RIT

Automatically Gather Address Specific Dwelling Images Using Google Street View

Chester F. Carlson CENTER for IMAGING SCIENCE

Authors:

Salman Khan

Carl Salvaggio

OUTLINE

- Significance - Industry Challenges
- Method
- Results

Current Major Challenges in Industry

- Building Inventory upkeep
- Existing public record databases incomplete on old buildings
- Current policies provide little to no incentive in sharing of privately owned data
- Expensive and lengthy process in developing accurate building inventory

AutoCrop

- **Cannot rely on images straight from the Google Street View API (Google SV API)**

AutoCrop

Raw Google SV API Images

Final Cropped Images

- **Employed 3 types of neural networks:**
  - Scene Classification
    - Pytorch-ResNet-50 pre-trained model - Places-365 data set
  - Object Detection
    - Keras-RetinaNet pre-trained model -Google’s OpenImagesV4 data set
  - Semantic Segmentation
    - Keras-PSPNet-101 pre-trained model – Cityscape data set

AutoCrop – Scene Classification

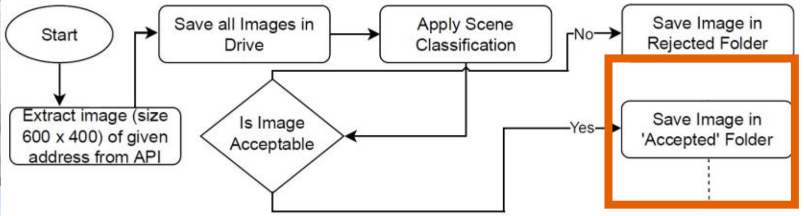
Campus

Alley

Construction

Bus Station

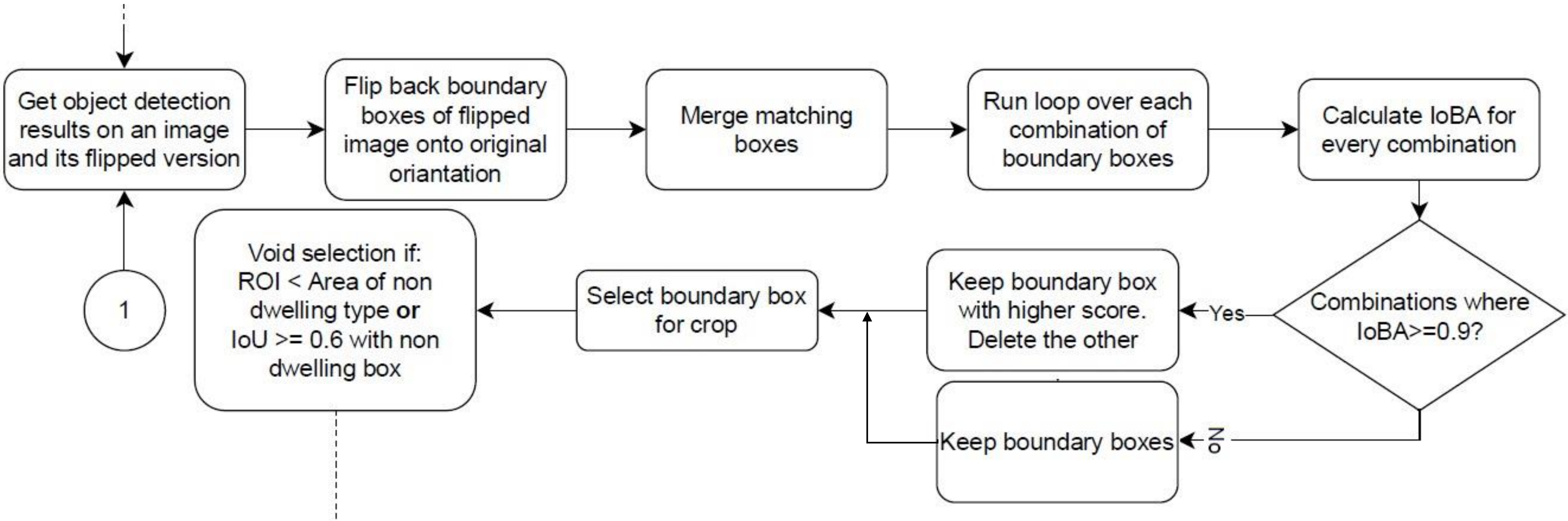
AutoCrop – Scene Classification



- **Accepted images from scene classification**

AutoCrop – Object Detection

- **Object Detection**



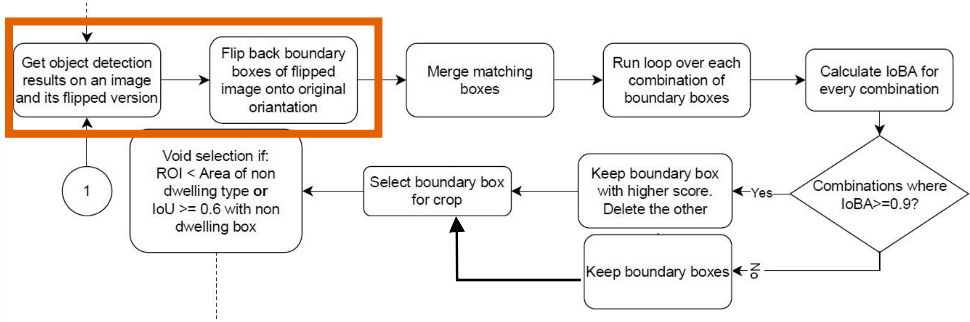
AutoCrop – Object Detection

(a) RetinaNet with ResNet50 backbone (b) RetinaNet with ResNet101 backbone (c) RetinaNet with ResNet152 backbone

Fig. 15: Object Detection Results of Networks Trained on *443* Labels Only (showing top 5 matches for each label to reduce clutter)

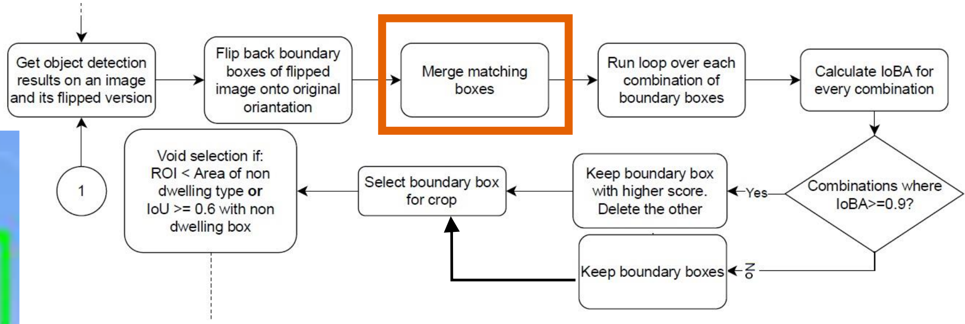
(a) RetinaNet with ResNet50 backbone (b) RetinaNet with ResNet101 backbone (c) RetinaNet with ResNet152 backbone

Fig. 16: Object Detection Results of Networks Trained on *All* Labels (showing top 5 matches for each label to reduce clutter)



