Understanding the Amazon Basin from Space

CS231N, Spring 2017

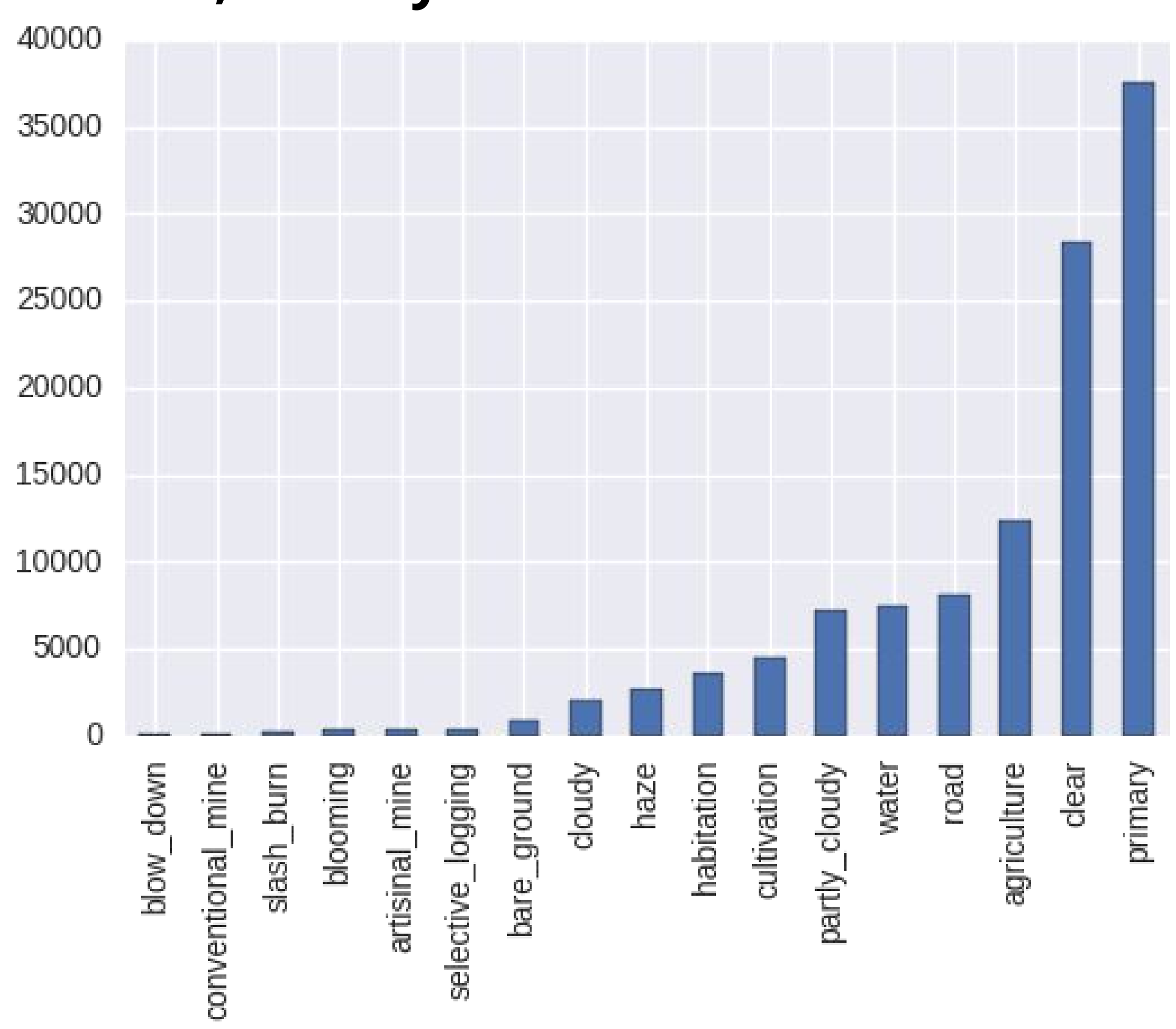
Bojiong Ni (bojiong@stanford.edu), FNU Budianto (budi71@stanford.edu), Nickolas Westman (nwestman@stanford.edu)

