# Beyond Domain Randomization

Josh Tobin
6/23/19

# Goals for this talk

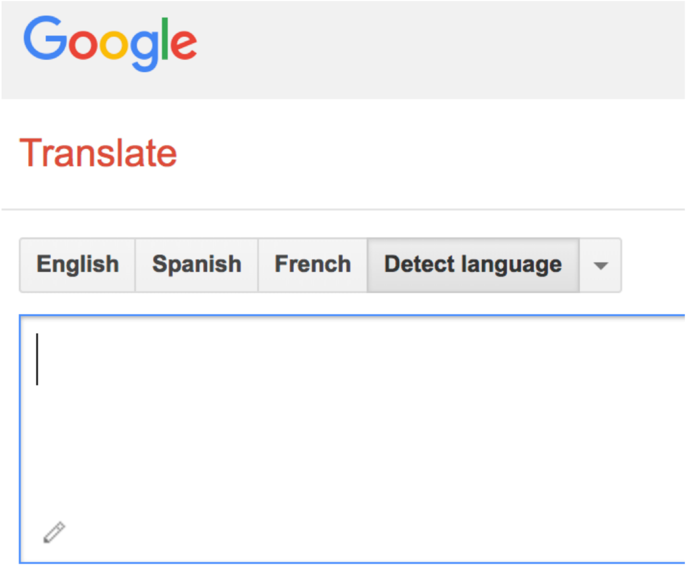
- Understand domain randomization & how it is being used today
- Discuss its limitations and what solutions could look like

# Deep learning is data-hungry…

## ImageNet

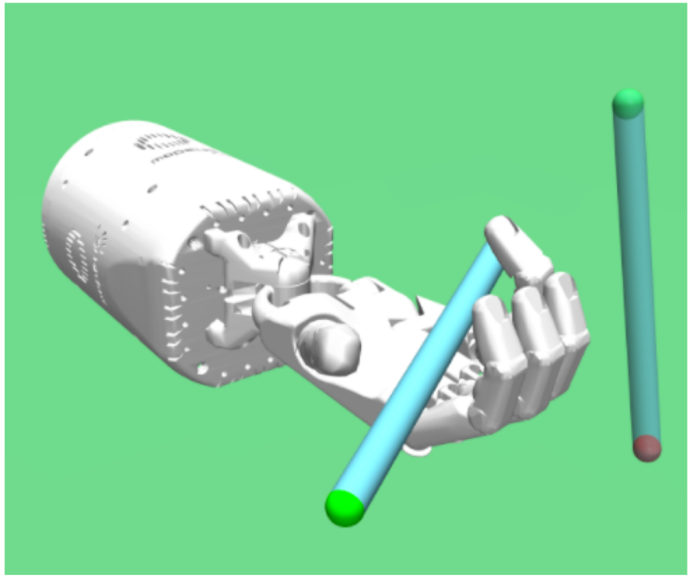
1.2M labeled images

## Machine Translation




36M sentence pairs (WMT En->Fr)
"Several orders of magnitude more" (production data)

## DeepRL

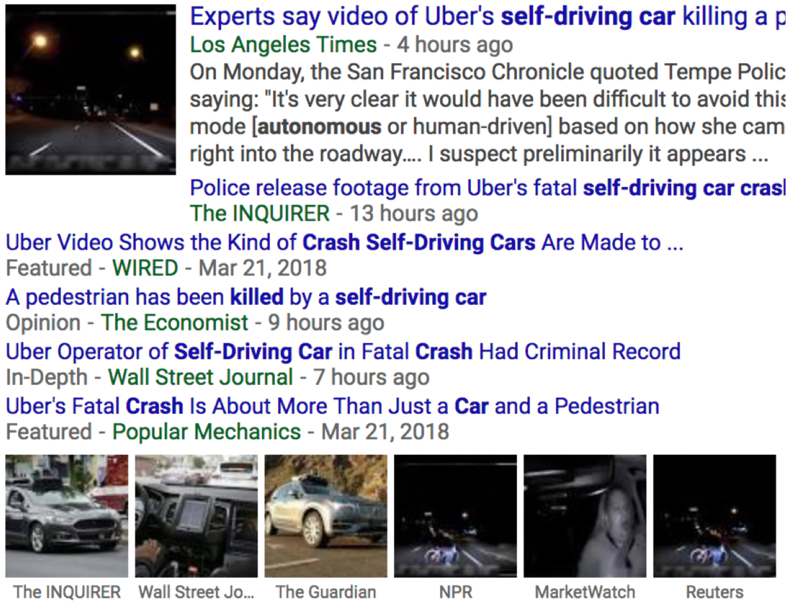


38M timesteps

# ...But robotic data is expensive

## Robot cost


## Safety




## Labeling

# Advantages of simulated data

Cheaper

Faster

Scalable

Labeled

# But does simulated data work?

***"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]

# How to bridge the gap?

- Better simulation

# Are better simulators enough?

## Models overfit to any difference

**Virtual KITTI Dataset**

**Multi-object tracking accuracy:**
**Sim: 63.7%**
**Real: 78.1%**

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

## High quality is expensive

**Jungle Book:**
**30M render hours**
**19 hours per frame**
**800 artist-years of effort**

**Jungle Book, 2016**

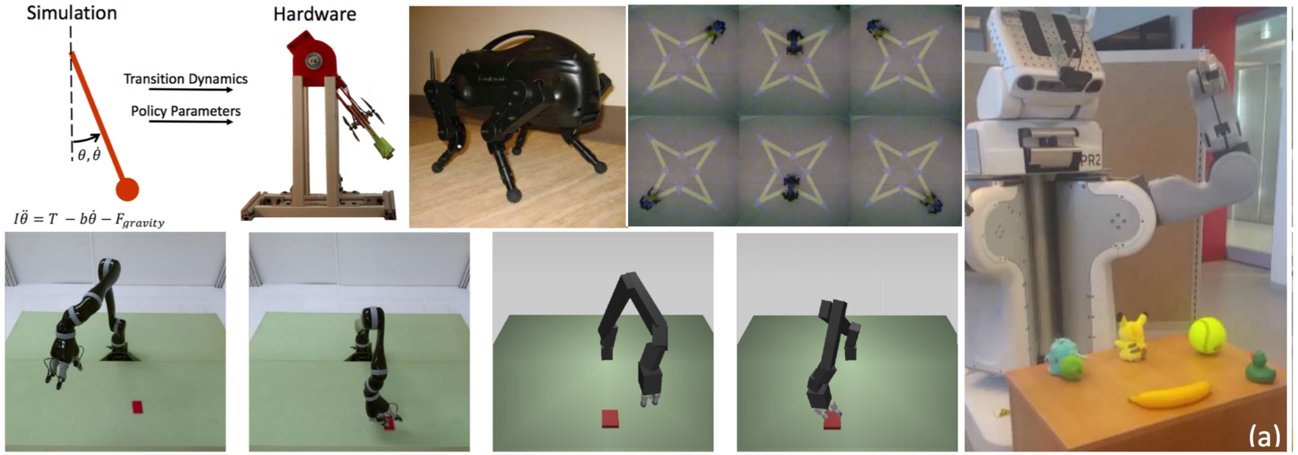
**Toward Understanding Stories From Videos**
[Sanja Fidler, NIPS Deep Learning Workshop 2016]