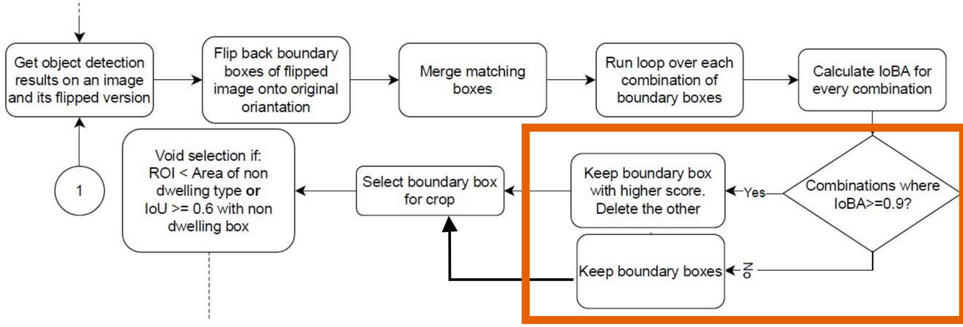
- **Individual Network Results**

AutoCrop – Object Detection



- Merged Result

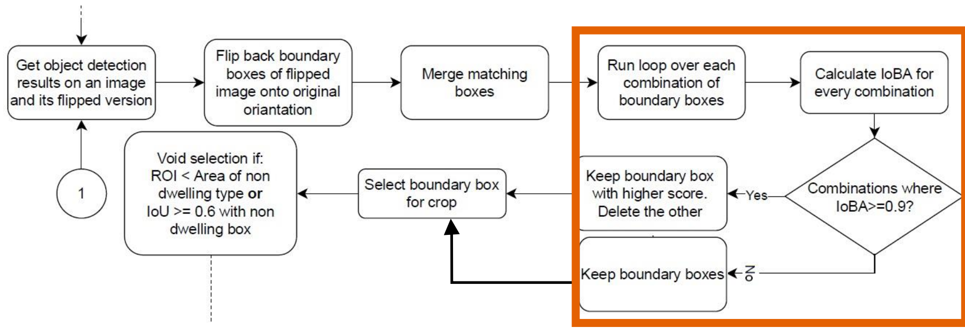
AutoCrop – Object Detection



- **After our NMS Strategy**

AutoCrop – Object Detection

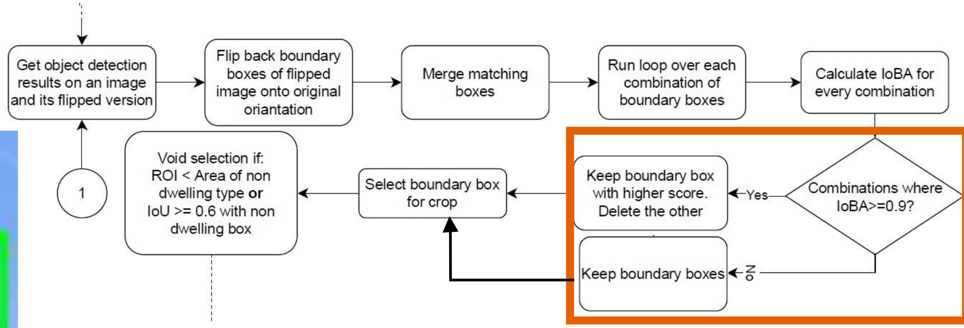
- **IoBA NMS Algorithm**



Algorithm 1 Bounding Box Reduction via Intersection over smaller Box Area (IoBA)

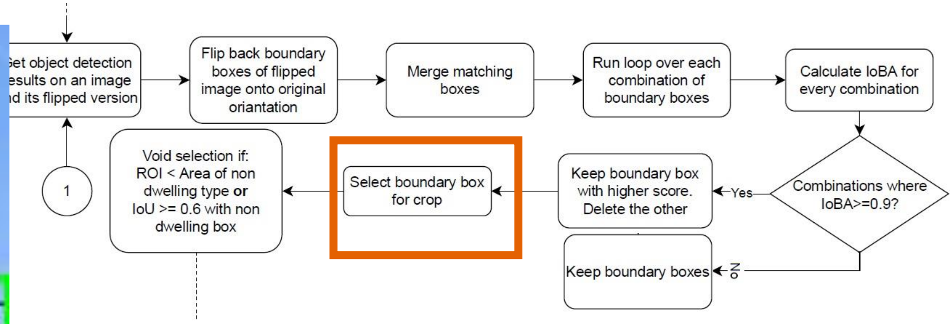
```
Read in Bounding Boxes(BB), Labels(L) and Scores (S)
for UniqueLabel do
    Combination(C) of every index
    if Del not in C then
        for Indices in C do
            Calculate area of boxA and boxB
            IoBA = Intersection/min(boxA,boxB)
            if IoBA >= 0.90 then
                Take BB with higher S
                Register deletion(Del) index
            end if
        end if
    end for
end for
```

AutoCrop – Object Detection



- **Per label, make combination of indices**
- **Per combination, calculate: IoBA = Area of Intersection/min(box-Area)**

