Salient object detection, objectness proposals, and potential applications

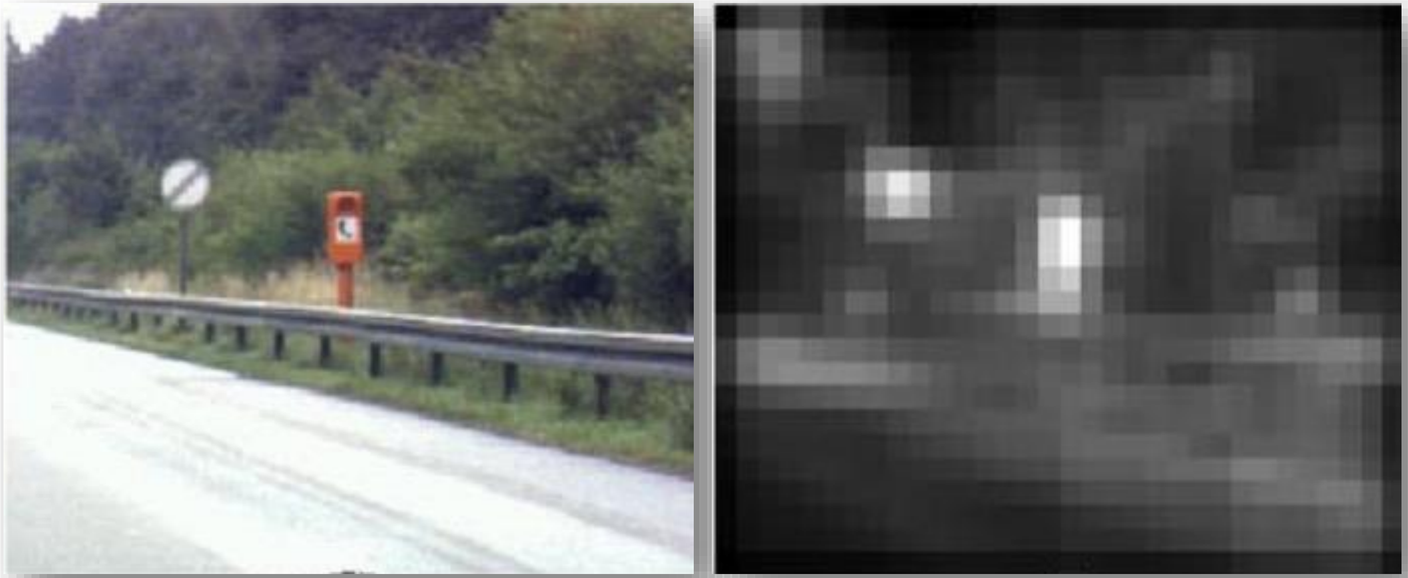
Slides: Ming-Ming Cheng

Presenter: Jian Li

Tutorial webpage: http://mmcheng.net/saltutorial/

Predict fixation → detect salient object

- Fixation prediction
  - Predicting saliency points of human eye movement



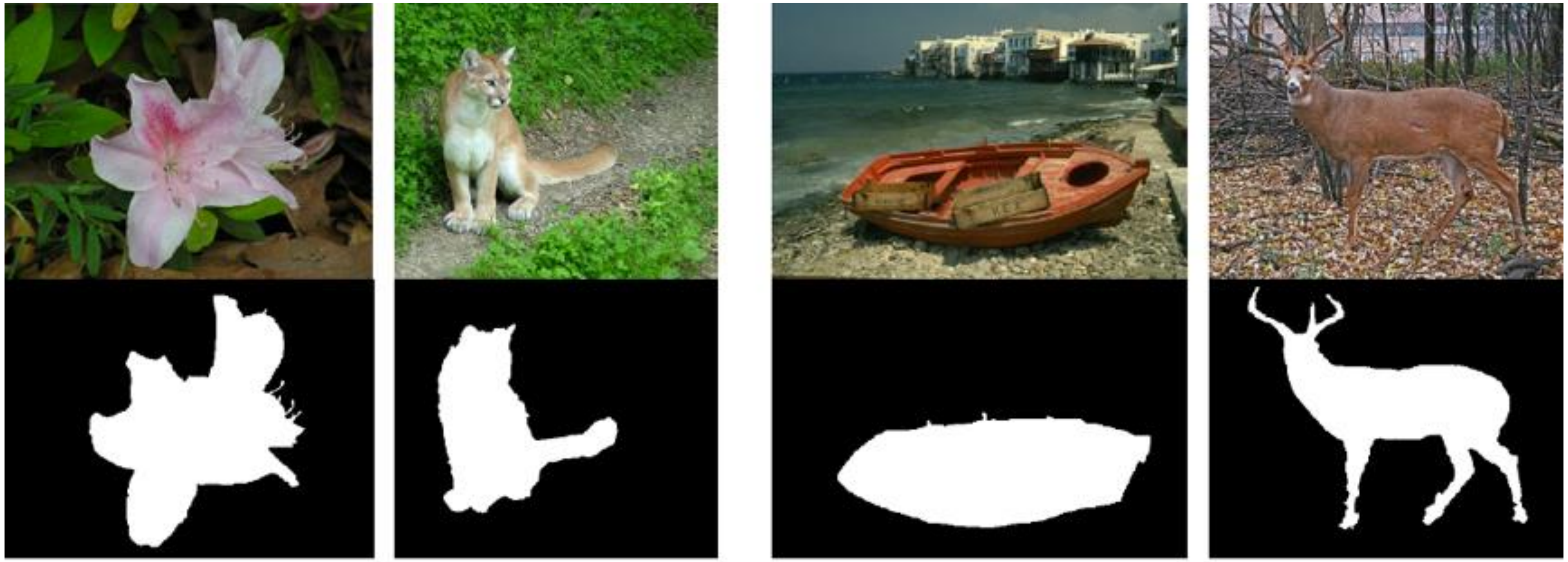
- A model of saliency-based visual attention for rapid scene analysis. PAMI 1998, Itti et al.
- Saliency detection: A spectral residual approach. CVPR 2007, Hou et. al.
- Graph-based visual saliency. NIPS, Harel et. al.
- Quantitative analysis of human-model agreement in visual saliency modeling: A comparative study, IEEE TIP 2012, Borji et. al.

Predict fixation → detect salient object

- Eye tracker
  - Cognitive psychology, neurobiology, etc.

Predict fixation → detect salient object

- Saliency detection as binary segmentation



(a) **MSRA10K** (b) **ECSSD**

- Learning to detect a salient object. CVPR 2007, Liu et. al.
- Frequency-tuned salient region detection, CVPR 2009, Achanta et. al.
- Global contrast based salient region detection, CVPR 2011, Cheng et. al.
- Salient object detection: a benchmark, arXiv 2015, Ali et. al.

Salient object detection

- How to define salient objects?
  - Match the human annotators’ behavior when they have been asked to pick a salient object in an image.

Global contrast based salient region detection, IEEE TPAMI 2015 (CVPR 2011), Cheng et. al.

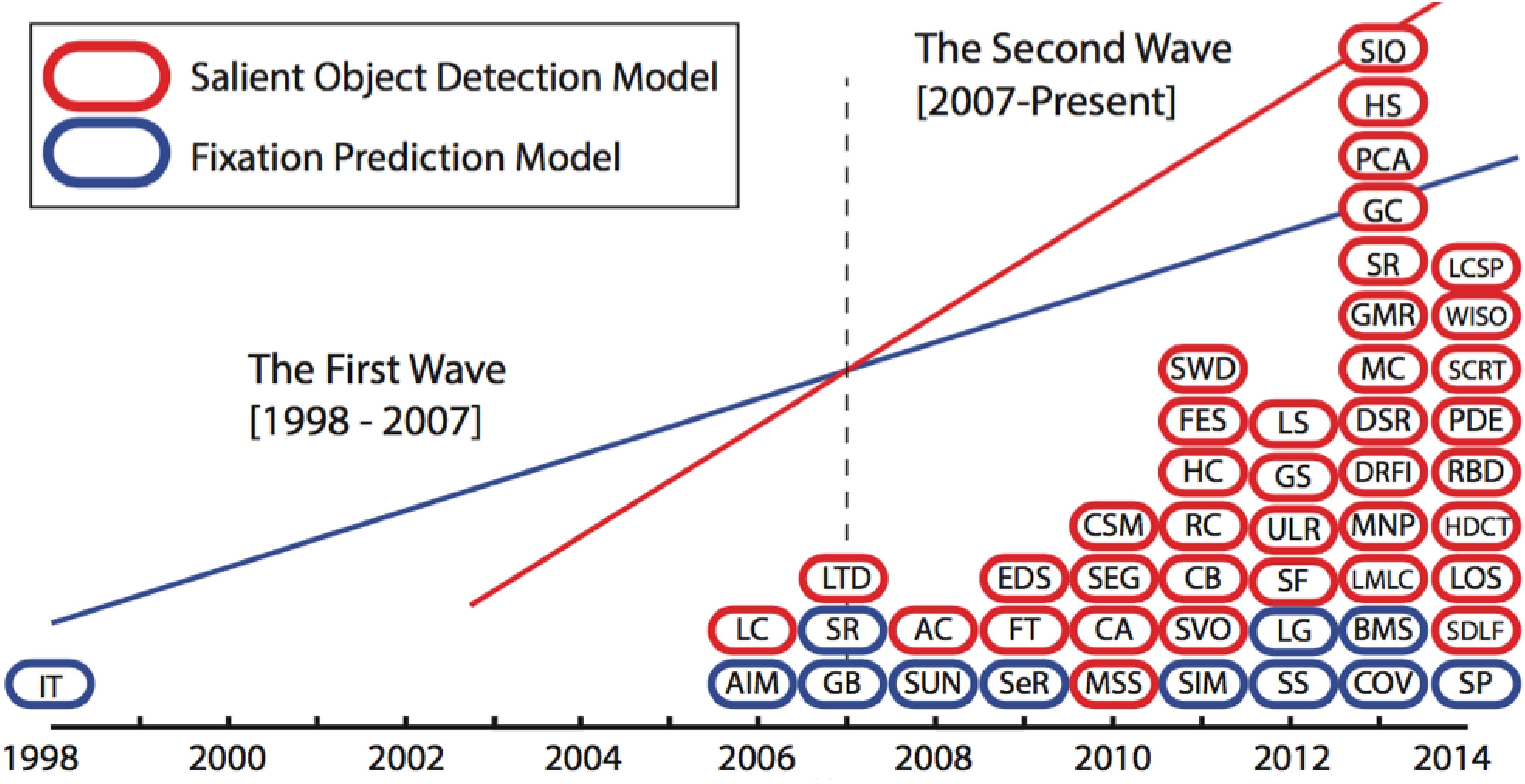
Salient object detection

- High consistency among labelers.

PASCAL-S dataset

The Secrets of Salient Object Segmentation, CVPR 2014, Li et. al.

Growing interest



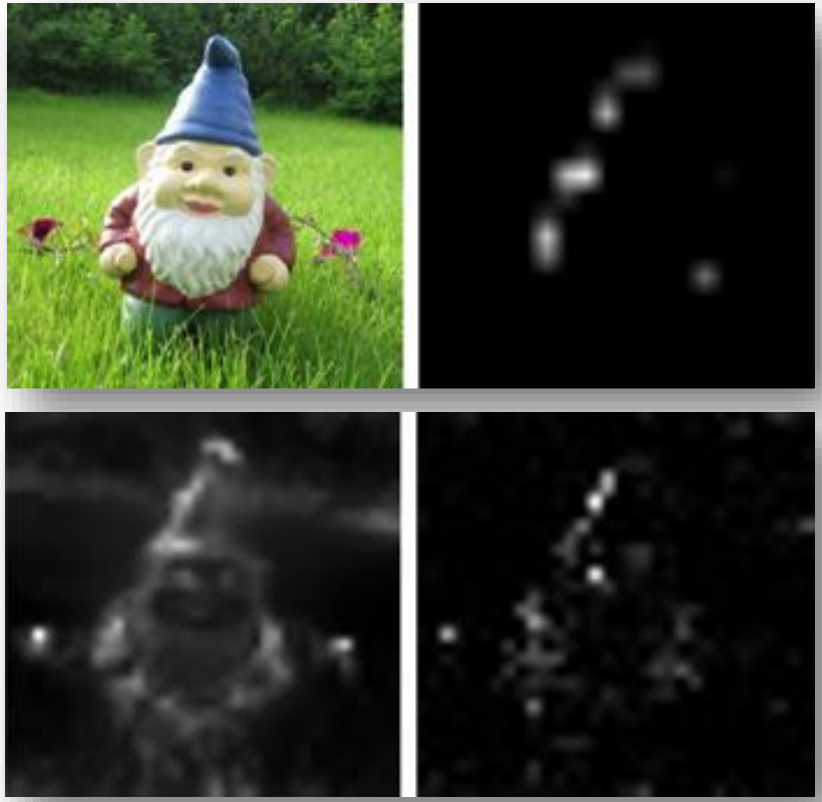
Salient object detection: a survey, arXiv 2014, Ali et. al.

