Doctoral theses at NTNU, 2022:67

Elnaz Namazi

Using Modern Vehicles as Mobile Sensors for Intelligent Traffic Awareness

Doctoral thesis

NTNU
Norwegian University of Science and Technology
Thesis for the Degree of Philosophiae Doctor
Faculty of Information Technology and Electrical Engineering
Department of Computer Science

NTNU
Norwegian University of Science and Technology

Elnaz Namazi

Using Modern Vehicles as Mobile Sensors for Intelligent Traffic Awareness

Thesis for the Degree of Philosophiae Doctor

Trondheim, March 2022

Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Computer Science

NTNU
Norwegian University of
Science and Technology

NTNU
Norwegian University of Science and Technology

Thesis for the Degree of Philosophiae Doctor

Faculty of Information Technology and Electrical Engineering
Department of Computer Science

ISBN 978-82-326-5727-8 (printed ver.)
ISBN 978-82-326-5631-8 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (online ver.)

Doctoral theses at NTNU, 2022:67

Printed by NTNU Grafisk senter

Abstract

Traffic management has become a critical problem with growing traffic congestion worldwide. As a result, the approaches to managing traffic tend to become smart, and Intelligent Traffic Management Systems (ITMSs) are becoming increasingly common. Accessing the traffic data is a key component in ITMSs. As vehicles and their sensing and connection capabilities are advancing, more studies are investigating the feasibility of using Modern Vehicles (MVs) in estimating and sharing traffic data. However, despite the identified potential and increasing attention given to ITMSs and MVs, only a few studies use an MV as a mobile sensor to collect traffic data, with the purpose of improving the ITMS's performance in a metropolitan area.

To address this gap, this thesis first identifies, by reviewing the literature, the existing methodologies and the corresponding traffic data types required by ITMSs and explores the potential research gaps in the field. Second, it addresses how a single MV and mounted low-cost sensors (i.e., a monocular camera with a built-in Global Positioning System [GPS] receiver) can be employed to estimate the traffic data of both MVs (i.e., geolocation) and the surrounding observed vehicles (i.e., number, type, relative position, distance, speed, lane, and geolocation) in a metropolitan area with a combination of Human-Driven Vehicle (HDV) and MV traffic. Third, it explores how the estimation error of a sensor mounted on a single MV can be mitigated to provide a more accurate picture of the traffic scene than what can be obtained by using data from only one MV, by fusing estimated traffic data (i.e., HDV's geolocations) of two observing MVs.

In the initial stage of this study, a Systematic Literature Review (SLR) is conducted. Then, the Design Science Research (DSR) methodology is applied. Case studies are performed to evaluate and validate the proposed approaches and algorithms in terms of acceptance, usability, and impact on the problem at stake.

This thesis aims to bring researchers to the forefront of this new interdisciplinary field. The thesis contributes new knowledge to both the Computer Science (CS) and Civil Engineering (CE) fields and guides the design and prototyping of traffic data estimation by MVs to generate a dynamic model of the traffic scene that can enhance the performance of ITMSs. The lessons learned by the author of this thesis provided knowledge about the feasibility of using MVs to enhance traffic awareness by generating a dynamic model of the traffic scene.

Preface

This doctoral thesis was submitted to the Norwegian University of Science and Technology (NTNU) in partial fulfillment of the requirements for the degree of Philosophiae Doctor.

The Ph.D. work was performed at the Department of Computer Science, NTNU, Trondheim. Associate Professor Jingyue Li (NTNU) was the main supervisor, and Professor Rudolf Mester (NTNU), Associate Professor Hallvard Trætteberg (NTNU), and Doctor Ørjan Tveit (Statens vegvesen: The Norwegian Public Roads Administration) were the co-supervisors.

The thesis has been supported by the Department of Computer Science, NTNU, Trondheim.

Acknowledgments

I would like to use this opportunity to express my sincere appreciation to all the people who supported, guided, and encouraged me during this Ph.D. journey.

First and foremost, I would like to express my gratitude to my main supervisor, Associate Professor Jingyue Li, for providing me with the opportunity to pursue a Ph.D. degree. It has been an honor to be his first Ph.D. student. So many thanks to him for his scientific advice, availability, and consistent support. He was an invaluable role model for me, and I was inspired by his hard work. His insightful feedback encouraged me to sharpen my thinking and gave me the confidence to learn various scientific skills to bring my research to the next level.

I would like to express my appreciation to my co-supervisors, Professor Rudolf Mester, Associate Professor Hallvard Trætteberg, and Doctor Ørjan Tveit, for their enthusiasm for the research area, immense knowledge, and invaluable guidance. They gave me generous support, constructive comments, and warm encouragement. Without their persistent help, this thesis would not have materialized.

My sincere thanks also go to Associate Professor Chaoru Lu, who was always willing to assist in any way from the early stage of this Ph.D. study. I have greatly benefited from his meticulous feedback and guidance.

I would like to extend my gratitude to Professor Nirmalie Wiratunga for being an amazing mentor and a great inspiration during the last year of my Ph.D. study in the “IDUN - from Ph.D. to professor” project.

