Learning from the Web

Presenter: Meng Song

4-20-2015

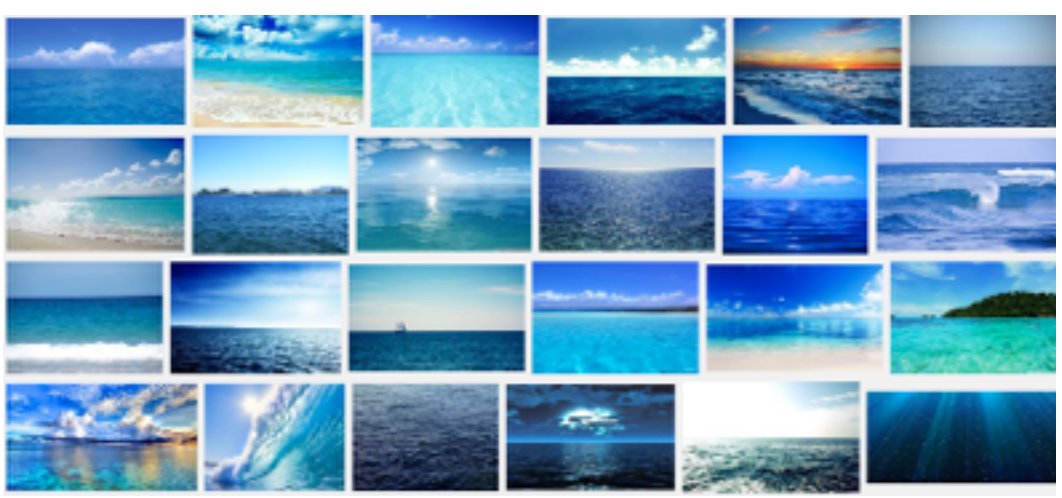
Web images

- Large-scale

Knowledge

- Knowledge: concepts + relationships
- Concept: an abstraction or generalization from experience
  - Objects, scenes, actions






Knowledge

- Relationships between concepts
    - is-a
    - is-a-part-of
    - is / has
    - …

Wheel is a part of car

Mapping between Concept and Images

- Images within the same semantic concept are not necessarily visually similar.
  - Intra-class variation is very large, sometimes greater than inter-class variation.

