Domain Adaptation with GANs

Claudia Drygala, Annika Mütze, Matthias Rottmann, Hanno Gottschalk
University of Wuppertal, IZMD, School of Mathematics and Natural Sciences, Stochastics Group

Abstract and Contribution

Aim:

- Automatically understanding complex visual scenes from RGB images
- → Semantic segmentation (pixel-wise classification of the image) with deep neural networks (DNNs)

Challenges:

- Neural networks need plenty of labeled images to generalize well on unseen scenes
- Manual label process is time and cost consuming

Solutions:

- Simulations of urban scenes were developed and improved
- Images generated by a simulation often come with labels for the semantic content for free [2]
- But: Domain gap to the real world – switching domains confuses the DNN
- Domain adaptation methods to mitigate the gap, e.g., Image-to-image translation

Our main contributions are:

- A semi-supervised domain adaptation method for semantic segmentation to guide the generator of a generative adversarial network (GAN) to downstream task awareness [3].
- Enrichment of synthetic data with photo-realistic appearance to increase the amount of training images for the supervised learning task and hence improve its performance.

GAN Concept

Computational cost

DNN	time per epoch
Deeplabv3-ResNet101 [4]	45 min
CycleGAN [5]	16 min

References

[1] Jülich Supercomputing Centre, "JUWELS: Modular Tier-0/1 Supercomputer at the Jülich Supercomputing Centre," *Journal of large-scale research facilities*, vol. 5, no. A135, 2019.

[2] A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, and V. Koltun, "CARLA: An open urban driving simulator," in *Proceedings of the 1st Annual Conference on Robot Learning* (S. Levine, V. Vanhoucke, and K. Goldberg, eds.), vol. 78 of *Proceedings of Machine Learning Research*, pp. 1–16, PMLR, 13–15 Nov 2017.

[3] A. Mütze, M. Rottmann, and H. Gottschalk, "Semi-supervised domain adaptation with cyclegan guided by a downstream task loss," *arXiv preprint arXiv:2208.08815 [cs]*, 2022.

[4] L.-C. Chen, G. Papandreou, F. Schroff, and H. Adam, "Rethinking Atrous Convolution for Semantic Image Segmentation," *arXiv:1706.05587 [cs]*, Dec. 2017.

[5] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *ICCV*, 2017.

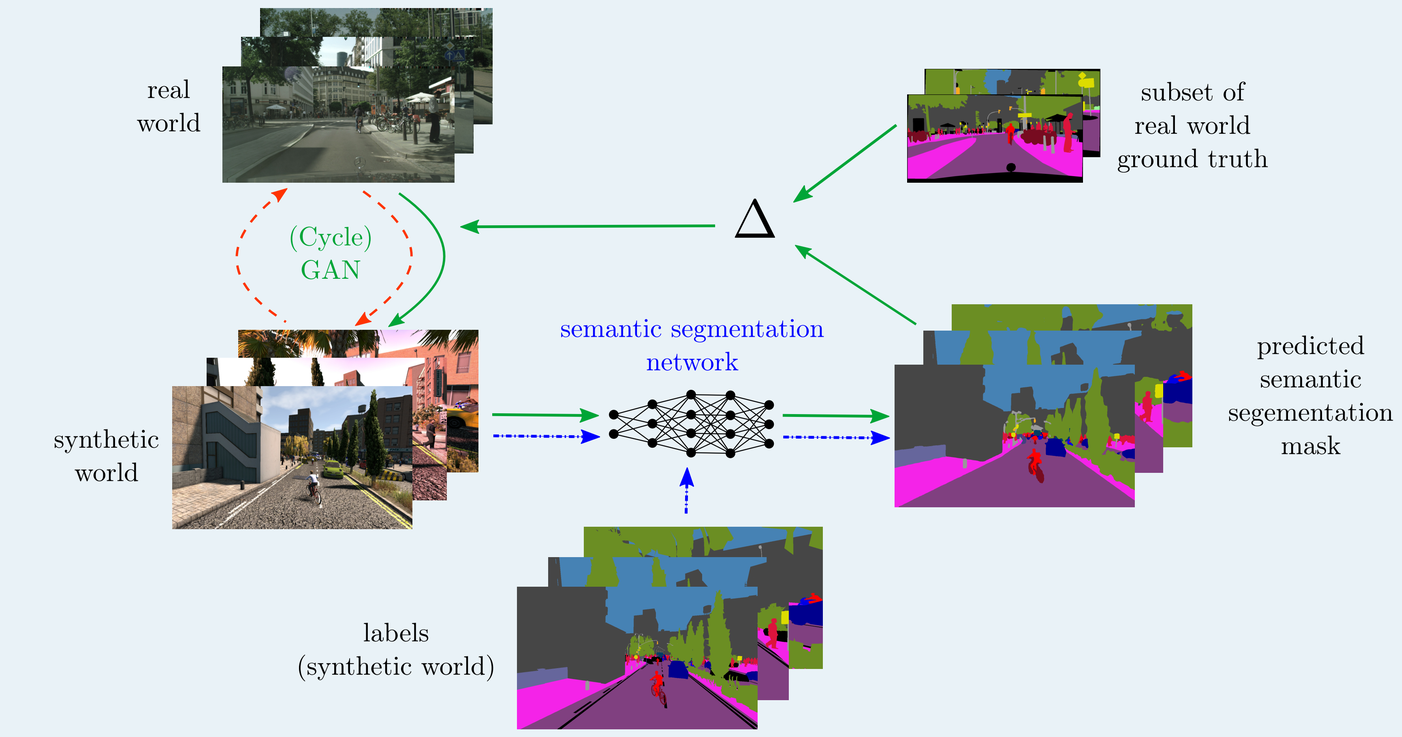
[6] T.-C. Wang, M.-Y. Liu, J.-Y. Zhu, A. Tao, J. Kautz, and B. Catanzaro, "High-resolution image synthesis and semantic manipulation with conditional gans," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018.

