360-Indoor: Towards Learning Real-World Objects in 360° Indoor Equirectangular Images

Shih-Han Chou¹, Cheng Sun¹, Wen-Yen Chang¹, Wan-Ting Hsu^{1,*}, Min Sun¹, Jianlong Fu² ¹National Tsing Hua University, Hsinchu² Microsoft Research, Beijing

In supplementary, We show more details about our dataset.

- Section 1: We show each category in 360-Indoor (Figure 1 to Figure 37).
- Section 2: We show the distribution of object categories in 360-Indoor.
- Section 3: The experiment results mAP by categories is provided.

1. Categories in 360-Indoor

For each category, we provide four images to illustrate the category annotations in 360-Indoor. The red boxes in the imagea are the Bounding Field of Views (BFoVs).



Figure 3: Bathtub



Figure 1: Air conditioner



Figure 4: Bed



Figure 2: Backpack



Figure 5: Board



Figure 6: Book



Figure 10: Chair



Figure 7: Bottle



Figure 11: Clock



Figure 8: Bowl



Figure 12: Computer



Figure 9: Cabinet



Figure 13: Cup



Figure 14: Door



Figure 18: Keyboard



Figure 15: Fan



Figure 19: Light



Figure 16: Fireplace



Figure 20: Microwave



Figure 17: Heater



Figure 21: Mirror



Figure 22: Mouse



Figure 26: Picture



Figure 23: Oven



Figure 27: Potted plant



Figure 24: Person



Figure 28: Refrigerator



Figure 25: Phone



Figure 29: Sink



Figure 30: Sofa



Figure 34: Vase



Figure 31: Table



Figure 35: Washer



Figure 32: Toilet



Figure 36: Window



Figure 33: TV



Figure 37: Wine glass

2. Distribution of object categories in 360-Indoor

In this section, we show the details of the object categories in the 360-Indoor in Figure 38.

3. Experiment results mAP by categories.

In this section, we show the detail results of the best model in each baselines proposed in Table 4 and Table 6 in the main paper. Among these category mAPs, the objects mostly in the middle of the images have better mAPs. For example, 'bed', 'bathtub', 'sink', 'table' are mostly in the vertical middle of the 360° image. On the other hand, the small objects, such as 'phone', 'backpack', 'wine', 'bowl', 'cup', 'bottle', 'mouse', do not perform well in these baselines. We regard that they are relatively small in the 360° images so that it is hard for model to detect them. Compare with Faster-RCNN and Faster R-CNN (SphereNet), the significant differences are small objects. In Faster R-CNN (SphereNet), the mAPs of small objects drops more severely than Faster R-CNN. Hence, we believe that 360° domain still has many unsolved issues and we hope the proposed 360-Indoor dataset would advantage this field.



Figure 38: Distribution of object categories in 360-Indoor.

| Model | mAP (%) | Category mAP (%) | | | | | | | | | | |
|-----------------------------|---------|-------------------|----------------|----------------------|-----------------|-----------------|-------------------------|-----------------|----------------|-------------------|--------------|------------------|
| YOLOv3 | 24.5 | toilet 10.9 | board 10.2 | mirror 23.2 | bed 33.8 | potted 24.1 | book 11.8 | clock 27.8 | phone 3.3 | keyboard 23.4 | tv 36.1 | fan 37.3 |
| | | backpack 6.2 | light 22.1 | refrigerator 28.4 | bathtub 22.3 | wine 3.6 | air conditioner 37.4 | cabinet 24.6 | sofa 39.2 | bowl 7.7 | sink 43.4 | 24.3 |
| | | cup 7.8 | bottle 9.2 | washer 41.6 | chair 37.1 | picture 44.2 | window 36.9 | door 30.2 | heater 11.3 | fireplace 31.9 | mouse 4.6 | oven 26.9 |
| | | microwave 29.3 | person 48.3 | vase 10.9 | table 32.2 | | | | | | | |
| Faster R-CNN | 30.2 | toilet 75.4 | board 9.4 | mirror 32.7 | bed 57.7 | potted 31.5 | book 6.5 | clock 13.6 | phone 3.6 | keyboard 7.2 | tv 53.8 | fan 46.7 |
| | | backpack 2.4 | light 25.5 | refrigerator 35.1 | bathtub 55.8 | wine 0.8 | air conditioner 41.0 | cabinet 37.0 | sofa 58.3 | bowl 1.7 | sink 48.6 | computer 20.4 |
| | | cup 4.7 | bottle 5.8 | washer 37.7 | chair 37.1 | picture 47.7 | window 49.8 | door 42.6 | heater 22.3 | fireplace 39.2 | mouse 0.3 | oven 33.0 |
| | | microwave 29.7 | person 52.8 | vase 3.0 | table 46.0 | | | | | | | |
| FPN | 33.6 | toilet 69.4 | board 18.1 | mirror 39.6 | bed 61.5 | potted 29.4 | book 11.6 | clock 20.2 | phone 4.3 | keyboard 14.8 | tv 56.8 | fan 56.2 |
| | | backpack 4.8 | light 31.2 | refrigerator 41.7 | bathtub 56.5 | wine 0.7 | air conditioner 34.8 | cabinet 42.6 | sofa 54.4 | bowl 6.9 | sink 48.6 | computer 21.6 |
| | | cup 5.7 | bottle 7.5 | washer 44.4 | chair 43.0 | picture 50.3 | window 50.5 | door 43.5 | heater 22.5 | fireplace 50.6 | mouse 6.3 | oven 34.2 |
| | | microwave 35.8 | person 56.9 | vase 10.7 | table 54.5 | | | | | | | |
| Faster R-CNN (SphereNet) | 21.7 | toilet 62.8 | board 3.8 | mirror 26.2 | bed 44.7 | potted 14.4 | book 2.5 | clock 10.2 | phone 0.8 | keyboard 1.7 | tv 36.1 | fan 44.3 |
| | | backpack 0.2 | light 23.3 | refrigerator 15.6 | bathtub 42.2 | wine 0.03 | air conditioner 18.8 | cabinet 30.0 | sofa 44.2 | bowl 2.6 | sink 29.9 | computer 15.2 |
| | | cup 1.2 | bottle 3.5 | washer 40.5 | chair 28.0 | picture 32.0 | window 40.1 | door 29.5 | heater 6.8 | fireplace 31.6 | mouse 0.1 | oven 23.1 |
| | | microwave 11.8 | person 41.1 | vase 1.1 | table 41.7 | | | | | | | |

Table 1: Detailed detection results on the 360-Indoor test set.