

Deep Reinforcement Learning for Robotic Grasping from Octrees

Learning Manipulation from Compact 3D Observations

Master's Thesis

June 25, 2021

Andrej Orsula

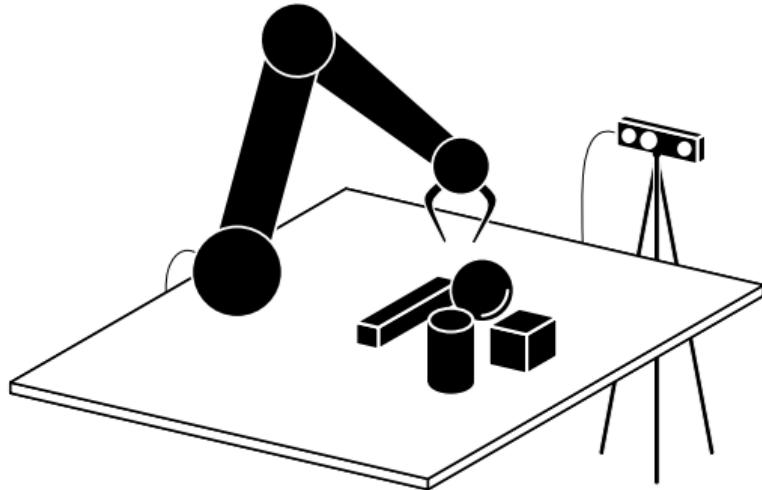
Aalborg University
Denmark



AALBORG UNIVERSITY



Vision-Based Robotic Grasping of Diverse Objects



Vision-Based Robotic Grasping of Diverse Objects

Approach



Approaches

- ▶ Analytical
- ▶ Empirical
 - ▶ Supervised learning
 - ▶ Imitation learning
 - ▶ Reinforcement learning

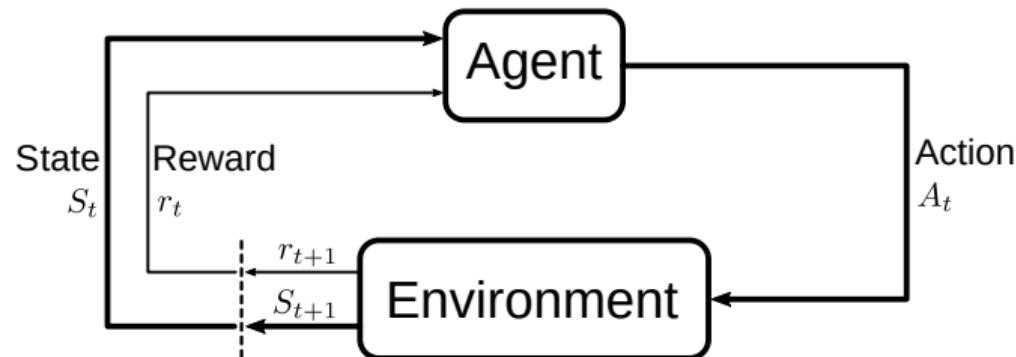


Vision-Based Robotic Grasping of Diverse Objects

Approach

Approaches

- ▶ Analytical
- ▶ Empirical
 - ▶ Supervised learning
 - ▶ Imitation learning
 - ▶ **Reinforcement learning**





Task Definition

Agent

- ▶ High-level controller
 - ▶ Gripper pose
 - ▶ Gripper action

Environment

- ▶ Objects
- ▶ Robot
 - ▶ Low-level controllers
- ▶ Physics and visuals

Episodic Task

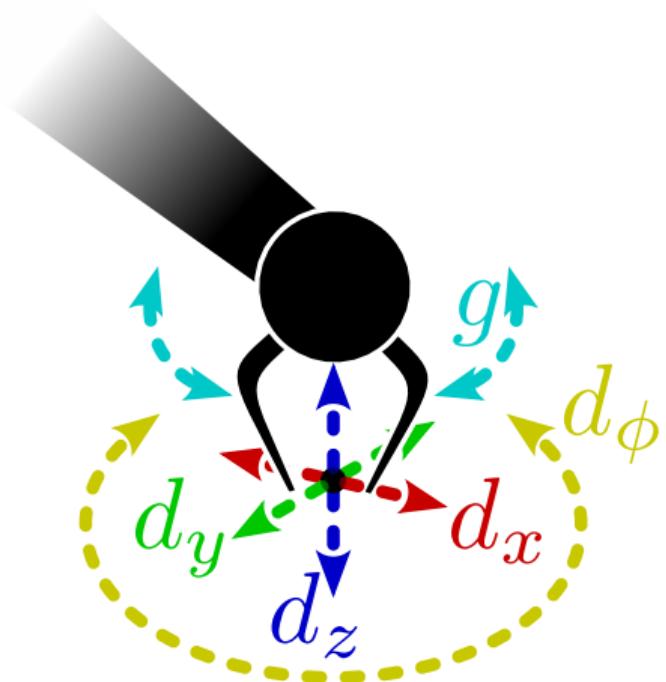
- ▶ Success
 - ▶ Lifting an object
- ▶ Failure
 - ▶ Pushing all objects away
- ▶ Max 100 time steps
 - ▶ ~40 s (simulation)

