Author list

Jordan Chipka

Pri Mudalige

Title

Synthetic-to-Real Domain Adaptation for Autonomous Driving

Abstract

The ultimate ambition of autonomous driving is that it would be able to transport us safely to every corner of the world in any condition. However, a major roadblock to this ambition is the inability of machine learning models to perform well across a wide range of domains. For instance, a computer vision model trained on data from sunny California will have diminished performance when tested on data from snowy Michigan. The goal of this challenge is to develop innovative techniques to enhance domain adaptation capabilities. The accompanying dataset spans both synthetic and real-world domains to present a unique challenge to the participant.

Keywords

Autonomous driving, computer vision, semantic segmentation, domain adaptation, synthetic data

Motivation

State of the art autonomous driving technology relies on robust artificial intelligence (AI) models. These models require vast amounts of diverse, high-quality driving data to be collected and annotated for training. Although these AI models demonstrate good performance in local environments, they often fail when met with new data from a different distribution. This problem, also known as distribution shift, can occur due to changes in lighting, weather, and sensor hardware/configuration.

A more drastic form of distribution shift is encountered when the AI model is trained on synthetic data but tested with real-world data. Although this is a very appealing strategy for AI model development, as it eliminates the need to collect and annotate large amounts of costly real-world driving data, distribution shift remains a difficulty that must first be resolved. Consequently, the research community has put considerable emphasis on domain adaptation techniques to resolve difficulties such as these.

Dataset

The dataset is comprised of both real and synthetic images from a vehicle's forward-facing camera. Each camera image is accompanied by a corresponding pixel-level semantic segmentation image.

The training data is split into two sets. The first set consists of synthetic RGB images collected with a wide range of weather and lighting conditions using the CARLA simulator [1]. The second set includes a small pre-selected subset of data from the Cityscapes training dataset – which is comprised of RGB-segmentation image pairs from driving scenarios in various European cities [2].

The testing data is split into three sets. The first set contains synthetic images with weather/lighting conditions that were not present in the training set. The second set is a subset of the Cityscapes testing dataset. Finally, the third set is an unknown testing set which will not be revealed to the participants until after the submission deadline.

Challenges

1. Data augmentation
   a. Determine the most effective data augmentation strategies to enhance the model's performance on the testing sets.
2. Model development
   a. Determine novel model architectures or training methodology that can be used to enhance performance on the testing sets.

Performance will be measured by the mean intersection-over-union (mIOU) across all classes for the segmentation results on all three testing sets.

Notes to participants

Participants are permitted to perform data augmentation by transforming the given training data (e.g. crop, mirror, brighten, etc.). However, they are not permitted to use additional data outside of the provided training data – either real or synthetic – to augment training (e.g. use additional Cityscapes training data outside of the pre-selected subset).

References

[1] Alexey Dosovitskiy, German Ros, Felipe Codevilla, Antonio Lopez, & Vladlen Koltun (2017). CARLA: An Open Urban Driving Simulator. In *Proceedings of the 1st Annual Conference on Robot Learning* (pp. 1–16).

[2] Cordts, M., Omran, M., Ramos, S., Rehfeld, T., Enzweiler, M., Benenson, R., Franke, U., Roth, S., & Schiele, B. (2016). The Cityscapes Dataset for Semantic Urban Scene Understanding. In *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.