I gratefully acknowledge the Department of Computer Science at NTNU for the funding that made my Ph.D. study possible.

From the bottom of my heart, I would like to say a big thank you to my colleagues and friends who helped me in this Ph.D. journey. Thanks to them for always believing in me, encouraging me to perform my best, and cheering me to achieve my goals.

Last but not least, I wish to express my heartfelt thanks to my lovely family. They are helping me to dream big and work hard to realize my ambitions. I could not have achieved this much without their endless support and encouragement. They are the source of the motivation and happiness in my life. A huge thank to them for their unconditional support and love all these years.

Contents

Glossary

ADAS Advanced Driver Assistance System. 19

AI Artificial Intelligence. 16, 41, 76, 87

AV Autonomous Vehicle. 4–7, 11, 16, 17, 19, 20, 23, 24, 31, 34, 39, 75, 87

CCF Camera Coordinate Frame. 54

CE Civil Engineering. iii

CIM Cooperative Intersection Management. 23

CNN Convolutional Neural Network. 25

COCO Common Objects in Context. 43, 47

CP Central Point. 48, 49

CS Computer Science. iii

DB Database. 16, 28, 77, 78

DSR Design Science Research. iii, 32–34, 36, 37, 87

DSSD Deconvolutional Single-Shot Multibox Detector. 21

EU European Union. 15

FoV Field of View. 26, 53, 55

FPS Frames Per Second. 21, 36, 37, 46, 58, 64

G-CNN Grid Convolutional Neural Network. 21

GDPR General Data Protection Regulation. 86

GNSS Global Navigation Satellite System. 27, 29, 79

GPRS General Packet Radio Service. 4

GPS Global Positioning System. iii, 4, 6, 7, 9, 10, 12, 13, 20, 27, 29, 31, 32, 36, 37, 42, 51, 52, 59–69, 72, 78, 79, 81–84, 86, 88, 89

HDV Human-Driven Vehicle. iii, 5, 40, 41, 76

HLF High-Level Fusion. 28

HT Hough Transform. 25

ID Identification. 40, 46

ILD Inductive Loop Detector. 4, 19

IMU Inertial Measurement Unit. 20, 83

IS Information System. 27

IT Information Technology. 36

ITMS Intelligent Traffic Management System. iii, 3–7, 11, 12, 15, 16, 23, 27, 29, 31, 33, 34, 39–42, 47, 61, 67, 75–82, 84–89

ITS Intelligent Transportation System. 5, 15

KITTI Karlsruhe Institute of Technology and Toyota Technological Institute. 47, 50, 52, 68

LDWS Lane Departure Warning System. 17

LiDAR Light Detection And Ranging. 4, 19, 20, 26, 28, 29, 41

LLF Low-Level Fusion. 28

MLF Mid-Level Fusion. 28

MV Modern Vehicle. iii, 4–13, 15, 16, 19, 20, 23, 24, 27, 29, 31–34, 36, 37, 39–43, 46–48, 50–54, 56–65, 67–72, 75–89

NED North-East-Down. 63

NPRA Norwegian Public Roads Administration. 5, 62

NSD Norwegian Centre for Research Data. 86

PCF Pixel Coordinate Frame. 54

PPHT Progressive Probabilistic Hough Transform. 43, 47, 80

R-CNN Region-Based Convolutional Neural Network. 21, 24, 25

RFID Radio-Frequency Identification. 16, 19

RMSE Root Mean Square Error. 60, 61

RNN Recurrent Neural Network. 25

RoI Region of Interest. 43, 47

RON Reverse connection with Objectness prior Networks. 21

RSS Received Signal Strength. 28

RSSI Received Signal Strength Indicator. 28

SAE Society of Automotive Engineers. 17

SLR Systematic Literature Review. iii, 11, 24, 34, 36, 37, 39–42, 61, 67, 75–77, 87

SSD Single-Shot Multibox Detector. 21

TMS Traffic Management System. 3, 15, 76, 85

ToA Time of Arrival. 27, 28

ToF Time of Flight. 19, 26

V2I Vehicle-to-Infrastructure. 5, 16, 20, 21, 29, 40, 85

V2P Vehicle-to-Pedestrian. 5

V2V Vehicle-to-Vehicle. 5, 16, 20, 29, 40, 85

V2X Vehicle-to-Everything. 5, 40, 41

WAAS Wide Area Augmentation System. 83

WIFI Wireless Fidelity. 4

WiMAX World interoperability for Microwave Access. 4

YOLO You Only Look Once. 21, 22, 24, 43, 46–48, 50, 52, 68, 80

Part I

Synopsis

Chapter 1

Introduction

This chapter is composed of five sections. Section 1.1 presents the problem statement and the subject area of the thesis. Section 1.2 lays out the research motivations. Section 1.3 describes the research questions that are addressed in this thesis. Section 1.4 gives an overview of the research outcomes, the research papers included in the thesis, and the research contributions. The structure of the thesis is presented in Section 1.5.

