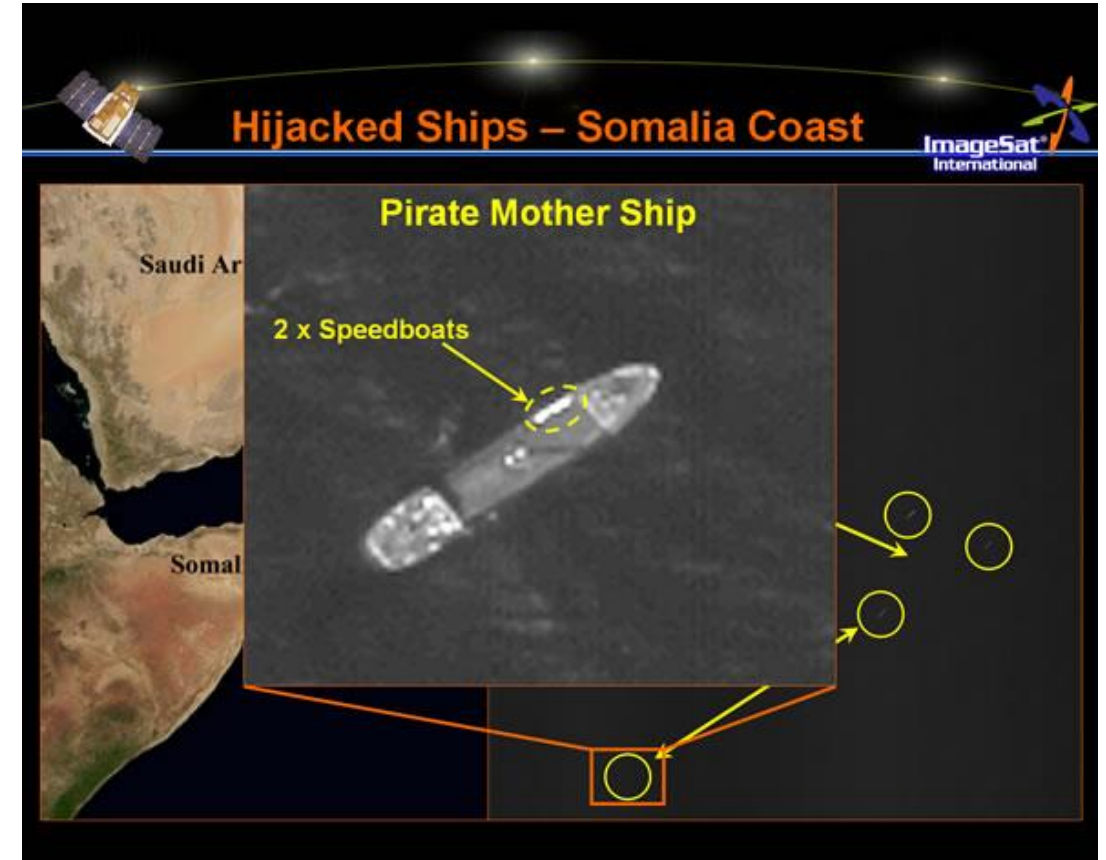# Identifying Ships in Satellite Imagery
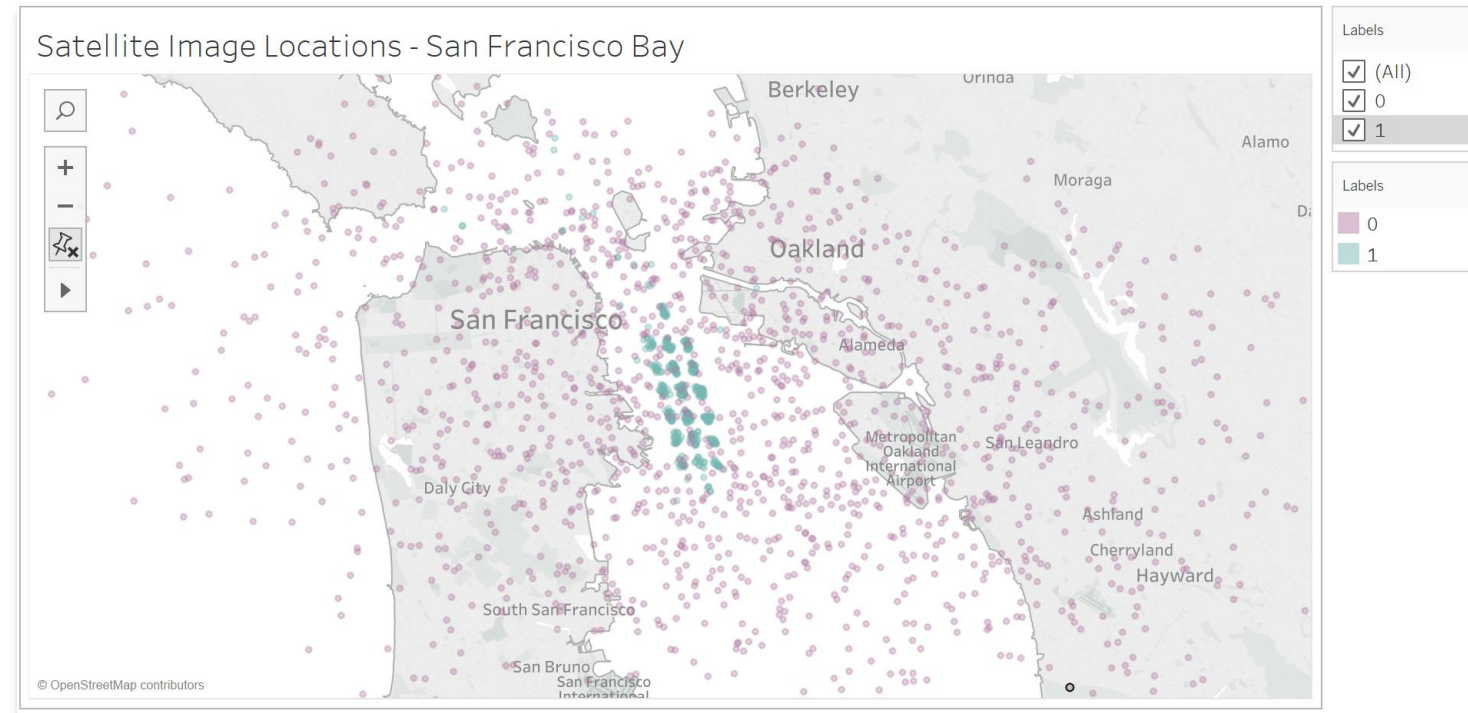
Business case