AutoCrop – Object Detection

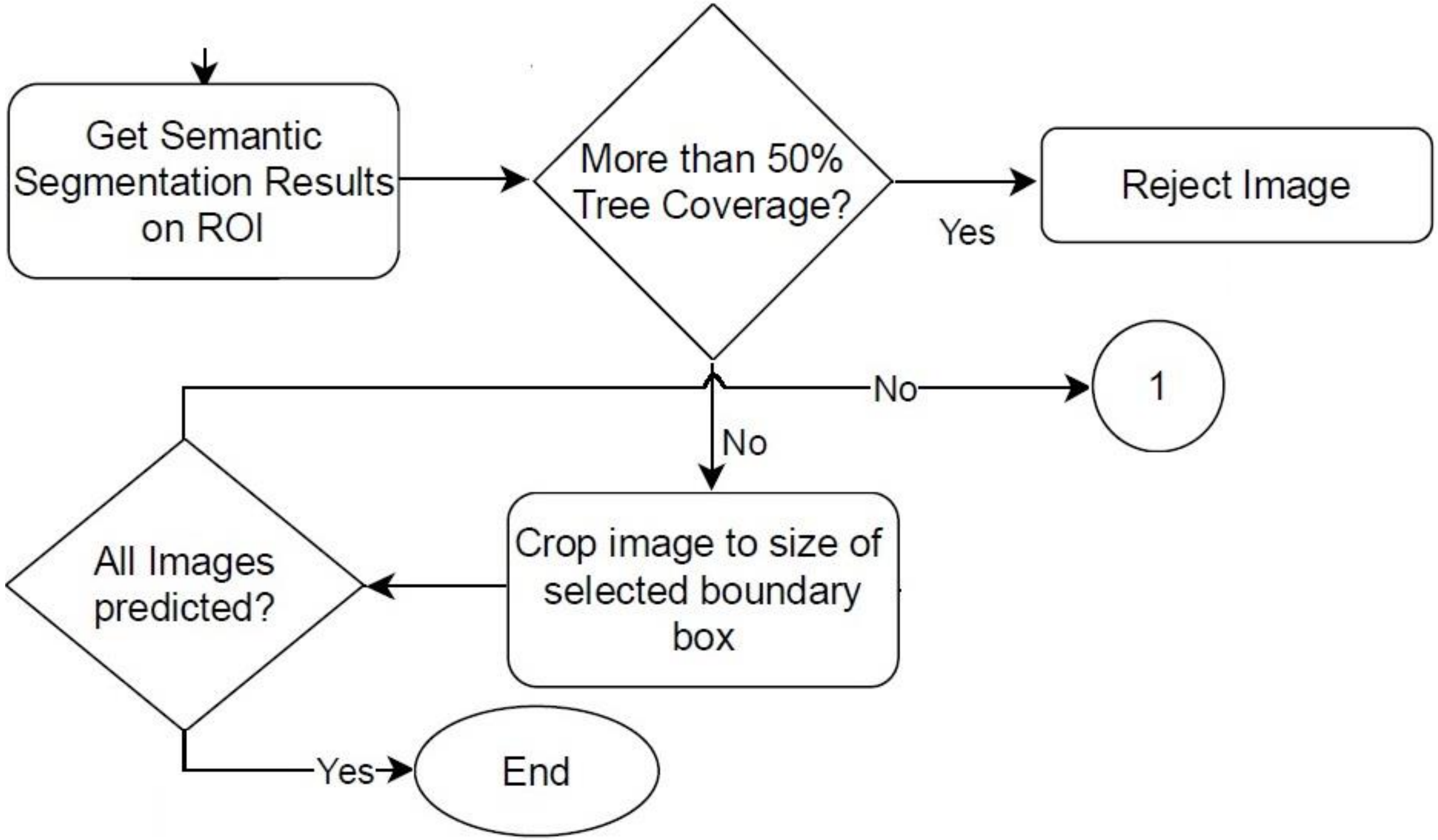


- **Selected Region of Interest**

3326 SE Milwaukie Ave, Portland, OR

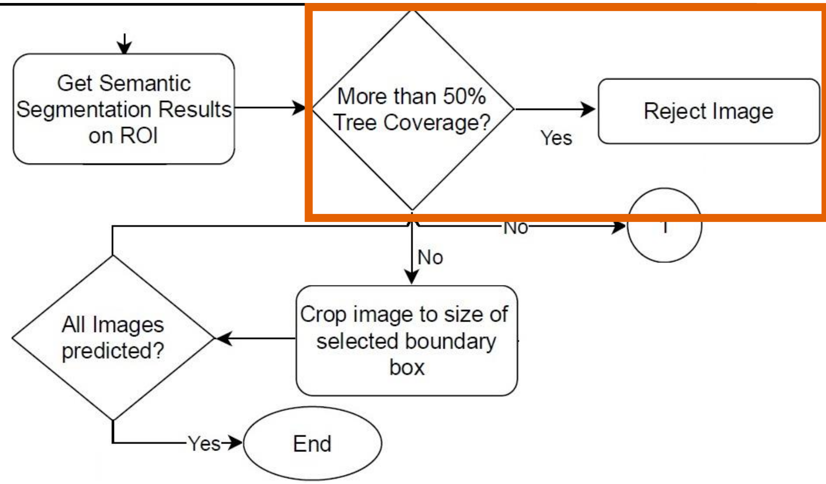
AutoCrop – Semantic Segmentation

- **Semantic Segmentation**



AutoCrop – Semantic Segmentation

- Images removed from final selection

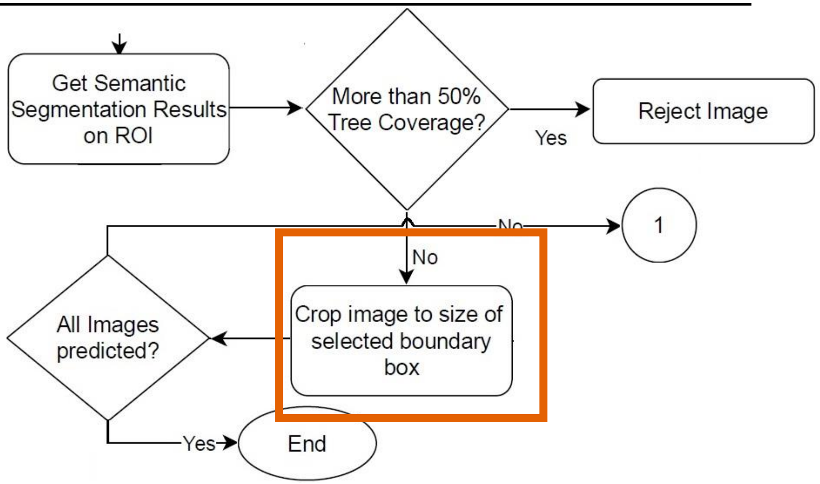


AutoCrop

(a) Merged Bounding Boxes (b) IoBA applied (c) 1921 Bonita Ave, Berkeley, CA

(d) Merged Bounding Boxes (e) IoBA applied (f) 1921 Delaware, Berkeley, CA

(g) Merged Bounding Boxes (h) IoBA applied (i) 2575 Le Conte Ave, Berkeley, CA



AutoCrop – Empirical Analysis

City	Total Images	Ac-cepted Scenes	Post NMS	Final Cropped Images	Not ROI	Err/ Prc/ Rcl
SF Bay Area	321	258	136	116	1	0.86%/ 99.14%/ 83.33%
Seattle	1106	749	453	405	1	0.25%/ 99.75%/ 84.34%
Portland	1620	993	587	579	3	0.52%/ 99.48%/ 74.32%
Santa Monica	1490	1191	322	257	4	1.57%/ 98.44%/ 71.07%
Vancouver	726	569	349	285	1	0.35%/ 99.65%/ 81.14%
Averaged Results:	**Avg Err:**	0.428%	**Avg Prc:**	99.29%	**Avg Rcl:**	78.84%

- **Final Empirical Results of AutoCrop**
- TP: `Final Cropped Images' - `Not ROI'
- FP: `Not ROI'
- FN: Misclassified images by Scene Classification Network + Images rejected by Semantic Segmentation but visually appear to have less than 50% tree occlusion
- Precision = TruePositives / (TruePositives + FalsePositives)