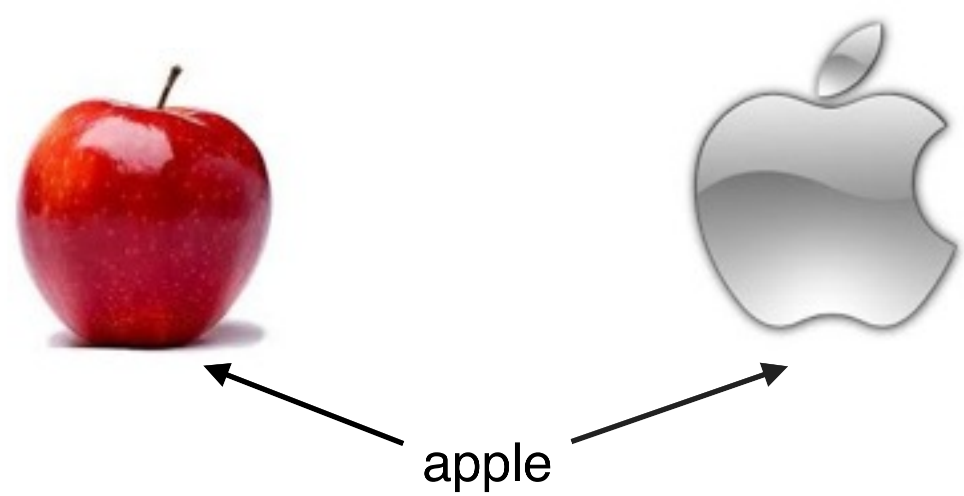
electric train

steam engine

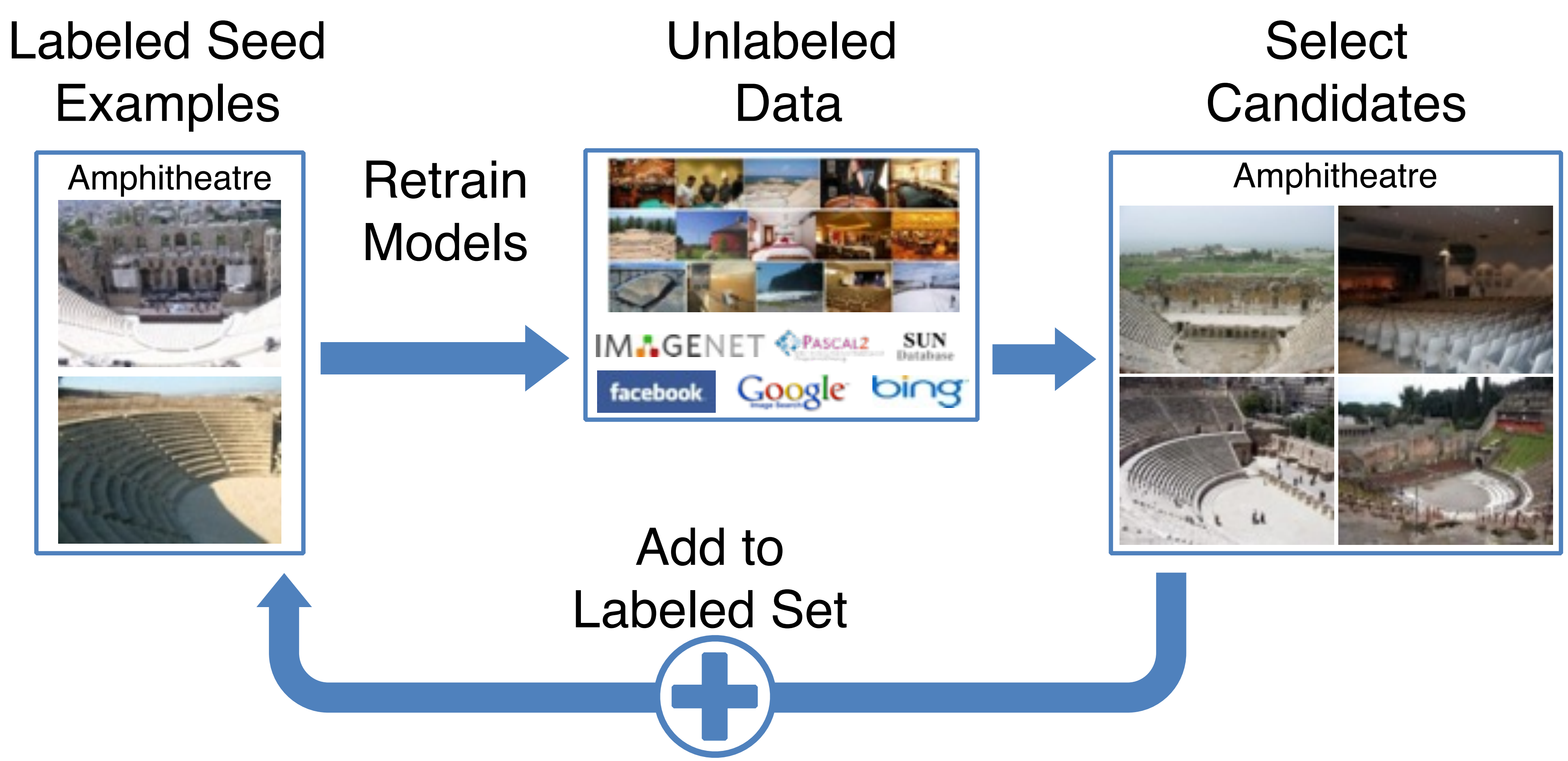
bus

Mapping between Concept and Image Cluster

- Images within the same semantic concept are not necessarily visually similar.
  - Polysemy (multiple meanings of a concept)



Self-learning (Bootstrapping)



Challenges

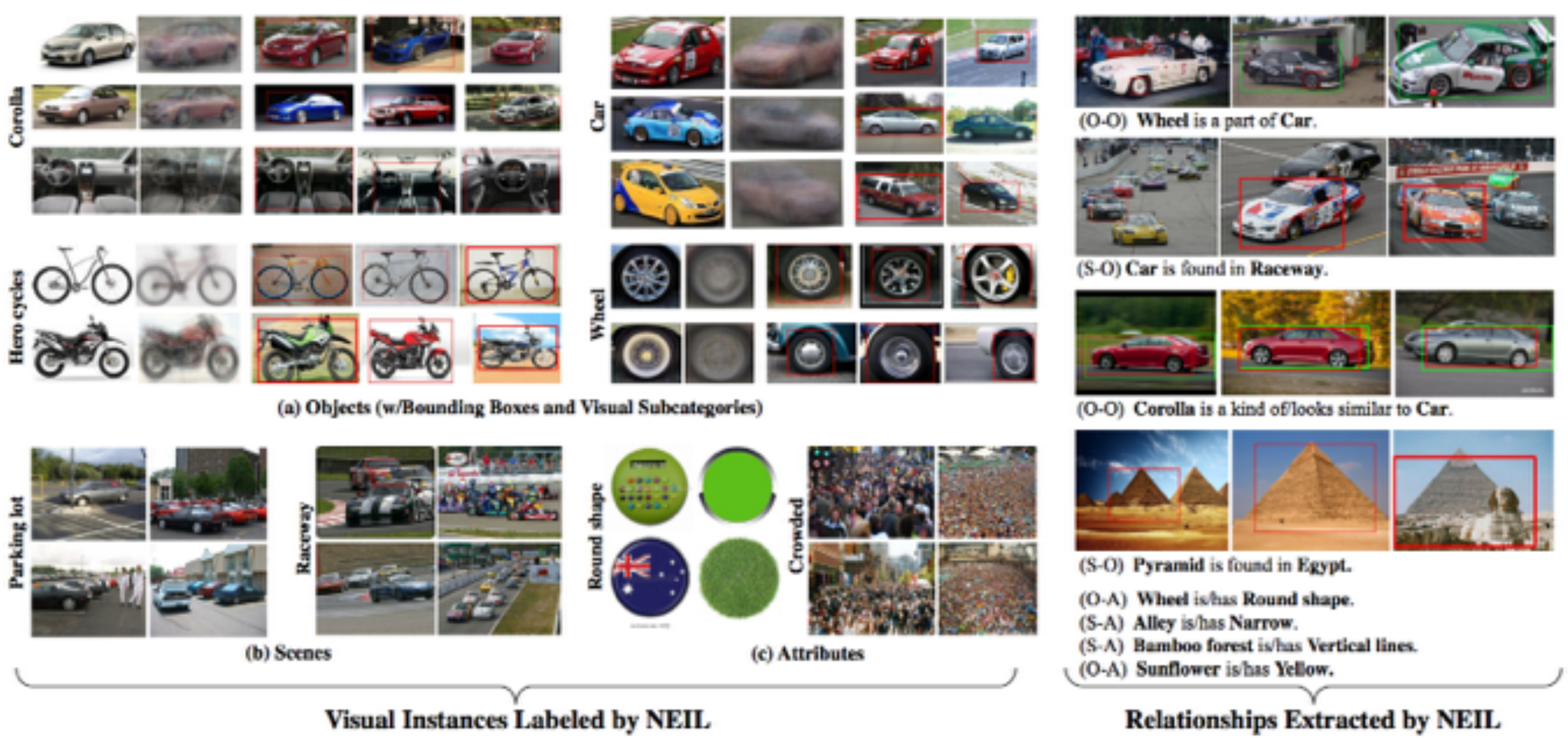
- Trade-off between purity and diversity
  - Keep intra-class diversity
  - Avoid semantic drift

