KI DELTA LEARNING
Scalable AI for Automated Driving

Final Event | March 09, 2023

Auxiliary Task-Guided CycleGAN for Black-Box Model Domain Adaptation

Michael Brunner

Neural network training needs a lot of data

- Task specific annotations are required for supervised learning
- Annotating data is a very time-consuming and therefore expensive task
- Very limited or even no data for dangerous and rare scenarios (corner cases)

https://cocodataset.org/#keypoints-2020
https://cocodataset.org/#panoptic-2020

Simulation

- These issues can be tackled with simulation
  - Fully observable environment
  - Ground truth available for free
  - Source of (almost) unlimited data
  - Corner cases can be explicitly addressed
- Sim vs real: Domain shift
- Domain adaptation to compensate the domain shift

Image source: J. Hoffman et al., "CyCADA: Cycle-Consistent Adversarial Domain Adaptation," in Proceedings of the 35th International Conference on Machine Learning, 2018.

Our contribution

- Usually, existing domain adaptation (DA) methods
  - are targeted at specific tasks
  - require access to source model's parameters

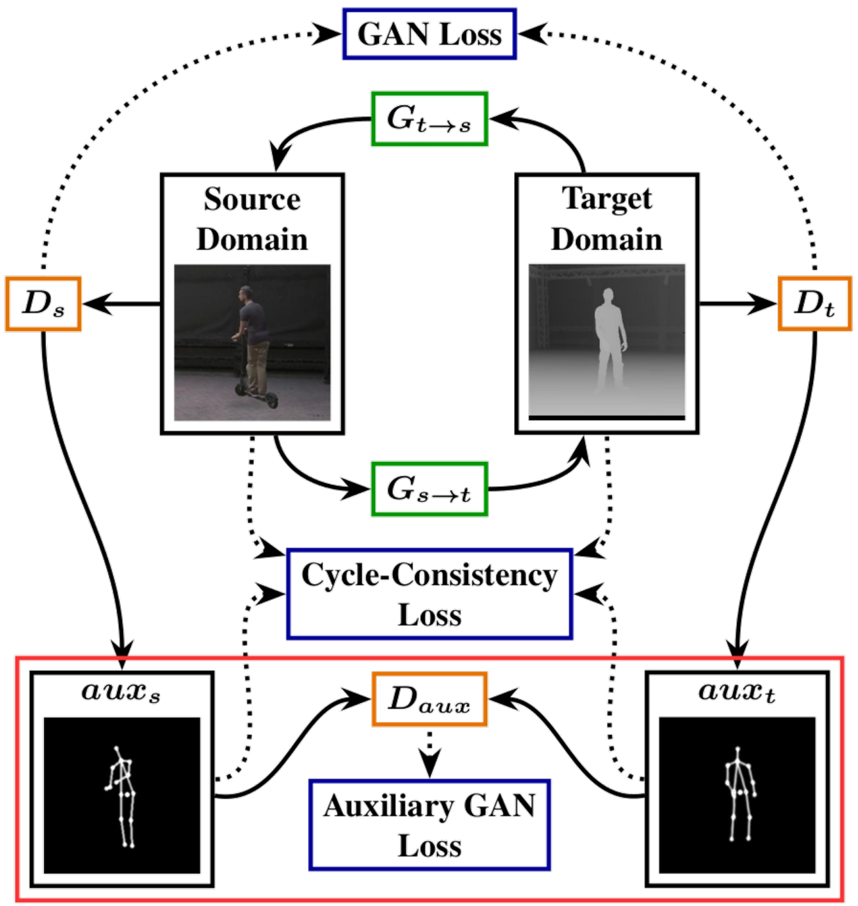
    → Major drawback when only a black-box model is available

- We perform unsupervised domain adaptation (UDA) for black-box models
  - Regression, i.e., human pose estimation, instead of classification as done in [1, 2]
  - Sim-to-real and cross-sensor transfer
  - Comparison with CycleGAN [3] and RegDA [4] under varying domain shifts

