Category-Independent Object-level Saliency Detection

by Yangqing Jia and Mei Han

Aysun Koçak

Outline

- Introduction
- Related Work
- Method
- Results
- References

Introduction

- Computer vision applications may benefit from understanding where humans focus given a scene

- Finding salient regions and objects in the images helps various tasks such as speeding up object detection, image and video compression...

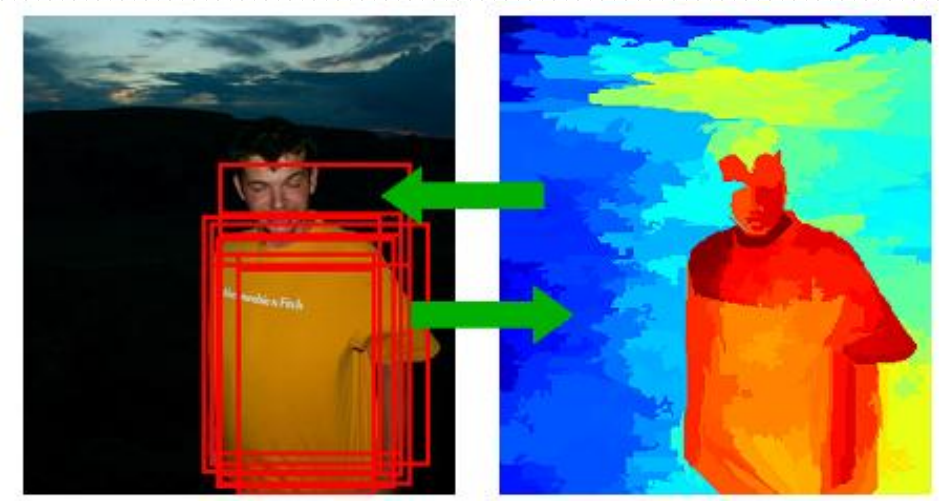
Introduction

- Saliency estimation methods can be implemented by following three distinct models: bottom-up model, top-down model and combination of both.

- The proposed approach fuses top-down object level information and bottom-up pixel appearances to obtain a final saliency map that identifies the most interesting regions in the image.

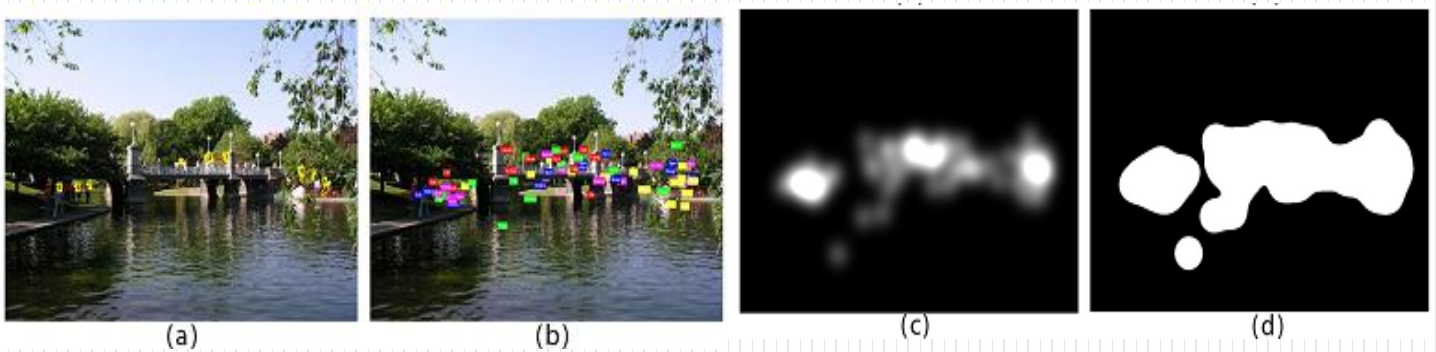
Related Work

- *Fusing Generic Objectness and Visual Saliency for Salient Object Detection(ICCV-2011)*: first to adopt a high-level object information as saliency prior



Related Work

- *Learning to Predict Where Humans Look(ICCV-2009)*:
  - method extracts low-, mid-, high-level features(33 features)
  - uses a learning approach to train a classifier directly from human eye tracking data.



