Human-Machine Collaboration for Fast Land Cover Mapping

Caleb Robinson, Anthony Ortiz, Kolya Malkin, Blake Elias, Andi Peng, Dan Morris, Bistra Dilkina, Nebojsa Jojic

calebrob6@gmail.com

Collaborators

Anthony Ortiz - University of Texas at El Paso

Kolya Malkin - Yale University

Blake Elias - Microsoft AI Resident

Andi Peng - Microsoft AI Resident

Dan Morris - Microsoft AI for Earth

Bistra Dilkina - University of Southern California

Nebojsa Jojic - Microsoft Research

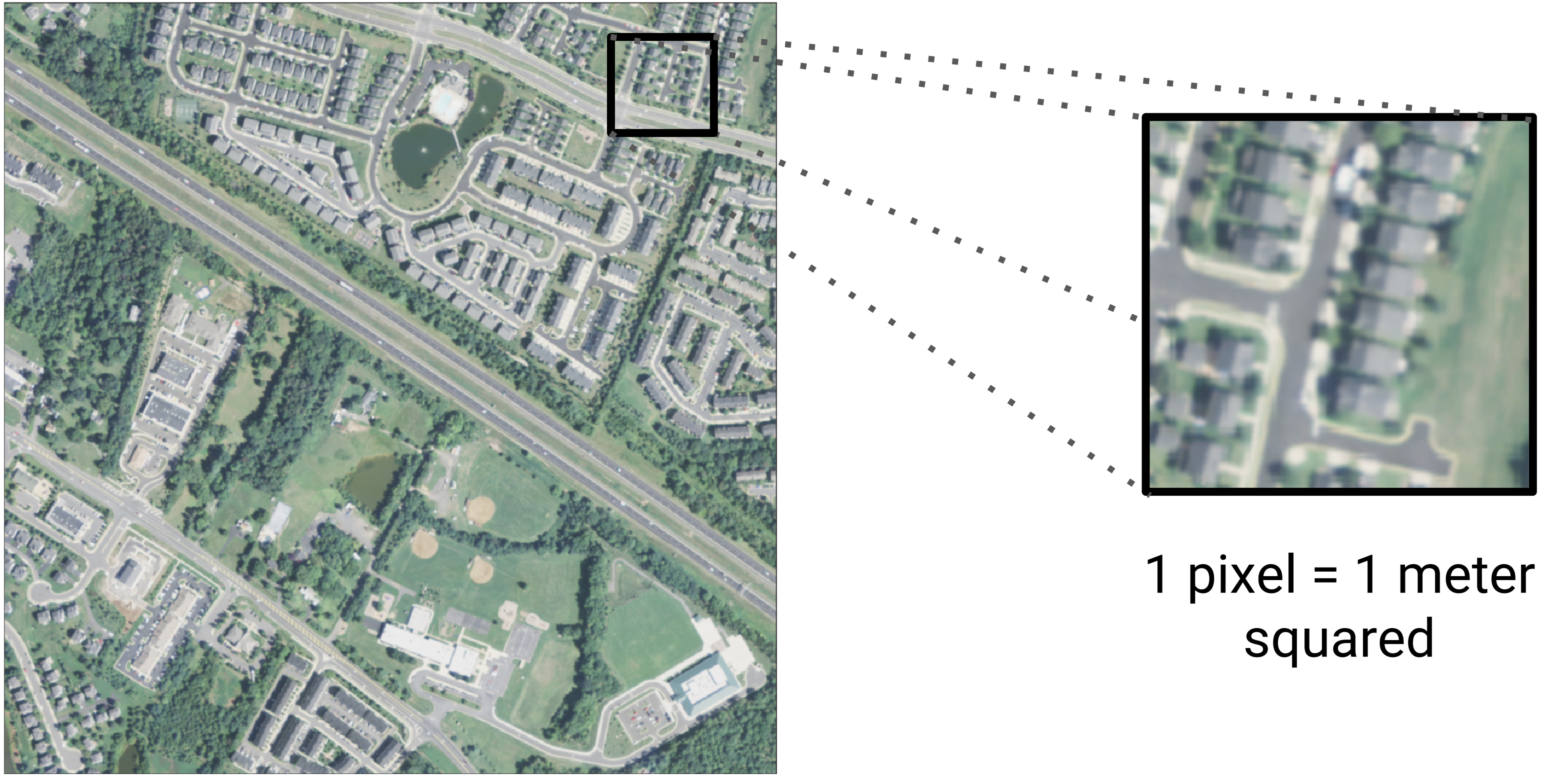
Georgia Tech College of Computing

USC Center for Artificial Intelligence in Society

UTEP

Microsoft

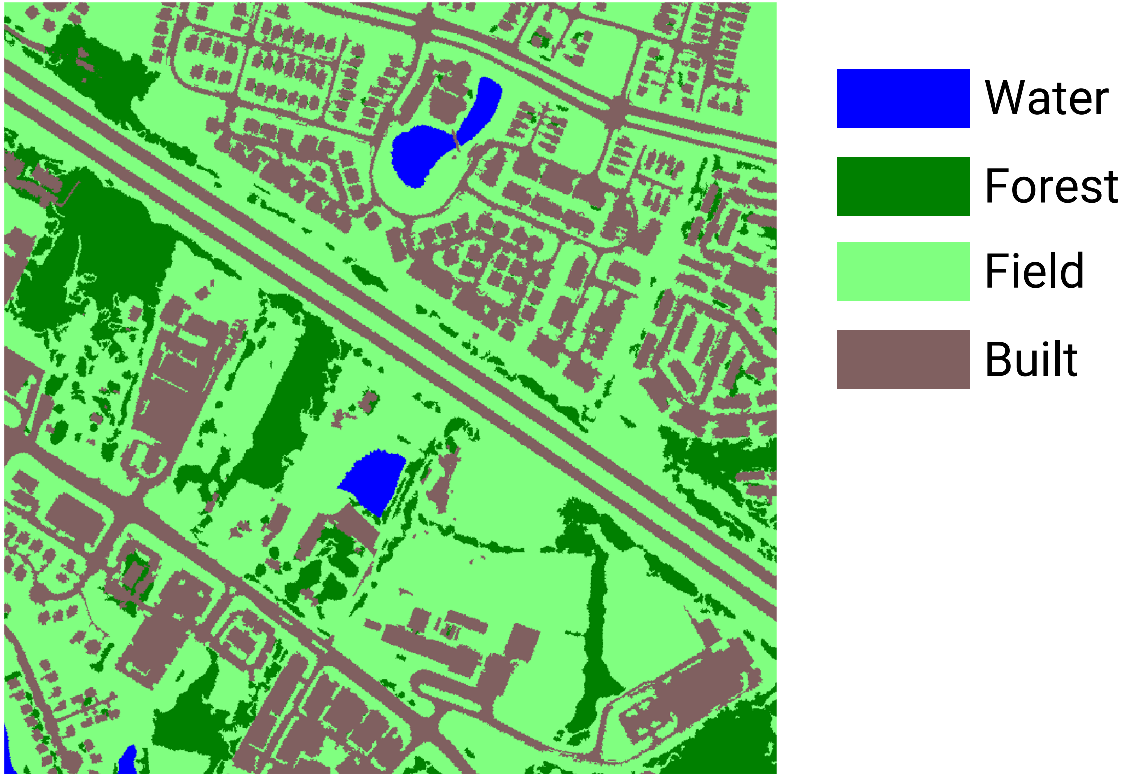
What is the land cover mapping problem?



1 pixel = 1 meter squared

High-Resolution Satellite/Aerial Imagery

NAIP 2013/2014



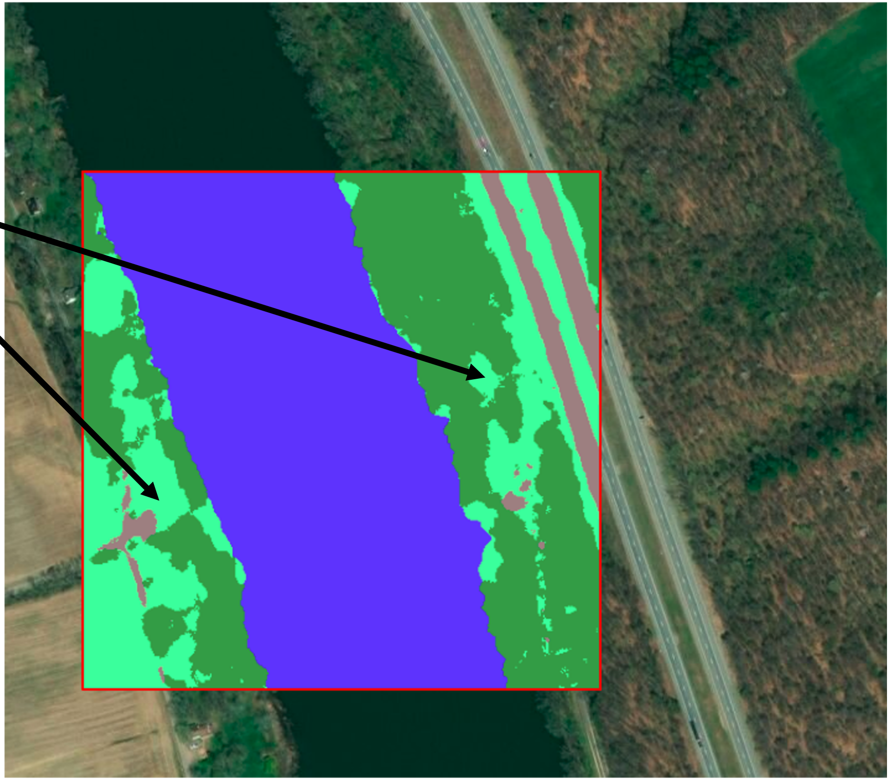
High-Resolution Land Cover Map

Chesapeake Conservancy

Why do we need high resolution land cover maps?

