Generalizing Models to a Diverse World

Judy Hoffman

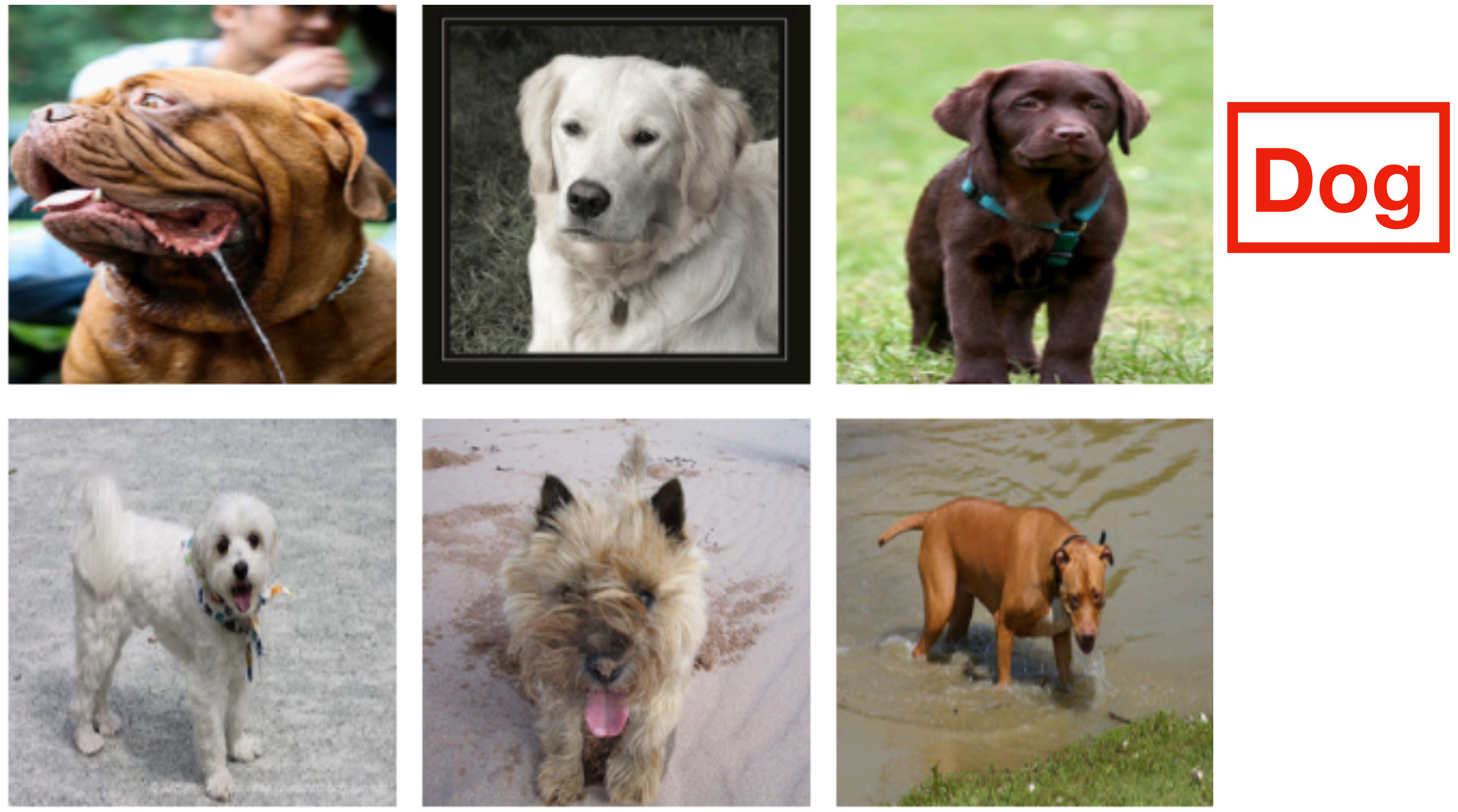
facebook
Artificial Intelligence Research

Georgia Institute of Technology

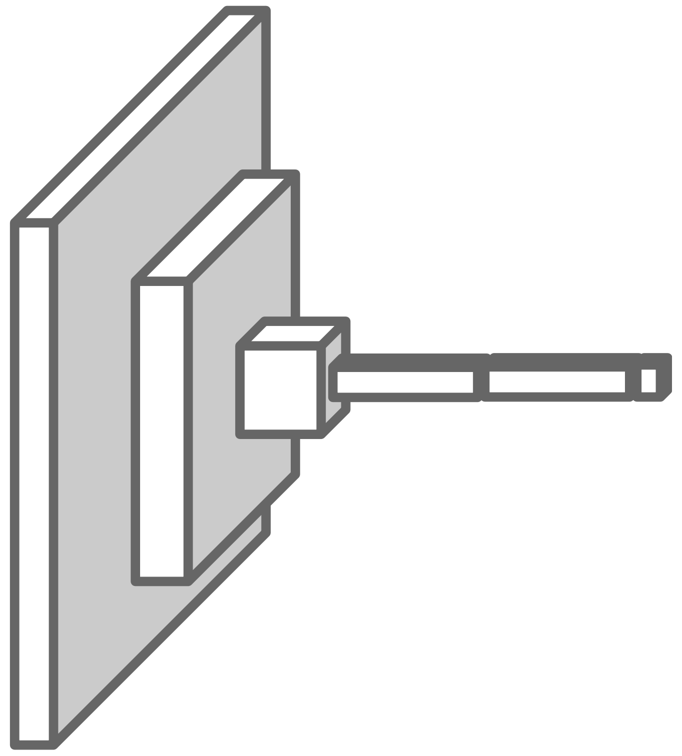
http://data-bdd.berkeley.edu/

Standard Deep Learning Pipeline

1. Collect Data



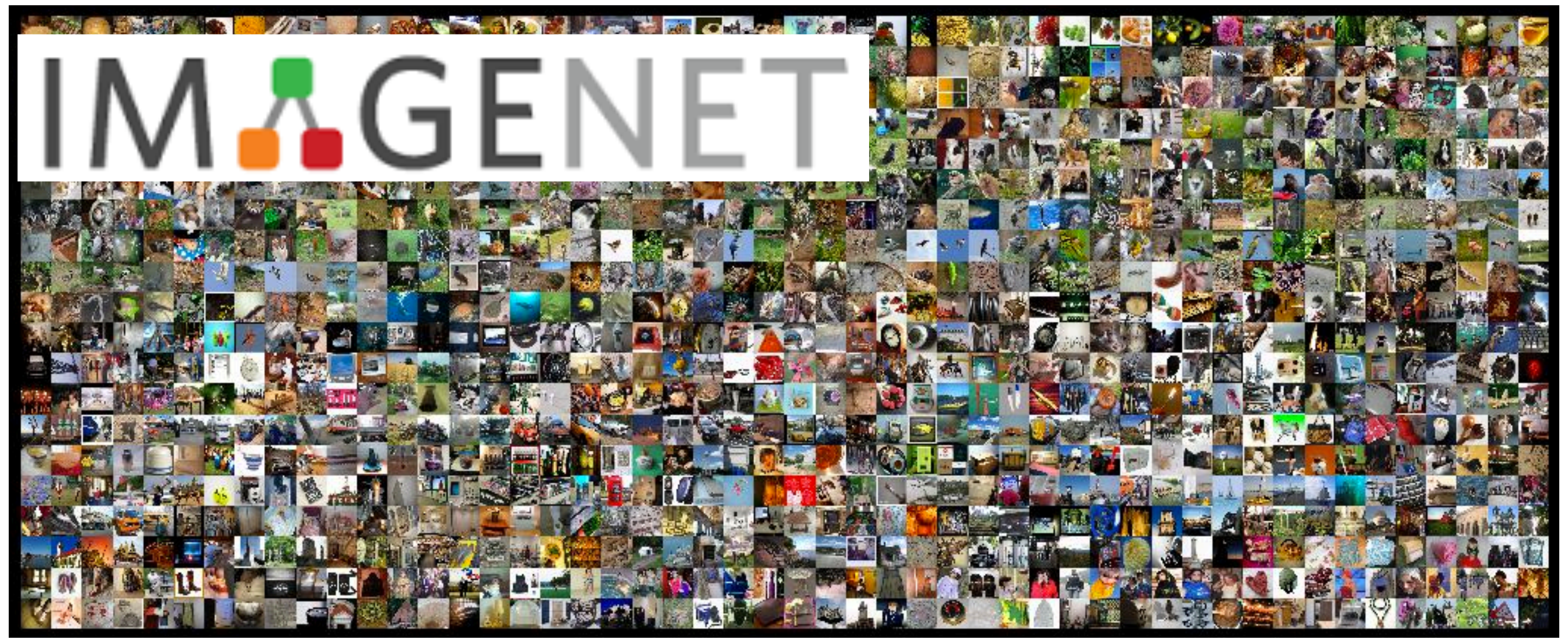
2. Annotate Data



3. Train Model

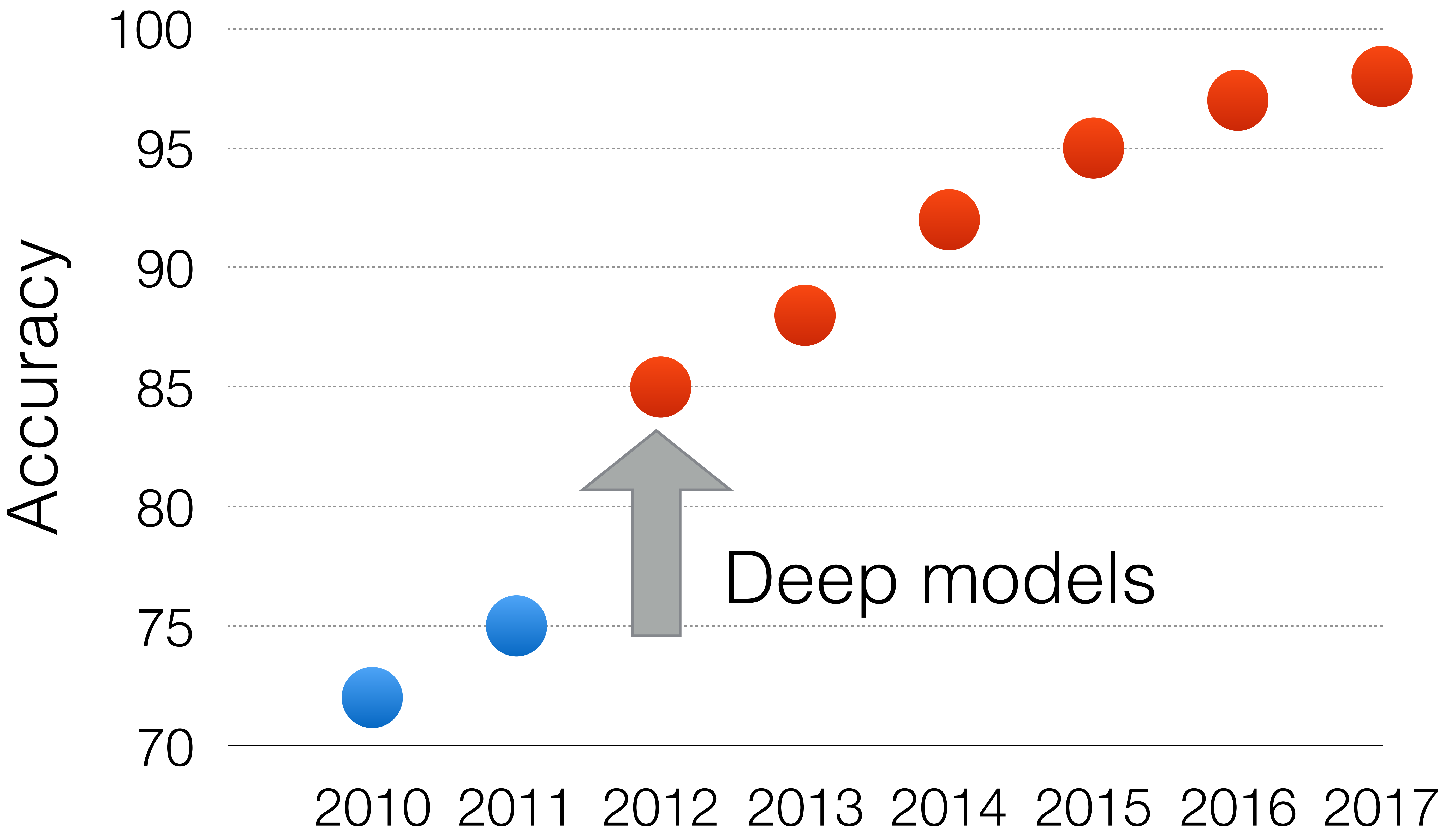
4. Validate Model

Benchmark Performance

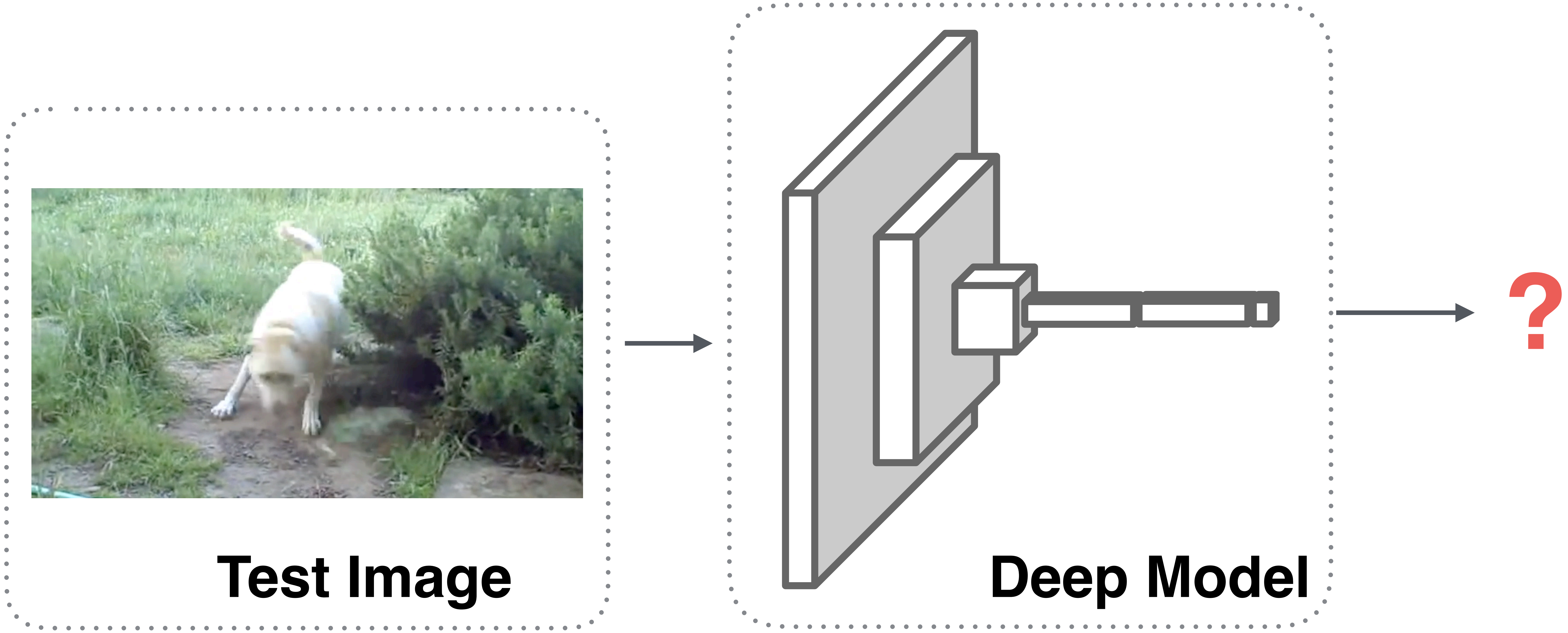


Millions of Images

Challenge to recognize 1000 categories



Dataset Bias

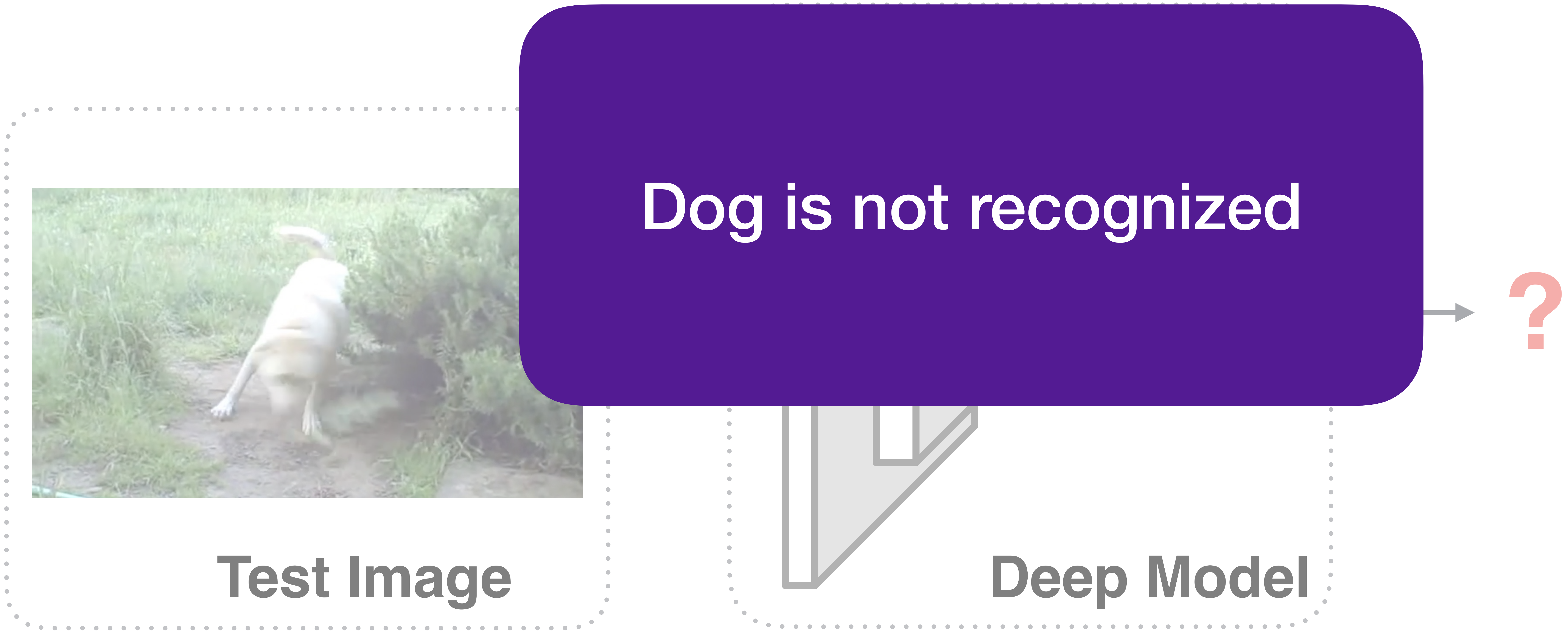


Test Image

Deep Model

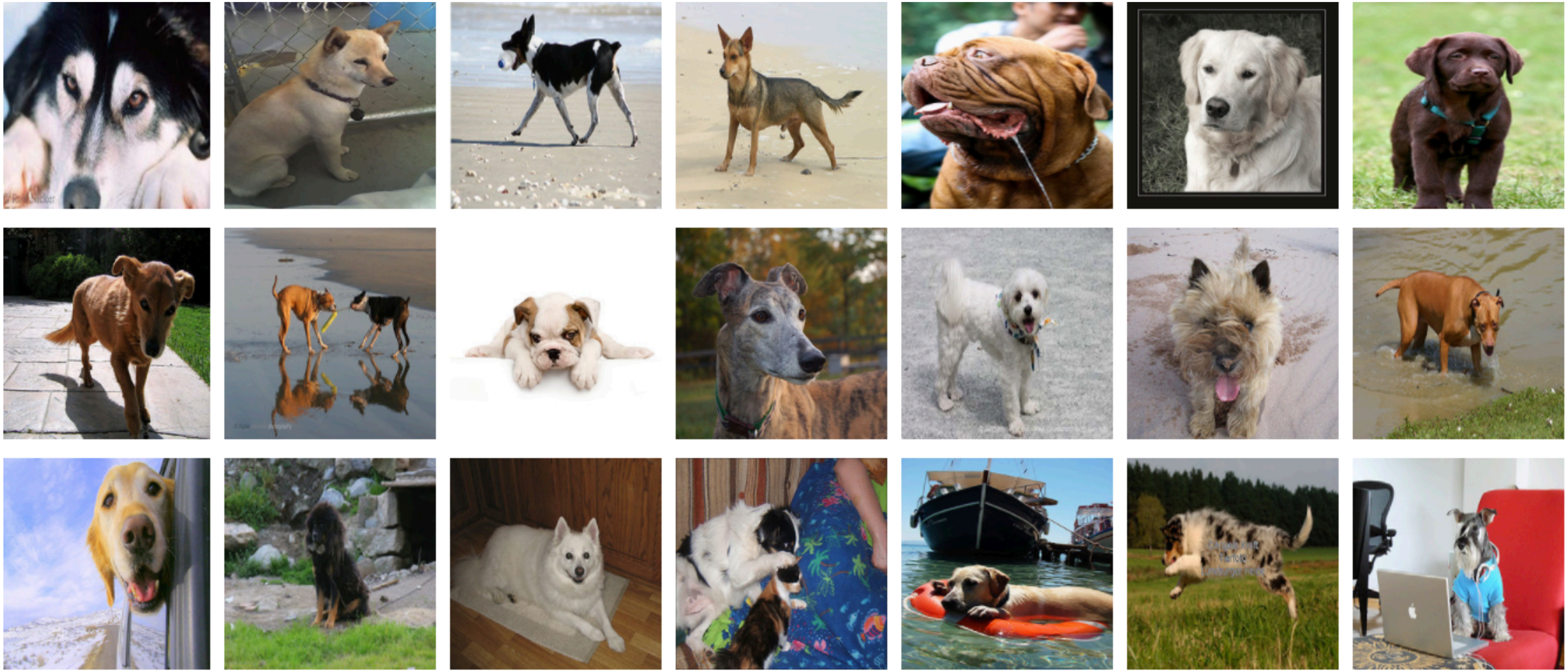
?

Dataset Bias



Dog is not recognized

Dataset Bias



IMAGENET

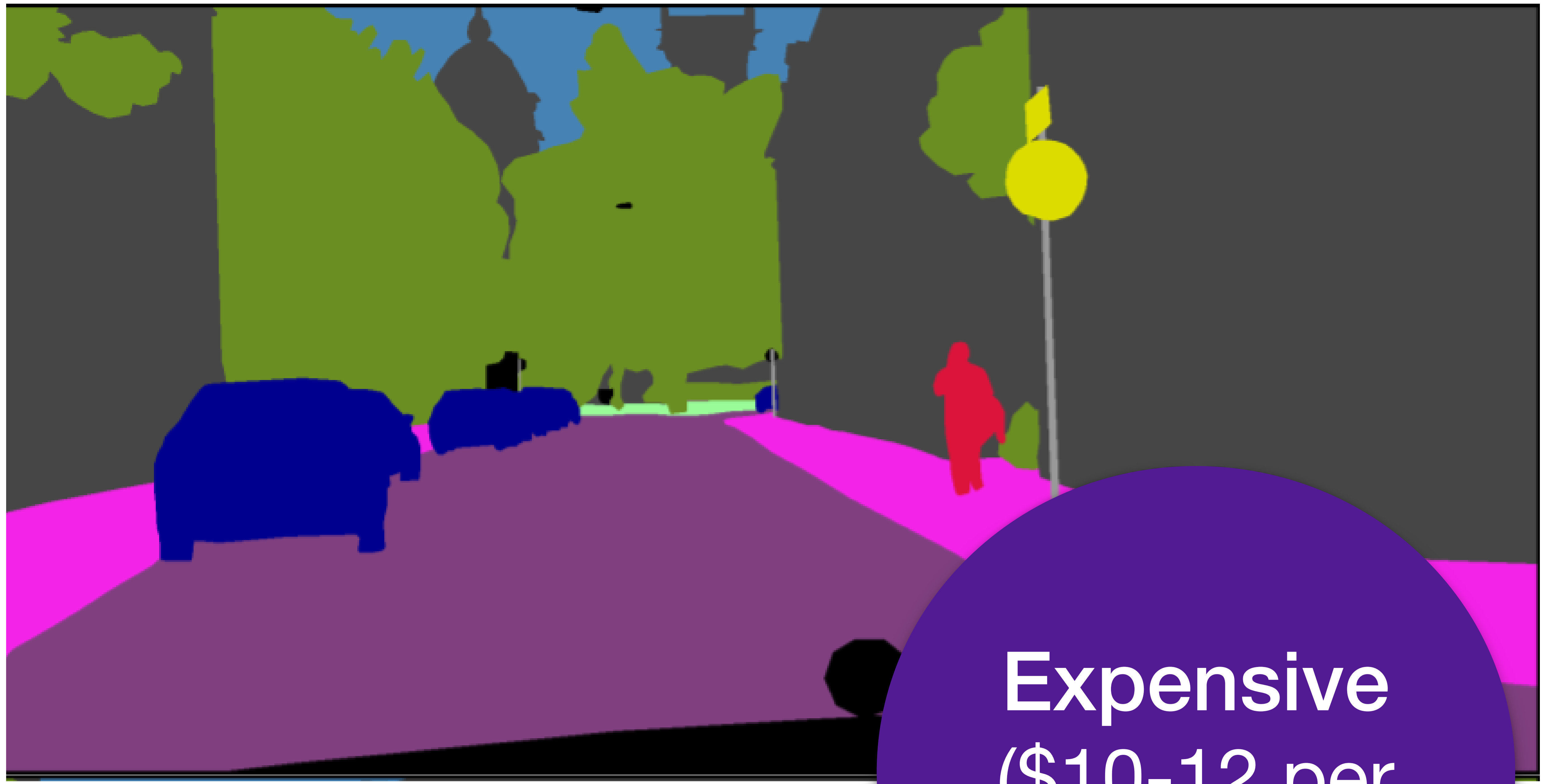
Dataset Bias

Low resolution

Motion Blur

Pose Variety

Diversity of the world



Expensive
($10-12 per image)

Large Potential for Change
Different: Weather, City, Car

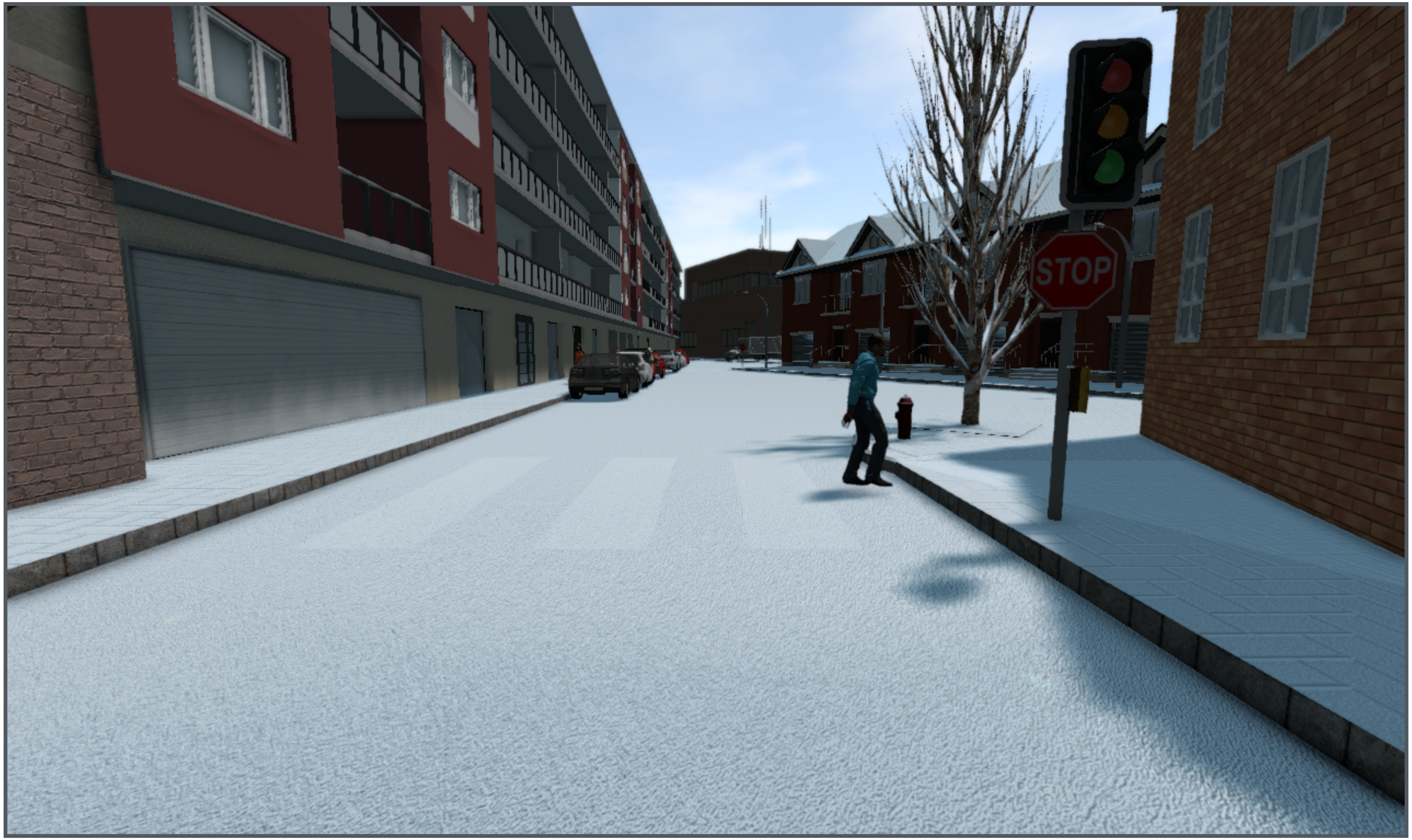
- 🟦 **Car**
- 🟪 **Road**
- 🟪 **Sidewalk**
- 🟥 **Person**
- 🟦 **Sky**
- 🟩 **Vegetation**
- 🟨 **Street Sign**
- ⬛ **Building**

Train in Sunny Weather

Hoffman, Tzeng, Park, Zhu, Isola, Saenko, Efros, Darrell, ICML 2018.

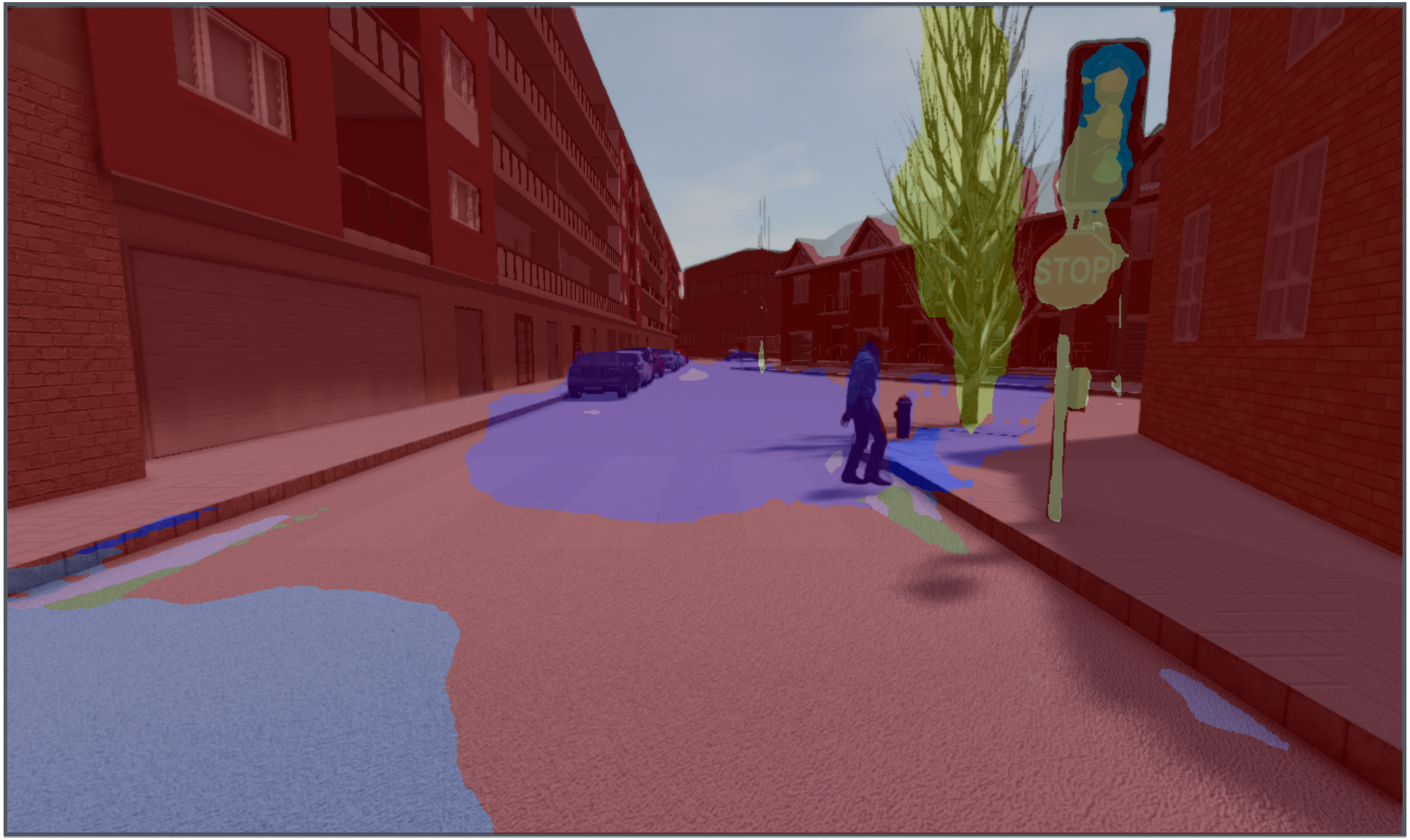
Robust to Weather Changes?

- 🟪 Car
- 🟪 Road
- 🟦 Sidewalk
- 🟪 Person
- ⬜ Sky
- 🟫 Vegetation
- 🟨 Street Sign
- 🟥 Building
- 🟦 Traffic Light

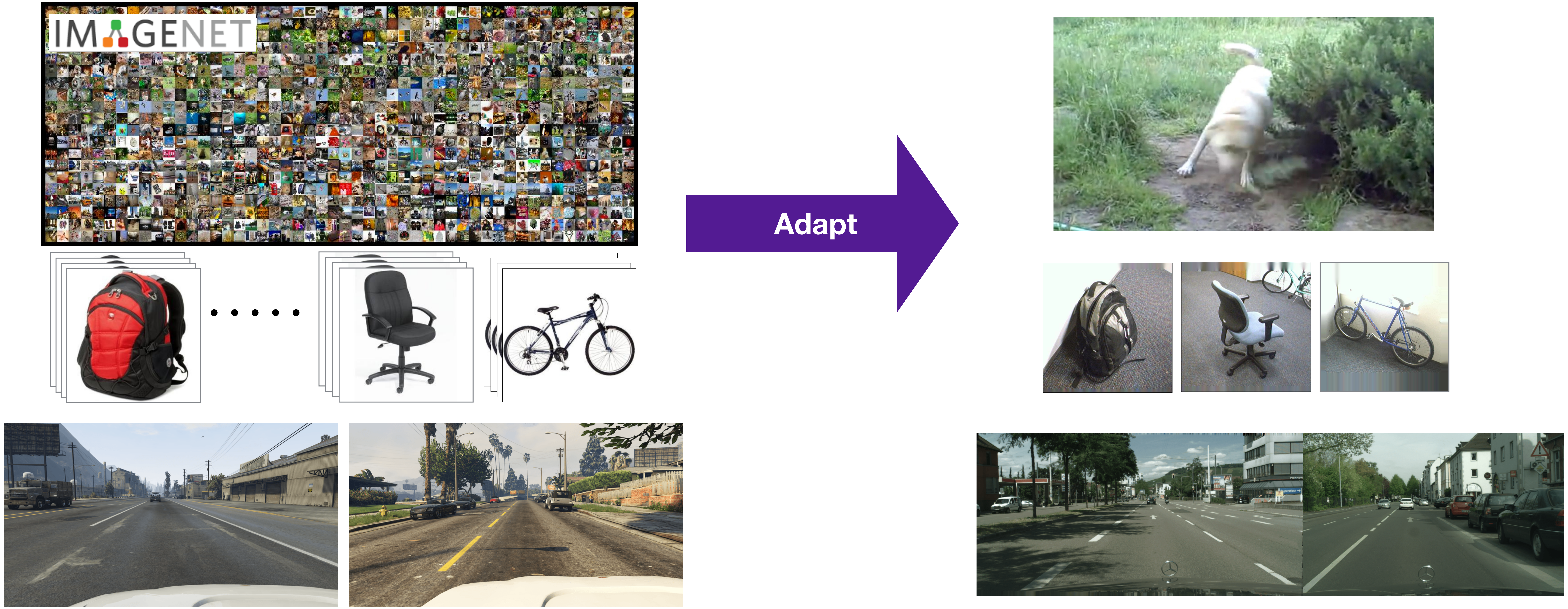


Robust to Weather Changes?

