Supplementary file for

Multi-Domain Conditional Image Translation: Translating Driving Datasets from Clear-Weather to Adverse Conditions

Vishal Vinod, K. Ram Prabhakar, R. Venkatesh Babu, Anirban Chakraborty
Department of Computational and Data Sciences,
Indian Institute of Science Bangalore, India

Abstract

This supplementary file contains additional results and more detail on the datasets and ablation experiments for our submission.

1. Datasets

As discussed in the main paper, we use a composite of images from the Cityscapes [1] and Indian Driving Dataset [9] for the source domain. We use a composite target dataset with three domains - Fog (Foggy-Cityscapes [2, 7]), Rain (Rainy-Cityscapes [4]) and Night-time (BDD100K Night Images [10]). For evaluation, we use the following: Fog (Foggy Zürich [3, 6], Foggy Driving [2], Foggy-Cityscapes [7]), Rain (ACDC-Rain [8], Rainy-Cityscapes [4]) and Night-time (Dark Zürich [3]).

Cityscapes includes 5,000 images with high quality annotations and 20,000 images with coarse annotations. It includes daytime scenes in mild weather conditions across several months (spring, summer, fall). The frames are manually selected with large number of dynamic objects and varying scene layout. It serves as a benchmark for Semantic Segmentation and Panoptic Segmentation.

Indian Driving Dataset (IDD) - High Resolution. A collection of 16,000 high-resolution road daytime scenes from Bangalore and Hyderabad cities in India. The images capture a wide variety of scenes, locations and varying inundations and annotations absent in most other driving datasets. Semantic segmentation annotations are available for a subset of the 16k high resolution images.

Foggy Zürich Test. A collection of 40 real-world foggy road scenes in the city of Zürich provided with semantic segmentation annotations for the 19 classes in Cityscapes.

Foggy-Cityscapes. The dataset includes 3475 Cityscapes images modified with 3 different fog intensity values (10,425 total images). Depth Completion with the L-R stereo pair to compute a Transmittance map and apply homogeneous synthetic fog based on the approximated atmospheric light (radiance). Foggy Cityscapes-refined has 550 high quality fog images.

ACDC-Rain. The Adverse Conditions Dataset with Correspondences (ACDC) dataset includes 400 train images, 100 validation images and 500 test images captured in rainy conditions in and around the canton of Zürich. Each adverse-condition image comes with a high-quality fine pixel-level semantic annotation, a corresponding image of the same scene taken under normal conditions, and a binary mask that distinguishes between intra-image regions of clear and uncertain semantic content. Since the images are captured with wet roads and light rain, the rain streaks are not as prominent as the Rainy-Cityscapes dataset used to train the proposed method.

Rainy-Cityscapes. The dataset consists of 295 Cityscapes images modified with 36 different rain streak variants. The synthetic rain and fog is applied using depth completion with the L-R stereo pair to compute the depth and distance of objects from the camera to include a more realistic degradation.

BDD100K. The dataset consists of semantic segmentation and object detection labels for 100,000 video clips in California. The dataset contains both daytime and night-time images in the 36728/5258 split with nearly 40% of images captured at night-time with low illumination and glare from oncoming traffic. We use 3145 night images from the BDD100K dataset for our night-domain.

Dark Zürich Validation. A collection of 50 real-world night-time road scenes in the city of Zürich with semantic segmentation annotations for the 19 classes in Cityscapes.

2. Additional Results

We show more qualitative results for the Indian Driving Dataset in Fig.(1) for single-domain translation to demonstrate the adverse translations. In Fig.(2), we present additional results for IDD in a multi-domain image translation setting. In Fig.(3), we show the respective style conditioning using the set of style images. Further, in Fig.(4), we show the qualitative results for the BDD100K dataset similar to the experiments in TSIT [5]. As the BDD100K images have minor weather degradations the same mild-weather reflect in the translations. For the experiment with BDD100K we use 2,500 images from each weather domain and only use BDD100K sunny images in the source input. In Fig.(5), we show additional results for the ablations to observe the degraded results without key components of the proposed method.

References

[1] Marius Cordts, Mohamed Omran, Sebastian Ramos, Timo Rehfeld, Markus Enzweiler, Rodrigo Benenson, Uwe Franke, Stefan Roth, and Bernt Schiele. The cityscapes dataset for semantic urban scene understanding. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3213–3223, 2016. 1

[2] Dengxin Dai, Christos Sakaridis, Simon Hecker, and Luc Van Gool. Curriculum model adaptation with synthetic and real data for semantic foggy scene understanding. *International Journal of Computer Vision*, 128(5):1182–1204, 2020. 1

[3] Dengxin Dai and Luc Van Gool. Dark model adaptation: Semantic image segmentation from daytime to nighttime. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, pages 3819–3824. IEEE, 2018. 1

