

## Supplementary Material

In the supplementary material, we show more quantitative and qualitative results in the following aspects:

1. In the paper, we retrain Amulet [11], DSS [2], NLDF [7] on the training set of DUTS dataset for fairly comparison. Hence we retrain the proposed CPD model on MSRA-B [6] dataset to compare with the original Amulet, DSS and NLDF models. In Section 1, we present the quantitative and qualitative results.
2. In the paper, the proposed model is compared with eight state-of-the-art algorithms. In section 2, we show the F-measure and PR curves of the proposed models and existing models. Besides, more qualitative results are added.
3. In Section 3, we present F-measure and PR curves of original models (BMPM, Amulet, NLDF) and optimized models (-CPD-A, -CPD) by the proposed framework. Besides, we add more visual comparison results.

### 1. Compared to original Amulet, DSS, NLDF

For fairly comparison, we retrain Amulet [11], DSS [2] and NLDF [7] on the training set of DUTS [8] dataset as the proposed CPD model and other state-of-the-art algorithms. Here we also train our model with the same dataset as these three deep aggregation models. More specially, DSS and NLDF are trained on MSRA-B [6] dataset, Amulet is trained on MSRA10K [1] dataset, and there is a large overlap between these two datasets. Besides, MSRA10K is more than twice bigger than MSRA-B dataset, models trained on MSRA10K have a competitive performance compared to those trained on the MSRA-B dataset [2]. Here we just utilize MSRA-B for training. Besides, DSS adopts CRF as post-preprocessing procedure. Hence we also utilize CRF to refine our final results. The quantitative results are shown in Table. 1. In each setting, the proposed model outperforms existing algorithms in all cases. Fig. 3 shows the qualitative comparison on challenging cases: small object and large object. Fig. 4 shows the qualitative comparison on challenging cases: multiple objects and complex scenes.

Method	Backbone	ECSSD [9]			HKU-IS [4]			DUT-OMRON [10]			DUTS [8]			PASCAL-S [5]		
		maxF	avgF	MAE	maxF	avgF	MAE	maxF	avgF	MAE	maxF	avgF	MAE	maxF	avgF	MAE
Amulet [11]	VGG16	0.915	0.870	0.059	0.894	0.838	0.053	0.744	0.648	0.097	0.779	0.672	0.085	0.832	0.764	0.097
NLDF [7]	VGG16	0.905	0.874	0.063	0.900	0.872	0.049	0.753	0.683	0.080	0.813	0.738	0.065	0.826	0.770	0.099
DSS [2]	VGG16	0.908	0.865	0.062	0.898	0.854	0.051	0.760	0.673	0.074	0.813	0.712	0.065	0.825	0.763	0.103
CPD-A	VGG16	0.916	0.883	0.059	0.905	0.865	0.048	0.775	0.694	0.069	0.818	0.727	0.064	0.836	0.783	0.096
CPD	VGG16	<b>0.922</b>	<b>0.895</b>	<b>0.052</b>	<b>0.912</b>	<b>0.876</b>	<b>0.043</b>	<b>0.786</b>	<b>0.722</b>	<b>0.064</b>	<b>0.824</b>	<b>0.750</b>	<b>0.061</b>	<b>0.846</b>	<b>0.797</b>	<b>0.090</b>
Post-processing with CRF [3](-C)																
DSS-C [2]	VGG16	0.921	0.904	0.052	0.911	0.896	0.041	0.781	0.740	0.063	0.825	0.790	0.057	0.835	0.805	0.095
CPD-C	VGG16	<b>0.926</b>	<b>0.911</b>	<b>0.048</b>	<b>0.917</b>	<b>0.902</b>	<b>0.038</b>	<b>0.792</b>	<b>0.753</b>	<b>0.060</b>	<b>0.829</b>	<b>0.796</b>	<b>0.056</b>	<b>0.847</b>	<b>0.816</b>	<b>0.087</b>

Table 1: Comparison of different models on five benchmark datasets. “-C” means using CRF [3] as the post-processing procedure. In each setting, the best score is shown in **red**.

