Depth-Aware Mirror Segmentation (Supplementary Material)

Haiyang Mei[1] Bo Dong[2,⋆] Wen Dong[1] Pieter Peers[3] Xin Yang[1,⋆] Qiang Zhang[1] Xiaopeng Wei[1,⋆]
[1] Dalian University of Technology [2] SRI International [3] College of William & Mary
https://mhaiyang.github.io/CVPR2021_PDNet/index

1. Overview

In this supplementary, we first give the mathematical definitions of the metrics used in the quantitative comparison in Section 2. Then, we show more statistics and examples of our constructed RGBD-Mirror dataset in Section 3. Finally, we present more visual comparison results of our PDNet against state-of-the-art segmentation methods in Section 4.

2. Evaluation Metrics

For a comprehensive evaluation, we adopt four widely used metrics for quantitatively assessing the mirror segmentation performance: intersection over union (IoU), weighted F-measure (F_β^w) [11], mean absolute error (MAE), and balance error rate (BER) [14].

The intersection over union (IoU) is widely used in the segmentation field, which is defined as:

$$IoU = \frac{\sum_{i=1}^{H}\sum_{j=1}^{W}(G(i,j) * P_b(i,j))}{\sum_{i=1}^{H}\sum_{j=1}^{W}(G(i,j) + P_b(i,j) - G(i,j) * P_b(i,j))}, \tag{1}$$

where G is the ground truth mask in which the values of the mirror region are 1 while those of the non-mirror region are 0; P_b is the predicted mask binarized with a threshold of 0.5; and H and W are the height and width of the ground truth mask, respectively.

We also adopt the weighted F-measure metric from the salient object detection field. F-measure (F_β) is a comprehensive measure on both the precision and recall of the prediction map. Recent studies [2, 3] have suggested that the weighted F-measure (F_β^w) [11] can provide more reliable evaluation results than the traditional F_β. Thus, we report F_β^w in the comparison.

The mean absolute error (MAE) metric is widely used in foreground-background segmentation tasks, which calculates the element-wise difference between the prediction map P and the ground truth mask G:

$$MAE = \frac{1}{H \times W}\sum_{i=1}^{H}\sum_{j=1}^{W}|P(i,j) - G(i,j)|, \tag{2}$$

where $P(i,j)$ indicates the predicted probability score at location (i,j).

The last metric is the balance error rate (BER), which is a standard metric in the shadow detection field, defined as:

$$BER = (1 - \frac{1}{2}(\frac{TP}{N_p} + \frac{TN}{N_n})) \times 100, \tag{3}$$

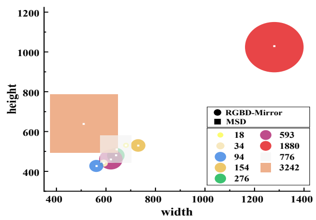
where TP, TN, N_p, and N_n represent the numbers of true positive pixels, true negative pixels, mirror pixels, and non-mirror pixels, respectively.

Note that for IoU and F_β^w, it is the higher the better, while for MAE and BER, it is the lower the better.

⋆ Xin Yang (xinyang@dlut.edu.cn) and Xiaopeng Wei are the corresponding authors. Xin Yang and Bo Dong lead this project.

Datasets	MSD [19]	GDD [12]	PMD [10]	STEREO [15]	NYUD-V2 [13]	KITTI-ROAD [6]	LFSD [9]	RGBD135 [1]	NLPR [16]	NJUD [8]	SSD [20]	DUT-RGBD [17]	SIP [4]	RGBD-Mirror
Publication	ICCV	CVPR	CVPR	CVPR	ECCV	ITSC	CVPR	ICIMCS	ECCV	ICIP	ICCVW	ICCV	TNNLS	Ours
Year	2019	2020	2020	2012	2012	2013	2014	2014	2014	2014	2017	2019	2020	2021
Number	4,018	3,916	6,461	797	1,449	289	100	135	1,000	1,985	80	1,200	929	3,049
Depth	×	×	×	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Table 1. Number statistics of relevant RGB-D segmentation datasets.



(a) resolution distribution

Datasets	RGBD135 [1]	NLPR [16]	DUT-RGBD [17]	SIP [4]	RGBD-Mirror
Target Region	0.224 / 0.053	0.250 / 0.077	0.314 / 0.105	0.117 / 0.075	0.414 / 0.266
Non-Target Region	0.405 / 0.188	0.462 / 0.224	0.602 / 0.184	0.630 / 0.310	0.343 / 0.137
Difference	-0.181 / -0.135	-0.212 / -0.147	-0.288 / -0.079	-0.513 / -0.235	0.071 / 0.129

(b) depth distribution (blue and orange numbers denote the mean and standard deviation, respectively)

Figure 1. Statistics of our dataset. We show that our RGBD-Mirror has reasonable property distributions in terms of resolution and depth.