E.g. to help inform conservation actions

Riparian buffers



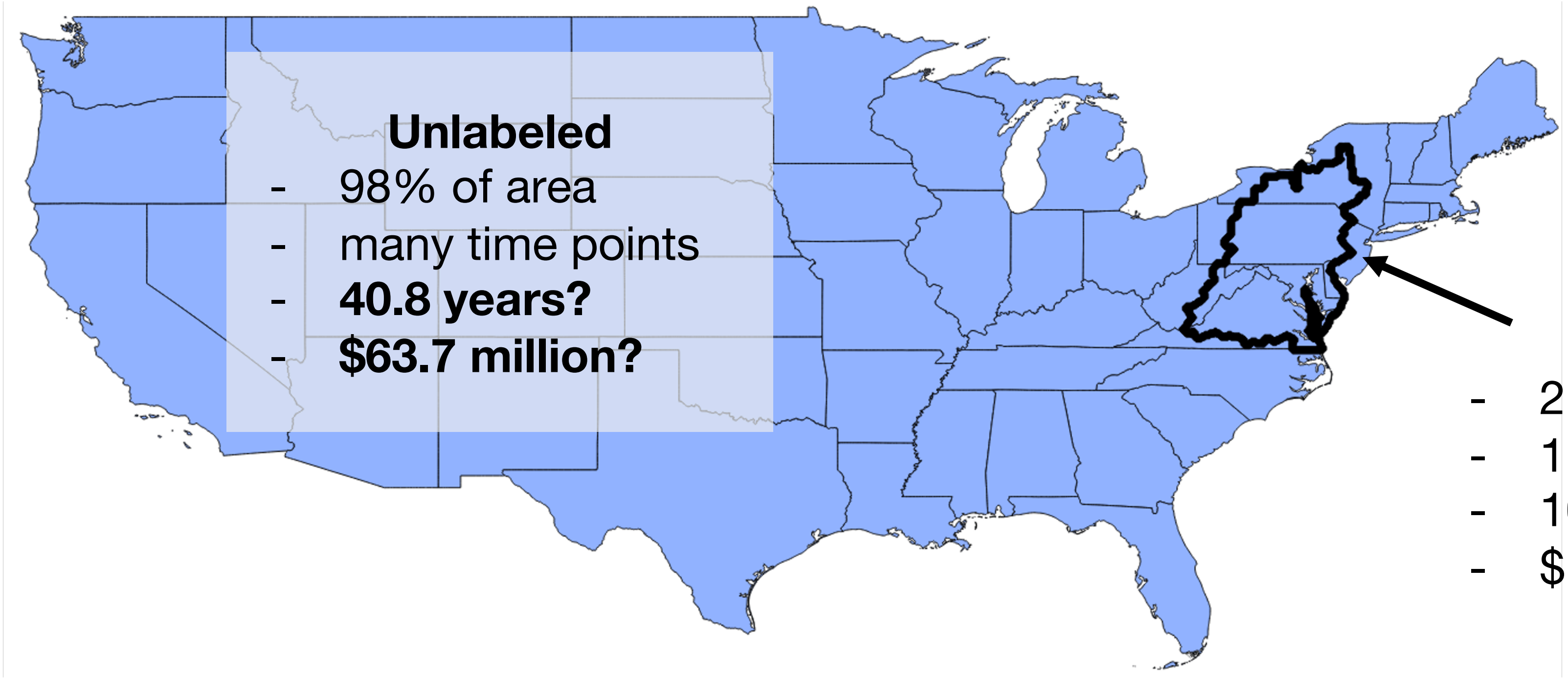
“[The Chesapeake Conservancy] **leverages** the combination of the enhanced flow path data and **high-resolution land cover data** to **identify** opportunity areas for planting **riparian forest buffers** within a specified distance of the flow paths.”

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

But…

(Semi-) Manual labeling is expensive

Chesapeake Conservancy



Unlabeled

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

Labeled

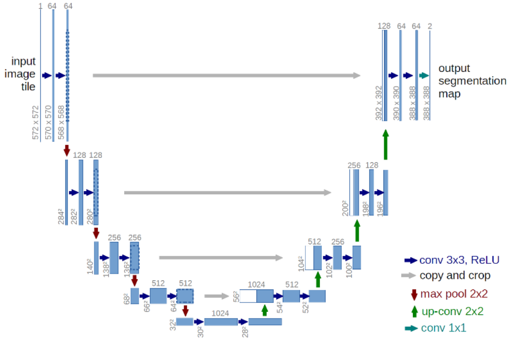
- 2% of area
- 1 time point
- 10 months
- $1.3 million

