Szczecin, September 17, 2024

Abstract of the Doctoral Dissertation

Processing and Integration of Multimodal Image Data to Support the Detection of Behaviors Related to Decreased Concentration Levels in Vehicle Users

Author: M.Sc. Eng. Anton Smoliński
Supervisor: D.Sc. Ph.D. Eng. Paweł Forczmański
Assistant Supervisor: Ph.D. Eng. Adam Nowosielski

The aim of this doctoral dissertation was to develop methods for the extraction and integration of features from multimodal facial images to detect indications of decreased concentration levels in vehicle users.

The thesis posits that appropriate extraction, processing, and integration of features from multimodal facial images allow for more effective detection of behaviors characteristic of decreased concentration levels in vehicle users compared to methods based on unimodal images.

This work presents original methods for the extraction, processing, and integration of multimodal image data to support the detection of behaviors related to decreased concentration levels in vehicle users. Experimental studies were conducted to evaluate different variants of data extraction and fusion at various stages of processing. Data from different image modalities were utilized, such as RGB images, thermal images, and depth maps, which extends traditional solutions based solely on single modalities. Both traditional methods of feature extraction and pattern recognition were analyzed, as well as techniques based on deep machine learning, including artificial neural networks like CNN and LSTM, enabling the analysis of dynamic changes in driver behavior.

Special attention was paid to the detection of signs of fatigue and distraction, such as blinking frequency, yawning, head dropping, and eye rubbing. The PEROPEN indicator was introduced to measure the level of fatigue based on the analysis of the degree of mouth opening. Additionally, the impact of face, eye, and mouth detection methods on the effectiveness of concentration level assessment was examined, contributing to the optimization of the detection process.

The research results demonstrated that the integration of multimodal image data significantly improves the effectiveness of detecting visual cues indicating driver fatigue and distraction compared to the analysis of single modalities, confirming the posed thesis.

As part of the conducted research, a test dataset was created and made available on the website *cvlab.zut.edu.pl*, containing both raw and processed data for research purposes. This allows other researchers to utilize these resources to test their own methods and algorithms.

The conclusions from the studies indicate that the developed methods can contribute to the creation of advanced and precise driver monitoring systems. They can be applied in both modern and older vehicles, thereby increasing road safety.

Keywords: computer vision, image processing, multimodal imaging, driver fatigue detection, machine learning, data fusion.

17.09.24 Anton Smoliński