Automating the Identification of Illegal Human Activities in the Amazon Rainforest

Computer Vision Algorithms for Satellite Imagery Analysis

Anisha Goel (anisha14), Jacky Lin (jackylin), Charles Zhang (masarain)

Predicting

- We wished to automate monitoring the loss of the Amazon forest.
- To do so, we experimented with a few different machine learning algorithms to analyze satellite imagery of the Amazon forest.
- These algorithms categorize each satellite image into three buckets: too cloudy to analyze, no destructive human activity, or some destructive human activity .

Data

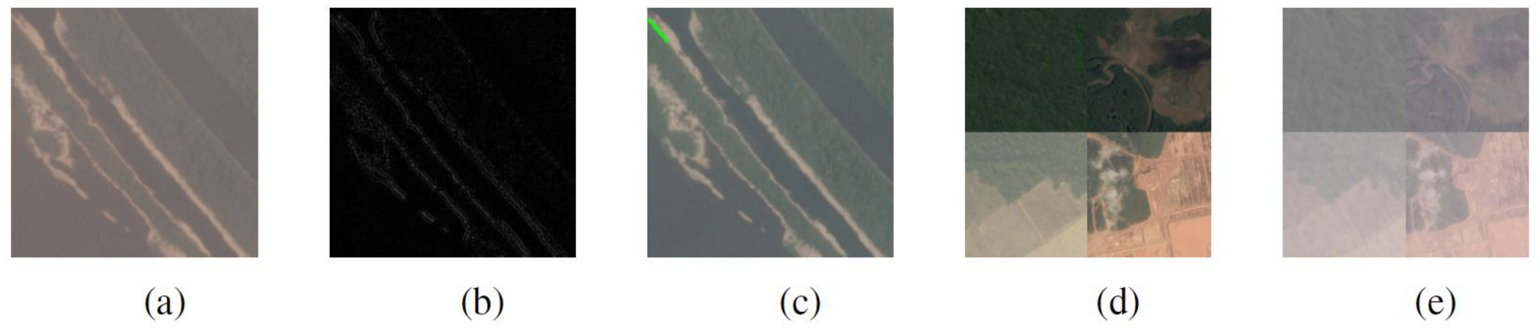
Our data consists of satellite images (each of size 256 x 256) of the Amazon Rainforest provided by Planet Labs, a San-Francisco private Earth imaging company and our labels are tags of what that image corresponds to (for instance, “habitation”, “agriculture”, “conventional mine”, etc.). The labels were crowd-sourced, so there is noise within the data set. Images divided into 36,000 train, 2,000 validation, and 2,000 test.

Features

We extracted the following features from the images:

- White Balance Ratios: mean(G)/mean(R) and mean(G)/mean(B)
  - To identify the relation between R, G, and B channels
  - For detecting clouds and haze in the image
- Mean luma and mean G: Y = 0.299 R + 0.587 G + 0.114 B and mean(G)
  - To distinguish green features like trees, pasture, and grasslands.
- Mean of Laplacian filter
  - To distinguish details such as trees, logging, mining, habitation etc
- Hough transform
  - To detect are the man made structures
- Stitching images as Maps
  - To increase our dataset for the final goal to predict the high risk areas

Note: RGB stands for Red, Green and Blue channels of the image.



(a) Original Image from training dataset. (b) Laplacian response of (a). (c) Hough Transform response of (a). (d) Concatenate 4 samples. (e) Blended with hazy image.

Models

Overview

First, we perform a softmax regression on the extracted features and compare that to the result from the K-Means. Then we run a NN on the actual satellite images and compare the results from the CNN.

Softmax Regression

Our first approach was to just try running a softmax regression on our dataset. We used the extracted features with an intercept term as the input. For the loss function, we used the log likelihood function from lecture. We ran gradient descent on the loss function to train our model.

