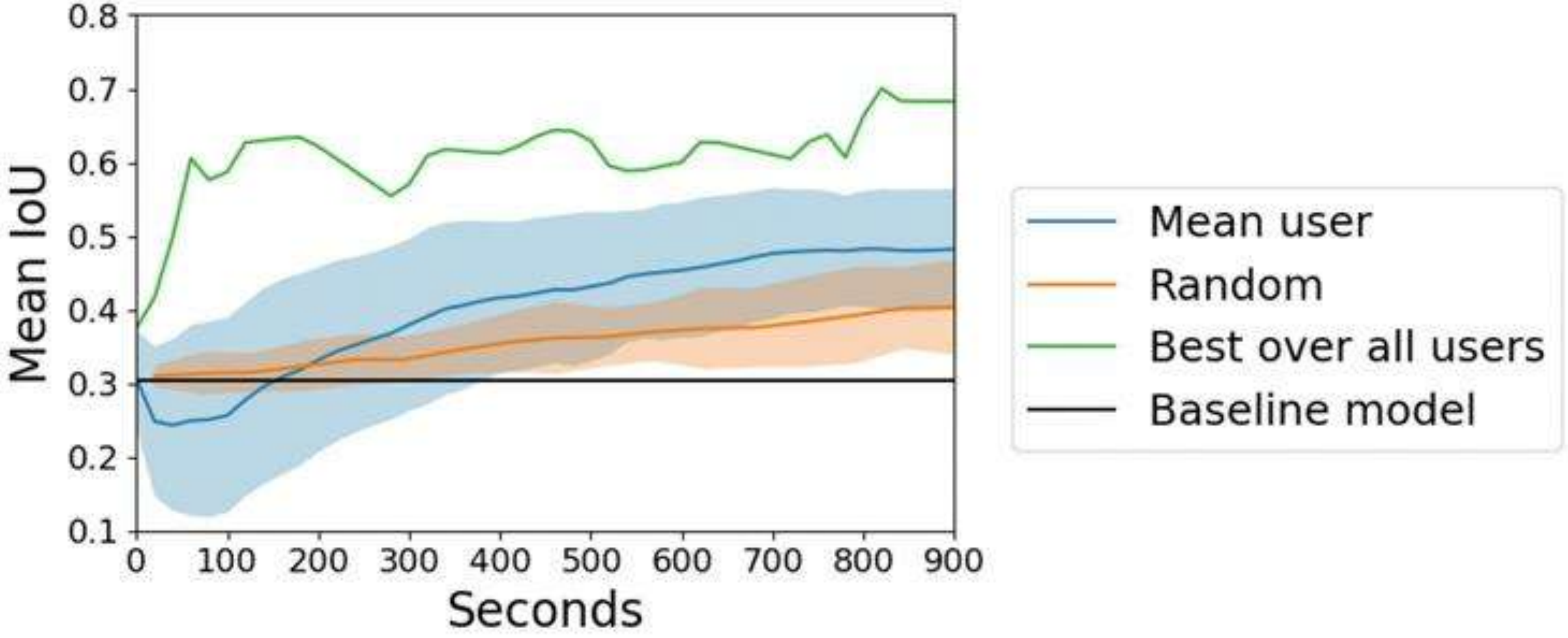


Last 2 Layers fine-tuning method
Different query methods

Our human-in-the-loop study

- 50 mechanical turkers with >95% approval
- Each working 15 minutes in each of the four different areas in NY state
- Always starting with the same base model pretrained in Maryland
- Two adaptation methods, last layer and last two layers fine tuning
  - first three tasks with one, and the fourth with the other
- Randomized orders

Humans Provide Non-Random Corrections, Achieving Better Performance



Skill differentiation and theory of mind

- Of the 25 top users in the second task, 17 are also in the top 25 in the third task
- The correlation breaks in the fourth area where the domain adaptation method switches
- In post-task interviews, the users describe the observed change in system behavior in the fourth task