- Increased availability of data

- Border security

- Identification of pirate vessels

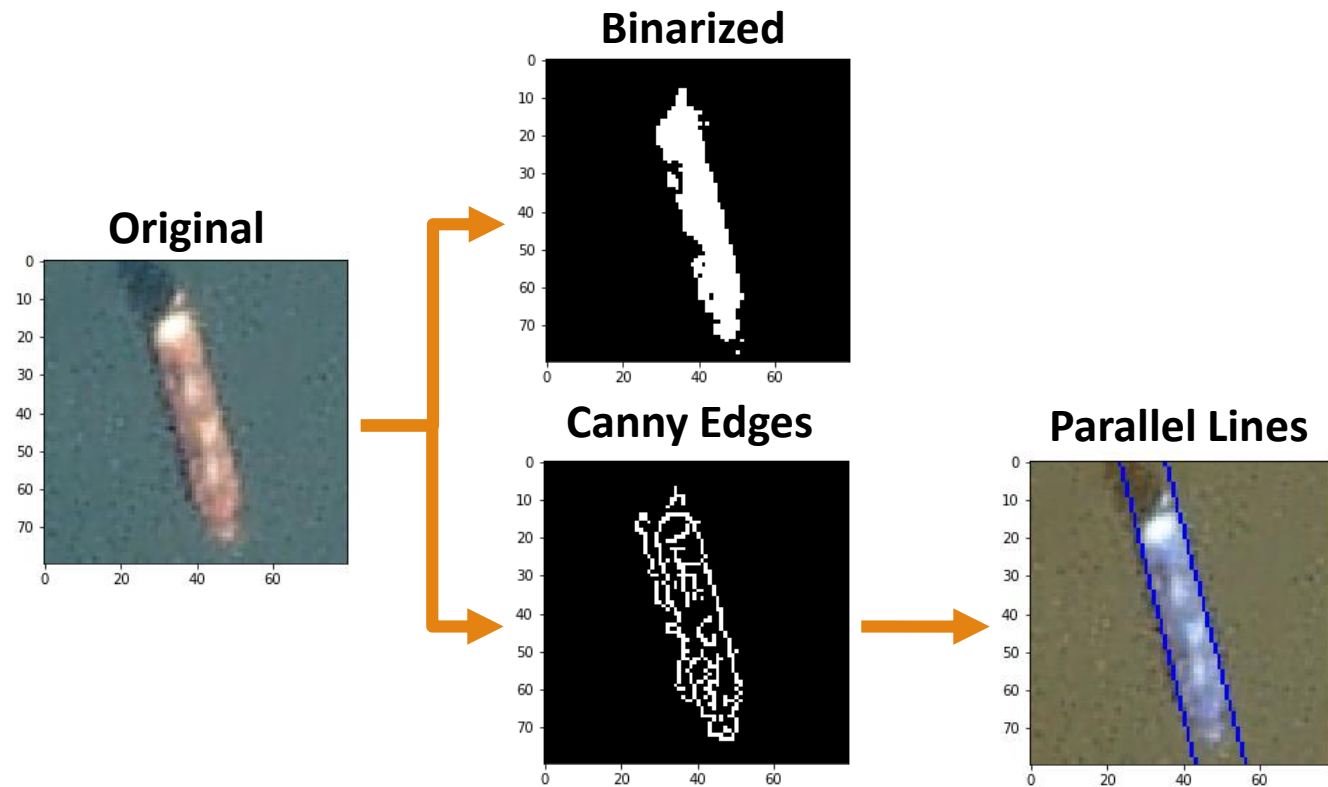- Environmental monitoring of ships

Dataset

- Imagery provided by Planet

- 2000 images, 500 labelled as ship images

- Cropped to size of ships

- Images made up of RGB layers, each 6400 pixels

Feature Extraction

Original

Binarized

Canny Edges

Parallel Lines

- Open CV for image processing
- Binary, edge ratios, and parallel lines
- RGB ratios
- Hu Moments
- PCA on raw pixels

Modeling

- Results similar for:
  - Gradient Boosting
  - AdaBoost
  - Extreme Gradient Boosting
  - Extra Trees
  - Random Forest

