

TRoVE: Transforming Road Scene Datasets into Photorealistic Virtual Environments (Supplementary)

Shubham Dokania¹, Anbumani Subramanian¹, Manmohan Chandraker², and C.V. Jawahar¹

¹ IIIT Hyderabad, Telangana, India

² UC San Diego, CA, USA

1 Code and Data

We release the code and data for the proposed framework at https://github.com/shubham1810/trove_toolkit and https://cvit.iiit.ac.in/images/Projects/trove_eccv22_data/raw_data_eccv.tar.gz. The processed files from the raw dataset are also available directly for the semantic segmentation experiments at : https://cvit.iiit.ac.in/images/Projects/trove_eccv22_data/eccv_train_data.tar.gz.

We provide the code toolkit for users to generate synthetic dataset using nuScenes annotations. However, as long as the configuration file can be generated in the same manner, we can produce data for any scenario. In the future, we shall release the processing scripts for other datasets as well to demonstrate the extension.

2 Additional Data Samples

While in the main paper, we perform experiments highlighting semantic segmentation as the main task, we shall show additional data samples from the generated data in the supplementary document. We note that semantic segmentation is considered as a good representative task in the the domain of computer vision for a variety of scene perception tasks, but we can easily extend the models trained using our data to depth estimation, surface normal estimation, 3D scene reconstruction, 2D/3D object detection etc. This amalgamation of different tasks shows the future direction of the proposed work.

References



Fig. 1. Additional samples. (Left) An example with traffic light (please zoom-in) during daytime. (Right) The same scene at night without additional lighting.

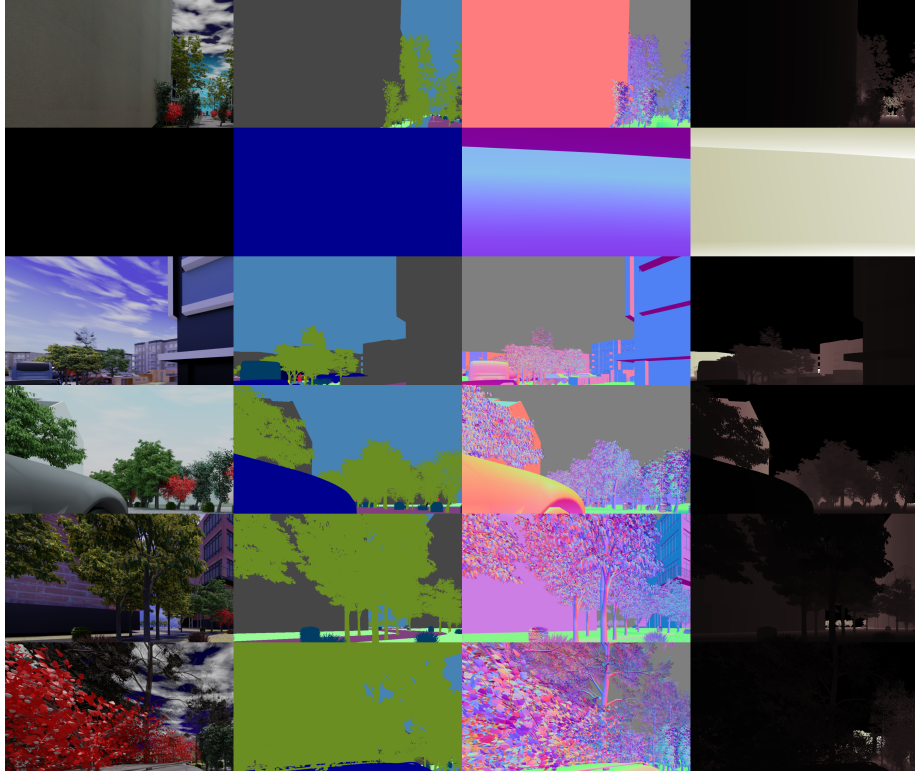


Fig. 2. We show additional data samples from the current datasets and the corresponding annotations. In this figure, we not only highlight the good samples but also the noisy cases where the generation process may fail sometimes due to wrong camera positioning. For example, the image in top row shows mostly a wall from a building and the data in second sample show a single class throughout the entire image. Cases like these may appear in the generative process and we recommend that these be filtered out using appropriate balancing methods such as pixel-density based filtering.

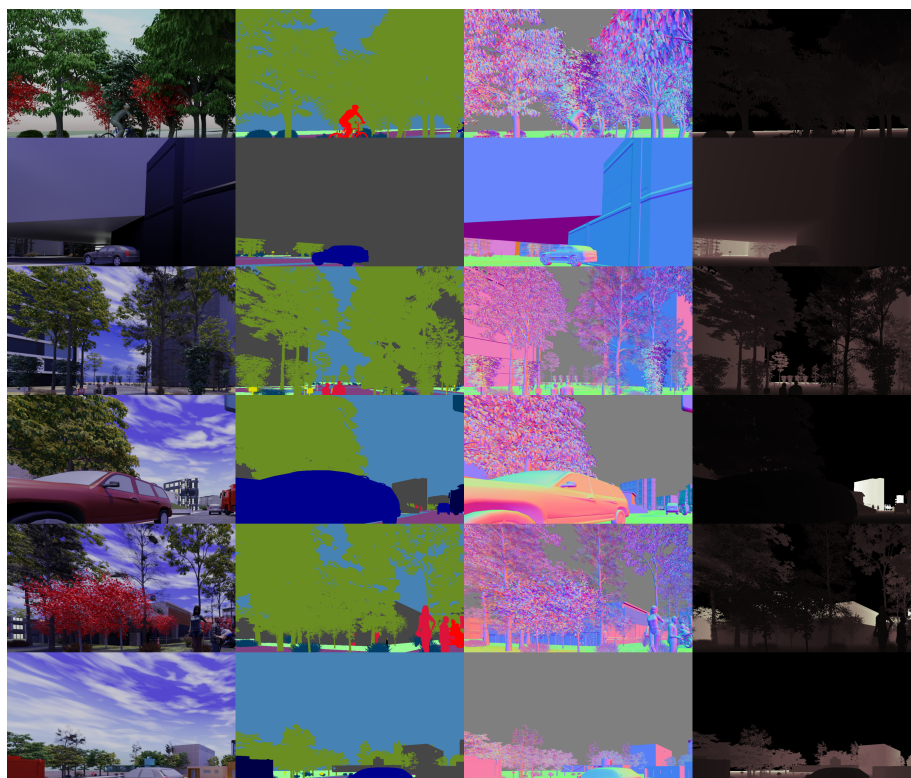


Fig. 3. Another set of samples where we show the dataset quality. The depth map is normalized for visualization purpose and raw depth data is available in the generated HDF files. For further steps, refer to the released code.