[1] H. Zhang et al., "Unsupervised domain adaptation of black-box source models," in 32nd British Machine Vision Conference (BMVC), 2021.
[2] J. Liang et al., "DINE: Domain adaptation from single and multiple black-box predictors," in 2022 IEEE/CVF Conference on Computer Vison and Pattern Recognition (CVPR), 2022.
[3] J.-Y. Zhu et al., "Unpaired image-to-image translation using cycle-consistent adversarial networks," in 2017 IEEE International Conference on Computer Vision (ICCV), 2017.
[4] J. Jiang et al., "Regressive domain adaptation for unsupervised keypoint detection," in 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021.

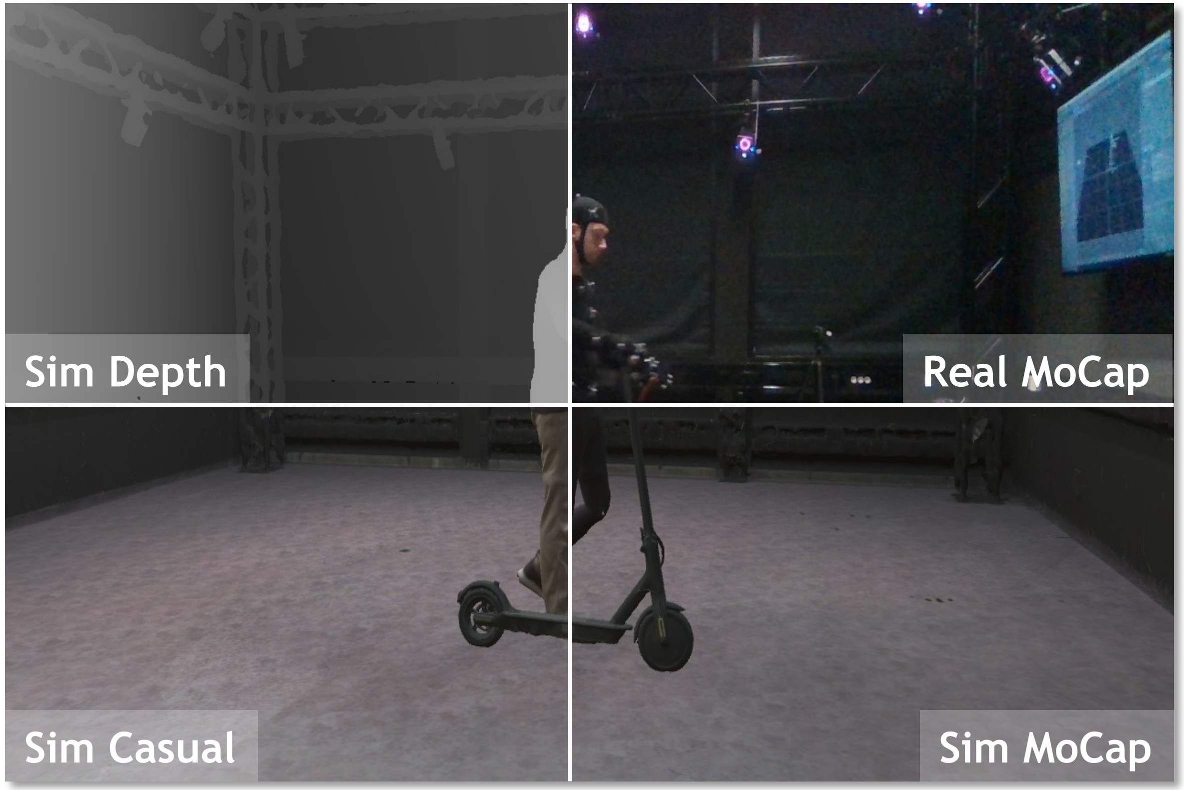
Our Method [1]