- 🟪 Car
- 🟪 Road
- 🟦 Sidewalk
- 🟪 Person
- ⬜ Sky
- 🟩 Vegetation
- 🟨 Street Sign
- 🟥 Building
- 🟦 Traffic Light



Domain Adaptation: Train on Source Test on Target



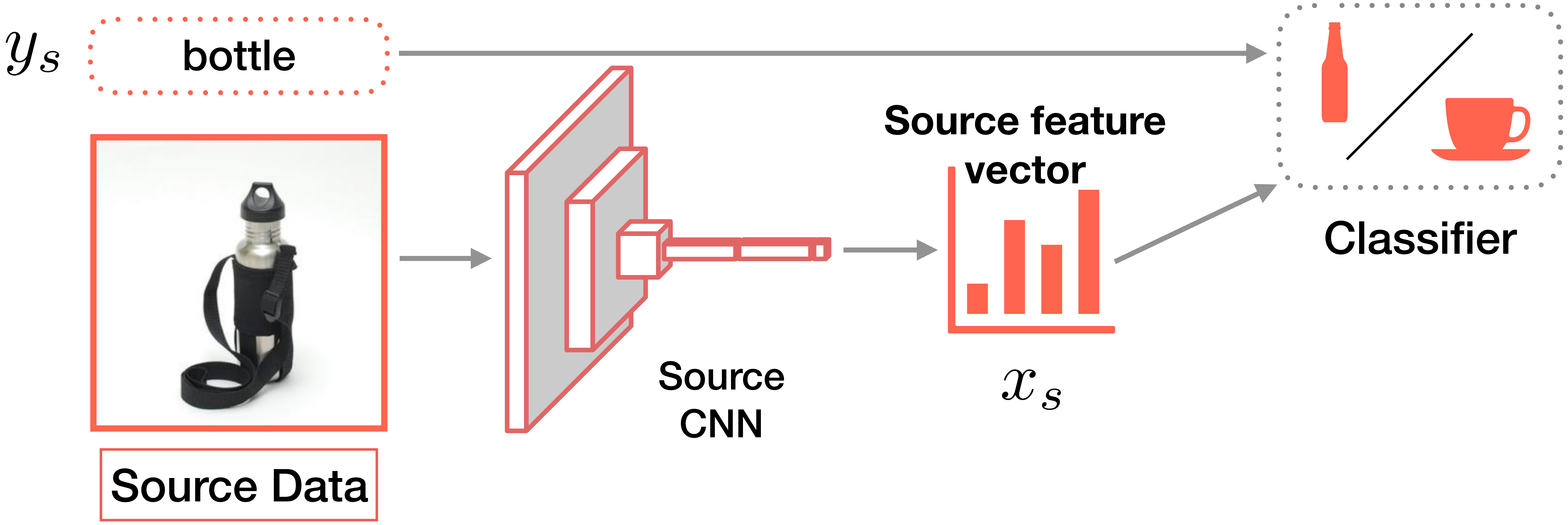
Source Domain $\sim P_S(X_S, Y_S)$

lots of **labeled** data

Target Domain $\sim P_T(X_T, Y_T)$

unlabeled or limited labels

Adversarial Domain Adaptation



y_s

bottle

Source feature vector

Classifier

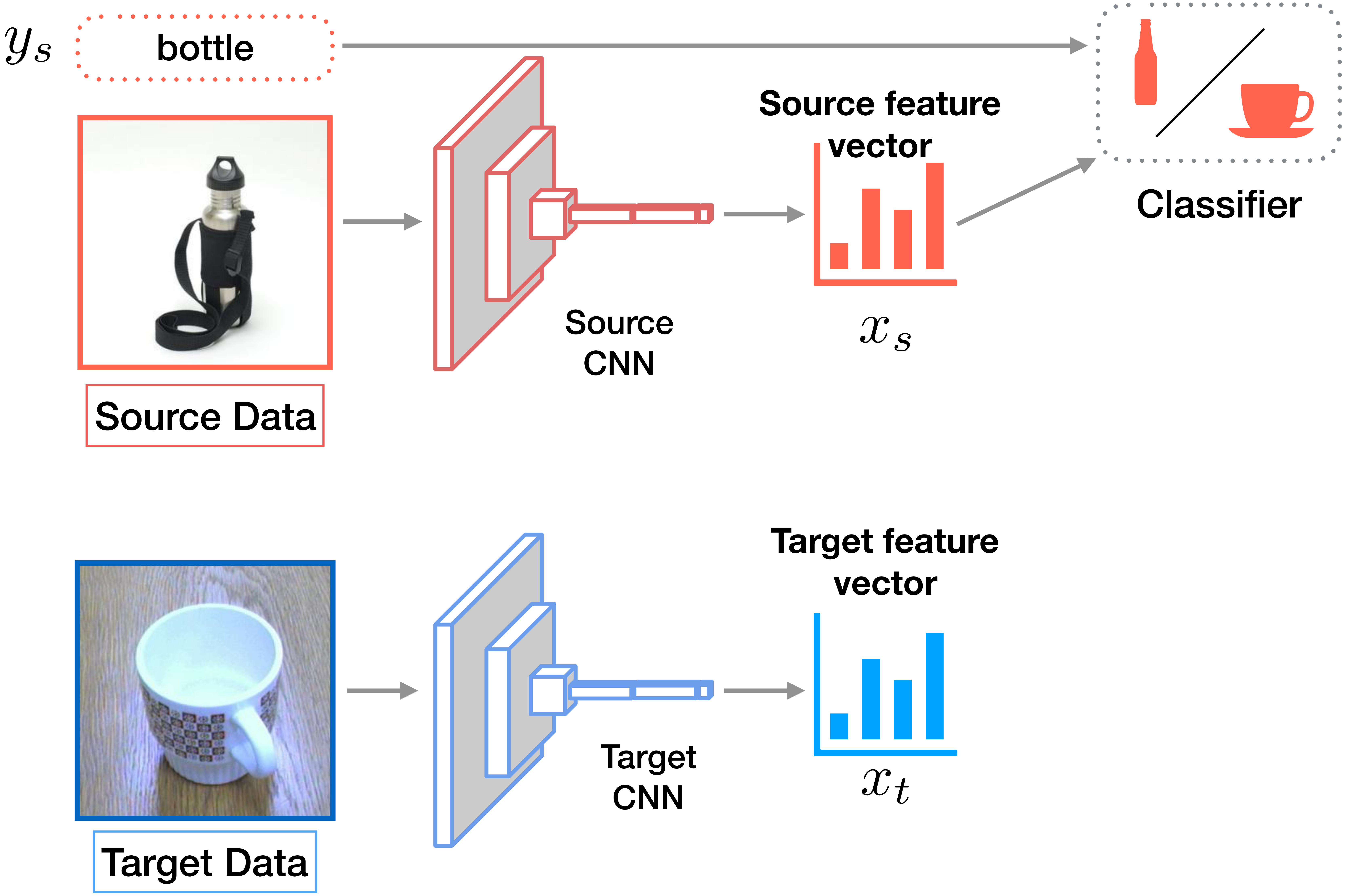
Source CNN

x_s

Source Data

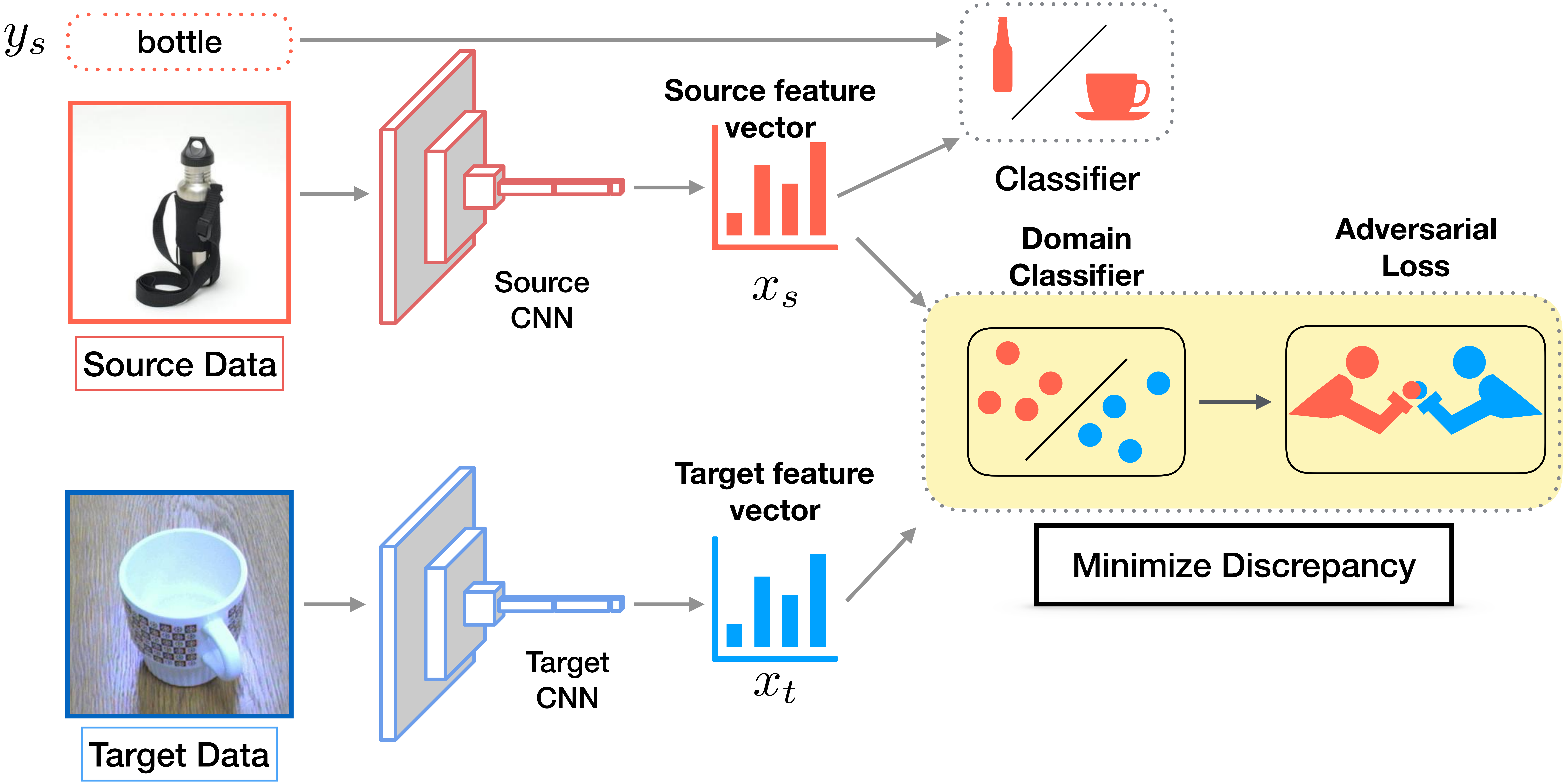
Ganin & Lempinsky, ICML 2015. Tzeng*, Hoffman*, Saenko, Darrell, ICCV 2015. Tzeng, Hoffman, Saenko, Darrell. *CVPR* 2017.

Adversarial Domain Adaptation



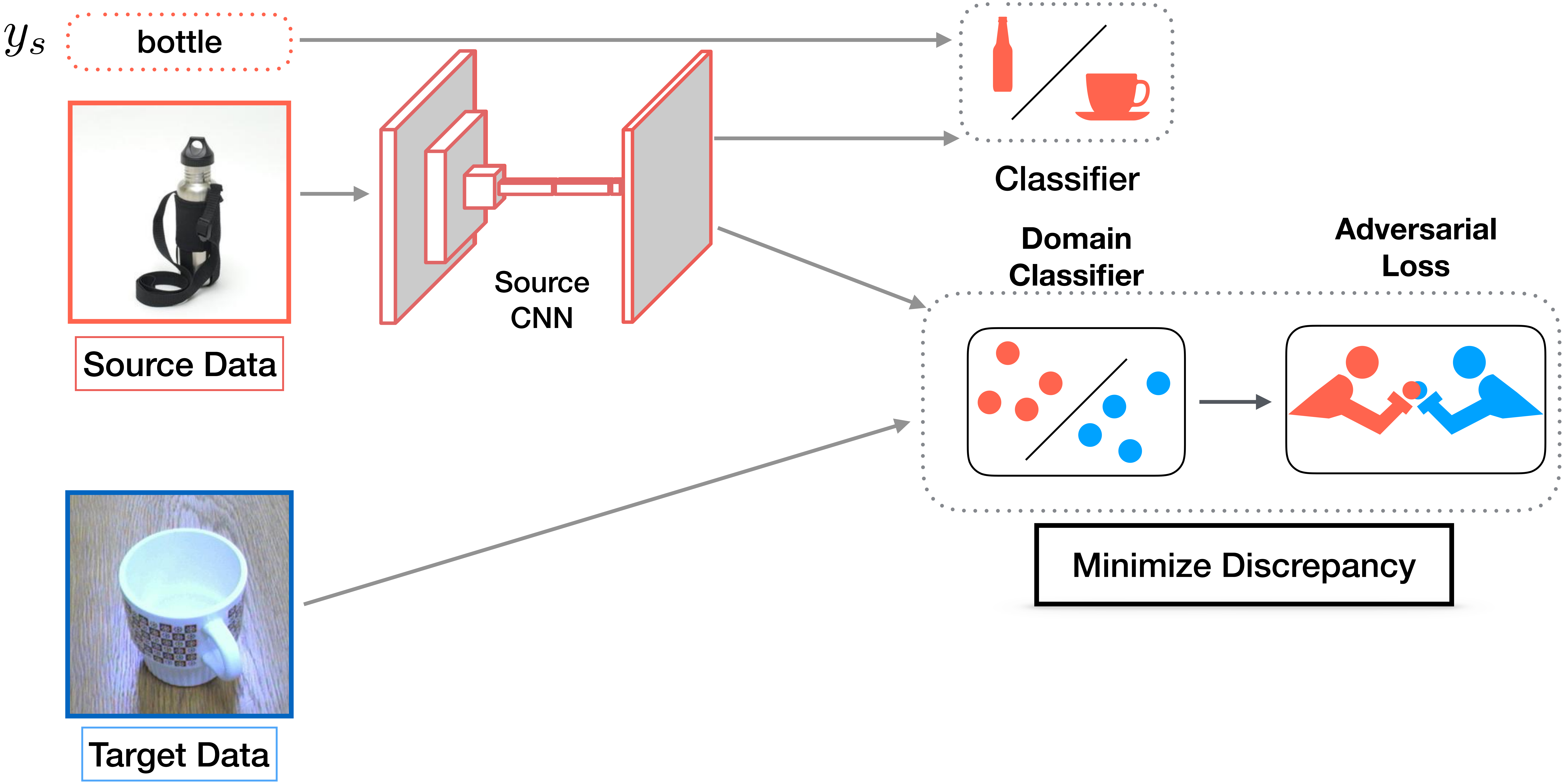
Ganin & Lempinsky, ICML 2015. Tzeng*, Hoffman*, Saenko, Darrell, ICCV 2015. Tzeng, Hoffman, Saenko, Darrell. *CVPR* 2017.

Adversarial Domain Adaptation



Ganin & Lempinsky, ICML 2015. Tzeng*, Hoffman*, Saenko, Darrell, ICCV 2015. Tzeng, Hoffman, Saenko, Darrell. *CVPR* 2017.

Adversarial Domain Adaptation



y_s

bottle

Source Data

Source CNN

Target Data

Classifier

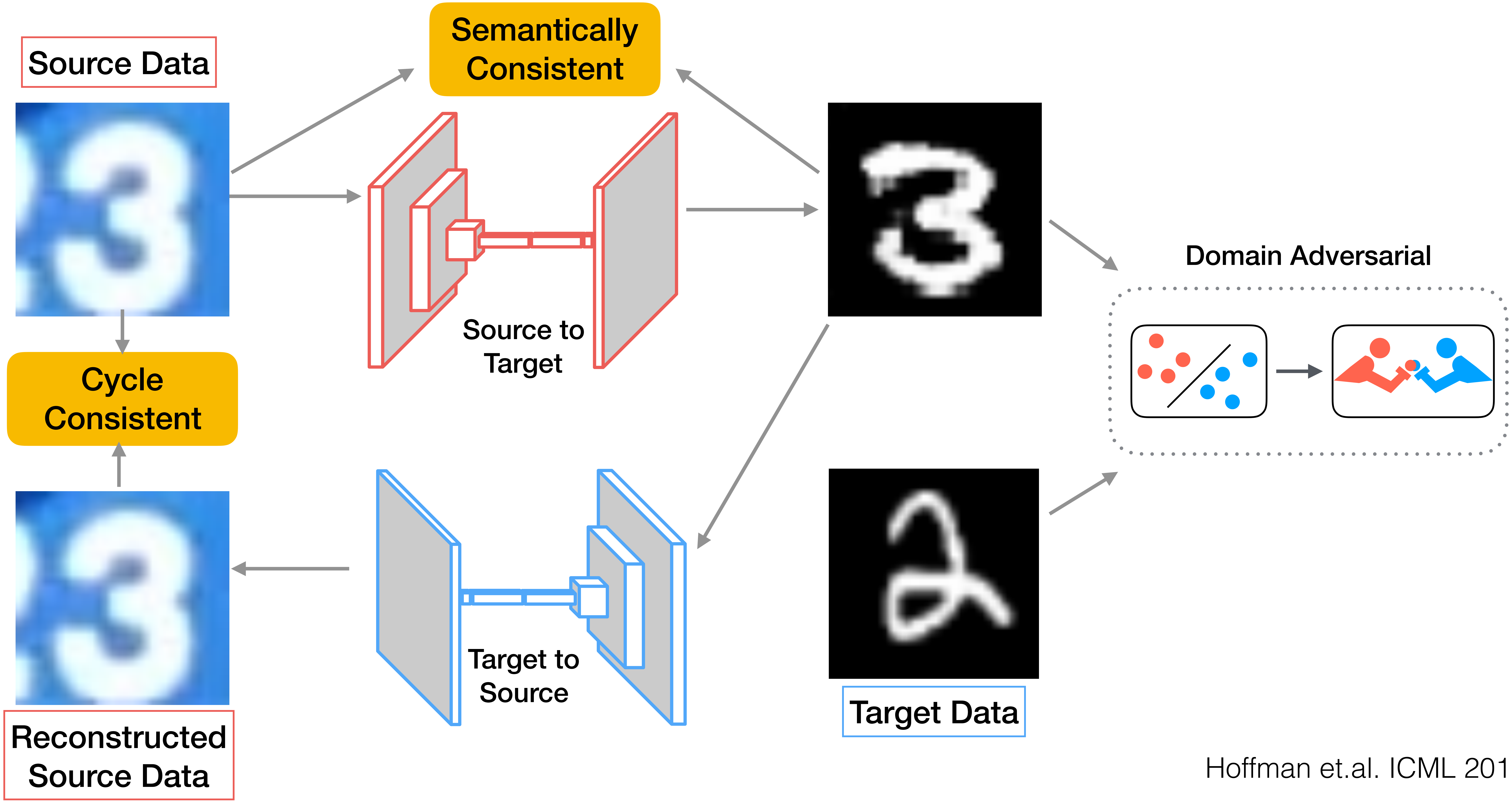
Domain Classifier

Adversarial Loss

Minimize Discrepancy

Liu 2016. Taigman 2016. Bousmalis 2017. Liu 2017. Kim 2017. Sankaranarayanan 2018. Hoffman 2018.

CyCADA: Cycle Consistent Adversarial DA

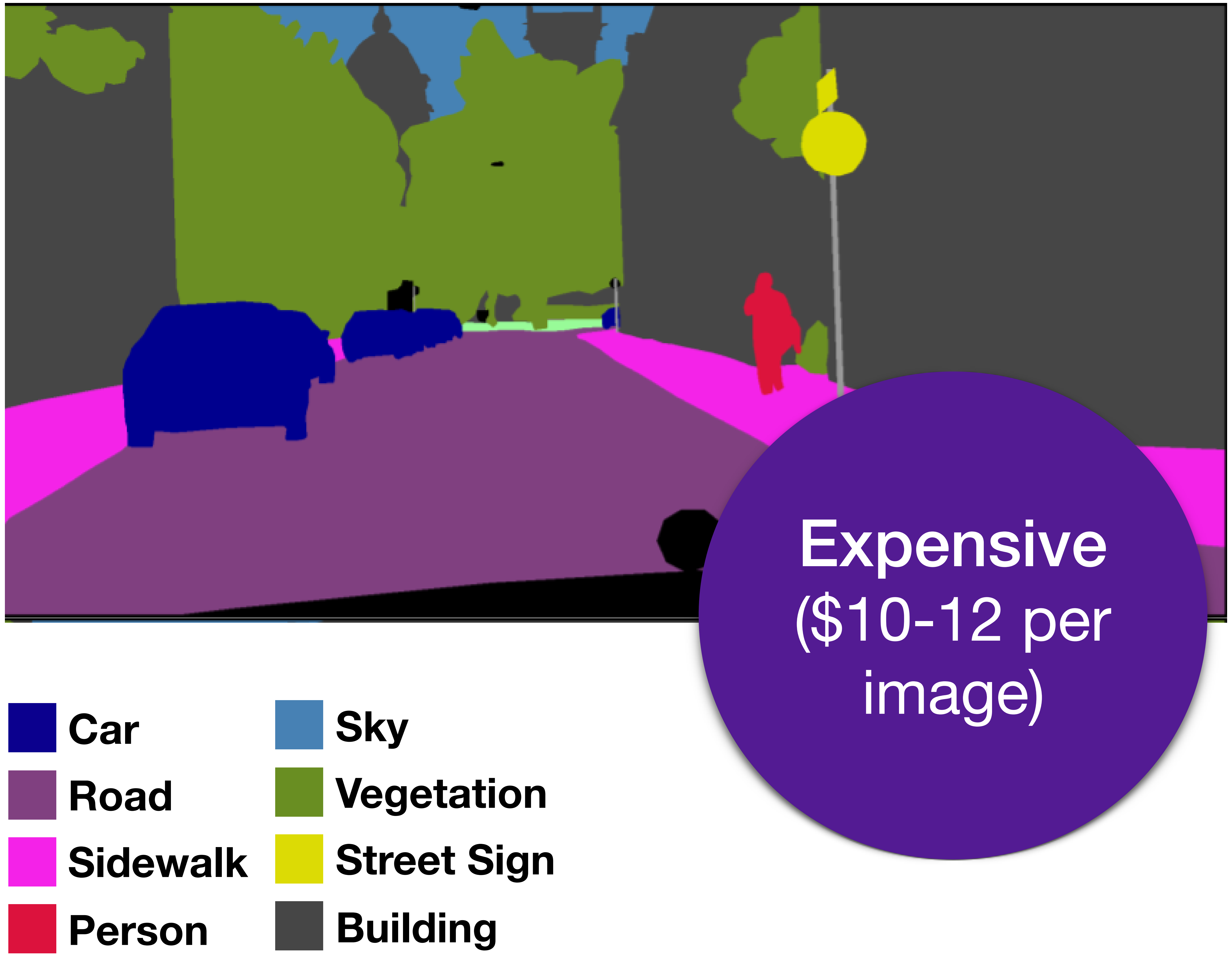


Hoffman et.al. ICML 2018

Adaptation of Semantic Segmentation

Large Potential for Change

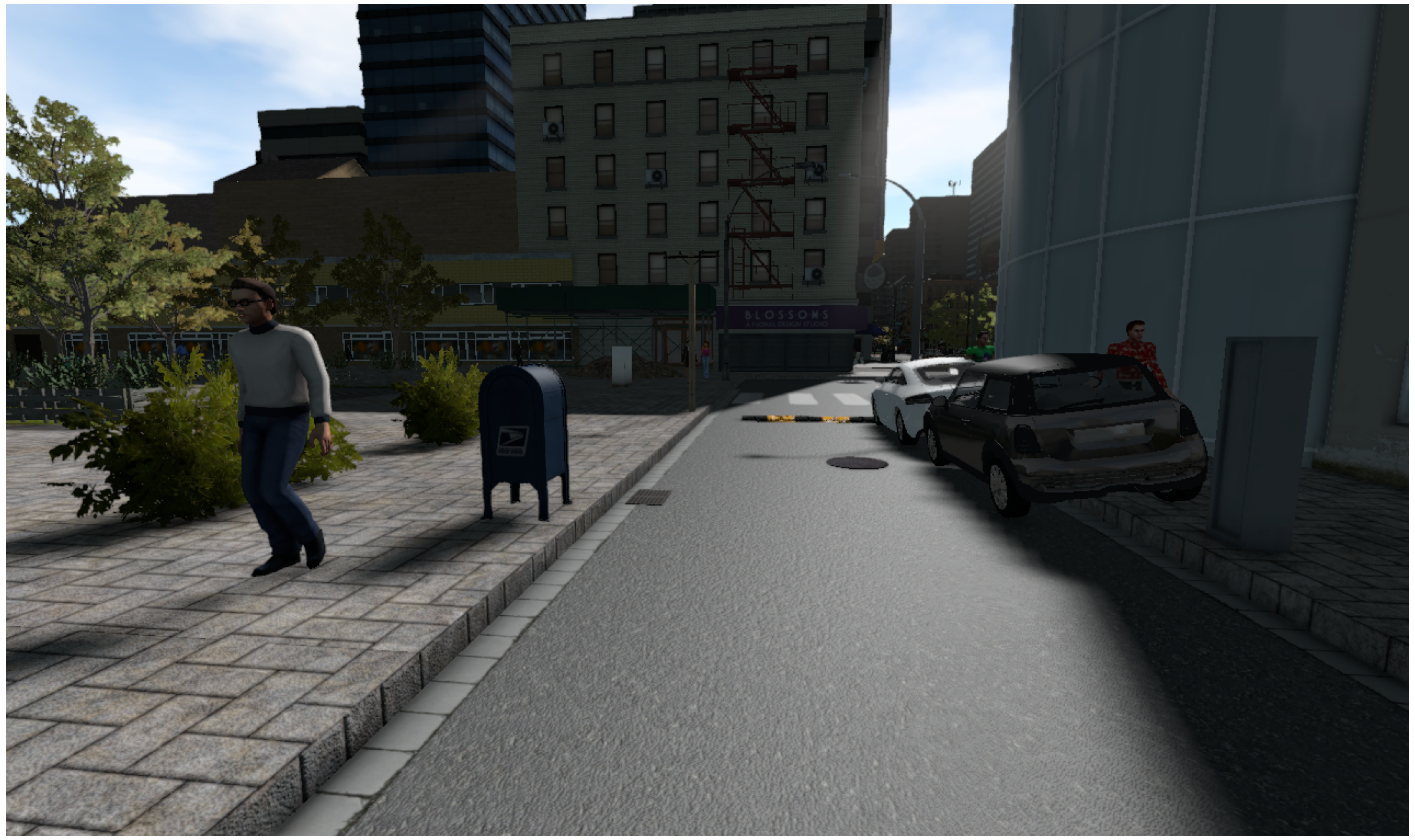
Different: Weather, City, Car



Expensive
($10-12 per image)

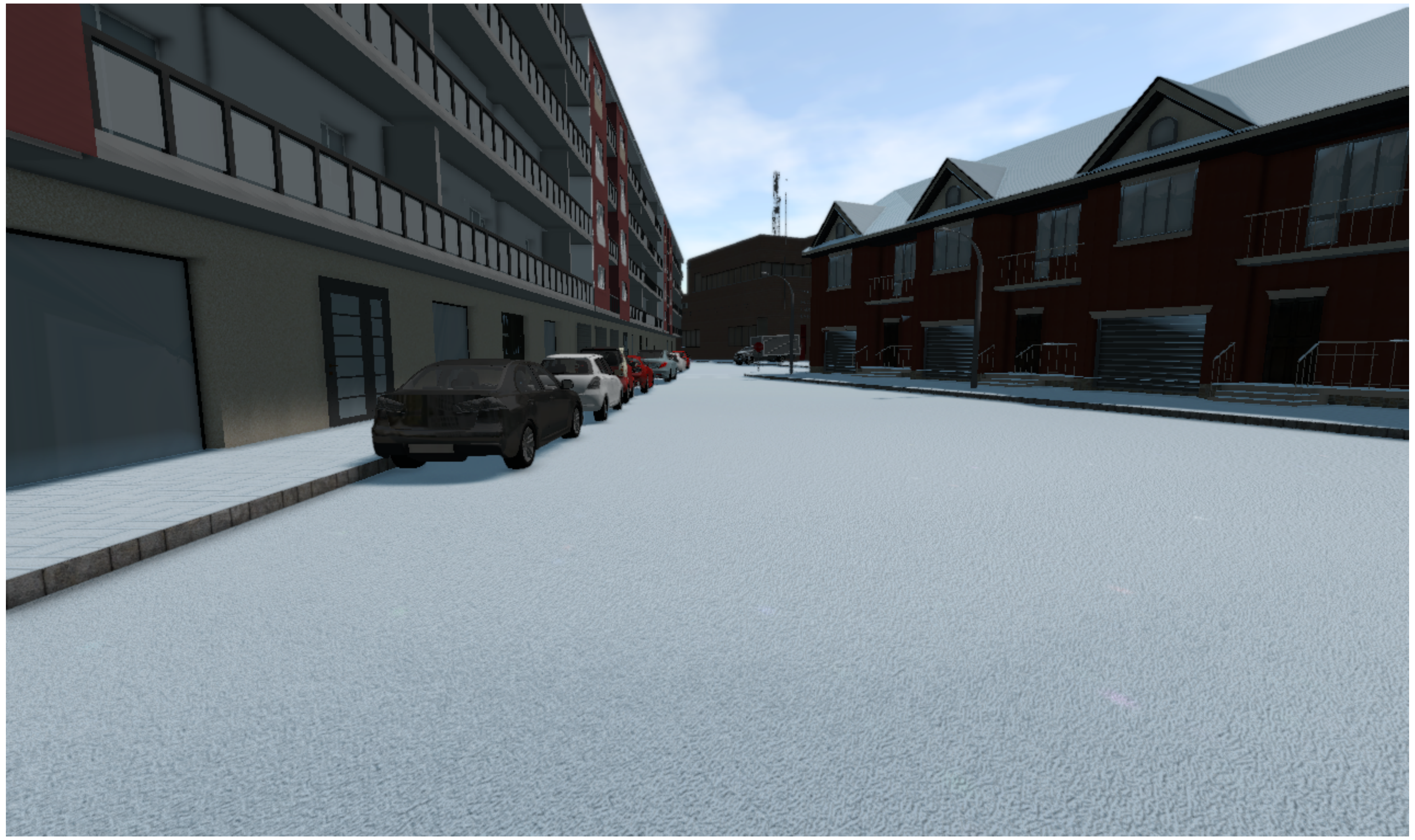
Cross Season Adaptation

Train



Fall Image

Test



Winter Image

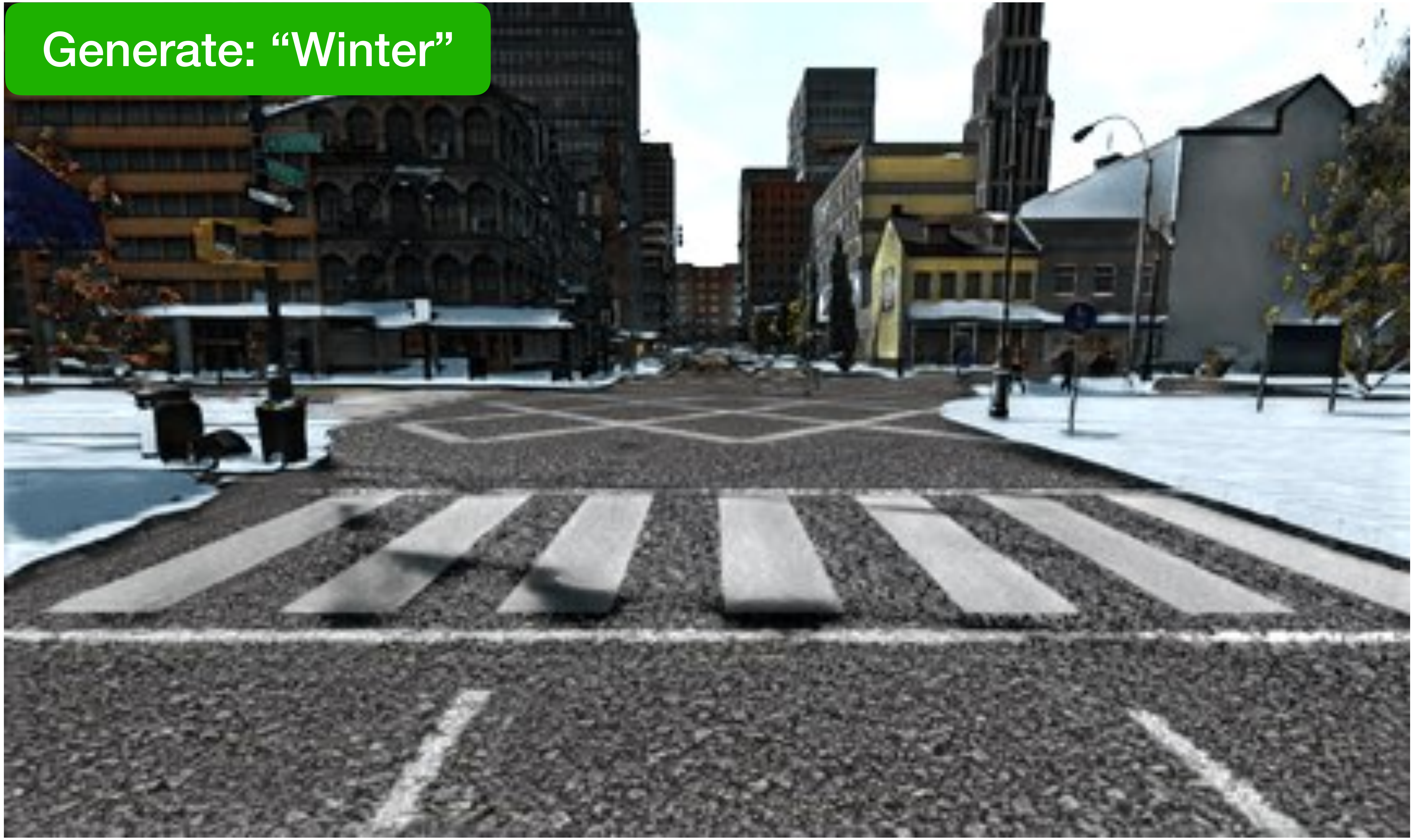
SYNTHIA Dataset

Hoffman, Wang, Yu, Darrell, arXiv 2017.
Hoffman, Tzeng, Park, Zhu, Isola, Saenko, Efros, Darrell, ICML 2018.

Cross Season Pixel Adaptation

Hoffman, Tzeng, Park, Zhu, Isola, Saenko, Efros, Darrell, ICML 2018.

Cross Season Pixel Adaptation

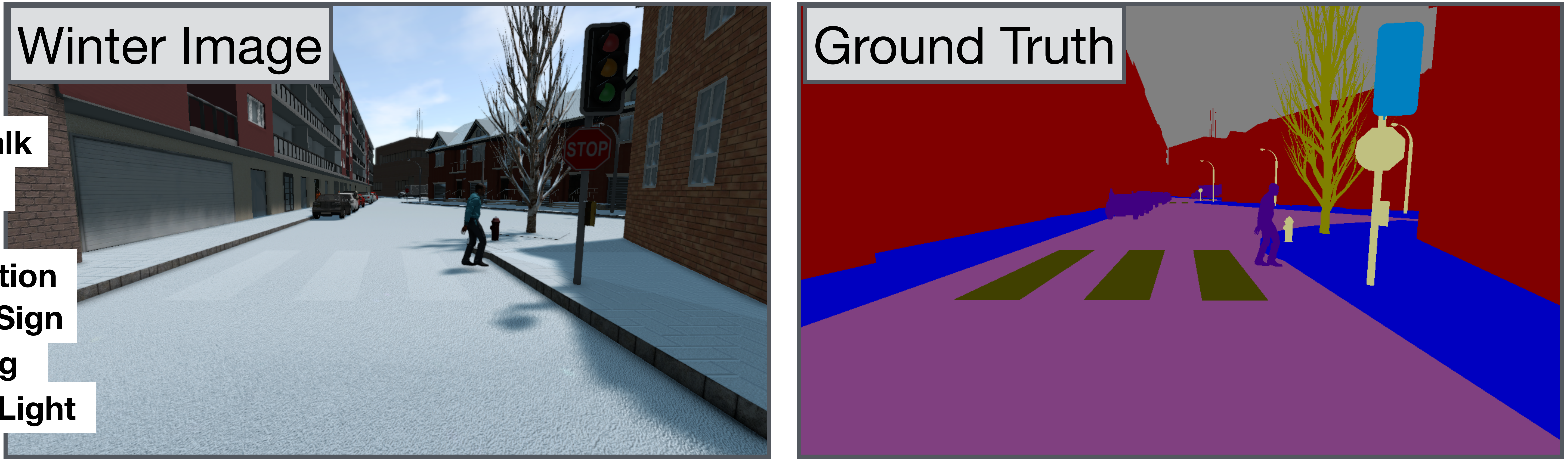


Generate: “Winter”

Hoffman, Tzeng, Park, Zhu, Isola, Saenko, Efros, Darrell, ICML 2018.

