Deep Contrast Learning for Salient Object Detection (Supplemental Materials)

Guanbin Li Yizhou Yu
Department of Computer Science, The University of Hong Kong
{gbli, yzyu}@cs.hku.hk

1. Quantitative Comparisons on PASCAL-S and SOD Datasets

PASCAL-S is a dataset built using the validation set of the PASCAL VOC 2010 segmentation challenge. It contains 850 images with the ground truth labeled by 12 subjects. We threshold the masks at 0.5 to obtain binary masks as suggested in [4]. The SOD dataset contains 300 images and it was originally designed for image segmentation. Many images in this dataset have multiple salient objects with low contrast.

We compare our saliency models (DCL and DCL$^+$) against eight recent state-of-the-art methods, including SF [6], GC [1], DRFI [2], PISA [8], BSCA [7], LEGS [9], MC [10] and MDF [3]. The last three are the latest deep learning based methods. For fair comparison, we use either the implementations or the saliency maps provided by the authors. In addition, we also train a fully convolutional neural network (FCN) (the FCN-8s network proposed in [5]) for comparison.

As shown in Fig. 1, our method achieves the highest PR curve on both PASCAL-S and SOD datasets. The average precision, recall and F-measure on these two datasets is shown in Fig. 2, sorted by F-measure. Our method achieves the best performance on the overall F-measure as well as significant improvements in both precision and recall. On the PASCAL-S dataset, our DCL with CRF (DCL$^+$) achieves 78.55% precision and 77.07% recall while the second best (MDF [3]) achieves 81.14% precision and 61.75% recall. On the SOD dataset, our method achieves 85.67% precision and 64.75% recall while the second best (MDF[3]) achieves 85.72% precision and only 58.32% recall. Although the precision of MDF is higher than ours, its recall is significantly lower. Thus it is much more likely for MDF to miss salient objects. This is also reflected in the lower F-measure and higher MAE achieved with MDF.

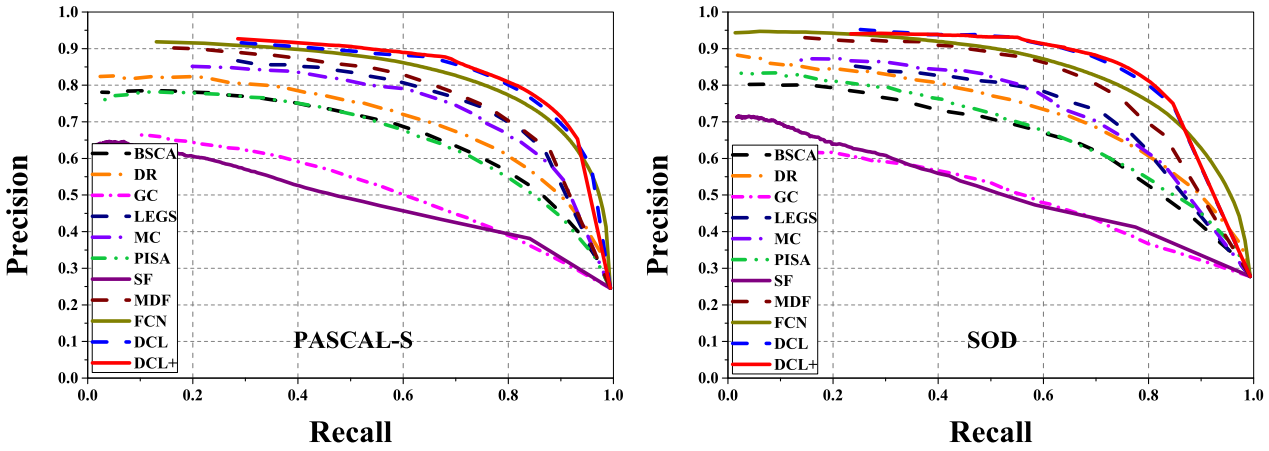

Figure 1: Comparison of precision-recall curves of 11 saliency detection models on PASCAL-S and SOD. Our DCL and DCL$^+$ (DCL with CRF) consistently outperform other methods on both datasets.

