

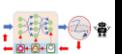
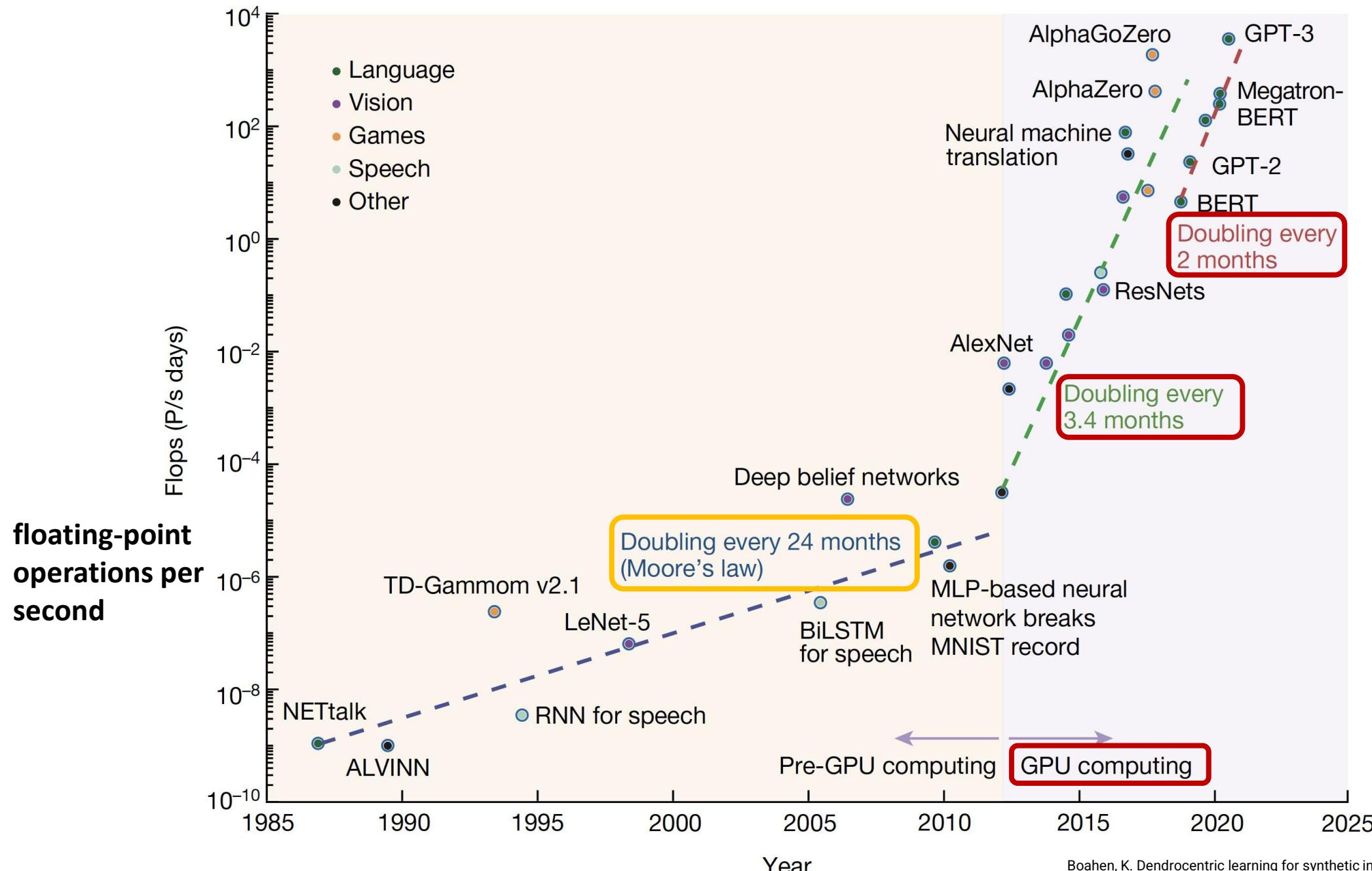
RBE474X/595-B01-ST: Deep Learning For Perception

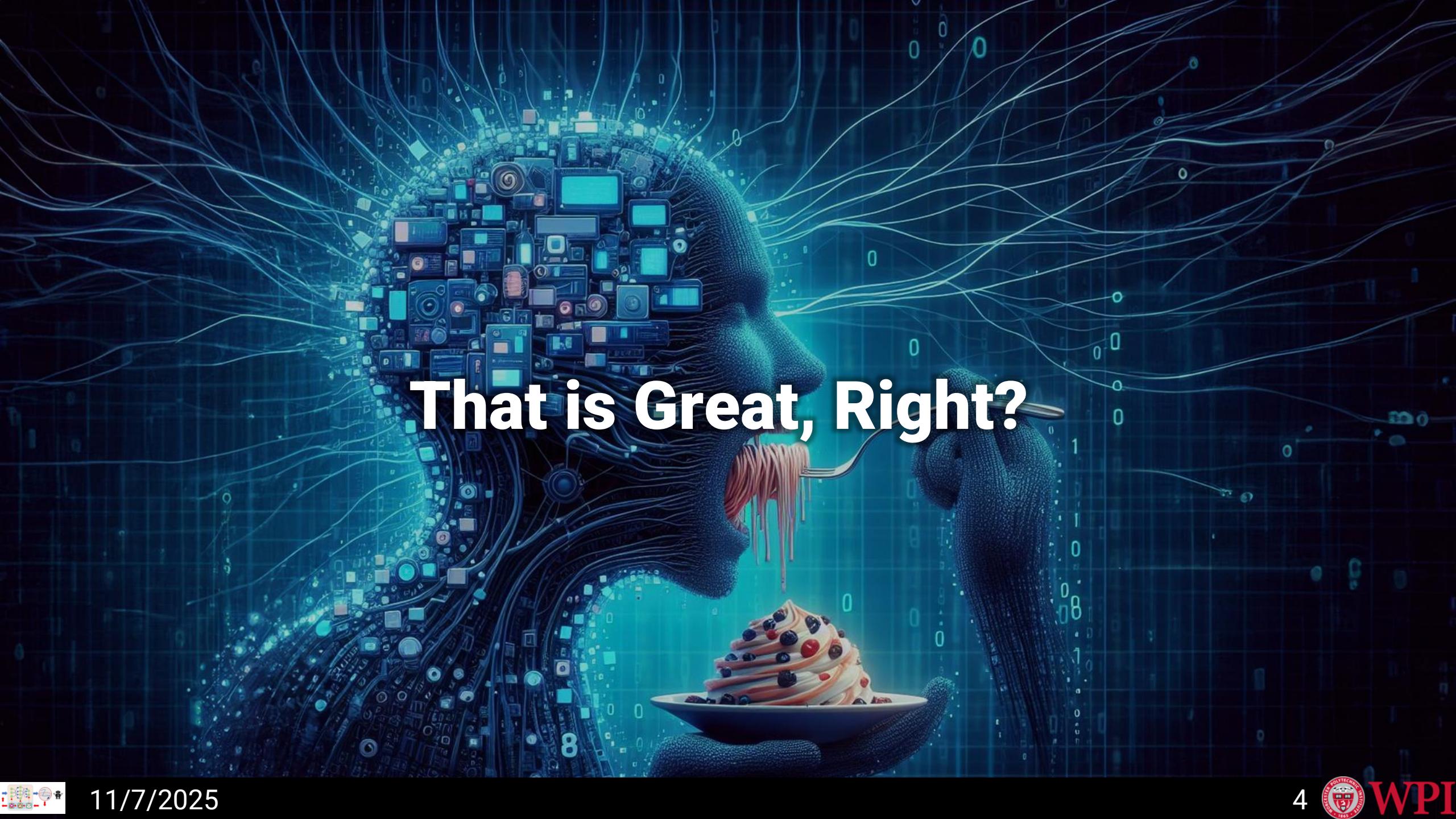
Class 5: Simulation for Data Generation And Sim2Real

Prof. Wei Xiao

Cluster Access







A large, complex brain composed of glowing blue and white circuit boards and binary code (0s and 1s) is shown against a dark background. A hand, also composed of binary code, holds a spoonful of pink frosting with blueberries over a bowl of similar frosting topped with berries. The scene is set against a grid pattern with floating binary digits.

That is Great, Right?



Data Is The Key!

Moore's Law for Machine Learning

"The amount of training data doubles every eighteen months."

- Danny Lange, Unity

Best Data

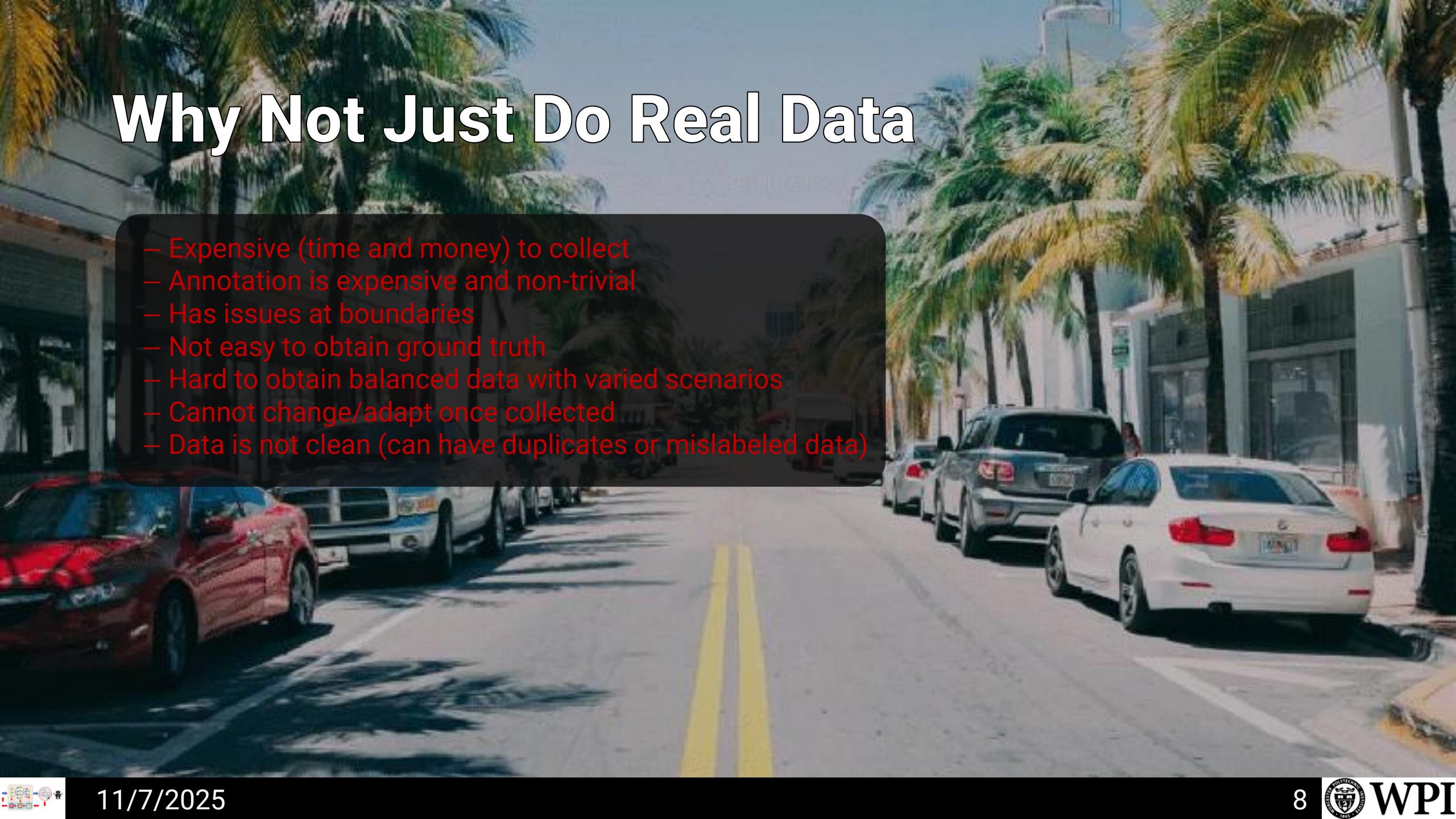
Real Data

- + Reflects what we want
- + No guesswork
- + Easier to verify performance and failure modes



Why Not Just Do Real Data

- Expensive (time and money) to collect
- Annotation is expensive and non-trivial
- Has issues at boundaries
- Not easy to obtain ground truth
- Hard to obtain balanced data with varied scenarios
- Cannot change/adapt once collected
- Data is not clean (can have duplicates or mislabeled data)



Tricks To Increase Data

Data Augmentation



Tricks To Increase Data

Data Augmentation



Tricks To Increase Data

Data Augmentation



Tricks To Increase Data

Data Augmentation



Tricks To Increase Data

Data Augmentation



Tricks To Increase Data

Data Augmentation



Tricks To Increase Data

Data Augmentation



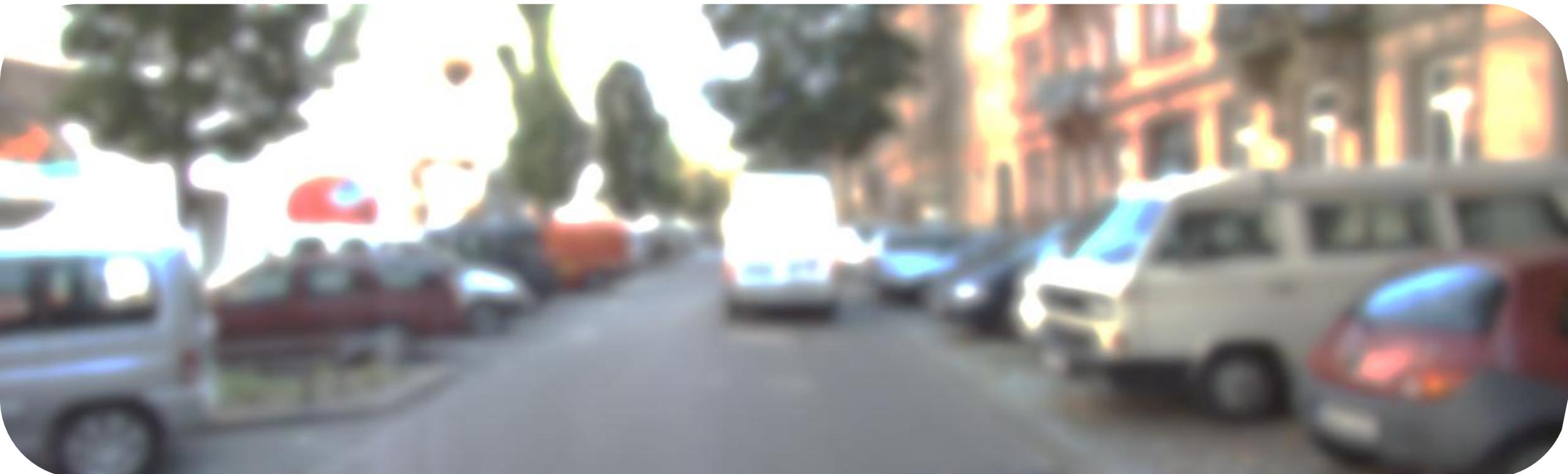
Tricks To Increase Data

Data Augmentation



Tricks To Increase Data

Data Augmentation



Tricks To Increase Data

Data Augmentation



Tricks To Increase Data

Data Augmentation



Tricks To Increase Data

Data Augmentation



Tricks To Increase Data

Data Augmentation



Tricks To Increase Data

Data Augmentation



Tricks To Increase Data

Data Augmentation



Tricks To Increase Data

Data Augmentation



SOURCE: ANOOP KHATRA VIA STORYFUL



storyful.