Cross Season Adaptation

- 🟪 Car
- 🟪 Road
- 🟦 Sidewalk
- 🟪 Person
- ⬜ Sky
- 🟩 Vegetation
- 🟨 Street Sign
- 🟥 Building
- 🟦 Traffic Light



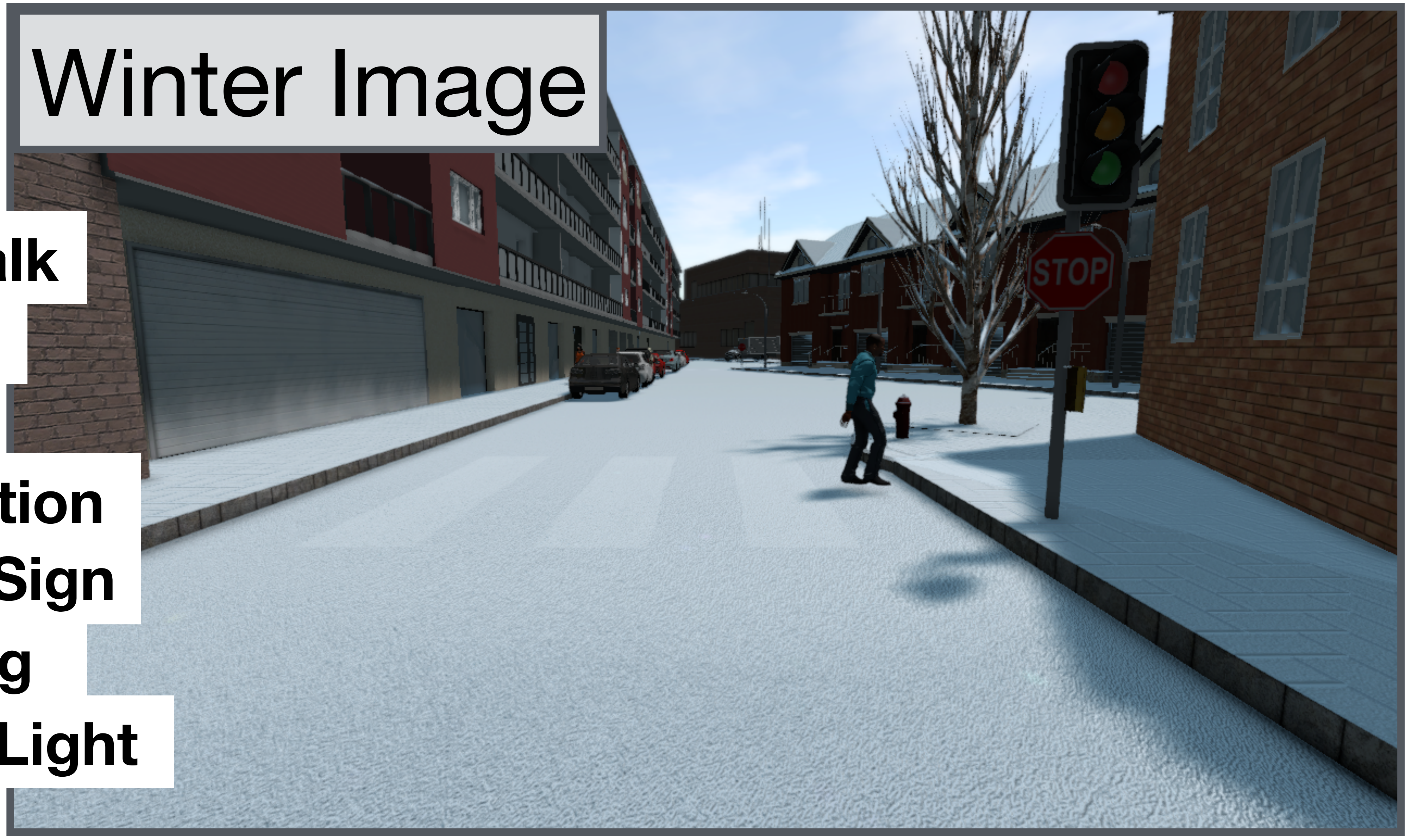
Winter Image

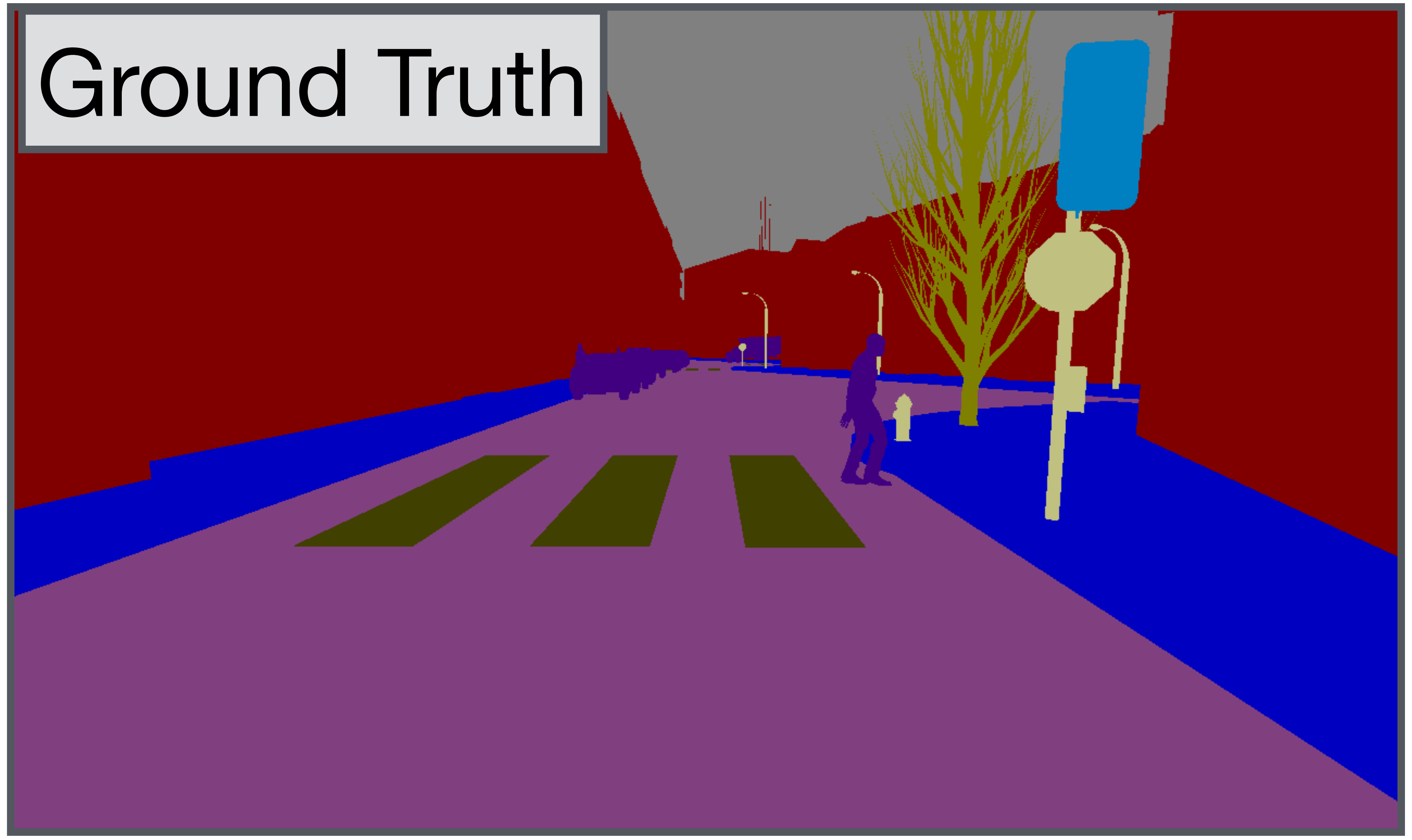
Ground Truth

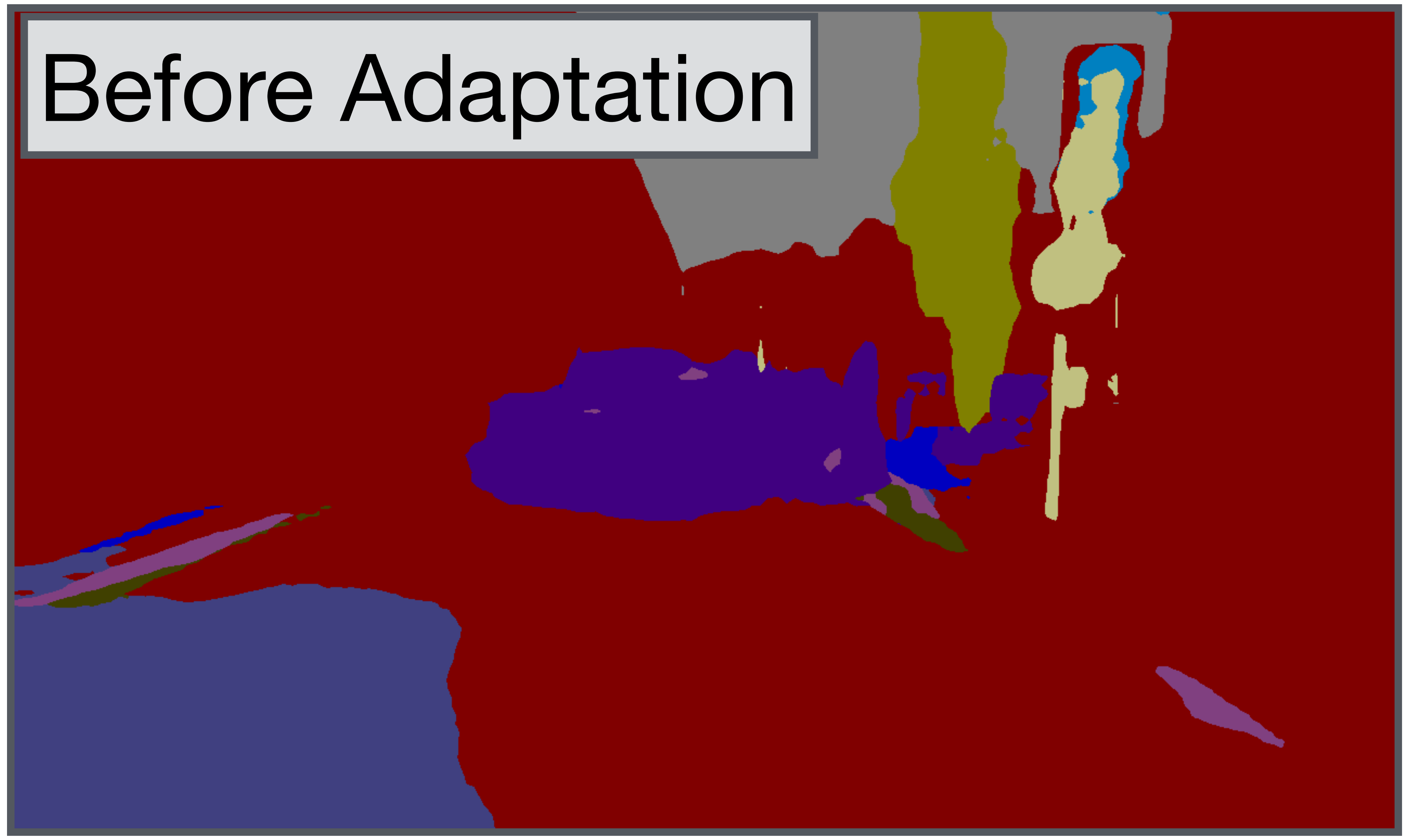
Hoffman, Wang, Yu, Darrell, arXiv 2017.

Cross Season Adaptation

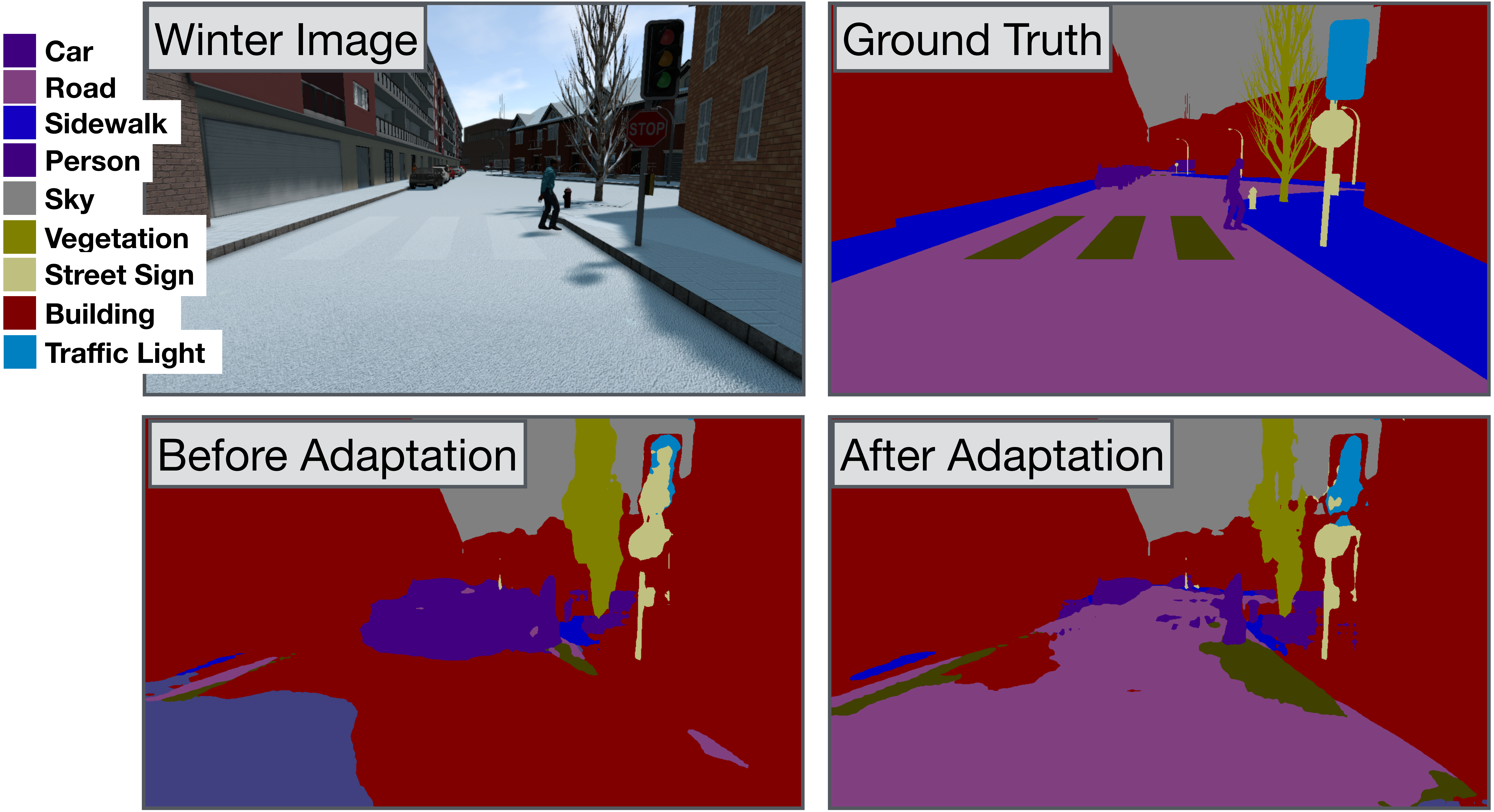
- 🟪 Car
- 🟪 Road
- 🟦 Sidewalk
- 🟪 Person
- ⬜ Sky
- 🟩 Vegetation
- 🟨 Street Sign
- 🟥 Building
- 🟦 Traffic Light


Winter Image


Ground Truth


Before Adaptation

Cross Season Adaptation

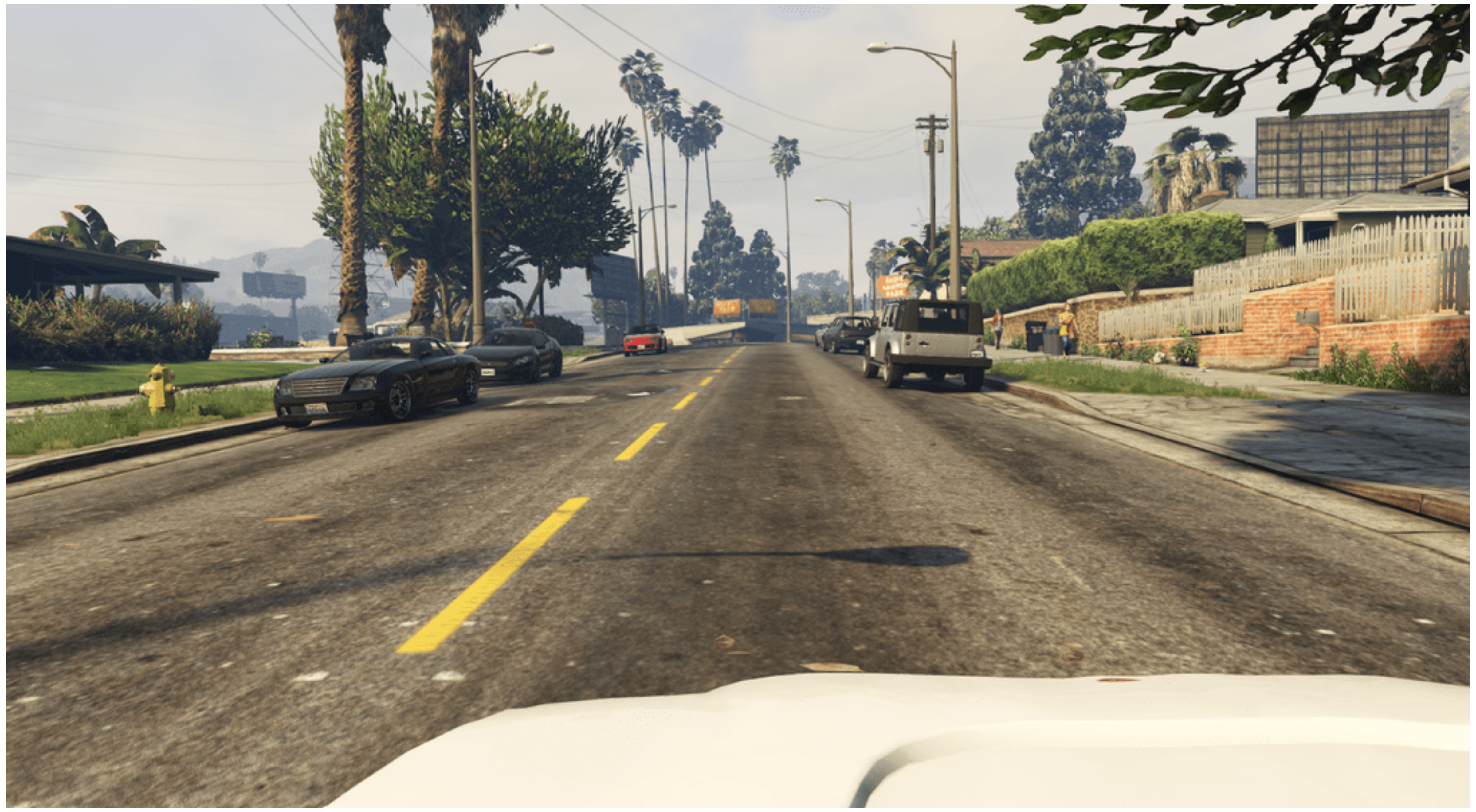


Day to Night Pixel Adaptation

Zhu*, Park*, Isola, Efros. ICCV 2017.

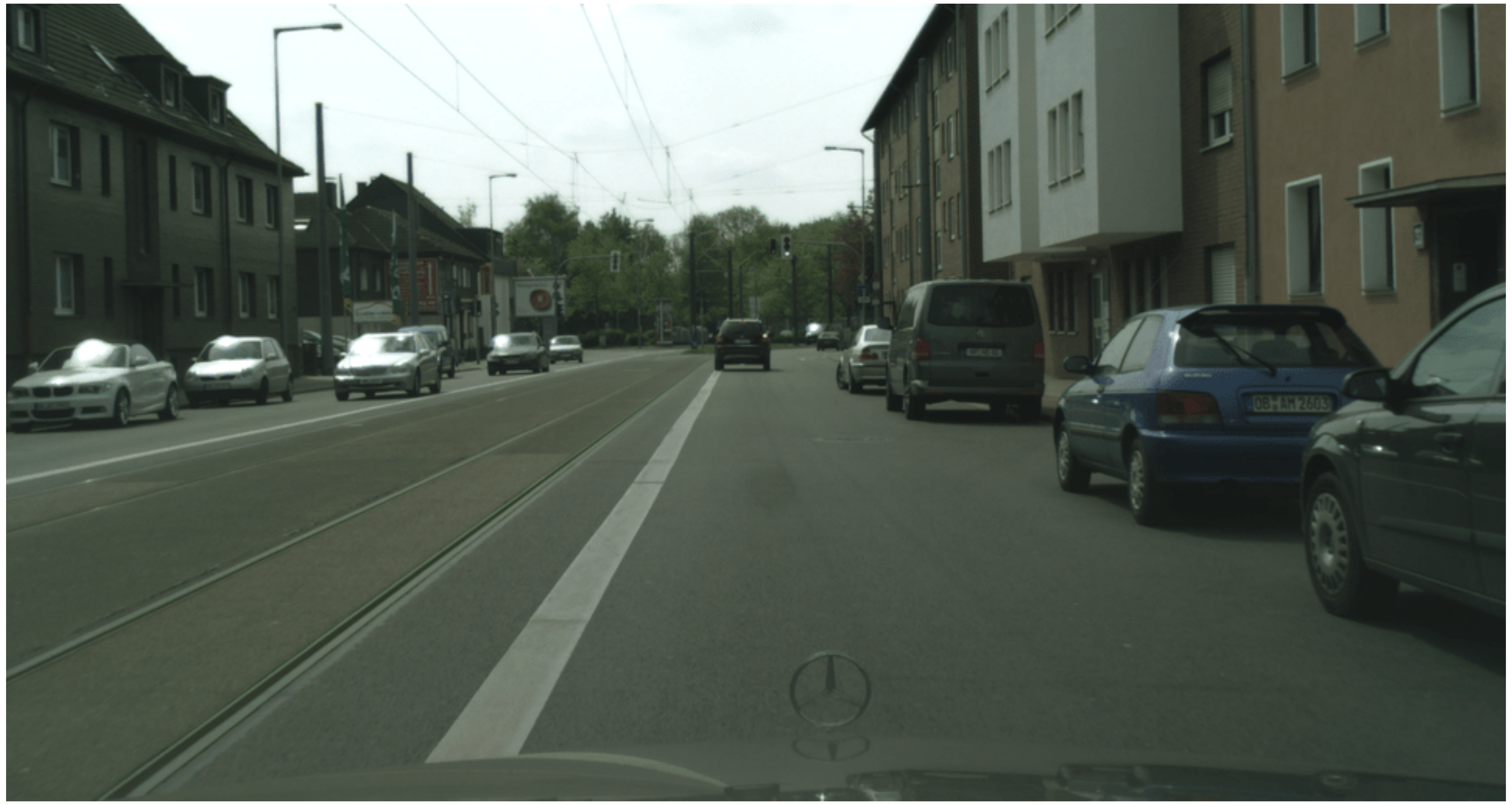
Synthetic to Real Pixel Adaptation

Train



GTA (synthetic)

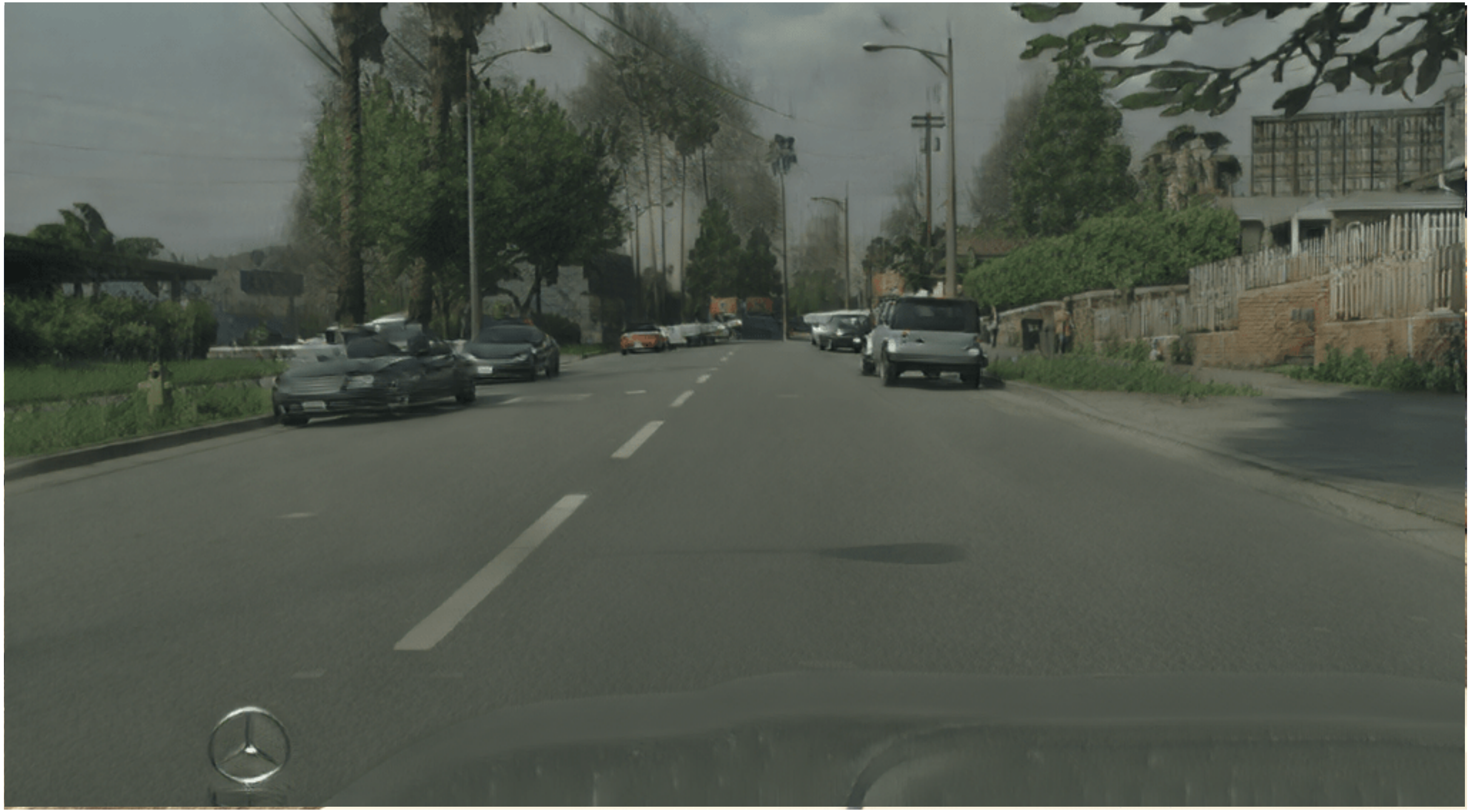
Test



CityScapes (Germany)

Hoffman et.al. ICML 2018

Synthetic to Real Pixel Adaptation

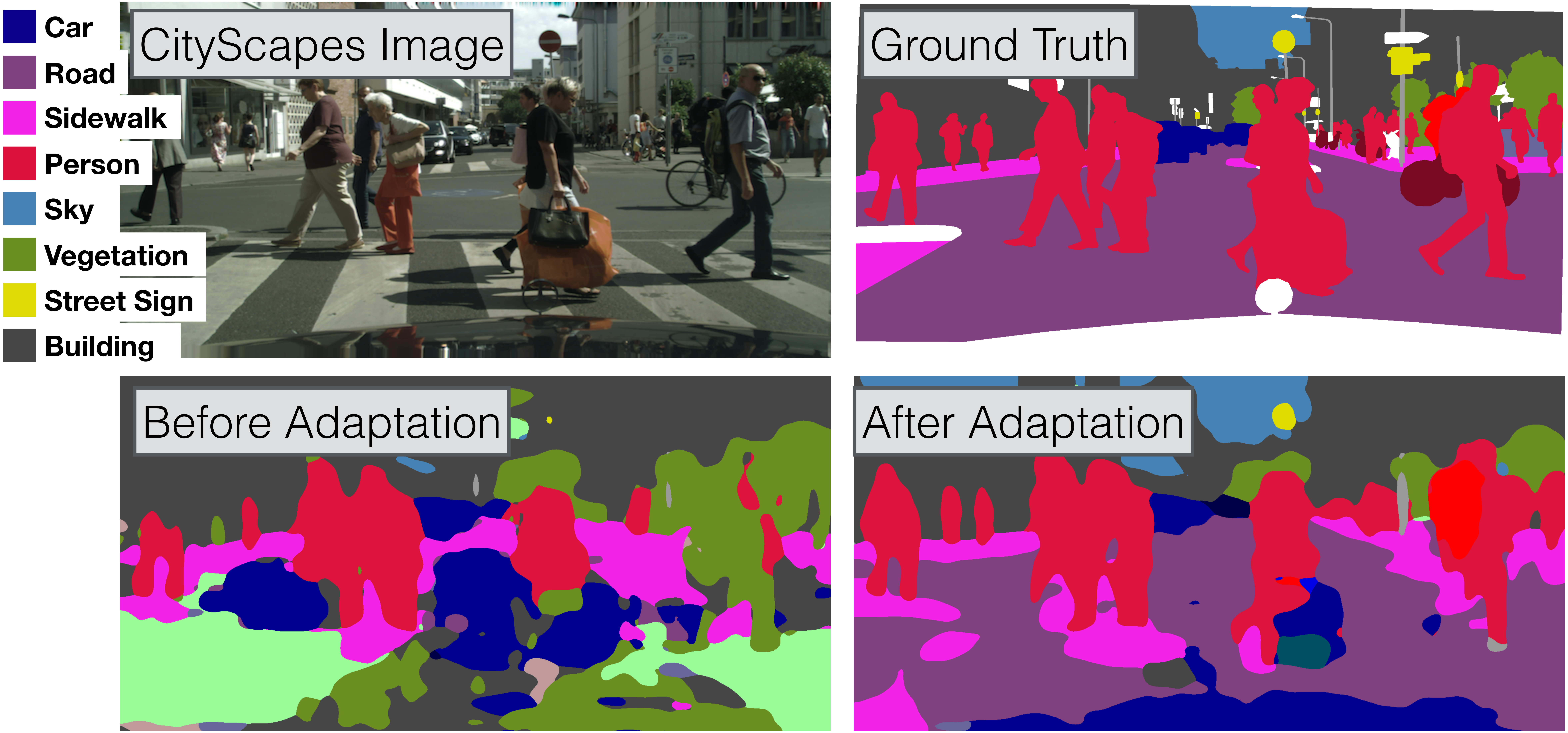


Hoffman et.al. ICML 2018

Synthetic to Real Pixel Adaptation

Zhu*, Park*, Isola, Efros. ICCV 2017.

CyCADA Results: CityScapes Evaluation



CityScapes Image

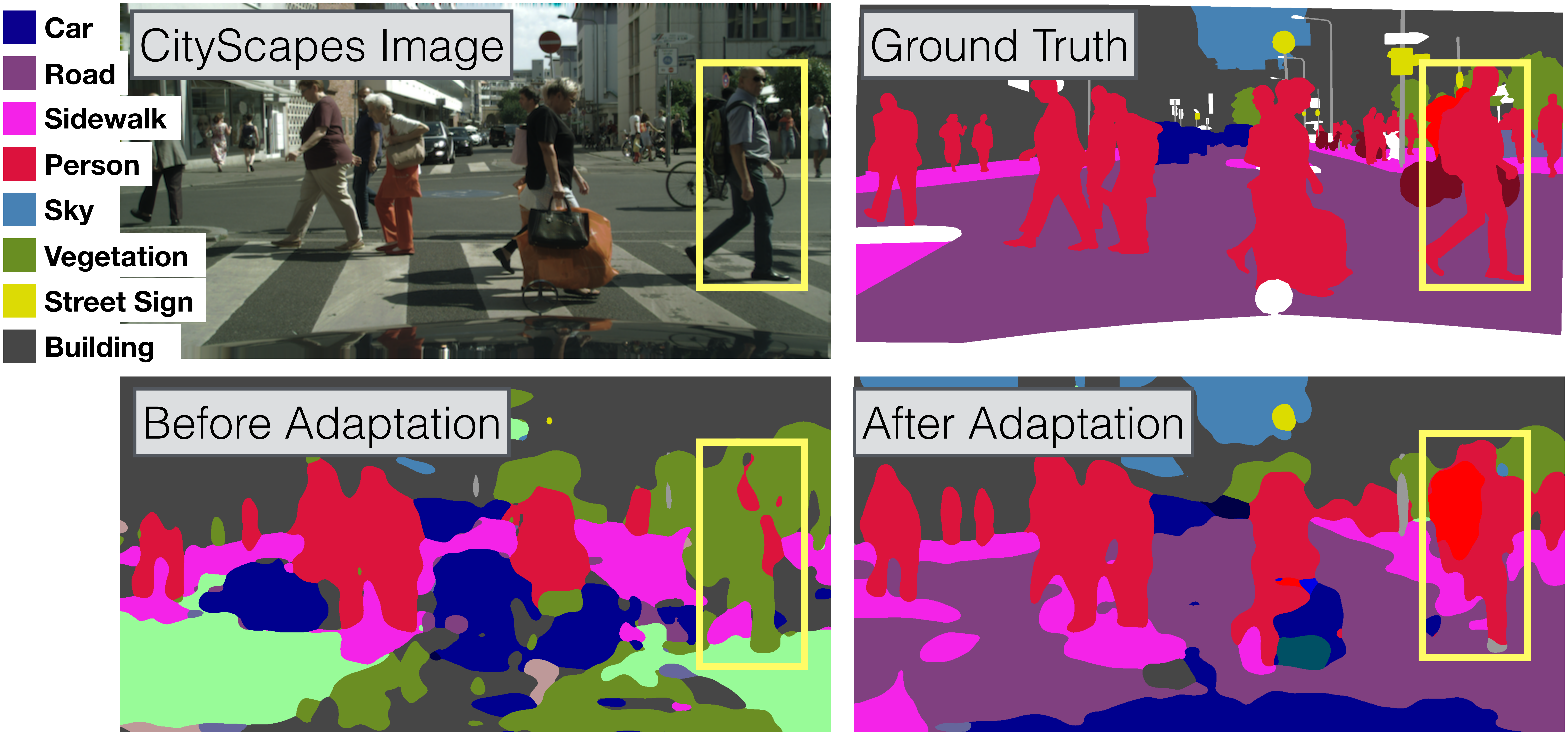
Ground Truth

Before Adaptation

After Adaptation

Hoffman et.al. ICML 2018

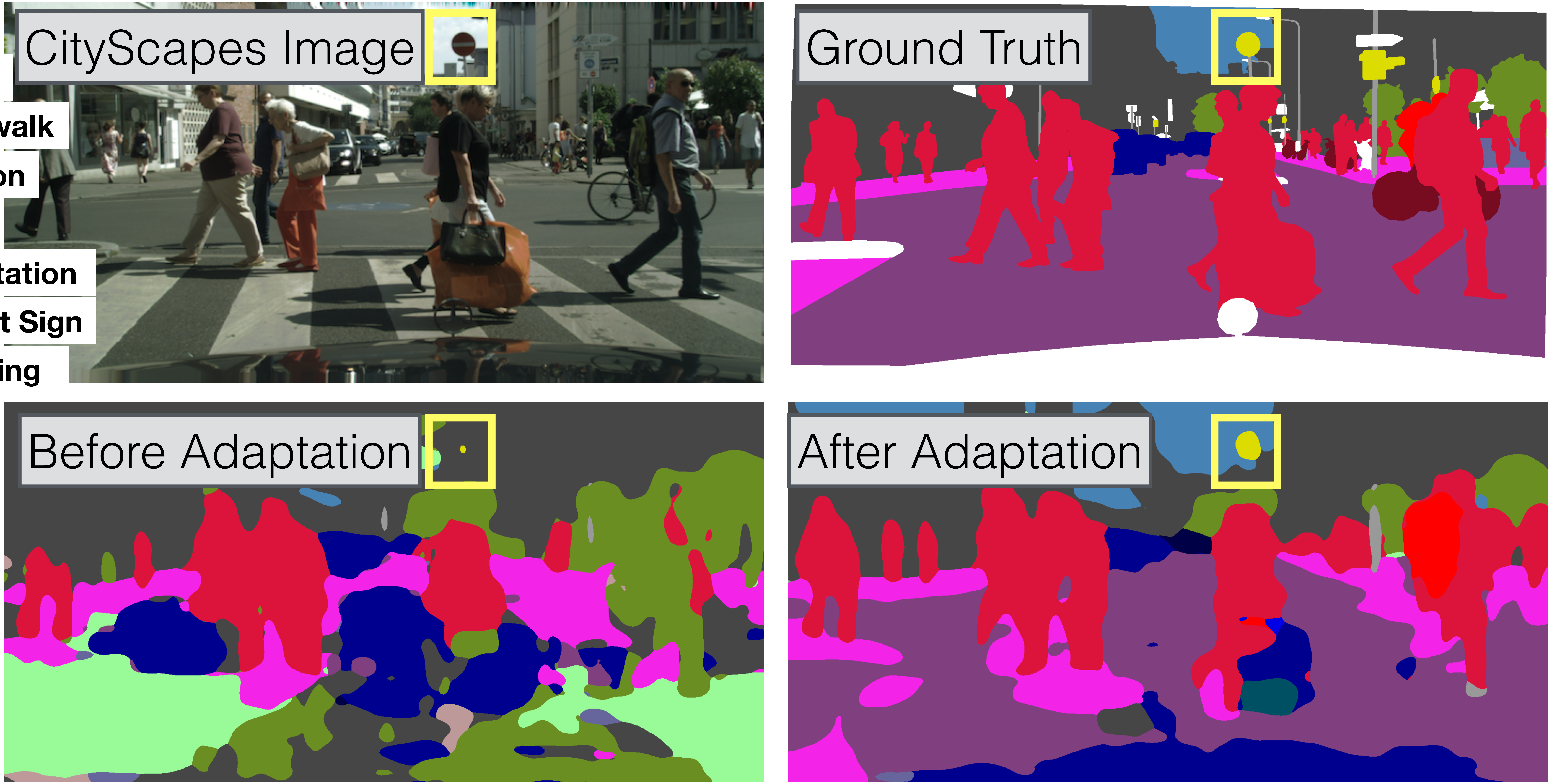
CyCADA Results: CityScapes Evaluation



Hoffman et.al. ICML 2018

CyCADA Results: CityScapes Evaluation

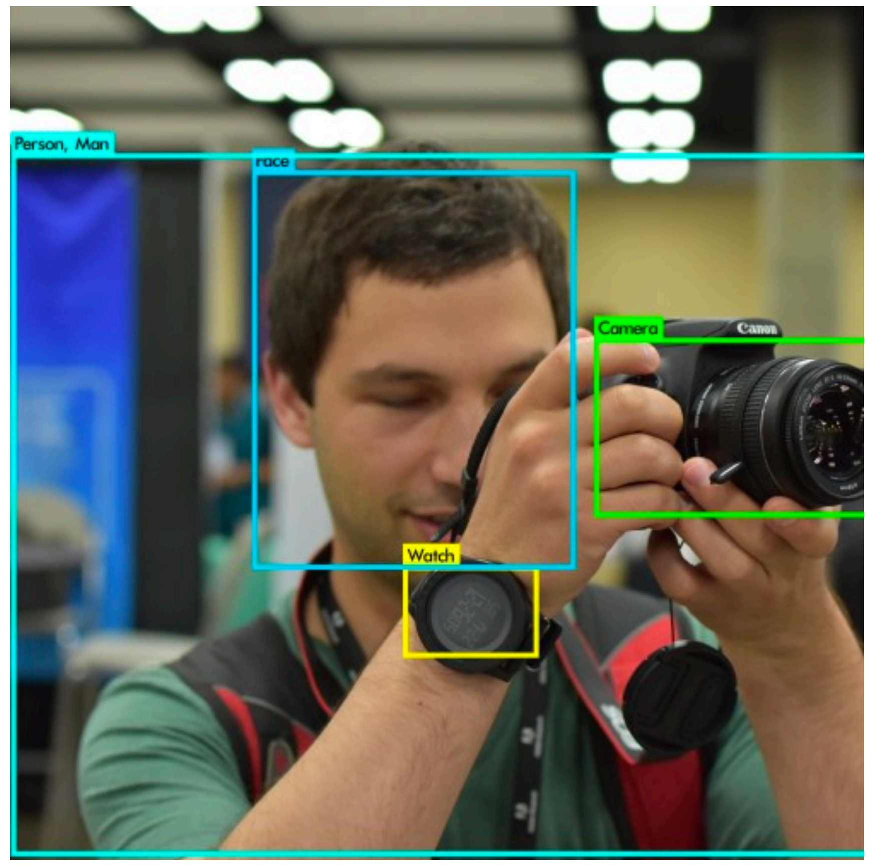
- 🟦 Car
- 🟪 Road
- 🟪 Sidewalk
- 🟥 Person
- 🟦 Sky
- 🟩 Vegetation
- 🟨 Street Sign
- ⬛ Building



Hoffman et.al. ICML 2018

Transfer for Embodied Tasks

SplitNet: Sim2Sim and Task2Task Transfer for Embodied Visual Navigation



Daniel Gordon
University of Washington

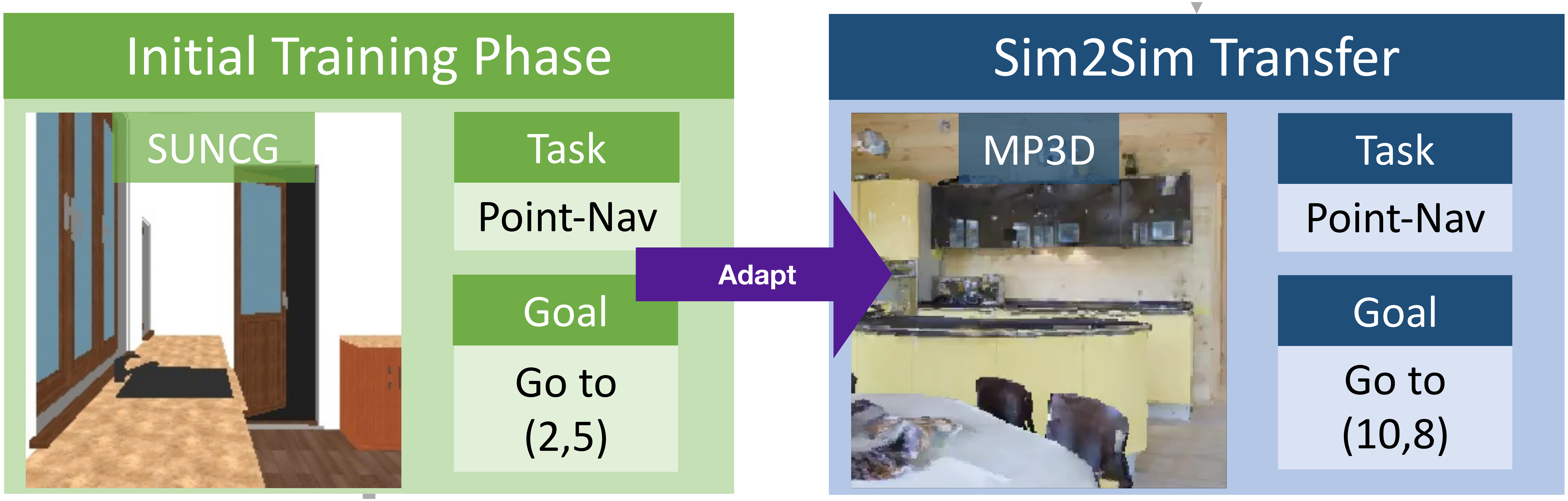
Abhishek Kadian
FAIR

Devi Parikh
Georgia Tech / FAIR

Dhruv Batra
Georgia Tech / FAIR

Gordon, Kadian, Parikh, Hoffman, Batra. arXiv 2019.

