清華大學
Tsinghua University

# Deep Reinforcement Learning

# **Lecture 10: Sim2real RL (succinct ver.)**

Huazhe Xu
Tsinghua University

# AI This Week


Robot Marathon

# In Lec10

1 Why sim2real?

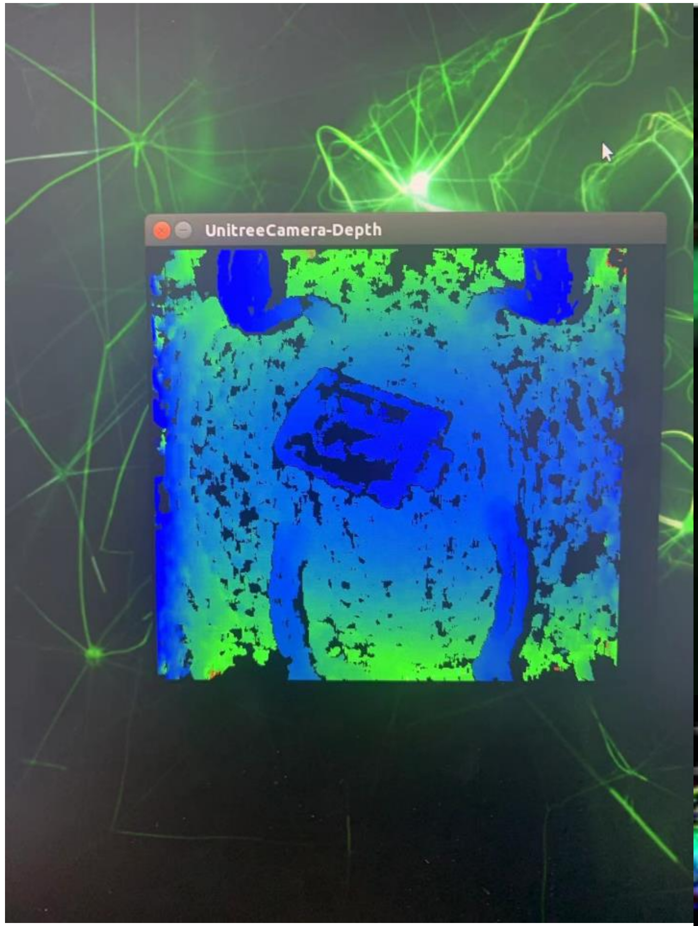
2 Domain Adaptation

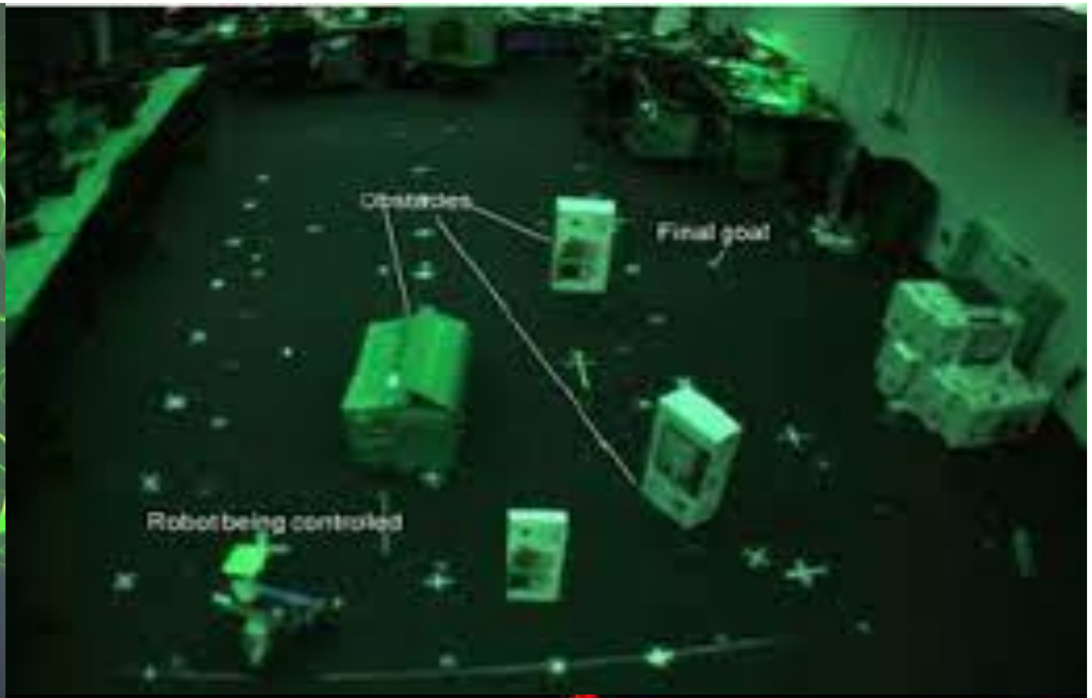
3 Domain Randomization

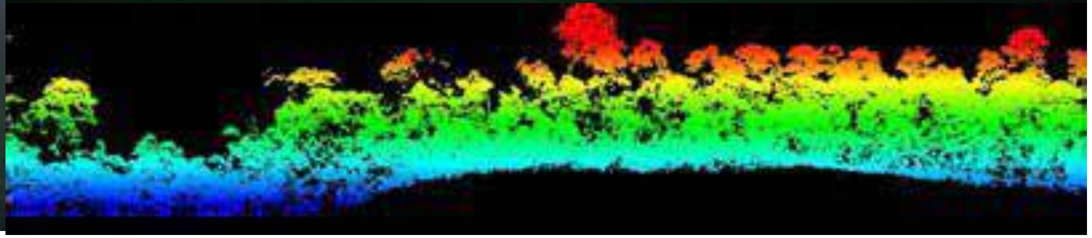