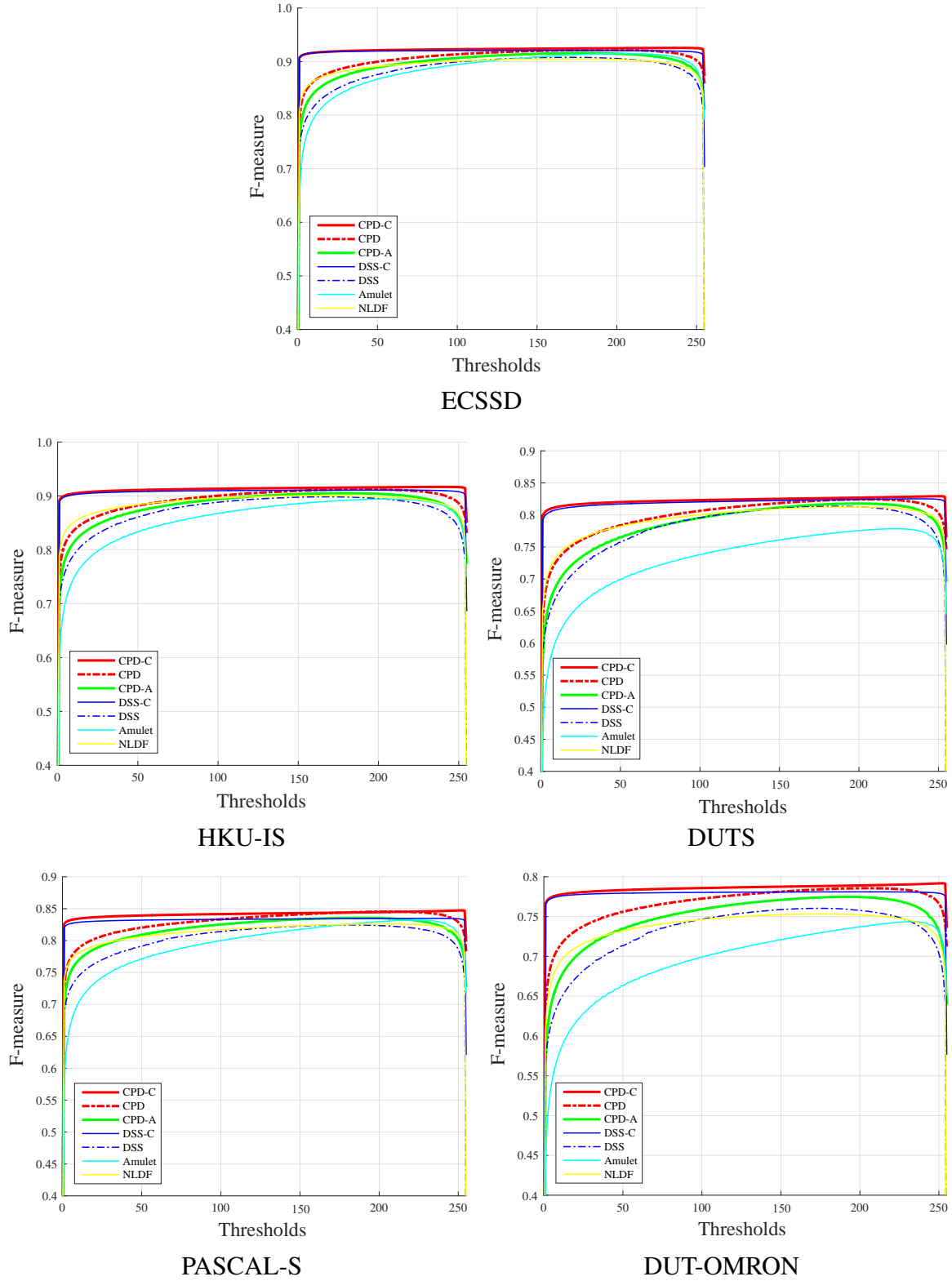


Figure 1: *F-measure curves of the proposed models and existing models (using MSRA-B or MSRA10K for training).*

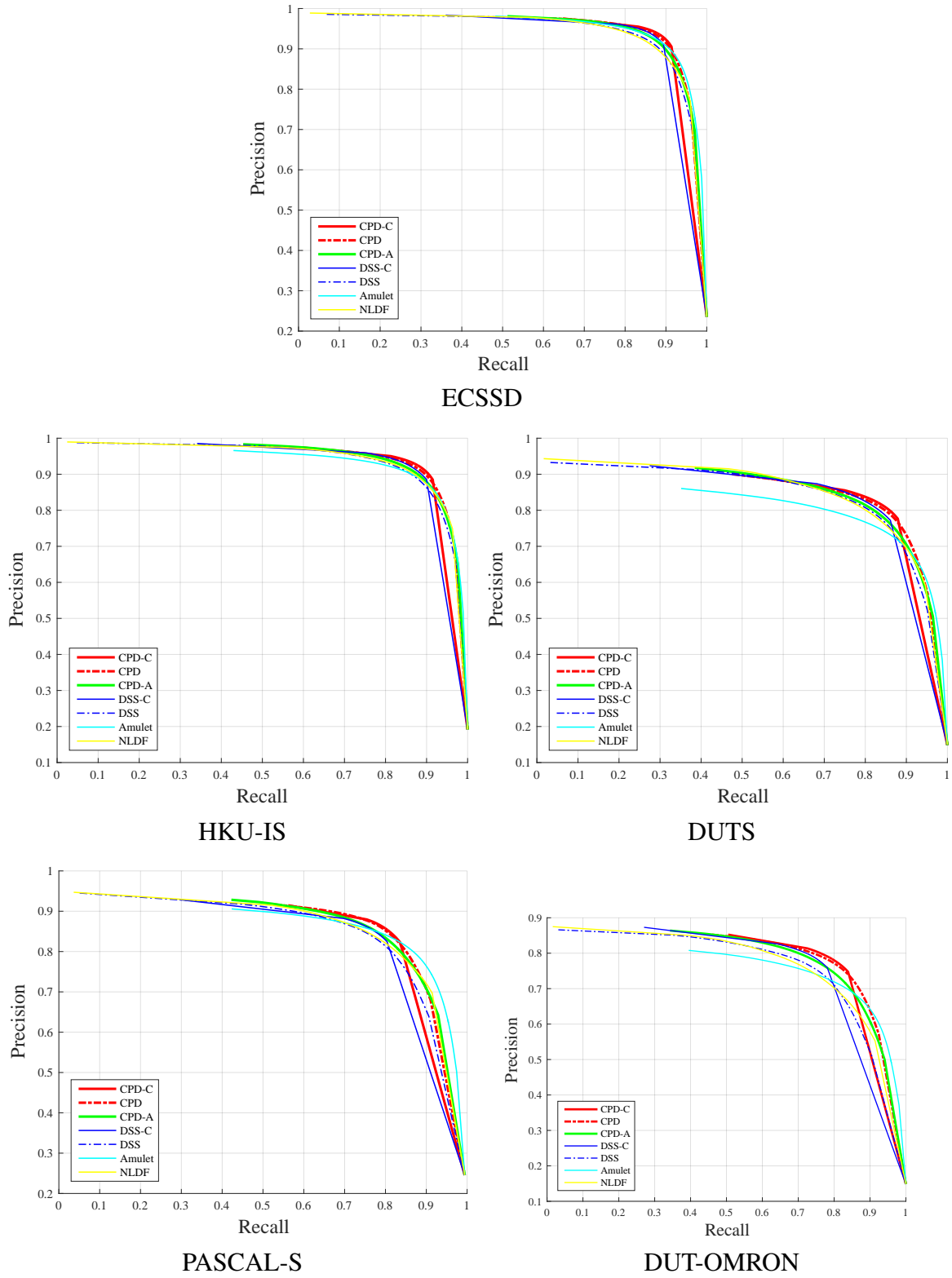


Figure 2: *PR curves of the proposed models and existing models (using MSRA-B or MSRA10K for training).*



Figure 3: Visual comparisons of different models in challenging cases: small object (top six rows), large object (bottom six rows).





