Summary of the doctoral dissertation

Automatic Analysis of Large Video Datasets for Detection of Chosen Objects and Weather Conditions Estimation

The aim of this doctoral thesis was to develop a system to improve the annotation and segregation of video recordings from car test drives. Image processing and deep learning methods were used to achieve this goal.

Annotating video recordings is a time-consuming activity that requires the involvement of a large group of people who mark all objects on subsequent frames of the film. From the point of view of the company ordering such a task, it is a considerable cost of employee salaries, renting space for an annotation laboratory and purchasing the necessary equipment. Improving the annotation process directly translates into cost reductions. Supervised learning requires a man-made database so that the target neural network can reproduce the human point of view. Therefore, the participation of people in the annotation process cannot be completely eliminated. However, the process can be improved, which was the main goal of starting this doctoral dissertation.

The annotation of recordings was to be facilitated thanks to a system of prompts generated by an auxiliary neural network detecting objects in the images. The research focused mainly on the R-CNN networks and their variants, which are currently the most efficient and accurate for this type of task. These networks, however, require very vast sets of training data, so a large part of the research was based on the so-called data augmentation methods - transformations of already existing data so as to increase the size of the training set without user intervention. Each method of data augmentation gives a potential increase of the dataset by 100 percent, which may result in an improvement in the quality of the detector. In practice, however, a drop in quality is possible if the augmentation method used is inadequate. The research has shown that the currently used augmentation methods do not exhaust all potential algorithms by proposing new augmentation methods.

Another related line of research has focused on transformations of training images, which may introduce a decline in the quality of neural network recognition. Several methods of image degradation and their impact on detector performance were tested in these studies. It was also taken into account to what extent the quality of the detections deteriorates, depending on the nature of the images themselves, i.e., the conditions under which they were recorded. Thanks to this, it was possible to demonstrate that the degradation of images from the optical sensor, which may be found in driver assistance systems, has an impact on the quality of detection and, depending on the type of distortion, deteriorates the recognition of objects to a varying degree.