



Zoom to Location:

Place Types

Required Amenities:

Only places visited in the last

Show Legend

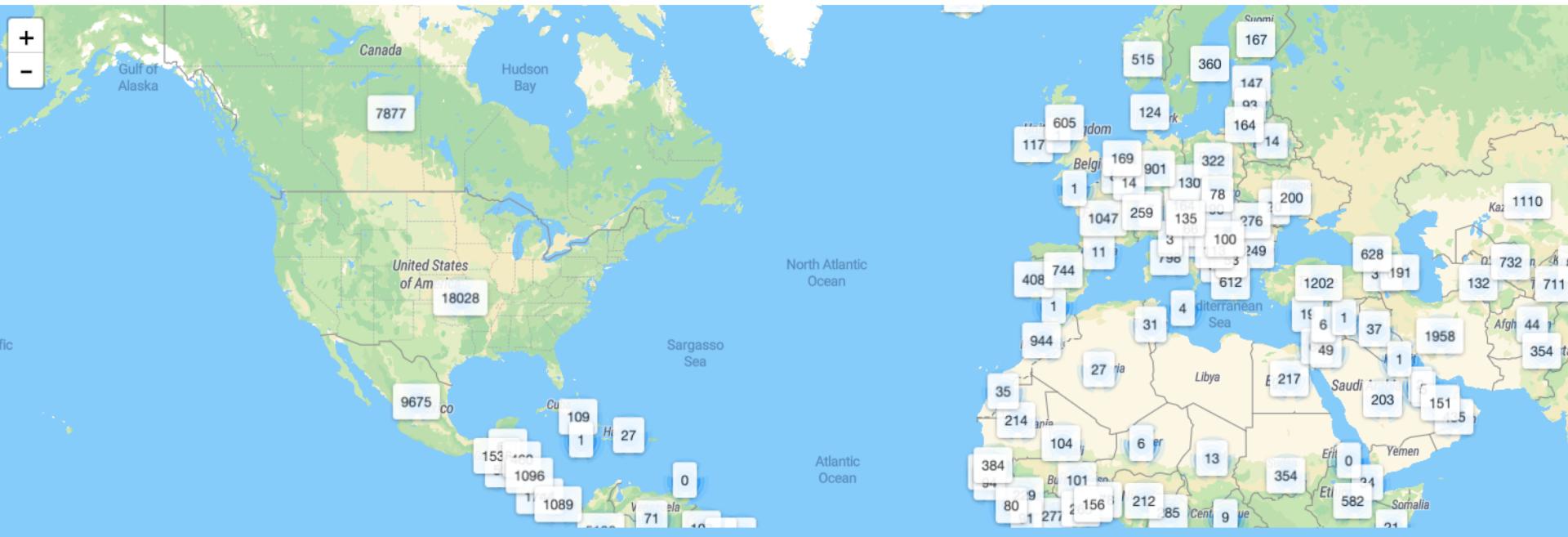
country, city or GPS coordinates

go

29 selected ▾

None selected ▾

None selected



Improved curation and classification of campsites with satellite imagery

Paul Tanger, Ph.D.

September 24, 2020

About me

- Background in plant genetics, bioinformatics
- Managed projects with USAID
- Focusing on my true love: data!



iOverlander: user contributed camping sites

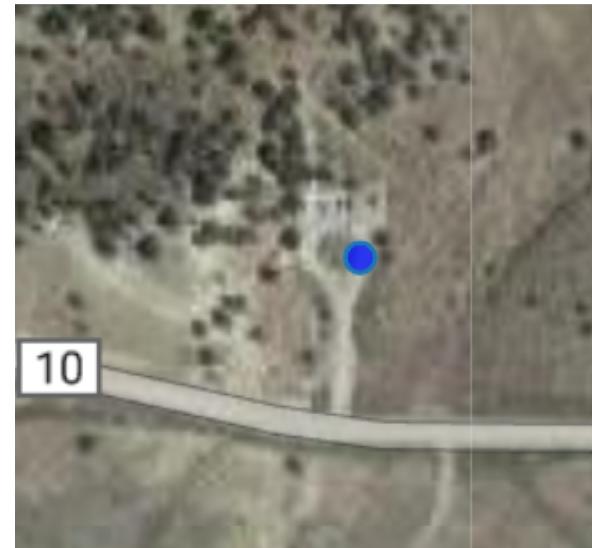
[Map](#) Place Details

 Ridgway Area Trails trailhead (RAT)
Informal Campsite
Open: YES
GPS: N 38.17434, W 107.73242
Altitude: 0 m
Contributed By: Paul
Distance: 352.4 km
Last Visited: 6/21/2019
Description:
Parking area for Ridgway area trailhead is BLM and stayed overnite and saw a couple others doing the same. There is a portapottie. Closed Dec-April 1. Quiet.

Place Website: <http://>

Amenities:



Motivation

[Map](#) Place Details

 Ridgway Area Trails trailhead (RAT)
Informal Campsite
Open: YES
GPS: N 38.17434, W 107.73242
Altitude: 0 m
Contributed By: Paul
Distance: 352.4 km
Last Visited: 6/21/2019
Description:
Parking area for Ridgway area trailhead is BLM and stayed overnite and saw a couple others doing the same. There is a portapottie. Closed Dec-April 1. Quiet.

Place Website: <http://>

Amenities:



- Talked about their challenges
- manually curate each site
- real impact from “illegal” sites and mis-categorized

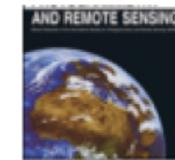
Given GPS, can we categorize site?