1.1 Problem Statement

An increasing population and the need for more vehicles have led to increasing traffic congestion. Traffic congestion can cause traffic costs due to travel delays, excessive fuel consumption, air/noise pollution, etc. Based on the INRIX report [59], "on average, Americans lost 99 hours a year due to congestion, costing them nearly $88 billion in 2019, an average of $1,377 per year. From 2017 to 2019, the average time lost by American drivers has increased by two hours as economic and urban growth continue nationally." This trend is expected to continue in the next years, with traffic congestion forecast to increase by 60% by 2030 [112].

Traffic Management Systems (TMSs) can greatly help to mitigate traffic congestion, besides enhancing performance. Over the past decades, TMSs have tended to become intelligent, and Intelligent Traffic Management Systems (ITMSs) have been introduced thanks to the advancement of technologies in hardware (such as sensors), software (such as computational technologies), and networks (such as wireless communications). ITMSs are mainly based on traffic data to make smart decisions and manage traffic intelligently. Against that background, modeling the traffic scene dynamically is among the most important factors for providing a safe and efficient transportation system [75]. Thus, a considerable amount of literature has been devoted to the theme of ITMSs. Studies on ITMSs have generally demonstrated positive effects on managing the traffic. An ITMS is able to mitigate traffic delays, air pollution, and traffic costs, besides enhancing the safety and optimizing the route planning [138].

Accessing traffic data is one of the key components of an ITMS. Traffic data play an important role in boosting awareness about the traffic scene and, as a result, improve the ITMS's performance. A growing body of literature aims to estimate traffic data (hereafter, collecting traffic data and estimating traffic data are used interchangeably). The existing approaches to estimating traffic data can be classified into two main groups: (1) stationary sensors (also known as roadway sensors, static sensors, road traffic sensors, and in-road sensors) and (2) mounted sensors on vehicles (also known as mobile sensors, on-board sensors, and in-vehicle sensors).

The first approach involves monitoring and estimating traffic data via stationary sensors directed toward road networks. Stationary sensors are able to estimate traffic data on a selected location, such as estimating the queue length at an intersection [46]. In this approach, various types of sensors, such as Inductive Loop Detectors (ILDs), radar sensors, ultrasonic sensors, and video cameras, are used to estimate different traffic data types. Each type of sensor has its strengths and weaknesses in collecting different types of data. For instance, radar sensors are efficient in estimating the direction of motion of a vehicle, vehicular volume, and speed. Radars are also used by applications for managing traffic lights [46]. However, radars cannot count vehicles that pass in a parallel direction, as well as low-speed or stationary vehicles in heavy traffic [36]. Although stationary sensors are widely used to collect traffic data these days, installing and maintaining these sensors to provide an acceptable coverage range on all roads might be costly [66].

The second approach estimates traffic data via mounted sensors on a Modern Vehicle (MV). This approach is introduced by inventing novel MVs, such as Autonomous Vehicles (AVs). These types of vehicles are equipped with various types of sensors, such as Global Positioning System (GPS) receivers, video cameras, radar sensors, and Light Detection And Ranging (LiDAR) sensors [46]. They are also able to build wireless communication, such as Wireless Fidelity (WIFI), General Packet Radio Service (GPRS), World interoperability for Microwave Access (WiMAX), and Bluetooth, to transmit and receive data [148]. Therefore, they are good resources to estimate and transmit/receive traffic data while driving. These sensors are mainly used for self-awareness purposes and to provide data to enable autopilot/autonomous transportation and mobility, for instance, monitoring the driver's body posture and scanning the road for frontal, side, and rear collisions. [46]. In addition, MVs are able to share these data with other MVs, infrastructure, and pedestrians by Vehicle-to-Vehicle (V2V), Vehicle-to-Infrastructure (V2I), and Vehicle-to-Pedestrian (V2P) communications, which are collectively called Vehicle-to-Everything (V2X) communications. The data collection and the transmitting/receiving of data of MVs have motivated researchers to consider MVs as mobile sensors. However, there are some challenges in this regard. For instance, most of the studies in this context considered MVs equipped with various types of advanced technology and high-quality sensors. These high-quality sensors give the possibility of collecting accurate data. However, we need to keep in mind that the cost of a vehicle is related to the number and type of mounted sensors on the vehicle. Therefore, many car manufacturers might tend to produce various vehicles with different features to fulfill their customers' requirements with different budget levels. Proposing an approach by considering only MVs with high-technology sensors (e.g., fully AVs) would not be generalizable and practical in reality. Therefore, it is important to consider a low-cost sensor that has a high probability of being mounted on MVs. In addition, studies show that only 50% of vehicles in the United States will have autonomy in Level 4 by 2050 [114] (different levels of automation are presented in Section 2.2). Also, it is predicted that by 2032 only half of all new vehicles will be autonomous [2]. Therefore, changing most vehicles into a modern version with advanced sensors takes time. Moreover, after this period, some people might still enjoy manual driving with traditional vehicles. Therefore, the near-future traffic would be a mixture of both Human-Driven Vehicles (HDVs) and MVs (hereafter called mixed traffic). So far, very little attention has been paid to mixed traffic. There remains a paucity of evidence and empirical studies on managing traffic intelligently in this context. Therefore, questions have been raised about the use of MVs equipped with low-cost and popular sensors as a mobile sensor for estimating the required traffic data about themselves and the surrounding vehicles to satisfy the needs of ITMSs in mixed traffic.