3. RGB-D Mirror Segmentation Dataset

Our first contribution is introducing a new RGB-D mirror segmentation dataset, named RGBD-Mirror, which contains 3,049 mirror images, depth maps, and corresponding ground truth mirror masks.

Number Statistics: the scale of a dataset plays an important role in providing diverse patterns for training a model. As shown in Table 1, our RGBD-Mirror offers the most RGB-D images among all compared RGB-D datasets.

Resolution Statistics: the images in our RGBD-Mirror dataset vary in size, as shown in Figure 1(a). Compared with the MSD [19], our dataset contains more images with high resolution (*i.e*., 1280×1024) and thus could provide more detailed information for accurate mirror segmentation.

Depth Statistics: Figure 1(b) presents the statistics of the depth inside and outside target regions in the existing RGB-D segmentation datasets. We observe that (i) the average depth of salient objects is lower than the ones of backgrounds, but mirror regions have a higher average depth value than non-mirror regions; (ii) the standard deviation in terms of the depth inside salient objects is typically small and lower than the ones of backgrounds. In contrast, depth varies dramatically inside mirror regions (*i.e*., the corresponding standard deviation is 0.266). The large depth variation inside the mirror leads to a great challenge for RGB-D mirror segmentation.

More examples of our RGBD-Mirror are shown in Figure 2, 3, 4, 5, 6, and 7.

4. Visual Comparison

We further qualitatively compare our PDNet with two prior mirror segmentation methods (*i.e*., MirrorNet [19] and PMD [10]) as well as the best approaches from other three categories (*i.e*., semantic segmentation method CCNet [7], salient object detection method F3Net [18], and RGB-D saliency detection method BBS-Net [5]). The results are shown in Figure 8, 9, 10, and 11.

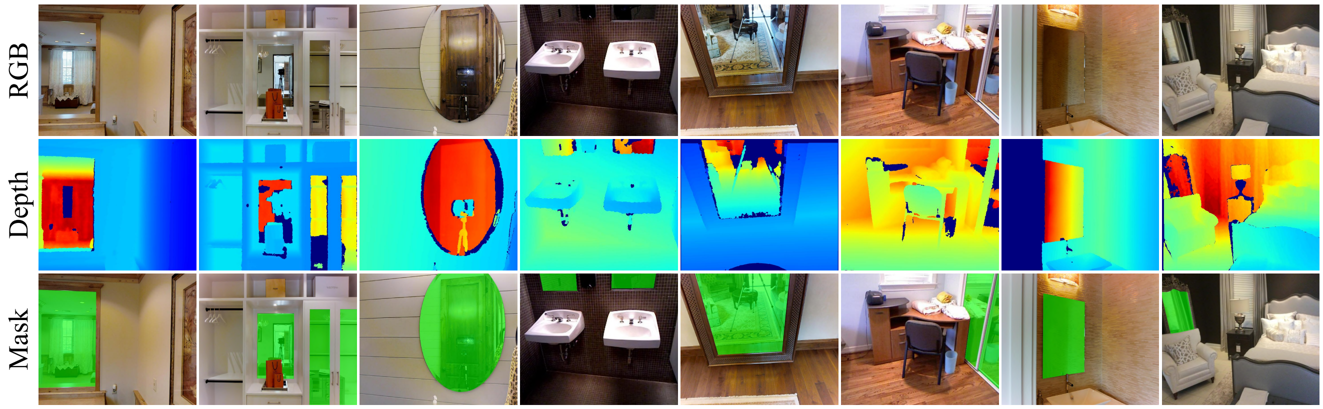


Figure 2. Visual examples of mirror image, depth map, and mirror mask triplets in our RGBD-Mirror dataset.

