Unsupervised Discovery of Mid-Level Discriminative Patches

Saurabh Singh, Abhinav Gupta, Alexei A. Efros
Carnegie Mellon University

Presented by Chuan Wang

Which is from Paris?

visual primitives: patches

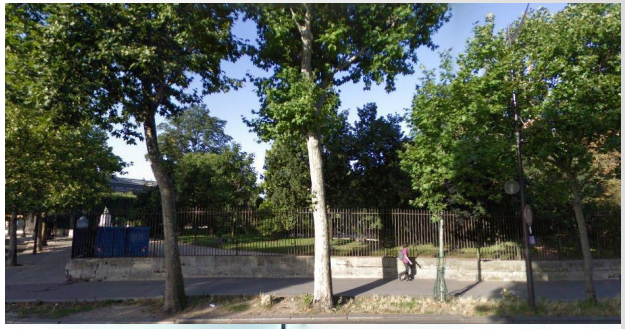
Discriminative Patches

~~representative~~
discriminative

Frequent

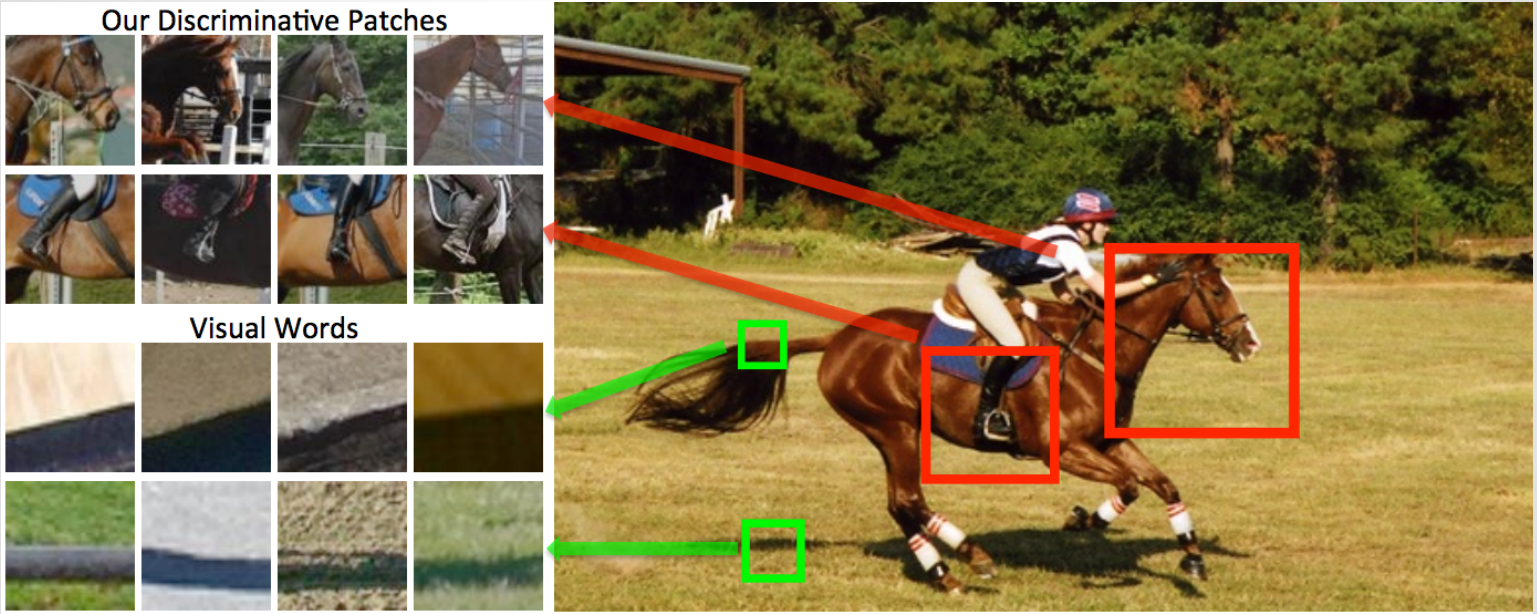
+

Discriminative



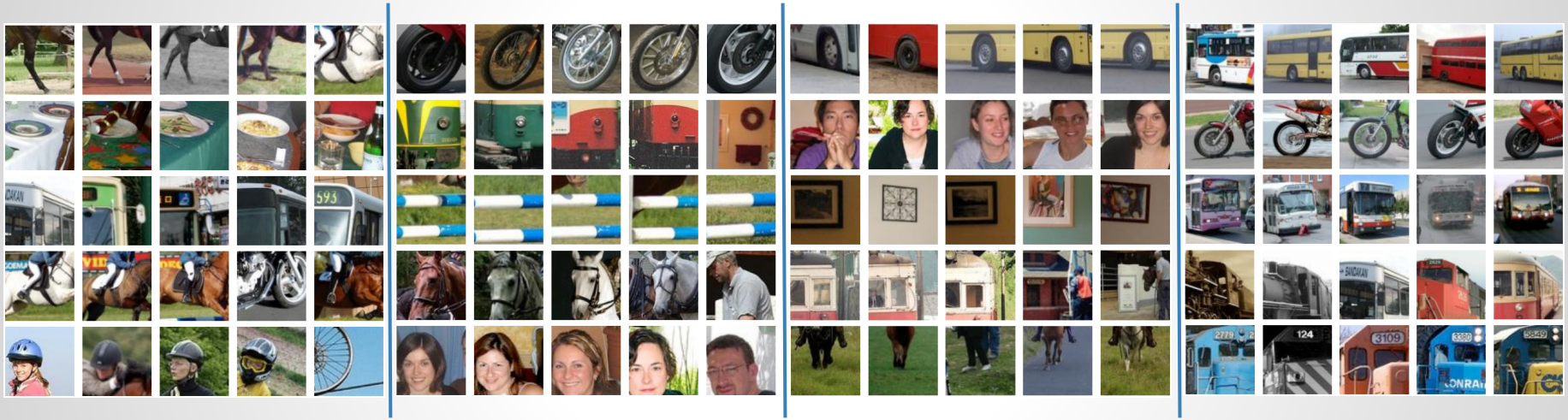
representative
~~discriminative~~

Mid-level Discriminative Patches



- Can be detected in a large number of images with high recall and precision
- Do not explicitly aim to discover whole semantic units
- Fully unsupervised: no separate training and test set

Mid-level Discriminative Patches

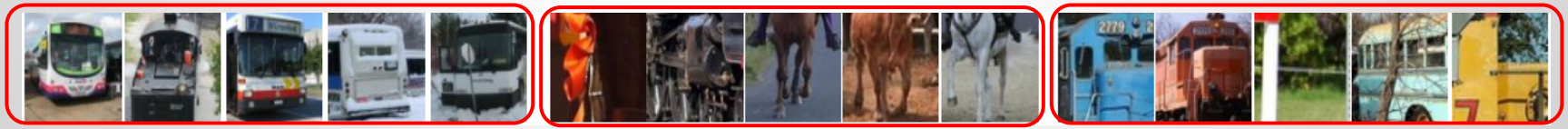


Highly Ranked Patches

Approach

- K-Means

- Discriminative Clustering
  - Input: discovery set $\mathcal{D}$ + “natural world” set $\mathcal{N}$
  - Make each cluster a detector
  - Train to find similar patches like those it already owns (linear SVM)
  - Not just against $\mathcal{D}$, but also against $\mathcal{N}$



Approach

Algorithm 1 Discover Top n Discriminative Patches

Require: Discovery set $\mathcal{D}$, Natural World set $\mathcal{N}$

$\mathcal{D} \Rightarrow \{D_1, D_2\}$; $\mathcal{N} \Rightarrow \{N_1, N_2\}$ ▷ Divide $\mathcal{D}, \mathcal{N}$ into equal sized disjoint sets

$S \Leftarrow rand\_sample(D_1)$ ▷ Sample random patches from D_1

$K \Leftarrow kmeans(S)$ ▷ Cluster patches using KMeans

while not $converged()$ **do**

  for all i such that $size(K[i]) \geq 3$ **do** ▷ Prune out small ones

    $C_{new}[i] \Leftarrow svm\_train(K[i], N_1)$ ▷ Train classifier for each cluster