# In Lec10

1. Why sim2real?
2. Domain Adaptation
3. Domain Randomization

# Real robot tasks are hard: the sensors are not reliable!



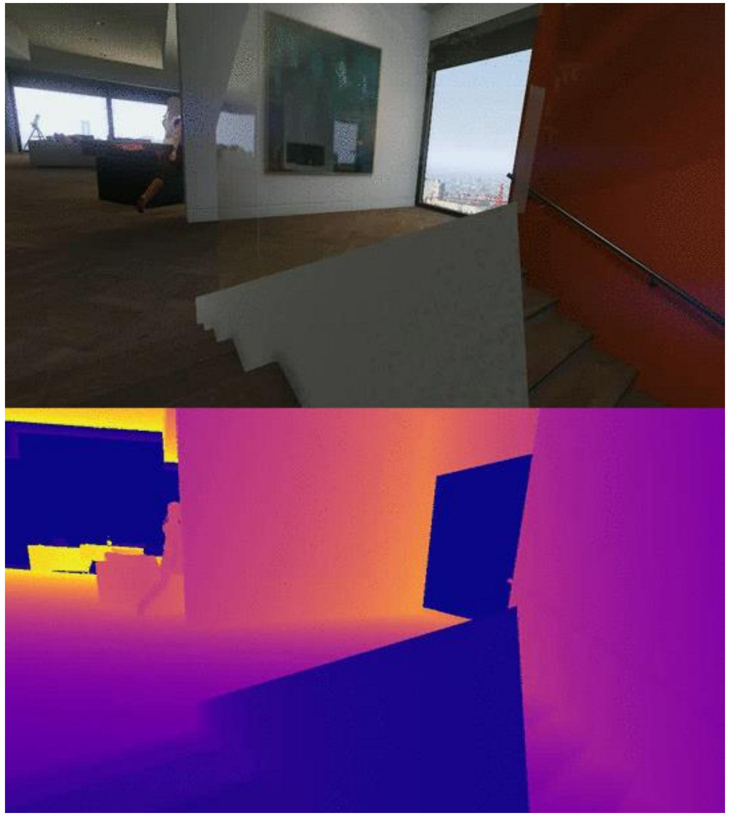
# Real robot tasks are hard: What else?

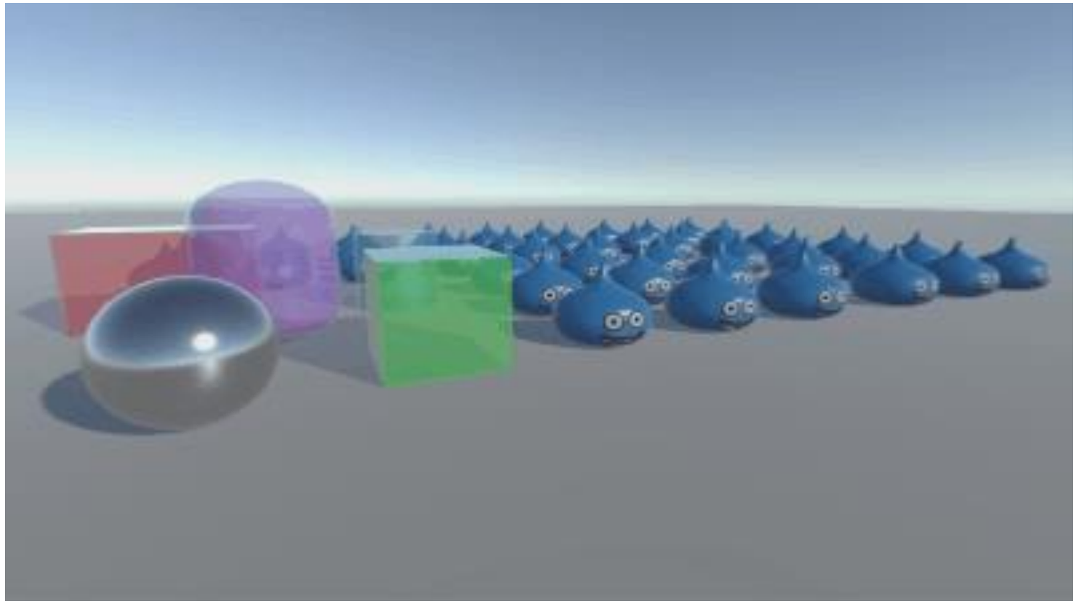
- Reward function is hard to extract
- The hardware is easily broken.
- …

