Smart Trash Net: Waste Localization and Classification

Vikram Sreedhar ~ Robel Mengistu ~ Oluwasanya Awe
vsreed@stanford.edu ~ robel@stanford.edu ~ oawe@stanford.edu
CS229, Stanford University

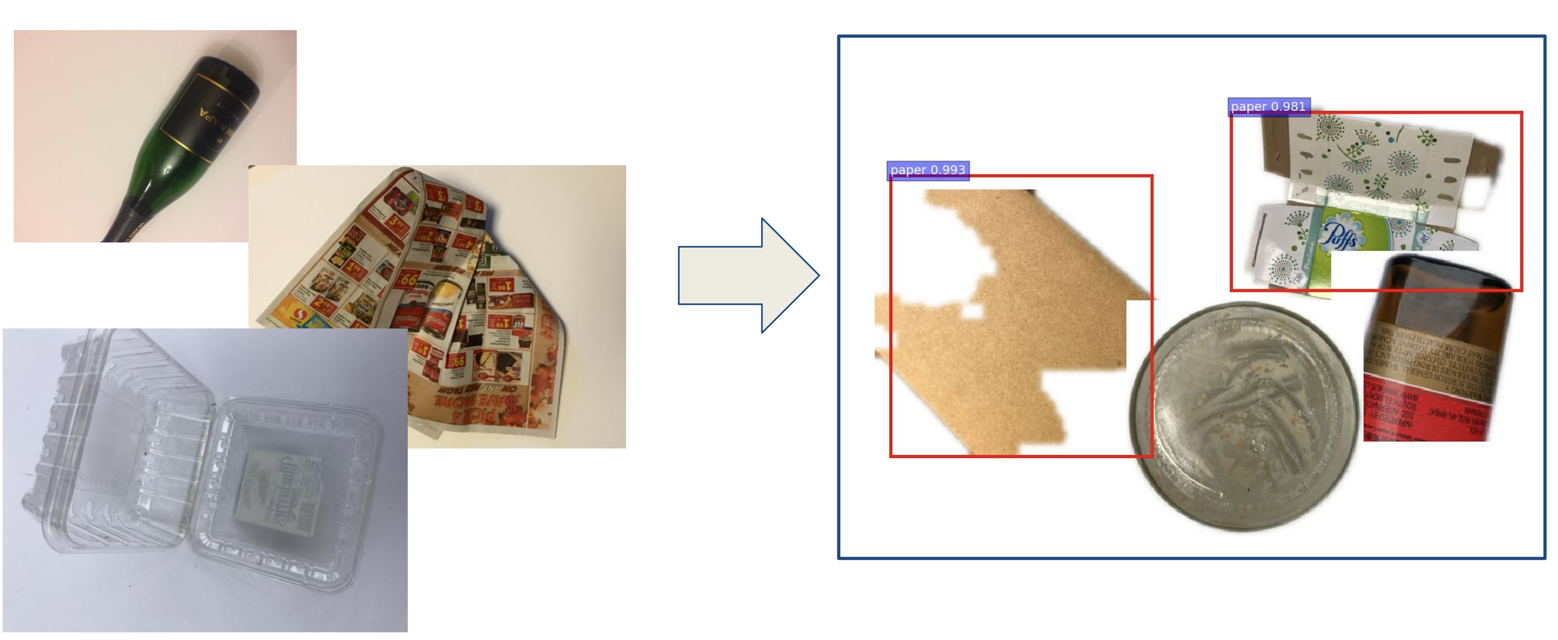
Background

Americans produce more than 250 million tons of waste every year. According to the Environmental Protection Agency, 75% of this waste is recyclable. However, currently, only about 30% of it is recycled. We want to increase this recycling rate by automating waste sorting.

Given an image of jumbled waste which contains two or more different pieces of waste of different types, we localized the image and classified the different forms of waste into three categories: recyclable, paper, and landfill. We utilized a fine-tuned Faster R-CNN to get region proposals and classify objects.