Sim2Sim Transfer



Initial Training Phase

SUNCG

Task

Point-Nav

Goal

Go to (2,5)

Adapt

Sim2Sim Transfer

MP3D

Task

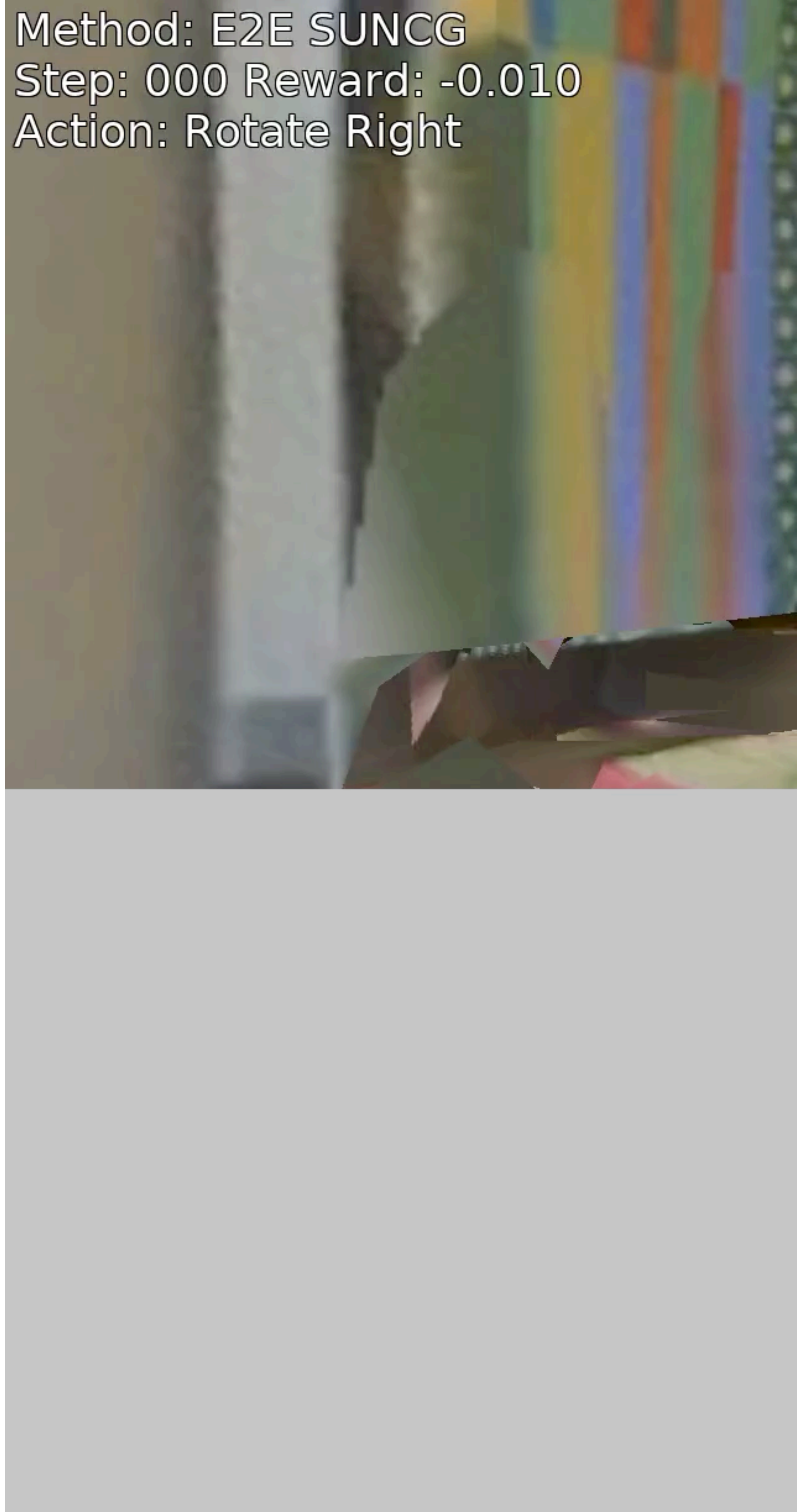
Point-Nav

Goal

Go to (10,8)

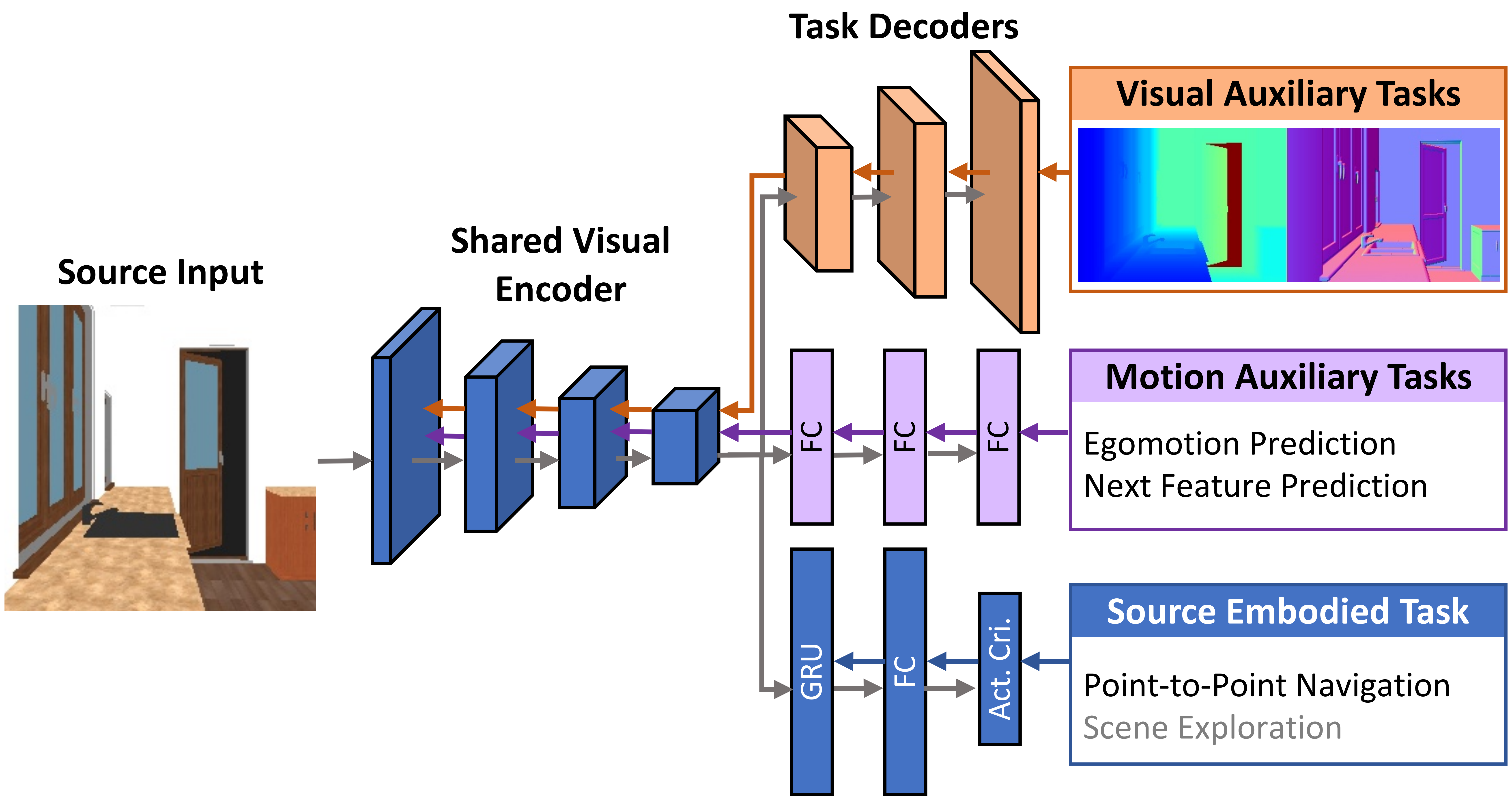
Gordon et. al., arXiv 2019

Train SUNCG
Test MP3D



Method: E2E SUNCG
Step: 000 Reward: -0.010
Action: Rotate Right

Source Training



Task Decoders

Source Input

Shared Visual Encoder

Visual Auxiliary Tasks

Motion Auxiliary Tasks

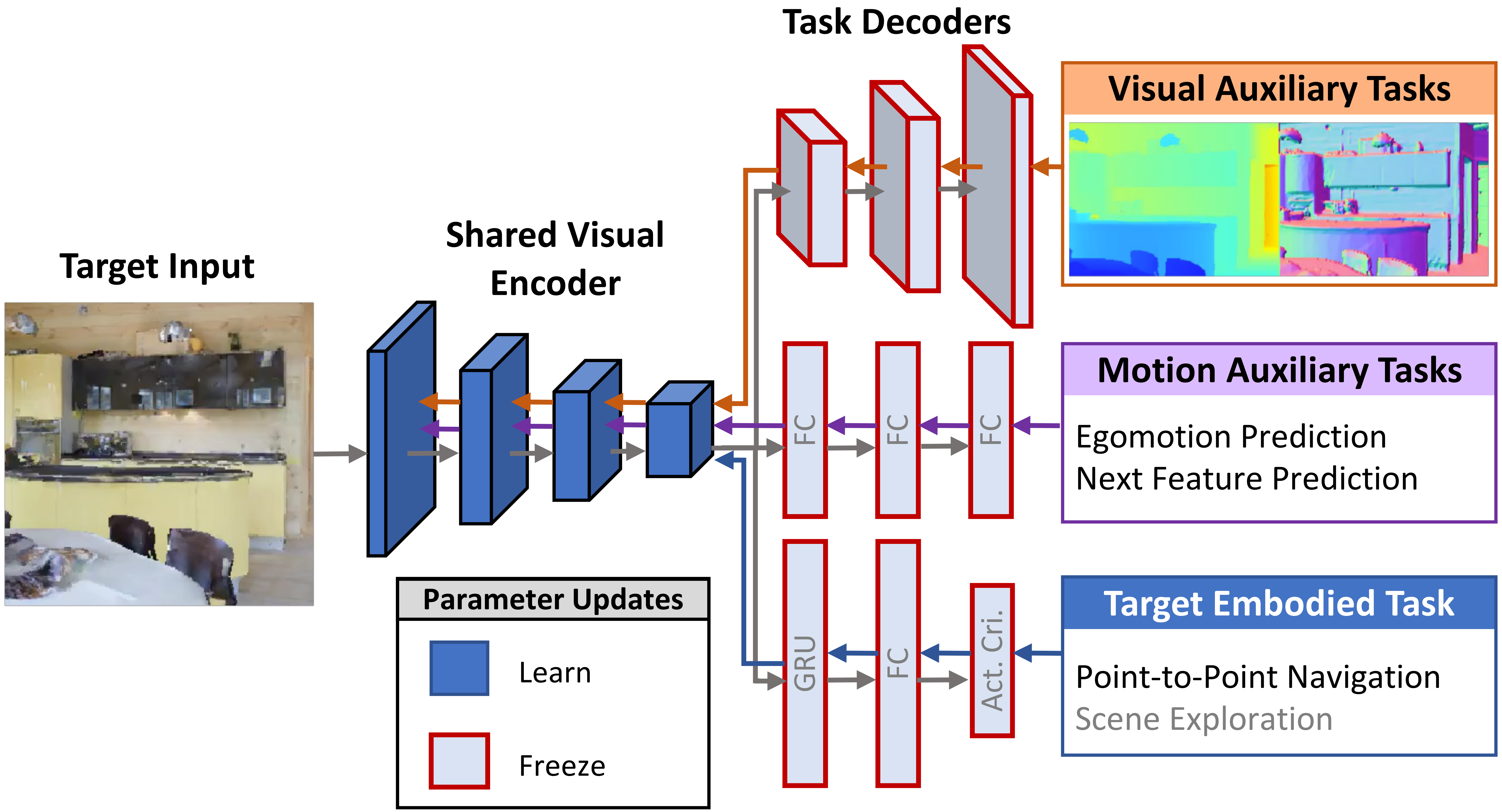
Egomotion Prediction
Next Feature Prediction

Source Embodied Task

Point-to-Point Navigation
Scene Exploration

Gordon et. al., arXiv 2019

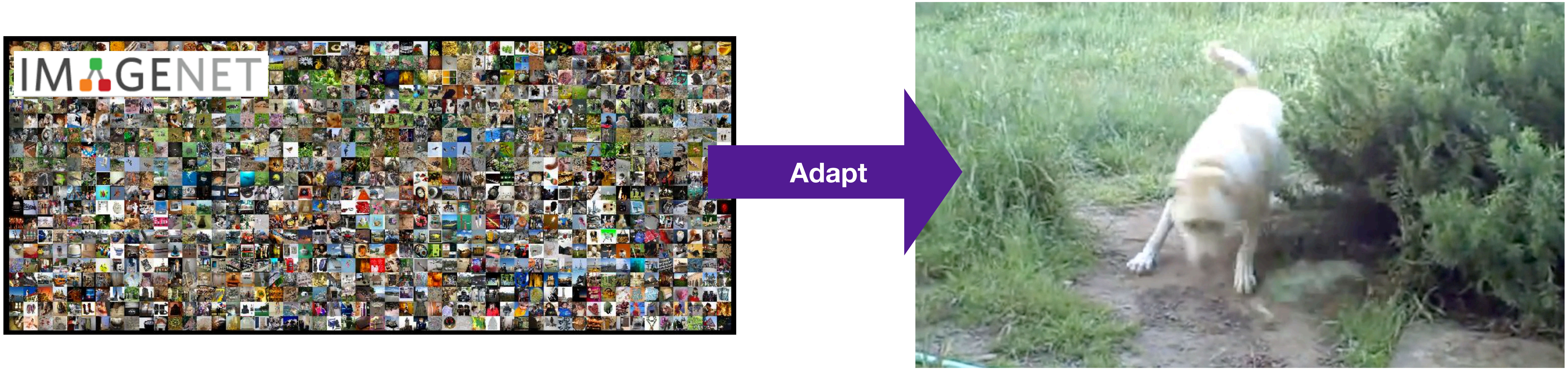
Sim2Sim Transfer



Gordon et. al., arXiv 2019

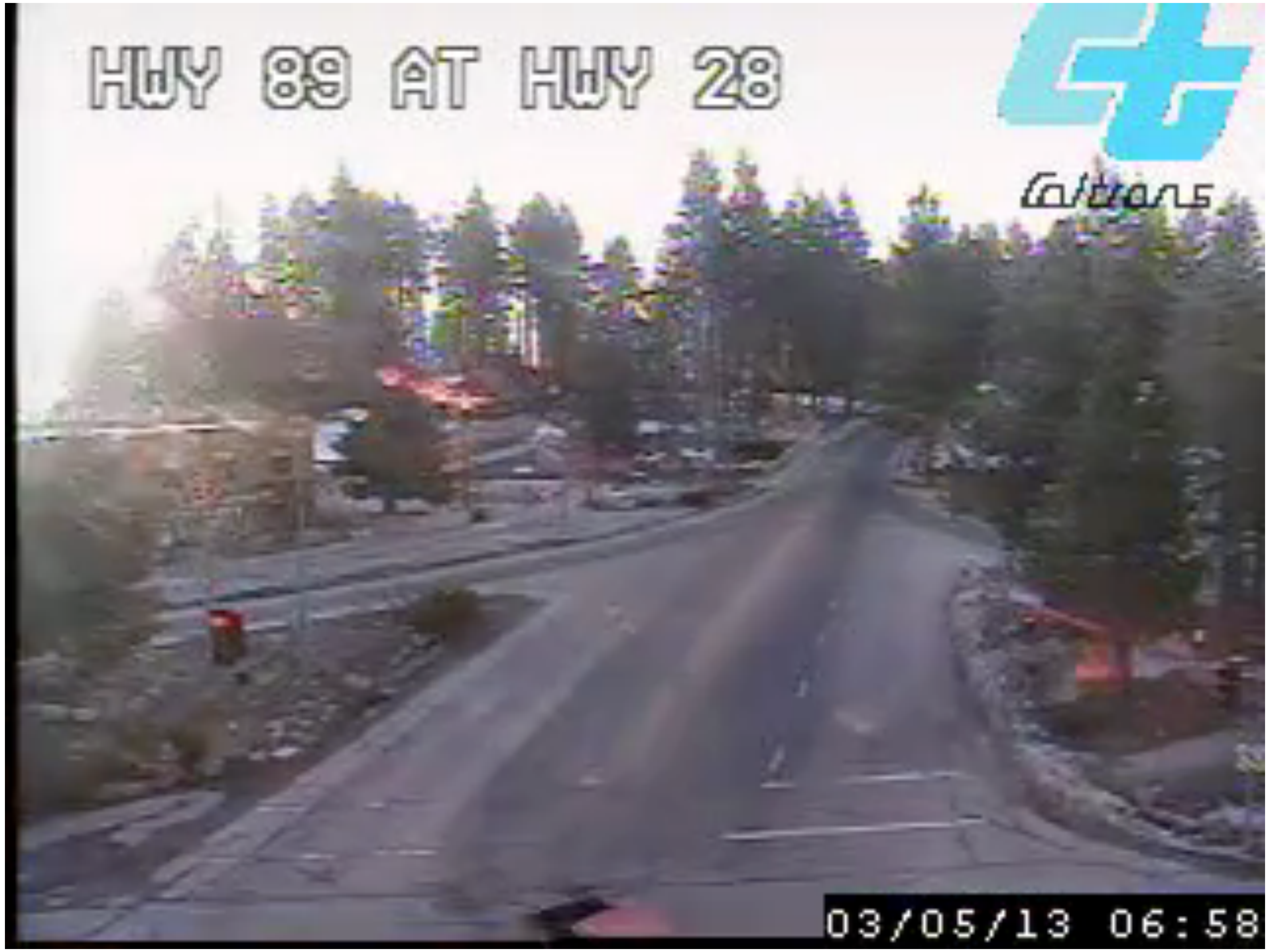
SIM2SIM - COMPARING MODELS WITH DIFFERENT AMOUNTS OF DATA FROM TARGET

Adapting from one data source to another



http://data-bdd.berkeley.edu/

Continuous Learning



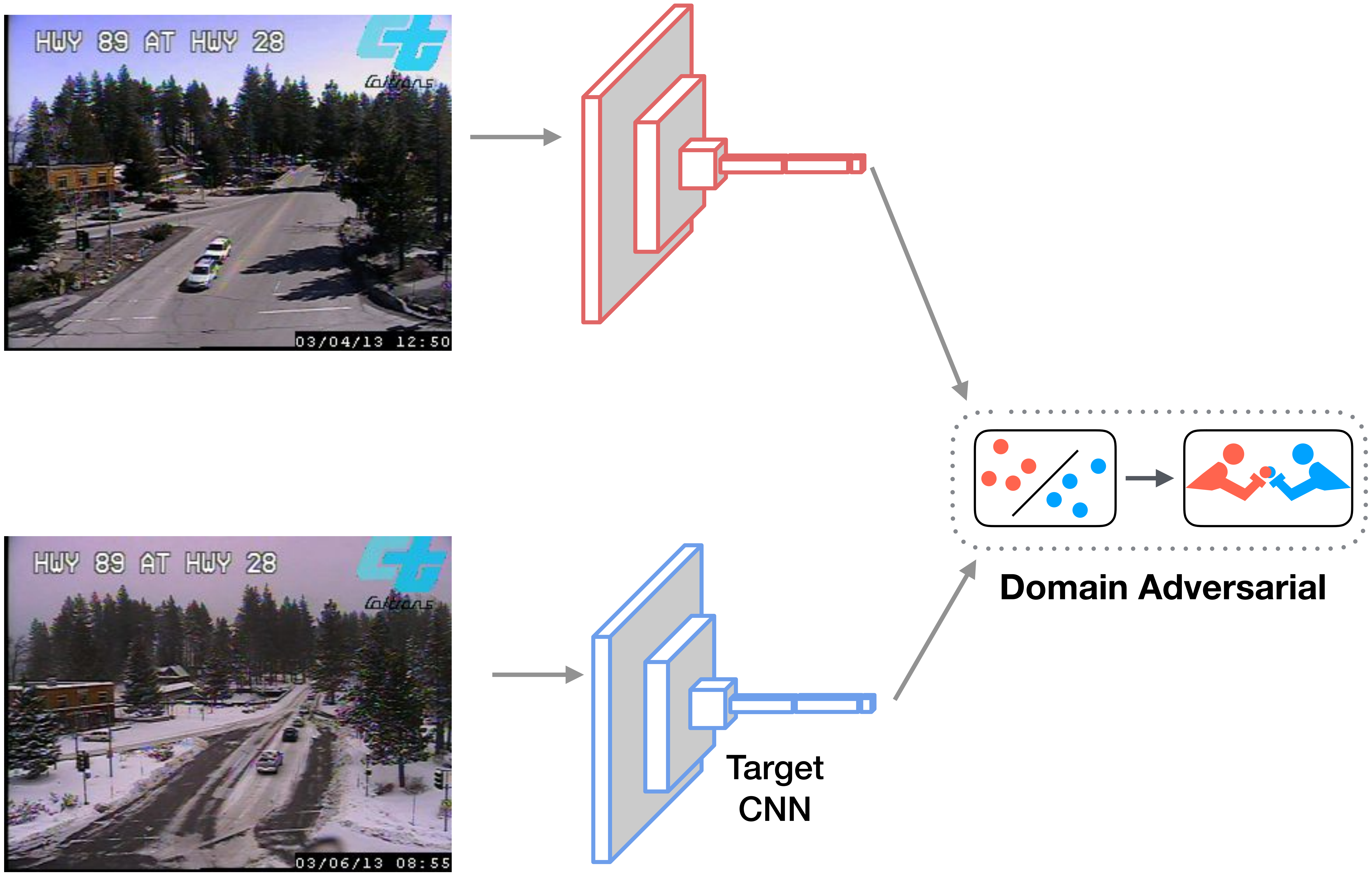
Continuous Manifold Based Alignment.
Hoffman, Darrell, Saenko, CVPR 14

Continuous Learning

Generalizable

Sample Efficient

Scalable

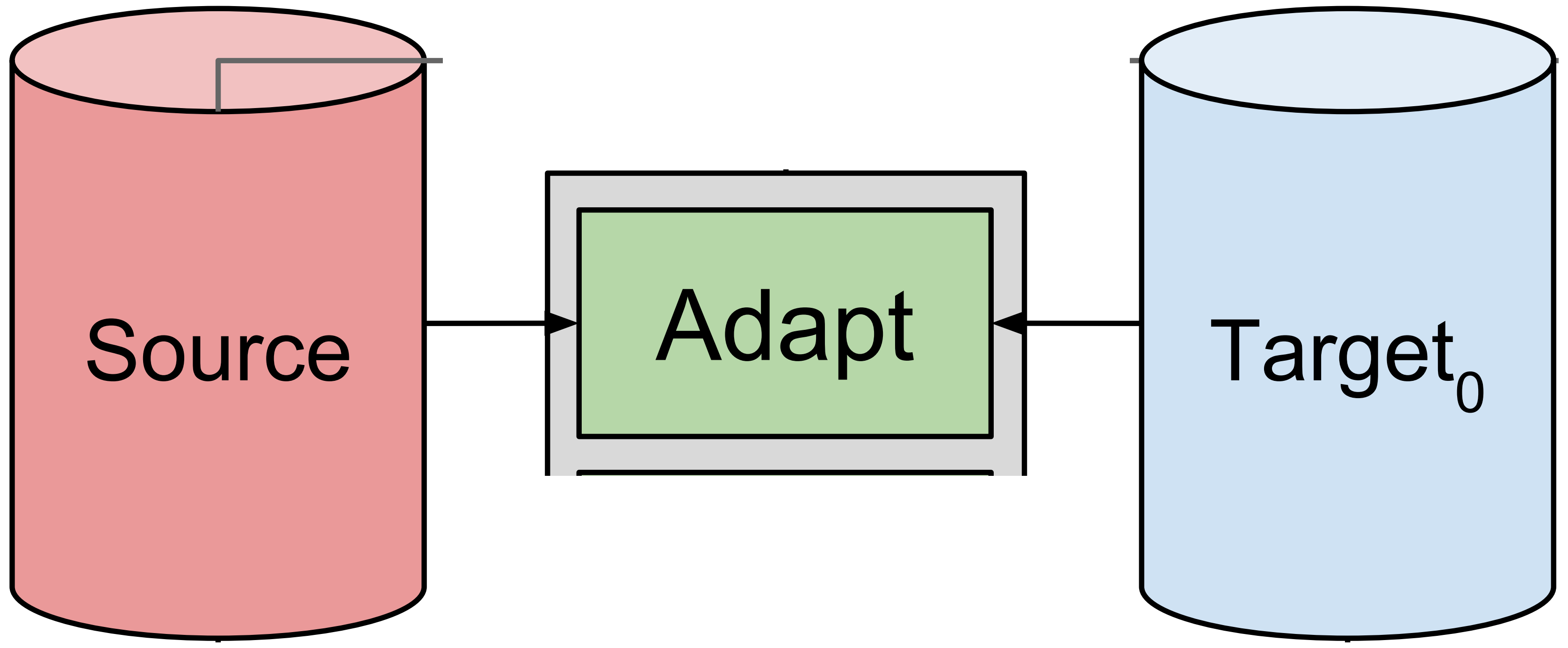


Continuous Unsupervised Adaptation

Source
Target$_0$

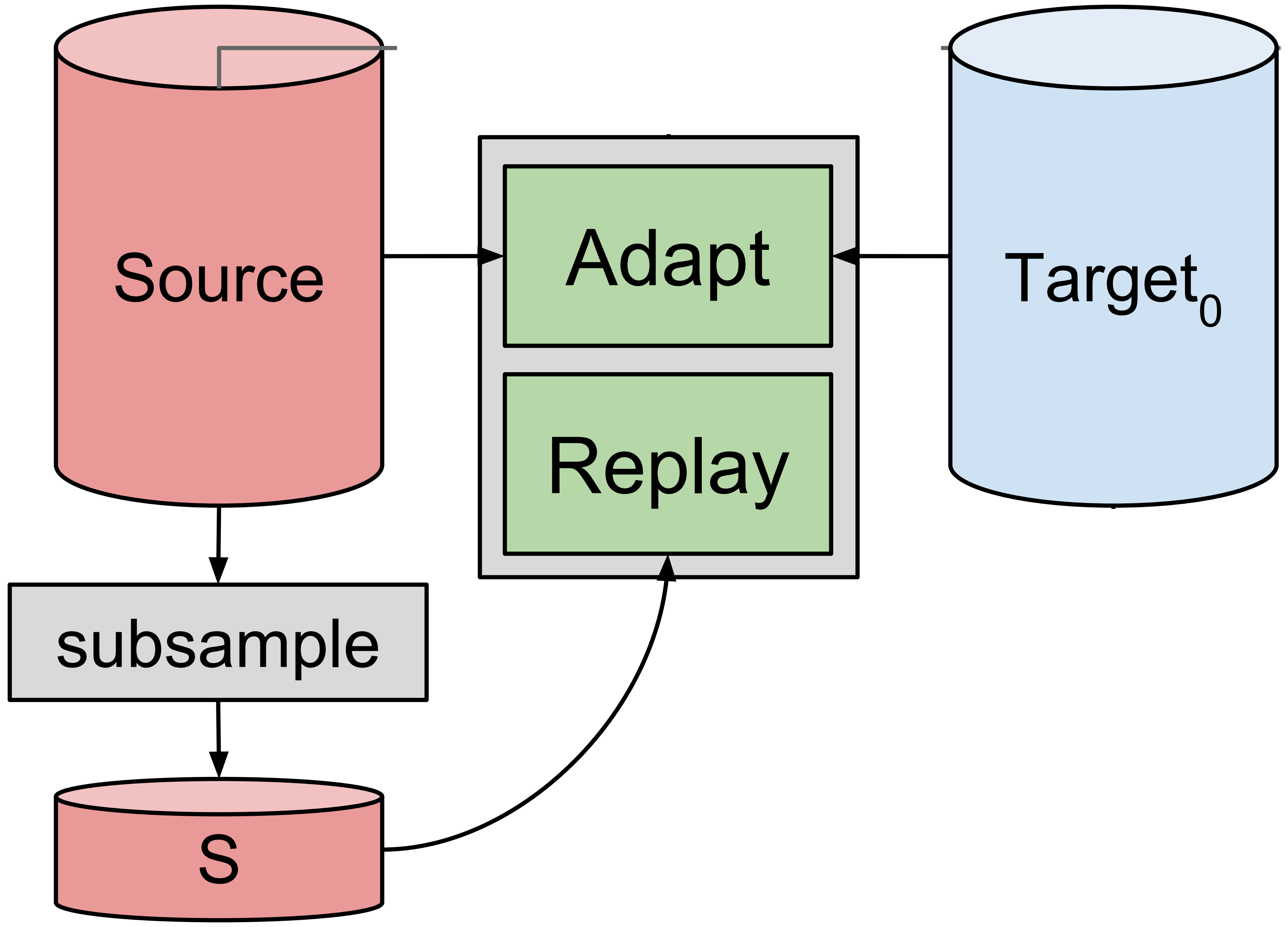
Bobu, Tzeng, Hoffman, Darrell. ICLR Workshop 2018.

Continuous Unsupervised Adaptation



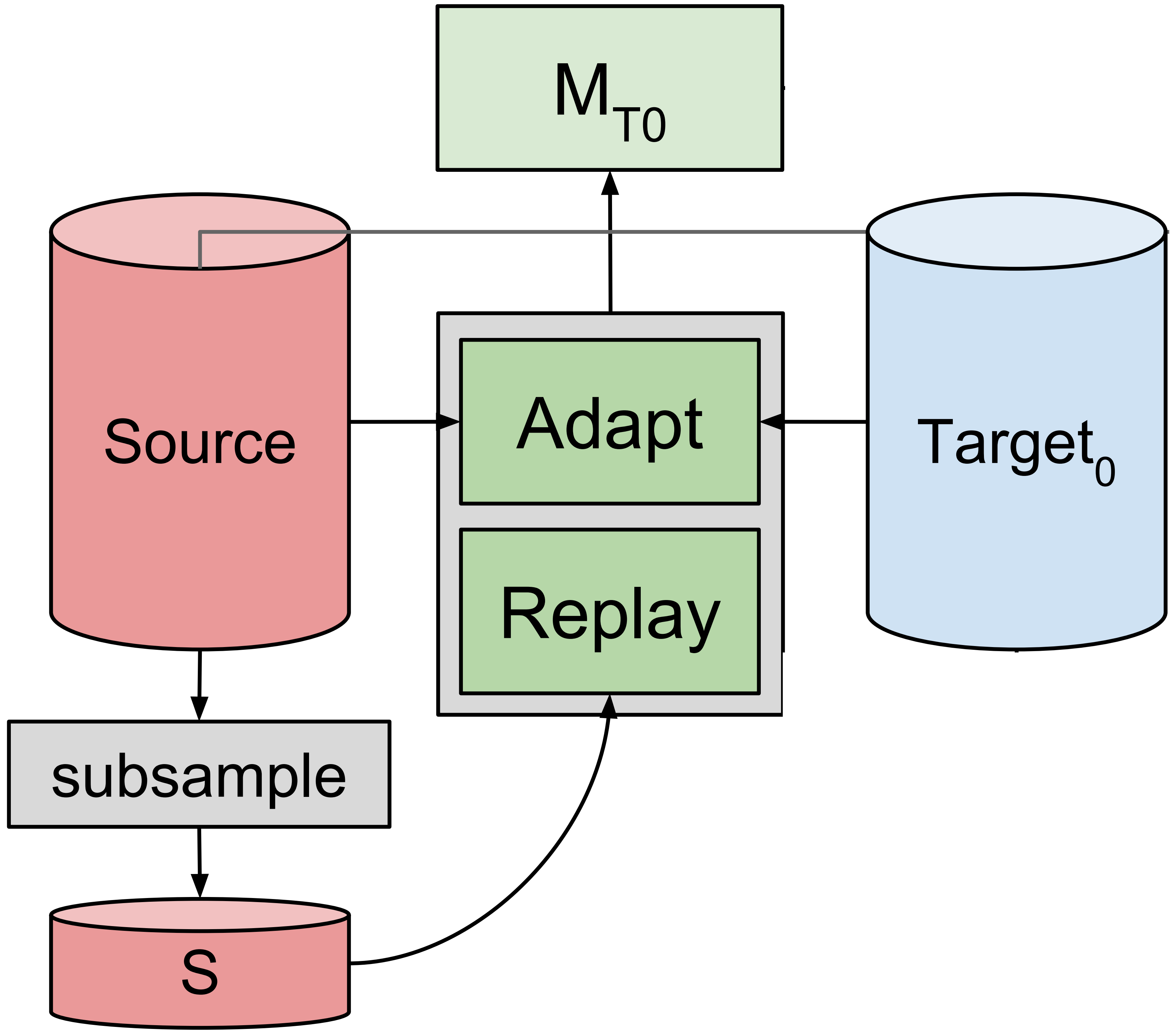
Bobu, Tzeng, Hoffman, Darrell. ICLR Workshop 2018.