Figure 4: Visual comparisons of different models in challenging cases: multiple objects (top six rows), complex scenes (bottom six rows).

## 2. More Comparison with State-of-the-arts

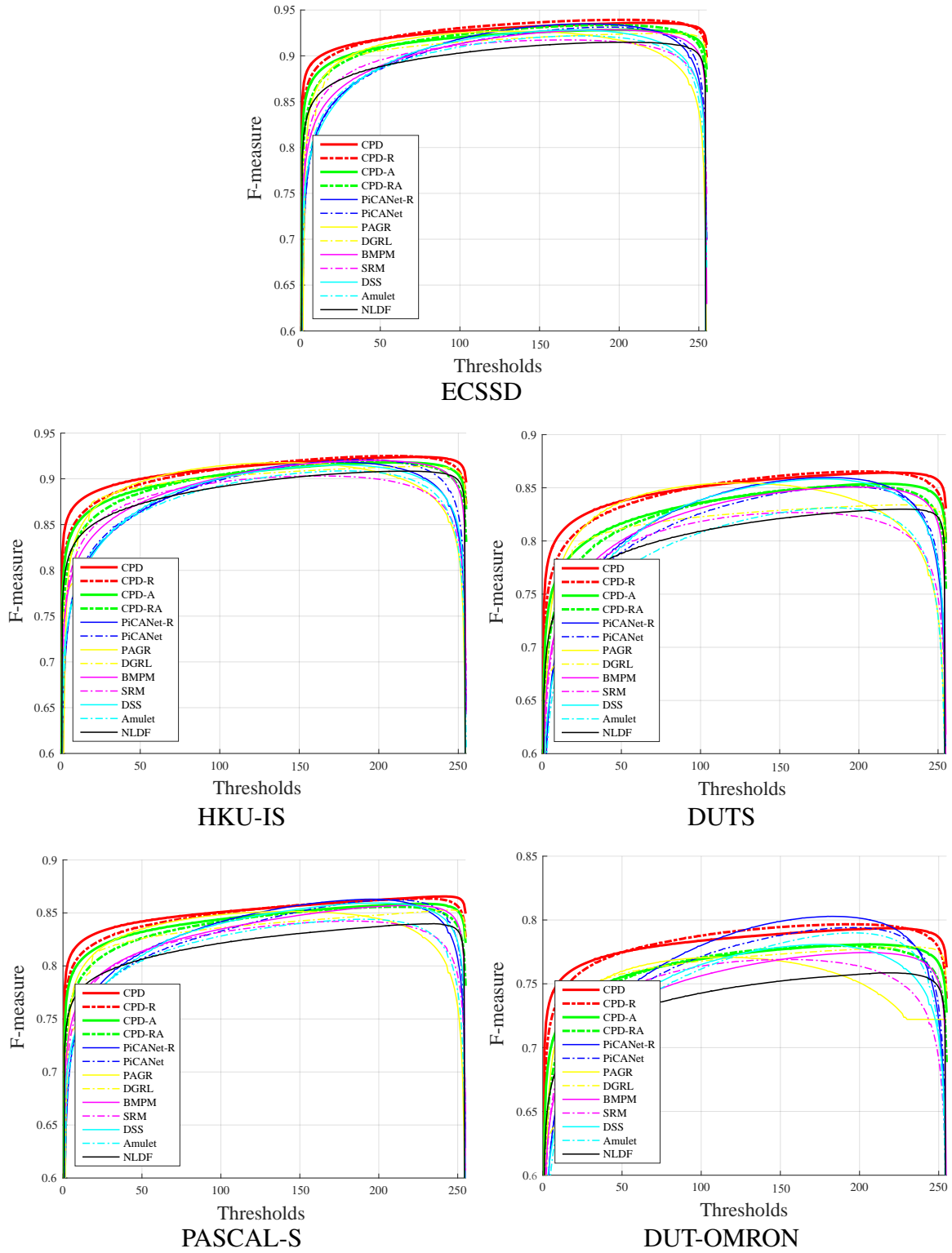


Figure 5: *F-measure curves of the proposed models and existing models (using training set of DUTS dataset).*

