

# Representational Challenges of Recognition, from Detection to Interpretation

NSF Frontiers in Computer Vision Workshop

Derek Hoiem

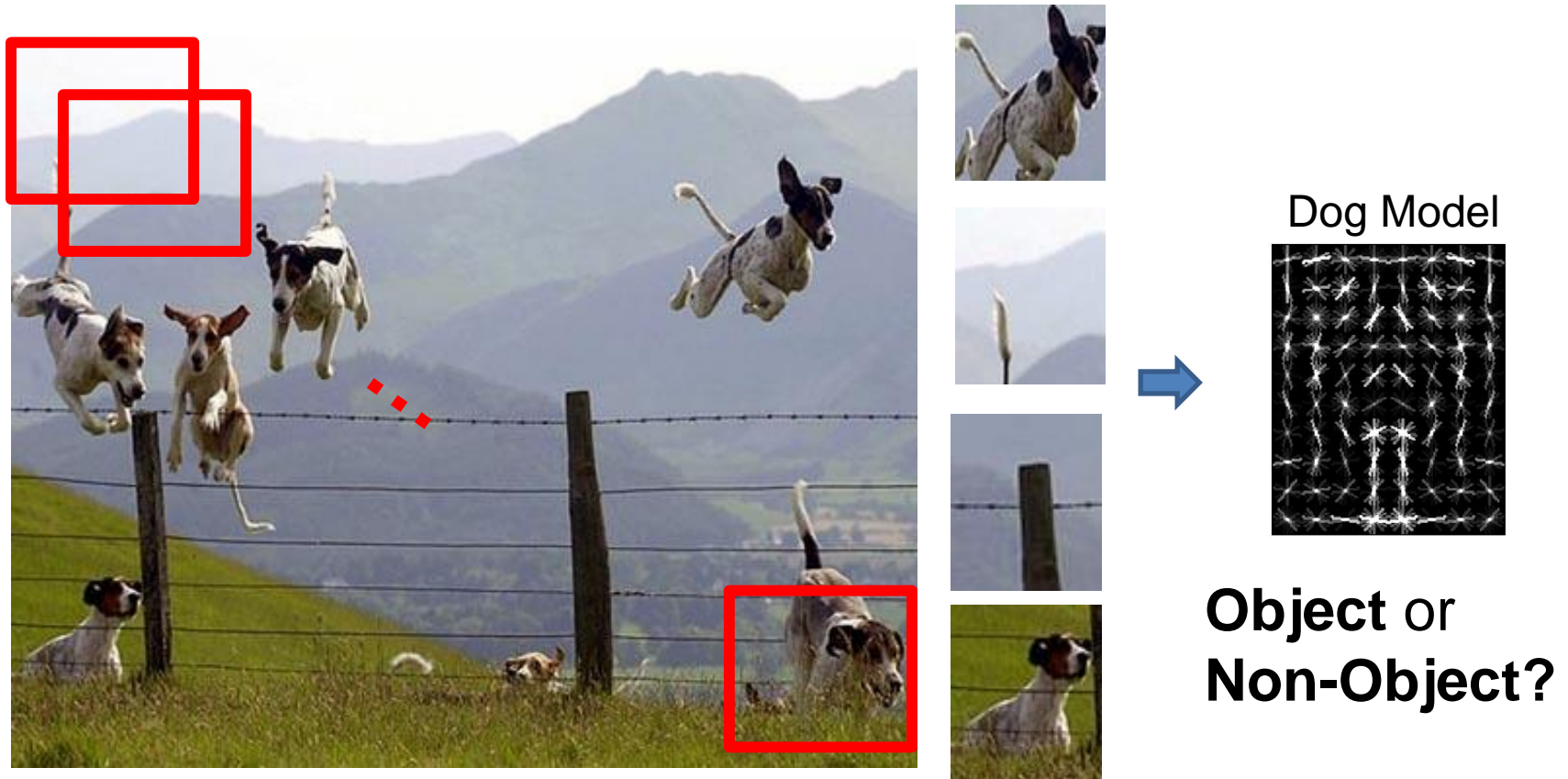
University of Illinois (UIUC)