1.2 Research Motivations

Studies show that understanding the traffic scene is a major area of interest within the field of ITMSs [8]. The performance of an ITMS depends on accurate and global awareness of the traffic scene achieved by estimating traffic data. For instance, the Norwegian Public Roads Administration (NPRA) tested Intelligent Transportation System (ITS) technology in the road from Skibotn, Norway, to Kilpisjärvi, Finland, to collect real-time information about the road surface conditions, traffic incidents, and weather and to provide warnings of wildlife or obstacles on the road [13]. In Gui'an, China, a project has been funded to develop a novel ITS to improve transportation safety and to reduce pollution and traffic congestion [34].

Furthermore, several studies have confirmed the effectiveness of image-based traffic data estimation via MVs. For instance, several methodologies are proposed by researchers in order to apply image-based vehicle detection and tracking (e.g., [15][80]), lane detection (e.g., [88][137][133]), target vehicle's distance estimation (e.g., [35][52]), and target vehicle's speed estimation (e.g., [113][64]).

Although much research has been carried out on traffic awareness and MVs' perception, the feasibility of using data estimated by such vehicles for ITMSs, with a future perspective of generating a dynamic model of the traffic scene to enhance ITMS performance, is not studied systematically and empirically, and there remains a paucity of evidence in that context.

Acknowledging the difficulties of generating a dynamic model of the traffic scene based on the collected traffic data from MVs, the main motivation of our research is defined as follows: to propose novel methodologies and develop new algorithms to explore the feasibility of using MV(s) equipped with low-cost sensors (i.e., a monocular camera with a built-in GPS receiver; as the monocular camera and a built-in GPS receiver can be a single unit, for instance, a GoPro Hero 7 camera, we use both terms interchangeably) as a mobile sensor to collect the required traffic data (i.e., vehicles' number, type, relative position, distance, speed, lane, and geolocation of the observed vehicle, which is called target vehicle or detected vehicle hereafter, although all three terms are used interchangeably). In other words, the motivation of this study is to take advantage of MVs' sensing capabilities and investigate their potential to be used as a replacement of stationary sensors, as well as a data resource for generating the digital twin of the traffic.

1.3 Research Questions

This thesis addresses the gaps in the current state of knowledge with the following Main Research Question (MRQ):

- **MRQ:** How can a mobile MV equipped with a low-cost monocular camera be used as a mobile sensor to estimate the ITMS's required traffic data in the context of mixed traffic in a metropolitan area?

  The MRQ is broken down into four Research Questions (RQs), as follows:

  - **RQ1:** What is the state of the art in managing an intersection intelligently in the context of both pure traffic of MVs (i.e., AVs) and mixed traffic at four-way signalized and unsignalized intersections?
  - **RQ2:** How can traffic data, such as the number, type, relative position, distance, speed, lane, and geolocation of multiple and mobile target

vehicles be estimated via a single mobile MV equipped with a front-facing monocular camera with a built-in GPS receiver in mixed traffic?

– **RQ3:** How can the self-localization accuracy of an MV be enhanced via two mounted low-cost built-in GPS receivers?

– **RQ4:** How can the accuracy of the estimated geolocation of the target vehicle be increased based on multiple sensor fusion techniques?

Firstly, RQ1 aims to ground the research by systematically reviewing the recent literature on ITMSs to understand the state of the art. RQ2 and RQ3 aim to propose new methodologies and develop new algorithms to estimate the most critical traffic data identified from RQ1 by a single MV. RQ4 explores how to fuse the estimated geolocations of a target vehicle observed by two MVs.

1.4 Research Outcomes

The RQs are addressed in six published/submitted research papers in peer-reviewed journals and conference proceedings.

1.4.1 Research Papers

The research papers that address the RQs are listed below. The connections between the research papers and the RQs are illustrated in Table 1.1. In addition, the coherence between the papers included in this thesis is illustrated in Figure 1.1.

- **Paper A:** Elnaz Namazi, Jingyue Li, and Chaoru Lu, "Intelligent Intersection Management Systems Considering Autonomous Vehicles: A Systematic Literature Review". In: *IEEE Access Journal 7 (2019)*, pp. 91946-91965, Status: published.

  Authors' contribution: Namazi led the paper writing and was the main author. Namazi was responsible for developing the research design and conceptualization. Li and Lu supervised this process by regular consensus meetings with Namazi. Namazi performed the keyword search process and selected papers based on the inclusion and exclusion criteria. Namazi extracted the data, thematically categorized the findings, and prepared the research results. Lu contributed to this process. All authors discussed the results. Namazi wrote the paper, and Li and Lu commented on the paper.

  Relevance to the thesis: This paper systematically reviews the recent literature on ITMS by focusing on both pure MVs (i.e., AVs) traffic and mixed traffic at four-way signalized and unsignalized intersections. The paper contributes to addressing RQ1. Our findings obtained in this paper help to formulate RQ2 - RQ4 and to conduct Paper B - Paper F. Paper A helps us with the following objectives:

- To identify the factors (goals) that were considered by other researchers regarding intelligent intersection management systems in terms of utilizing MVs.
- To gain knowledge about the related proposed methodologies and to consider traffic data types in the existing literature and categorize them.
- To explore the potential research gaps.
- To design and conceptualize our research and to formulate our RQs related to this Ph.D. journey.