Method

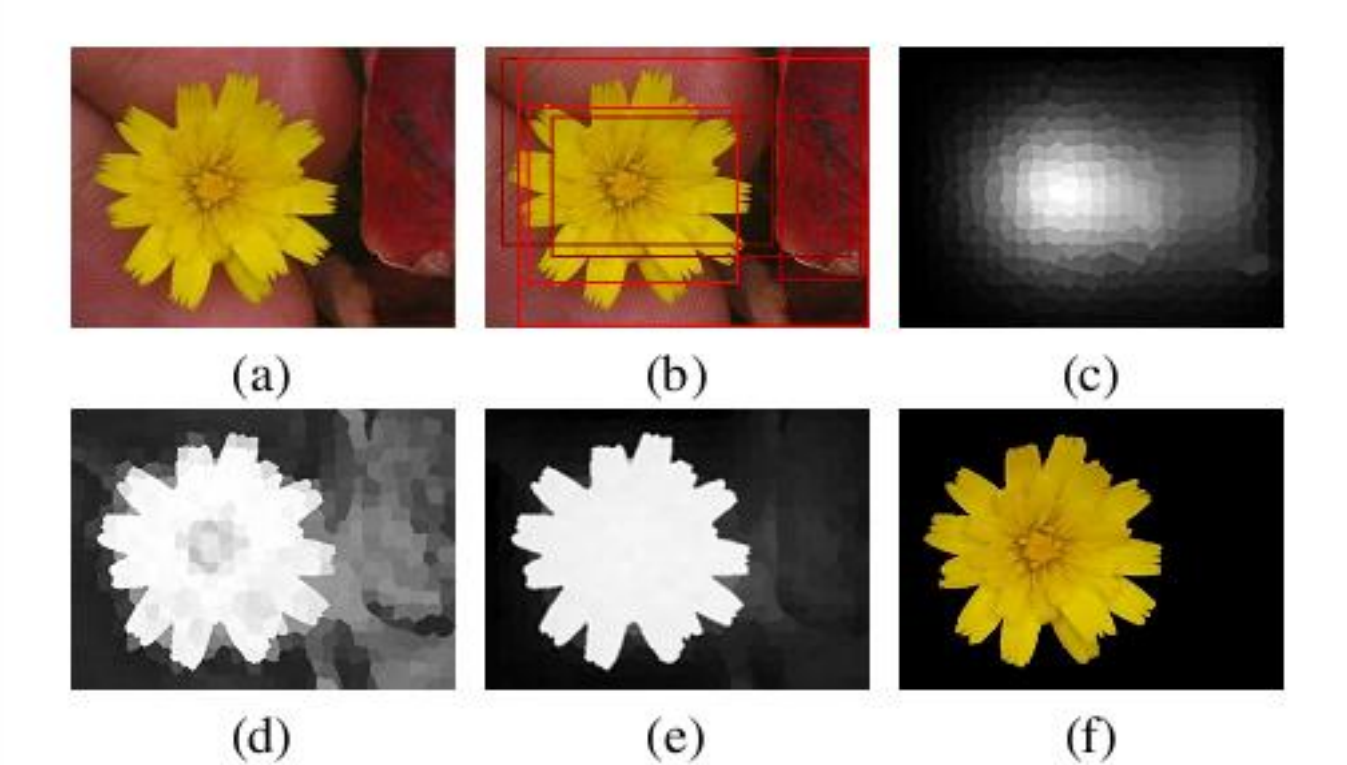
- The proposed method gets high-level object information and low-level pixel appearance then fuse them with a fully-connected Markov random field

- Two steps;

