RL Topic HW2

1. Training curve and testing result of DQN Pacman:
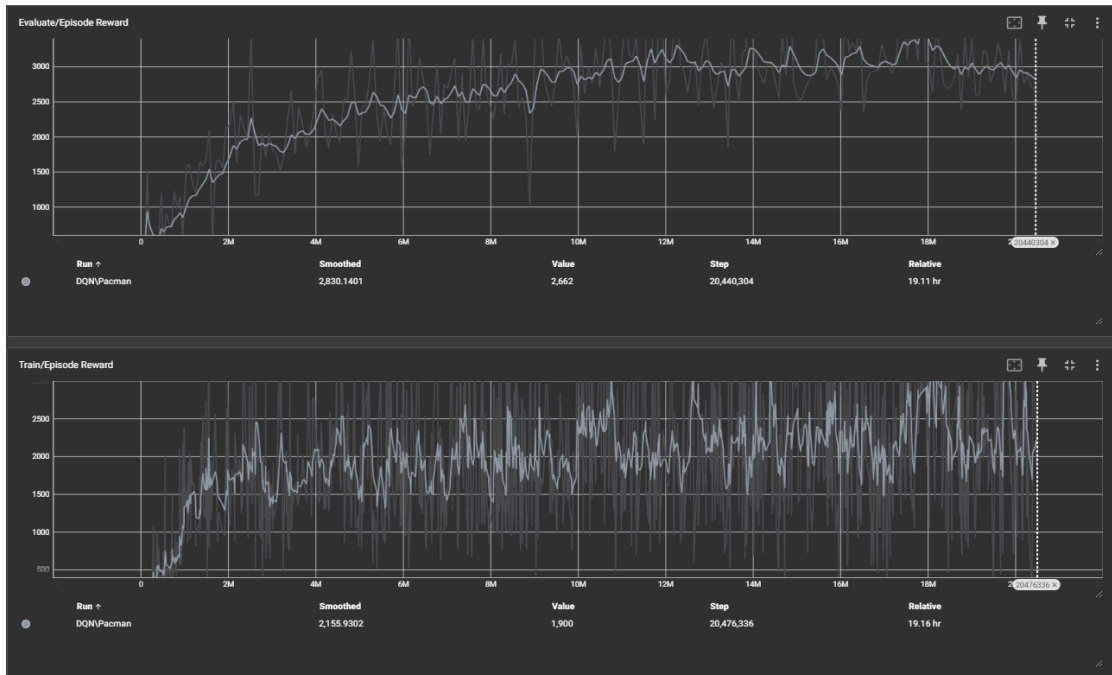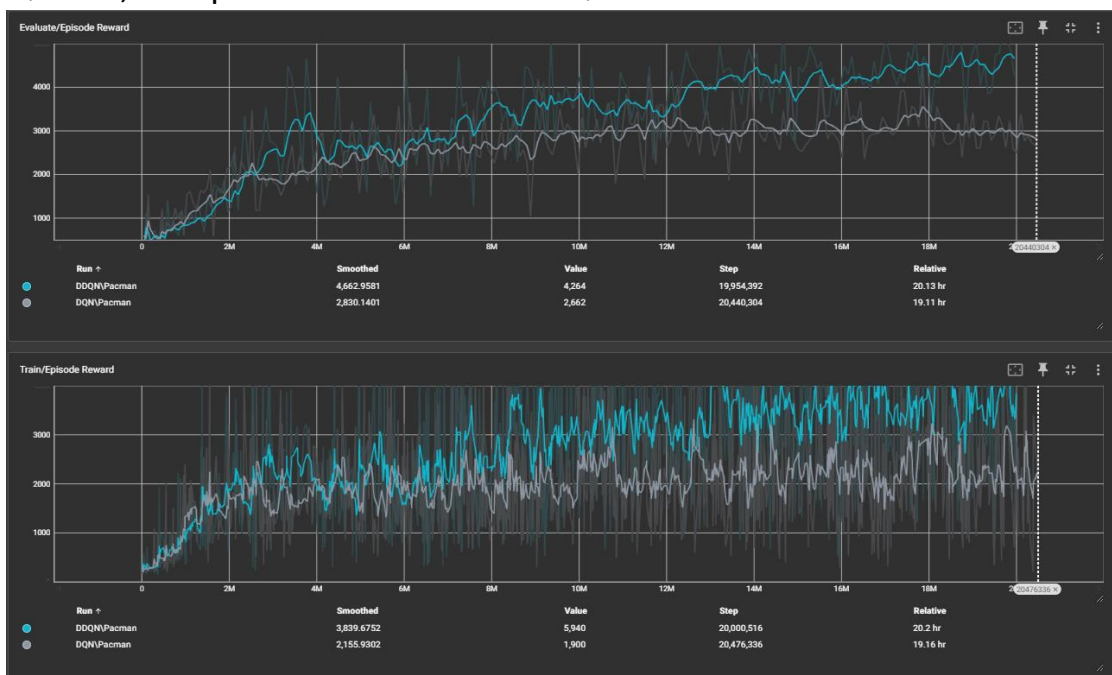


```
================================================
Evaluating...
E:\Code\RLTopics\.venv\lib\site-packages\gym\utils\
  if not isinstance(terminated, (bool, np.bool8)):
episode 1 reward: 3830.0
episode 2 reward: 2960.0
episode 3 reward: 3850.0
episode 4 reward: 4720.0
episode 5 reward: 3020.0
average score: 3676.0
================================================
```

