

	Phrase Detection	Relationship Detection	Predicate Detection
Visual Phrases	0.03	-	0.71
Joint CNN	0.05	0.04	1.02
Ours - V only	0.93	0.84	6.42
Ours - L only	0.08	0.08	8.94
Ours - V + naive FC	1.21	1.19	11.05
Ours - V + L only	1.74	1.32	16.31
Ours - V + L + Reg.	1.78	1.40	17.95
Ours - V + L + K	2.07	1.52	29.47