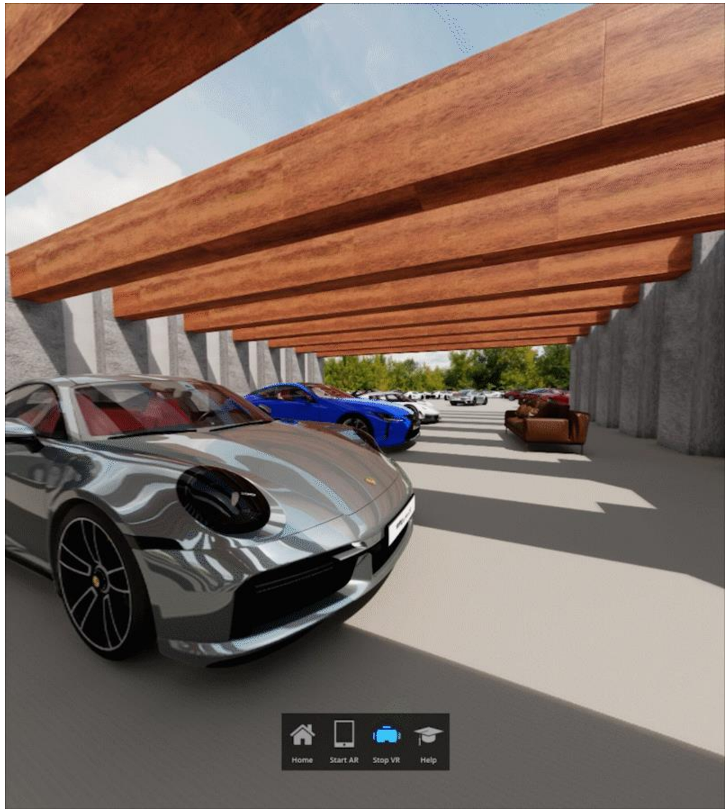
# Traditional robotics methods fail…

“Manipulation breaks all the rigorous/reliable approaches I know for control.” --- Russ Tedrake, MIT

# Simulators are amazing!





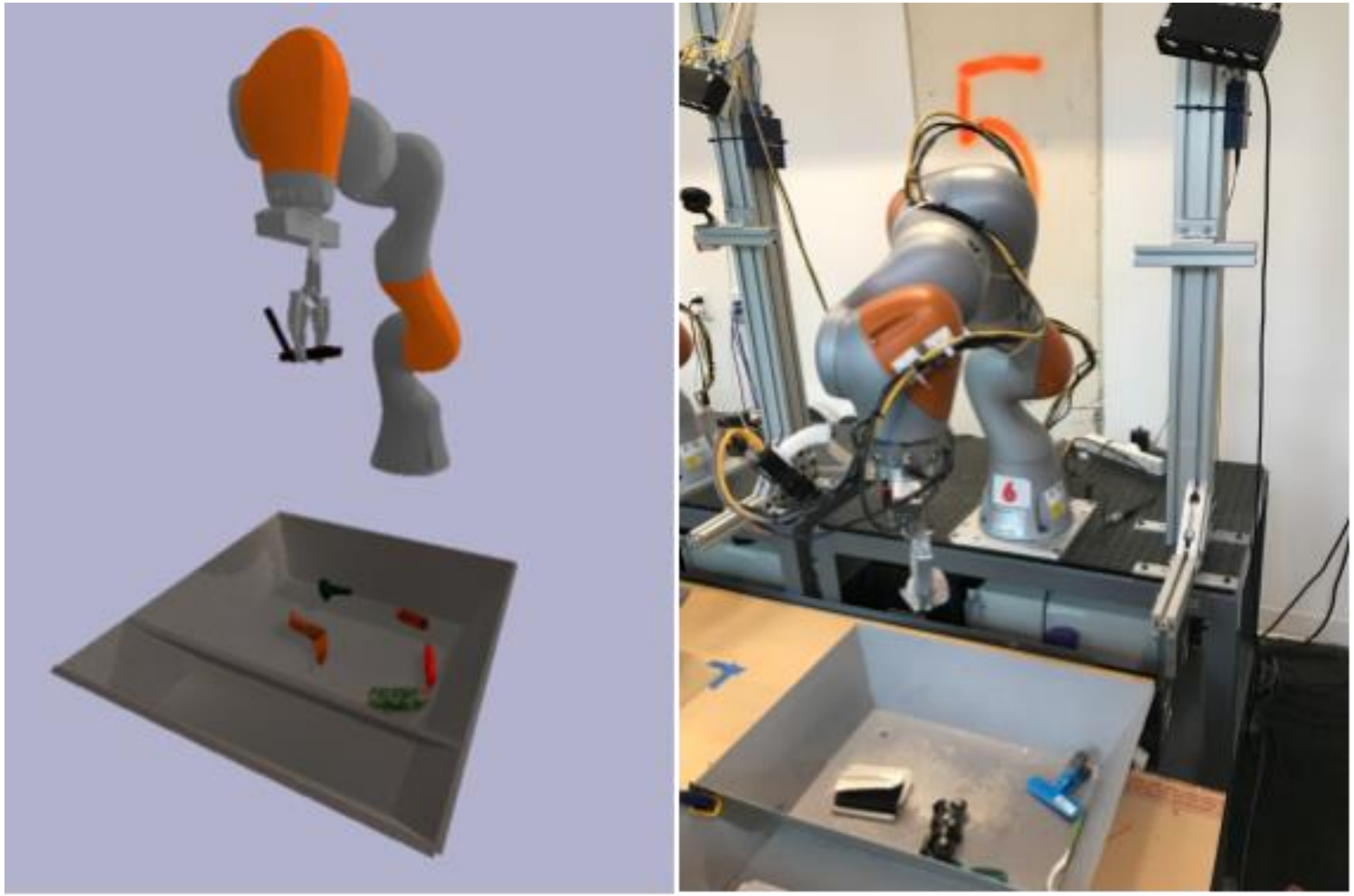


# Is everything solved by learn in sim and deploy in real?

- "There is a real danger (in fact, a near certainty) that programs which work well on simulated robots will completely fail on real robots because of the differences in real world sensing and actuation - it is very hard to simulate the actual dynamics of the real world."