OPTIMOL: Automatic Online Picture Collection via Incremental Model Learning

- Only modeling **isolated object categories**
- Use thresholds to control intra-category diversity and avoid semantic drift
- Datasets are small and lack richness

NEIL: Never Ending Image Learning

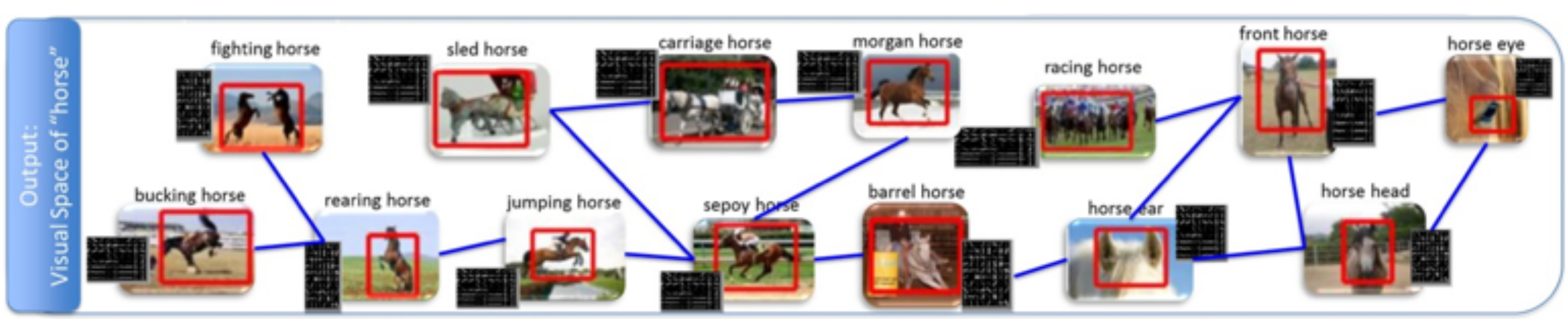
- Scaling up
  - 2.5 month: 400K instances, 1700 relationships
- Automatically[1] extract common sense **relationships**
- Use relationships as constraints to avoid semantic drift