Dataset

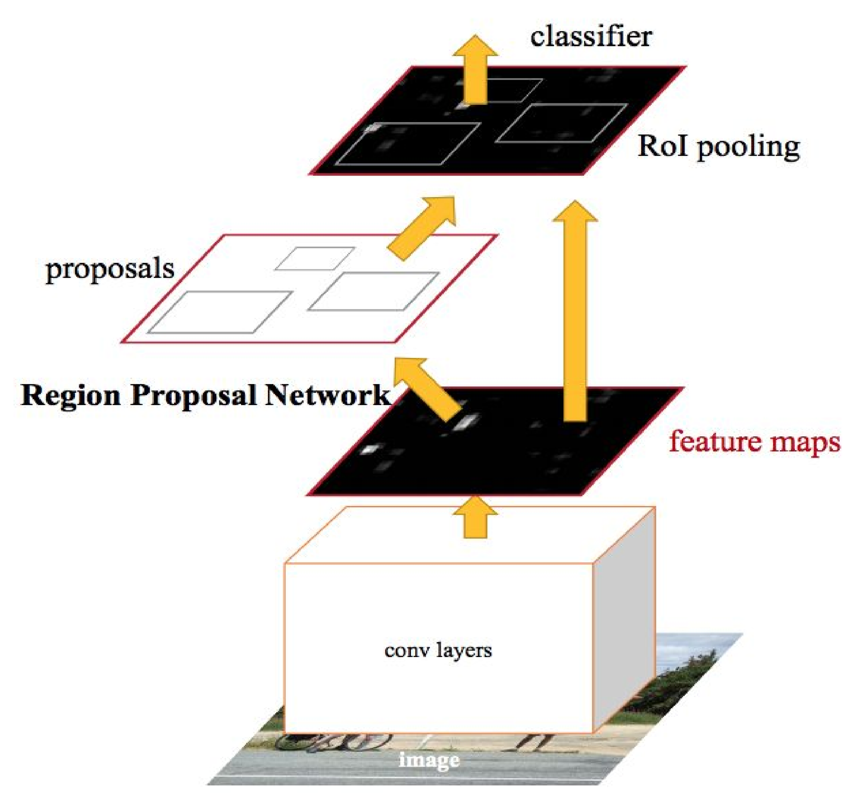
- 10,000+ computer generated images of piles of trash
- Created by augmenting Mindy Yang and Gary Thung's dataset of single pieces of trash
  - Background removal, Merge, and Labeling
  - Labels: bounding box coordinates for each trash



Faster R-CNN

Architecture:

- Has a Region Proposal Network (RPN) that allows nearly cost-free region proposals and share con layers with the detection network.
- RPN simultaneously predicts object bounds and objectness scores at each position



Loss Function:

$$L(\{p_i\},\{t_i\}) = \frac{1}{N_{cls}}\sum_i L_{cls}(p_i, p_i^*) + \lambda\frac{1}{N_{reg}}\sum_i p_i^* L_{reg}(t_i, t_i^*).$$

Fine Tuning

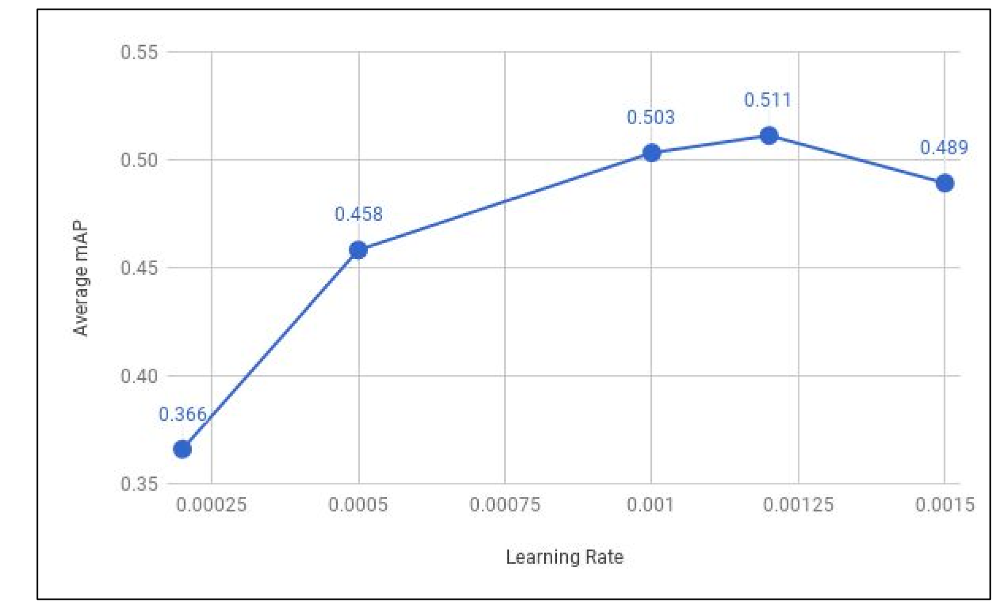
- Pre-trained lower level features, and image pixels
- Tuning the last fully connected layers — clsscore and bboxpred
- Trained for ~1000 iterations to find optimal hyperparameters