Aug 2011



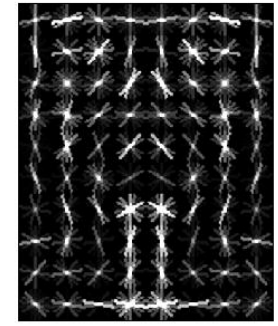
# Recognition in last 15 years

- Focus on object search: “Where is it?”
- Build templates that quickly differentiate object patch from background patch



# Template Matching Problem

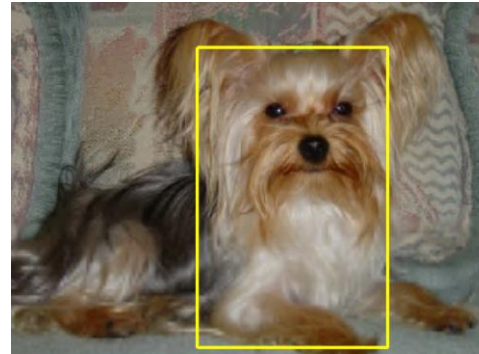
Dog Model



True  
Detections



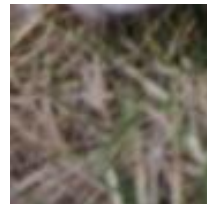
Bad  
Localization



Confused with  
Similar Object



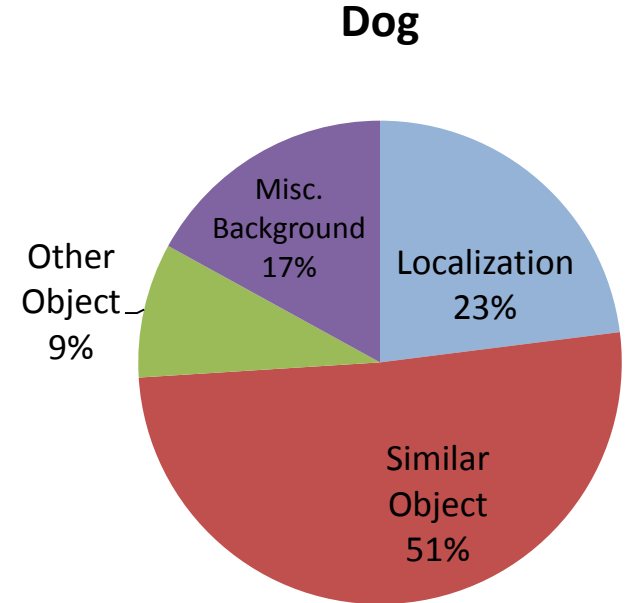
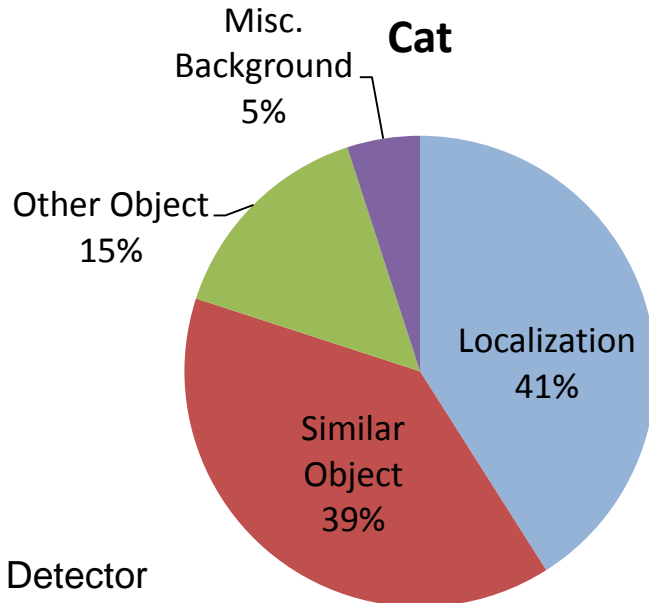
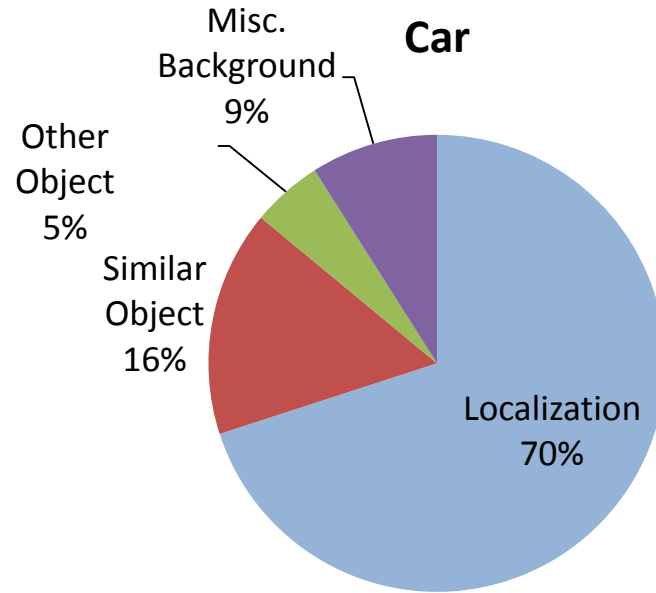
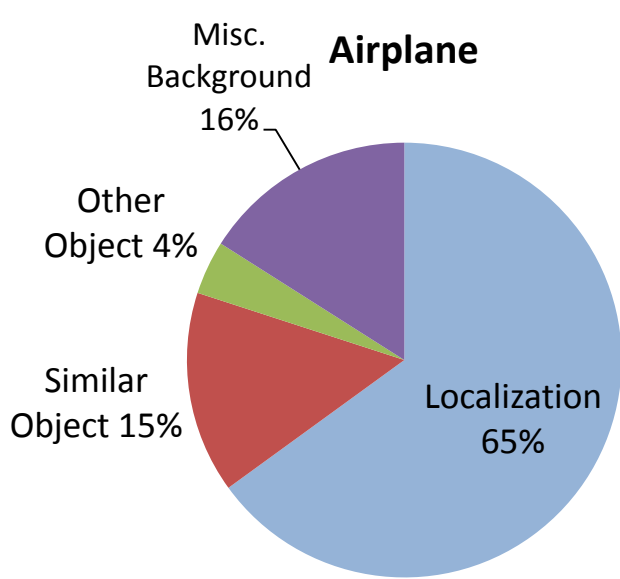
Misc. Background