Artificial Life and Real Robots [Rodney Brooks, 1992]

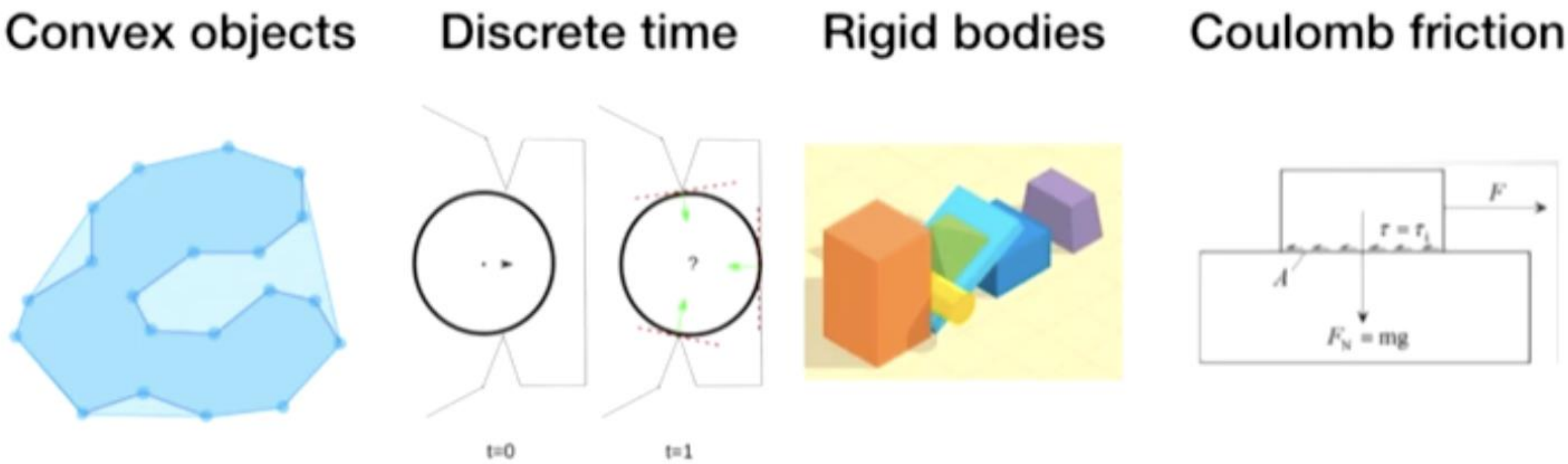
# The reality gap are sometimes huge!



# Why the gap exists?

- **Sensors and physical systems are hard to model**
- (Small) error accumulation

# Physics simulators make big assumption




Source: Josh Tobin Talk

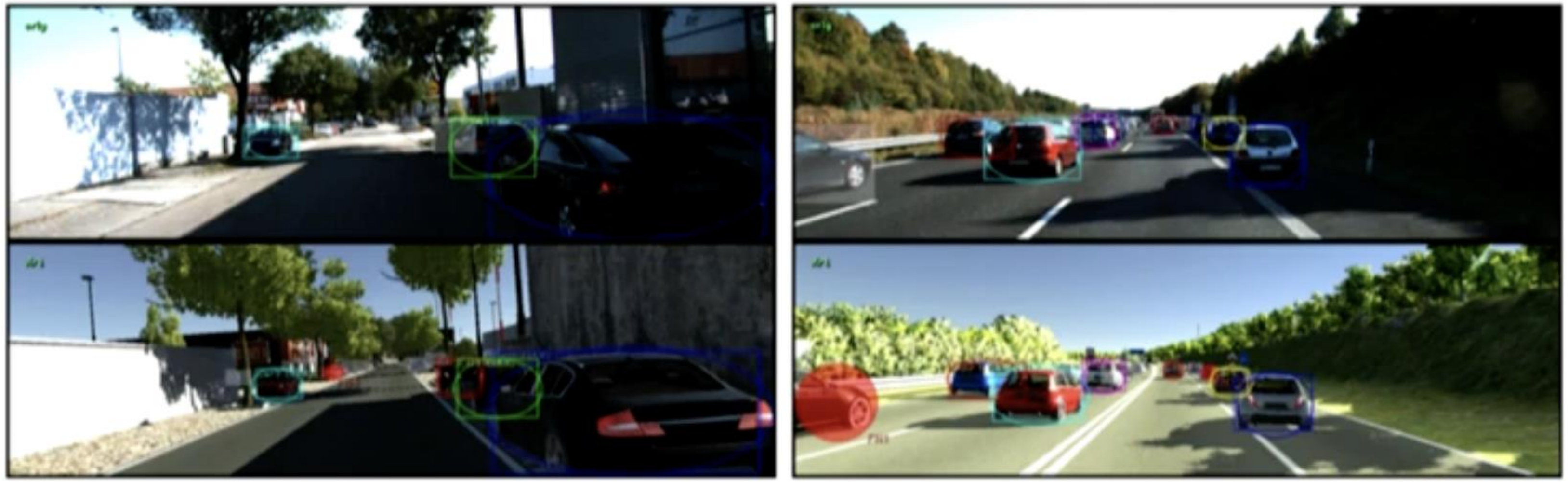
# Photorealistic sensor simulation is expensive