[1] A. Shrivastava and S. Singh and A. Gupta, Constrained Semi-Supervised Learning using Attributes and Comparative Attributes, ECCV 2012

LEVAN: Learning Everything about Anything

- **Intra-concept** modeling
  - Learning a model that exhaustively cover all appearance variances of a concept
- Vocabulary is discovered from Google-book n-grams
  - Extensive and concept-specific

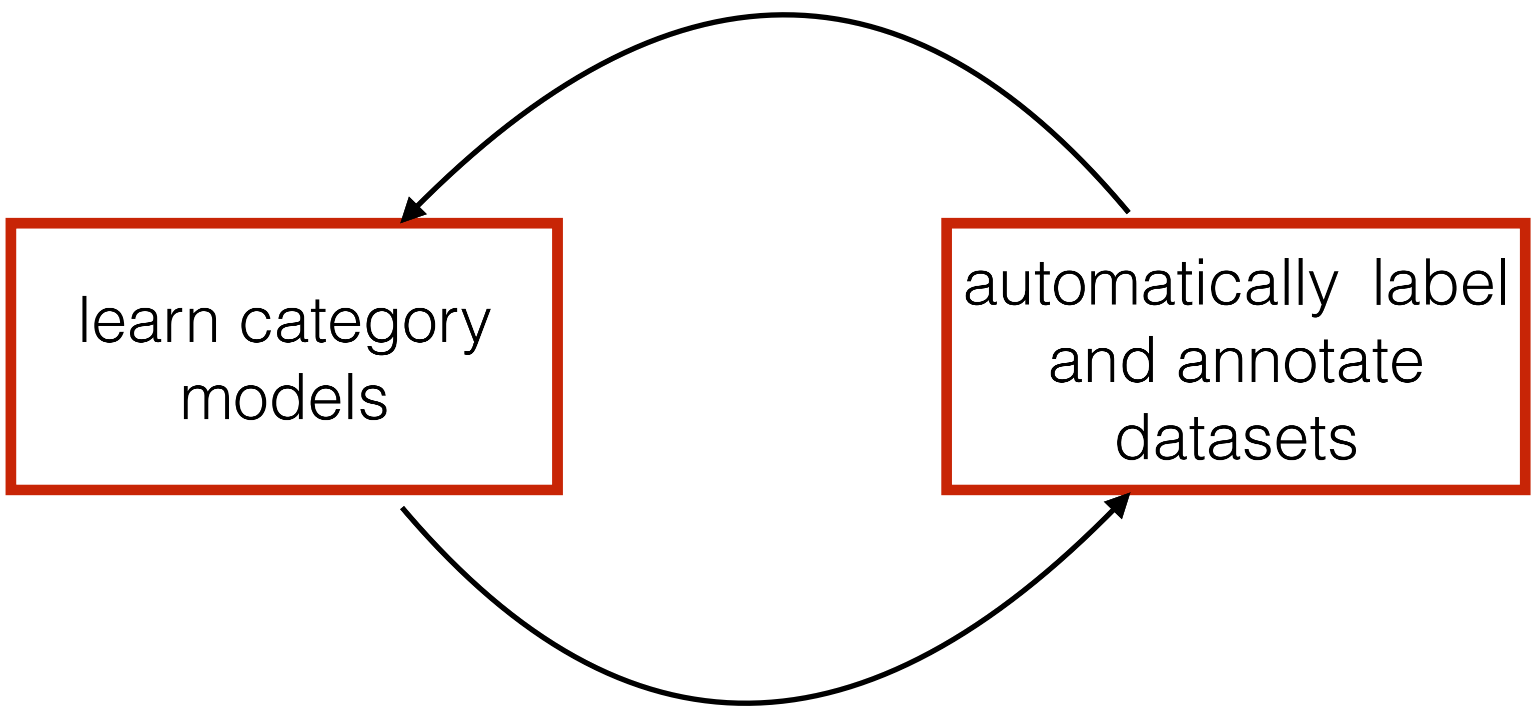


OPTIMOL: Automatic Online Picture Collection via Incremental Model Learning

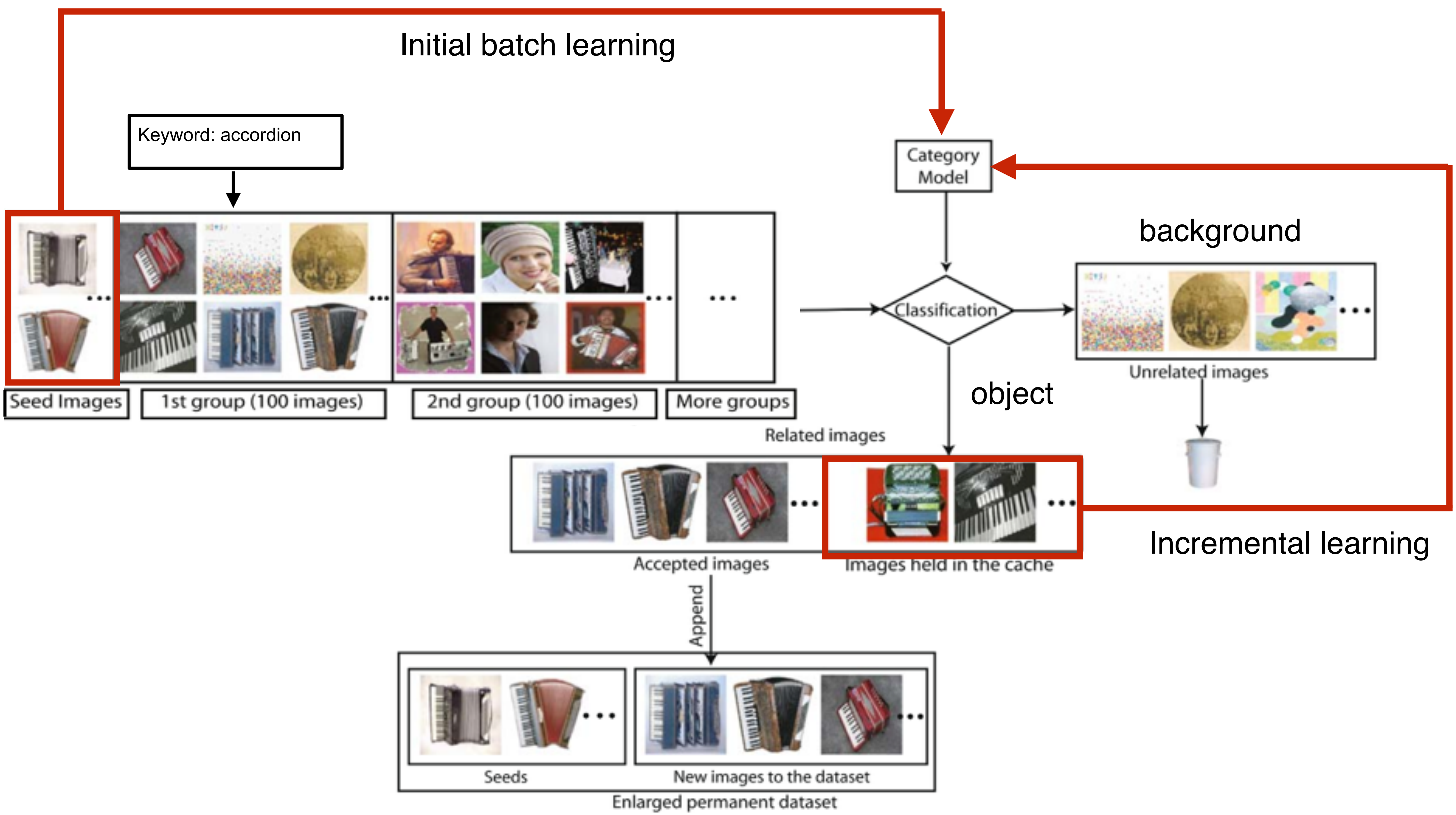
Jia Li, Fei-fei Li

Motivation

- Automatically collect a larger and more diverse dataset



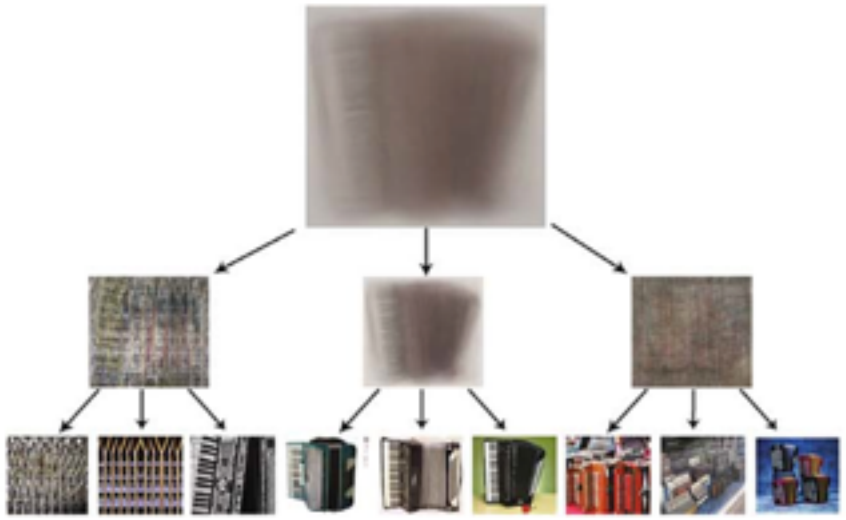
Framework



Initial batch learning
Keyword: accordion
Category Model
background
Classification
Seed Images
1st group (100 images)
2nd group (100 images)
More groups
Unrelated images
object
Related images
Accepted images
Images held in the cache
Incremental learning
Append
Seeds
New images to the dataset
Enlarged permanent dataset

Modeling Categories

- Generative graphical model
  - Directly model data: p(x), p(x|y), p(x,y)
- For object image and background image, build a HDP (Hierarchical Dirichlet Process) model respectively
  - The number of clusters can be adjusted to accommodate new data
  - This model provides natural clustering of data
    - Sub-category clustering
    - Solving polysemy problem