Cost Savings Over Semi-Manual Labeling

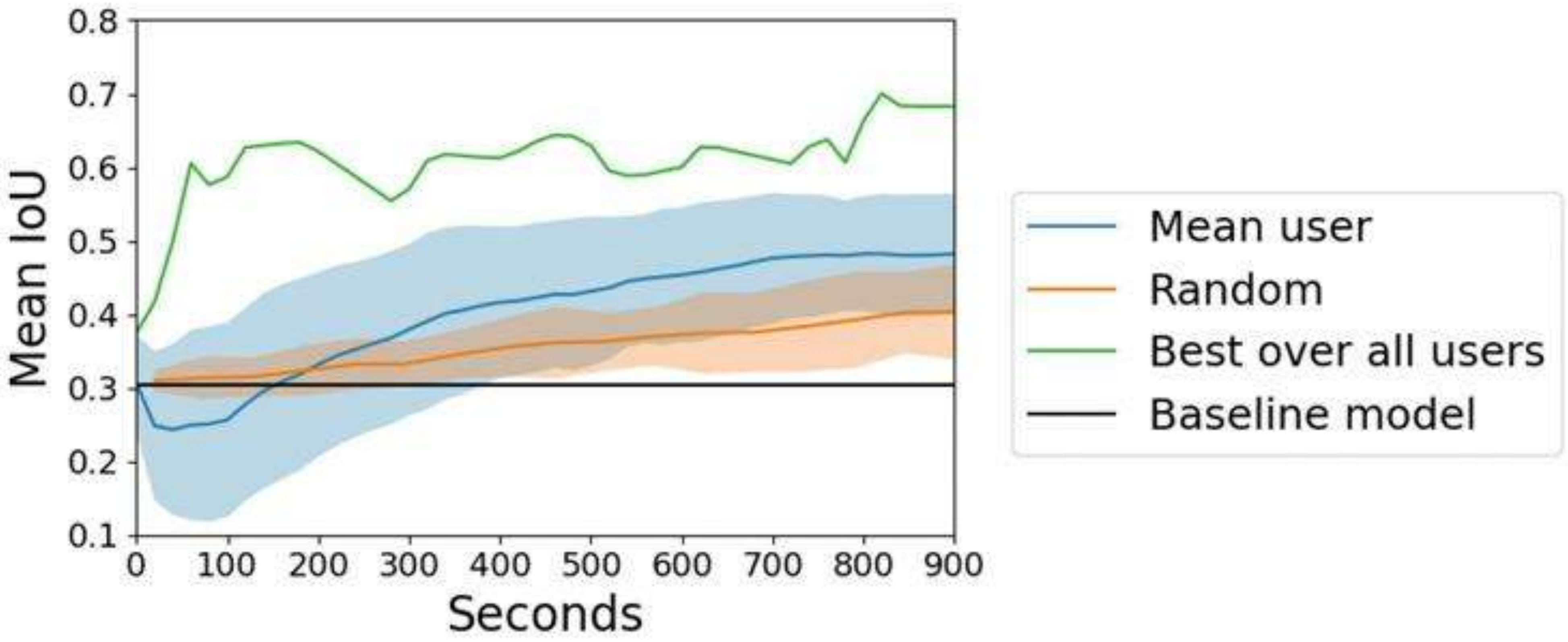
$1.3 M	**$18.5 K***
10 months	925 hours*
90-95% accuracy	89% accuracy*
Chesapeake Conservancy	AI4E

* Estimated based on performance of top user in 1-hour trial

Next for this summer

- Collaborative mapping
  - Displaying various heat maps to users so they see what other users are doing
- Spatial ensembles
  - M models possibly with some trainable parameters
  - A spatial mapping function that maps lat/long to a weight distribution over models
  - Mixing models “before or after” the log (in param space or in predictions)
  - Backprop through the whole ensemble
- Collaboration challange

