Invasive Species Identifier

Andres Hernandez andresh@stanford.edu

Elliott Jobson emjobson@stanford.edu

Background

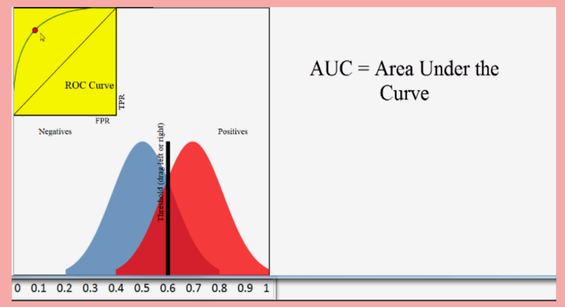
Invasive species of plants can have a negative impact on the environment, the economy, and health.

By identifying them, scientists can more easily alert local governments and NGO efforts; however, scientists cannot possibly classify the vast swaths of flora in the Amazon.

Problem Statement

Leveraging algorithms based on computer vision techniques, convolutional neural networks, and data preprocessing, we will develop a classification model that, given an input image of foliage, will determine whether or not there is an invasive species within.

Evaluation Methods



http://www.dataschool.io/roc-curves-and-auc-explained/

First Model Findings

Final Evaluation Accuracy: 63.2%

The following were incorrectly classified:

(Small flowers)

(Low contrast)

Potential fixes: Maintaining high-res images
Augment with varied contrast

Datasets

The training set contains 2,295 images taken in a Brazilian national forest in addition to 1,531 in the test set.

Each numbered image will be associated with a probability that a species is present.

Some images only have one flower. Also, the dataset is only approximately twice the size of the test set.

Non-invasive

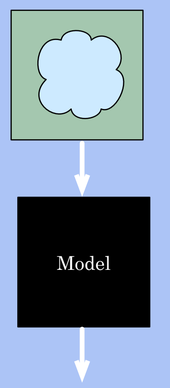
Invasive

Methods

1st model: Written “from scratch.”
2nd model: Train on augmented set.
3rd model: Repurpose VGG-16.

Our model:

```
Padded Conv.
Max Pool
Padded Cov.
Max Pool
Padded Conv.
Padded Conv.
Max Pool
Padded Conv.
Padded Conv.
Max Pool
FC
Dropout
FC
```



Considerations

Pitfalls:

Trained on limited dataset (2,295 vs 1,531)

Input size 150x200, too small

Flowers with low contrast to surroundings

Extensions:

Transfer Learning

Dataset Augmentation by: Contrast adjustment
Jitter
Flipping