    1. Object detection for high-level information

    2. Saliency computation with MRF

Method



Method

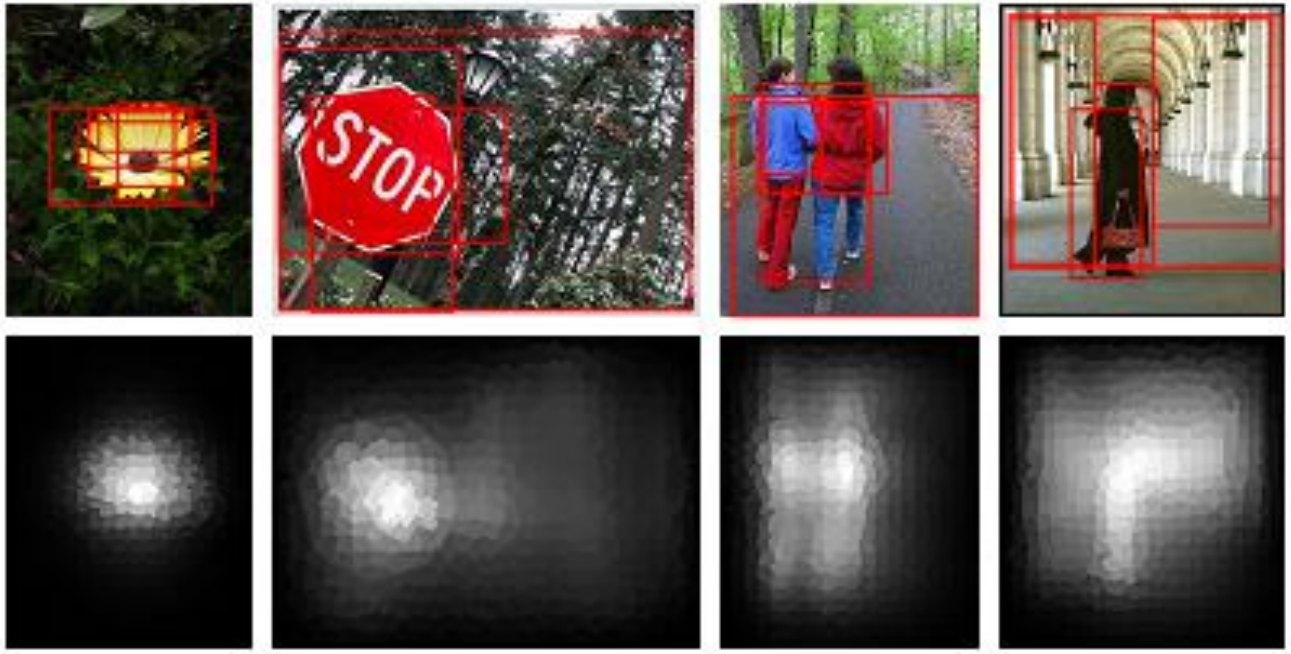
1. Object Detection

- An objectness algortihm is used for detection: *What is an object?, (Bogdan Alexe, Thomas Deselaers, Vittorio Ferrari – CVPR 2010)*

- *It adopts four different low-level cues: Multiscale saliency, color contrast, edge density, superpixel straddling.*

Method

1. Object Detection



Method

1. Object Detection

- Pixel level objectness scores

$$s_p = \left[\sum_{i=1}^{N} b_i^2 I(p \in B_i) \exp\{-\lambda d(p, B_i)\} \right]^{1/2}$$

