Robust Dexterous Grasping of General Objects

Hui Zhang*, Zijian Wu*, Linyi Huang, Sammy Christen, Jie Song
*Equal Contribution

ETH zürich
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)

Overview

1 Motivation

Goal: robustly grasp "any" object with a dexterous hand under single-view visual inputs

Q1: Why grasping?

- The most fundamental skill for different manipulation tasks

Q2: Why dexterous hand?

- Feasible for more complex manipulation tasks
- Easier to incorporate and learn from human manipulation data

Q3: Why single-view?

- Most common cases, with more data available
- Easier to deploy

2 Challenges

Challenge 1: Generalization

Question: Real objects with diverse shapes, materials, and weights?

Challenge 2: Robustness

Question: Robust grasping under noisy single-view visual inputs?

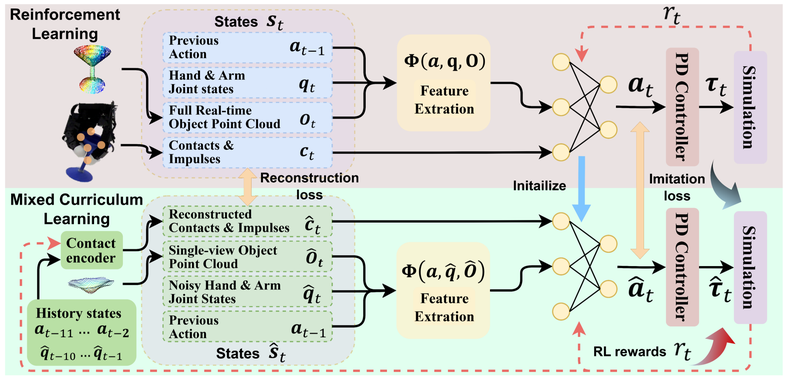
Challenge 3: Adaptation

Question: Adaptive actions under disturbances without tactile?

Method

1 Pipeline

Trained with **35** objects in **simulation**

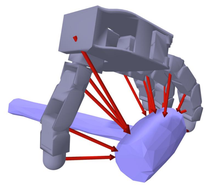


2 Sparse Hand-centric Object Shape Representation

Formulation: Vectors from each joint to its nearest point on the object surface

Advantage:

- A specific on contact-relevant local shape
- Capturing overall manifold instead of fine-grained shape details
- Leading to better generalization and robustness to observation noise



3 Mixed Curriculum Learning

Initially train with pure IL for fast grasping skill distillation

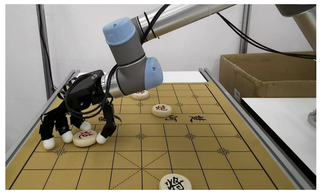
Adaptively transfer to RL for exploration of adaptive motions under disturbances

Train with 35 objects in simulation
Generalize to 500+ real objects with 94.6% success rate

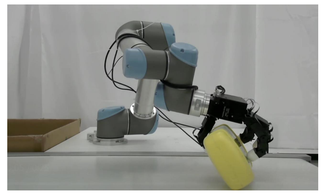
Applications

1 Downstream tasks enabled by our grasping capability

Grasping under cluttered scenes



Playing chess



Grasping moving objects on a conveyor belt

Task-driven long-term manipulation

Project Page

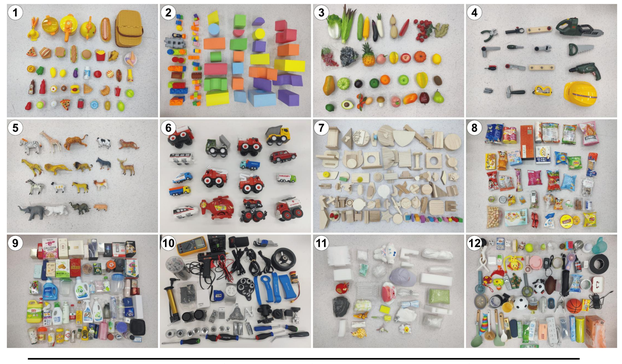
Personal Website

Code Released

Correspondence
huizhang@ethz.ch

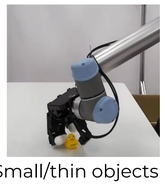
Results

1 Generalization on 500+ Real Objects

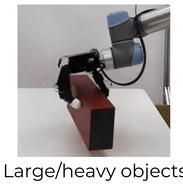


Category	Suc. Rate	Category	Success Rate
Picnic Models	0.902	Building Blocks	0.963
Fruit & Vegetable Models	0.962	Tool Models	0.875
Animal Models	0.907	Toy Cars	0.979
Wooden Models	0.940	Snacks	0.974
Bottles & Boxes	0.970	Real Tools	0.893
Deformable Objects	0.957	Other Daily Objects	0.971
Average	0.946		

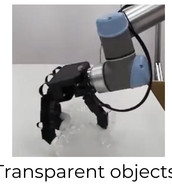
2 Grasping Challenging Objects



Small/thin objects



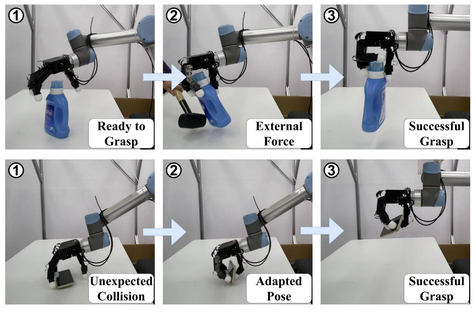
Large/heavy objects



Transparent objects

Deformable objects

3 Robustness and Adaptation to Disturbances

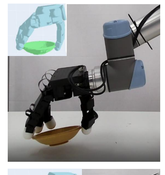
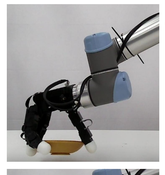
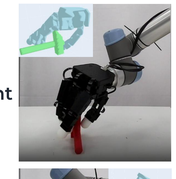
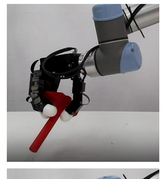


Maintain robust grasping under large external forces and recover from object slippage

Adaptively adjust grasping poses when unexpected collision happens due to observation or execution noise

3 Method Comparison

DexGraspNet [Wang et al., 2023]

	Pose-based	Ours	
Thin Object			• Pose-based methods suffer from thin objects as the potential collision with the table • While our method can dynamically deal with the table collision in real time
Object Movement			• Pose-based methods assume fixed objects, and suffer when the objects are moved during grasping • While our method can deal with object movement according to real-time feedback
Round/ slippery Object			• Pose-based methods suffer from round objects due to the challenge of controlling the fingers for a force-closure grasp with zero net force • While our method can adaptively adjust the finger torques to keep a force-closure grasping