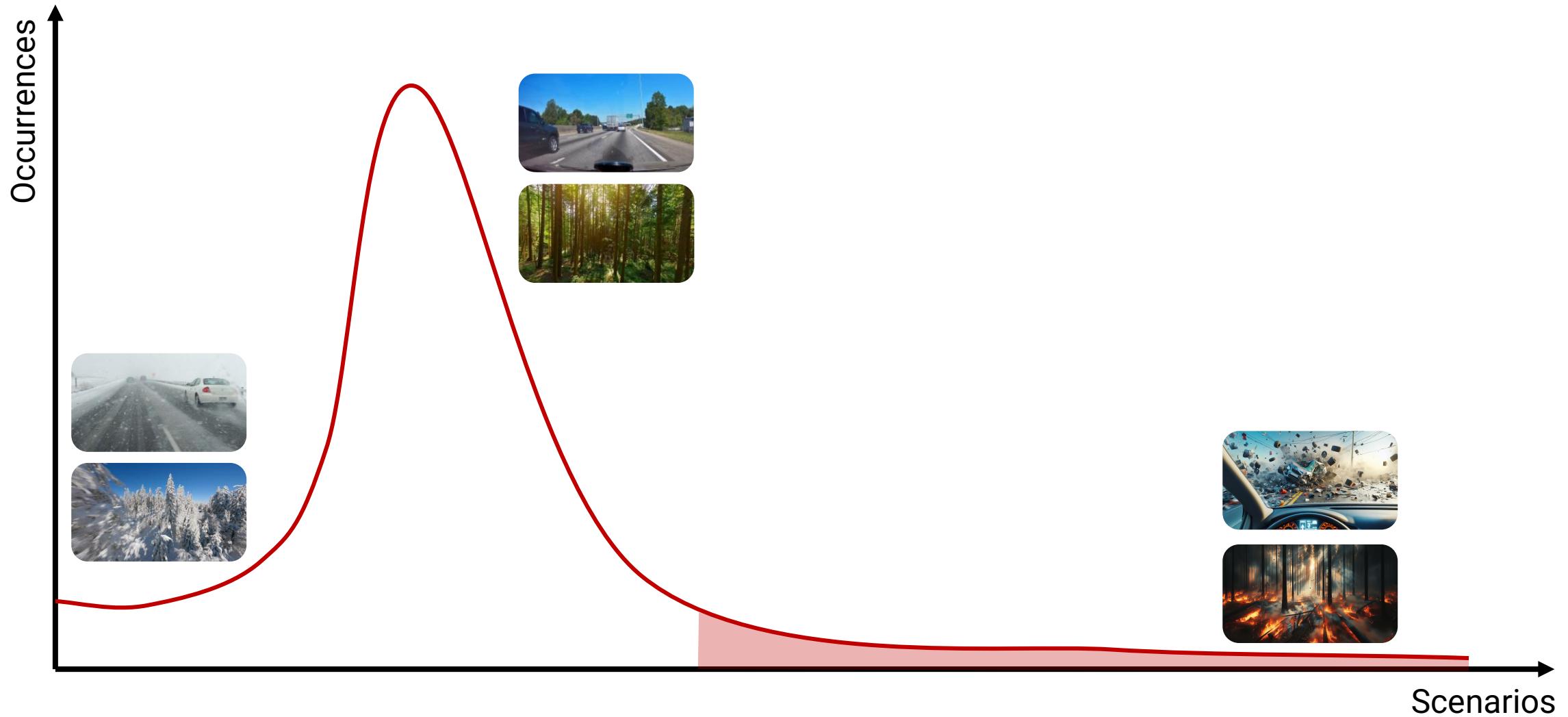
CTV
NEWS

The
Guardian





Data Distribution



IMPLAUSIBLE



EXPENSIVE



DESTRUCTIVE



ARDUOUS



IMPLAUSIBLE



EXPENSIVE



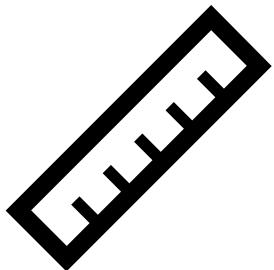
DESTRUCTIVE



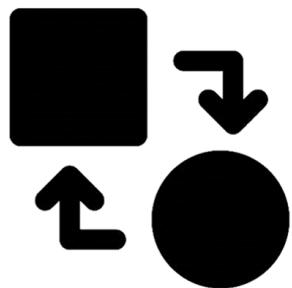
ARDUOUS



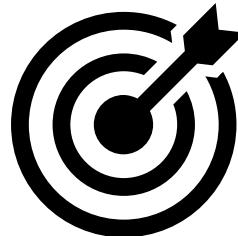
Why Deep Learning?



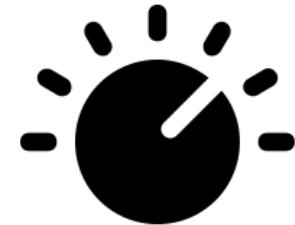
Scalable



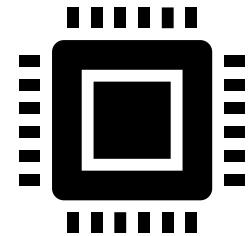
Adaptive



Accurate

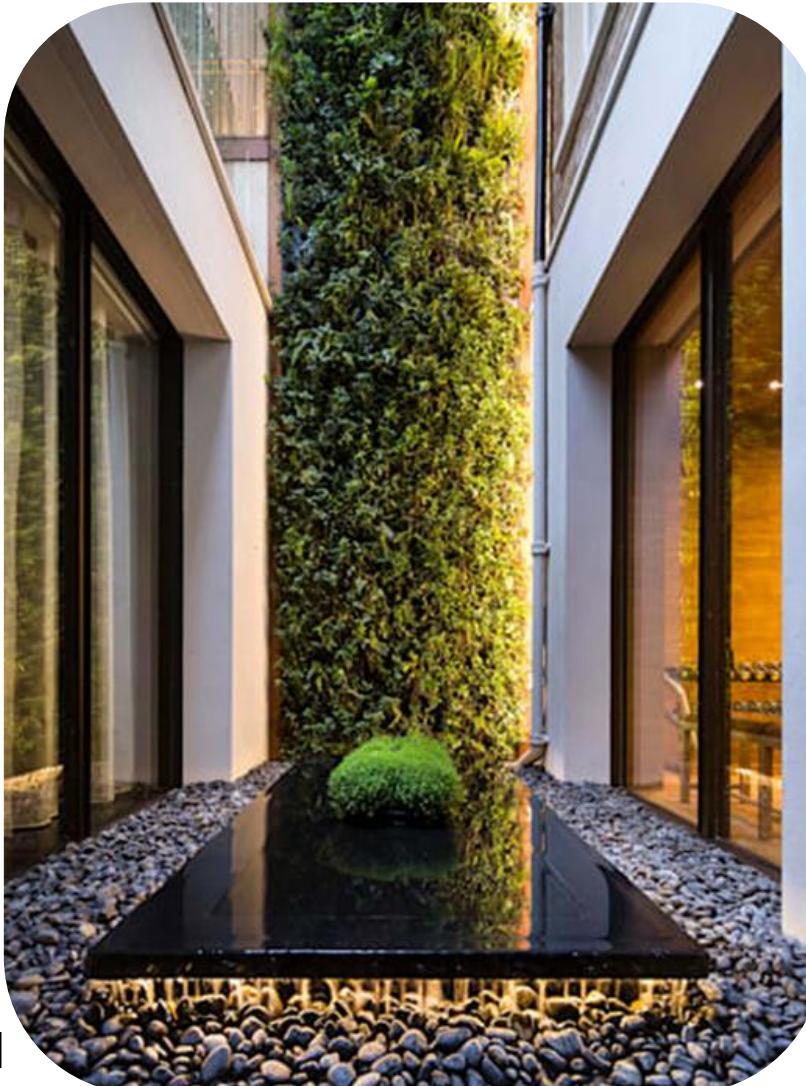


Accelerated



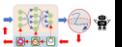
Tunable

Can We Get “Free” Data?

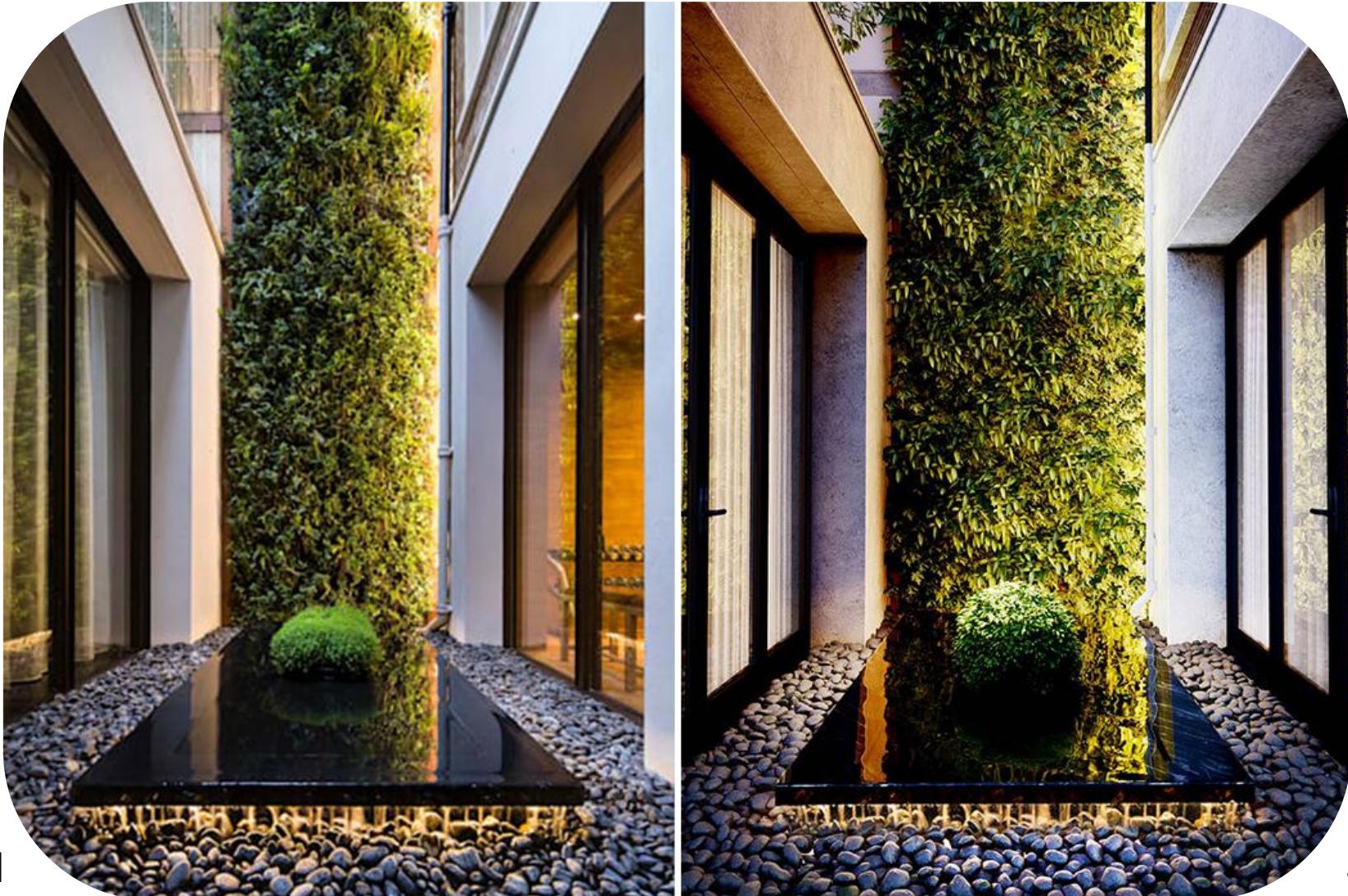


World → Image

Real



Can We Get “Free” Data?



World → Image

Real

Image → World

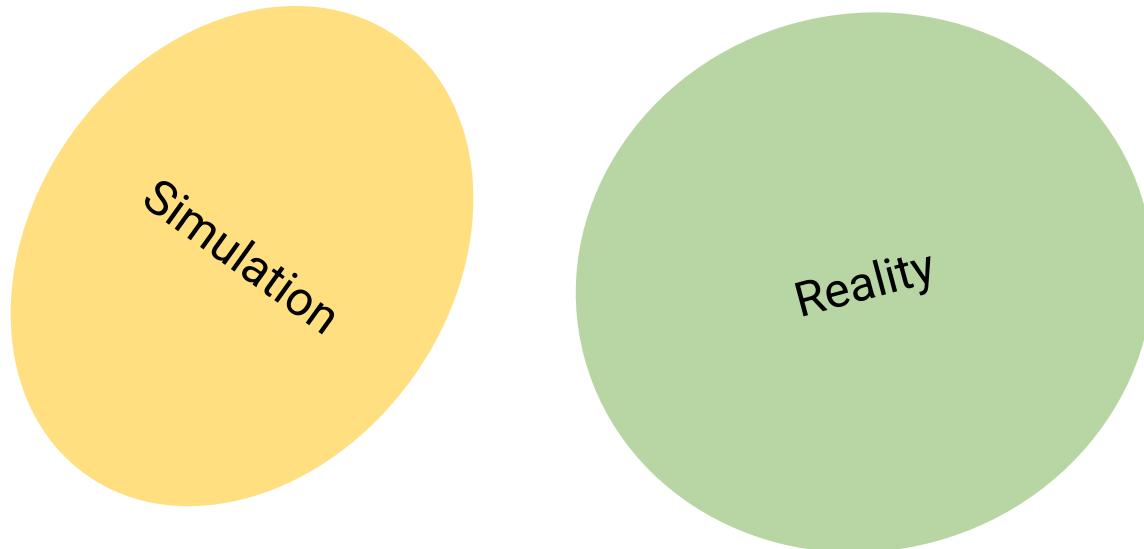
Simulation

Tricks To Increase Data

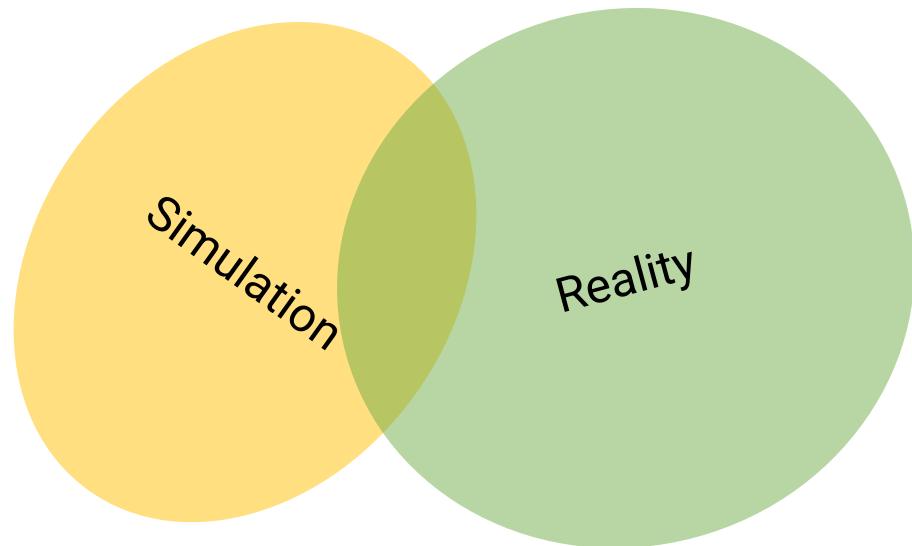
Data Augmentation



Sim2Real Gap



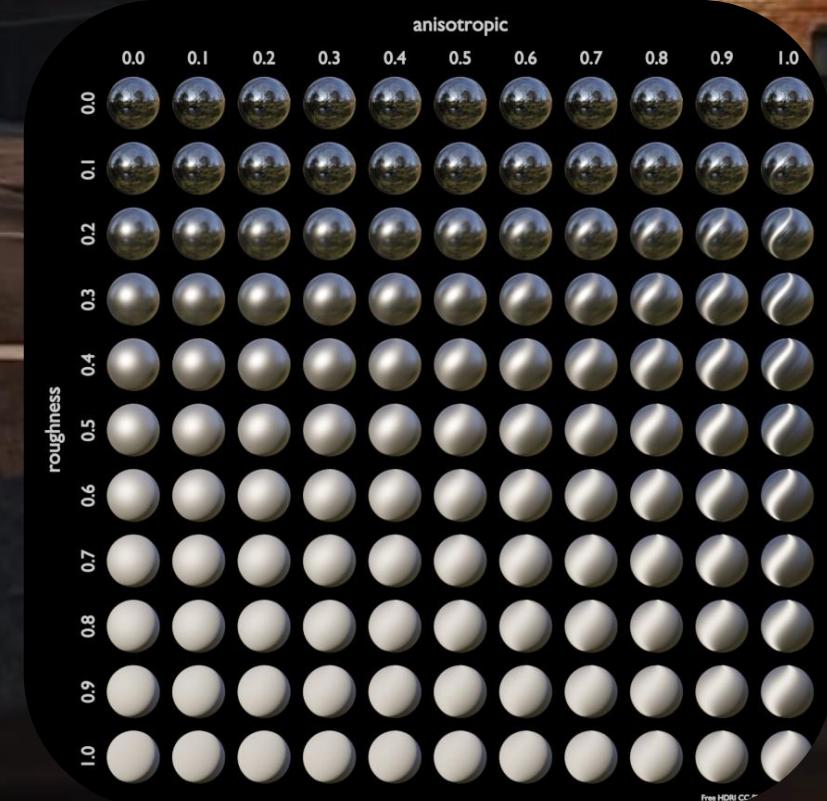
Sim2Real Gap



Simple Way To Fix Sim2Real Gap?

