Manipulating Handles in Domestic Environments

R&D Defense

Anirudh Narasimamurthy Jayasimha
March 24rd, 2020

Supervisors
Prof. Dr. Paul G. Plöger
Alex Mitrevski

Hochschule
Bonn-Rhein-Sieg
University of Applied Sciences

Bonn-Aachen
International Center for
Information Technology

Introduction

Handles in domestic environment

- Handles are articulated objects.
- Used to manipulate an underlying attached object.
- Applications of handle manipulation :
  - Access new areas in robot environment.
  - Helping humans in performing tasks eg. Opening drawer and bringing something.

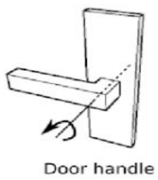
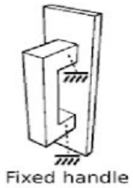
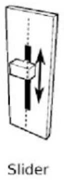
Revolute joints	Fixed joints	Prismatic joints
		

Figure 1. Common types of joints and their kinematic models found in domestic environments. [1]

Key Constraints and Assumptions

- **No locking mechanism involved.**
- Small subset of common domestic handles.
- No collisions with the environment.
- No self collisions.
- Stable when manipulating the handle.
- Learning based algorithms.

Figure 2. The drawers have no locks. The silver marking on the floor denotes the initial position to run the handle manipulation pipeline.

Key Constraints and Assumptions

- No locking mechanism involved.
- **Small subset of common domestic handles.**
- No collisions with the environment.
- No self collisions.
- Stable when manipulating the handle.
- Learning based algorithms.

Figure 3. Different types of handles found in domestic environments. {Fixed handle, Round handle, Lever handle}.[2]

Key Constraints and Assumptions

- No locking mechanism involved.
- Small subset of common domestic handles.
- **No collisions with the environment.**
- **No self collisions.**
- Stable when manipulating the handle.
- Learning based algorithms.

No damage to the robot or the environment

Key Constraints and Assumptions

- No locking mechanism involved.
- Small subset of common domestic handles.
- No collisions with the environment.
- No self collisions.
- **Stable when manipulating the handle.**
- Learning based algorithms.

Source : Toyota

Source : Boston Dynamics

Figure 4. Humanoid robots have a smaller “support polygon” and hence are less statically stable.

Key Constraints and Assumptions

- No locking mechanism involved.
- Small subset of common domestic handles.
- No collisions with the environment.
- No self collisions.
- Stable when manipulating the handle.
- **Learning based algorithms.**

Contributions of the Project

- ***State of the art (Systematic literature survey).***
  - **Classification of subtasks.**
  - **Advantages and Disadvantages of algorithms.**
- *Manipulation of handles in DoorGym simulation environment.*
  - Adapting Toyota HSR URDF to MuJoCo environment.
  - Training handle manipulation models.
- *Development of the pipeline on Toyota HSR.*
  - Deep Learning model for handle detection.
  - MLE model for sampling best grasp pose.
  - Complete handle manipulation pipeline.

Contributions of the Project

- *State of the art (Systematic literature survey).*
  - Classification of subtasks.
  - Advantages and Disadvantages of algorithms.
- ***Manipulation of handles in DoorGym simulation environment.***
  - **Adapting Toyota HSR URDF to MuJoCo environment.**
  - **Training handle manipulation models.**
- *Development of the pipeline on Toyota HSR.*
  - Deep Learning model for handle detection.
  - MLE model for sampling best grasp pose.
  - Complete handle manipulation pipeline.

Contributions of the Project

- *State of the art (Systematic literature survey).*
  - Classification of subtasks.
  - Advantages and Disadvantages of algorithms.

- *Manipulation of handles in DoorGym simulation environment.*
  - Adapting Toyota HSR URDF to MuJoCo environment.
  - Training handle manipulation models.

- ***Development of the pipeline on Toyota HSR.***
  - **Deep Learning model for handle detection.**
  - **MLE model for sampling best grasp pose.**
  - **Handle manipulation pipeline.**

State of the Art

General Pipeline

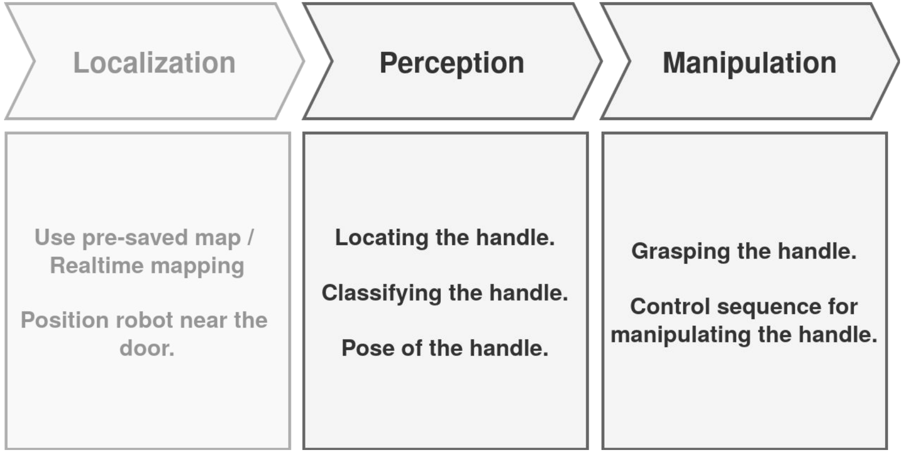


Figure 5. Division of the task of “Manipulation of Handles” into its logical subtasks.

- The task of “Manipulating handles” can be divided based on the logical subtasks [5]:
  - Localization
  - Perception
  - Manipulation

Perception - State of the Art

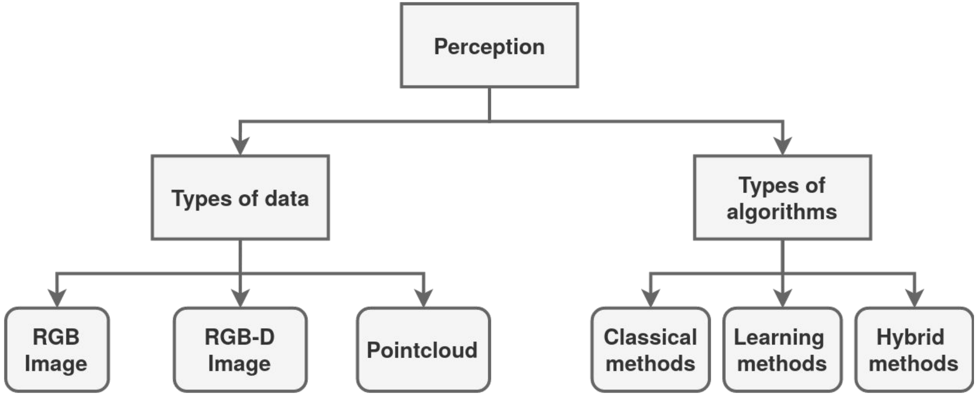


Figure 6. Perception algorithms divided based on the algorithms used and type of data obtained by the sensor.

- Detecting the handles :
  - RGB provides better results in detection accuracy and speed.
  - Learning based methods provide state of the art results.
- Pose detection :
  - Pointcloud data is preferred.
  - Classical methods are more often used.