- Based on CycleGAN → unpaired data
- Auxiliary task to support transfer across domains
- No target domain labels → unsupervised DA
- Auxiliary task is learned by the discriminators
- D_{aux} to support transfer of auxiliary task from source to target domain



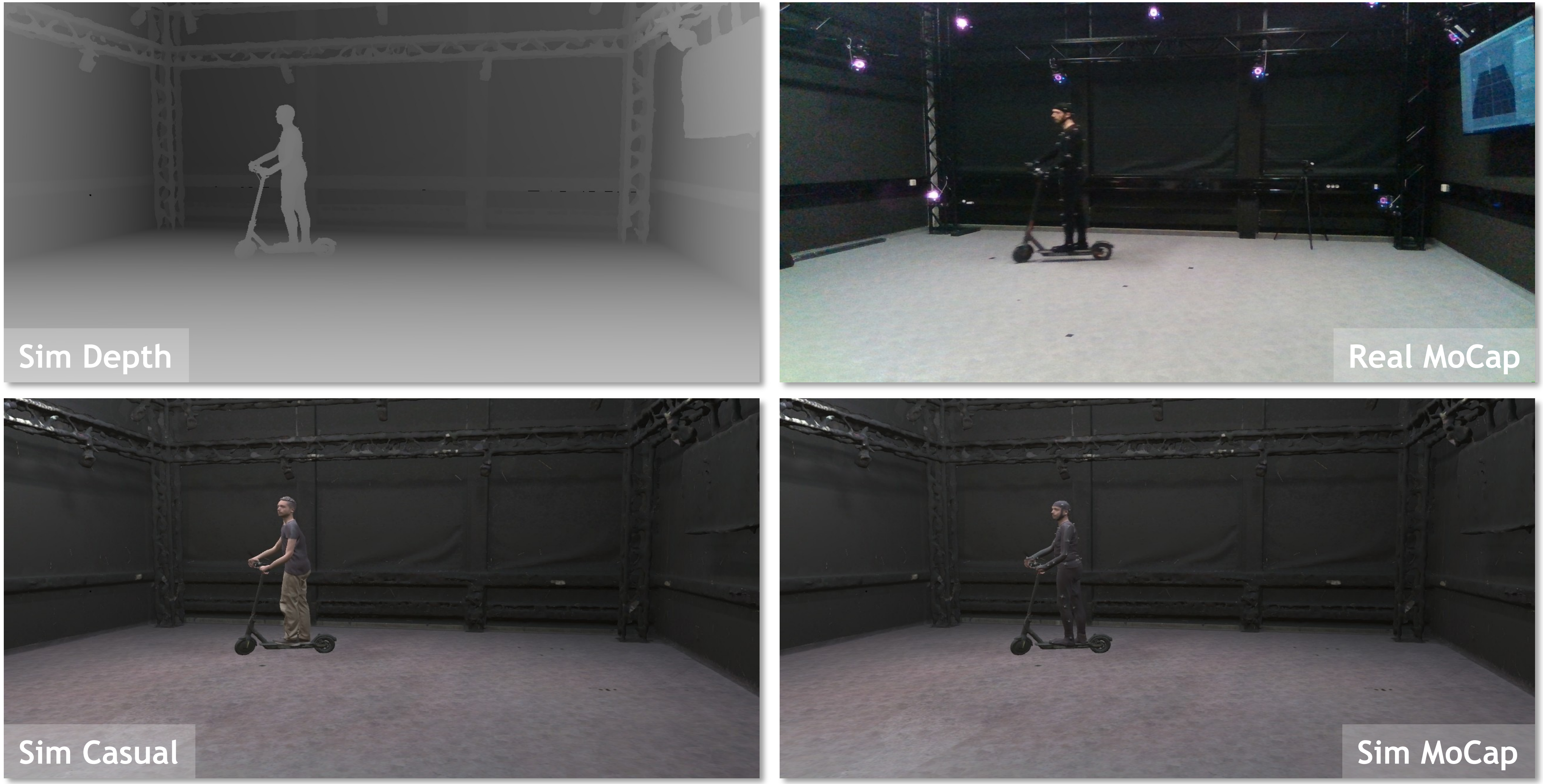
[1] M. Essich, M. Rehmann, and C. Curio, "Auxiliary Task-Guided CycleGAN for Black-Box Model Domain Adaptation," in Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), 2023.