# Why the gap exists?

- Sensors and physical systems are hard to model
- **(Small) error accumulation**

# Neural network overfits to tiny differences

- Neural network are very lazy
- Sim ~63% Real ~78%



Virtual Worlds as Proxy for Multi-Object Tracking Analysis [Gaidon*, Wang", Cabon, Vig. 2016]

# In Lec10

1 Why sim2real?

2 Domain Adaptation

3 Domain Randomization

# Domain Adaptation

- Domain Adaptation is a technique to improve the performance of a model on a target domain containing insufficient annotated data by using the knowledge learned by the model from another related domain with adequate labeled data.

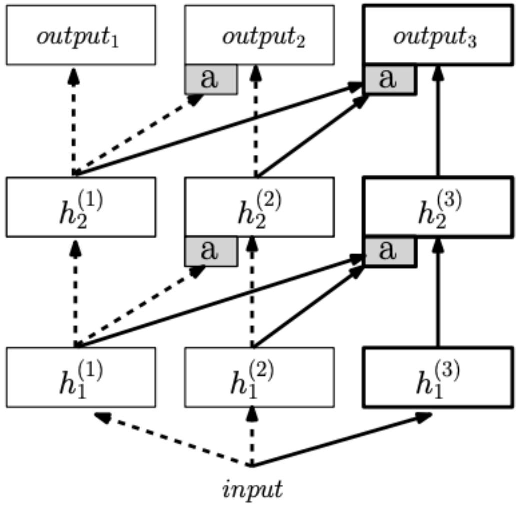
# Domain Adaptation Examples

## Progressive Neural Networks

**Andrei A. Rusu*, Neil C. Rabinowitz*, Guillaume Desjardins*, Hubert Soyer, James Kirkpatrick, Koray Kavukcuoglu, Razvan Pascanu, Raia Hadsell**
* These authors contributed equally to this work

Google DeepMind
London, UK

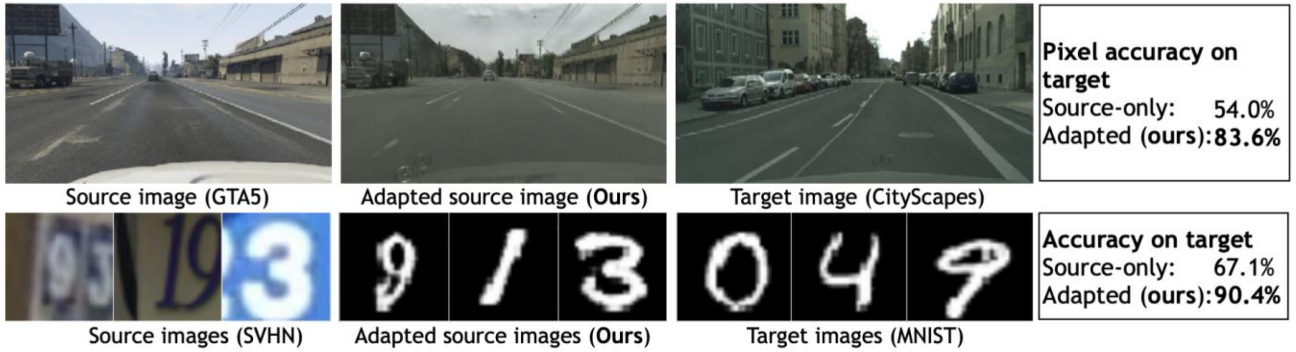
`{andreirusu, ncr, gdesjardins, soyer, kirkpatrick, korayk, razp, raia}@google.com`



# Domain Adaptation Examples

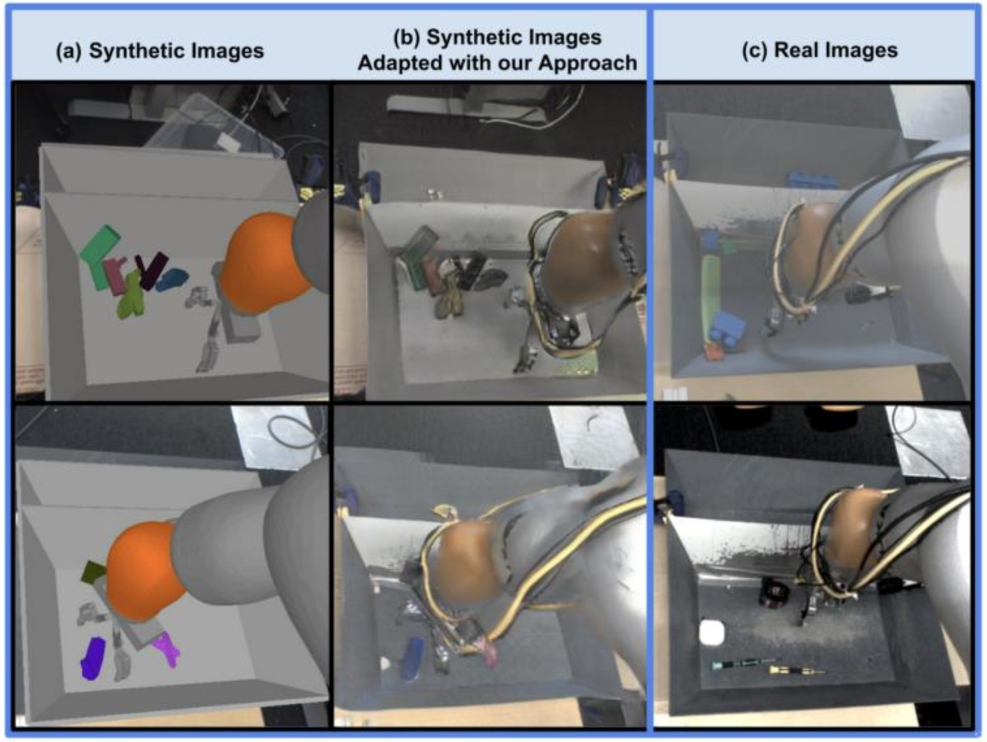
## CyCADA: Cycle-Consistent Adversarial Domain Adaptation

