

A APPENDIX

Table 6: Results from tests on Adv.{RR, RPD, RPS, AR, APD, APS} according to the experimental sets in Table 3. Every four rows in the data correspond to a group of experiments. According to adversarial attack iteration, it is divided into three groups in 9 columns, representing 1, 50, 100 respectively. The bold parts in the table indicate the attack success rate corresponding to different L_∞ constraints under the same Adv Iter and the same Case.

		Adv Iter = 1			Adv Iter = 50			Adv Iter = 100		
		10	20	30	10	20	30	10	20	30
Adv.RR	Case I	88.18	88.18	90.15	93.60	94.09	95.57	94.09	95.57	96.06
	Case II	83.74	83.74	84.73	92.61	92.61	93.10	93.10	93.10	94.09
	Case III	56.16	54.68	49.26	76.35	72.41	75.37	76.85	74.88	76.35
	Case IV	55.17	52.71	47.78	76.35	69.95	69.46	76.35	71.92	70.44
Adv.RPD	Case I	87.62	89.16	89.66	93.07	96.06	95.57	93.56	96.55	95.57
	Case II	83.17	84.24	82.76	91.58	93.60	93.10	92.08	94.09	93.60
	Case III	62.87	51.72	52.22	77.23	73.40	71.92	77.72	75.37	74.88
	Case IV	61.88	50.25	50.25	75.74	69.95	69.46	76.24	72.41	71.92
Adv.RPS	Case I	87.68	89.66	88.18	93.10	94.58	95.07	94.09	96.06	95.57
	Case II	83.74	85.22	82.76	91.63	92.61	93.10	93.10	94.09	94.09
	Case III	57.64	55.17	49.26	74.38	71.92	73.89	76.35	74.38	76.85
	Case IV	56.65	52.71	47.78	73.89	70.44	70.44	74.88	71.92	72.91
Adv.AR	Case I	87.68	88.67	89.16	92.61	96.06	94.58	93.10	97.04	96.06
	Case II	84.24	82.76	83.25	91.13	93.60	92.61	91.63	94.58	94.09
	Case III	58.62	50.25	50.25	74.38	73.89	73.89	74.38	74.88	76.35
	Case IV	56.65	47.78	49.26	72.91	71.92	69.95	72.91	72.91	71.92
Adv.APD	Case I	87.19	89.16	88.67	91.63	95.07	95.57	93.60	95.07	95.57
	Case II	83.25	83.25	82.76	90.15	93.60	93.10	92.61	93.60	93.10
	Case III	59.11	52.71	48.28	74.88	73.40	74.38	76.35	76.85	74.88
	Case IV	57.64	50.25	45.81	73.89	70.44	69.95	75.37	72.41	70.94
Adv.APS	Case I	86.70	88.67	91.13	92.61	95.07	95.07	92.61	96.06	96.06
	Case II	81.77	83.25	84.24	90.64	92.61	93.10	90.64	94.09	94.58
	Case III	55.67	51.72	50.74	75.37	74.38	72.91	75.86	75.86	74.88
	Case IV	54.19	49.26	48.77	74.88	69.95	68.47	75.37	70.94	69.95

Table 7: Ablation study results, with $L_\infty \leq 10$ as the constraint. Each row in the table represents the results of the adversarial attack iteration from 10 to 100 under the same Adv.setting. The bolded parts are the items with the highest attack success rate among different Adv. at the same adversarial attack iteration.

	1	10	20	30	40	50	60	70	80	90	100
Adv.RR	18.67	45.57	52.85	56.33	60.13	62.66	62.97	63.92	64.56	64.87	65.19
Adv.RPD	19.62	46.20	54.43	57.28	59.18	60.44	62.66	63.92	64.24	64.24	65.51
Adv.RPS	17.72	42.41	51.58	56.33	58.23	58.86	61.08	62.66	63.92	65.19	66.46
Adv.AR	17.09	40.51	47.78	54.75	61.08	62.66	64.56	65.19	65.82	66.14	66.14
Adv.APD	18.67	45.25	52.53	56.33	58.54	59.49	62.03	64.56	64.87	66.46	67.41
Adv.APS	18.35	42.72	49.68	52.53	56.33	59.18	60.44	62.34	62.97	63.61	65.82