Task Definition

Action Space

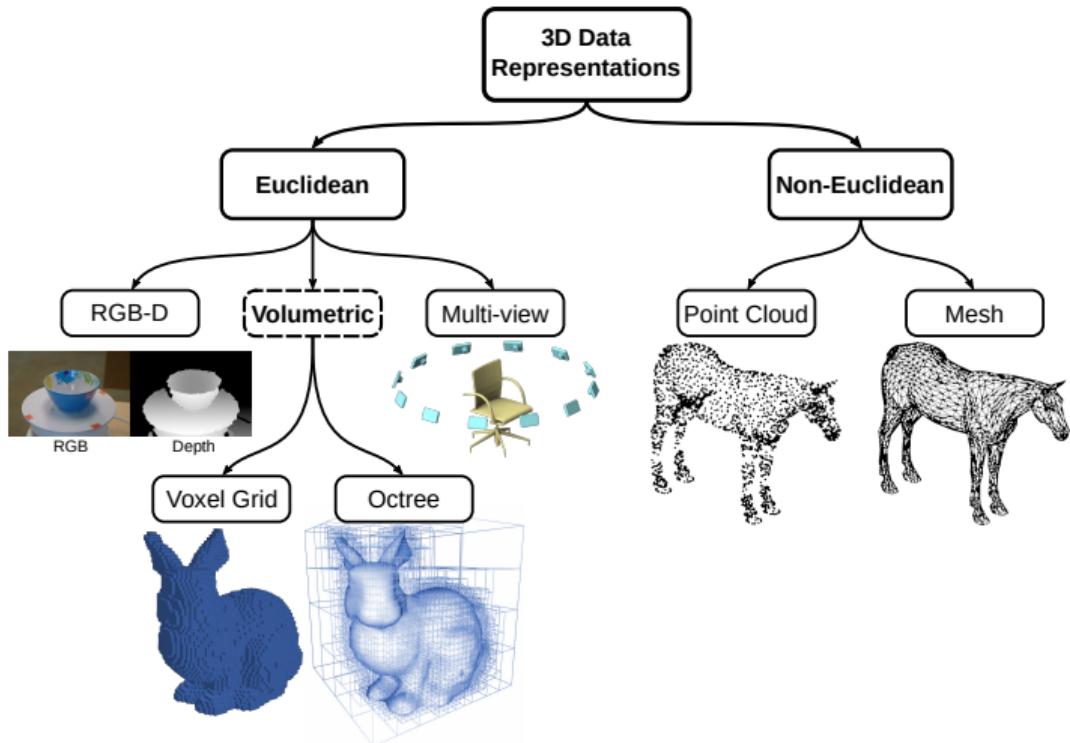
Actions in Cartesian Space

- ▶ Translational displacement
 - ▶ d_x
 - ▶ d_y
 - ▶ d_z
- ▶ Gripper rotation
 - ▶ d_ϕ
- ▶ Gripper actions (open/close)
 - ▶ g