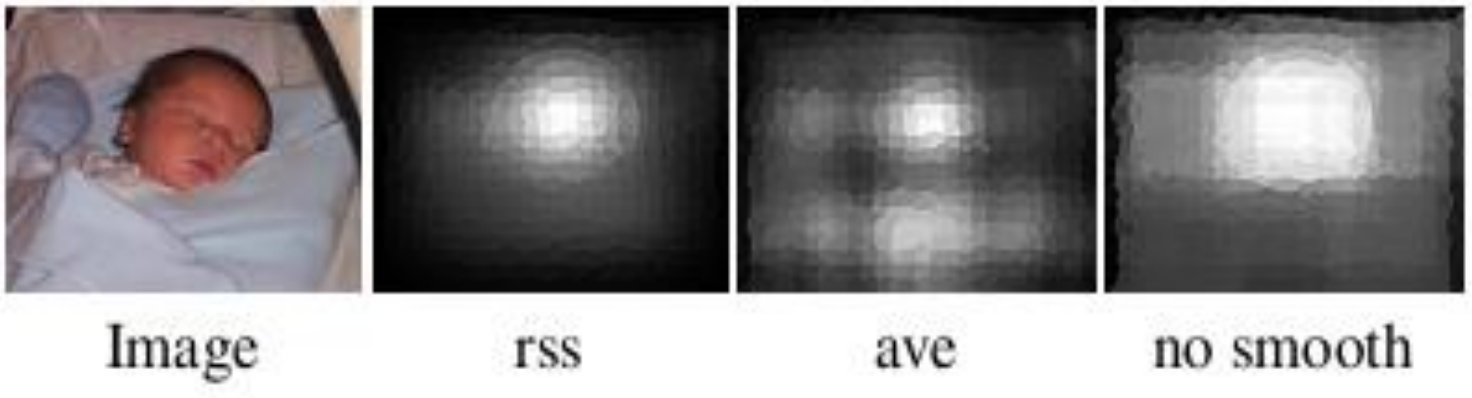
- To reduce the computation cost for subsequent steps they adopted the idea of superpixels and averaged the saliency values of pixels inside each superpixel.

- Turbopixel algorithm is used to product superpixel that have similar sizes.

Method

1. Object Detection

- Pixel level objectness scores



Method

2. Saliency Computation

- objectness bounding boxes are often over-complete, and the saliency map is often very coarse
- low-level appearance based information to be helpful in refining the saliency maps.
- The proposed method extracts features for individual pixels, and use a MRF to enforce agreement between salient regions in the image, based on the similarities between pixel level features.

Method

2. Saliency Computation

- They use fully connected MRF where any two super-pixels are connected :

$$W_{ij} = \exp\left(\frac{-\|p_i, p_j\|_2^2}{2\sigma^2}\right)$$

$$\mathrm{DD}_i = \sum_j \Big(W_{ij} s_j + (1 - W_{ij})(1 - s_j)\Big)$$

Method

$$\mathrm{DD}_i \leftarrow \left(\frac{\mathrm{DD}_i}{\max_j \mathrm{DD}_j} \right)^{\gamma}$$