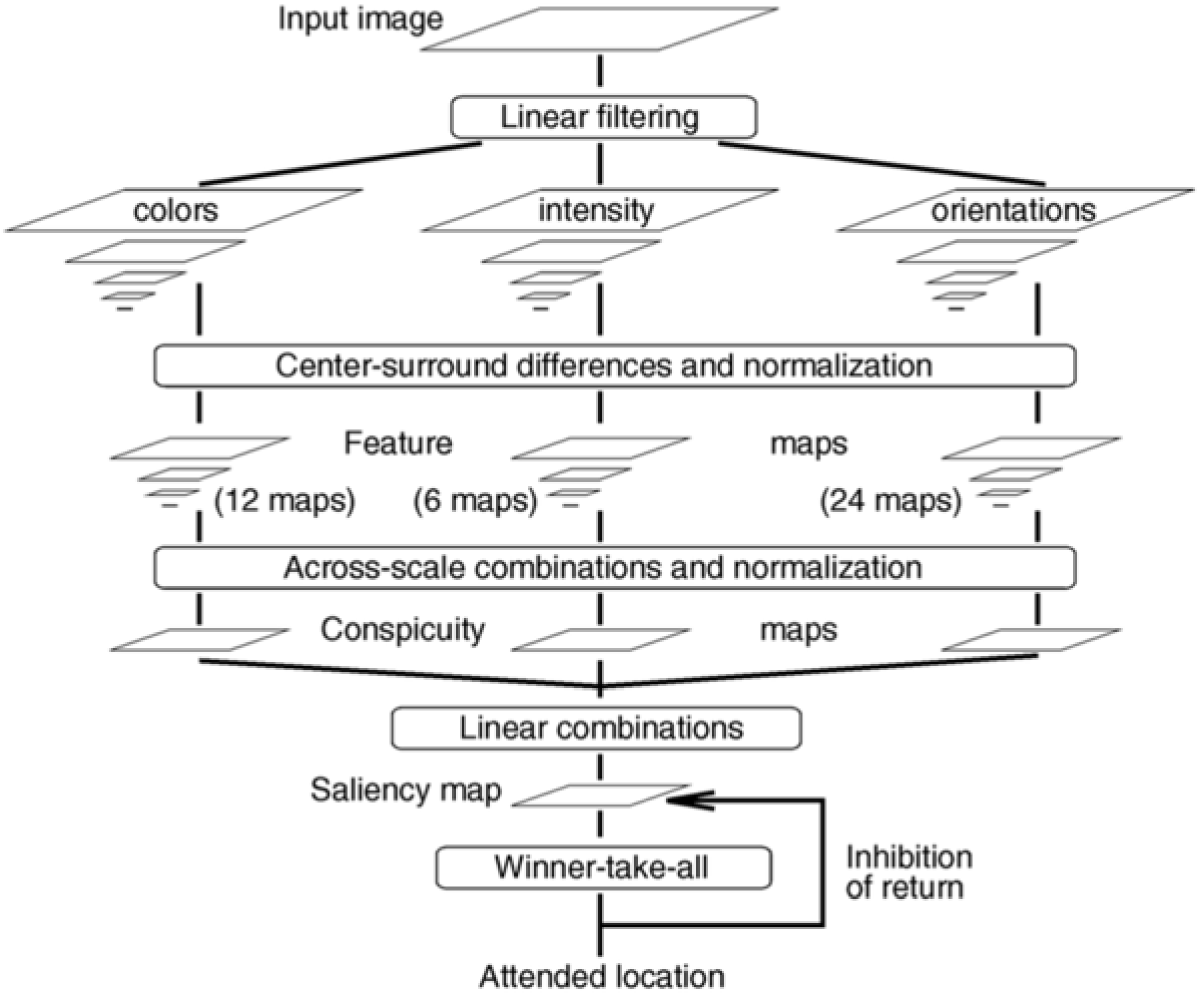
What makes an object salient?

Salient object detection: a survey, arXiv 2014, Ali et. al.

Hypothesis for Salient object detection

- Local contrast

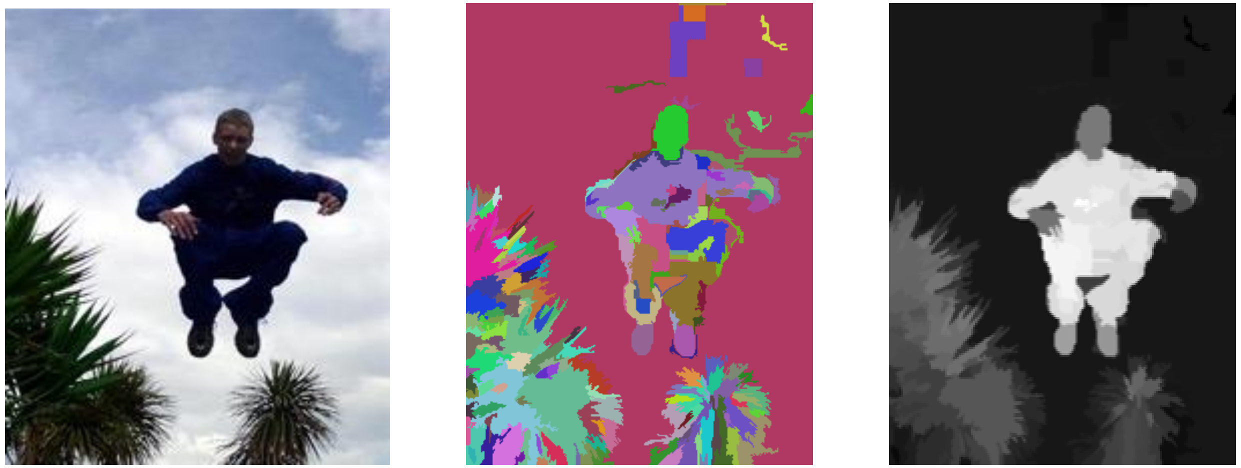




A Model of Saliency-Based Visual Attention for Rapid Scene Analysis, IEEE TPAMI 1998, Itti et al.

Hypothesis for Salient object detection

- Global contrast



Spatial weighting

Region size

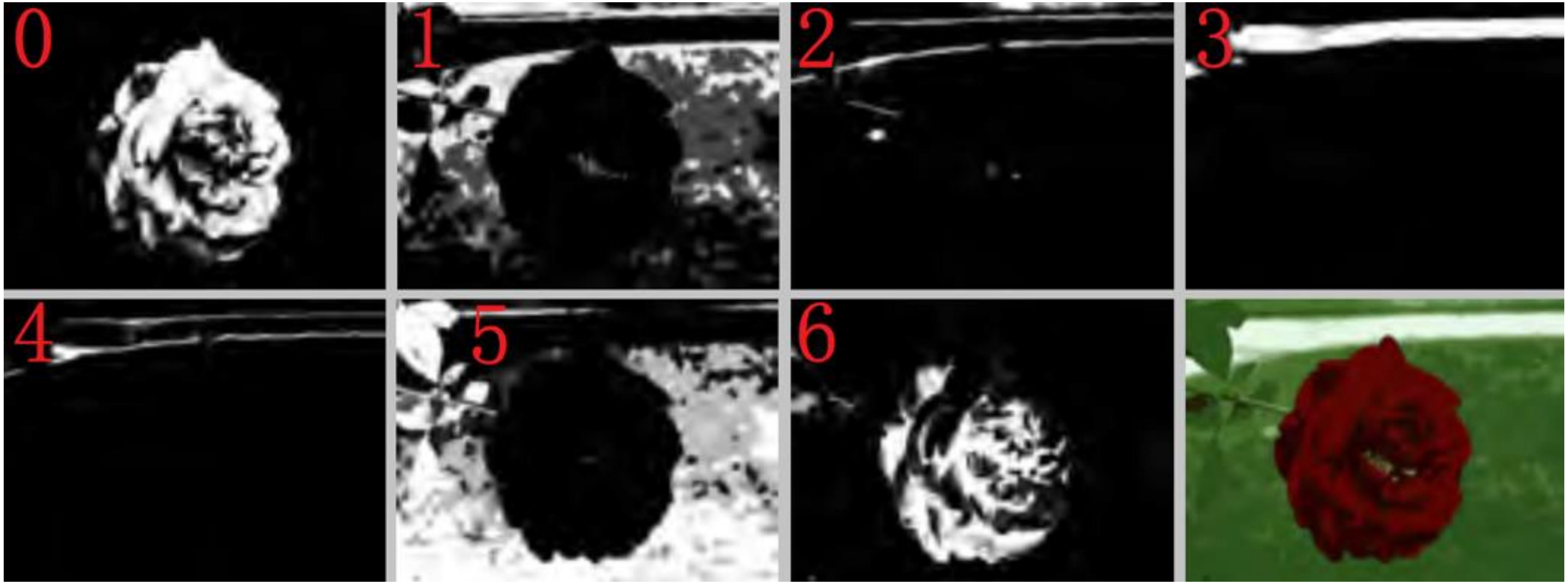
$$S(r_k) = \sum_{r_k \neq r_i} \exp\left(-D_s(r_k, r_i)\right) w(r_i) D_r(r_k, r_i)$$

Region contrast by sparse histogram comparison.

Global Contrast based Salient Region detection. IEEE TPAMI 2015 (CVPR 2011), Cheng et al.

Hypothesis for Salient object detection

- Spatial distribution



- Learning to Detect a Salient Object, IEEE TPAMI 2011 (CVPR 2007), Liu et al.
- Efficient Salient Region Detection with Soft Image Abstraction, ICCV 2013, Cheng et al.

Hypothesis for Salient object detection

- Sparse noises
  - Background residing in a low dimensional space with salient objects as sparse noises

A Unified Approach to Salient Object Detection via Low Rank Matrix Recovery, CVPR 2012, Shen et. al.