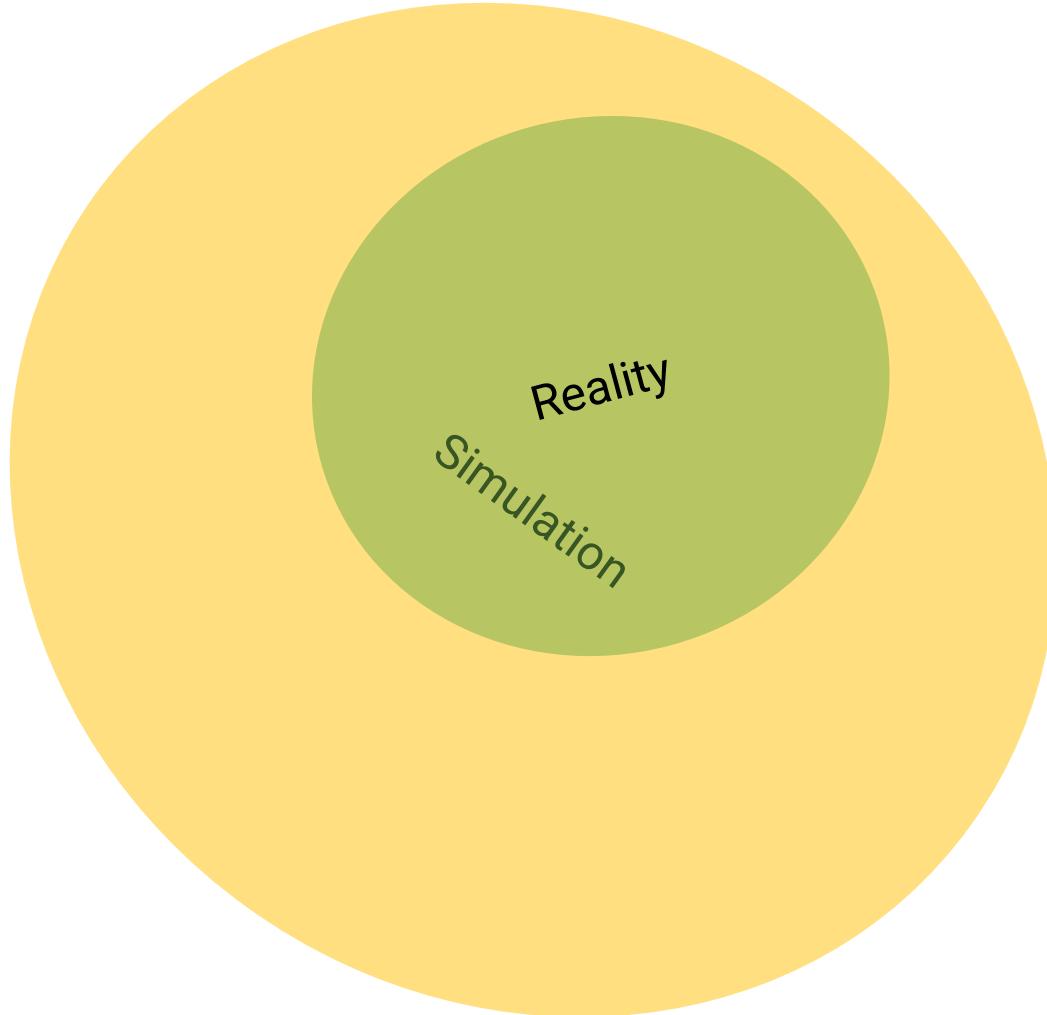
Photorealism

- Better material properties
 - Reflection
 - Refraction
 - Transmission
- Higher quality models
 - More detailed texture = more accurate
 - More variations in model = better variability!
- Model tiny to tiny details
 - Dust
 - Scratches
 - Lens flares
 - Motion blur
 - Diffraction



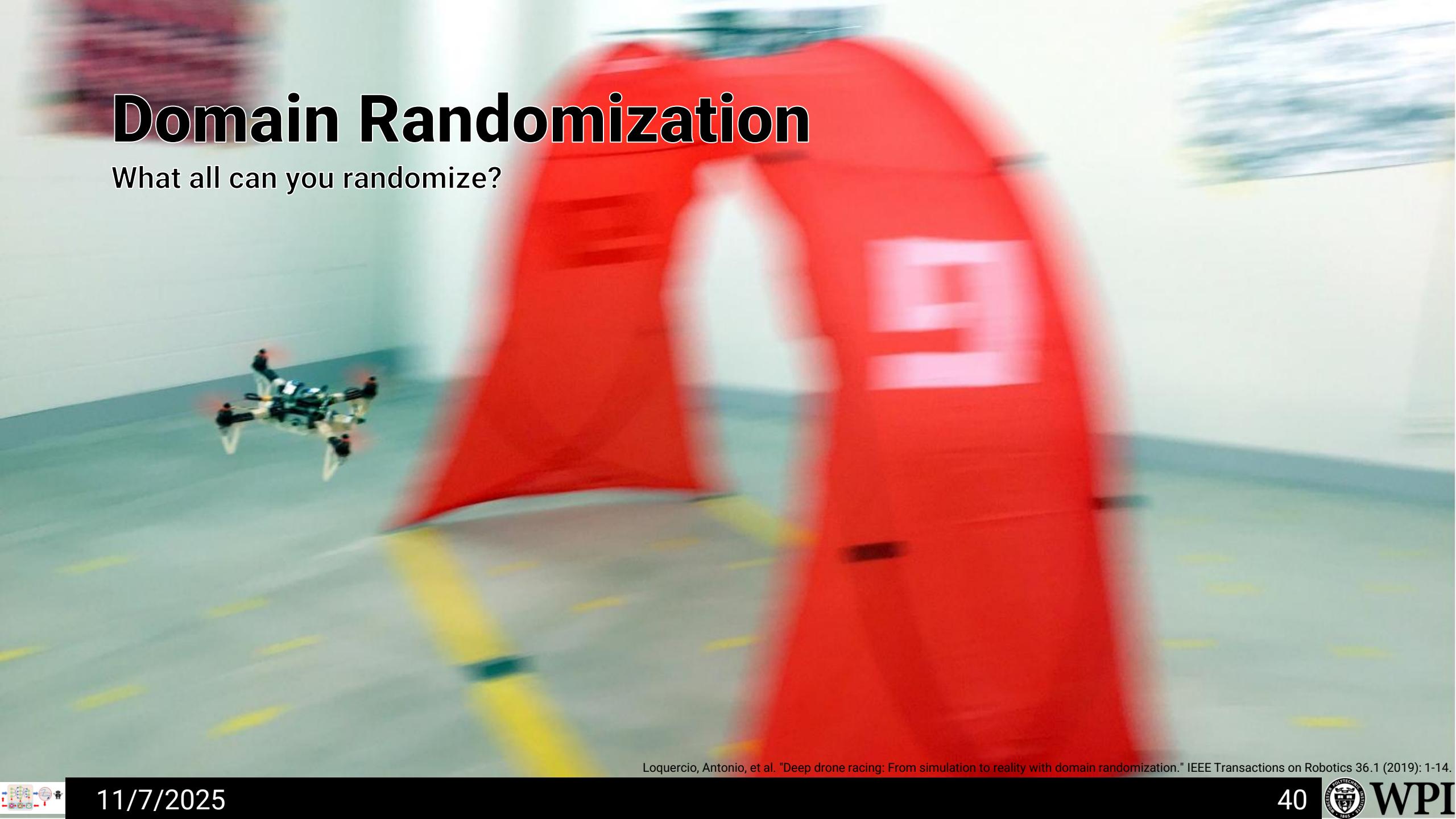
Sim2Real Gap

Domain Randomization



Domain Randomization

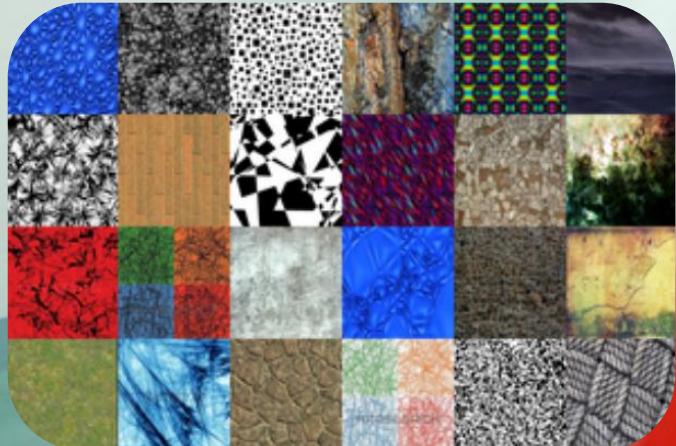
What all can you randomize?



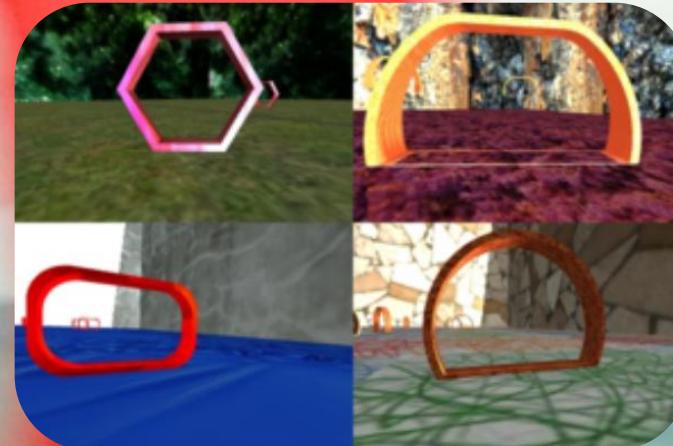
Loquercio, Antonio, et al. "Deep drone racing: From simulation to reality with domain randomization." *IEEE Transactions on Robotics* 36.1 (2019): 1-14.

Domain Randomization

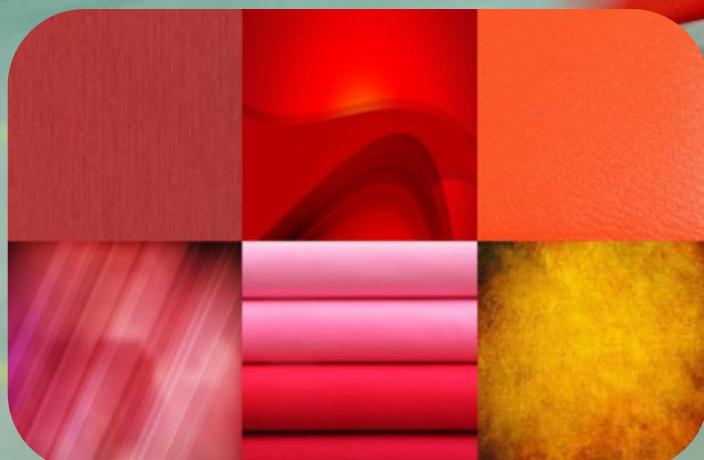
What all can you randomize?



Background textures!



Gate shapes!



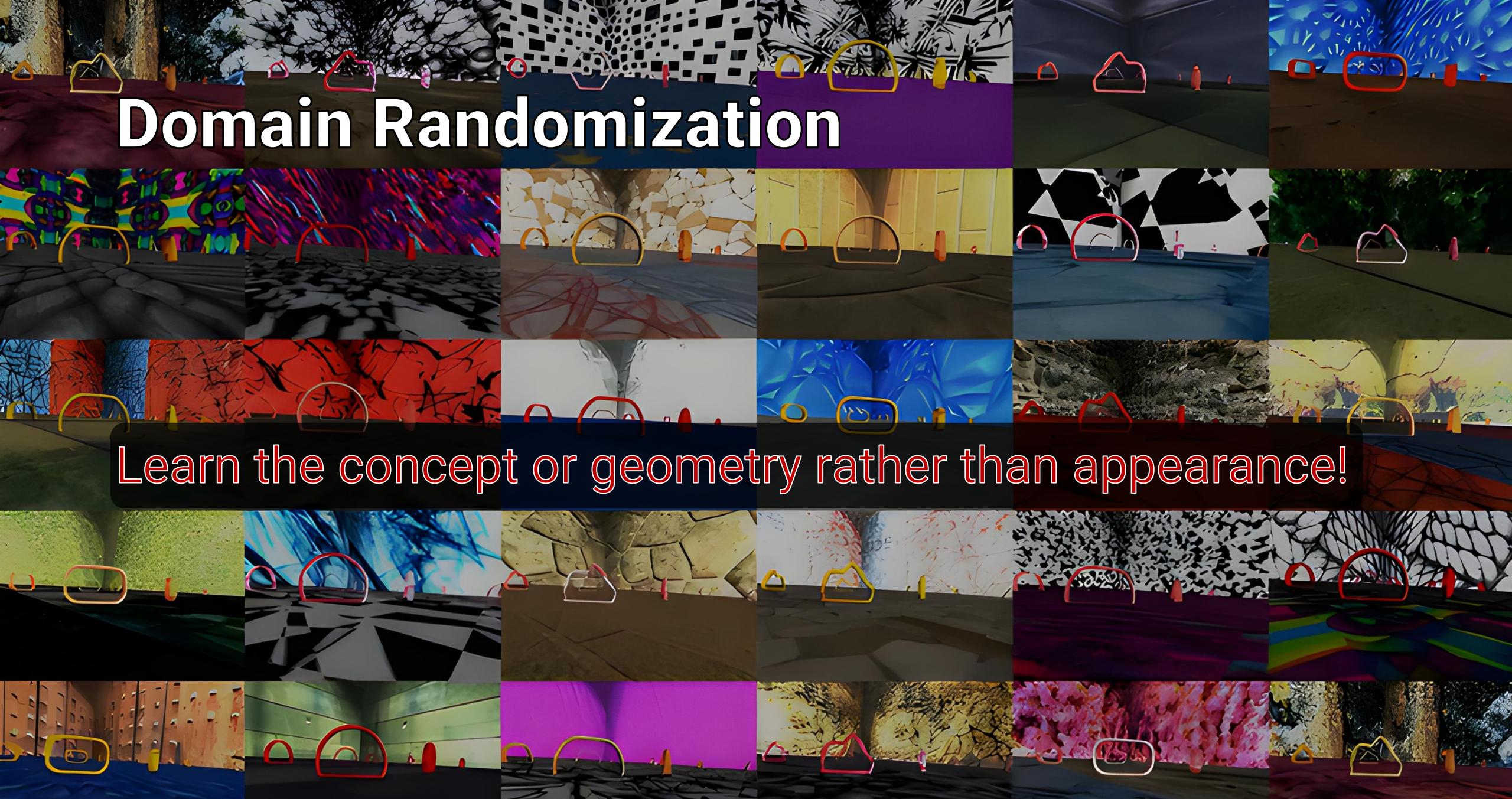
Gate textures!

What more?

Loquercio, Antonio, et al. "Deep drone racing: From simulation to reality with domain randomization." IEEE Transactions on Robotics 36.1 (2019): 1-14.

Domain Randomization

Learn the concept or geometry rather than appearance!