K-Means

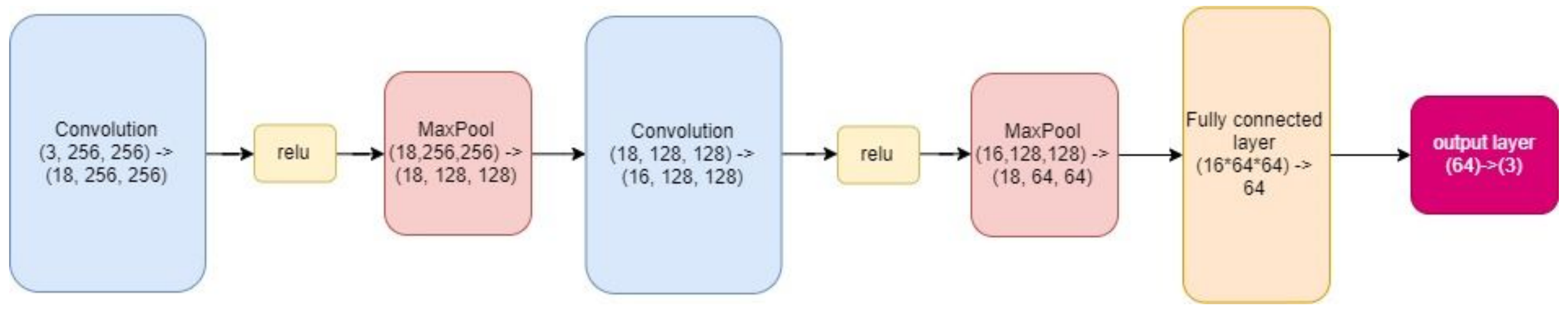
Our next approach was to use the filter and separate the dataset with K-Means into 3 clusters, assigning each cluster to one of the three classes. We conducted K-Means on the features extracted (of which there were 7), rather the images themselves, and ran K-Means 20 times with different random initialization and took the best result.

Standard Feed-Forward Network

We ran a feed-forward network with 5 hidden layers on the flatten images of size 256 x 256 x 3. We used Cross Entropy Loss for our output layer, which had 3 neurons, one corresponding to each class. All others neurons were activated with a sigmoid function.

CNN

The network is composed of two convolution layer 24 and 16 filters respectively with one sigmoid fully connected hidden layer. We used Cross Entropy Loss for our output layer, which had 3 neurons, one corresponding to each class. We used L2 regularization to prevent overfitting.



