Detecting Coconut Trees from Aerial Photographs
Amber Thomas and Tywan Challenger

Abstract
- Natural disasters in the tropics can threaten food security.
- As a result, the World Bank issued a challenge to the global AI community to develop computer vision algorithms to count and locate standing food trees, like coconut trees, from aerial photographs.
- With an R-CNN using ResNet101 base classifier pre-trained on the kiti object detection dataset and trained using an in situ aerial photograph taken over the island of Tonga, we trained a detection model that returned 94% AP using the PASCAL VOC evaluation metric.

Dataset and Data Cleaning
- Our data set was a single 25,006 by 17,761 aerial photograph with a spatial resolution of 90m, covering a roughly 50km² region. It came with georeferenced map points of coconut trees. We paddled and broke up the image into more manageable 300 by 300 segments, yielding 5040 sub images and 11,330 labeled coconut trees in total. We then randomly shuffled these sub images to separate train (74%), 3,832 images, dev (17%), 6,64 images, and test (7%), 6,64 images sets. In addition we augmented the data by randomly flipping the images half the time.
- During our first round of testing and analysis, we found that there was a high incidence of unlabeled trees. After a round of clean up, we correctly labeled over a thousand more trees.

Model Attributes
- Batch Size: 1 (SGD)
- Regularization: L2
- Initialization: Variance scaling initializer
- IOU Threshold: 0.6
- Optimizer: Momentum
- Starting α: 0.0001
- Learning rate decay: Hand adjusted
- Data Augmentation: Random horizontal flip

Error Analysis
- In 160 dev images (multiple trees per image), the model made errors:
  - 10 times it found a banana tree, they have similar fronds.
  - 7 times it did not bound a coconut tree that was present.
  - 4 times it found the shadow of a tree.
- The training data were not well labeled, some banana trees were mislabeled coconut trees. A more extensive analysis is needed to determine the best way to decrease loss further. Some possibilities include:
  - Removing positive labels from banana trees.
  - Training longer will help with.
  - Not labeling shadows or correctly labeling coconut trees.
- Many unlabeled coconut trees were missed, the fronds were not like spokes.
- More training examples with labeled fronds.

Discussion
- We needed a detection algorithm in order to locate and count coconut trees. We initially trained on a Single Shot Detector with an inception base network in the hopes of developing a quick-to-train model without compromising on accuracy. However, after being unable to achieve the desired accuracy scores we decided to change architectures to a Faster-R-CNN, which performs better than SSD when detecting smaller to mid sized objects.
- One advantage to lightweight networks such as mobilenet is that they can be run on mobile devices. This could be a boon for relief efforts. To train such a network we would likely need more training data with more reliable labeling.
- Nonetheless, our object detector using a faster R-CNN accomplishes the challenge set out by the World Bank, an example of deep learning's applicability.

References