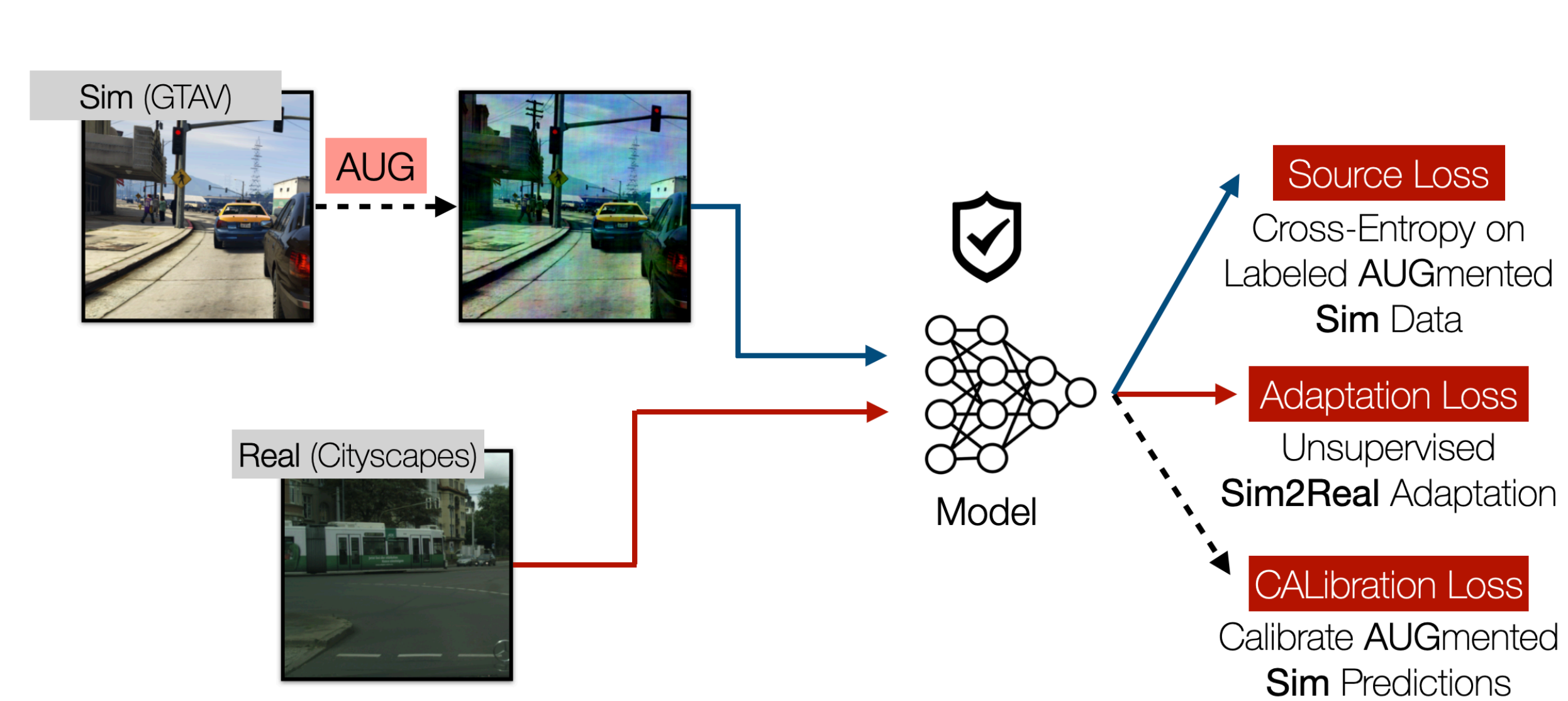


AUGCAL: Improving Sim2Real Adaptation by Uncertainty Calibration on Augmented Synthetic Images [ICLR 2024]



Prithvijit Chattopadhyay



Bharat Goyal



Bogi Ecsedi

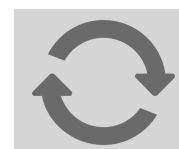


Viraj Prabhu



Judy Hoffman





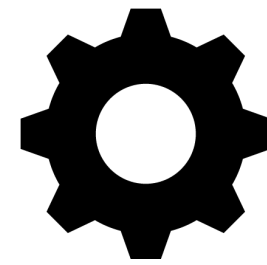
(Visual) Sim2Real Transfer

Goal: Sim2Real Adaptation



(Visual) Sim2Real Transfer

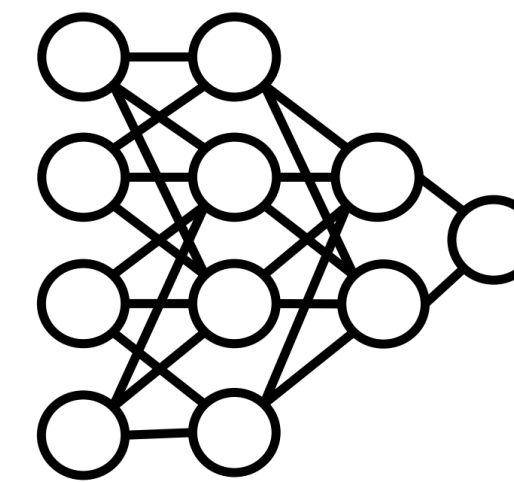
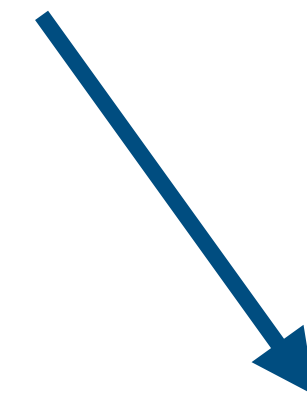
Goal: Sim2Real Adaptation



GTAV



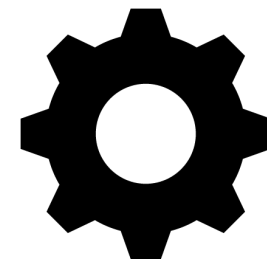
Labeled "Synthetic" Data



Model

(Visual) Sim2Real Transfer

Goal: Sim2Real Adaptation



GTA V



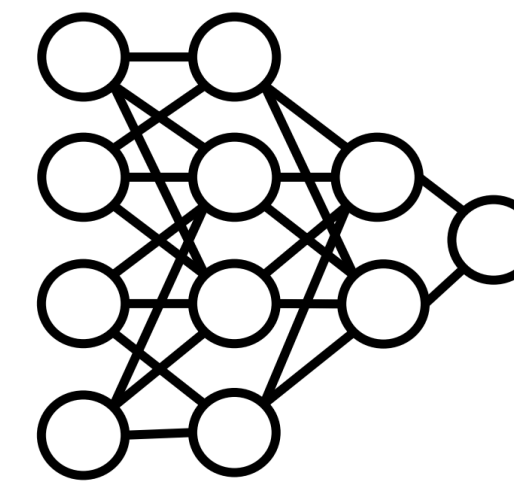
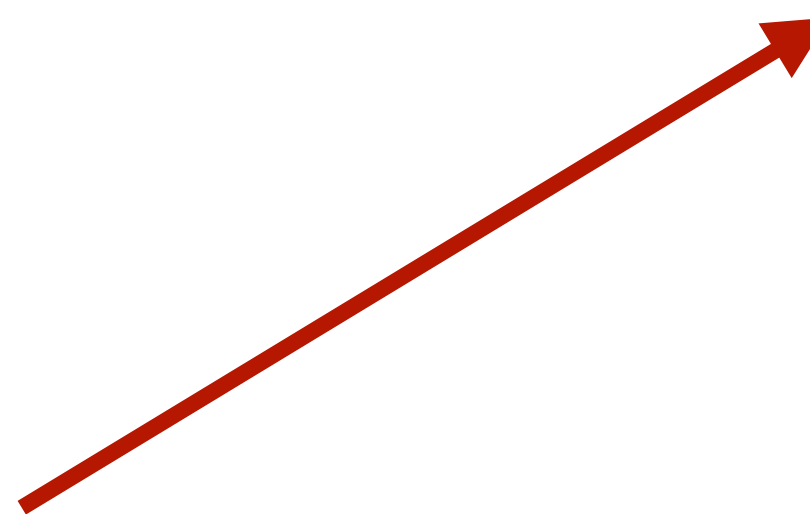
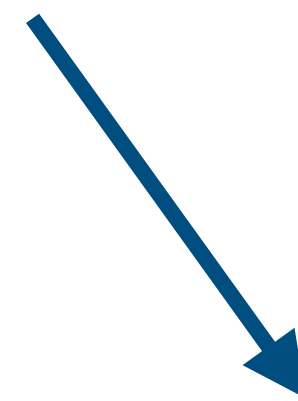
Labeled "Synthetic" Data



Cityscapes



Unlabeled "Real" Data



Model



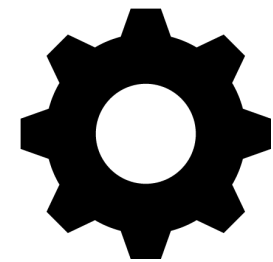
Cityscapes



Test on "Real" Data

(Visual) Sim2Real Transfer

Goal: (Reliable) Sim2Real Adaptation



GTA V



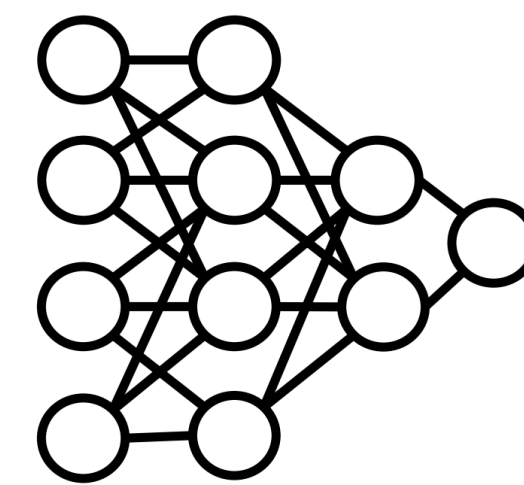
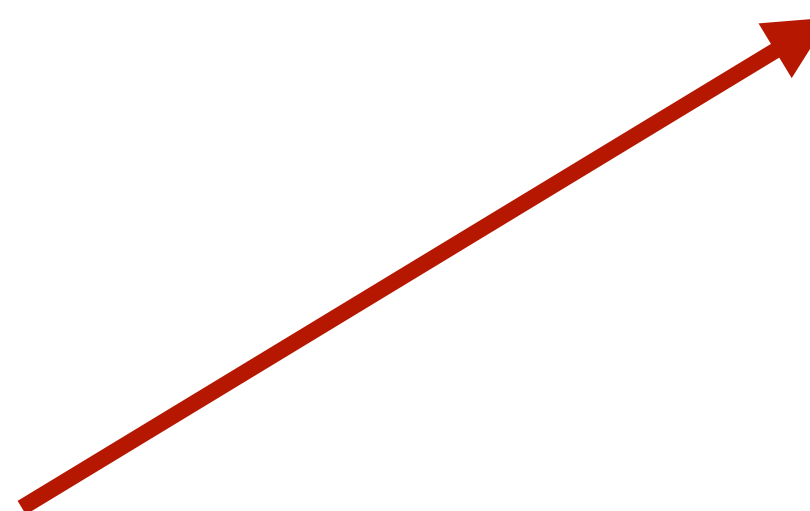
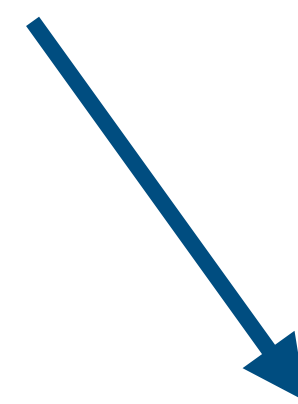
Labeled "Synthetic" Data



