Bi-directional Object-Context Prioritization Learning for Saliency Ranking — Supplementary Material

Xin Tian[1,2,3] Ke Xu[2,†] Xin Yang[1,†] Lin Du[3] Baocai Yin[1] Rynson W.H. Lau[2,‡]
[1]Dalian University of Technology [2]City University of Hong Kong
[3]AI Application Research Center (AARC), Huawei Technologies Co., Ltd.
{xin.tian.831, kkangwing}@gmail.com, {xinyang, ybc}@dlut.edu.cn,
dulin09@huawei.com, Rynson.Lau@cityu.edu.hk

In this supplemental, we first provide more quantitative experiments to analyze our model design. Specifically, we investigate how the parameter settings of R in the SOS module and P in the OCOR module affect the performance in Table 1 and Table 2, respectively. According to Table 1 and Table 2, we set R to 4 and P to 8 in our implementation. We also evaluate the backbone choices for our model in Table 3. Note that even if we use the ResNet-50 backbone to initialize our model, it can still achieve the best performance against other state-of-the-art methods on the SA-SOR and SOR metrics (refer to Table 1 in the main submission for comparison). We choose the Swin-L backbone as it provides the best performance.

We then provide more qualitative comparisons of our method *vs*. existing saliency ranking methods (RSDNet [4], ASSR [8], and IRSR [7]) and the adapted baseline methods (BlendMask [2], CenterMask [5], SOLOv2 [9], Cascade R-CNN [1], CBNetv2 [6], and QueryInst [3]) in Figure 1, 2, 3, 4, 5, and 6, respectively. These images cover diverse daily scenarios and the comparisons generally verify the superiority of our method.

Table 1. Evaluation on different parameter settings for R in SOS module.

Settings of S	SA-SOR ↑	SOR↑	MAE↓
$R = 16$	0.724	0.895	0.084
$R = 8$	0.729	0.890	0.082
$R = 2$	0.735	0.899	0.081
$R = 4$ (Ours)	**0.738**	**0.904**	**0.078**

Table 2. Evaluation on different parameter settings for P in OCOR module.

Settings of P	SA-SOR ↑	SOR↑	MAE↓
$P = 1$	0.727	0.888	0.081
$P = 2$	0.731	0.892	0.081
$P = 4$	0.736	0.900	0.080
$P = 16$	0.737	**0.905**	0.080
$P = 8$ (Ours)	**0.738**	0.904	**0.078**

Table 3. Evaluation on different backbone choices for our model.

Backbone	SA-SOR ↑	SOR↑	MAE↓
ResNet-50	0.723	0.877	0.085
ResNet-101	0.727	0.885	0.084
Swin-T	0.726	0.890	0.084
Swin-B	0.730	0.896	0.080
Swin-L (Ours)	**0.738**	**0.904**	**0.078**

References

[1] Zhaowei Cai and Nuno Vasconcelos. Cascade r-cnn: Delving into high quality object detection. In *CVPR*, 2018. 1, 2, 3, 4, 5, 6, 7

[2] Hao Chen, Kunyang Sun, Zhi Tian, Chunhua Shen, Yongming Huang, and Youliang Yan. Blendmask: Top-down meets bottom-up for instance segmentation. In *CVPR*, 2020. 1, 2, 3, 4, 5, 6, 7

[3] Yuxin Fang, Shusheng Yang, Xinggang Wang, Yu Li, Chen Fang, Ying Shan, Bin Feng, and Wenyu Liu. Queryinst: Parallelly supervised mask query for instance segmentation. In *ICCV*, 2021. 1, 2, 3, 4, 5, 6, 7

[4] Md Amirul Islam, Mahmoud Kalash, and Neil DB Bruce. Revisiting salient object detection: Simultaneous detection, ranking, and subitizing of multiple salient objects. In *CVPR*, 2018. 1, 2, 3, 4, 5, 6, 7

[5] Youngwan Lee and Jongyoul Park. Centermask: Real-time anchor-free instance segmentation. In *CVPR*, 2020. 1, 2, 3, 4, 5, 6, 7

[6] Tingting Liang, Xiaojie Chu, Yudong Liu, Yongtao Wang, Zhi Tang, Wei Chu, Jingdong Chen, and Haibin Ling. Cbnetv2: A

† Ke Xu and Xin Yang are joint corresponding authors.
‡ Rynson Lau leads this project.