Task Definition

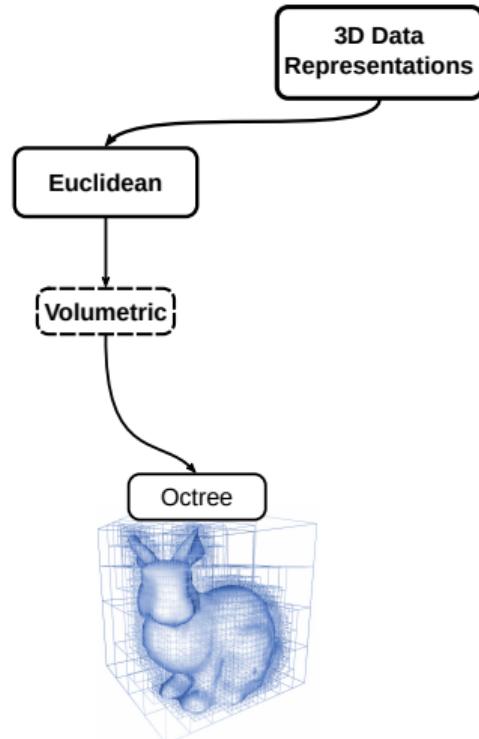
Observation Space





Task Definition

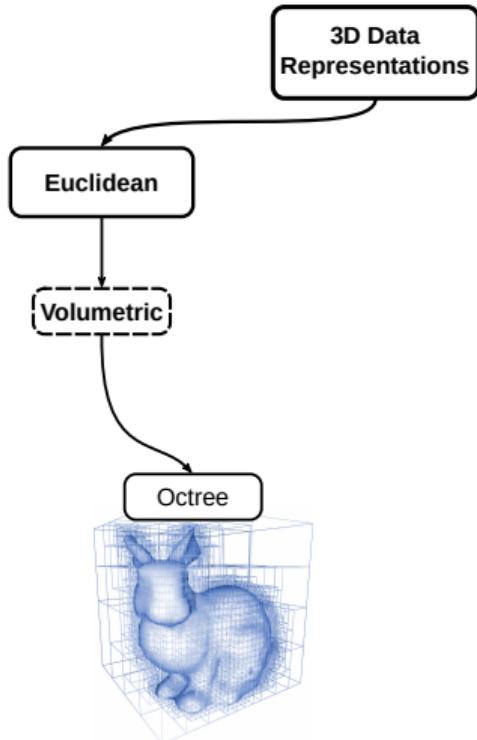
Observation Space





Task Definition

Observation Space



Proprioceptive Observations

- ▶ Gripper position
- ▶ Gripper rotation
- ▶ Gripper state



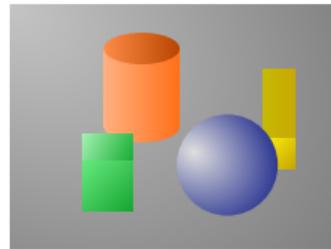
Task Definition

Observation Space - Construction of Octree

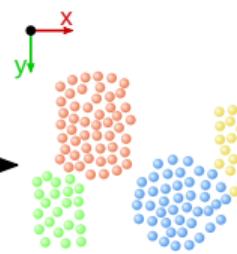
Depth Map



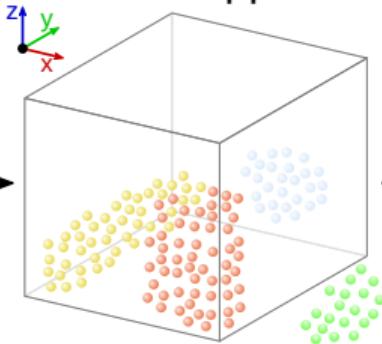
RGB Image



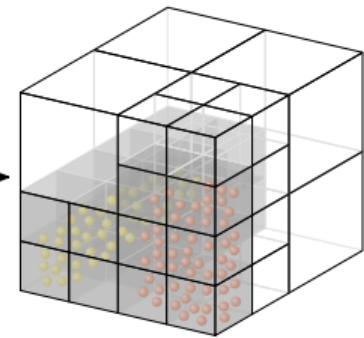
Point Cloud



Transformed
and Cropped



Octree





Task Definition

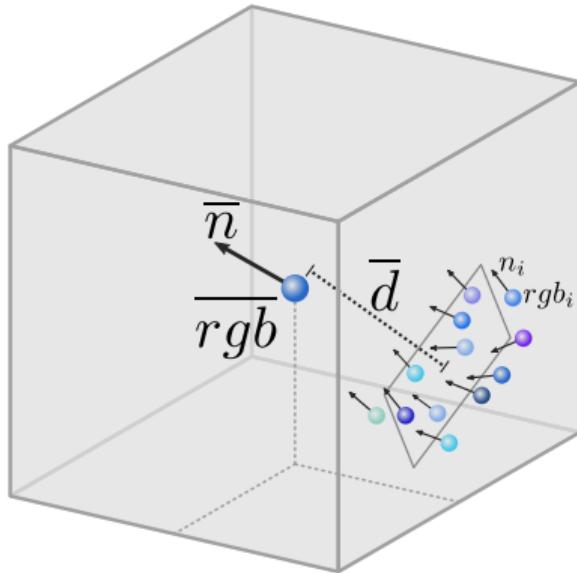
Observation Space - Features and Stacks

Features

- ▶ Spatial
 - ▶ Average normal vector \bar{n}
 - ▶ Average distance to points \bar{d}
- ▶ Colour
 - ▶ Average intensity of RGB channels \overline{rgb}

Observation Stacking

- ▶ Three consecutive observations





Task Definition

Reward Function

Composite Reward

- ▶ Reach
 - ▶ $+1 (7^0)$
- ▶ Touch
 - ▶ $+7 (7^1)$
- ▶ Grasp
 - ▶ $+49 (7^2)$
- ▶ Lift
 - ▶ $+343 (7^3)$

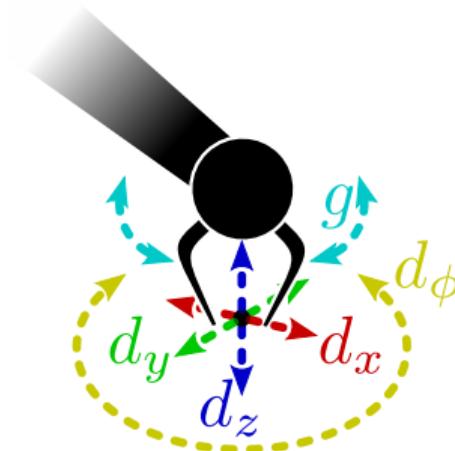
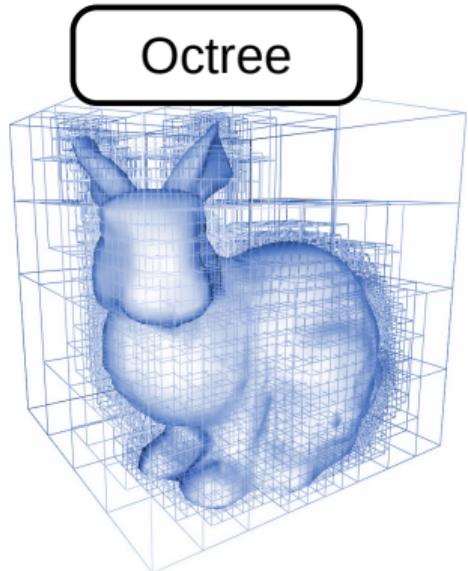
Recurring Reward