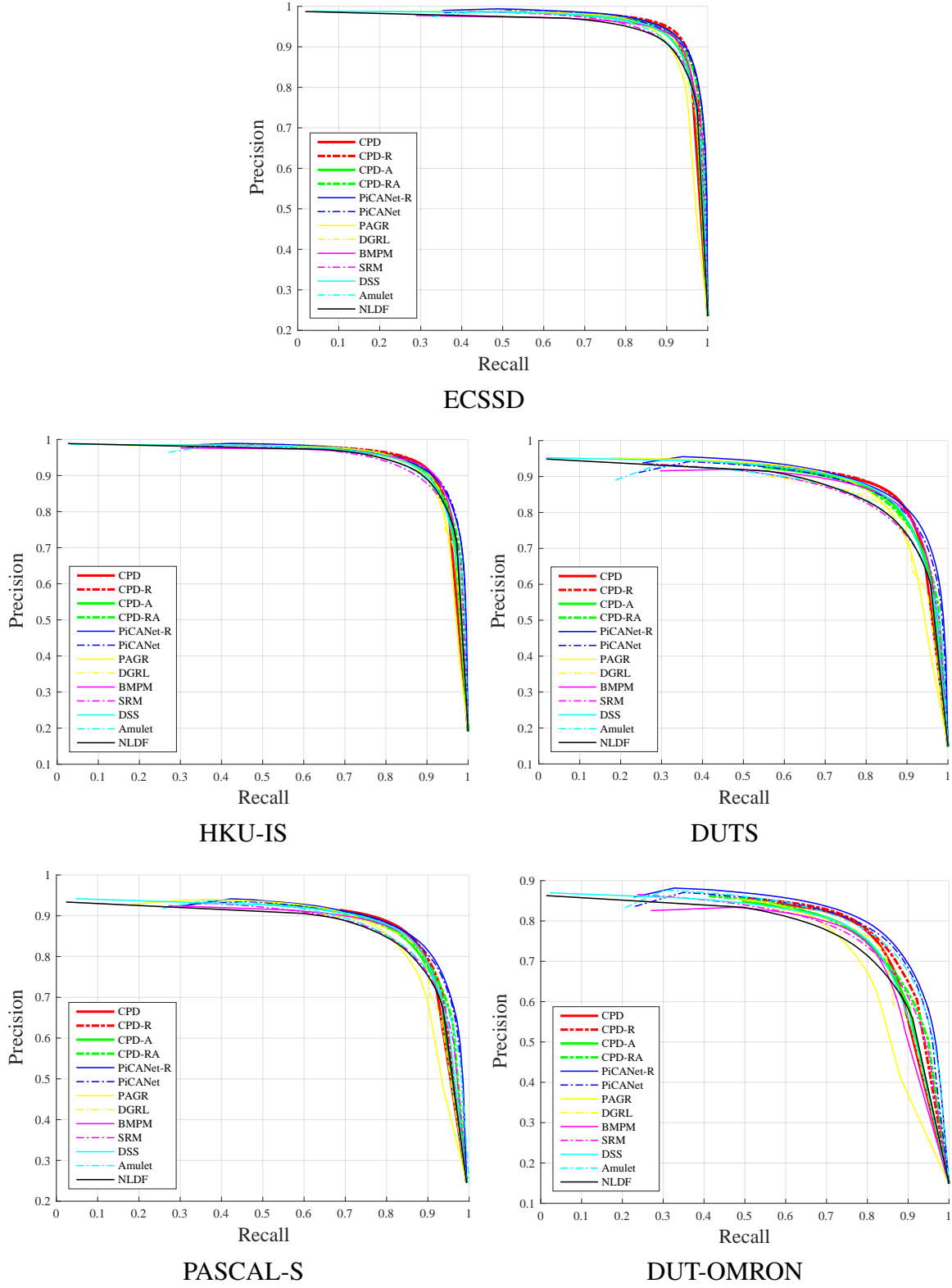


Figure 6: *PR curves of the proposed models and existing models (using training set of DUTS dataset).*

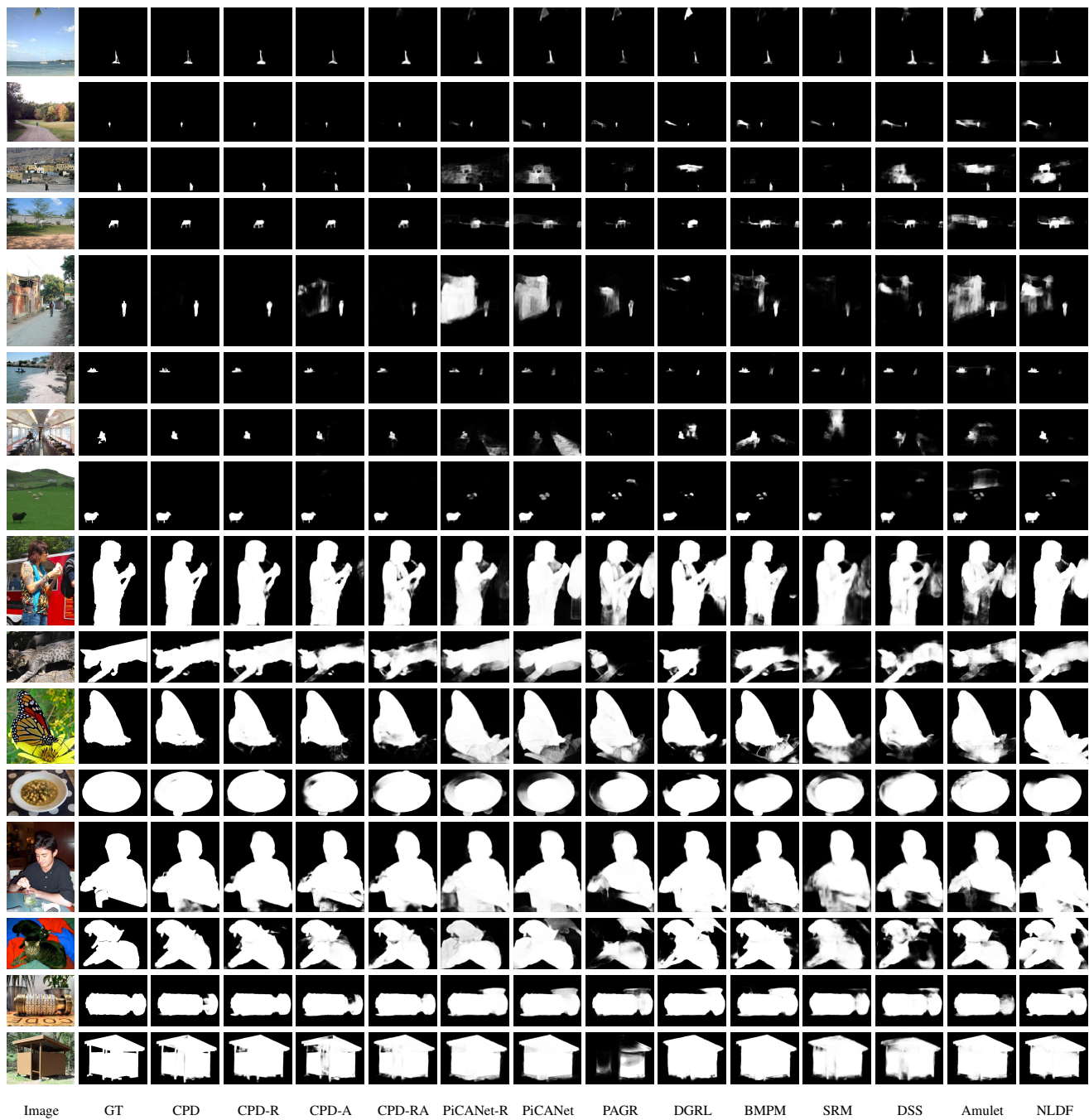


Figure 7: Visual comparisons with the existing methods in some challenging cases: small object (top eight rows) and large object (bottom eight rows).

