

BMVC
2022

The 33rd British Machine Vision Conference
21st - 24th November 2022, London, UK

542

RORD: A Real-world Object Removal Dataset

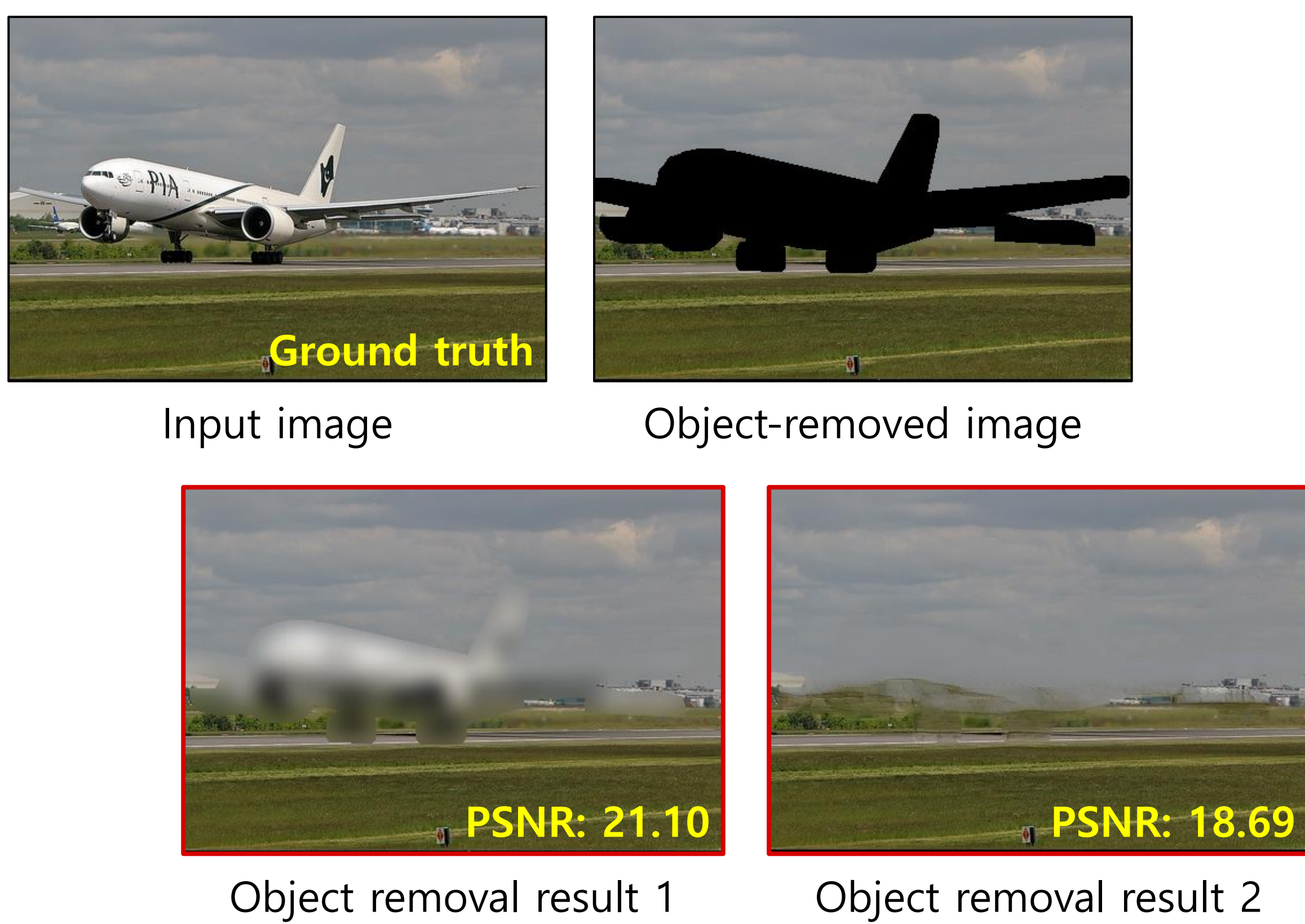
Min-Cheol Sagong, Yoon-Jae Yeo, Seung-Won Jung, and Sung-Jea Ko

Korea University, Seoul, Korea

<https://github.com/Forty-lock/RORD>

1. Introduction

- Toy example for object removal



Which one is BETTER ?

