Multi-label Classification of Amazon Forest Satellite Images

Rajanie Prabha Chinmay Prabhakar Min-An Chao

TECHNISCHE UNIVERSITÄT MÜNCHEN

DL4CV
WS17/18

Introduction

- Given the dataset of satellite images of Amazon forests, the goal is to analyze the changes in patterns, especially those related to illegal deforestation, by classifying the data with respect to atmospheric conditions and various land cover types. Data are provided by Planet (`www.planet.com`) used as a Kaggle competition.

Input Dataset

- 40,480 training samples and 61,192 test samples are provided. Types include JPG (3 channels, 8-bit RGB) and TIFF (4 channels, 16-bit RGB plus Near-Infrared, or NIR).
- Total 17 labels to be classified, 4 are weather conditions, others are geographic features and signs of human activities, including rare labels such as illegal mining and burndown, etc.
- Each image is labeled with exactly 1 weather condition, and 0, 1, or more labels presenting land features and human activities in it.

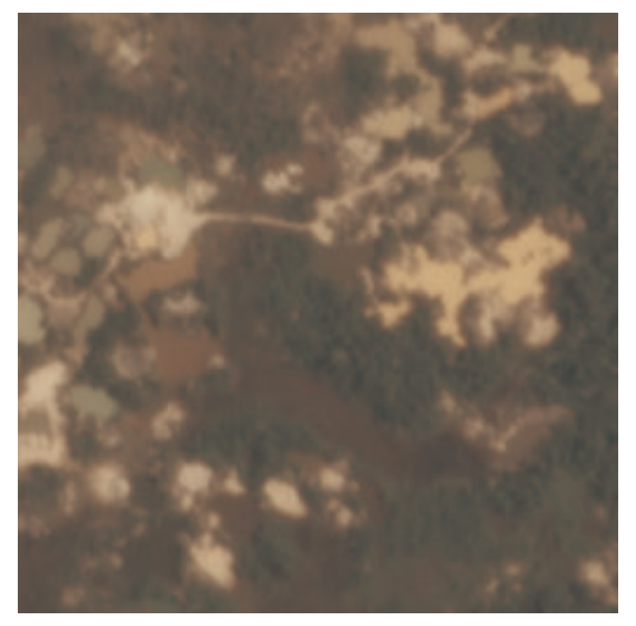


Image with multiple activities: agriculture, artisan mining, habitation, road, water.

Hazy image with water.

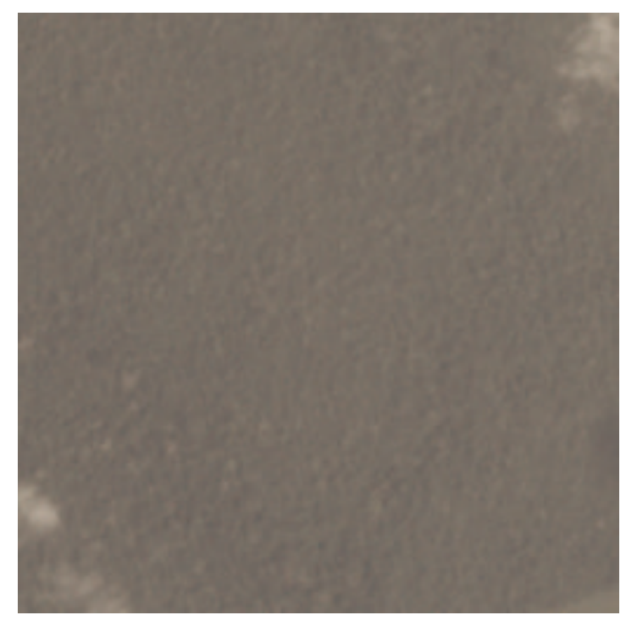


Image labeled clear and primary forest, but somehow hazy.