Hypothesis for Salient object detection

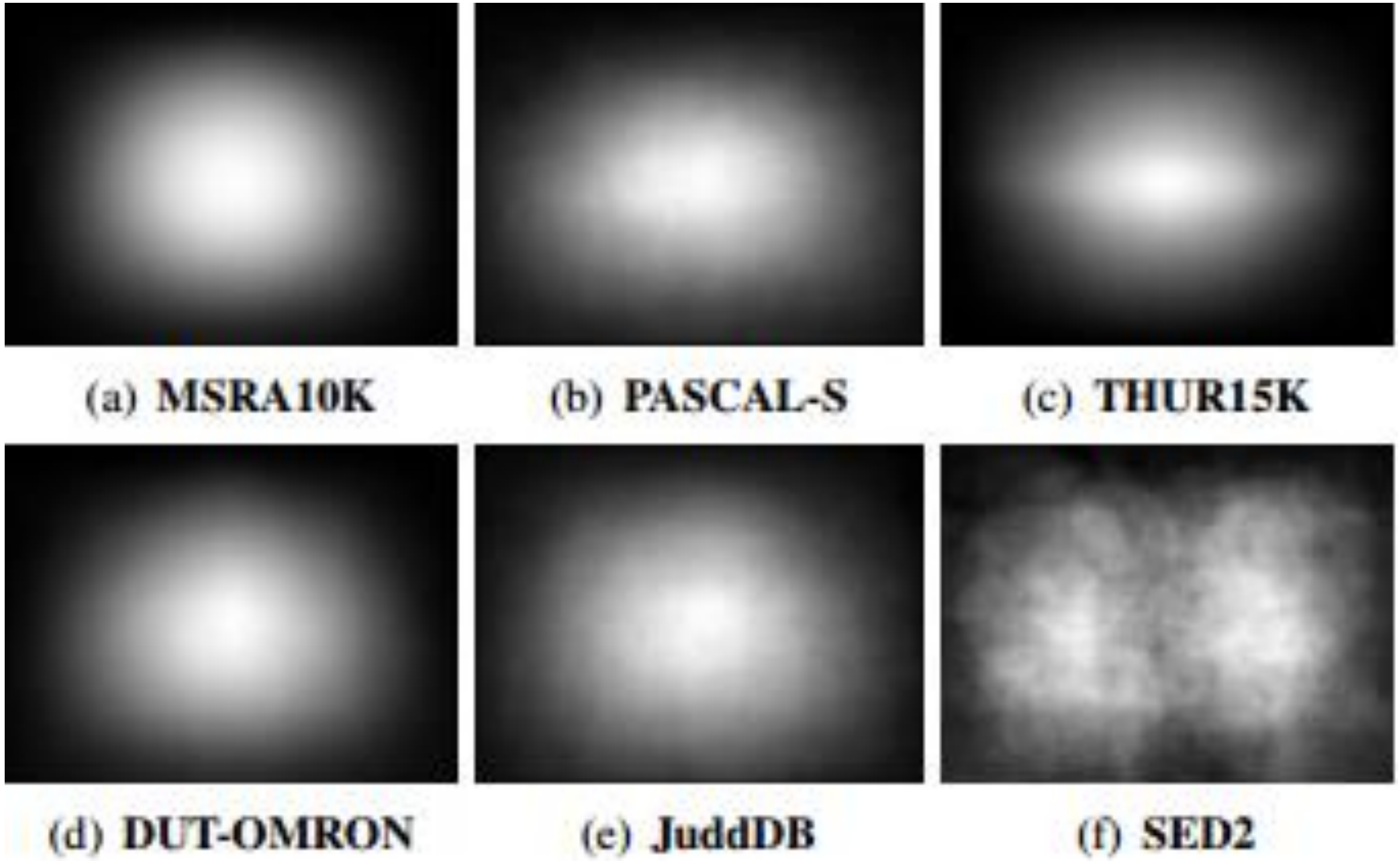
- Focusness

Image Uniqueness Focusness

Salient Region Detection by UFO: Uniqueness, Focusness and Objectness, ICCV 2013, Jiang et al.

Hypothesis for Salient object detection

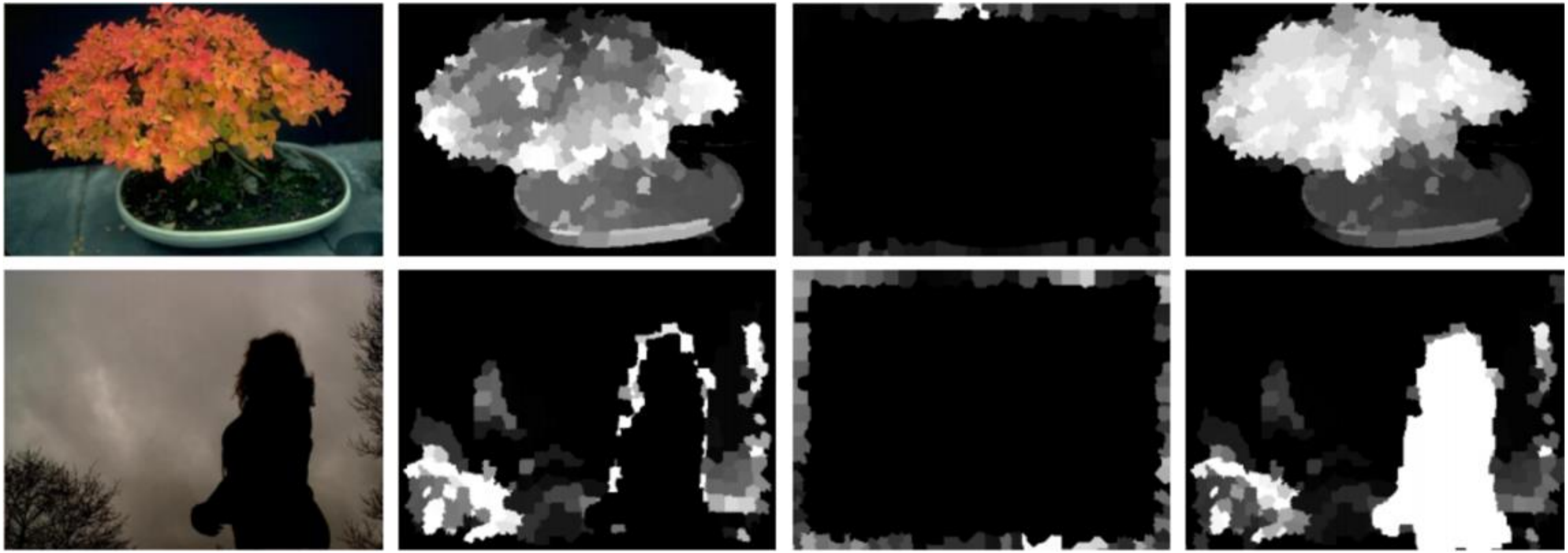
- Center prior/bias



Salient object detection: a benchmark, arXiv 2015, Ali et. al.

Hypothesis for Salient object detection

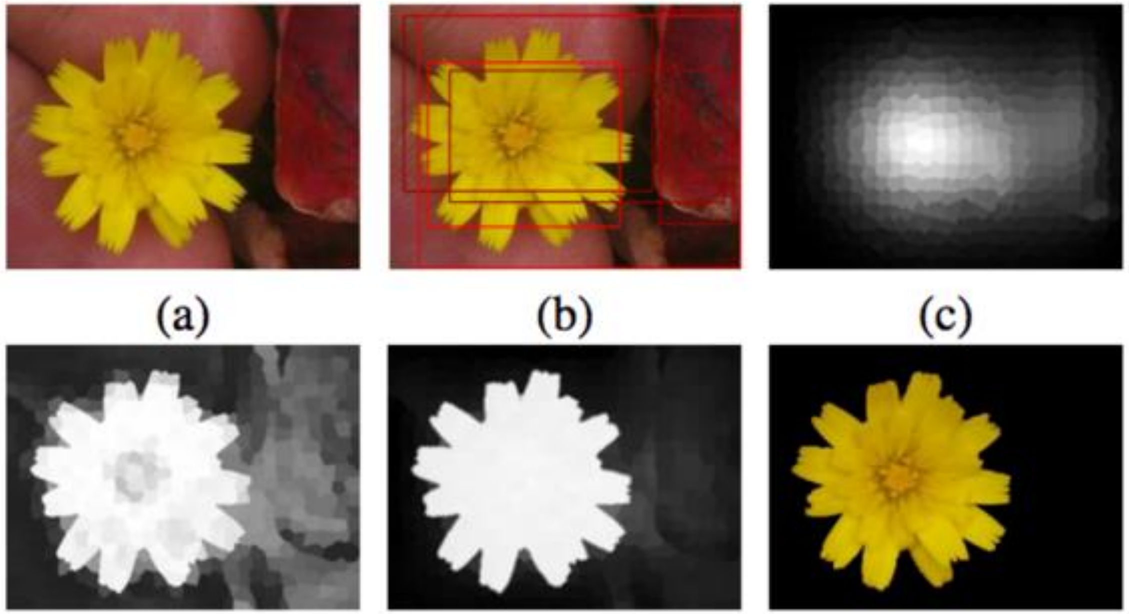
- Backgroundness



Geodesic saliency using background priors, ECCV, 2012, Wei et al.

Hypothesis for Salient object detection

- Objectness



(a) (b) (c)

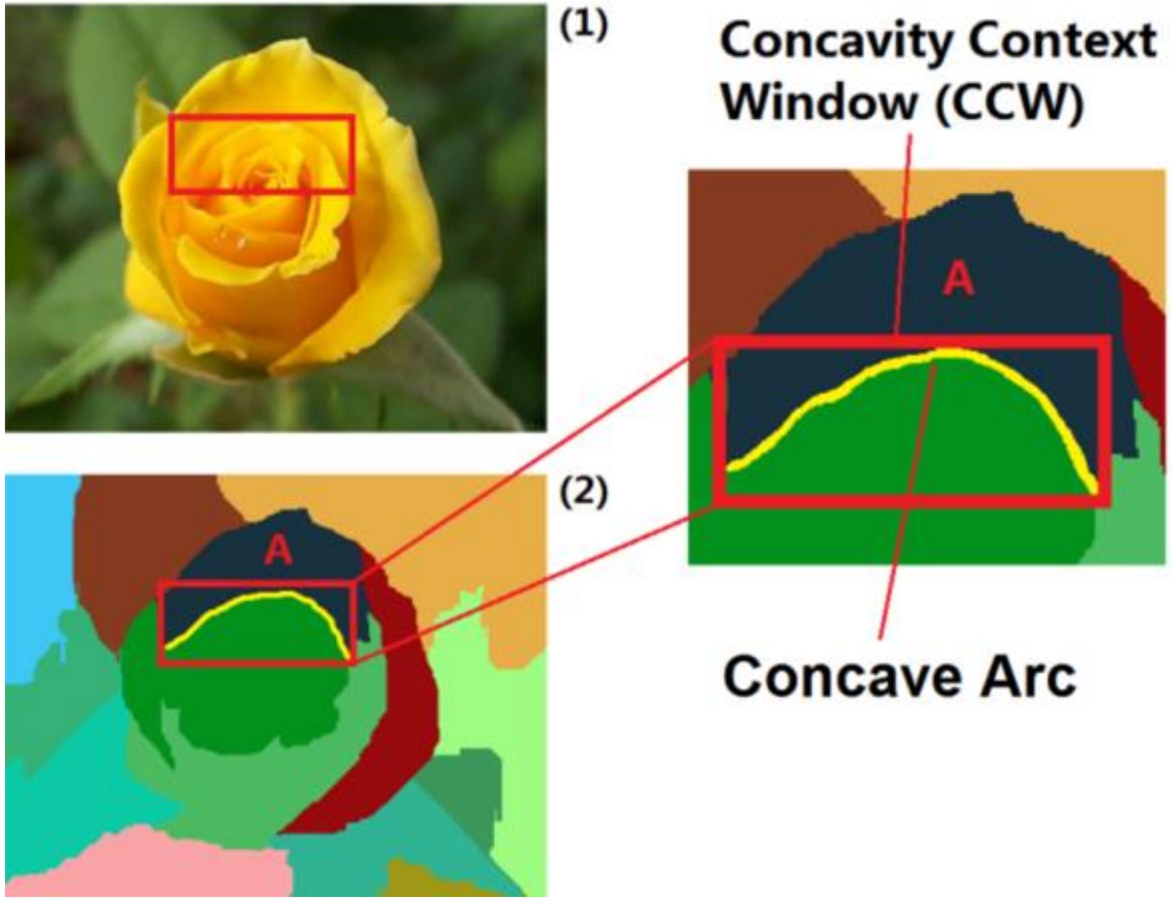
(d) (e) (f)

Figure 1. An illustration of our approach from images to the final saliency map: (a) Input Image (b) objectness detections, (c) saliency prior from objectness, (d) diverse density scores for pixels, (e) the final saliency map, and (f) the segmented object.

Category-independent object-level saliency detection, ICCV 2013, Jia et. al.

Hypothesis for Salient object detection

- Convexity



Salient object detection using concavity context, ICCV 2011, Lu et al.

Block-based vs. region based

- Pixels/patches → regions/super-pixels
  - Efficiency consideration
  - Abstract unnecessary details
  - Region contain complemental cues

Constantly shifting to region based analysis for robustness.

