Salient Object Subitizing

Jianming Zhang[1] Shugao Ma[1] Mehrnoosh Sameki[1] Stan Sclaroff[1] Margrit Betke[1] Zhe Lin[2] Xiaohui Shen[2] Brian Price[2] Radomír Měch[2]

[1]Boston University [2]Adobe Research

Dataset & CNN Models Available at:
http://www.cs.bu.edu/groups/ivc/Subitizing/

What Is Subitizing?

People can immediately and precisely identify that an image contains 1, 2, 3 or 4 items by a simple glance. This phenomenon, known as *Subitizing*, is speculated as an in-born numeric capacity that can help humans and animals make prompt decisions in basic tasks like navigation, searching and choice making.

Problem Definition

Salient Object Subitizing:

To predict the existence and the number of salient objects in a scene using only holistic cues, without resorting to any object localization process.

Potential Applications

It may benefit many applications, if computers can "subitize" significant objects without localizing and counting each of them.

Having a quick estimate of the number of objects can help limit or adapt processing. Moreover, knowing that an image contains no prominent objects can yield a clue not to process that image, or give hints for classifying the image.

A few example application scenarios:

- Salient object detection
- Weakly supervised learning
- Robot vision
- Image indexing and retrieval
- Automated photo-editing

SOS Dataset

Images were taken from several well known datasets. Mechanical Turk was used to collect annotations and further clean up the image set based on the consensus of the annotators. The SOS dataset has 6900 images.

Sample images from our SOS dataset.

Annotation Consistency

Are MTurk workers' annotations consistent with the subitizing phenomenon?

To answer this question, we conducted a controlled offline experiment for salient object subitizing. We found that human subitizing results consistently match MTurk annotations.

Accuracy of MTurk workers' annotations in matching human subitizing results.

	sbj.1	sbj.2	sbj.3	Avg.
Accuracy	90%	92%	90%	91%

Computational SOS

Salient Object Subitzing task is formulated as an image classification problem.

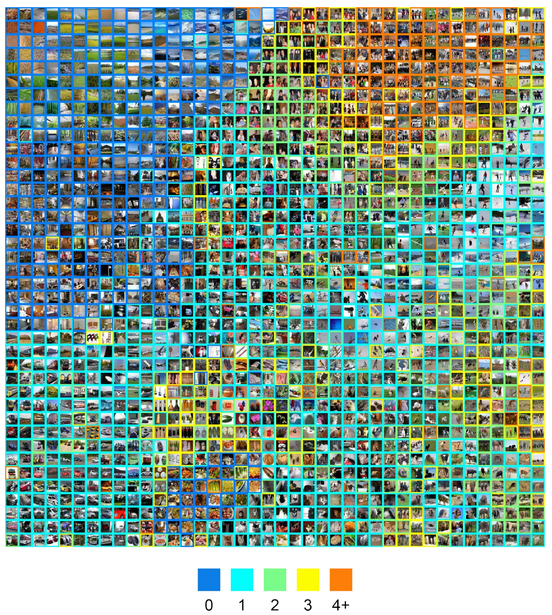
We evaluated several baselines using different image features. The fine-tuned CNN features significantly outperform the other features tested.

Average precision scores of baselines.

	0	1	2	3	4+	mean
Chance	.28	.48	.19	.12	.07	.23
SalCount	-	.55	.21	.16	.11	-
SalPyr	.41	.62	.36	.21	.09	.34
HOG	.65	.62	.32	.29	.14	.40
GIST	.69	.66	.32	.23	.22	.42
IFV	.84	.69	.32	.24	.44	.50
CNN_wo_FT	.92	.82	.34	.31	.56	.59
CNN	.93	.90	.51	.48	.65	.69

Feature Visualization

2D embedding of SOS datasets images using the fc7 layer output of the CNN model.

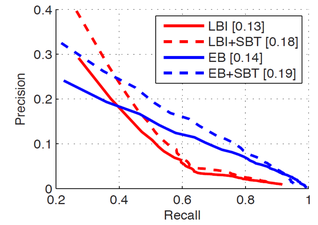


Applications

Application 1: Salient Object Detection

Experiments with new Multi-Salient-Object (MSO) dataset, where about half of the images contain either no object or multiple objects.

Precision-Recall for MSO Dataset

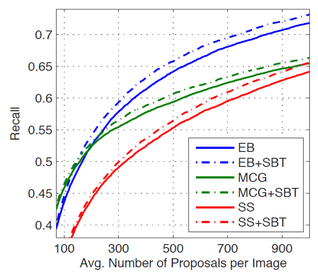


Our CNN subitizing model can help to suppress salient object detection on pure background images, which leads to improvement for state-of-the-art methods.

Application 2: Object Proposal

SOS can be used to retrieve a smaller number of proposals if only a couple of objects exist in an image. This yields consistent improvement for various object proposal methods on the VOC07 dataset.

Recall vs # of Proposals for VOC07 Dataset



References

[EB] Zitnick and Dollar, Edge Boxes: Locating object proposals from edges, ECCV 14.

[LBI] Siva, *et al.*, Looking Beyond the Image: Unsupervised learning for object saliency detection, CVPR 13.

[MCG] Arbelaez, *et al.*, Multiscale combinatorial grouping, CVPR 2014.

[SS] Uijlings, *et al.*, Selective Search for Object Recognition, IJCV 2013.