HDP: Hierarchical Dirichlet Process

- Model each image as unordered bag of visual words

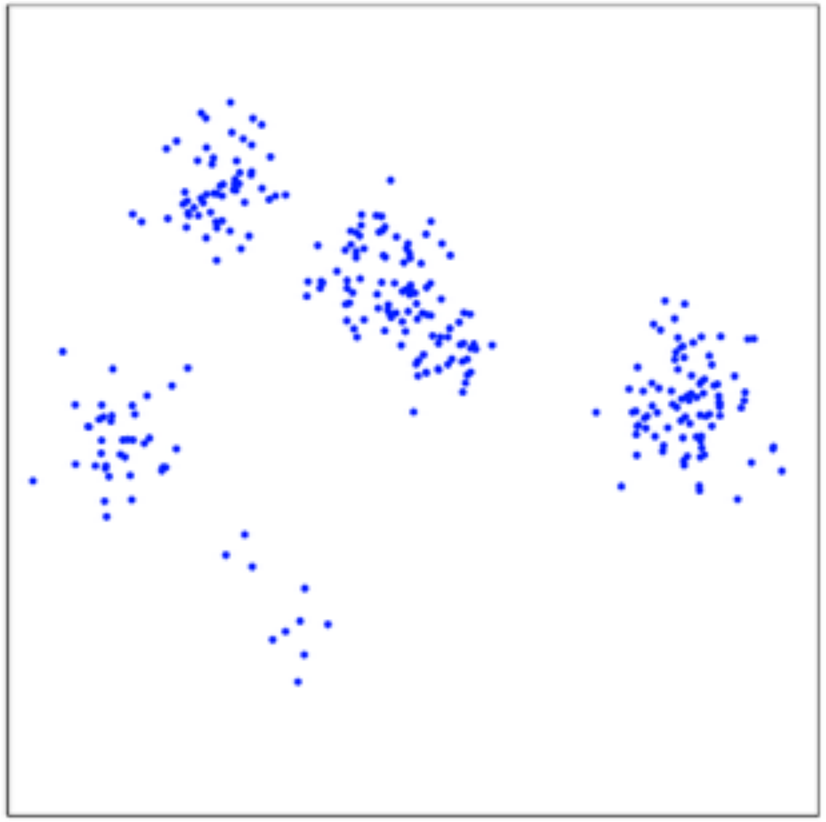
Image patches: x

Topics: clusters of image patches

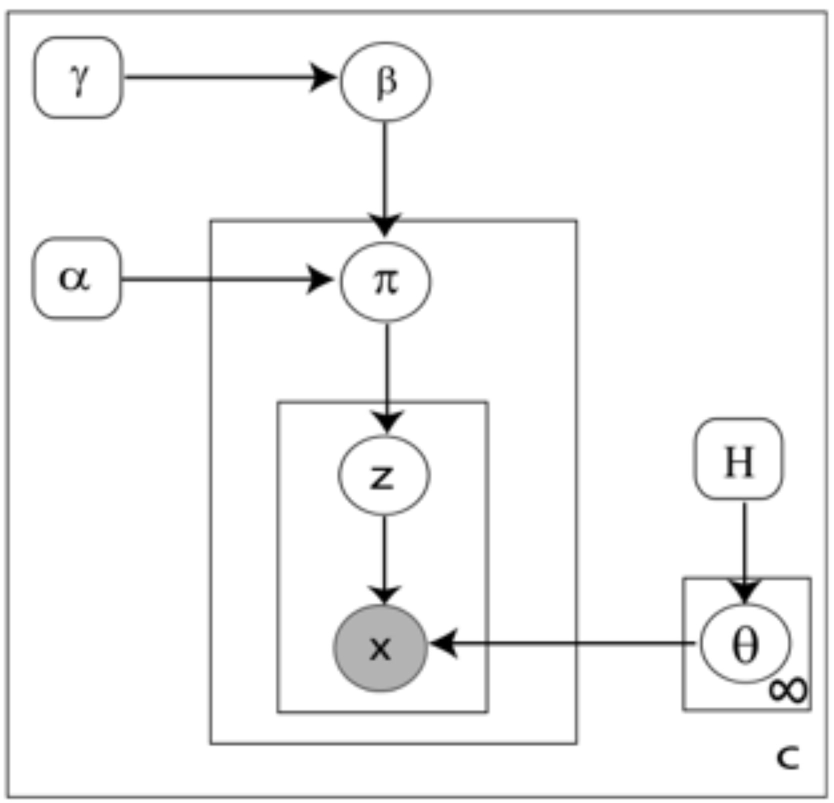
View Clustering from Mixture Model Perspective

- Data $\mathbf{x} = \{x_1, x_2, \ldots, x_n\}$
- Each cluster is modeled by a distribution
- For each data item x_i
  - Pick a cluster z_i from clusters
  - Generate x_i from corresponding distribution

$$p(\mathbf{x}|\theta) = \prod_{i=1}^{n} \sum_{z_i=1}^{K} \underbrace{\boxed{\pi_{z_i}}}_{\text{mixture proportion}} \underbrace{\boxed{p(x_i|z_i, \theta_{z_i})}}_{\text{component distribution}}$$