Disclaimer

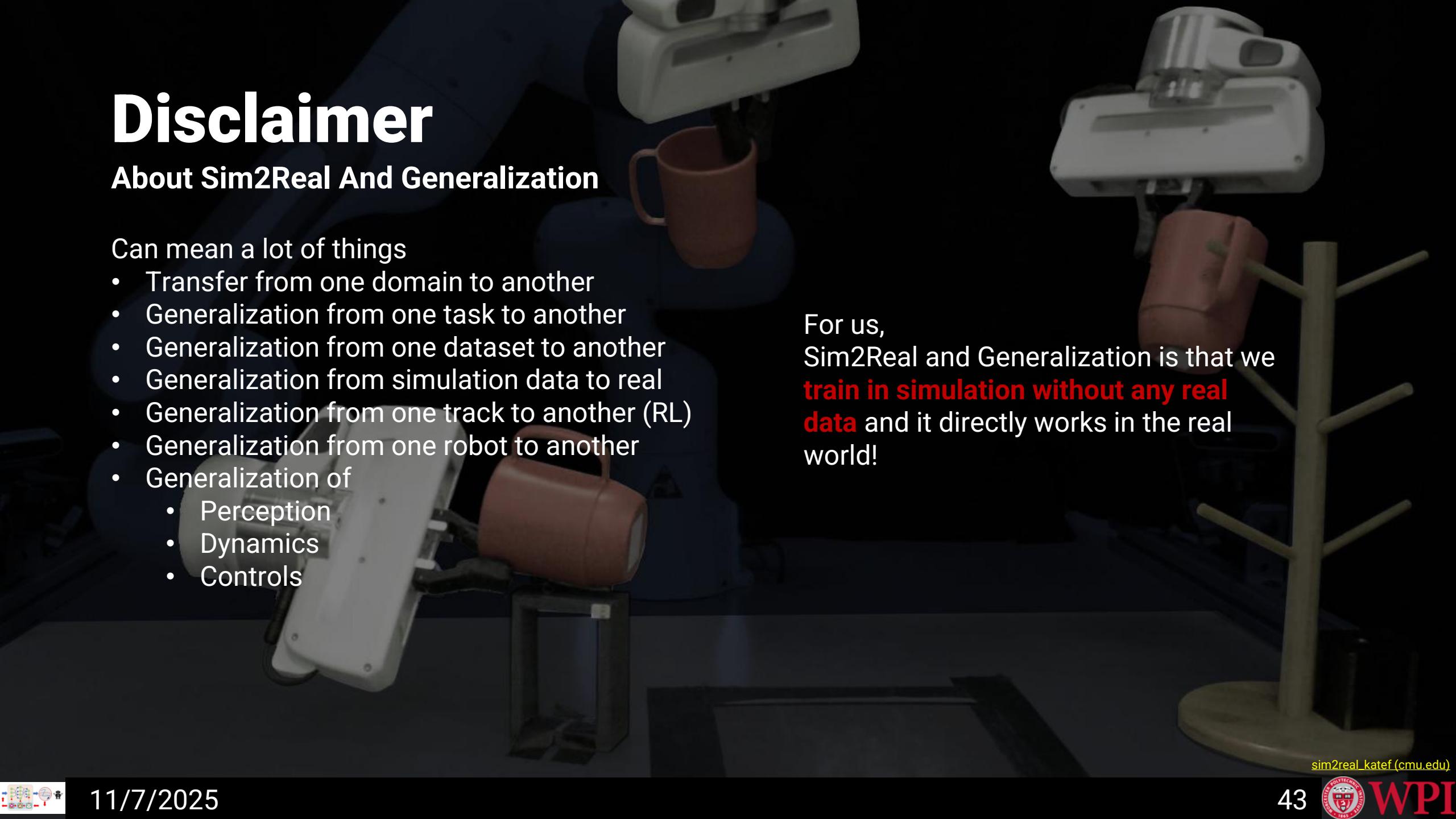
About Sim2Real And Generalization

Can mean a lot of things

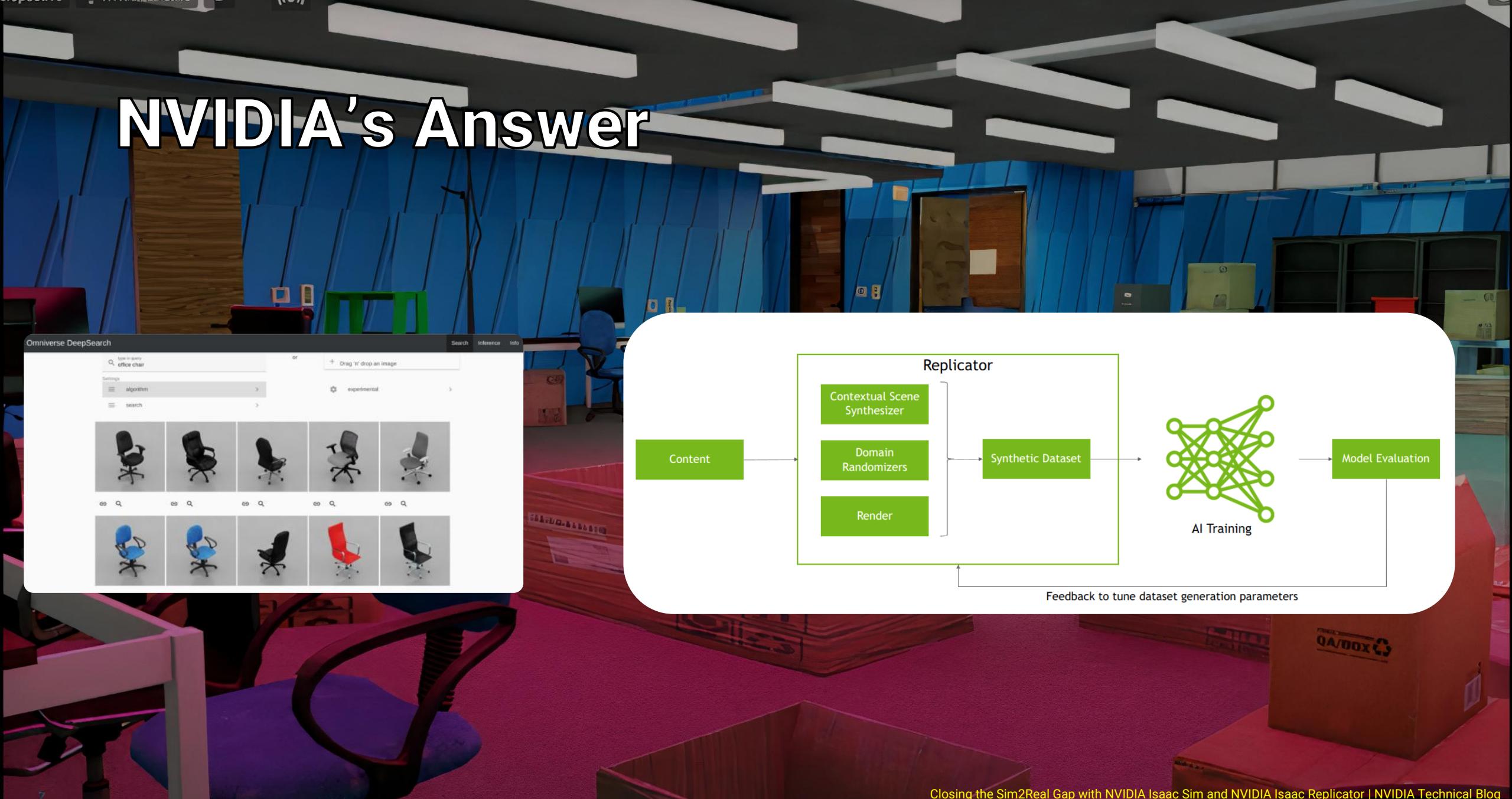
- Transfer from one domain to another
- Generalization from one task to another
- Generalization from one dataset to another
- Generalization from simulation data to real
- Generalization from one track to another (RL)
- Generalization from one robot to another
- Generalization of
 - Perception
 - Dynamics
 - Controls

For us,

Sim2Real and Generalization is that we
**train in simulation without any real
data** and it directly works in the real
world!



NVIDIA's Answer



[Closing the Sim2Real Gap with NVIDIA Isaac Sim and NVIDIA Isaac Replicator | NVIDIA Technical Blog](#)



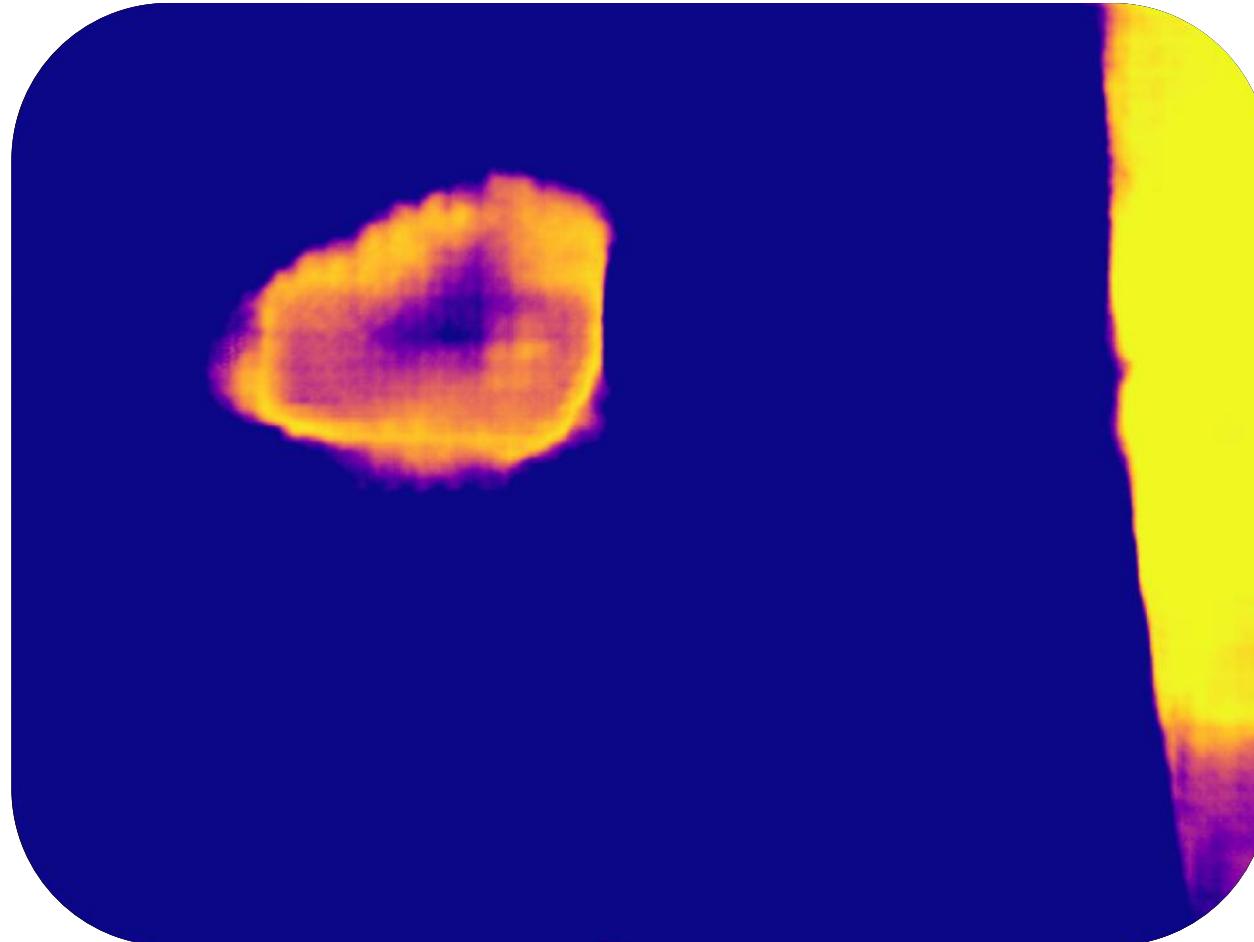
Can This Be Made Simpler?

Use Agnostic Representations!



Can This Be Made Simpler?

Use Agnostic Representations!



Occlusion or Flow Over Monocular Depth





RSS, 2021

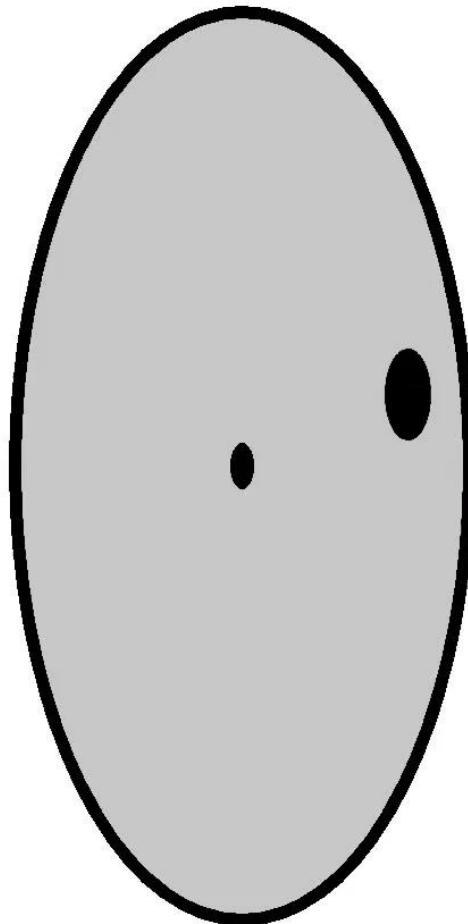
Can This Be Made Simpler?

Use Agnostic Sensors!

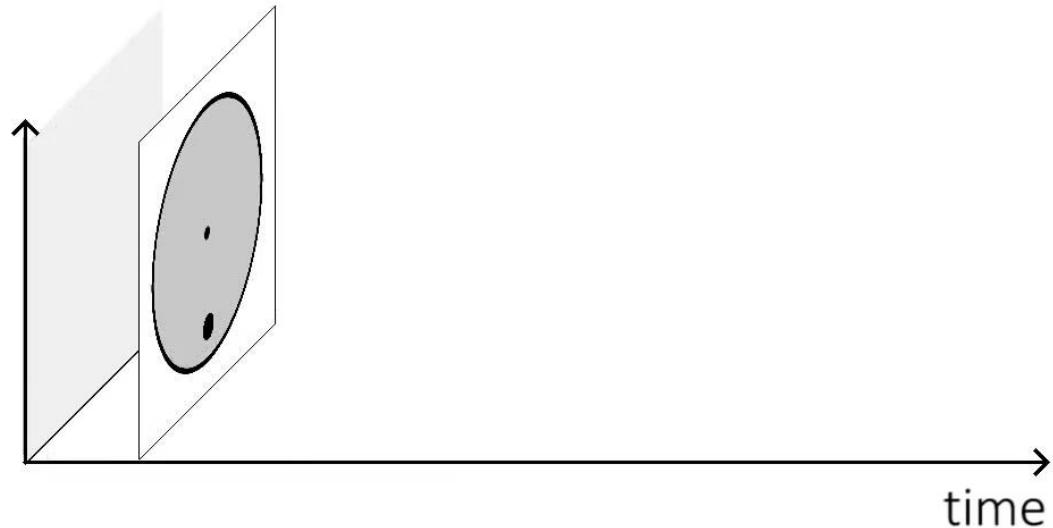
In collaboration with
TU Delft MAVLab™



Event Cameras



standard
camera
output:









IEEE ICRA, 2020

EV DodgeNet

Deep Dynamic Obstacle Dodging with Event Cameras

Applicable to Quadrotor sizes
of 250 mm and below

Featured in
Mashable Futurism



11/7/2025

In collaboration with

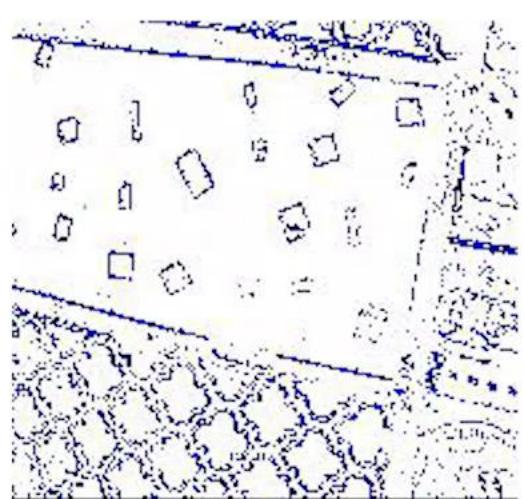
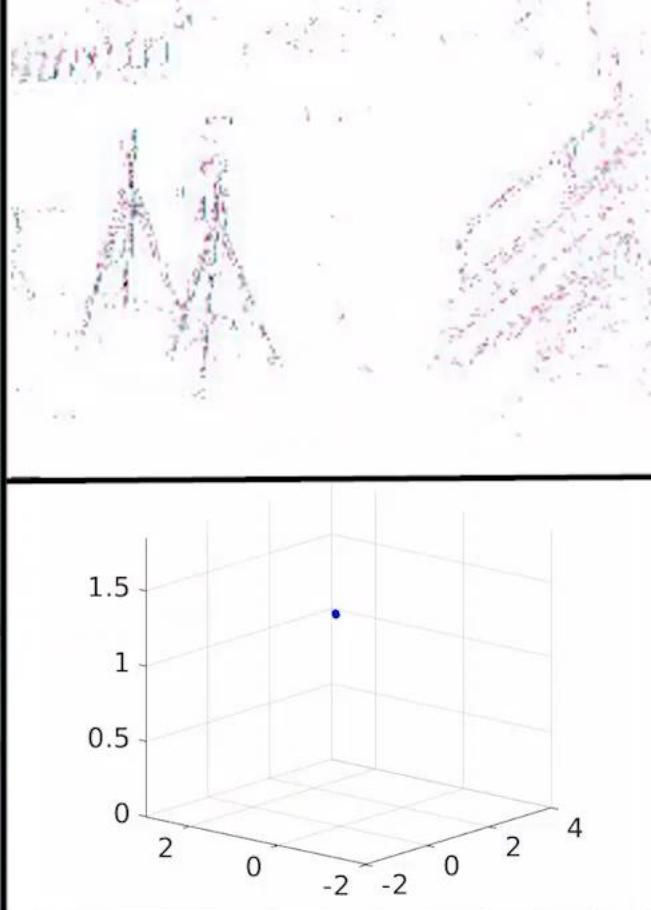


ETH zürich



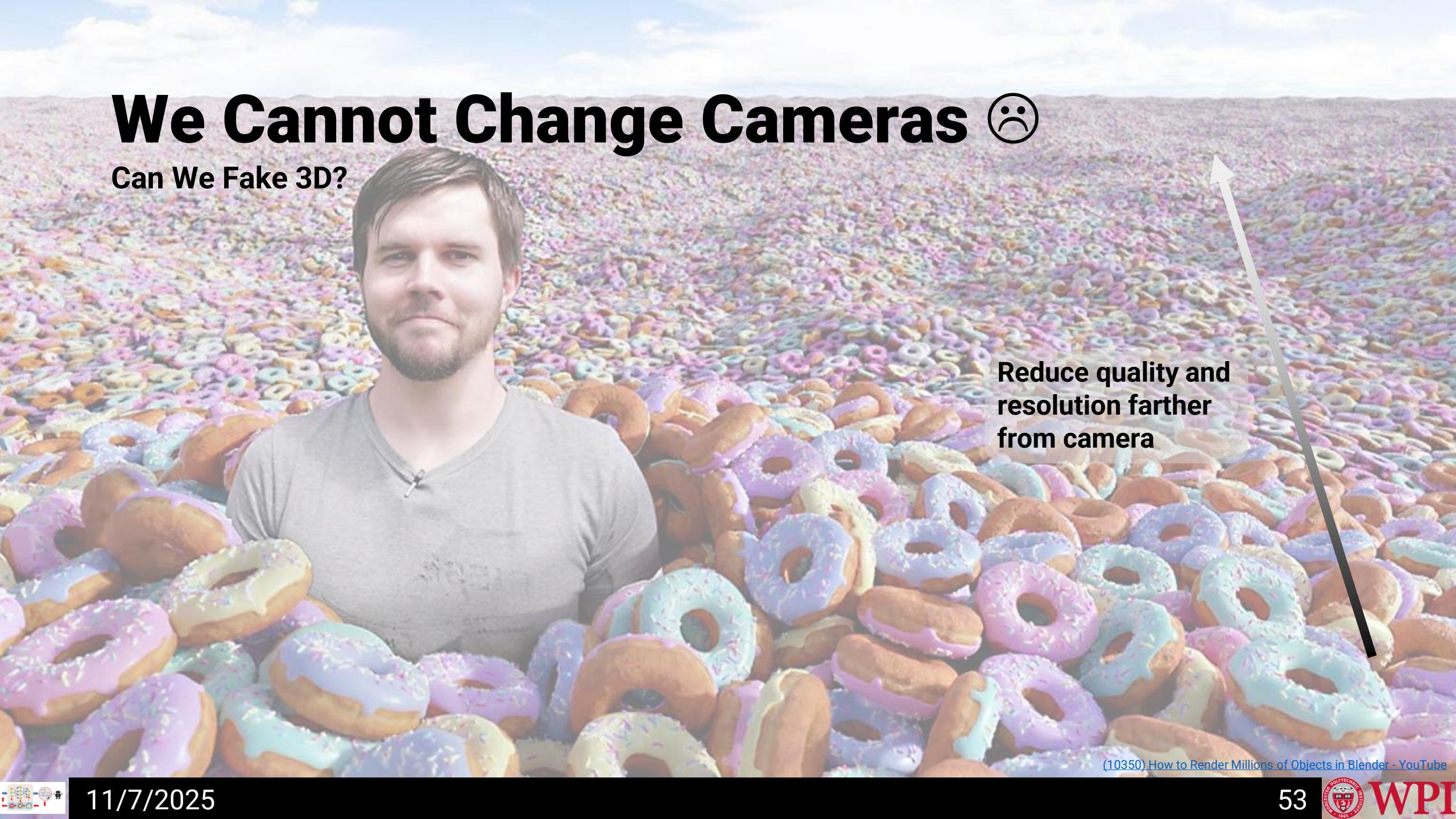
51

WPI



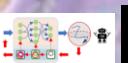
We Cannot Change Cameras 😞

Can We Fake 3D?

