Biased Experiment

Analyze attack trends

Table 1: We run a logistic model regressing success against log(number of factors) in the randomized attack experiment. Success rates increase with the number of factors combined to select target and perturb objects for all models and attacks. Table headers are explained in Appendix **??**.

Group	Regression							
Attack	term	sig	estimate	std.error	statistic	p.value	conf.low	conf.high
YOLOv3								
Vanishing	num_cri	*	1.144	0.077	14.871	0	0.996	1.298
Mislabeling	num_cri	*	1.179	0.078	15.094	0	1.029	1.335
Untargeted	num_cri	*	1.007	0.073	13.700	0	0.865	1.153
SSD								
Vanishing	num_cri	*	0.749	0.065	11.549	0	0.624	0.878
Mislabeling	num_cri	*	0.684	0.064	10.752	0	0.561	0.810
Untargeted	num_cri	*	0.678	0.065	10.497	0	0.552	0.806
RetinaNet								
Vanishing	num_cri	*	0.546	0.086	6.315	0	0.378	0.717
Mislabeling	num_cri	*	0.586	0.126	4.657	0	0.342	0.836
Untargeted	num_cri	*	0.951	0.071	13.302	0	0.813	1.093
Faster R-CNN								
Vanishing	num_cri	*	0.558	0.088	6.319	0	0.387	0.733
Mislabeling	num_cri	*	0.771	0.107	7.202	0	0.564	0.984
Untargeted	num_cri	*	1.228	0.077	16.021	0	1.080	1.381
Cascade R-CNN								
Vanishing	num_cri	*	0.694	0.078	8.847	0	0.542	0.849
Mislabeling	num_cri	*	0.765	0.089	8.623	0	0.594	0.942
Untargeted	num_cri	*	0.948	0.075	12.714	0	0.804	1.096

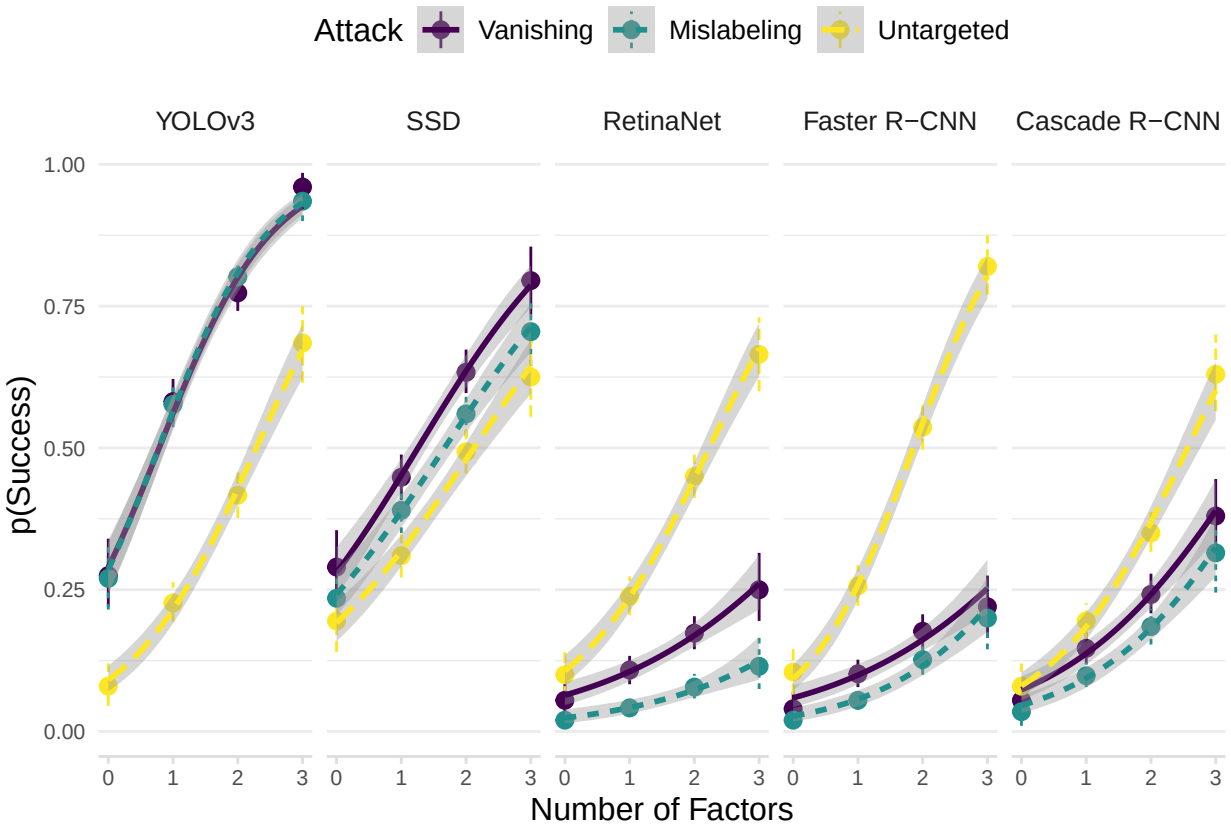

Figure 1: Success factors can be exploited in combination to significantly increase success rates We sampled target and perturb objects based on three validated success factors in Table **??** by targeting objects with low predicted confidence, perturbing large objects and selecting target and perturb objects close to one another. The binned summaries and regression trendlines graph success proportion against number of factors in the deliberate attack experiment. Errors are 95% confidence intervals and every point aggregates success over 200 images. Success rates significantly increase as the number of factors combined increases. Significance is determined at $\alpha < 0.05$ using a Wald z-test on the logistic estimates. Full details are given in Section **??**.