Figure 1. More visual comparison of our method to existing saliency ranking methods and adapted baseline methods.

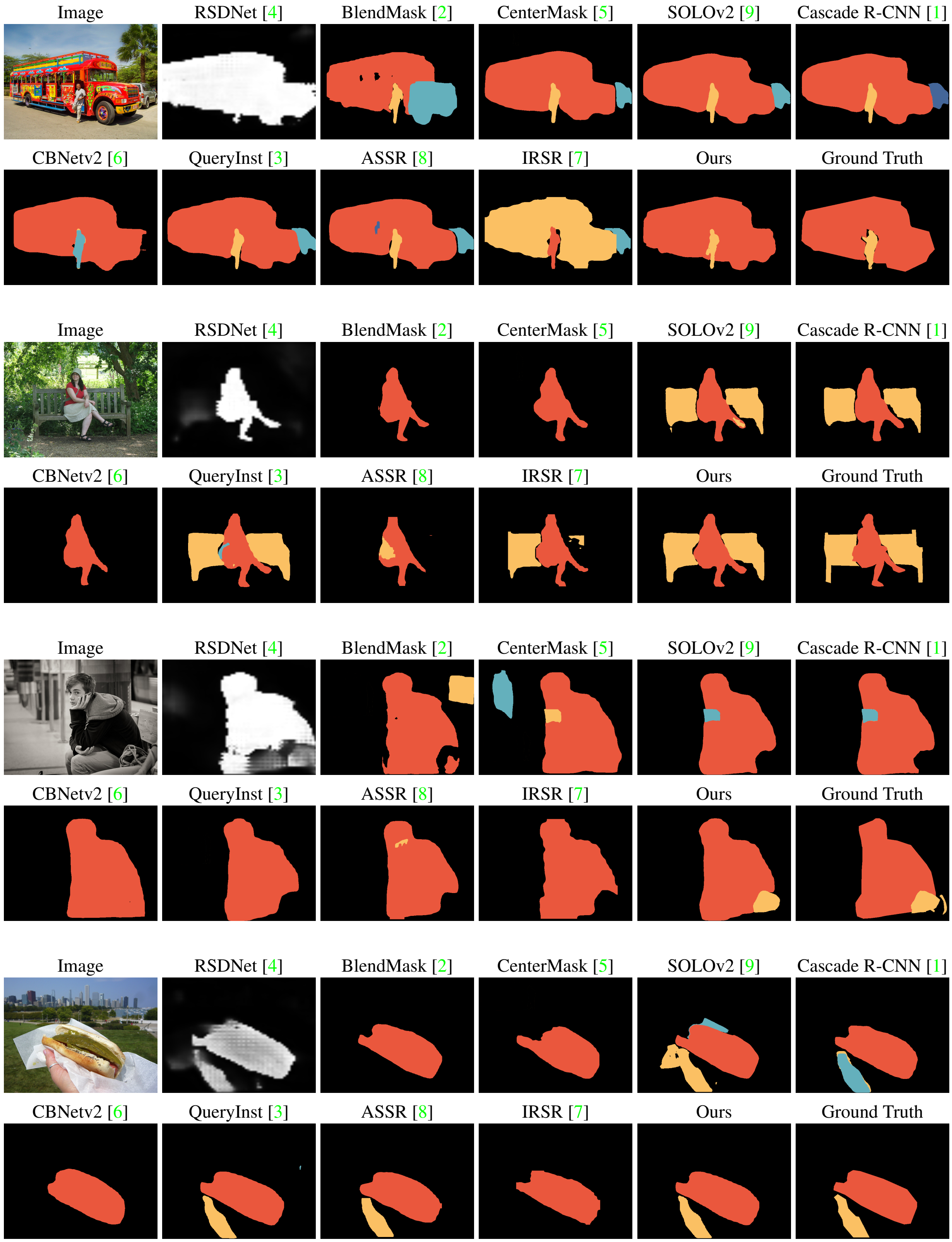

Figure 2. More visual comparison of our method to existing saliency ranking methods and adapted baseline methods.