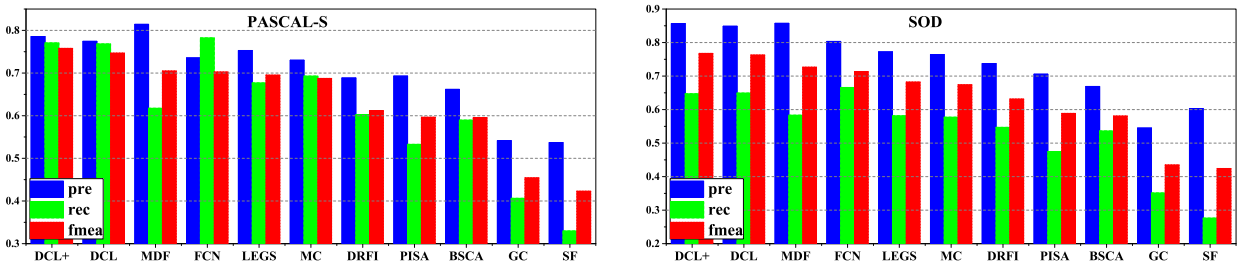

Figure 2: Comparison of precision, recall and F-measure (computed using a per-image adaptive threshold) among 11 different models on PASCAL-S and SOD.

2. Visual Comparisons

Fig. 3 and Fig. 4 show more visual comparisons of saliency maps generated from 10 different models, including ours (DCL and DCL$^+$). The ground truth (GT) is shown in the last column. DCL and DCL$^+$ consistently produce saliency maps closest to the ground truth. We compare DCL and DCL$^+$ against SF [6], GC [1], DRFI [2], PISA [8], BSCA [7], LEGS [9], MC [10] and MDF [3].

References

[1] M. Cheng, N. J. Mitra, X. Huang, P. H. Torr, and S. Hu. Global contrast based salient region detection. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 37(3):569–582, 2015. 1, 2

[2] P. Jiang, H. Ling, J. Yu, and J. Peng. Salient region detection by ufo: Uniqueness, focusness and objectness. In *Computer Vision (ICCV), 2013 IEEE International Conference on*, pages 1976–1983. IEEE, 2013. 1, 2

[3] G. Li and Y. Yu. Visual saliency based on multiscale deep features. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2015. 1, 2

[4] Y. Li, X. Hou, C. Koch, J. M. Rehg, and A. L. Yuille. The secrets of salient object segmentation. In *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on*, pages 280–287. IEEE, 2014. 1

[5] J. Long, E. Shelhamer, and T. Darrell. Fully convolutional networks for semantic segmentation. *arXiv preprint arXiv:1411.4038*, 2014. 1

[6] F. Perazzi, P. Krähenbühl, Y. Pritch, and A. Hornung. Saliency filters: Contrast based filtering for salient region detection. In *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, pages 733–740. IEEE, 2012. 1, 2

[7] Y. Qin, H. Lu, Y. Xu, and H. Wang. Saliency detection via cellular automata. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 110–119, 2015. 1, 2

[8] K. Wang, L. Lin, J. Lu, C. Li, and K. Shi. Pisa: Pixelwise image saliency by aggregating complementary appearance contrast measures with edge-preserving coherence. *Image Processing, IEEE Transactions on*, 24(10):3019–3033, Oct 2015. 1, 2

[9] L. Wang, H. Lu, X. Ruan, and M.-H. Yang. Deep networks for saliency detection via local estimation and global search. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3183–3192, 2015. 1, 2

[10] R. Zhao, W. Ouyang, H. Li, and X. Wang. Saliency detection by multi-context deep learning. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1265–1274, 2015. 1, 2