Continuous Unsupervised Adaptation



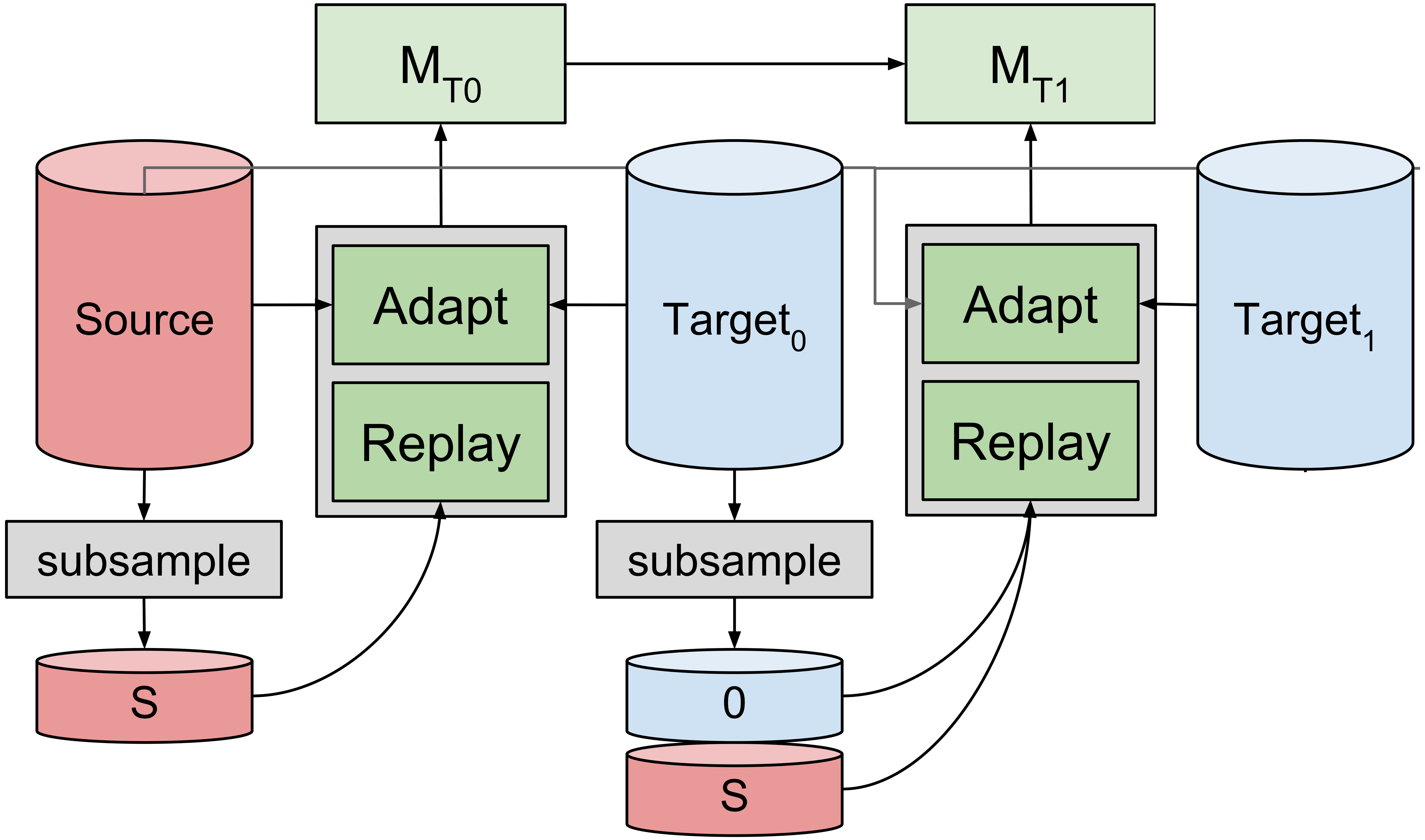
Bobu, Tzeng, Hoffman, Darrell. ICLR Workshop 2018.

Continuous Unsupervised Adaptation



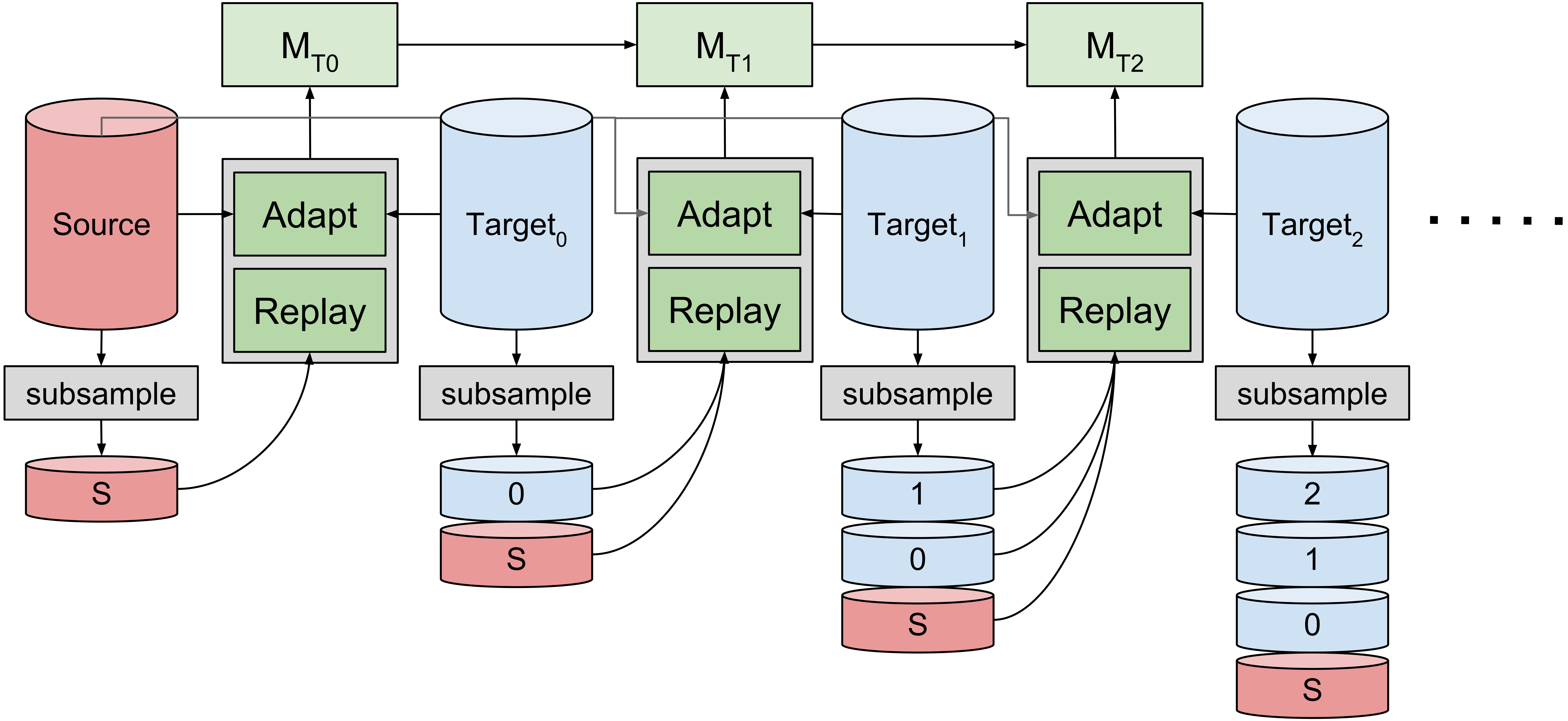
Bobu, Tzeng, Hoffman, Darrell. ICLR Workshop 2018.

Continuous Unsupervised Adaptation



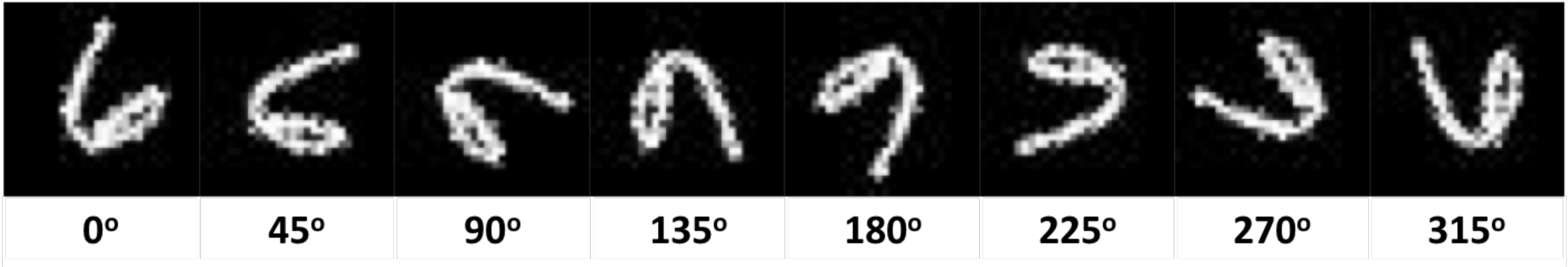
Bobu, Tzeng, Hoffman, Darrell. ICLR Workshop 2018.

Continuous Unsupervised Adaptation



Bobu, Tzeng, Hoffman, Darrell. ICLR Workshop 2018.

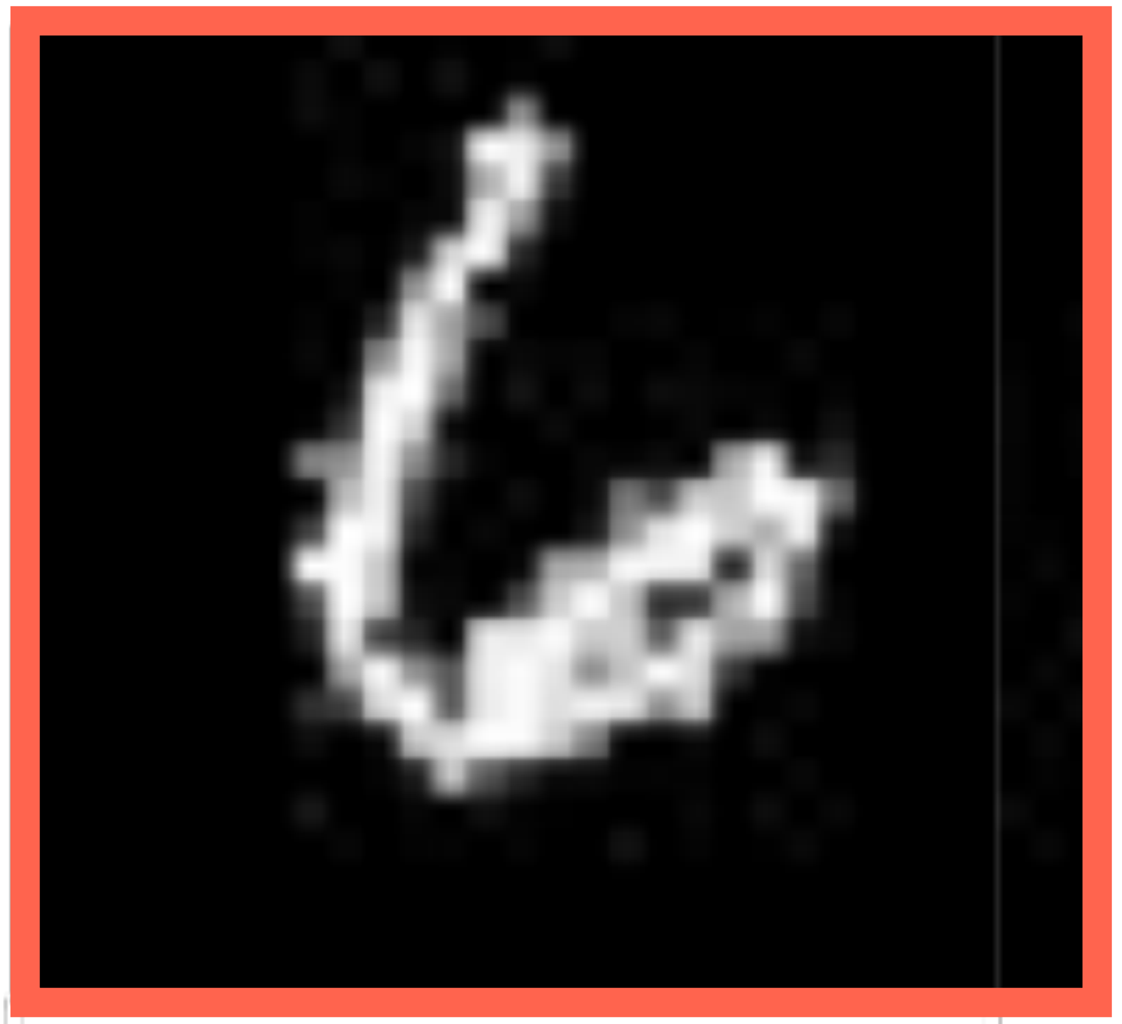
Experiment: MNIST Rotations



Bobu, Tzeng, Hoffman, Darrell. ICLR Workshop 2018.

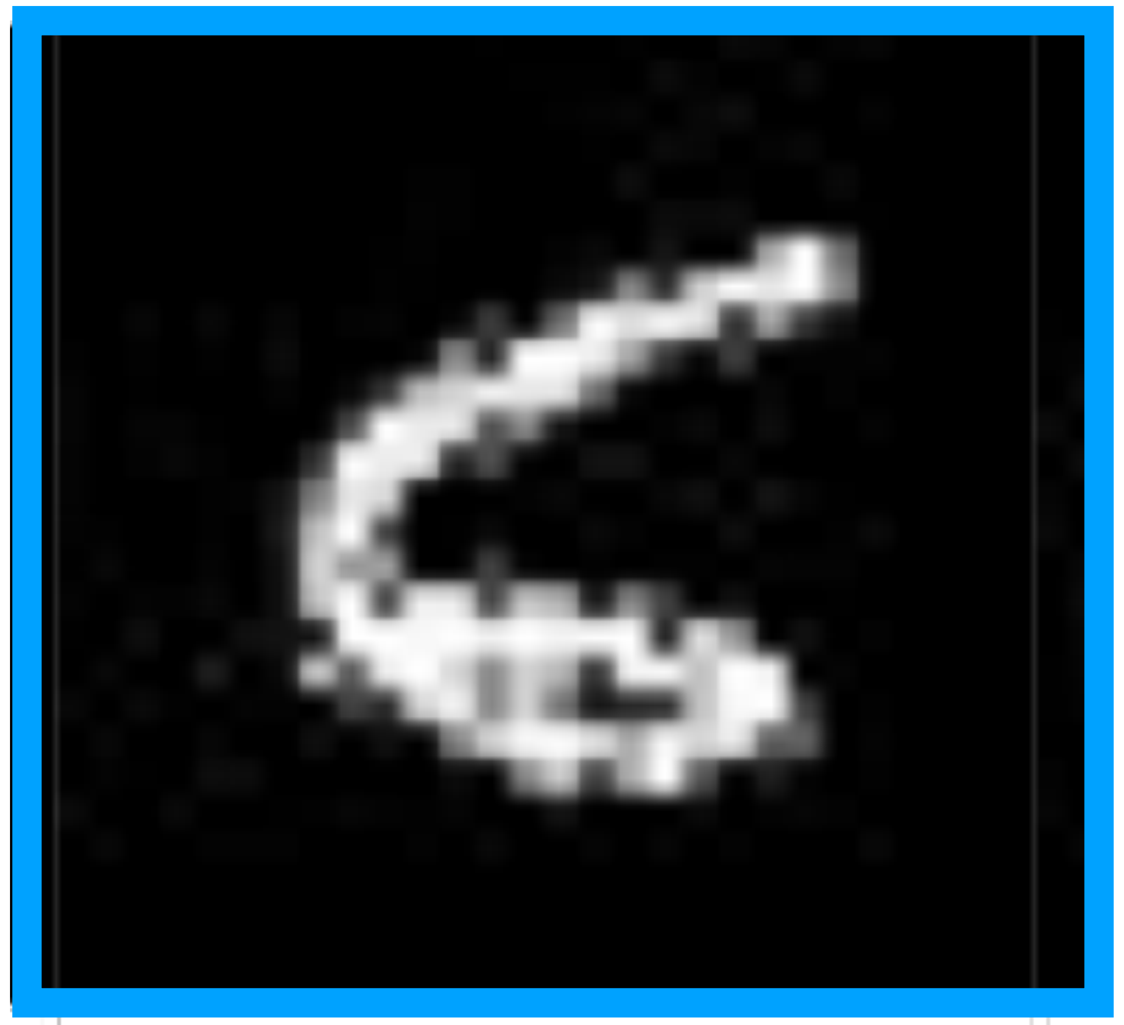
Experiment: MNIST Rotations

Labeled Source



0°

Unlabeled Target (test)

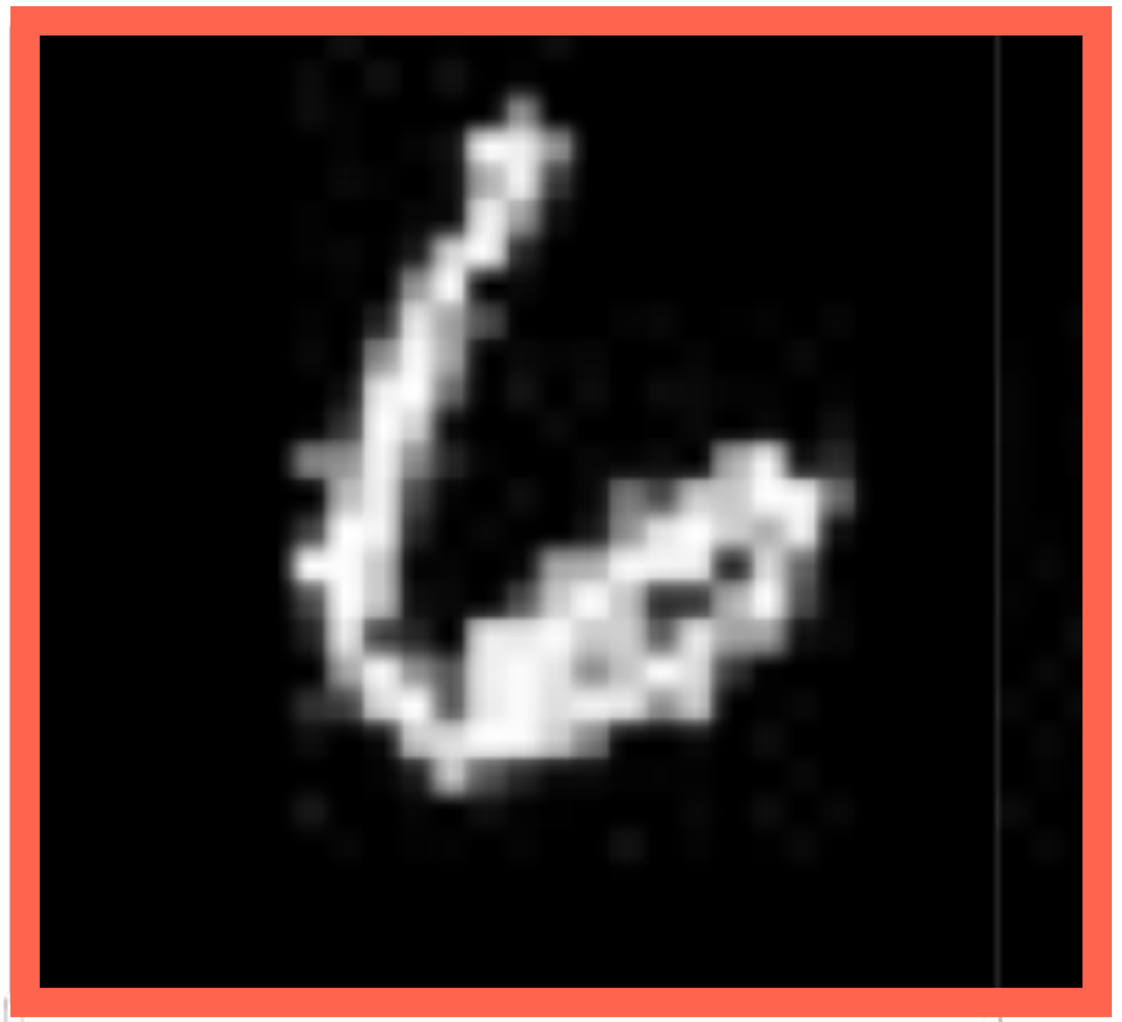


45°

Bobu, Tzeng, Hoffman, Darrell. ICLR Workshop 2018.

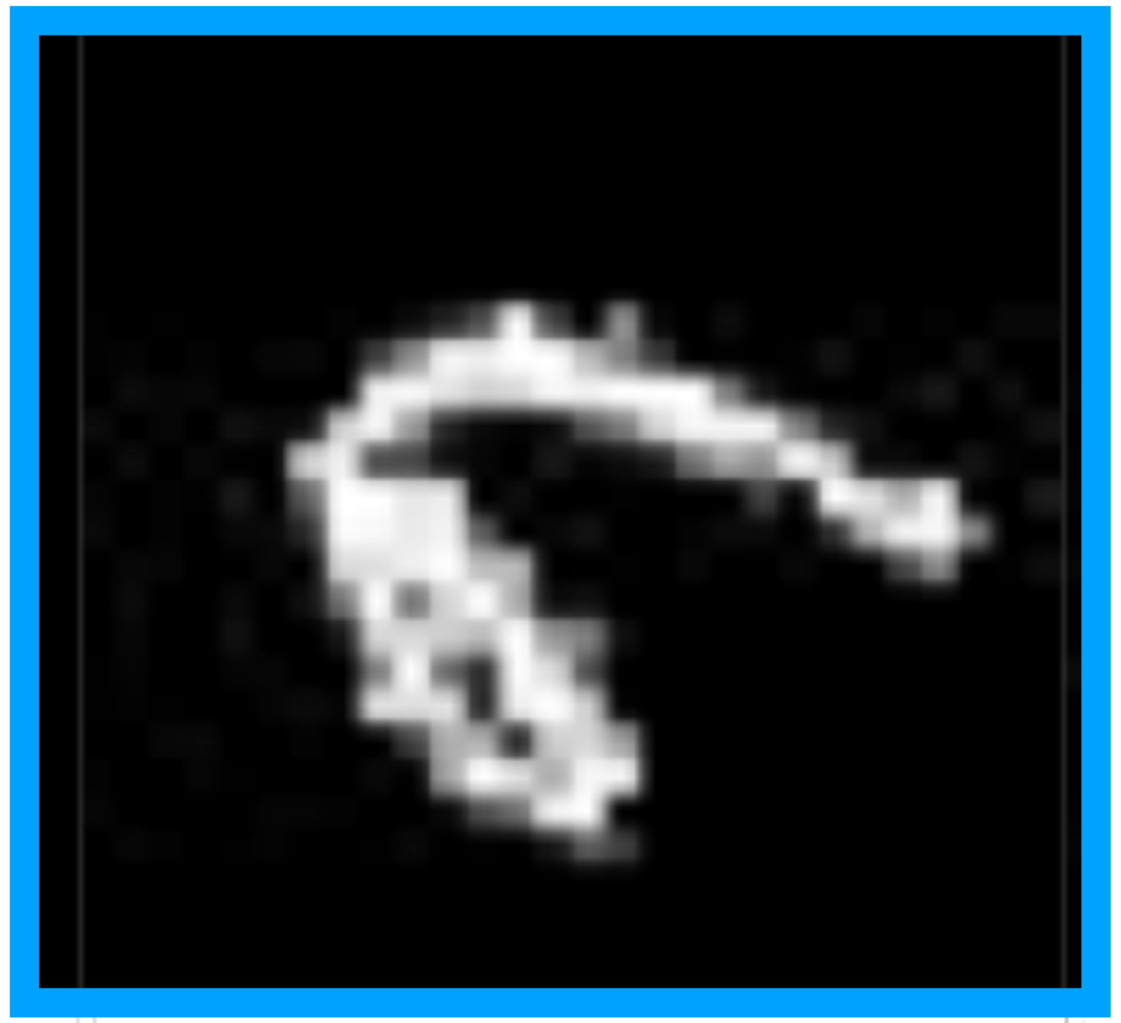
Experiment: MNIST Rotations

Labeled Source



0°

Unlabeled Target (test)

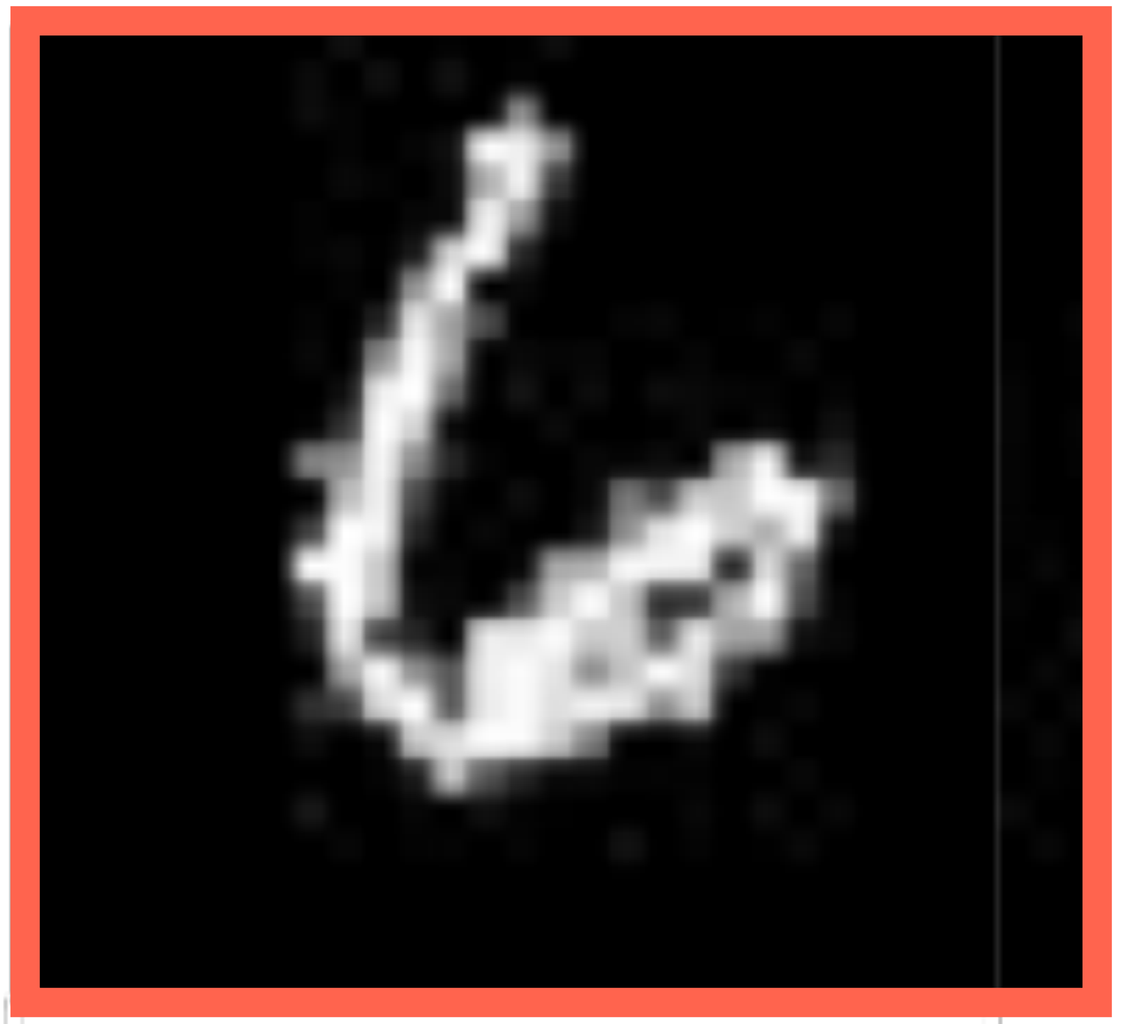


90°

Bobu, Tzeng, Hoffman, Darrell. ICLR Workshop 2018.

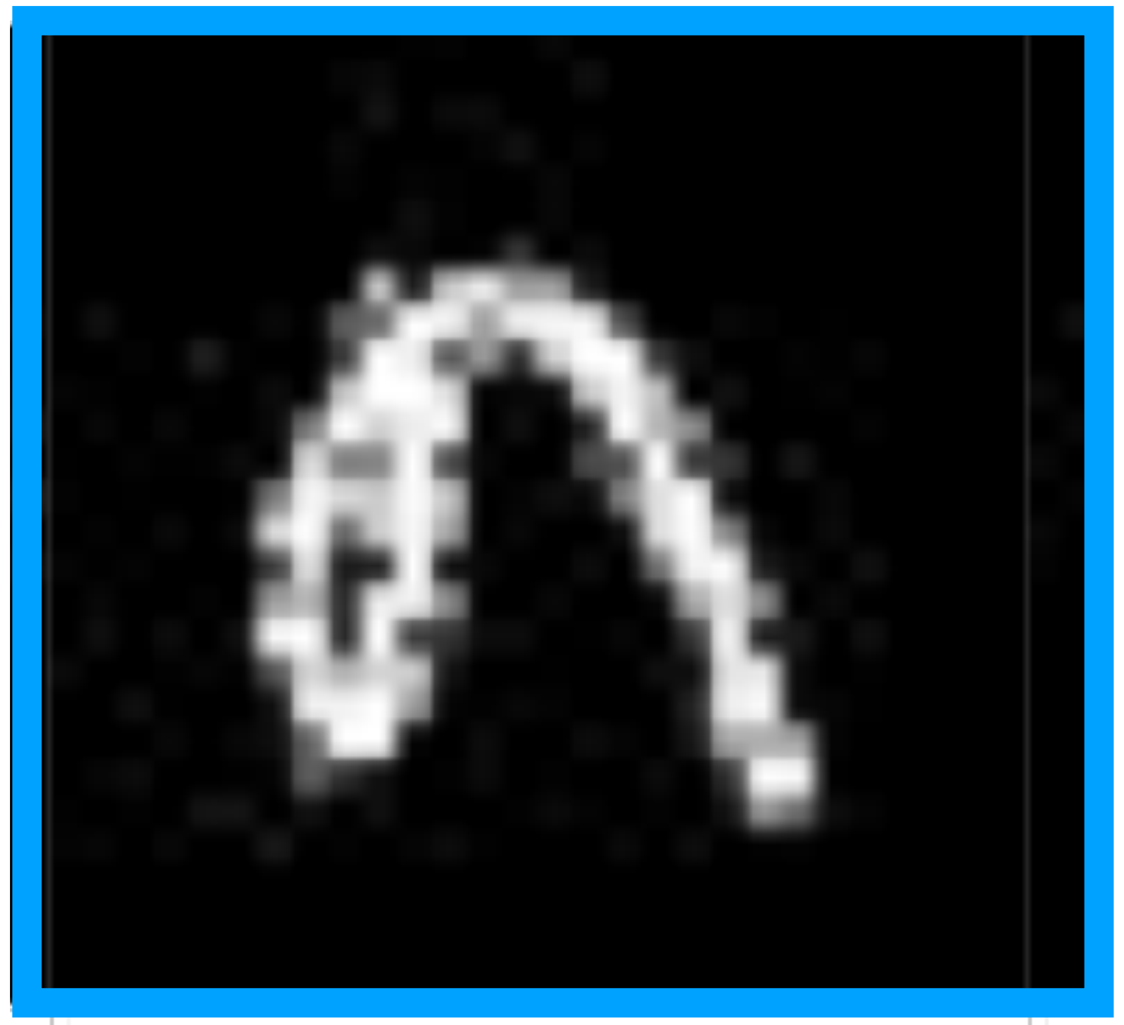
Experiment: MNIST Rotations

Labeled Source



0°

Unlabeled Target (test)



135°

Bobu, Tzeng, Hoffman, Darrell. ICLR Workshop 2018.

Experiment: MNIST Rotations

Labeled Source

0°

Unlabeled Target (test)

180°

Bobu, Tzeng, Hoffman, Darrell. ICLR Workshop 2018.

Experiment: MNIST Rotations

Labeled Source

0°

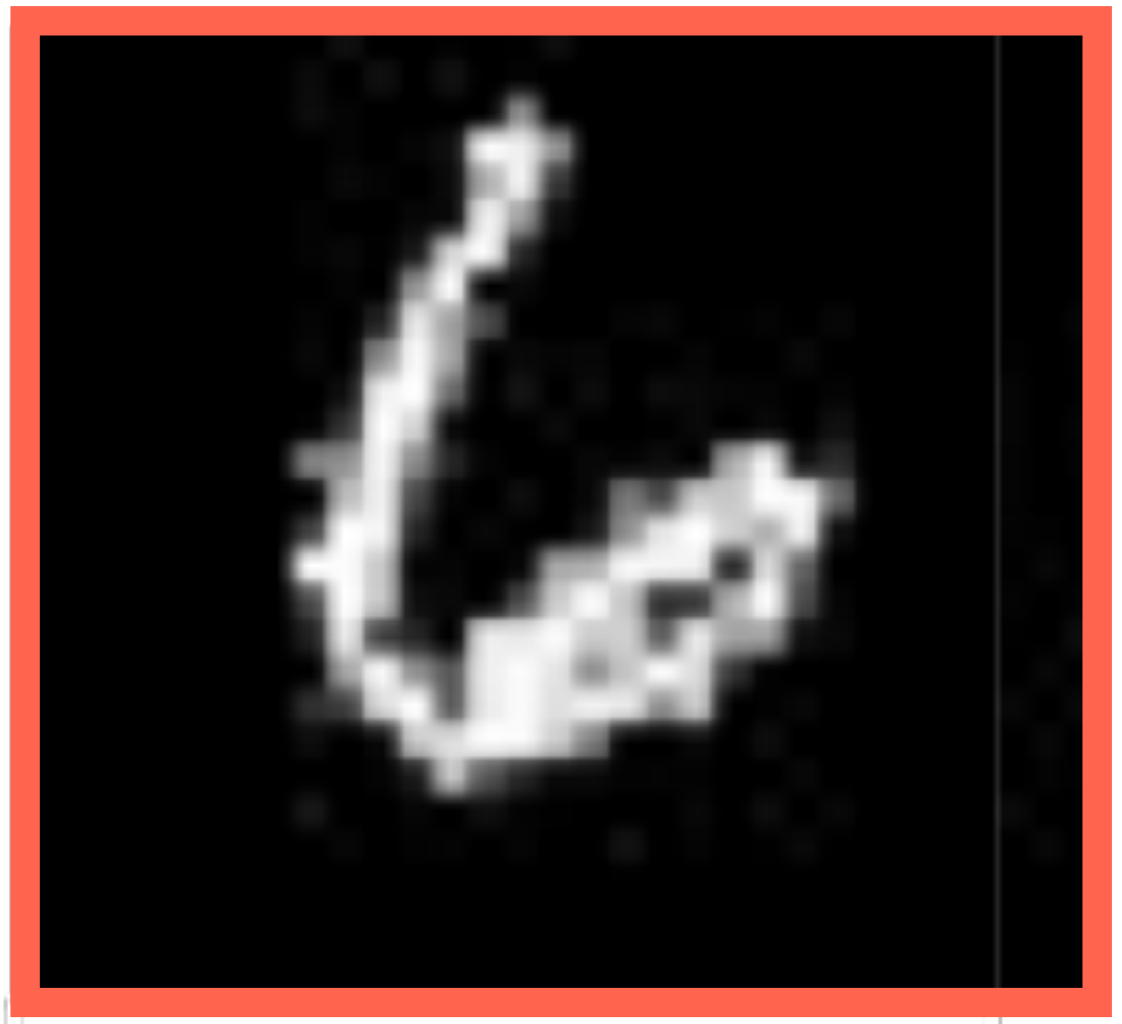
Unlabeled Target (test)

225°

Bobu, Tzeng, Hoffman, Darrell. ICLR Workshop 2018.