Cityscapes



Unlabeled "Real" Data



Model



Cityscapes



Test on "Real" Data

Ensure models trained on Synthetic images make reliable predictions on Real images

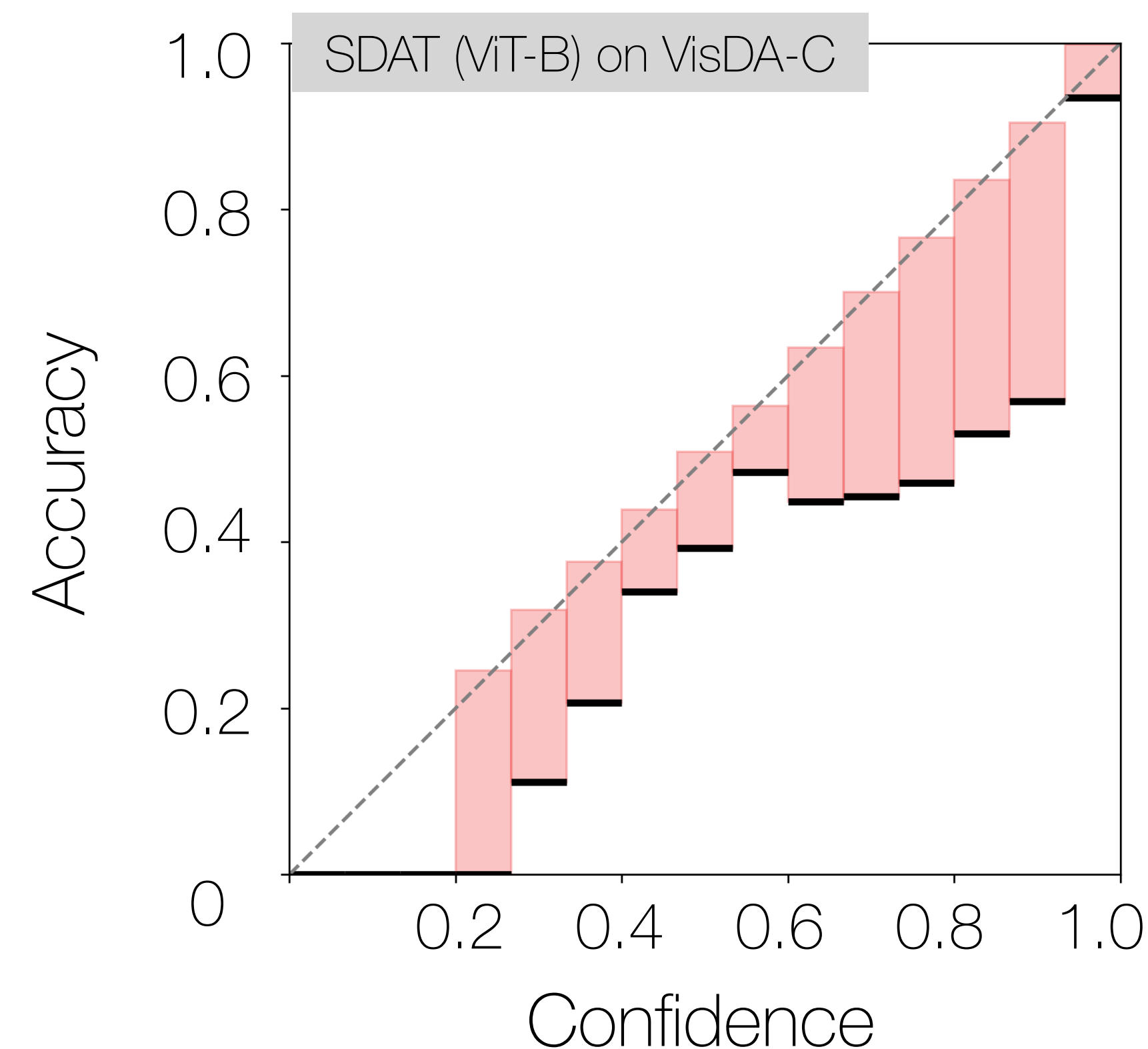
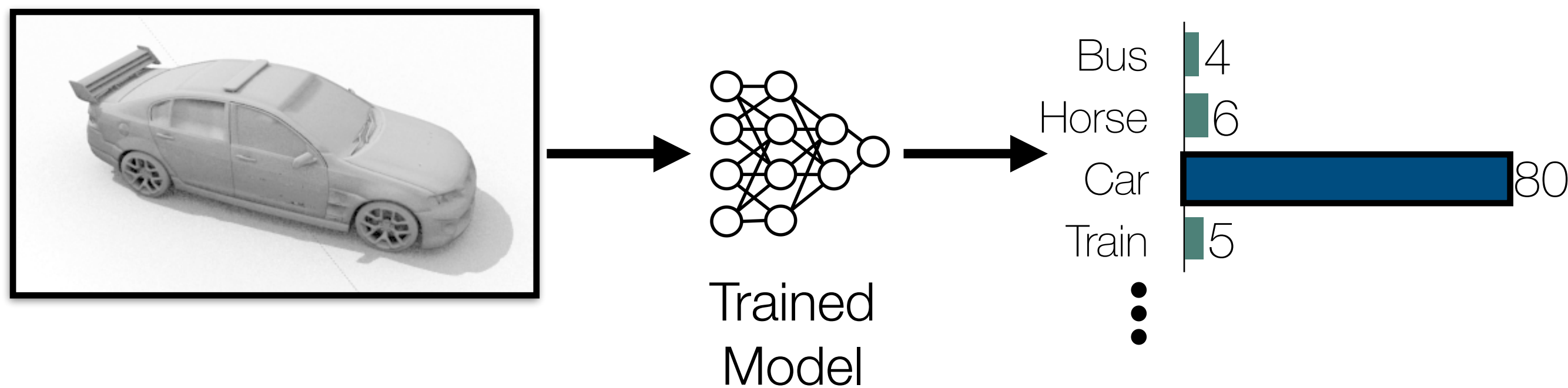




(Visual) Sim2Real Transfer

Reliability: Calibrated Sim2Real Predictions

■ Confidence
■ Accuracy



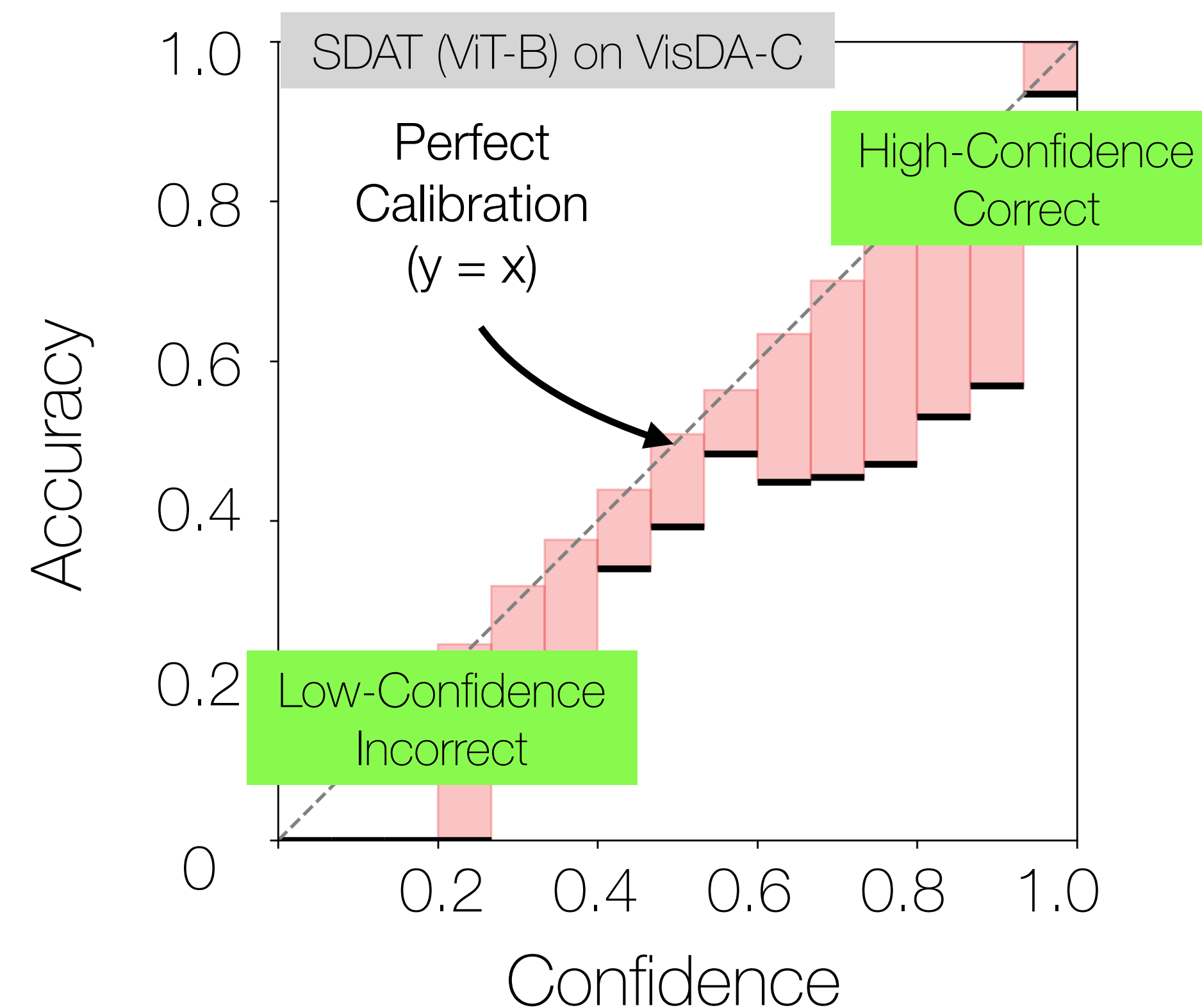
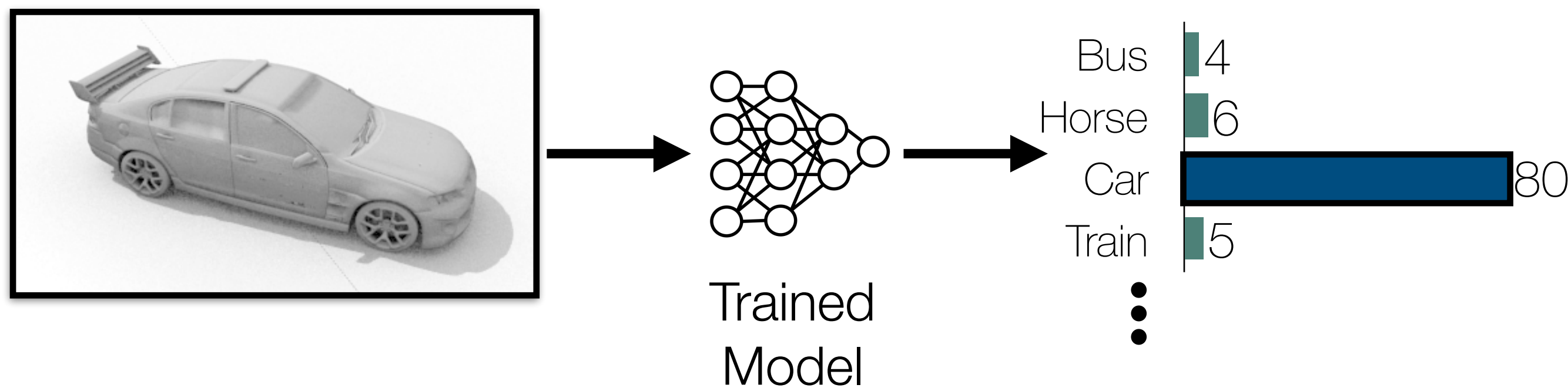
- ✓ Accuracy: Models should make “correct” predictions
- ✓ Calibration: Confidence in predictions should align with true likelihood of correctness



