

CS 5180 Final Project

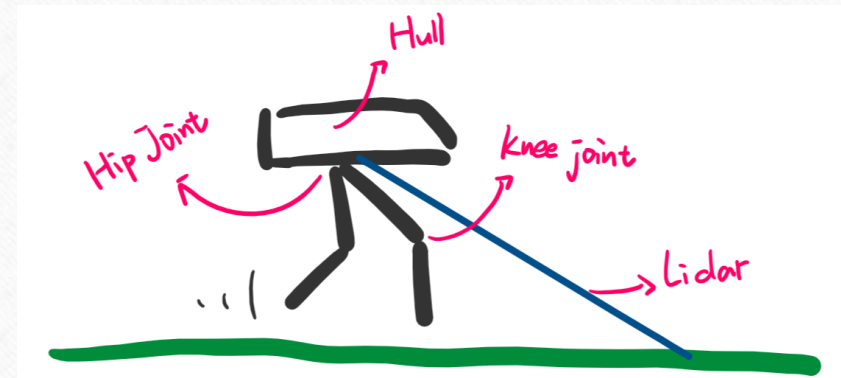
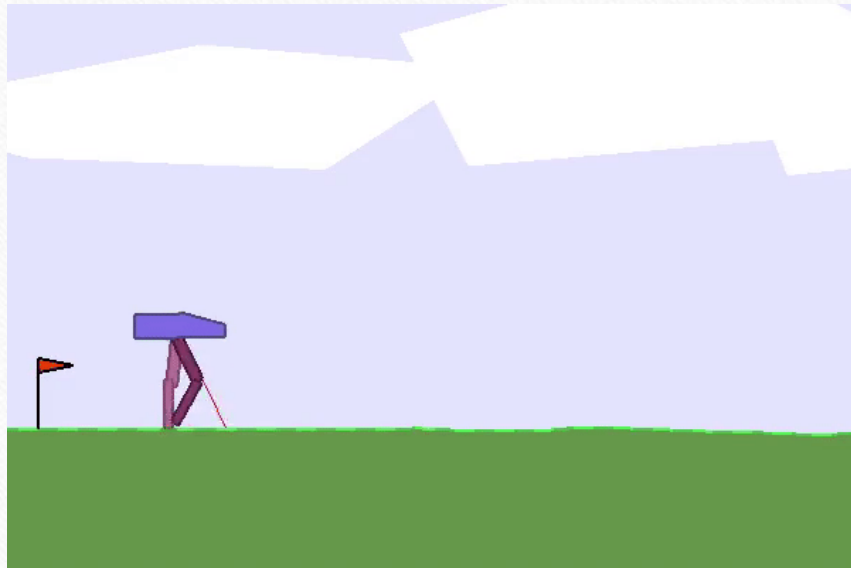
- Bipedal Walker -

Presented by
Meishan Li, Zhida Zhang

Project Problem

BipedalWalker-v3:

The main problem to be solved in this project is to train the robot to choose the correct action to stand up and start walking a certain distance in an environment with terrain that is not perfectly horizontal.