    $K_{new}[i] \Leftarrow detect\_top(C[i], D_2, m)$ ▷ Find top m new members in other set

  end for

  $K \Leftarrow K_{new}$; $C \Leftarrow C_{new}$

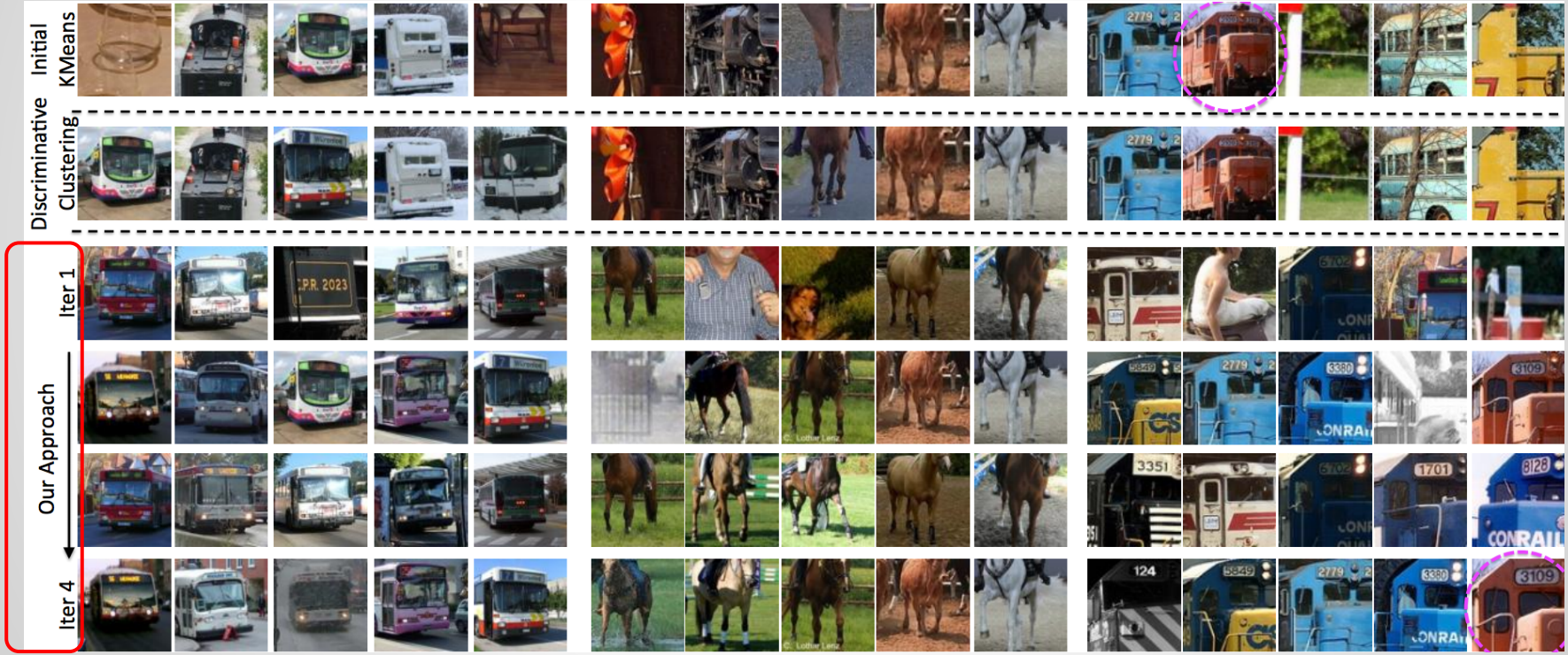
  $swap(D_1, D_2)$; $swap(N_1, N_2)$ ▷ Swap the two sets

end while

$A[i] \Leftarrow purity(K[i]) + \lambda \times discriminativeness(K[i])\ \forall\ i$ ▷ Compute scores

return $select\_top(C, A, n)$ ▷ Sort according to scores and select top n patches

Approach


Initial KMeans
Discriminative Clustering
Our Approach
Iter 1
Iter 4

Ranking Patches

- Purity
  - Summing up the SVM detection scores of top r cluster members (r>m)
- Discriminativeness
  - The number of firings on $\mathcal{D}$ / the number of firings on $\mathcal{D} \cup \mathcal{N}$
- Rank: linear combination of the two terms