Synchronized and paired sim-to-real and cross-sensor dataset



- Targeted at sim-to-real and cross-sensor DA
- Synchronized and paired data was recorded with our motion capture system
- Pose distributions are kept the same across domains
- Paired data is only used for validation

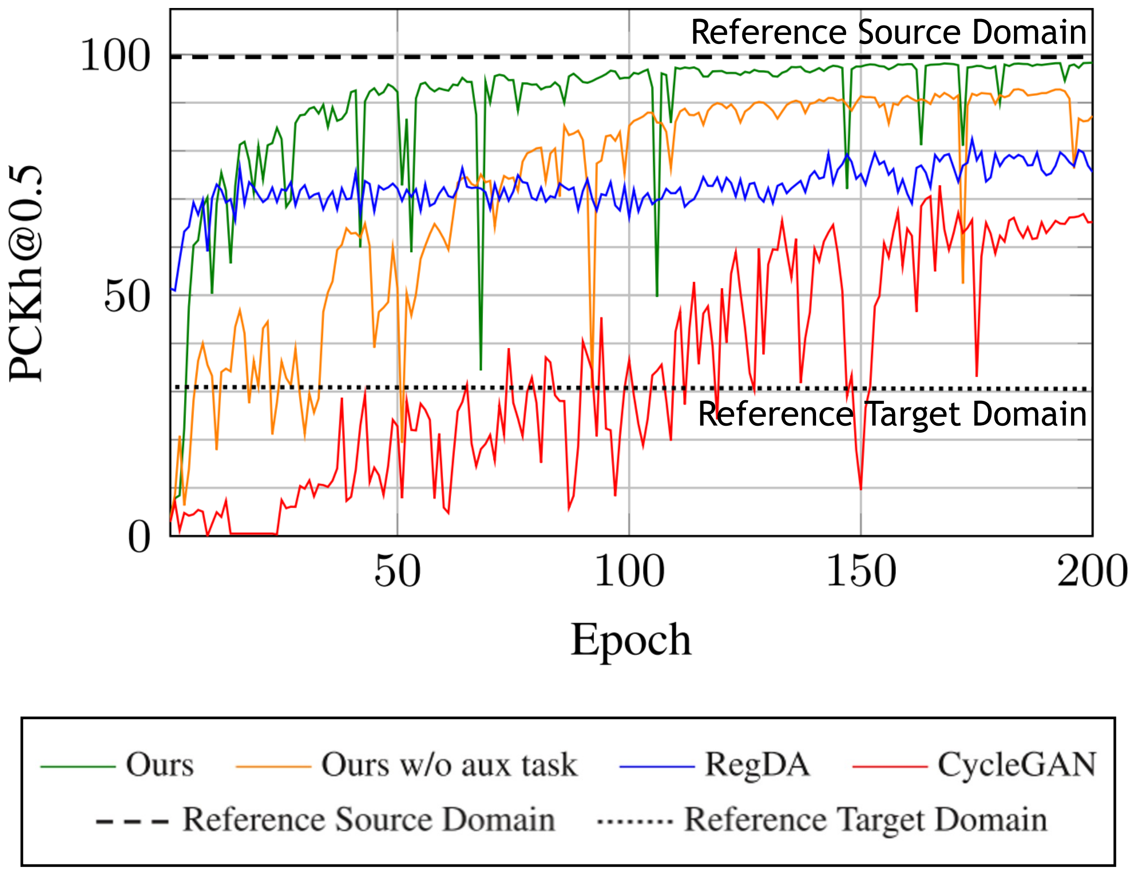