(Visual) Sim2Real Transfer

Reliability: Calibrated Sim2Real Predictions

■ Confidence
■ Accuracy

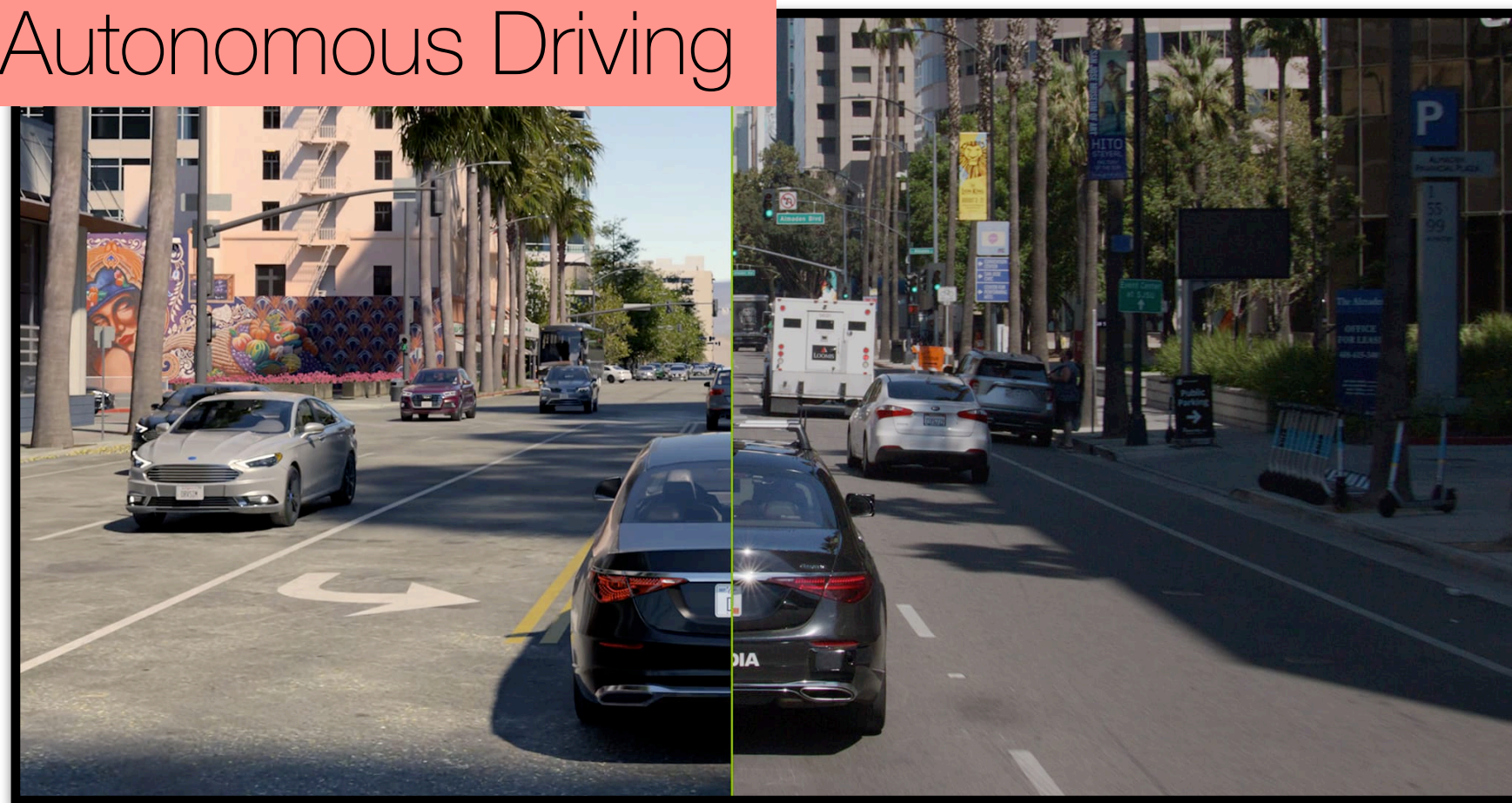


- ✓ Accuracy: Models should make “correct” predictions
- ✓ Calibration: Confidence in predictions should align with true likelihood of correctness

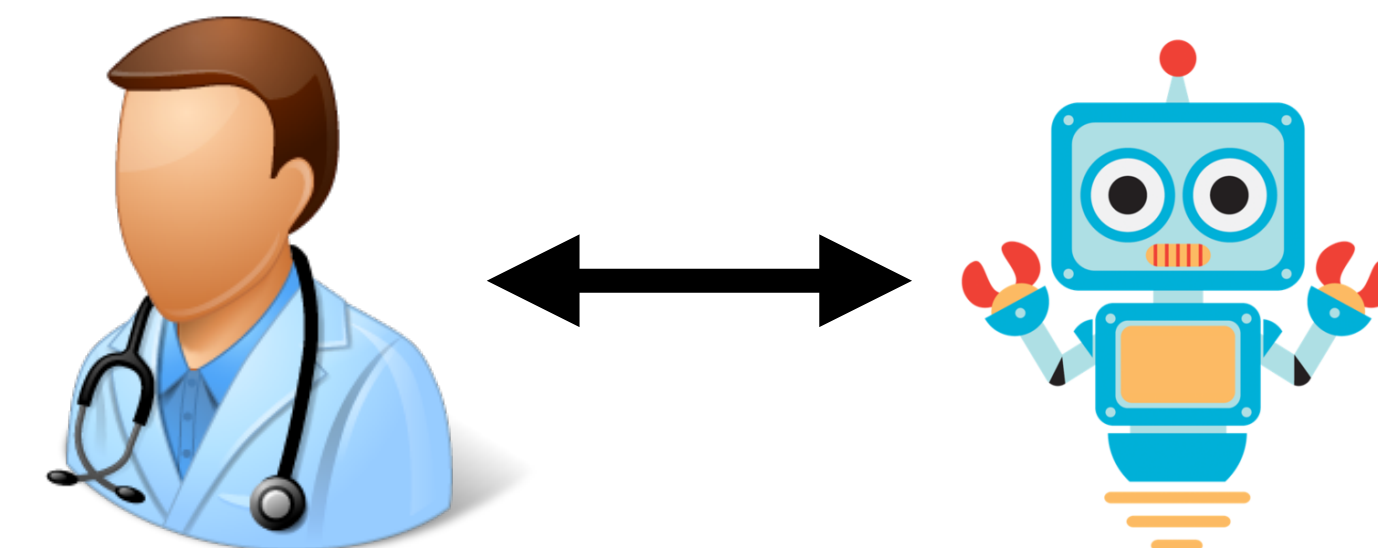
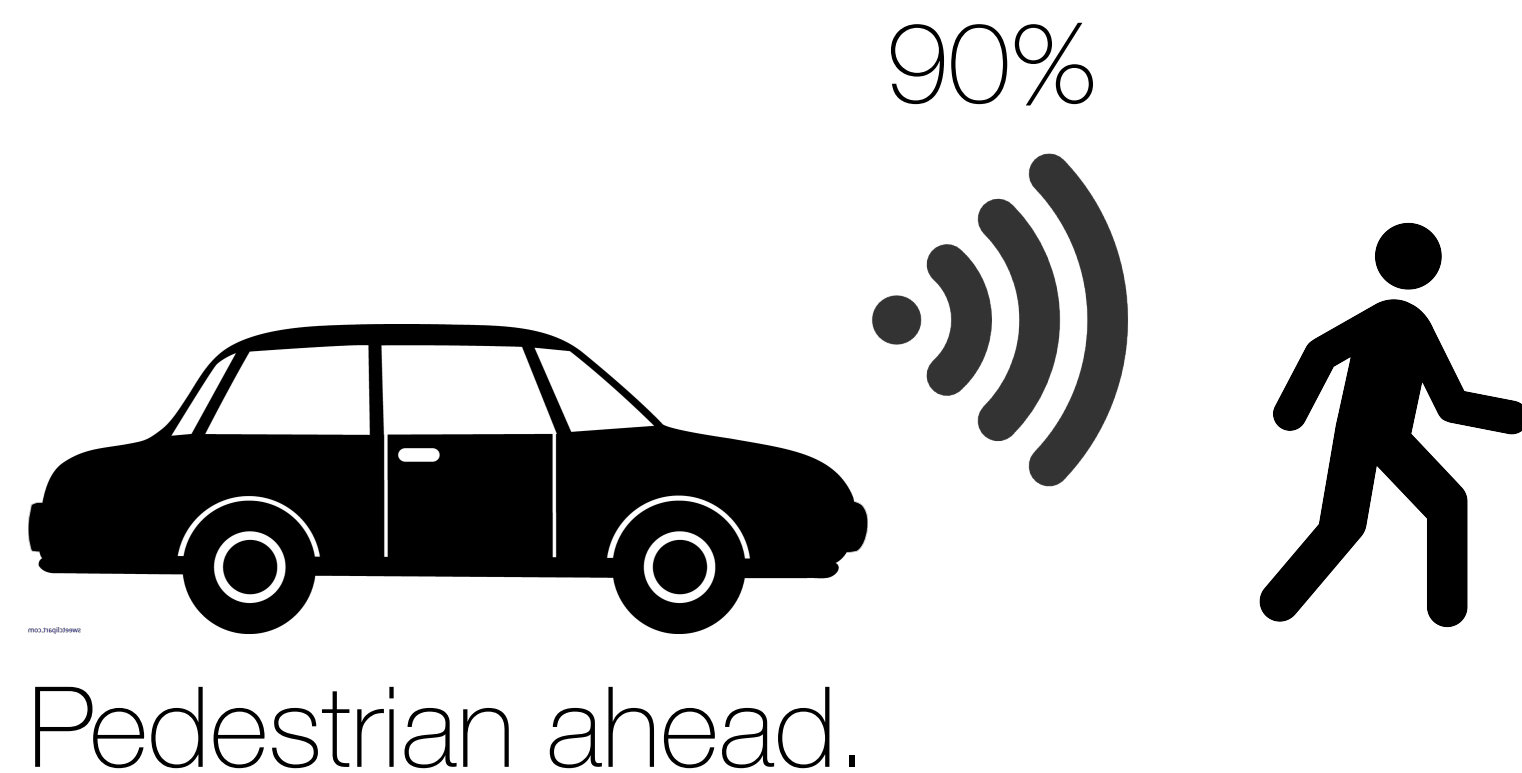
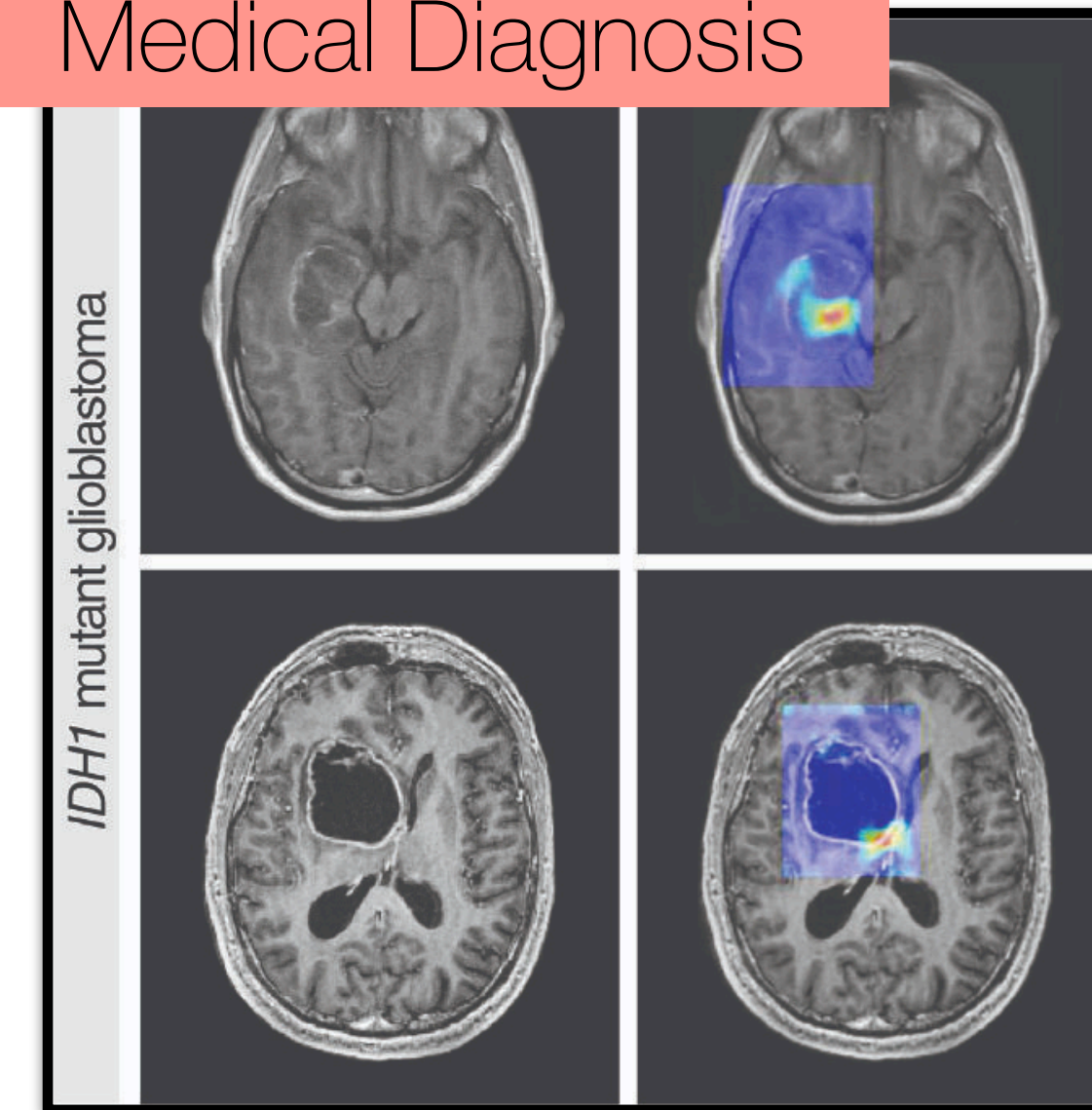
(Visual) Sim2Real Transfer