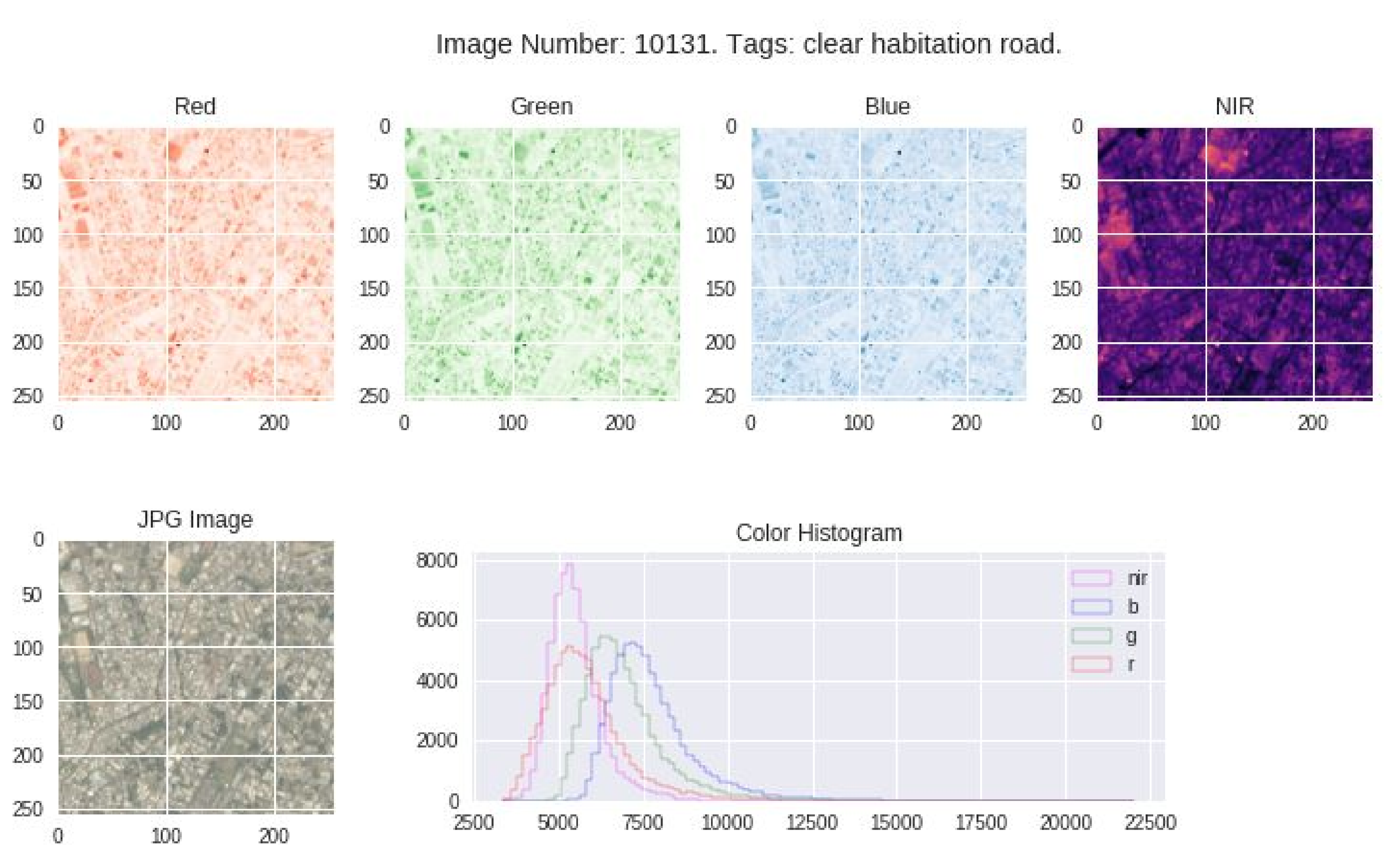
Background

- **Amazon deforestation issue: climate change, habitat loss, reduced biodiversity.**
- **Kaggle competition hosted by Planet.**
- **Lots of machine learning models for other satellite imagery, but no robust method yet to differentiate human vs natural causes of deforestation for Planet imagery.**
- **Potential impact is to help global community to understand and respond to deforestation.**

Dataset

- **256 x 256 x (3,4) satellite images of Amazon (JPG and GeoTIFF)**
- **40k train images, 61k test images**
  - **Remove 5k train for a validation set**
- **17 labels, heavily skewed.**





Problem Statement

- **Given satellite imagery of Amazon basin, classify a scene.**
- **This is a multi-label classification task with 17 labels. Labels are created by crowdworkers, so might be noisy.**
- **Evaluated using F2 score.**

$$F_2 = 5 * \frac{Precision * Recall}{4 * Precision + Recall}$$

Experiment Results

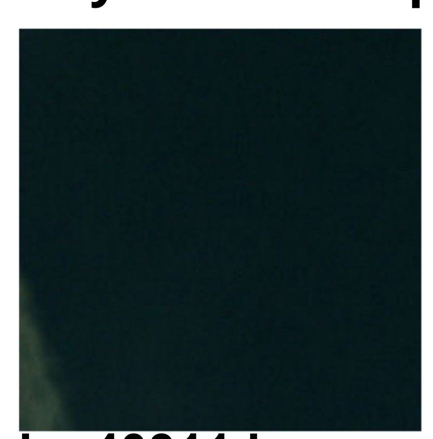
- **1-2 hours train time for most models using Google Cloud GPUs (k80), all with JPG images.**