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

Deep learning approach to land cover mapping

High-resolution input

CNN



High-resolution predictions

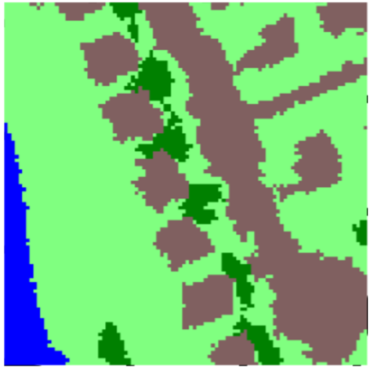
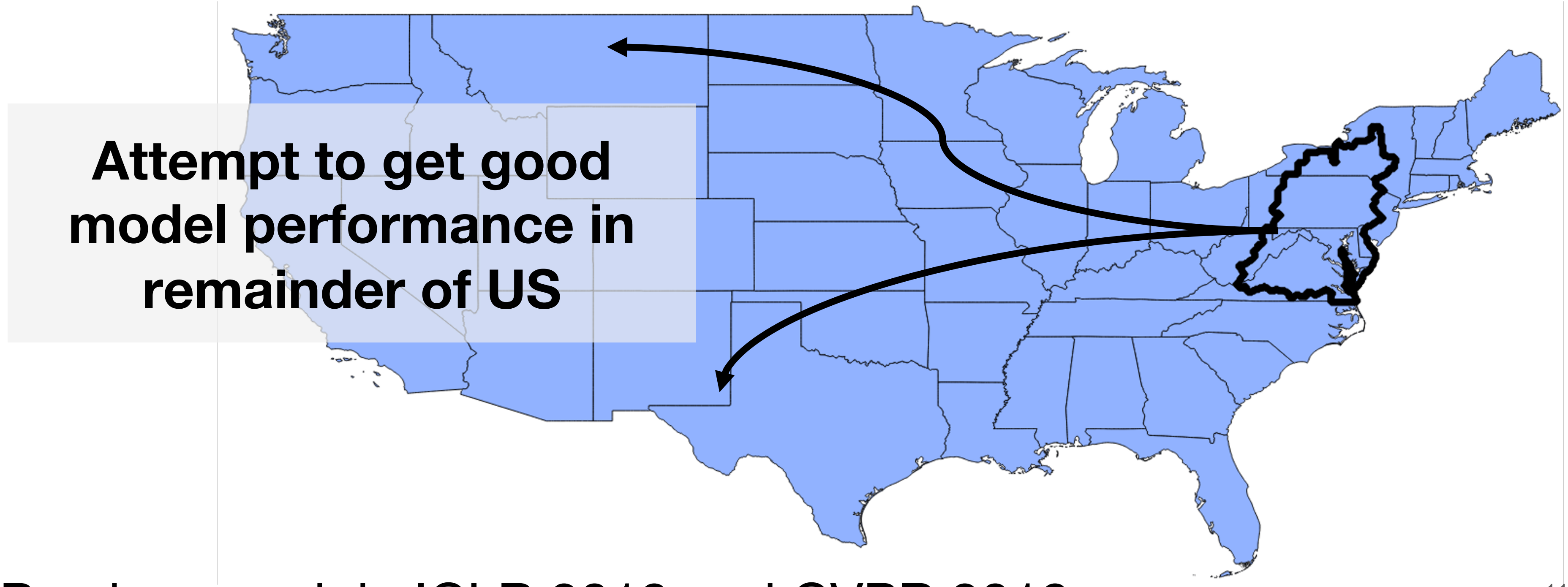
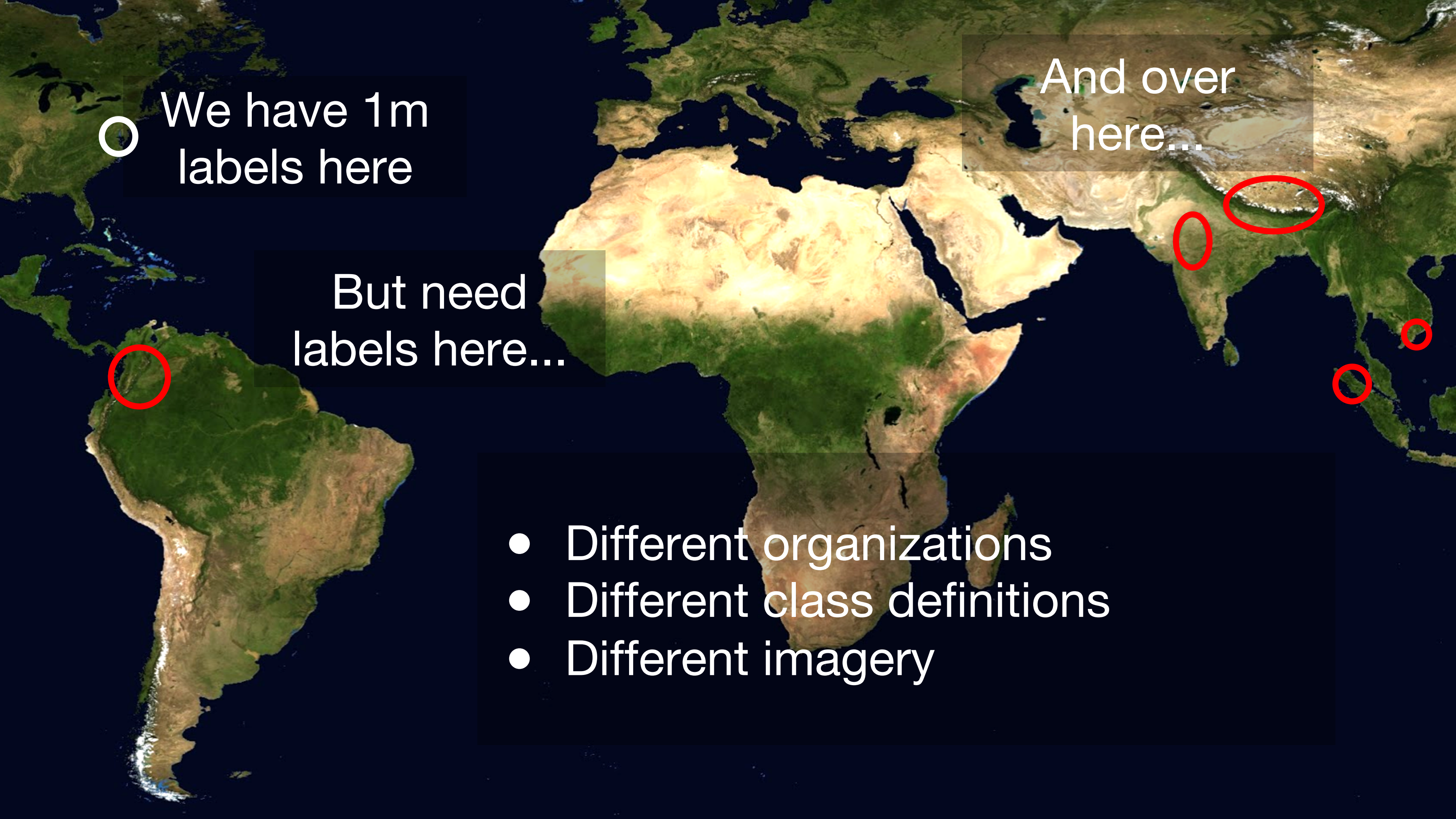


Image from: "U-net: Convolutional networks for biomedical image segmentation."

Problems in generalization



Previous work in ICLR 2019 and CVPR 2019



We have 1m labels here

But need labels here...

And over here...

- Different organizations
- Different class definitions
- Different imagery

Potential Approaches

1. Revisit assumptions
   - Try different modeling approaches
   - Retrain model with different hyperparameters
   - Retrain model with different data
   - ...

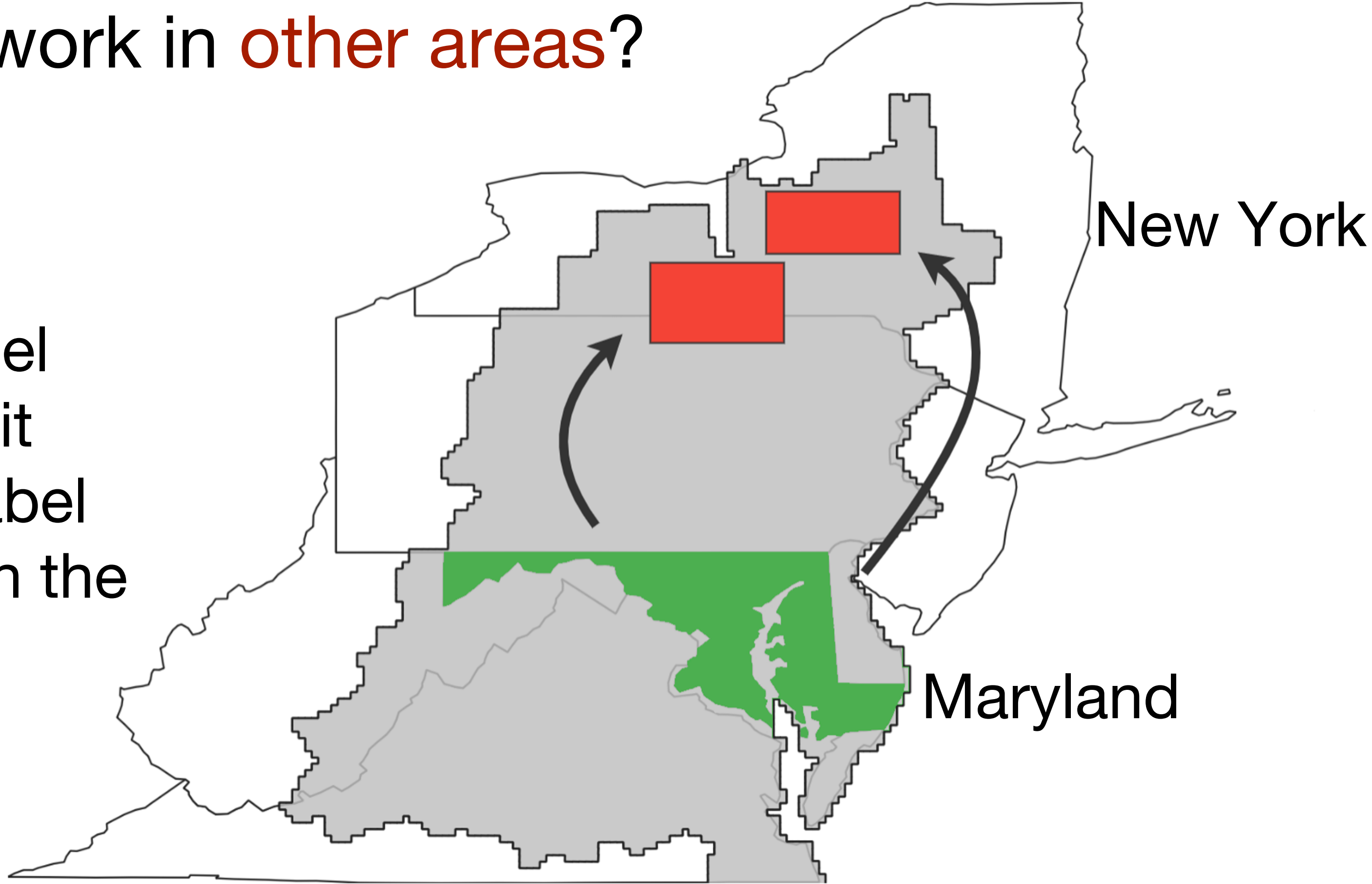
Local stakeholders **can not** do this (not scalable)