#	Model	Pub	Year	Elements
1	**FG** [10]	MM	2003	PI
2	**RSA** [77]	MM	2005	PA
3	**RE** [11]	ICME	2006	mPI + RE
4	**RU** [85]	TMM	2007	RE
5	**AC** [9]	ICVS	2008	mPA
6	**FT** [81]	CVPR	2009	PI
7	**ICC** [80]	ICCV	2009	PI
8	**EDS** [79]	PR	2009	PI
9	**CSM** [95]	MM	2010	PI + PA
10	**RC** [86]	CVPR	2011	RE
11	**HC** [86]	CVPR	2011	RE
12	**CC** [96]	ICCV	2011	mRE
13	[illegible] [82]	ICCV	2011	mPA
14	[illegible] [73]	ICCV	2011	PA + RE
[illegible]	[illegible] [97]	BMVC	2011	mRE
[illegible]	**SF** [54]	CVPR	2012	RE
[illegible]	[illegible]	[illegible]	2012	RE
[illegible]	[illegible]	[illegible]	2012	PA/RE
[illegible]	[illegible]	TIP	2013	RE
[illegible]	[illegible] [87]	CVPR	2013	hRE
[illegible]	[illegible] [101]	CVPR	2013	RE
[illegible]	[illegible] [94]	CVPR	2013	RE
[illegible]	[illegible]	CVPR	2013	RE
[illegible]	**PCA** [84]	CVPR	2013	PA + PE
[illegible]	**GU** [89]	ICCV	2013	RE
[illegible]	[illegible] [89]	ICCV	2013	RE
[illegible]	[illegible] [83]	ICCV	2013	PA + mRE
28	[illegible]	ICCV	2013	mRE
29	**MC** [103]	ICCV	2013	RE
30	**UFO** [104]	ICCV	2013	RE
31	**CIO** [105]	ICCV	2013	RE
32	**SLMR** [106]	BMVC	2013	RE
33	**LSMD** [107]	AAAI	2013	RE
34	**SUB** [90]	CVPR	2013	RE
35	**PDE** [108]	CVPR	2014	RE
36	**RBD** [109]	CVPR	2014	RE

Salient object detection: a survey, arXiv 2014, Ali et. al.

Intrinsic cues vs. Extrinsic cues

- Intrinsic cues only from input image itself
  - Contrast, spatial distribution, center prior, etc.
- Extrinsic cues of similar images
  - User annotations, depth map, or statistical information of similar images

Heuristics vs. feature learning

- Use supervised learning approach to map the regional feature vector to a saliency score

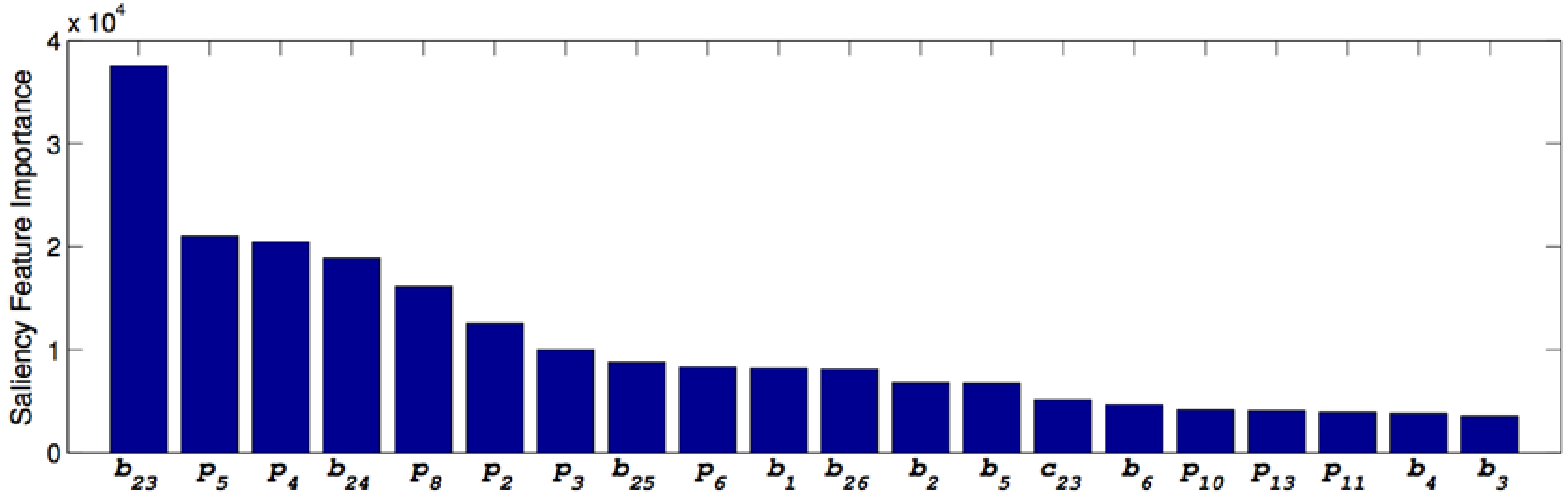


Figure 3. The most important 20 regional features. See Table 1 and Table 2 for the description of the features.

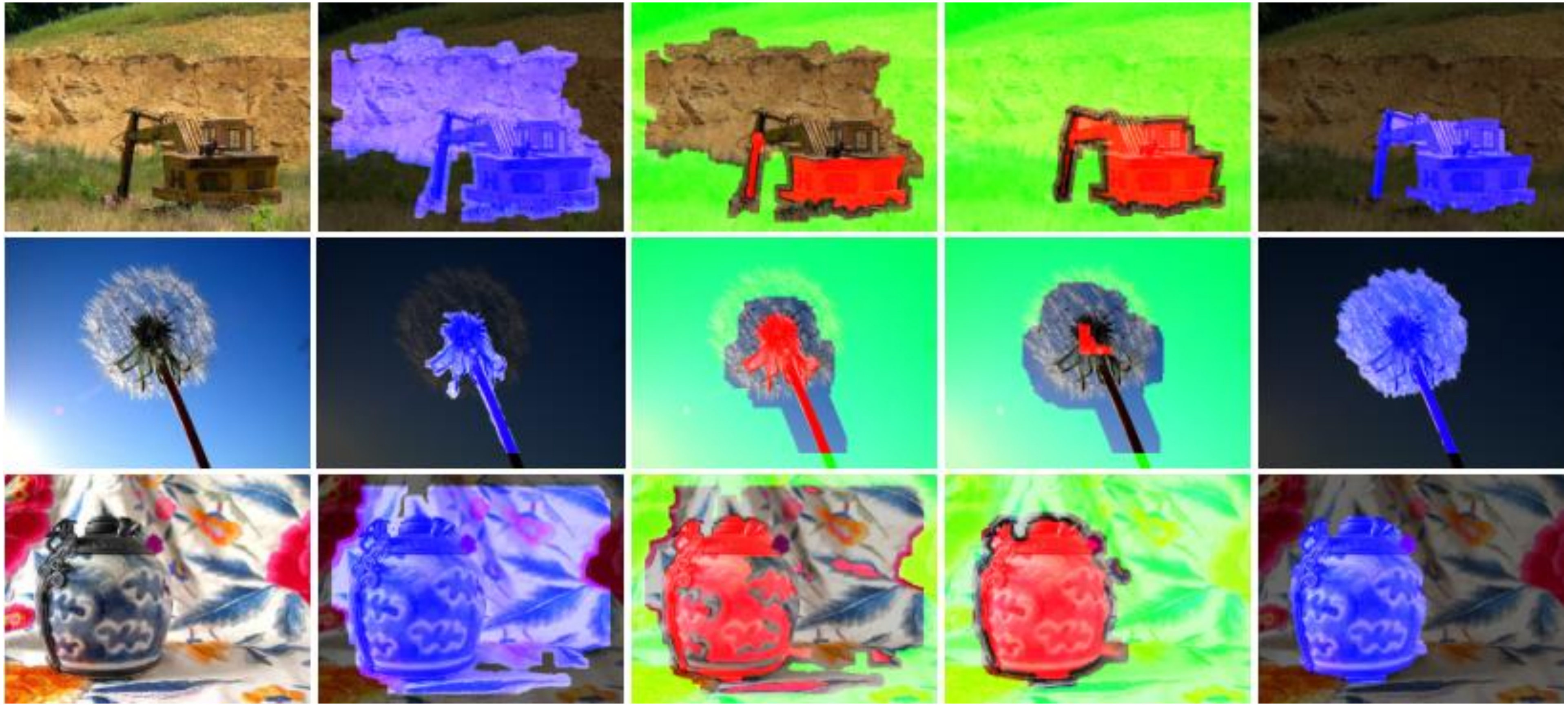
Salient Object Detection: A Discriminative Regional Feature Integration Approach, CVPR 2013, Jiang et al.

Saliency maps aggregation/optimization

- Linear
- Non-Linear
- Adaptive
- Hierarchical
- Bayesian
- Energy minimization
- Least square solver
- Gaussian MRF

How to optimally combine different saliency maps is still an open problem.

SaliencyCut

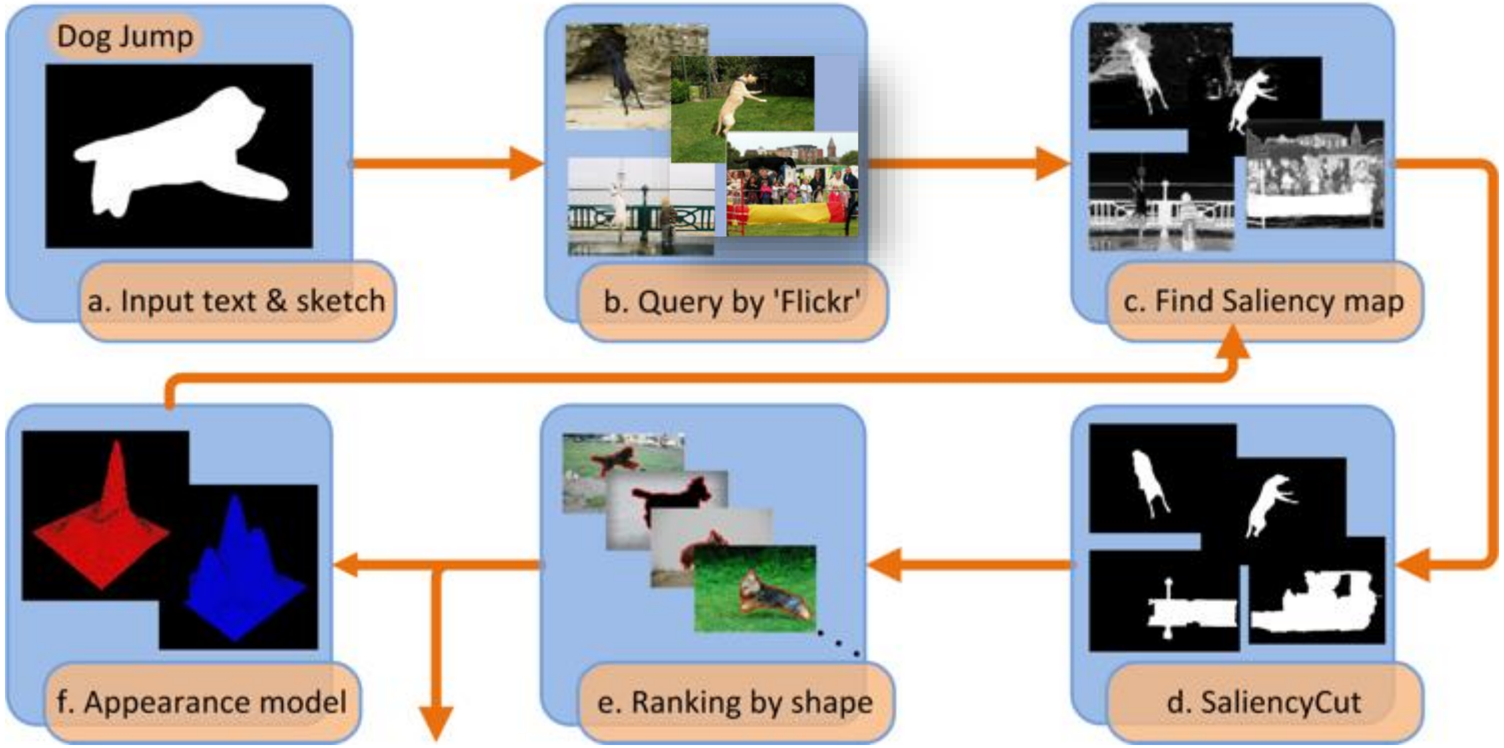


Enables automatic initialization provided by salient object detection.

