

Object Detection Meets Knowledge Graphs



Yuan Fang, Kingsley Kuan, Jie Lin, Cheston Tan and Vijay Chandrasekhar

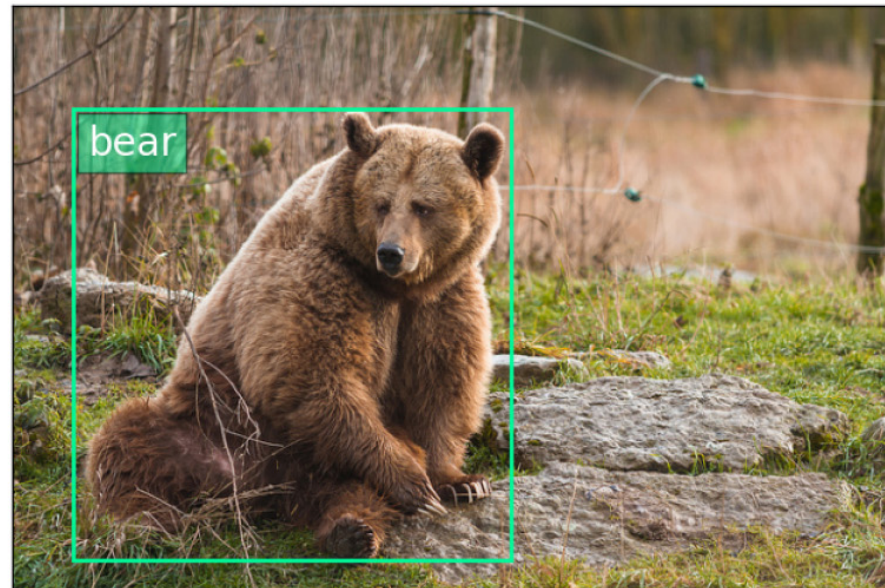
Institute for Infocomm Research, A*STAR, Singapore

