Amazon Rainforest Image Classification Challenge

Yixin Cai
Department of Computer Science, Stanford University

Introduction

- Based on a Kaggle challenge - Planet: Understanding the Amazon from Space
- Track human activities in the Amazon rainforest
- Differentiate human causes and natural causes on deforestation
- Previous approaches cannot make use of newly collected high resolution images
- Find robust methods for processing imagery

Problem Statement

- Multilabel Classification
- Input – 256*256 satellite image
- Output – 17 possible label (not exclusive)
  agriculture, artisinal_mine, bare_ground, blooming, blow_down, clear, cloudy, conventional_mine, cultivation, habitation, haze, partly_cloudy, primary, road, selective_logging, slash_burn, water
- Tools
  - Different CNN architectures with increasing number of convolution layers
  - Try existing models such as VGG
- Evaluation Metric
  - F2 score for each image

$$(1+\beta^2)\frac{pr}{\beta^2 p+r}, p=\frac{tp}{tp+fp}, r=\frac{tp}{tp+fn}, \beta=2$$

  - Mean Accuracy
  - Precision and Recall for each label

Models

- 4–layer CNN
  - 2 convolutional layers, each with 32 3*3 filters
  - Dropout and max pooling is used
  - 2 fully connected layer
- 6–layer CNN
  - 2 additional convolutional layers with one 2*2 filters
  - Image down sampling
- Existing model
  - VGG

Experiments

- Initial accuracy with 4–layer CNN – 0.89
- Problem 1: VGG accuracy – 0.905, f2 score – 0.64
  - All images are classified as clear and primary
  - High bias in each class
- Sigmoid cross entropy is a terrible loss function
- SVM (hinge loss) leads to a much better result
- Problem 2: Inconsistent f2 score for training set
  - Training – 0.95
  - Test – 0.75
- Batch normalization behaves differently in training and test time
- Decision boundary for SVM is no longer valid
- Remove batch normalization
- Slow training but better results

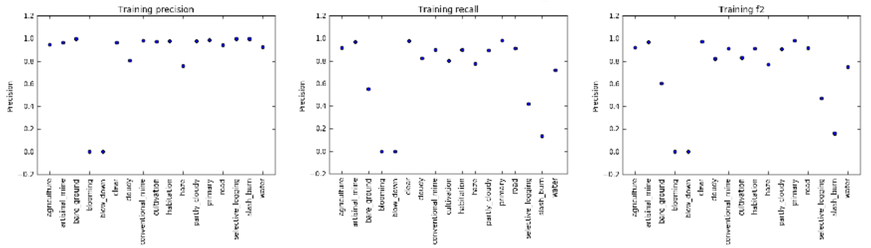
Results

- The benchmark f2 is 0.67
- Using SVM loss function
- Batch normalization is not used
- Training becomes longer
- 6-layer CNN is still training

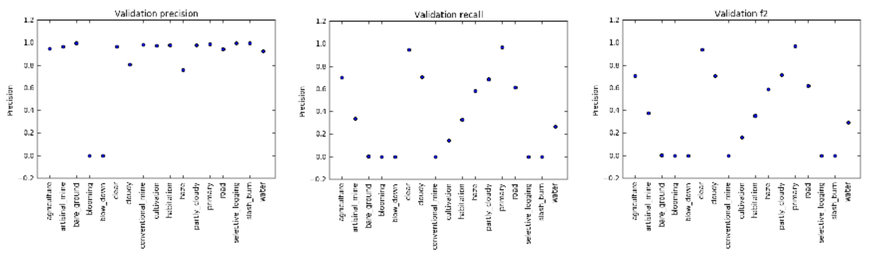
	Training f2	Training accuracy	Val f2 & Test f2
4-layer+sigmoid	0.60	0.89	0.59
VGG+sigmoid	0.64	0.91	0.64
4-layer+svm 10 epochs	0.87	0.96	0.80
4-layer+SVM 20 epochs	0.92	0.98	0.81

Statistics and Analysis

Training Precision, Recall and F2 score



Validation Precision, Recall and F2 score



Datasets

- Provided by Kaggle
- 40479 training images
  - 80% as training set
  - 10% as validation set
  - 10% as test set
- 61191 test images
- Training set may contain mislabeled images

Figure 1 Sample Images with labels

Conclusion and Future Work

- Train deeper networks
- Existing network will not be used - dataset are too different
- HOG feature and color histogram
- Precision and recall is 0 for blooming and blown_down
  - Too much regularization
  - Change to a lower C value in SVM - reduce false negative
  - Maybe add more weight to false negative in loss function
- Examine misclassification examples