Informative Object Annotations - *“Tell Me Something I Don’t Know”*

Lior Bracha [1], Gal Chechik [1,2]

[1]Bar-Ilan University, Israel [2]Nvidia Research

Code & Paper

CVPR LONG BEACH CALIFORNIA June 16-20, 2019

The Problem of Informative Labeling

People effortlessly decide what an image is “about”, extracting informative relevant labels.

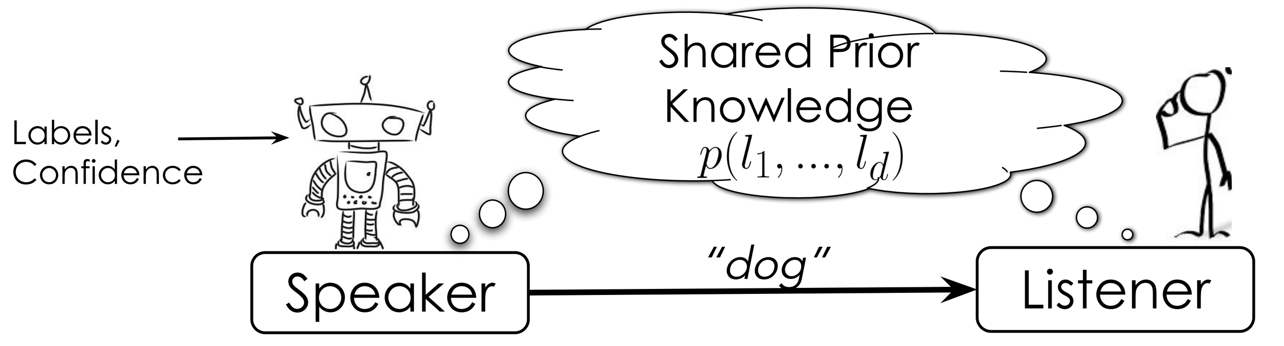
Automated classifiers can produce many technically correct labels, which may not be useful for communicating with people.

The Learning Setup

A speaker receives many machine-generated labels, selects k labels to communicate to a listener.

Multiclass Classifier

```
Label, Confidence
Grass,0.8      Green,0.7    Tennis ball,0.8
Animal,1       Lawn,0.7     Golden retriever,0.9
Dog toy,0.6    Dog,1        Pet,0.9
Mammal,0.9     Nature,0.9   Goldendoodle,0.6
```



The key idea: the speaker aims to **reduce the uncertainty of the listener.**

We look for an ***unsupervised*** method, from first principles.

Theories of Relevant Communication

- **Basic Level Categories** Rosch, 1976
  People prefer “intermediate” categories.
- **Cooperative Principles (Grice Maxims)** Grice, 1975
  Be as informative as possible; give as much information as is needed, and no more.
- **Relevance Theory** Wilson & Sperber,1995
  Be compatible with the communicator's abilities and preferences.

Goal: A ***quantitative*** theory of relevant communication that can be applied to concrete problems with real-world data.

Measures of Informative Communication

Reduce Uncertainty

The **entropy** over labels reflects what the listener “doesn’t know” about the image. Transmitting a label i reduces the entropy about the remaining labels.

$$\Delta H(i) = H[p(l_1,\dots,l_d)] - H[p(l_1,\dots,l_d \mid l_i = true)]$$

Entropy reduction favors fine-grained labels (dog over animal). The confidence of the prediction is important.

$$cw\text{-}\Delta H(i) = q(l_i \mid I)\Delta H(i)$$

Confidence scores $q(l_i)$ are **calibrated** and reflect the true fraction of correct labels.