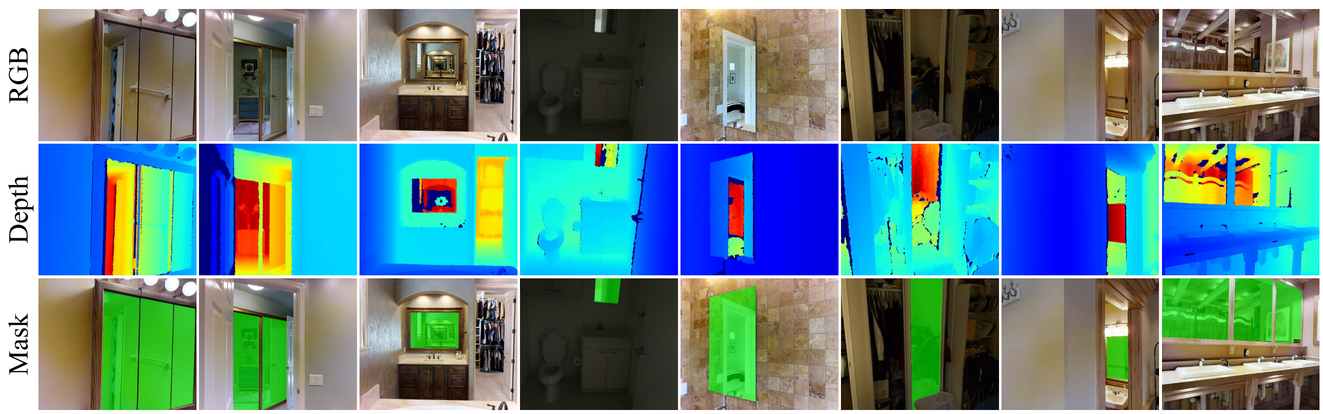


Figure 3. Visual examples of mirror image, depth map, and mirror mask triplets in our RGBD-Mirror dataset.

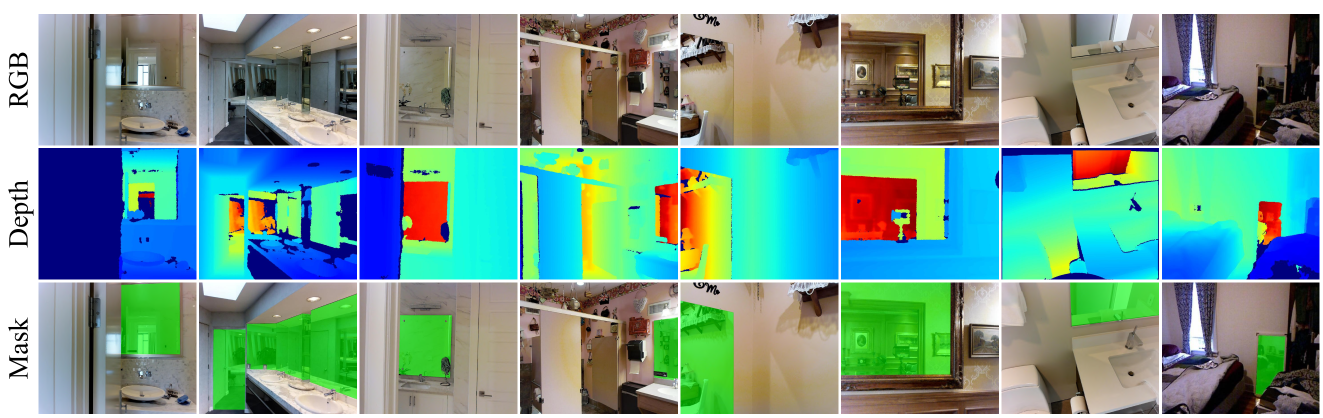


Figure 4. Visual examples of mirror image, depth map, and mirror mask triplets in our RGBD-Mirror dataset.

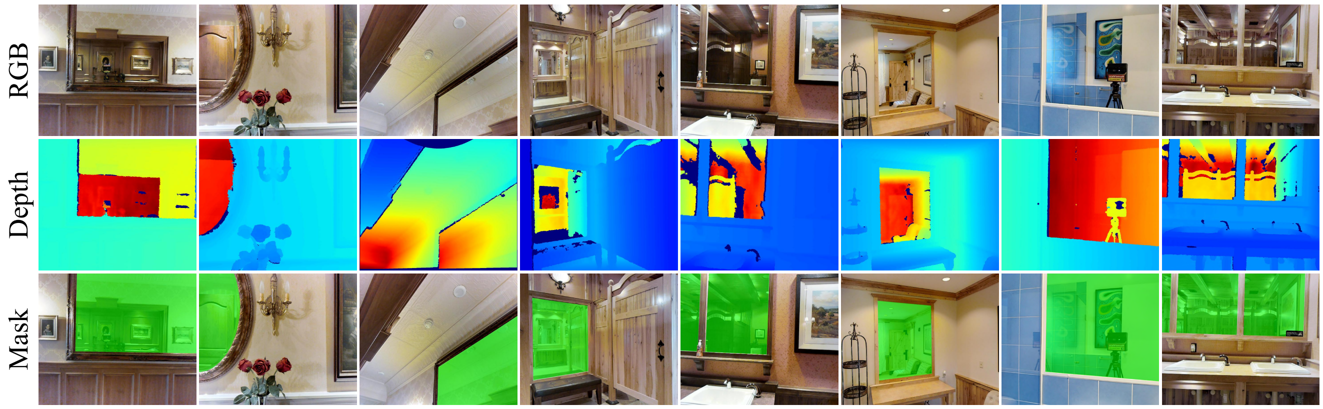


Figure 5. Visual examples of mirror image, depth map, and mirror mask triplets in our RGBD-Mirror dataset.

