University at Buffalo The State University of New York

CSE4/510 Project

UAV LEARNING TO FLY

Rugved Hattekar

Nanditha Nandanavanam

Chen Zeng

University at Buffalo The State University of New York

1846

Objectives

This project aims to apply a reinforcement learning pipeline on a quadcopter dynamics to function as the controller that is capable of performing takeoff, landing, and stable hovering.

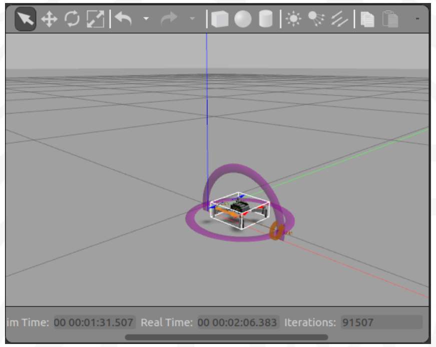
- Based on a commercial platform (Crazyflie 2.0)
- Simulated flight dynamics in ROS environment
- DoE & RL training for taking off and hovering (separately)

The Platform

Hardware: Crazyflie 2.0

- Weights below 35g
- Integrated sensors and controllers
- Open-source controller and SDK

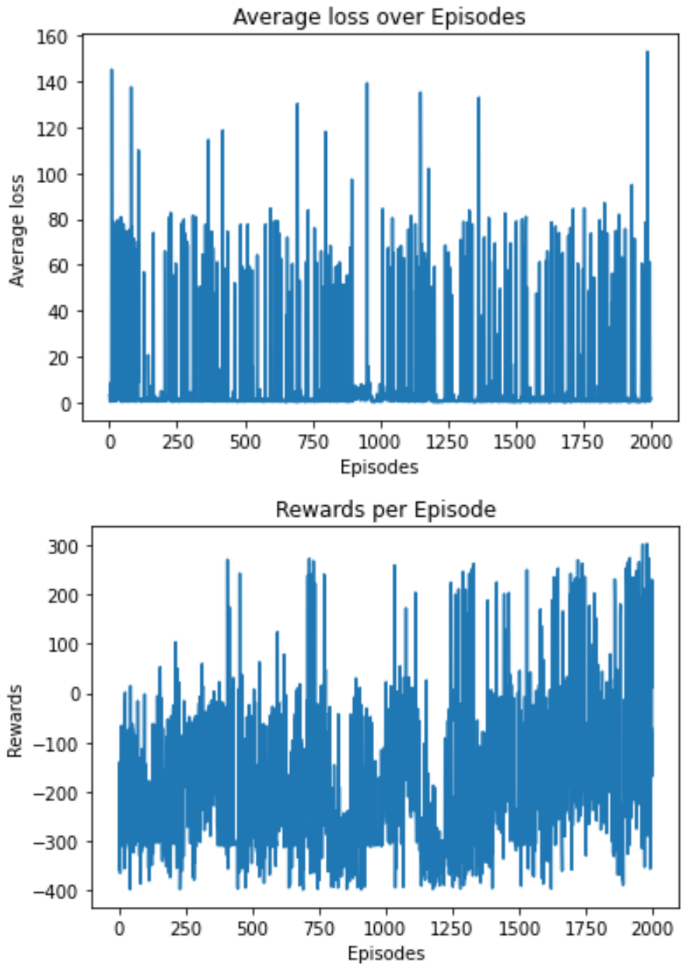
Software (planned): Gazebo Simulator



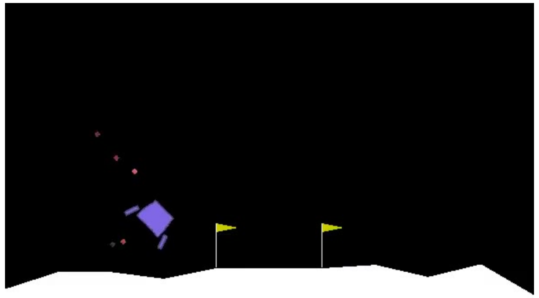
- Works in ROS environment
- Existing packages (CrazyS) for dynamics simulation of Crazyflie