# How to bridge the gap?

- Better simulation
- Domain adaptation

# Supervised domain adaptation

## Fine-tuning



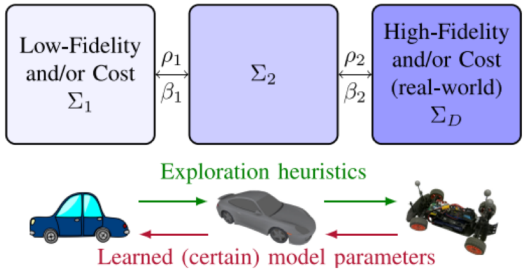

**Learning Omnidirectional Path Following Using Dimensionality Reduction** [Kolter, Ng, 2003]

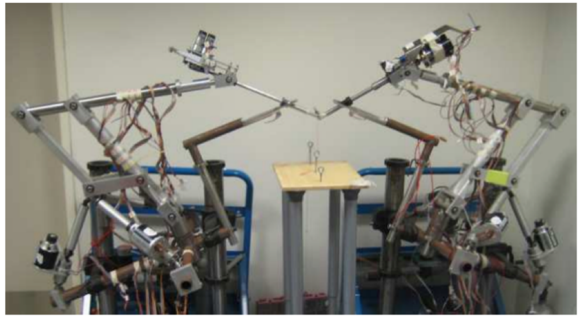
**Efficient Reinforcement Learning for Robotics using Informative Simulated Priors** [Cutler, How, 2015]

**Sim-to-Real Robot Learning from Pixels with Progressive Nets** [Rusu et al. 2016]

**Deep Predictive Policy Training using Reinforcement Learning** [Ghadirzadeh, Maki, Kragic, Bjorkman, 2017]

## Iterative learning control



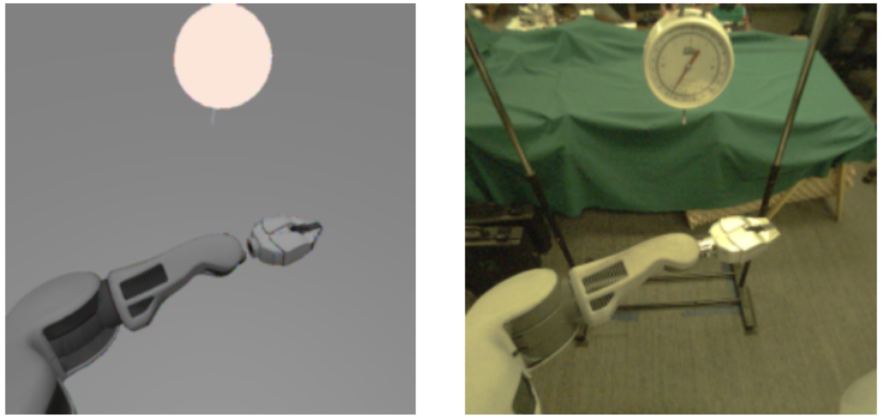
**Using inaccurate models in reinforcement learning** [Abbeel, Quigley, Ng, 2006]

**Reinforcement learning with multi-fidelity simulators** [Cutler, Walsh, How 2014]

**Superhuman performance of surgical tasks by robots using iterative learning from human-guided demonstrations** [Van Den Berg, Miller, Duckworth, Hu, Wan, Fu, Goldberg, Abbeel, 2010]

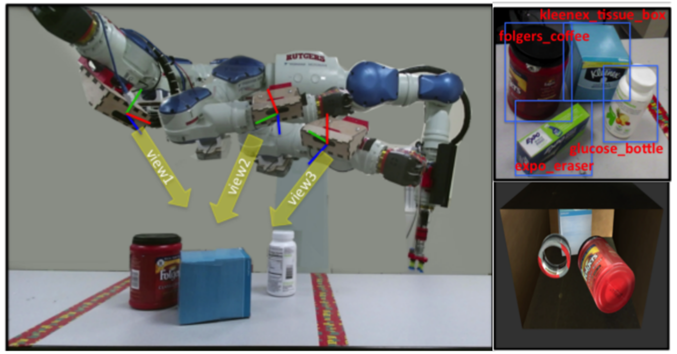
# (Less) supervised domain adaptation

## Weakly Supervised



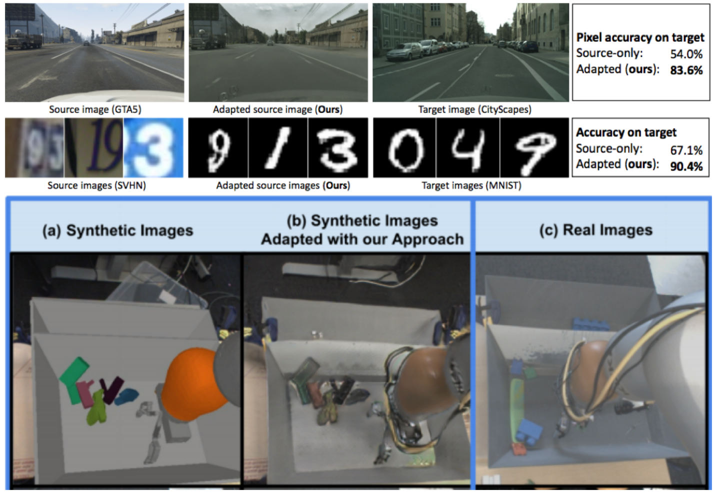
**Adapting Deep Visuomotor Representations with Weak Pairwise Constraints** [Tzeng, Devin, Hoffman, Finn, Abbeel, Levine, Saenko, Darrell, 2016]

## Self-Supervised




**A Self-supervised Learning System for Object Detection using Physics Simulation and Multi-view Pose Estimation** [Mitash, Bekris, Boularias, 2017]

## Unsupervised

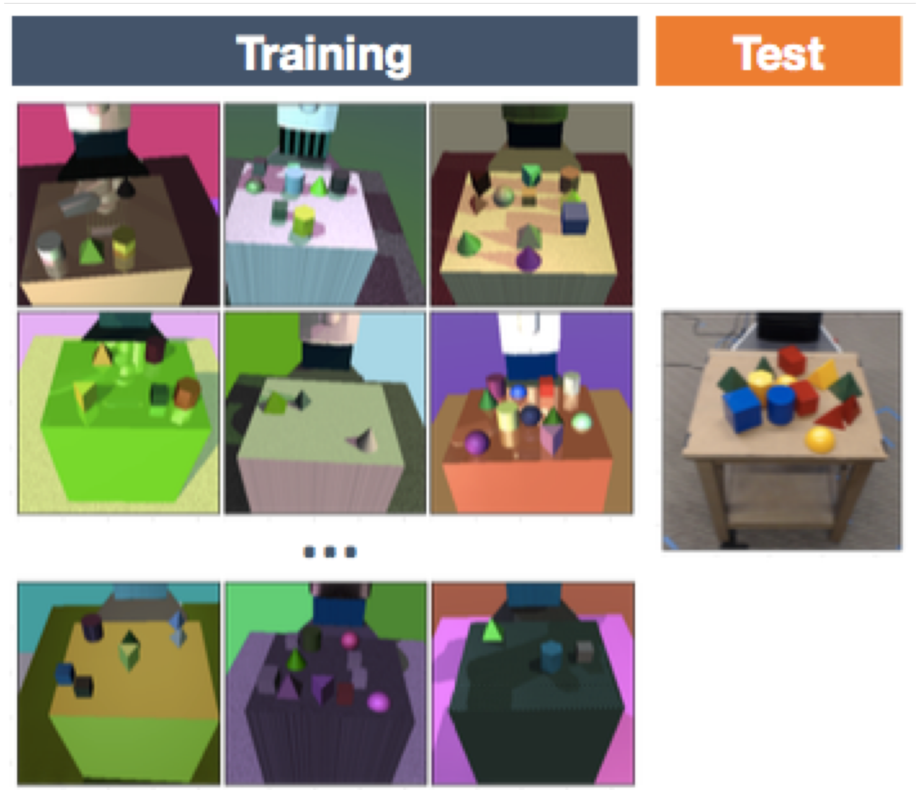



**CyCADA** [Hoffman, Tzeng, Park, Zhu, Isola, Saenko, Efros, Darrel, 2017]
**Using Simulation and Domain Adaptation to Improve Efficiency of Deep Robotic Grasping** [Bousmalis et al., 2017]