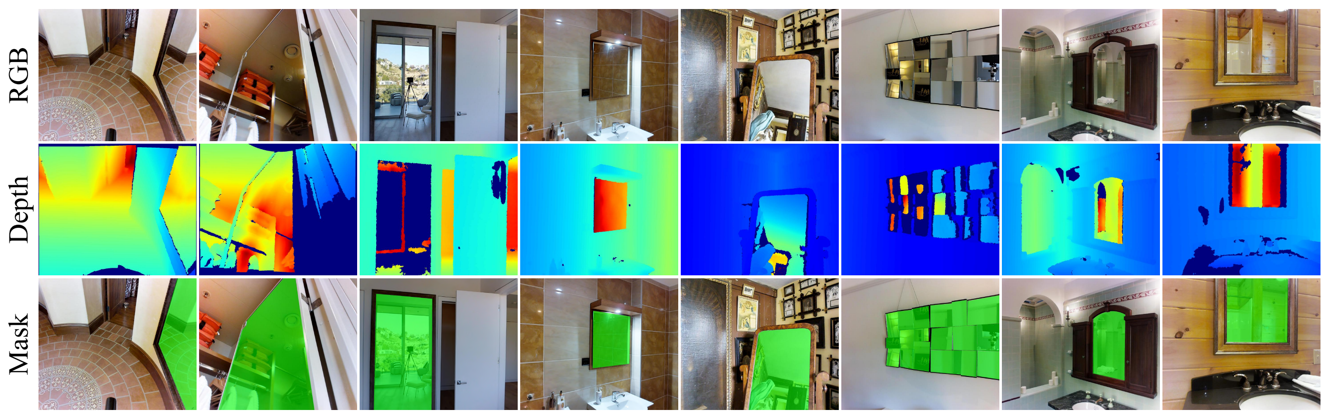


Figure 6. Visual examples of mirror image, depth map, and mirror mask triplets in our RGBD-Mirror dataset.

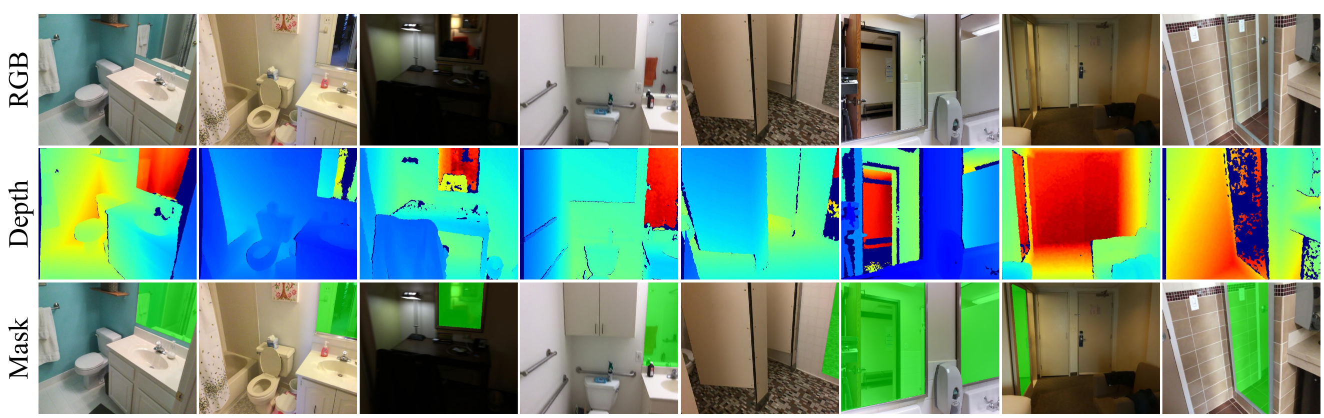


Figure 7. Visual examples of mirror image, depth map, and mirror mask triplets in our RGBD-Mirror dataset.