[7] C. Drygala, B. Winhart, F. di Mare, and H. Gottschalk, "Generative modeling of turbulence," *Physics of Fluids*, vol. 34, no. 3, p. 035114, 2022.

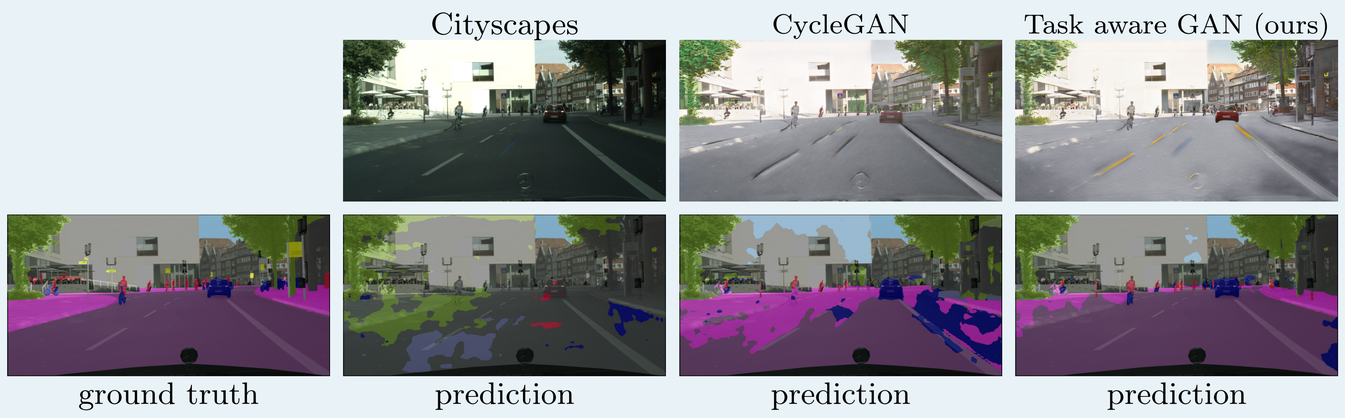
[8] A. Tao, K. Sapra, and B. Catanzaro, "Hierarchical multi-scale attention for semantic segmentation," *arXiv:2005.10821 [cs]*, 2020.

[9] M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. Benenson, U. Franke, S. Roth, and B. Schiele, "The cityscapes dataset for semantic urban scene understanding," in *CVPR*, 2016.

Concept of Task Aware Generator [3] (Real2Sim)

