Classification of natural landmarks and human footprints of Amazon using satellite data

Kaifeng Chen[1], Yao Zhou[2]

[1]Department of Applied Physics, Stanford University

[2]Department of Material Science Engineering, Stanford University

Objective

The understanding of natural landmarks and human footprints of the Amazon is of great importance for the preservation of the forest and habitat of Amazon. Such understanding requires not only the knowledge of the nature itself, but also the allocation of natural resources and the effects of human activities. Thus for this purpose, it is useful to classify the different regions of Amazon with correct labels.

In our project, we use Kaggle Amazon satellite image data, and develop deep learning algorithms to correctly classify all the images into in total 17 classes. We use SVM as the baseline, and use convolutional neural network along with transfer learning using VGG-16 for further improvements. We present the obtained accuracies as well as the F_2 scores.

Dataset

Kaggle Amazon satellite image dataset contains in total 40479 images with corresponding labels. Each image is of size 256x256, and can belong to one or many classes among the in total 17 classes. Fig. 1 is a collection of examples. The numbers of occurrences of each label are not evenly distributed (Fig. 2). Here we use 80% of the data as the training data, and the rest as the validation data.

Figure 1: Examples of the Dataset

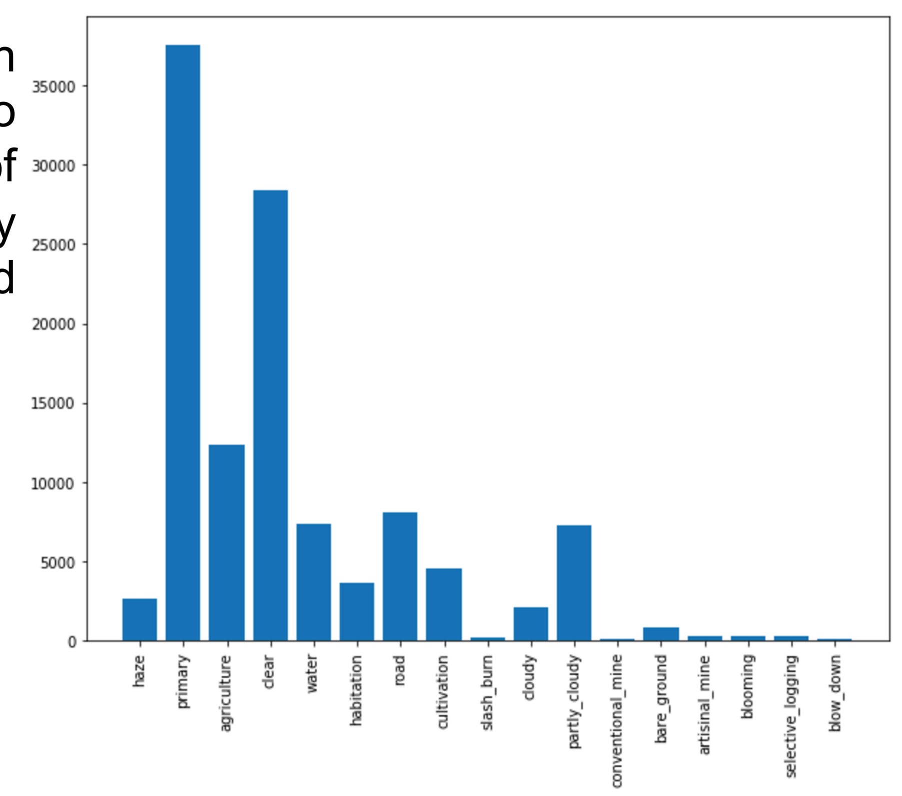


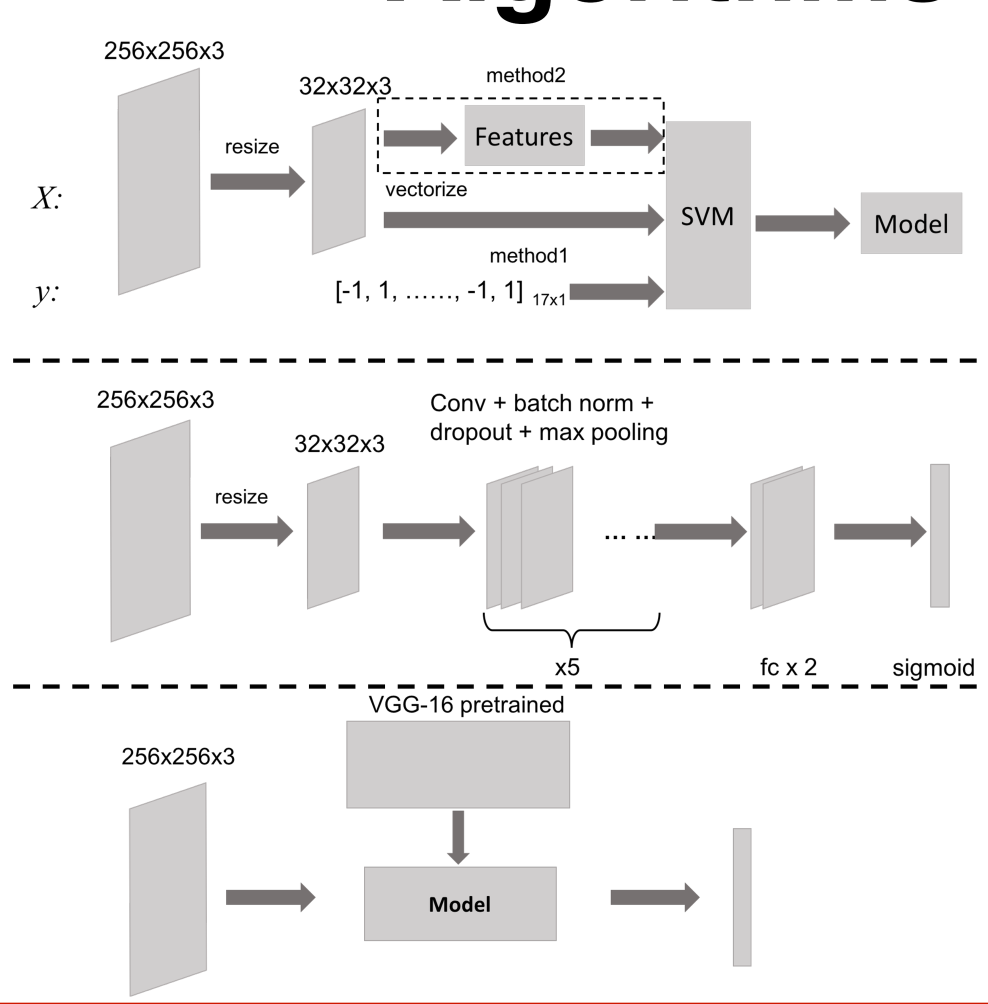
Figure 2: Distribution of the Labels

Algorithms

1. Baseline (SVM):

2. Convolutional neural network:

3. Transfer learning:



Loss function:

$$L = \frac{1}{N}\sum_{N}\sum_{C}\max\left(0, 1 - y_{ij}x_i w_j\right) + \lambda\|W\|^2$$

$$L = -\frac{1}{N}\sum_{N}\sum_{C} y_{ij}\log\sigma(s_{ij}) + (1-y_{ij})\log\left(1-\sigma(s_{ij})\right)$$

F_2 score:

$$(1+\beta^2)\frac{pr}{\beta^2 p + r}$$

where $\beta = 2, p = \frac{t_p}{t_p + f_p}, r = \frac{t_p}{t_p + f_n}$

Results

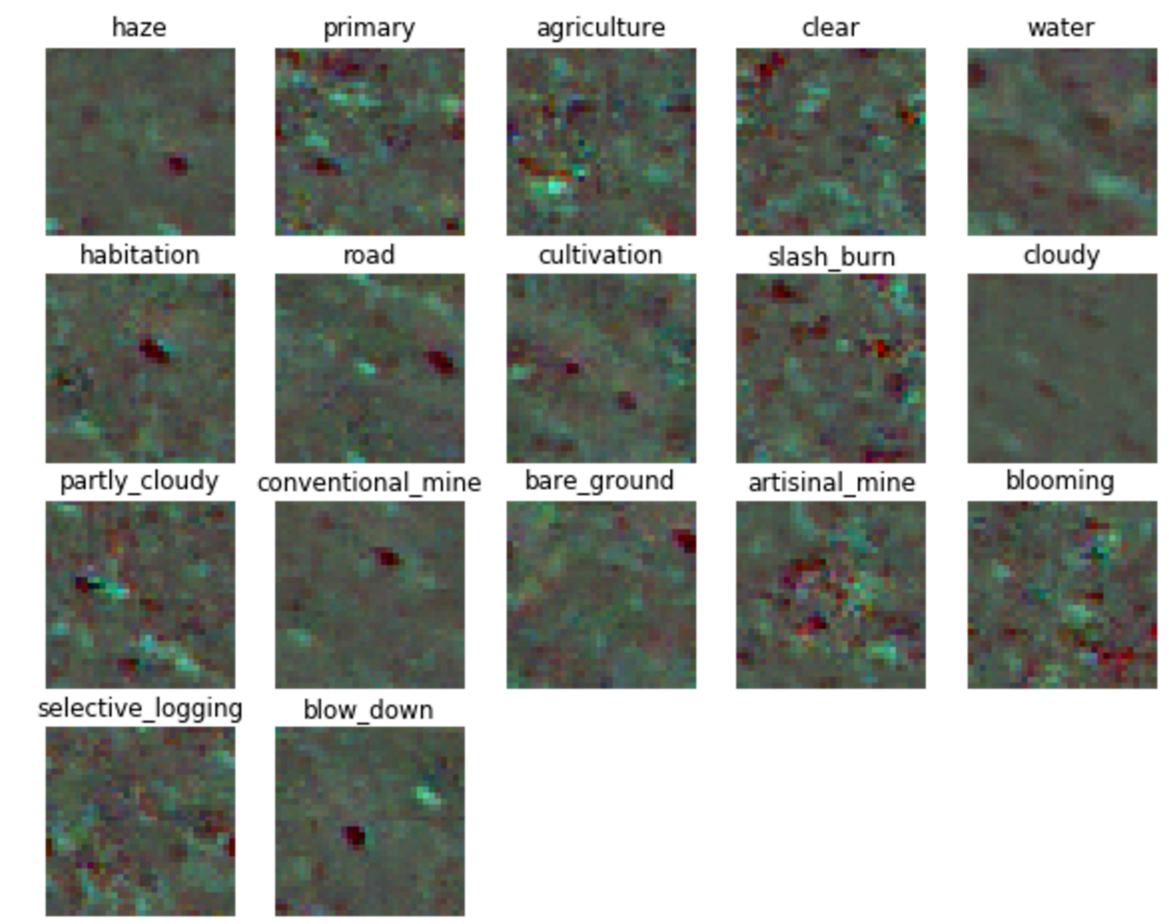
Table 1: Validation Accuracy by Class

	SVM	SVM (features)	CNN	CNN (VGG16)
haze	0.9350	0.9350	0.9321	0.9526
primary	0.9307	0.9316	0.9527	0.9667
agriculture	0.6948	0.7294	0.8125	0.8829
clear	0.7020	0.7227	0.8714	0.9405
water	0.8221	0.8223	0.6950	0.8943
habitation	0.9119	0.9119	0.8489	0.9377
road	0.8005	0.8109	0.7720	0.8961
cultivation	0.8871	0.8871	0.8531	0.8697
slash_burn	0.9942	0.9942	0.9938	0.9941
cloudy	0.9520	0.9527	0.9704	0.9718
partly_cloudy	0.8150	0.8179	0.9171	0.9597
conventional_mine	0.9979	0.9979	0.9979	0.9981
bare_ground	0.9758	0.9758	0.9562	0.9717
artisinal_mine	0.9915	0.9915	0.9894	0.9970
blooming	0.9918	0.9918	0.9918	0.9907
selective_logging	0.9911	0.9911	0.9911	0.9902
blow_down	0.9975	0.9975	0.9975	0.9975

Table 2: Overall Validation Accuracy and F_2 Score

	SVM	SVM (features)	CNN	CNN (VGG16)
Accuracy (all labels correct)	0.3365	0.3491	0.4134	0.5488
F_2 score	0.6465	0.6770	0.8520	0.9177

Figure 3: Label Visualization of CNN



Conclusion

- We preprocessed data and implemented SVM classifiers using raw pixels or extracted features of compressed 32x32 images as our baseline. We achieved our baseline overall validation accuracy 35% and F_2 score 0.68.
- We implemented a convolutional neural network on compressed 32x32 images and achieved overall validation accuracy 41% and F_2 score 0.85.
- We implemented transfer learning with VGG16 on original images and achieved overall validation accuracy 55% and F_2 score 0.92.

Future Work

- To learn the correlation among different labels, more sophisticated loss function or architecture could be used.
- To improve F_2 score, threshold values used for each label could also be incorporated into the learning process.

1. G. M. Foody and A. Mathur. A relative evaluation of multiclass image classification by support vector machines. *IEEE Transactions on Geoscience and Remote Sensing*, 42(6):1335–1343, June 2004.
2. A. Mathur and G. M. Foody. Multiclass and binary svm clas- sification: Implications for training and classification users. *IEEE Geoscience and Remote Sensing Letters*, 5(2):241–245, April 2008 .
3. Wei, Y., Xia, W., Huang, J., Ni, B., Dong, J., Zhao, Y., and Yan, S. CNN: Single-label to multi-label. CoRR, abs/1406.5726, 2014.
4. Simonyan, K. & Zisserman, A. Very deep convolutional networks for large-scale image recognition. In *Proc. International Conference on Learning Representations* (2014).