- **Paper B:** Elnaz Namazi, Rein Nisja Holthe-Berg, Christoffer Skar Lofsberg, and Jingyue Li, "Using Vehicle-Mounted Camera to Collect Information for Managing Mixed Traffic". In: *International Conference on Signal-Image Technology & Internet-Based Systems (SITIS), IEEE (2019)*, Status: published.

  Authors' contribution: This paper was mainly conducted based on the master's thesis of Holthe-Berg and Lofsberg, supervised by Li, and Namazi contributed to the co-supervising process. Namazi was responsible for conceptualizing the research, including identifying the research motivation, scope, and RQs based on the findings obtained in Paper A. Holthe-Berg and Lofsberg focused on the technical parts, including data collection, algorithm development, and running experiments. Regular consensus meetings of all authors approved each step of this research process. All authors discussed the results, and Namazi wrote the paper based on the findings. Li commented on the paper. Namazi attended the conference and presented the paper.

  Relevance to the thesis: This paper proposes a new methodology and provides a new system by developing new algorithms that integrate with the state of the art. This paper contributes to addressing one part of RQ2 (i.e., estimating the traffic data, such as the number, type, relative position, distance, and speed of the target vehicle). Our findings obtained in this paper help to address some parts of RQ2 (i.e., estimating the target vehicle's lane and geolocation) and RQ4 and to conduct Paper C, Paper E, and Paper F. Paper B helps us with the following objective:

  - To explore the feasibility of using a mobile MV equipped with a front-facing low-cost monocular camera to estimate the required data (i.e., estimating the number, type, relative position, distance, and speed of the target vehicle) in mixed traffic in a metropolitan area.

- **Paper C:** Elnaz Namazi, Jingyue Li, Rudolf Mester, and Chaoru Lu, "Identifying and Counting Vehicles in Multiple Lanes by Using a Low-Cost Vehicle-Mounted Sensor for Intelligent Traffic Management Systems". In: *International Conference on Hybrid Artificial Intelligence Systems (HAIS), Springer, Cham (2020)*, Status: published.

  Authors' contribution: Namazi led the paper writing and was the main author. Namazi was responsible for developing the research design and conceptualization. Li, Mester, and Lu supervised this process by regular consensus meetings with Namazi. Namazi proposed a novel methodology, and Mester contributed to this process. Namazi developed the new algorithms, designed and performed the experiments, and analyzed the findings. All authors discussed the results, and Namazi wrote the paper based on the findings. Li, Mester, and Lu commented on the paper. Namazi presented the paper at the online conference.

  Relevance to the thesis: This paper proposes a new methodology and provides a new system by developing new algorithms that integrate with the state of the art. This paper contributes to addressing one part of RQ2 (i.e., estimating the target vehicle's lane). Paper C helps us with the following objective:

  – To explore the feasibility of using a mobile MV equipped with a front-facing low-cost monocular camera to estimate the lane the target vehicle is in, relative to the MV lane in a multiple-lane street in mixed traffic in a metropolitan area.

- **Paper D:** Elnaz Namazi, Rudolf Mester, Chaoru Lu, Markus Metallinos Log, and Jingyue Li, "Improving Vehicle Localization with Two Low-cost GPS Receivers". In: *Smart City Applications (SCA) International Conference (2021)*, Status: published.

  Authors' contribution: Namazi led the paper writing and was the main author. Namazi was responsible for developing the research design and conceptualization. Li, Mester, and Lu supervised this process by regular consensus meetings with Namazi. Namazi designed the data collection process from real traffic, and Li supervised this step. Namazi proposed the novel methodology, and Mester and Log contributed to this process. Namazi developed the new algorithms, designed and performed the experiments, and analyzed the findings. All authors discussed the results, and Namazi wrote the paper based on the findings. Li, Log, Mester, and Lu commented on the paper. Namazi presented the paper at the online conference.

  Relevance to the thesis: This paper proposes a new methodology and provides a new system by developing new algorithms that integrate with the state of the art. This paper contributes to addressing RQ3. Our findings obtained in this paper help to address one part of RQ2 (i.e., estimating the target vehicle's geolocation) and RQ4 and to conduct Paper E and Paper F. Paper D helps us with the following objective:

  – To address the low-cost GPS receiver uncertainty in MV self-location, when the GPS signal is noisy by keeping the sensor cost low.

• **Paper E:** Elnaz Namazi, Rudolf Mester, Chaoru Lu, and Jingyue Li, "Geolocation Estimation of Target Vehicles Using Image Processing and Geometric Computations". In: *Neurocomputing Journal, Elsevier (2021)*, Status: accepted.

