

Scene	ResNet50 (%)	OTS (%)
Balcony interior	87	89
Bathroom	88	88
Bedroom	87	91.5
Closet	87	88
Dining room	66	82
Garage indoor	86	90
Home office	70	81
Home theater	83	88
Kitchen	70	83
Laundromat	90	92
Living room	60	72
Playroom	86	89
Staircase	93	91
Wet bar	60	73
Total	80.0	85.9

Table 1. Comparison of the proposed OTS and ResNet50 on the Places365-14classes.

Scene	ResNet50 (%)	[1] (%)	OTS (%)
Bathroom	90	91	93
Bedroom	83	90	98
Corridor	93	96	95
Dining room	73	79	88
Kitchen	84	87	92
Living room	71	80	80
Office	88	94	91
Total	83.1	88.1	91.0

Table 2. Comparison of OTS, DEDUCE and ResNet50 on the Places365-7classes.

Dataset	Backbone	FA	POTO	GRA	Accuracy (%)
Places14	ResNet50	✓			76.6
Places14	ResNet50	✓	✓		82.0
Places14	ResNet50	✓	✓	✓	85.9

Table 3. Ablation studies on the feature aggregation (FA) module, portable object-to-object (POTO) module, and global relation aggregation (GRA) module.

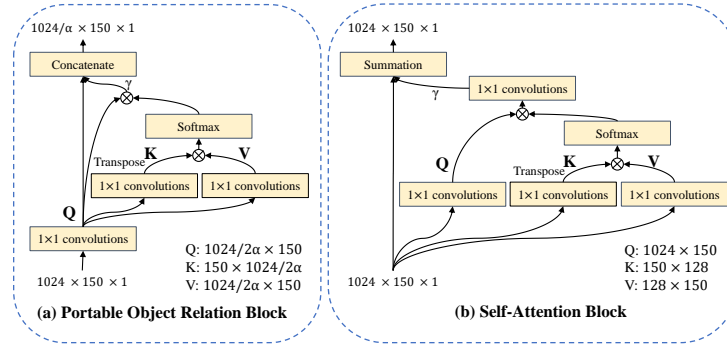


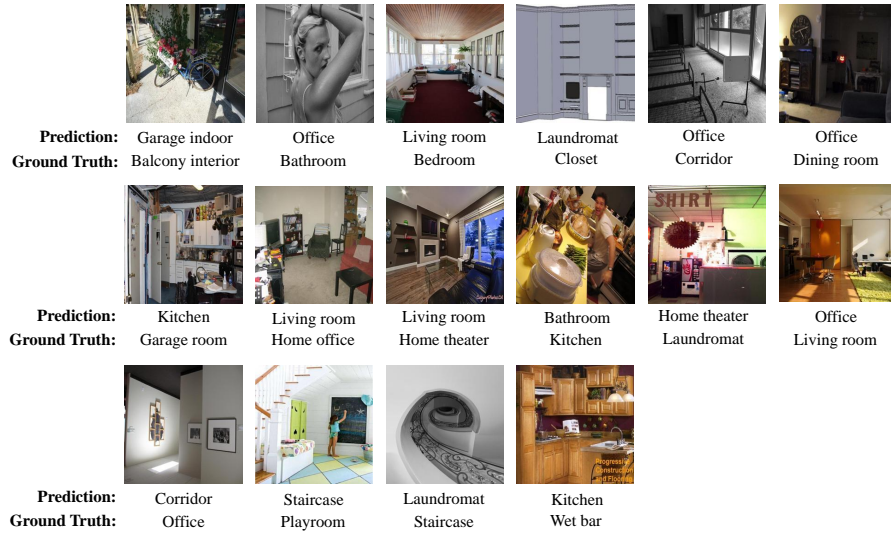
Fig. 1. Structure of the proposed portable object relation block and self-attention block.

References

1. Pal, A., et al.: DEDUCE: Diverse scEnE Detection methods in Unseen Challenging Environments. IROS, 4198–4204, 2019

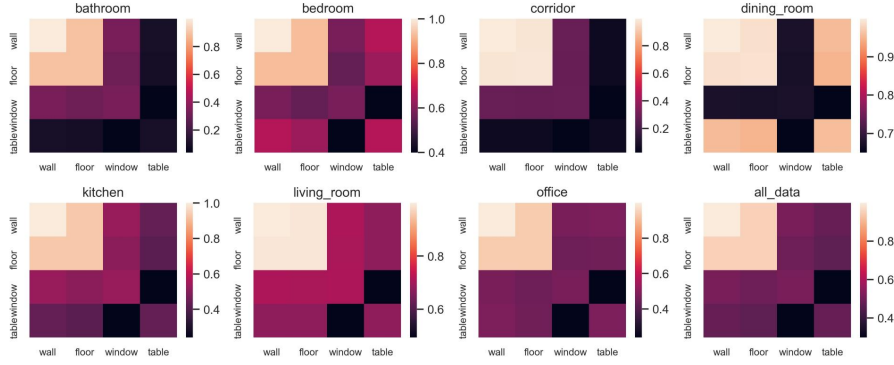


(a) Images that ResNet50 misclassified but OTS correctly classified.