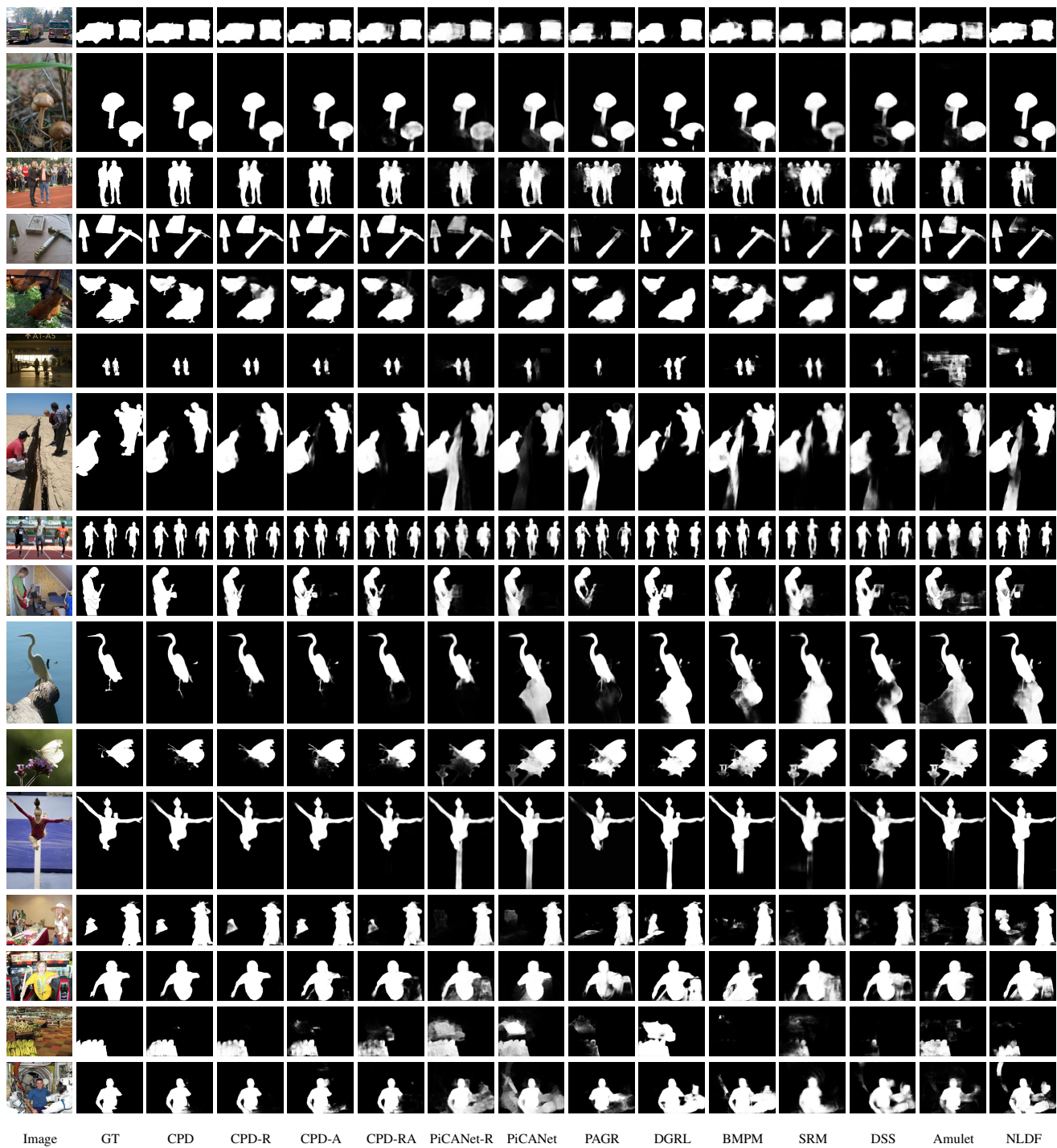


Figure 8: Visual comparisons with the existing methods in some challenging cases: multiple objects (top eight rows) and complex scenes (bottom eight rows).

