CS6301

Intro to Robot Manipulation and Navigation

Team 13

RL-based Top-Down Grasping with a focus on Selective Shapes

1. Abhishek Chauhan
2. Shreeprasad Sonar
3. Vedang Wartikar

About the Project

- **Top-Down Grasping**: Involves the robot positioning itself above the target object for manipulation where the robot's camera system is oriented downward, allowing it to analyze the scene from an overhead viewpoint

- **Reinforcement Learning**: A type of machine learning concept where an **agent *(DQN/DDPG)*** learns by interacting with an **environment *(robot arm + target object)***, receiving feedback in the form of rewards or penalties for its **actions *(grasps)***

- **Selective Shape Manipulation**: Extend/Modify the RL-based approach to grasp specific shapes using the Robot's arm. Tweak the rewards to ***incentivize the grasp of a specific object*** that user specifies

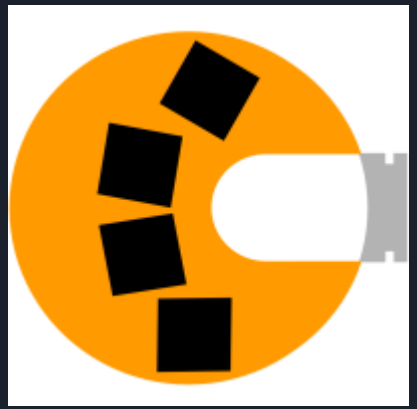
Simulation Environment

OpenAI Gym

\+ **stable-baselines**

Forked from
openai/baselines

⟺



PyBullet

Deep Q - Network (DQN)

1. Reinforcement Learning Framework: In reinforcement learning, an agent learns to interact with an environment by taking actions to **maximize a cumulative reward.**

2. Q-Learning: Q-learning is a classic reinforcement learning technique that aims to learn the quality (Q-value) of taking a specific action in a given state. The **Q-value represents the expected cumulative reward** the agent will receive by taking that action and following the optimal policy thereafter.

3. Deep Q-Network: DQN uses a neural network to approximate the Q-value function. The neural network takes the **environment's state as input** and **outputs Q-values** for each possible action. This allows the agent to estimate the best action to take in a given state.

Deep Deterministic Policy Gradient (DDPG)

1. Continuous Action Spaces: Continuous action spaces **(DDPG)** are spaces where the **action space is a continuous set of values**, such as the position of an object. Discrete action spaces **(DQN)** are spaces where the **action space is a set of discrete values**, such as the buttons that an agent can press.

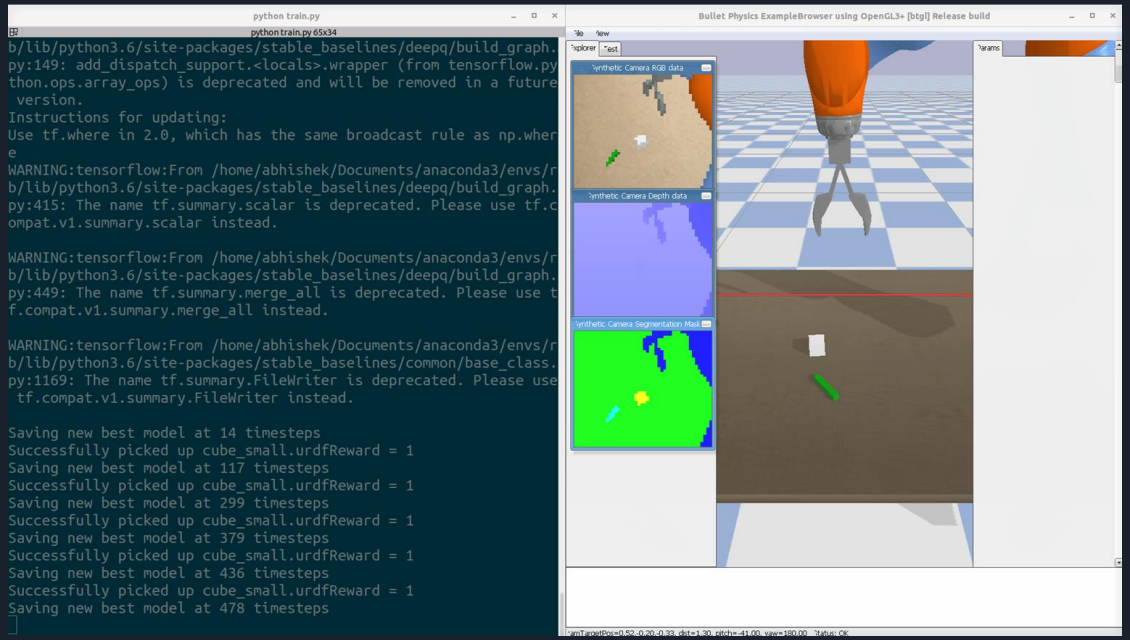
2. Actor-Critic Architecture:
Actor Network: The actor is responsible for l**earning the policy directly**, mapping states to specific actions.
Critic Network: The critic evaluates the actions chosen by the actor by **estimating the Q-value function.**