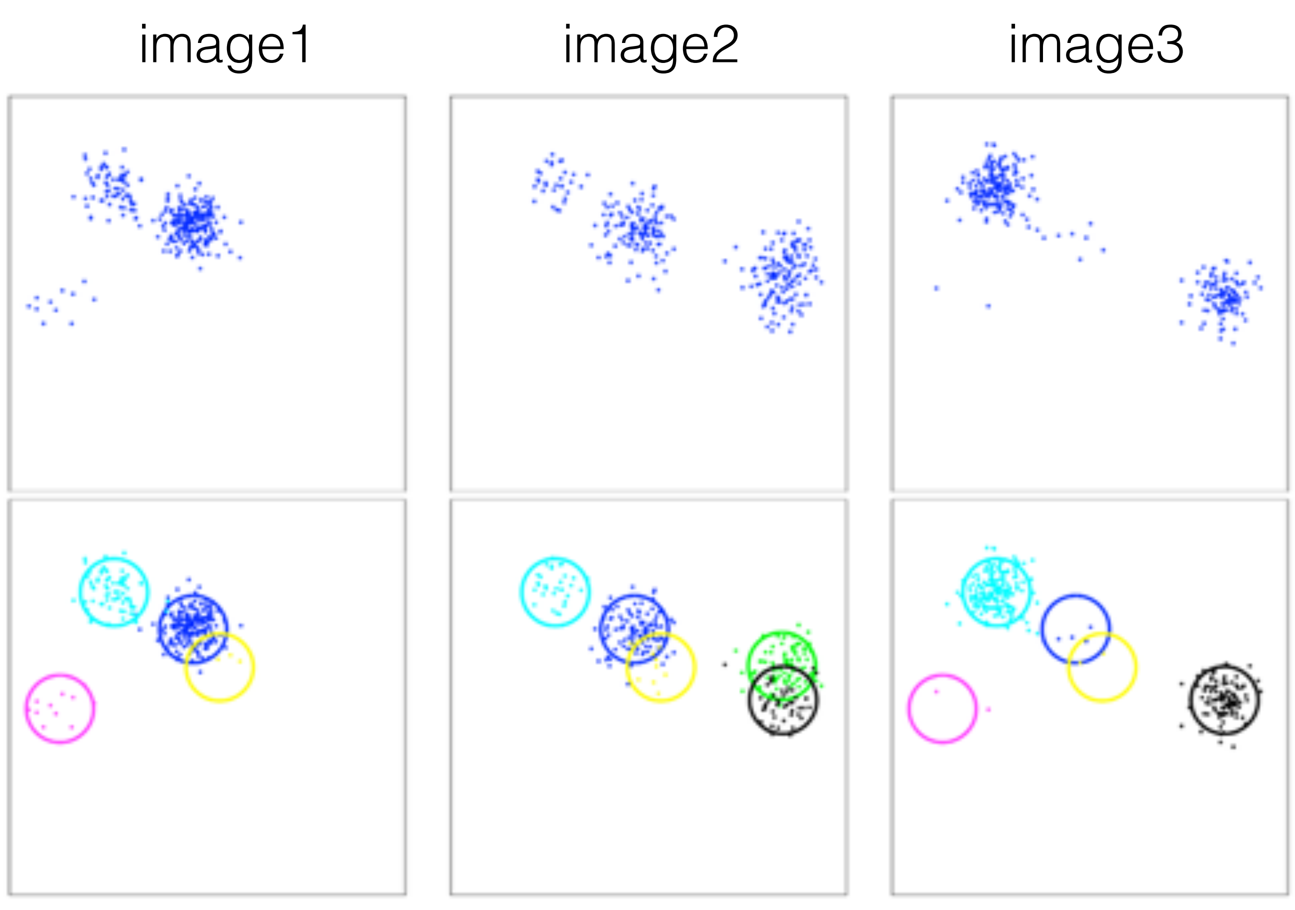
HDP: Hierarchical Dirichlet Process



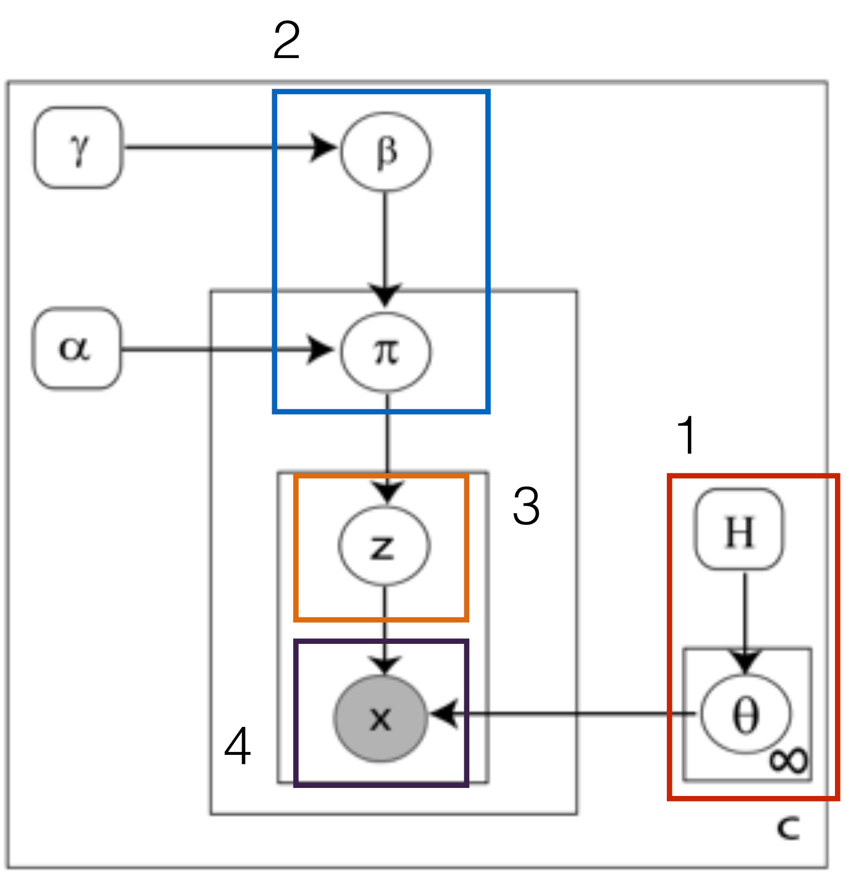
- Mixture model with unbounded number of clusters $K \to \infty$
  - Encode prior belief that data arise from complex processes that cannot be described with a finite mixture model.

HDP: Hierarchical Dirichlet Process

- Hierarchical model:
  - Sharing global topics across images
  - Image-specific mixture proportion



How to generate an image patch from HDP



- Sample component distributions from Dirichlet process
- For image j, sample mixture proportion

$$\pi_j = (\pi_{j1}, \pi_{j1}, \ldots, \pi_{j\infty})$$

- For each patch x_{ji}
  - Sample topic index z_{ji}
  - Sample x_{ji} from multinomial distribution $F(\theta^c_{z_{ji}})$

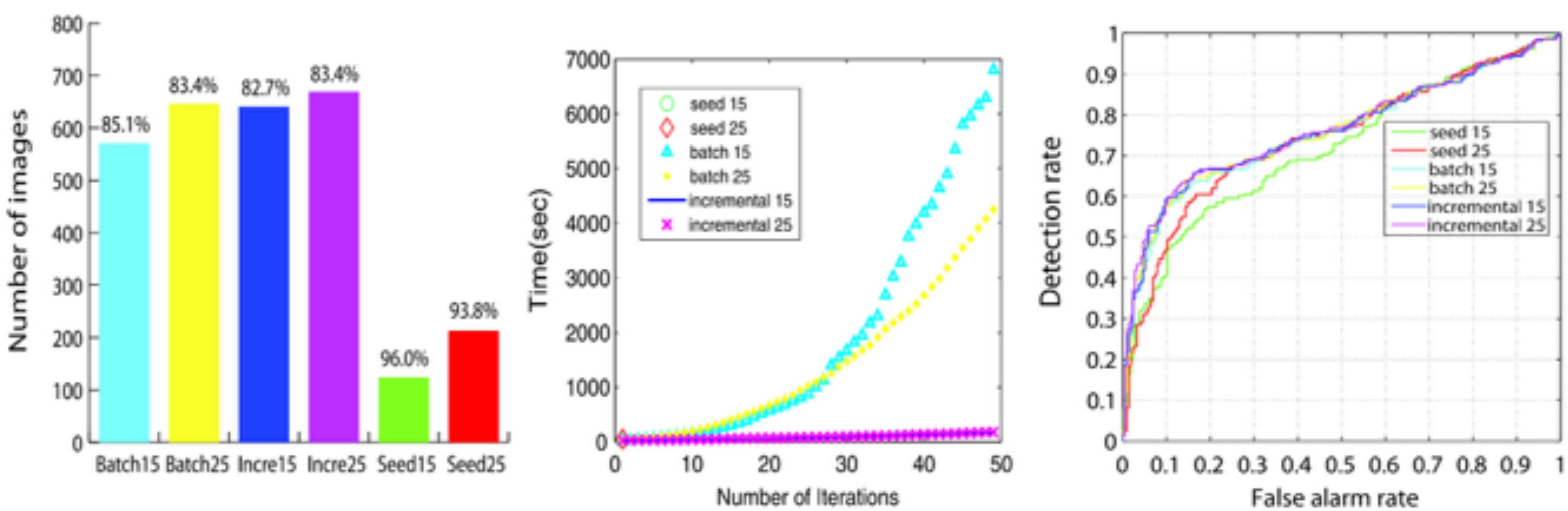
Learning Model Parameters

- Fit the data to the model
- The model parameters and hidden variables are updated iteratively
- Inference: Gibbs sampling
    - Sample local clusters
        - Choose an existing cluster
        - Sample a new cluster
    - Sample global topics
        - Choose an existing topic
        - Sample a new topic

Learning Model Parameters

- Incremental learning
  - Instead of using both new data and old data to re-train the model, only the images of current iteration are used.

$$z_j \sim p(z|\Theta^{j-1}, I_j) \quad \Theta^j \sim p(\Theta|z_j, \Theta^{j-1}, I_j)$$



Batch vs. Incremental learning

Image Classification

- Choose the category model with higher probability of generating the image

$$P(I|c) = \prod_{i} \sum_{z} P(x_i|z, c) P(z|c)$$

- Penalize false positives
  - Only accept an image when the risk of accepting it is lower than rejecting it.