A photograph of a man with a beard and short hair standing in a vast field of colorful donuts. He is wearing a light gray t-shirt. The donuts are numerous, filling the frame, and are covered in various frostings like pink, blue, and white. A large white arrow points from the text "Reduce quality and resolution farther from camera" towards the horizon where the donuts meet the sky.

Reduce quality and
resolution farther
from camera

(10350) How to Render Millions of Objects in Blender - YouTube



We Cannot Change Cameras 😞

Can We Fake 3D?



We Cannot Change Cameras 😞

Can We Fake 3D?



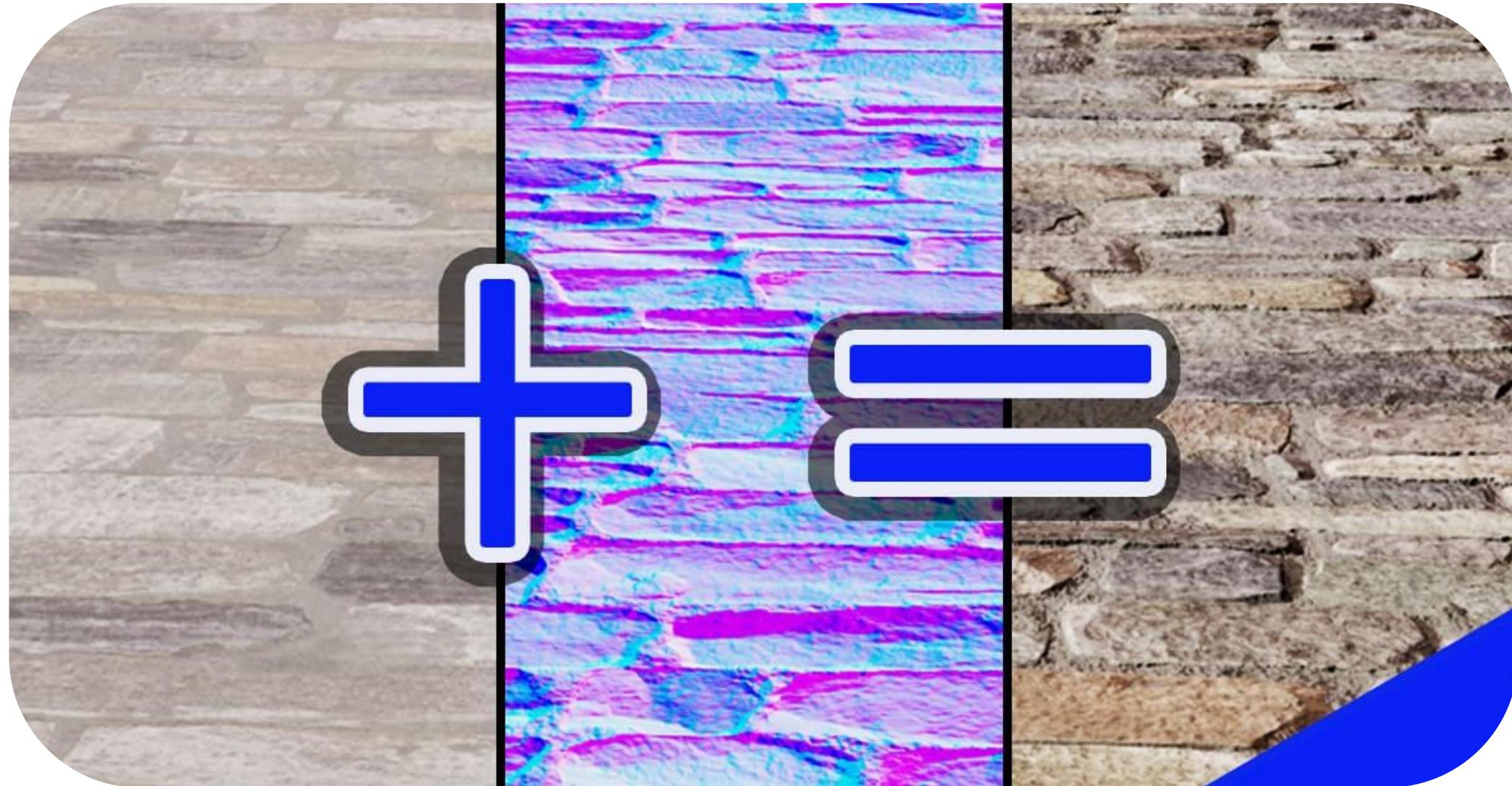
We Cannot Change Cameras 😞

Can We Fake 3D?



What Did I Do?

Normal Map

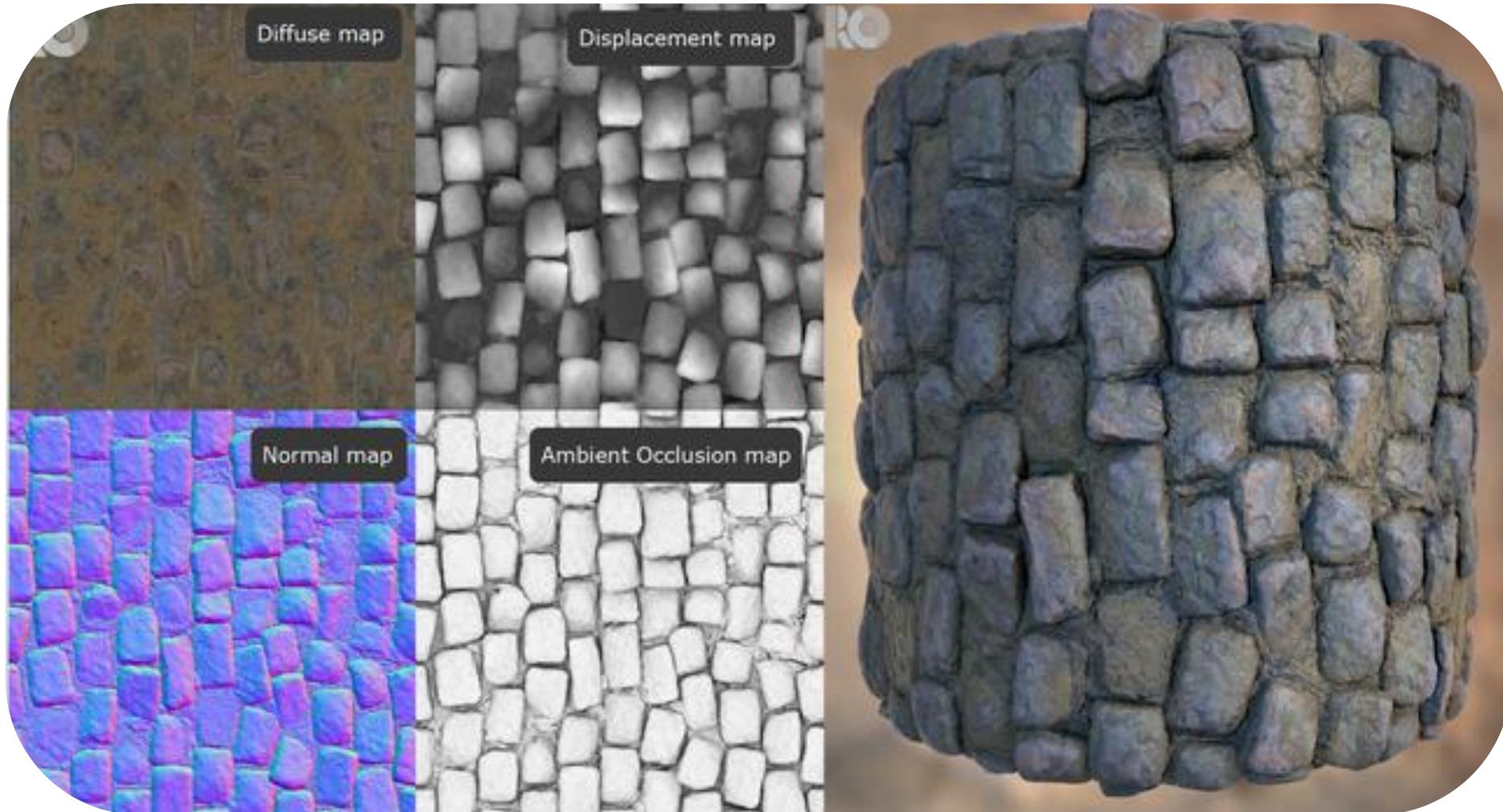


Normal Maps



Best Of All Worlds

Displacement + Normal Maps



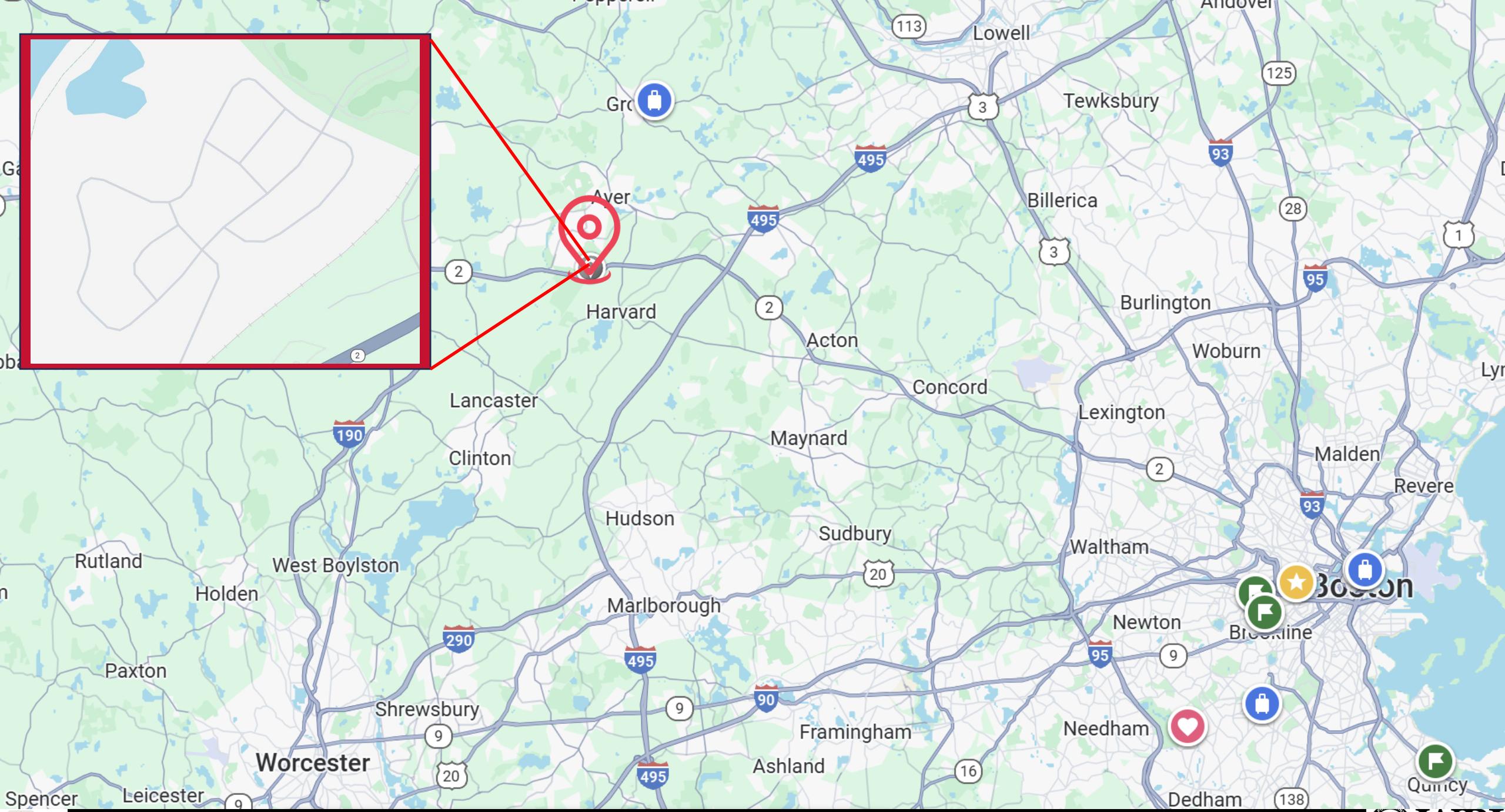
What If You Have Little Real Data?

Low Data Regime



Real to Sim to Real







Data Collection (for model generalization)



Data Augmentation



Original data



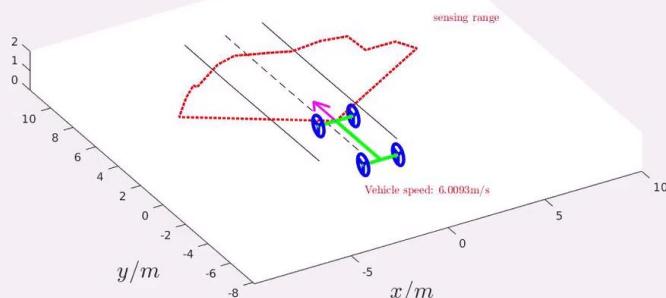
Augmented Reality



- 49 traces for training
- 1 trace for testing
- 40 randomly generated obstacles for each trace

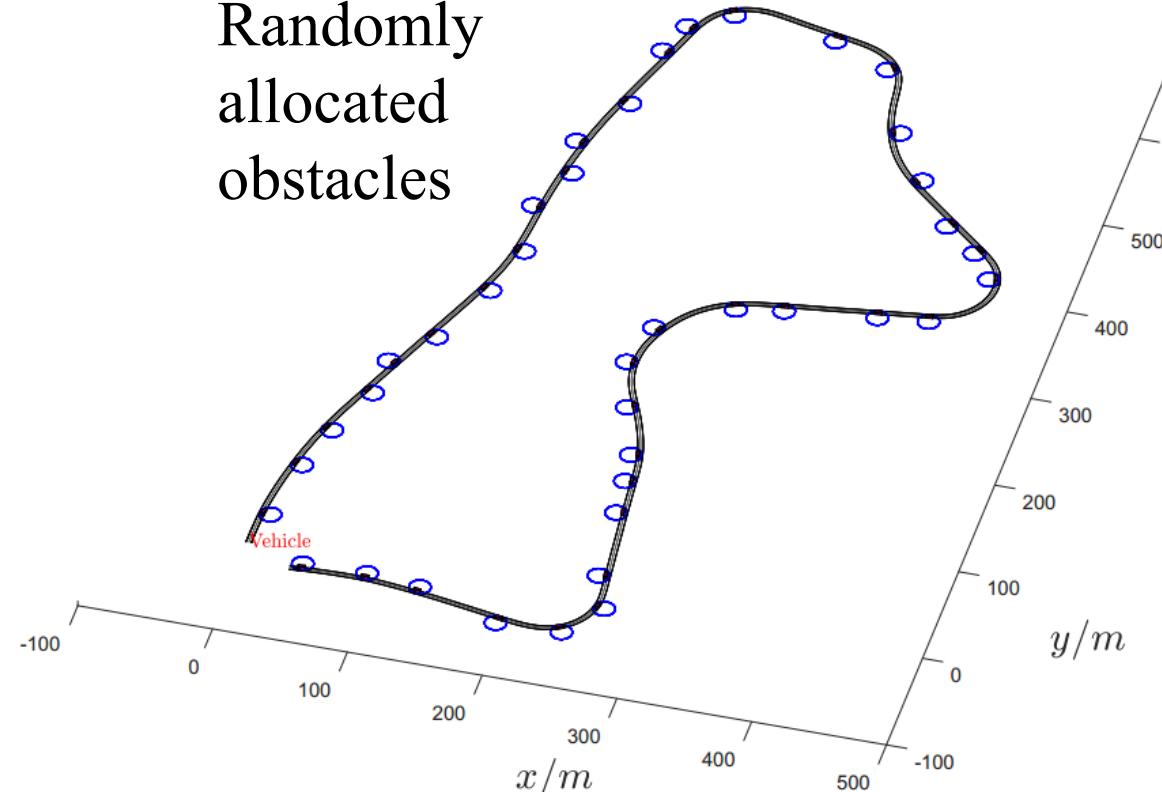
Ground truth controls

System states and obstacle locations are all known



Nominal controller: NMPC (tractable offline)

Randomly allocated obstacles



Front-view images
labelled with controls





What If Labelling Is Hard?