# How to bridge the gap?

- Better simulation
- Domain adaptation
- Domain randomization

# Domain Randomization




If the model sees enough simulated variation, the real world may look like just the next simulator

# Domain Randomization

- **History**
- Appearance randomization
- Scene / object randomization
- Physics randomization
- Frontiers

# Radical Envelope of Noise Hypothesis

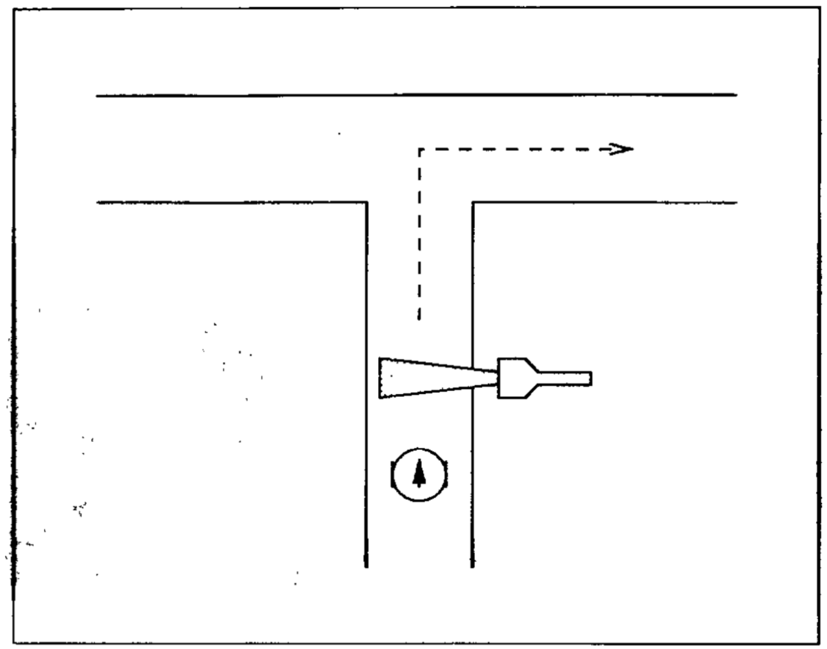
## Create a "minimal simulation" consisting of:

### 1. Base Set

- Aspects of the simulator that are "sufficient to underlie the behavior we want"
- These will be measured and then randomized a bit for robustness

### 2. Implementation aspects

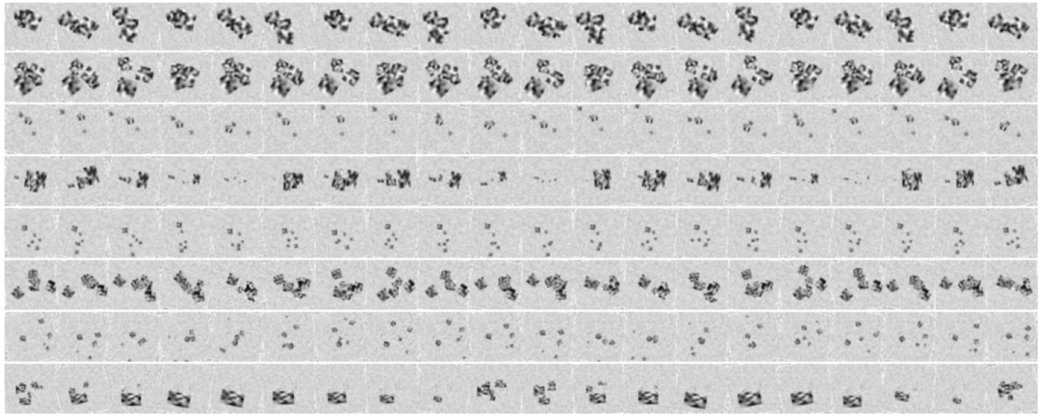
- All other aspects, which do not have a basis in reality in the simulator
- These will be randomized enough so successful controllers "ignore each implementation aspect entirely"



**Evolutionary Robotics and the Radical Envelope of Noise Hypothesis**
[Nick Jakobi, 1997]

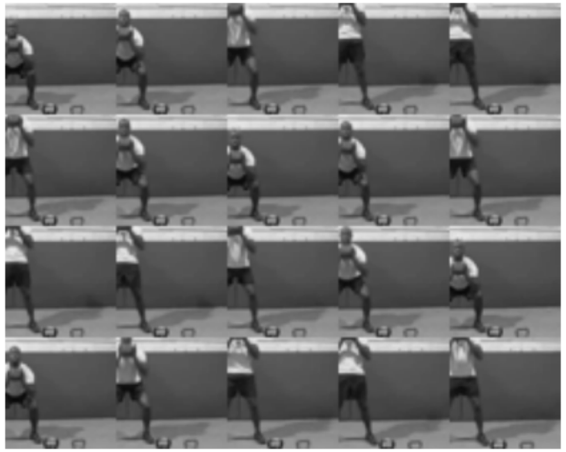
# Live Repetition Counting

## Training



**Predict cycle length of periodic random noise**

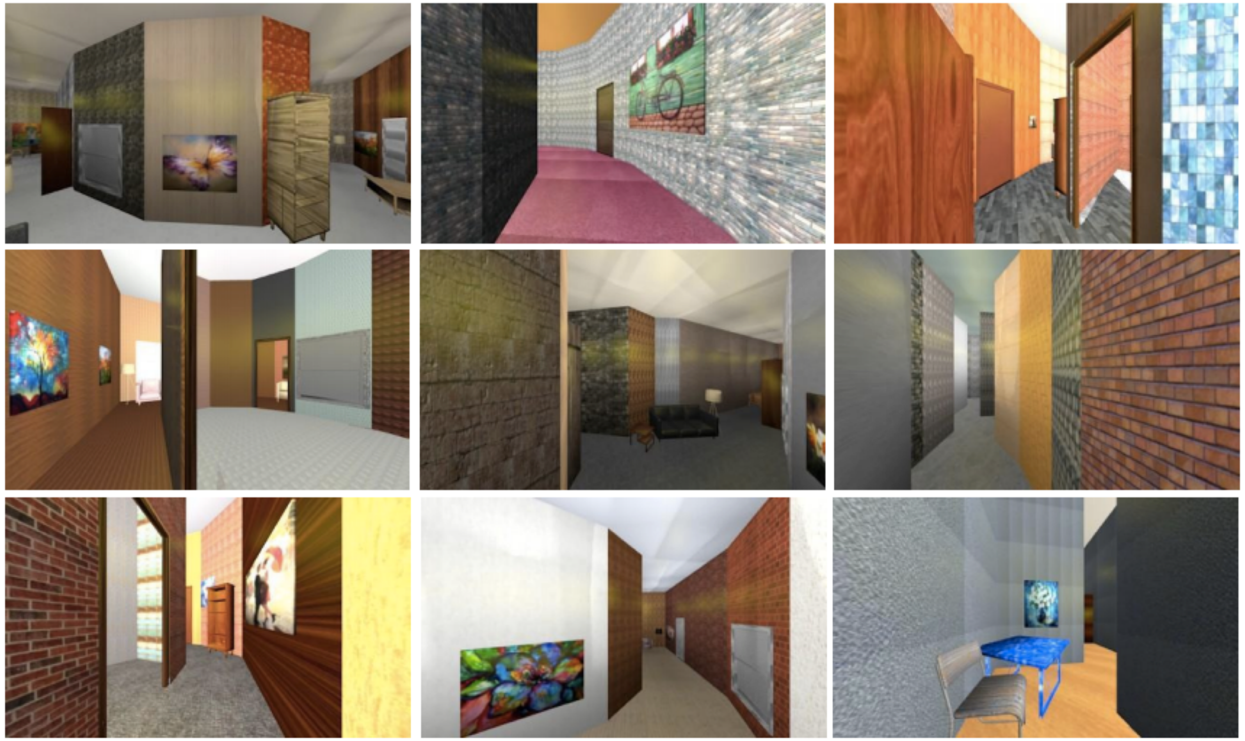
## Test



**Count repetitive behavior by integrating the predicted period**

**Live Repetition Counting**
[Levy & Wolf, 2015]

# CAD$^2$ RL



- **Quadcopter collision avoidance**
- **~500 semi-realistic textures, 12 floorplans**
- **~40-50% of 1000m trajectories are collision-free**