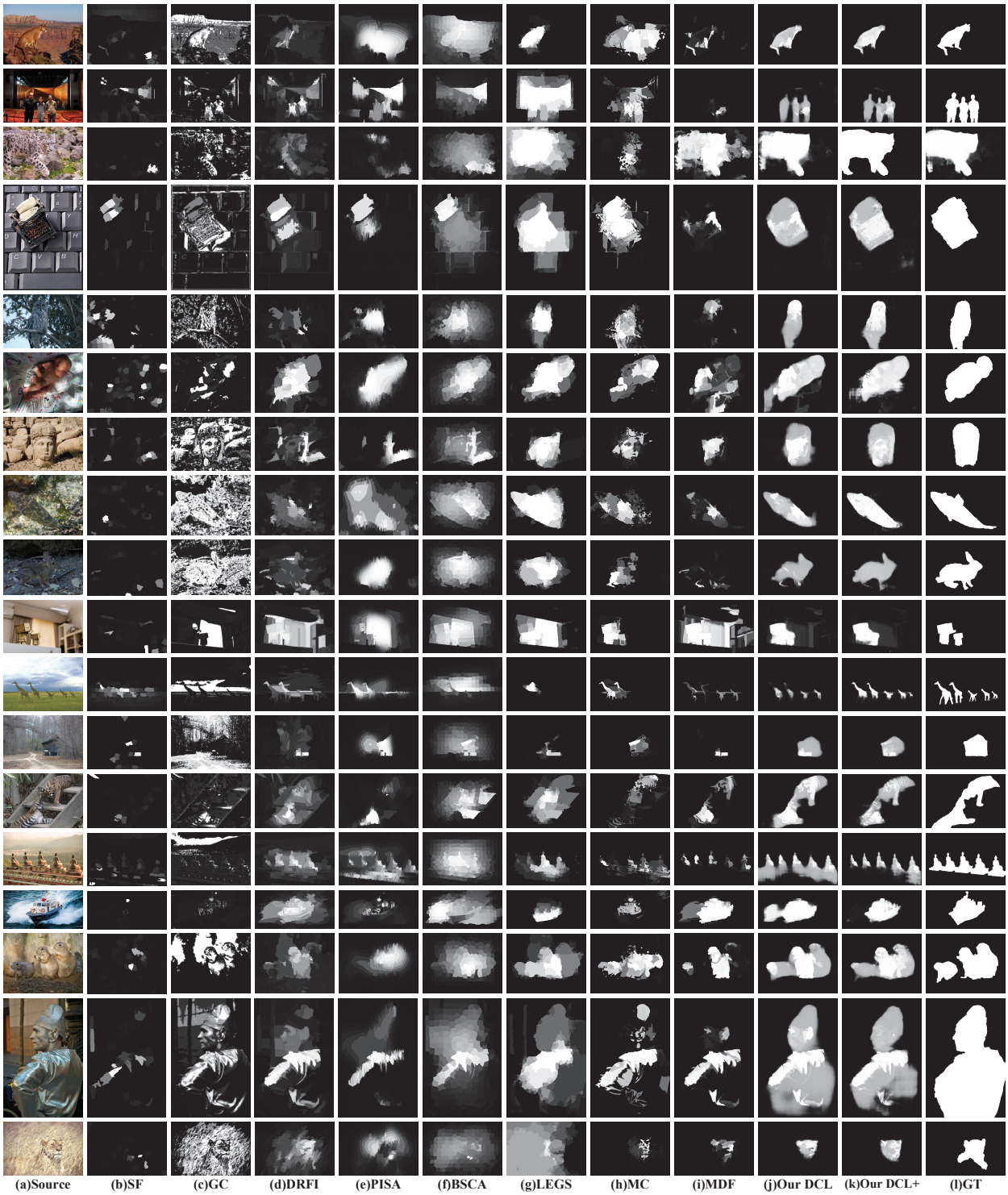

Figure 3: Visual comparison of saliency maps generated from state-of-the-art models, including our DCL and DCL$^+$. The ground truth (GT) is shown in the last column. DCL and DCL$^+$ consistently produce saliency maps closest to the ground truth.