Results

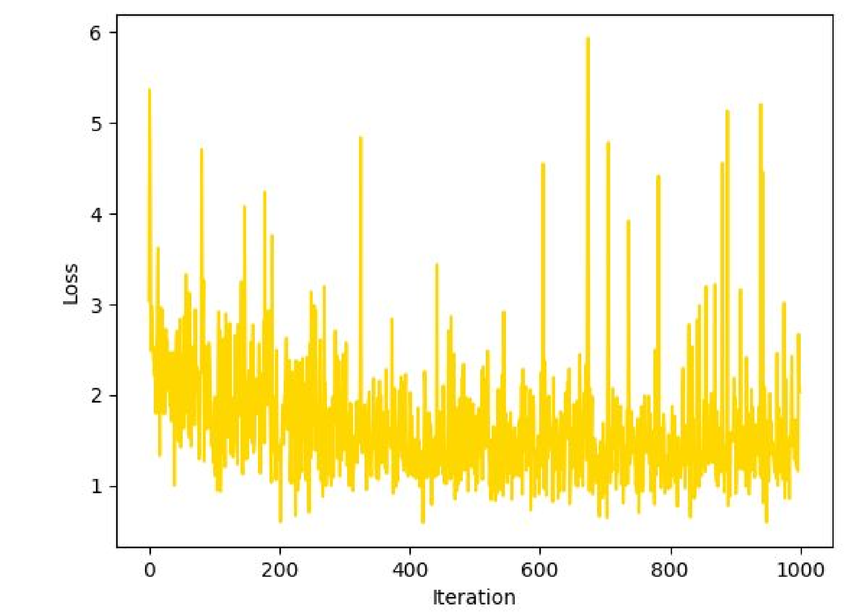
After training our model on our training data, we computed the mean Average Precision (mAP) to evaluate our performance.

Class	Baseline 2 images	AP : RPN	AP: learning rate	Final 2-6 images
Landfill	0.837	0.725	0.514	0.699
Paper	0.652	0.596	0.433	0.607
Recycling	0.849	0.767	0.585	0.744
mAP	0.7795	0.696	0.511	0.683

Results from hyperparameter tuning on baseline and final model. The baseline included a maximum of 2 images which explains the reason for the higher validation average.



Average mAP scores over the 3 classes for different learning rates



Loss curve for final model over 10K iterations

Discussion

Hyperparameters:

Some of the hyperparameters we optimized:

- RPN Batch Size
- Batch Size (Number of Regions of Interest)
- Learning Rate
- Model: RPN/SS + Fast R-CNN

Interpretation:

We had hoped to get a higher accuracy for each category, however this is due to our baseline using a larger dataset from PASCAL VOC rather than our custom dataset.

Looking Ahead

- Train from Scratch
- More hyperparameter search
- Test our model on real images of trash

References

Mark Everingham, Luc Van Gool, Christopher KI Williams, John Winn, and Andrew Zisser- man. The pascal visual object classes (voc) challenge. International journal of computer vision, 88(2):303–338, 2010.

Gary Thung Mindy Yang. Classification of trash for recyclability status. CS229 Project Report 2016, 2016.

Shaoqing Ren, Kaiming He, Ross B. Girshick, and Jian Sun. Faster r-cnn: Towards real-time object detection with region proposal networks. IEEE Transactions on Pattern Analysis and Machine Intelligence, 39:1137–1149, 2015.