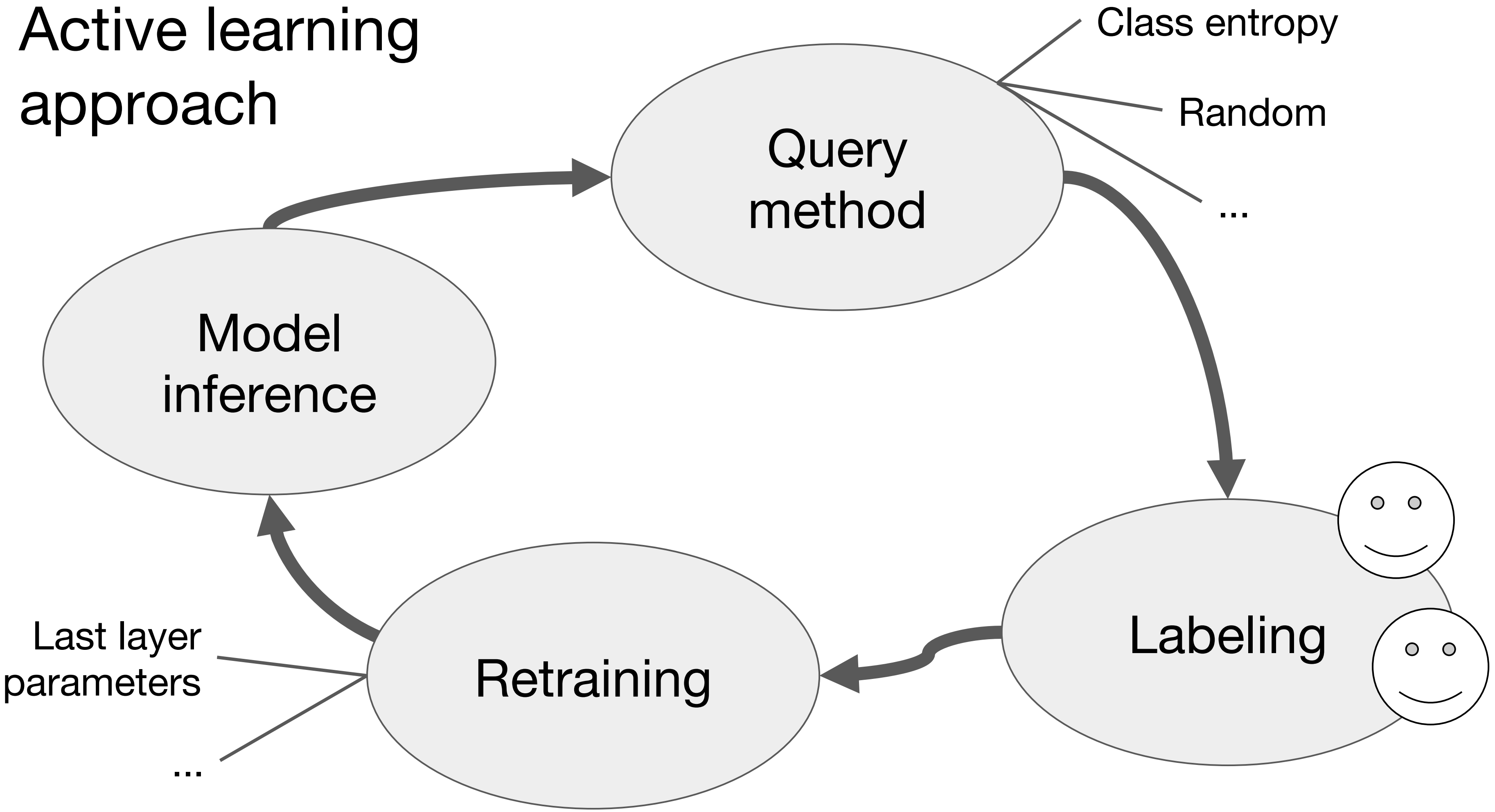
2. Fine-tune existing model with new data
   - Query labelers for new data
   - Adapt model accordingly

Local stakeholders **can** do this (scalable)

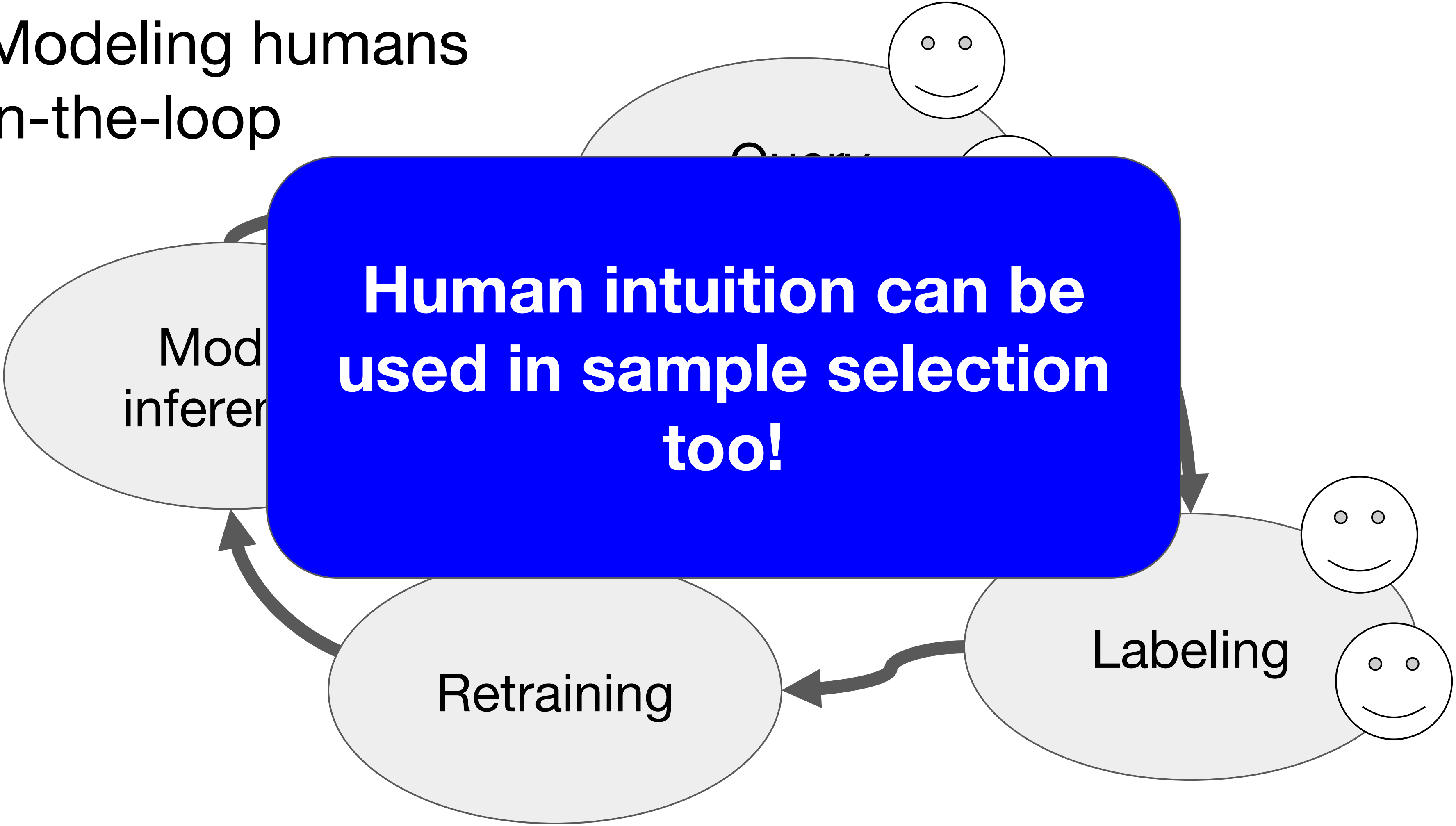
How can models trained in one area be quickly adapted to work in other areas?

Assumption:
- We have an existing model
- We can solicit humans to label data points in the other areas




Active learning
approach
Query
method
Class entropy
Random
...
Model
inference
Labeling
Retraining
Last layer
parameters
...