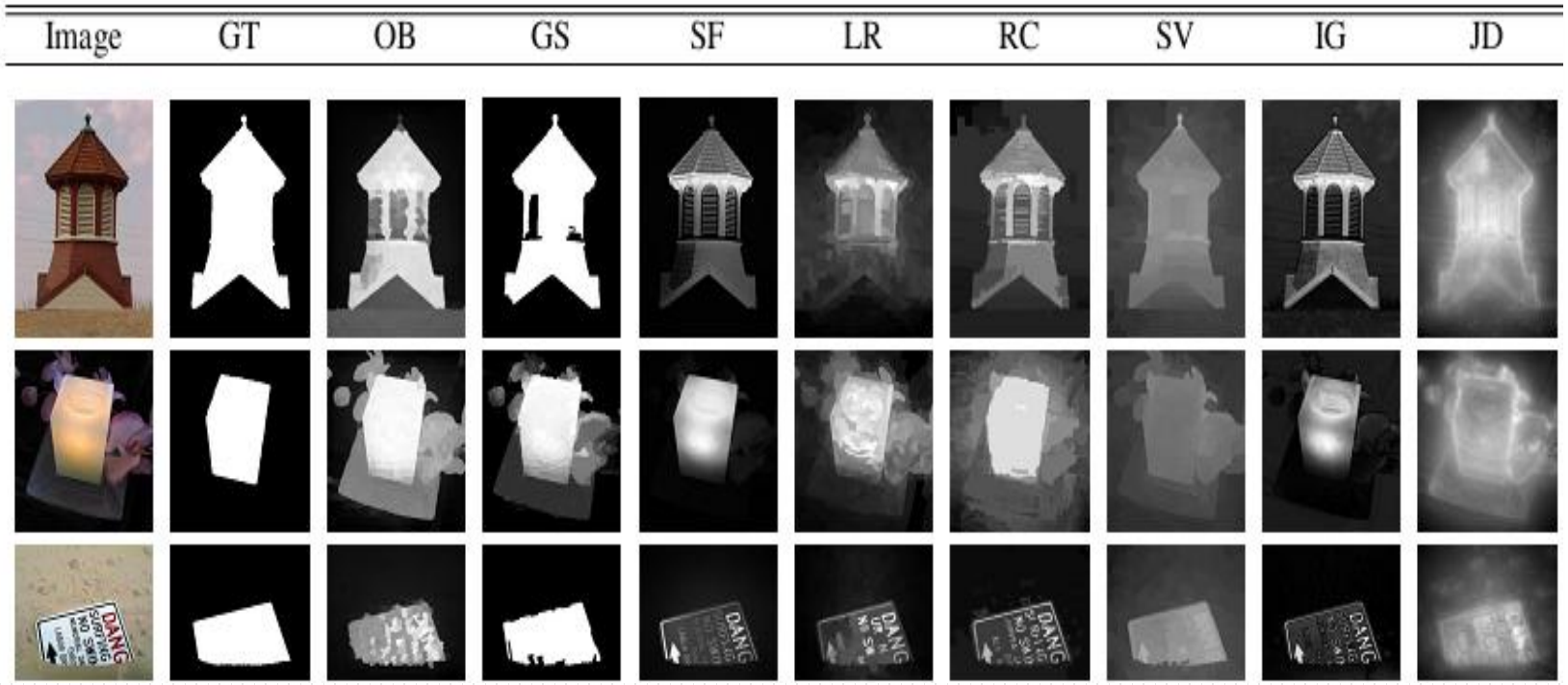
Method

2. Saliency Computation

$$G_{ij} = \frac{\mathrm{DD}_i + \mathrm{DD}_j}{2} W_{ij}$$

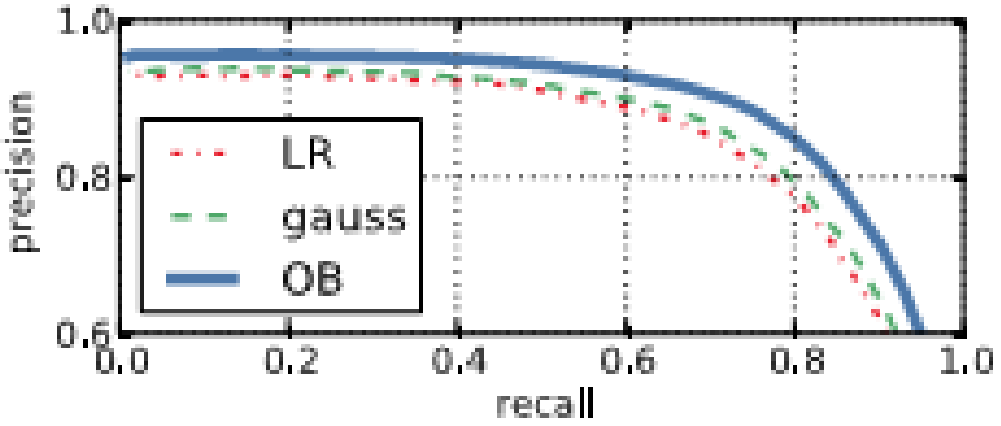
$$\hat{\mathbf{s}} = \left(\mathrm{diag}(\mathbf{G1}) - \mathbf{G}\right)^{+} \begin{bmatrix} \mathbf{s} & 1-\mathbf{s} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Method