**(cad)^2 RL: Real Single-Image Flight Without a Single Real Image**
[Sadeghi & Levine, 2016]

# Other related work


Yair Movshovitz-Attias, Takeo Kanade, and Yaser Sheikh. How useful is photo-realistic rendering for visual learning? In Computer Vision– ECCV 2016 Workshops, pages 202–217. Springer, 2016.

Su, Hao, et al. "Render for cnn: Viewpoint estimation in images using cnns trained with rendered 3d model views." Proceedings of the IEEE International Conference on Computer Vision. 2015.

Shafaei, Alireza, James J. Little, and Mark Schmidt. "Play and learn: Using video games to train computer vision models." arXiv preprint arXiv:1608.01745 (2016).

Richter, Stephan R., et al. "Playing for data: Ground truth from computer games." European Conference on Computer Vision. Springer, Cham, 2016.

M. Johnson-Roberson, C. Barto, R. Mehta, S. N. Sridhar, Karl Rosaen,and R. Vasudevan, “Driving in the matrix: Can virtual worlds replace human-generated annotations for real world tasks?,” in IEEE International Conference on Robotics and Automation, pp. 1–8, 2017.

Hao Su, Charles R Qi, Yangyan Li, and Leonidas J Guibas. Render for cnn: Viewpoint estimation in images using cnns trained with rendered 3d model views. In Proceedings of the IEEE International Conference on Computer Vision, pages 2686–2694, 2015.

David G Lowe. Three-dimensional object recognition from single two-dimensional images. Artificial intelligence, 31(3):355–395, 1987

Baochen Sun and Kate Saenko. From virtual to reality: Fast adaptation of virtual object detectors to real domains. In BMVC, volume 1, page 3, 2014.

Gaidon, Adrien, et al. "Virtual worlds as proxy for multi-object tracking analysis." arXiv preprint arXiv:1605.06457 (2016).

Xingchao Peng, Baochen Sun, Karim Ali, and Kate Saenko. Learning deep object detectors from 3d models. In Proceedings of the IEEE International Conference on Computer Vision, pages 1278–1286, 2015.

Ramakant Nevatia and Thomas O Binford. Description and recognition of curved objects. Artificial Intelligence, 8(1):77–98, 1977.

McCormac, John, et al. "SceneNet RGB-D: 5M Photorealistic Images of Synthetic Indoor Trajectories with Ground Truth." arXiv preprint arXiv:1612.05079 (2016).

de Souza, César Roberto, et al. "Procedural Generation of Videos to Train Deep Action Recognition Networks." arXiv preprint arXiv:1612.00881 (2016).

Ros, German, et al. "The synthia dataset: A large collection of synthetic images for semantic segmentation of urban scenes." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016.

Mahendran, A., et al. "ResearchDoom and CocoDoom: Learning Computer Vision with Games." arXiv preprint arXiv:1610.02431 (2016).

Vazquez, David, et al. "Virtual and real world adaptation for pedestrian detection." IEEE transactions on pattern analysis and machine intelligence 36.4 (2014): 797-809.

# Our Approach: More Variability, More Data, Less Fidelity

**100K highly randomized scenes with unrealistic textures**

Tobin, Josh, et al. "Domain randomization for transferring deep neural networks from simulation to the real world." *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2017.

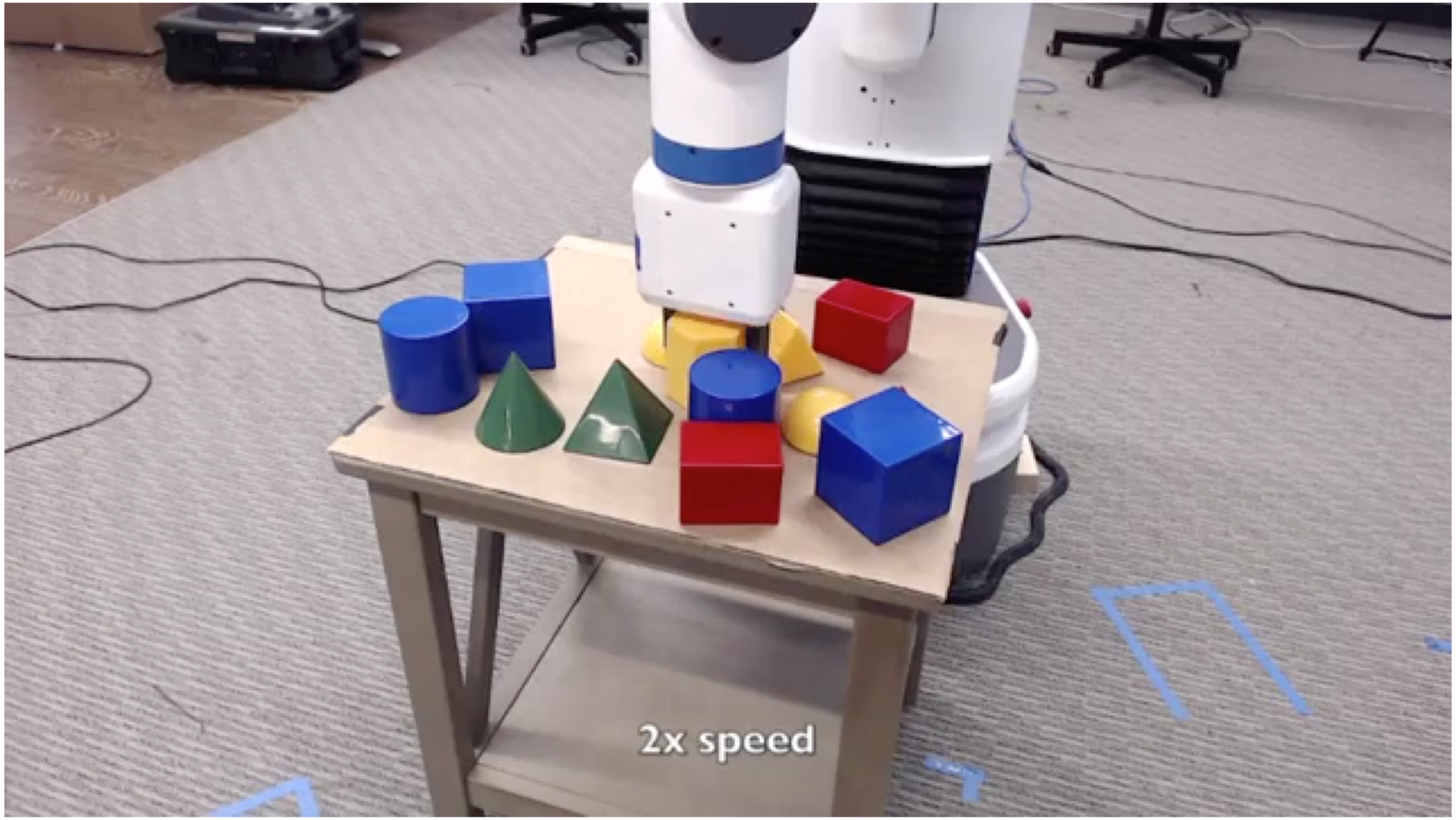
# Domain Randomization

- History
- **Appearance randomization**
- Scene / object randomization
- Physics randomization
- Frontiers