- ▶ Collision with ground/table
 - ▶ -1
- ▶ Incentive to act quickly
 - ▶ -0.005



Task Definition

Summary



Reward Function

- ▶ Composite
 - ▶ Reach
 - ▶ Touch
 - ▶ Grasp
 - ▶ Lift
- ▶ Collision with ground/table
- ▶ Incentive to act quickly



Reinforcement Learning

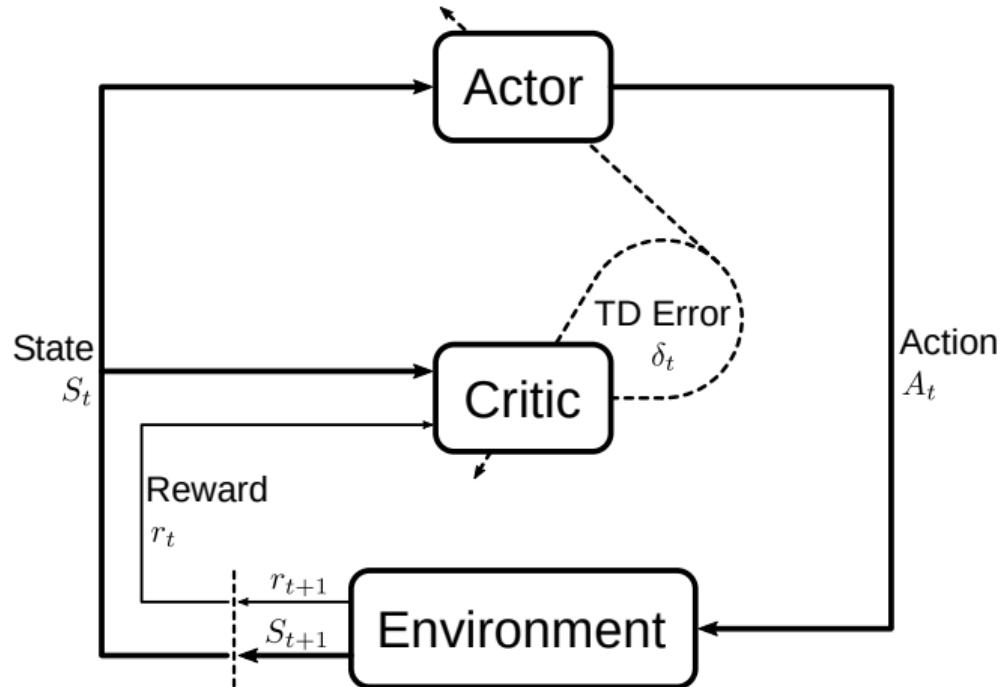
Algorithms

Actor-Critic Algorithms

- ▶ TD3
- ▶ SAC
- ▶ TQC

Implementation

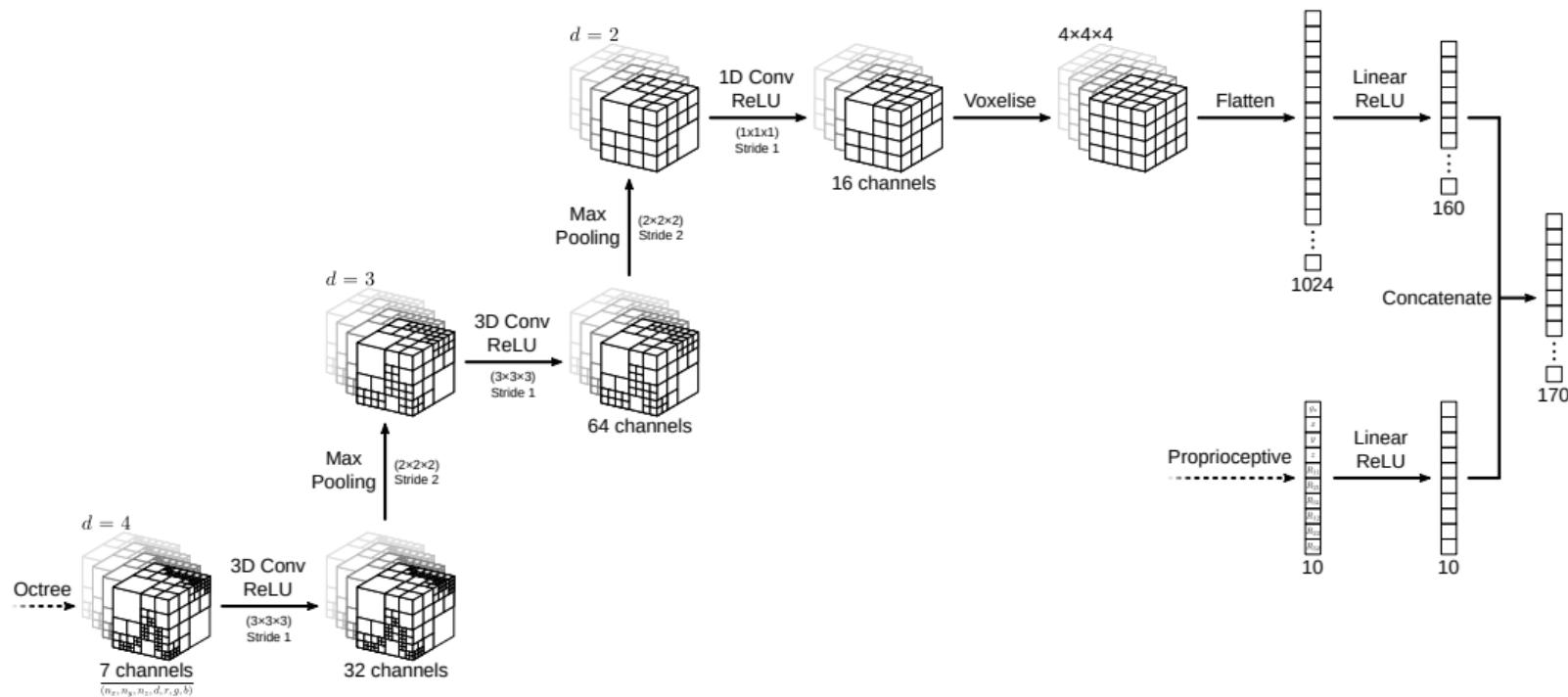
