

# Neutral Network Agents for Control Tasks in OpenAI Gym

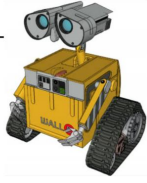
Shengjun (Sophia) Qin<sup>1</sup> and Junkai Jiang<sup>2</sup>

<sup>1</sup> Department of Electrical Engineering, Stanford University <sup>2</sup> Department of Electrical and Computer Engineering, UC Santa Barbara



## Introduction

- Reinforcement learning (RL) has a wide range of applications including robotics, data center management, and the well-known AlphaGo!
- The OpenAI Gym offers various scenarios to benchmark RL algorithms, where we design our NN agents for different control tasks to achieve desired performance.



## Environments & Networks

(a) Atari-Breakout: Discrete Action, Observation, Neural Network Agent

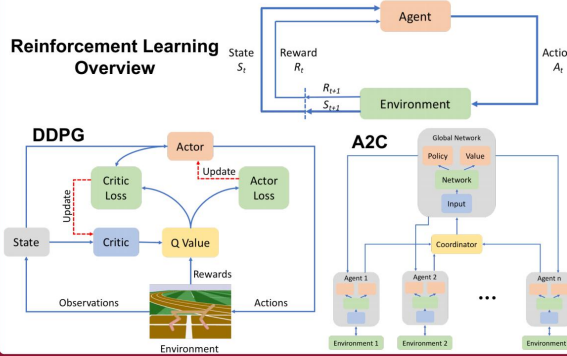
(b) Roboschool HalfCheetah, (c) Roboschool Ant, (d) Roboschool Walker2d

**DNN** (Deep Neural Network) architecture diagram

**CNN** (Convolutional Neural Network) architecture diagram: Conv-1 (Filter: 8 x 8 x 32, Stride: 4), Conv-2 (Filter: 4 x 4 x 64, Stride: 2), Conv-3 (Filter: 3 x 3 x 64, Stride: 1), FC-1 (512)

## RL Algorithms

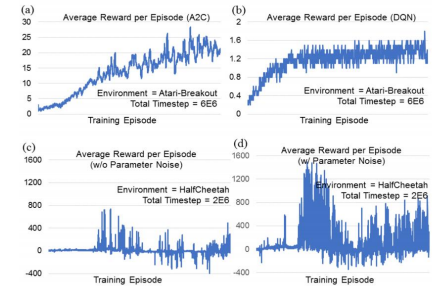
- RL trainings under Markov decision process assumption
- Search for Q-function to determines the action that maximizes rewards



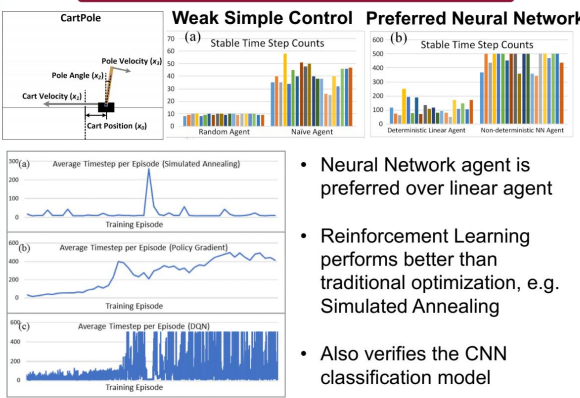
## Results and Discussions

### RL Algorithm Dependency

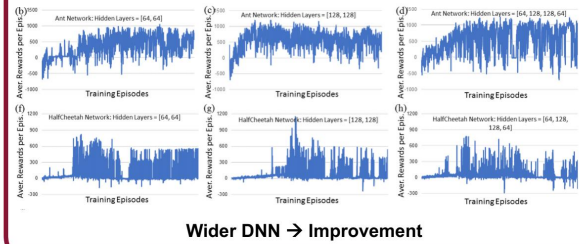
- A2C > DQN (Unoptimized hyperparameters)
- DDPG Parameter noise offers huge improvement in agent performance



## CartPole



### Neural Network Architecture Dependency



Wider DNN → Improvement

## Conclusions & Future Work

- Reinforcement learning is sensitive to hyperparameter tuning
- Wider DNN shows advantages in improving agent performance
- Future work:
  - Improve training convergence and agent performance
  - Apply reinforcement learning in deepmind/pysc2

