

Calibra Targeted Benchmark -- lerobot/pusht
(n_seeds=10, n_epochs=120, tail=15% by action-space density)