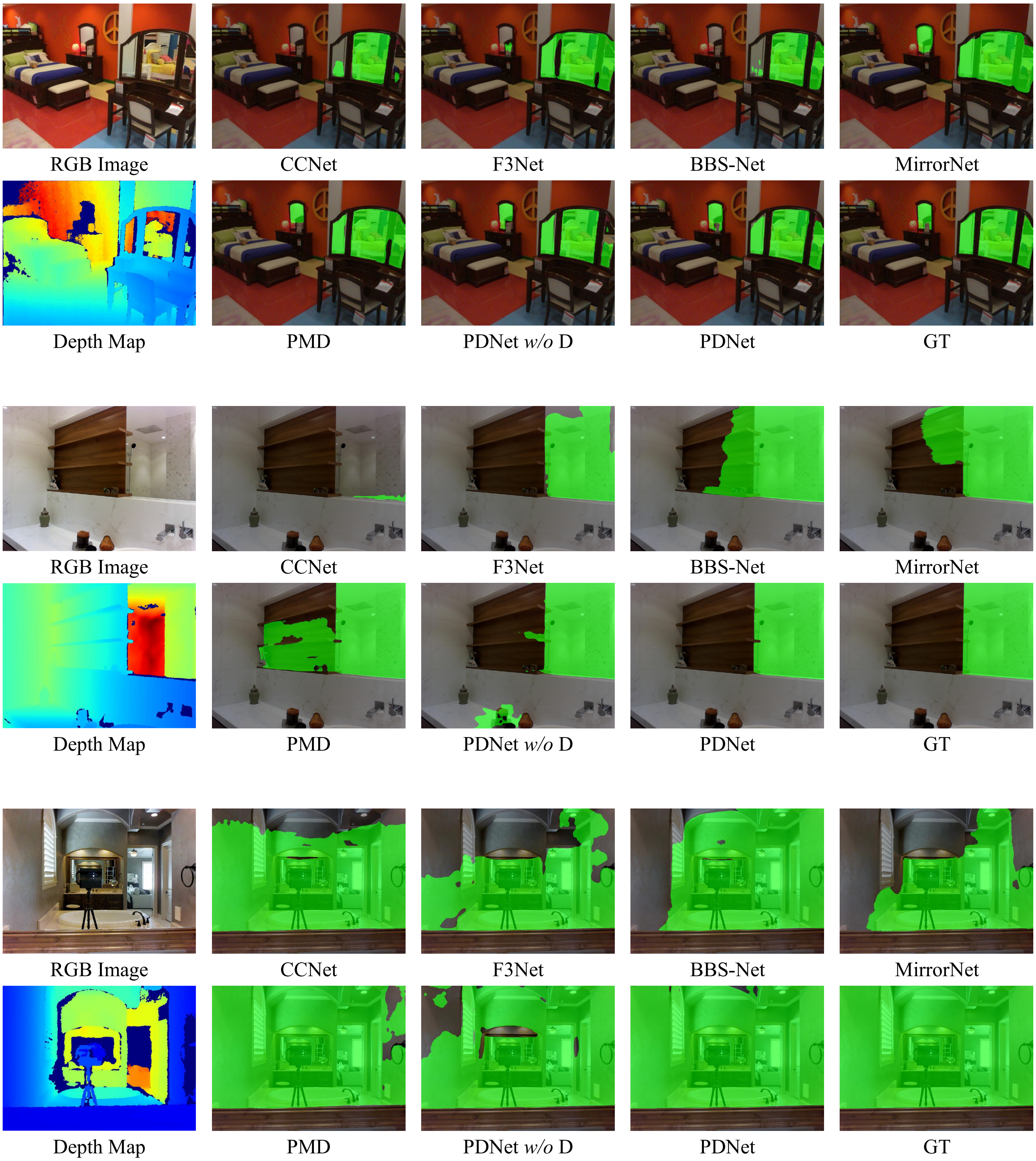


Figure 8. Visual comparison of PDNet against state-of-the-art segmentation methods retrained on the RGBD-Mirror dataset.