**Judy Hoffman** [1] **Eric Tzeng** [1] **Taesung Park** [1] **Jun-Yan Zhu** [1] **Phillip Isola** [1 2] **Kate Saenko** [3] **Alexei A. Efros** [1] **Trevor Darrell** [1]




## Using Simulation and Domain Adaptation to Improve Efficiency of Deep Robotic Grasping

Konstantinos Bousmalis[*,1], Alex Irpan[*,1], Paul Wohlhart[*,2], Yunfei Bai[2], Matthew Kelcey[1], Mrinal Kalakrishnan[2], Laura Downs[1], Julian Ibarz[1], Peter Pastor[2], Kurt Konolige[2], Sergey Levine[1], Vincent Vanhoucke[1]



# Domain Adaptation Examples

LEARNING CROSS-DOMAIN CORRESPONDENCE FOR CONTROL WITH DYNAMICS CYCLE-CONSISTENCY

**Qiang Zhang**
Shanghai Jiao Tong University
zhangqiang2016@sjtu.edu.cn

**Tete Xiao**
UC Berkeley
txiao@eecs.berkeley.edu

**Alexei A. Efros**
UC Berkeley
efros@eecs.berkeley.edu

**Lerrel Pinto**
New York University
lerrel@cs.nyu.edu

**Xiaolong Wang**
UC San Diego
xiw012@ucsd.edu

# In Lec9

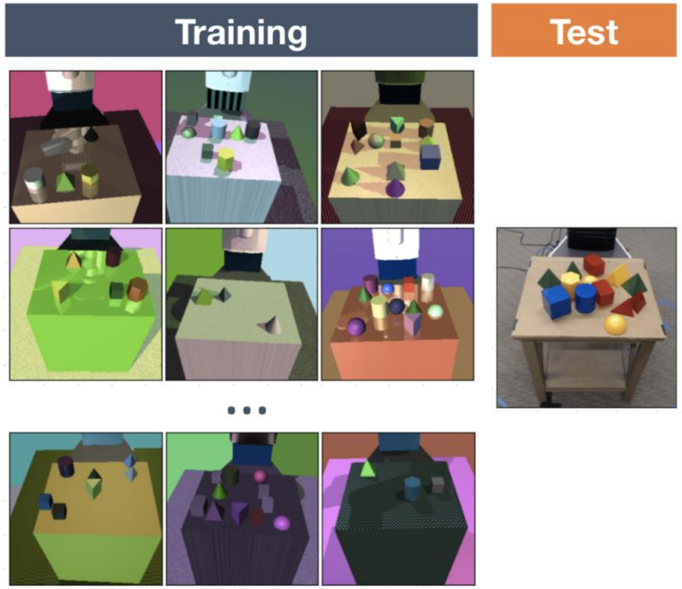
1 Why sim2real?

2 Domain Adaptation

3 Domain Randomization

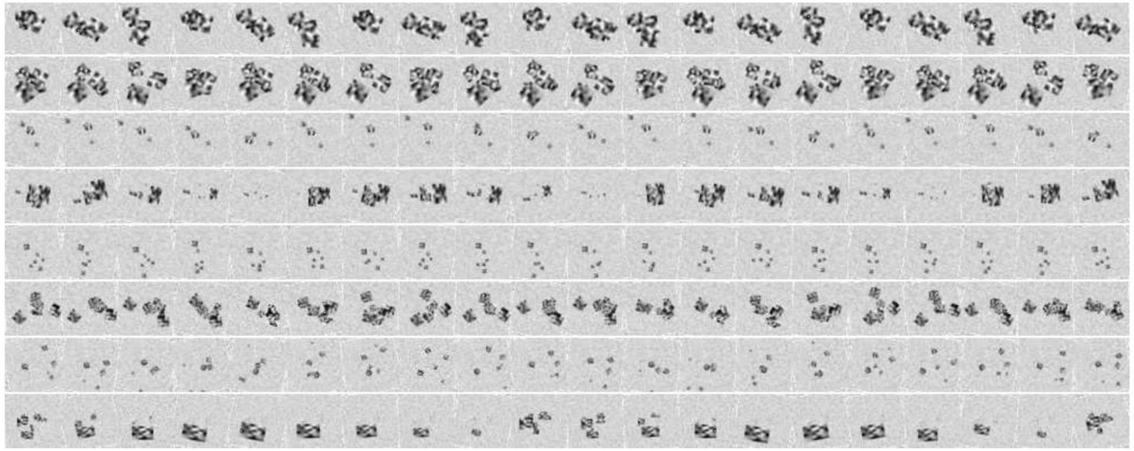
# Domain Randomization --- World is just another simulator ☺