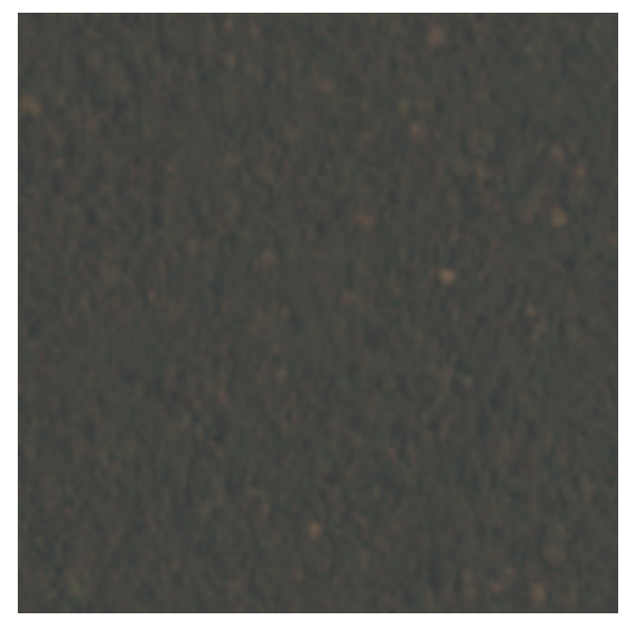
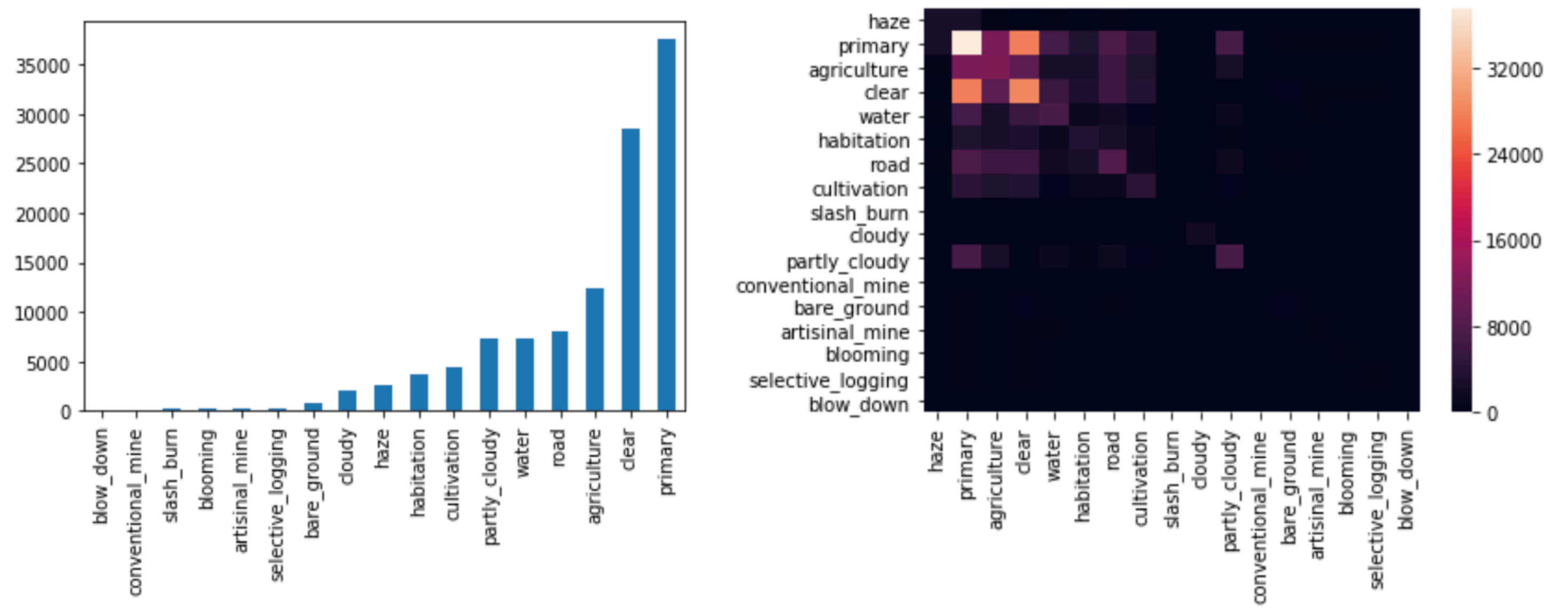


Image labeled clear and blooming, but really hard to recognize them.