IEEE IROS, 2023

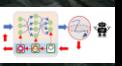
Olive the Above

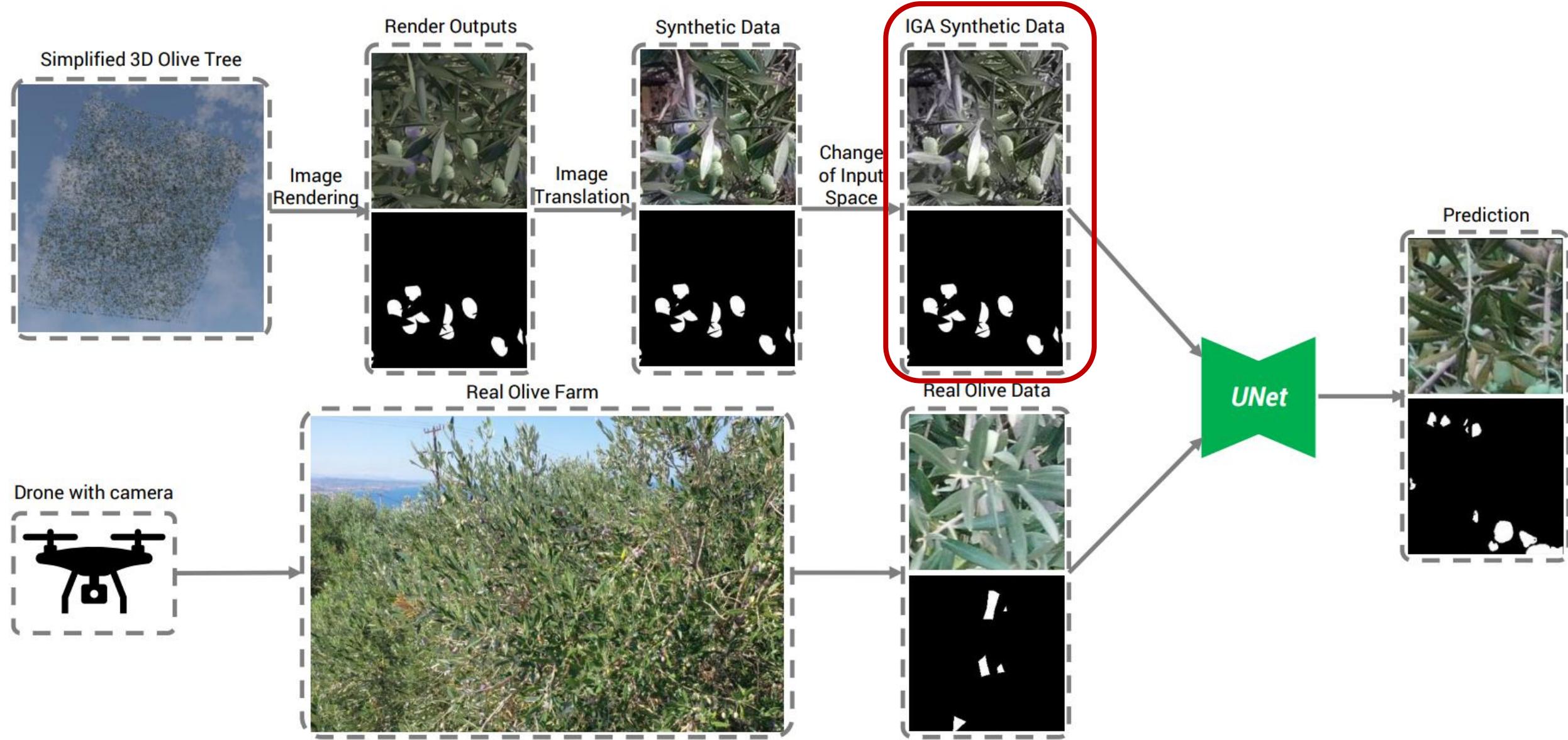
Detecting Olives with Synthetic or Real Data?

In collaboration with



PRG







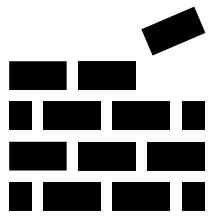
Generate Branches/Trees







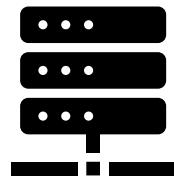
There is a need for



Creating



Scalable



Data

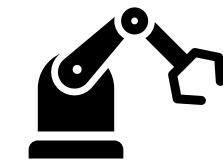
Using
Open-source



Maps



and
Models



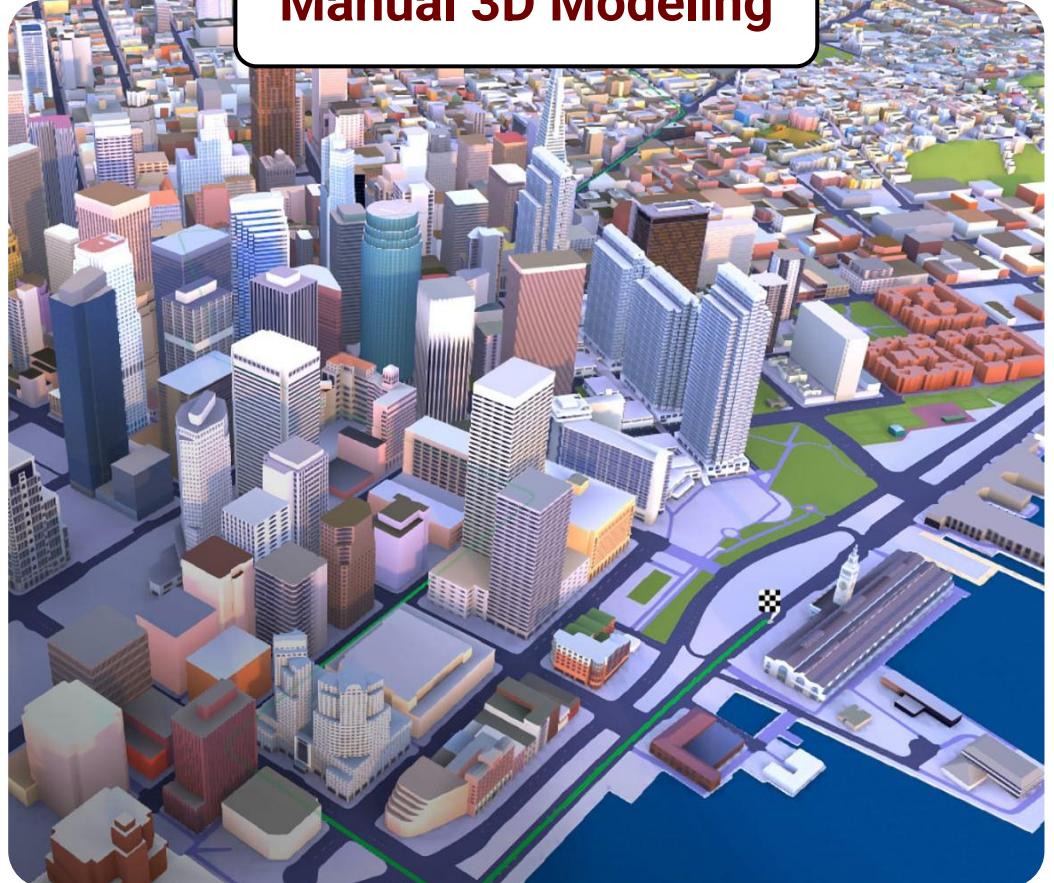
Automatically

Generative Data to Minimize

Manual Annotations



Manual 3D Modeling



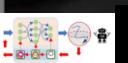


IEEE ICRA, 2023

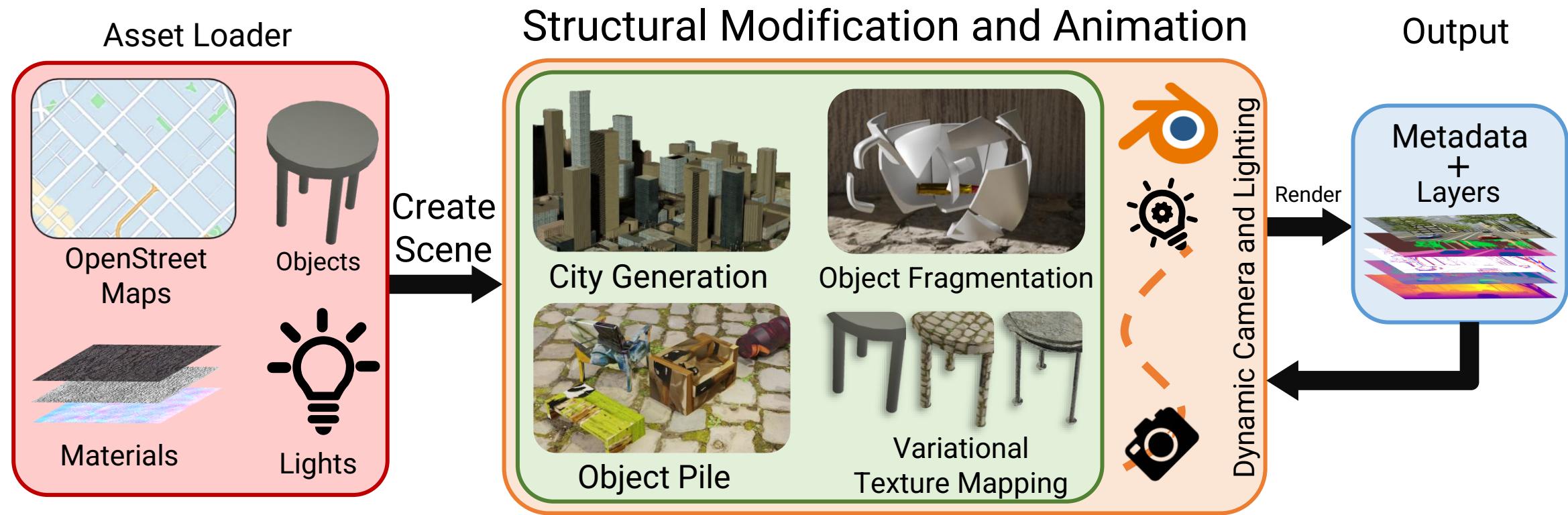
WorldGen

A Large Scale Generative Simulator

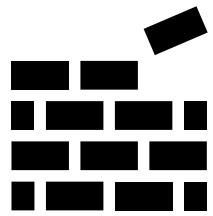
In collaboration with



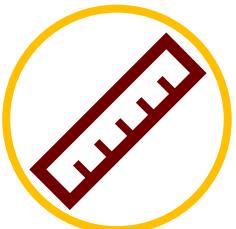
WorldGen



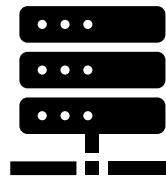
WorldGen



Creating



Scalable



Data

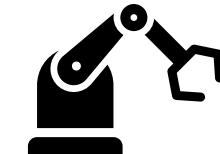
Using
Open-source



Maps



Models



Automatically

WorldGen: A Large Scale Generative Simulator

Chahat Deep Singh, Riya Kumari, Cornelia Fermüller, Nitin J. Sanket, Yiannis Aloimonos



prg.cs.umd.edu



WPI

WorldGen: A Large Scale Generative Simulator

Chahat Deep Singh, Riya Kumari, Cornelia Fermüller, Nitin J. Sanket, Yiannis Aloimonos

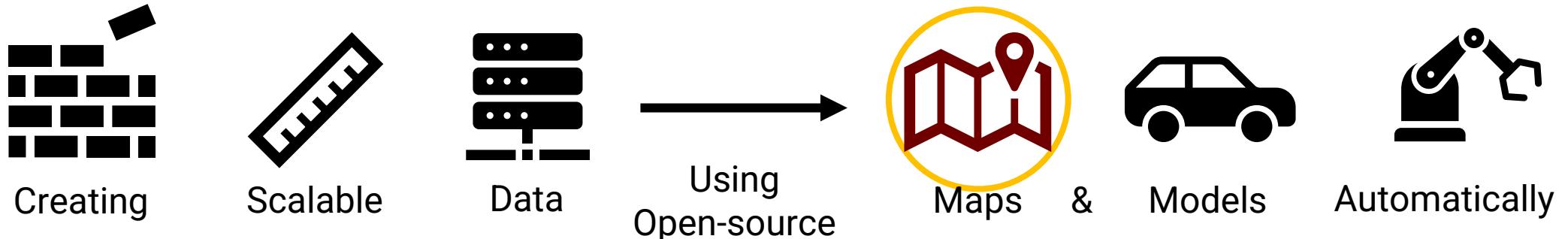


prg.cs.umd.edu



WPI

WorldGen

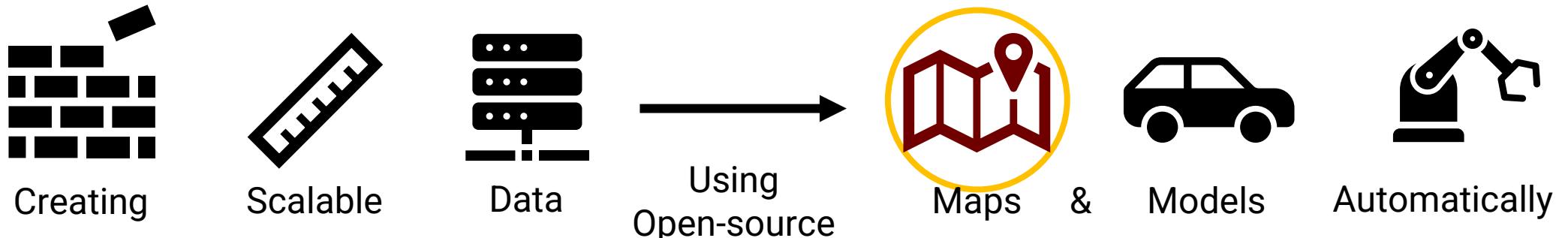


WorldGen: A Large Scale Generative Simulator

Chahat Deep Singh, Riya Kumari, Cornelia Fermüller, Nitin J. Sanket, Yiannis Aloimonos



WorldGen



WorldGen: A Large Scale Generative Simulator

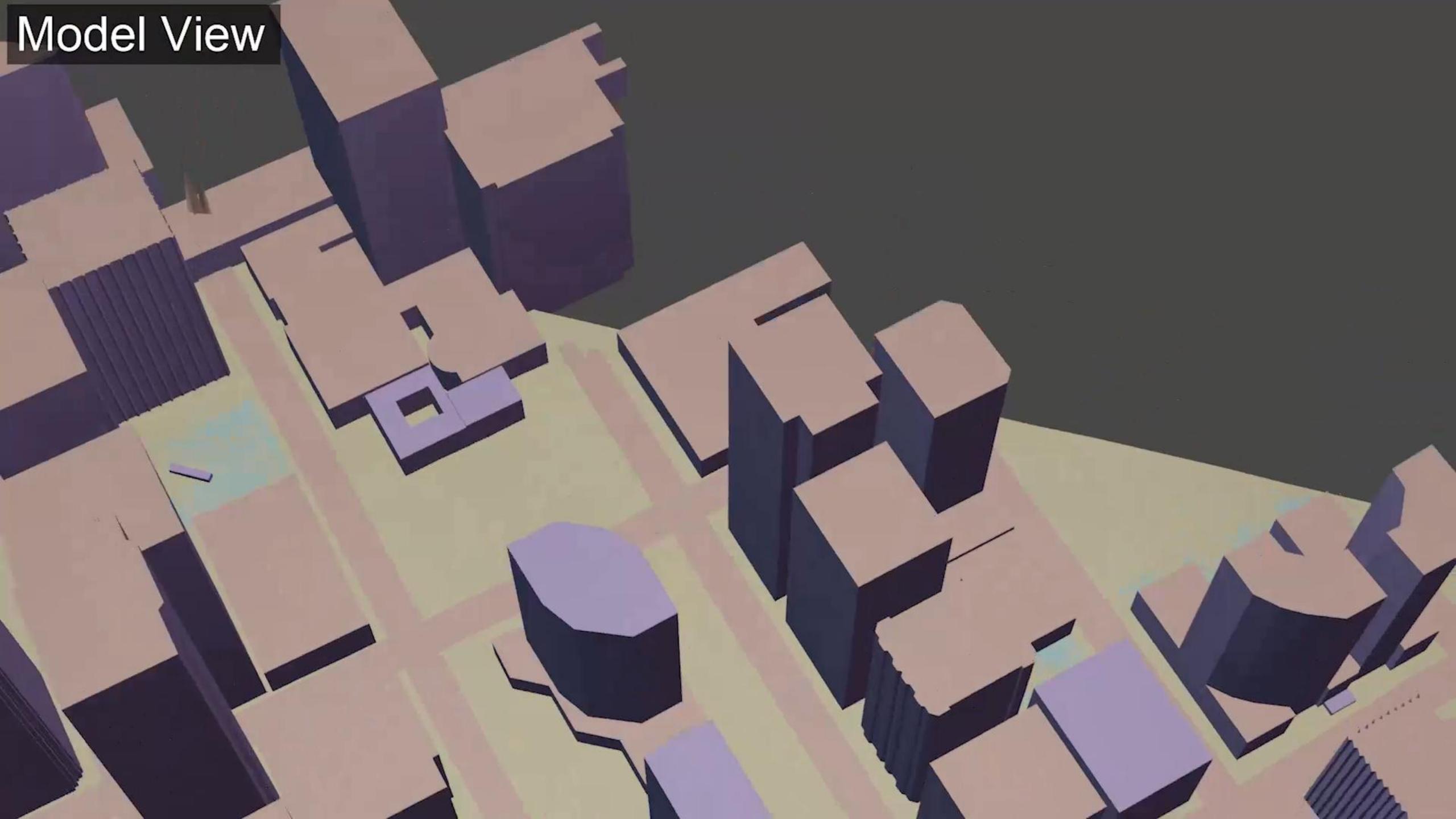
Chahat Deep Singh, Riya Kumari, Cornelia Fermüller, Nitin J. Sanket, Yiannis Aloimonos



Multiple Scenarios



Model View



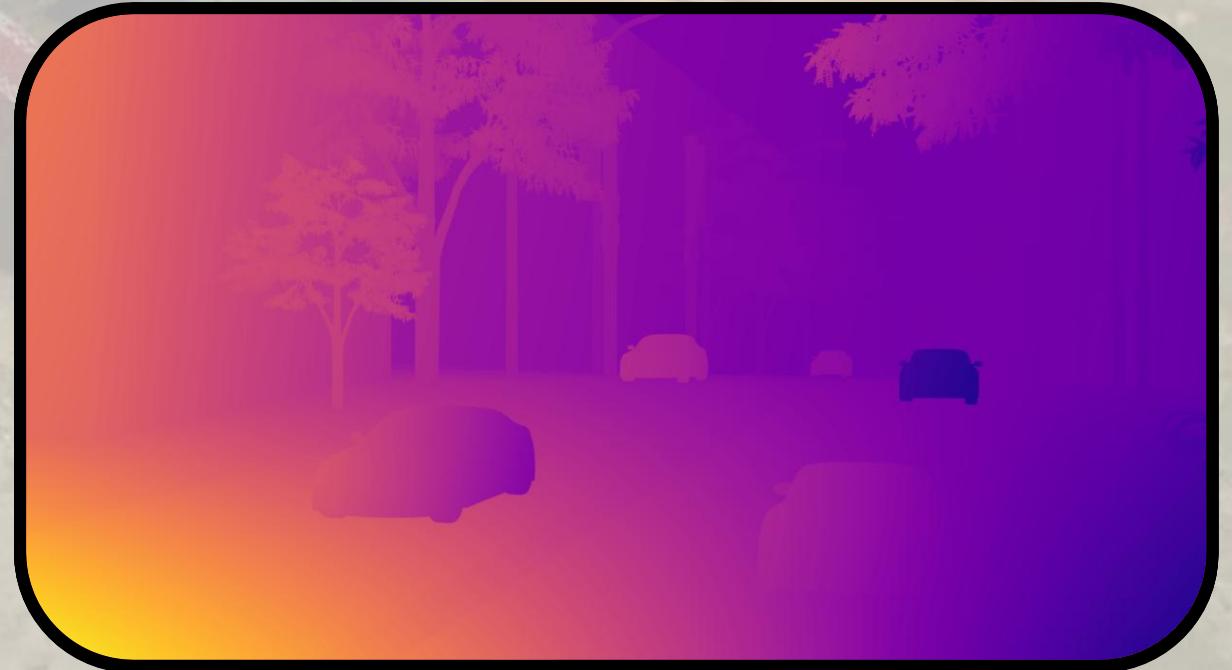
With Camera Photorealism



Data is Useful for Generalization

- Creating Digital Twins
- Depth maps
- Semantics
- Optical Flow

... and much more





How Do We Use These On Robots?