Global Contrast based Salient Region detection. IEEE TPAMI 2015 (CVPR 2011), Cheng et al.

Applications

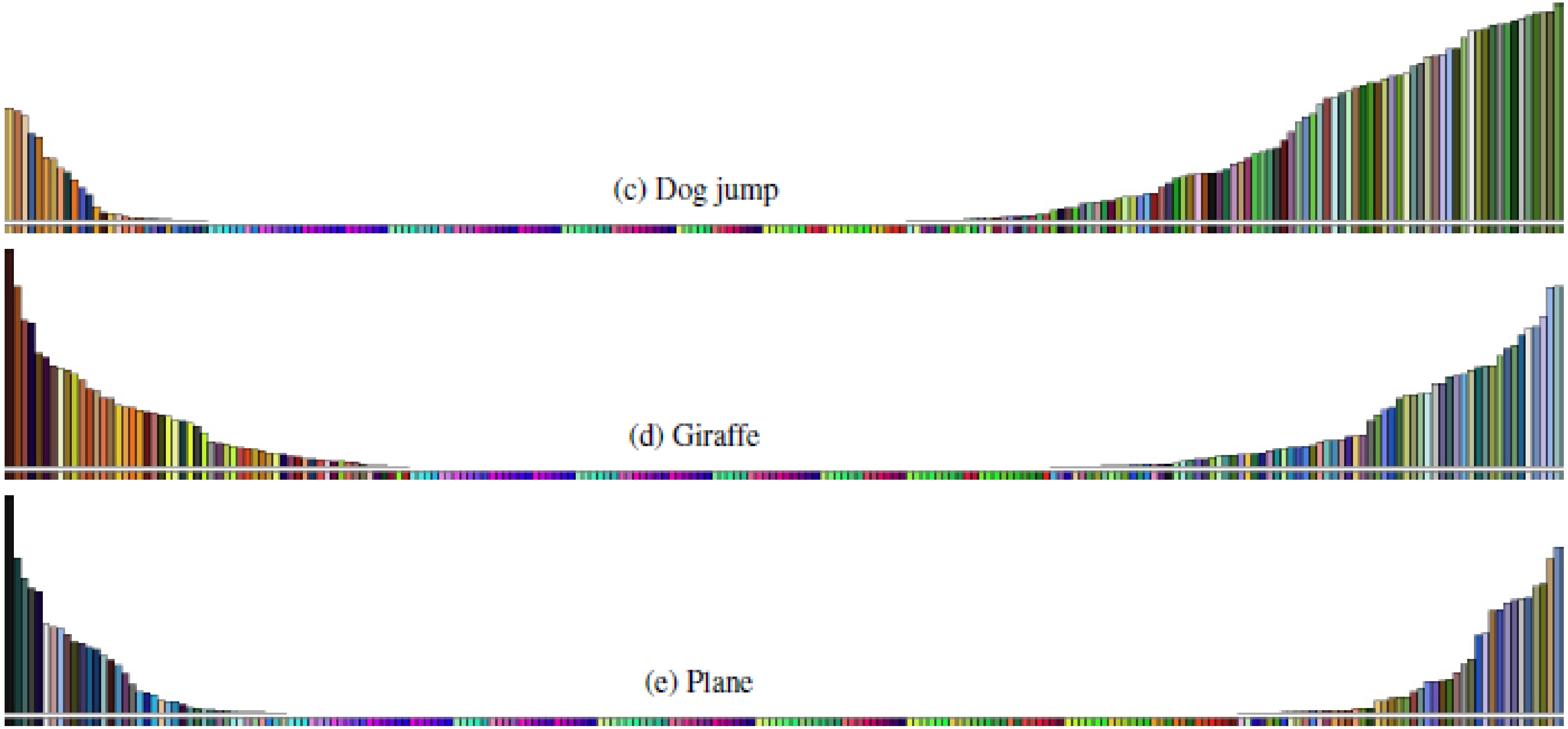
- Is salient object detection for 'simple' images useful?



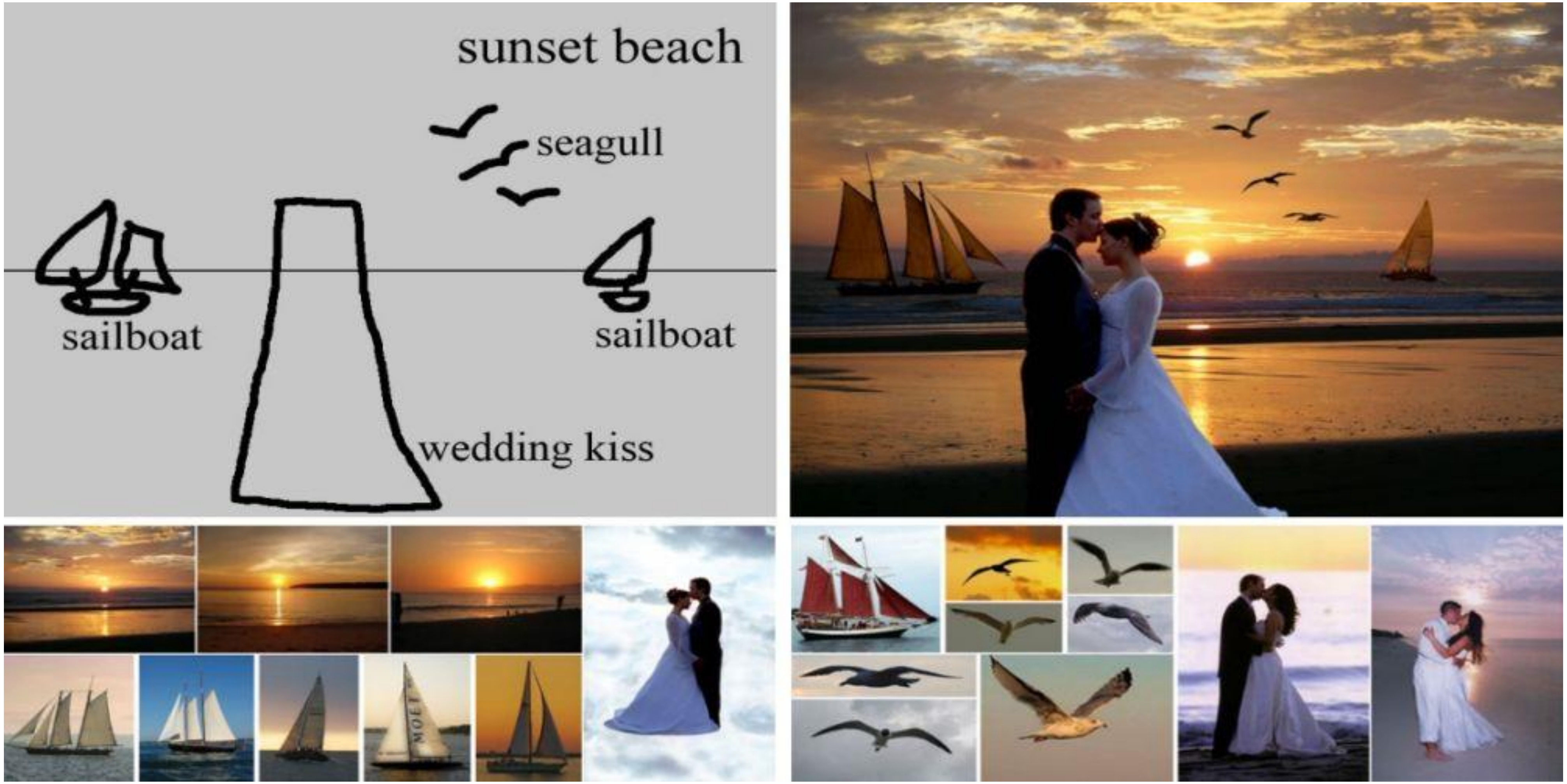
SalientShape: Group Saliency in Image Collections, The Visual Computer 2014. Cheng et. al.

Applications

- Illustration of learned appearance models
  - Accords with our understanding of these categories

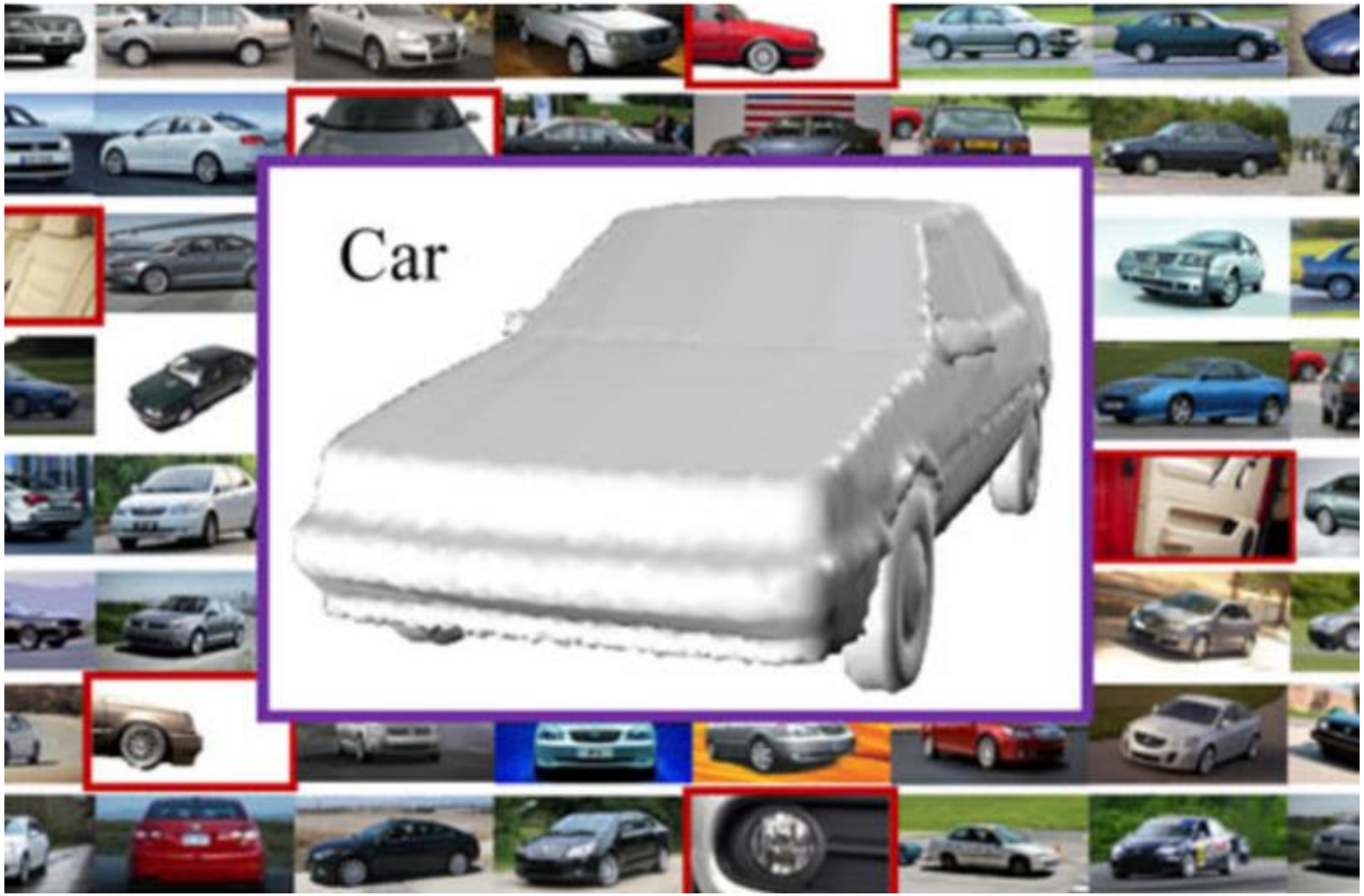


Applications



- Sketch2Photo: internet image montage, ACM TOG 2009. Chen et. al.