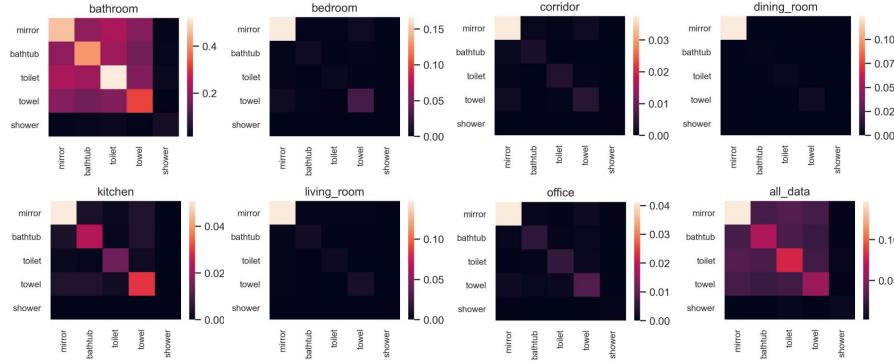


(b) Images that OTS misclassified but ResNet50 correctly classified.

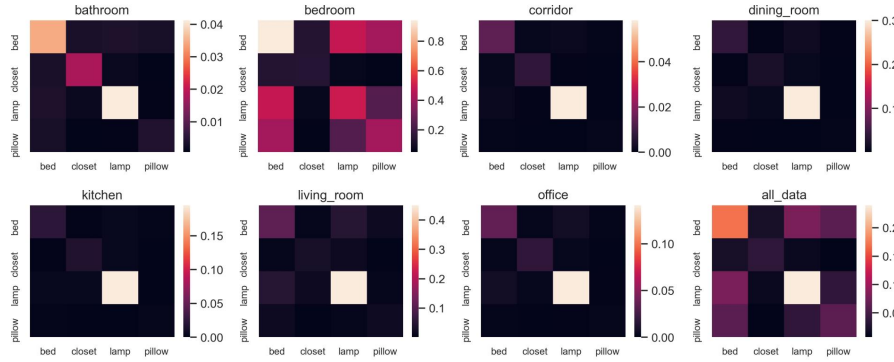
Fig. 2. Bad case of OTS and ResNet50. (a) Images that ResNet50 misclassified but OTS correctly classified. (b) Images that OTS misclassified but ResNet50 correctly classified.



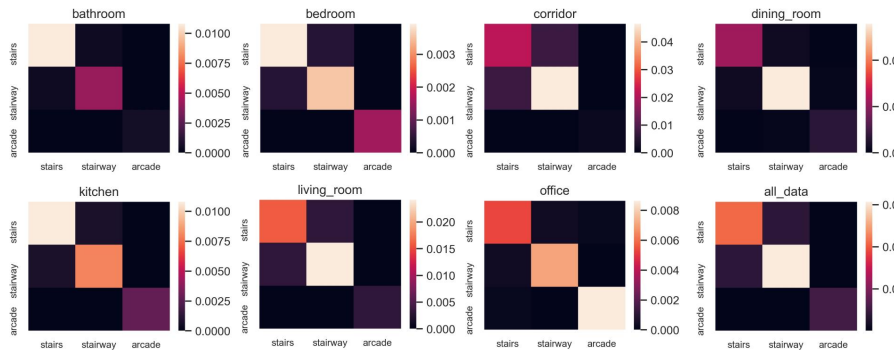
(a) Most frequently appeared objects/object pairs in all scenes