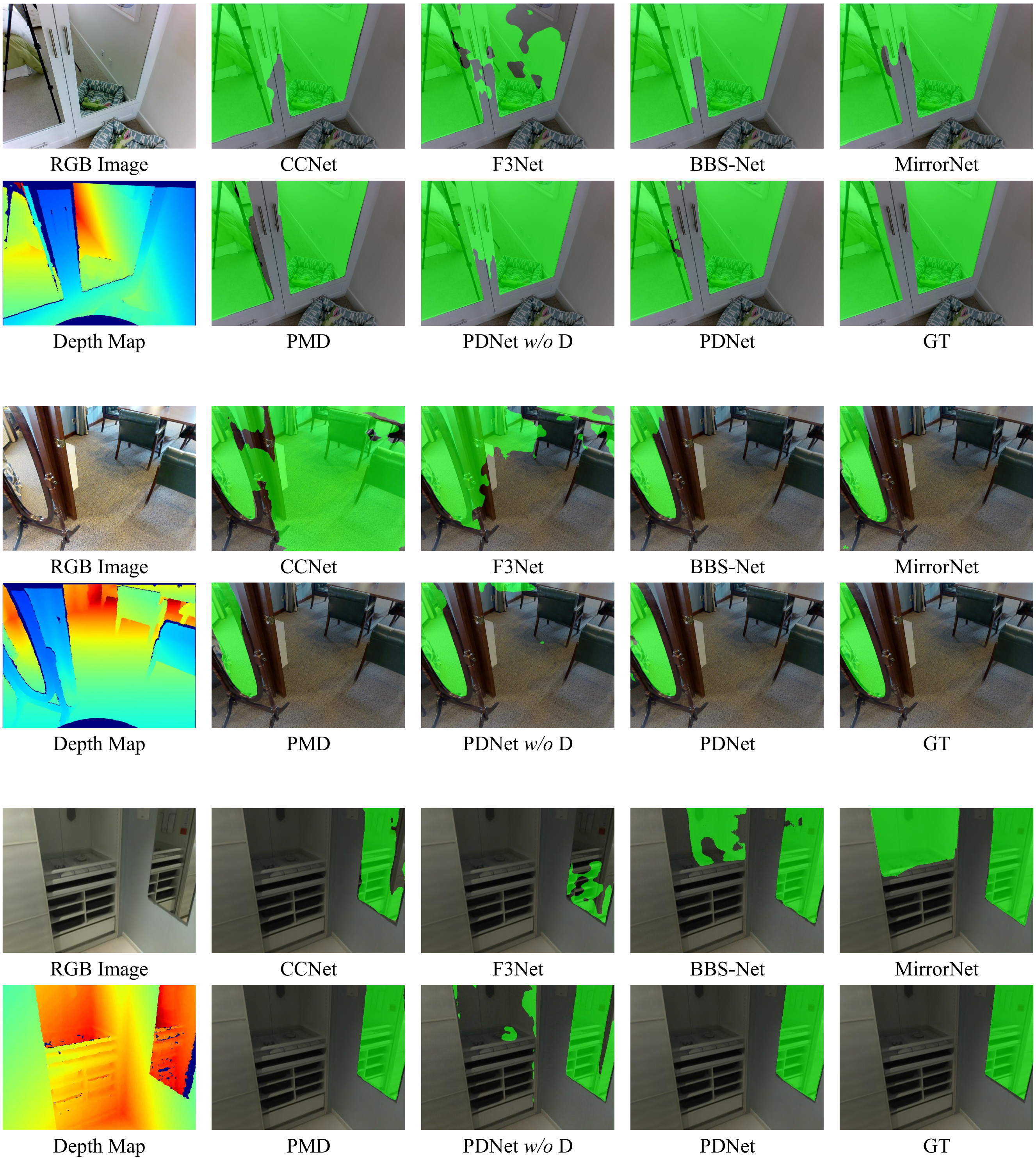


Figure 9. Visual comparison of PDNet against state-of-the-art segmentation methods retrained on the RGBD-Mirror dataset.