Manipulation - State of the Art

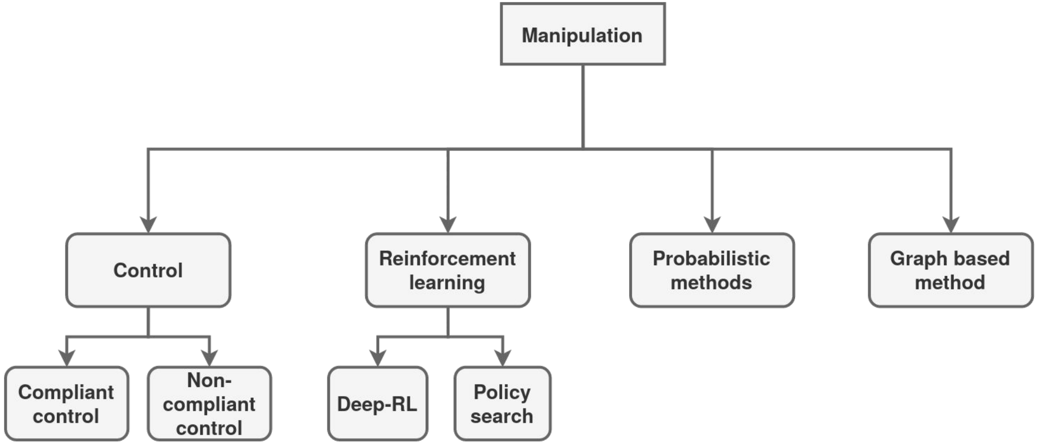


Figure 7. Different classes of Manipulation algorithms used for the task of “Manipulating handles”.

- Characteristics of the algorithms :
  - Learning / not-learning
  - Modifications to be made to the robot environment.
  - Number of rollouts.
  - Complexity of the algorithm.

Manipulation of Handles in DoorGym Simulation Environment

Simulation for Reinforcement Learning

- **Problems with using RL on real robot :**
  - **Real world is not a perfect MDP.**
  - **Difficulty in collecting rollouts.**
  - **Impossible to set same initial state.**
  - **Partial/noisy state prediction.**

- Advantages of simulation :
  - Faster collection of rollouts.
  - Approximate model of the environment.
  - Repeatable and reproducible.
  - Domin randomization

Simulation for Reinforcement Learning

- Problems with using RL on real robot :
  - Real world is not a perfect MDP.
  - Difficulty in collecting rollouts.
  - Impossible to set same initial state.
  - Partial/noisy state prediction.

- **Advantages of simulation :**
  - **Faster collection of rollouts.**
  - **Approximate model of the environment.**
  - **Repeatable and reproducible.**
  - **Domain randomization**

DoorGym Environment

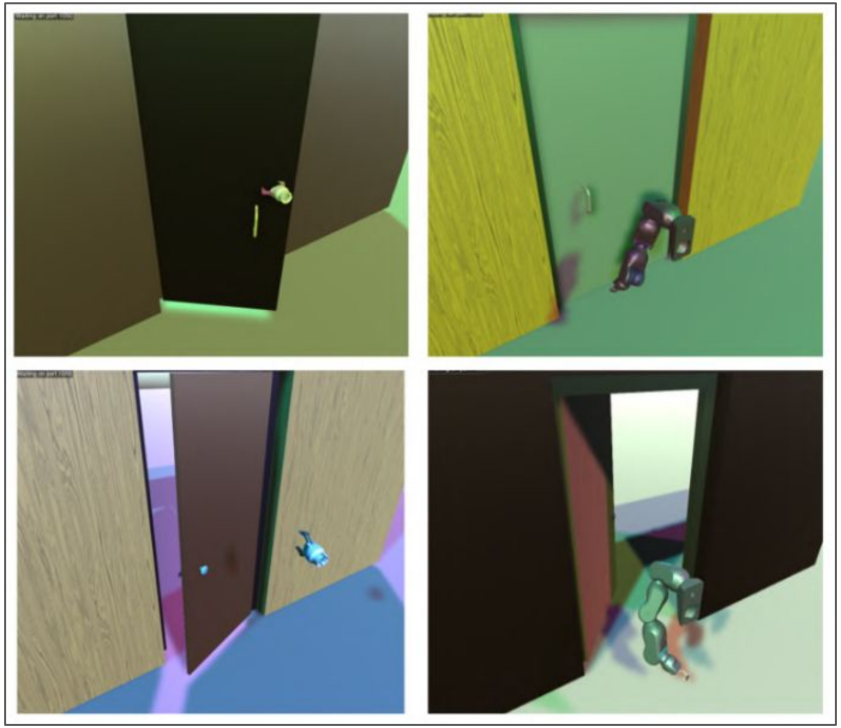


Figure 8. Randomized environment generated by DoorGym.
(https://github.com/PSVL/DoorGym)

- DoorGym [2] : Specialized MuJoCo simulation environment.
- Different handles are modeled :
  - Pull handle
  - Lever handle
  - Round handle
- Domain randomization by changing the lighting, material and environment properties.
- Pretrained baseline RL models.

Robot properties in DoorGym

- Continuous action space.
  - All possible motor commands.
- Continuous state space.
  - {Robot orientations, joint angles, ..} $\in$ Internal states
  - {All environment states} $\in$ External states
- High dimensional (8 DoF)

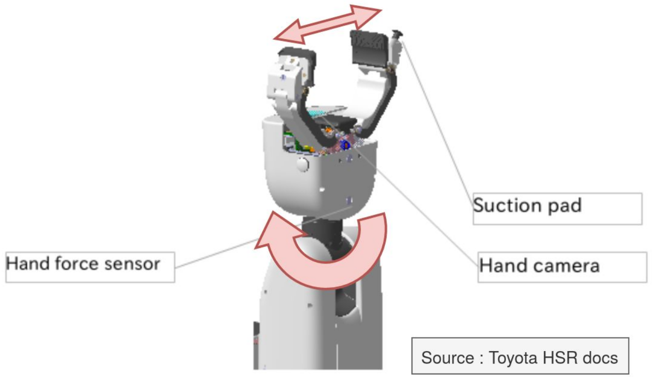


Figure 9. Toyota HSR arm and wrist. Red arrows denotes the rotation and translation action space.