### 3. More Comparison of Optimized models

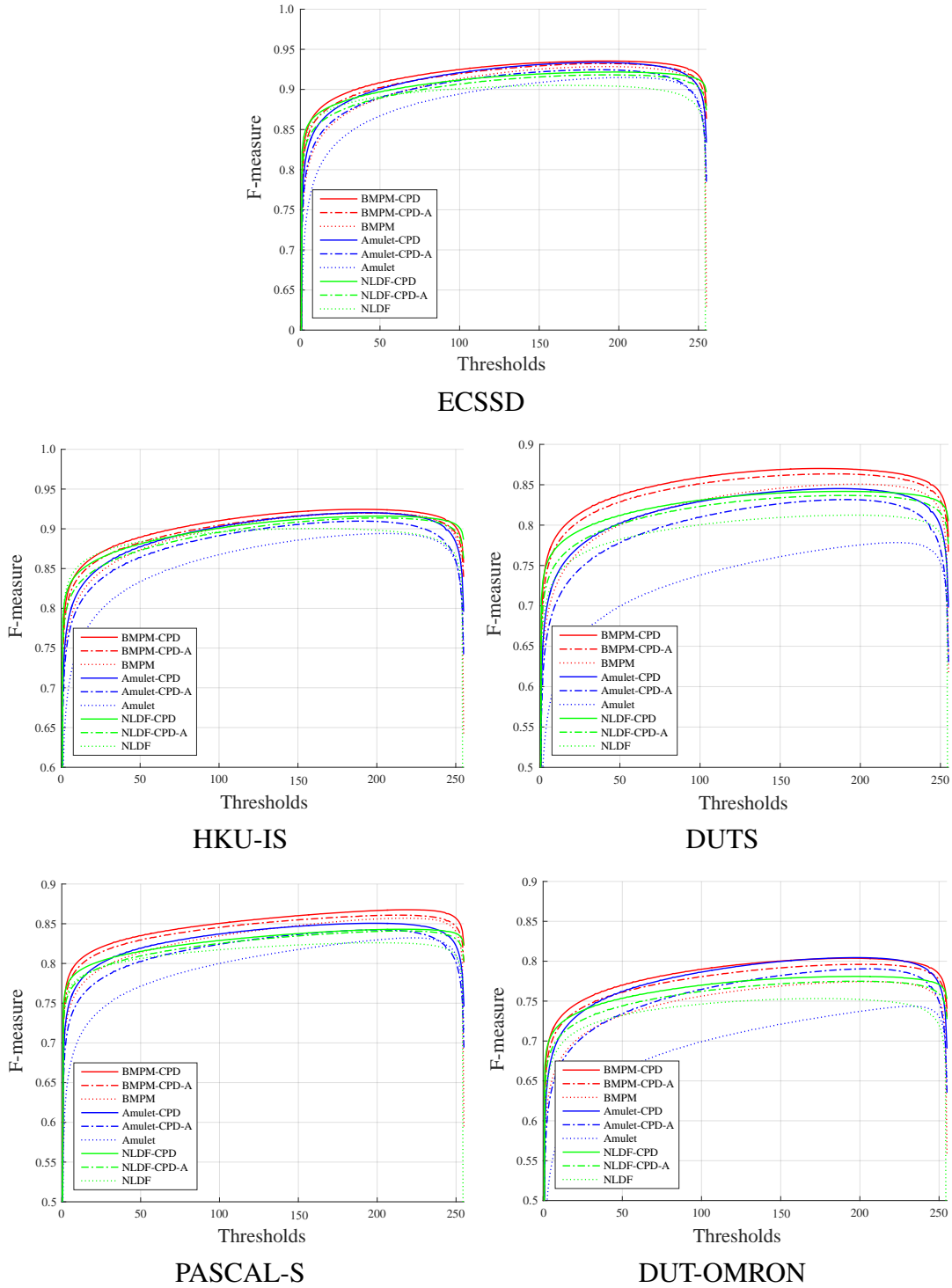


Figure 9: *F*-measure curves of original models (BMPM, Amulet, NLDF) with optimized models (-CPD-A, -CPD) by the proposed framework.