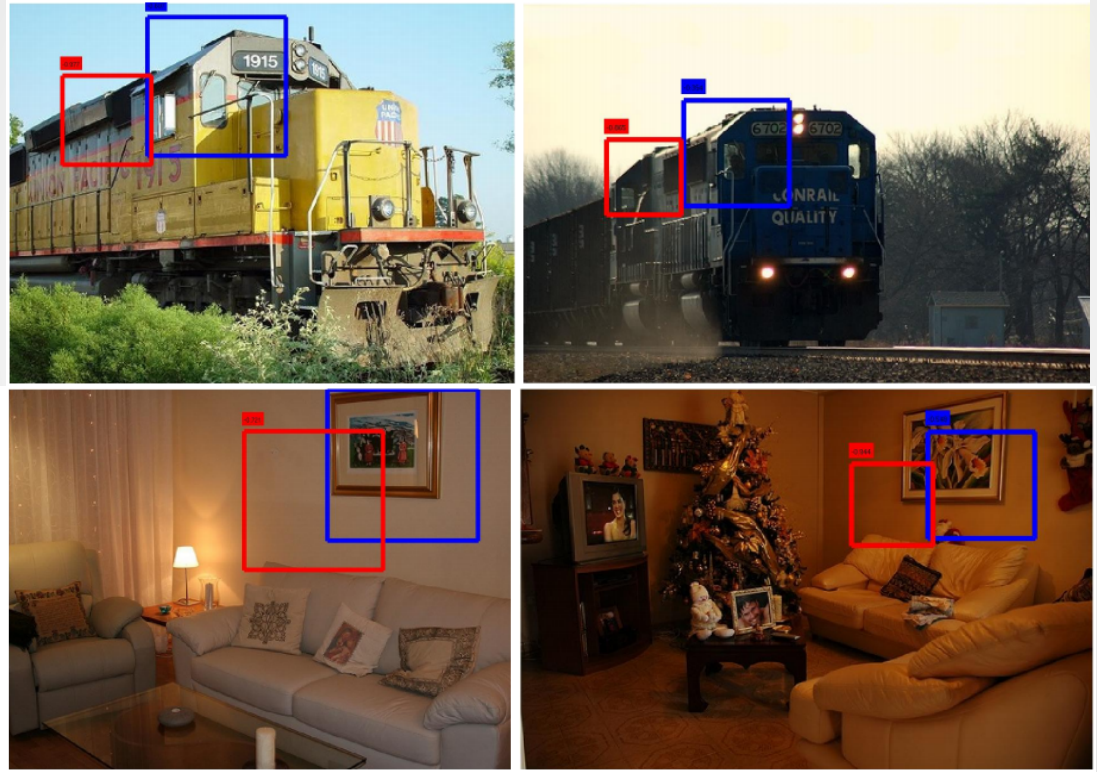
Doublets

- Cluster converges to two or more concepts sometimes (b)

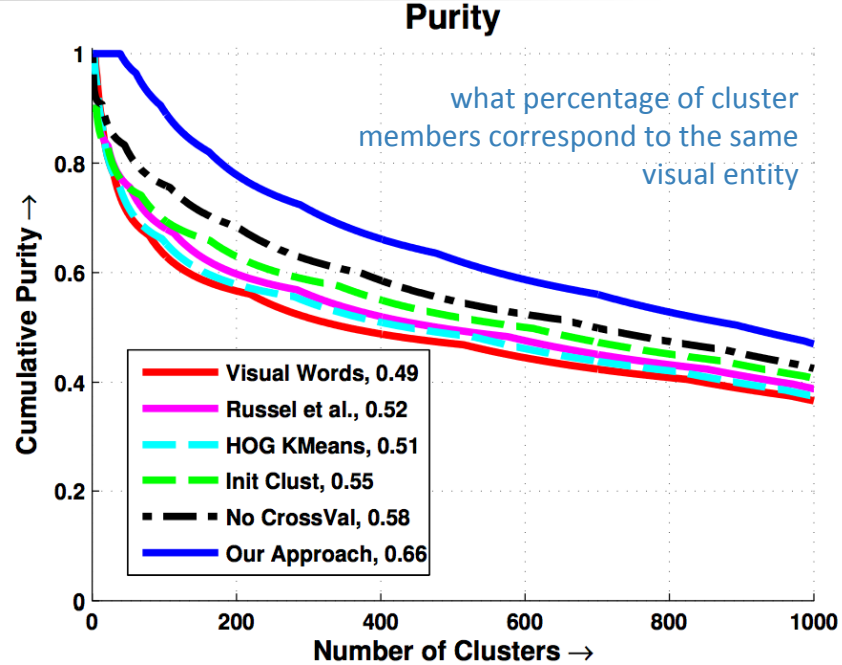
(a) Clean cluster (b) Noisy Cluster (c)Cleaned-up