Development applications of image classification

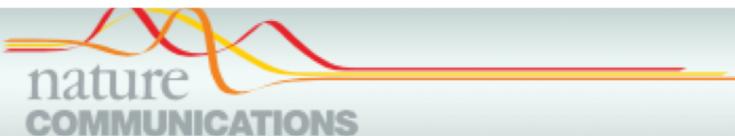


ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs



Detection of banana plants and their major diseases through aerial images
and machine learning methods: A case study in DR Congo and Republic
of Benin



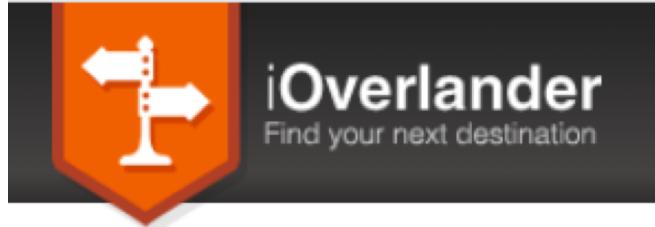
ARTICLE

<https://doi.org/10.1038/s41467-020-16185-w>

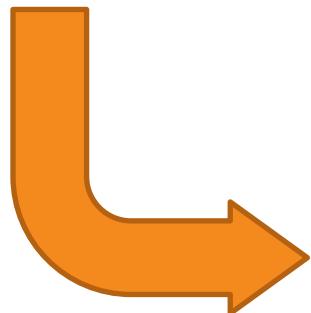
OPEN

Using publicly available satellite imagery
and deep learning to understand economic
well-being in Africa

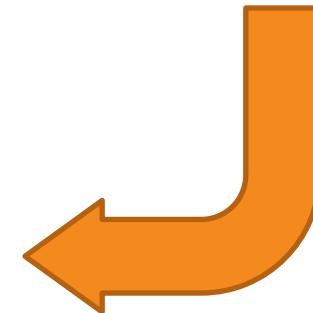
Data overview



Google APIs



Reverse geocode
Obtain satellite image



10,600+ sites in the US

Focus on site “Category”

-  Established Campground
-  Informal Campsite
-  Wild Camping

Category Description:
Campgrounds that advertise camping,
irrespective of amenities provided.

-  Established Campground
-  Informal Campsite
-  Wild Camping

Category Description:
A camping spot not linked to a business
or institution irrespective of amenities or
formal permissions.

Established Campground



Wild Camping



Supervised image classification with CNN



~ 8m trainable params

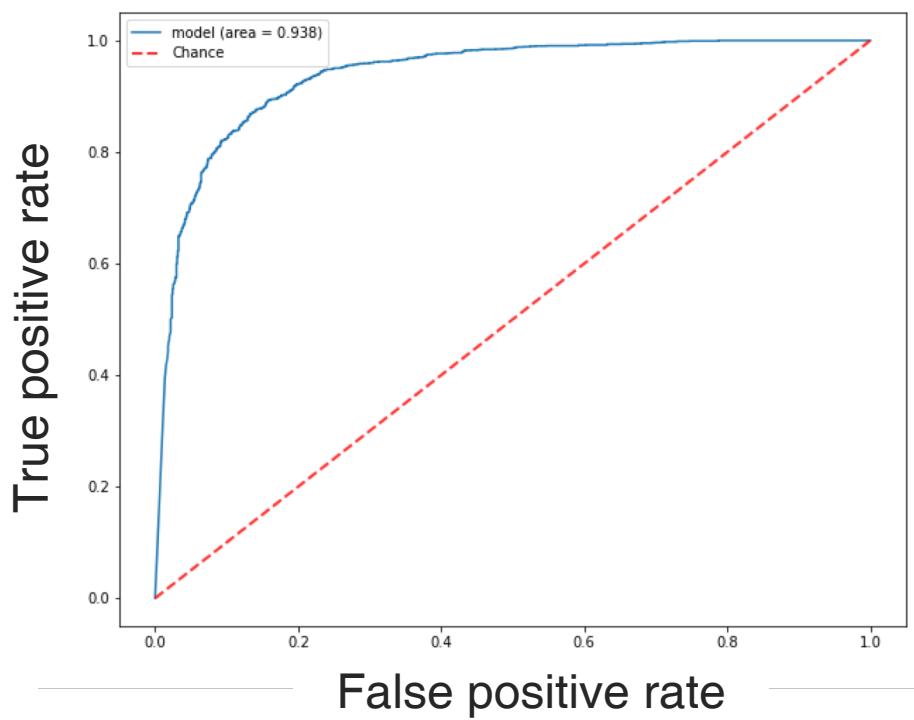
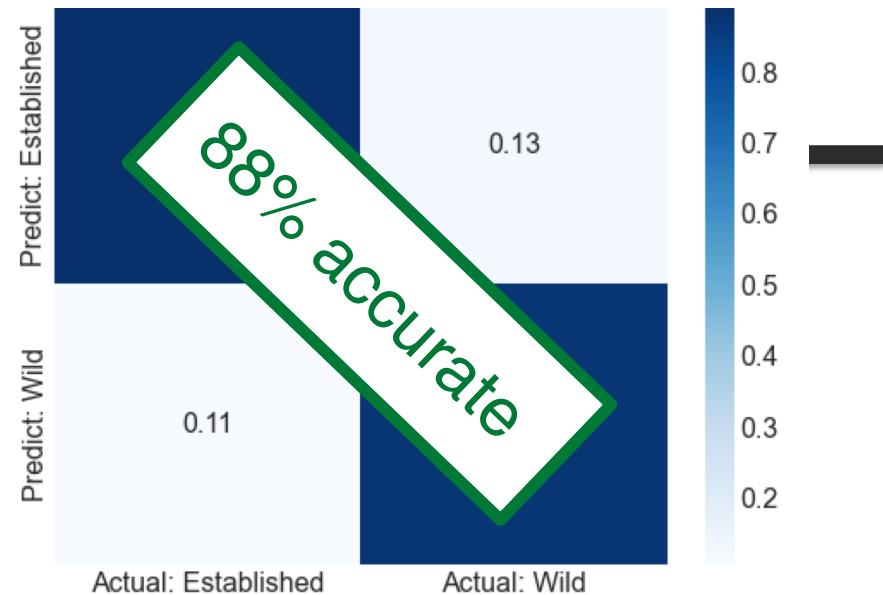
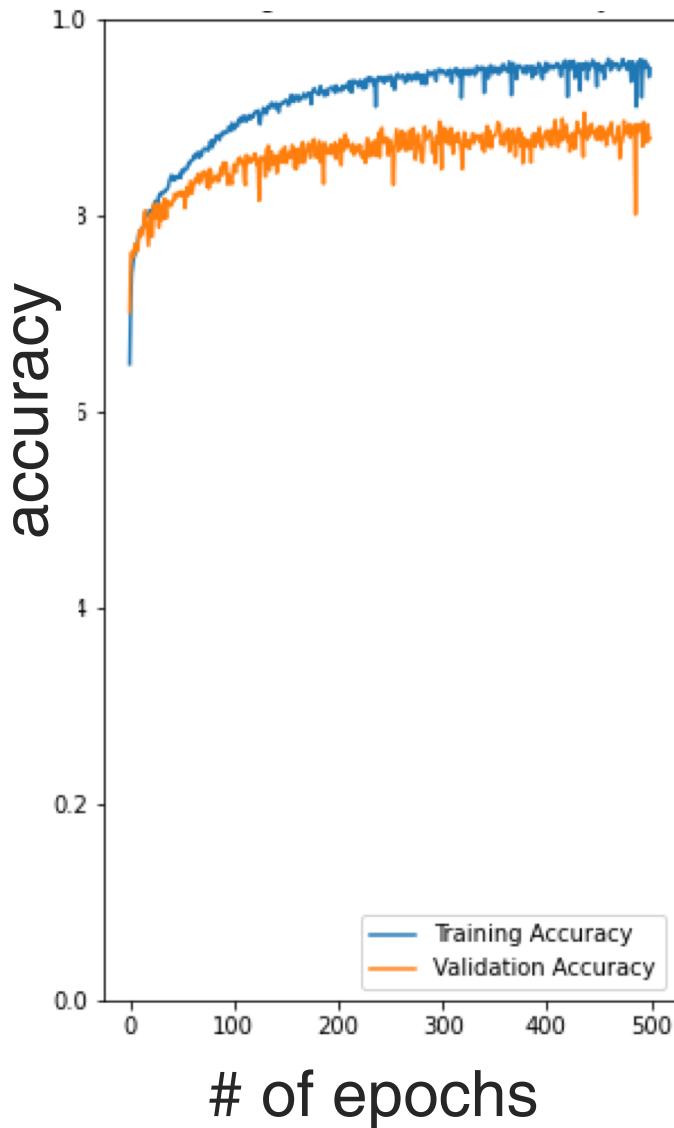