DoorGym - Toyota HSR

Simulation Steps

- Convert URDF into Mujoco MJCF format
- Adding controls to individual joints
- Setting collisions to robot parts
- Increasing damping parameters of handle
- Place the robot into the correct position

URDF Adaption: Convert URDF into Mujoco MJCF format; Adding controls to individual joints; Setting collisions to robot parts

Environment Modelling: Increasing damping parameters of handle; Place the robot into the correct position

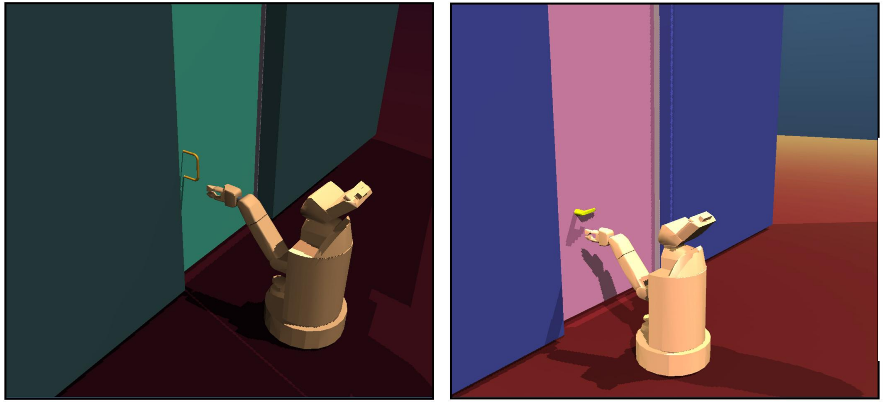
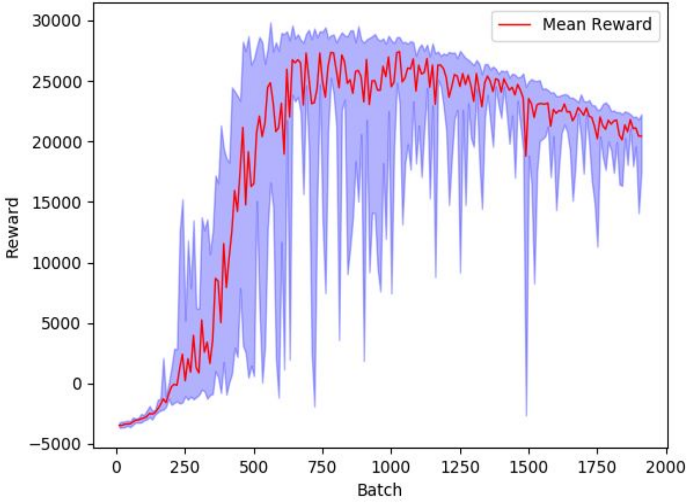


Figure 10. Toyota HSR ported into randomized DoorGym environments.

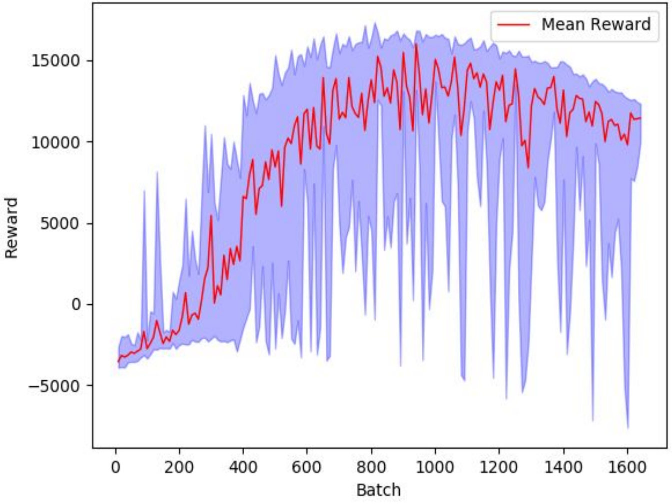
Reward Function

$$r_t = \underbrace{-a_0 d_t}_{\text{fingertip distance}} - \overbrace{a_1 log(d_t + \alpha)}^{\text{log distance}} - \underbrace{a_2 o_t}_{\text{required end effector orientation}} - a_3 \overbrace{\sqrt{u_t^2}}^{\text{control input}} + \underbrace{a_4(\phi_t) + a_5 \psi_t}_{\text{door and knob angle}}$$

Proximal Policy Optimization (PPO) - Results



a) Lever handle (lr = 10-4)



b) Fixed handle (lr = 10-4)

Figure 11. Policy for manipulating handles learnt in the DoorGym environment using the Toyota HSR URDF. The continuously improving reward shows that the model is learning the task.

Conclusion

- Adapting a Toyota HSR model to the MuJoCO environment.
- Modifications to the DoorGym environment.
- Trained RL policy for the task of “Manipulating Handles”.
- Robot learn to manipulate the handles but few actions taken by the robot were not realistic.

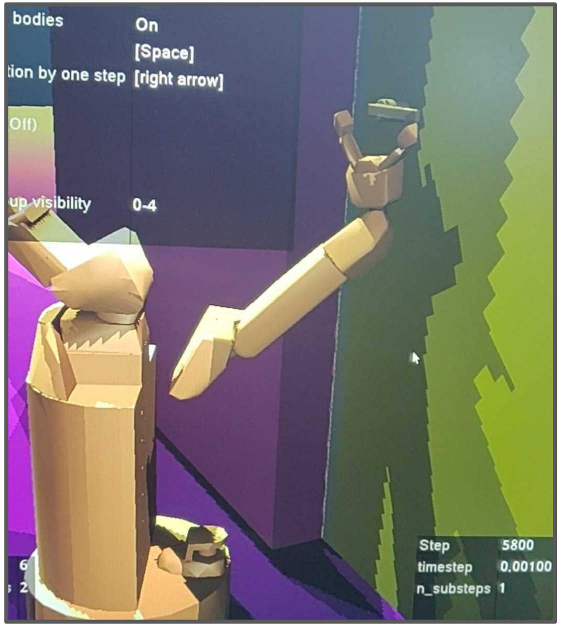


Figure 12. The robot performs unrealistic maneuvers to grasp the handle.

Future Work

- DoorGym Simulation environment
  - Modify simulation parameters for realistic grasping.
  - Add cabinets and refrigerator to the simulation environment.
  - Modeling locking mechanism.
  - *‘Sim to Reality’* transfer.