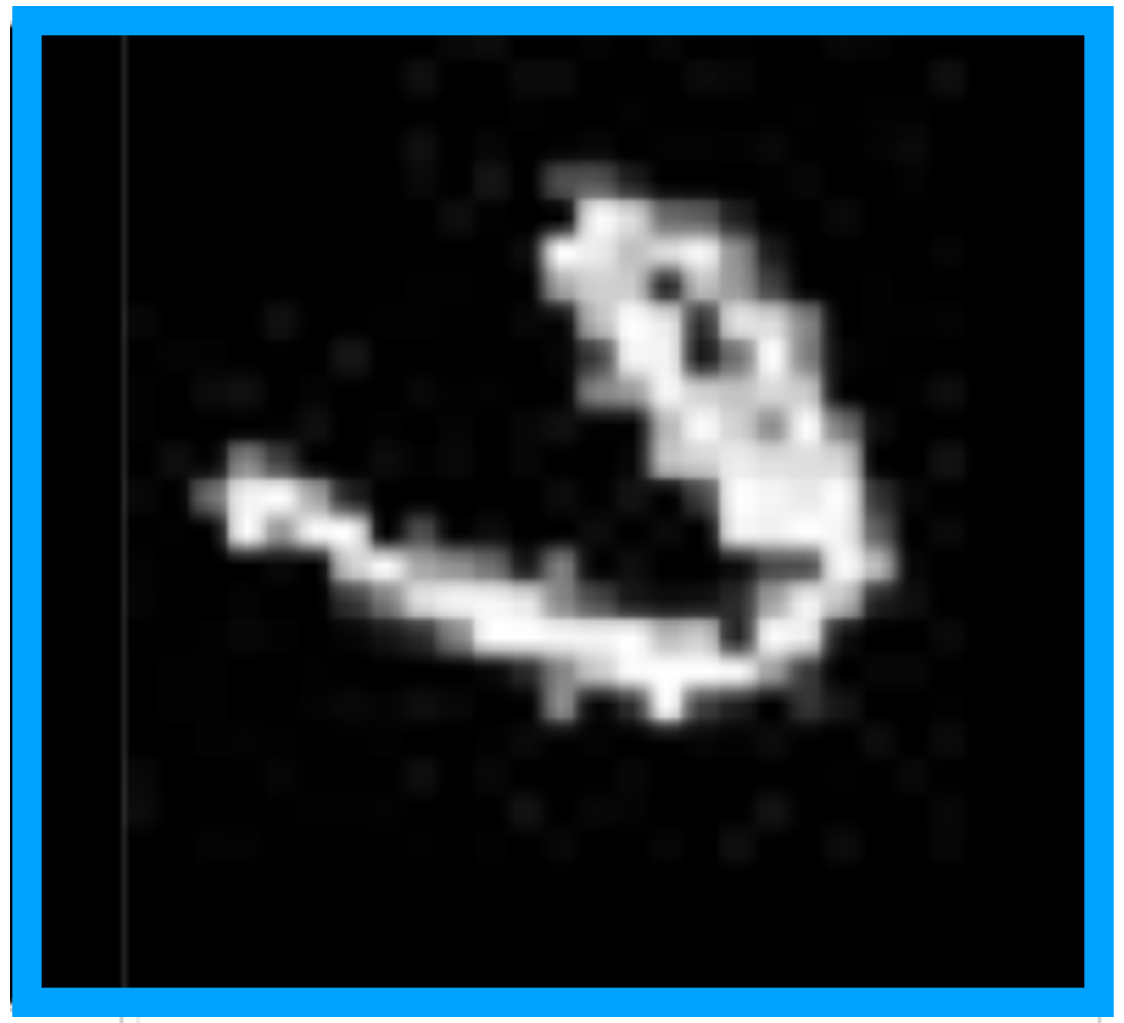
Experiment: MNIST Rotations

Labeled Source



0°

Unlabeled Target (test)

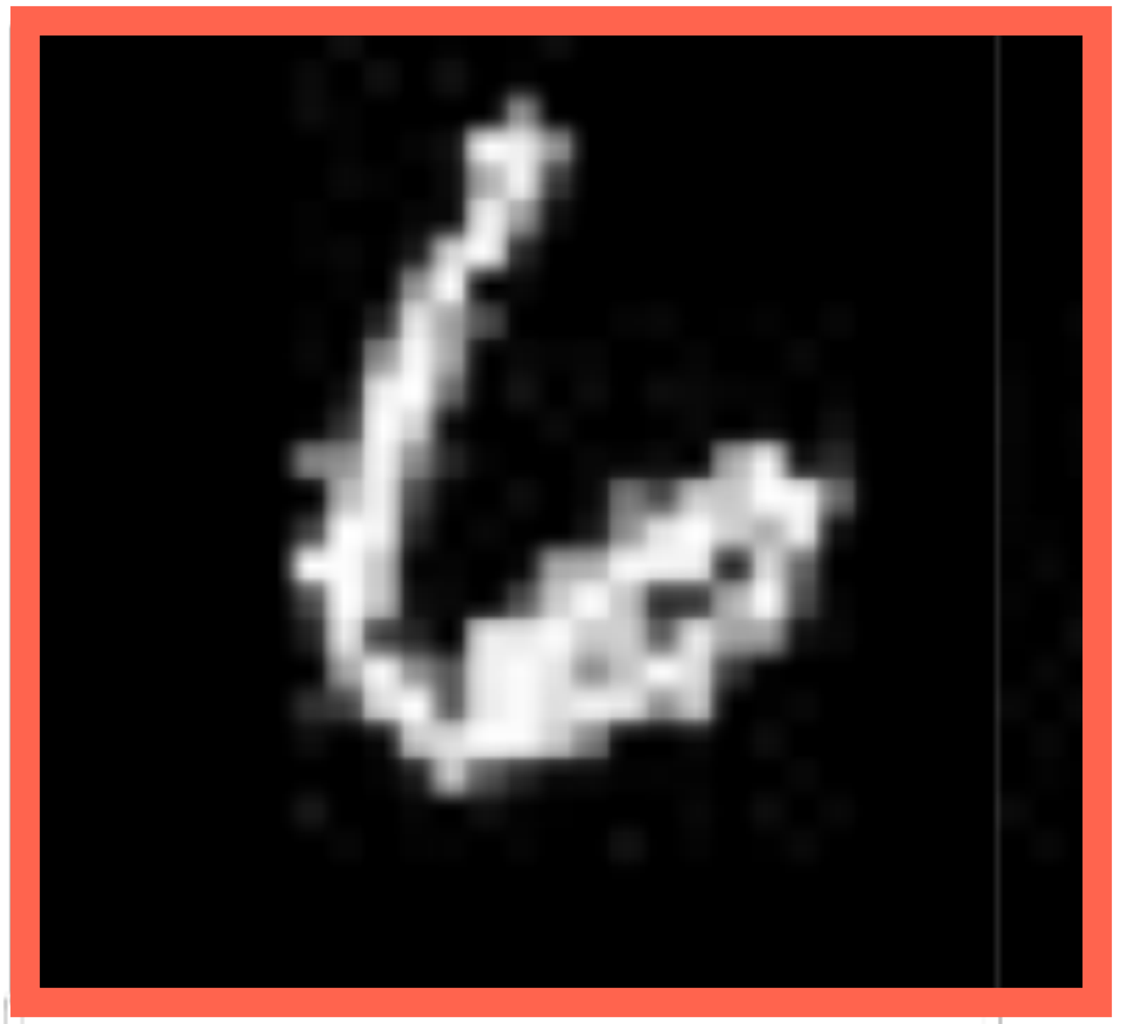


270°

Bobu, Tzeng, Hoffman, Darrell. ICLR Workshop 2018.

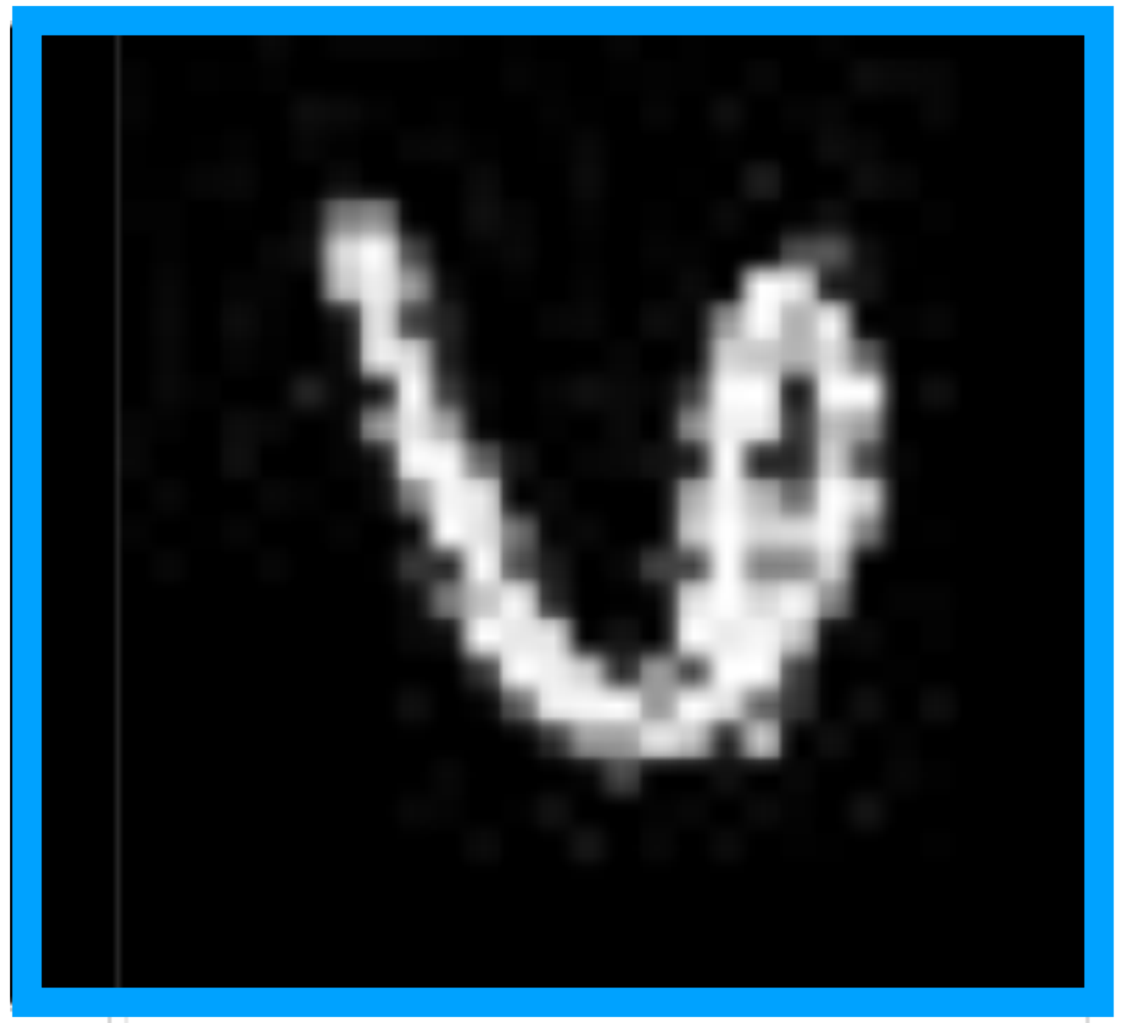
Experiment: MNIST Rotations

Labeled Source



0°

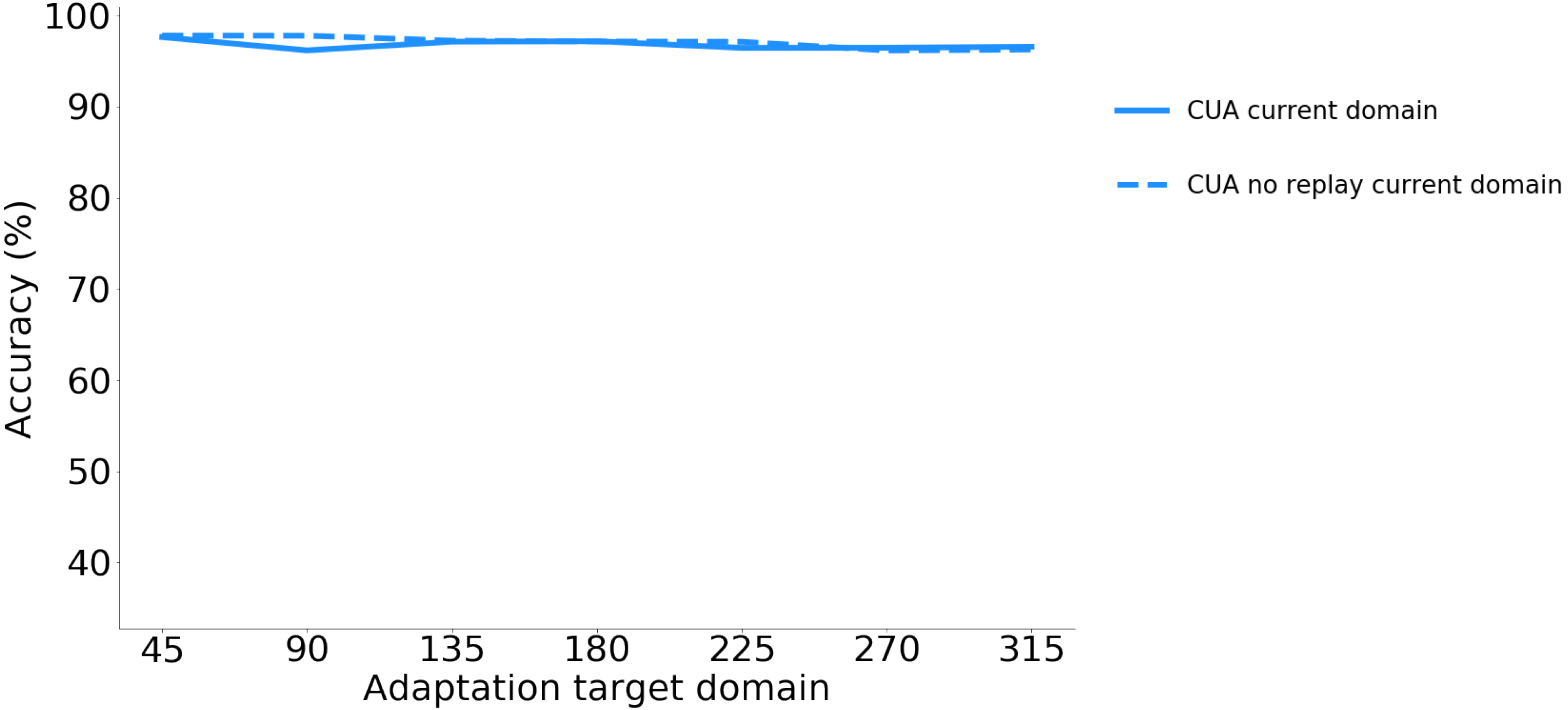
Unlabeled Target (test)



315°

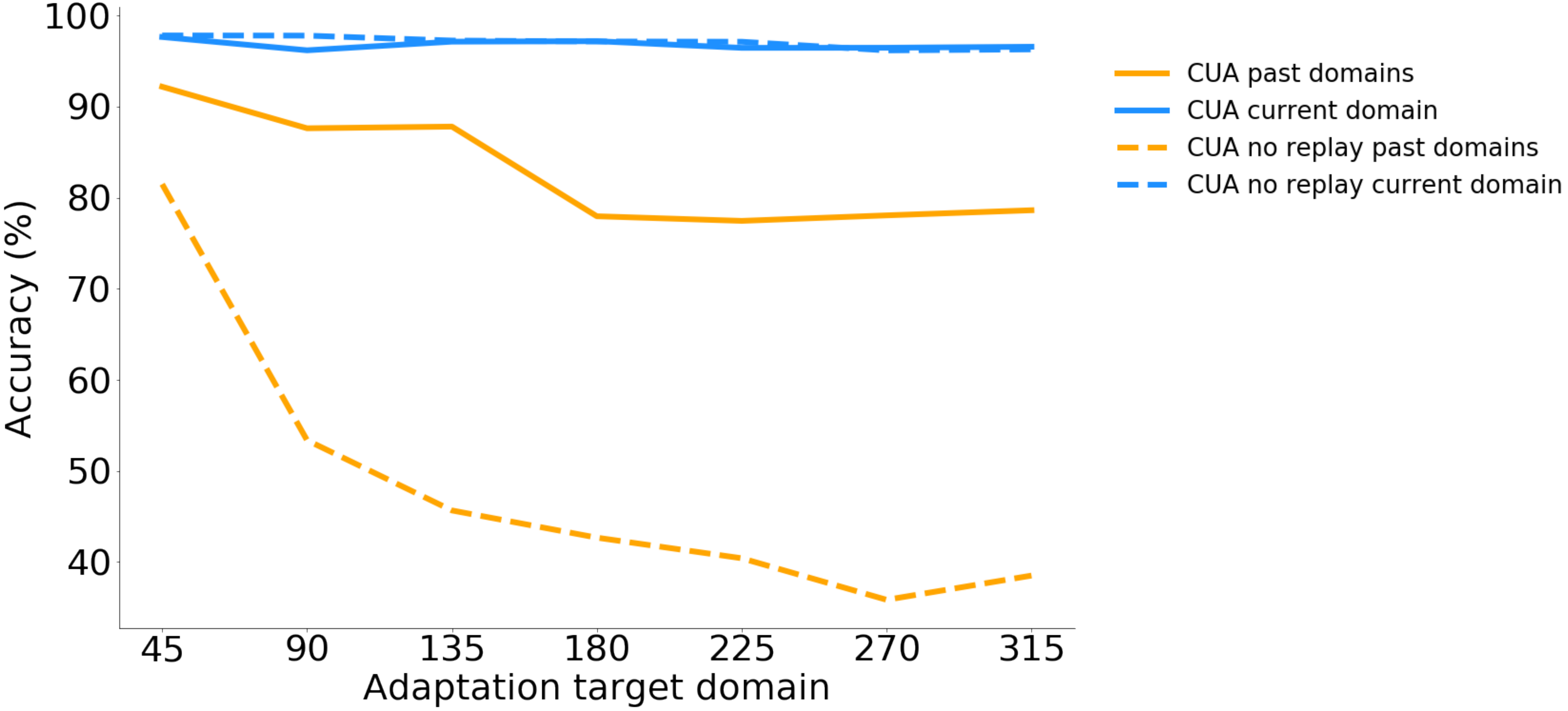
Bobu, Tzeng, Hoffman, Darrell. ICLR Workshop 2018.

Replay to Remember: MNIST Rotations



Bobu, Tzeng, Hoffman, Darrell. ICLR Workshop 2018.

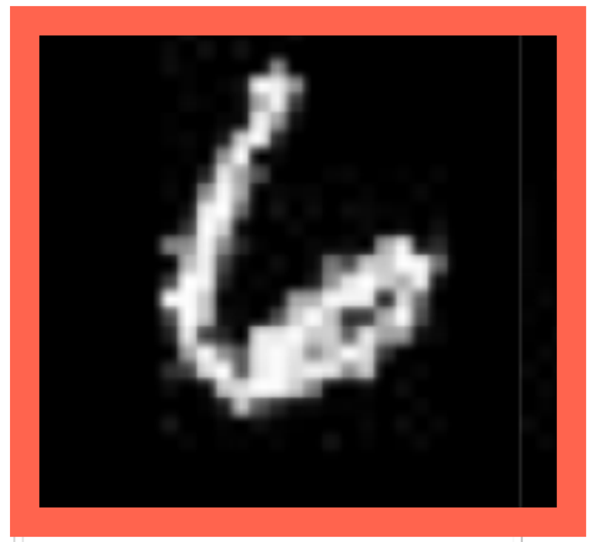
Replay to Remember: MNIST Rotations



Bobu, Tzeng, Hoffman, Darrell. ICLR Workshop 2018.

Evaluate MNIST 135 after all rotations

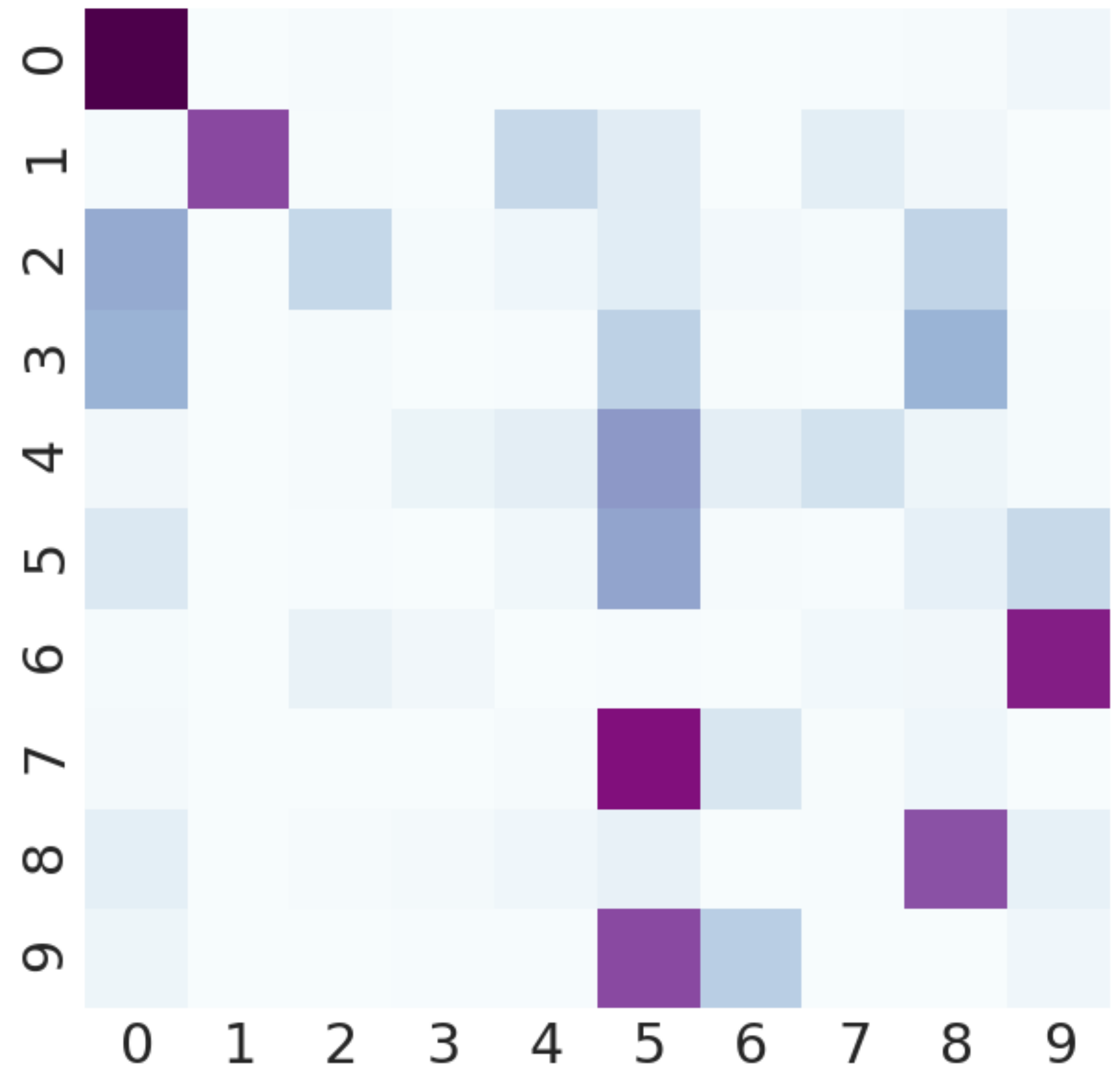
Labeled Source



0º

Unlabeled Target (test)

135º

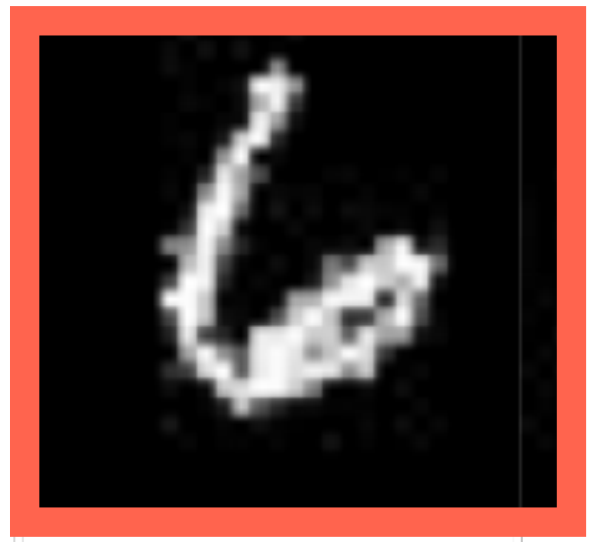


(a) Source

Bobu, Tzeng, Hoffman, Darrell. ICLR Workshop 2018.

Evaluate MNIST 135 after all rotations

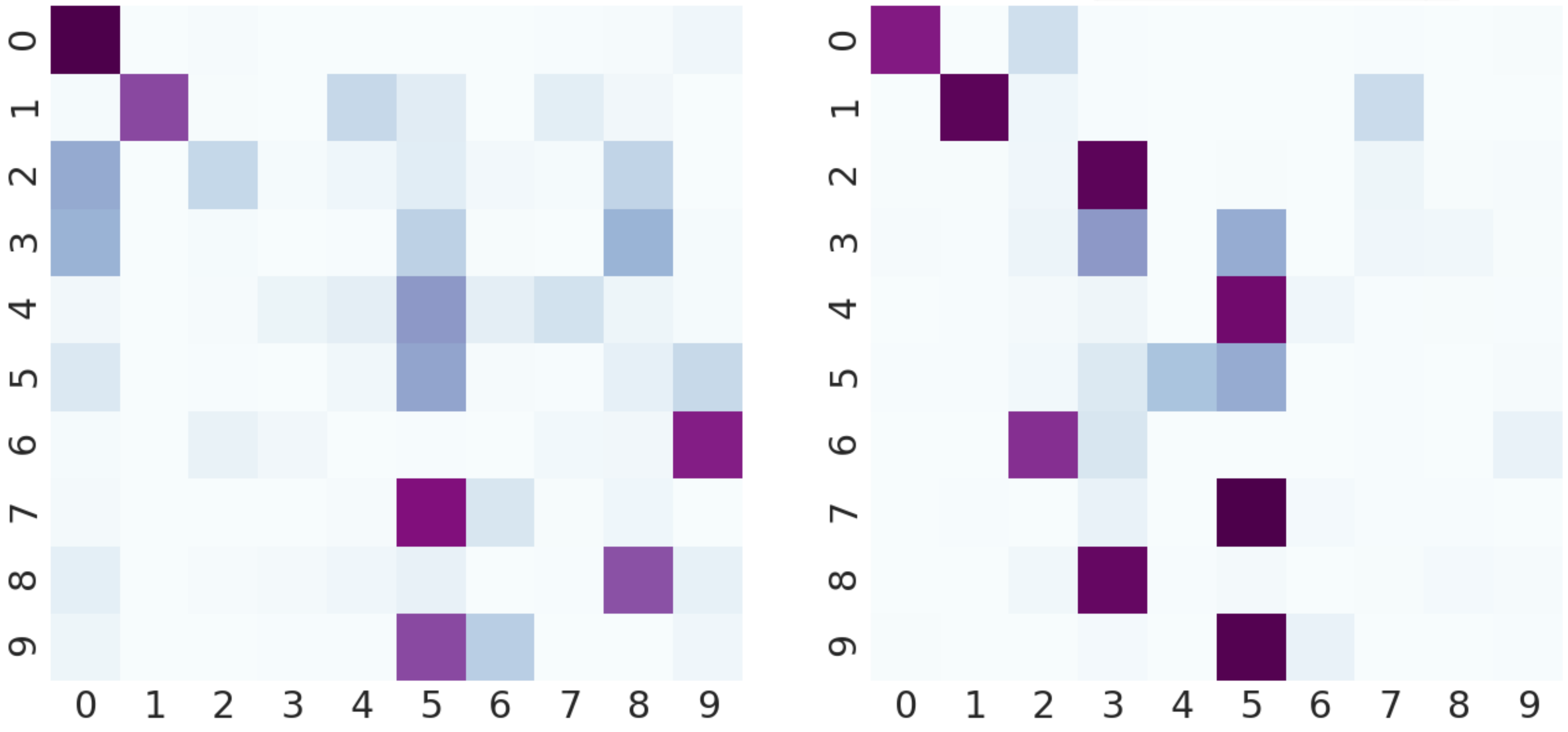
Labeled Source



0º

Unlabeled Target (test)

135º



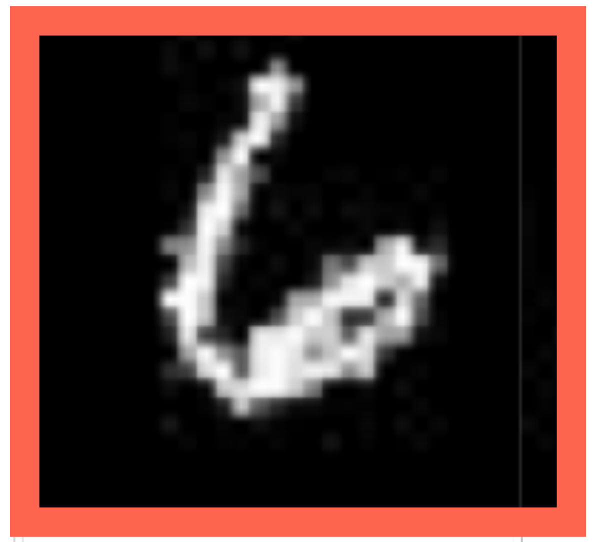
(a) Source

(b) Adapt Batch

Bobu, Tzeng, Hoffman, Darrell. ICLR Workshop 2018.

Evaluate MNIST 135 after all rotations

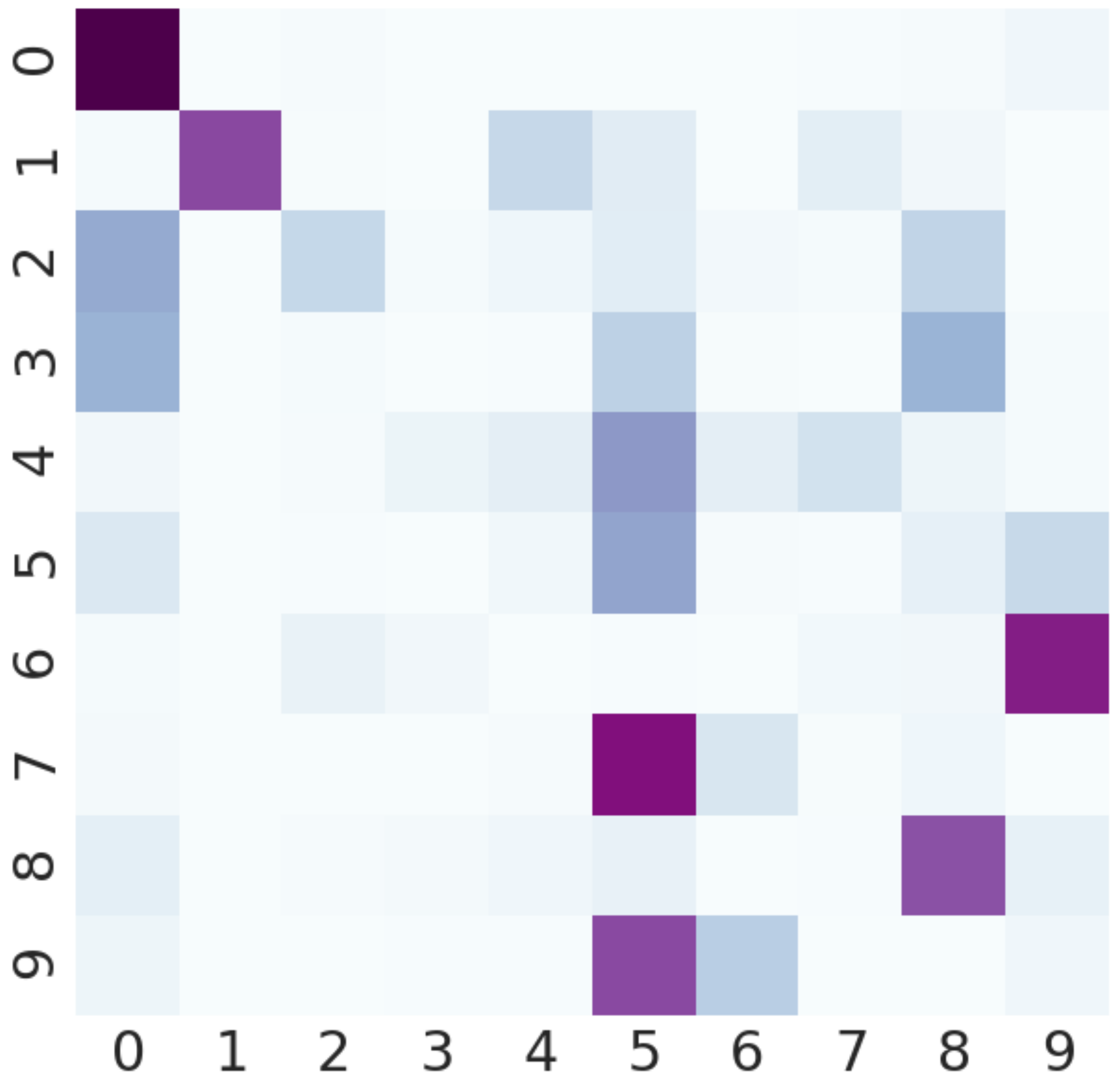
Labeled Source



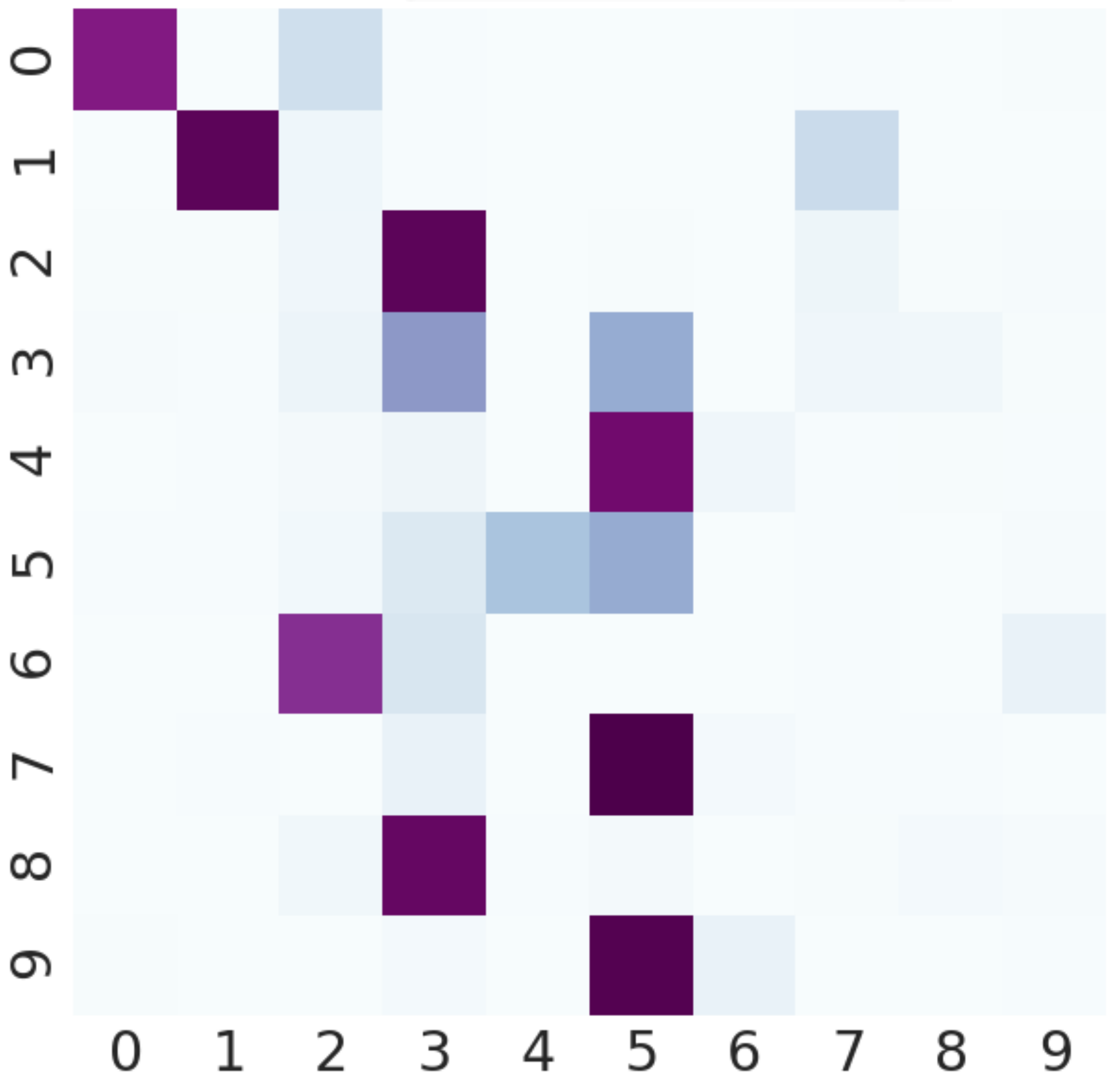
0º

Unlabeled Target (test)

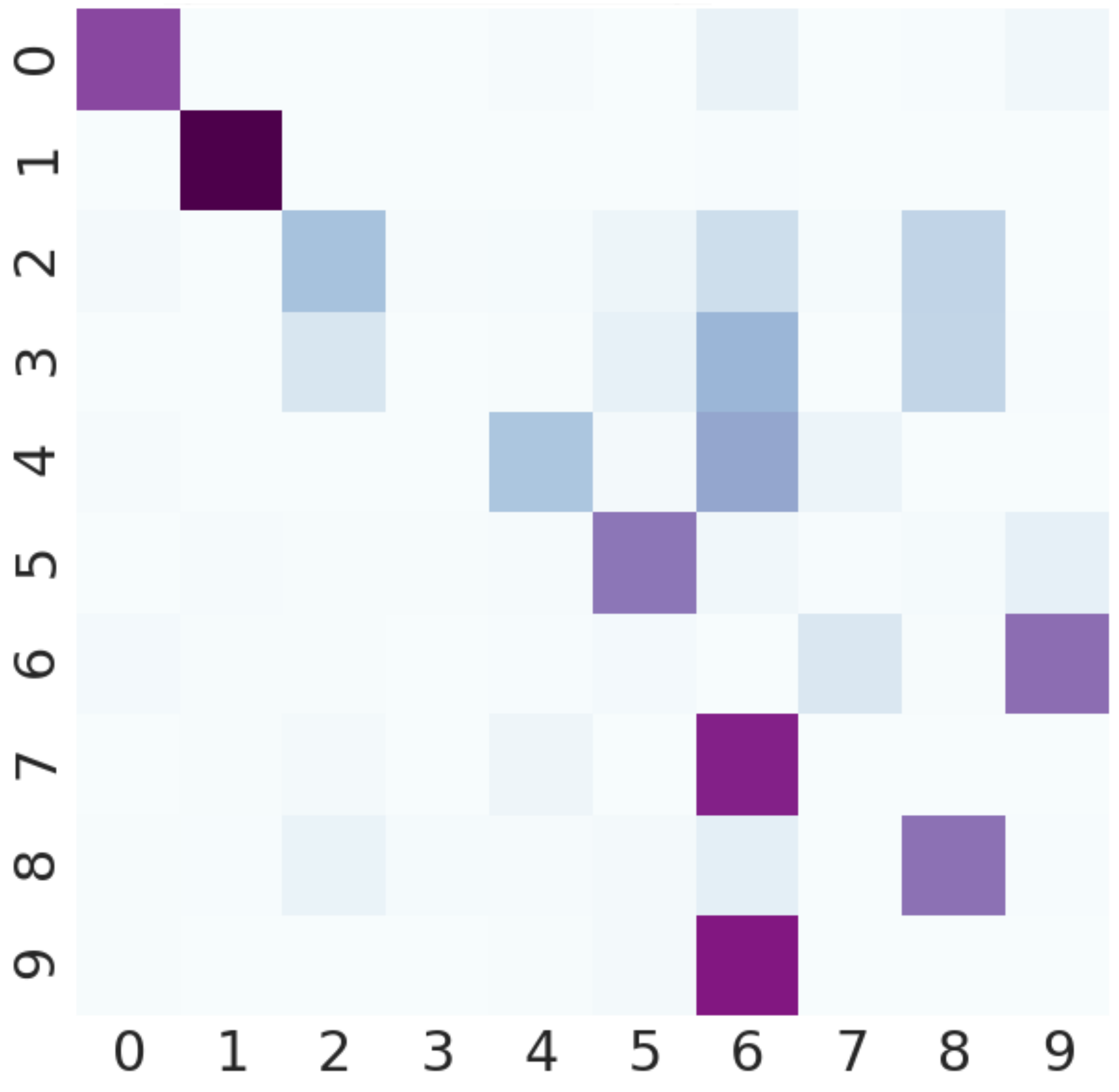
135º



(a) Source



(b) Adapt Batch

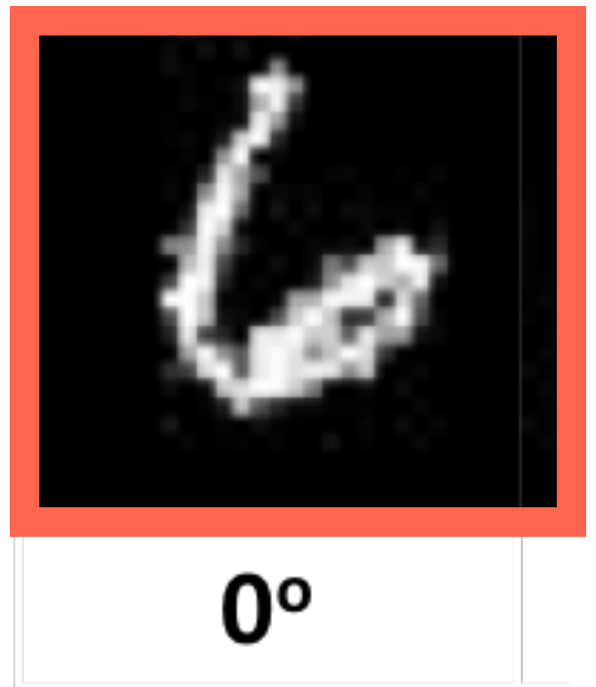


(c) CUA (no replay)

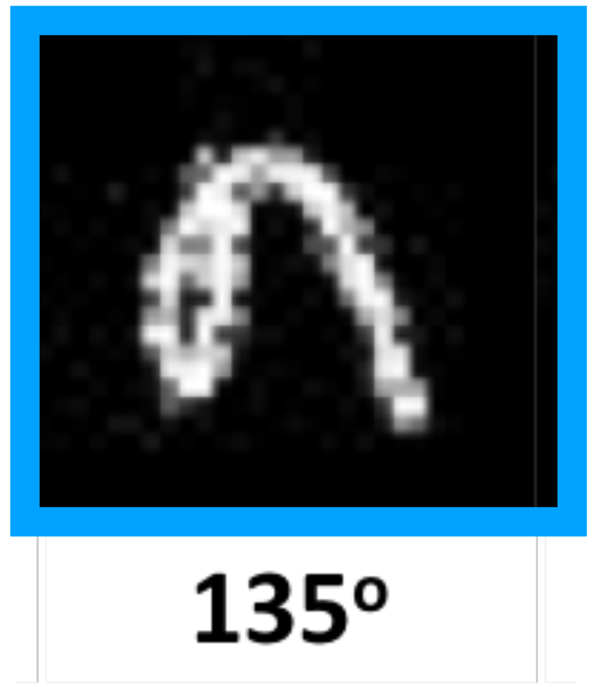
Bobu, Tzeng, Hoffman, Darrell. ICLR Workshop 2018.