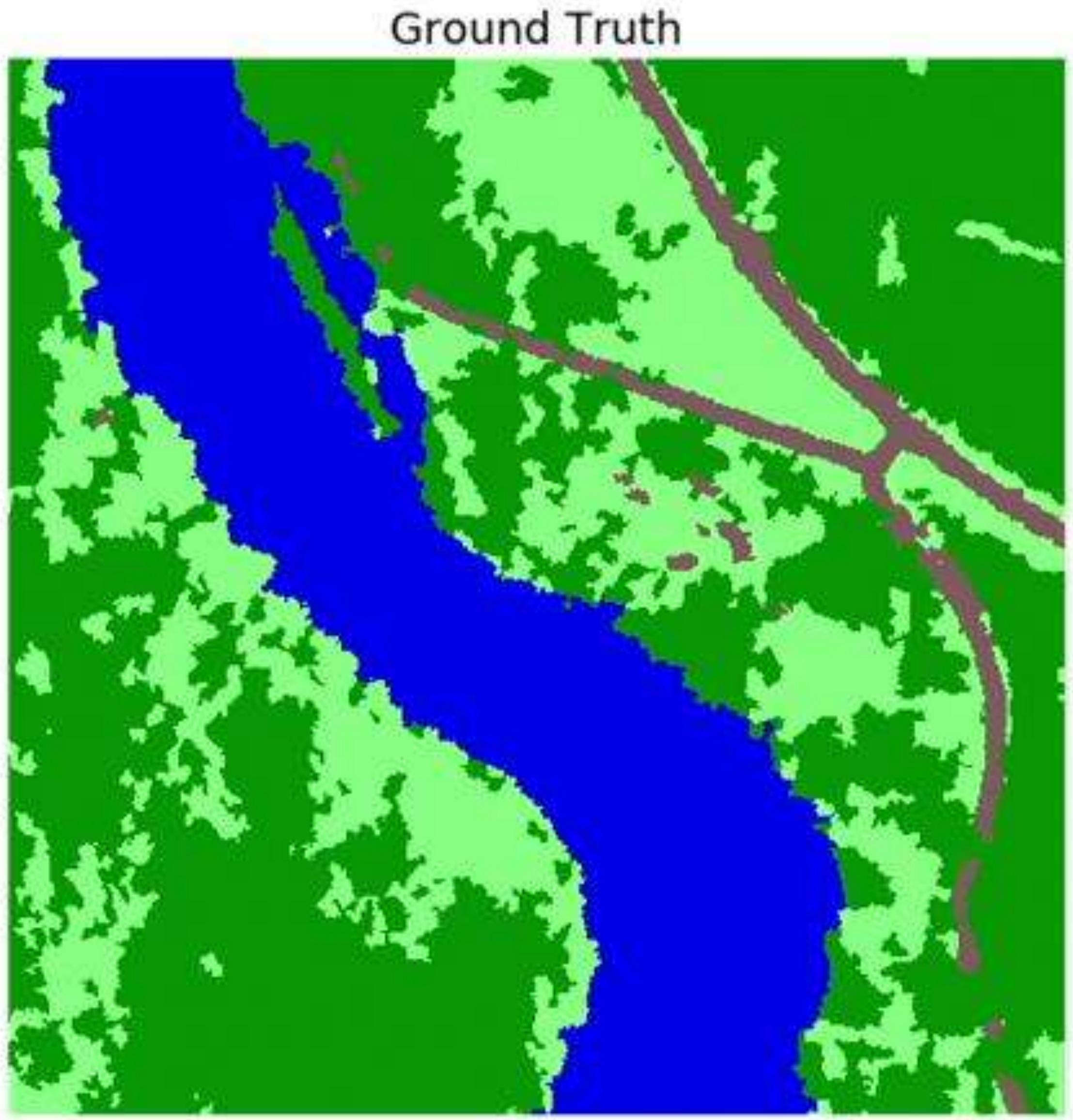
Humans Provide Non-Random Corrections, Achieving Better Performance



Next for this summer

- Collaborative mapping
  - Displaying various heat maps to users so they see what other users are doing
- Spatial ensembles
  - M models possibly with some trainable parameters
  - A spatial mapping function that maps lat/long to a weight distribution over models
  - Mixing models “before or after” the log (in param space or in predictions)
  - Backprop through the whole ensemble
- Collaboration challange

