A system for detecting unusual situations within vehicles using images and depth maps based on deep learning

Summary

The doctoral dissertation undertook an analysis of creating a system capable of detecting unusual situations around a vehicle using images and depth maps. To implement this system, a hybrid approach combining image processing, depth maps, synchronization, and artificial intelligence was employed. Several research experiments were conducted in this regard.

The studies focused on key aspects of the system's operation:

- synchronization of data from multiple sensors,
- object detection using images,
- object detection in depth maps,
- effectiveness of the warning system.

Initially, efforts concentrated on developing a system for analyzing and collecting synchronized data. To effectively utilize data from different sensors, synchronization was crucial to ensure they reflect the same situation. The implemented system allows data synchronization at the data level with defined limits of deviation between expected and actual data. In case of discrepancies exceeding these limits, missing data is generated accordingly, such as duplicating the last sample or artificially generating data. Experiments demonstrated the validity of synchronization for both same-frequency and different-frequency sensor data.

The system was used to create custom datasets from drives with a research vehicle, incorporating data from 16-beam and 64-beam LiDARs and cameras. These datasets were
used to train selected artificial neural network models and to test the developed warning system.

Another series of studies involved analyzing available solutions for object detection in images. The operation, effectiveness, and speed of detection of multiple artificial neural network models were analyzed. Based on the results, the YOLOv7 architecture was chosen, with custom object labels defined and training and testing processes conducted. This resulted in a model achieving an average

accuracy of 95% across all defined object classes. This network was employed in the developed warning system for camera data.

Further research focused on solutions for object detection in depth maps. Point clouds generated from available devices may contain up to 2.5 million points. Analysis of available network architectures and their application tests identified two potential models: PillarPoints and PointRCNN. Both networks underwent training on custom datasets, with the analysis of performance leading to the selection of PointRCNN as the final model.

Subsequent studies concentrated on data fusion from LiDAR and camera sources. Throughput of models was a critical consideration, leading to the selection of the Frustum PointNet network. This network operates in two stages: initially detecting objects in images and subsequently refining detections by analyzing significant segments of depth maps.

The final studies focused on the implemented warning system. Using available data (images, depth maps, and data fusion), the system analyzes the vehicle’s surroundings. Three scenarios triggering driver alerts were defined: approaching animals, and analyzing pedestrian crossing areas with suspicion of a pedestrian or cyclist entering the lane. Simulations conducted using collected data demonstrated high accuracy in detecting these predefined hazardous situations.

Based on the conducted research experiments, it was concluded that data fusion effectively enables monitoring of the vehicle’s surroundings and detection of defined unusual situations, making it an excellent solution for autonomous mobility applications.