Pipeline Development and Evaluation on Toyota HSR for Fixed Handles

Pipeline for Manipulating the Handle

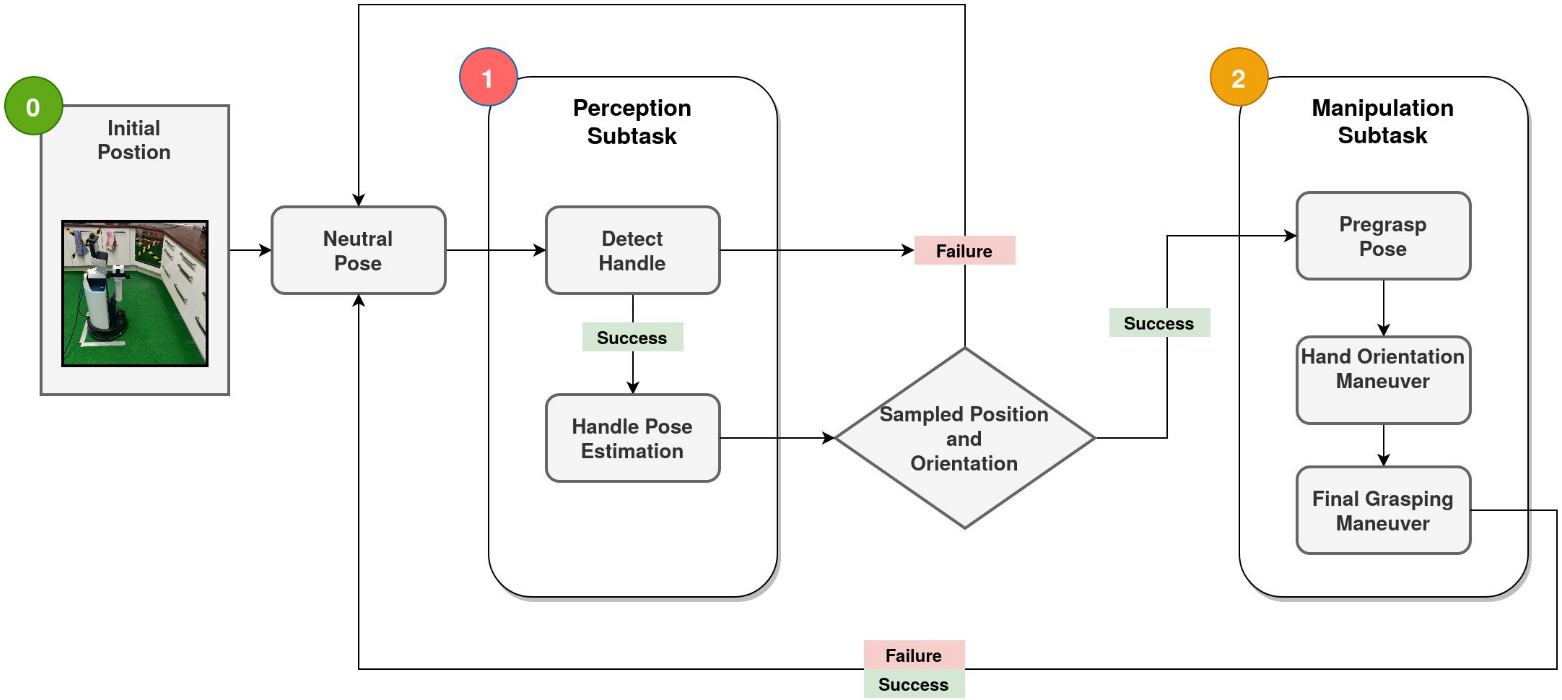
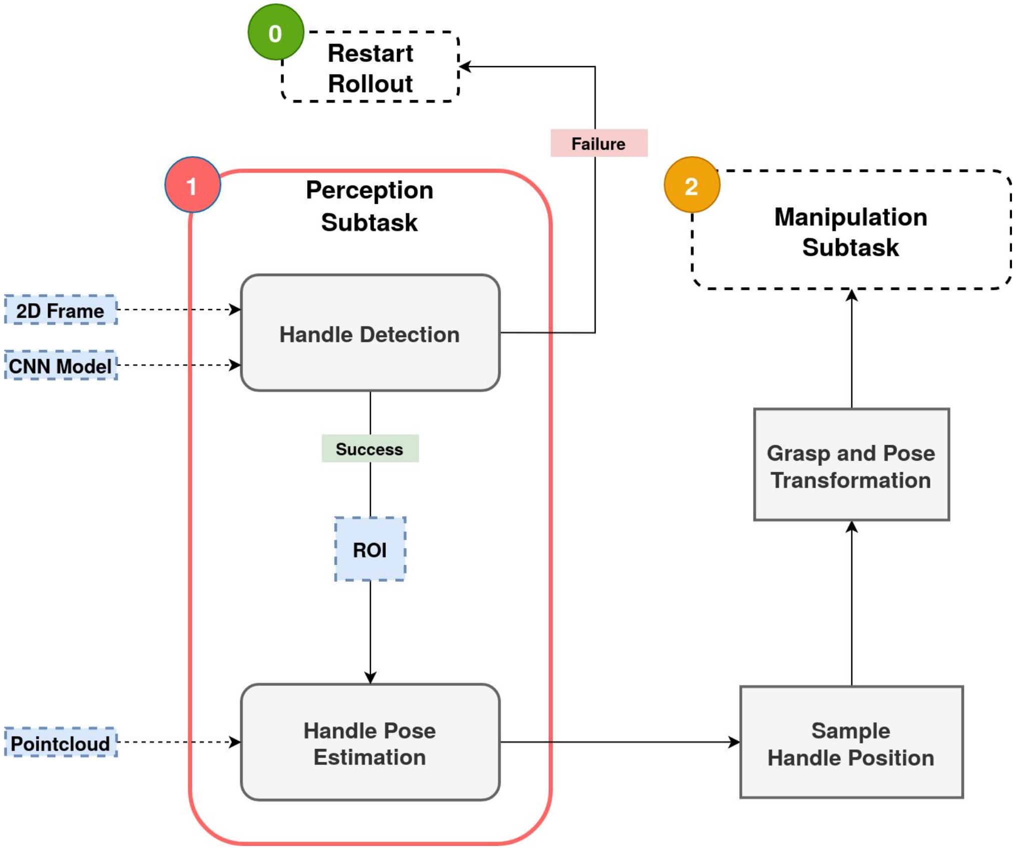


Figure 12. Pipeline showing the initialization of the robot and perception subtask followed by the manipulation subtask.

Perception - Pipeline



- Three models **(FRCNN, YOLO, SSD)** were selected.
- Selection of models based on the following criterias :
  - Detection accuracy
  - Ease of training (Code, Pretrained model)
  - Inference time
- Transfer learning is used for training.
- The pretrained model were trained on COCO and OIDv4 datasets.

Perception - Dataset Collection

Figure 13. Sample images from the dataset showing the different types of handles.

- Classes in collected dataset :
  - Lever handle
  - Fixed handle
  - Round handle
- Dataset details :
  - Total 200 images.
  - {150 /50 } Training/Test dataset split.
  - Augmentations for increasing generalizability.

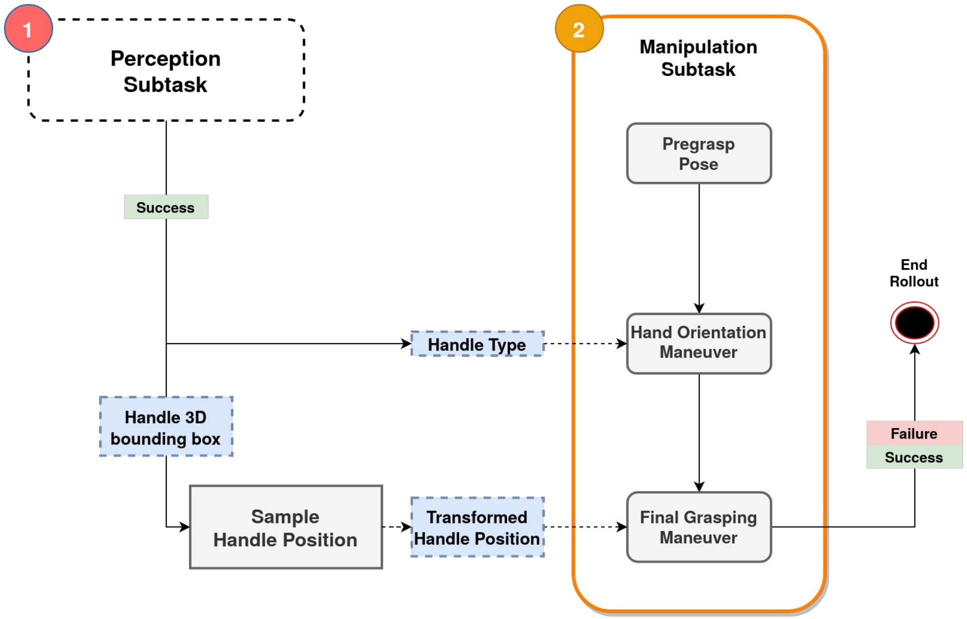
Perception - 2D Detection Results

Figure 14. Handles detected by the trained CNN models.