Histogram of labels (left) and correlation matrix (right) show data are extremely imbalanced and labels have no strong dependence but are not purely independent.

Measuring Metric

- Because of data imbalance, F2 score is designated by Kaggle competition instead of total accuracy.

$$F2 = (1+\beta^2)\frac{PR}{\beta^2 P + R}, \text{where } P = \frac{TP}{TP+FP}, R = \frac{TP}{TP+FN}, \beta = 2$$

- Since F2 score cannot be modeled directly as loss function, we use Sigmoid function for each label and compute F2 after each epoch. Early stopping happens if decreasing loss is not helping F2 score.

Thresholding Adjustment

- Because of multi-labeling problem and F2 score, we have to adjust the decision boundary for each label separately. This is done after all epochs with the original un-adjusted F2 scores. We sweep from 0 to 1 for each label to find a point maximizing F2. F2 scores shown in this poster are all after thresholding adjustment.

Neuron Network Architectures

- Several pretrained neuron network models on ImageNet, such as AlexNet, VGG, ResNet, DenseNet can be directly applied to such task. We use SGD with 0.9 momentum and JPG images as the first benchmark. Learning rate shown below refer to FC layers, while feature layers are updated with 0.1 of it.
- We also tried GoogLeNet (`Inception_v3`) on PyTorch but stuck at some problems, resulting in slow convergence and low F2 (0.72).