Reliability: Calibrated Predictions

Autonomous Driving



Medical Diagnosis





(Visual) Sim2Real Transfer

Goal: (Reliable) Sim2Real Adaptation

(Weak) EntMin



Sim2Real Adapt. Predictions

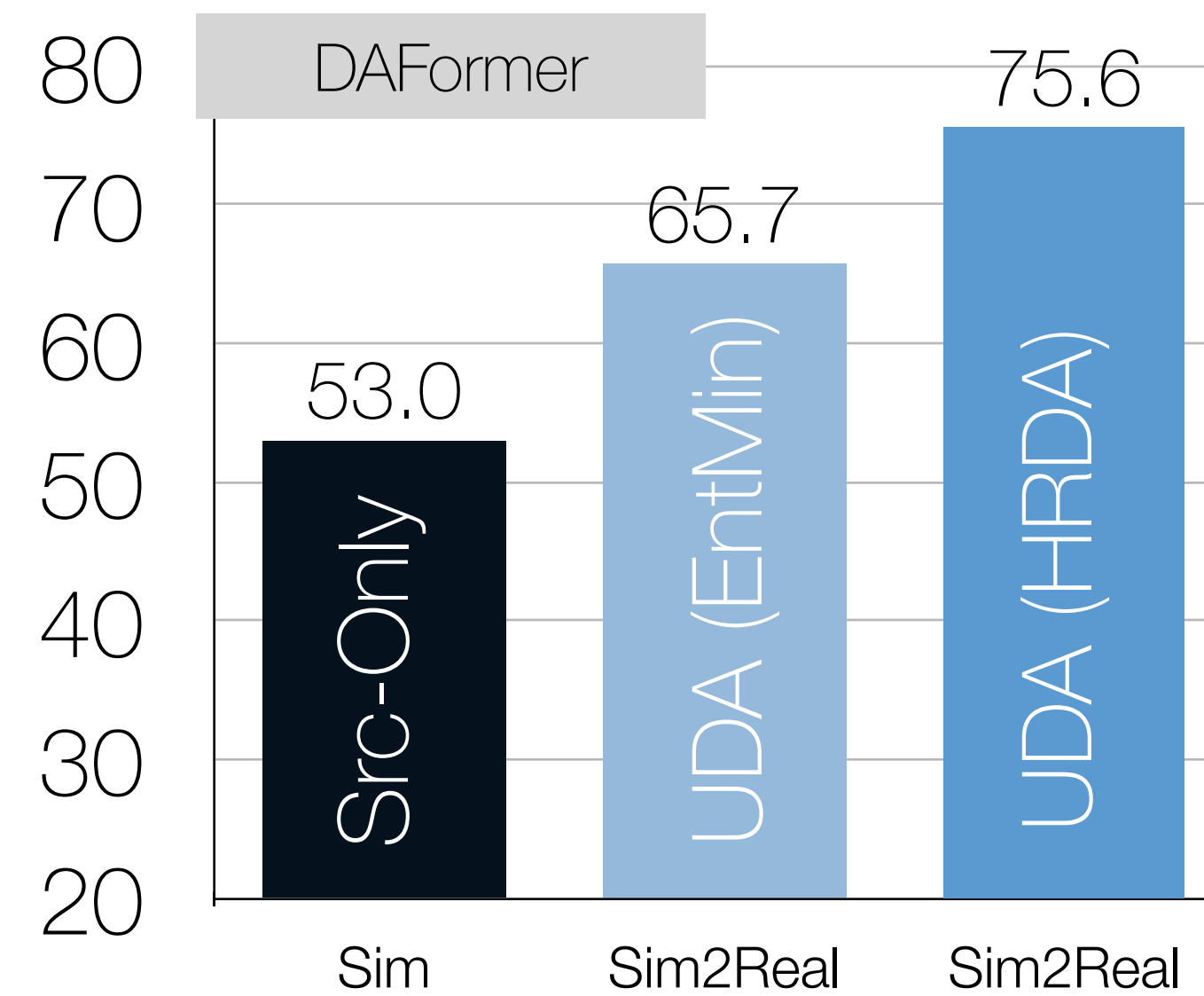


(Strong) HRDA



GTA V → City. (mIoU) ↑

Performance



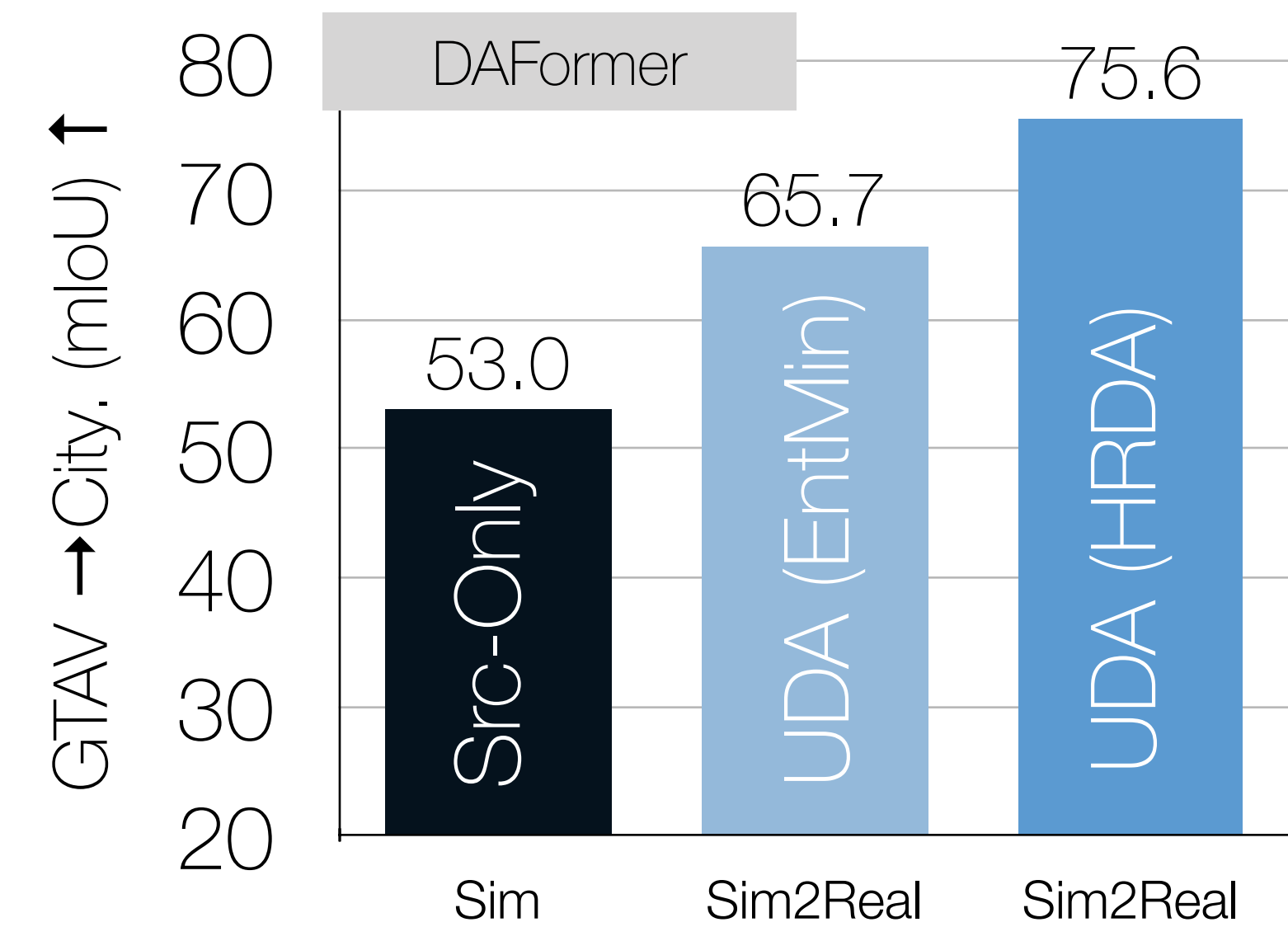
(Desirable) Better Sim2Real adaptation methods have improved performance