- Three models (FR-CNN, YOLO, SSD) were trained.
- mAP comparison (More is better):
  - FR-CNN > SSD > YOLO
- Inference timing (Less is better) :
  - YOLO > SSD > FR-CNN

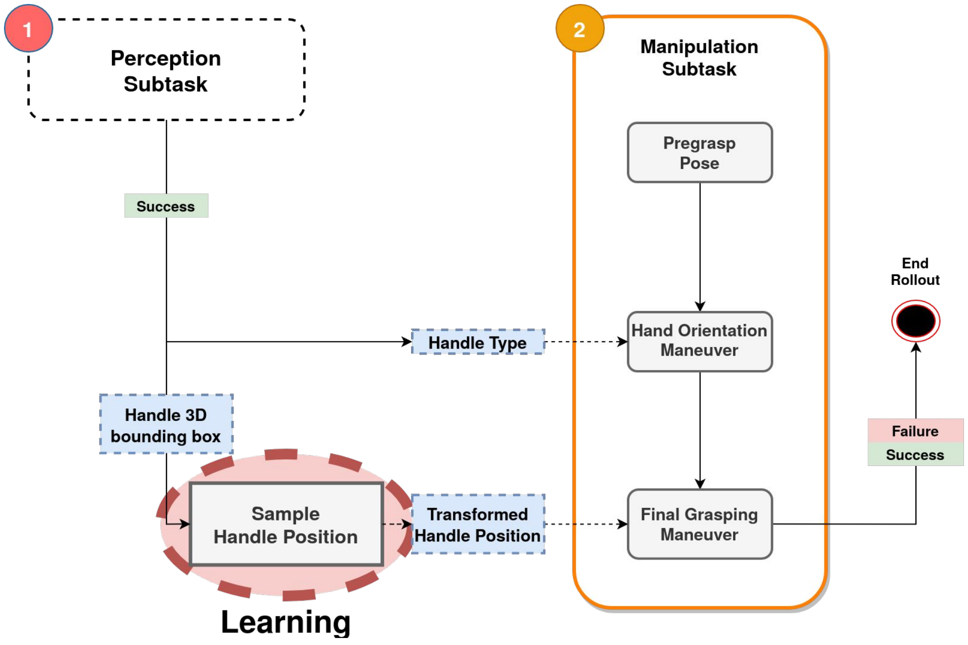
	AP	AP-0.5	AP-0.75	AP small	AP medium	AP large	AR max=1	AR max=10	AR max=100	AR small	AR medium	AR large
YOLO	0.64	-	-	-	-	-	-	-	-	-	-	
FR-CNN	0.904	0.980	0.977	0.646	0.894	0.936	0.574	0.923	0.924	0.700	0.911	0.956
SSD	0.702	0.96	0.781	-	0.715	0.708	0.503	0.752	0.752	-	0.757	0.751

Manipulation - Pipeline



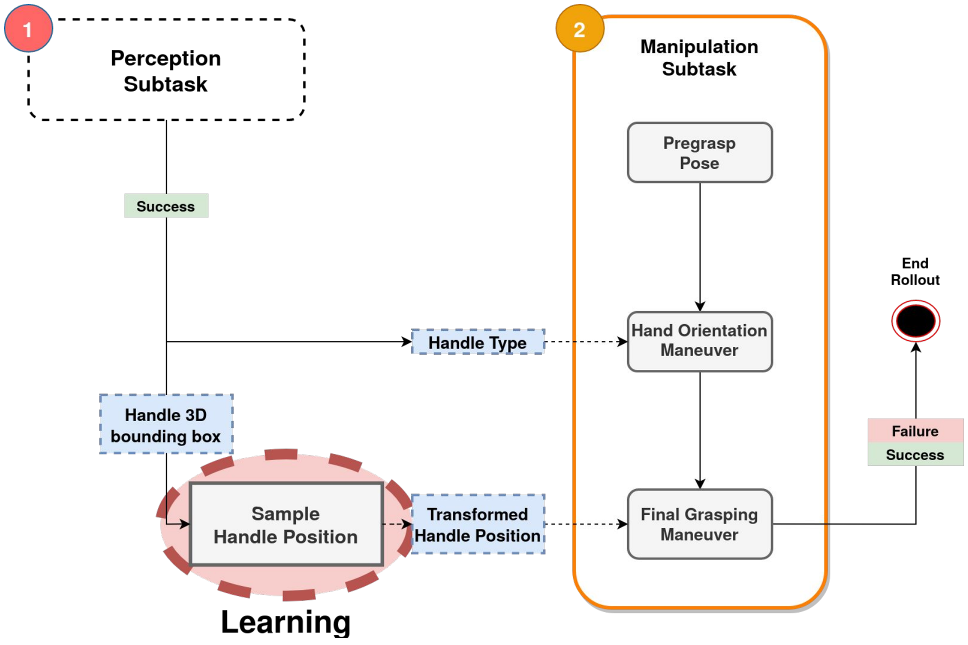
- Outputs from the perception :
  - 3D bounding box around the handle.
  - Detected handle class.
- Major components of the manipulation subtask :
  - Best position to grasp the handle.
  - Fixed grasping maneuver based on the types of handle.

Manipulation - Pipeline

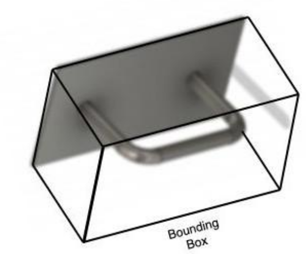


- Outputs from the perception :
  - **3D bounding box around the handle.**
  - Detected handle class.
- Major components of the manipulation subtask :
  - **Best position to grasp the handle.**
  - Fixed grasping maneuver based on the types of handle.

Manipulation - Sampling best position for grasping



- Increase probability of grasp positions that leads to success.
- Decrease probability of unsuccessful grasps.
- Initially sample grasp position from uniform distribution.



Manipulation - Sampling position for grasping

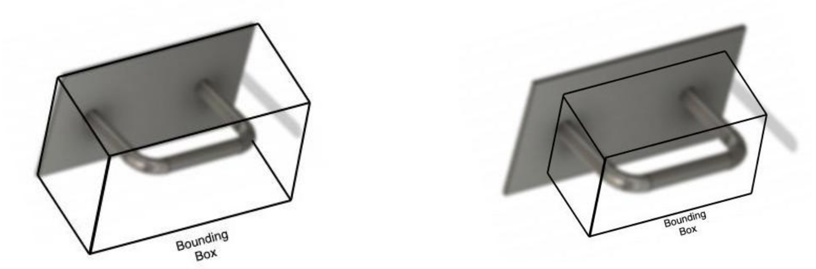


Figure 14. The sampling region effectively changes after MLE to improving grasping.[2]