- No uplift using ensemble modeling

- Selected Gradient Boosting

ROC Curves
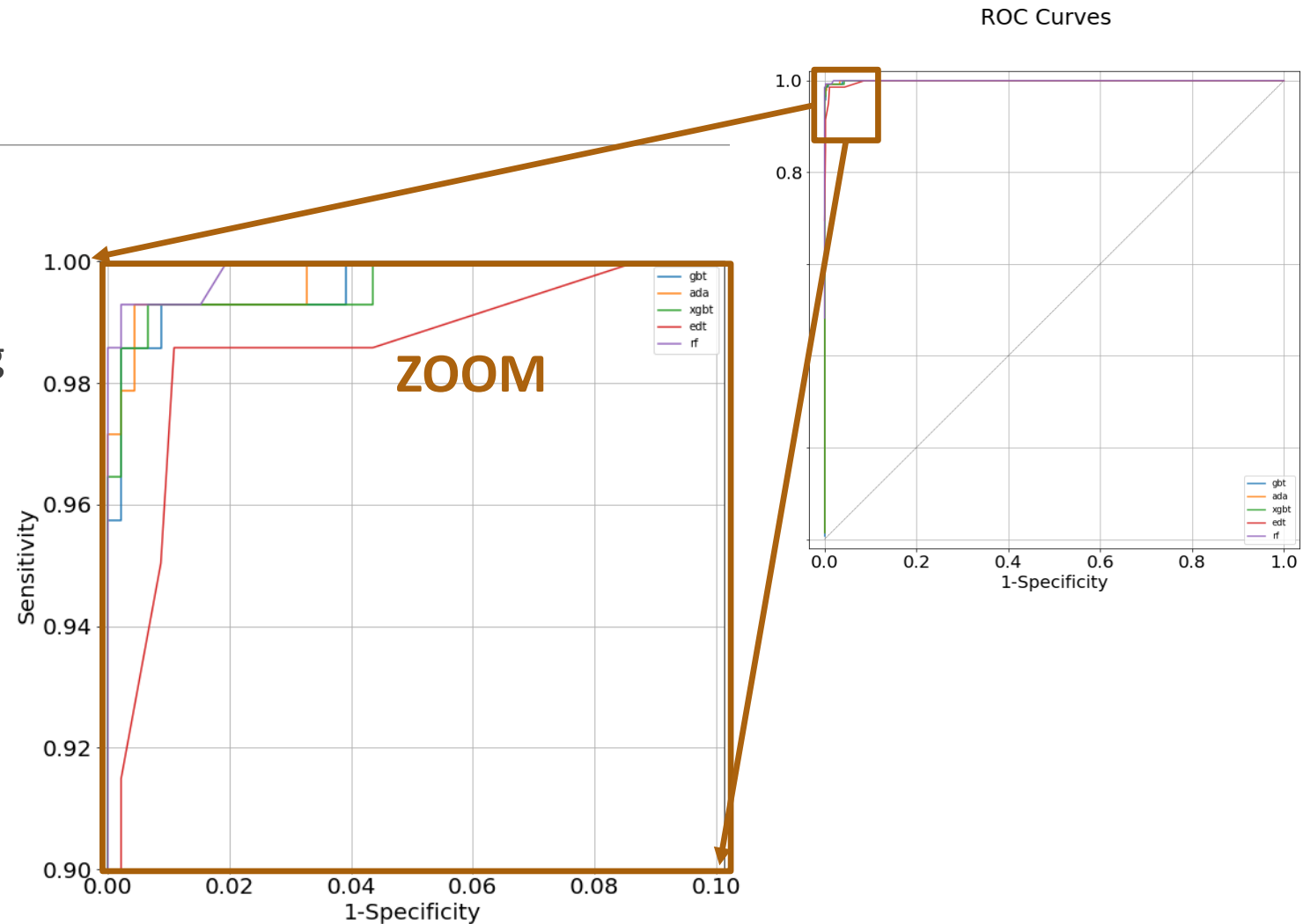
Gradient Boosting Evaluation

- Accuracy .996, recall .993

- PCA, edge, and binary ratio most influential features

- Scale invariant features, not as strong influence

Confusion Matrix