Applications



Web-image driven best views of 3d shapes, The Visual Computer 2012. Liu et al.

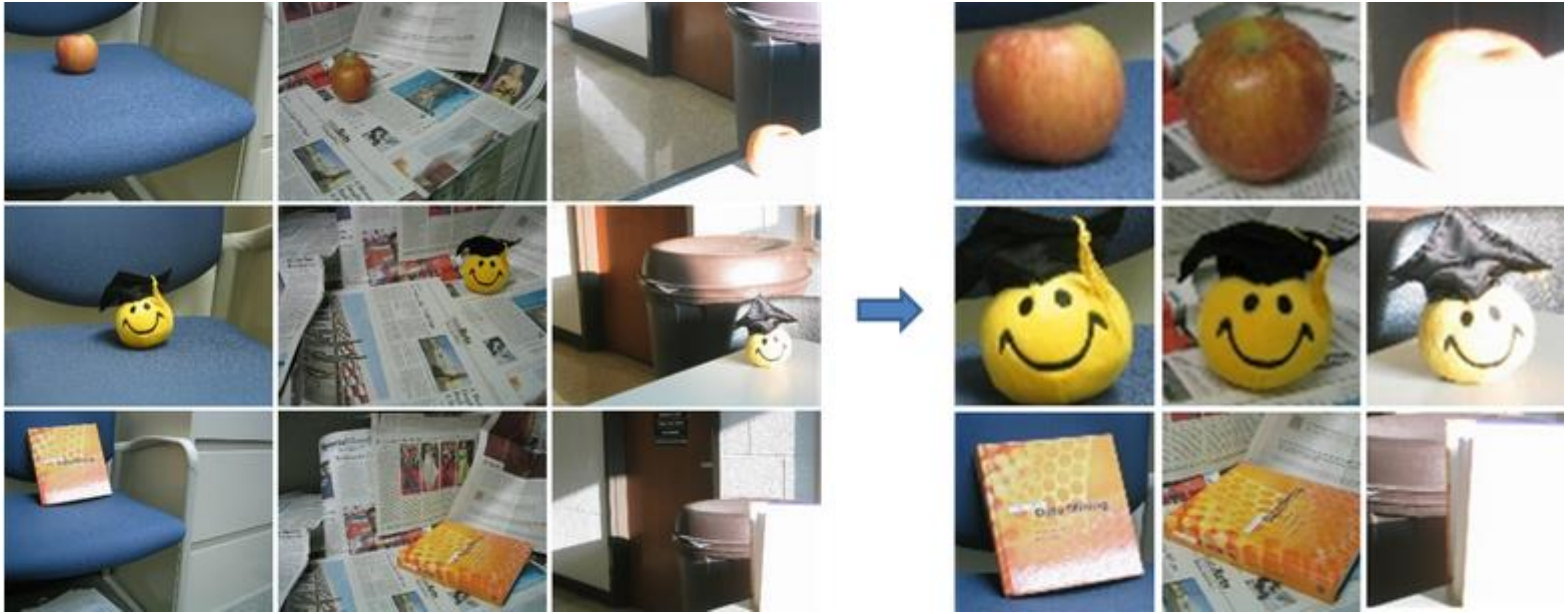
Applications

Saliency for image Manipulation, The Visual Computer 2013, Margolin et al.

Applications

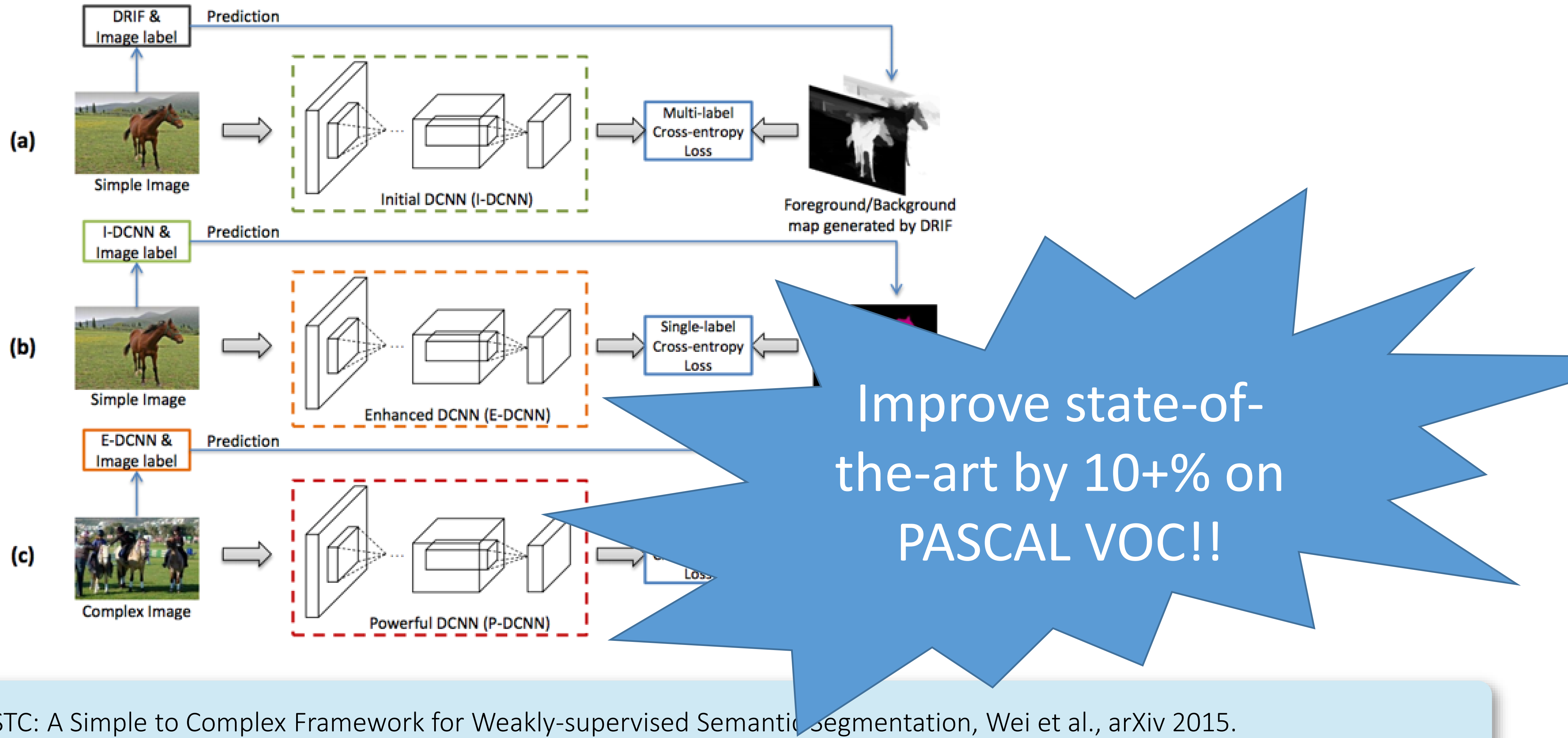
Semantic Colorization with Internet Images, ACM TOG 2011. Chia et al.

Applications



Unsupervised Object Discovery via Saliency-Guided Multiple Class Learning, IEEE CVPR 2012. Zhu et al.

Applications



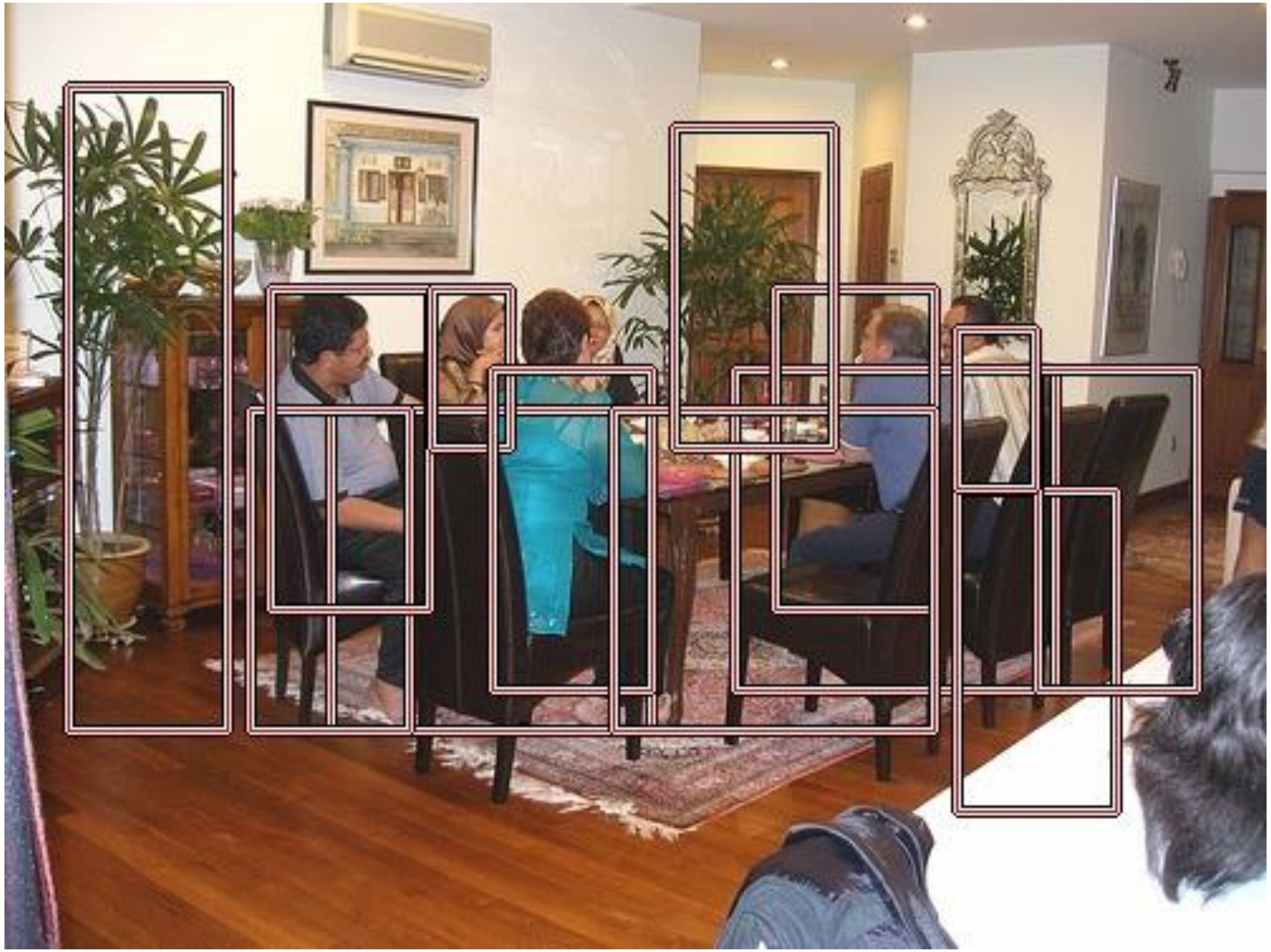
STC: A Simple to Complex Framework for Weakly-supervised Semantic Segmentation, Wei et al., arXiv 2015.

Applications: what we learnt?

Don't ask what segments can do for you, ask what you can do for the segments.

