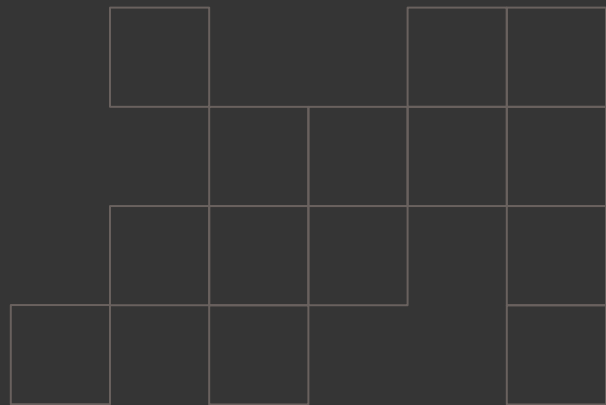


CSCI 166
Fall 2025

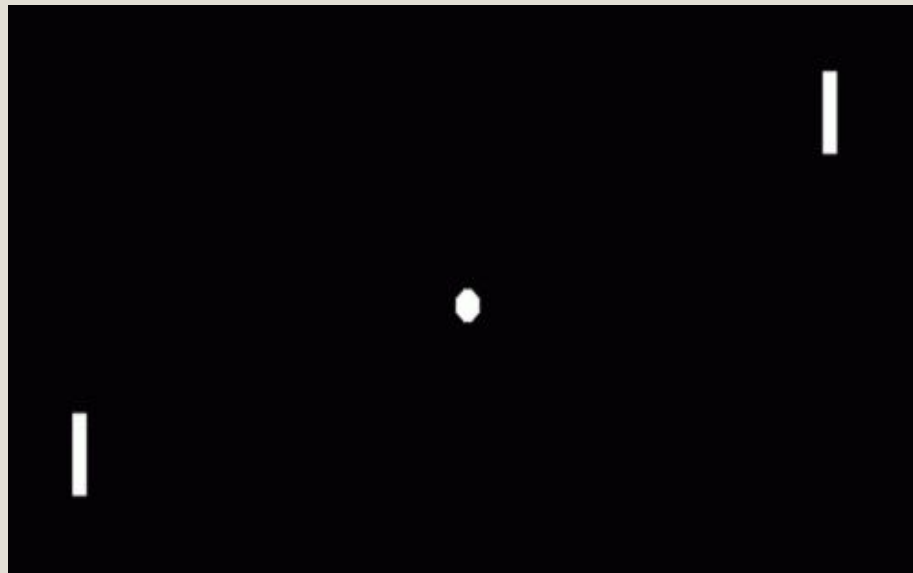
DQN vs Double DQN on Atari Pong

Isac Lopez



Motivation

- Pong is a classic RL benchmarking environment
- Simple discrete actions
- Ideal for comparing DQN variants



Environment

- Observation: (4, 84, 84) stacked frames
- Actions: NOOP, UP, DOWN
- Sparse rewards: +1 / -1



Baseline DQN

- ConvNet \rightarrow FC \rightarrow $|A|$ -
- Experience replay
- Target network
- Epsilon-greedy exploration



Double DQN

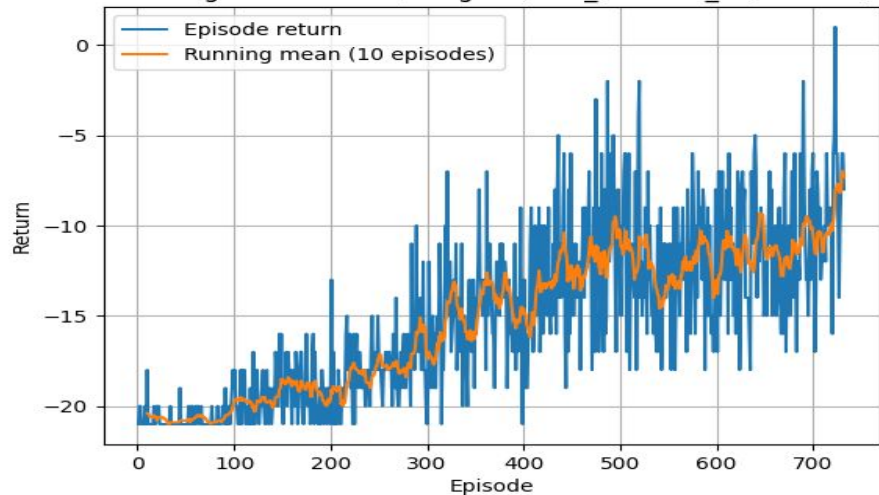
- Fixes Q-value overestimation
- Online net selects action
- Target net evaluates action
- More stable learning