Established
campground

Wild
camping

Model validation

Validation data - orange



Can you tell the difference?

The model predicted this site: **Established Campground**
and the actual category label was: **Wild Camping**



The model predicted this site: **Wild Camping**
and the actual category label was: **Established Campground**



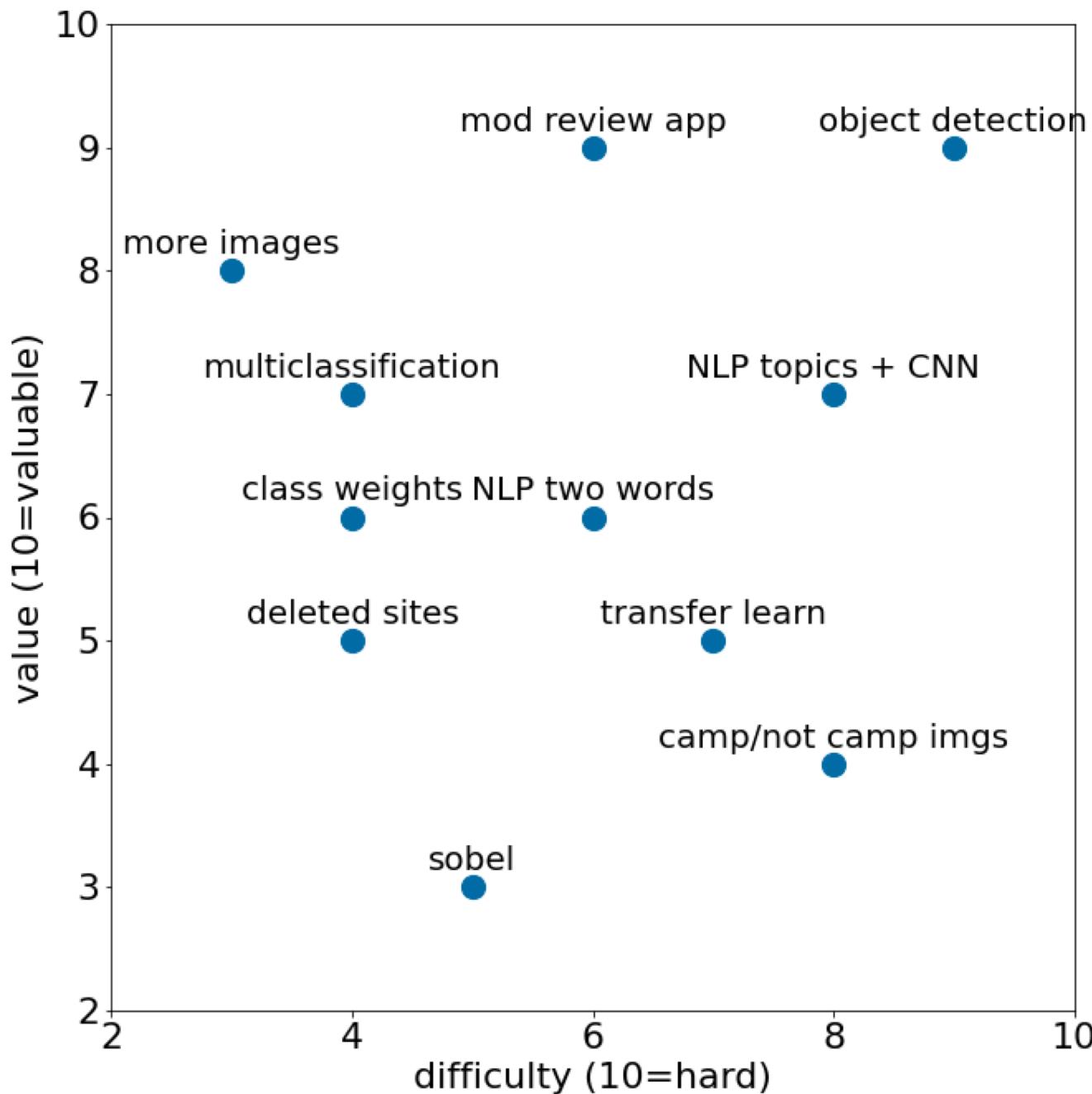
Can I improve:

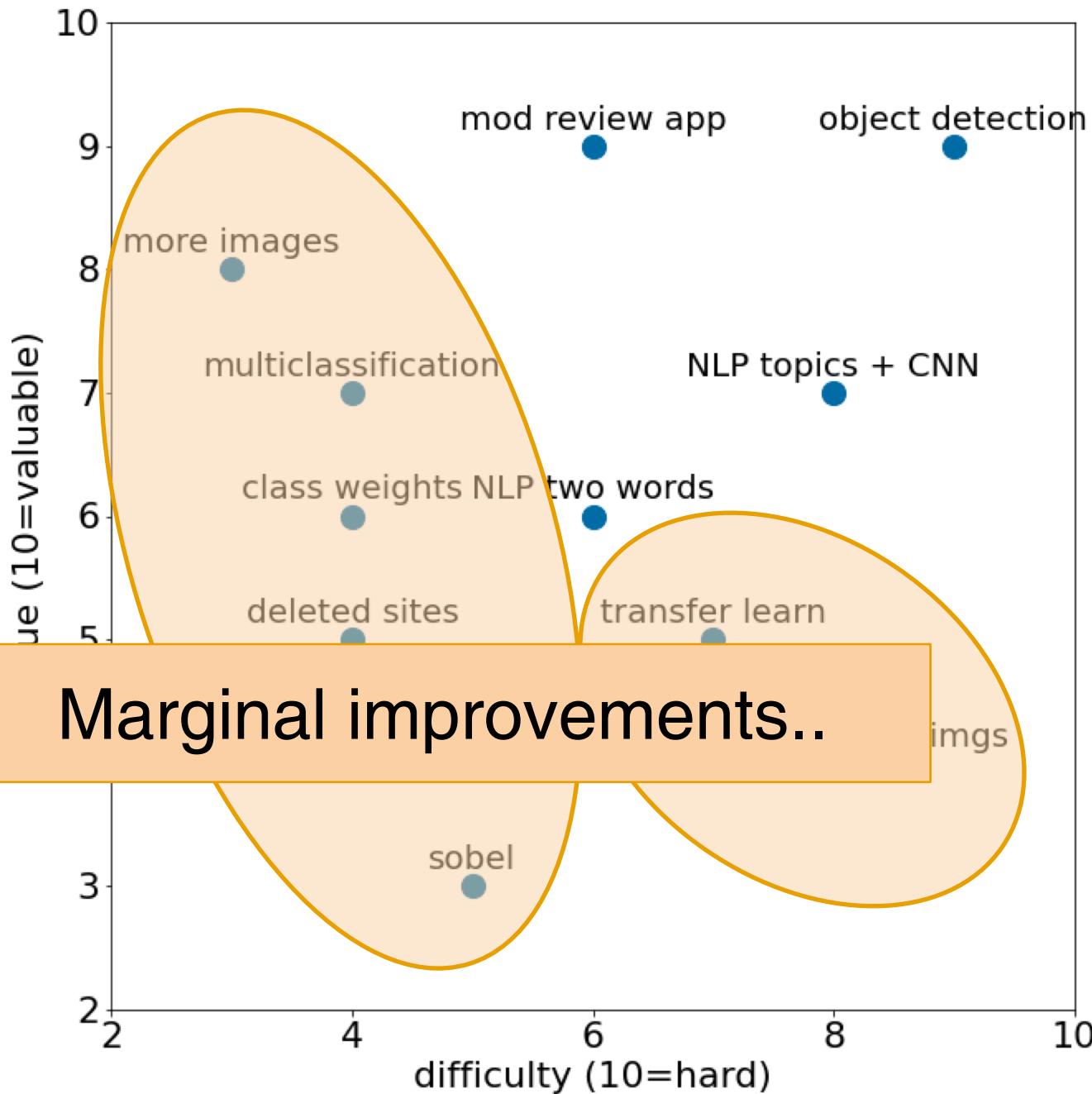


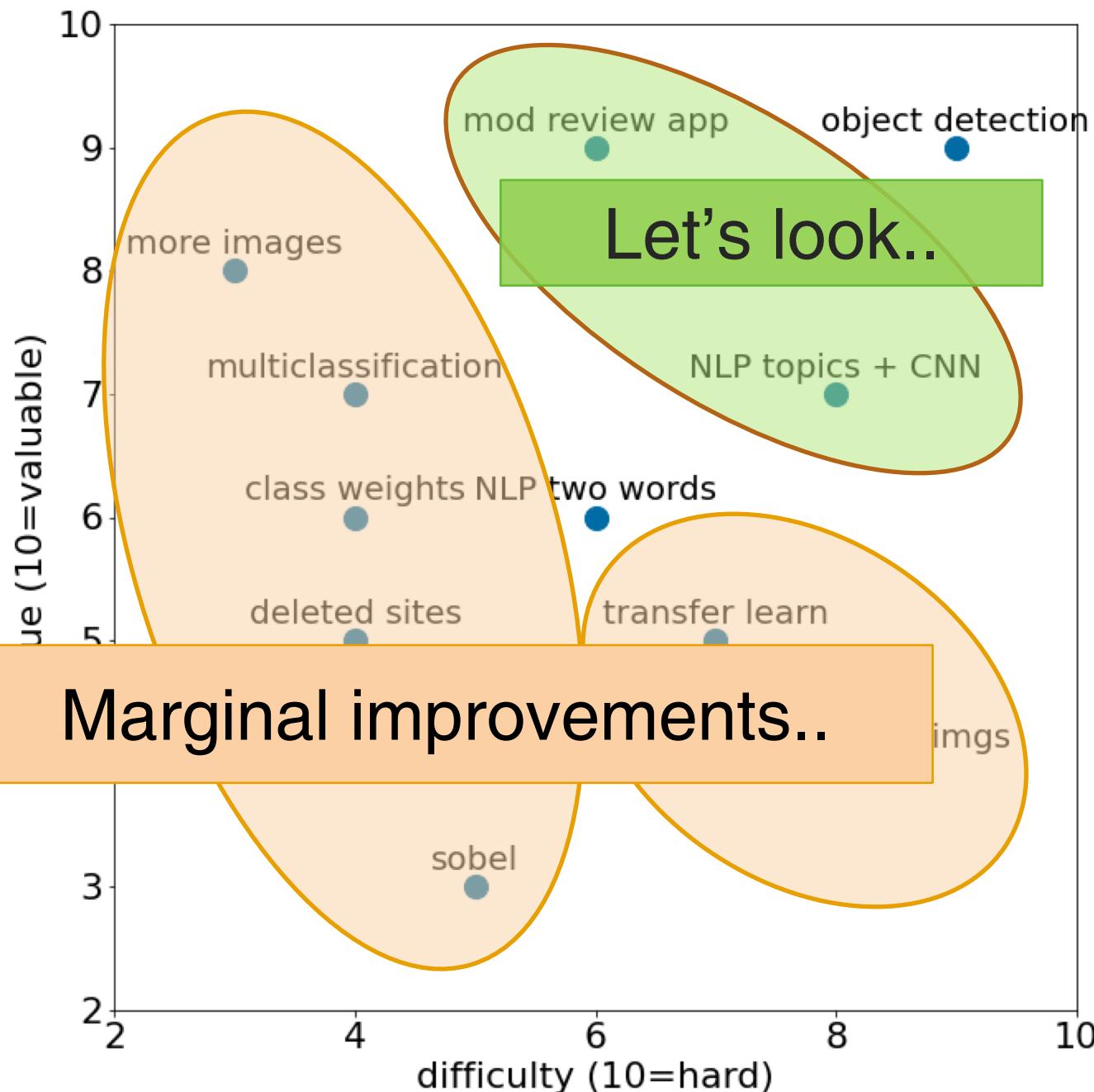
data



model







Feature engineering of improved labels

The model predicted this site: **Established Campground** and the actual category label was: **Wild Camping**



This is a very small campground for cars and RVs. Paved lot right by the ocean with toilets (a bit scary), and a picnic site. Maximum stay is 3 nights and fees were \$23/night.

