Vision-Based Ego Vehicle Speed Estimation

3D Vision Project Proposal
Supervised by: Sandro Lombardi
March 13, 2022

GROUP MEMBERS

Jens Eirik Saethre Christopher Raffl Livio Schläpfer

I. DESCRIPTION OF THE PROJECT

Over-speeding is a major cause of a large number of road traffic accidents. Vision-based ego vehicle speed estimation therefore is a useful tool for insurance companies as the vehicle speed can be estimated e.g. from a dashcam [4], [6]. The goal of this project is to develop such a vision-based speedometer for ego vehicle speed estimation. Publicly available datasets [5], [2], [1] which include groundtruth speed measurements will be used for learning the speed measurements and validating our predictions.

II. WORK PACKAGES AND TIMELINE

We split our project into four work packages. Within the first, we want to explore different datasets and setup the development environment on ETH Zurich's *Euler* cluster. In the second step, we will implement a baseline algorithm that is able to perform vision-based ego vehicle speed estimation. Afterwards, we resolve potential problems of this first working version. In our final work package we aim to improve the performance of our model by resolving potential shortcomings and exploring additional data and techniques. In all our work packages, we make use of ETH Zurich's *Euler* cluster and implement our algorithm in Python3 and PyTorch. Every group member will contribute equally to all work packages.

A. Work Package 1: Explore datasets and setup of environment

- **Timeline:** Now - End of March.
- **Tasks:** This work package consists mainly of the following tasks:
  - In-depth exploration of several datasets that provide dashcam video sequences with corresponding ego-motion data, such as vehicle speed and vehicle heading. These include (but are not limited to) Cityscapes [2], KITTI [5], and DBNet [1].
  - Setting up a development environment on Euler to start developing a computer vision pipeline based on PyTorch that is able to deal with the large amounts of video data (e.g. in excess of 350 GB for Cityscapes).
- **Challenges:** Data access, i.e. we are still waiting for access credentials to the DBNet dataset, which requires manual approval by the authors; getting access to Euler.
- **Outcome:** This work package should have two main outcomes: First, we select one or multiple datasets to use in our project. This decision will be made based on availability, ego-motion data, and the sustainability for our PyTorch pipeline. Second, our development is ready to start with work package 2.

B. Work Package 2a: Implement baseline algorithm

- **Timeline:** End of March - Midterm presentations (April 25)
- **Tasks:**
  - PyTorch data loading pipeline for the selected dataset / datasets.
  - PyTorch implementation of a baseline model using a deep neural network based on optical flow and single-view depth prediction methods as presented in [6].
- **Challenges:** Setting calibration parameters of vision pipeline (e.g. best performing model for KITTI dataset mostly based on improved calibration parameters [3]; Application of paper on a different dataset.
- **Outcome:** Baseline implementation of a model with reasonable performance.

C. Work Package 2b: Mitigate challenges

- **Timeline:** Midterm presentations (April 25) - Beginning of May
- **Tasks:**
  - Resolution of problems related to baseline implementation.
- **Challenges:** *to be determined.*
- **Outcome:** Baseline implementation of a model with performance comparable to [6].

D. Work Package 3: Enhancements to Baseline

- **Timeline:** Beginning of May - Final project presentation (May 30)
- **Tasks:**
  - Performance enhancements by resolving potential shortcomings of the baseline.
  - Performance enhancements by including additional data sources. Options are to include data sources such as Stereo Camera data, LiDAR data, which may support our Vision model in producing better ego-vehicle speed estimation. Note: The used additional datasources at this stage of the project will be dependent on the chosen dataset in work package 1.
- **Challenges:** *to be determined.*
- **Outcome:** Implementation of a model that outperforms the outcome of work package 2b.

III. Outcomes and Demonstration

The expected outcome of our project is a working PyTorch implementation of a vision-based speedometer for ego vehicle speed estimation. We will compare our implementation to the results provided in [6]. As demonstration, we will present video sequences of the above mentioned datasets together with the outputs of our vision pipeline.

References

[1] Yiping Chen, Jingkang Wang, Jonathan Li, Cewu Lu, Zhipeng Luo, HanXue, and Cheng Wang. Lidar-video driving dataset: Learning driving policies effectively. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018.

[2] 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. *CoRR*, abs/1604.01685, 2016.

[3] Igor Cvišić, Ivan Marković, and Ivan Petrović. Recalibrating the kitti dataset camera setup for improved odometry accuracy. In *2021 European Conference on Mobile Robots (ECMR)*, pages 1–6, 2021.

[4] David Fernández Llorca, Antonio Hernández Martínez, and Iván García Daza. Vision-based vehicle speed estimation: A survey. *arXiv preprint arXiv:2101.06159*, 2021.

[5] Moritz Menze and Andreas Geiger. Object scene flow for autonomous vehicles. In *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.

[6] Róbert Adrian Rill. Speed estimation evaluation on the KITTI benchmark based on motion and monocular depth information. *CoRR*, abs/1907.06989, 2019.