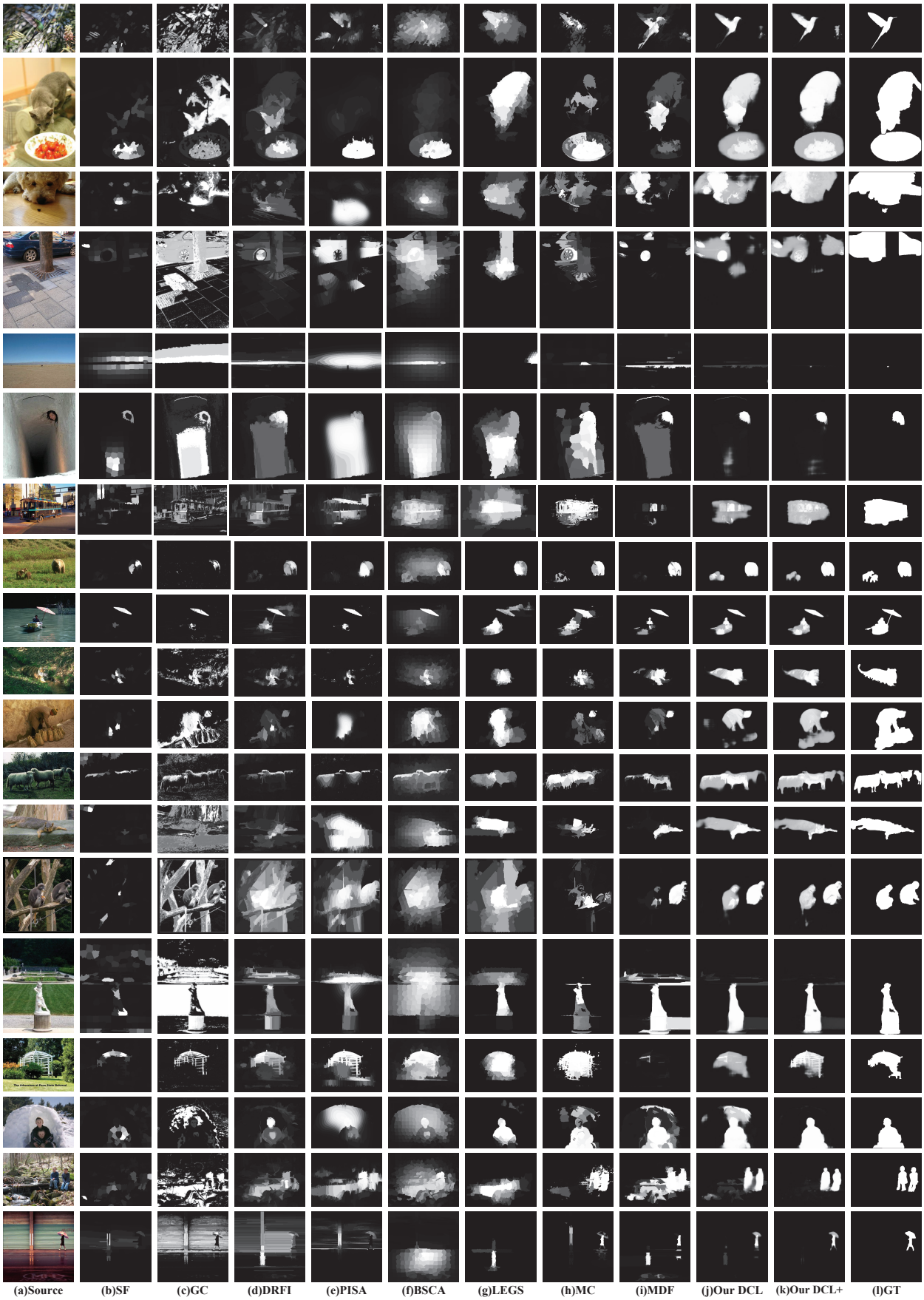

Figure 4: Visual comparison of saliency maps generated from state-of-the-art models, including our DCL and DCL$^+$. The ground truth (GT) is shown in the last column. DCL and DCL$^+$ consistently produce saliency maps closest to the ground truth.