Deep Learning Strong Parts for Pedestrian Detection

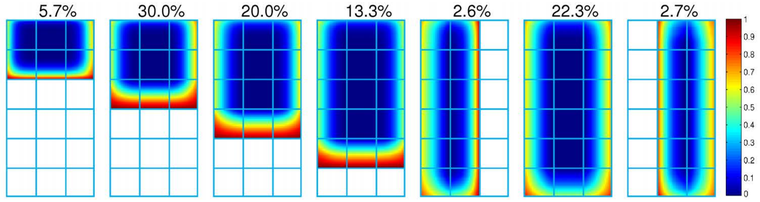
Yonglong Tian, Ping Luo, Xiaogang Wang, and Xiaoou Tang
The Chinese University of Hong Kong
{ty014, pluo, xtang}@ie.cuhk.edu.hk, xgwang@ee.cuhk.edu.hk

ICCV15
International Conference on Computer Vision

1. Motivation

Background

- **Occlusion distribution of Caltech Pedestrian dataset.** Over 97% of occluded pedestrians belong to a small subset of hundreds of possible occlusion types.

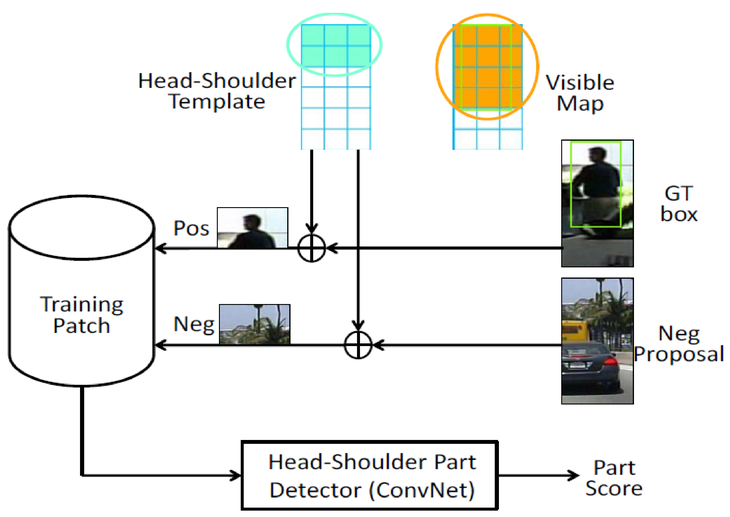


- **Occlusion types may vary in different scenes.**

Idea

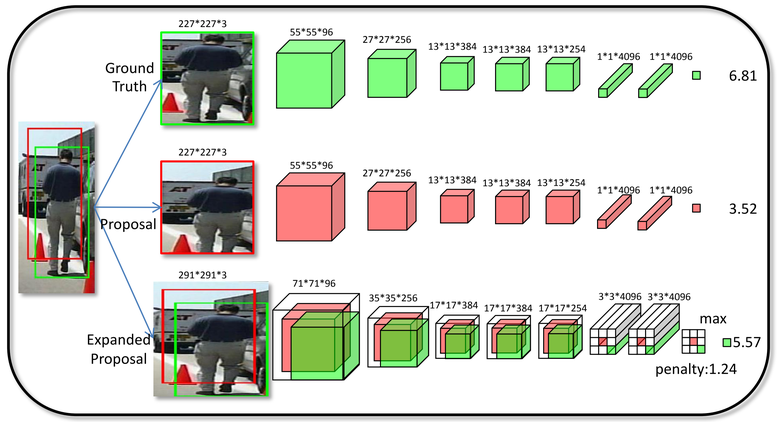
- **Part Detector Pool.** We construct a detector pool which includes part detectors of all sizes and positions.
- **Part Selecting.** We select significant and complement part detectors in different scenes or datasets.

2.1 Training Part Detectors



- **LDCF Proposals.** Use LDCF detector to propose candidates
- **Part Patches.** Crop positive and negative part patches from positive and negative candidates, respectively.