Free with two vault toilets, nothing else exists
Campsite fires are required to be in some sort of metal container.

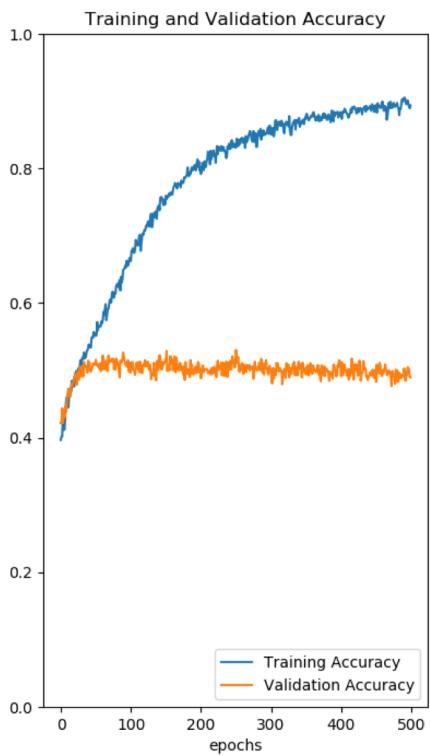


K means - assign each site cluster as label

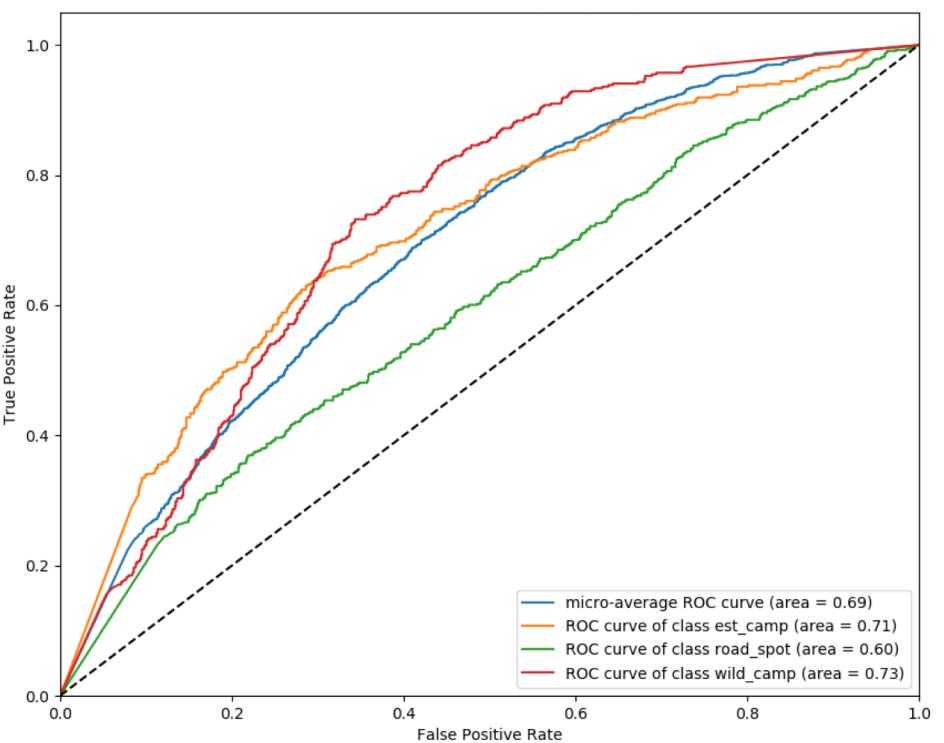
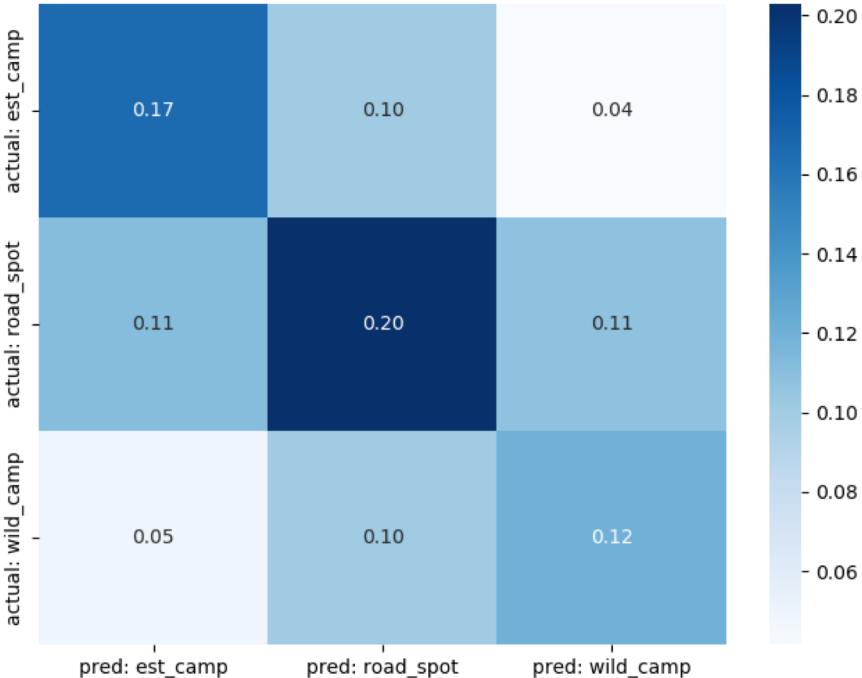
Three clusters split “wild camping”

