Deep Learning on Images and Signals

Nils Fahrni

BSc Data Science Student

November 25, 2024

Contents

Chapter 1

Task

1.1 Scientific Question

How does the performance of a U-Net semantic segmentation model differ between scenes of city streets and non-city streets in the BDD100K dataset?

1.2 Dataset

I decided for the BDD100K dataset. The BDD100K dataset is the largest driving video dataset, featuring 100,000 videos and supporting 10 tasks for evaluating and advancing multitask learning in autonomous driving. It offers diverse geographic, environmental, and weather conditions, making it a benchmark for studying heterogeneous multitask learning and training robust computer vision models [1]. For this Mini Challenge I use the "10k" subset, which is made up of 10,000 images and are sampled from the 100,000 videos' frames. This subset is intended for semantic segmentation tasks. These 10,000 images have already been pre-partitioned into a train, validation and test partition. The train partition consists of 8000 images, the validation 1000 and the test 1000. This smaller subset does unfortunately not have scene attributes but the larger video dataset does. Since the semantic segmenatiton subset is derived from the video dataset, I retrieve the scene attributes through the larger datasets metadata JSON. The issue here is that not all images in the semantic segmentation subset seem to be in the video dataset. I will therefore only use the small overlap of images that has scene attributes and exists both in the video and semantic segmentation dataset.

1.3 Methodology and Procedure

To answer the scientific question, I will train a U-Net model on the BDD100K dataset. I will then evaluate the model on the test set and compare the performance between city street and non-city street scenes. I will use the mean Intersection over Union (mIoU) as the evaluation metric. The mIoU is a common metric for semantic segmentation tasks. It is calculated as the mean of the Intersection over Union (IoU) for each class.

Chapter 2
Discussion

Bibliography

[1] F. Yu, H. Chen, X. Wang, *et al.*, *BDD100K: A Diverse Driving Dataset for Heterogeneous Multitask Learning*, Apr. 2020. arXiv: 1805.04687. (visited on 11/25/2024).