# Saliency Detection by Multi-Context Deep Learning

Rui Zhao[1,2], Wanli Ouyang[2], Hongsheng Li[2,3], Xiaogang Wang[1,2]
[1]Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences
[2]Department of Electronic Engineering, The Chinese University of Hong Kong
[3]School of Electronic Science, University of Electronic Science and Technology of China

Low-level saliency cues or priors do not produce good enough saliency detection results especially when the salient object presents in a low-contrast background with confusing visual appearance. This issue raises a serious problem for conventional approaches. In this paper, we tackle this problem by proposing a multi-context deep learning framework for salient object detection. We employ deep Convolutional Neural Networks to model saliency of objects in images. Global context and local context are both taken into account, and are jointly modeled in a unified multi-context deep learning framework. As shown in Figure 1, the upper branch (global-context modeling) of our saliency detection pipeline is a deep CNN architecture with global and coarse context. The local-context model takes an input with a similar form as in the global-contex model, but with one third of the scope of context. The local-context model shares the same deep structure with the global-context model, but with independent parameters. Parameters in our multi-context model are jointly optimized by backpropagating the classification loss.

To provide a better initialization for training the deep neural networks, we investigate different pre-training strategies, and a new task-specific scheme with pre-training strategies based on superpixel-level and object-level annotation are proposed for the global- and local-context modeling. Superpixel-level annotation aligns spatial location of objects when pre-training the global-context model, and it is consistent with the input format of the global-context model. Features pre-trained with object-level annotation are sensitive to the location of objects, and they provide more appropriate pre-training information for the local-context model. Furthermore, recently proposed contemporary deep models in the ImageNet Image Classification Challenge are tested, and their effectiveness in saliency detection are investigated. Our approach is extensively evaluated on five public datasets, and experimental results show significant and consistent improvements over the state-of-the-art methods.

In Table 1, we compare our approach with nine latest state-of-the-art methods, including IS [5], GBVS [4], SF [10], GC [3], CEOS [8], PCAS [9], GBMR [12], HS [11], and DRFI [6]. Our approach significantly outperforms all the state-of-the-art salient object segmentation algorithms. Also, we qualitatively compare our saliency maps with those by other methods in Figure 2. It is obvious that our approach is able to highlight the salient object parts more coherently, and has a better prediction especially in complex scene with confusing background, such as the cases in the $6^{th}$ and $7^{th}$ rows in Figure 2. More comparisons can be found at our project website http://www.ee.cuhk.edu.hk/~rzhao/.

| | ASD | SED1 | SED2 | ECSSD | PASCAL-S |
|---|---|---|---|---|---|
| IS [5] | 0.5943 | 0.5540 | 0.5682 | 0.4731 | 0.4901 |
| GBVS [4] | 0.6499 | 0.7125 | 0.5862 | 0.5528 | 0.5929 |
| SF [10] | 0.8879 | 0.7533 | 0.7961 | 0.5448 | 0.5740 |
| GC [3] | 0.8811 | 0.8066 | 0.7728 | 0.5821 | 0.6184 |
| CEOS [8] | 0.9020 | 0.7935 | 0.6198 | 0.6465 | 0.6557 |
| PCAS [9] | 0.8613 | 0.7586 | 0.7791 | 0.5800 | 0.6332 |
| GBMR [12] | 0.9100 | 0.9062 | 0.7974 | 0.6570 | 0.7055 |
| HS [11] | 0.9307 | 0.8744 | 0.8150 | 0.6391 | 0.6819 |
| DRFI [6] | 0.9448 | 0.9018 | 0.8725 | 0.6909 | 0.7447 |
| **Ours** | **0.9548** | **0.9295** | **0.8903** | **0.7322** | **0.7930** |

Table 1: The F-measure scores of benchmarking approaches on five public datasets, including ASD [1], SED1 & SED2 [2], ECSSD [11], and PASCAL-S [7].

This is an extended abstract. The full paper is available at the Computer Vision Foundation webpage.