3. Evaluation

- Quantitative results
 - Trained on MS-COCO, Places2, or RORD
 - Evaluated on PASCAL-VOC, MS-COCO, or RORD

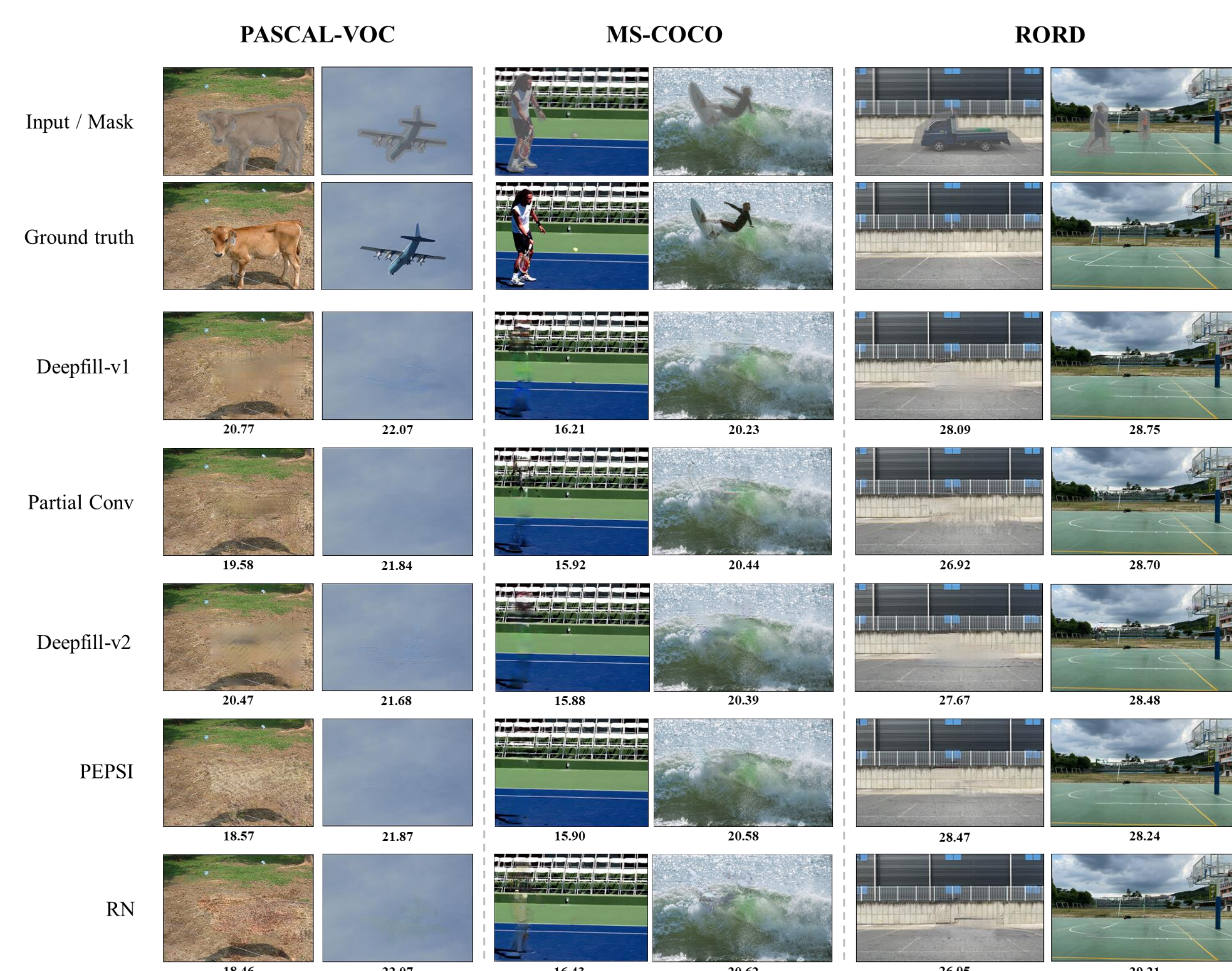
Cross-validation test

| | | MS-COCO | | Places2 | | RORD | |
|--------------|------------|---------|-------|---------|-------|--------------|--------------|
| | | PSNR | SSIM | PSNR | SSIM | PSNR | SSIM |
| Deepfill-v1 | PASCAL-VOC | 17.01 | 0.774 | 17.10 | 0.777 | 16.87 | 0.773 |
| | MS-COCO | 19.17 | 0.832 | 19.25 | 0.835 | 19.07 | 0.831 |
| | RORD | 24.43 | 0.864 | 24.49 | 0.868 | 24.81 | 0.869 |
| Partial Conv | PASCAL-VOC | 16.59 | 0.745 | 17.03 | 0.770 | 16.78 | 0.755 |
| | MS-COCO | 18.81 | 0.812 | 19.24 | 0.830 | 18.81 | 0.814 |
| | RORD | 24.00 | 0.854 | 24.39 | 0.860 | 24.65 | 0.861 |
| Deepfill-v2 | PASCAL-VOC | 16.88 | 0.774 | 17.01 | 0.776 | 16.90 | 0.770 |
| | MS-COCO | 19.08 | 0.832 | 19.10 | 0.834 | 19.08 | 0.828 |
| | RORD | 24.46 | 0.865 | 24.43 | 0.868 | 24.80 | 0.866 |
| PEPSI | PASCAL-VOC | 16.81 | 0.772 | 17.03 | 0.781 | 16.69 | 0.767 |
| | MS-COCO | 19.06 | 0.831 | 19.23 | 0.838 | 18.96 | 0.828 |
| | RORD | 24.45 | 0.865 | 24.80 | 0.838 | 24.87 | 0.868 |
| RN | PASCAL-VOC | 16.31 | 0.759 | 16.44 | 0.744 | 16.54 | 0.757 |
| | MS-COCO | 18.53 | 0.823 | 18.59 | 0.813 | 18.68 | 0.820 |
| | RORD | 23.32 | 0.856 | 22.26 | 0.831 | 23.91 | 0.858 |