(Visual) Sim2Real Transfer

Reliability: Calibrated Sim2Real Predictions

Performance



(Desirable) Better Sim2Real adaptation methods have improved performance

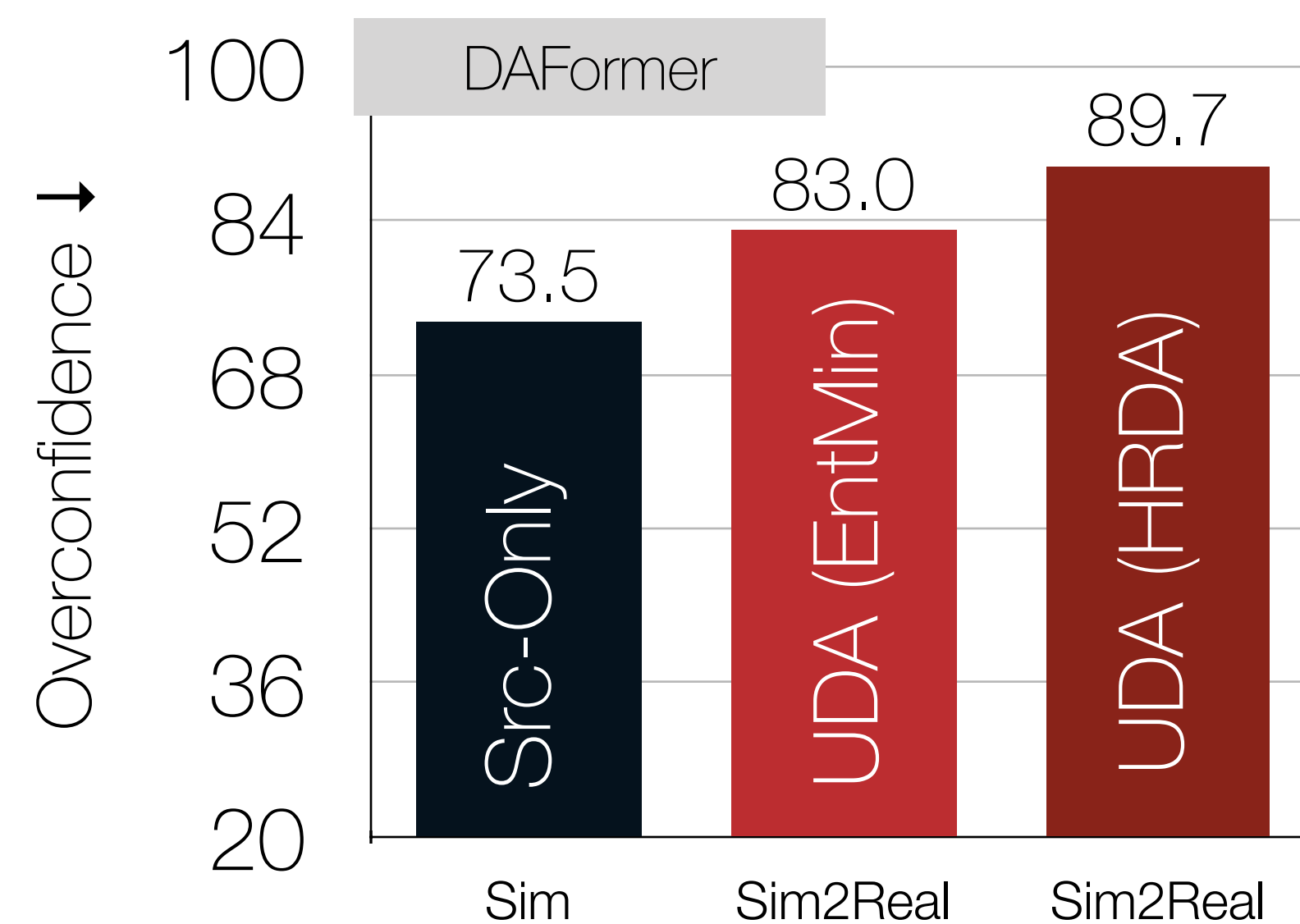
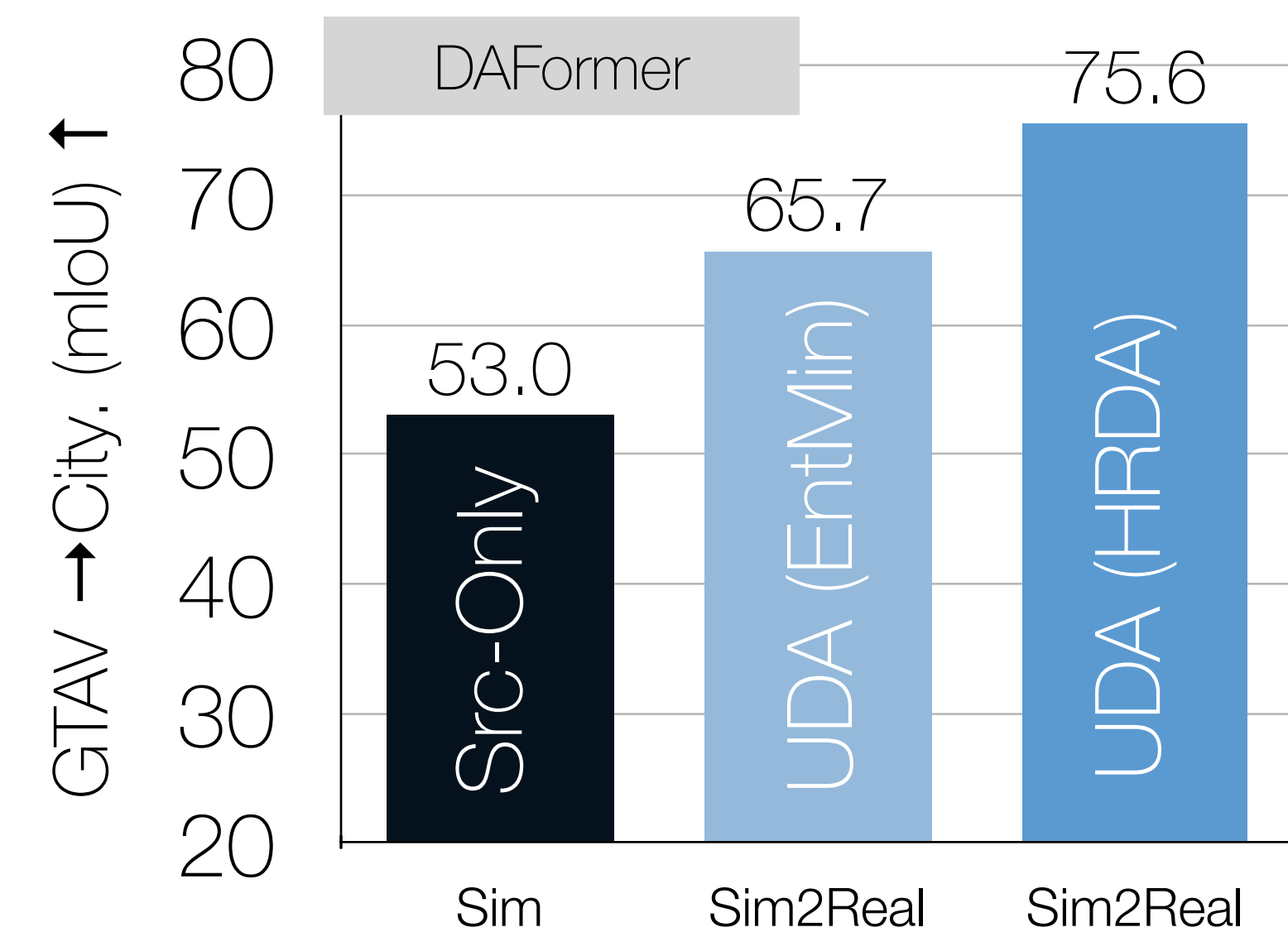