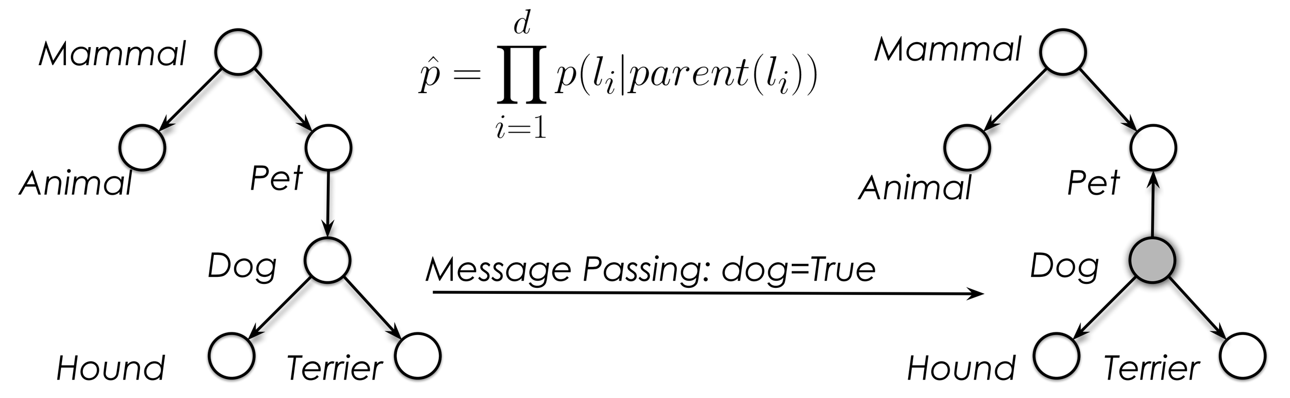
Modelling the Distribution Efficiently

Every joint distribution can be written as

$$p(l_1,\dots,l_d) = p(l_1)p(l_2|l_1)\cdots p(l_d|l_{d-1},\dots,l_1)$$

We approximate it using conditional independence

$$p(l_1,\dots,l_d) \approx p(l_1)p(l_2|l_1)\cdots p(l_d|l_{d-1})$$



The entropy factorizes over the tree, and can be computed efficiently

$$H[p(l_1,\dots,l_d)] = H\left[\prod_{i=1}^{d} p(l_i|Pa(l_i))\right] = \sum_{i=1}^{d} H[p(l_i|Pa(l_i))]$$

Other Measures of Informative Annotations

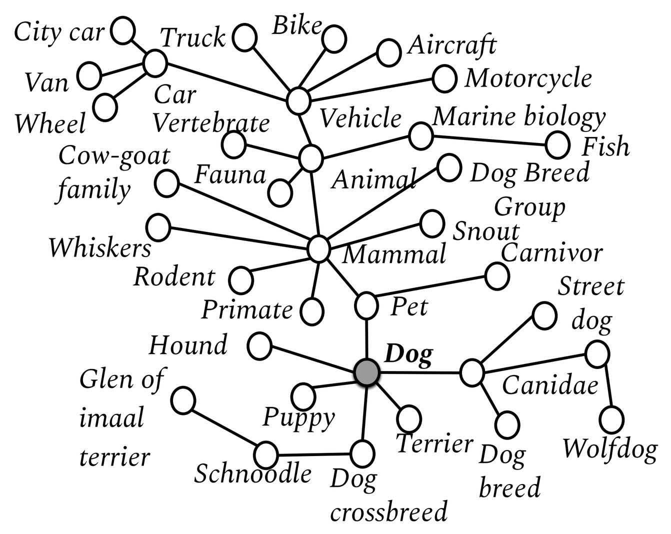
- **Probabilistic surprise**: changes to the label distribution
  $cw\text{-}D_{KL}(i) = q(l_i|I)D_{KL}(p(l_1,\dots,l_d|l_i)||p(l_1,\dots,l_d))$
- **Information about images** (reference game)
  $cw\text{-}Image\Delta H(i) = q(l_i|I)\log f(l_i)$
- **Entropy-of isolated labels** (Singleton)
  $cw\text{-}Singleton(i) = q(l_i|I)H(l_i)$

Chow-Liu Tree

The model captures semantic relations from co-occurrence data.

Predefined hierarchical models (WordNet , OID KG) were inferior.

The subtree around “dog”



Evaluation Protocol

The Data for modelling label distribution: The *Open Images Dataset* test set: 125K images with 1.5M labels.