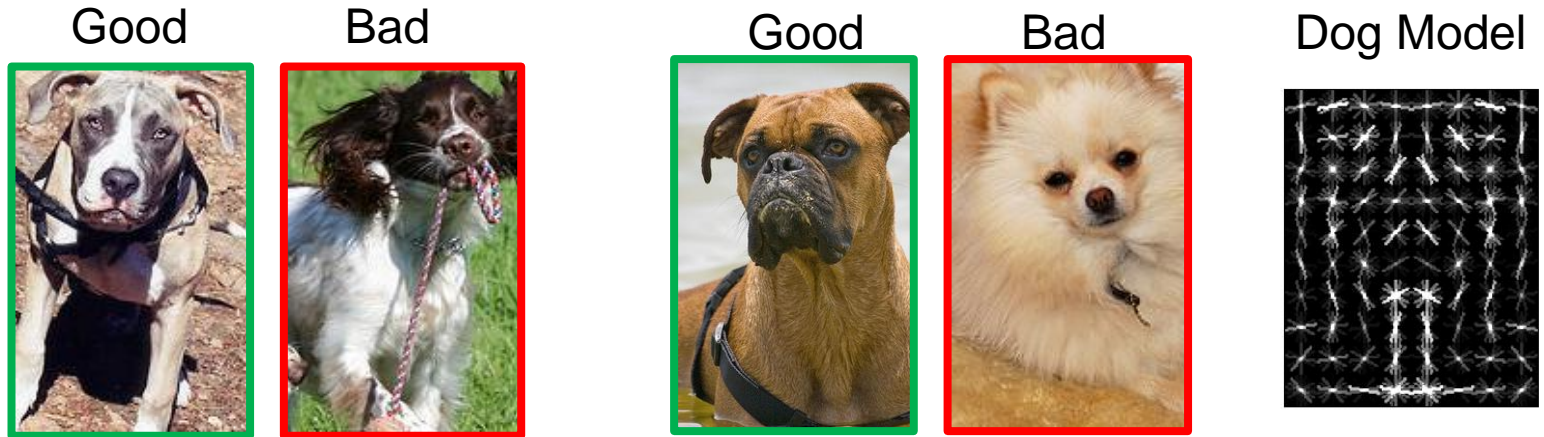
Confused with  
Dissimilar Object



# Breakdown of top 100 false positives



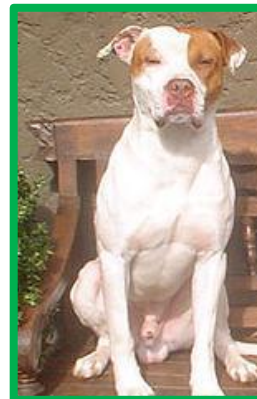
# Key Challenge: localize the object from a detection



Need good category-sensitive segmentation methods

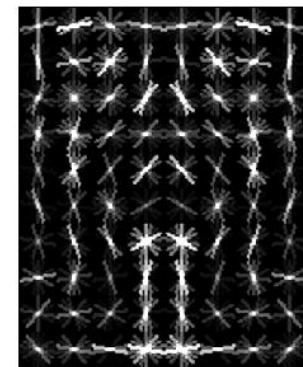