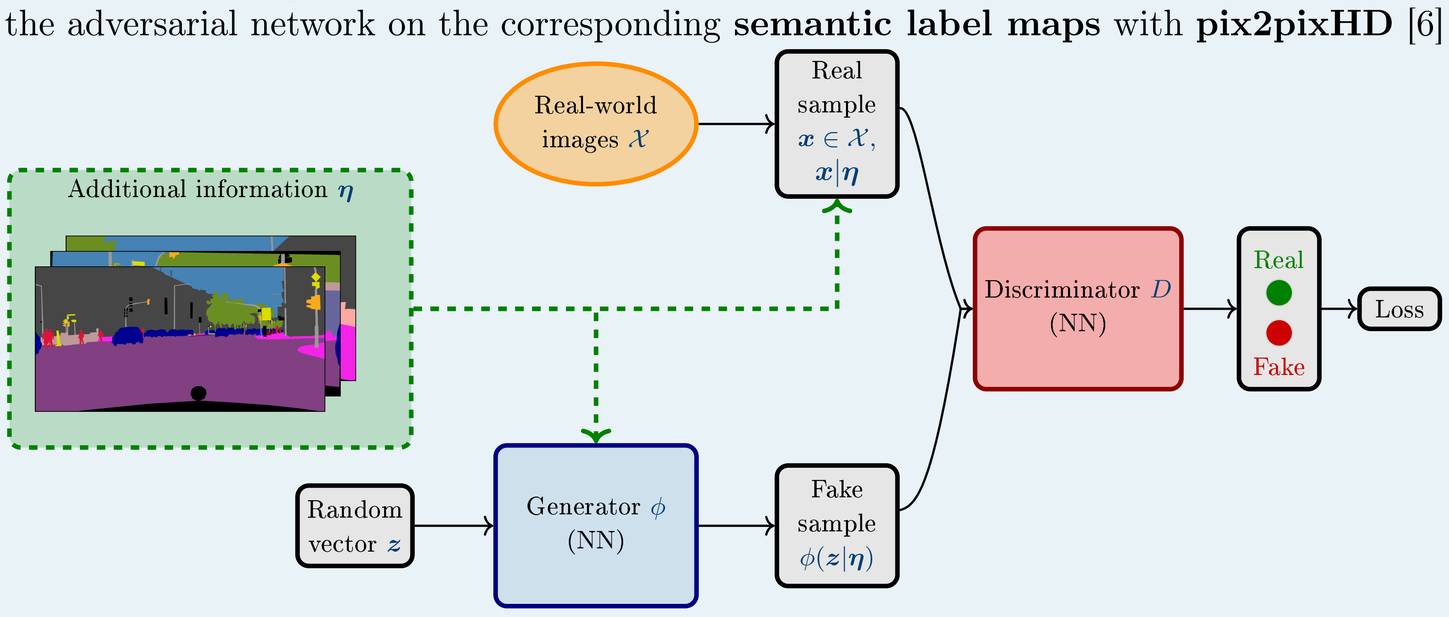


Real2Sim – Results on Semantic Segmentation



Enhancement of synthetic images (Sim2Real)

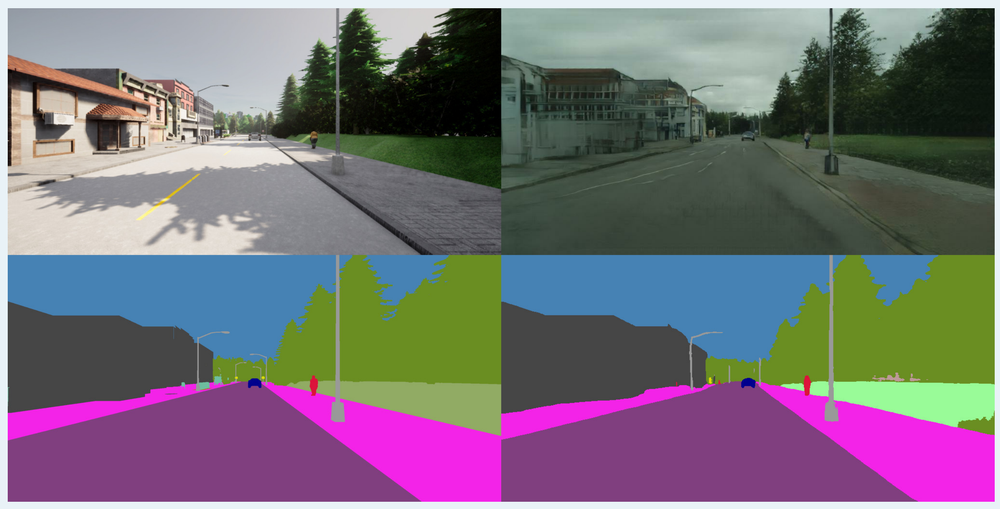
Generation of **high-resolution photo-realistic** images by **conditioning** the input of the adversarial network on the corresponding **semantic label maps** with **pix2pixHD** [6]



Architecture of conditional GAN. According to: [7].

Sim2Real – Results on Semantic Segmentation

Network: Deeplab V3+ with WideResNet38 backbone [8] trained on Cityscapes [9]



Acknowledgments. This work is funded by the German Federal Ministry for Economic Affairs and Climate Action.
The authors gratefully acknowledge the Gauss Centre for Supercomputing e.V. (www.gauss-centre.eu) for funding this project by providing computing time through the John von Neumann Institute for Computing (NIC) on the GCS Supercomputer JUWELS [1] at Jülich Supercomputing Centre (JSC).
Contact. {drygala,muetze,rottmann,hgottsch}@uni-wuppertal.de