Modeling humans in-the-loop



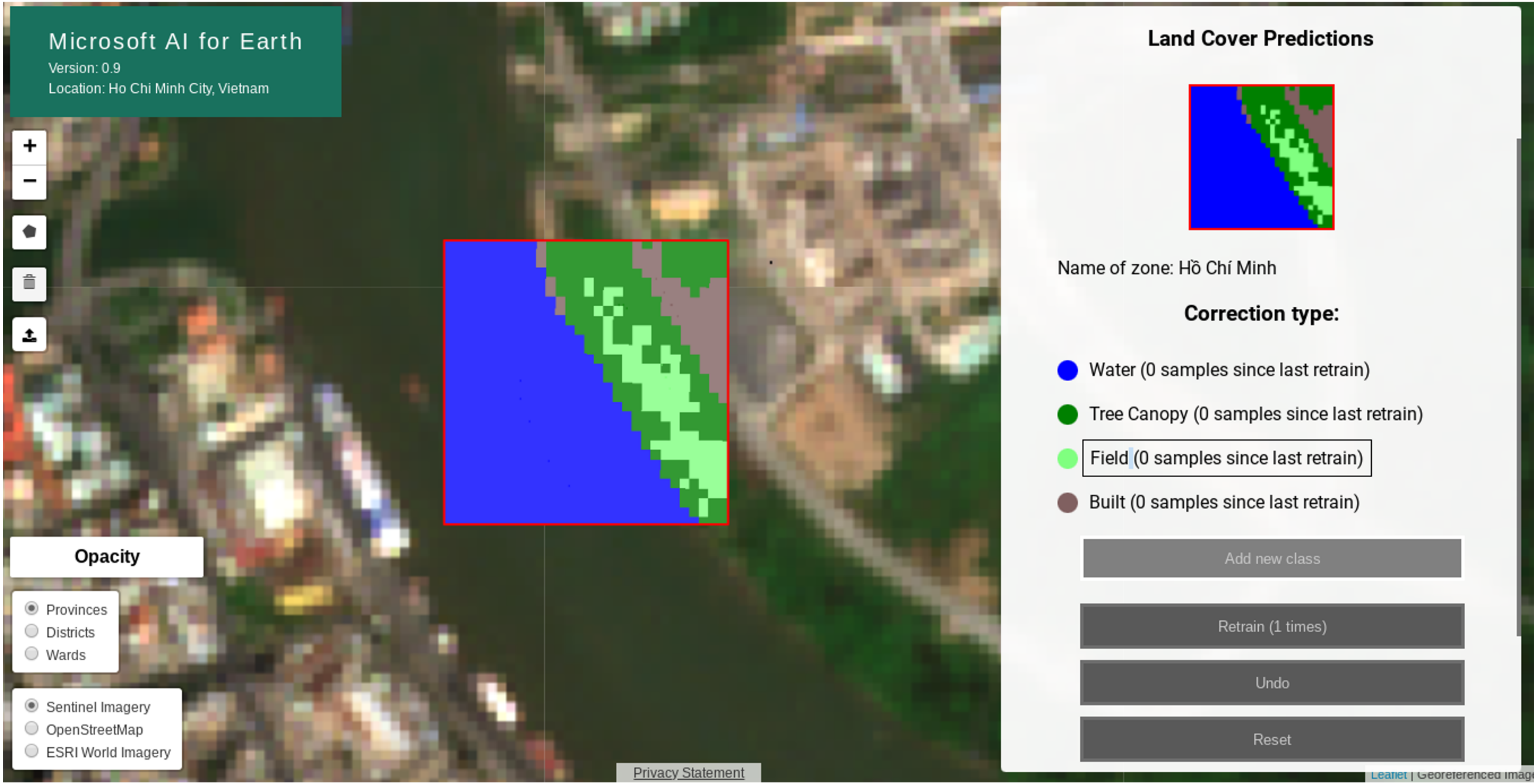
Human intuition can be used in sample selection too!

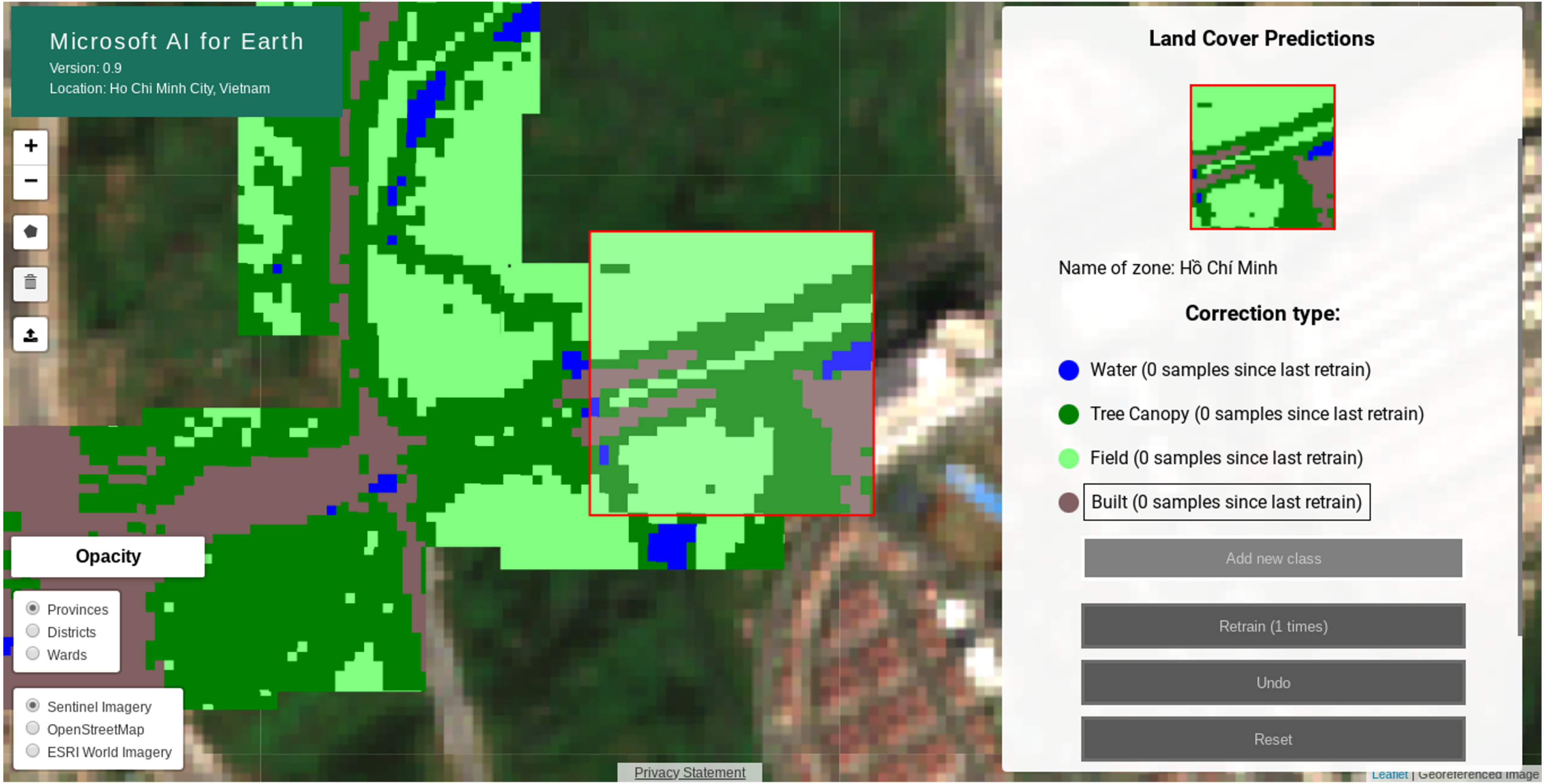
Implementation of humans-in-the-loop

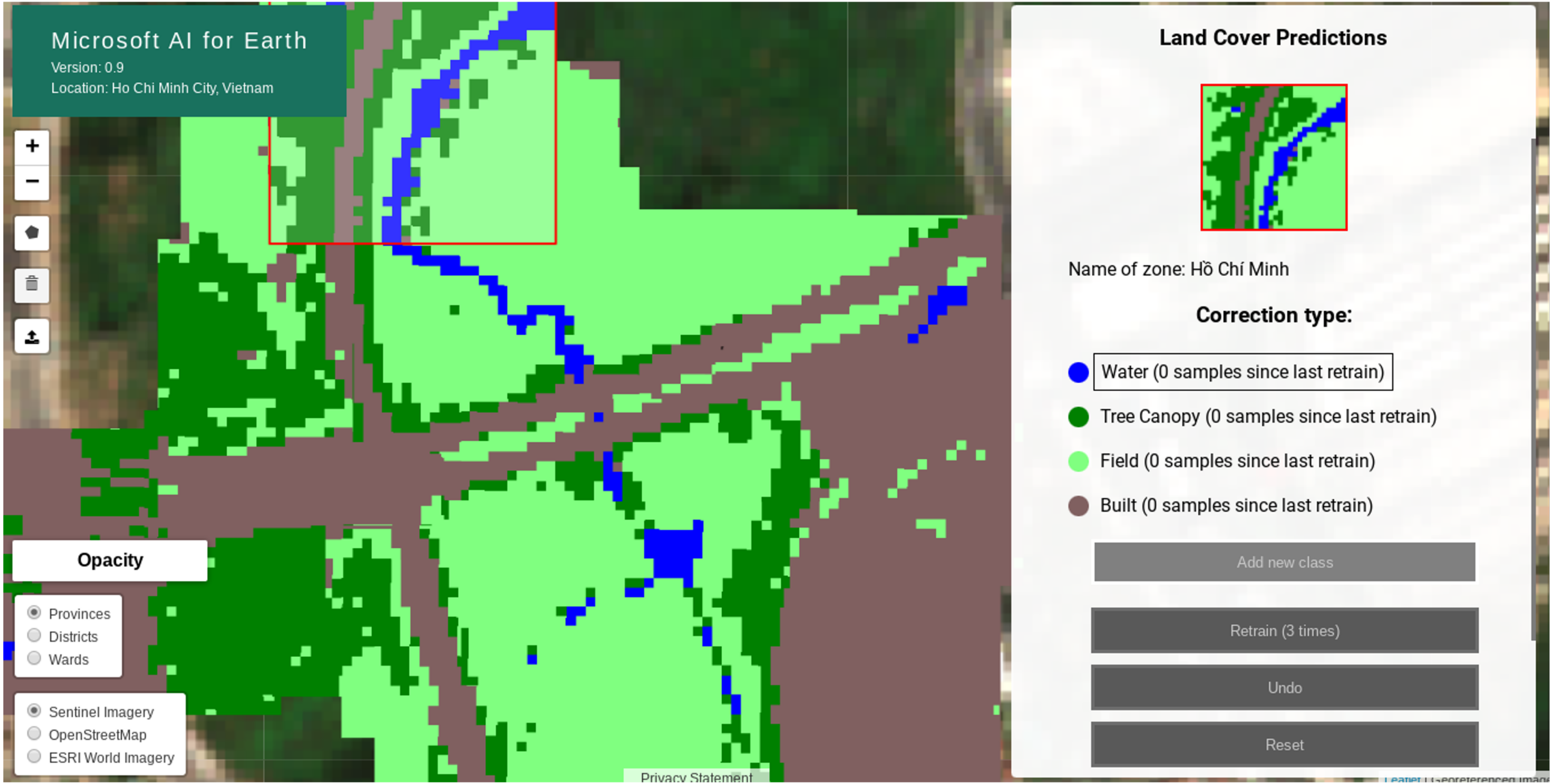
http://msrcalebubuntu1.eastus.cloudapp.azure.com:8080/