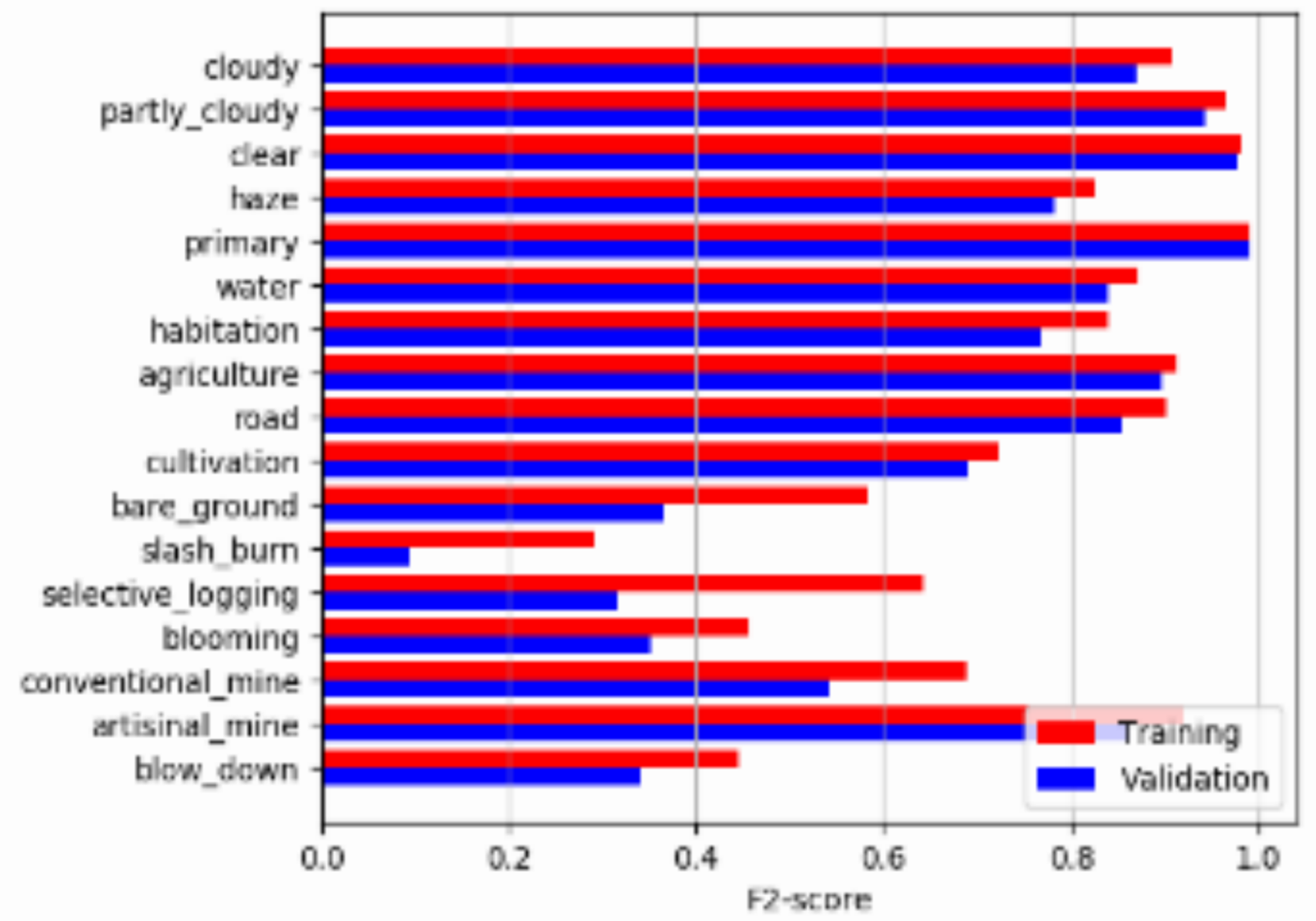
Model	Lrn. rate	Batch Size	F2 score	Trn. time (hr)
AlexNet	0.2	320	0.9082	0.5
VGG-16	0.05	100	0.9224	6
ResNet18	0.3		0.9276	1.5
ResNet34	0.3			
DenseNet121	0.1	80	0.9257	7
DenseNet169	0.1	64	0.9259	7.5
DenseNet201	0.1	52	0.9260	12.5

Data Augmentation

- Without any labeling adjustment, flip and rotate is the only helpful way for labeled satellite images. Results in this poster include this.

Result Analysis

- Labeling noise. There are certain portion of images with features distinguishable even by experts. This can be found for significant jitter between training and validation sets. Leaderboard on Kaggle also show very close scores (from 0.9332 to 0.9322) among top 10.
- Inability to serve as screening for some illegal activities. Although F2 prefers recall than precision, some rare labels like blow down, slash burn, and blooming, still have low recall, i.e., high miss detection.