(b) Most frequently appeared objects/object pairs in bathroom

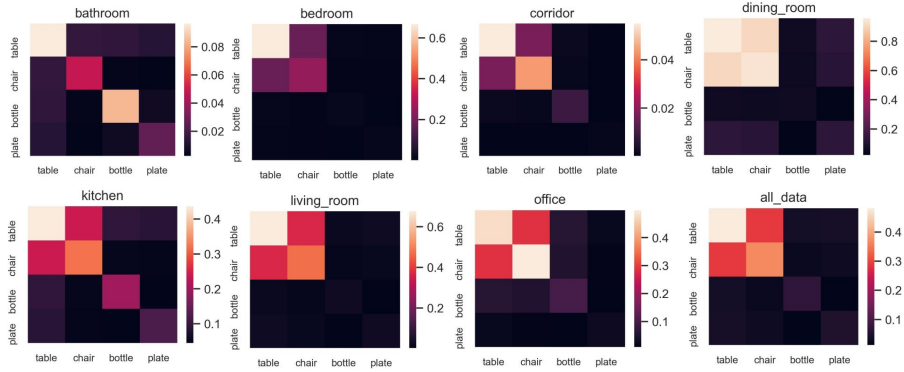


(c) Most frequently appeared objects/object pairs in bedroom

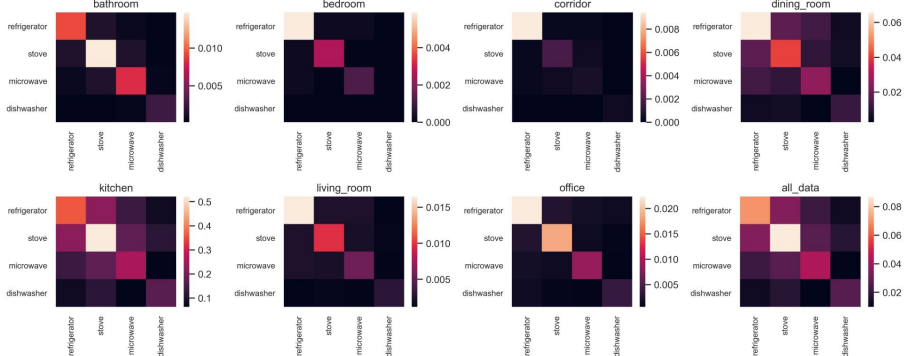


(d) Most frequently appeared objects/object pairs in corridor

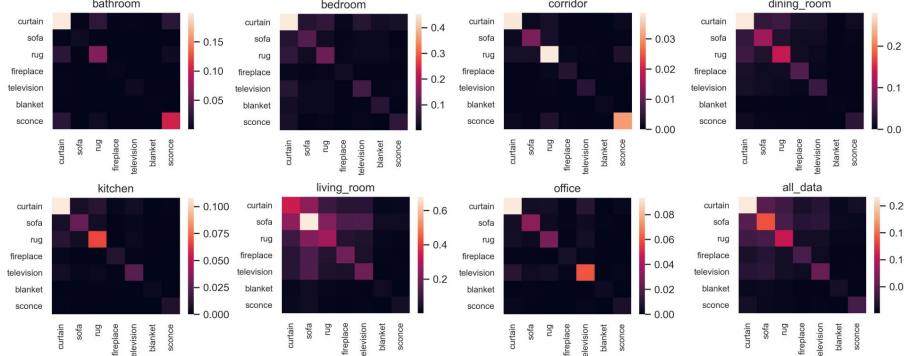
Fig. 3. Frequency of each object/object pairs in different Scenes Part 1. The frequency is calculated by diving number of occurrences by the number of images.



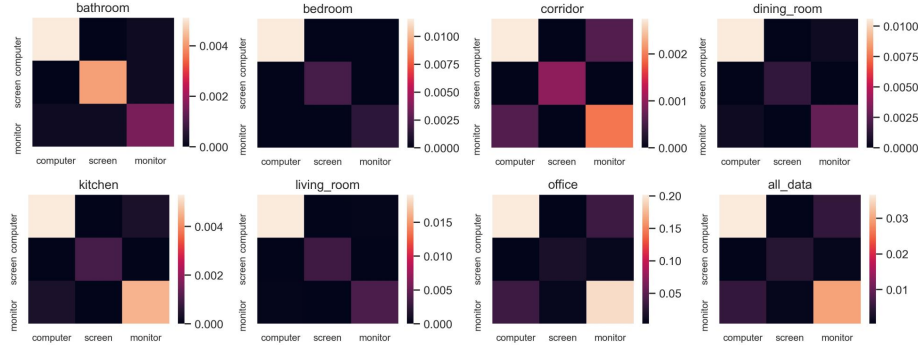
(e) Most frequently appeared objects/object pairs in dining room



(f) Most frequently appeared objects/object pairs in kitchen



(g) Most frequently appeared objects/object pairs in living room



(h) Most frequently appeared objects/object pairs in office

Fig. 4. Frequency of each object/object pairs in different Scenes Part 2. The frequency is calculated by diving number of occurrences by the number of images.