# What do we randomize?

- Texture & material properties of all objects, table, background, robot
  - Textures are colors, color gradients, or texture patterns
- Position of cameras (within a small range)
- Lighting position, orientation, color, and specular properties
- Distractor objects in the scene

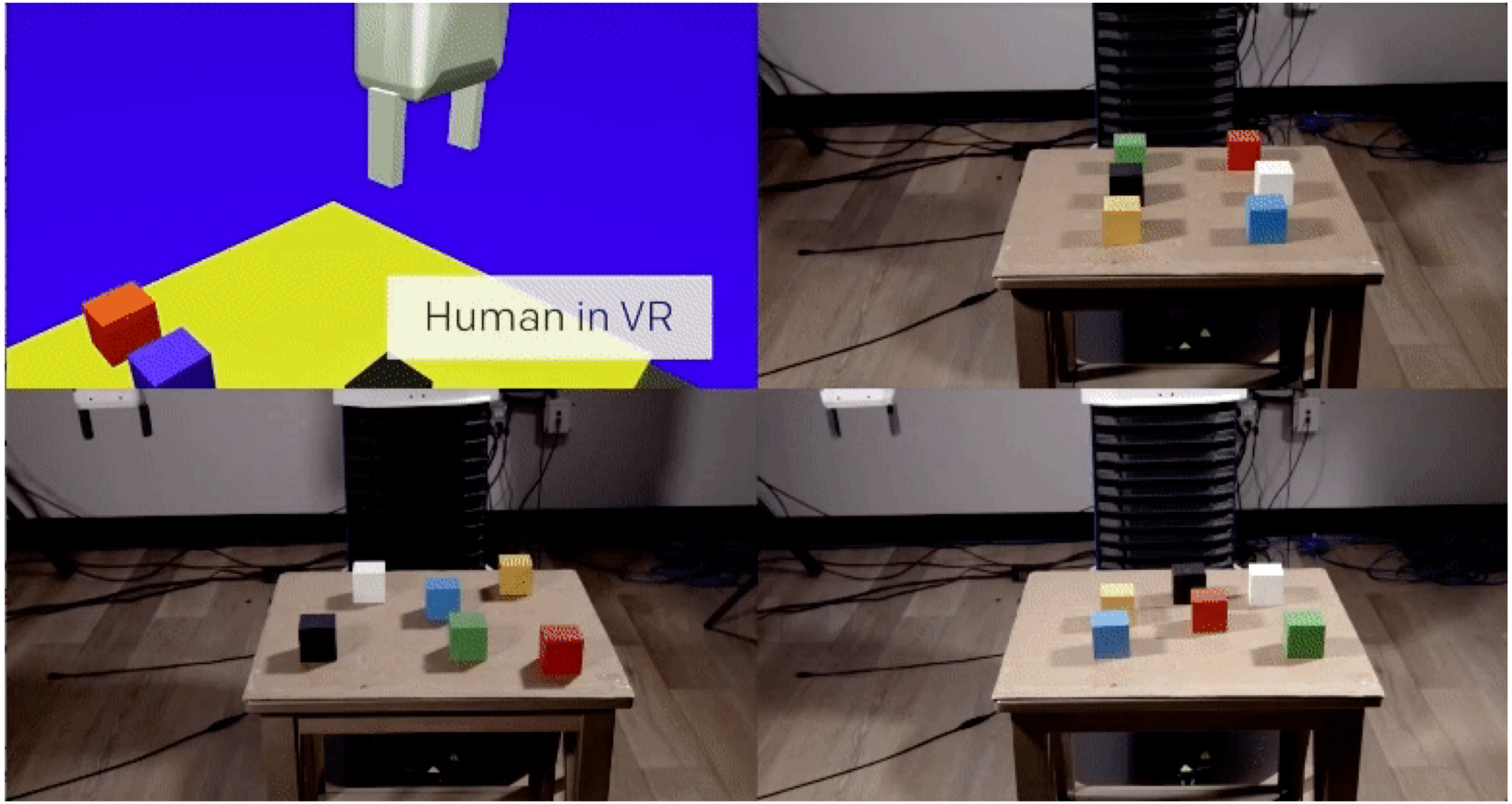
# Applications



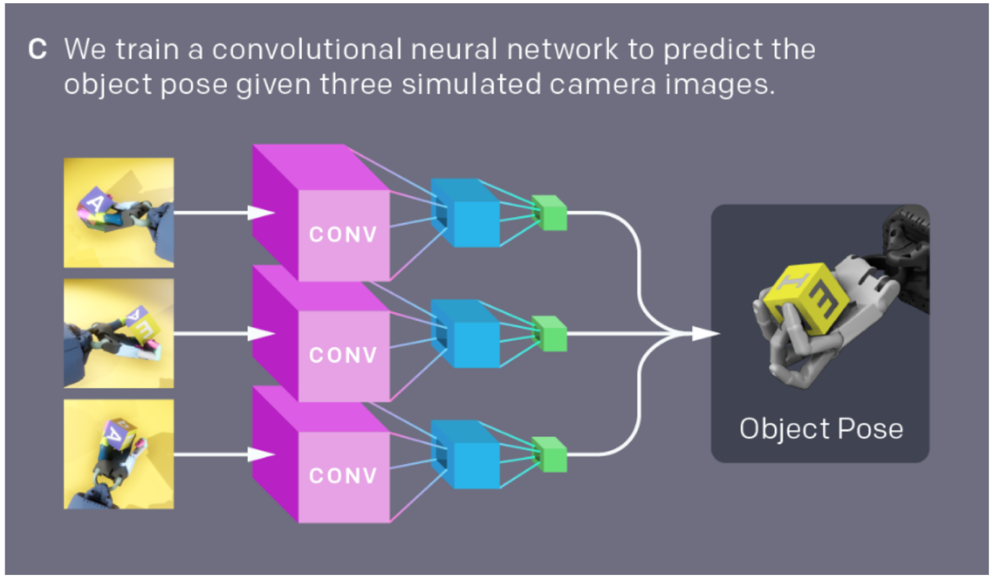

Tobin, Josh, et al. "Domain randomization for transferring deep neural networks from simulation to the real world." *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2017.