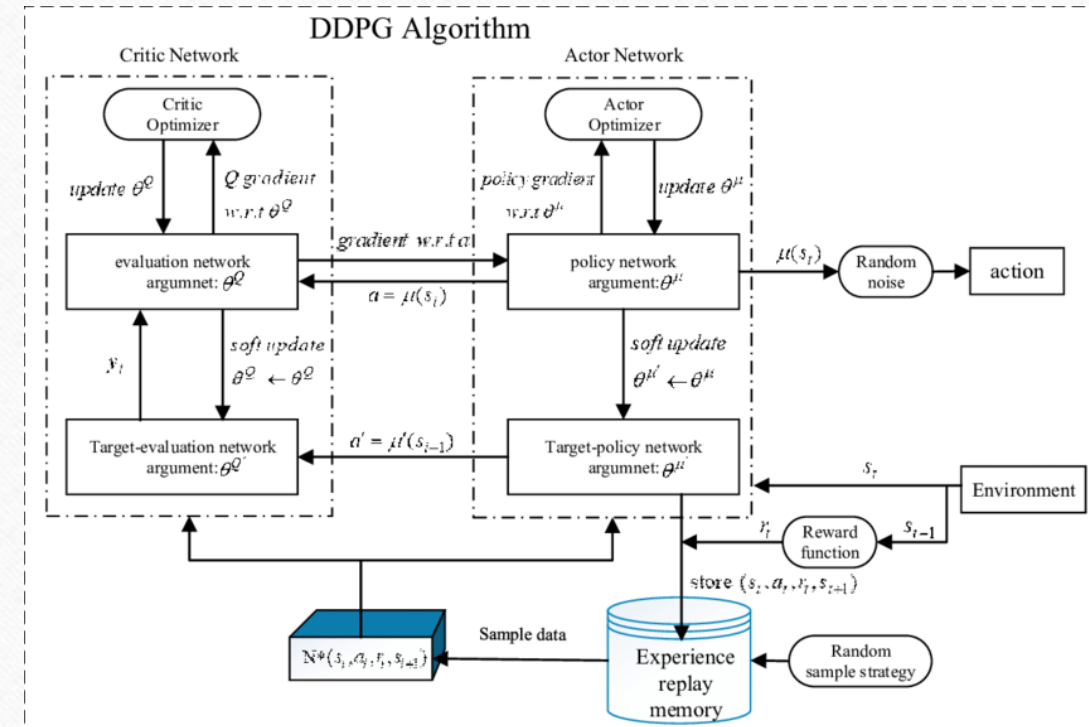
Project Approach/Algorithms

- Deep deterministic policy gradient algorithm(DDPG)

DDPG is a reinforcement learning algorithm based on the AC(Actor-Critic) framework, which maximizes the total reward by predicting a deterministic policy by updating the policy in a single step through the off-policy method.

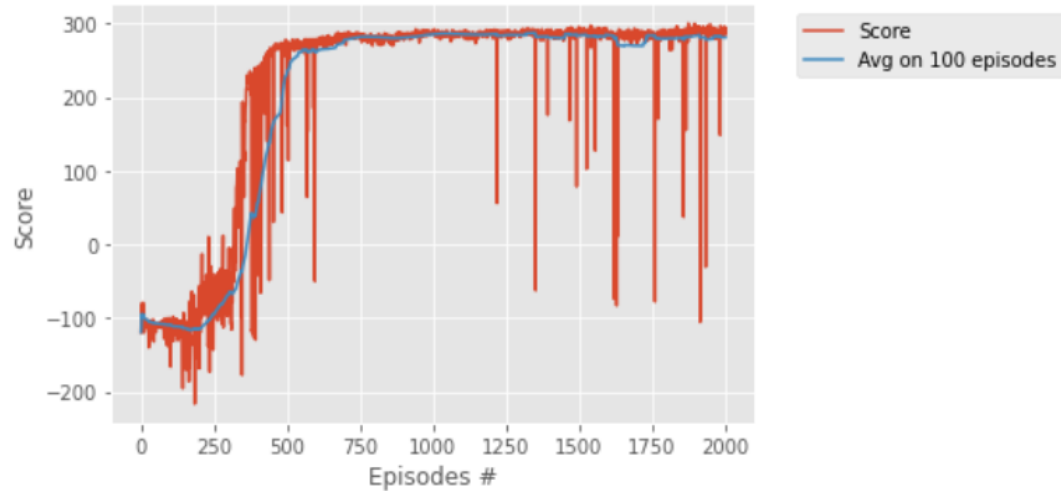
- Twin Delayed DDPG(TD3)

In TD3, two sets of Critic networks are used to represent different Q values and to suppress persistent overestimation by selecting the smallest one as the Target Q Value.



Project Experiments/comparisons/results

length of scores: 2000 , len of avg_scores: 2000



Results of TD3 algorithm

100% | ██████████ | 1990/2000 [3:14:36<00:55, 5.59s/it]

Ep. 1991, Timestep 1903222, Ep.Timesteps 885, Score: 288.01, Avg.Score: 281.06, Time: 03:14:36

100% | ██████████ | 1991/2000 [3:14:41<00:50, 5.61s/it]

Ep. 1992, Timestep 1904107, Ep.Timesteps 885, Score: 291.61, Avg.Score: 281.10, Time: 03:14:42