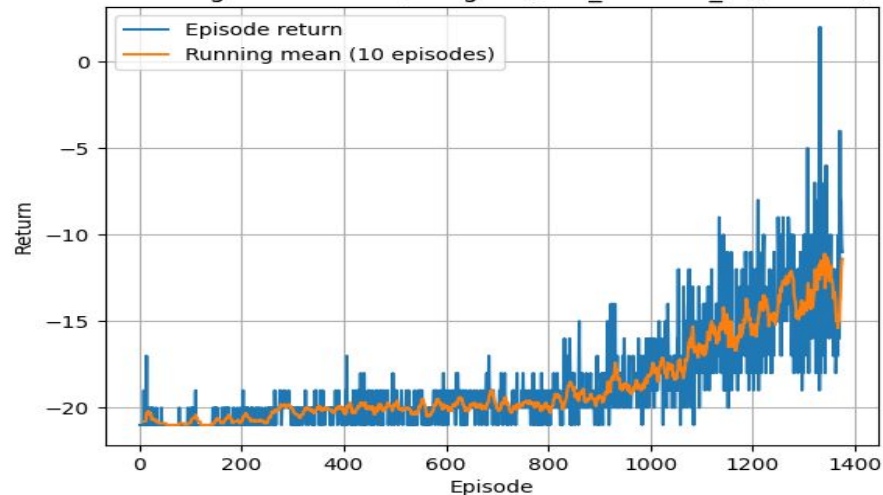
Results

- Baseline: noisy, unstable Double DQN Results
- Double DQN: smoother, faster improvement
- Better Q-value stability

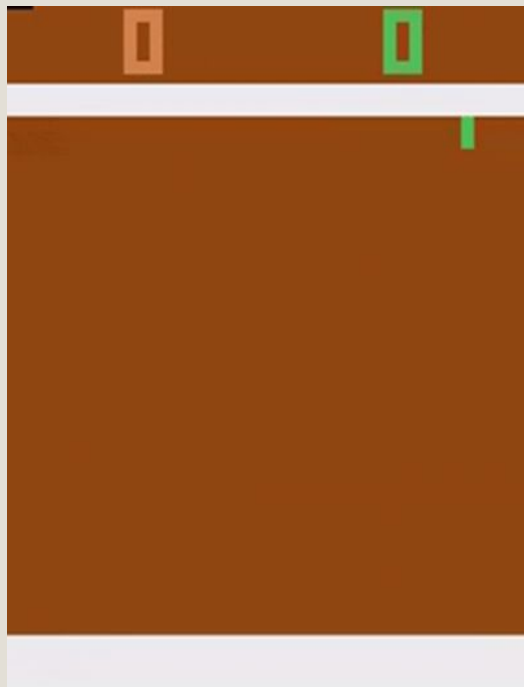
Learning curve for ALE/Pong-v5 (USE_DOUBLE_DQN=False)



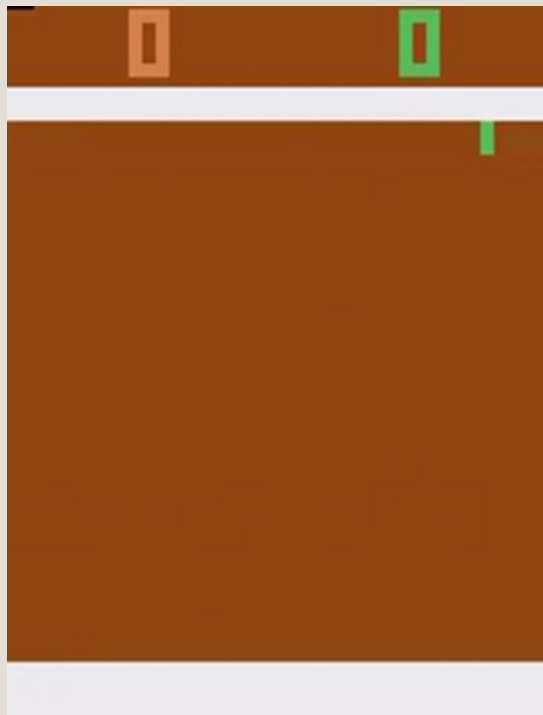
Learning curve for ALE/Pong-v5 (USE_DOUBLE_DQN=True)



Baseline DQN: Early



Double DQN: Early



Early policy: random paddle movement

Baseline DQN: Learned



Double DQN: Learned



Learned policy: controlled, tracks ball

Reflection

- Sparse rewards were challenging
- Double DQN improved stability
- Future: PER, N-step, Dueling, sticky actions



Conclusion

- Implemented DQN + Double DQN
- Compared learning curves & gameplay
- Clear benefits to Double DQ