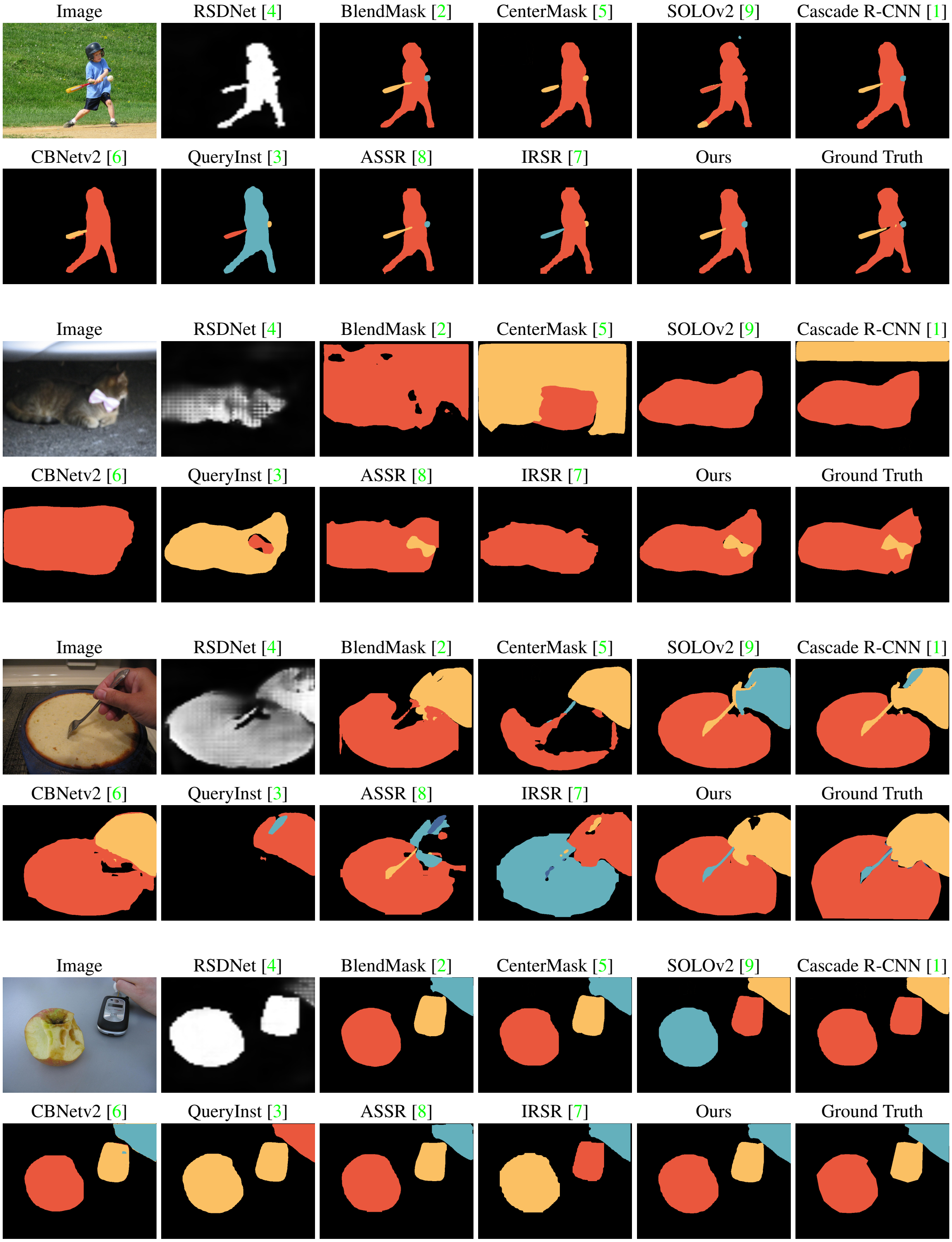

Figure 3. More visual comparison of our method to existing saliency ranking methods and adapted baseline methods.

