TABLE III: Best validation AP_{50}, AP_{75}, with different levels of added noise (σ) and systematic bounding box offset (Δ) and at different sizes of training data. Baseline ($\sigma = 0$, $\Delta = 0$) in each case shown as absolute value in bold, other cases shown as percent change. Mean and standard deviation of 5 different datasets, at half-resolution: *seals*, *apples*[2] and *penguins*; at full-resolution: *scott base*, *branches*.

(a) 100% training data

		$\Delta = 0\%$	$\Delta = 4\%$	$\Delta = 8\%$	$\Delta = 16\%$	$\Delta = 32\%$
AP_{50}	$\sigma = 0\%$	$\mathbf{95.1 \pm 2.7}$	$-0.6 \pm 0.6\%$	$-0.7 \pm 0.7\%$	$-2.8 \pm 2.1\%$	$-10.1 \pm 10.6\%$
	$\sigma = 4\%$	$-0.3 \pm 0.6\%$	$-0.4 \pm 0.6\%$	$-1.3 \pm 1.4\%$	$-2.8 \pm 1.7\%$	$-10.4 \pm 10.3\%$
	$\sigma = 8\%$	$-1.1 \pm 0.8\%$	$-1.3 \pm 1.7\%$	$-2.4 \pm 1.5\%$	$-4.0 \pm 2.3\%$	$-12.6 \pm 9.8\%$
	$\sigma = 16\%$	$-6.9 \pm 4.7\%$	$-8.4 \pm 5.3\%$	$-8.0 \pm 4.4\%$	$-10.2 \pm 3.7\%$	$-23.8 \pm 11.7\%$
	$\sigma = 32\%$	$-32.6 \pm 6.8\%$	$-32.1 \pm 7.1\%$	$-36.1 \pm 4.4\%$	$-42.2 \pm 4.8\%$	$-61.5 \pm 6.6\%$
AP_{75}	$\sigma = 0\%$	$\mathbf{84.0 \pm 8.5}$	$-2.7 \pm 1.5\%$	$-9.1 \pm 3.4\%$	$-27.3 \pm 21.5\%$	$-24.1 \pm 17.0\%$
	$\sigma = 4\%$	$-2.1 \pm 1.1\%$	$-2.9 \pm 1.1\%$	$-10.5 \pm 4.0\%$	$-26.8 \pm 21.0\%$	$-24.3 \pm 15.8\%$
	$\sigma = 8\%$	$-7.5 \pm 7.7\%$	$-8.6 \pm 5.2\%$	$-17.0 \pm 7.7\%$	$-42.5 \pm 14.9\%$	$-32.7 \pm 12.7\%$
	$\sigma = 16\%$	$-23.8 \pm 16.4\%$	$-30.1 \pm 16.7\%$	$-42.4 \pm 14.3\%$	$-71.0 \pm 8.9\%$	$-63.4 \pm 11.7\%$
	$\sigma = 32\%$	$-81.3 \pm 7.2\%$	$-83.4 \pm 6.0\%$	$-86.6 \pm 3.2\%$	$-94.9 \pm 2.5\%$	$-97.0 \pm 1.3\%$