(Visual) Sim2Real Transfer

Reliability: Calibrated Sim2Real Predictions

Performance

(Mispredictions) Calibration



Models routinely overestimate their capabilities

(Desirable) Better Sim2Real adaptation methods have improved performance

(Undesirable) Better Sim2Real adaptation methods make overconfident mistakes



(Visual) Sim2Real Transfer

Reliability: Calibrated Sim2Real Predictions

(Weak) EntMin



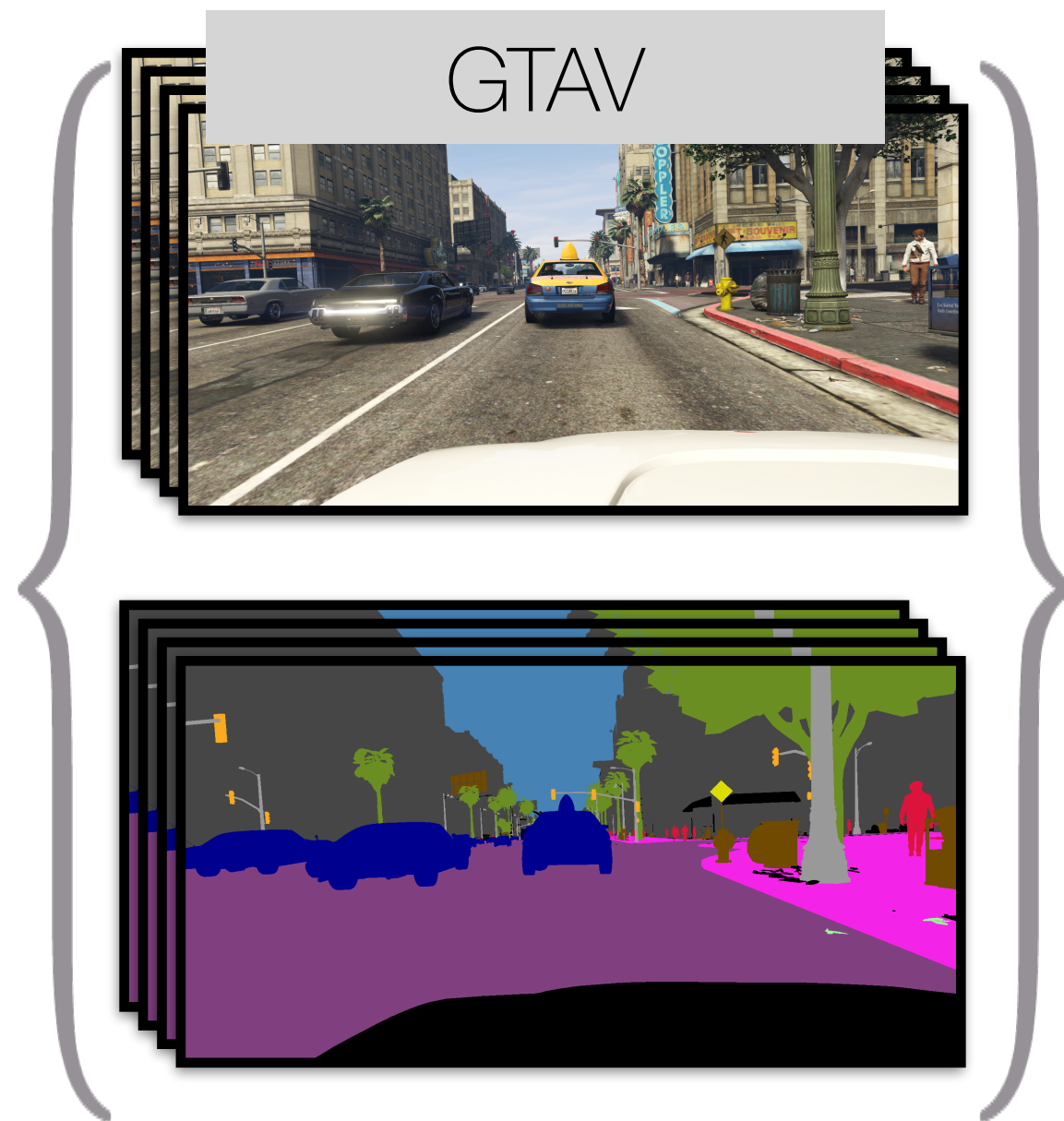
(Strong) HRDA





(Visual) Sim2Real Transfer

Reliability: Why do we need interventions to fix this?



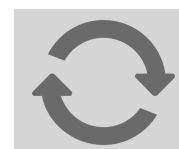
GTA V



Cityscapes

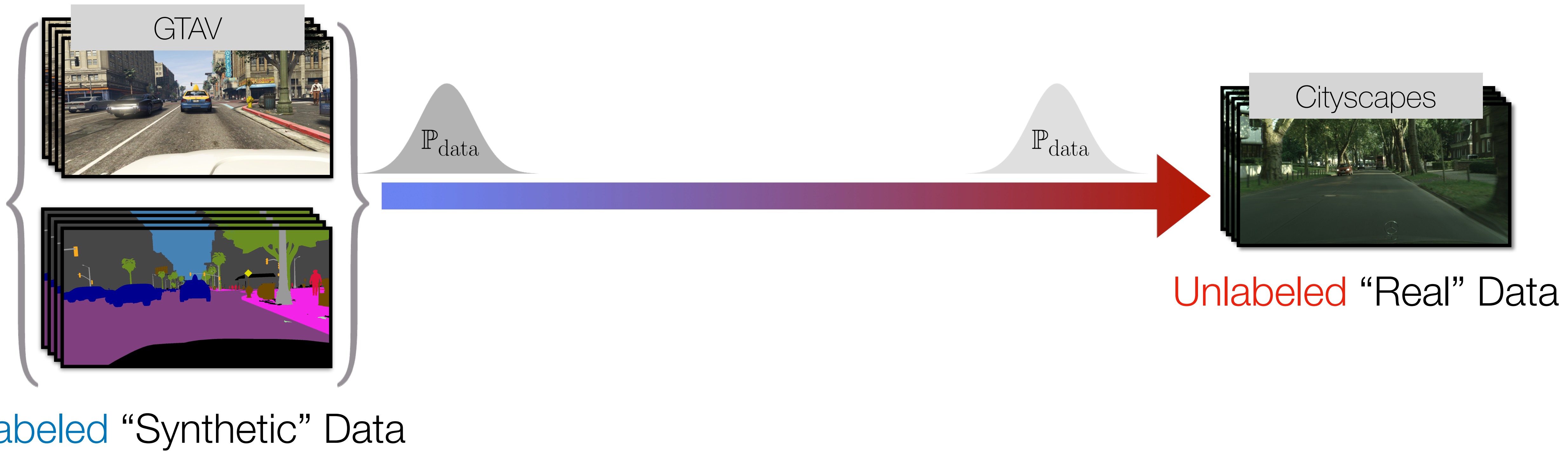
Unlabeled "Real" Data

Labeled "Synthetic" Data



(Visual) Sim2Real Transfer

Reliability: Why do we need interventions to fix this?





(Visual) Sim2Real Transfer

Reliability: Why do we need interventions to fix this?

Calibrating on labeled “synthetic” data is easy because ***you have labels!***



Labeled “Synthetic” Data

Temp. Scaling

Cal. Losses

(Visual) Sim2Real Transfer

Reliability: Why do we need interventions to fix this?

Data you care about (1) **has no labels** and (2) **is dissimilar!**





(Visual) Sim2Real Transfer

Reliability: Why do we need interventions to fix this?



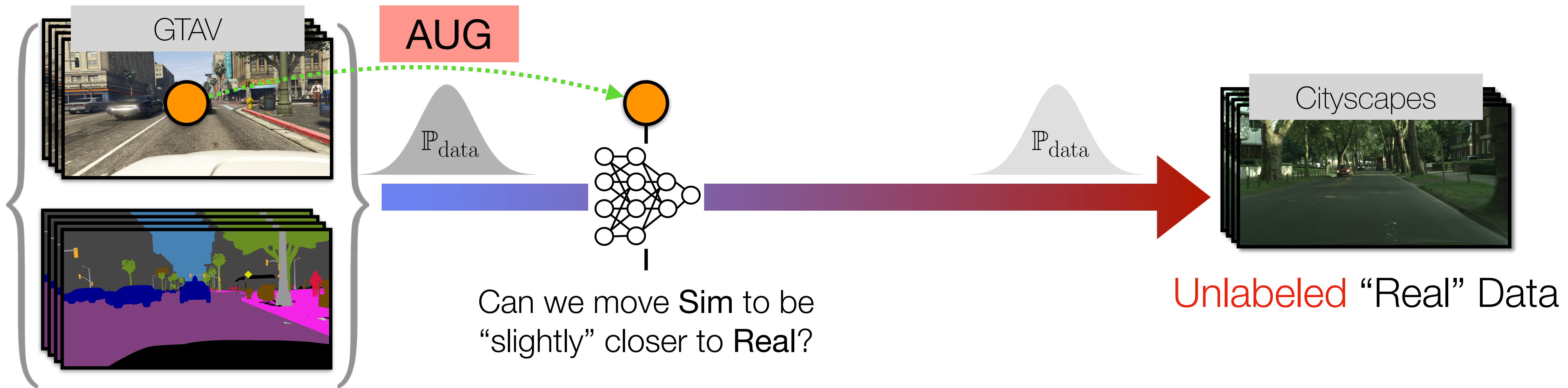
Labeled “Synthetic” Data

$$\underbrace{\mathbb{E}_{P^R} [\mathcal{L}_{\text{CAL}}(x, y)]}_{\text{Real Calibration Error}} \leq \underbrace{\frac{1}{2} d_2(P^R(x) || P^S(x))}_{\text{Syn2Real Distance}} + \underbrace{\mathbb{E}_{P^S} [\mathcal{L}_{\text{CAL}}(x, y)^2]}_{\text{Syn. Calibration Error}}$$