Grasp position coordinates collected for training :

$$\mathcal{Y} = [y_1, y_2, y_3, \ldots.]$$

Likelihood function for a Normal distribution :

$$\mathcal{L} = \mathcal{N}(\mathcal{Y}|\mu, \sigma^2)$$

Algorithm Weighted Maximum Likelihood Estimation

Data: List of grasp position coordinates, $y_i \in \mathcal{Y}$ where $y_i = \{x, y, z\}$
Data: Success of each sample in $\mathcal{Y}$, $\mathcal{S}$
Data: Weighted list of samples $\overline{\mathcal{Y}}$
Result: The parameters of gaussian distribution ($\mathcal{N}$), μ and σ^2
for *Each sample* $y_i \in \mathcal{Y}$ **do**
  if y_i *leads to a successful grasp or successful pull* **then**
    if y_i *leads to only a successful grasp* **then**
      Add 3 instances of y_i to $\overline{\mathcal{Y}}$;
    else
      Add 5 instances of y_i to $\overline{\mathcal{Y}}$;
    end
  else
    Consider the axis along which the failure occurred.
    Replace the value along that axis to zero in y_i.
    Add the modified y_i to $\overline{\mathcal{Y}}$.
  end
end
Use the $\overline{\mathcal{Y}}$ to perform MLE and find parameters of $\mathcal{N}$

Manipulation - Sampling position for grasping

	Sampling	Successful grasp	Successful pull	Total Success (out of 50 rollouts)	Total Success (percentage)
Drawer handle	Uniform distribution	14	7	14	28%
	Gaussian distribution	30	11	30	60%
Fridge handle	Uniform distribution	13	6	13	26%
	Gaussian distribution	20	7	20	40%

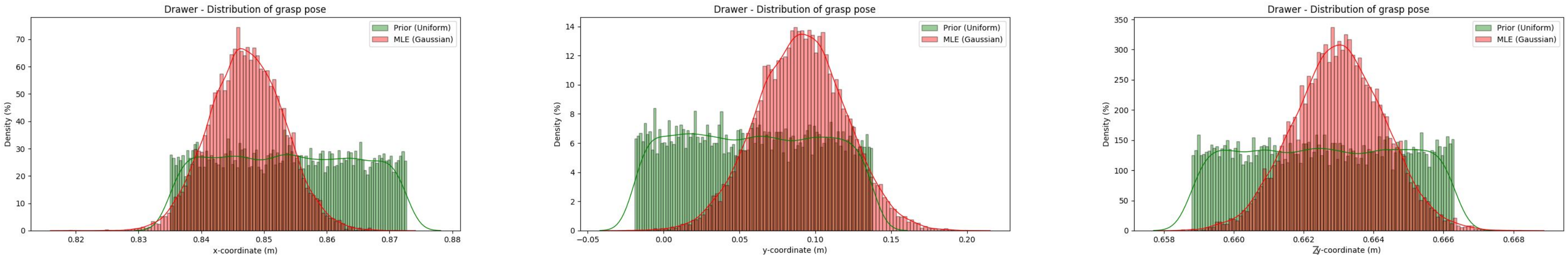


Figure 15. Graphs showing the grasp position distributions before (Uniform distribution) and after (Normal distribution) performing MLE.

Complete Pipeline - Results

Figure 16. Toyota HSR running the pipeline for manipulating the handles. (Speed 4X)

Conclusion

- Design of a general pipeline.
- Perception subtask
  - Dataset generation and training the 2D object detection.
  - Point cloud for pose detection.
- Manipulation subtask
  - Quantifiably improved sampling best grasp positions by using MLE.
- Coupling Perception and Manipulation subtasks.
- Running the complete pipeline on the Toyota HSR.

Future Work

- Pipeline Development and Evaluation on Toyota HSR
  - Learn other models like *Gaussian Processes*, RBFs etc.
  - Extend pipeline for lever type handles.
  - Apply failure detection and recovery.

References

1. Prats, Mario, Angel P. Del Pobil and Pedro J. Sanz. “Robot Physical Interaction through the combination of Vision, Tactile and Force Feedback - Applications to Assistive Robotics.” Springer Tracts in Advanced Robotics (2012).
2. Urakami, Yusuke, Alec Hodgkinson, Casey Carlin, Randall Leu, Luca Rigazio and Pieter Abbeel. “DoorGym: A Scalable Door Opening Environment And Baseline Agent.” ArXiv abs/1908.01887 (2019): n. Pag.
3. Bullock, Ian M. and Aaron M. Dollar. “Classifying human manipulation behavior.” 2011 IEEE International Conference on Rehabilitation Robotics (2011): 1-6.
4. Sutton, Richard S. and Andrew G. Barto. “Reinforcement Learning: An Introduction.” IEEE Transactions on Neural Networks 16 (1988): 285-286.
5. Meeussen, Wim, Melonee Wise, Stuart Glaser, Sachin Chitta, Conor McGann, Patrick Mihelich, Eitan Marder-Eppstein, Marius Muja, Victor Eruhimov, Tully Foote, John M. Hsu, Radu Bogdan Rusu, Bhaskara Marthi, Gary R. Bradski, Kurt Konolige, Brian P. Gerkey and Eric Berger. “Autonomous door opening and plugging in with a personal robot.” 2010 IEEE International Conference on Robotics and Automation (2010): 729-736.
6. Sutton, Richard S. and Andrew G. Barto. “Reinforcement Learning: An Introduction.” IEEE Transactions on Neural Networks 16 (1988): 285-286.

References

7. Prats, Mario, Angel P. Del Pobil and Pedro J. Sanz. “Robot Physical Interaction through the combination of Vision, Tactile and Force Feedback - Applications to Assistive Robotics.” Springer Tracts in Advanced Robotics (2012).
8. Maitin-Shepard, Jeremy, Marco Cusumano-Towner, Jinna Lei and Pieter Abbeel. “Cloth grasp point detection based on multiple-view geometric cues with application to robotic towel folding.” 2010 IEEE International Conference on Robotics and Automation (2010): 2308-2315.
9. Bollini, Mario, Stefanie Tellex, Tyler Thompson, Nicholas Roy and Daniela Rus. “Interpreting and Executing Recipes with a Cooking Robot.” ISER (2012).
10. Kober, Jens, J. Andrew Bagnell and Jan Peters. “Reinforcement learning in robotics: A survey.” The International Journal of Robotics Research 32 (2012): 1238 - 1274.
11. Tobin, Joshua, Rachel H Fong, Alex Ray, Jonas Schneider, Wojciech Zaremba and Pieter Abbeel. “Domain randomization for transferring deep neural networks from simulation to the real world.” 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (2017): 23-30.
12. Spinning up openai.URL : https://spinningup.openai.com/en/latest/algorithms/ppo.html.

Thank you

Learning Based Algorithms

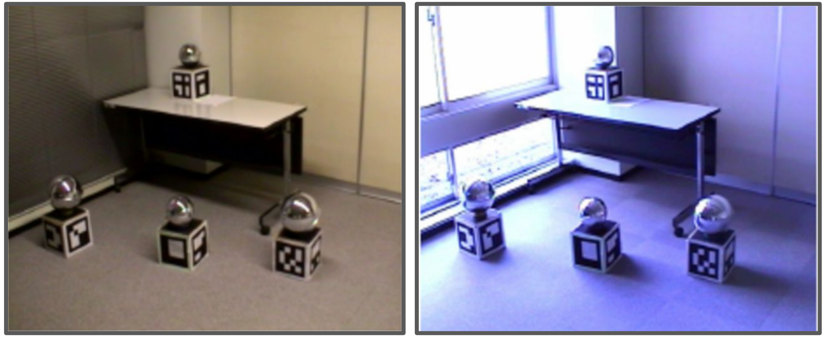


Figure 17. The same scene under different lighting conditions can cause problem in the perception subtask. [7]

- Characteristic of domestic environments :
  - Unstructured
  - Conditions change with time
  - Dynamic environment
- Difficult to hardcode controllers for manipulation.
- Variable environment conditions and object design makes perception very difficult.