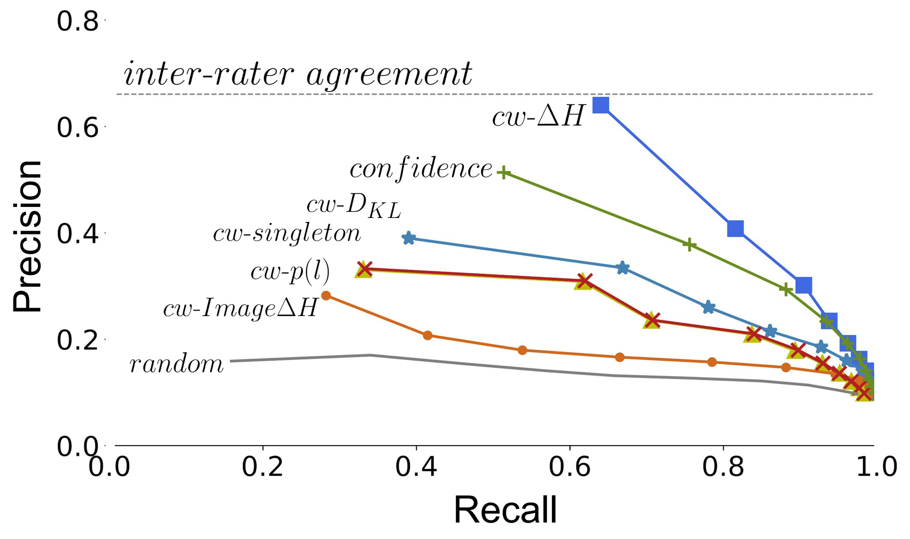
Evaluation Ground-truth: New set: 10K images, 3 raters/image.

Metrics: For each image, we rank labels by the above metrics, and compute precision and recall.

	confidence	ΔH	$cw\text{-}\Delta H$
Animal	1	36.37	36.37
Small to medium-sized cats	0.8	**40.67**	32.54
Pet	1	39.95	39.95
Whiskers	0.7	40.63	28.44
Carnivore	0.5	38.95	19.47
Cat	1	40.24	**40.24**
Mammal	1	38.5	38.5

Results

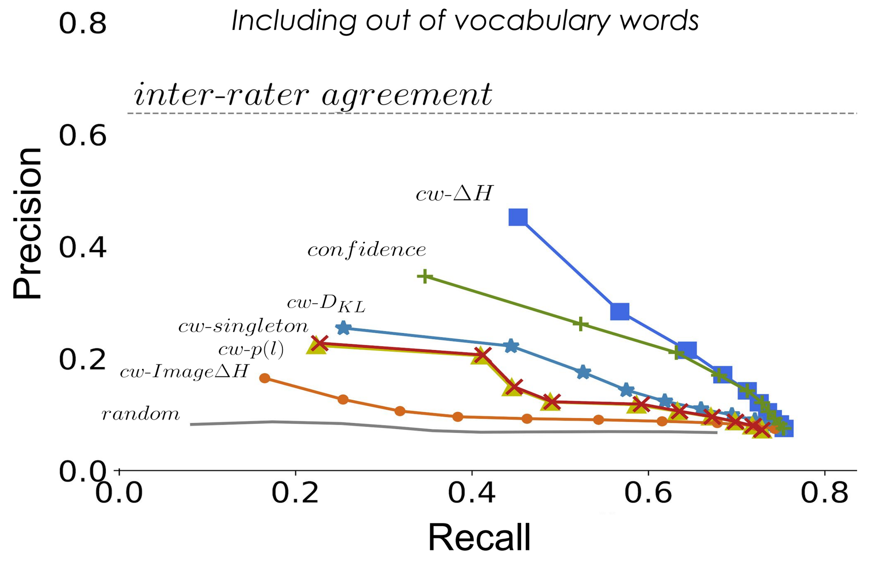
The expected entropy reduction $cw\text{-}\Delta H$ achieves precision@1 almost as good as inter-rater agreement.



No training data about relevance was used.

Pairwise information is much better that isolated labels (singletons).

Using confidence alone is also highly predictive.



Qualitative Examples			
Confidence $q(l)$	**Shoe**, footwear, purple	**Leaf**, plant, tree, nature, yellow, green	land vehicle
$cw\text{-}\Delta H$	**Shoe**	**Leaf**	**Car**
$cw\text{-}D_{KL}$	**Shoe**	Autumn	Mercedes-benz
$cw\text{-}Image\Delta H$	Violet	Season	Mercedes-benz
$cw\text{-}p(l)$	Purple	Plant	Vehicle
$cw\text{-}Singleton$	Purple	Plant	Vehicle