Model	Val F2 (%)	Test F2 (%)
Naive	64.6	N/A
MLP	64.6	N/A
DenseNet	66.98	N/A
ConvNet	88	88
InceptionV3 Transfer	88	N/A
Resnet Transfer	90.4	90.6
Resnet (single model)	92.6	92.25
Resnet (ensemble)	93.03	92.796

Per label F2 using Resnet (ensemble):

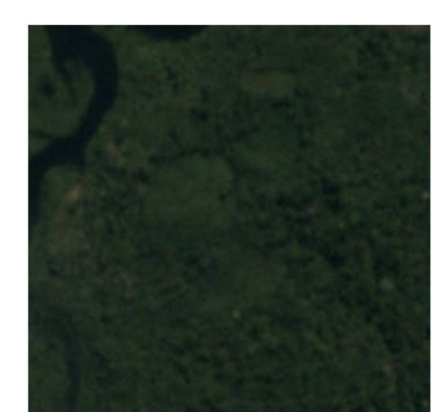
Label	Val F2 (%)	Label	Val F2 (%)
primary	99.15	habitation	78.49
clear	97.69	cultivation	68.56
partly_cloudy	94.2	conventional_mine	46.51
agriculture	90.4	bare_ground	42.96
cloudy	87.64	selective_logging	29.41
road	86.63	blooming	19.89
artisinal_mine	85.71	slash_burn	3.16
water	85.53	blow_down	0
haze	79.07		

Noisy label example:



train_40311.jpg
Actual: cloudy
Predicted: clear primary water

train_40390.jpg
Actual: cloudy
Predicted: cloudy



train_39604.jpg
Actual: clear primary water
Predicted: clear primary water

Findings and Future Work

- **As of 06/05/2017, got rank 26 out of 370. Top of leaderboard is at 93.296% F2.**
- **Resnet performs the best on this image classification task.**
- **There are some errors in the ground truth labels: ~10-15 mislabeled.**
- **Low F2 scores for rare labels even with weighted loss.**
- **Averaging performs better than majority vote for models ensembling, probably because the thresholds are optimized after averaging.**
- **Dataset augmentation by rotation helped to increase the F2 score by ~0.2%, additional augmentation might provide additional increase.**
- **Models trained using GeoTIFF images perform worse than using JPG images.**
  - **GeoTIFF Images are not corrected for sun angle and distance.**
  - **Lack of available multi-spectral tools (preprocessing, model zoos)**

Methods and Models

- **Weighted sigmoid cross entropy loss**
- **Per label threshold optimized for Val F2**
- **Dataset augmentation by rotation**
- **Ensemble of 13 Resnet(Average and Majority Vote)**
- **Exponential Moving Average of Weights**
- **Batch size: 32. Trained for ~10 epochs**
- **Implemented with Keras and Tensorflow**

MLP

- **FC(1024) ➔ Drop(0.5) ➔ FC(17)**
- **Adam: 1e-4 learning rate, 0.001 decay**

ConvNet

- **2x (Conv3-32 ➔ Conv3-32 ➔ Pool2 ➔) FC(1024) ➔ Drop(0.5) ➔ FC(17)**
- **Adam: 1e-4 learning rate, 0.001 decay**

DenseNet

- **25 layers, initial filter: 8, growth rate: 8**
- **Dropout: 0.2**
- **Adadelta: 0.5 learning rate, 0.001 decay**
- **Did not learn well**

InceptionV3

- **Transfer learning**
- **InceptionV3 ➔ FC(2048) ➔ Drop(0.5) ➔ FC(1024) ➔ Drop(0.5) ➔ FC(17)**
- **Adam: 5e-4 learning rate, 0.001 decay**

ResNet

- **Transfer learning and full retrain**
- **ResNet50 ➔ FC(2048) ➔ Drop(0.5) ➔ FC(1024) ➔ Drop(0.5) ➔ FC(17)**
- **Adadelta: 0.5 learning rate, 0.002 decay**

Acknowledgement

Special thanks to Google for providing Compute Engine credits to train the models.

References

[1] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. CoRR, abs/1512.03385, 2015.
[2]] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna. Rethinking the inception architecture for computer vision. CoRR, abs/1512.00567, 2015.
[3] F. Chollet et al. Keras. https://github.com/fchollet/keras, 2015