TU WIEN Informatics

Ao.Univ.Prof. Dr. Robert Sablatnig
Head of Institute
TU Wien, Institute of Visual Computing & Human-Centered Technology
Computer Vision Lab

Favoritenstraße 9-11/193-1
1040 Vienna, Austria
+43-1-58801-18361
sab@cvl.tuwien.ac.at
https://cvl.tuwien.ac.at/

To whom it may concern

Thesis Evaluation

Jakub ŠPAŇHEL

Steps towards Improvements of
Computer Vision Methods for Traffic Analysis

The thesis submitted constitutes a collection of articles, known as a compilation thesis or thesis by publication, centered on the common theme of traffic analysis using computer vision. The candidate has chosen five papers, three of which he is the principal author (one being highly cited) and two highly cited papers where he is the second and third author, respectively. Each article serves as a chapter of the thesis (Chapters 5-9), supplemented with an introduction, conclusion, and a unifying framework including a comprehensive literature review to connect the articles. This format enables candidates to present their research output and contributions to the field while fulfilling the requirements for a doctoral dissertation. The topics covered can be categorized into "Datasets for Traffic Analysis," "Re-Identification of Vehicles from Image and Video," and "Improvements in License Plate Recognition," all integral fields within the realm of intelligent transport systems empowered by computer vision. This integration enhances transportation systems' efficiency, safety, and functionality. The thesis touches upon aspects such as Traffic Monitoring and Management, Vehicle Detection and Tracking, Driver Assistance Systems, and Vehicle Classification. However, it does not cover Pedestrian Detection and Safety, Automated Parking Systems, or Smart Infrastructure Maintenance. Integrating computer vision into intelligent transport systems facilitates real-time monitoring, analysis, and decision-making, resulting in safer, more efficient, and sustainable transportation networks. It enables early hazard detection, optimized traffic flow, and enhanced accessibility by providing real-time information on transit options, parking availability, and multimodal connections.

Drawing from an extensive review of literature, Mr. ŠPAŇHEL endeavors to organize various topics in Chapter 2, providing a comprehensive overview of the state of the art in fields pertinent to the five selected papers. This includes detailed discussions on Camera Calibration for Vehicle Speed Measurement, Object and Vehicle Detection and Recognition, and Re-Identification Methods in both general and Traffic Surveillance contexts. Additionally, since two papers focus on License Plate Recognition, a thorough examination of this topic is included. Consequently, the thesis offers a unifying framework with an extensive literature review. However, the structure becomes somewhat unclear in Part II, dedicated to Datasets for Traffic Analysis, as Chapters 5 and 6 are already covered in selected papers, while Chapters 3 and 4 are presented later in Chapters 7 and 8, leading to a fragmented structure and content replication across multiple chapters. Part V of the thesis attempts to illustrate the practical application of the research presented in the selected papers within the realm of NVIDIA AI City Challenges, organized as the Computer Vision and Pattern Recognition Workshop since 2017. Nonetheless, the structure lacks clarity, making it challenging to correlate the results with the papers in the collection. Chapter 11 seems disconnected from the core topic as it delves into driver assistance, straying from the thesis's central focus without clear integration. Finally, Chapter 12 concludes the thesis, but a more detailed discussion of the achieved findings and future avenues for research could have been provided.

Overall, the thesis demonstrates a commendable level of technical depth, showcasing original research within the domain of computer vision traffic analysis. The selected papers are well-cited and present significant contributions to the field. However, the thesis structure could have been improved to more effectively highlight Mr. ŠPAŇHEL's overall contributions. His noteworthy contributions encompass Vehicle Fine-Grained Recognition, Vehicle Re-Identification, License Plate Recognition, and Monocular Vehicle Speed Measurement. The thesis has the following contributions:

1. **Datasets**: Throughout his research, several comprehensive datasets were introduced: BoxCars116k for Vehicle Fine-Grained Recognition, CarsReId74k for Vehicle Re-Identification, and ReId, HDR, and CamCar6k datasets for License Plate Alignment/Recognition. Each dataset serves as a benchmark for advancing research in its respective area, providing valuable resources for ITS and Computer Vision. Large, publicly available datasets can inspire new research directions and fuel innovation in computer vision. By providing researchers with access to a rich source of data, datasets encourage the development of novel algorithms and techniques that push the boundaries of what is possible in computer vision. These datasets provide the raw material for training machine learning models, which are essential components of computer vision applications. By providing examples of images with corresponding labels or annotations, datasets enable algorithms to learn patterns and features that are relevant to the task at hand and are, therefore, valuable contributions.
2. **Vehicle Re-Identification:** An approach known as LFTD was introduced to aggregate visual features in the temporal domain, enhancing the consolidation of visual features extracted from multiple observations, such as various video frames of vehicle

trajectories. This method significantly bolsters the robustness of re-identification tasks. Additionally, significant contributions were made to Vehicle Fine-Grained Recognition, including the proposal of multiple augmentation techniques for neural network training, refinement of 3D bounding box estimation from a single image, and the meticulous curation and expansion of the vehicle fine-grained recognition dataset in an unconstrained environment.

3. **License Plate Recognition:** A method for recognizing low-quality license plates holistically using a machine learning model was presented and extended by suggesting license plate alignment as a pre-processing step for LPR using neural networks.

The candidate's contribution to monocular vehicle speed measurement was primarily organizational, involving the coordination of data acquisition and participation in the evaluation phase. However, significant contributions were made in other areas, including novel methods and augmentation techniques for vehicle fine-grained recognition, an innovative approach for aggregating visual features for vehicle re-identification, and an original method for license plate recognition using alignment and holistic recognition. Additionally, a valuable contribution was made through the meticulous curation of datasets in the domains of license plate recognition, vehicle re-identification, vehicle fine-grained recognition, and monocular vehicle speed measurement.

The consistency of English writing quality varies throughout the thesis. While some sections, particularly the papers, demonstrate good quality, others exhibit issues, ranging from stylistic inconsistencies, such as overly general statements like, "*Transportation has become an inextricable part of our modern existence, with thousands of vehicles hitting the roads daily*", found in the Introduction, to duplicated sentences, as seen in Chapter 2 at the outset, and grammatical errors like, "*Besides evaluation datasets covered in Table 2.2 on the next page, multiple new datasets were for automatic camera calibration were published subsequently*", in Chapter 2. These examples highlight the need for closer attention to language consistency and editing across the thesis.

Mr. ŠPAŇHEL demonstrates familiarity with the broader field of Traffic Surveillance and presents innovative ideas for enhancing existing techniques. His three main contributions underscore his ability to generate novel scientific findings within the realm of Computer Vision. Moreover, his profile on SCOPUS reflects his status as a well-established researcher, boasting 20 documents, over 300 citations, and an impressive h-index of 7 for a Ph.D. student. However, it is notable that completing his Ph.D. thesis took a considerable amount of time, considering his first publication in 2017.

In summary, the thesis offers a commendable contribution to new research in its field. The author demonstrates a capacity for scientific inquiry and has already established a solid profile within the scientific community. Consequently, this dissertation is deemed acceptable, as the author has demonstrated proficiency in conducting research and producing scientific results.

TU WIEN Informatics

Ao.Univ.Prof. Dr. Robert Sablatnig

Head of Institute

TU Wien, Institute of Visual Computing &
Human-Centered Technology
Computer Vision Lab

13 February 2024