- Recall = TruePositives / (TruePositives + FalseNegatives)

AutoCrop – Empirical Analysis

City	Total Images	Accepted Scenes	Post NMS	Final Cropped Images	Not ROI	Err/ Prc/ Rcl
SF Bay Area	321	258	256	217	14	6.45%/ 93.55%/ 89.82%
Seattle	1106	749	673	475	23	4.84%/ 95.15%/ 84.96%
Portland	1620	993	954	704	59	8.34%/ 91.61%/ 73.13%
Santa Monica	1490	1191	1124	860	191	22.20%/ 77.79%/ 86.43%
Vancouver	726	569	558	348	14	4.02%/ 95.97%/ 82.2%
Averaged Results:	Avg Err:	9.17%	Avg Prc:	90.81%	Avg Rcl:	83.31%

- **If no checks and balances are applied**
  - Occlusion threshold

AutoCrop – Empirical Analysis

- **Empirical analysis on thresholds**

City	$Occ\_Thr_{0.6}$ & $IoBA_{0.85}$	$Occ\_Thr_{0.6}$ & $IoBA_{0.95}$	$Occ\_Thr_{0.55}$ & $IoBA_{0.9}$	$Occ\_Thr_{0.65}$ & $IoBA_{0.9}$	$\mathbf{Occ\_Thr_{0.6}}$ & $\mathbf{IoBA_{0.9}}$
SF Bay Area	$1/127 =$ 0.79%	$1/127 =$ 0.79%	$1/122 =$ 0.79%	$1/122 =$ 0.79%	$1/116 =$ 0.86%
Seattle	$16/411 =$ 3.89%	$5/403 =$ 1.20%	$6/411 =$ 1.46%	$6/411 =$ 1.46%	$1/405 =$ 0.25%
Portland	$4/548 =$ 0.73%	$4/540 =$ 0.74%	$4/545 =$ 0.74%	$4/545 =$ 0.74%	$3/579 =$ 0.52%
Santa Monica	$9/279 =$ 3.22%	$12/278 =$ 4.31%	$9/279 =$ 3.22%	$9/279 =$ 3.22%	$4/257 =$ 1.57%
Vancouver	$3/303 =$ 0.98%	$3/303 =$ 0.99%	$3/308 =$ 0.97%	$3/308 =$ 0.97%	$1/285 =$ 0.35%
Avg Error	1.92%	1.61%	1.44%	1.44%	0.428%