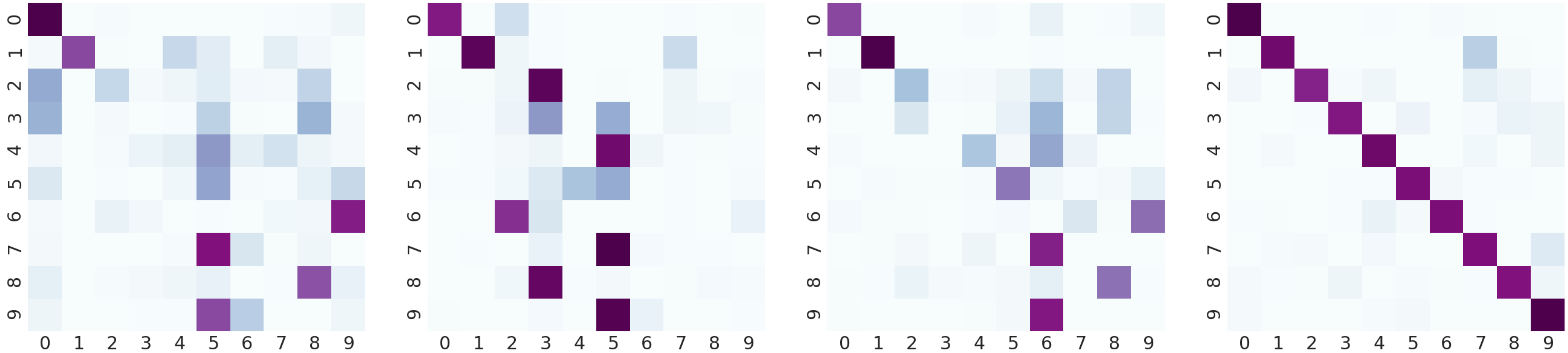
Evaluate MNIST 135 after all rotations

Labeled Source



Unlabeled Target (test)





(a) Source

(b) Adapt Batch

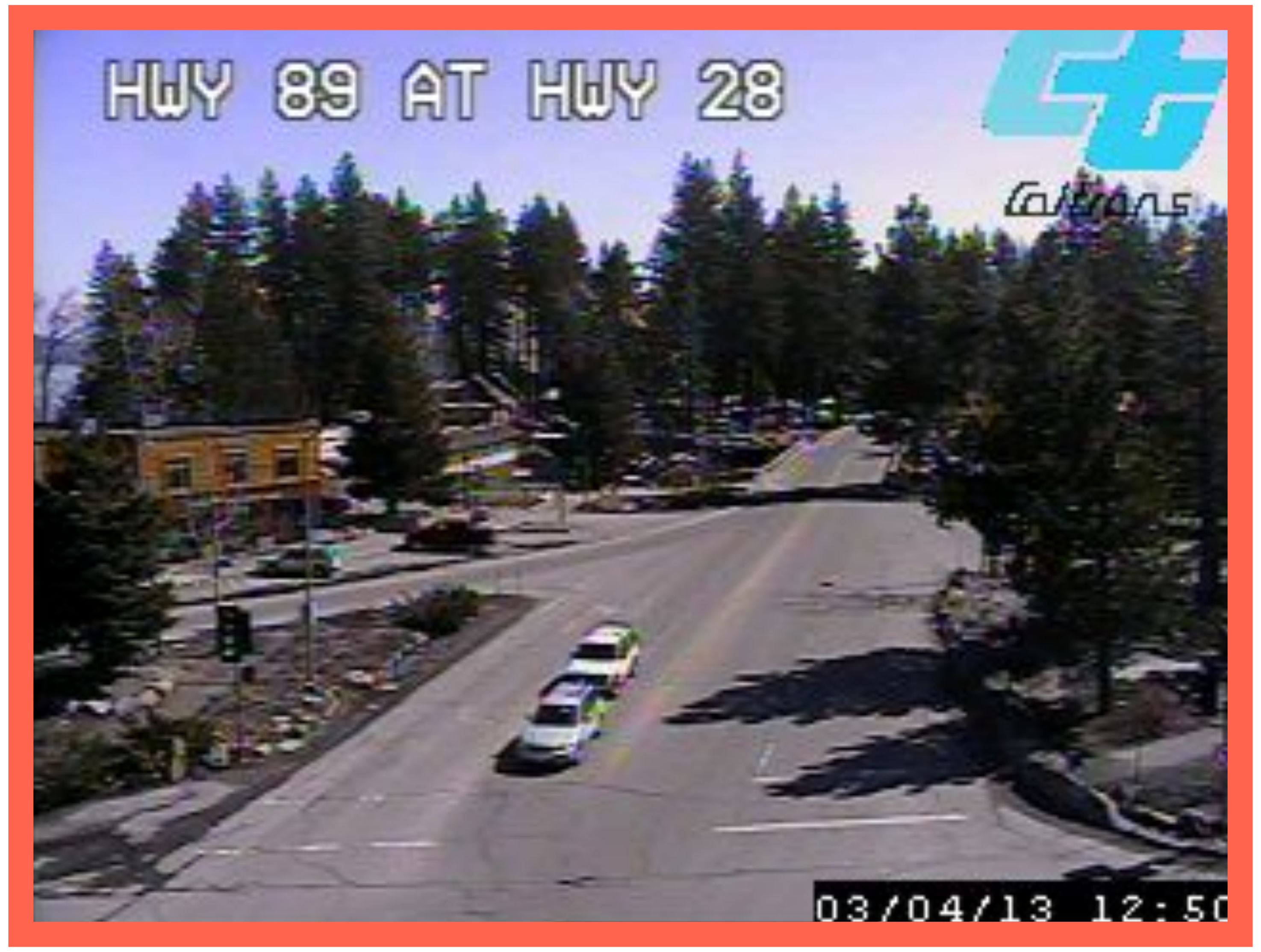
(c) CUA (no replay)

(d) CUA (Full Model)

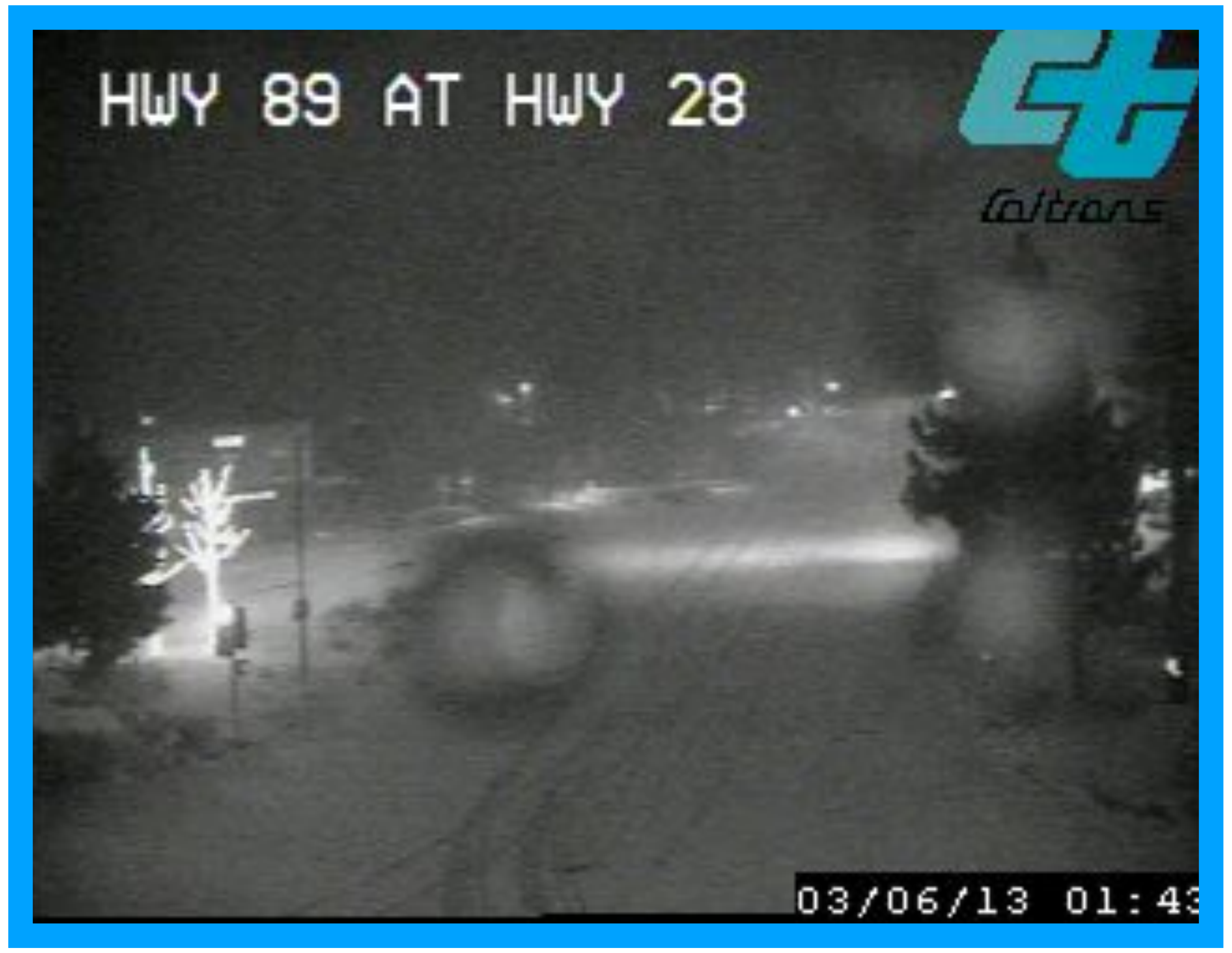
Bobu, Tzeng, Hoffman, Darrell. ICLR Workshop 2018.

Summary Batch Adaptation

Labeled Source

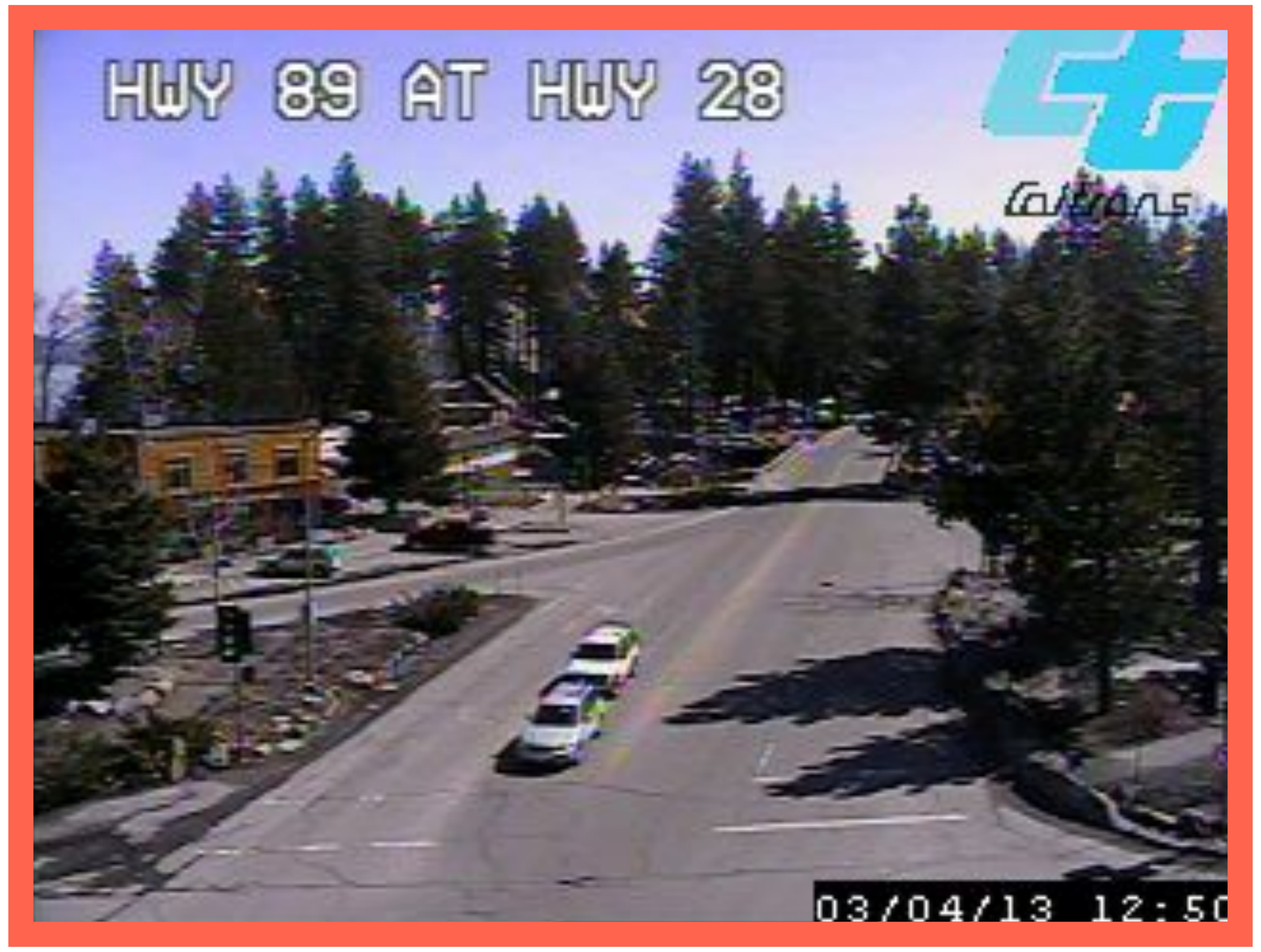


Unlabeled Target (test)

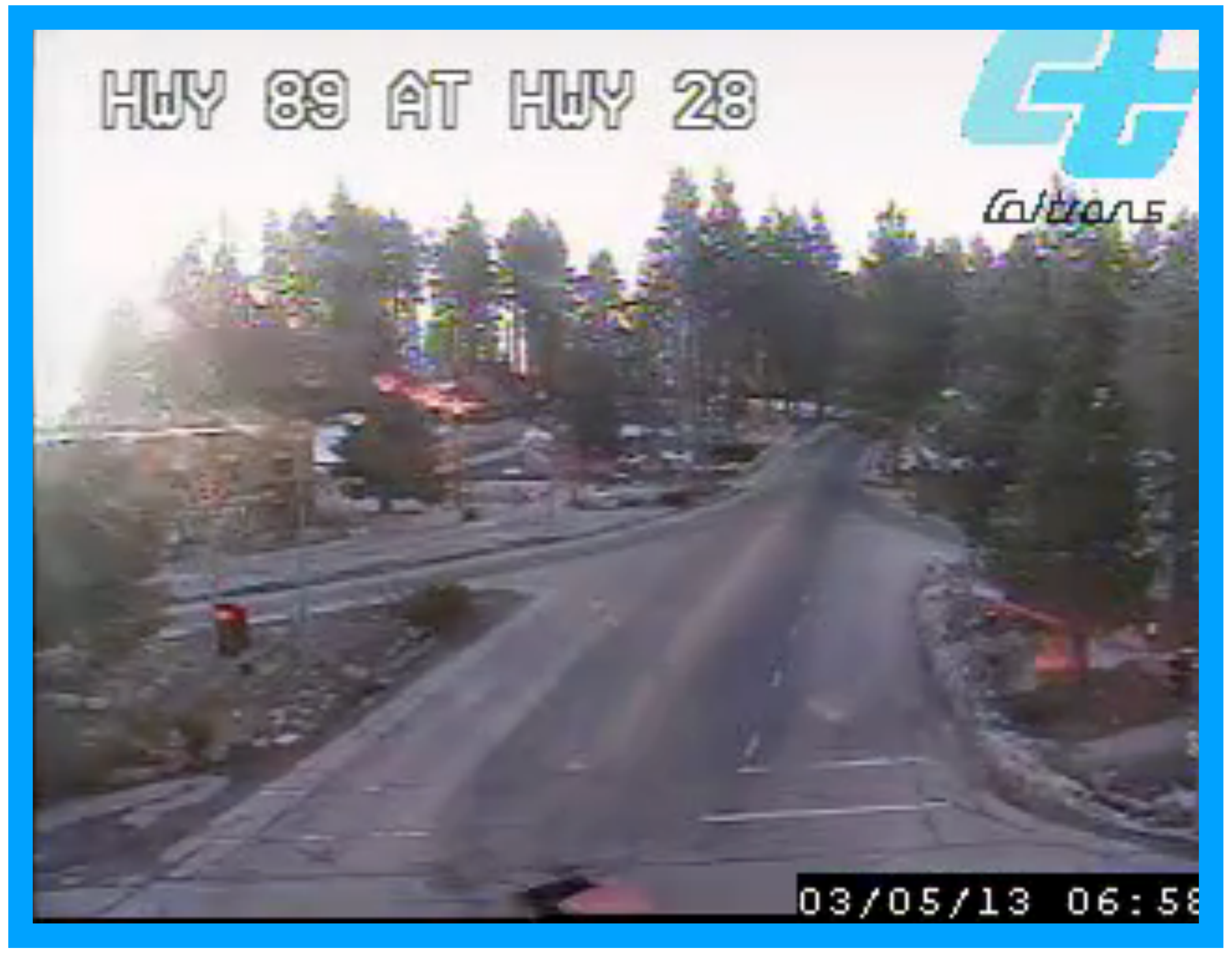


Summary Continuous Adaptation

Labeled Source



Unlabeled Target (test)



Adaptation vs Robustness

Robust Learning with Jacobian Regularization

Dan Roberts
Diffeo

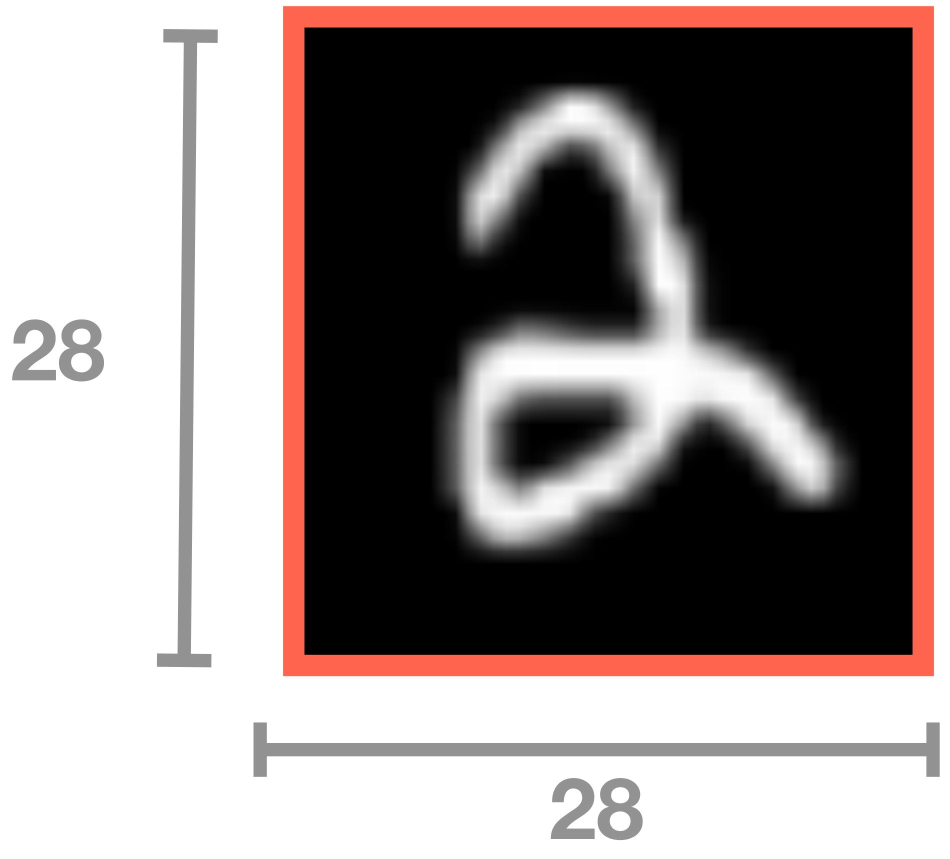
Sho Yaida
FAIR

Hoffman, Roberts, Yaida, In Submission, 2019.

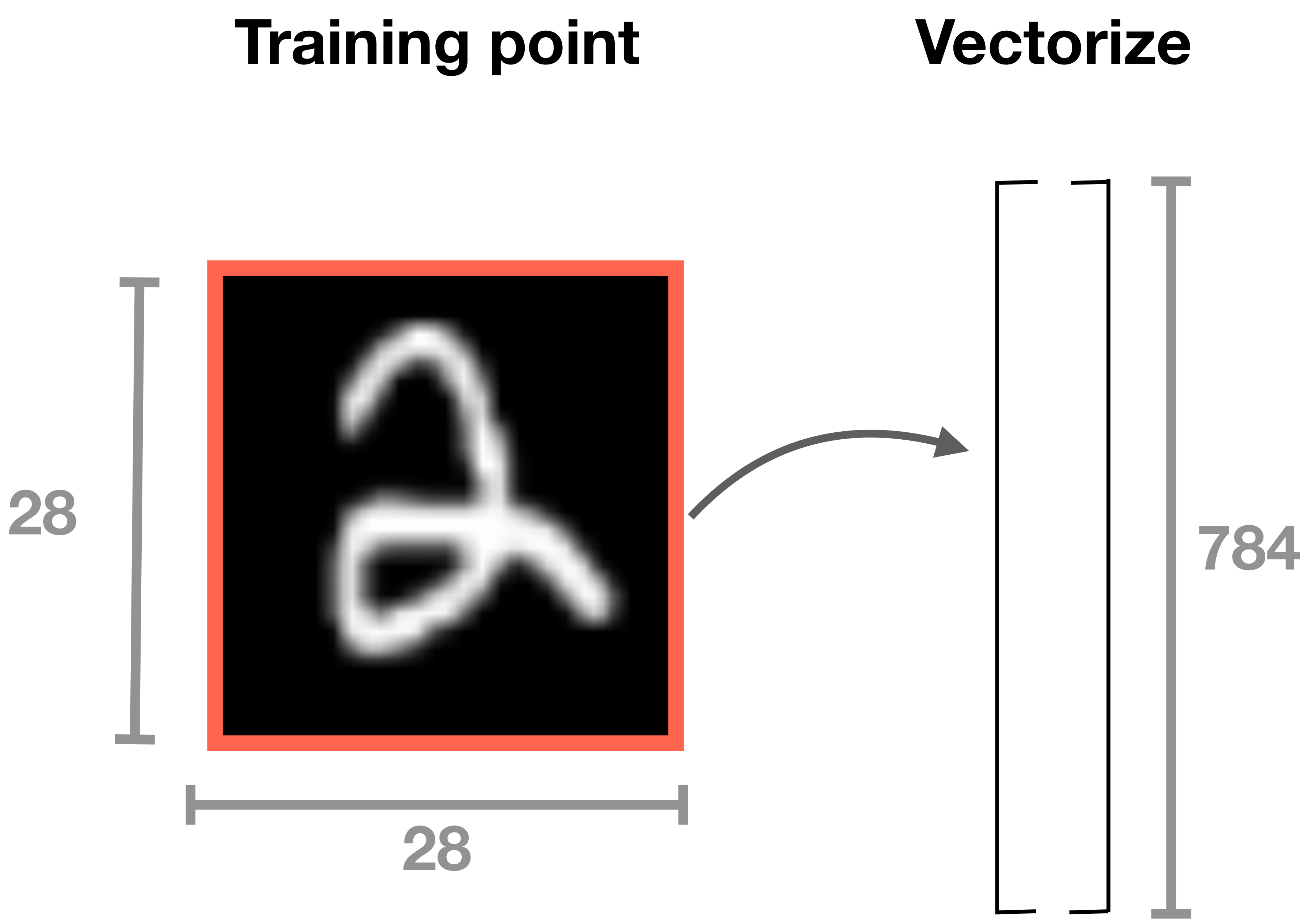
Visualize Perturbation Space

Visualize Perturbation Space

Training point



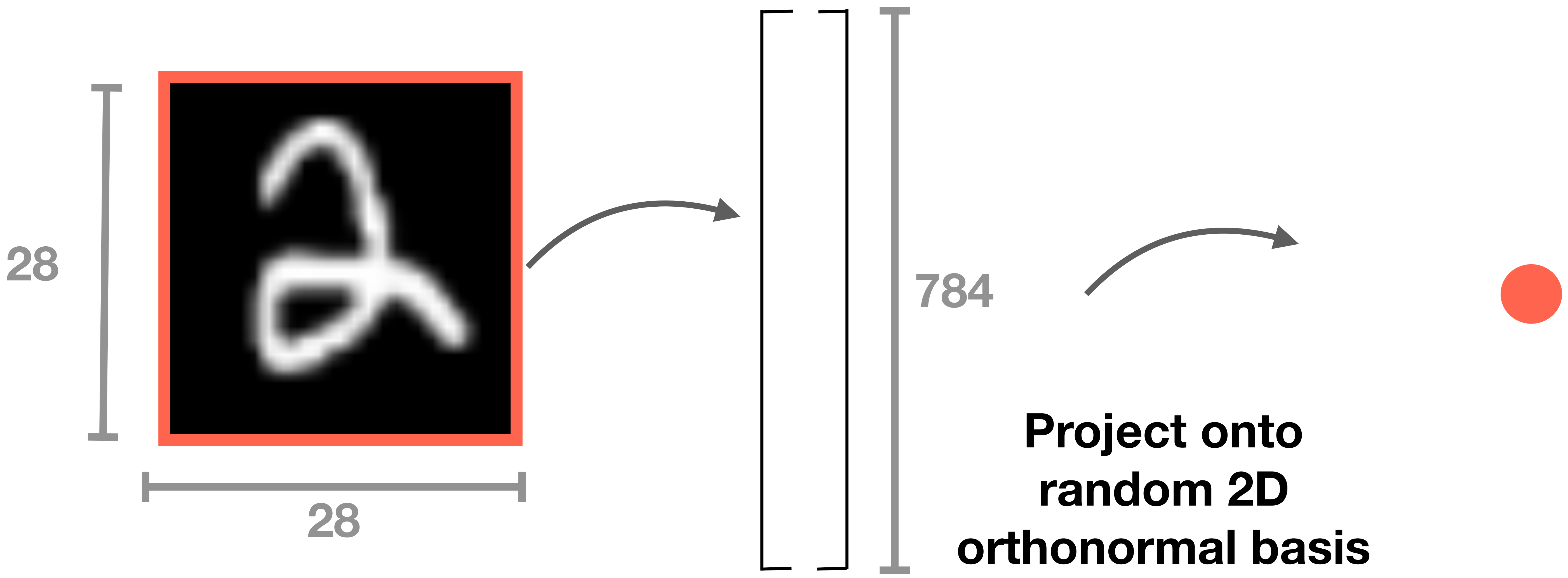
Visualize Perturbation Space



Visualize Perturbation Space

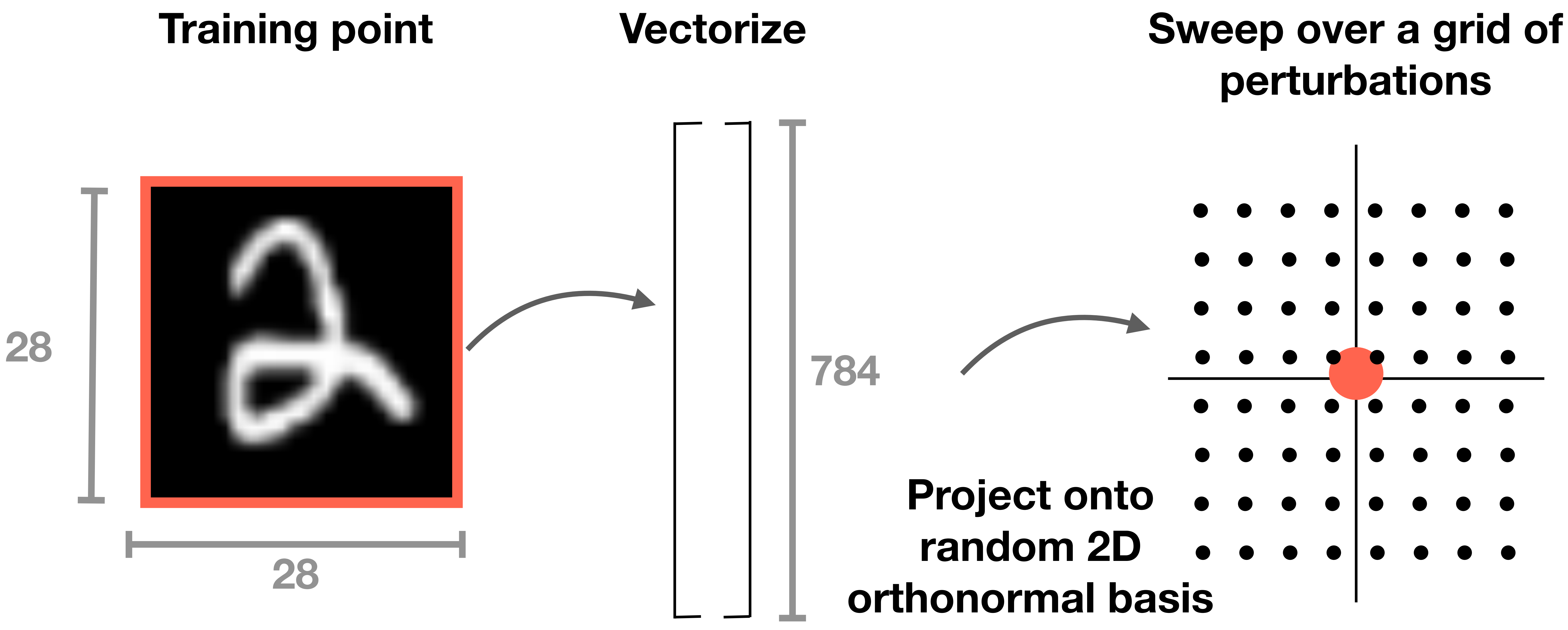
Training point

Vectorize



Project onto random 2D orthonormal basis

Visualize Perturbation Space



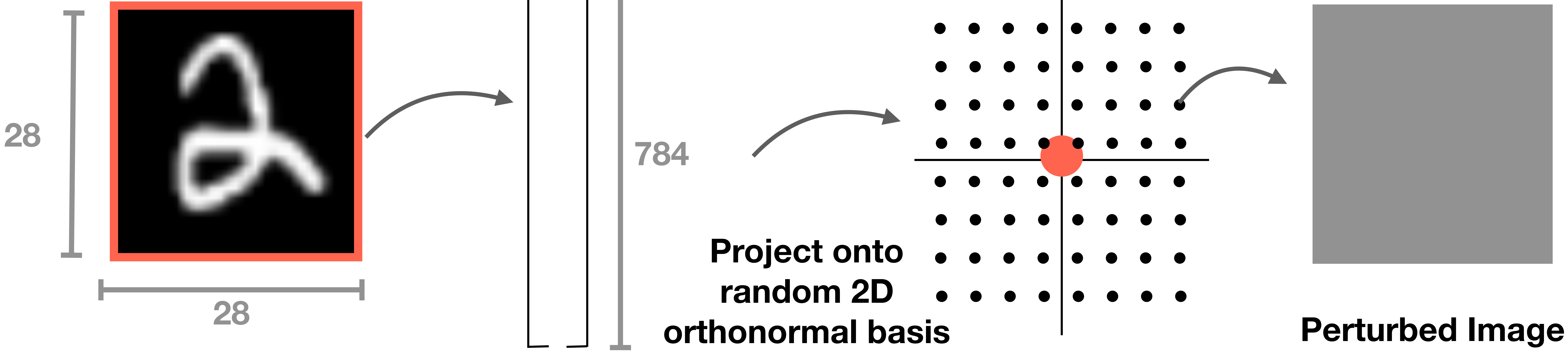
Training point

Vectorize

Sweep over a grid of perturbations

Project onto random 2D orthonormal basis

Visualize Perturbation Space



Training point

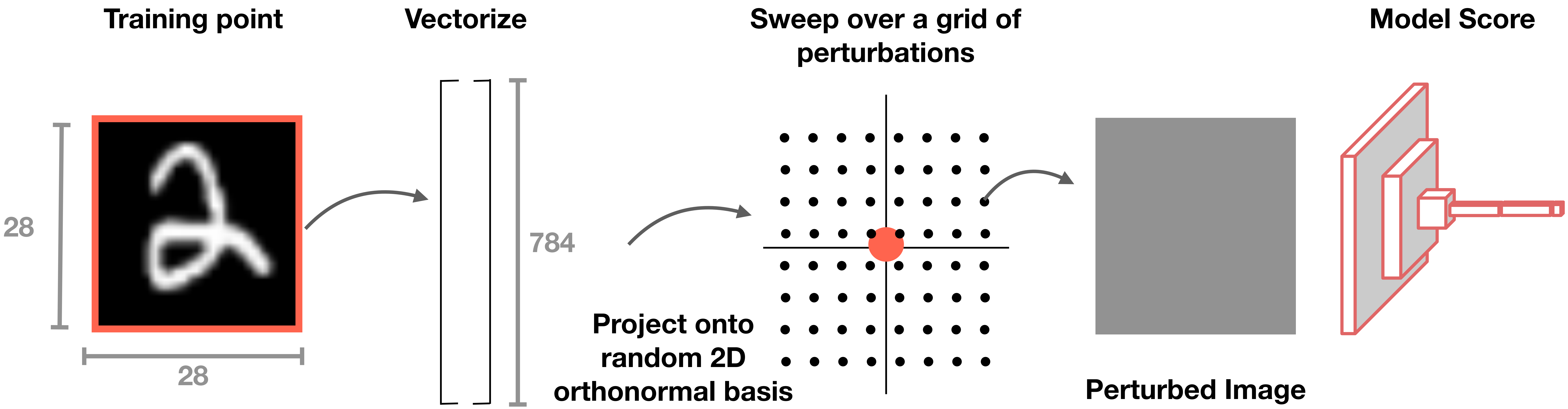
Vectorize

Sweep over a grid of perturbations

Project onto random 2D orthonormal basis

Perturbed Image

Visualize Perturbation Space



Training point

Vectorize

Sweep over a grid of perturbations

Model Score

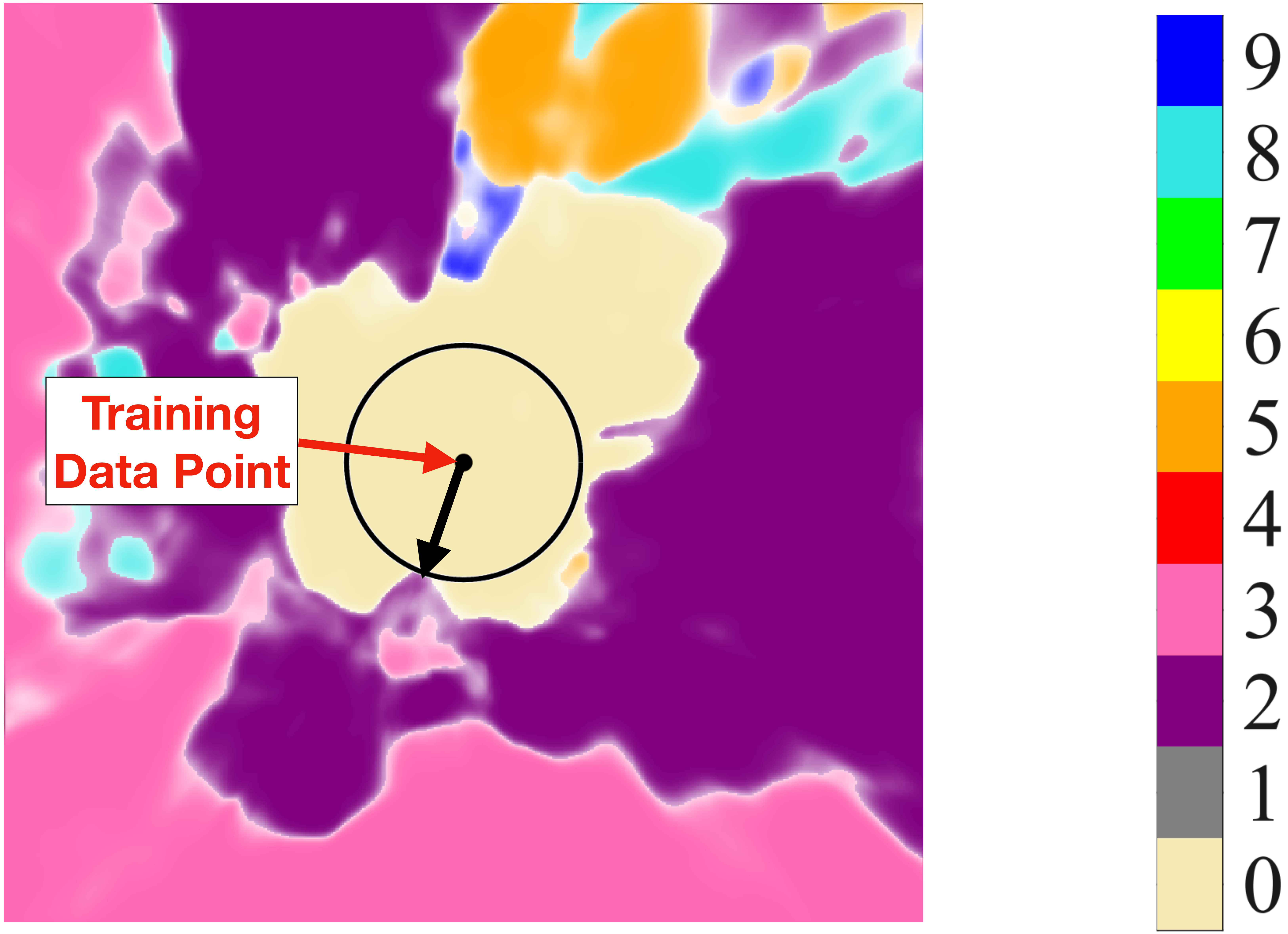
Project onto random 2D orthonormal basis