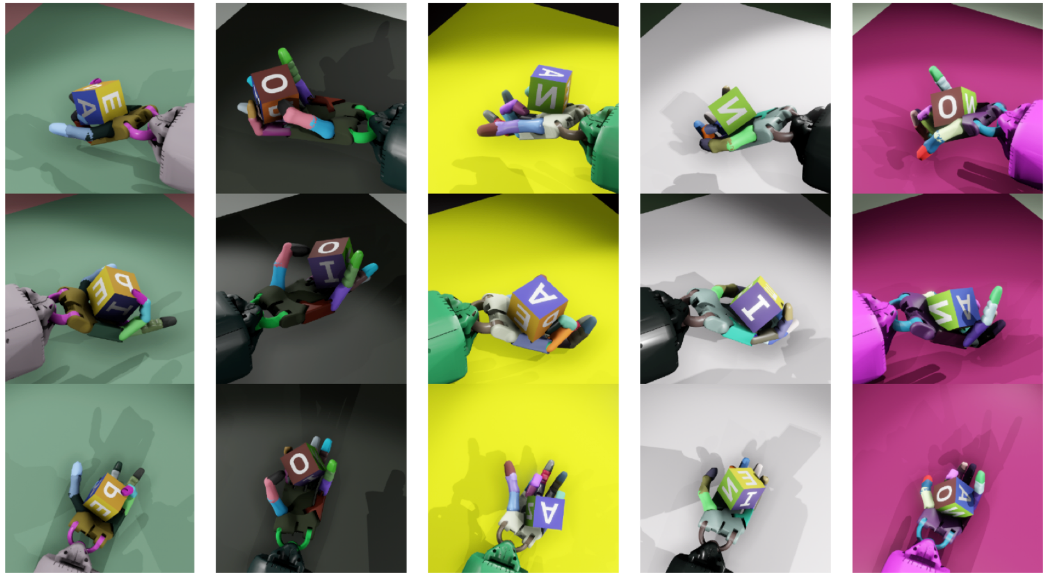
# Applications

# Applications



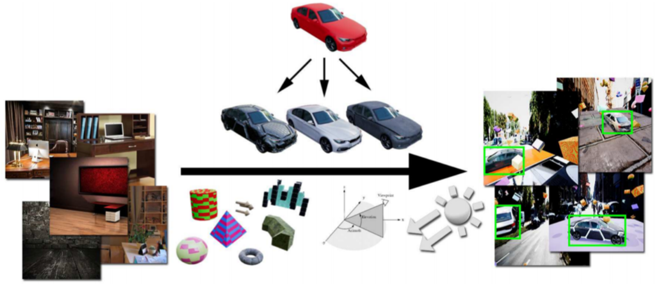
# Applications



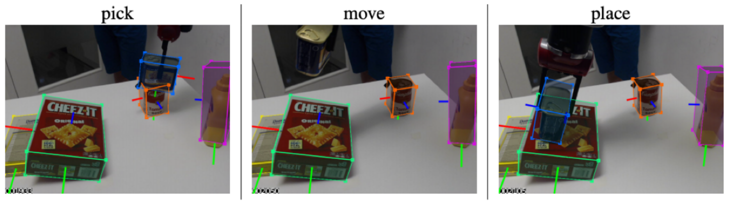
# Selected additional applications

Autonomous vehicles

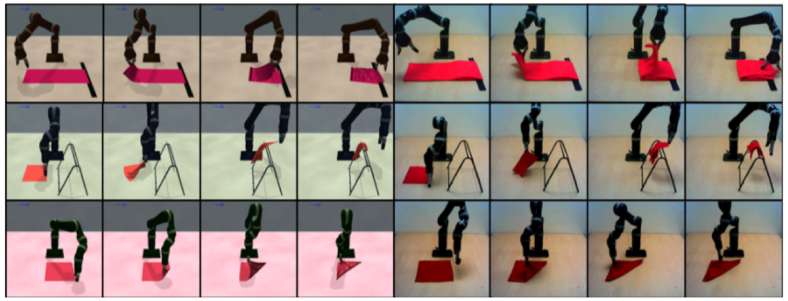


Face tracking

Manipulation

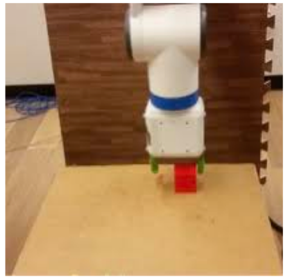



Cloth manipulation

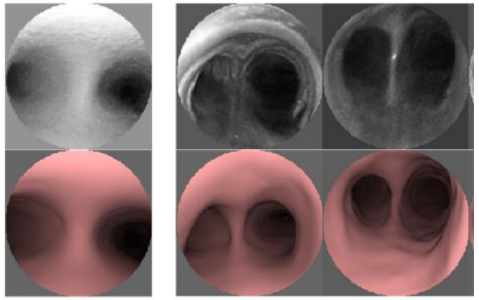


Shiny / reflective objects

Visuomotor policies



Surgical robotics

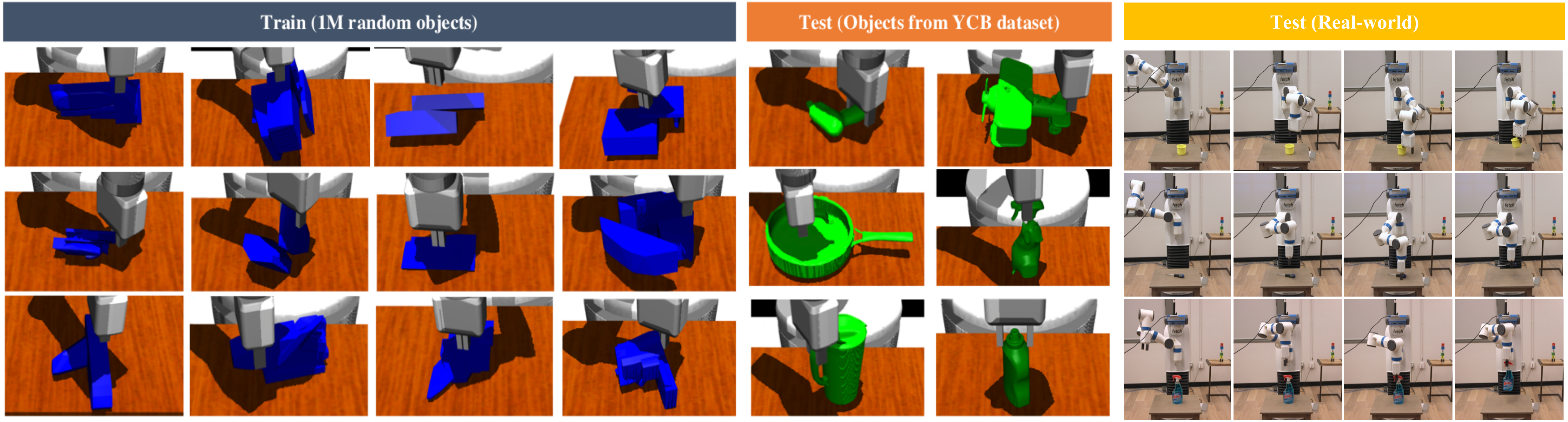


Jonathan Tremblay et al. “Training deep networks with synthetic data: Bridging the reality gap by domain randomization”. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. 2018, pp. 969–977.
Jonathan Tremblay et al. “Deep object pose estimation for semantic robotic grasping of household objects”. In: arXiv preprint arXiv:1809.10790 (2018).
Mikko Ronkainen et al. “Dense tracking of human facial geometry-aware”. In: (2017).
Jan Matas, Stephen James, and Andrew J Davison. “Sim-to-real reinforcement learn- ing for deformable object manipulation”. In: arXiv preprint arXiv:1806.07851 (2018).
Jonatan S Dyrstad and John Reidar Mathiassen. “Grasping virtual fish: A step to- wards robotic deep learning from demonstration in virtual reality”. In: 2017 IEEE In- ternational Conference on Robotics and Biomimetics (ROBIO). IEEE. 2017, pp. 1181– 1187.
Lerrel Pinto et al. “Asymmetric actor critic for image-based robot learning”. In: arXiv preprint arXiv:1710.06542 (2017).
Sganga, Jake, et al. "Deep Learning for Localization in the Lung." arXiv preprint arXiv:1903.10554 (2019).

