Lehrstuhl für Fahrzeugtechnik
Fakultät für Maschinenwesen
Technische Universität München

BA/SA/IDP

VLM-based Scene Understanding for Automated Parking

Reliable and safe automated parking is a key feature for modern vehicles and a critical step towards full autonomy. Current systems often depend on short-range sensors like ultrasonic and simple camera-based logic, which primarily detect immediate obstacles and basic lane markings. However, this approach often falls short in complex, unstructured, and dynamic parking environments. It struggles to interpret ambiguous markings, differentiate between various types of obstacles (e.g., a curb vs. a shopping cart), or understand the occupancy status of potential parking slots in challenging lighting and weather conditions. This lack of a comprehensive scene understanding limits the system's robustness and ability to operate in diverse real-world scenarios.

This thesis aims to address these limitations by developing and evaluating a set of deep learning models for holistic scene understanding in parking scenarios. The core idea is to create a rich, semantic representation of the environment by simultaneously performing multiple perception tasks from camera data. This includes semantic segmentation to identify drivable areas, parking lines, and physical boundaries; object detection to locate other vehicles, pedestrians, and static obstacles; and parking slot detection to identify and classify available spaces. By integrating this information, the system can build a comprehensive model of the scene, enabling more intelligent, safe, and efficient path planning for automated parking maneuvers.

Work packages:

- **Literature review:** Datasets for Scene Understanding, QA and Autonomous Driving. Latest VLM architectures and available Models

- **Evaluation design:** KPI definition for both datasets and models. What makes a dataset good? How can i evaluate a model for the parking task?

- **Design and implementation:** Implementation of several VLM's for scene understanding, potentially including a baseline and a more advanced multi-task learning approach.

- **In-depth evaluation and comparison:** Rigorous testing and comparison of the implemented models on key performance metrics.

Requirements:

- Very good programming skills in Python and Pytorch

- High personal motivation and independent working style.

- Knowledge of Computer Vision and Deep Learning (Must), VLMs or LLMs (Desired)

- Very good language proficiency in German, English or Spanish.

TUT

The thesis should document the individual work steps in a clear form. The candidate commits him/herself to carry out the study independently and to indicate the scientific aids used by him/her.
The submitted paper remains the property of the chair as an examination document.

Prof. Dr.-Ing. M. Lienkamp   Betreuer: Esteban Rivera, M. Sc.

Ausgabe:_____   Abgabe: _____