Can free up detectors to focus on discriminative pieces

# Key challenge: differentiate between similar categories



Robustness through learned abstraction (e.g., shape), rather than hand-coded invariance

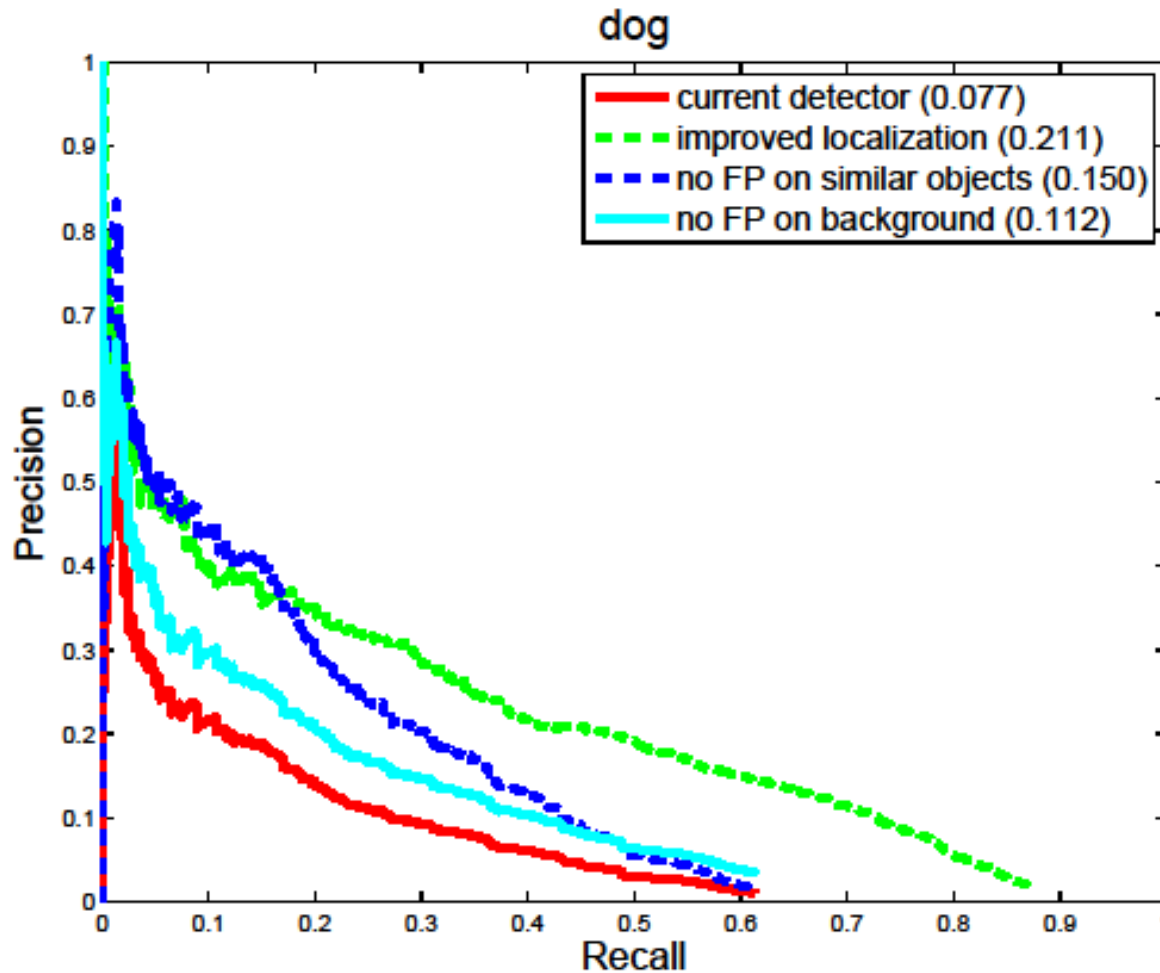
Compare details, rather than holistic appearance