# Domain Randomization

- History
- Appearance randomization
- **Scene / object randomization**
- Physics randomization
- Frontiers

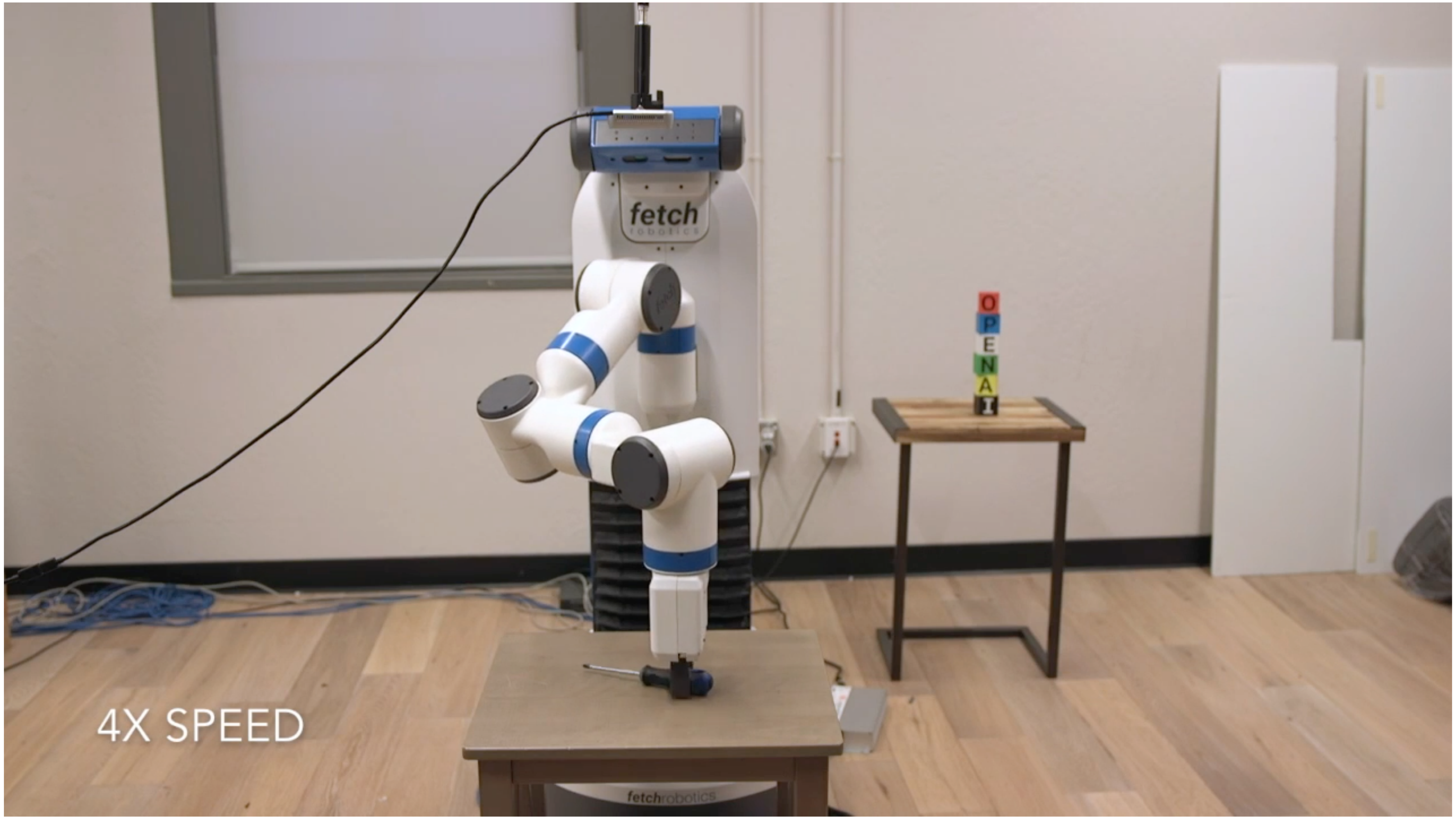
# How to avoid building object models?




**Hypothesis: If the model sees a wide enough range of (unrealistic) objects during training, at test time it will generalize to realistic objects**

Tobin, Josh, et al. "Domain randomization and generative models for robotic grasping." *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2018.

# Application



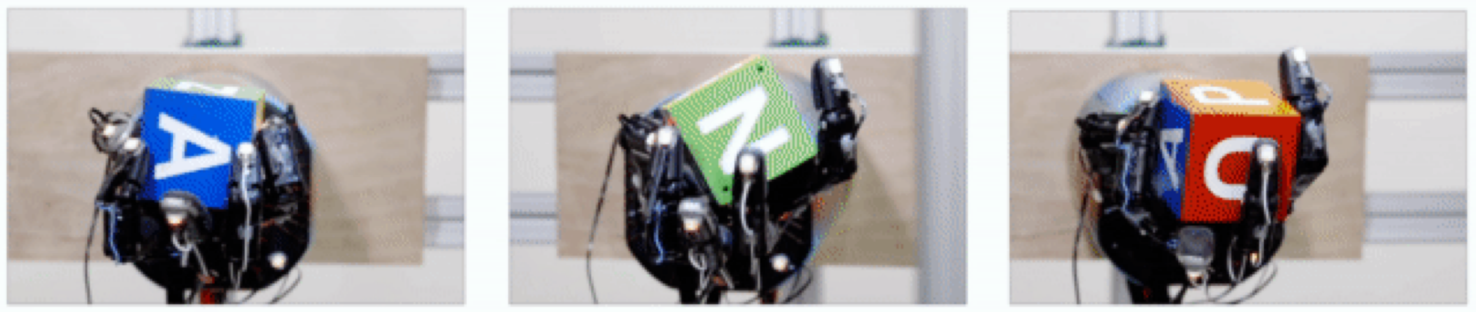
# Domain Randomization

- History
- Appearance randomization
- Scene / object randomization
- **Physics randomization**
- Frontiers

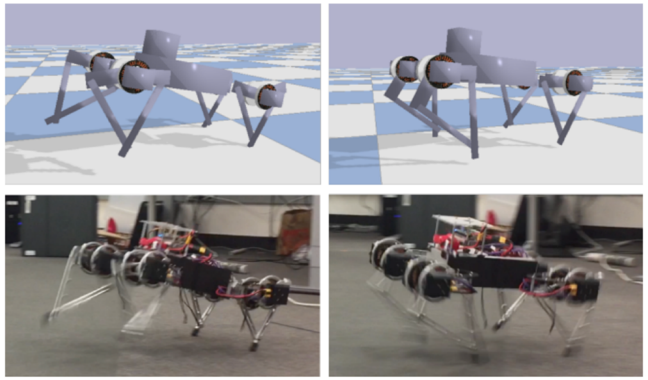
# What do we randomize?

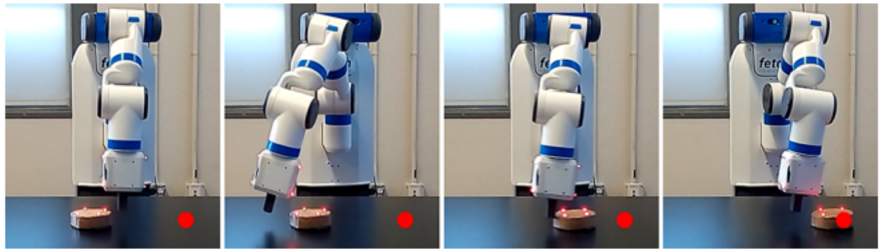
- Dimensions
- Masses
- Friction
- Damping
- Actuator gains
- Joint limits
- Gravity

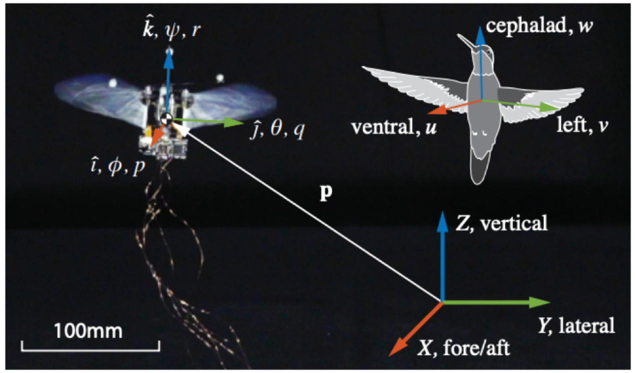
# Applications



FINGER PIVOTING    SLIDING    FINGER GAITING








Jie Tan et al. “Sim-to-real: Learning agile locomotion for quadruped robots”. In: arXiv preprint arXiv:1804.10332 (2018).

