Finding the Weakest Link in Person Detectors

Devi Parikh (TTIC) and Larry Zitnick (Microsoft Research)

Introduction

Person detection: Useful, Difficult

We've come a long way...

Fischler and Elschlager, 1973

Still a ways to go...

Dollar et al., 2009

How does it work?

Feature selection → Part detection → Spatial model → NMS / context

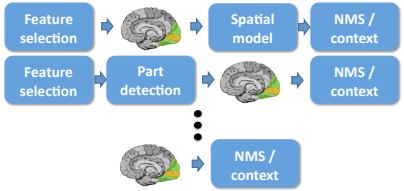
Felzenszwalb et al., 2005

Hoiem et al., 2006

How can humans help?

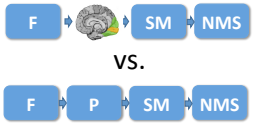
- Supply training data: 100,000s labeled images
- We design the algorithms: Going on 40 years
- **Can we use humans to debug?**

Human Debugging

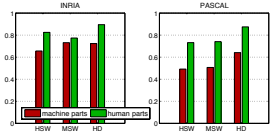


All machine experiments use Felzenszwalb et al. 2010

Parts



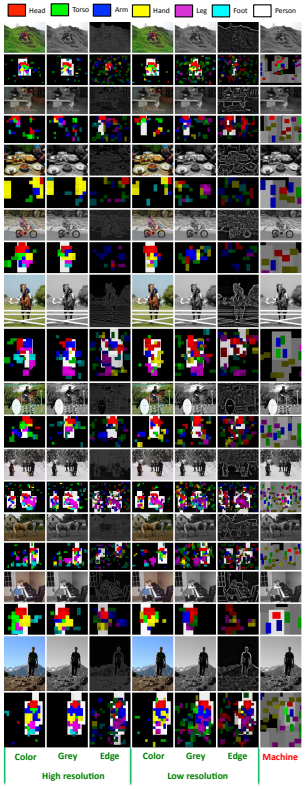
Human experiments



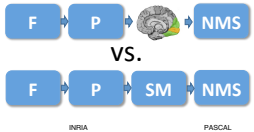
Results

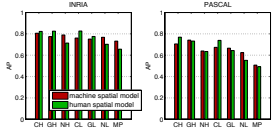
Part Patch Dataset (available online!)

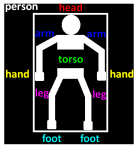
- Labeled patches in high and low resolution, color, grey-scale and edges.
- Six parts (head, torso, arm, hand, leg, foot) and root.
- 10 human subjects labeled each patch.
- 50 images from INRIA dataset and 100 images from PASCAL dataset. (0.3 million patches total)

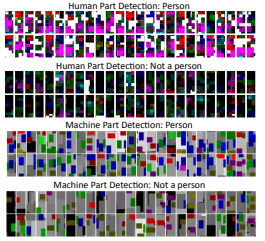


Spatial model





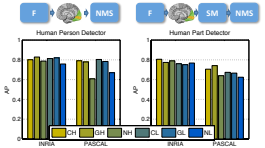




Training data used for human studies

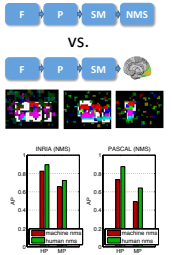
Features

What information is useful for recognition?

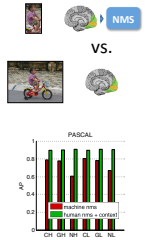


NMS / Context

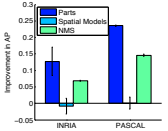
NMS



NMS + Context



Contributions



- Part detection is the weakest link in person detection, followed by NMS.
- Spatial models were not found to be crucial.
- Part Patch Dataset is available online.
- Human Debugging
  - Applicable to other visual recognition problems
  - Useful for other problem domains and AI in general