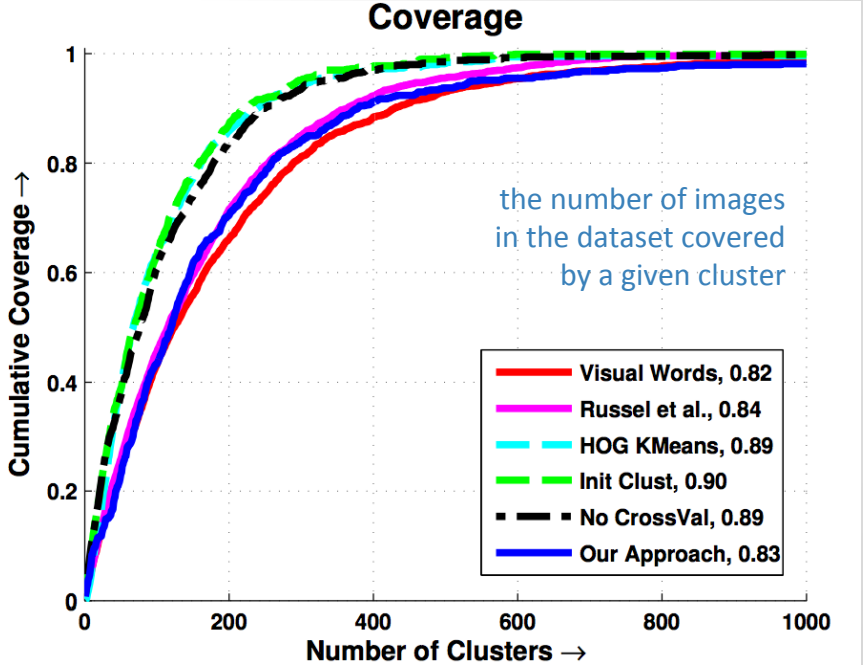
- Discover second-order spatial relationships and cleaning up

Highly Ranked Doublets

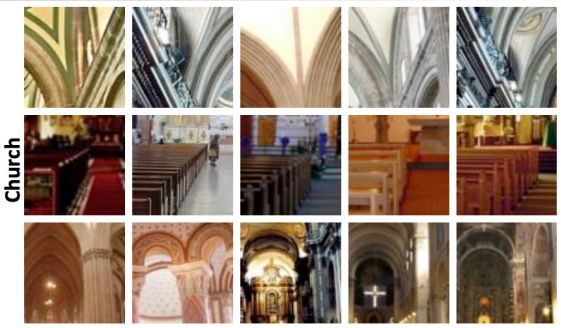


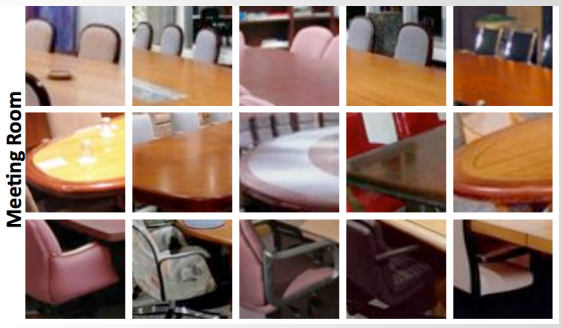
Evaluation (PASCAL VOC2007)