Original labels	Proposed labels	Majority	top words
Wild Camping	Roadside spot Wild campground	73 %	parking, place, night, quiet, area, lot, nice, camping, park, free, good, overnight, great, near, spot, street, big, just, view, river
		89 %	road, spot, forest, spots, dirt, nice, just, service, good, small, river, right, camping, quiet, creek, site, camp, gravel, great, little
Established Campground	Established Campground	95 %	campground, sites, water, nice, showers, park, night, lake, toilets, site, rv, clean, free, tent, tables, pit, picnic, hot, camping, available

NLP model performance



label	%
established	29%
Road spot	30%
wild	25%



Moderator review application development

- Display predictions with other information
- Receive corrections from moderators
- Train a better ensemble model

camp.servehttp.com:33507



Thank you!



linkedin.com/in/paultanger

github.com/paultanger



Future improvement plans

- Supervised learning approach via app
- Include additional binary cols
 - Such as “bathrooms” etc
- Object detection of actual camp location
- Add flags for language (English)
- Flags for public land or not



Overview

Review predictions

View samples

Predict new

Review Dashboard selection

Here you can examine the model predictions versus the category "Wild Camping" or "Established Campground" as entered by the user who submitted the site to iOverlander. There is currently 7,834 that the model has predicted on. You can either view sites that the model correctly predicted, or more interestingly, when the model got it wrong. Functionality will be added to enable you to edit the category label.

right now, you can only choose 1 sample at a time.

choose if you want to view correct or incorrect predictions:

Incorrect Predictions ▾

enter number of samples to view:

once you have selected the options, click below:

view results

Results

The model predicted this site: **Established Campground** and the actual category label was: **Wild Camping**



details of this site:

Name	Category	Description	State
Wildcamp at the river	Wild Camping	Nice little spot just on the river bed. Follow the dirt track for like 100 meters and you will find three spots for camping or standard cars. Nice view along the river into the mountains. There is also a trail for like 9 miles along the river. Spots are sheltered by lot of bushes. Good protection against the wind. No mosquitoes in August. Perfect for motorbike and tent. Not for RV or bigger Campers/Trucks. \n\n	AK

Results

The model predicted this site: **Wild Camping**

Predict new site location

Enter the GPS coordinates in simple decimal degrees format (s -106.945057):

38.497862, -106.9450

click below to get model prediction:

[view results](#)



GPS coordinates of this site:

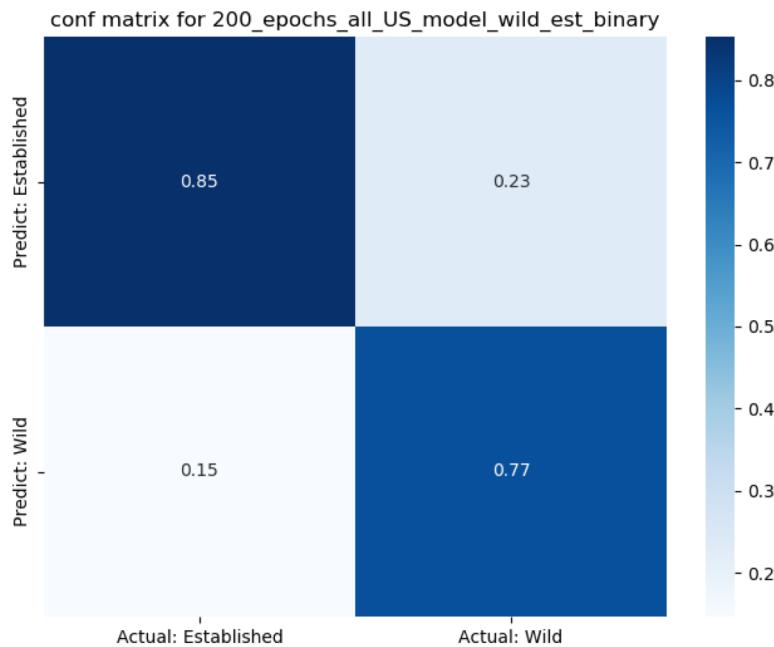
38.497862, -106.945057

Model: "sequential"