2.2. Handling Shifted Proposals



- **Fully Convolutional Neural Network.** Reformulate the fully connected layers as convolutional layers.
- **Expanded Proposal.** Given a proposal (x_{min}, y_{min}, w, h), crop it as $(x'_{min}, y'_{min}, w', h')$, where

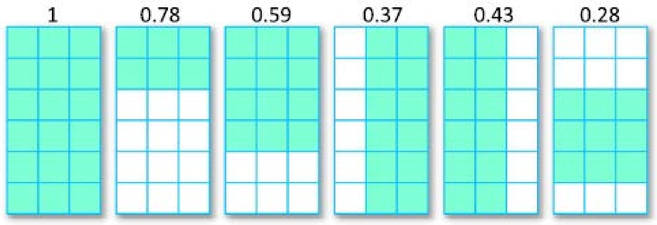
$$x'_{min} = x_{min} - \frac{16n}{227} \times w, \;\; y'_{min} = y_{min} - \frac{16n}{227} \times h,$$
$$w' = (1 + \frac{32n}{227}) \times w, \;\; h' = (1 + \frac{32n}{227}) \times h.$$

- **Score Adjustment.** The final score of this proposal is adjusted by

$$s = \max_{1 \le i,j \le n+1} \{S_{i,j} - P_{i,j}\}$$
$$P_{i,j} = a \times (|i - \frac{n+2}{2}| + |j - \frac{n+2}{2}|) \times \frac{32}{227}$$
$$+ b \times (|i - \frac{n+2}{2}|^2 + |j - \frac{n+2}{2}|^2) \times (\frac{32}{227})^2$$

2.3. Complementary Parts

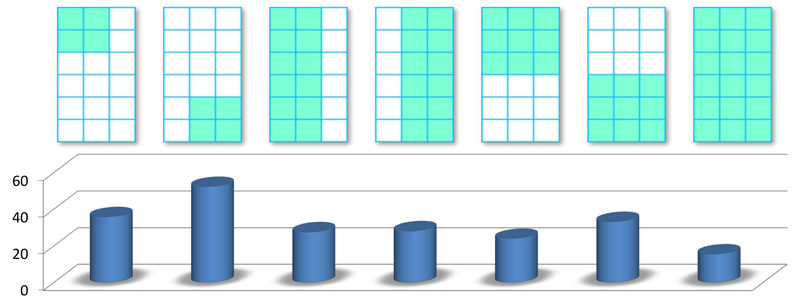
- **Complementary Parts.** Combine part scores; Select significant parts according to combinational weights; Re-learn the complementary weights.



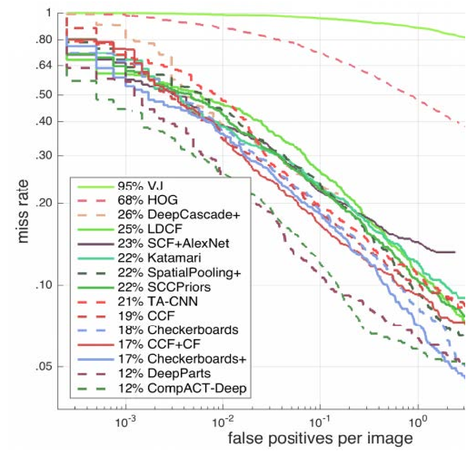
Parts' weights/importance on Caltech Pedestrian

3 Results

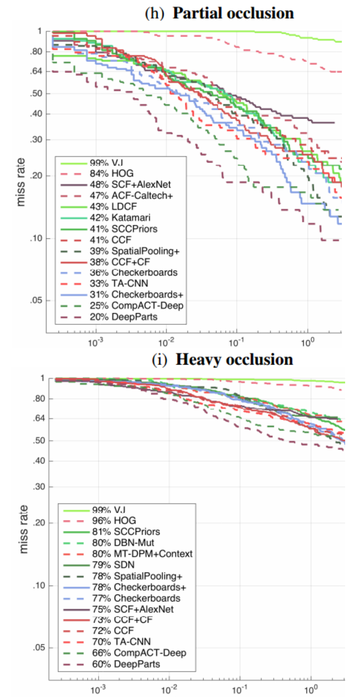
- **Effectiveness of Part Detectors**



- **Overall Results**



Reasonable Subset



- **KITTI**