  Authors' contribution: Namazi led the paper writing and was the main author. Namazi was responsible for developing the research design and conceptualization. Li, Mester, and Lu supervised this process by regular consensus meetings with Namazi. Namazi and Mester contributed to proposing a novel methodology and developing a mathematical model. Namazi developed the new algorithms, designed and performed the experiments, and analyzed the findings. All authors discussed the results, and Namazi wrote the paper based on the findings. Li, Mester, and Lu commented on the paper.

  Relevance to the thesis: This paper proposes a new methodology and provides a new system by developing new algorithms that integrate with the state of the art. This paper contributes to addressing one part of RQ2 (i.e., estimating the target vehicle's geolocation). Our findings obtained in this paper help to address RQ4 and to conduct Paper F. Paper E helps us with the following objective:

  – To explore the feasibility of a mobile MV equipped with a low-cost monocular camera to estimate the geolocation of multiple target vehicles in a GPS coordinate system.

• **Paper F:** Elnaz Namazi, Rudolf Mester, Jingyue Li, Chaoru Lu, Meng Tang, and Ying Xiong, "Traffic Awareness Through Multiple Mobile Sensor Fusion" In: *IEEE Sensors Journal (2021)*, Status: submitted.

  Authors' contribution: Namazi led the paper writing and was the main author. Namazi was responsible for developing the research design and conceptualization. Li, Mester, and Lu supervised this process by regular consensus meetings with Namazi. Tang and Xiong contributed to the data collection process. The same data were used to run experiments in Paper D and Paper E, for which Namazi designed the data collection process from real traffic, and Li supervised this step. Namazi and Mester contributed to proposing a novel methodology and developing a mathematical model. Namazi developed the new algorithms, designed and performed the experiments, and analyzed the findings. Namazi, Li, Mester, Lu discussed the results, and Namazi wrote the paper based on the findings. Li, Mester, Lu, and Xiong commented on the paper.

Relevance to the thesis: This paper proposes a new methodology and provides a new system by developing new algorithms that integrate with the state of the art. This paper contributes to addressing RQ4. Paper F helps us with the following objective:

– To use multiple sensor fusion techniques to integrate the estimated geolocations of a specific target vehicle by two MVs equipped with a low-cost monocular camera. The goal is to improve the accuracy of the target vehicle's geolocalization estimation and to provide a more comprehensive picture of the traffic scene than what can be obtained by using data from only one MV.

	Paper A	Paper B	Paper C	Paper D	Paper E	Paper F
RQ1	●					
RQ2		●	●		●	
RQ3				●		
RQ4						●

Table 1.1: Mapping of research papers and RQs.

1.4.2 Research Contributions

This thesis is highly interdisciplinary and makes contributions to establishing a link between MVs' capabilities to collect traffic data and ITMS's required data. The connections between the research contributions and the RQs are illustrated in Table 1.2. The connections between the research papers, the RQs, and the research contributions are illustrated in Figure 1.2.

This thesis has four major contributions:

- **C1** *Generated new knowledge by systematically reviewing, summarizing, and conceptualizing the state of the art in managing an intersection intelligently with a focus on (1) signalized intersections for both pure MVs (i.e., AVs) and mixed traffic and (2) unsignalized intersections if the traffic includes pure MVs.* The Systematic Literature Review (SLR) proposes new

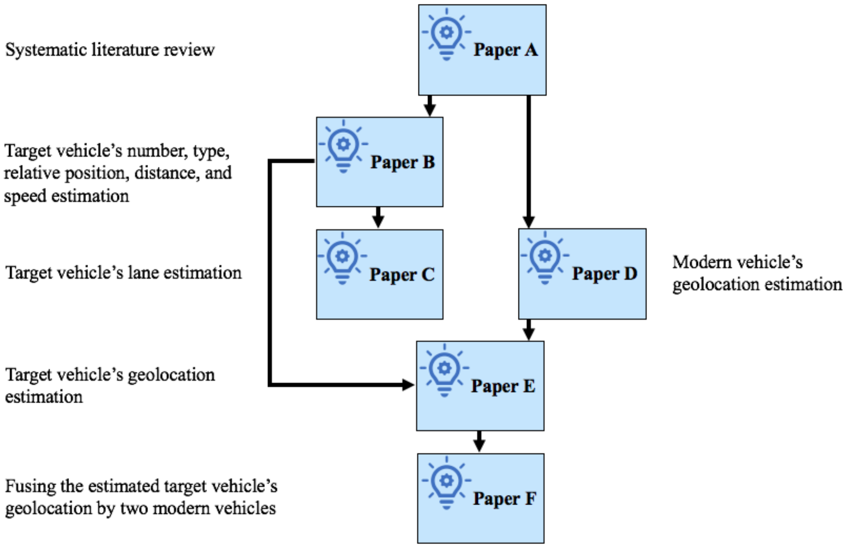


Figure 1.1: Coherence between our research papers.

taxonomies to categorize and summarize the state of the art in proposed methodologies, considered factors (goals), and collected traffic data types. Also, it explores potential research gaps to formulate the conceptual framework of this Ph.D. study.