Synchronized Real and Simulated Data



Experiment Setting 1

Small variation in avatar appearance

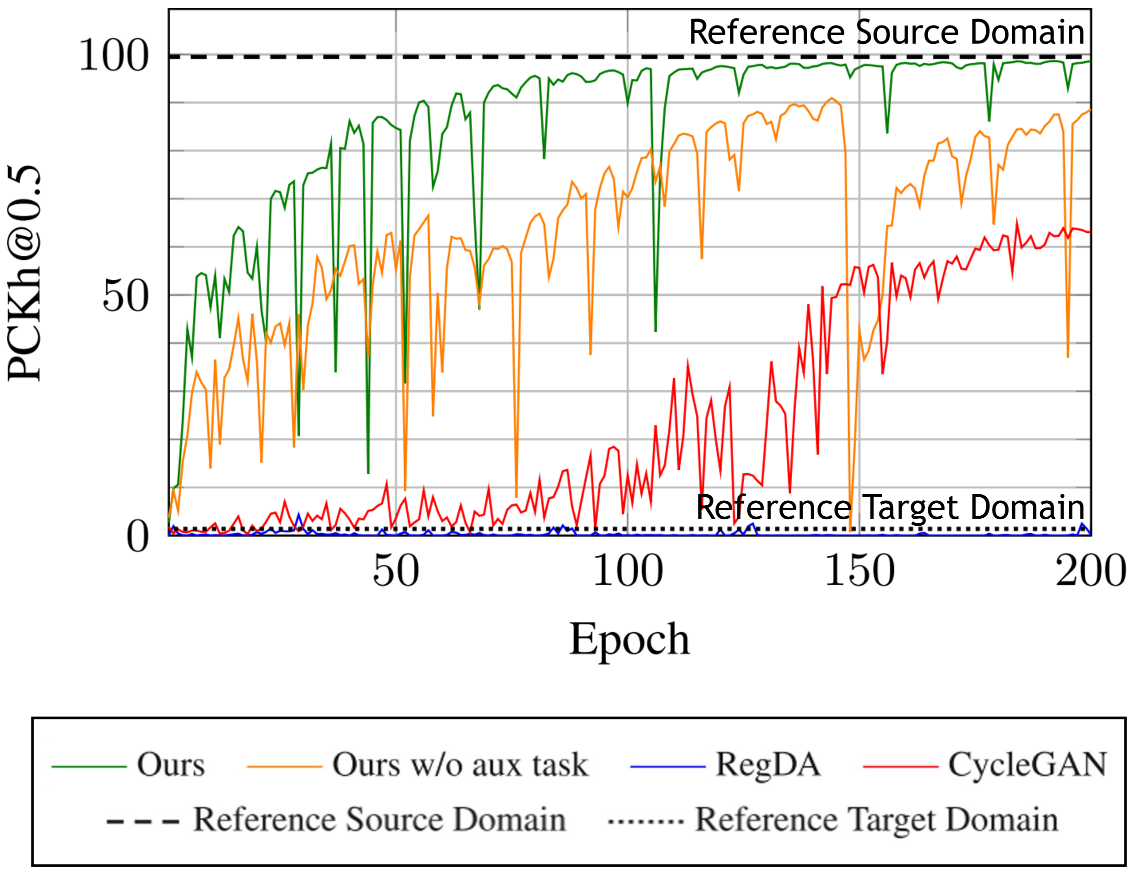
- A cyclical learning rate increases DA performance with CycleGAN (orange)
- Our (green) auxiliary task
  - further increases DA performance
  - leads to a more stable training process



