Model	Threshold		Score	
	Sound	Object	Temporal (AUC)	Spatial (pixel)
Center	NA	NA	NA	37.2
Sound model	NA	NA	0.881	NA
Object model	NA	NA	0.712	33.3
Video tag as target (VT)	NA	NA	0.790	33.9
Sound as target (ST01)	0.1	NA	0.824	32.7
Sound as target (ST03)	0.3	NA	0.827	32.2
Sound as target (ST05)	0.5	NA	0.827	33.7
Sound as target (ST07)	0.7	NA	0.830	33.0
Sound as target (ST09)	0.9	NA	0.824	33.7
Object as target (OT01)	NA	0.1	0.680	45.8
Object as target (OT03)	NA	0.3	0.809	29.4
Object as target (OT05)	NA	0.5	0.809	29.5
Object as target (OT07)	NA	0.7	0.798	29.2
Object as target (OT09)	NA	0.9	0.788	31.0
S×O as target (SOT0501)	0.5	0.1	0.828	29.4
S×O as target (SOT0503)	0.5	0.3	**0.834**	29.1
S×O as target (SOT0505)	0.5	0.5	0.827	**28.9**
S×O as target (SOT0507)	0.5	0.7	0.820	29.5
S×O as target (SOT0509)	0.5	0.9	0.803	30.7