$$R_i(A|I) < R_i(R|I)$$

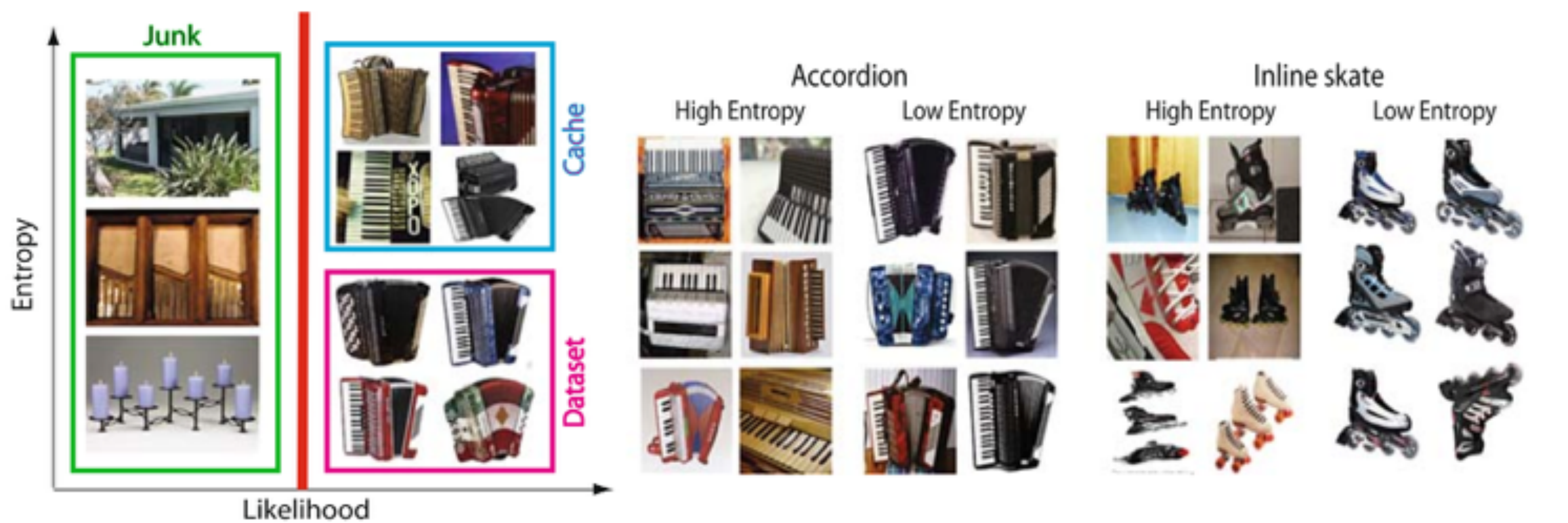
$$\frac{P(I|c_f)}{P(I|c_b)} > \boxed{\frac{\lambda_{Ac_b} - \lambda_{Rc_b}}{\lambda_{Rc_f} - \lambda_{Ac_f}}} \frac{P(c_b)}{P(c_f)}$$

inline skate

Increasing Class Diversity

- Only use **cache set** for incremental learning
  - Images with high entropy are more likely to have new topics

$$H(I) = -\sum_{z} p(z|I) \ln p(z|I)$$

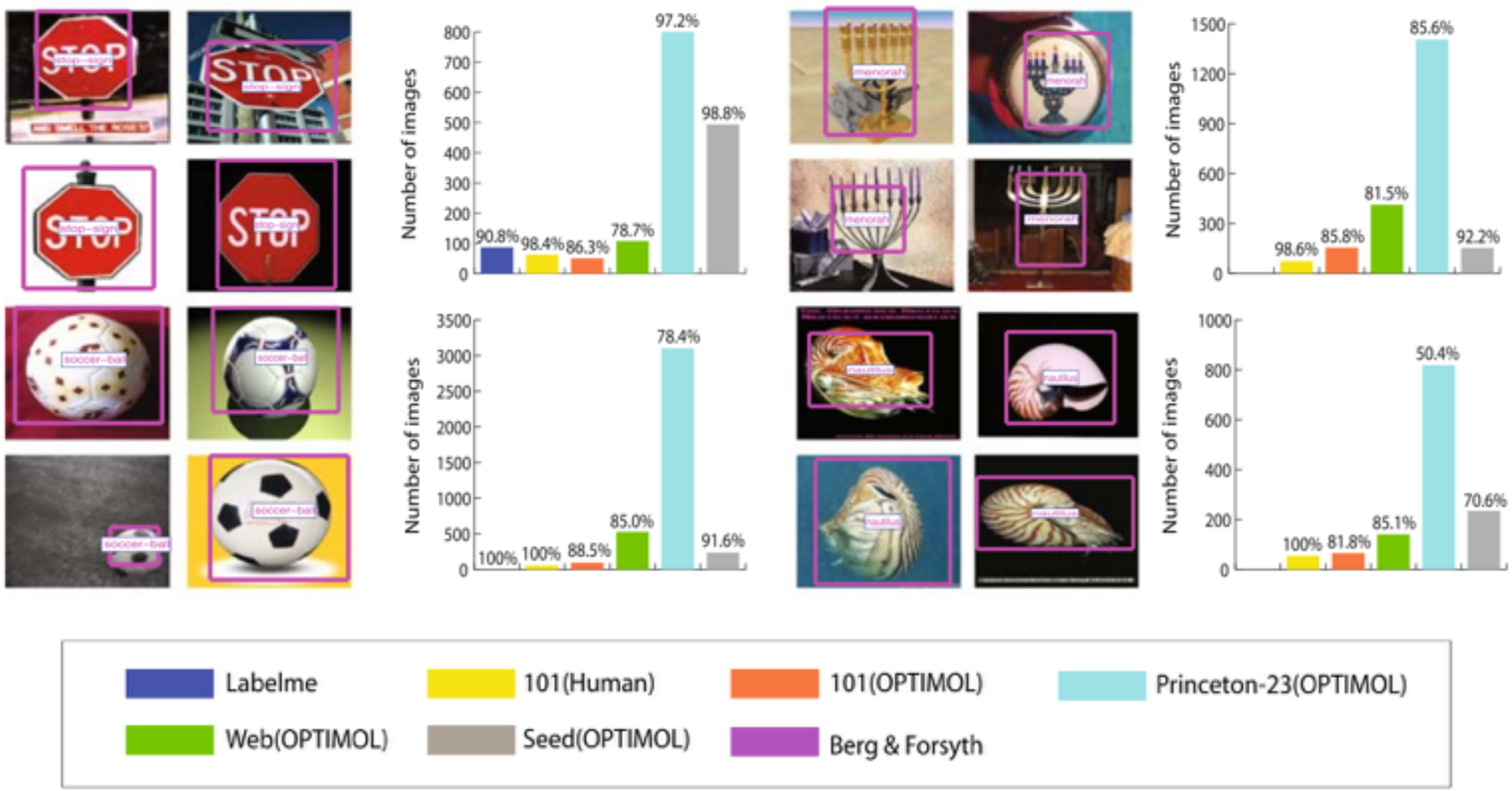


Experiments and Results

- Datasets:
  - Caltech 101
  - Web-23
  - Princeton-23
  - Fergus ICCV'05 dataset
- The number of categories: 7 to101
- The number of images in each category: ~100 to ~1000

