Project report

Collaboration and Competition

Learning Algorithm:

For this project DDPG (**D**eep **D**eterministic **P**olicy **G**radients) is used which performs good in continuous action space. The network is being updated 10 times after 2-time steps.

The Deep Neural Network has following layers:

Actor Network

- A hidden layer with 128 units and RELU activation.
- Batch normalization.
- Second hidden layer with 64 units and RELU activation.
- Batch normalization.
- Fully connected Layer.
- Tanh.

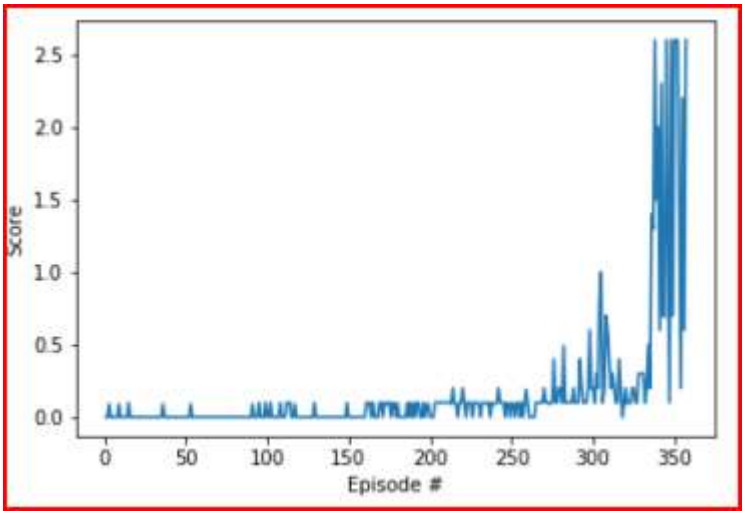
Critic Network

- Batch normalization on input.
- A hidden layer with 128 units and RELU activation.
- Second hidden layer with 64 units and RELU activation.
- Fully connected Layer.

Parameters used in DDPG Algorithm:

```
BUFFER_SIZE = int(1e6)  # replay buffer size
BATCH_SIZE = 512        # minibatch size
GAMMA = 0.99            # discount factor
TAU = 1e-3              # for soft update of target parameters
LR_ACTOR = 5e-4         # learning rate of the actor
LR_CRITIC = 1e-3        # learning rate of the critic
WEIGHT_DECAY = 0        # L2 weight decay
```

Result:



The task was completed in 357 episodes.

Ideas for future work

1. Will try to solve this project using other algorithms like PPO, A3C OR MADDPG.

Acknowledgements:

Model Visualization: https://github.com/szagoruyko/pytorchviz