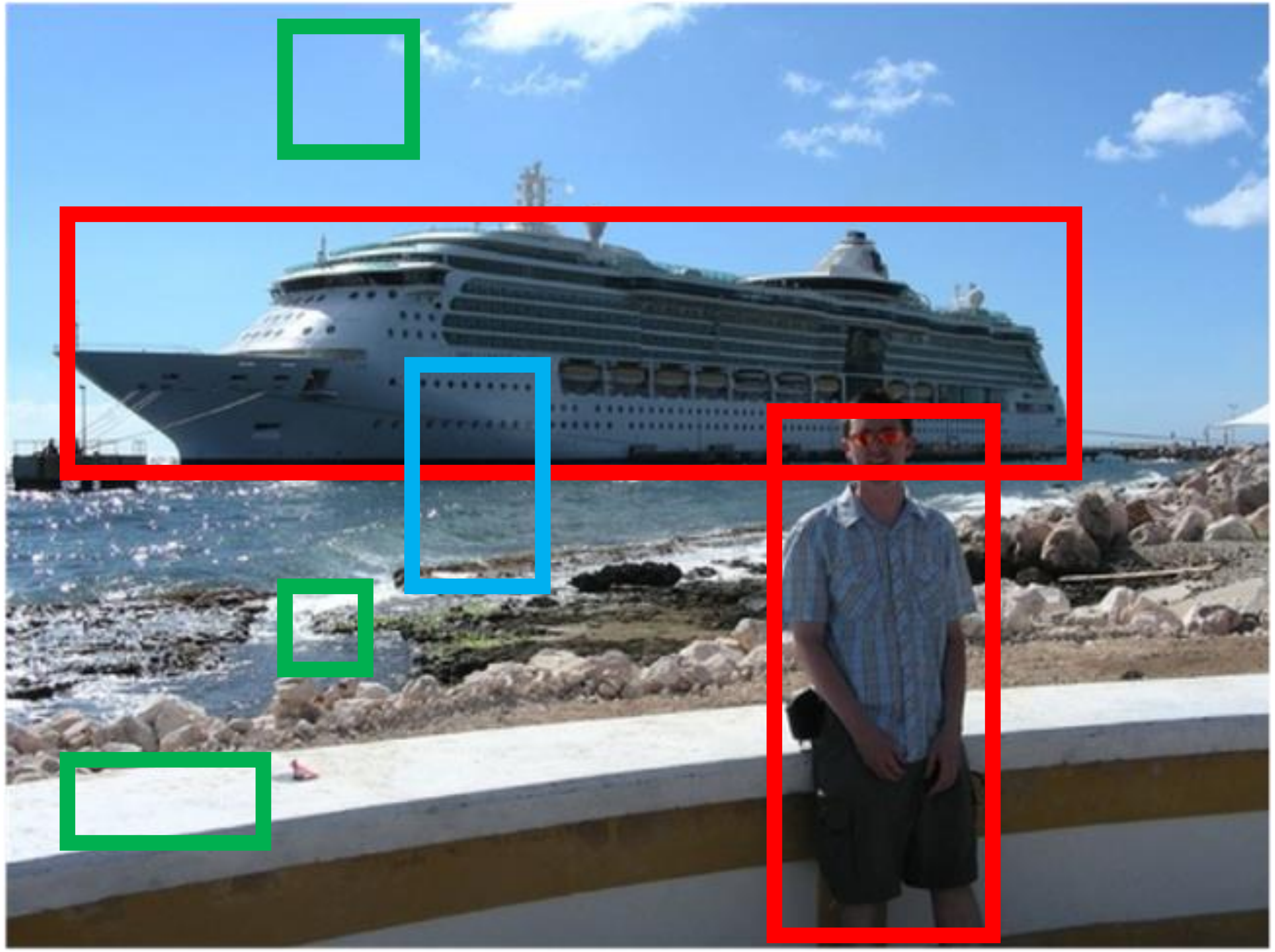
— Jitendra Malik

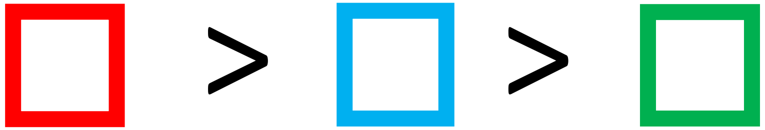
How about complicated images?



Objectness proposals

Motivation: What is an object?





Motivation: What is an object?

- An objectness measure
  - A value to reflects how likely an image window covers an object of any category.
- What's the benefits?
  - Improve computational efficiency, reduce the search space
  - Allowing the usage of strong classifiers during testing, improve accuracy

Measuring the objectness of image window, IEEE TPAMI 2012, Alexe et. al.

Motivation: What is an object?

- What is a good objectness measure?
  - Achieve **high object detection rate (DR)**
    - Any undetected objects at this stage cannot be recovered later
  - Produce a **small number of proposals**
    - Reducing computational time of subsequent detectors
  - Obtain **high computational efficiency**
    - The method can be easily involved in various applications
    - Especially for realtime and large-scale applications;
  - Have **good generalization ability** to unseen object categories
    - The proposals can be reused by many category specific detectors
    - Greatly reduce the computation for each of them.

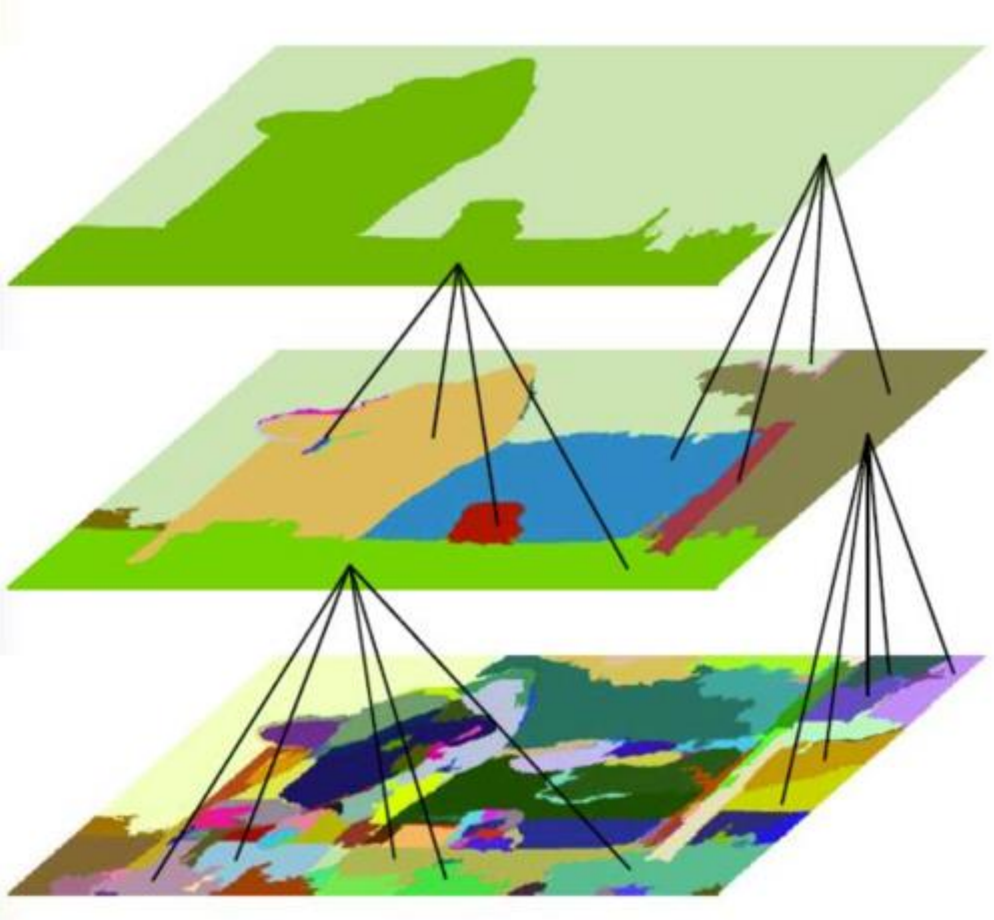
A feature integration approach

- Objectness proposal generation
  - A small number (e.g. 1K) of category-independent proposals
  - Expected to cover all objects in an image

Measuring the objectness of image windows. PAMI 2012, Alexe, et. Al.

Region merging & Diversification

- Region merging
  - Merge two most similar regions based on region similarity.
  - Update similarities between the new region and its neighbors.
- Diversification



[Selective Search for Object Recognition](), IJCV 2013, Uijlings et. Al.

Local & global search

- Local search
  - Unsuitable for object with distinct parts
- Global search
  - Initialize with foreground/background seeds
  - A global optimization function for each parameter set


Generating object segmentation proposals using global and local search, CVPR 2014, Rantalankila et al.

BING method

- Our observation: a small interactive demo
  - Take you pen and paper and draw an object which is current in your mind.
  - What if we resize it to a tiny fixed size?
    - E.g. 8x8. Not only changing the scale, but also aspect ratio.

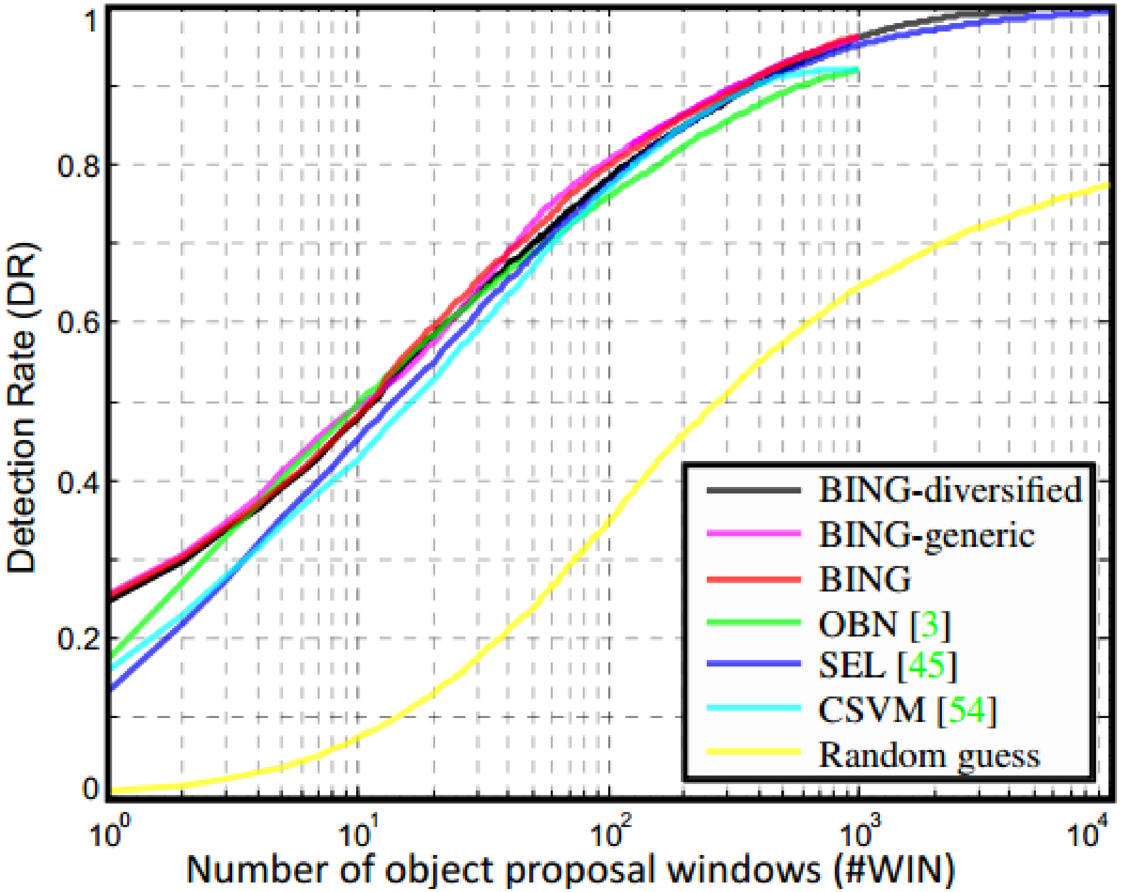
BING: Binarized Normed Gradients for Objectness Estimation at 300fp, IEEE CVPR 2014 (Oral), M.M. Cheng, et. al.

BING method

- Objects are stand-alone things with well defined **closed boundaries** and centers.

  - Finding pictures of objects in large collections of images. Springer Berlin Heidelberg, 1996, Forsyth et. al.
  - Using stuff to find things. ECCV 2008, Heitz et. al.
  - Measuring the objectness of image window, IEEE TPAMI 2012, Alexe et. al.

