Open problem: What should we infer or label?

© 2019 Philipp Krähenbühl and Chao-Yuan Wu

Computer vision tasks

- Proxy for down stream applications
- Few are useful by themselves
- How good are our proxies?

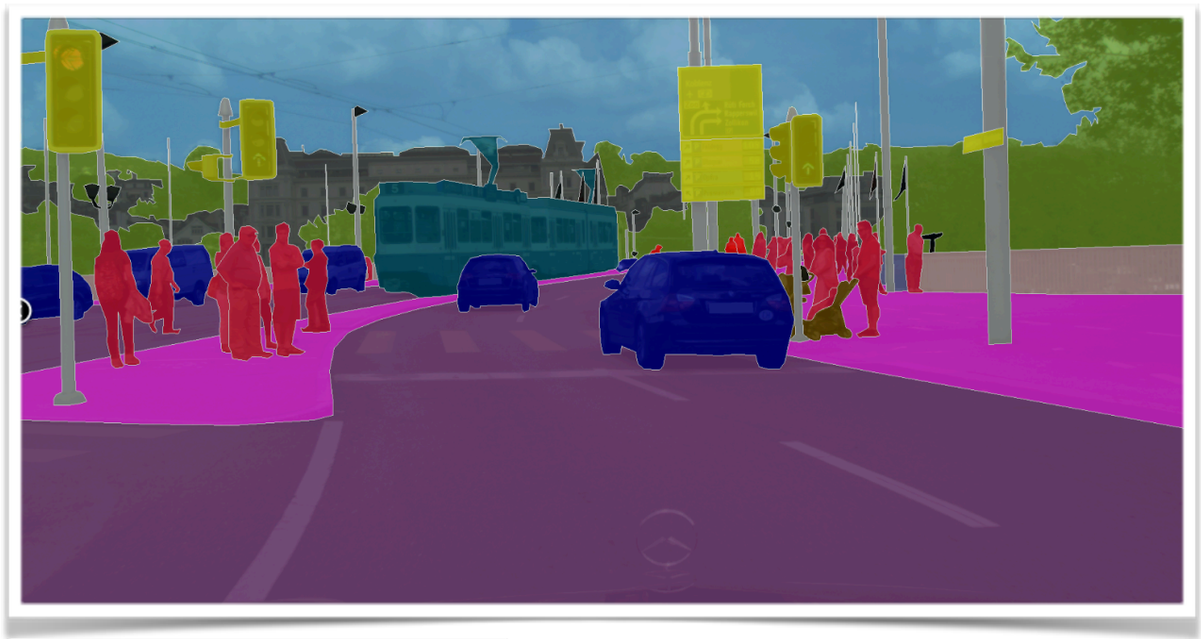
Example: Assistive technologies

- Object detection
  - Goal: How many elderly people do you detect?

- Down stream goal
  - How quickly and accurately do you detect a dangerous situation?
  - How much more efficient do you make the caretaker?

Example: Autonomous driving

- Semantic segmentation
  - Goal: Label every pixel with a class
- Down stream task:
  - Get from A to B without crashing

What do computer vision tasks provide?

- An efficient proxy to design and tune models
- Pre-training
- A compact and generalizable representation