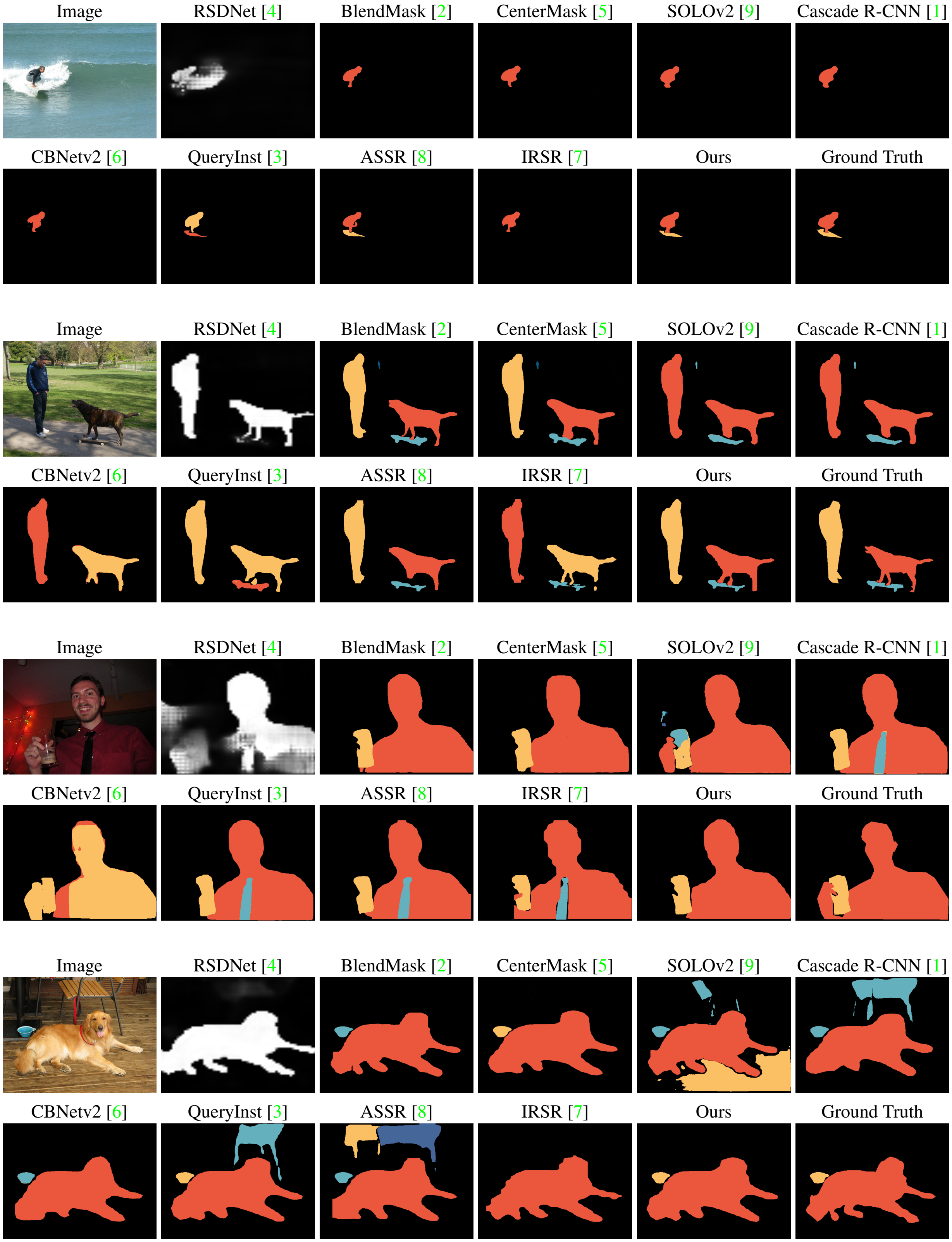

Figure 4. More visual comparison of our method to existing saliency ranking methods and adapted baseline methods.

Figure 5. More visual comparison of our method to existing saliency ranking methods and adapted baseline methods.