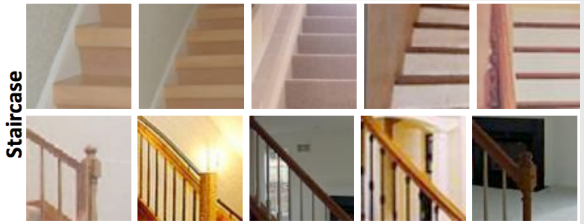




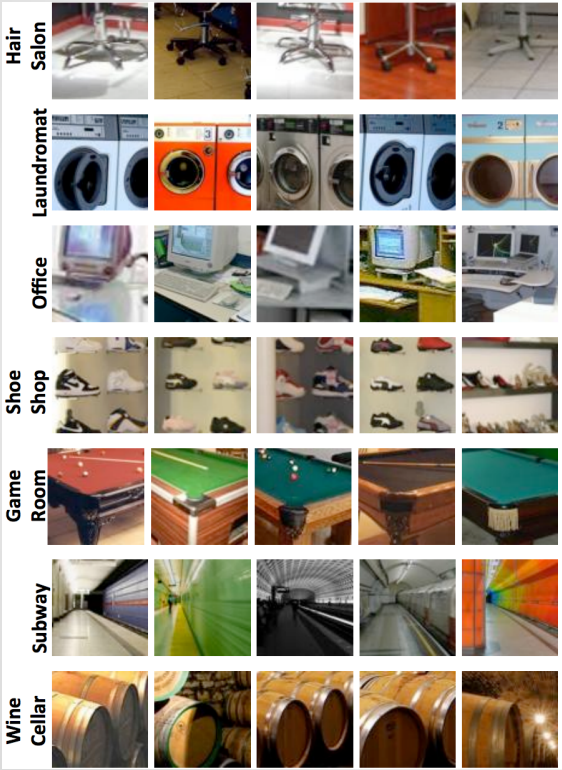
Works Even Better if Weakly Supervised

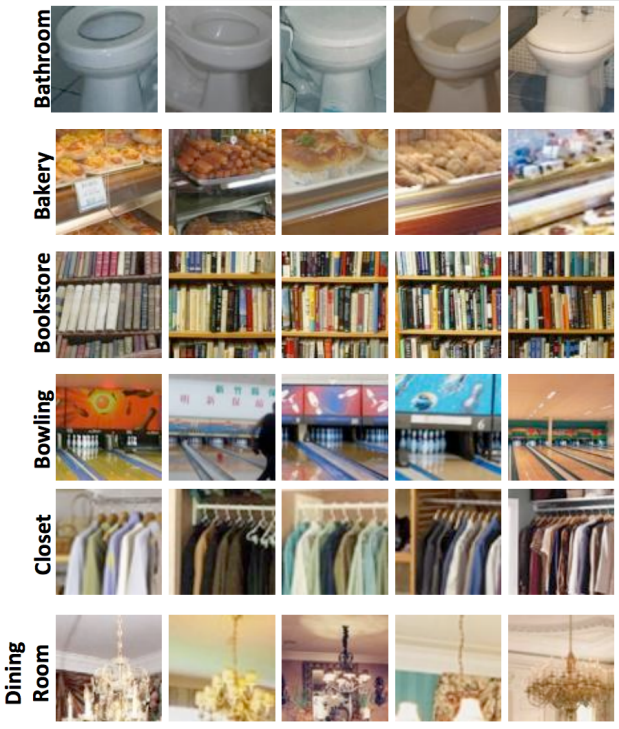






Works Even Better if Weakly Supervised





Quantitative Evaluation (MIT Indoor-67)

GIST	29.7
Spatial Pyramid HOG ($SPHOG$)†	29.8
Spatial Pyramid SIFT (SP)†	34.4
ROI-GIST [5]	26.5
Scene DPM [4]	30.4
MM-Scene [37]	28.0
Object Bank [3]	37.6
Ours	**38.1**

Ours+GIST	44.0
Ours+SP	46.4
Ours+GIST + SP	47.5
Ours+DPM	42.4
Ours+GIST+DPM	46.9
Ours+SP+DPM	46.4
GIST+SP+DPM [4]	43.1*
Ours+GIST+SP+DPM	**49.4**

Discussion

- Strong Points
  - Unsupervised: train classifier against natural world
  - Mid-level patches: semantically meaningful and discriminative
  - Cross validation: simple but effective
- Weak Points
  - m=5: quite a small number and reason?
  - Perspective changes / deformation: difficult ?
- Discussion
  - What if shuffle picking subset during the iterative step?

Thank You!
Question?