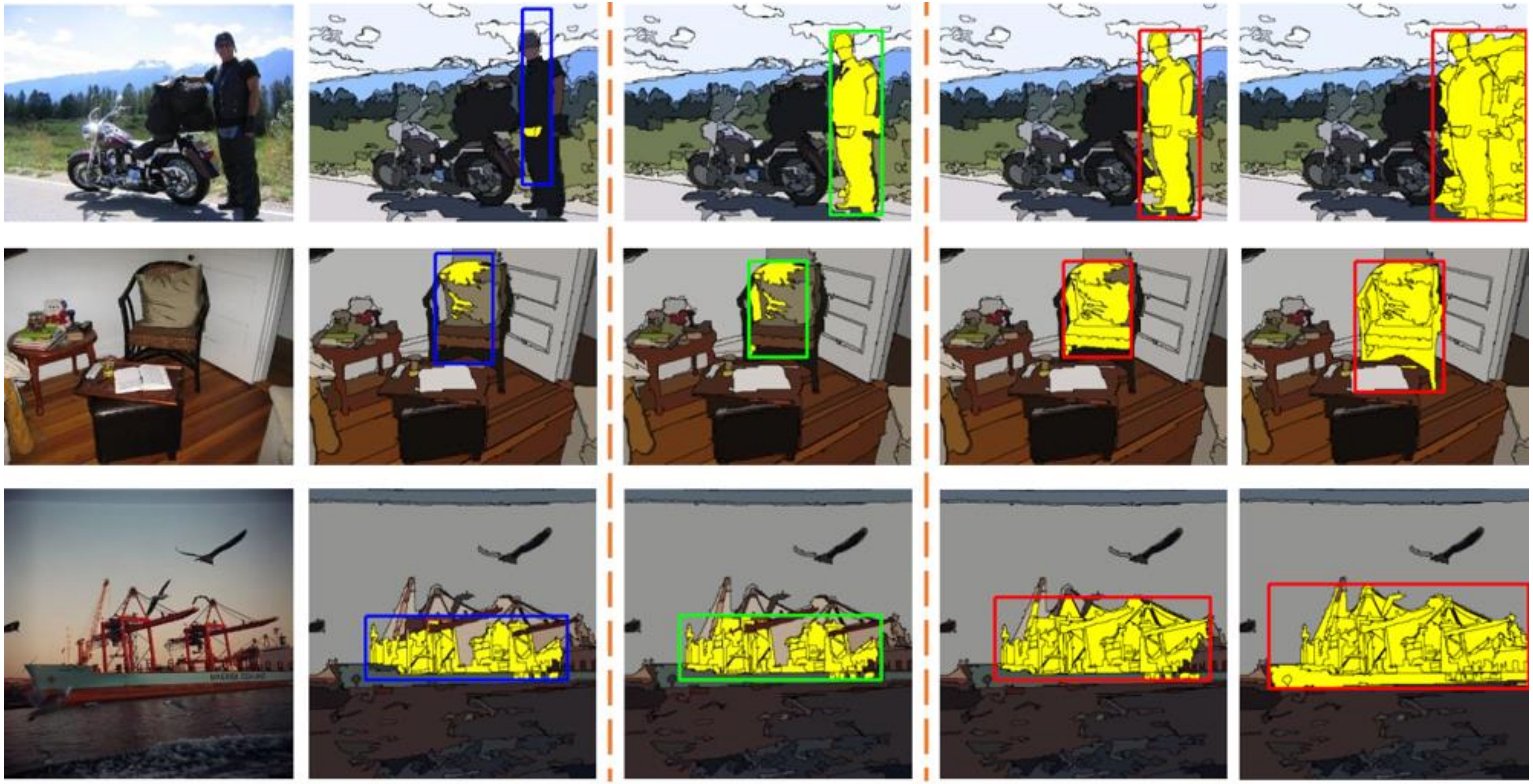
Experimental results of BING method



Method	[22]	OBN [3]	CSVM [57]	SEL [48]	Our BING
Time (seconds)	89.2	3.14	1.32	11.2	**0.003**

Table 1. Average computational time on VOC2007.

Combining boxes and regions

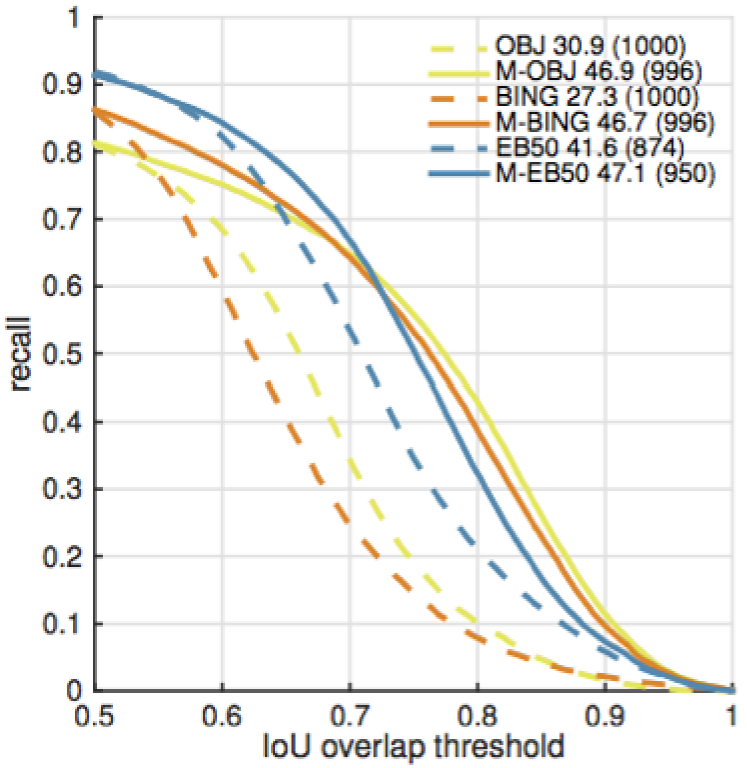


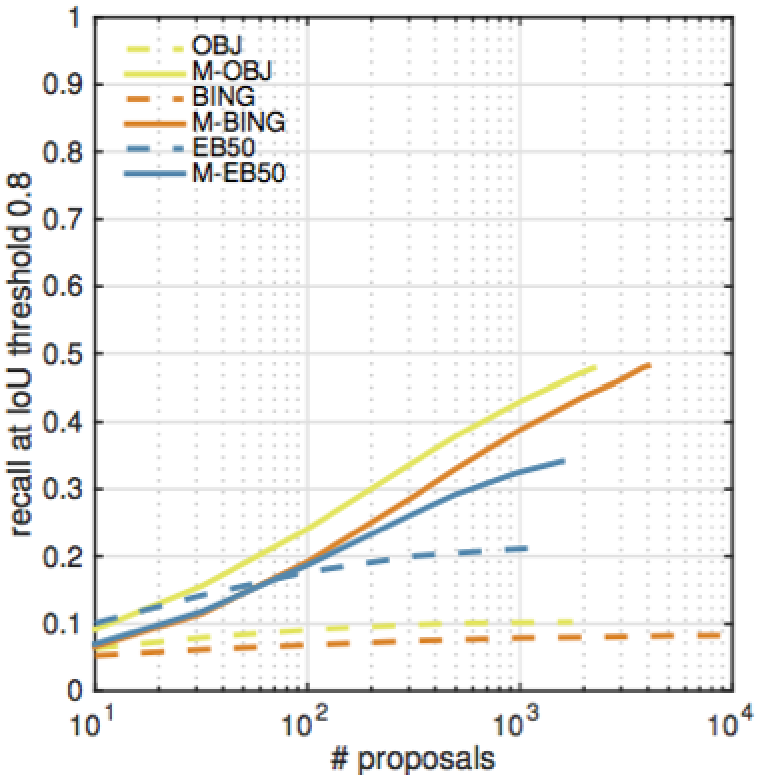
(a) Input (b) Initial boxes (c) Box Alignment (d) $\delta = 0.7$ (e) $\delta = 0.3$

Improving Object Proposals with Multi-Thresholding Straddling Expansion, IEEE CVPR 2015, Chen, et. al.

Combining boxes and regions

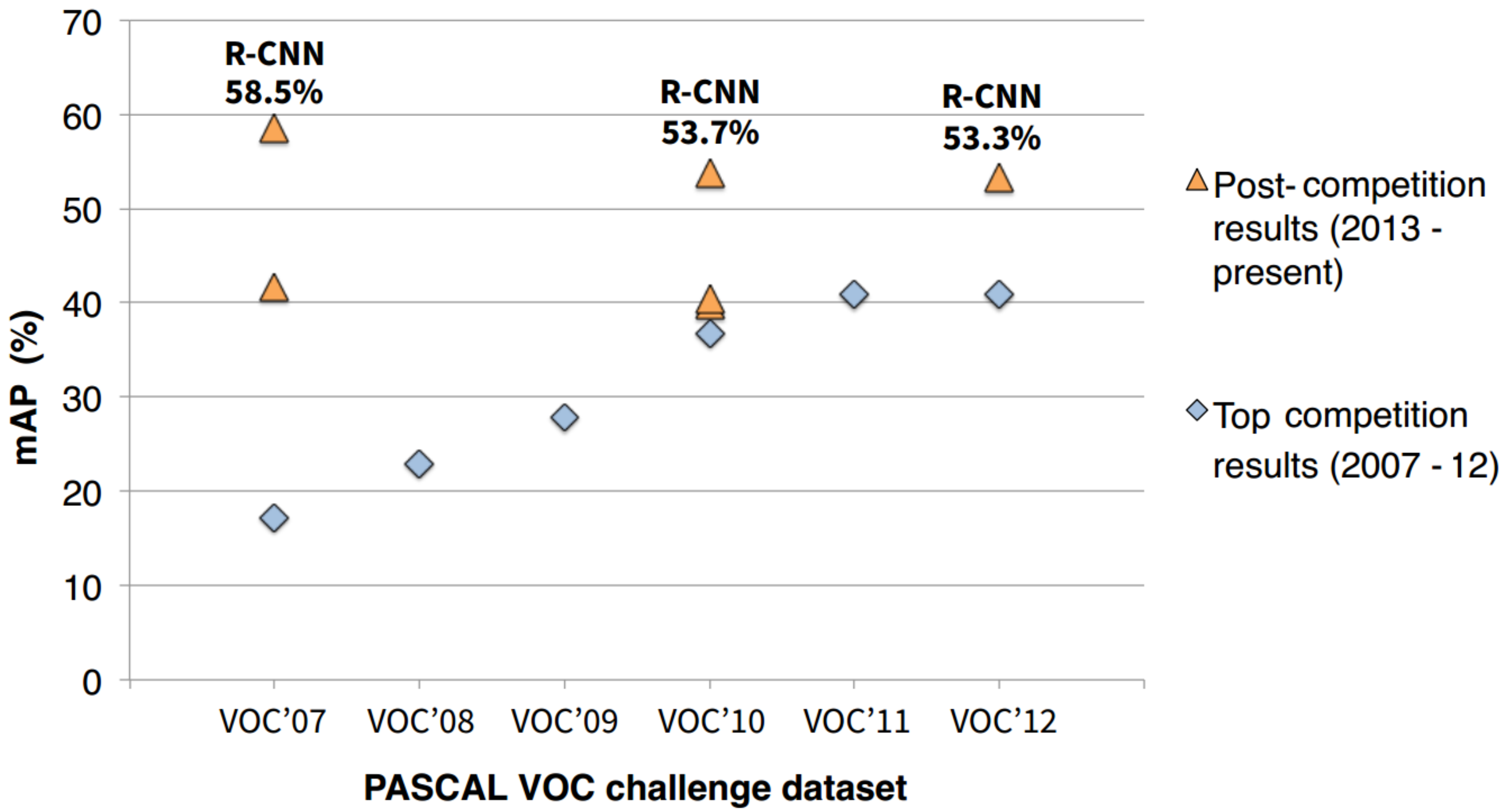
- Experimental results





Improving Object Proposals with Multi-Thresholding Straddling Expansion, IEEE CVPR 2015, Chen, et. al.

Applications



Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation, IEEE CVPR 2014, Girshick et al.

Future work in objectness proposals

- High detection rate under large IoU
- Running speed
- Small number of proposals
- Exploring more applications

Thanks!

Tutorial webpage: http://mmcheng.net/saltutorial/