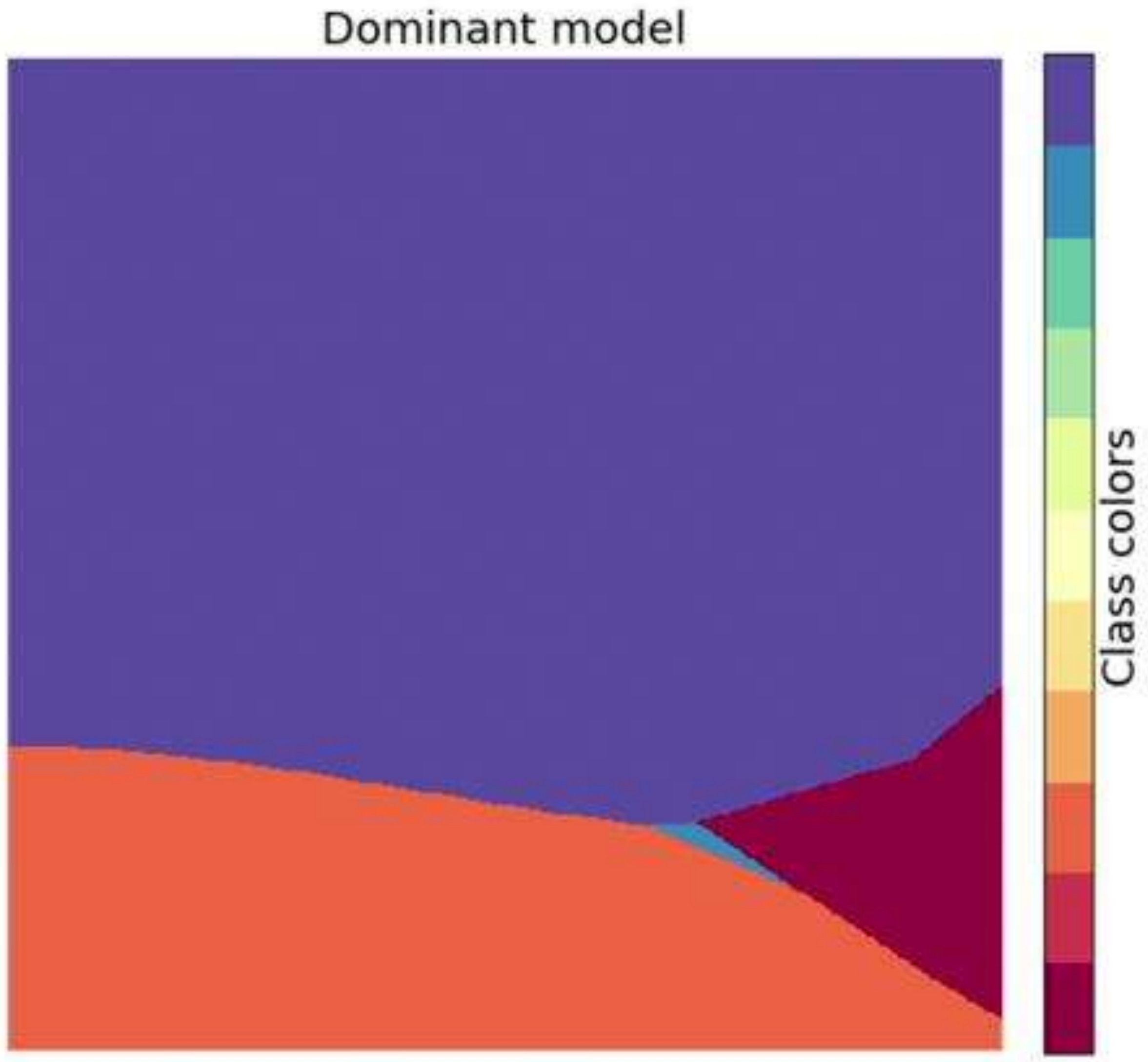
Segmentation



Input tile

Imagery

Tile specific model jurisdiction division (a lat/long -> model mapping)