3. Deterministic Policy Gradient: Allows the actor network to **directly learn the optimal policy in continuous action spaces** without the complexities involved in learning a stochastic policy.

Reward Function

The reward function in reinforcement learning serves as a crucial **guiding principle**.

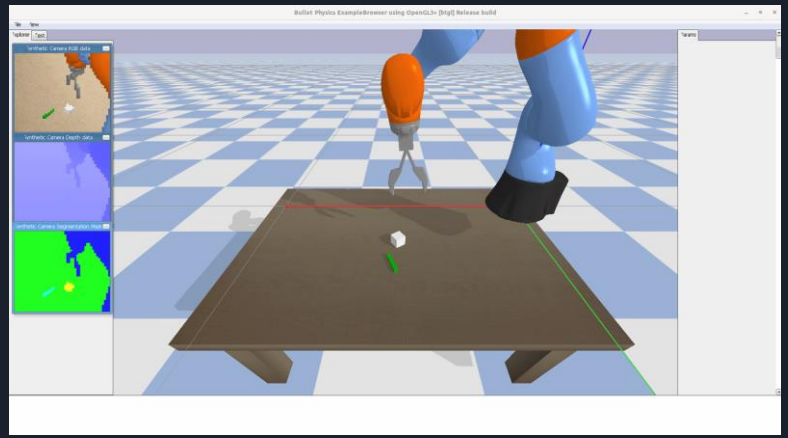
If it successfully picks up the correct object, then reward given is 1,
otherwise if it grabs the incorrect object, then -1 reward is given.



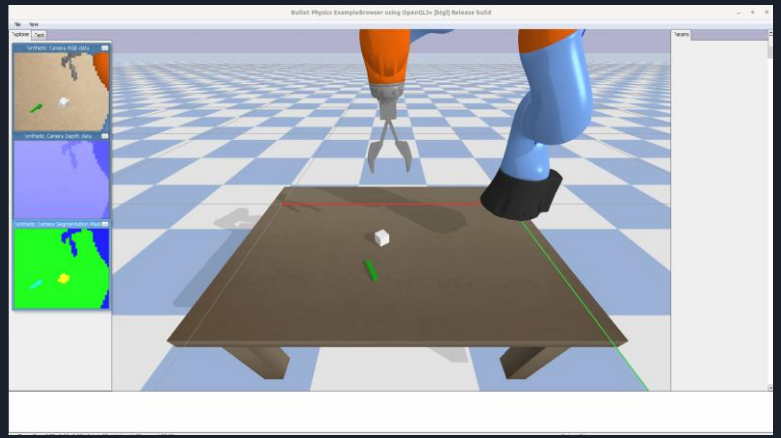
Training

For the environment, we added 2 shapes to the environment , a cuboid and a cube . The intended behaviour is to pick up the cube as specified by the user.

Then, we've trained our reinforcement learning (RL) model using two algorithms, DQN and DDPG. Our model employs RGB as the perception layer, the custom reward function and learns for a minimum of 10000 timesteps.



Successfully picking up the cube among different objects



Failed to pick up the cube among different objects

Demo Video

The demo video for the project is listed here : https://youtu.be/KAYMQ7sokI4

Current Status of our Work

- **Completed:**
  - Create a simulation environment
  - Train DQN and DDPG algorithms for grasping using RGB sensor system
  - Test the model on different target objects (cube, cuboid, sphere)
- **In Progress:**
  - Modify the reward function to incentivize grasping of select shapes
  - Fine-tune parameters of the RL algorithms to optimize the model

– End of presentation –