Preliminary Progress (Checkpoint)

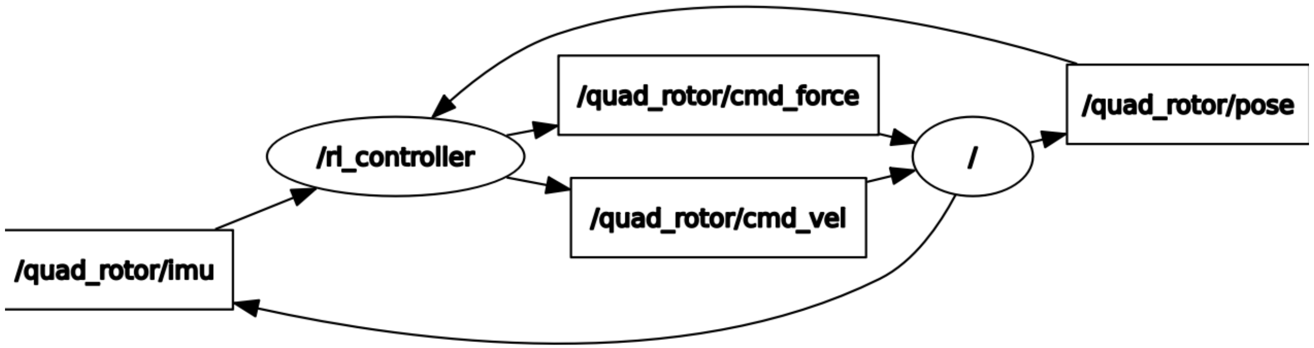
- Setting up Gazebo environment is more time-consuming …
- Lunar Landing Problem chosen as the preliminary test problem
- DQN applied, results not promising

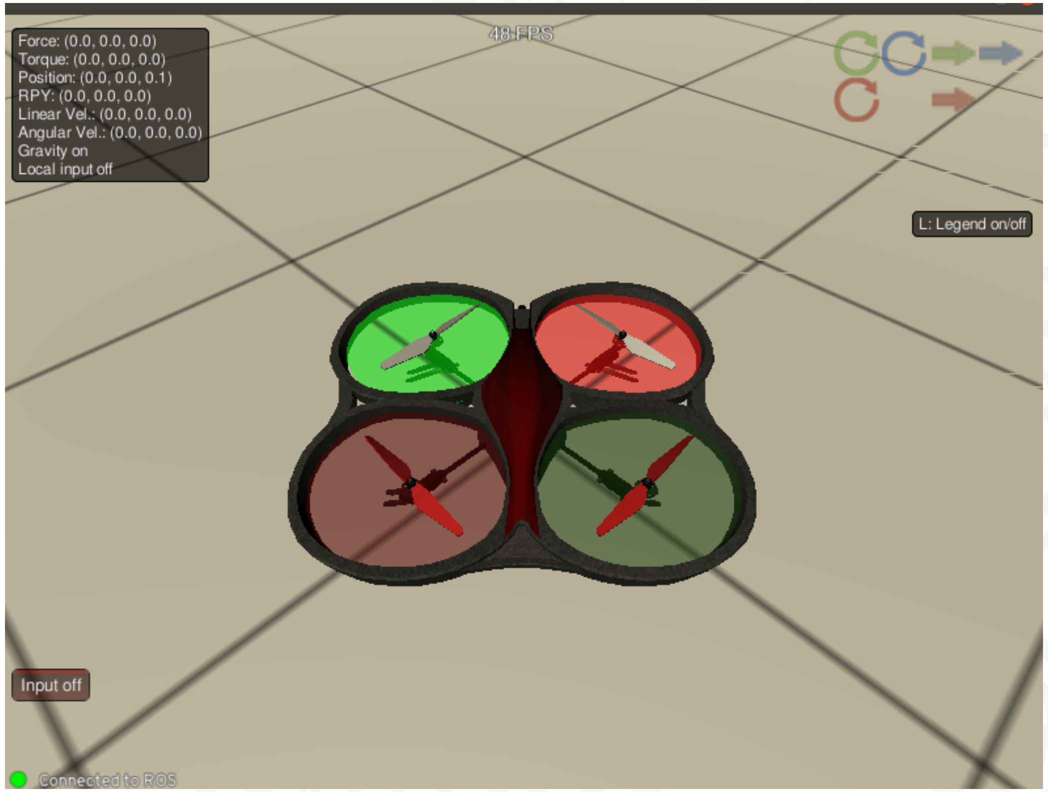


- Learning rate = 0.001
- Epsilon = 0.7
- Epsilon decay = 0.995
- Number of episodes = 5000
- Discount factor = 0.98