(Visual) Sim2Real Transfer

AUGCAL: "Proactive" Sim2Real Calibration

AUG helps you make the hop



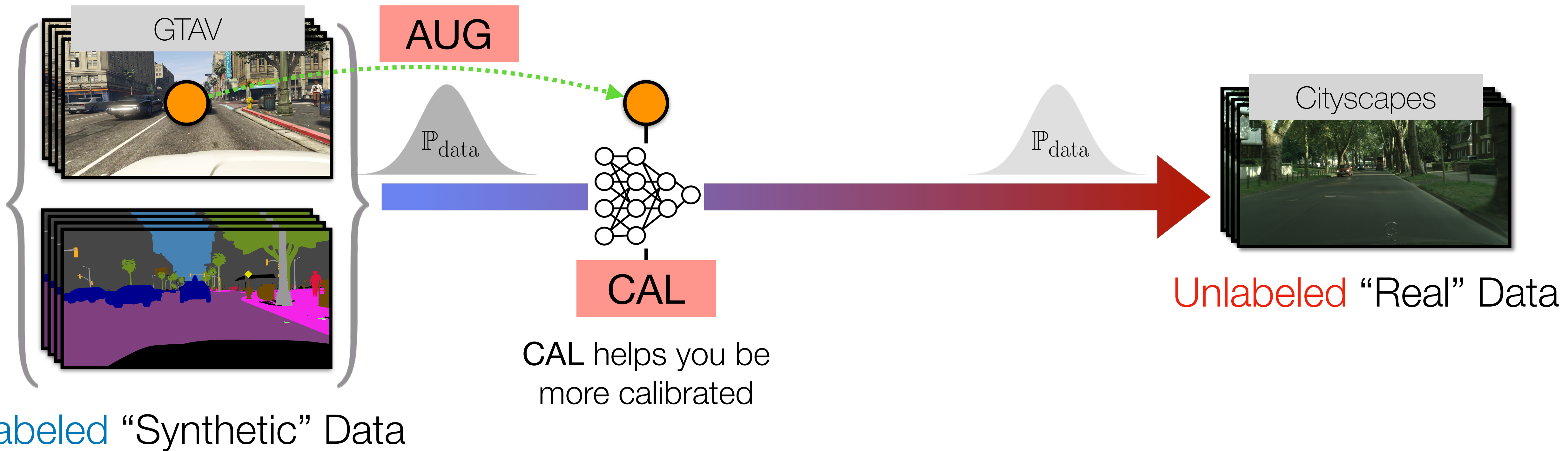
Labeled "Synthetic" Data

Unlabeled "Real" Data

(Visual) Sim2Real Transfer

AUGCAL: "Proactive" Sim2Real Calibration

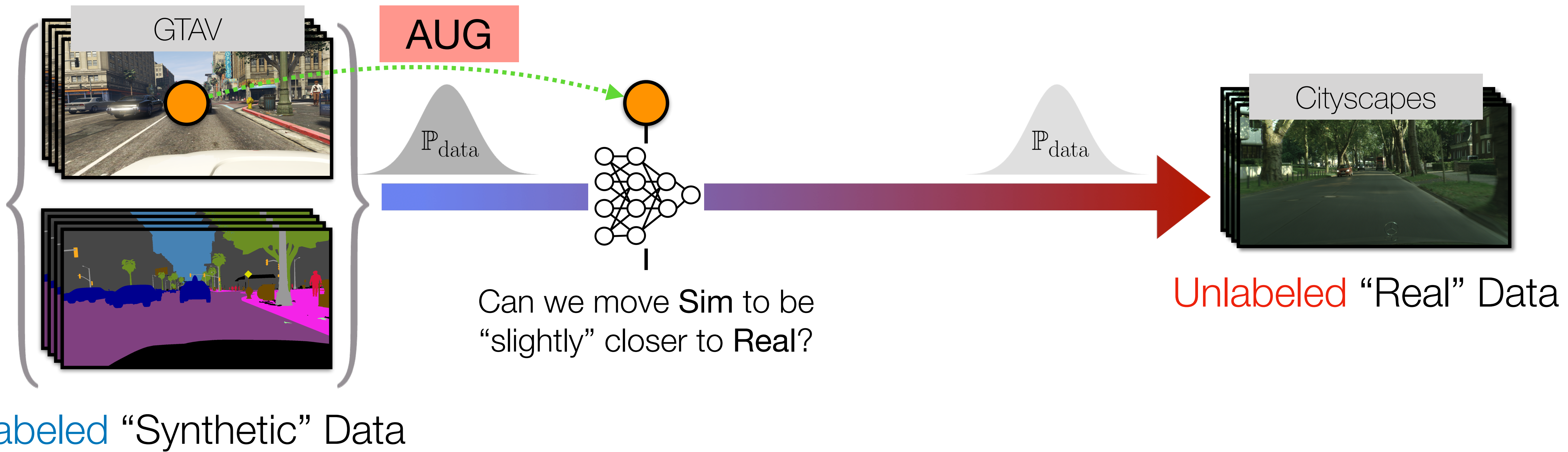
AUG helps you make the hop



(Visual) Sim2Real Transfer

AUGCAL: "Proactive" Sim2Real Calibration

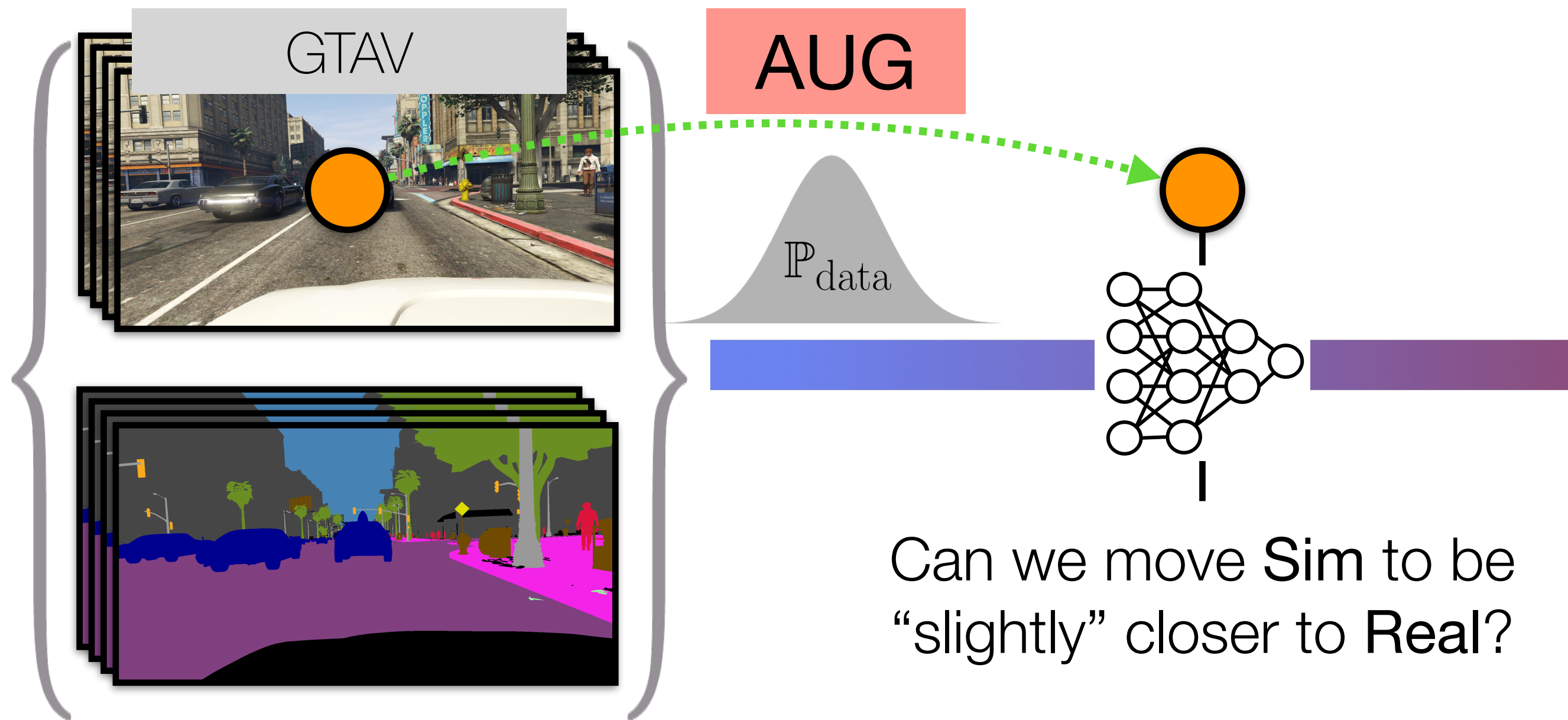
AUG helps you make the hop



(Visual) Sim2Real Transfer

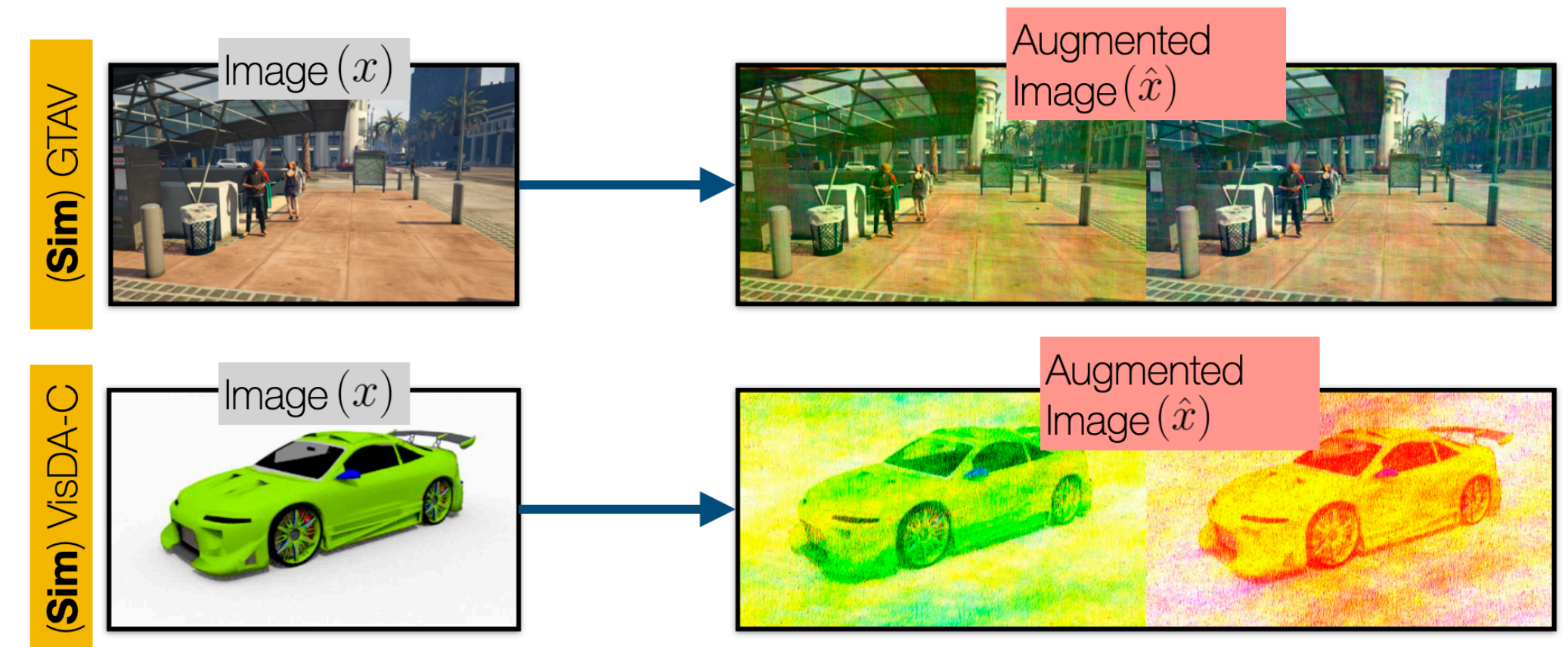
AUGCAL: "Proactive" Sim2Real Calibration

AUG helps you make the hop



Can we move Sim to be "slightly" closer to Real?

Labeled "Synthetic" Data



Perhaps an augmentation that aids Sim2Real generalization

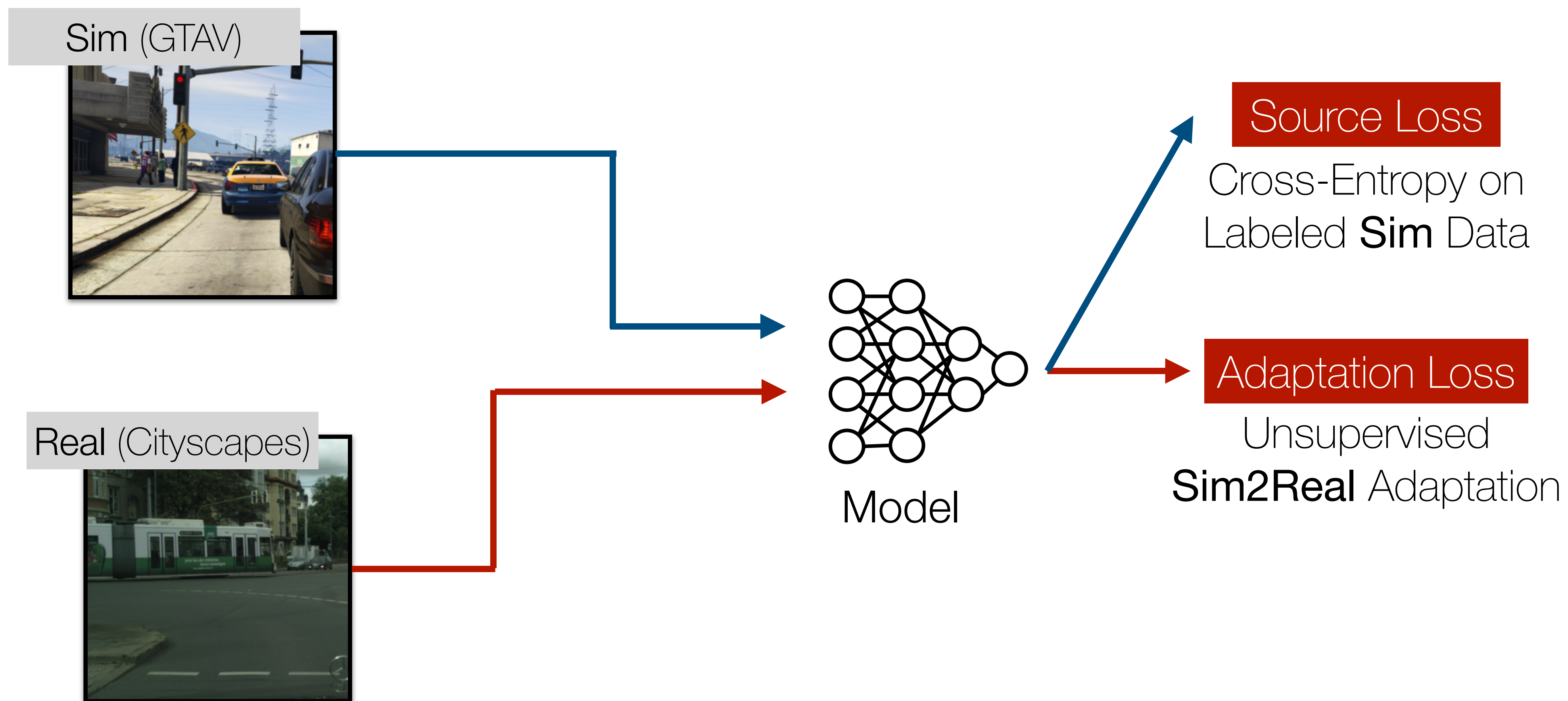
PASTA [C*S*VH ICCV'23]