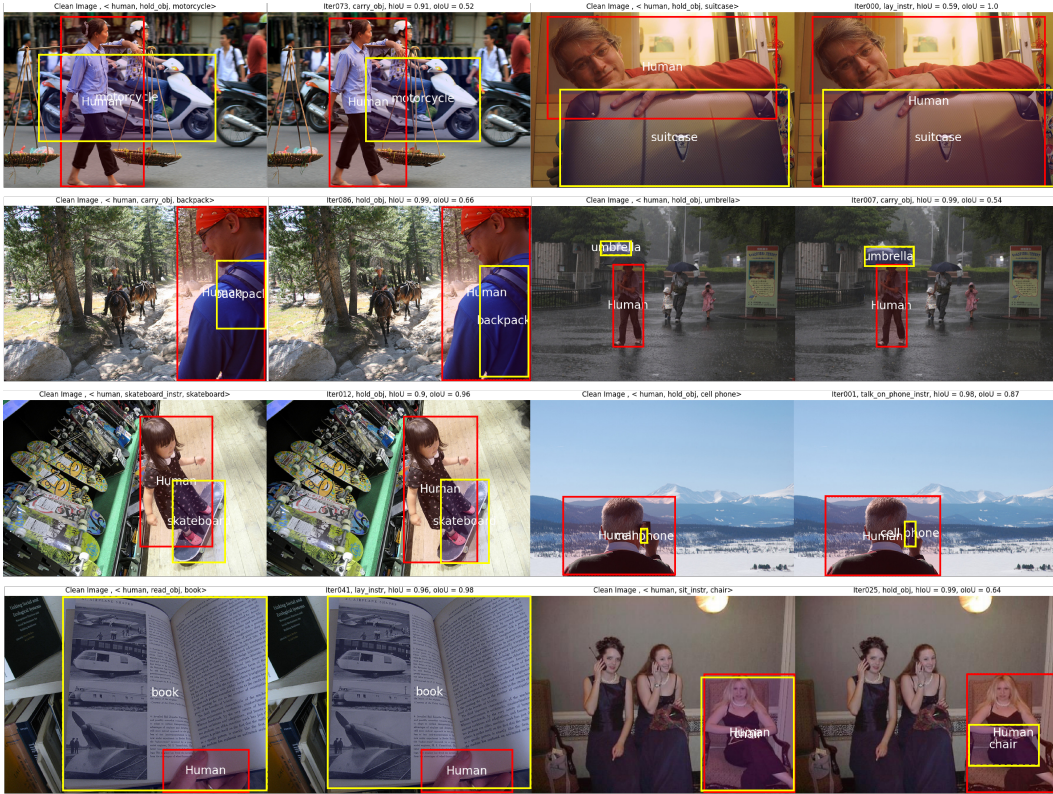


Figure 3: Here is the visualization of the attack results. Each group comprises two images: the clean image on the left and the adversarial image on the right. The title of each image includes the original prediction result, the IoU before and after the attack, and the interaction after the attack.

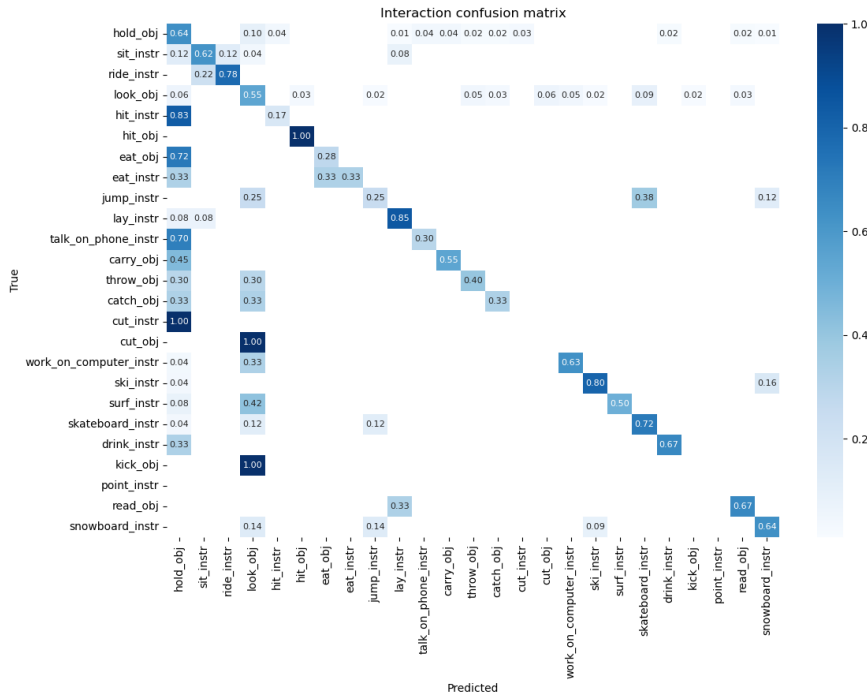


Figure 4: The confusion matrix analysis for interactions. The vertical axis represents the original interactions, while the horizontal axis represents the predicted interactions.

