LEARNING TO GRASP USING SIMULATION

Yunfei Bai
yunfeibai@google.com

Collaboration between X (Formally Google[X]), Robotics at Google (Research), and DeepMind

Jun 23, 2019
Outline

- Learning Vision Based Grasping
  - Self-Supervised Learning
  - Deep Reinforcement Learning
- Improving Data Efficiency
  - With Sim-to-Real
  - With Sim-to-Sim
- Learning New Tasks
- Learning New Object Representation
Learning Vision Based Grasping with Self-Supervised Learning

“Learning hand-eye coordination for robotic grasping with deep learning and large-scale data collection”, Sergey Levine, Peter Pastor, Alex Krizhevsky, Julian Ibarz, Deirdre Quillen
Goal: Learn to grasp arbitrary objects

- RGB monocular camera input
- Camera positioned “over the shoulder”
- Poor / non-existent camera calibration

Assumptions:
- Overhead grasps

Reward function:
- Gripper angle
- Image subtraction
1,100 objects used for training
Grasp Success Prediction Model

- Inference using Cross-Entropy Method (CEM)
- Replan every 300 to 500ms
- Hand-eye coordination
Grasp Success Rates with Increasing Amounts of Data
Two Directions for Improvement

- Break the ceiling of the grasp success rate to make it close to 100%.
- Improve real world data efficiency and reduce the data collection time.
Learning Vision Based Grasping with Deep Reinforcement Learning

“QT-Opt: Scalable Deep Reinforcement Learning for Vision-Based Robotic Manipulation”,
Dmitry Kalashnikov, Alex Irpan, Peter Pastor, Julian Ibarz, Alexander Herzog, Eric Jang, Deirdre Quillen, Ethan Holly,
Mrinal Kalakrishnan, Vincent Vanhoucke, Sergey Levine
Q Learning for Grasping

- Supervised learning
  - Optimize for the next step
- Reinforcement learning
  - Predict a few steps ahead
Q Learning for Grasping: QT-OPT

- Offline data 580K grasps
- Replay buffers
  - off-policy
  - on-policy
  - train
- Bellman Updater
- Model weights
- Training Worker
- Cross Entropy Method
- Q Learning for Grasping: QT-OPT
Q Network: Action Space Includes Open/Close Gripper and Termination

Real World Grasping Data

- Initial Image
- Current Image
- Motor Command
- Gripper Open/Close
- Episode Terminate

Grasping Value Function $Q(s, a)$
QT-Opt achieves 96% grasp success, with higher data efficiency

less data, 78% -> 96% grasp success
3

Leveraging Sim-to-Real for Data Efficiency

“Using Simulation and Domain Adaptation to Improve Efficiency of Deep Robotic Grasping”,
Konstantinos Bousmalis, Alex Irpan, Paul Wohlhart, Yunfei Bai, Matthew Kelcey, Mrinal Kalakrishnan, Laura Downs,
Julian Ibarz, Peter Pastor, Kurt Konolige, Sergey Levine, Vincent Vanhoucke
Motivation for Using Simulation

- 608,000 real-world grasps to achieve best performance
  - 7 KUKA robots running for 2-3 months
- Use simulation!
  - Easy to parallelize, reset, safe exploration, access to ground-truth for exploration policy, …

Real world  
Simulation
Reality Gap

Real world

Dynamic model discrepancy

Uncertain environment

Erroneous sensing

Numerical error

Latency

Others

Simulation
Methods for Sim-to-Real Transfer

System Identification

Use statistical methods to build mathematical models of dynamical systems from measured data.

Domain Randomization

Vary texture, background, lighting, color, object shape, and dynamics in simulation.

Feature-level Domain Adaptation

Train features to be domain-invariant yet expressive, by using an adversarial loss.

Pixel-level Domain Adaptation

Train a generator network that converts simulated images to real images, by using an adversarial loss.
Sim-to-Real Transfer for Physics - System Identification

Reality gap

Sim  real

<table>
<thead>
<tr>
<th>Successful grasp</th>
<th>Unsuccessful grasp</th>
</tr>
</thead>
<tbody>
<tr>
<td>855</td>
<td>56</td>
</tr>
<tr>
<td>29</td>
<td>2080</td>
</tr>
</tbody>
</table>

Sim  Real
Sim-to-Real Transfer for Physics - System Identification

<table>
<thead>
<tr>
<th></th>
<th>Sim</th>
<th>Real</th>
</tr>
</thead>
<tbody>
<tr>
<td>Successful grasp</td>
<td>855</td>
<td>56</td>
</tr>
<tr>
<td>Unsuccessful grasp</td>
<td>29</td>
<td>2080</td>
</tr>
</tbody>
</table>

Copyright 2019 X Development LLC
Sim-to-Real Transfer for Perception - Domain Randomization
• 51,300 object models from ShapeNet.

Chang et al., CoRR 2015
● 1,000 procedurally generated object models.
• Procedural objects with random textures.
Sim-to-Real Transfer for Perception - Feature Level Domain Adaptation

- Domain Adversarial Neural Network. Train features to be domain-invariant yet expressive, by using an adversarial loss. Learn features that confuse the domain classifier.