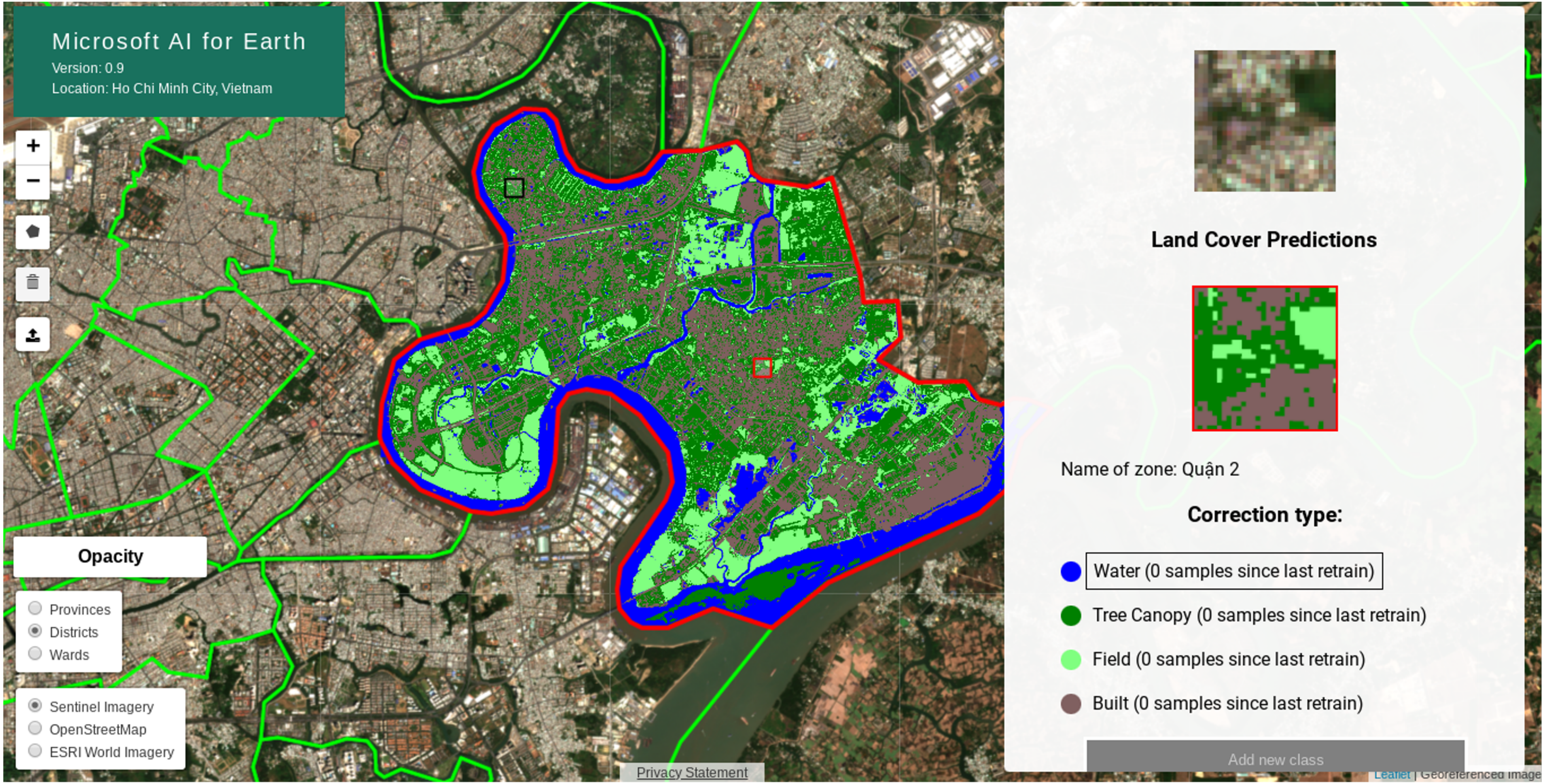

Microsoft AI for Earth
Version: 0.9
Location: Ho Chi Minh City, Vietnam
Opacity
Provinces
Districts
Wards
Sentinel Imagery
OpenStreetMap
ESRI World Imagery
Privacy Statement
Land Cover Predictions
Name of zone: Hồ Chí Minh
Correction type:
Water (0 samples since last retrain)
Tree Canopy (0 samples since last retrain)
Field (0 samples since last retrain)
Built (0 samples since last retrain)
Add new class

Microsoft AI for Earth
Version: 0.9
Location: Ho Chi Minh City, Vietnam
Opacity
Provinces
Districts
Wards
Sentinel Imagery
OpenStreetMap
ESRI World Imagery
Privacy Statement
Land Cover Predictions
Name of zone: Hồ Chí Minh
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)
Add new class
Retrain (1 times)
Undo
Reset
Leaflet | Georeferenced Image


Microsoft AI for Earth
Version: 0.9
Location: Ho Chi Minh City, Vietnam
Opacity
Provinces
Districts
Wards
Sentinel Imagery
OpenStreetMap
ESRI World Imagery
Privacy Statement
Land Cover Predictions
Name of zone: Hồ Chí Minh
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)
Add new class
Retrain (1 times)
Undo
Reset


Microsoft AI for Earth
Version: 0.9
Location: Ho Chi Minh City, Vietnam
Opacity
Provinces
Districts
Wards
Sentinel Imagery
OpenStreetMap
ESRI World Imagery
Privacy Statement
Land Cover Predictions
Name of zone: Hồ Chí Minh
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)
Add new class
Retrain (1 times)
Undo
Reset


Microsoft AI for Earth
Version: 0.9
Location: Ho Chi Minh City, Vietnam
Opacity
Provinces
Districts
Wards
Sentinel Imagery
OpenStreetMap
ESRI World Imagery
Privacy Statement
Land Cover Predictions
Name of zone: Hồ Chí Minh
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)
Add new class
Retrain (3 times)
Undo
Reset


Microsoft AI for Earth
Version: 0.9
Location: Ho Chi Minh City, Vietnam
Opacity
Provinces
Districts
Wards
Sentinel Imagery
OpenStreetMap
ESRI World Imagery
Privacy Statement
Land Cover Predictions
Name of zone: Quận 2
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)
Add new class

