Real-time Object Detection
Zibo Gong, Tianchang He, Ziyi Yang
Department of Electrical Engineering, Stanford University

Introduction
- Object detection: one of the classic problems in computer vision.
- Convolution Neural Networks (CNN): one of the most powerful tools now for Artificial Intelligence and Machine Learning problem.
- We used Faster Region-based Convolutional Neural Network method (Faster R-CNN) to detect certain objects in photos and matched with one object in our database. The matching of objects adopts features learned by Faster R-CNN or conventional CV features, for instance, histogram of oriented gradients (HOG). As a result, we can realize the detection and matching of objects (e.g. cars) in pictures.

Faster R-CNN
Training the Faster R-CNN.
- Train the region proposal network (RPN) end-to-end by back-propagation and stochastic gradient descent.
- Train a separate detection network by Fast R-CNN using the proposals generated by the step-1 RPN for 20k mini-batches.
- Use the detector network to initialize RPN training for 40k mini-batches.
- Keeping the shared convolutional layers fixed, fine-tune the unique layers of Fast R-CNN for 20k mini-batches.
R-CNN output: bounding box with confidence level.

Object Detection
- Detection with p(object | box) >=0.8
- Color detected
- Detection of car in different perspective

Matching
- Features
  - Color
  - Keypoint regions (SIFT)
  - Histogram of oriented Gradients (HoG)
  - ...
- Algorithm
  - K-nearest neighbor (K-NN)
  - Distance: Euclidean/Cosine

Application Pipeline
- Input video/image
- Feature Extraction
- Detection using neural network
- Output matches from database

Perspective | Average Precision
---|---
Front | 81.57%
Back | 86.40%
Left | 87.56%
Right | 79.31%

Reference