ROS Architecture Schematic:





Basic Architecture:

- IMU node : Position, Velocity, Acceleration

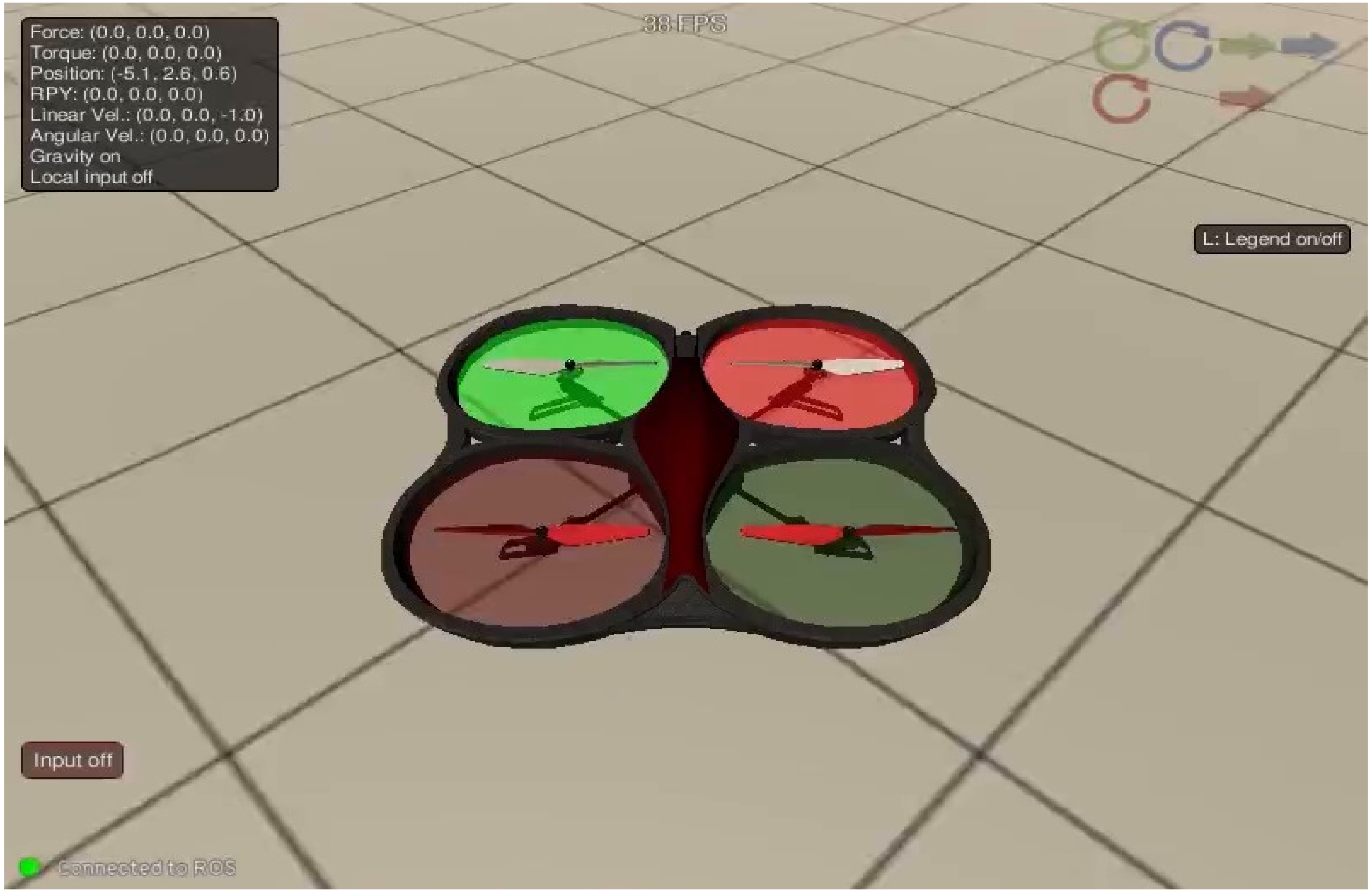
- RL Node: Thrust Force, Linear Velocity