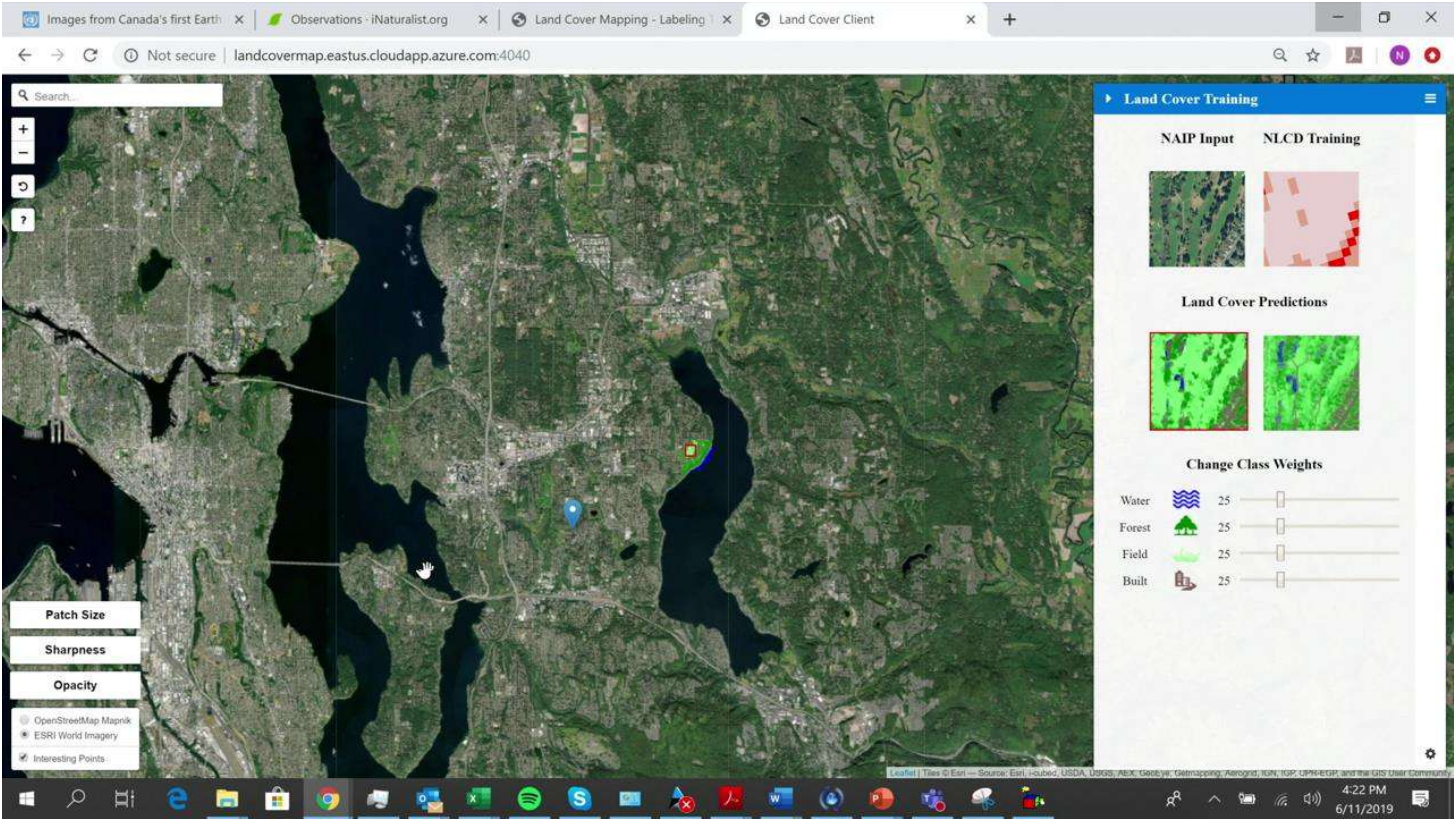
Collaborations

- Chesapeake Conservancy/NOAA and longitudinal wetlands mapping
- World Bank and Yangon and Ho Chi Minh municipal governments
- Various internal product/CSS collaborations