- ▶ Stable Baselines3





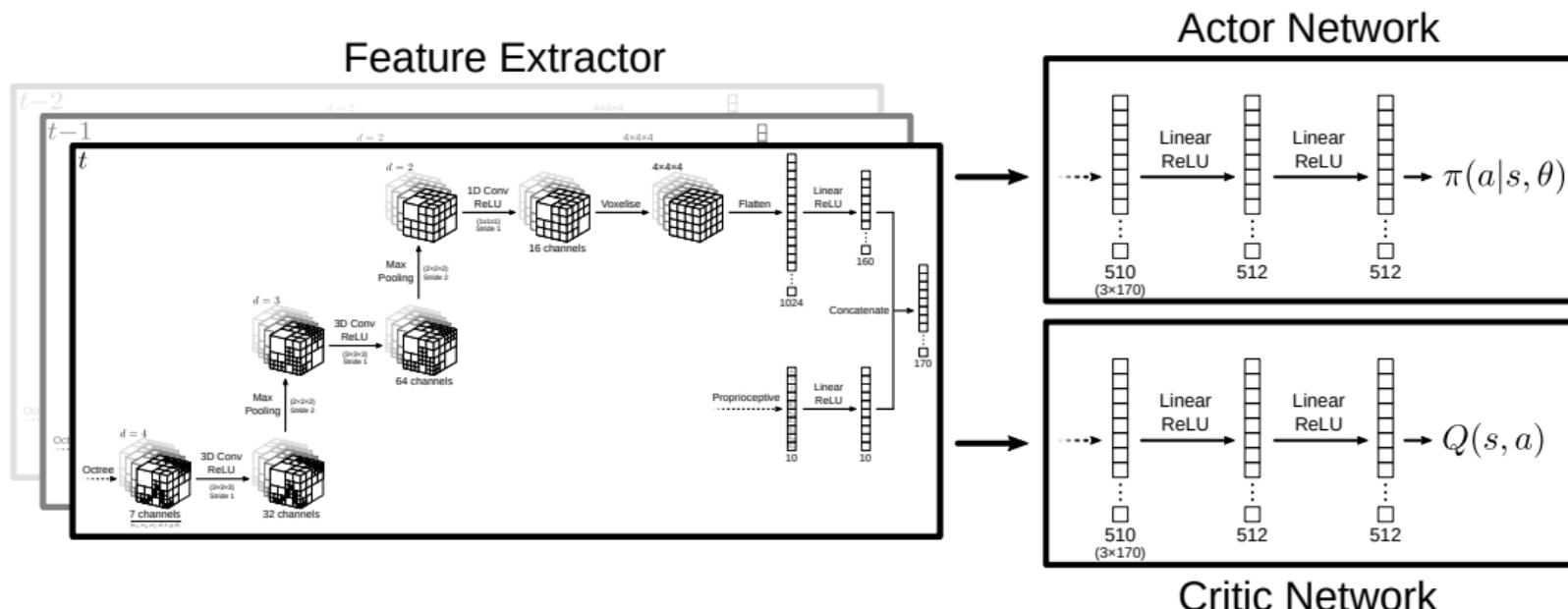
Deep Reinforcement Learning

Octree-Based Feature Extractor



Deep Reinforcement Learning

Full Actor-Critic Network Architecture





Simulation Environment

Selection

Simulators

- ▶ MuJoCo
- ▶ PyBullet
- ▶ Gazebo Classic
- ▶ Ignition Gazebo
- ▶ Isaac
- ▶ Webots
- ▶ Unreal Engine
- ▶ Unity
- ▶ Unigine
- ▶ RaiSim
- ▶ ...