Method

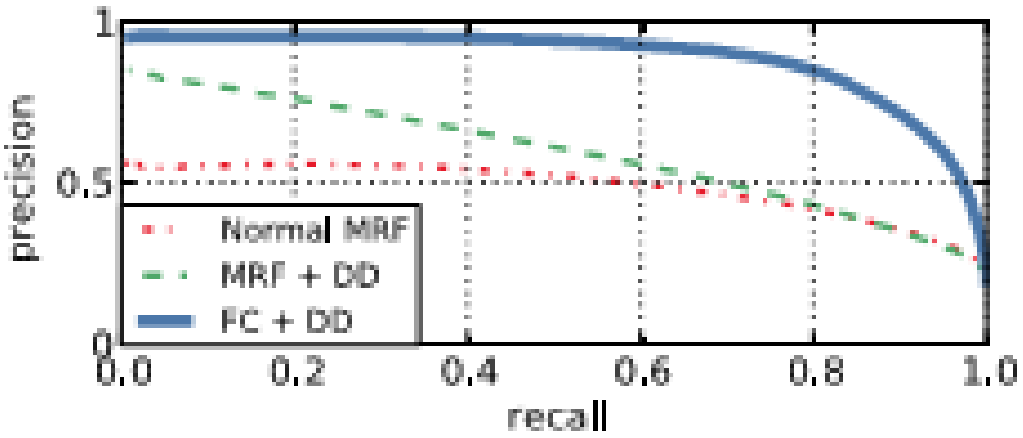
- The prior effect



(a) Different Prior

Method

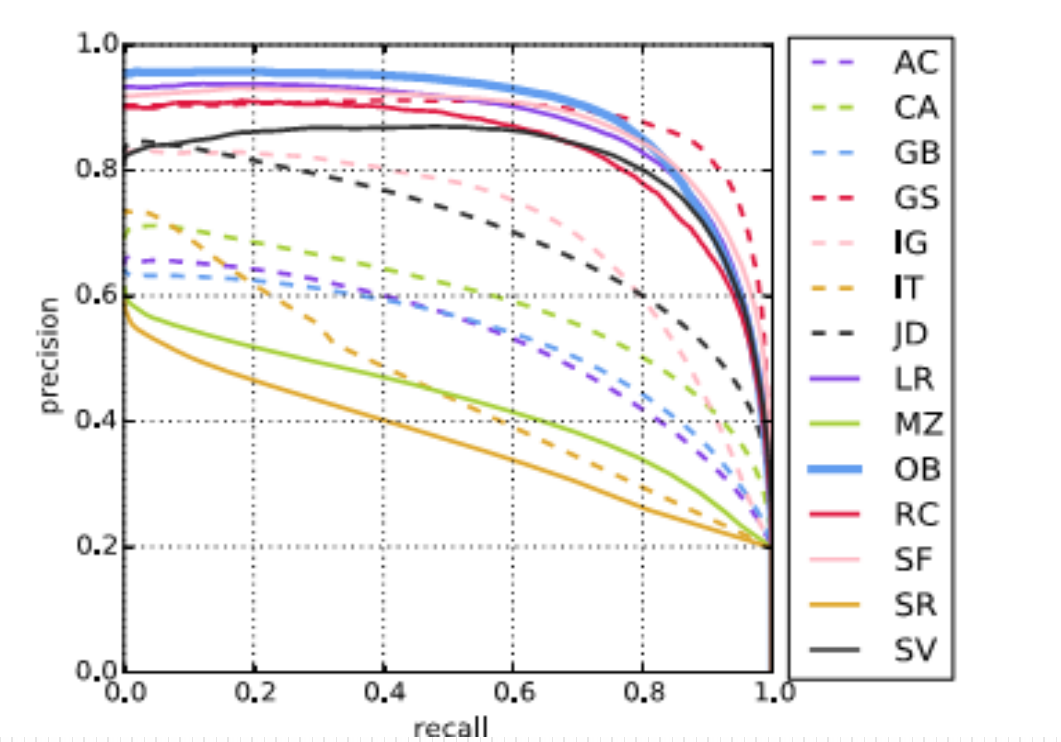