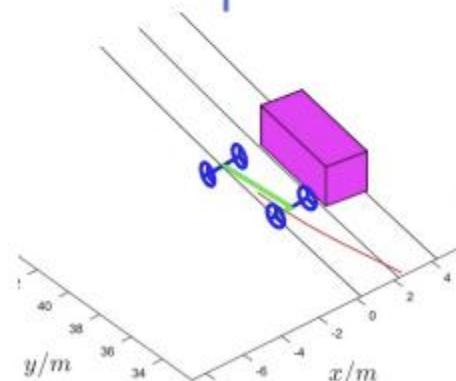
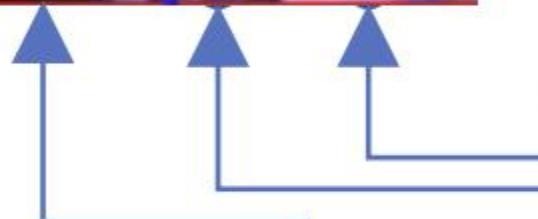
An End-to-End Learning System with a BarrierNet

Front view image



Neural Nets +
Differentiable CBFs
with safety guarantees

Acceleration
Steering rate



Simulation

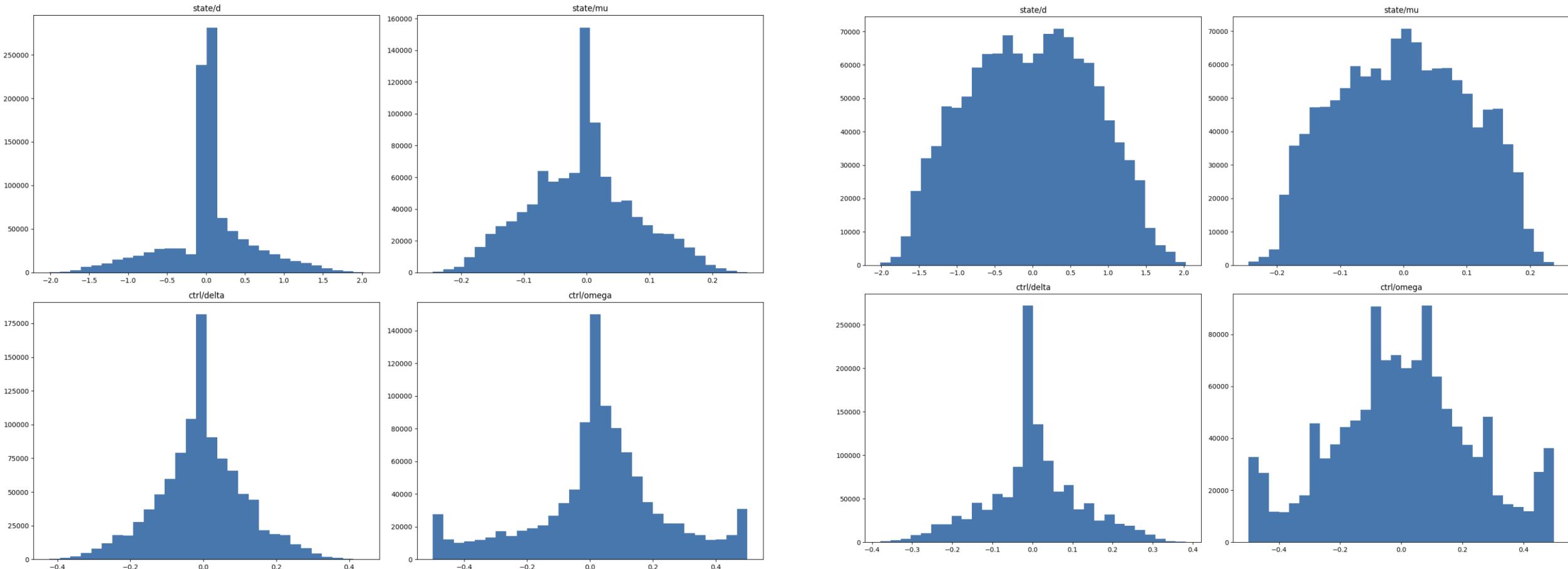


Sim-to-Real

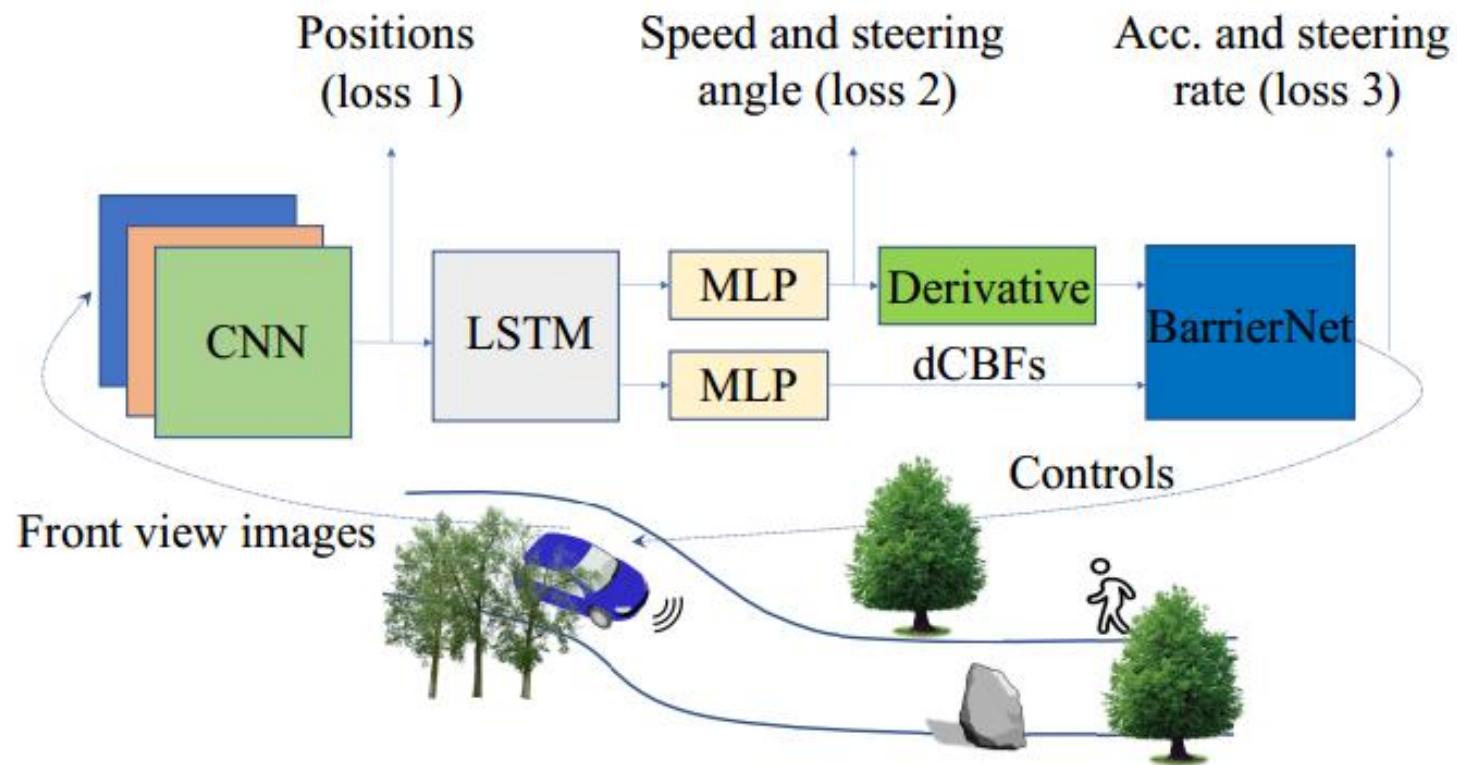


Real

Check Data Distribution



Proper Model and Loss Design



Simulation Testing

without dCBF

76dd2c: camera_front



Top-down View



with dCBF

fc36ce: camera_front

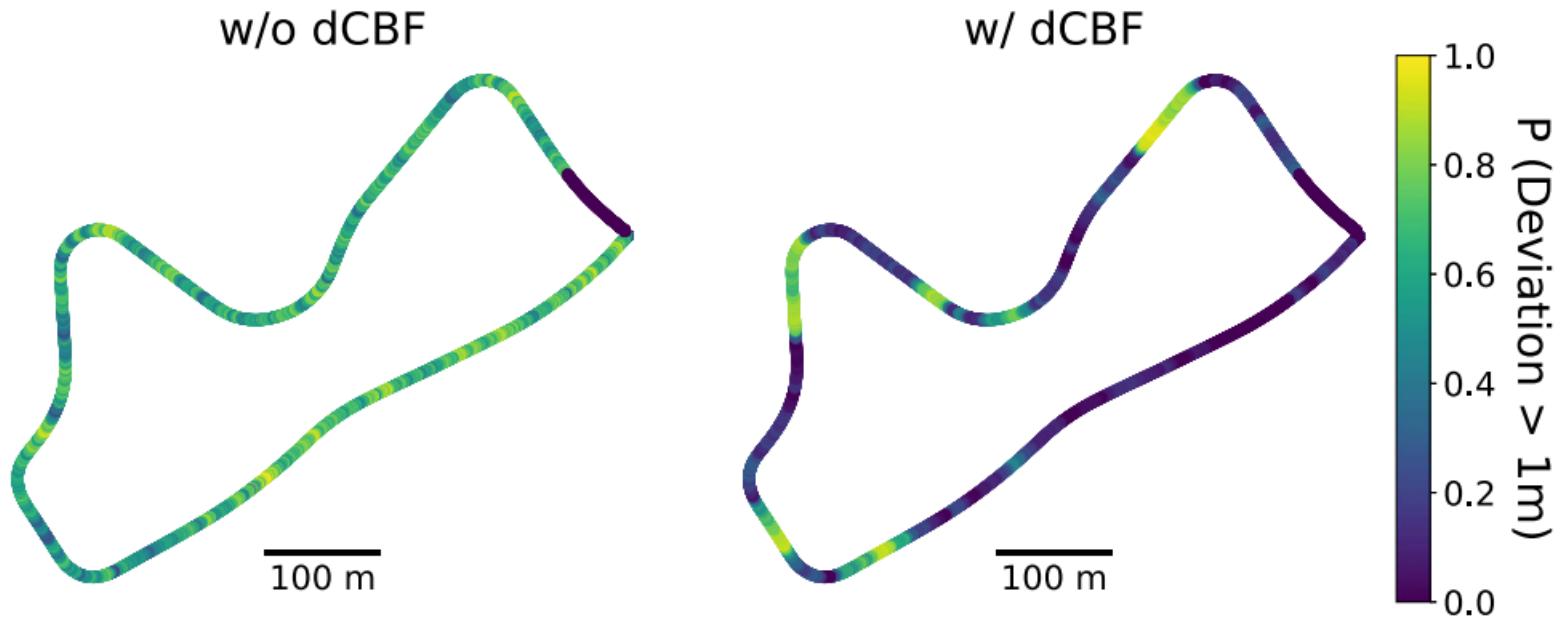


Top-down View



Simulation Testing

Testing in **sim-to-real** environments: lane following



Real Car Experiments



Scenario Gen

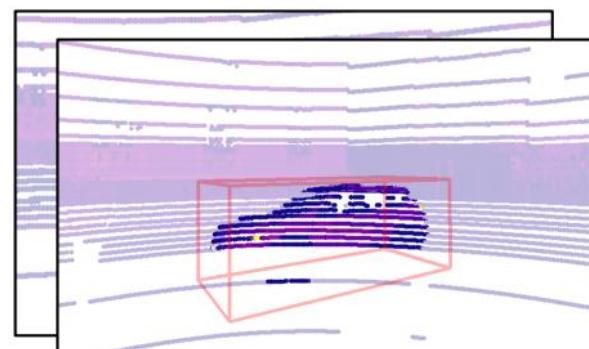
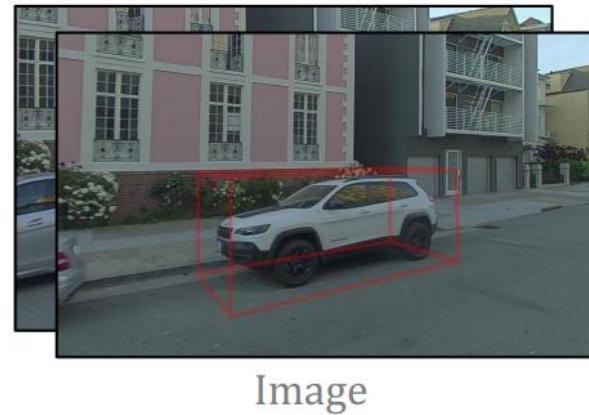


Simulate Sensors

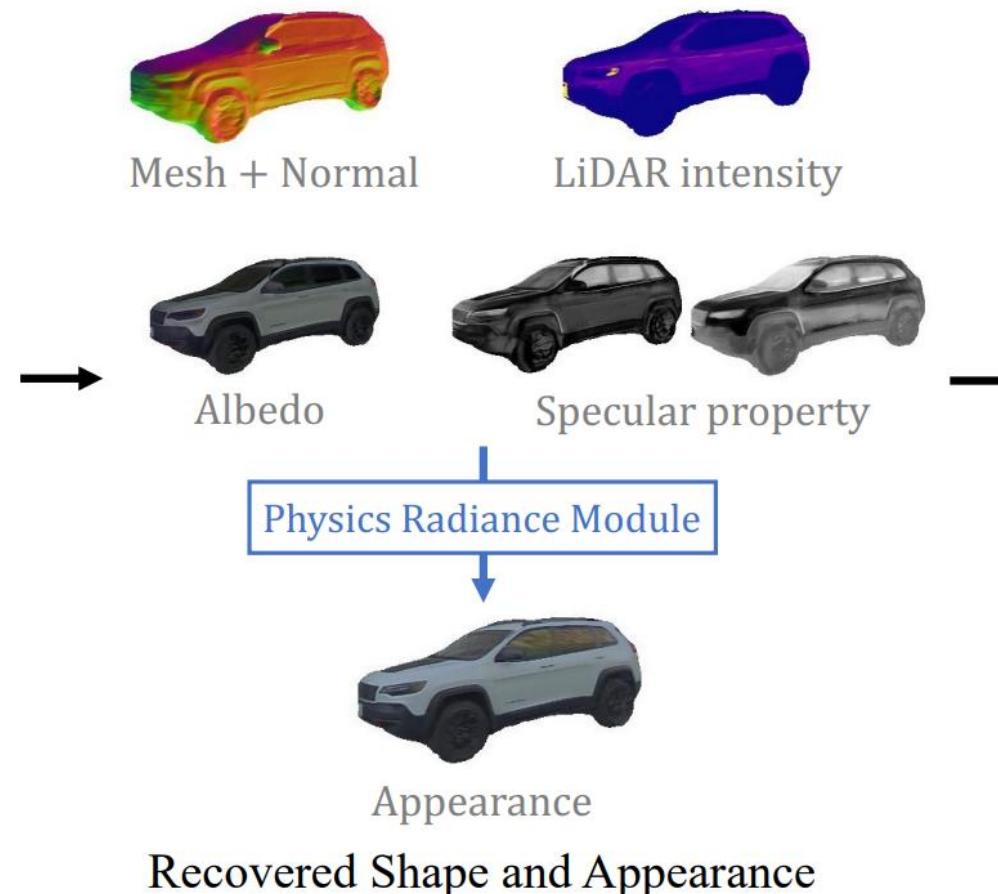
Uber ATG 



Raquel Urtasun



Multi-sensor Captures



Yang, Ze, et al. "Reconstructing objects in-the-wild for realistic sensor simulation." ICRA, 2023.