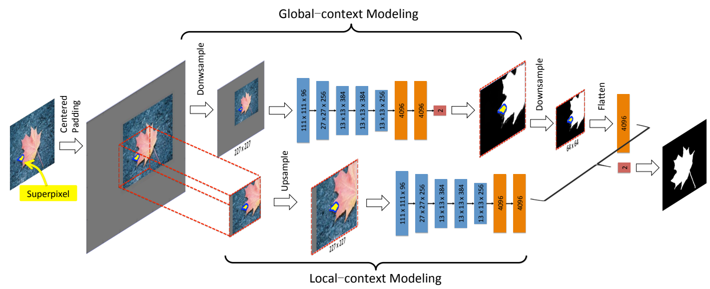


Figure 1: Upper branch: Deep CNN-based global-context modeling for saliency detection with a superpixel-centered window padded with mean pixel value. Lower branch: local-context modeling with a closer-focused superpixel-centered window, and global-context saliency detection results are combined into finally fully-connected layer in the local-context model.

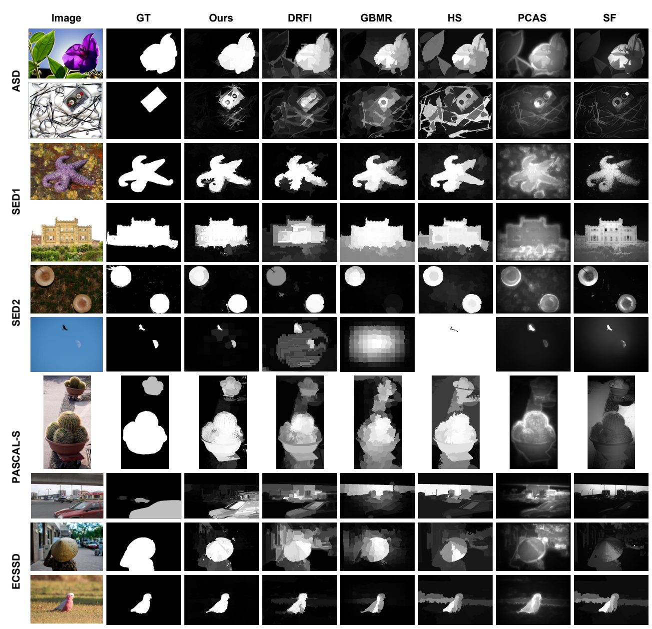


Figure 2: Example images from five datasets and the saliency maps of compared methods.

[1] Radhakrishna Achanta, Sheila Hemami, Francisco Estrada, and Sabine Susstrunk. Frequency-tuned salient region detection. In *CVPR*, 2009.

[2] Sharon Alpert, Meirav Galun, Ronen Basri, and Achi Brandt. Image segmentation by probabilistic bottom-up aggregation and cue integration. In *CVPR*, 2007.

[3] Ming-Ming Cheng, Jonathan Warrell, Wen-Yan Lin, Shuai Zheng, Vibhav Vineet, and Nigel Crook. Efficient salient region detection with soft image abstraction. In *ICCV*, 2013.

[4] Jonathan Harel, Christof Koch, and Pietro Perona. Graph-based visual saliency. In *NIPS*, 2006.

[5] Xiaodi Hou, Jonathan Harel, and Christof Koch. Image signature: Highlighting sparse salient regions. *IEEE Trans. on PAMI*, 34(1):194–201, 2012.

[6] Huaizu Jiang, Jingdong Wang, Zejian Yuan, Yang Wu, Nanning Zheng, and Shipeng Li. Salient object detection: A discriminative regional feature integration approach. In *CVPR*, 2013.

[7] Yin Li, Xiaodi Hou, Christof Koch, J Rehg, and A Yuille. The secrets of salient object segmentation. In *CVPR*, 2014.

[8] Rotem Mairon and Ohad Ben-Shahar. A closer look at context: From coxels to the contextual emergence of object saliency. In *ECCV*. 2014.

[9] Ran Margolin, Ayellet Tal, and Lihi Zelnik-Manor. What makes a patch distinct? In *CVPR*, 2013.

[10] Federico Perazzi, Philipp Krahenbuhl, Yael Pritch, and Alexander Hornung. Saliency filters: Contrast based filtering for salient region detection. In *CVPR*, 2012.

[11] Qiong Yan, Li Xu, Jianping Shi, and Jiaya Jia. Hierarchical saliency detection. In *CVPR*, 2013.

[12] Chuan Yang, Lihe Zhang, Huchuan Lu, Xiang Ruan, and Ming-Hsuan Yang. Saliency detection via graph-based manifold ranking. In *CVPR*, 2013.