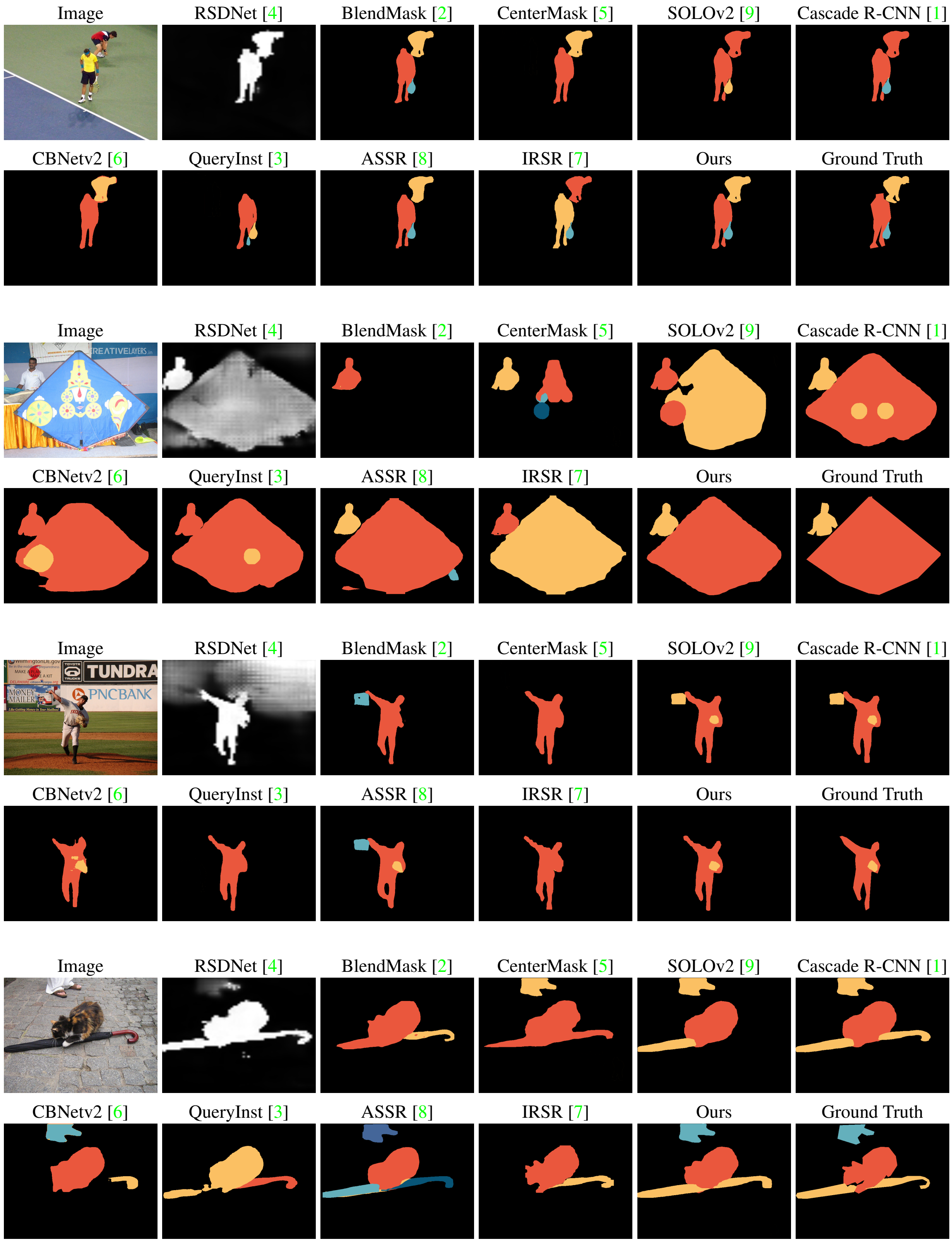

Figure 6. More visual comparison of our method to existing saliency ranking methods and adapted baseline methods.

composite backbone network architecture for object detection. *arXiv:2107.00420*, 2021. 1, 2, 3, 4, 5, 6, 7

[7] Nian Liu, Long Li, Wangbo Zhao, Junwei Han, and Ling Shao. Instance-level relative saliency ranking with graph reasoning. *TPAMI*, 2021. 1, 2, 3, 4, 5, 6, 7

[8] Avishek Siris, Jianbo Jiao, Gary KL Tam, Xianghua Xie, and Rynson WH Lau. Inferring attention shift ranks of objects for image saliency. In *CVPR*, 2020. 1, 2, 3, 4, 5, 6, 7

[9] Xinlong Wang, Tao Kong, Chunhua Shen, Yuning Jiang, and Lei Li. Solo: Segmenting objects by locations. In *ECCV*, 2020. 1, 2, 3, 4, 5, 6, 7