Experiment Setting 2

Different sensors

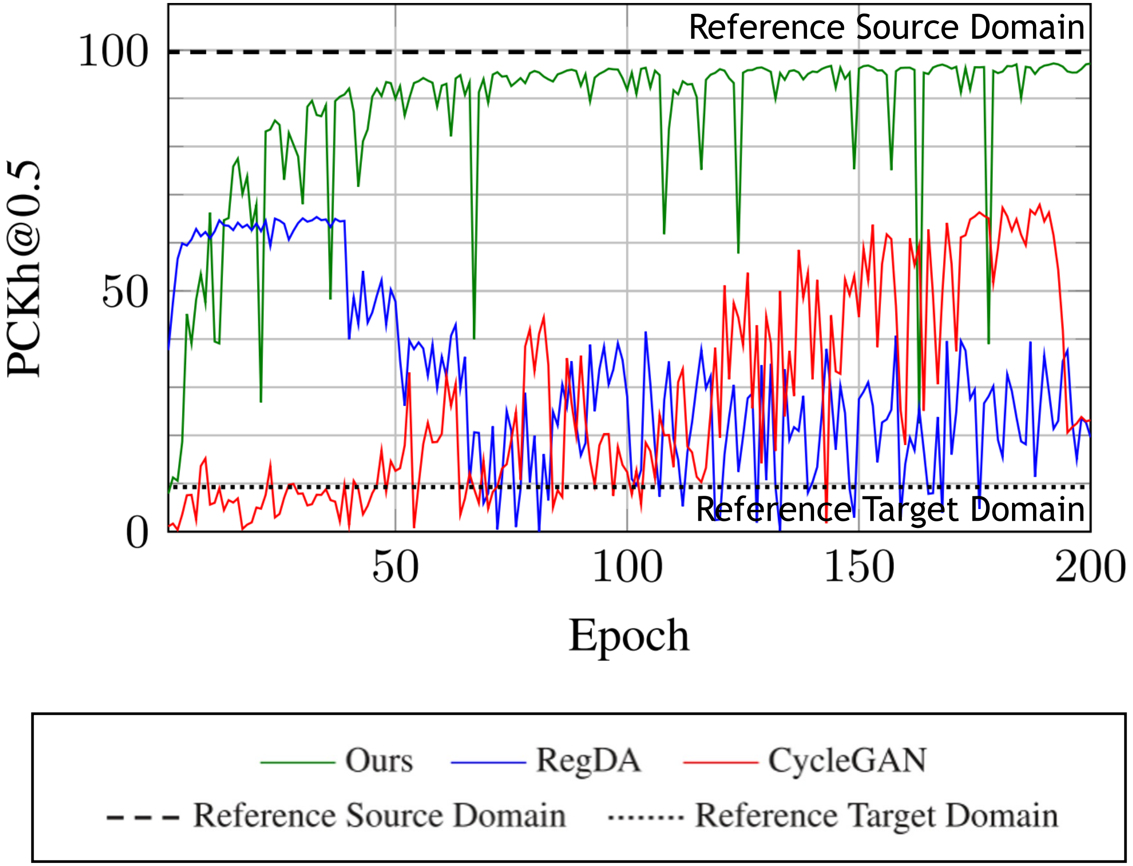
- RegDA (blue) requires the pose estimation's predictions on target domain
- RegDA (blue) struggles with the large sensor domain shift



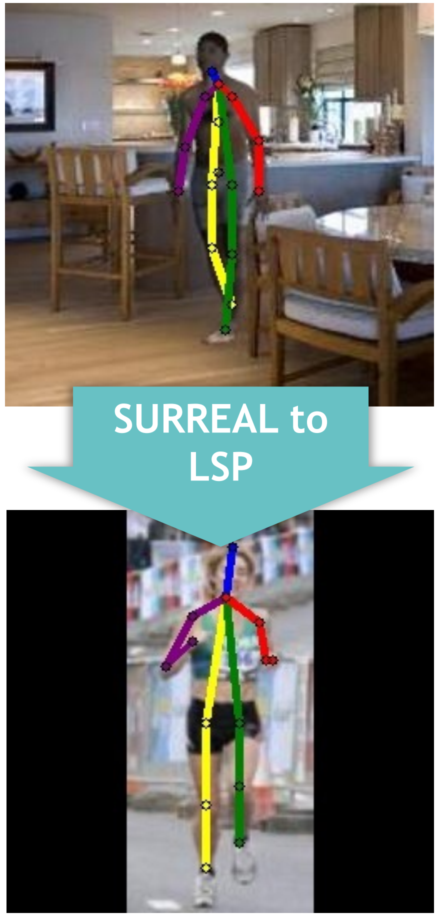
Experiment Setting 3
Sim-to-real

- RegDA's performance (blue) depends on the pose estimation's predictions on target domain