Perturbed Image

MNIST LeNet Decisions Around Training Point

Non-smooth Decision Boundary

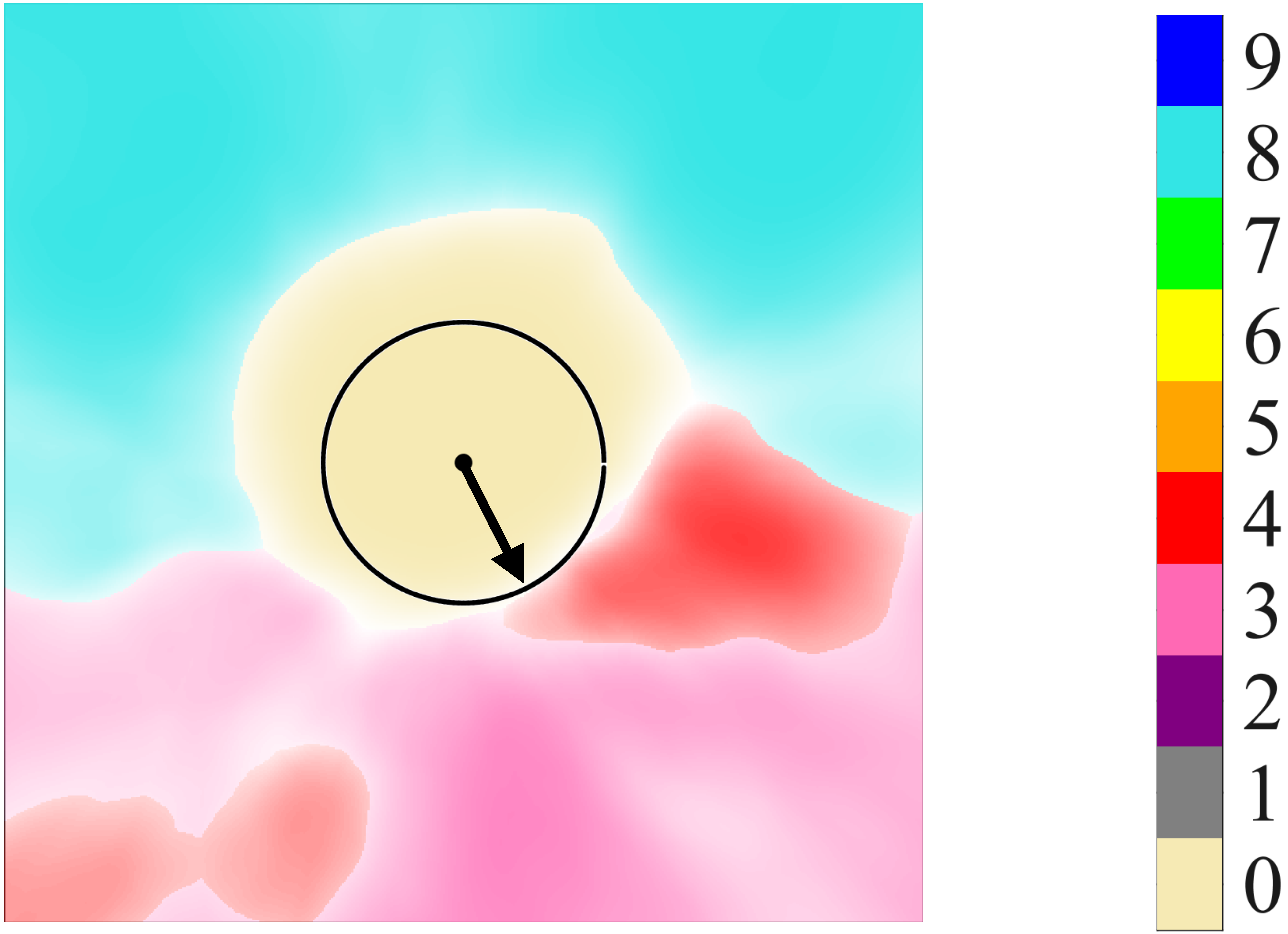
Small perturbations lead to new outputs



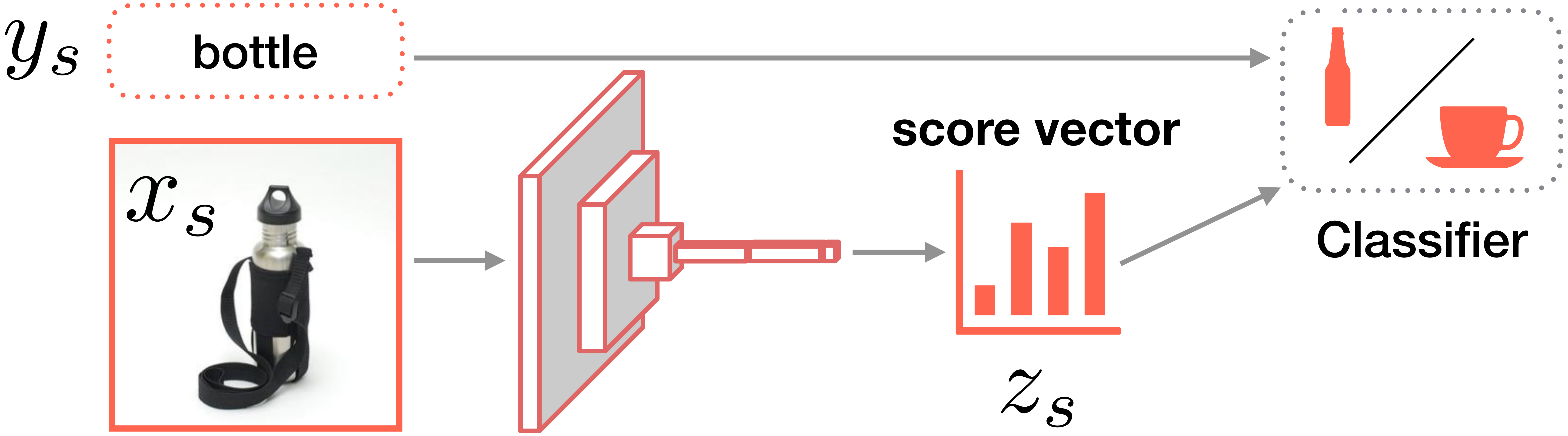
MNIST LeNet with L2 Regularization

Smooth Decision Boundary

Small perturbations lead to new outputs



Jacobian Regularization



Input-output Jacobian matrix

$$J_{c,i} = \frac{\partial z_c}{\partial x_i}$$

Minimize Frobenius Norm

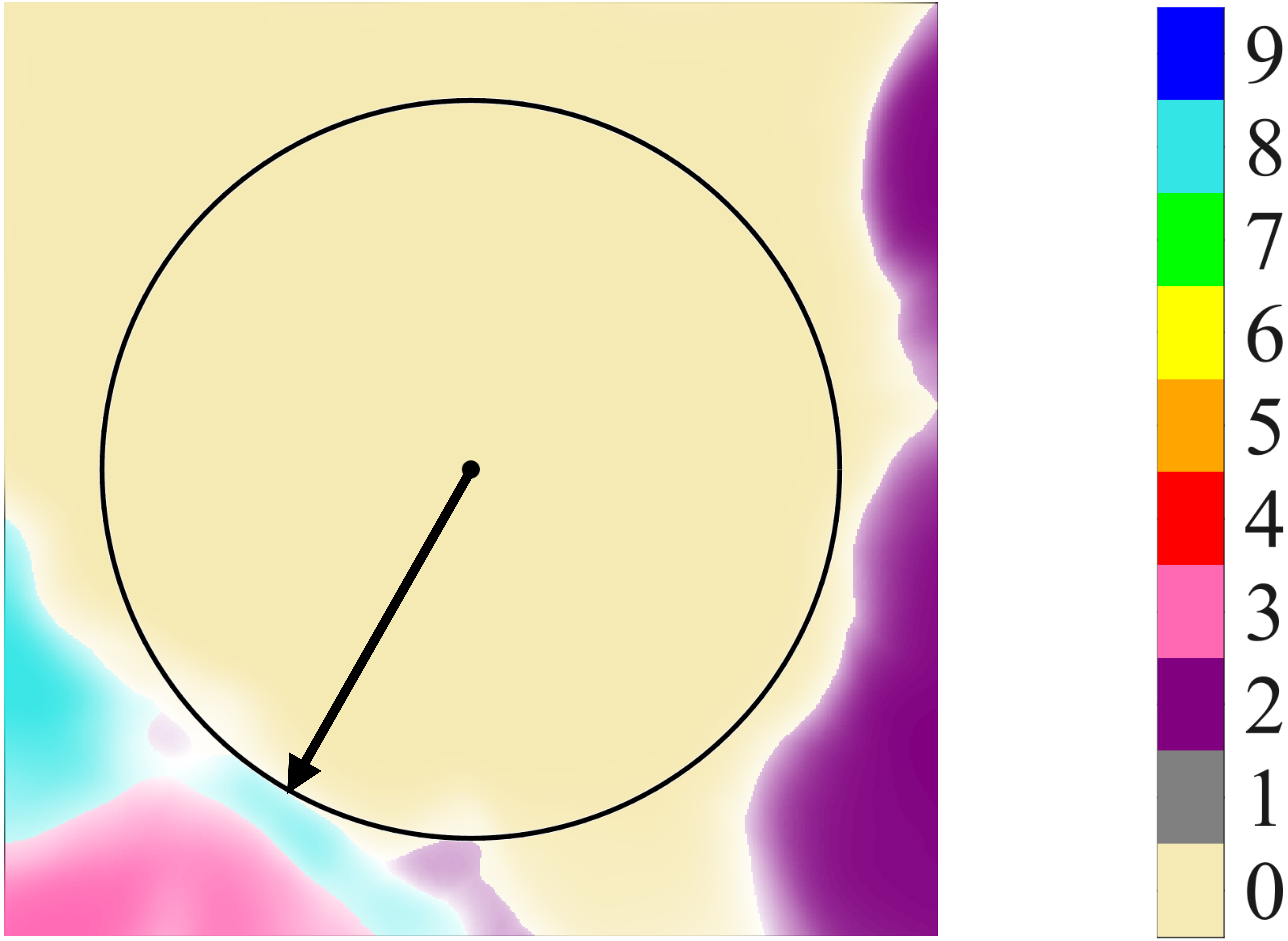
$$||J||_F^2$$

Hoffman, Roberts, Yaida, In Submission, 2019.

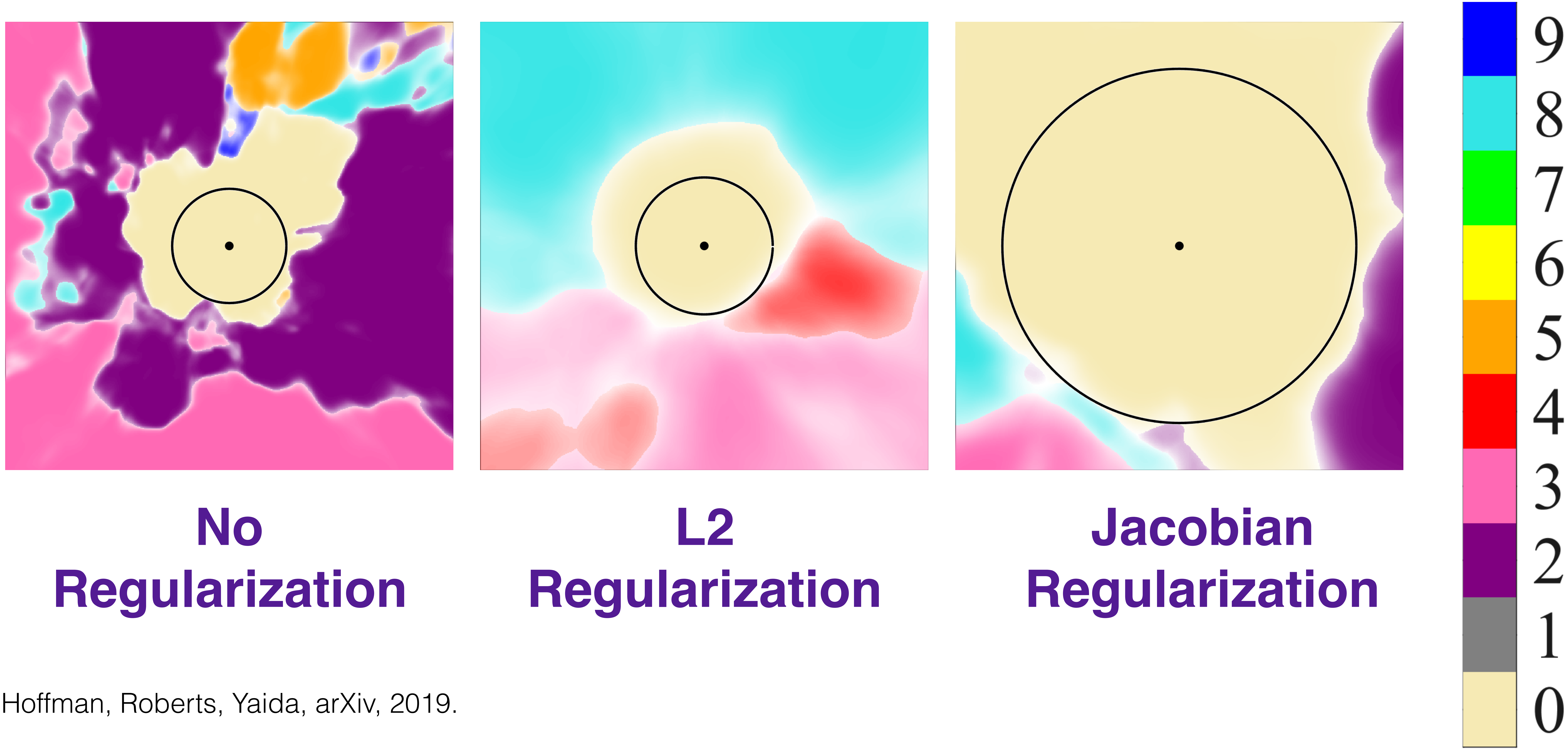
MNIST LeNet with Jacobian Regularization

Mostly Smooth Decision Boundary

Larger perturbations needed to lead to new outputs



Decision Boundary Comparison



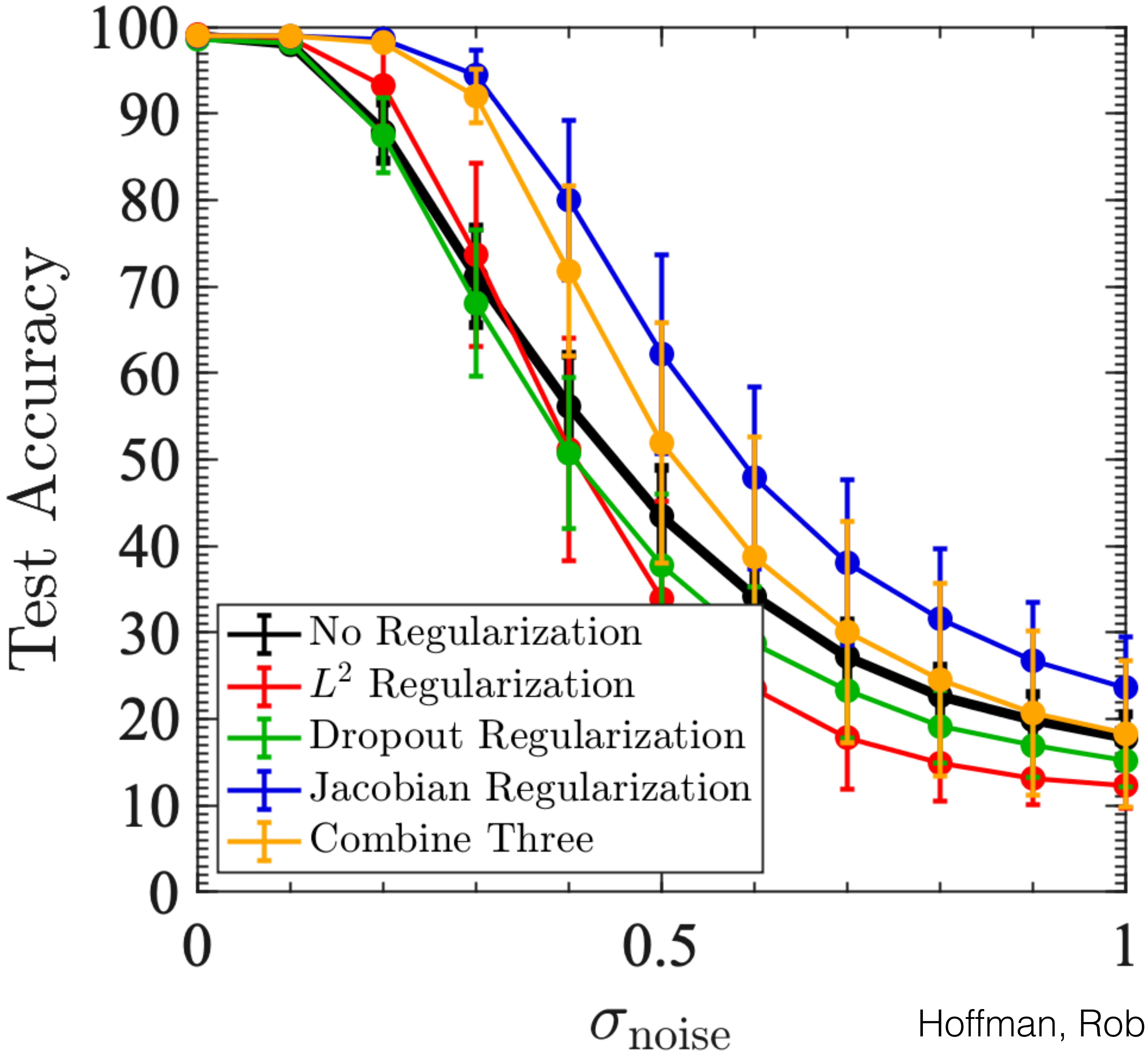
No Regularization

L2 Regularization

Jacobian Regularization

Hoffman, Roberts, Yaida, arXiv, 2019.

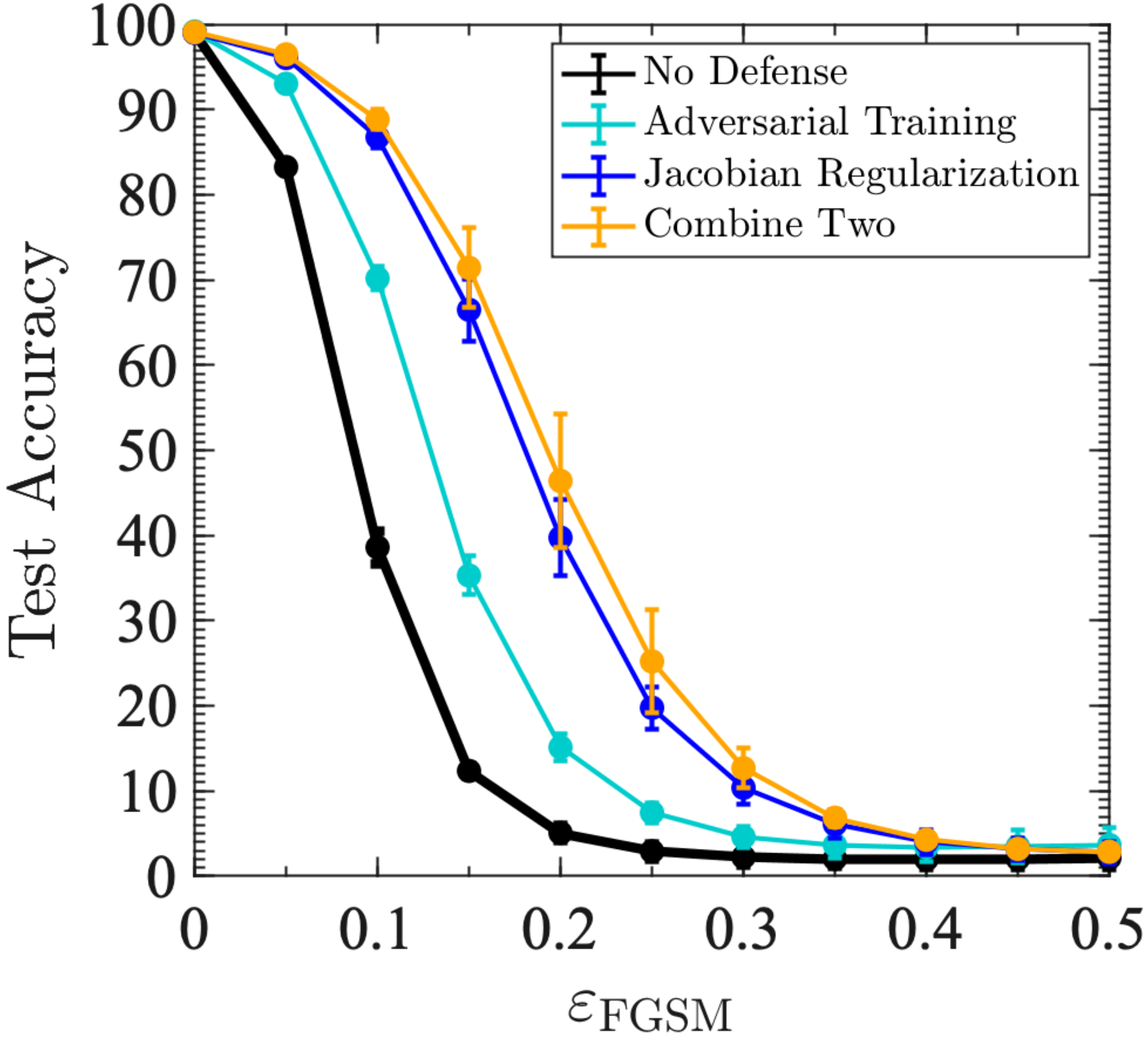
Robustness to Random Perturbations

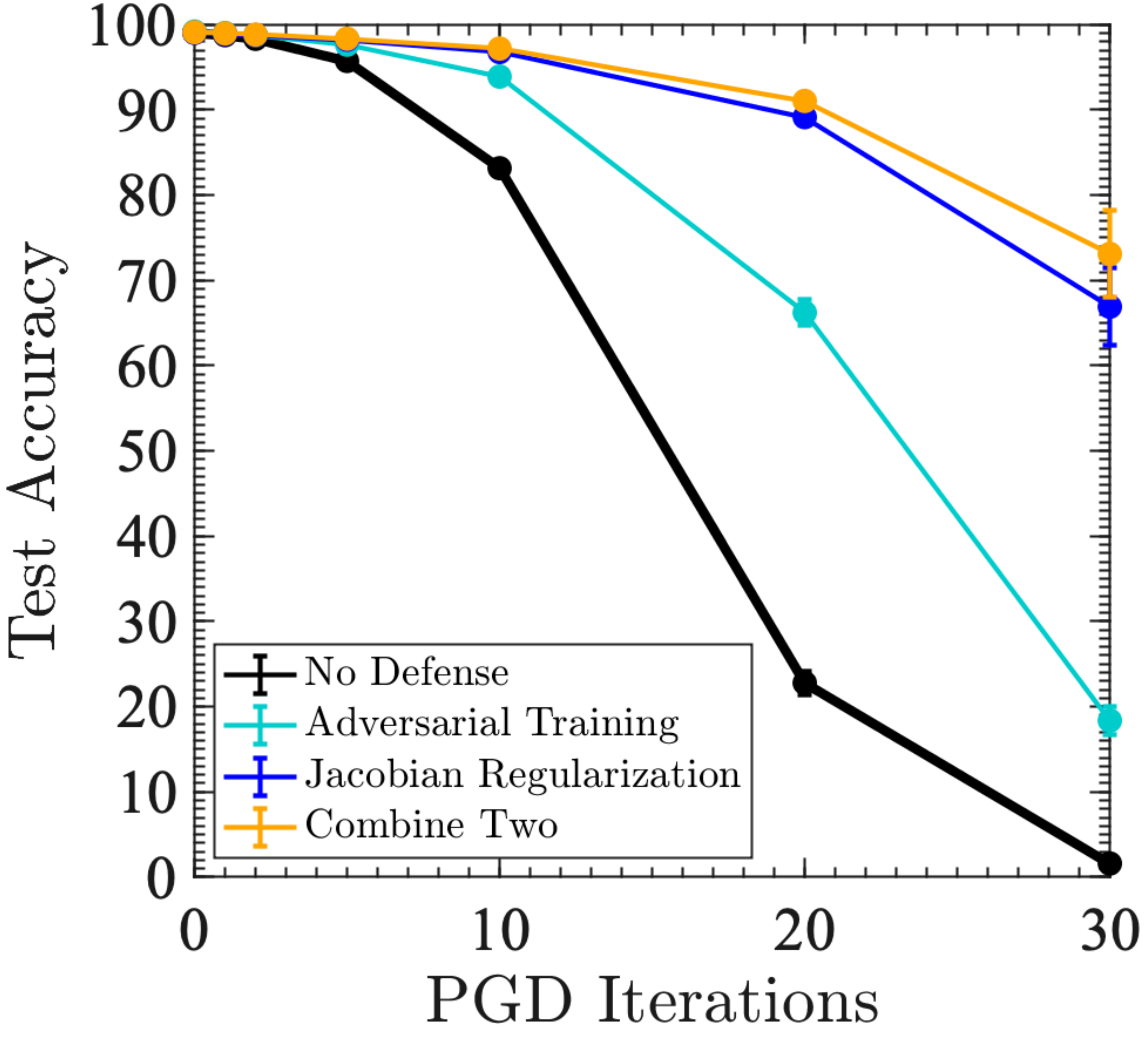


MNIST
LeNet Model

Hoffman, Roberts, Yaida, In Submission, 2019.

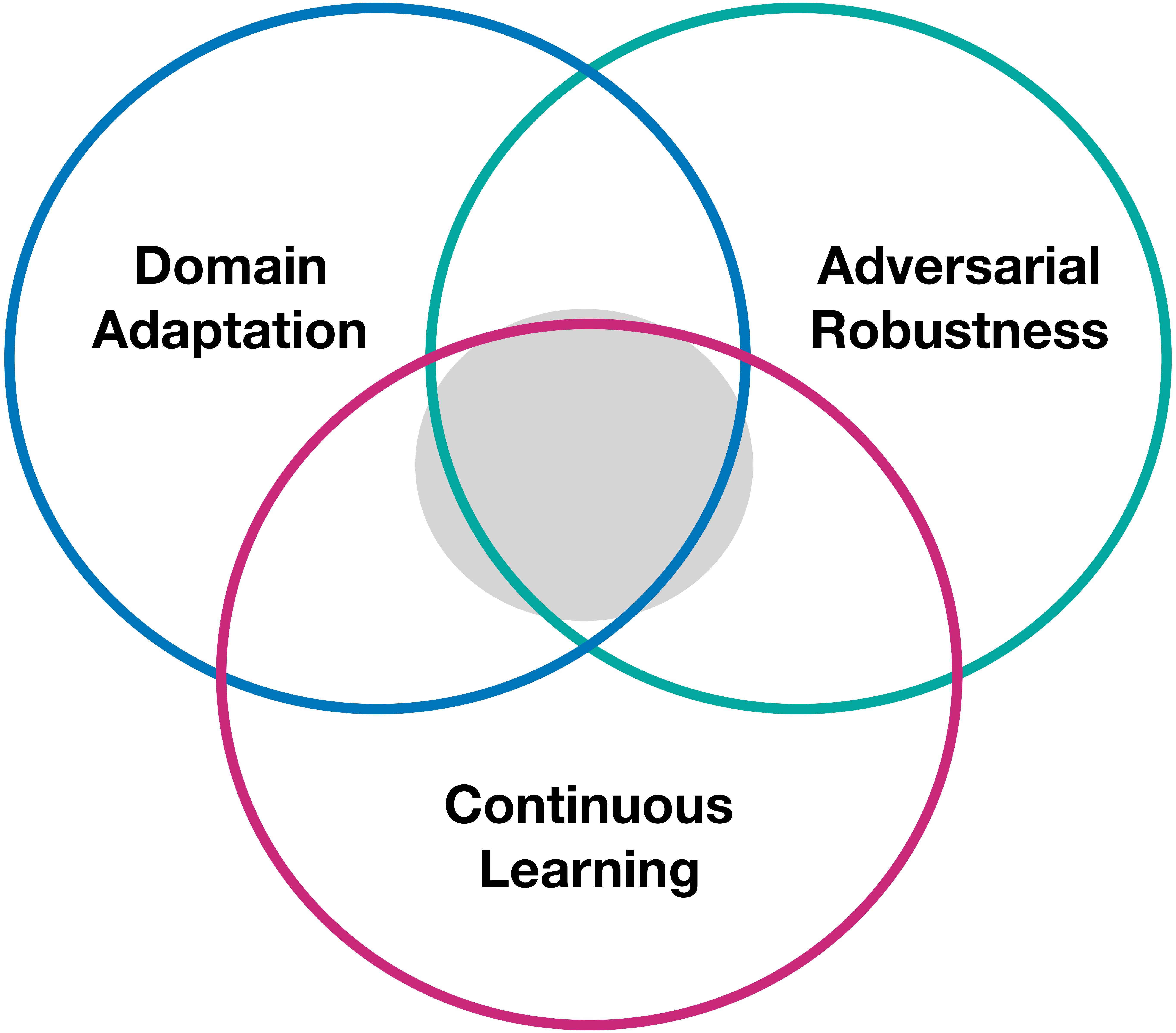
Robustness to Adversarial Perturbations





Hoffman, Roberts, Yaida, In Submission, 2019.

Next Steps

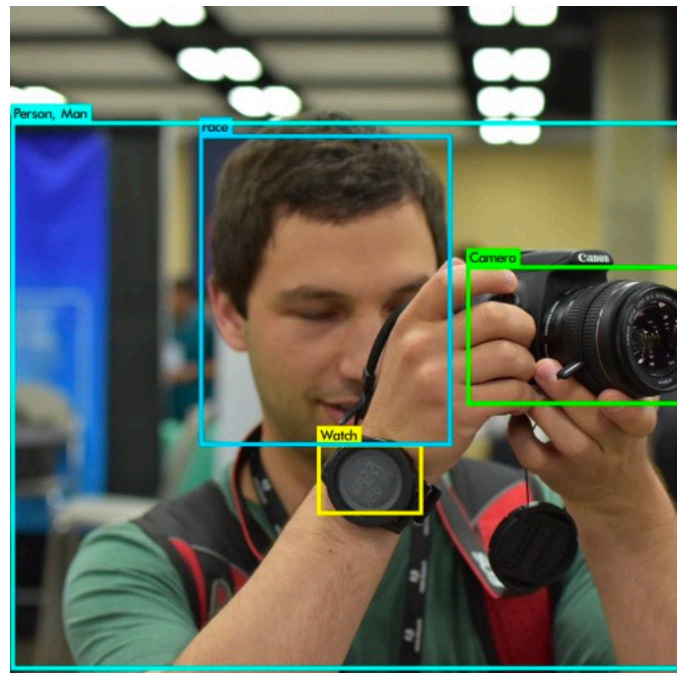


Robustness regularizers as unsupervised adaptive loss?

Adaptation to an adversarial domain?

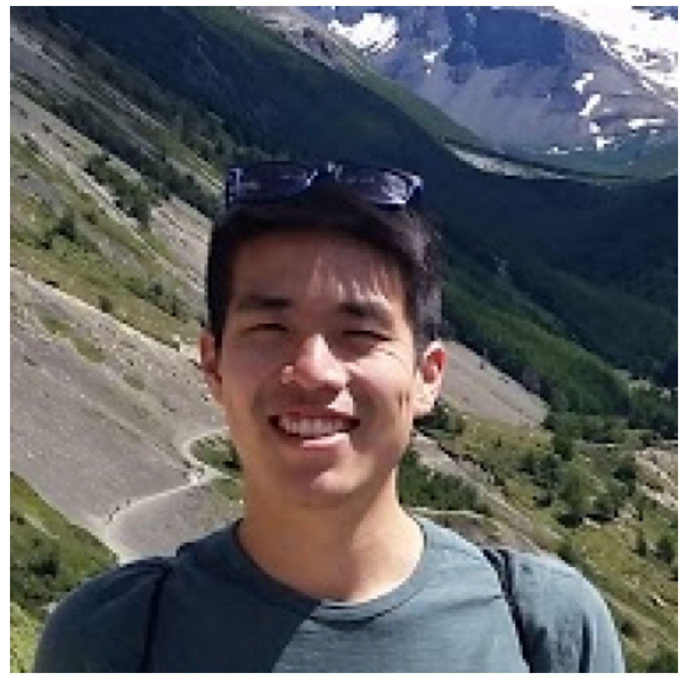
Continuous adaptation to adversarial images

Thank you

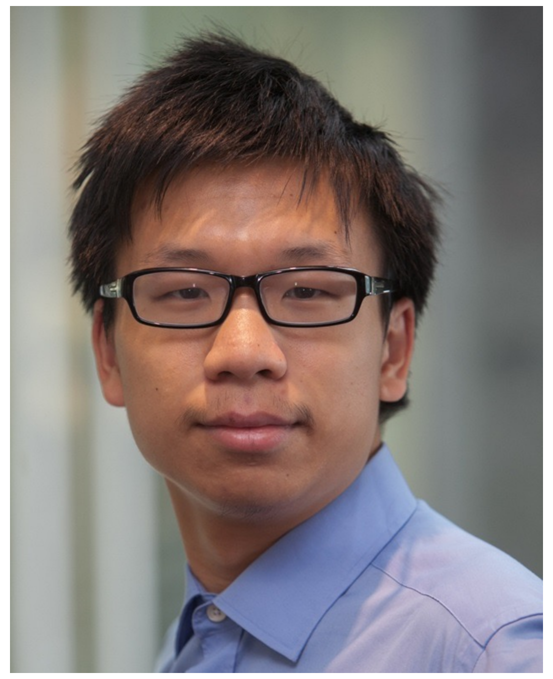


Daniel Gordon
UW

Taesung Park
UC Berkeley



Eric Tzeng
UC Berkeley



Jun-Yan Zhu
MIT

Dan Roberts
Diffeo

Devi Parikh
Georgia Tech / FAIR

Phil Isola
MIT

Kate Saenko
Boston University

Trevor Darrell
UC Berkeley

Alyosha Efros
UC Berkeley

Sho Yaida
FAIR

Dhruv Batra
Georgia Tech / FAIR

Judy Hoffman
judyhoffman.io