

C. Atari 100k

The Atari 100k benchmark (Kaiser et al., 2020) features Atari games (Bellemare et al., 2013) with diverse mechanics. The specificity of this benchmark is the hard constraint on the number of interactions, namely one hundred thousand per environment. Compared to the standard Atari benchmark, this constraint results in a dramatic drop in real-time experience, from 900 hours to 2 hours.

Regarding baselines, we consider four model-based RL agents learning in imagination: SimPLe (Kaiser et al., 2020), DreamerV3 (Hafner et al., 2023), STORM (Zhang et al., 2023), and IRIS (Micheli et al., 2023). We note that the current best performing methods for Atari 100k resort to other approaches, such as lookahead search for EfficientZero (Ye et al., 2021), or self-supervised representation learning with periodic resets for BBF (Schwarzer et al., 2023).

The usual metric of interest is the HNS, the human-normalized score, based on the performance of human players with similar experience. A negative HNS indicates worse than random performance whereas an HNS above 1 signifies superhuman performance. We evaluate Δ -IRIS by computing an average over 100 episodes collected at the end of training for each game (5 seeds). For the baselines, we report the published results.

Table 8 displays returns across games and aggregate metrics (Agarwal et al., 2021). Δ -IRIS achieves higher aggregate metrics than IRIS, while training in 26 hours, a 5-fold speedup.

Table 8. Returns on the 26 games of Atari 100k after 2 hours of real-time experience, and human-normalized aggregate metrics.

Game	Random	Human	SimPLe	DreamerV3	STORM	IRIS	Δ -IRIS (ours)
Alien	228	7128	617	959	984	420	599
Amidar	6	1720	74	139	205	143	51
Assault	222	742	527	706	801	1524	1435
Asterix	210	8503	1128	932	1028	854	2001
BankHeist	14	753	34	649	641	53	1206
BattleZone	2360	37188	4031	12250	13540	13074	10365
Boxing	0	12	8	78	80	70	56
Breakout	2	31	16	31	16	84	226
ChopperCommand	811	7388	979	420	1888	1565	1101
CrazyClimber	10781	35829	62584	97190	66776	59324	70920
DemonAttack	152	1971	208	303	165	2034	884
Freeway	0	30	17	0	34	31	31
Frostbite	65	4335	237	909	1316	259	287
Gopher	258	2413	597	3730	8240	2236	9349
Hero	1027	30826	2657	11161	11044	7037	6235
Jamesbond	29	303	101	445	509	463	345
Kangaroo	52	3035	51	4098	4208	838	1573
Krull	1598	2666	2205	7782	8413	6616	6392
KungFuMaster	259	22736	14863	21420	26182	21760	25159
MsPacman	307	6952	1480	1327	2674	999	1175
Pong	-21	15	13	18	11	15	15
PrivateEye	25	69571	35	882	7781	100	100
Qbert	164	13455	1289	3405	4523	746	3438
RoadRunner	12	7845	5641	15565	17564	9615	10622
Seaquest	68	42055	683	618	525	661	895
UpNDown	533	11693	3350	9234	7985	3546	8091
#Superhuman	0	N/A	1	9	10	10	12
Mean	0.00	1.00	0.33	1.10	1.27	1.05	1.39
Interquartile Mean	0.00	1.00	0.13	0.50	0.64	0.50	0.68