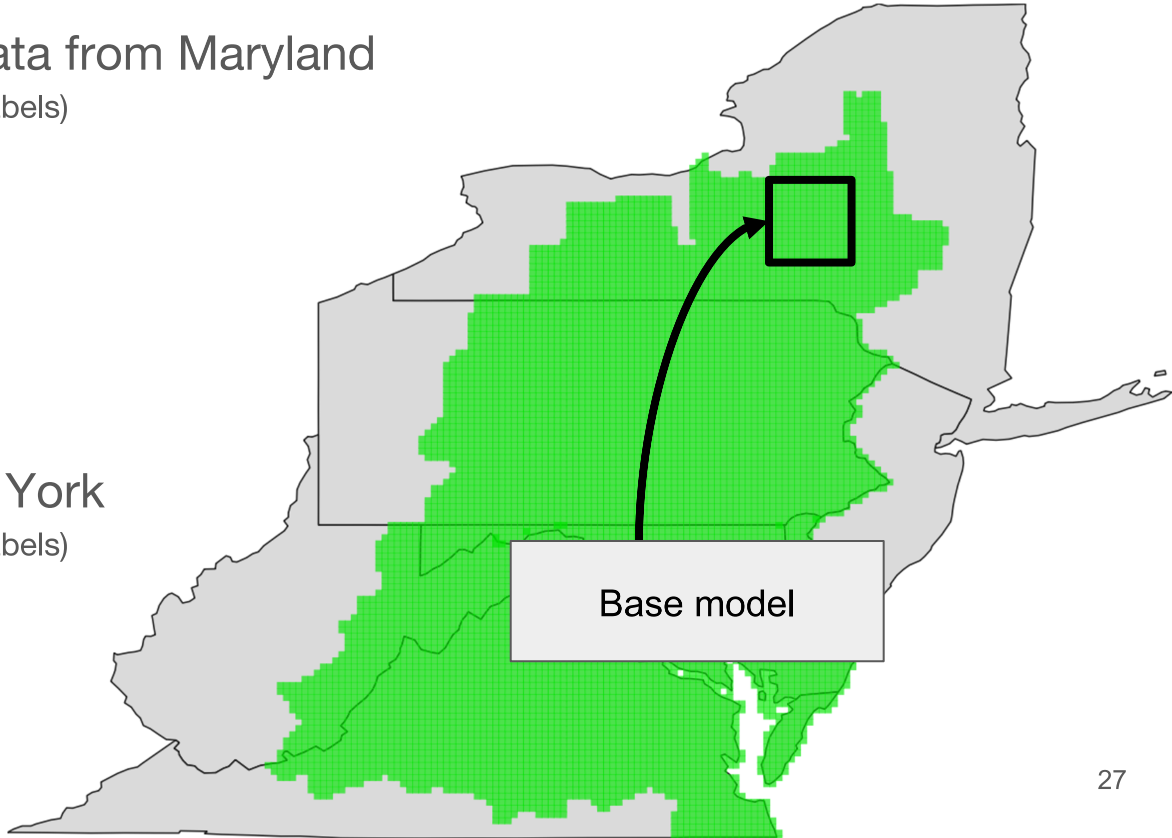
Experimental Setup

Base UNet model trained on data from Maryland

(where we have high-resolution ground truth labels)

4 different 84km^2 *areas* in New York

(where we have high-resolution ground truth labels)



Experiment Setup

- Offline study
  - Compare a variety of {**active learning**} x {**fine-tuning methods**} for adapting a model to a new area
- Online study with crowdsourced workers
  - Compare best(ish) active learning strategy against human labelers in our tool

Methods - All

Query methods:

- Random
- Entropy (where model is uncertain about the class)
- Min-margin (where model is uncertain about the class)
- Mistakes (where model makes mistakes)
- Human (where a human labeler wants)

Fine-tuning methods:

- Last-*k*-layers
- Group norm parameters
- Dropout

Which combination of **query method** and **fine-tuning** method is best?

Methods - Offline study

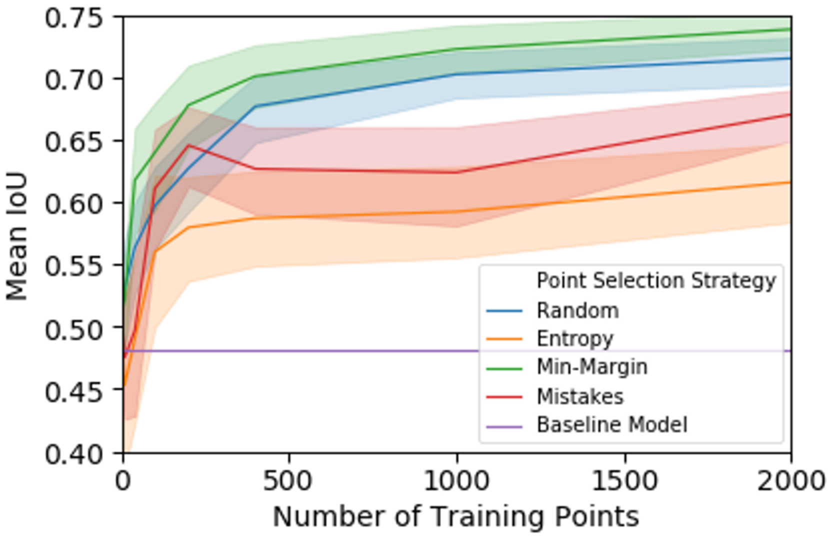
Query methods:

- Random
- Entropy (where model is uncertain about the class)
- Min-margin (where model is uncertain about the class)
- Mistakes (where model makes mistakes)
- ~~Human (where a human labeler wants)~~

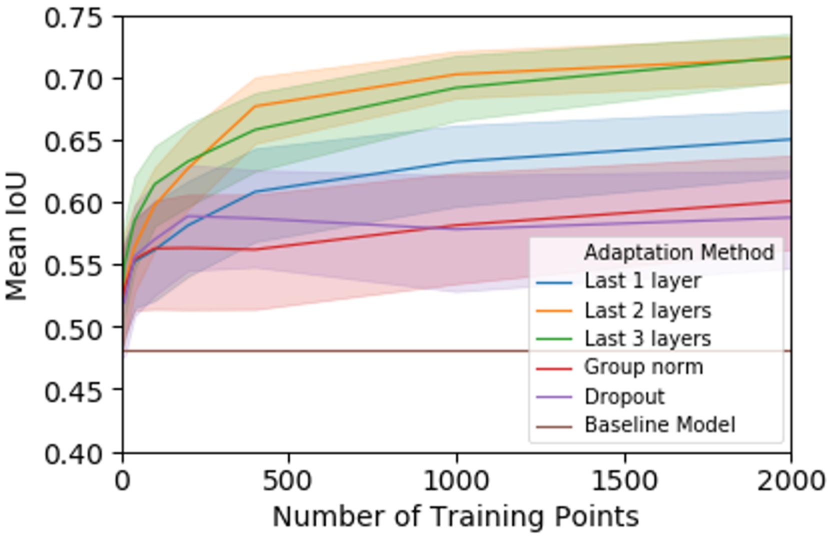
Fine-tuning methods:

- Last-*k*-layers
- Group norm parameters
- Dropout

Results - Offline study



With **Last 2 layers** fine-tuning method



With **Random** query method

- All methods are showing improvements with additional points added
- Random and Min-Margin are the best performing query methods
- Last-k-layers is the best performing fine-tuning method

Methods - Online study

Query methods:

- Random
- ~~Entropy (where model is uncertain about the class)~~
- ~~Min-margin (where model is uncertain about the class)~~
- ~~Mistakes (where model makes mistakes)~~
- Human (where a human labeler wants)

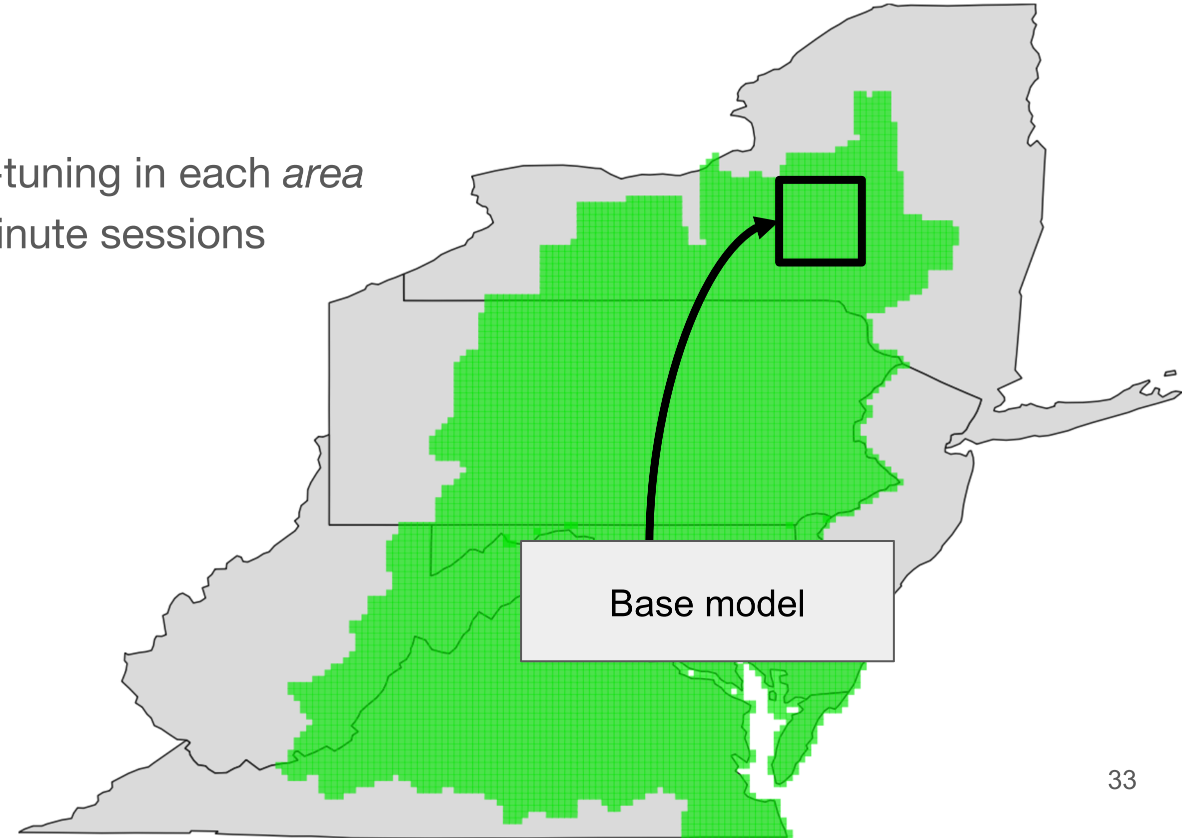
Fine-tuning methods:

- Last-*{1,2}*-layers
- ~~Group norm parameters~~
- ~~Dropout~~

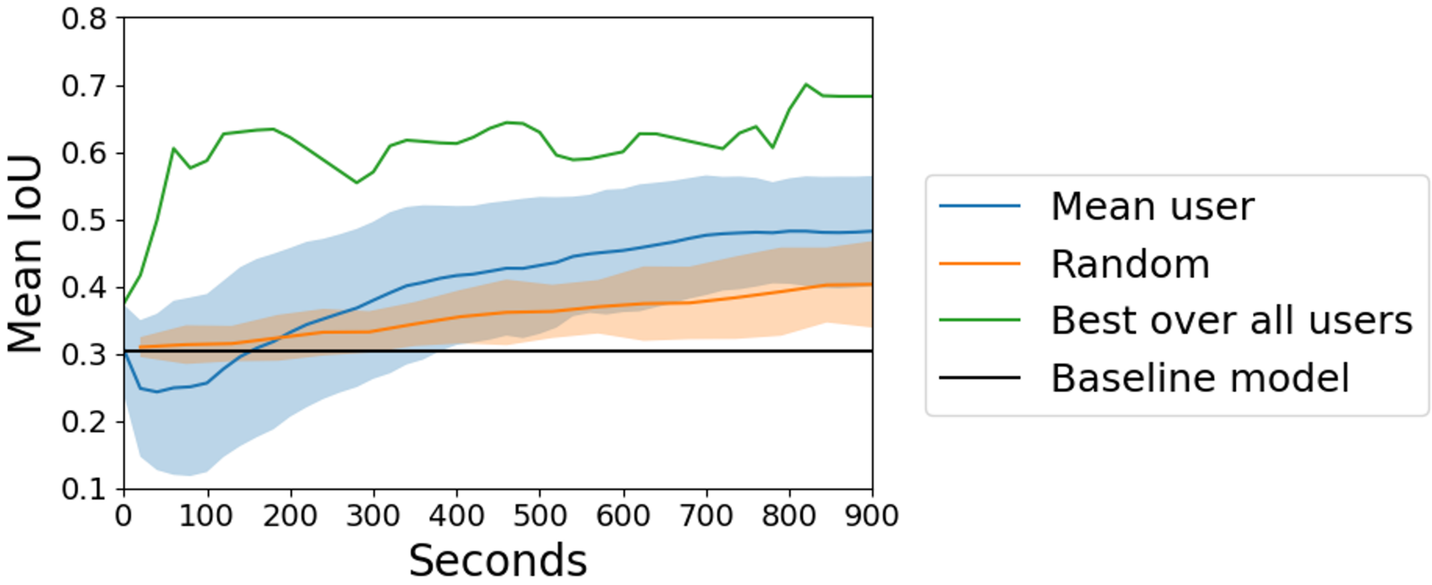
Experimental Setup - Online study

For a Human

- Randomly order the 4 *areas*
- User spends 15 minutes fine-tuning in each *area*
- Model is reset between 15 minute sessions



Results - Online study

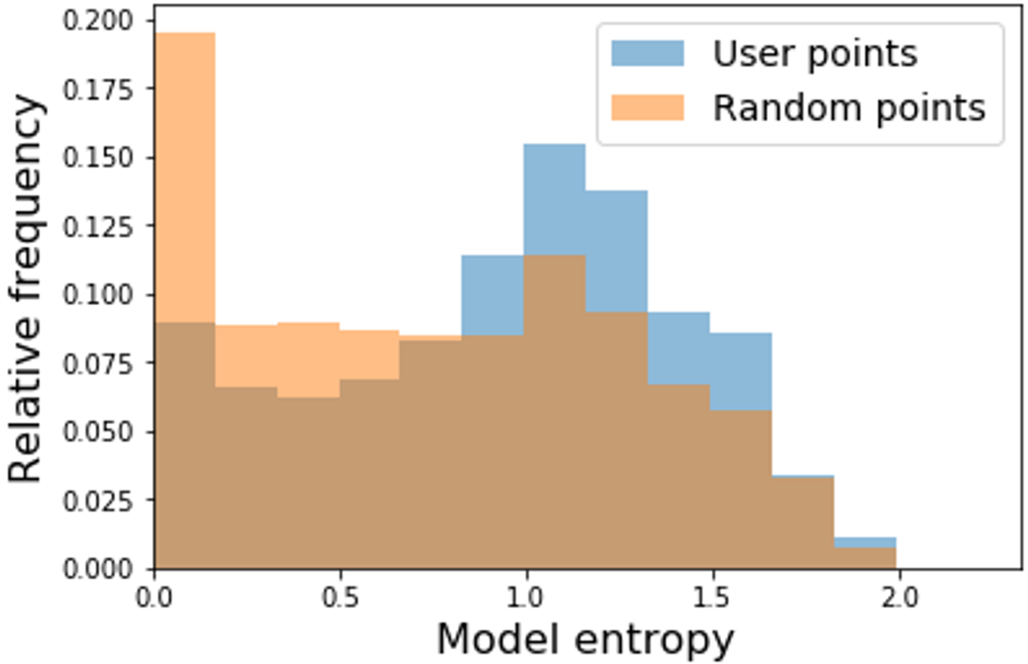


- On average users are outperforming *Random* selection of fine-tuning points
- Some users are much better

User behavior

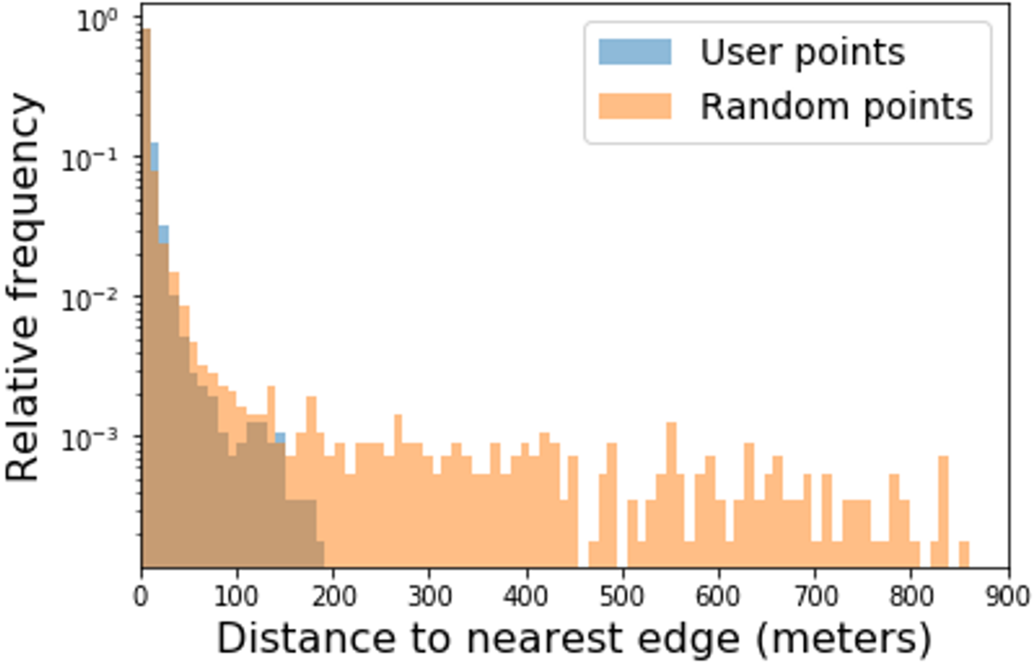
Users pick more points from mid-model entropy ranges

Users pick fewer points from low-model entropy ranges



User behavior

Users always pick points that are close to an edge in the imagery



Summary

- Proposed modeling human-in-the-loop methods in an active learning framework
- Compared different query methods and fine-tuning methods for adapting land cover models to new areas
- Performed an online study comparing Human query method to Random selection
- We find that users outperform random selection and behave distinctively different from other query strategies
- Local stakeholders can use our interface and methodology to tune existing models to new areas that they care about[*]

People / Papers / Code / Data

https://aka.ms/landcovermapping

Publications

- *Label Super-Resolution Networks*. ICLR 2019.
- *Large Scale High-Resolution Land Cover Mapping with Multi-Resolution Data*. CVPR 2019.
- *Human-Machine Collaboration for Fast Land Cover Mapping*. AAAI 2020.