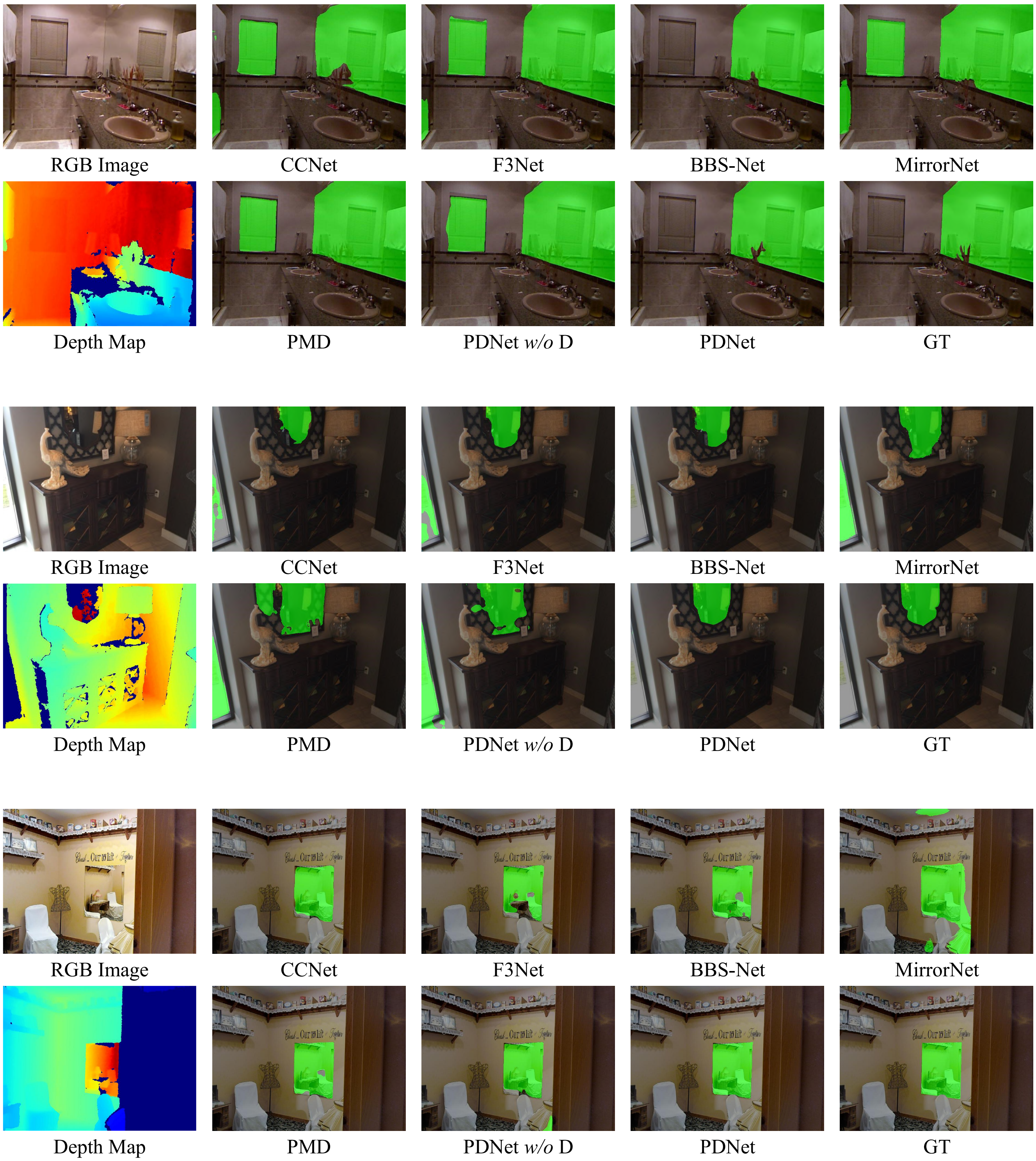


Figure 10. Visual comparison of PDNet against state-of-the-art segmentation methods retrained on the RGBD-Mirror dataset.