- **C2** *Proposed new methodologies and algorithms to estimate the required traffic data of the target vehicle via a single MV equipped with a low-cost monocular camera in mixed traffic.* This contribution includes reviewing the literature, proposing a new methodology, developing new algorithms, experimenting on real traffic data, and analyzing findings. This contribution helps us to build a bridge between MVs' sensing capabilities and the traffic data required by ITMSs. This contribution is mainly based on object detection, lane detection, image processing, and geometric computation techniques. This contribution investigates the feasibility of an MV equipped with a low-cost sensor (i.e., a monocular camera with a built-in GPS receiver) to be utilized as a mobile sensor to collect the required traffic data in mixed traffic. The low-cost sensor is considered with the aim of proposing a new approach that is compatible with existing ITMSs, current advanced vehicles, and future MVs to enhance our proposed approach's generalizability, potential usability, and practicality in the real world.

- **C3** *Proposed a new methodology and algorithms to enhance the self-localization accuracy of an MV by using low-cost GPS receivers.* This contribution includes reviewing the literature, proposing a new methodology, developing new algorithms, experimenting on real traffic data, and analyzing findings. This contribution is mainly based on integrating cross-GPS validation, interpolation/best-fit, and map-matching techniques to localize an MV in the presence of GPS signal noise. This contribution investigates the feasibility of using two low-cost GPS receivers on the same MV with a known distance from each other to enhance the MV's localization accuracy. The key point of this contribution is to enhance the self-localization accuracy while keeping the cost of the sensor receiver low.

- **C4** *Proposed a new methodology and algorithms to fuse the estimated geolocations of the observed target vehicle via two MVs equipped with a low-cost monocular camera by considering sensor estimation uncertainty in mixed traffic.* This contribution includes reviewing the literature, proposing a new methodology, developing algorithms, experimenting on real traffic data, and analyzing findings. This contribution is mainly based on re-identification and multiple sensor fusion techniques. As in C2, one of the most significant points of this contribution is considering a low-cost monocular camera on MVs as a mobile sensor and considering mixed traffic.

	C1	**C2**	**C3**	**C4**
RQ1	●			
RQ2		●		
RQ3			●	
RQ4				●

Table 1.2: Mapping of contributions and RQs.

1.5 Structure of the Thesis

The thesis is composed of Parts I and II, as follows:

Part I: This part presents an introduction to the research work and provides an overview of the background, related work, research methodology, results, discussion, and conclusion and future work.

Chapter 2: Gives the background related to the research concepts.

Chapter 3: Presents the related work based on the RQs.

Chapter 4: Presents the applied research methodology and research strategies to address each RQ.

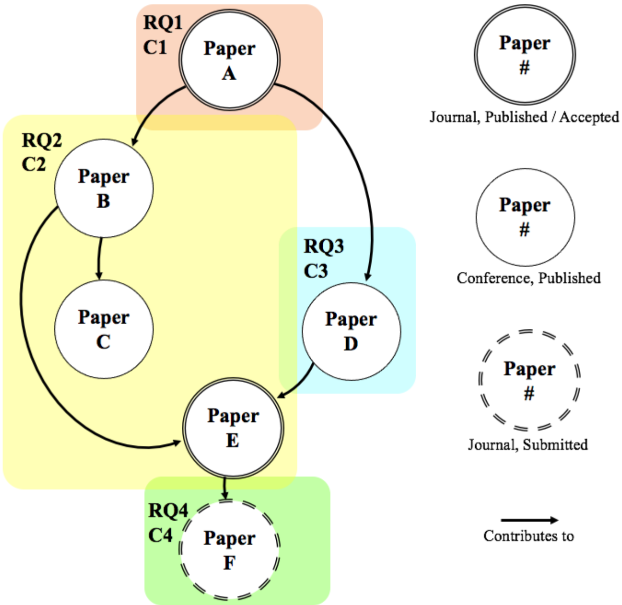


Figure 1.2: A schema of the research papers, RQs, and research contributions.

Chapter 5: Summarizes the research findings based on the RQs by focusing on the included papers in this thesis.

Chapter 6: Discusses the results of the RQs in terms of a comparison with related work, implications for academia, implications for practitioners, and threats to validity, as well as the ethical issues of the research.

Chapter 7: Concludes and gives suggestions for potential future studies.

Part II: Contains the collection of the six research papers in full length that constitute this thesis.

Chapter 2

Background

This chapter introduces the concepts used in this thesis. It includes five sections. Section 2.1 gives a brief overview of the ITS and ITMS concepts. The next sections present the concepts of MVs, sensors, vehicle communications, and object detection using deep learning algorithms, which are relevant to this thesis.

2.1 Intelligent Transportation System and Intelligent Traffic Management System

According to a definition provided by the European Union (EU):

"Intelligent Transport Systems (ITS) are advanced applications which without embodying intelligence as such aim to provide innovative services relating to different modes of transport and traffic management and enable various users to be better informed and make safer, more coordinated and 'smarter' use of transport networks" [30].