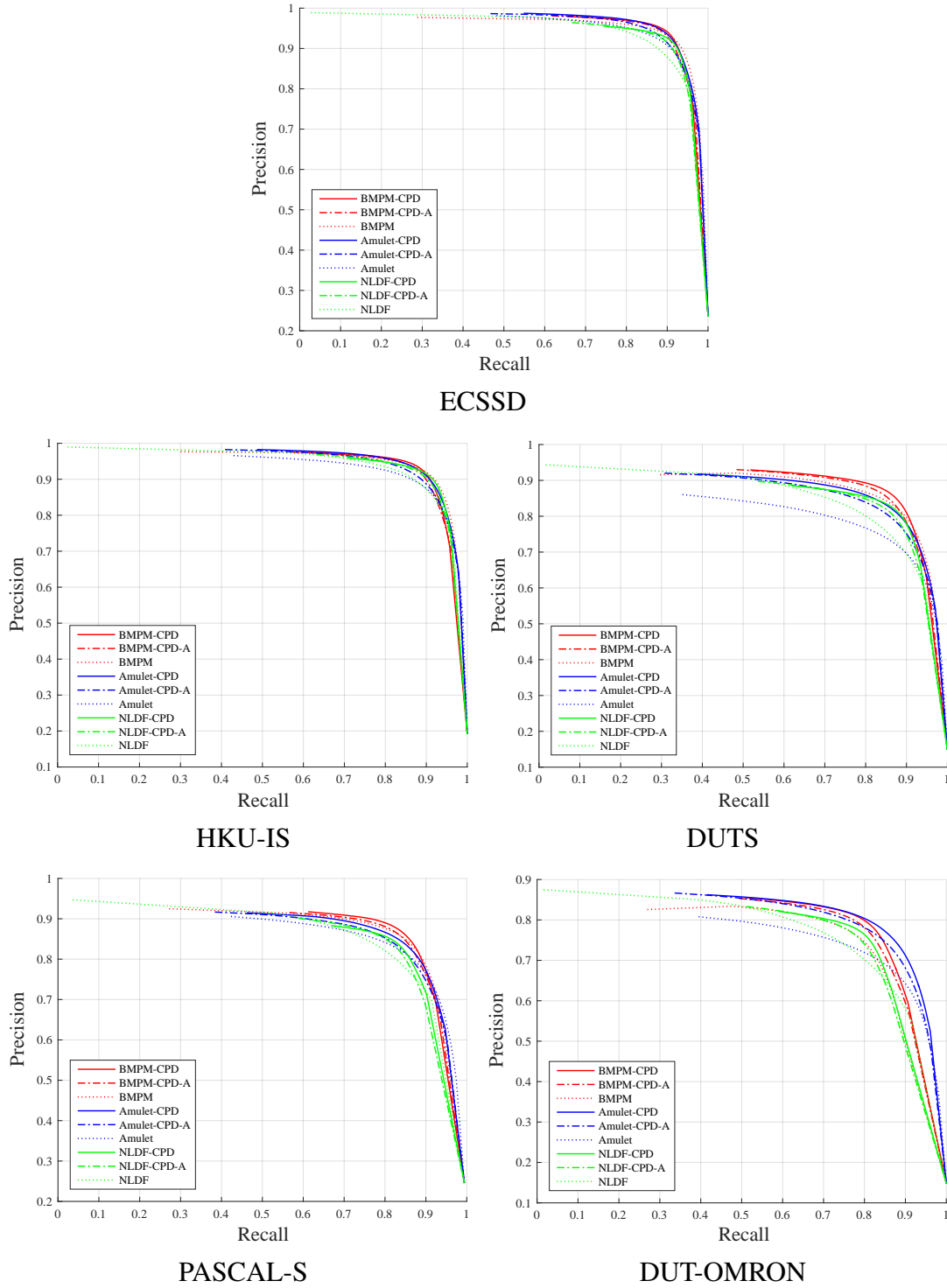


Figure 10: PR curves of original models (BMPM, Amulet, NLDF) with optimized models (-CPD, -CPD-A) by the proposed framework.



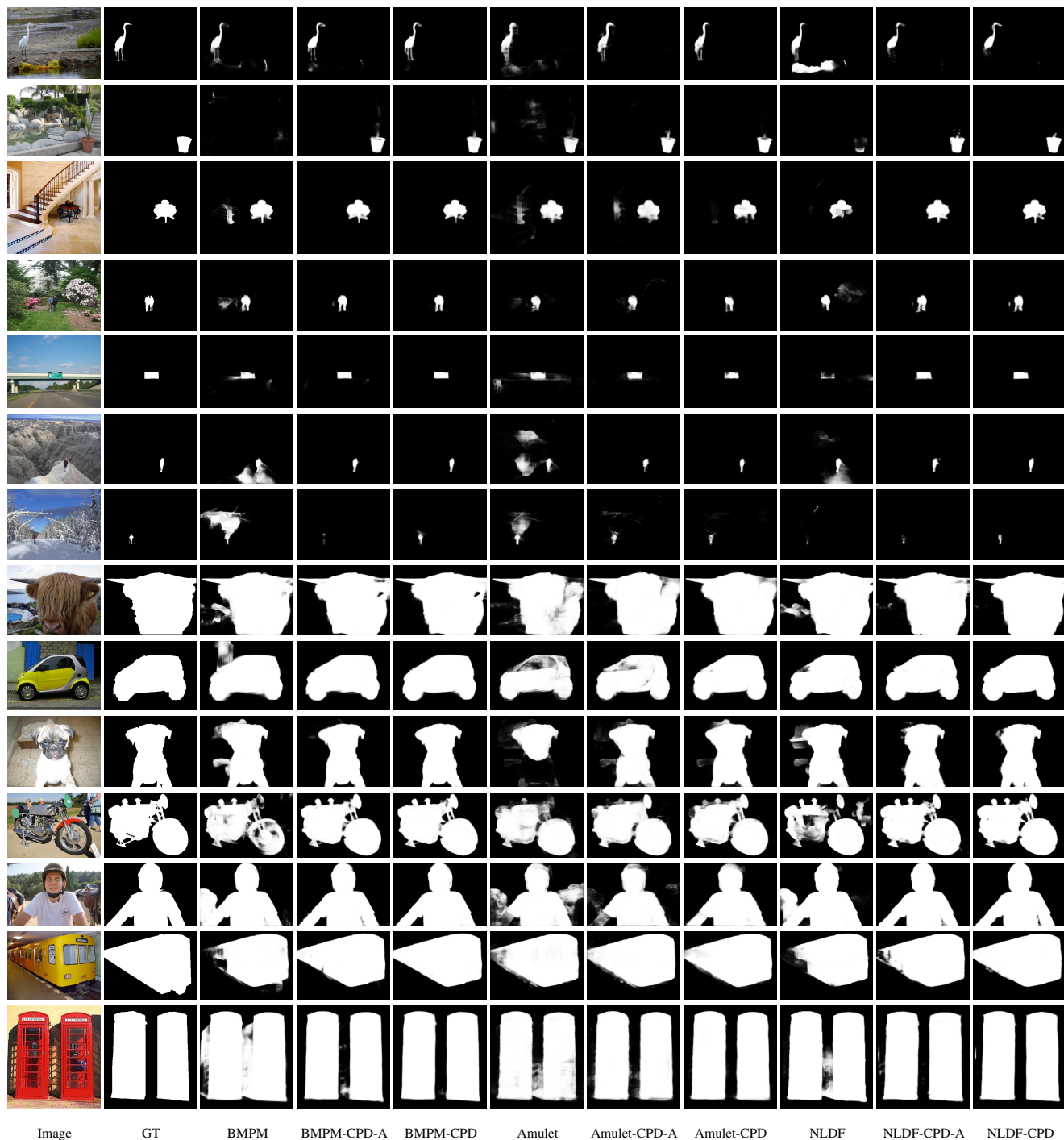


Figure 11: Visual comparisons of original models (BMPM, Amulet, NLDF) with optimized models (-CPD-A, -CPD) by the proposed framework in challenging cases: small object (top seven rows) and large object (bottom seven rows).