Based on Ganin et al., JMLR 2016
- Learn a generator ($G$) which converts synthetic images to real looking (fake) images.
Sim-to-Real Transfer for Perception - Pixel Level Domain Adaptation
**GraspGAN** outperforms other techniques, providing more than 50x data efficiency.
GraspGAN outperforms other techniques, providing more than 50x data efficiency. 2% of the real world data, same performance.
Solve Sim-to-Real via Sim-to-Sim

“Sim-to-Real via Sim-to-Sim: Data-efficient Robotic Grasping via Randomized-to-Canonical Adaptation Networks”,

Stephen James, Paul Wohlhart, Mrinal Kalakrishnan, Dmitry Kalashnikov, Alex Irpan, Julian Ibarz,

Sergey Levine, Raia Hadsell, Konstantinos Bousmalis
Randomized-to-Canonical Adaptation Networks

- **RCAN** is a real-to-sim image translator trained with domain randomization:
  - We define a “canonical” version of simulation and randomizations
Randomized-to-Canonical Adaptation Networks

- **RCAN** is a real-to-sim image translator trained with domain randomization:
  - We define a "canonical" version of simulation and randomizations
  - We train a pix2pix model to convert randomized sim images to equivalent canonical versions

\[
\mathcal{L}_{GAN}(G, D) = \mathbb{E}_x[\log D(x)] + \mathbb{E}_x[\log(1 - D(G_x(x)))]
\]
Randomized-to-Canonical Adaptation Networks

- **RCAN** is a real-to-sim image translator trained with domain randomization:
  - We define a “canonical” version of simulation and randomizations
  - We train a pix2pix model to convert randomized sim images to equivalent canonical versions
  - In the real world, RCAN will then also be able to translate real images to canonical sim versions
RCAN achieves the similar success rate (94% vs 96%), but without using 580k real word data.
Learning New Tasks

“Multi-Task Domain Adaptation for Deep Learning of Instance Grasping from Simulation”,
Kuan Fang, Yunfei Bai, Stefan Hinterstoisser, Silvio Savarese, Mrinal Kalakrishnan
Sim-to-Real Transfer: Applying to a More Challenging Task

Initial Image

Target Mask
Instance Grasping Framework: Multi-Task Domain Adaptation

Real World Indiscriminate Grasping Data
- Initial Image
- Current Image
- Motor Command

Simulated Indiscriminate Grasping Data
- Initial Image
- Current Image
- Motor Command

Domain Classifier
- Adversarial Loss
- Indiscriminate Grasp Prediction Loss

Indiscriminate Grasp Prediction Loss

Copyright 2019 X Development LLC
Instance Grasping Framework: Multi-Task Domain Adaptation

Real World Indiscriminate Grasping Data
- Initial Image
- Current Image
- Motor Command

Simulated Indiscriminate Grasping Data
- Initial Image
- Current Image
- Motor Command

Simulated Instance Grasping Data
- Initial Image
- Current Image
- Target Mask
- Motor Command

Domain Classifier
- Adversarial Loss

Indiscriminate Grasp Prediction Loss
- Simulated Indiscriminate Grasping Data

Instance Grasp Prediction Loss
- Simulated Instance Grasping Data
Instance Grasping Framework: Multi-Task Domain Adaptation
Evaluation in the Real World

Initial Image

Target Mask
Learning New Object Representation

“Data-Efficient Learning for Sim-to-Real Robotic Grasping using Deep Point Cloud Prediction Networks”,
Xinchen Yan, Mohi Khansari, Jasmine Hsu, Yuanzheng Gong, Yunfei Bai, Soren Pirk, Honglak Lee
Shape Prediction

visualizations of point clouds generated with our point prediction network
Learning to Grasp Using Object Point Cloud

visualizations of point clouds generated with our point prediction network
Grasping Evaluations

6X speed
Conclusion
Summary

• Learn vision based grasping through self-supervised learning and deep reinforcement learning.

• Simulation helps reduce real world data requirements by 100x, by solving sim-to-real transfer and sim-to-sim transfer.

• Simulation also enables us to learn new related tasks, and good object representation can facilitate sim-to-real transfer.
1. Learning Hand-Eye Coordination for Robotic Grasping with Deep Learning and Large-Scale Data Collection
   S. Levine et al. IJRR 2017.

2. QT-Opt: Scalable Deep Reinforcement Learning for Vision-Based Robotic Manipulation

3. Using Simulation and Domain Adaptation to Improve Efficiency of Deep Robotic Grasping
   K. Bousmalis et al. ICRA 2018.

4. Sim-to-Real via Sim-to-Sim: Data-efficient Robotic Grasping via Randomized-to-Canonical Adaptation Networks

5. Multi-Task Domain Adaptation for Deep Learning of Instance Grasping from Simulation
   K. Fang et al. ICRA 2018.

6. Data-Efficient Learning for Sim-to-Real Robotic Grasping using Deep Point Cloud Prediction Networks
   Under review
### Credits

<table>
<thead>
<tr>
<th>X</th>
<th>Google Brain</th>
<th>DeepMind</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yunfei Bai</td>
<td>Laura Downs</td>
<td>Konstantinos Bousmalis</td>
</tr>
<tr>
<td>Kuan Fang</td>
<td>Ethan Holly</td>
<td>Raia Hadsell</td>
</tr>
<tr>
<td>Alexander Herzog</td>
<td>Julian Ibarz</td>
<td></td>
</tr>
<tr>
<td>Stefan Hinterstoisser</td>
<td>Alex Irpan</td>
<td></td>
</tr>
<tr>
<td>Stephen James</td>
<td>Eric Jang</td>
<td></td>
</tr>
<tr>
<td>Mrinal Kalakrishnan</td>
<td>Dmitry Kalashnikov</td>
<td></td>
</tr>
<tr>
<td>Peter Pastor</td>
<td>Matthew Kelcey</td>
<td></td>
</tr>
<tr>
<td>Paul Wohlhart</td>
<td>Kurt Konolige</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Alex Krizhevsky</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sergey Levine</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Deirdre Quillen</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Vincent Vanhoucké</td>
<td></td>
</tr>
</tbody>
</table>
Thank you!

Yunfei Bai
yunfeibai@google.com