[4] Xiaowei Hu, Chi-Wing Fu, Lei Zhu, and Pheng-Ann Heng. Depth-attentional features for single-image rain removal. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 8022–8031, 2019. 1

[5] Liming Jiang, Changxu Zhang, Mingyang Huang, Chunxiao Liu, Jianping Shi, and Chen Change Loy. Tsit: A simple and versatile framework for image-to-image translation. In *European Conference on Computer Vision*, pages 206–222. Springer, 2020. 2

[6] Christos Sakaridis, Dengxin Dai, and Luc Van Gool. Guided curriculum model adaptation and uncertainty-aware evaluation for semantic nighttime image segmentation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 7374–7383, 2019. 1

[7] Christos Sakaridis, Dengxin Dai, Simon Hecker, and Luc Van Gool. Model adaptation with synthetic and real data for semantic dense foggy scene understanding. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 687–704, 2018. 1

[8] Christos Sakaridis, Dengxin Dai, and Luc Van Gool. Acdc: The adverse conditions dataset with correspondences for semantic driving scene understanding. *arXiv preprint arXiv:2104.13395*, 2021. 1

[9] Girish Varma, Anbumani Subramanian, Anoop Namboodiri, Manmohan Chandraker, and CV Jawahar. Idd: A dataset for exploring problems of autonomous navigation in unconstrained environments. In *2019 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pages 1743–1751. IEEE, 2019. 1

[10] Fisher Yu, Haofeng Chen, Xin Wang, Wenqi Xian, Yingying Chen, Fangchen Liu, Vashisht Madhavan, and Trevor Darrell. Bdd100k: A diverse driving dataset for heterogeneous multitask learning. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 2636–2645, 2020. 1

Figure 1. Single-domain translation qualitative results for the Indian Driving Dataset.

Figure 2. multi-domain translation qualitative results for the Indian Driving Dataset.

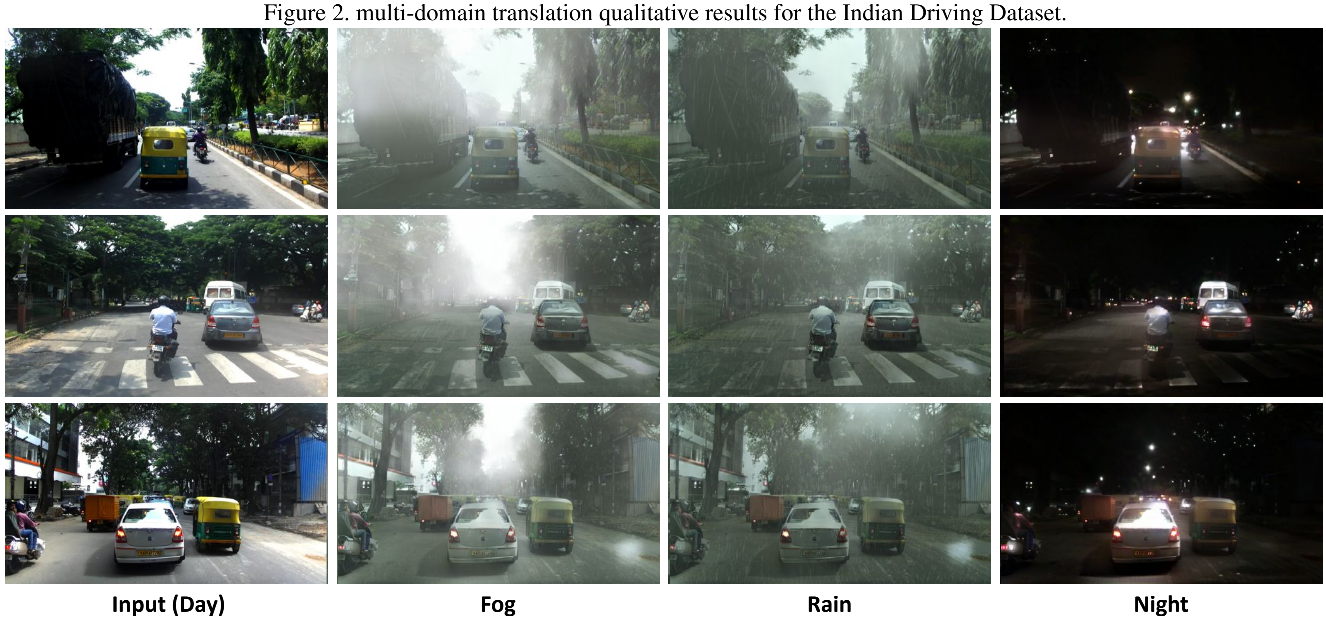

Figure 3. Qualitative results for the Indian Driving Dataset with the respective style conditioning images for the same content image.

Figure 4. Qualitative results for the BDD100k dataset with target domain images from BDD100k only.

Figure 5. Qualitative result for ablation experiments.