Extracting Adaptive Contextual Cues From Unlabeled Regions

Congcong Li (Cornell) Devi Parikh (TTIC) Tsuhan Chen (Cornell)

Motivation

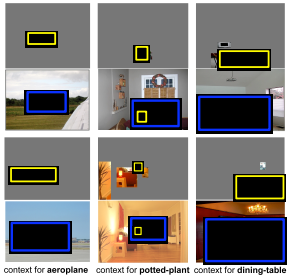
Most works use labeled data to learn contextual cues.

Drawback 1: **unlabeled regions are neglected.**

Drawback 2: contextual cues are often of **fixed-granularity**.

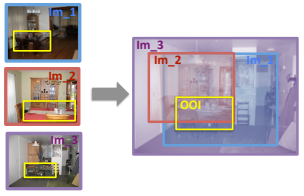
Goal

Extract contextual cues at **adaptive-granularity** from **unlabeled** regions



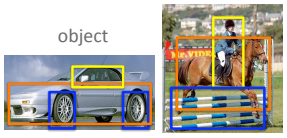
Approach Overview

Observation 1: objects provide information for scene alignment.

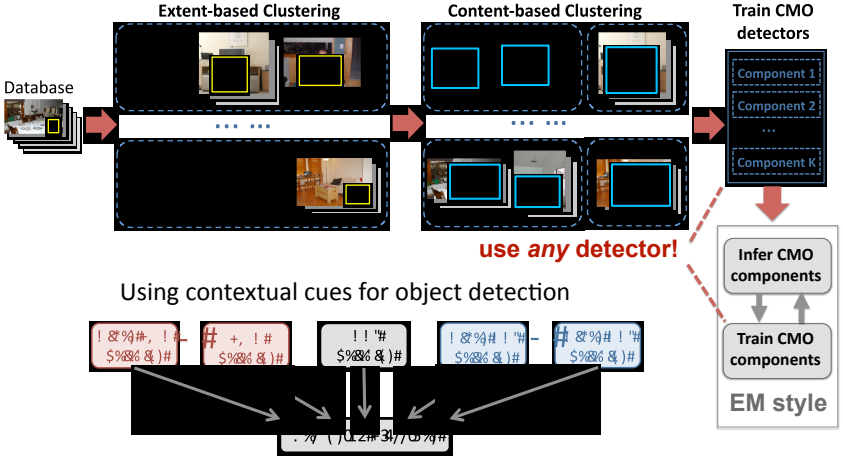


Observation 2: extent and content of the context varies.

Observation 3: objects are to scenes as parts are to objects



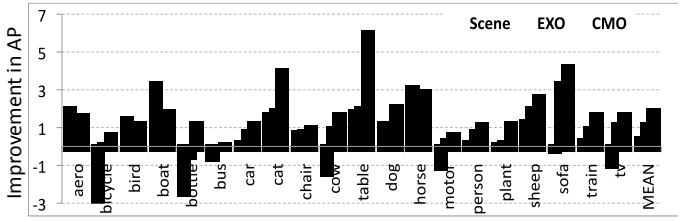
Extracting contextual cues from unlabeled regions



Using contextual cues for object detection

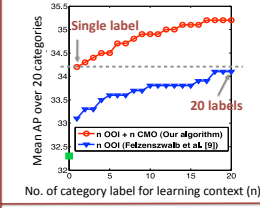
Quantitative Results

Adaptive granularity helps!

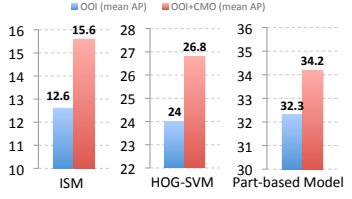


Unlabeled regions help!

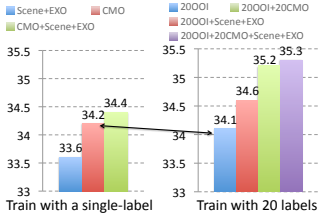
CMO with single labeled category ~ state-of-art using 20 labeled categories



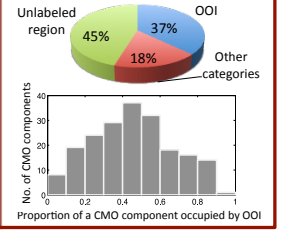
Can employ any object detector!



Complementary contextual cue



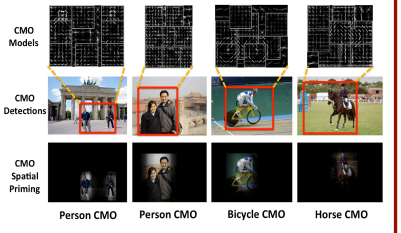
Content in CMO



Qualitative Results

Adaptive Granularity ⟷ **Multi-level Interactions**

CMO provides **spatial prior** for OOI detection



Contributions

- Capturing contextual interactions at varying granularities: **Scene, Inter-object, Intra-object**
- Extracting contextual regions by learning "object" models using **any** object detector
- Improving object detection performance of multiple detectors
- Intelligently leveraging existing detectors: **easily accessible** to community

Contextual-Meta Object (CMO)