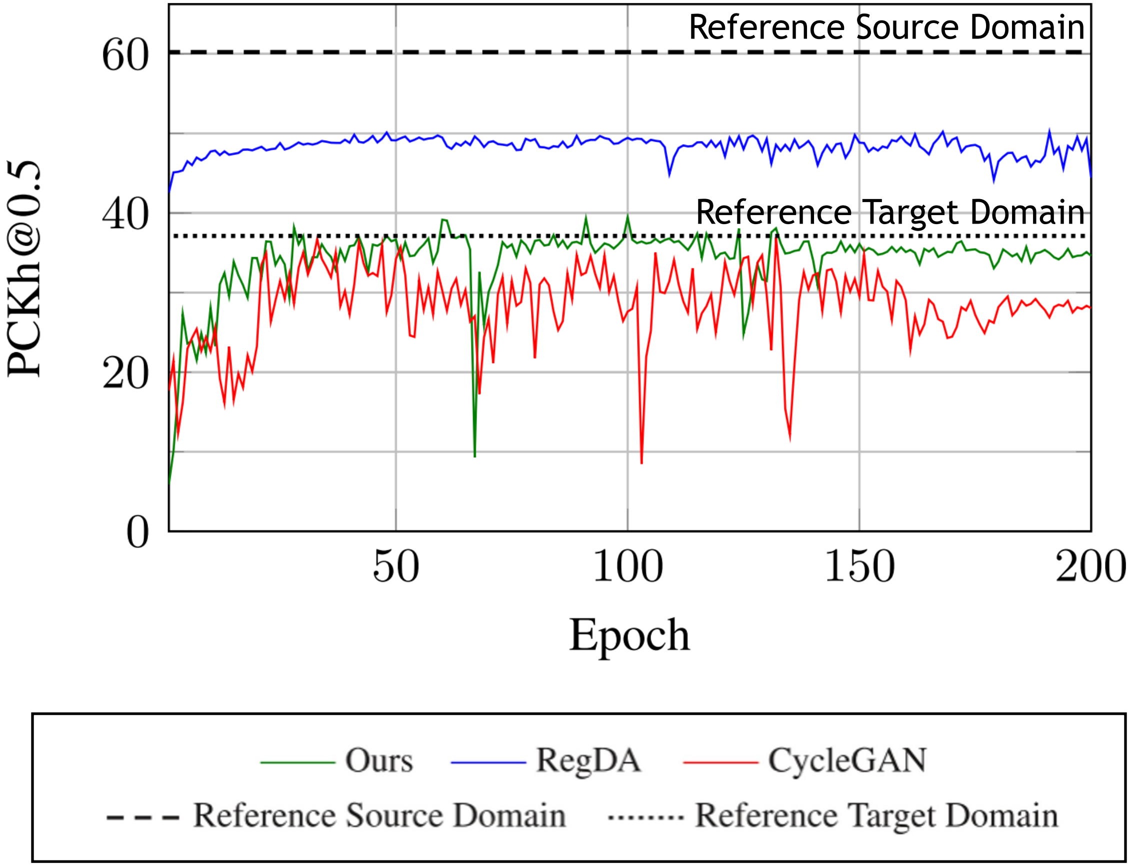
- Sim-to-real adaptation successfully handled with our approach (green)