Dog Model

To get large improvements, we need to solve the “mid-level” problems

## Potential Gains in Precision-Recall



# Object Recognition Challenges

- Last 15 years: object detection
  - Good methods to detect objects, ignore background
  - Better segmentation and mid-level representations are crucial for further improvement
- Next 10+ years: object interpretation
  - How do we represent objects themselves?





Key Challenge: How do we deal with objects that we can't categorize?



How to localize objects without categorization?

How to build representations that apply to novel objects?

Key Challenge: build/infer representations that encode physical context



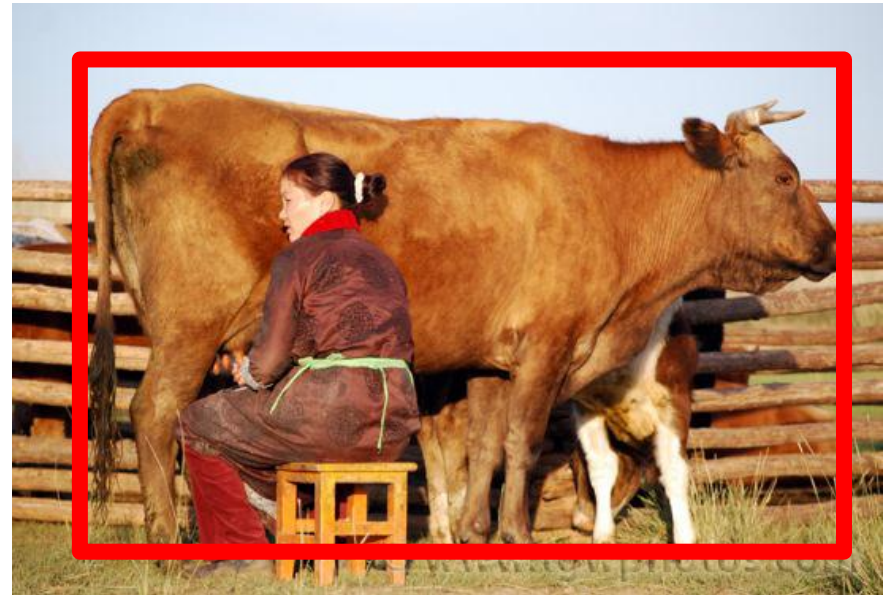
How to infer physical relations (contact, engagement, etc.)?

How to interpret an object's role in the scene?