Xue Bin Peng et al. “Sim-to-real transfer of robotic control with dynamics randomiza- tion”. In: 2018 IEEE International Conference on Robotics and Automation (ICRA). IEEE. 2018, pp. 1–8.

Fan Fei et al. “Learning extreme hummingbird maneuvers on flapping wing robots”. In: arXiv preprint arXiv:1902.09626 (2019).

# Domain Randomization

- History
- Appearance randomization
- Scene / object randomization
- Physics randomization
- **Frontiers**

# Where might DR not work?

- Contact-rich manipulation
- Hard-to-simulate objects like cloth
- Highly varied environments

# Where might DR not work?

- Contact-rich manipulation
- Hard-to-simulate objects like cloth
- Highly varied environments

# Where might DR not work?

- ~~Contact-rich manipulation~~ (OpenAI et al, 2018)
- Hard-to-simulate objects like cloth
- Highly varied environments

# Where might DR not work?

- ~~Contact-rich manipulation~~
- ~~Hard-to-simulate objects like cloth~~ (Matas et al, 2018)
- Highly varied environments

# Where might DR not work?

- ~~Contact-rich manipulation~~
- ~~Hard-to-simulate objects like cloth~~ (Matas et al, 2018)
- Highly varied environments
    - Procedural generation
    - Massive object databases
    - Drive an RC car around your campus?

# How to make DR work better?

# How does DR work in practice?

- Build a simulated world
- Calibrate it to the environment
- Design randomizations to “cover” real-world variability
- Train a model and evaluate in real
- Examine failure modes and add randomization

# How does DR work in practice?

- Build a simulated world
- Calibrate it to the environment
- Design ra… world variability
- Train a m…
- Examine failure modes and add randomization

Highly manual, lots of human knowledge / intuition

# Automatically build worlds

## Current

- Scene graphs (e.g., Prakash et al, 2018, Kar et al, 2019)
- SFM / SLAM
- Inverse graphics

## Future

- Learn scene graphs from scratch
- Real – sim – real via inverse graphics

Prakash, Aayush, et al. "Structured Domain Randomization: Bridging the Reality Gap by Context-Aware Synthetic Data." arXiv preprint arXiv:1810.10093 (2018).
Kar, Amlan, et al. "Meta-Sim: Learning to Generate Synthetic Datasets." arXiv preprint arXiv:1904.11621 (2019).

# Calibrate & choose randomizations

## Current

- Minimize distance between real trajectories & sim trajectories (Chebotar et al, 2018)
- Choose simulations that make behaviors on a held-out environment look the same as in training (Mehta et al, 2019)
- Choose adversarial randomizations (Zakharov et al, 2019)
- Choose randomizations that aid transfer on current task (i.e., architecture search over randomizations) (Ruiz et al, 2019)

## Future

- Maximally entropic randomizations s.t. task performance doesn't degrade
- Efficient neural architecture search (E-NAS, population-based augmentation) with / without task performance
- Less constrained adversarial randomizations (e.g., with a GAN). How to ensure task remains solvable?
- Tools


Yevgen Chebotar et al. "Closing the sim-to-real loop: Adapting simulation random- ization with real world experience". In: arXiv preprint arXiv:1810.05687 (2018).
Bhairav Mehta et al. "Active Domain Randomization". In: arXiv preprint arXiv:1904.04762 (2019).
Sergey Zakharov, Wadim Kehl, and Slobodan Ilic. "DeceptionNet: Network-Driven Domain Randomization". In: arXiv preprint arXiv:1904.02750 (2019).
Ruiz, Nataniel, Samuel Schulter, and Manmohan Chandraker. "Learning to simulate." arXiv preprint arXiv:1810.02513 (2018).
Pham, Hieu, et al. "Efficient neural architecture search via parameter sharing." arXiv preprint arXiv:1802.03268 (2018).
Ho, Daniel, et al. "Population Based Augmentation: Efficient Learning of Augmentation Policy Schedules." arXiv preprint arXiv:1905.05393 (2019).

# Train and evaluate

## Current

- Successfully solve a wider range of simulation parameters in simulation
  - Search over sim params at test time (Yu et al, 2018)
- Estimate real-world performance inexpensively
  - Estimate transfer performance (Muratore et al, 2018)
- Which model structures work best? (James et al., 2018)

## Future

- Successfully solve a wider range of simulation parameters in simulation
  - Search over a lower-dimensional space at test time like (Kolter, Ng 2007)
  - From 0-shot to few-shot eval via meta learning
- Estimate real-world performance inexpensively
  - Cheaper transfer metrics
  - Learn an informative sampling policy
- More investigation of architectures & policies

Yu, Wenhao, C. Karen Liu, and Greg Turk. "Policy transfer with strategy optimization." arXiv preprint arXiv:1810.05751 (2018).
Muratore, Fabio, et al. "Domain Randomization for Simulation-Based Policy Optimization with Transferability Assessment." Conference on Robot Learning. 2018.
Kolter, J. Zico, and Andrew Y. Ng. "Learning omnidirectional path following using dimensionality reduction." Robotics: Science and Systems. 2007.
Stephen James et al. “Sim-to-Real via Sim-to-Sim: Data-efficient Robotic Grasping via Randomized-to-Canonical Adaptation Networks”. In: arXiv preprint arXiv:1812.07252 (2018).

# Update randomizations

## Current

- Model-based RL (e.g., Clavera et al, 2018)
- Iterative learning control (e.g., Cutler et al, 2014, Chebotar et al, 2018)

## Future

- Explore ILC framework with an explicit notion of incorporating randomness
- Better baselines – e.g., combine DR with domain adaptation techniques
- Explicitly learn an unsupervised domain adaptation strategy via meta-learning (e.g., see Finn et al, 2017)

Clavera, Ignasi, et al. "Model-based reinforcement learning via meta-policy optimization." arXiv preprint arXiv:1809.05214 (2018).
Mark Cutler, Thomas J Walsh, and Jonathan P How. "Reinforcement learning with multi-fidelity simulators". In: Robotics and Automation (ICRA), 2014 IEEE Interna- tional Conference on. IEEE. 2014, pp. 3888–3895.
Yevgen Chebotar et al. "Closing the sim-to-real loop: Adapting simulation random- ization with real world experience". In: arXiv preprint arXiv:1810.05687 (2018).
Finn, Chelsea, et al. "One-shot visual imitation learning via meta-learning." arXiv preprint arXiv:1709.04905 (2017).

# Thanks to my collaborators

Pieter Abbeel

Woj Zaremba

**And:**
Marcin Andrychowicz, Lukas Biewald, Rocky Duan, Rachel Fong, Ankur Handa, Vikash Kumar, Bob McGrew, Alex Ray, Jonas Schneider, Peter Welinder

# Questions?

josh@openai.com

@josh_tobin_