Problem: Object Detection in Images

(a) Detecting cat and table



(b) Detecting bear

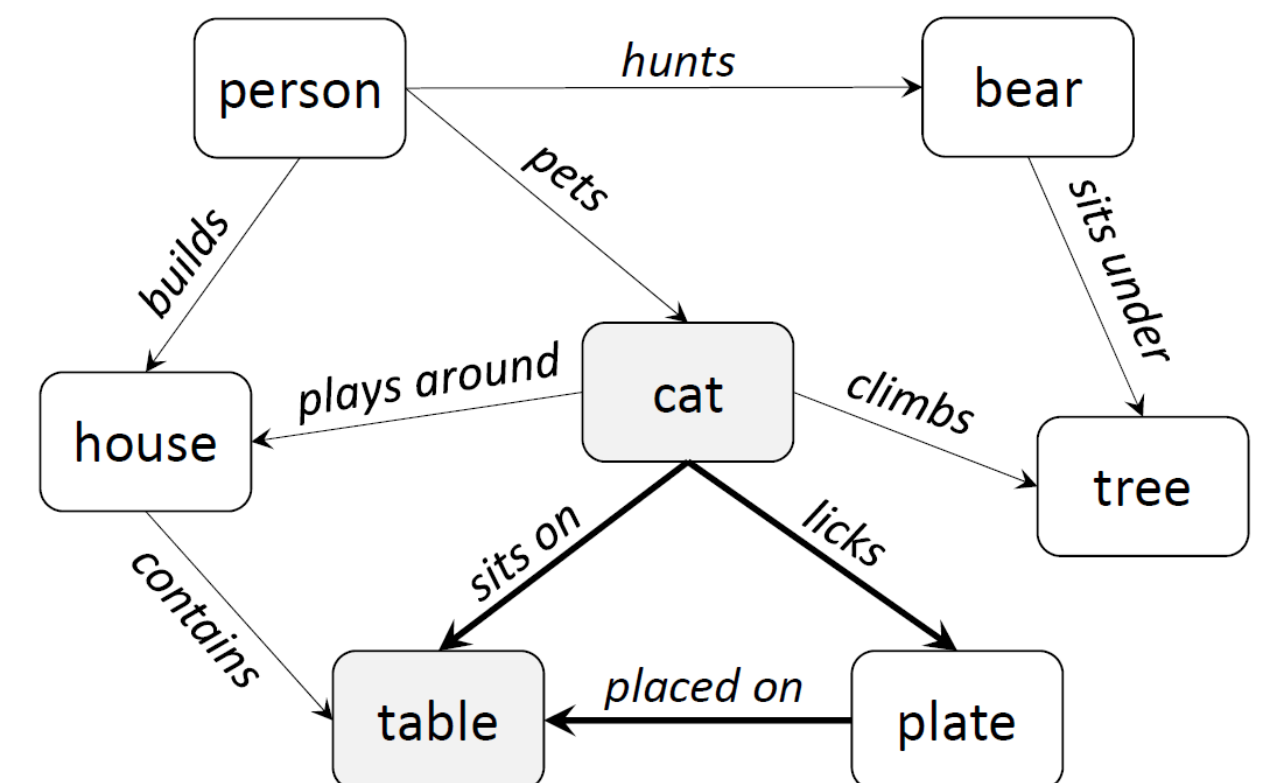


Motivation: External knowledge

Commonsense

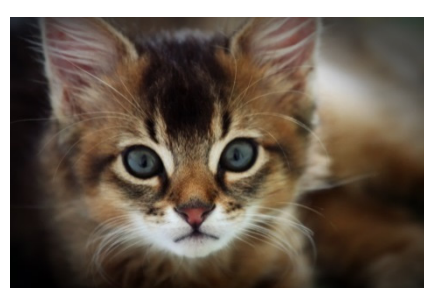
- cat sits on table ✓
- bear sits on table ?
- bear sits on rock ✓

Knowledge graph (KG)

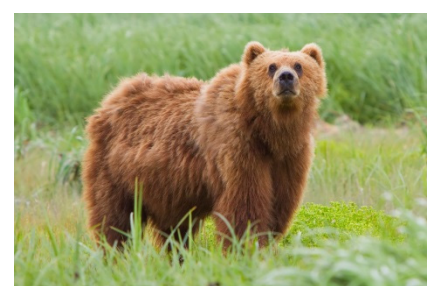


Formalization: Semantic Consistency

semantically *consistent*



semantically *inconsistent*



Semantic Consistency Matrix S

$$S_{\text{cat}, \text{table}} \gg S_{\text{bear}, \text{table}}$$

$$\begin{aligned} |p(\text{cat}|b) - p(\text{table}|b')| &\approx 0 \\ |p(\text{bear}|b) - p(\text{table}|b')| &\gg 0 \end{aligned}$$

Frequency of co-occurrences

$$S_{l,l'} = \max \left(\log \frac{n(l,l')N}{n(l)n(l')}, 0 \right) \quad \text{no generalization}$$

Random walk on KG

$$R_{l,l'} = \lim_{t \rightarrow \infty} P(v_t = l' | v_0 = l)$$

$$S_{l,l'} = \sqrt{R_{l,l'} R_{l',l}}$$

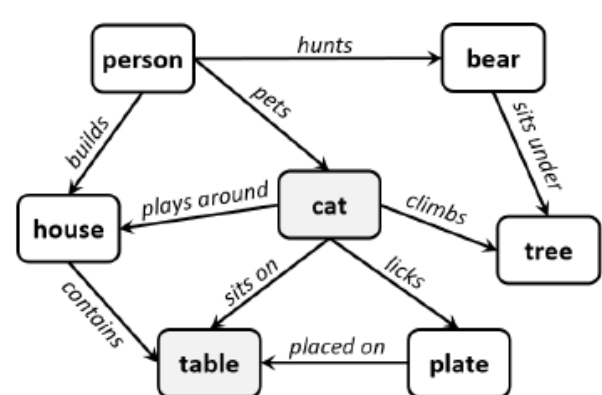
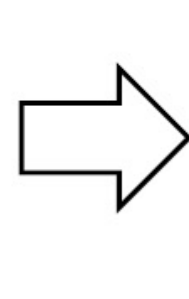
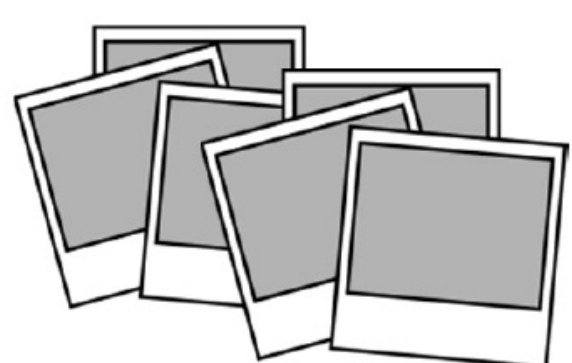
generalization & robustness

