Appendix

Fig. 1 visualizes PD-Net's advantage in deciphering the polysemy problem of verbs on V-COCO. The performance gain by PD-Net compared with SH-VCM reaches 3.17% in AP for the "carry backpack" category.



Fig. 1. Visualization of PD-Net's advantage in deciphering the verb polysemy problem. We randomly select three verbs affected by the polysemy problem: "carry" (top row), "ride" (middle row), and "hold" (bottom row). The green number and red number denote the AP of SH-VCM and PD-Net respectively for the same HOI category.