2. Screenshot of DQN Enduro:
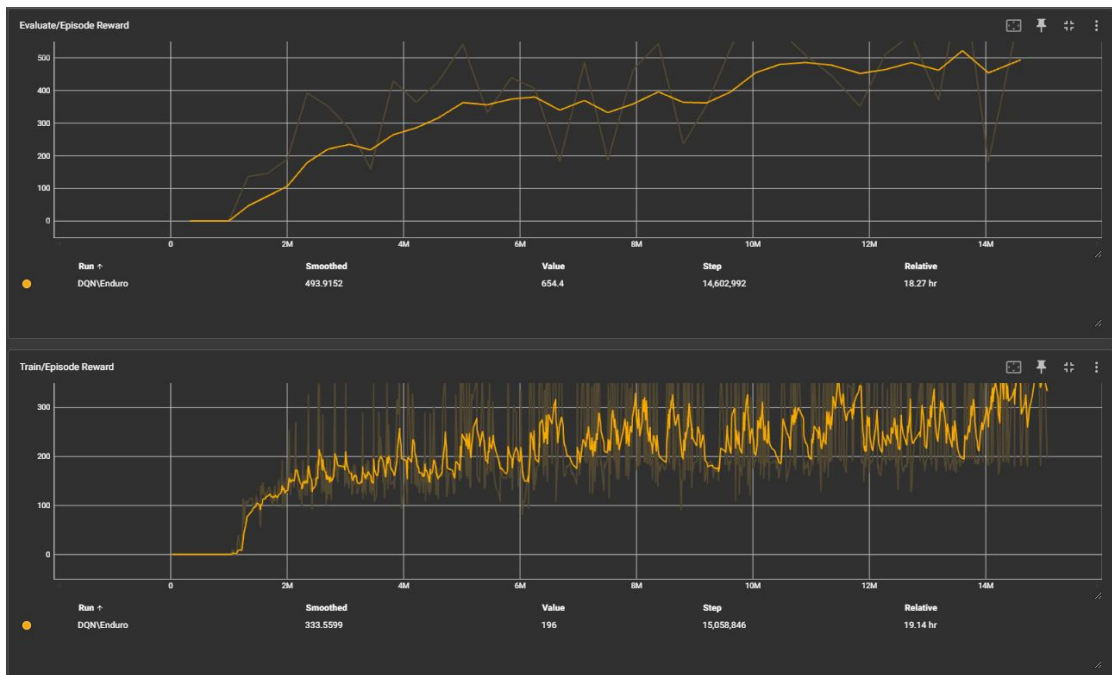
3. Screenshot of DQN and DDQN: DDQN can prevent over-estimating a Q value, so it performs better than DQN.

4. Screenshot of DQN and dueling DQN (PC restarted so it only has 15M steps): Dueling DQN uses an advantage head to learn a single action value, and another value head to adjust all action value, so it can adjust overall action values without testing through every action.

Evaluate/Episode Reward

Run ↑	Smoothed	Value	Step	Relative
DQN\Pacman	2,830.1401	2,662	20,440,304	19.11 hr
DuelingDQN\Pacman	3,366.6164	2,786	15,186,238	17.23 hr

Train/Episode Reward

Run ↑	Smoothed	Value	Step	Relative
DQN\Pacman	2,155.9302	1,900	20,476,336	19.16 hr
DuelingDQN\Pacman	2,845.2449	2,250	15,256,090	17.32 hr