Experiment Setting 4
Different datasets



- Source model already achieves a high performance on target domain
  → RegDA (blue) performs well
- One/many-to-many mappings occur in this setting
  → Limitation of our approach (green) and CycleGAN (red)



Conclusion

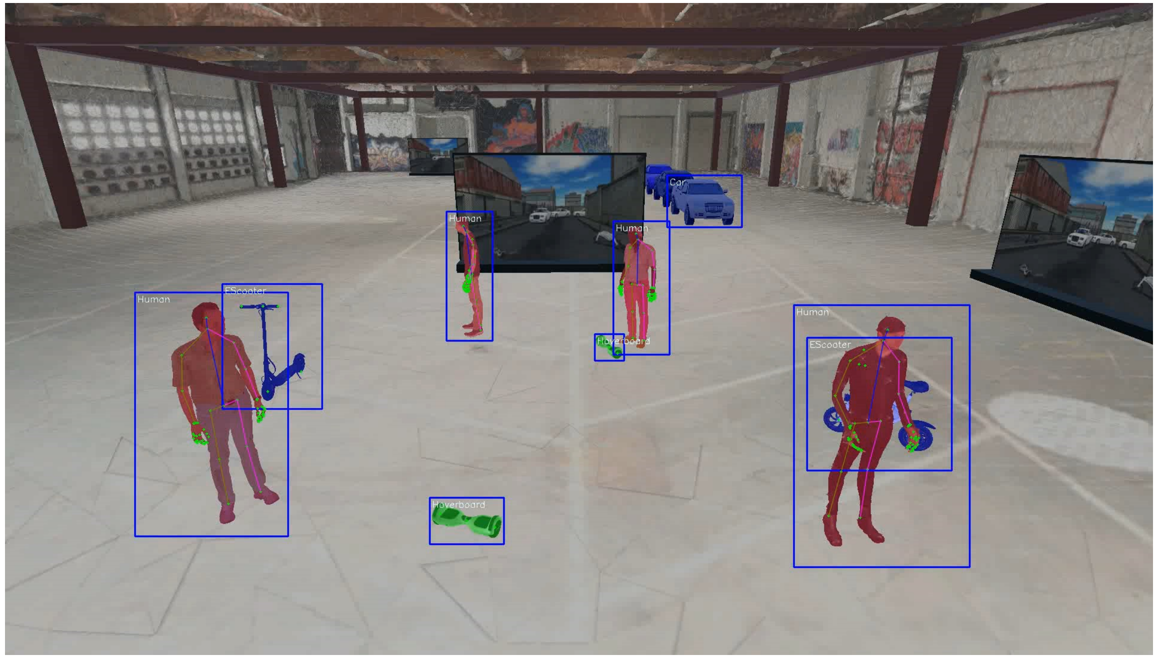
- We compared our approach with RegDA and CycleGAN
- Four settings with varying domain shift
- We emphasize the necessity for explicitly addressing sensor domain shift

Outlook

- One/many-to-many mappings are a current limitation of our approach
  → Addressable by, e.g., disentangling style and content [1]
- Follow-up work: Sim-to-real research in large-scale motion capture space

[1] X. Huang, M.-Y. Liu, S. Belongie, and J. Kautz, “Multimodal Unsupervised Image-to-Image Translation,” in Computer Vision – ECCV 2018, Cham, 2018, pp. 179–196.

Human-centered Interactive
Artificial Intelligence Data-Incubation Center (AIDA)



Co-funded by the European Union

Baden-Württemberg

KIDELTA LEARNING
Scalable AI for Automated Driving

Michael Brunner | Reutlingen University

michael.brunner@reutlingen-university.de

KI Delta Learning is a project of the KI Familie. It was initiated and developed by the VDA Leitinitiative autonomous and connected driving and is funded by the Federal Ministry for Economic Affairs and Climate Action.

KI FAMILIE

Supported by:
Federal Ministry for Economic Affairs and Climate Action
on the basis of a decision by the German Bundestag

www.ki-deltalearning.de @KI_Familie KI Familie