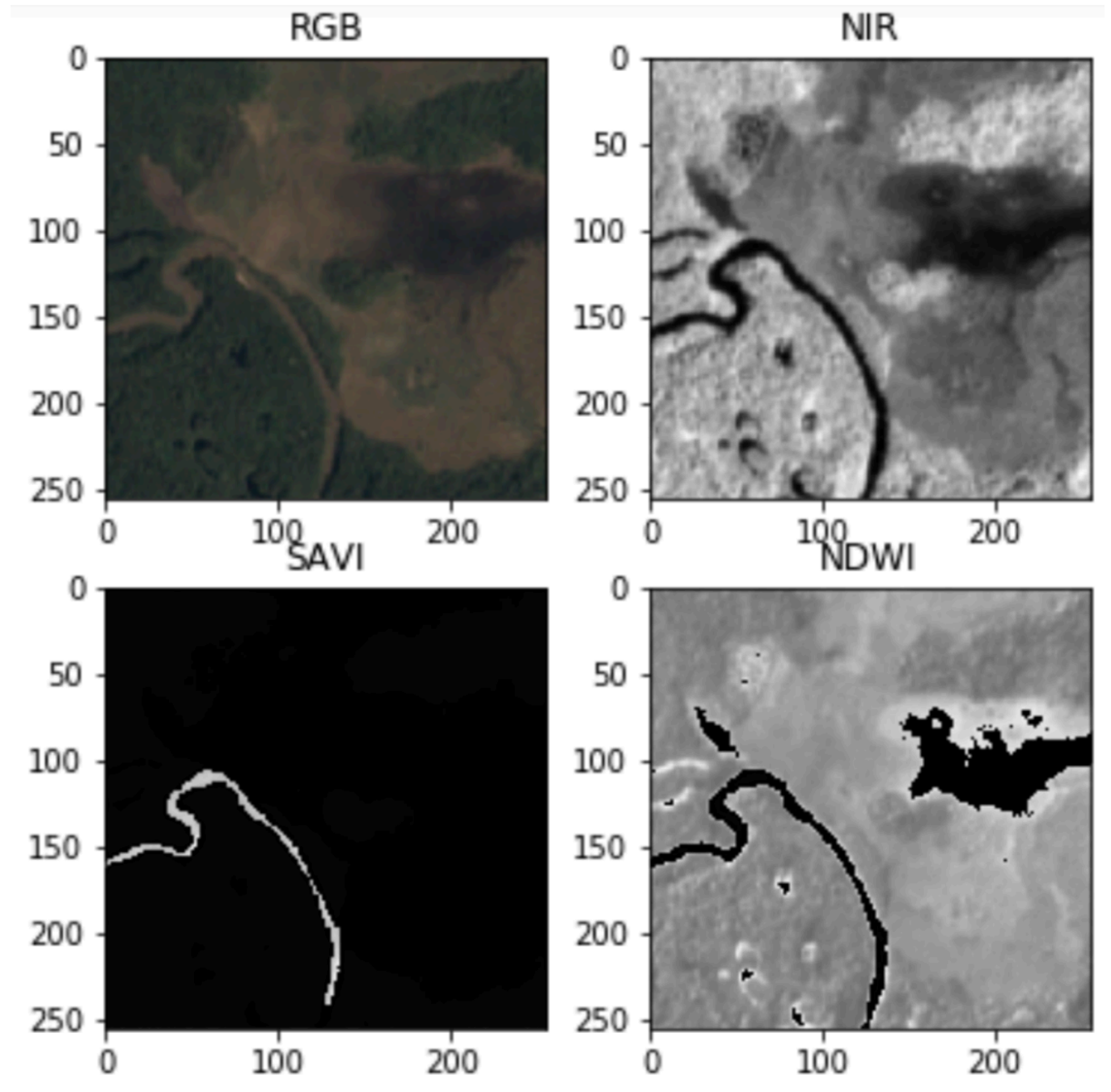


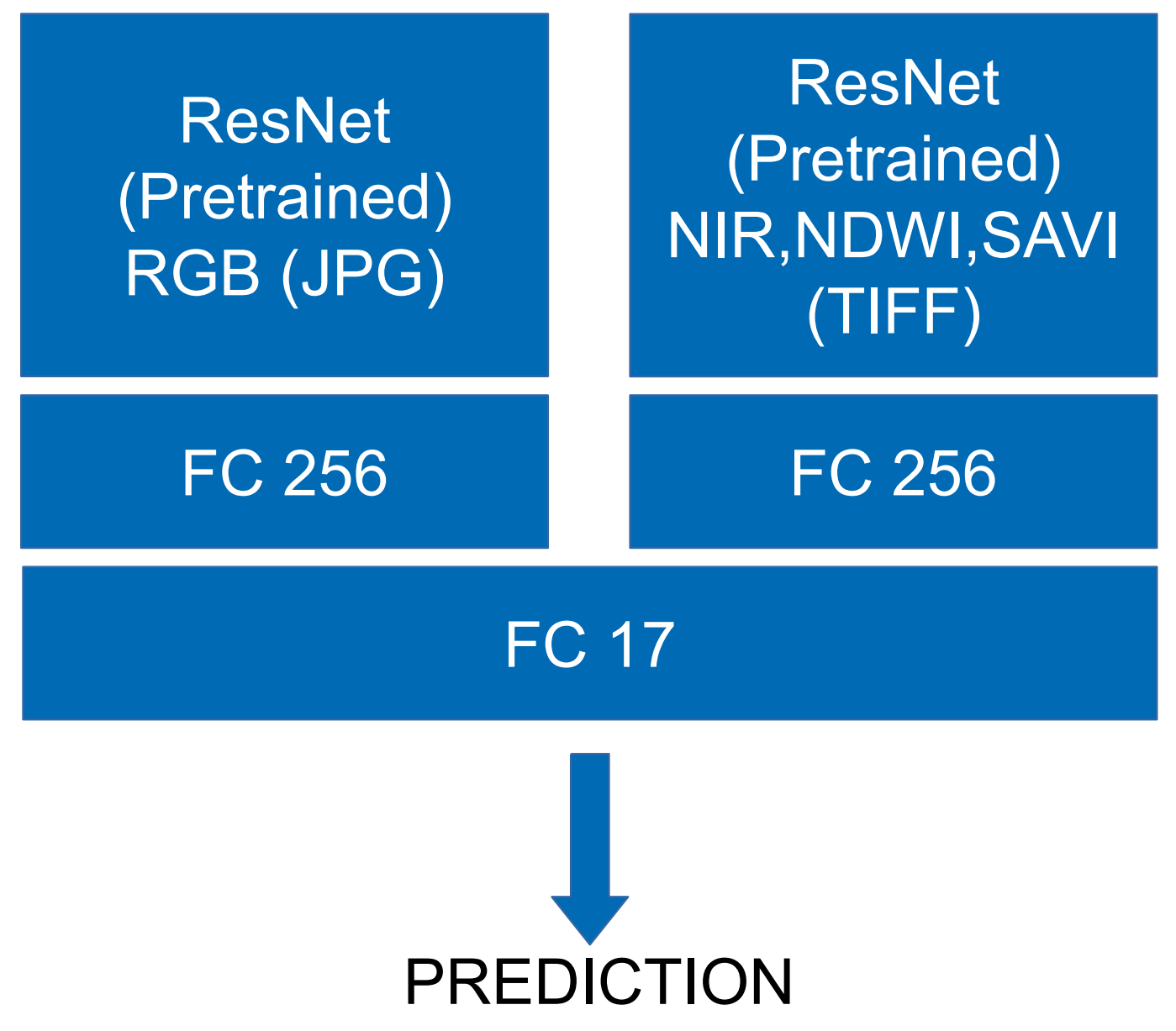
Label	N_{sample}/N	Prec.	Recall
Blow down	101/40480	1.000	0.125
Slash burn	209/40480	0.412	0.175
Blooming	332/40480	0.385	0.156
Conv. mine	100/40480	0.722	0.619
Artis. mine	339/40480	0.724	0.887

For some rare labels, there are mismatch between training and validation sets (left), and some with small recall (up) even though F2 prefers recall.

Mix-Model with NIR information

$$\text{SAVI} = \frac{(1+L)(\text{NIR}-\text{R})}{\text{NIR}+\text{R}+L}, L = 0.5, \text{NDWI} = \frac{\text{G}-\text{NIR}}{\text{G}+\text{NIR}}$$





ResNet (Pretrained) RGB (JPG)

ResNet (Pretrained) NIR,NDWI,SAVI (TIFF)

FC 256

FC 256

FC 17

PREDICTION