(b) 25% training data

		$\Delta = 0\%$	$\Delta = 4\%$	$\Delta = 8\%$	$\Delta = 16\%$	$\Delta = 32\%$
AP_{50}	$\sigma = 0\%$	$\mathbf{86.9 \pm 7.7}$	$-0.4 \pm 1.4\%$	$-2.0 \pm 1.7\%$	$-7.2 \pm 6.2\%$	$-18.3 \pm 16.5\%$
	$\sigma = 4\%$	$0.4 \pm 2.6\%$	$0.5 \pm 1.5\%$	$-1.6 \pm 2.5\%$	$-8.1 \pm 7.4\%$	$-19.8 \pm 14.8\%$
	$\sigma = 8\%$	$-1.2 \pm 1.3\%$	$-1.9 \pm 1.7\%$	$-3.1 \pm 2.7\%$	$-9.1 \pm 7.6\%$	$-23.1 \pm 16.7\%$
	$\sigma = 16\%$	$-9.0 \pm 7.7\%$	$-9.2 \pm 6.6\%$	$-11.9 \pm 9.1\%$	$-20.4 \pm 10.9\%$	$-38.3 \pm 12.6\%$
	$\sigma = 32\%$	$-41.3 \pm 8.6\%$	$-43.5 \pm 7.8\%$	$-46.3 \pm 8.1\%$	$-55.8 \pm 7.1\%$	$-75.6 \pm 3.8\%$
AP_{75}	$\sigma = 0\%$	$\mathbf{72.8 \pm 16.4}$	$-6.3 \pm 6.4\%$	$-18.6 \pm 9.0\%$	$-39.5 \pm 17.5\%$	$-29.2 \pm 19.2\%$
	$\sigma = 4\%$	$-6.1 \pm 5.0\%$	$-8.7 \pm 7.4\%$	$-24.7 \pm 11.6\%$	$-38.2 \pm 20.8\%$	$-34.8 \pm 17.6\%$
	$\sigma = 8\%$	$-19.1 \pm 17.8\%$	$-21.7 \pm 14.8\%$	$-33.2 \pm 14.7\%$	$-56.3 \pm 14.8\%$	$-51.1 \pm 18.9\%$
	$\sigma = 16\%$	$-39.5 \pm 24.5\%$	$-48.3 \pm 24.8\%$	$-59.3 \pm 19.5\%$	$-82.8 \pm 12.6\%$	$-76.1 \pm 15.3\%$
	$\sigma = 32\%$	$-86.9 \pm 7.9\%$	$-87.2 \pm 9.1\%$	$-92.3 \pm 5.5\%$	$-97.6 \pm 1.4\%$	$-98.7 \pm 0.5\%$

(c) 6.25% training data

		$\Delta = 0\%$	$\Delta = 4\%$	$\Delta = 8\%$	$\Delta = 16\%$	$\Delta = 32\%$
AP_{50}	$\sigma = 0\%$	$\mathbf{75.1 \pm 23.2}$	$-3.9 \pm 7.7\%$	$-4.2 \pm 6.5\%$	$-15.4 \pm 11.8\%$	$-31.7 \pm 19.6\%$
	$\sigma = 4\%$	$-6.7 \pm 9.9\%$	$-2.1 \pm 3.2\%$	$-4.0 \pm 2.0\%$	$-15.7 \pm 14.1\%$	$-31.6 \pm 19.8\%$
	$\sigma = 8\%$	$-2.2 \pm 2.7\%$	$-4.6 \pm 6.1\%$	$-8.8 \pm 9.1\%$	$-18.9 \pm 10.1\%$	$-42.8 \pm 17.6\%$
	$\sigma = 16\%$	$-18.7 \pm 9.8\%$	$-21.0 \pm 15.5\%$	$-21.9 \pm 8.3\%$	$-30.2 \pm 11.6\%$	$-58.8 \pm 12.7\%$
	$\sigma = 32\%$	$-68.6 \pm 12.2\%$	$-70.8 \pm 10.6\%$	$-71.6 \pm 8.4\%$	$-76.6 \pm 9.0\%$	$-87.1 \pm 5.3\%$
AP_{75}	$\sigma = 0\%$	$\mathbf{61.5 \pm 26.7}$	$-8.8 \pm 6.4\%$	$-29.3 \pm 9.8\%$	$-55.1 \pm 13.6\%$	$-46.9 \pm 14.1\%$
	$\sigma = 4\%$	$-10.1 \pm 5.0\%$	$-16.7 \pm 8.9\%$	$-40.6 \pm 14.1\%$	$-58.7 \pm 19.1\%$	$-51.8 \pm 14.4\%$
	$\sigma = 8\%$	$-25.7 \pm 15.2\%$	$-32.9 \pm 18.1\%$	$-51.4 \pm 12.9\%$	$-76.8 \pm 10.9\%$	$-76.5 \pm 6.7\%$
	$\sigma = 16\%$	$-63.7 \pm 20.6\%$	$-67.9 \pm 23.2\%$	$-77.0 \pm 15.2\%$	$-92.4 \pm 7.7\%$	$-93.4 \pm 5.1\%$
	$\sigma = 32\%$	$-97.4 \pm 1.8\%$	$-97.2 \pm 1.8\%$	$-98.0 \pm 1.2\%$	$-98.7 \pm 0.7\%$	$-99.6 \pm 0.2\%$