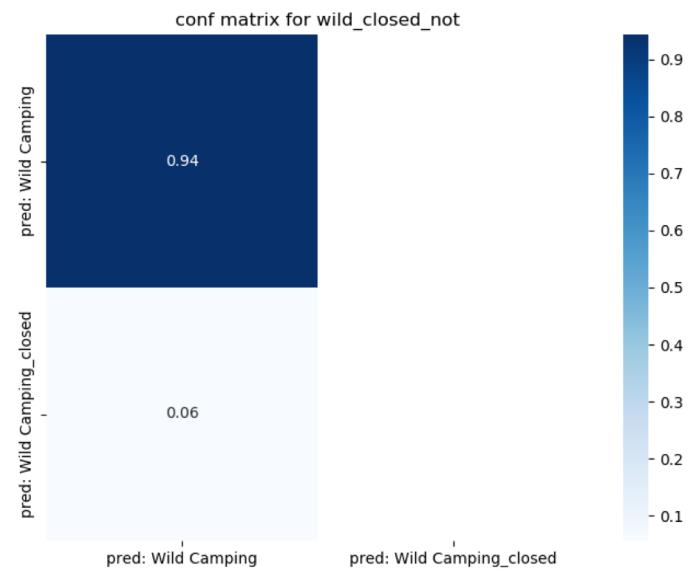
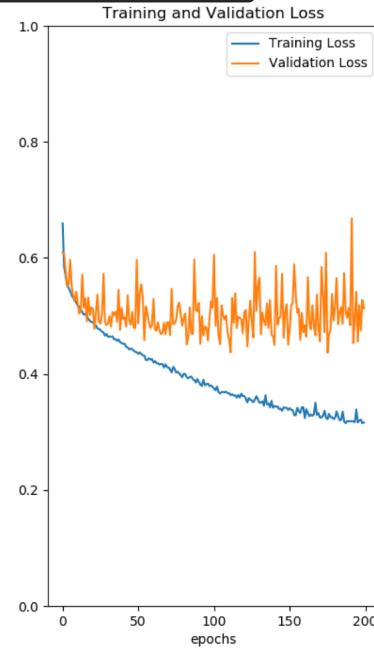
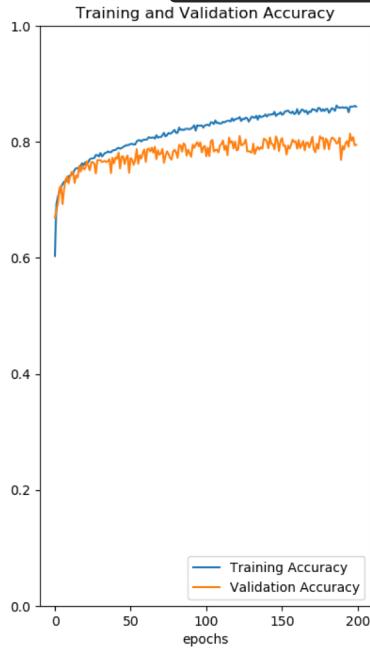
Layer (type)	Output Shape	Param #
<hr/>		
rescaling (Rescaling)	(None, 256, 256, 3)	0
random_flip (RandomFlip)	(None, 256, 256, 3)	0
random_rotation (RandomRotation)	(None, 256, 256, 3)	0
random_zoom (RandomZoom)	(None, 256, 256, 3)	0
conv2d (Conv2D)	(None, 256, 256, 32)	416
max_pooling2d (MaxPooling2D)	(None, 128, 128, 32)	0
conv2d_1 (Conv2D)	(None, 128, 128, 64)	8256
max_pooling2d_1 (MaxPooling2D)	(None, 64, 64, 64)	0
conv2d_2 (Conv2D)	(None, 64, 64, 96)	24672
max_pooling2d_2 (MaxPooling2D)	(None, 32, 32, 96)	0
conv2d_3 (Conv2D)	(None, 32, 32, 128)	49280
max_pooling2d_3 (MaxPooling2D)	(None, 16, 16, 128)	0
flatten (Flatten)	(None, 32768)	0
dense (Dense)	(None, 256)	8388864
dropout (Dropout)	(None, 256)	0
dense_1 (Dense)	(None, 1)	257
<hr/>		
Total params: 8,471,745		
Trainable params: 8,471,745		
Non-trainable params: 0		

9 new models... scaled up on AWS GPU instances

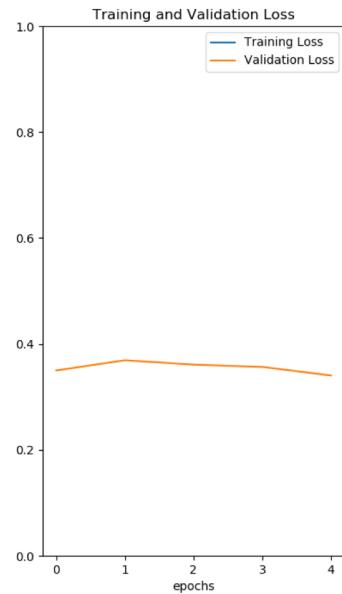
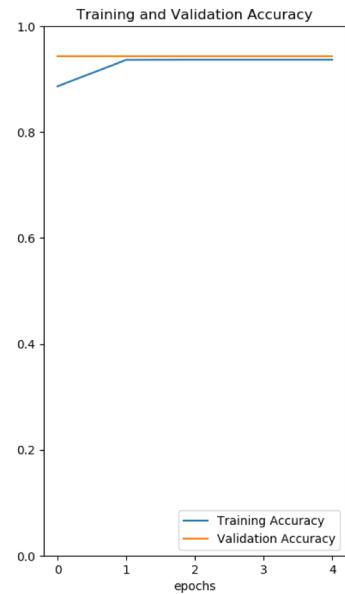
model	train set	type	unbalanced data	acc	performance	notes
original 500 epoch	wild and established from western US	binary	not really	88%	good even on unseen data	
200 epoch	wild established all US	binary	not really	75%		
"closed" vs "not closed"	wild camping, very unbalanced (6%)	binary	yes	flat	always predicted majority class even with bias and class weights	
3 category multiclass	wild camping, established, informal	multiclass	not really	plateaus at 55%		could be optimized
Xception on 3 cat multiclass	wild camping, established, informal	multiclass	not really	plateaus at 59%		could be optimized
Xception for binary (wild est)	wild established all US	binary	not really		overfits on train data after opening layers 106 + (before that each training plateaus around 55%)	could be optimized
sobel binary	wild established all US	binary	not really	75%	trains in 50 epochs instead of 200	
using NLP cluster labels	wild established and roadside	multiclass	not really	50%		
using NLP cluster labels	bin into wild and not wild	binary	yes	flat	predicts majority class	