100% | ██████████ | 1992/2000 [3:14:47<00:44, 5.62s/it]

Ep. 1993, Timestep 1904981, Ep.Timesteps 874, Score: 293.97, Avg.Score: 281.12, Time: 03:14:48

100% | ██████████ | 1993/2000 [3:14:53<00:39, 5.61s/it]

Ep. 1994, Timestep 1905901, Ep.Timesteps 920, Score: 281.15, Avg.Score: 281.05, Time: 03:14:53

100% | ██████████ | 1994/2000 [3:14:59<00:34, 5.69s/it]

Ep. 1995, Timestep 1906805, Ep.Timesteps 904, Score: 288.70, Avg.Score: 280.94, Time: 03:14:59

100% | ██████████ | 1995/2000 [3:15:04<00:28, 5.72s/it]

Ep. 1996, Timestep 1907668, Ep.Timesteps 863, Score: 294.81, Avg.Score: 281.00, Time: 03:15:05

100% | ██████████ | 1996/2000 [3:15:10<00:22, 5.65s/it]

Ep. 1997, Timestep 1908553, Ep.Timesteps 885, Score: 289.06, Avg.Score: 280.98, Time: 03:15:11

100% | ██████████ | 1997/2000 [3:15:15<00:16, 5.66s/it]

Ep. 1998, Timestep 1909412, Ep.Timesteps 859, Score: 295.10, Avg.Score: 281.04, Time: 03:15:16

100% | ██████████ | 1998/2000 [3:15:21<00:11, 5.61s/it]

Ep. 1999, Timestep 1910319, Ep.Timesteps 907, Score: 284.18, Avg.Score: 280.98, Time: 03:15:22

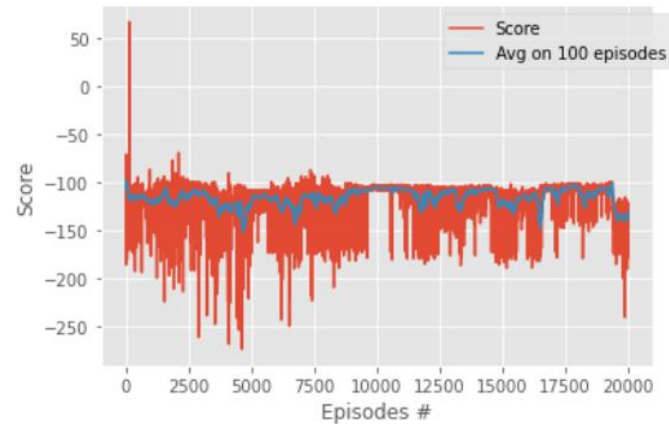
100% | ██████████ | 1999/2000 [3:15:27<00:05, 5.66s/it]

Ep. 2000, Timestep 1911208, Ep.Timesteps 889, Score: 288.87, Avg.Score: 280.90, Time: 03:15:27

100% | ██████████ | 2000/2000 [3:15:32<00:00, 5.87s/it]

Project Experiments/comparisons/results

length of scores: 20000 , len of avg_scores: 20000



Only one critic network(1)