Inject New Actors



Albedo



Specular Albedo



Specular Shininess



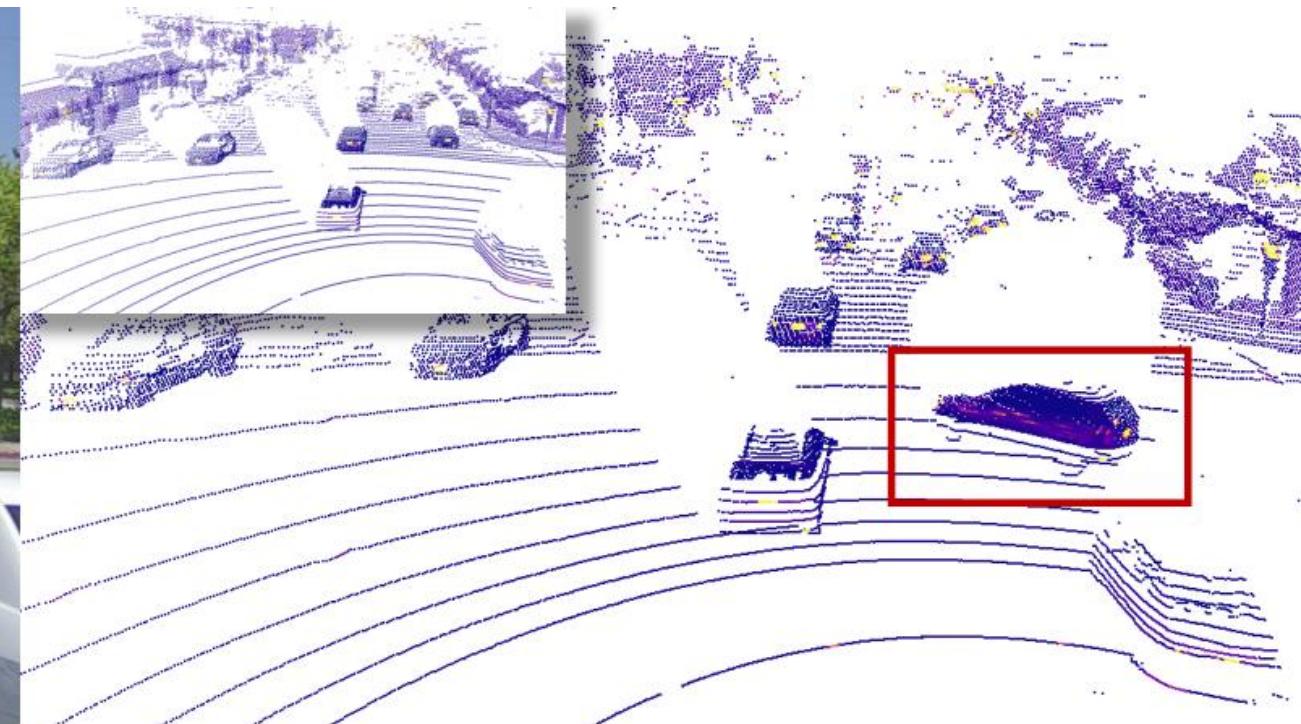
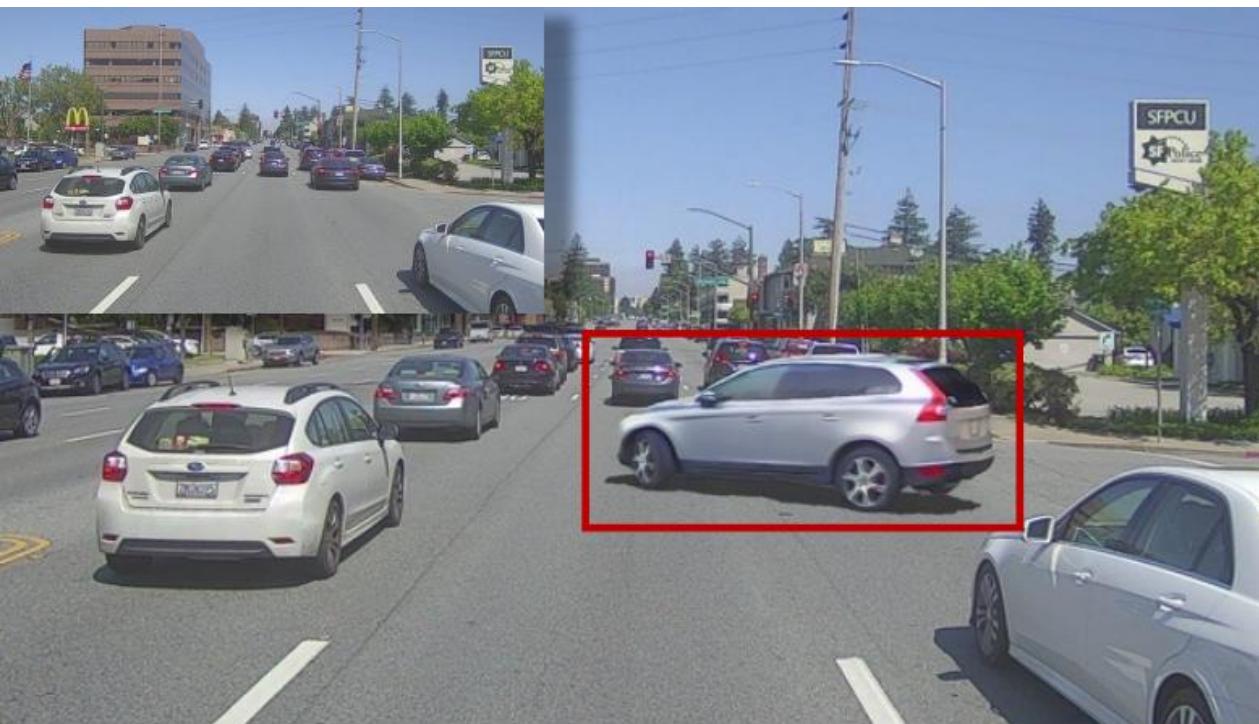
Mesh



Shaded Mesh



Volume Rendering



Manipulate Actors

UniSim



Server: 60 FPS
Client: 62 FPS

Vehicle: Lincoln Mkz 2020
Map: Tel Aviv
Simulation time: 00:50

Speed: 0 km/h
Compass: 0°
Accel: 0 m/s²
Gyros: 0°/s
Location: 37.5188, 12.4701
GNSS: 0.0000, 0.0000
Height: 0.0000

Throttle: 0.0000
Steering: 0.0000
Brake: 0.0000
Reverse: 0.0000
Handbrake: 0.0000
Manual gear: 0.0000
Gear: 0.0000

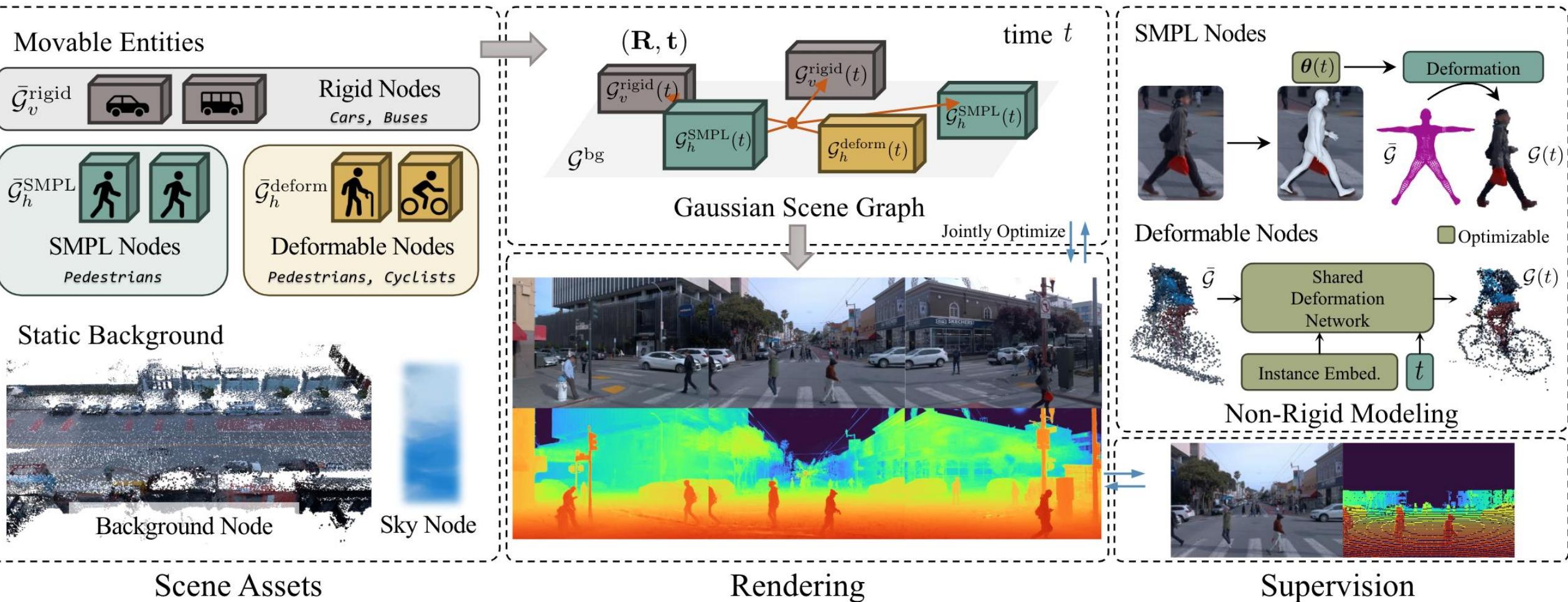
Collision:

Number of vehicles: 1

MATLAB RoadRunner



OmniRe



Chen, Ziyu, Jiawei Yang, Jiahui Huang, Riccardo de Lutio, Janick Martinez Esturo, Boris Ivanovic, Or Litany et al. "OmniRe: Omni Urban Scene Reconstruction." arXiv preprint arXiv:2408.16760 (2024).

OmniRe(Ours)



Can Do Fun Stuff!



Can Do Fun Stuff!



Lessons We Have Learned

To Generate And Use Data



Generate Random Geometries For
Better Generalization



VSAIT
Inference



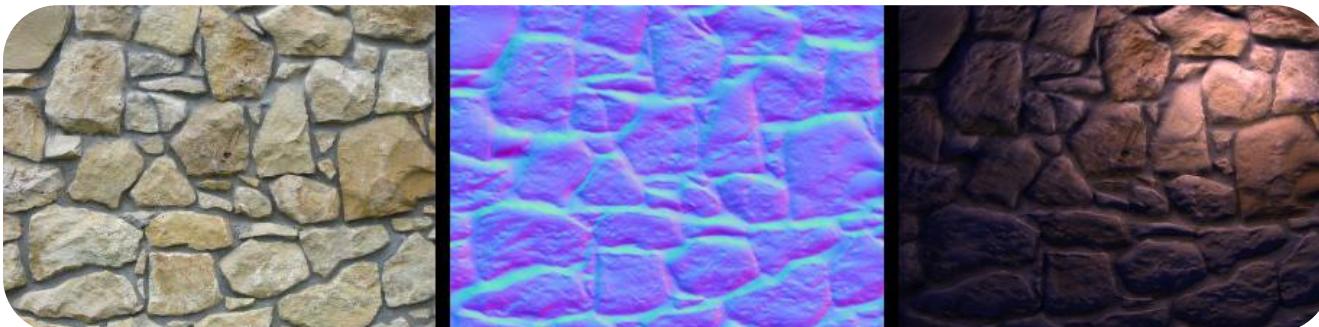
Use Image Translation For Better
Sim2Real



Use Synthetic + Real Data For Better
Performance

Lessons We Have Learned

To Speedup Data



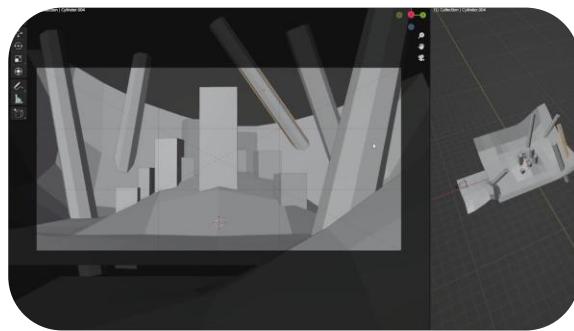
Use Normal Maps To Fake 3D



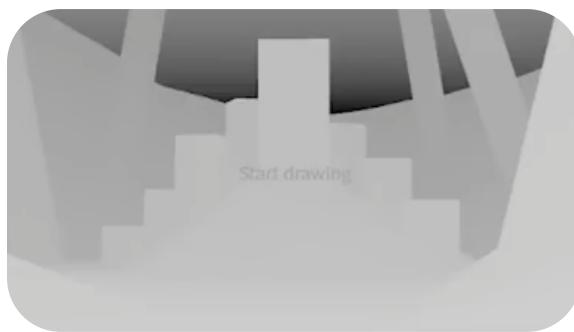
Use Approximate Renderers

Harnessing GenAI

AKA Diffusion Models



Simple Geometry



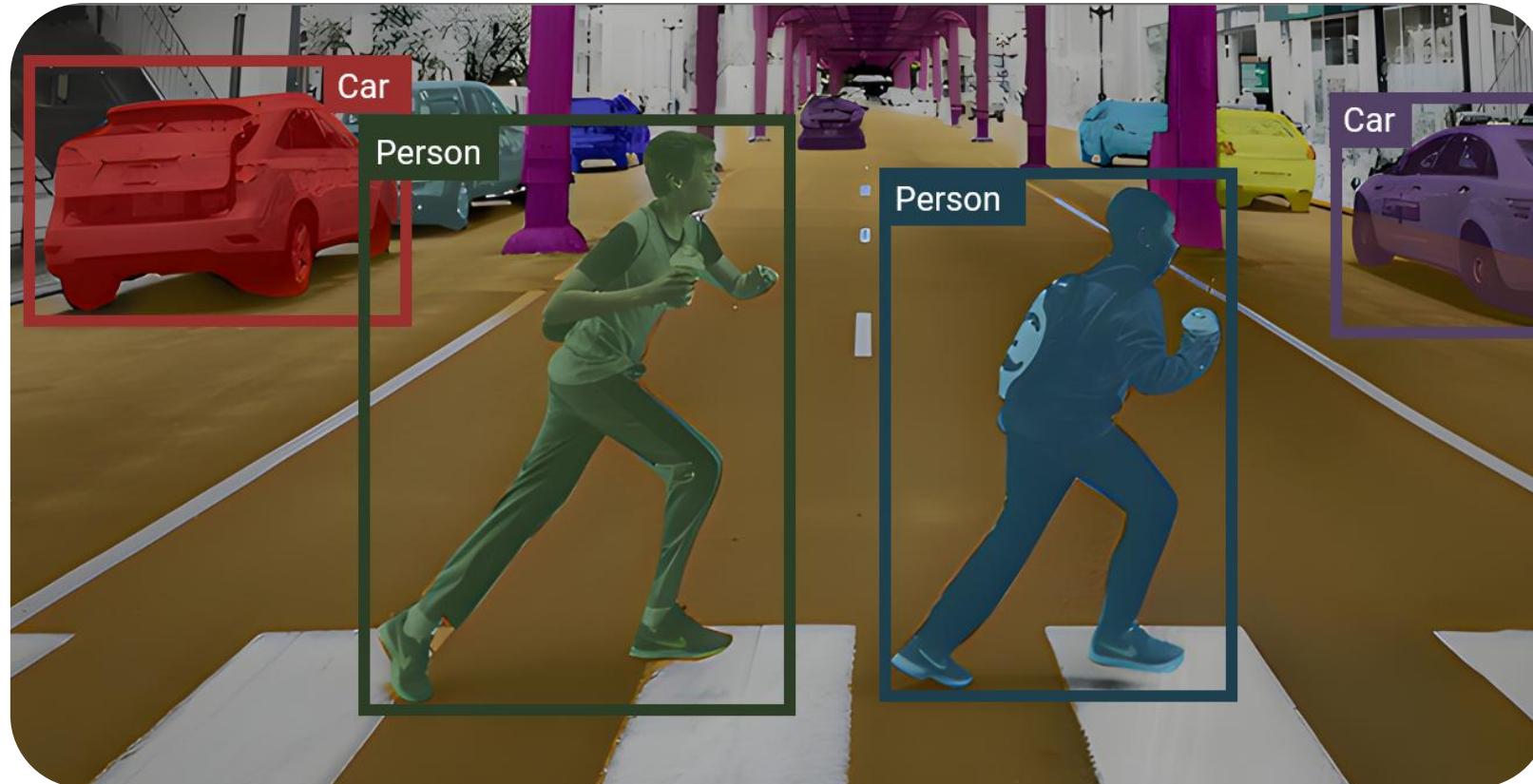
Depth Map

Can also do Semantic Maps



(16229) Use AI to create amazing 3D Animations! [Blender + Stable Diffusion] - YouTube

Next Class!



Object Detection And Segmentation