RandAugment (Cubuk et. al, CVPRW 2020)



(Visual) Sim2Real Transfer

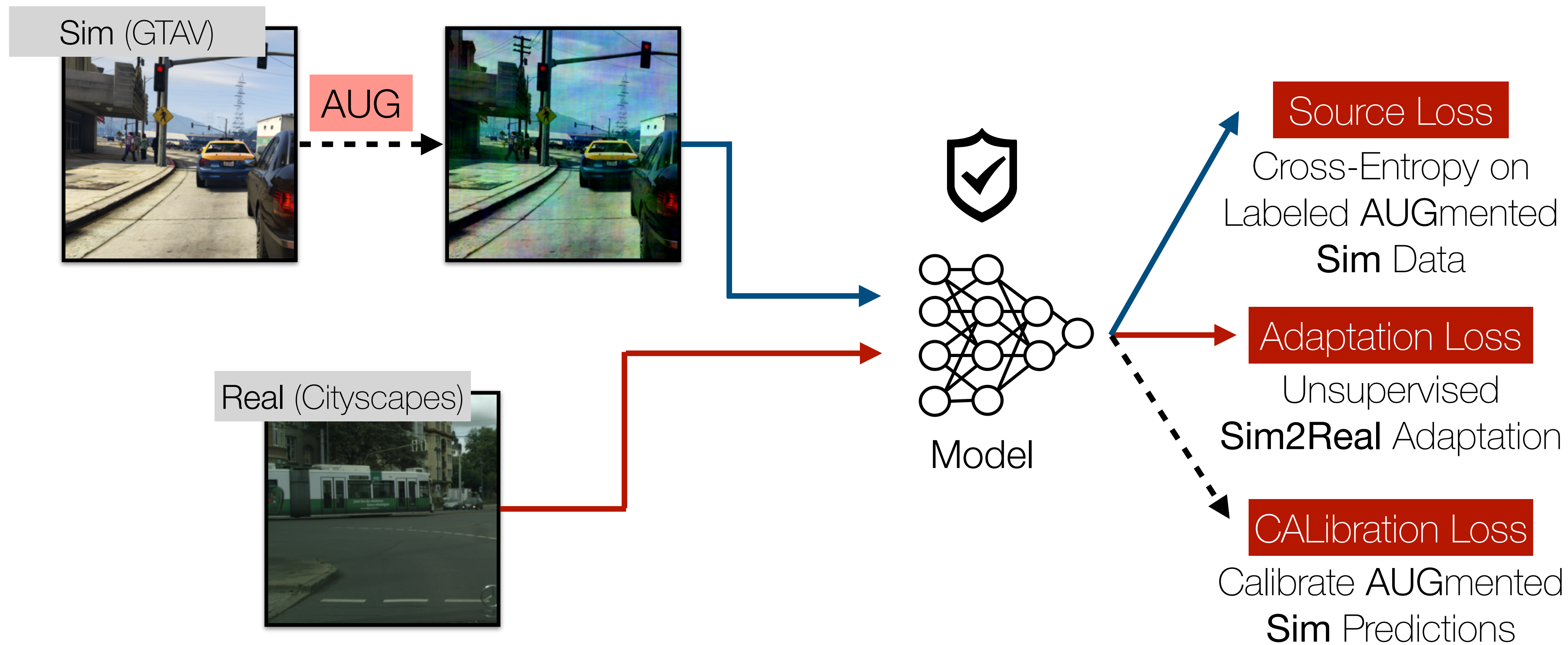
AUGCAL: "Proactive" Sim2Real Calibrated Adaptation





(Visual) Sim2Real Transfer

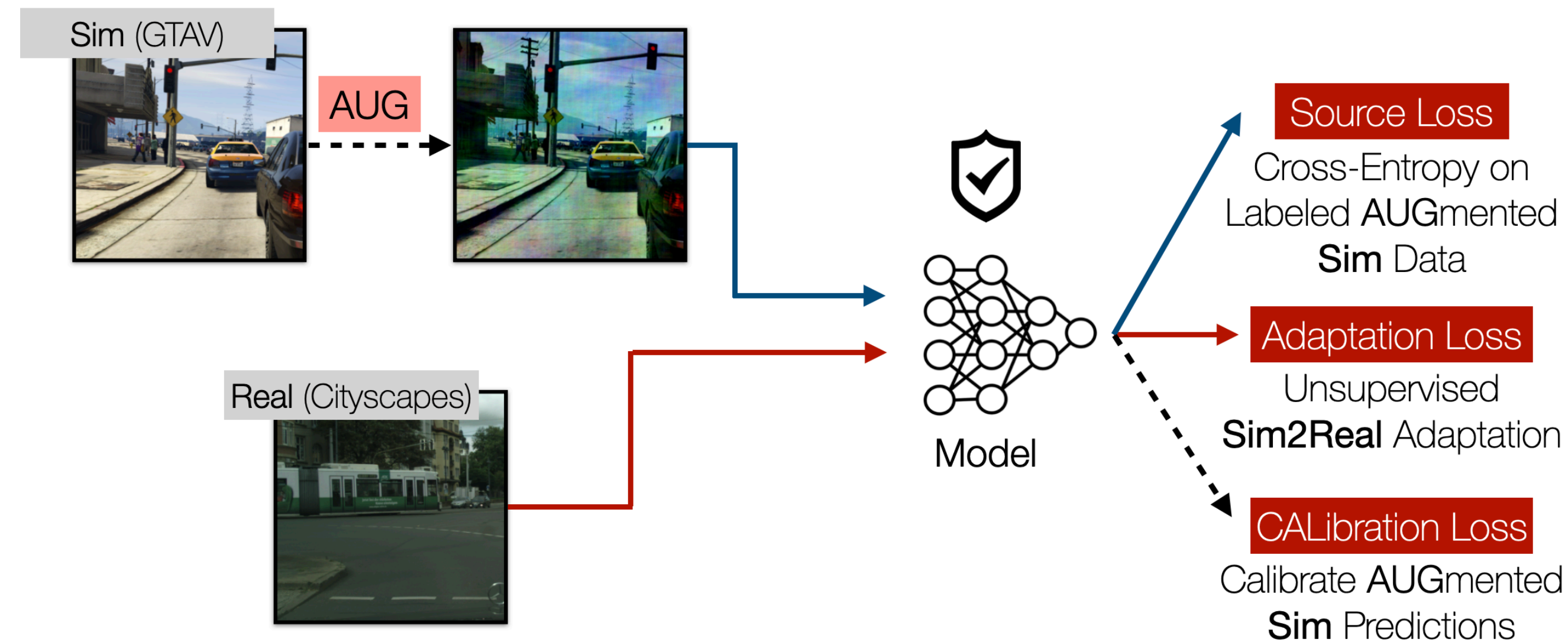
AUGCAL: "Proactive" Sim2Real Calibrated Adaptation



(Visual) Sim2Real Transfer

AUGCAL: "Proactive" Sim2Real Calibrated Adaptation

General Pipeline



Calibration Losses

Modulate your classification scores at a minibatch level to be more calibrated



DCA (Liang et. al, BMVC 2020)

MDCA (Hebbalogue et. al, CVPR 2022)

MbLS (Liu et. al, CVPR 2022)



(Visual) Sim2Real Transfer

AUGCAL: "Proactive" Sim2Real Calibrated Adaptation

(Weak) EntMin



w/o AUGCAL



Accurate and Certain Predictions

(Strong) HRDA



(Visual) Sim2Real Transfer

AUGCAL: "Proactive" Sim2Real Calibrated Adaptation

(Weak) EntMin



w/o AUGCAL



w AUGCAL



Accurate and Certain Predictions

(Strong) HRDA





(Visual) Sim2Real Transfer

AUGCAL: “Proactive” Sim2Real Calibrated Adaptation

VisDA-C Syn → Real



Object Recognition

GTAV → Cityscapes



Semantic Segmentation

- **Retained or Improved (~5 mIoU for SemSeg)** performance for adaptation methods!
- **Reduced Expected Calibration Error (ECE)** — more calibrated predictions!
- **Reduced overconfidence** in mispredictions!



(Visual) Sim2Real Transfer

AUGCAL: “Proactive” Sim2Real Calibrated Adaptation

VisDA-C Syn → Real



Object Recognition

GTAV → Cityscapes



Semantic Segmentation

- **Retained or Improved (~5 mIoU for SemSeg)** performance for adaptation methods!
- **Reduced Expected Calibration Error (ECE)** — more calibrated predictions!
- **Reduced overconfidence** in mispredictions!
- BONUS: **Better mis-classification detection** with AUGCAL!



(Visual) Sim2Real Transfer

AUGCAL: “Proactive” Sim2Real Calibrated Adaptation

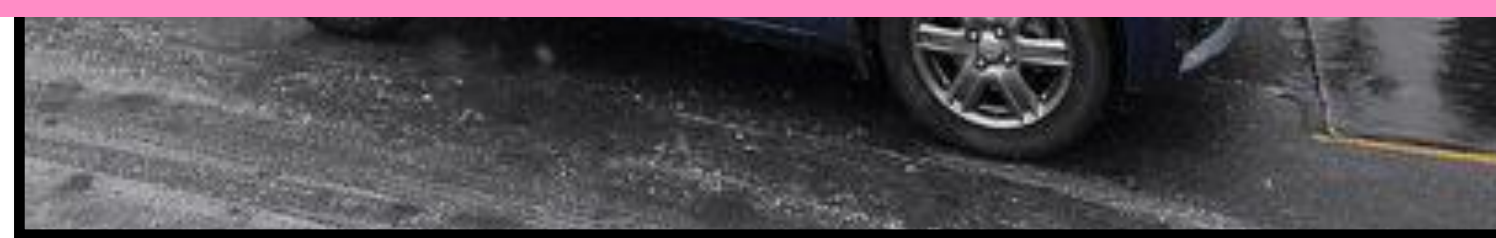
VisDA-C Syn → Real



GTAV → Cityscapes



AUGCAL, a training time patch to improve Sim2Real adaptation

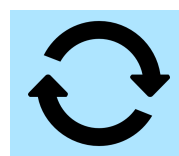


Object Recognition



Semantic Segmentation

- ✓ Performance Retention: Sim2Real adaptation performance should be unaffected
- ✓ Reduced Miscalibration & Overconfidence: Less overconfident mistakes
- ✓ Improved Reliability: Can calibrated confidence scores guide misclassification detection



Thank You! Questions?