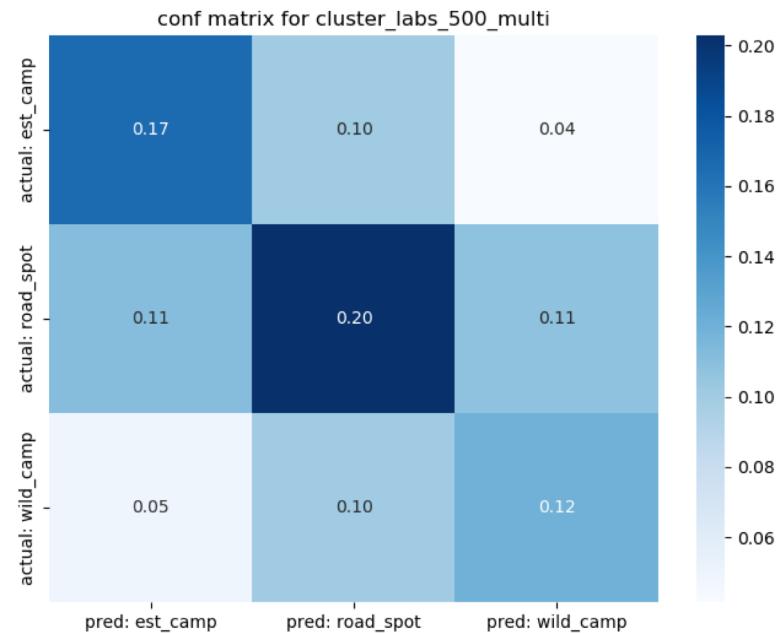
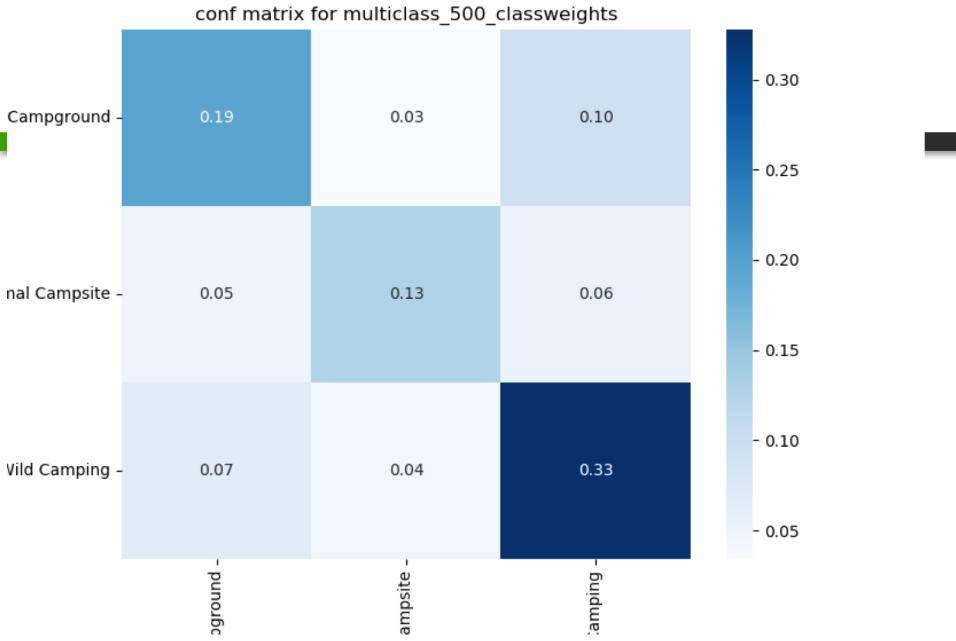


Wild vs est all data

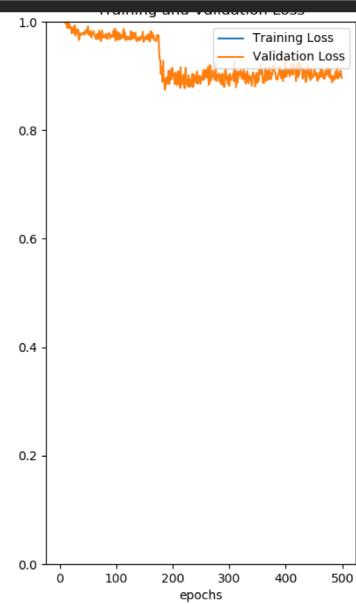
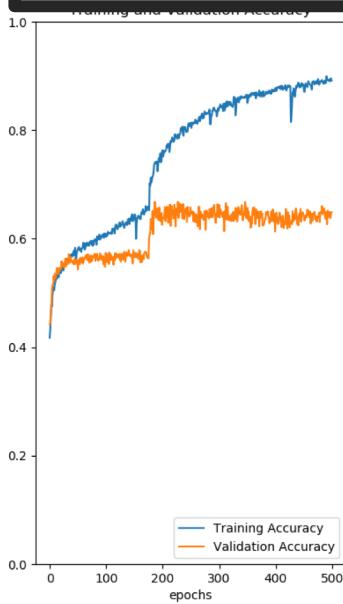


Wild closed vs wild open

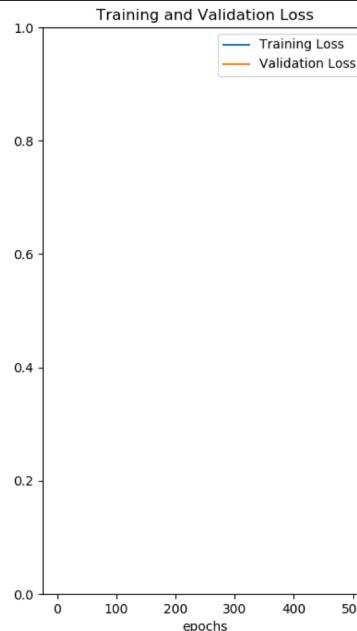
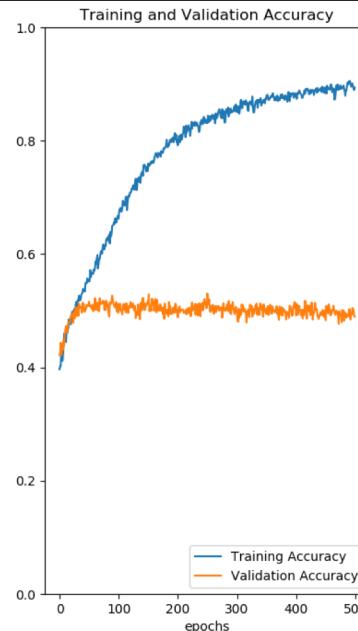




Wild, est, informal multiclass



NLP labels multiclass



Discussion of model bias and recall

- Trained on sites in western US
- Biased but appears better due to unbalanced classes?