Characteristics of Handle Manipulation Task

- Contact based.
- Prehensile manipulation.
- Closed grasp.
- No movement within the hand.
- Medium precision and dexterity.

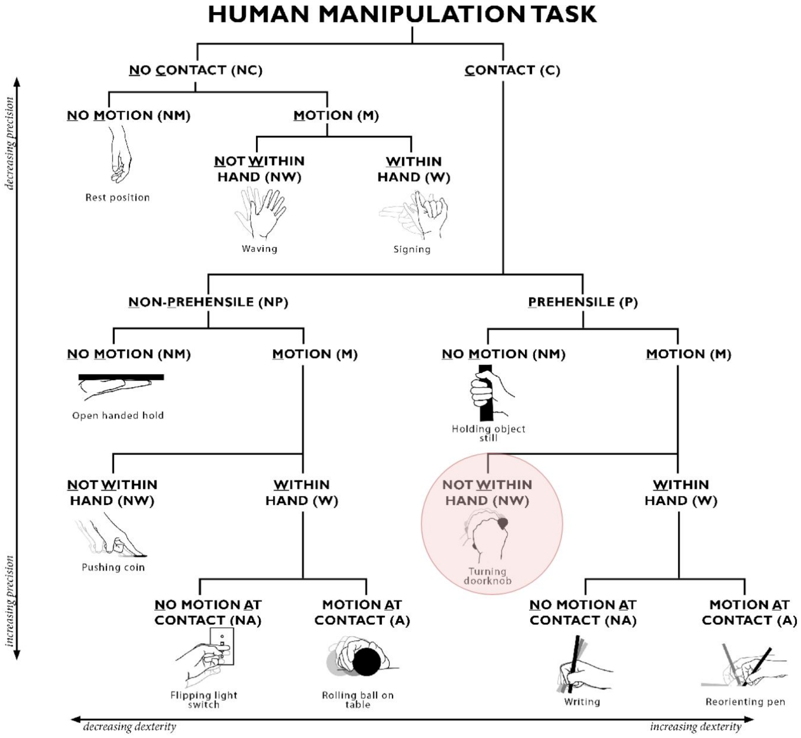


Figure 18. Human manipulation taxonomy.[4]

Reinforcement Learning

- ”Learning to map situations to actions to maximize a numerical signal..” - Sutton and Barto [6].
- Manipulation problems can naturally be formulated as an RL problem.
- “Rather than understanding your environment, simply collect a lot of experience and let the algorithm handle the rest” [11]

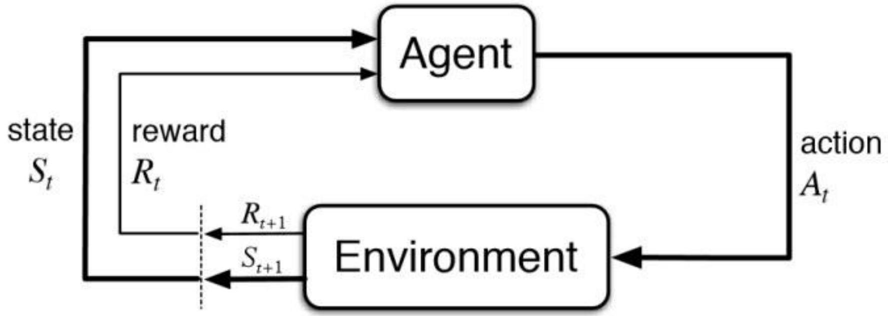


Figure 19. General framing of a RL problem. [6]

Real World as a MDP

States

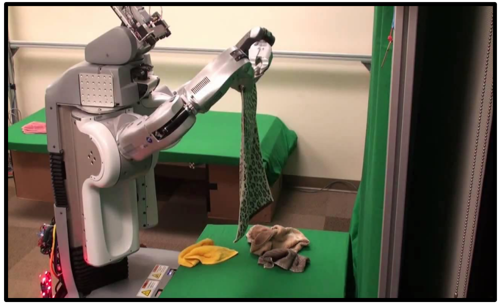
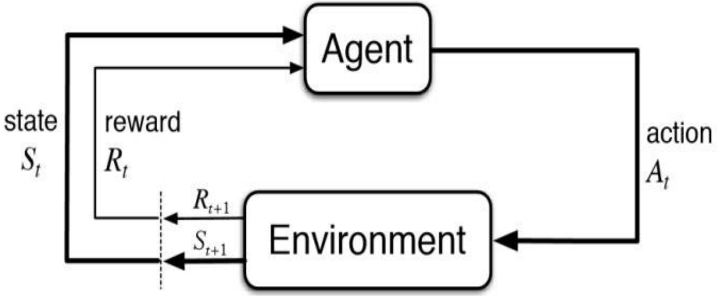


Figure 20. Ambiguous states occurring in the real world. [8]



Rewards

Figure 21. Specifying reward function for many tasks are difficult. Eg. Cooking. [9]

Actions

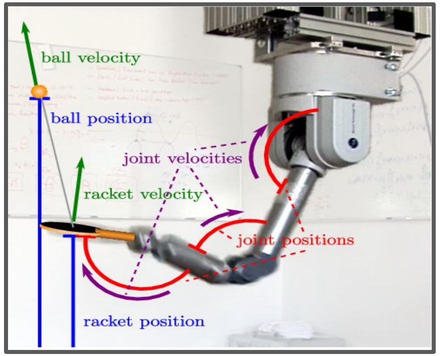


Figure 22. Robot manipulator may degrade over time. [10]

Proximal Policy Optimization

- “How to take biggest step in the policy using the data without steeping too far..”[12]

- Advantages of PPO :
  - Sample efficient
  - Ease of implementation
  - Ease of tuning hyperparameters