Simulation Environment Selection

Simulators

- ▶ MuJoCo
- ▶ PyBullet
- ▶ Gazebo Classic
- ▶ **Ignition Gazebo**
- ▶ Isaac
- ▶ Webots
- ▶ Unreal Engine
- ▶ Unity
- ▶ Unigine
- ▶ RaiSim
- ▶ ...



Simulation Environment

Ignition Gazebo



Physics



Rendering





Simulation Environment

Ignition Gazebo

Physics



Gym-Ignition

- ▶ Interface for Ignition Gazebo
- ▶ Tooling for creation of OpenAI Gym environments

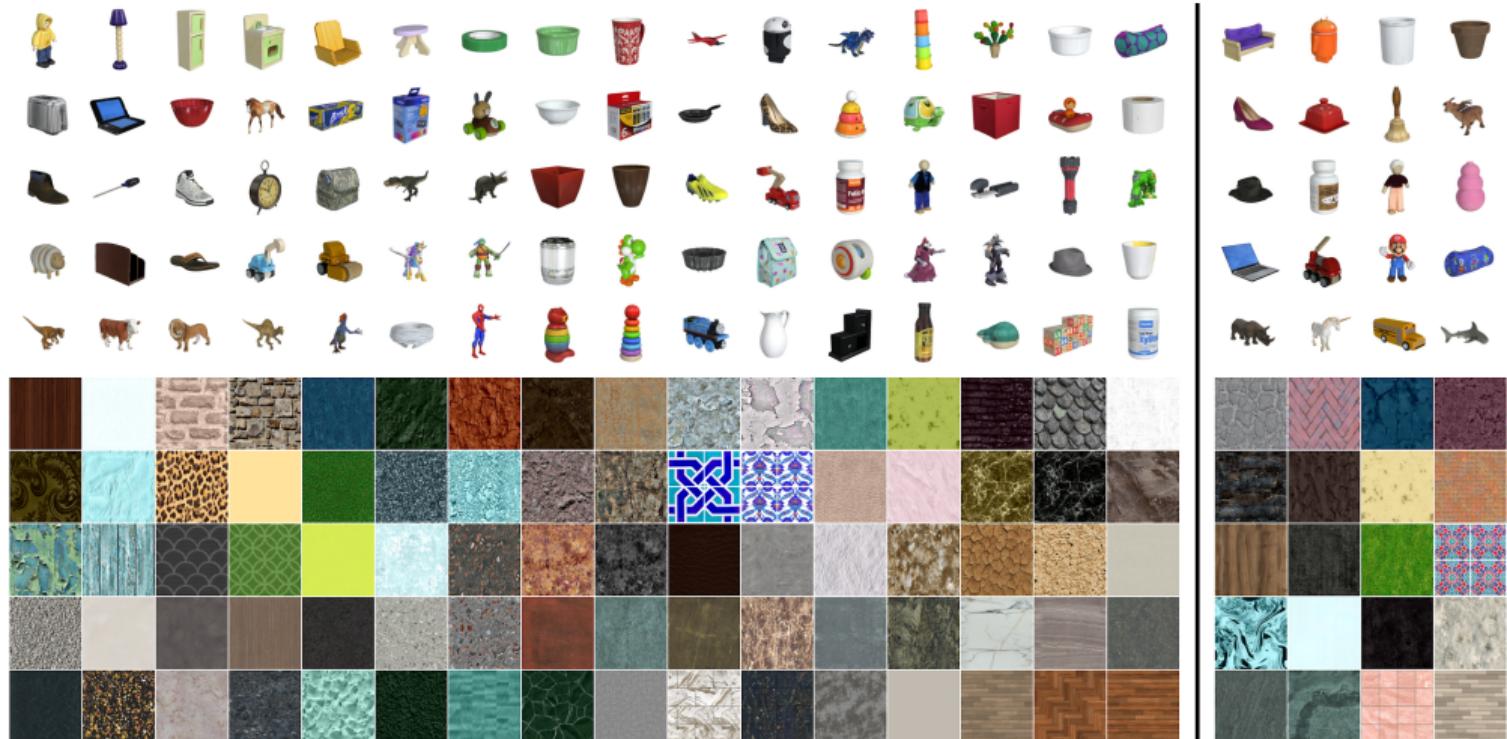
Rendering





Simulation Environment

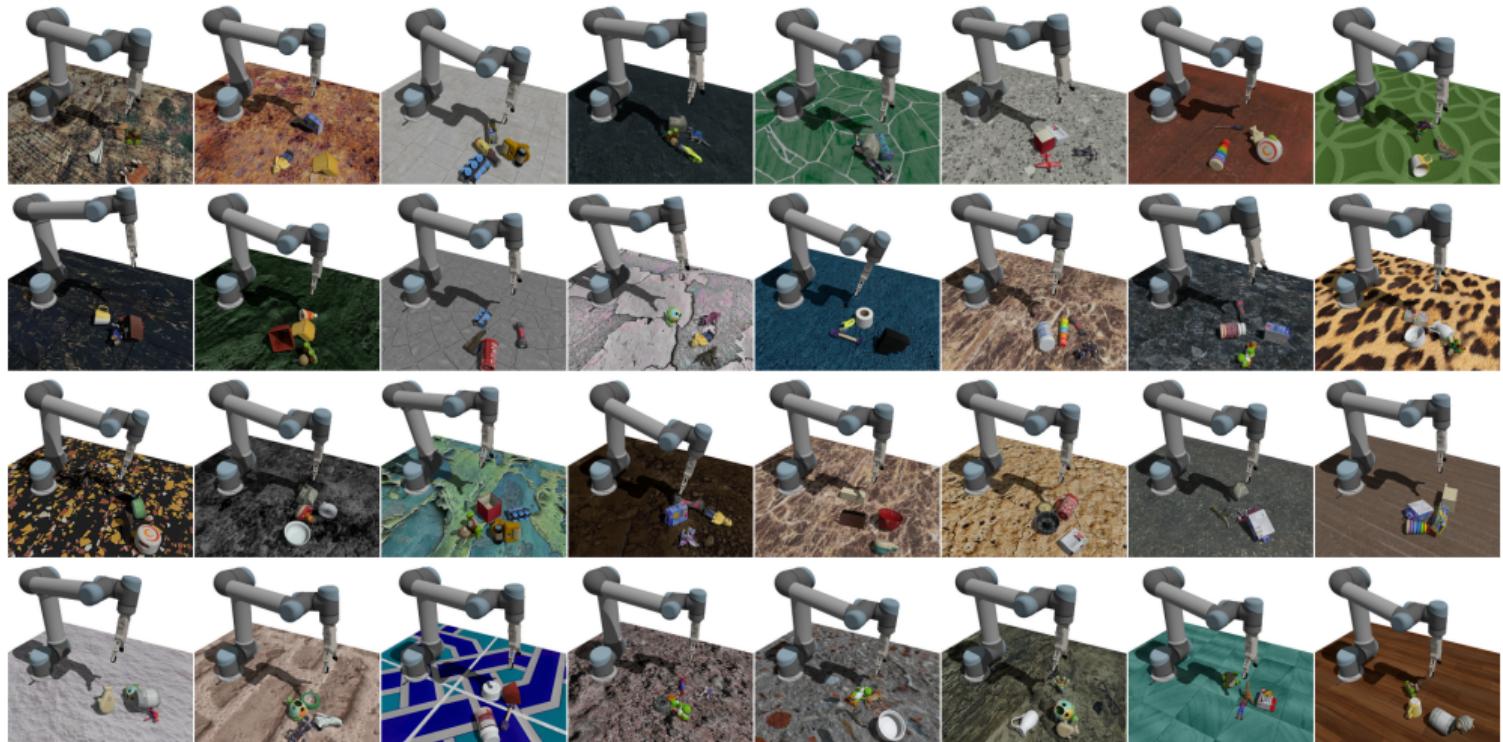
Domain Randomization





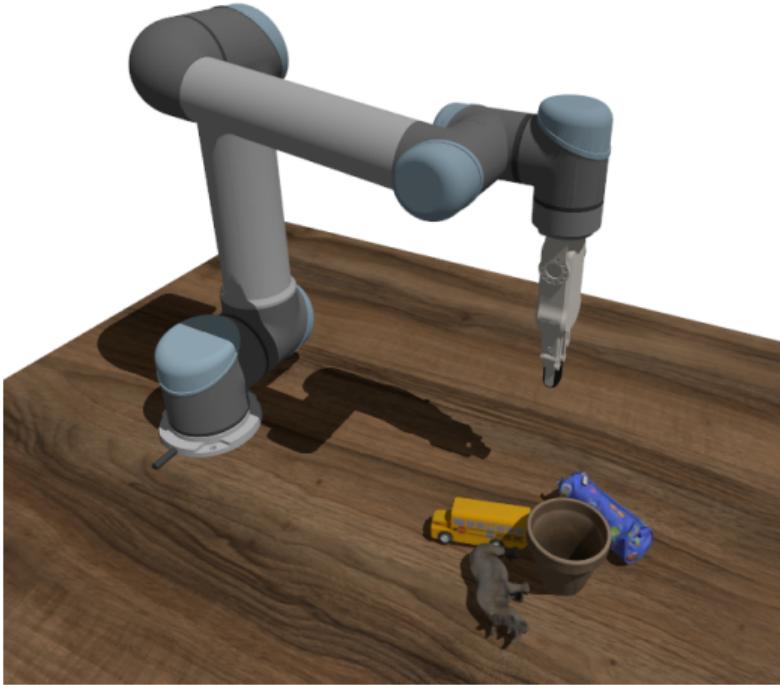
Simulation Environment