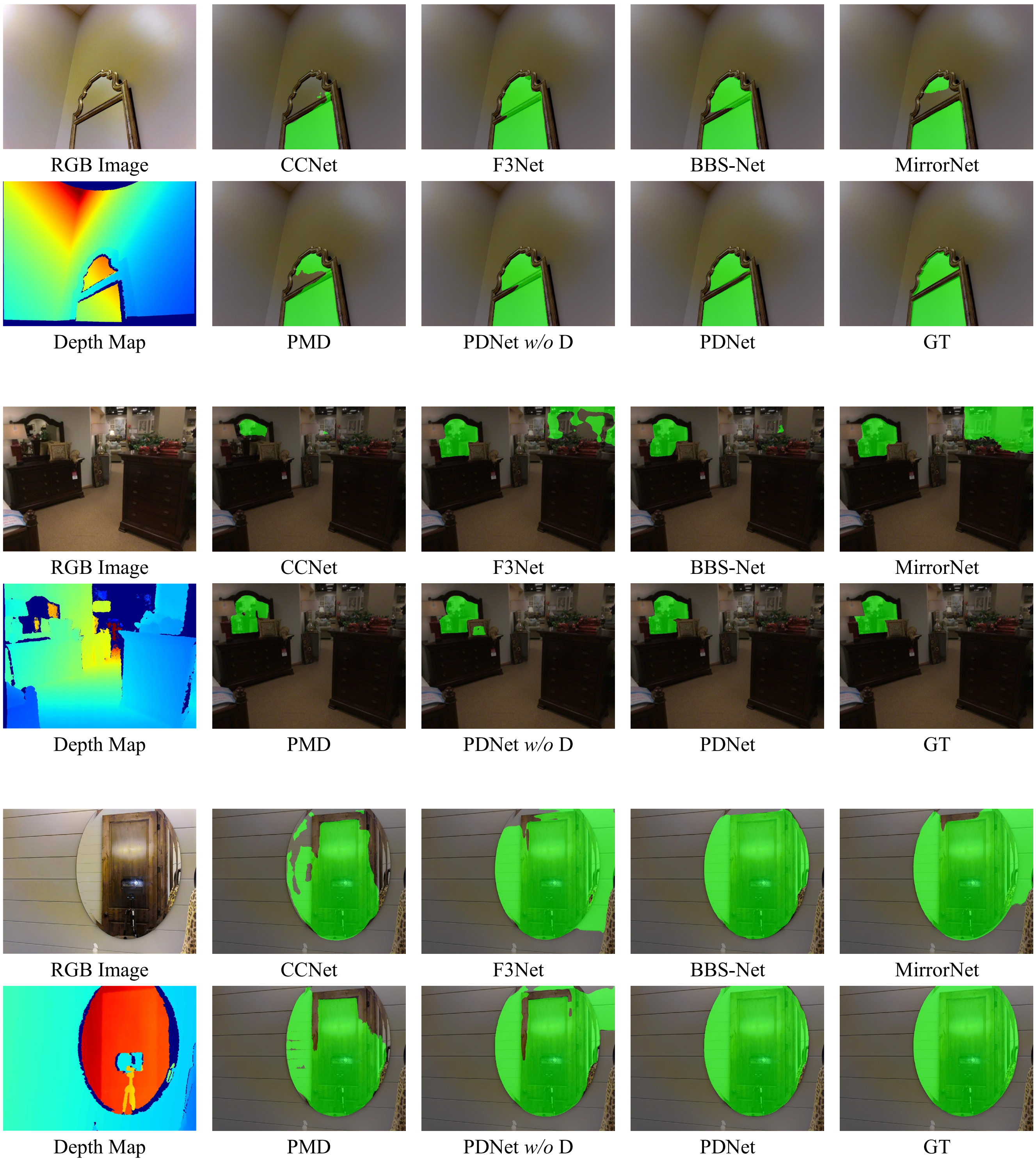


Figure 11. Visual comparison of PDNet against state-of-the-art segmentation methods retrained on the RGBD-Mirror dataset.

References

[1] Yupeng Cheng, Huazhu Fu, Xingxing Wei, Jiangjian Xiao, and Xiaochun Cao. Depth enhanced saliency detection method. In *ICIMCS*, 2014. 2

[2] Deng-Ping Fan, Ming-Ming Cheng, Yun Liu, Tao Li, and Ali Borji. Structure-measure: A new way to evaluate foreground maps. In *ICCV*, 2017. 1

[3] Deng-Ping Fan, Cheng Gong, Yang Cao, Bo Ren, Ming-Ming Cheng, and Ali Borji. Enhanced-alignment measure for binary foreground map evaluation. In *IJCAI*, 2018. 1

[4] Deng-Ping Fan, Zheng Lin, Zhao Zhang, Menglong Zhu, and Ming-Ming Cheng. Rethinking RGB-D salient object detection: Models, datasets, and large-scale benchmarks. *IEEE TNNLS*, 2020. 2

[5] Deng-Ping Fan, Yingjie Zhai, Ali Borji, Jufeng Yang, and Ling Shao. Bbs-net: Rgb-d salient object detection with a bifurcated backbone strategy network. In *ECCV*, 2020. 2

[6] J. Fritsch, T. Kühnl, and A. Geiger. A new performance measure and evaluation benchmark for road detection algorithms. In *ITSC*, 2013. 2

[7] Zilong Huang, Xinggang Wang, Lichao Huang, Chang Huang, Yunchao Wei, and Wenyu Liu. Ccnet: Criss-cross attention for semantic segmentation. In *ICCV*, 2019. 2

[8] Ran Ju, Ling Ge, Wenjing Geng, Tongwei Ren, and Gangshan Wu. Depth saliency based on anisotropic center-surround difference. In *ICIP*, 2014. 2

[9] Nianyi Li, Jinwei Ye, Yu Ji, Haibin Ling, and Jingyi Yu. Saliency detection on light field. In *CVPR*, 2014. 2

[10] Jiaying Lin, Guodong Wang, and Rynson W.H. Lau. Progressive mirror detection. In *CVPR*, 2020. 2

[11] Ran Margolin, Lihi Zelnik-Manor, and Ayellet Tal. How to evaluate foreground maps? In *CVPR*, 2014. 1

[12] Haiyang Mei, Xin Yang, Yang Wang, Yuanyuan Liu, Shengfeng He, Qiang Zhang, Xiaopeng Wei, and Rynson W.H. Lau. Don't hit me! glass detection in real-world scenes. In *CVPR*, 2020. 2

[13] Pushmeet Kohli Nathan Silberman, Derek Hoiem and Rob Fergus. Indoor segmentation and support inference from rgbd images. In *ECCV*, 2012. 2

[14] Vu Nguyen, Tomas F Yago Vicente, Maozheng Zhao, Minh Hoai, and Dimitris Samaras. Shadow detection with conditional generative adversarial networks. In *ICCV*, 2017. 1

[15] Yuzhen Niu, Yujie Geng, Xueqing Li, and Feng Liu. Leveraging stereopsis for saliency analysis. In *CVPR*, 2012. 2

[16] Houwen Peng, Bing Li, Weihua Xiong, Weiming Hu, and Rongrong Ji. Rgbd salient object detection: A benchmark and algorithms. In *ECCV*, 2014. 2

[17] Yongri Piao, Wei Ji, Jingjing Li, Miao Zhang, and Huchuan Lu. Depth-induced multi-scale recurrent attention network for saliency detection. In *ICCV*, 2019. 2

[18] Jun Wei, Shuhui Wang, and Qingming Huang. F3net: Fusion, feedback and focus for salient object detection. In *AAAI*, 2020. 2

[19] Xin Yang, Haiyang Mei, Ke Xu, Xiaopeng Wei, Baocai Yin, and Rynson W.H. Lau. Where is my mirror? In *ICCV*, 2019. 2

[20] Chunbiao Zhu and Ge Li. A three-pathway psychobiological framework of salient object detection using stereoscopic technology. In *ICCVW*, 2017. 2