Proposed Approach: Re-optimization Framework

Training images

Existing model (e.g., Faster R-CNN)

Test image



Knowledge

semantic consistency

re-optimization

$$P = \begin{bmatrix} 0.6 & 0.4 \\ 0.2 & 0.8 \end{bmatrix} \Rightarrow \hat{P} = \begin{bmatrix} 0.4 & 0.6 \\ 0.1 & 0.9 \end{bmatrix}$$

Existing model output

Knowledge-aware output

$$\begin{aligned} E(\hat{P}, P) = & (1 - \epsilon) \sum_{b=1}^B \sum_{\substack{b'=1 \\ b' \neq b}}^B \sum_{\ell=1}^L \sum_{\ell'=1}^L S_{\ell, \ell'} \left(\hat{P}_{b, \ell} - \hat{P}_{b', \ell'} \right)^2 \\ & + \epsilon \sum_{b=1}^B \sum_{\ell=1}^L B \|S_{\ell, *}\|_1 \left(\hat{P}_{b, \ell} - P_{b, \ell} \right)^2 \end{aligned}$$

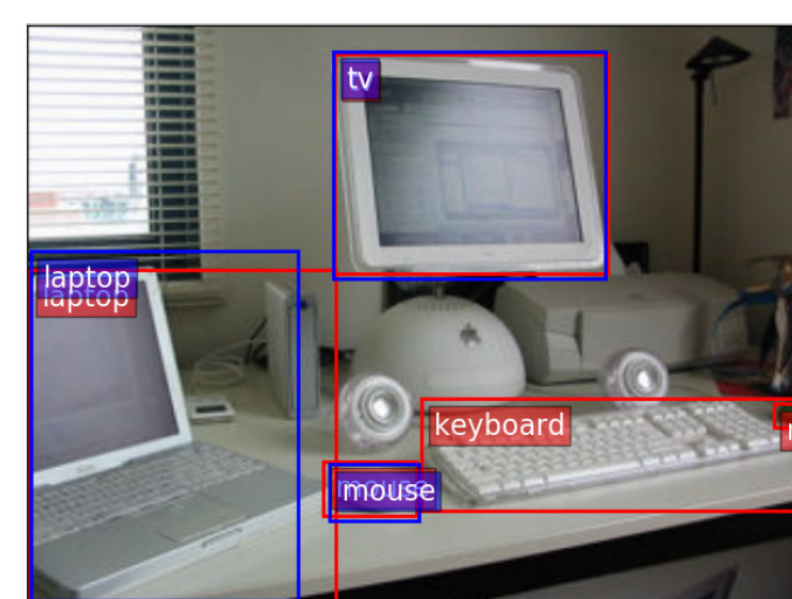
Main Results

MS COCO Dataset

	mAP @ 100	Recall @ 100	Recall @ 10	Recall@100 by area		
				small	medium	large
minival-4k						
FRCNN	24.5	35.9	35.2	14.2	41.5	55.6
KF-500	24.4	37.1	35.6	14.3	42.8	57.3
KF-AII	24.5	37.9	36.2	14.6	43.9	58.6
KG-CNet	24.4	38.9	36.6	14.4	45.2	60.0
test-dev						
FRCNN	24.2	34.6	34.0	12.0	38.5	54.4
KF-500	24.3	37.4	35.9	13.7	42.1	58.0
KF-AII	24.3	38.2	36.4	14.2	43.0	59.2
KG-CNet	24.2	39.2	36.9	14.5	44.0	60.7
test-std						
FRCNN	24.2	34.7	34.1	11.5	38.9	54.4
KG-CNet	24.1	39.2	37.0	14.2	44.4	60.5

Case study

FRCNN



FRCNN + KG