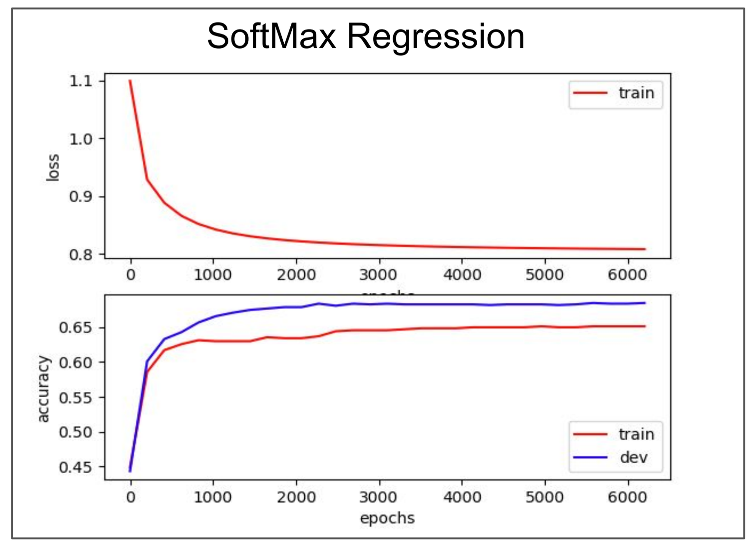
CNN architecture

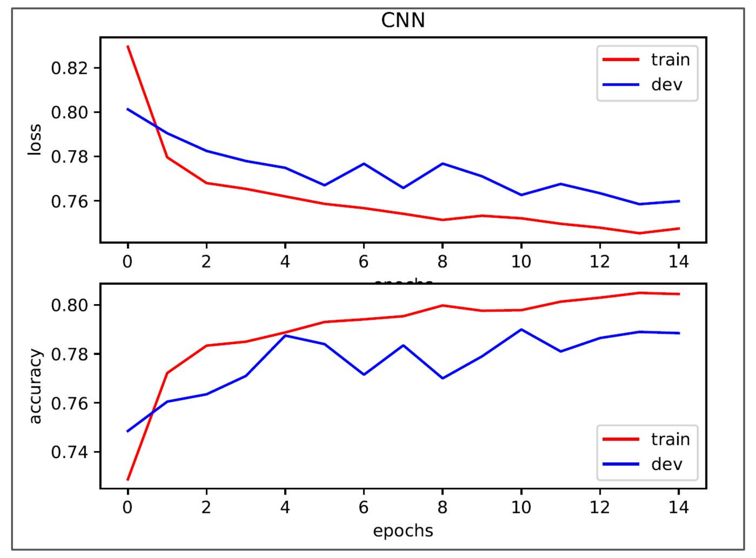
Future

- Generalize our model to other regions of the world
  - Does our model have more predictive power in areas with certain geographies or landscapes?
- Model improvement:
  - Compare against CNN trained on stitched images and originals panels.
  - Compute model with a linear combination of more than one label.
- Robustness checks for overfitting, error analysis, and parameter tuning

Results

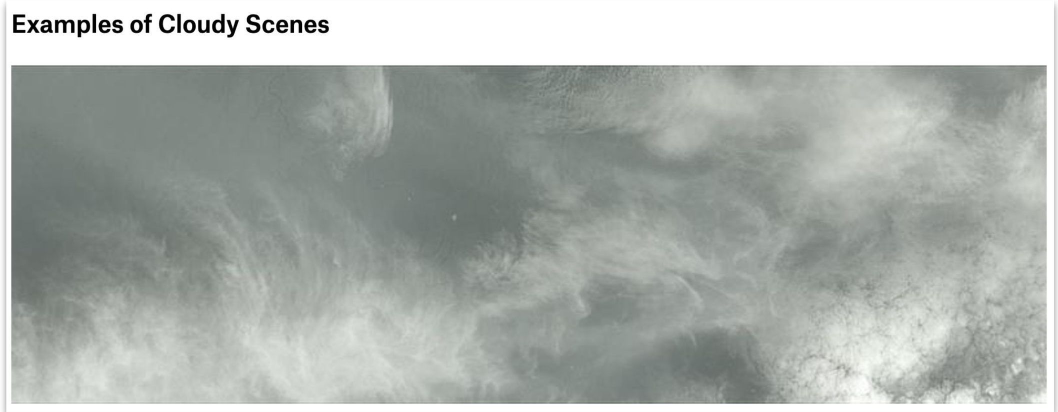
Model	**Training accuracy**	**Test accuracy**	**Stitched Test accuracy**
Softmax Regression	**0.651**	**0.685**	**NA**
K-Means	**NA**	**0.72**	**NA**
Standard Feed-Forward Network	**0.61**	**0.71**	**NA**
CNN	**0.80**	**0.81**	**0.76**





Discussion

- Deriving meaningful features was a difficult and time consuming process.
- The K-Means and the Standard Feed-Forward Neural Network performed better than random, but it was highly biased by the prior distribution of the labeled data.
- The CNN performed better with two convolution layer and less neurons in hidden layer. Adding more neurons in fully connected layer caused it to overfit on training data.
- The CNN model performed better when partly cloudy images were not categorized in cloud class.



References

Jean, N., Burke, M., Xie, M., Davis, W. M., Lobell, D. B., & Ermon, S. (2016). Combining satellite imagery and machine learning to predict poverty. *Science, 353*(6301), 790-794.

Basu, S., Ganguly, S., Mukhopadhyay, S., DiBiano, R., Karki, M., & Nemani, R. (2015, November).Deepsat: a learning framework for satellite imagery. In Proceedings of the 23rd SIGSPATIAL interna-tional conference on advances in geographic information systems (p. 37). ACM.