Ep. 959, Timestep 449171, Ep.Timesteps 1030, Score: 268.99, Avg.Score: 47.35, Time: 00:40:44
48% [██████] 959/2000 [40:49:1:06:14, 3.82s/it]
Ep. 960, Timestep 449279, Ep.Timesteps 108, Score: -107.74, Avg.Score: 47.72, Time: 00:40:49
48% [██████] 960/2000 [40:50:49:19, 2.85s/it]
Ep. 961, Timestep 449538, Ep.Timesteps 259, Score: -57.87, Avg.Score: 48.80, Time: 00:40:50
48% [██████] 961/2000 [40:51:42:32, 2.46s/it]
Ep. 962, Timestep 450828, Ep.Timesteps 1290, Score: 250.53, Avg.Score: 53.69, Time: 00:40:53
48% [██████] 962/2000 [40:59:1:08:17, 3.95s/it]
Ep. 963, Timestep 451911, Ep.Timesteps 1083, Score: 263.69, Avg.Score: 57.42, Time: 00:41:00
48% [██████] 963/2000 [41:05:1:21:25, 4.71s/it]
Ep. 964, Timestep 452159, Ep.Timesteps 248, Score: -72.12, Avg.Score: 56.20, Time: 00:41:06
48% [██████] 964/2000 [41:07:1:04:14, 3.72s/it]
Ep. 965, Timestep 452436, Ep.Timesteps 277, Score: -108.53, Avg.Score: 56.18, Time: 00:41:07
48% [██████] 965/2000 [41:08:53:02, 3.08s/it]
Ep. 966, Timestep 452957, Ep.Timesteps 521, Score: -151.96, Avg.Score: 55.70, Time: 00:41:09
48% [██████] 966/2000 [41:11:52:49, 3.06s/it]
Ep. 967, Timestep 453308, Ep.Timesteps 351, Score: -152.04, Avg.Score: 55.08, Time: 00:41:12
48% [██████] 967/2000 [41:13:47:44, 2.77s/it]
Ep. 968, Timestep 453639, Ep.Timesteps 331, Score: -156.01, Avg.Score: 54.64, Time: 00:41:14
48% [██████] 968/2000 [41:15:43:15, 2.51s/it]
Ep. 969, Timestep 455239, Ep.Timesteps 1600, Score: -122.44, Avg.Score: 54.07, Time: 00:41:17
48% [██████] 969/2000 [41:25:1:17:26, 4.51s/it]
Ep. 970, Timestep 455508, Ep.Timesteps 269, Score: -83.49, Avg.Score: 51.60, Time: 00:41:25
48% [██████] 970/2000 [41:26:1:02:14, 3.63s/it]
Ep. 971, Timestep 456281, Ep.Timesteps 773, Score: -189.89, Avg.Score: 47.49, Time: 00:41:27

Only one critic network(2)

Ep. 17465, Timestep 12077844, Ep.Timesteps 702, Score: 150.69, Avg.Score: 185.60, Time: 15:58:28

87% [██████] | 17465/20000 [15:58:31:1:55:17, 2.73s/it]

Ep. 17466, Timestep 12078615, Ep.Timesteps 771, Score: 292.10, Avg.Score: 185.68, Time: 15:58:32

87% [██████] | 17466/20000 [15:58:35:2:17:01, 3.24s/it]

Ep. 17467, Timestep 12079404, Ep.Timesteps 789, Score: 292.74, Avg.Score: 189.12, Time: 15:58:36

87% [██████] | 17467/20000 [15:58:40:2:31:52, 3.60s/it]

Ep. 17468, Timestep 12080182, Ep.Timesteps 778, Score: 292.72, Avg.Score: 189.58, Time: 15:58:40

87% [██████] | 17468/20000 [15:58:44:2:37:29, 3.73s/it]

Ep. 17469, Timestep 12080949, Ep.Timesteps 767, Score: 294.69, Avg.Score: 189.54, Time: 15:58:45

87% [██████] | 17469/20000 [15:58:48:2:42:00, 3.84s/it]

Ep. 17470, Timestep 12081723, Ep.Timesteps 774, Score: 293.70, Avg.Score: 189.52, Time: 15:58:49

87% [██████] | 17470/20000 [15:58:52:2:45:15, 3.92s/it]

Ep. 17471, Timestep 12082470, Ep.Timesteps 747, Score: 296.60, Avg.Score: 189.58, Time: 15:58:53

87% [██████] | 17471/20000 [15:58:56:2:44:53, 3.91s/it]

Ep. 17472, Timestep 12083243, Ep.Timesteps 773, Score: 293.44, Avg.Score: 191.53, Time: 15:58:57

87% [██████] | 17472/20000 [15:59:00:2:47:37, 3.98s/it]

Ep. 17473, Timestep 12083990, Ep.Timesteps 747, Score: 294.88, Avg.Score: 194.67, Time: 15:59:01

87% [██████] | 17473/20000 [15:59:04:2:48:12, 3.99s/it]

Ep. 17474, Timestep 12084744, Ep.Timesteps 754, Score: 293.55, Avg.Score: 197.84, Time: 15:59:05

87% [██████] | 17474/20000 [15:59:08:2:47:58, 3.99s/it]

Ep. 17475, Timestep 12085507, Ep.Timesteps 763, Score: 293.45, Avg.Score: 197.97, Time: 15:59:09