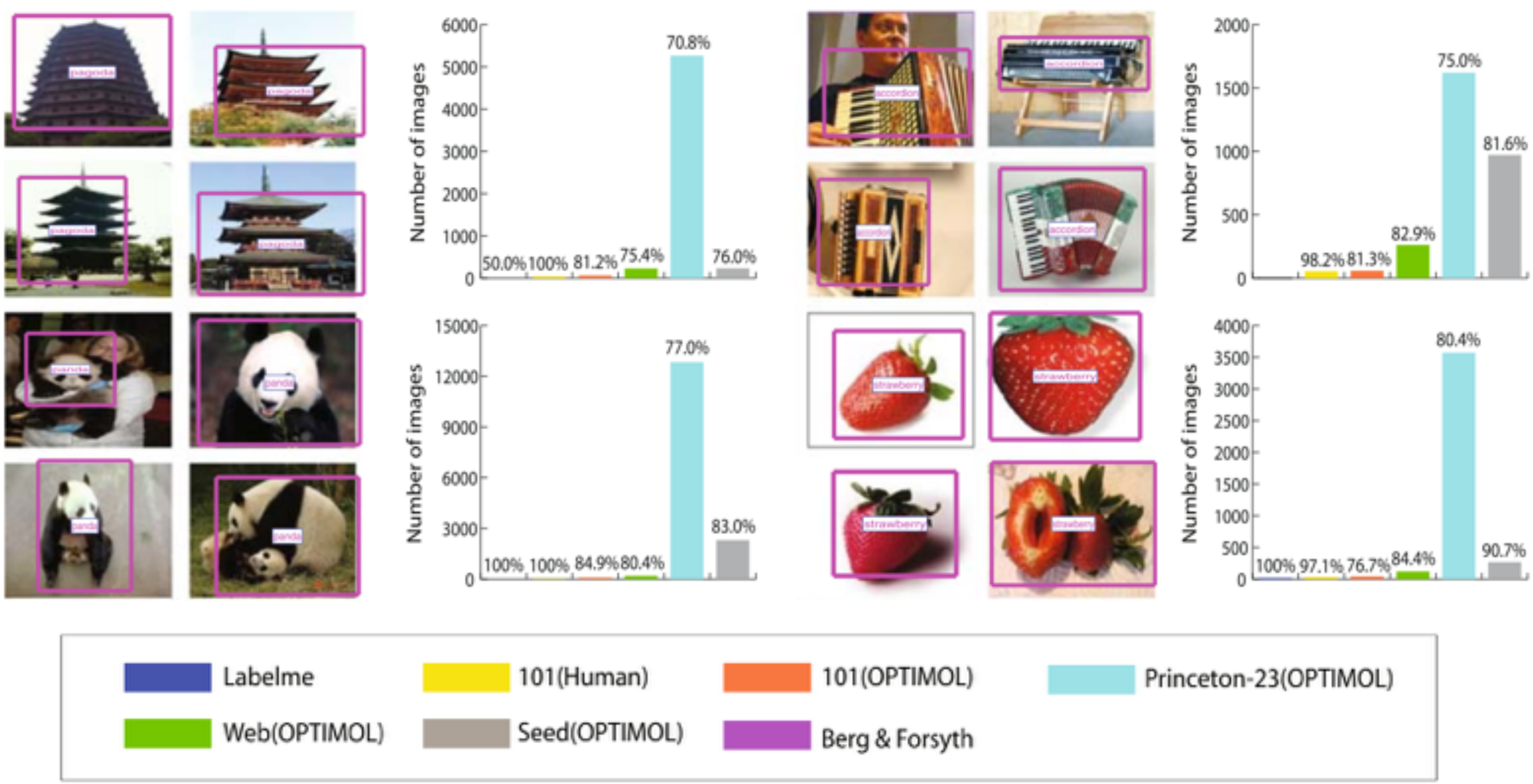
Experiments and Results

- Image collection and annotation results
  - Collected more or similar related images than other datasets



Experiments and Results

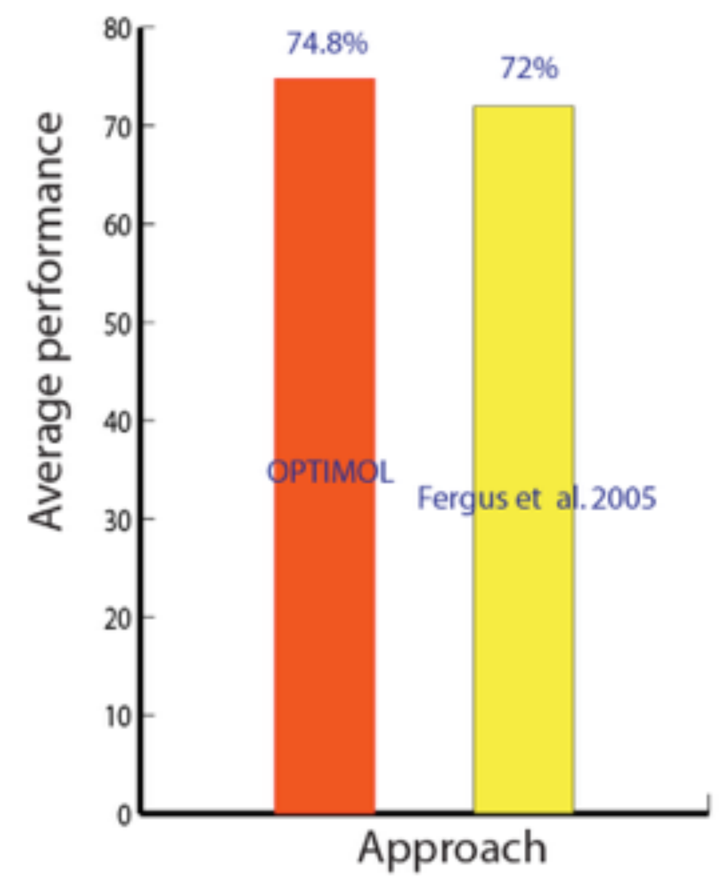
- Image collection and annotation results
    - Collected more related images than other datasets



Experiments and Results

- Classification results
  - Learned good model

	a	c	f	g	l	m	w
airplane	76.0	14.0	0.3	5.3	0.3	0.3	4.8
car	1.0	94.5	0.3	4.5	0.3	0.3	0.3
face	0.5	1.4	82.9	3.7	0.5	0.5	11.5
guita	2.2	4.9	5.6	60.4	13.3	0.2	13.3
leopard	1.0	2.0	1.0	5.0	89.0	1.0	2.0
motorbike	0.3	5.5	0.3	5.5	1.0	67.3	20.5
watch	1.7	5.5	17.7	11.0	5.5	5.0	53.6



Q&A

Thank you !