

Object Detection with R-CNN

Ross Girshick, Jeff Donahue, Trevor Darrell, and Jitendra Malik
University of California, Berkeley

Project Presentation By
Albert Haque and Fahim Dalvi

June 3, 2015

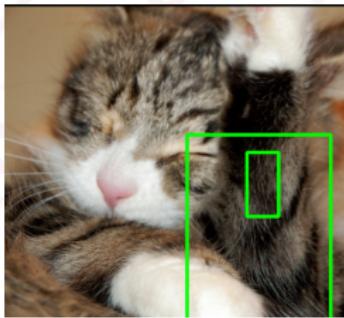
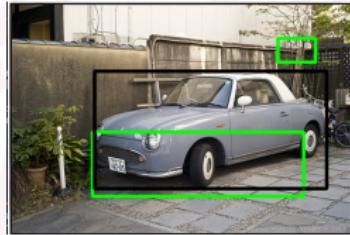
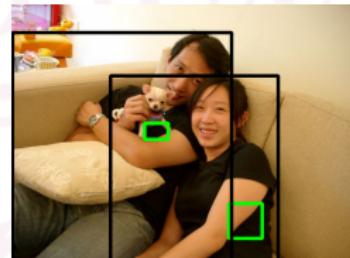
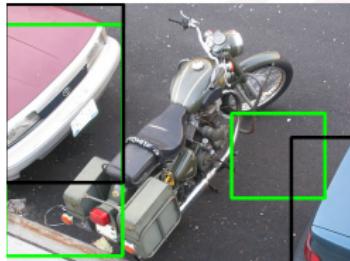
Outline

- ▶ Baseline Results
 - ▶ Iterative SVM
 - ▶ Full SVM
- ▶ Extensions
 - ▶ SGD based implementation
 - ▶ VGG Network

Baseline Results

- ▶ Setup
 - ▶ Validation set (100 images)
 - ▶ Zero-mean unit-variance normalization
 - ▶ ℓ_2 regularization loss
 - ▶ Number of negatives: 30,000 for iterative SVM
 - ▶ Positives weighted 10 times more than negatives
- ▶ Results
 - ▶ Full SVM gives no improvement over Iterative SVM
 - ▶ 65% accuracy on the validation set

Baseline Results



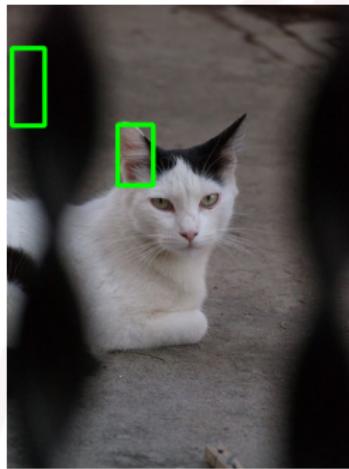
Baseline Results



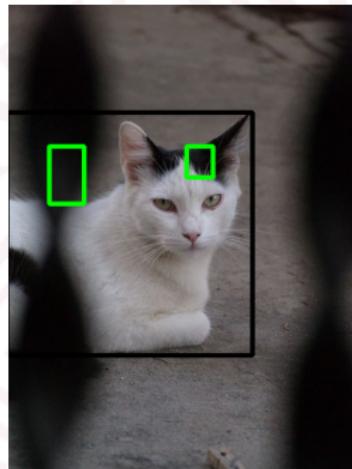
Extensions

- ▶ SGD Classifier
 - ▶ Trained iteratively using small sets of images
 - ▶ Same setup as SVM
- ▶ Results
 - ▶ SGD classifier does half as worse as the SVM
 - ▶ 55% accuracy on the validation set

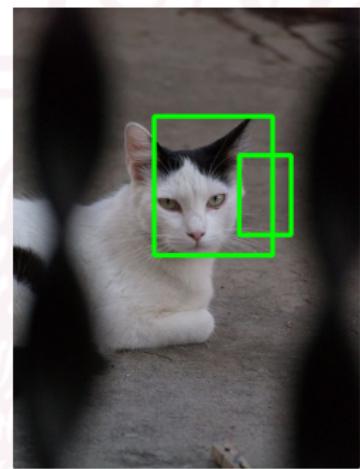
Extensions



Car



Cat



Person

Extensions

- ▶ VGG Network
 - ▶ Features extraction is complete (6 hours on 20 GPUs)
 - ▶ Currently using fc7 non-rectified features
 - ▶ SVM training is in progress