

	LVIS	ObjNet	COCO	BDD	Avg.	
CLIP alone	.63	.64	.90	.74	.72	all queries
+ multiscale representation	.70	.64	.95	.76	.76	
+ logistic reg. on feedback	.67	.59	.87	.68	.70	
+ CLIP alignment objective	.75	.69	.96	.77	.79	
+ DB alignment objective	.76	.70	.96	.79	.80	
CLIP alone	.19	.28	.27	.02	.19	hard subset
+ multiscale representation	.32	.28	.58	.10	.32	
+ logistic reg. on feedback	.34	.28	.57	.07	.31	
+ CLIP alignment objective	.42	.39	.74	.20	.44	
+ DB alignment objective	.44	.40	.75	.24	.46	