- The graph effect



(b) Different Graph

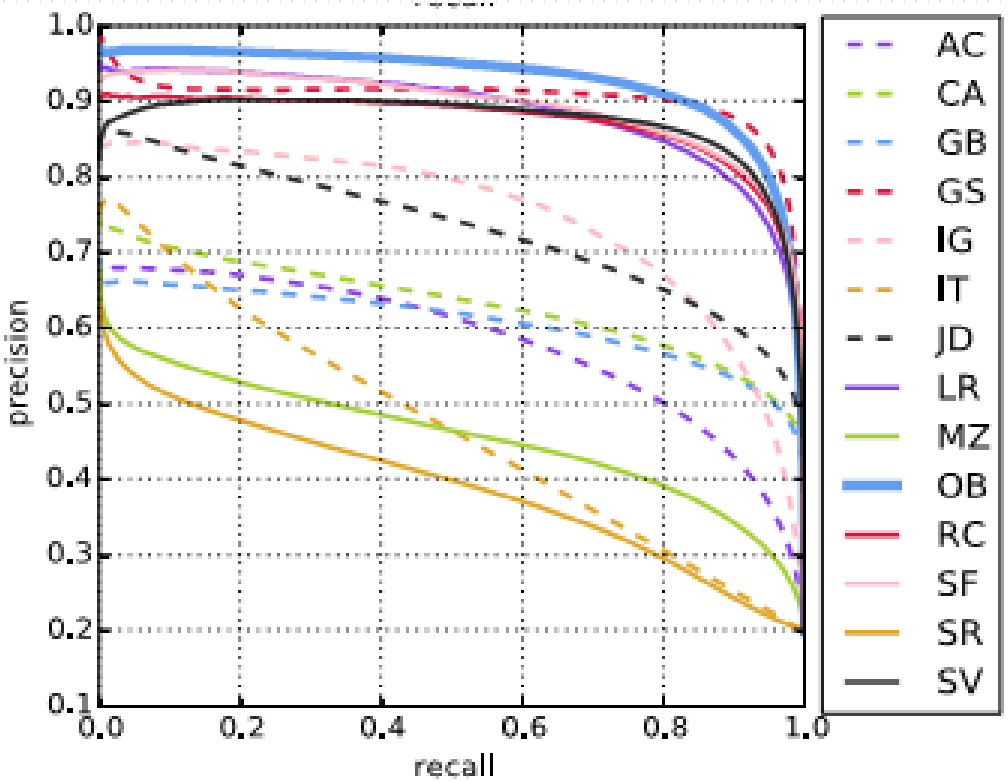
Results

- MSRA Dataset
  - PR-overall:



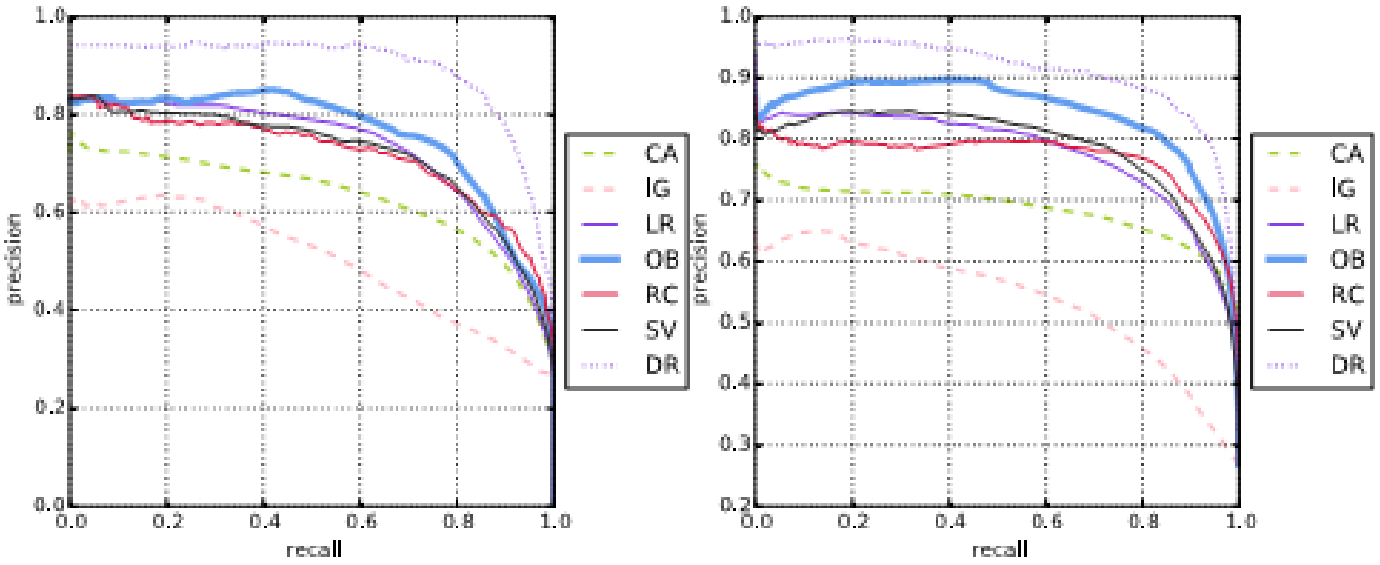
Results

- MSRA Dataset
  - PR- individual :



Results

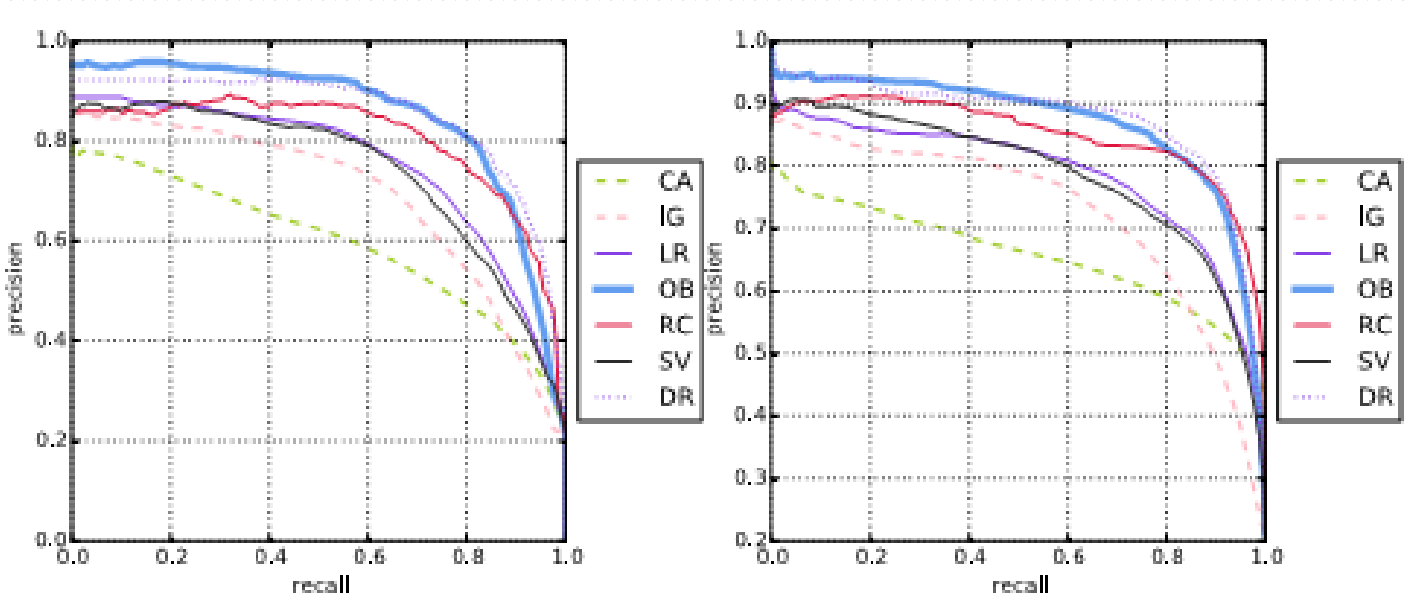
- Weizmann Dataset



(a) Weizmann 1 Object Dataset

Results

- Weizmann Dataset



(b) Weizmann 2 Object Dataset

References

- B. Alexe, T. Deselaers, and V. Ferrari. What is an object? In CVPR, 2010
- K.Y. Chang, T.L. Liu, H.T. Chen, and S.H. Lai. Fusing generic objectness and visual saliency for salient object detection. In CVPR. IEEE, 2011
- T. Judd, K. Ehinger, F. Durand, and A. Torralba. Learning to predict where humans look. In ICCV, 2009

Thanks