- Action_space: Thrust Force (From DQN)

- Linear Velocity: Constant Throughout

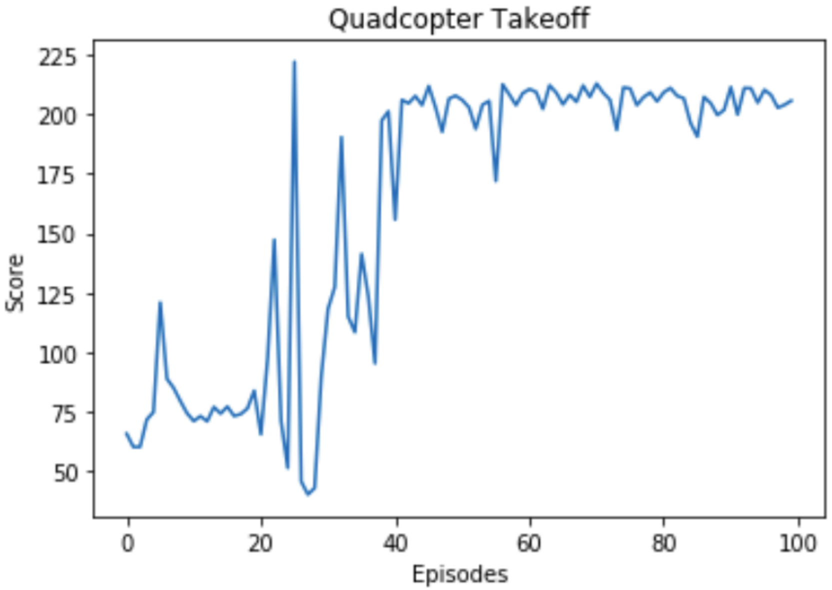

```
---
force:
  x: 0.0
  y: 0.0
  z: 17.0
torque:
  x: 0.0
  y: 0.0
  z: 0.0
---
force:
  x: 0.0
  y: 0.0
  z: 25.0
torque:
  x: 0.0
  y: 0.0
  z: 0.0
---
```

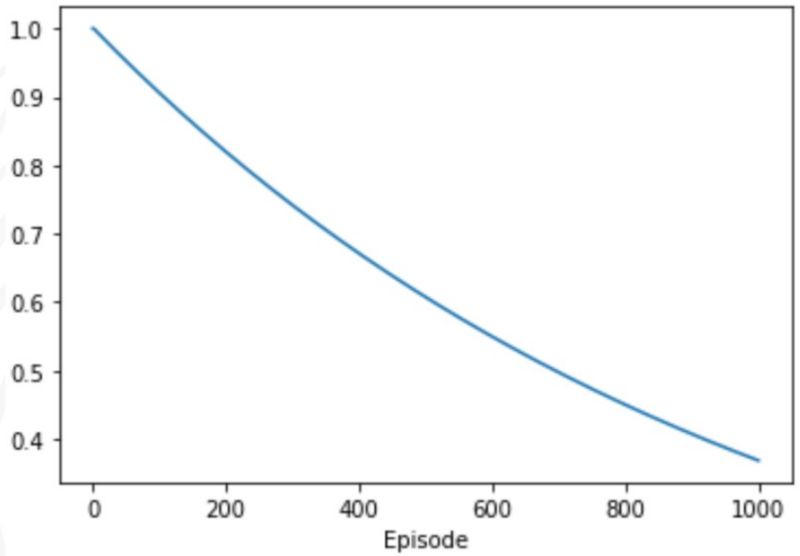
Simulator Demo:



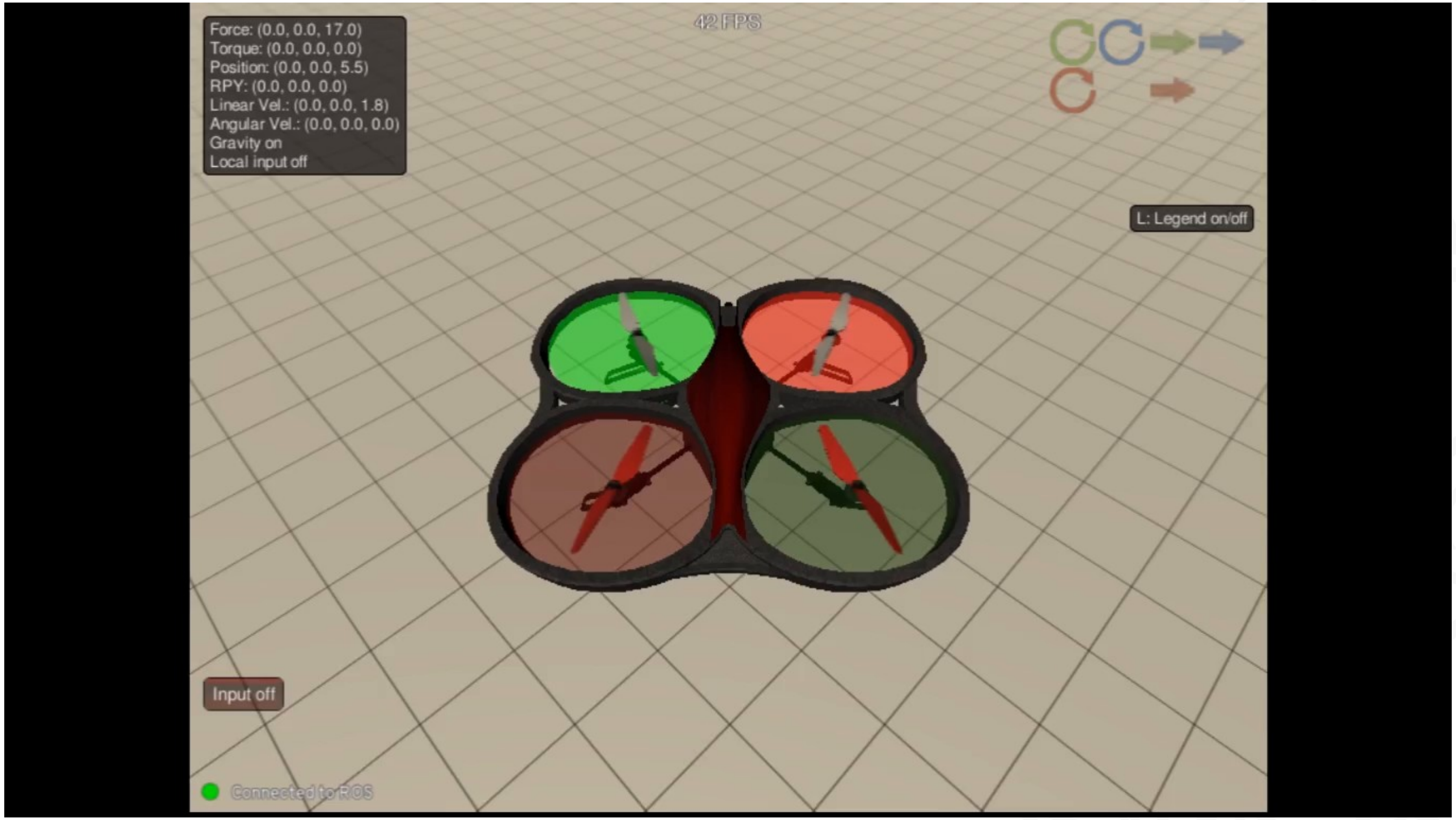
DQN learning for Takeoff

Learning rate = 0.001
Epsilon = 1
Epsilon decay = 0.995
Batch size = 64
Number of episodes = 64
Discount factor = 0.9