Feature Importance
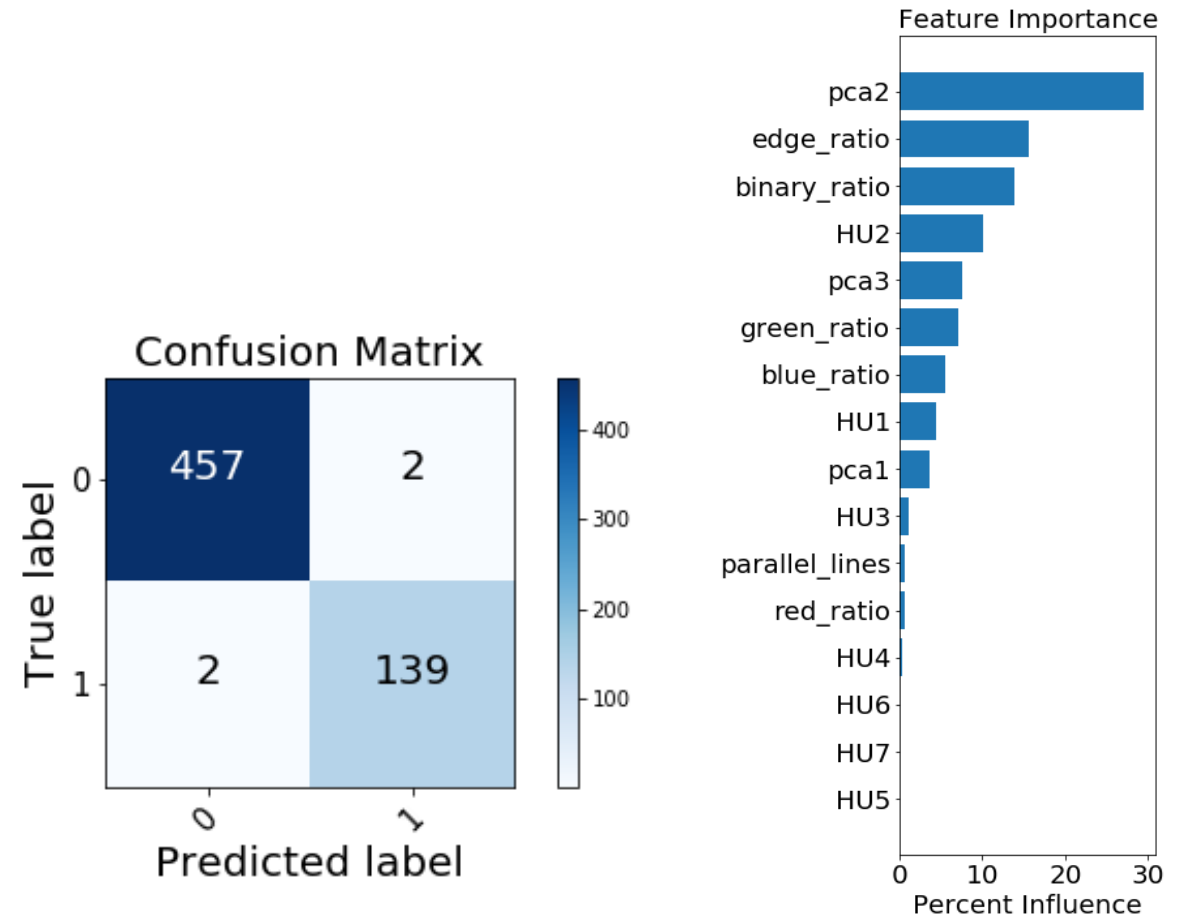
Future Work and Limitations

- Model may not generalize well to non-cropped satellite data

- Should be tested on smaller ships

- Higher resolution datasets available

- With timing of satellite capture one could potentially link AIS data with satellite capture