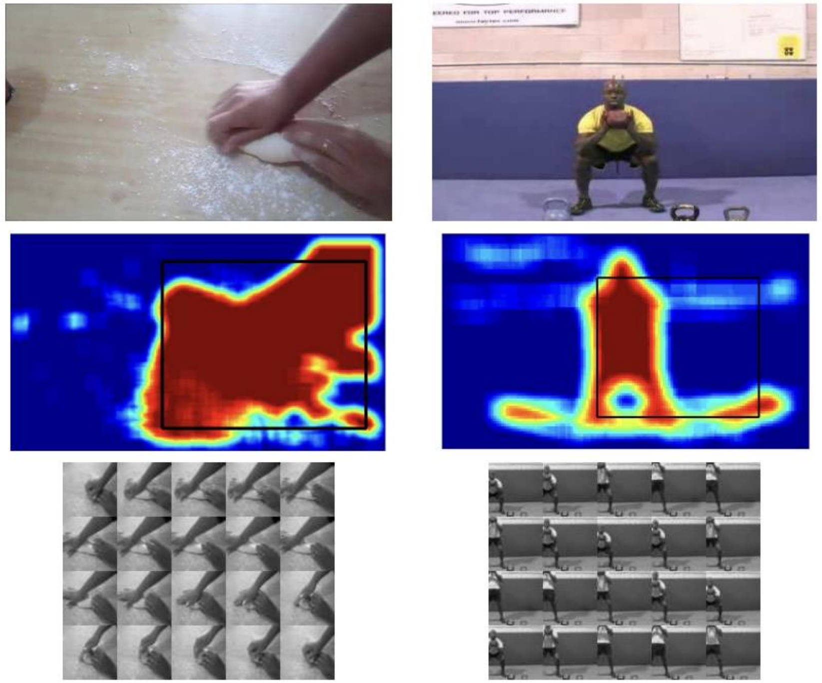
- Applying random transformations or perturbations, such as changes in lighting conditions, camera angles, object textures, or background noise, among others.
- These modifications introduce variability and complexity into the training data, which can help the model learn to be more robust and adaptable to different environments.



# History: Live Repetitive Counting

- Train on repetitive random images
- Test on repetitive real images





You randomize anything that should be ignored.

