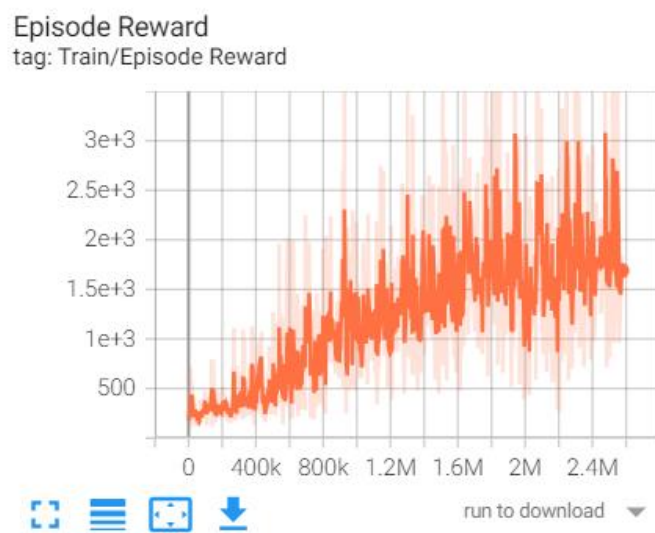


DQN

Testing result:

```
episode 1 reward: 4150.0  
episode 2 reward: 1570.0  
episode 3 reward: 1990.0  
episode 4 reward: 5530.0  
episode 5 reward: 2210.0  
average score: 3090.0  
=====
```

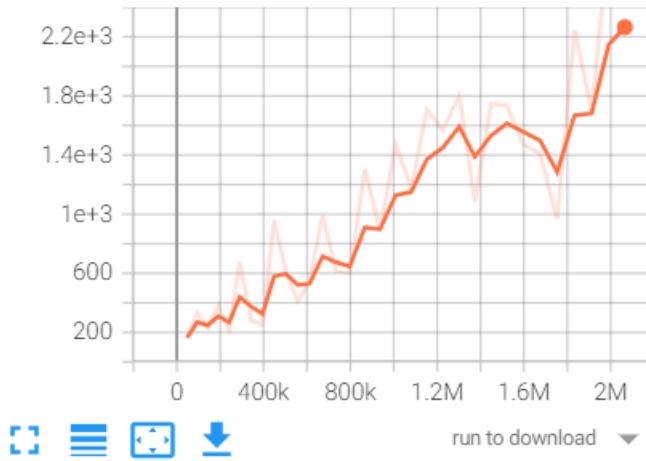
Training and evaluating curve:



Evaluate

Episode Reward

tag: Evaluate/Episode Reward



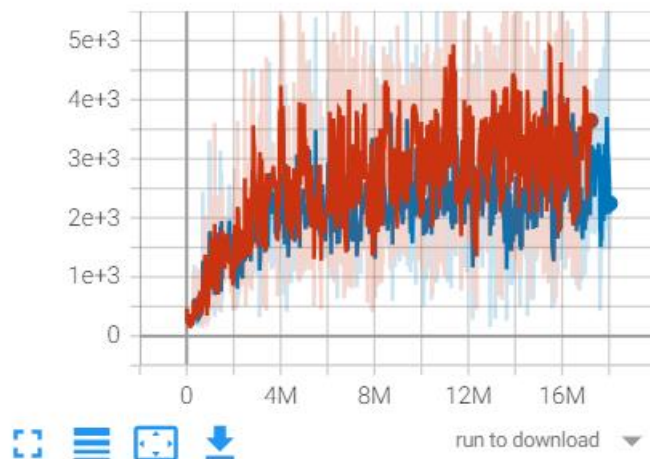
Training and evaluating curve on DDQN(red) and DQN(blue):

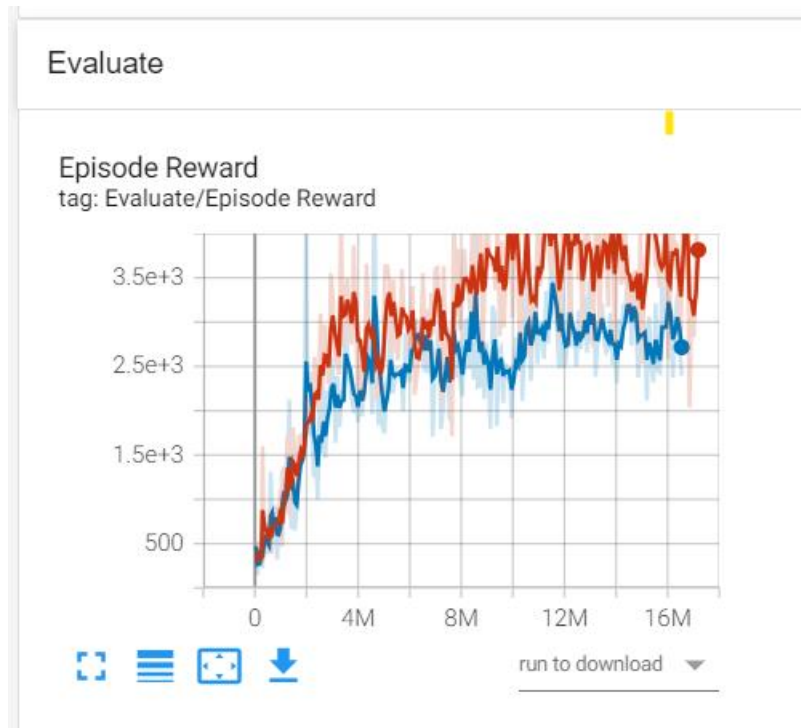
DDQN testing result:

```
episode 1 reward: 4580.0
episode 2 reward: 3950.0
episode 3 reward: 920.0
episode 4 reward: 3060.0
episode 5 reward: 2970.0
average score: 3096.0
```

Episode Reward

tag: Train/Episode Reward





Difference between DQN and DDQN:

DQN uses target network to select the next best action which will result in over-optimistic result. DDQN, on the other hand, select the action which has the greatest Q value first, and using that action to predict the target value on target network, which is more accurate to the real value.