Domain Randomization



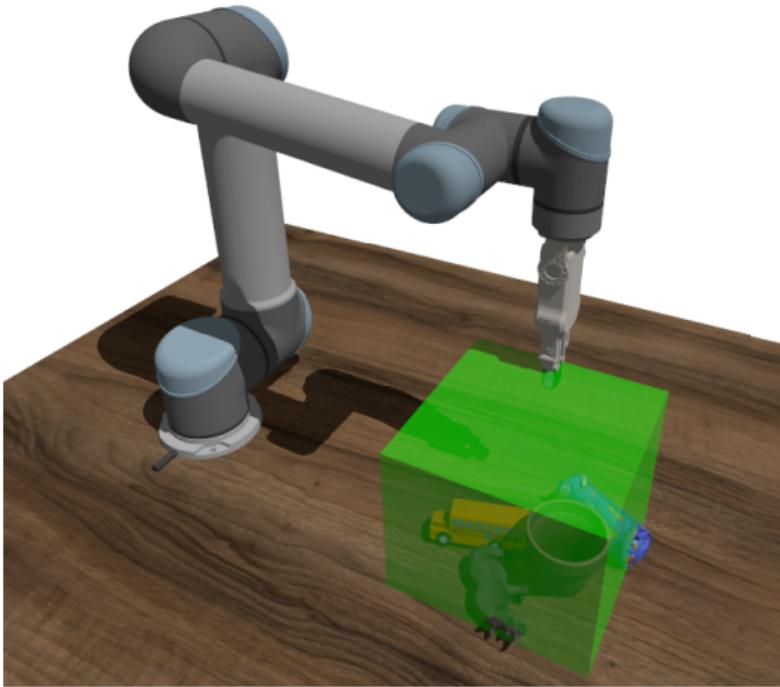
Simulation Environment

Environment for Grasping



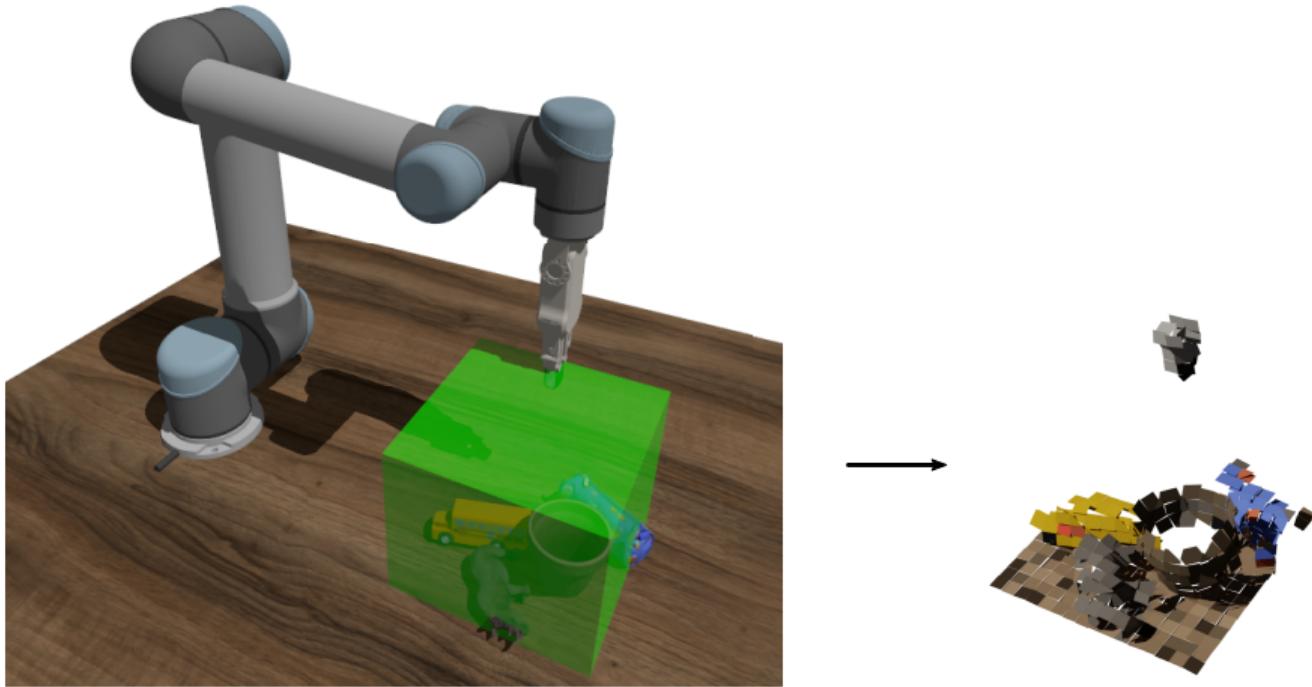
Simulation Environment

Environment for Grasping



Simulation Environment

Environment for Grasping



Thank you for your time



AALBORG UNIVERSITY