DoorGym - Wrong Results

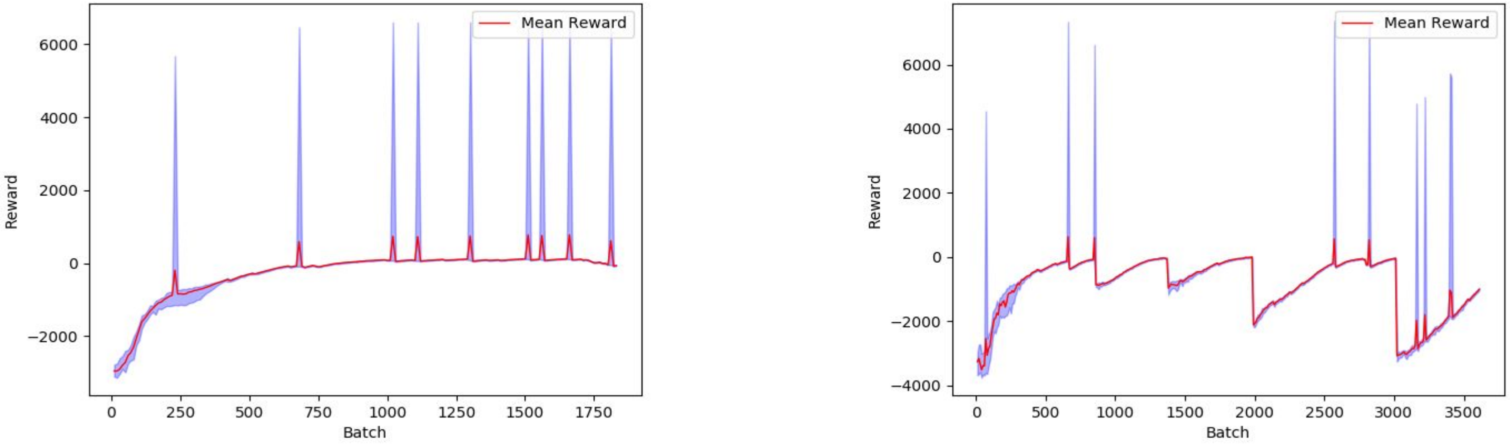


Figure 23. Results obtained from DoorGym environment when the environment parameters are not set correctly.

Manipulation sampling results

Handle_X (meters)	Handle_Y (meters)	Handle_Z (meters)	Success	Notes on Success/Failure	Grasp	Opening
0.8419	0.1211	0.6625	0	y - failure	-	-
0.8480	0.1058	0.6615	0	x - failure	-	-
0.8382	0.1378	0.6636	1	Slipped	✓	✗
0.8506	0.0830	0.6663	1	Drawer opened	✓	✓
0.8508	0.1000	0.6633	0	y - failure	-	-
0.8483	0.1006	0.6640	1	Slipped	✓	✗
0.8466	0.0924	0.6637	0	x - failure	-	-
0.8549	0.0524	0.6622	0	x - failure	-	-
0.8475	0.0865	0.6627	1	Drawer opened	✓	✓
0.8443	0.0648	0.6630	0	x - failure	-	-

Figure 24. Tabulation of the first ten rollouts for training a model to grasp a horizontal drawer handle.

Sampling graphs

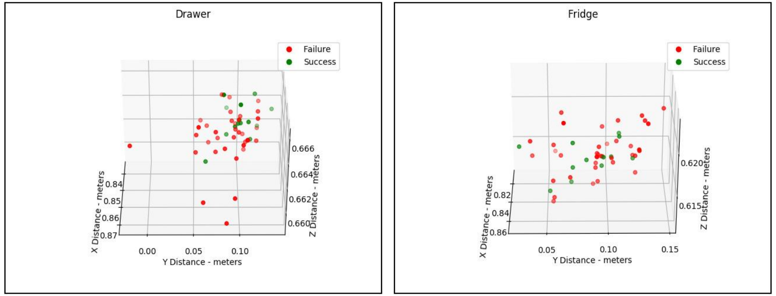


Figure 25. Graph showing the successful (Green) and unsuccessful (Red) grasp positions when sampled from a 3D uniform distribution within the limits of the 3D bounding box obtained by perception.

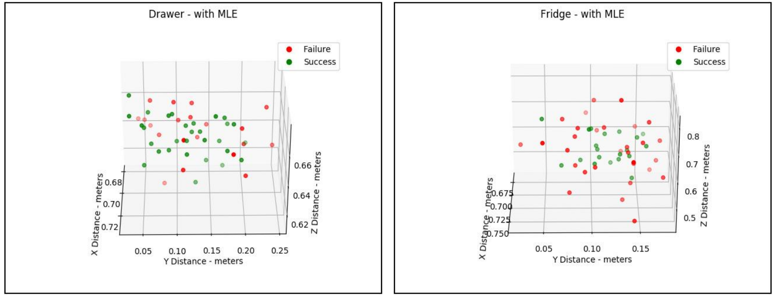


Figure 26. Graph showing the successful (Green) and unsuccessful (Red) grasp positions when sampled from a 3D normal distribution within the limits of the 3D bounding box obtained by perception. The normal distribution for sampling is obtained after training a MLE model to improve probability of successful grasps.

Manipulation - Handle Slip

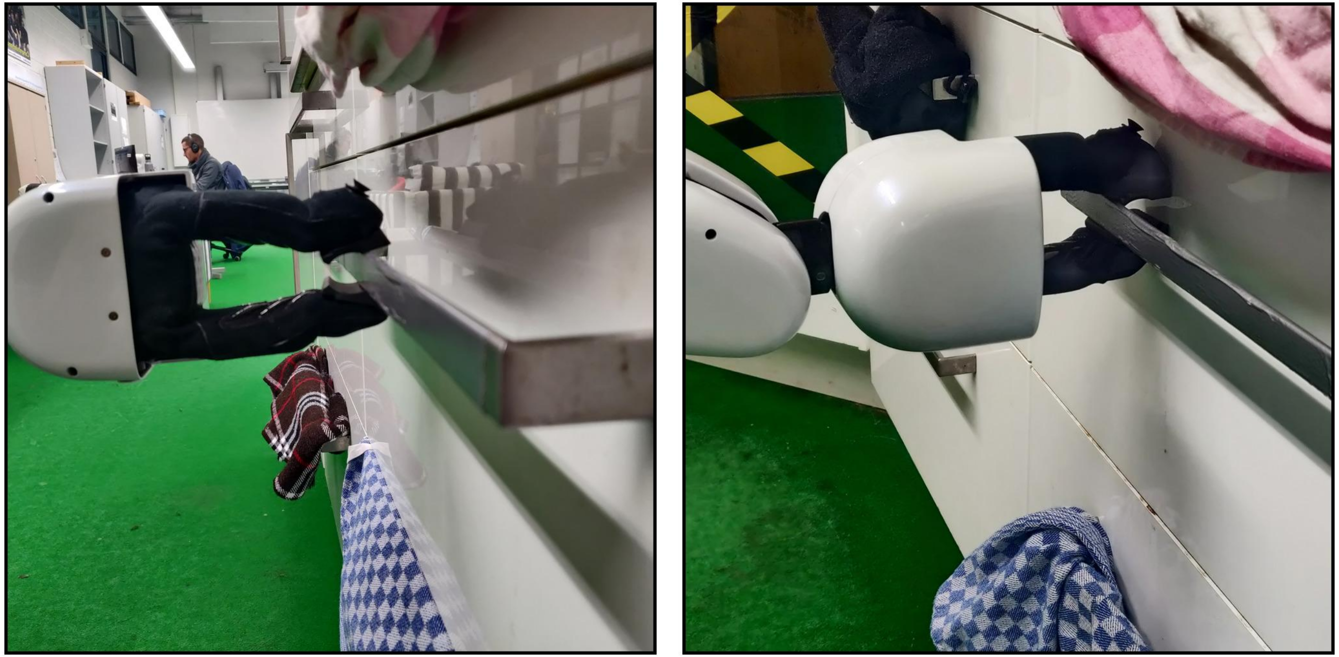


Figure 27. The handle design leads failure in grasping if the fingers don't wrap the handle completely. The (right) image shows the successful grasp whereas the (left) image leads to an unsuccessful grasp and pull.