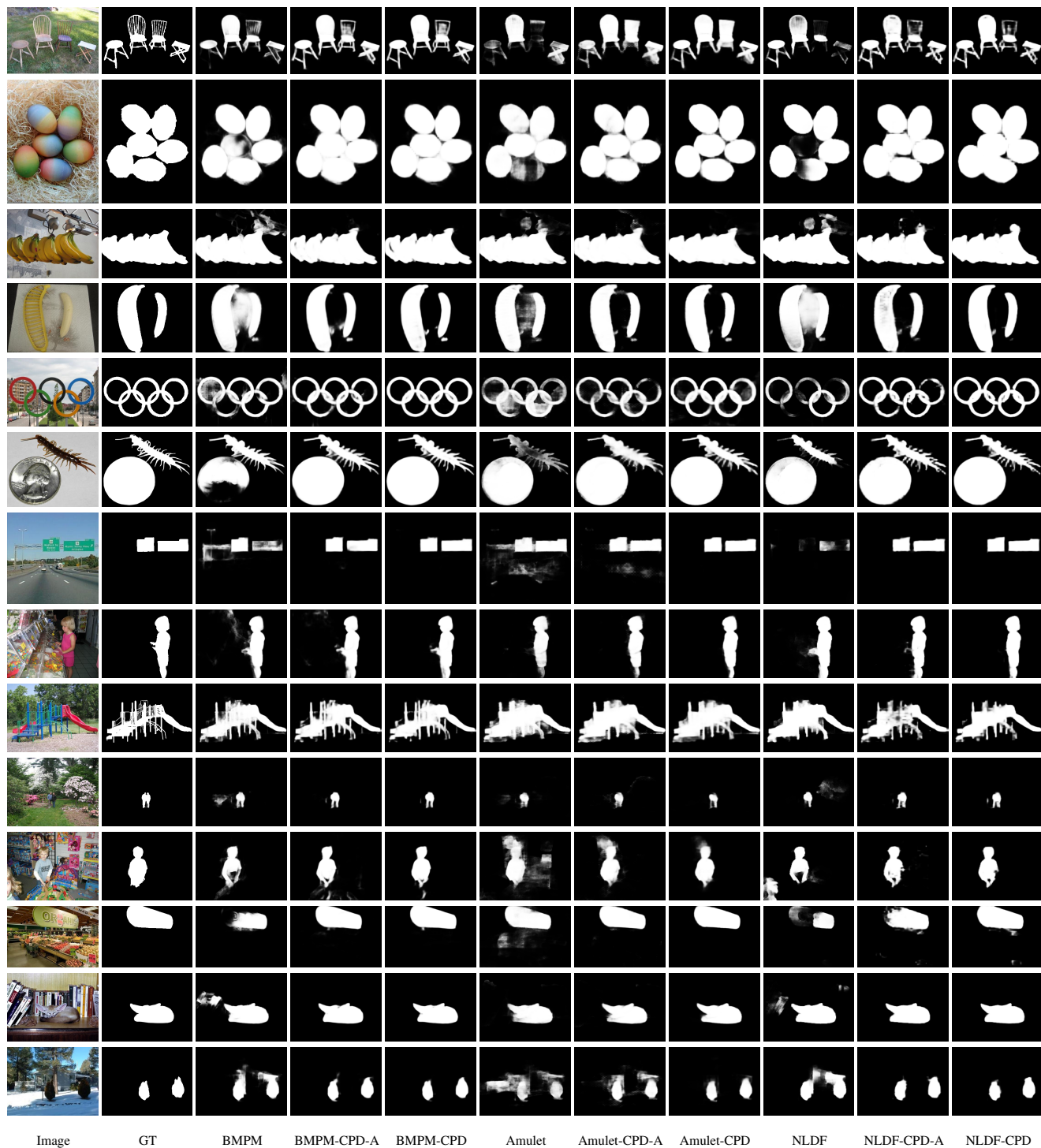


Figure 12: Visual comparisons of original models (BMPM, Amulet, NLDF) with optimized models (-CPD-A, -CPD) by the proposed framework in challenging cases: multiple objects (top seven rows) and complex scenes (bottom seven rows).

## References

- [1] M.-M. Cheng, N. J. Mitra, X. Huang, P. H. Torr, and S.-M. Hu. Global contrast based salient region detection. *IEEE TPAMI*, 37(3):569–582, 2015.
- [2] Q. Hou, M.-M. Cheng, X. Hu, A. Borji, Z. Tu, and P. Torr. Deeply supervised salient object detection with short connections. *IEEE TPAMI*, 2018.
- [3] P. Krähenbühl and V. Koltun. Efficient inference in fully connected crfs with gaussian edge potentials. In *NIPS*, pages 109–117, 2011.
- [4] G. Li and Y. Yu. Visual saliency based on multiscale deep features. In *CVPR*, pages 5455–5463, 2015.
- [5] Y. Li, X. Hou, C. Koch, J. M. Rehg, and A. L. Yuille. The secrets of salient object segmentation. In *CVPR*, pages 280–287, 2014.
- [6] T. Liu, Z. Yuan, J. Sun, J. Wang, N. Zheng, X. Tang, and H.-Y. Shum. Learning to detect a salient object. *IEEE TPAMI*, 33(2):353–367, 2011.
- [7] Z. Luo, A. K. Mishra, A. Achkar, J. A. Eichel, S. Li, and P.-M. Jodoin. Non-local deep features for salient object detection. In *CVPR*, volume 2, page 7.
- [8] L. Wang, H. Lu, Y. Wang, M. Feng, D. Wang, B. Yin, and X. Ruan. Learning to detect salient objects with image-level supervision. In *CVPR*, pages 136–145, 2017.
- [9] Q. Yan, L. Xu, J. Shi, and J. Jia. Hierarchical saliency detection. In *CVPR*, pages 1155–1162, 2013.
- [10] C. Yang, L. Zhang, H. Lu, X. Ruan, and M.-H. Yang. Saliency detection via graph-based manifold ranking. In *CVPR*, pages 3166–3173, 2013.
- [11] P. Zhang, D. Wang, H. Lu, H. Wang, and X. Ruan. Amulet: Aggregating multi-level convolutional features for salient object detection. In *CVPR*, pages 202–211, 2017.