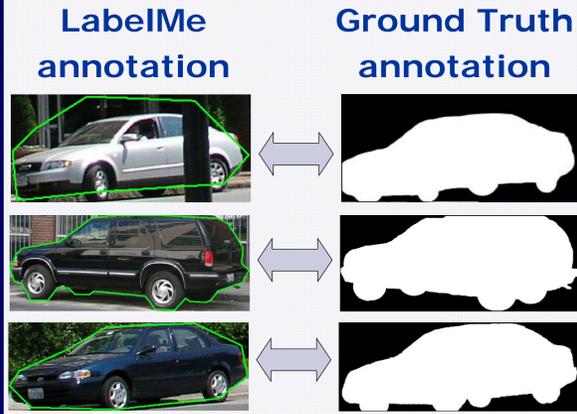


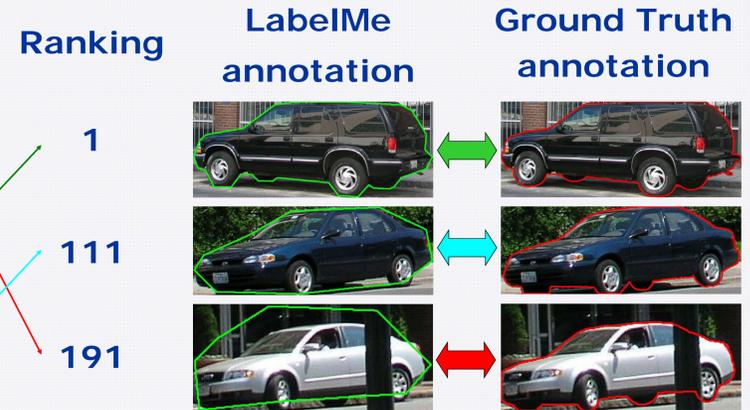
# Quality Assessment for Crowdsourced Object Annotations

Sirion Vittayakorn and James Hays, Brown University

## Database Creation



## "Ground Truth" Annotation Quality Ranking



Ranking based on **Overlap score** and **Boundary agreement**.

Consider two annotations  $B_u$  and  $B_v$  and its corresponded points  $(X_i, Y_i) \in B_u$  and  $(x_i, y_i) \in B_v$

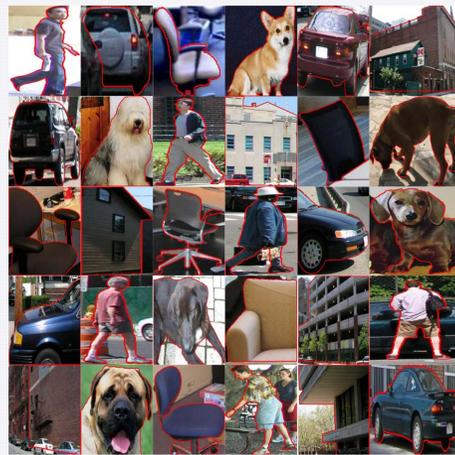
• **Overlap Score:**  $score = \frac{area(B_u \cap B_v)}{area(B_u \cup B_v)}$

where  $B_u \cap B_v$  denotes the intersection of two Annotations and  $B_u \cup B_v$  their union.

• **Euclidean Distance Score:**  $dist = \sum_i \sqrt{(X_i - x_i)^2 + (Y_i - y_i)^2}$   
**(Boundary agreement)**  $score = 1 - \frac{dist}{max(dist)}$

where  $max(dist)$  is the maximum Euclidean distance of that category.

## Overview of the Database



Ground truth rigorously defined for 200 objects in 5 categories.

## Ranking Agreement between Scoring Functions and "Ground Truth"

Category	Spearman's Rank Correlation					
	Points	Size	Edge	Bayesian	Proposal	Final
Car	0.5216	0.4356	0.5972	0.3848	0.0817	<b>0.5999</b>
Chair	0.6758	0.6519	0.6132	0.6780	0.0190	<b>0.6947</b>
Building	-0.3874	0.4271	0.4055	0.2030	0.0386	<b>0.5214</b>
Person	0.5503	0.4386	0.5716	0.7036	0.0394	<b>0.7072</b>
Dog	0.6070	0.2367	0.6932	0.6503	0.0468	<b>0.7689</b>
Average	0.3935	0.4380	0.5761	0.5239	0.0232	<b>0.6584</b>

## Annotation Scoring Function and Proposed Ranking

### Baseline: #points

Score proportional to #points in an annotation



### Baseline: size

Score proportional to annotation area



### Image edge agreement

Score proportional to the degree of erosion or dilation necessary to achieve the best possible overlap of image edges and annotation boundary.



### Bayesian matting agreement

Three regions: "background", "object" and "unknown" are created based on the annotation. Given these regions, Bayesian matting returns the fractional opacities,  $\alpha$ , which can be interpreted as a confidence of each pixel belonging to the object. Then, the score is  $\sum \alpha_{in} / area_{in} - \sum \alpha_{out} / area_{out}$  where  $\alpha_{in}$  and  $\alpha_{out}$  are the opacity of each pixel within the unknown region inside and outside the annotation respectively while  $area_{in}$  and  $area_{out}$  are the area of the unknown region inside and outside the annotation.



### Object proposal probability

The method of Endres and Hoiem creates a ranked list of possible objects in the current image. The score is proportional to the rank of the object that is most similar to the user annotation.



### Final ranking

The final score is the combination of the Bayesian matting score and edge agreement score with equally weight.