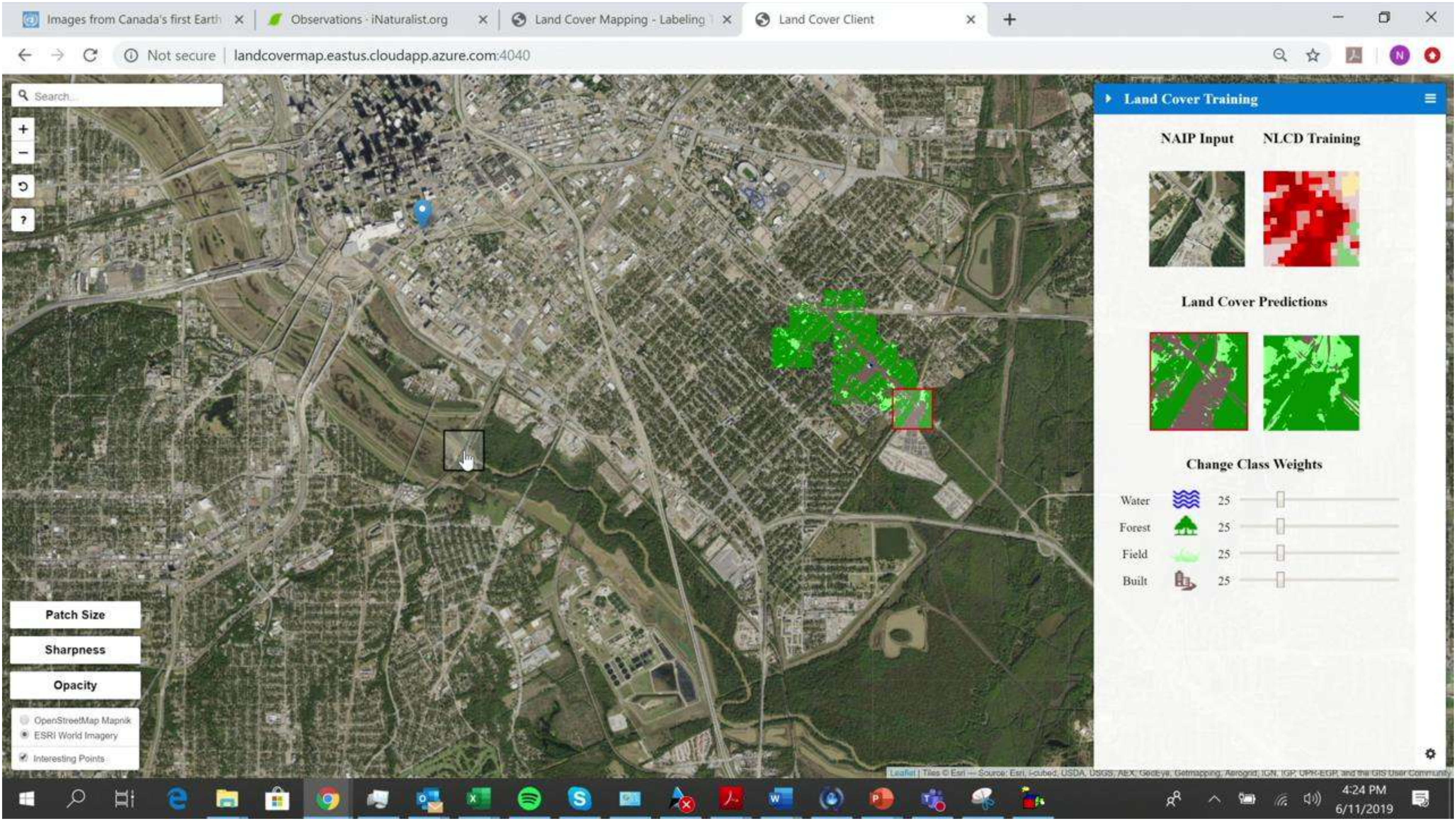
Open problems

- Collaborative mapping, humans in the loop

- Models
  - Unsupervised pretraining
  - Spatial ensembles
  - Clustering models
  - Meta learning



Images from Canada's first Earth
Observations · iNaturalist.org
Land Cover Mapping - Labeling
Land Cover Client
Not secure
landcovermap.eastus.cloudapp.azure.com:4040
Search...
Land Cover Training
NAIP Input
NLCD Training
Land Cover Predictions
Change Class Weights
Water 25
Forest 25
Field 25
Built 25
Patch Size
Sharpness
Opacity
OpenStreetMap Mapnik
ESRI World Imagery
Interesting Points
4:22 PM
6/11/2019


Images from Canada's first Earth
Observations · iNaturalist.org
Land Cover Mapping - Labeling
Land Cover Client
Not secure
landcovermap.eastus.cloudapp.azure.com:4040
Search...
Land Cover Training
NAIP Input
NLCD Training
Land Cover Predictions
Change Class Weights
Water 25
Forest 25
Field 25
Built 25
Patch Size
Sharpness
Opacity
OpenStreetMap Mapnik
ESRI World Imagery
Interesting Points
4:24 PM
6/11/2019