# Key Challenge: build/infer representations that depend on task context



**Big animal ahead,  
moving left**

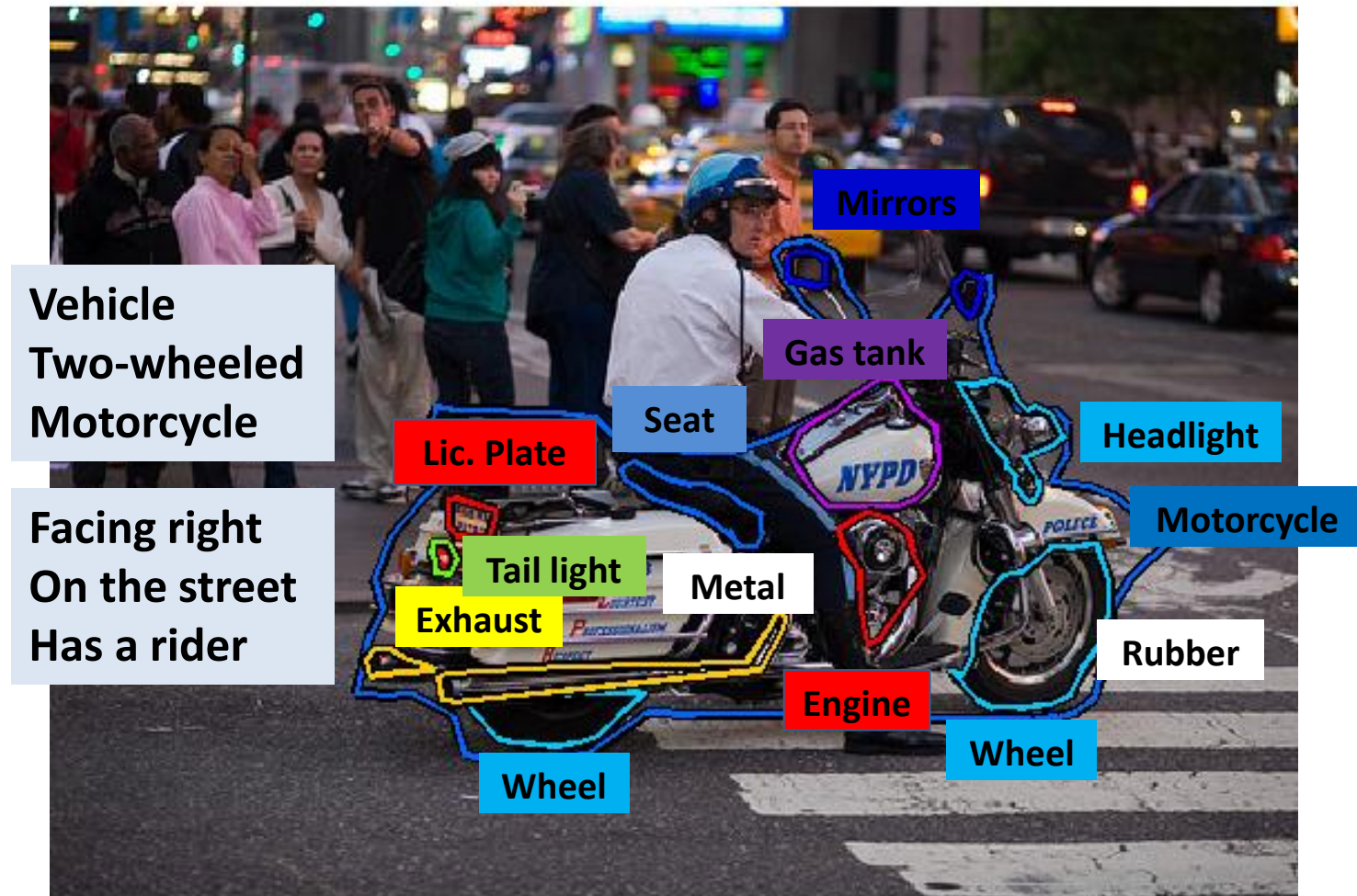


**Cow**

Which objects are relevant, and how are they relevant?

# We need complex, multi-faceted representations

- Categories, pose, material, unusual characteristics, etc.



# Summary

- In object detection, key challenges are object segmentation and fine differentiation
- Object interpretation is a wide-open problem, and we need new object representations
  - Unfamiliar objects
  - Situational context
  - Task context

Thank you