Takeoff demonstration

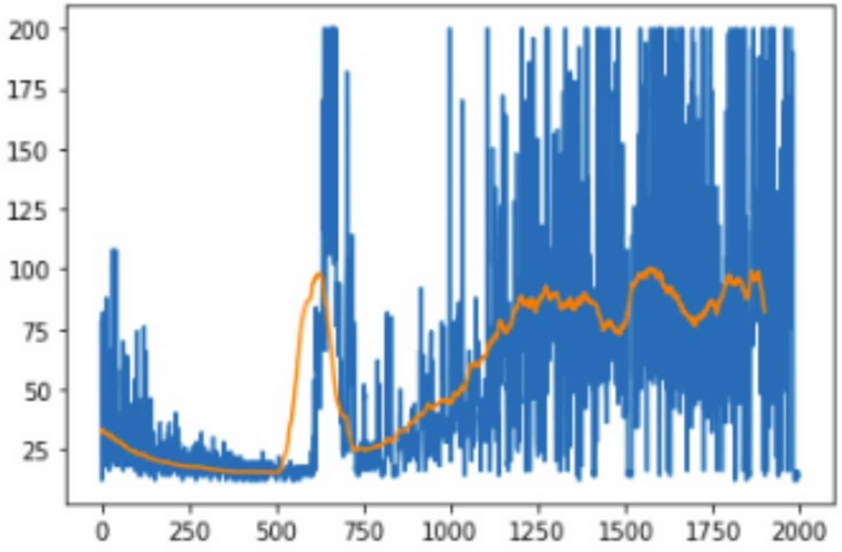


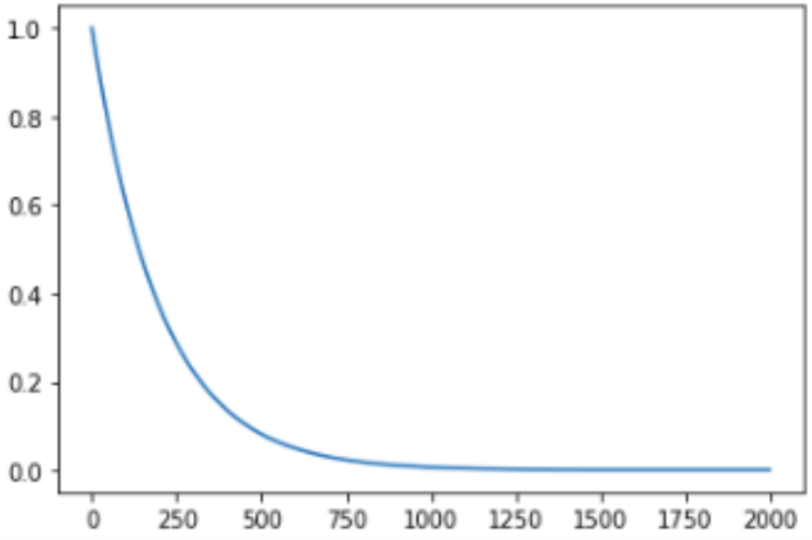
Hovering

- State : Quadcopter position and linear acceleration from simulation environment
- Action : Control command to quadcopter as thrust force
- Reward: Negative absolute distance from quadcopter position till the target height(10 m) minus linear acceleration
- Bonus: 10 when distance less than 3 m
- Episode termination: 5 s elapse(penalty -10)

DQN learning for Hovering

Learning rate = 0.001
Epsilon = 1
Epsilon decay = 0.995
Batch size = 64
Number of episodes = 64
Discount factor = 0.9





Hovering demonstration

Conclusion

- DQN is clearly capable of learning direct thrust (force) control to perform takeoff and landing in 1 DoF dynamics.
- The simulated environment is computationally frugal to apply extensive learning in complicated experiments (in the future)

Future Works

- Gradually introduce more degree-of-freedoms to increase the fidelity of the dynamics
- Import more accurate inertia models, introduce aerodynamics and energy models
- Advanced RL algorithm (PPO) scheduled to replace DQN