- Qualitative results

- Trained with RORD

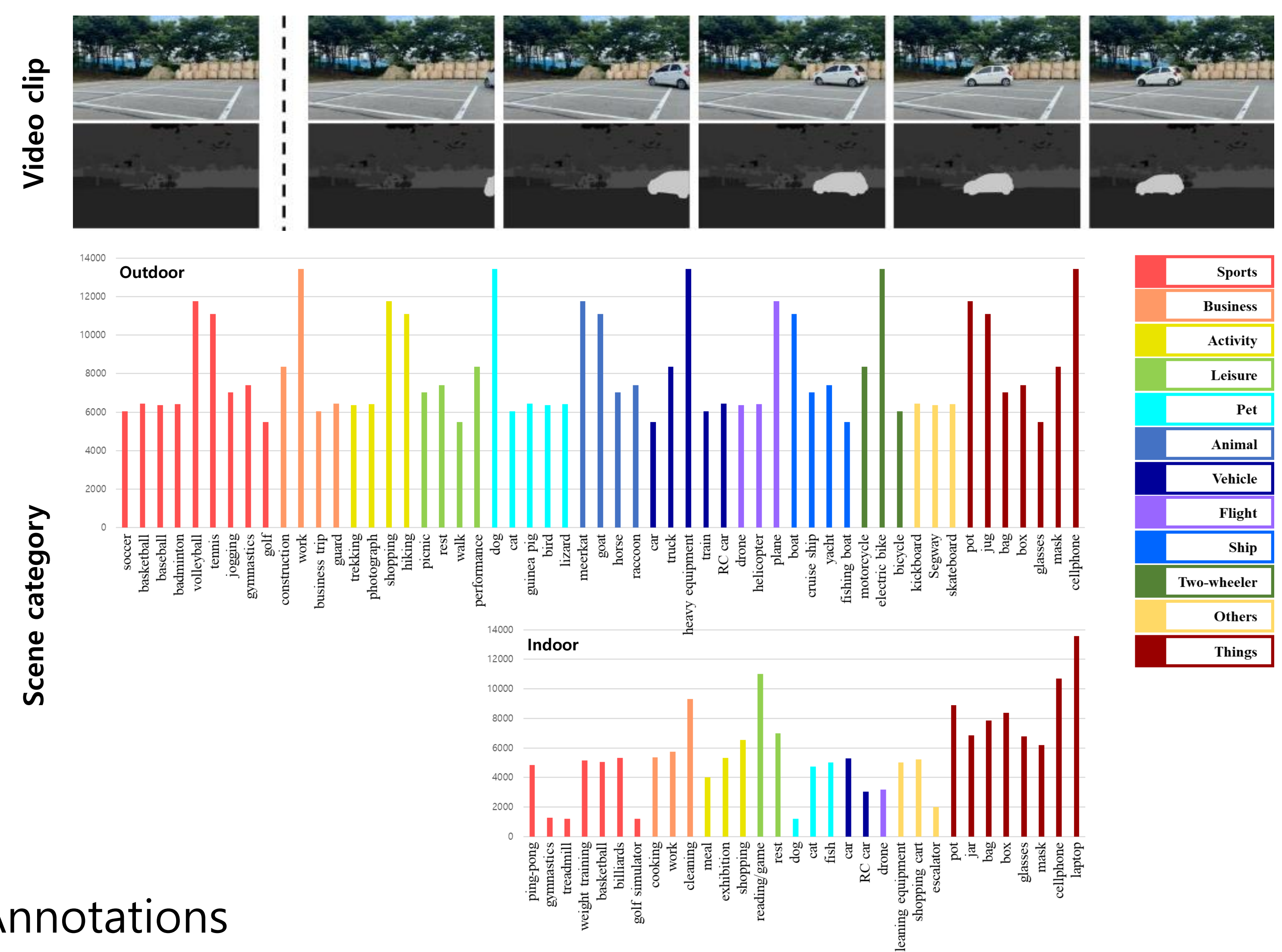
Visual comparison



2. Proposed Dataset: RORD

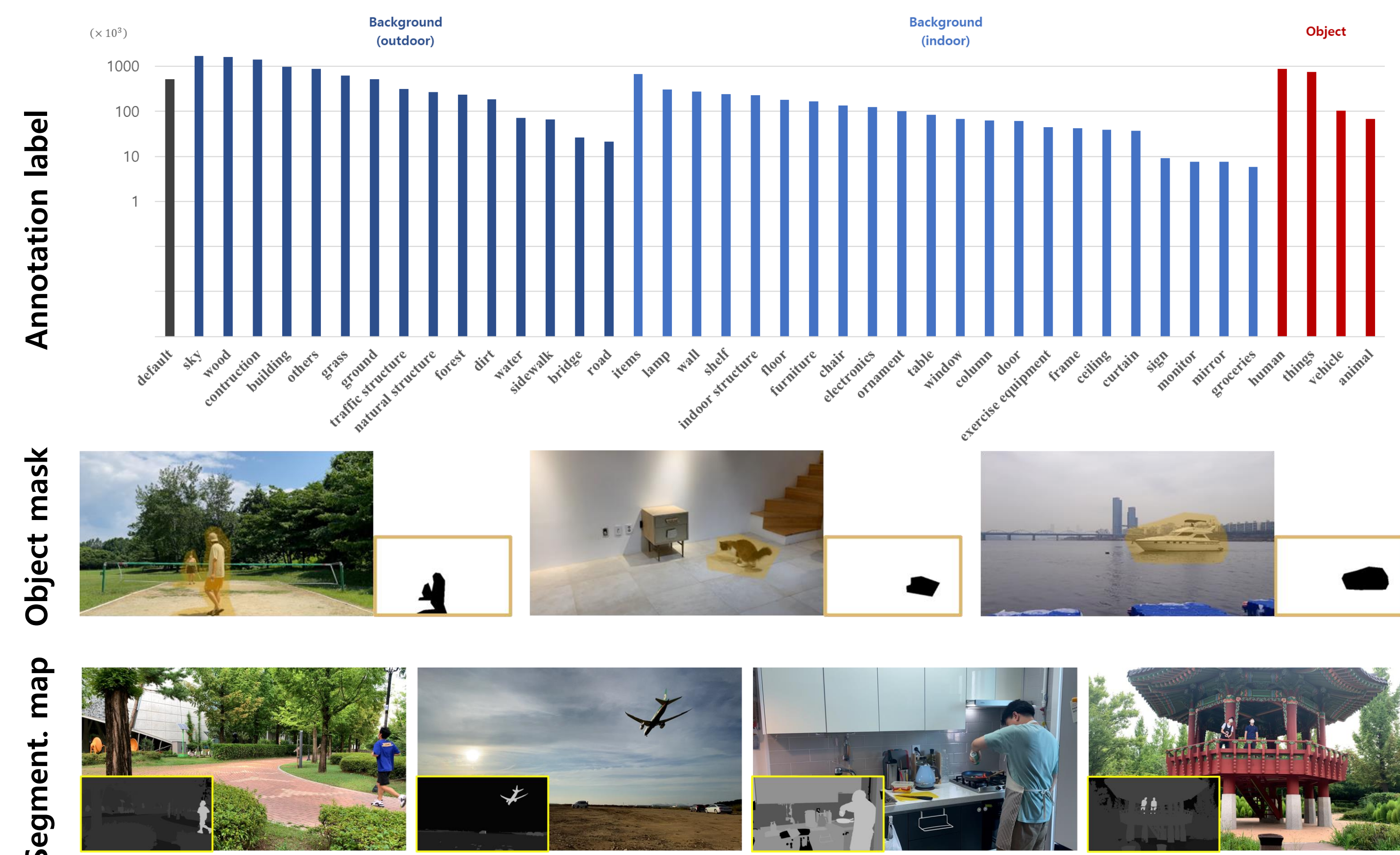
- Dataset statistics

- 516,705 images captured under 3,447 unique scenes
- HD videos (1920×1080)
- 334,275 outdoor images: 55 subcategories for 12 scene categories
- 182,329 indoor images: 32 subcategories for 9 scene categories



- Annotations

- Binary object mask (object and background)
- 42 semantic classes



4. Conclusion

- We release a Real-world Object Removal Dataset (RORD), containing a pair of images with and without objects.
- RORD supports numerous images captured under a wide variety of indoor and outdoor real-world environments.
- Two types of pixel-accurate annotations, *i.e.*, object masks and segmentation maps, are provided.