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
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-resolution” matters.

High res at 1m vs. Low res at 30m.
Chesapeake high res (HR) labels (1m)
National Aggricuture 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...

= 300M km²/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

## Projected market growth

<table>
<thead>
<tr>
<th>Service</th>
<th>2018</th>
<th>2023</th>
</tr>
</thead>
<tbody>
<tr>
<td>Geospatial analytics:</td>
<td>$41 billion</td>
<td>$86 billion</td>
</tr>
<tr>
<td>Cloud:</td>
<td>$272 billion</td>
<td>$623 billion</td>
</tr>
<tr>
<td>AI software:</td>
<td>$9.5 billion</td>
<td>$71 billion</td>
</tr>
<tr>
<td>Wine:</td>
<td>$108 billion</td>
<td>$450 billion</td>
</tr>
</tbody>
</table>
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
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

\[ \rightarrow \]

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

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

Cross Entropy (pixel-wise comparison)
Super-Resolution Loss

High-resolution input → High-resolution predictions

30m block - 1 label
Low-resolution ground truth

30m block - 900 predictions
High-resolution predictions

Super-Resolution Loss (block-wise comparison)
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

<table>
<thead>
<tr>
<th>Low resolution classes</th>
<th>High resolution classes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Water</td>
</tr>
<tr>
<td>Developed, Open Space</td>
<td>0%</td>
</tr>
<tr>
<td>Developed, Low Intensity</td>
<td>1%</td>
</tr>
<tr>
<td>Developed, Medium Intensity</td>
<td>1%</td>
</tr>
</tbody>
</table>
**Super-Resolution Loss**

*(block-wise comparison)*

<table>
<thead>
<tr>
<th>Low resolution classes</th>
<th>Water</th>
<th>Forest</th>
<th>Field</th>
<th>Built</th>
</tr>
</thead>
<tbody>
<tr>
<td>Developed, Open Space</td>
<td>0%</td>
<td>42%</td>
<td>46%</td>
<td>12%</td>
</tr>
<tr>
<td>Developed, Low Intensity</td>
<td>1%</td>
<td>30%</td>
<td>34%</td>
<td>35%</td>
</tr>
<tr>
<td>Developed, Medium Intensity</td>
<td>1%</td>
<td>14%</td>
<td>21%</td>
<td>64%</td>
</tr>
</tbody>
</table>

**Developed, Open Space:**
- **On Average** - Should contain 0% water labels, 42% forest labels, ... (+/- something)
Super-Resolution Loss (block-wise comparison)

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

<table>
<thead>
<tr>
<th>Low resolution classes</th>
<th>Water</th>
<th>Forest</th>
<th>Field</th>
<th>Built</th>
</tr>
</thead>
<tbody>
<tr>
<td>Developed, Open Space</td>
<td>0%</td>
<td>42%</td>
<td>46%</td>
<td>12%</td>
</tr>
<tr>
<td>Developed, Low Intensity</td>
<td>1%</td>
<td>30%</td>
<td>34%</td>
<td>35%</td>
</tr>
<tr>
<td>Developed, Medium Intensity</td>
<td>1%</td>
<td>14%</td>
<td>21%</td>
<td>64%</td>
</tr>
</tbody>
</table>

Low-resolution ground truth, High-resolution predictions
Super-Resolution Loss
(block-wise comparison)

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)

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

Ground Truth: Expected distribution

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

Compare with a differentiable distribution based measure
Super-Resolution Loss
(block-wise comparison)

Low-resolution ground truth

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

Ground Truth:
Expected distribution

High-resolution predictions

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, ...

<table>
<thead>
<tr>
<th>Low resolution classes</th>
<th>Water</th>
<th>Forest</th>
<th>Field</th>
<th>Built</th>
</tr>
</thead>
<tbody>
<tr>
<td>Developed, Open Space</td>
<td>0%</td>
<td>42%</td>
<td>46%</td>
<td>12%</td>
</tr>
<tr>
<td>Developed, Low Intensity</td>
<td>1%</td>
<td>30%</td>
<td>34%</td>
<td>35%</td>
</tr>
<tr>
<td>Developed, Medium Intensity</td>
<td>1%</td>
<td>14%</td>
<td>21%</td>
<td>64%</td>
</tr>
</tbody>
</table>

High resolution classes

Low-resolution ground truth

High-resolution predictions
Super-Resolution Loss
(block-wise comparison)

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

Ground Truth:
Expected distribution

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

Predicted distribution

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

Compare with a differentiable distribution based measure
Label super resolution

- Given ONLY pairs

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
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)

Consistency with NLCD:
- 0.0 - 0.1
- 0.1 - 0.2
- 0.2 - 0.3
- 0.3 - 0.4
- 0.4 - 0.5
- 0.5 - 0.6
- 0.6 - 0.7
- 0.7 - 0.8
- 0.8 - 0.9
- 0.9 - 1.0
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
Found in a small group of a dozen salmon; staying in deeper pools of the creek
Sightings of Haliaeetus (Sea eagles 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.
• 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
Remember to check how your fine-tuned model is performing locally.
Labelling tutorial

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
10 months
90-95% accuracy
Chesapeake Conservancy

$18.5 K*
925 hours*
89% accuracy*
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 challenge
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 challenge
Segmentation
Input tile
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