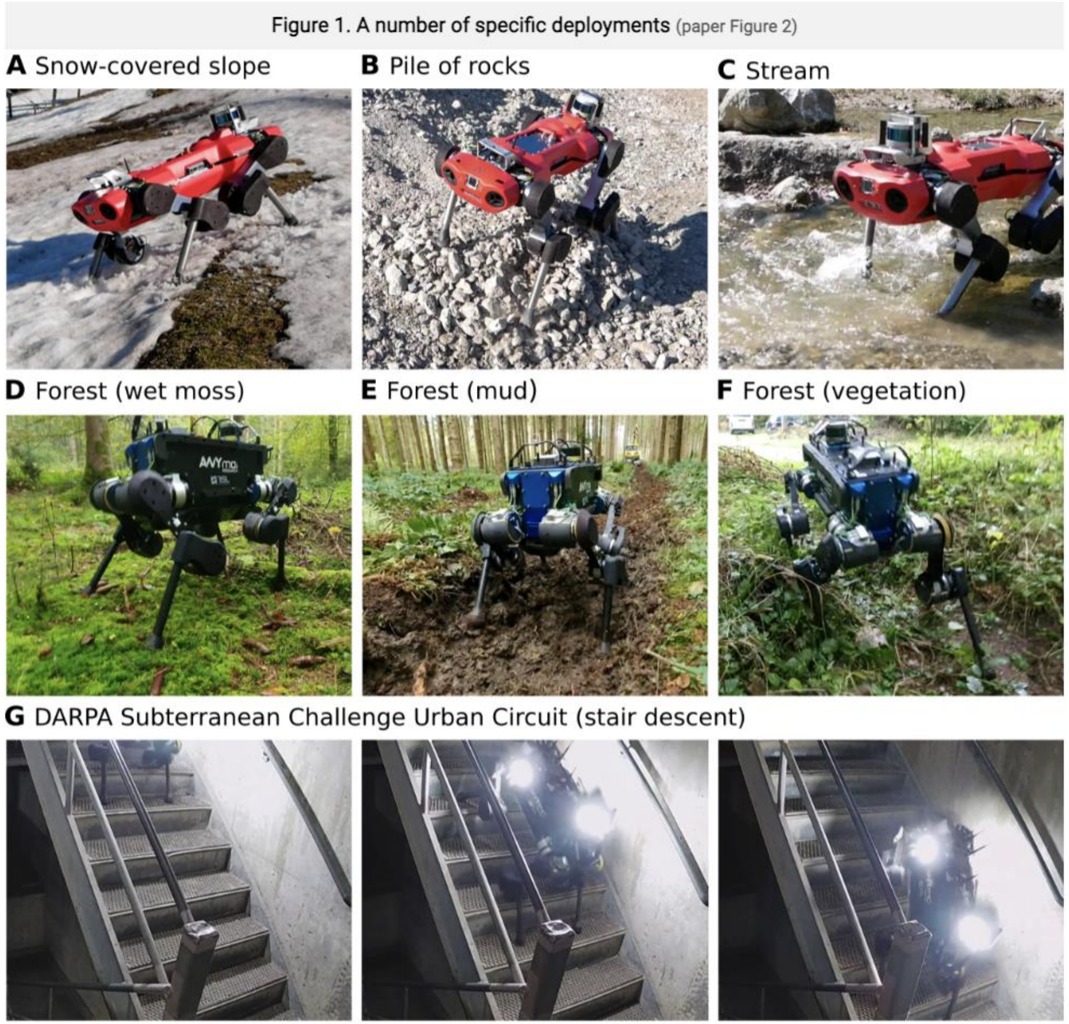
# DR in dexterous manipulation

# DR in dexterous manipulation

- Deep RL training (PPO)
- Automatic Domain Randomization
  - Physical parameters
  - Noise to policy inputs
  - Sensor dropout
  - Physics discretization steps
  - Backlash
  - Force
  - Visual appearance

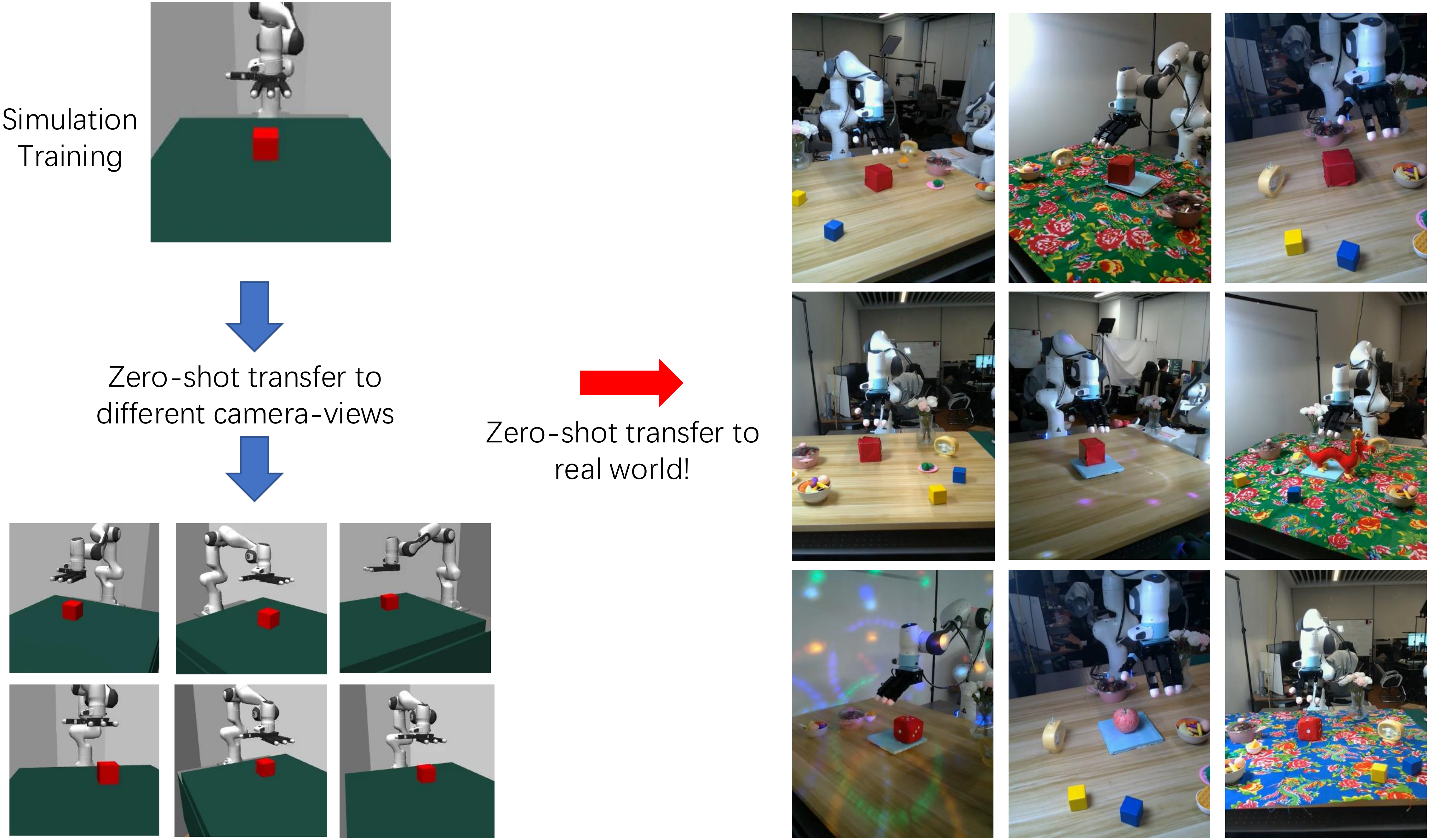
# DR in locomotion

- Randomized terrain
- Teacher-student architecture



# In Lec10

1. Why sim2real?
2. Domain Adaptation
3. Domain Randomization


Simulation
Training
Zero-shot transfer to
different camera-views
Zero-shot transfer to
real world!