Collaboration and Competition Report

This report describes the details of my implementation of the **Collaboration and Competition** project from the Udacity's Deep Reinforcement Learning Nanodegree program, along with ideas for future work.

Learning Algorithm

As the action space is continuous, I decided to use Deep Deterministic Policy Gradients Algorithm to train the agents to play tennis. There are two neural networks I have used: the actor network and the critic network.

The Actor network

The Actor network contains two hidden layers of 128 and 64 units respectively. ReLU activation has been applied on both layers. On the final fully connected layer, tanh activation has been applied. Batch normalization has also been applied on the input to actor network.

The Critic network

The Critic network has also two hidden layers of 128 and 64 units with reLU activation on both layers. On the final fully connected layer, no activation has been applied. Same as the actor network, batch normalization has also been applied on the input to critic network.

Hyperparameters

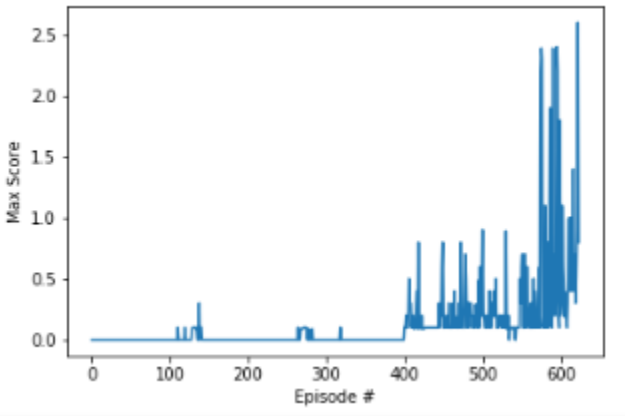
Below are the hyperparameters used:

1. BUFFER_SIZE = int(1e6) # replay buffer size
2. BATCH_SIZE = 512 # minibatch size
3. GAMMA = 0.99 # discount factor
4. TAU = 1e-3 # for soft update of target parameters
5. LR_ACTOR = 5e-4 # learning rate of the actor
6. LR_CRITIC = 1e-3 # learning rate of the critic
7. WEIGHT_DECAY = 0.0000 # L2 weight decay

Rewards

Based on the plot below, the environment is solved in 624 episodes.

Average Score: 0.50



Ideas for future work

The agent's performance could be further improved by using standard techniques for better neural network training:

- Changing the Actor and Critic network sizes in DDPG
- Trying algorithms other than DDPG such as PPO, A3C, and D4PG.