Developing modern technologies such as sensors, network bandwidth, cloud computing, big data analytics, and computer vision are leveraged to develop advanced TMSs [138], named ITMSs. An ITMS is an important component of smart cities, which is based on communication between vehicles, devices, and other individual entities in integrated networks [138]. Such systems, by using a variety of sensors, are able to track and monitor the flow of vehicles through road networks, optimize route planning by their awareness of unforeseen road events (e.g., accidents, vehicle breakdowns, or roadblocks), plan and allocate parking spaces based on demand, give priority to emergency vehicles, and install automated toll management [138]. Overall, ITMSs decrease travel duration, traffic jams, management costs, financial losses, and air/noise pollution. In addition, ITMSs improve safety and pave the way for sustainable and cleaner smart cities [138].

Based on [87][93], Alsrehin et al. in [6] identified the following general steps to develop intelligent transportation and control systems.

- Collection: The first step is to collect traffic data via various methods, such as image/video-based methods, sensor-based methods (e.g., photoelectric sensors, ultrasonic sensors, Radio-Frequency Identifications [RFIDs], lasers, radar, and vehicle probe data; some of the most popular sensors are briefly described in Section 2.3), V2V and V2I communications (which are briefly described in Section 2.4), and hybrid-based methods that combine two or more of the aforementioned methods [6].
- Pre-processing: The collected raw data from the above methods is subject to noise, missing values, and inaccurate data. Therefore, pre-processing approaches, such as data cleaning, dimensionality reduction, sparsity analysis, and data fusion, are required [6].
- Analysis: To provide meaningful information (i.e., traffic density in a specific roadway segment on a specific day of the year) requires analyzing data with special tools. These tools are mainly based on machine learning, data mining, and Artificial Intelligence (AI) algorithms [6].
- Storage: To store big traffic data requires suitable storage, such as cloud storage.
- Communication: To use and share traffic data with the purpose of studying, planning, designing, constructing, operating, and monitoring traffic systems requires communication [6].
- Maintenance and archiving: Data maintenance is needed to include the ongoing correction and verification of data analysis results. Data archiving deals with moving less frequently used data from active systems and databases DBs to specialized archival systems to enhance the performance of intelligent transportation and control systems [6].

2.2 Modern Vehicles

Inventing MVs, such as AVs, has a significant effect on ITMSs. Such vehicles sense the surrounding environment through various sensors and collect data. These data can be used in self-awareness, the driving process, or managing the traffic. Figure 2.1 [2] shows the technical evolution of AVs over the years. This process started with developing a modern cruise control system in 1948. This process continued with mechanical antilock braking, electronic cruise control, electronic stability control, and laser-based adaptive cruise control. In the early 2000s, developments in Lane Departure Warning Systems (LDWSs), pre-crash mitigation, DARPA challenges, and active parking assistance played critical roles in this development process. This development process of AVs is still evolving. It is anticipated that fully automated AVs with no driver backup will be available by 2030 [2].

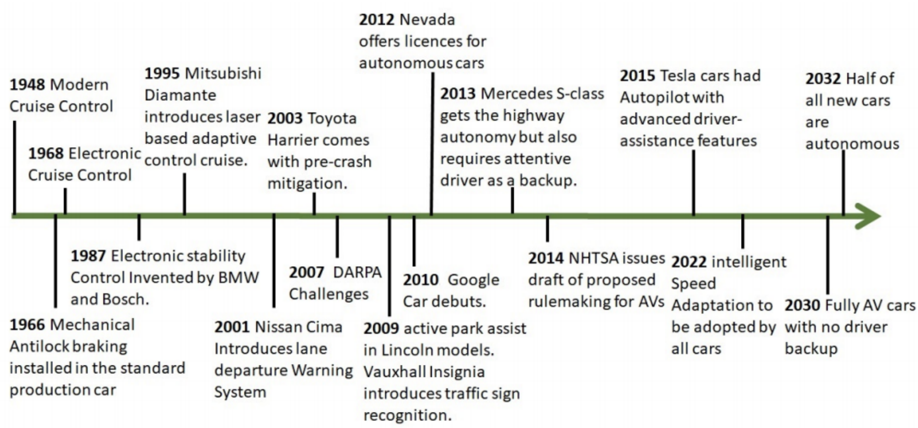


Figure 2.1: Technical evolution of AVs. Figure from [2].

The Society of Automotive Engineers (SAE) classified vehicles into six different levels of automation with its $J3016^{(TM)}$ "Levels of Driving Automation" standard, from no automation to full automation [122]. This classification is presented in Figure 2.2 as a visual chart.

The detailed features of each level of automation are as follows [2][122]:

- Level 0: No automation. At this level, all tasks are performed by the driver. This level provides only limited driver support features, such as warnings and momentary assistance.

- Level 1: Driver assistance. This level provides driver support features, such as steering or brake/acceleration support.

- Level 2: Partial automation. This level provides driver support features, such as steering and brake/acceleration support. However, the driver is responsible for many safety-critical actions.

- Level 3: Conditional automation. This level provides conditional driving automation, with automated driving features driving the vehicle under limited conditions. At this level, the driver is not responsible for safety-critical issues.
- Level 4: High automation. This level provides conditional driving automation, with automated driving features driving the vehicle under limited conditions, and if the automated situation turns unsafe, then the driver holds control.
- Level 5: Full automation. This level is completely automatic and can drive the vehicle under all conditions, and there is no need for human intervention.

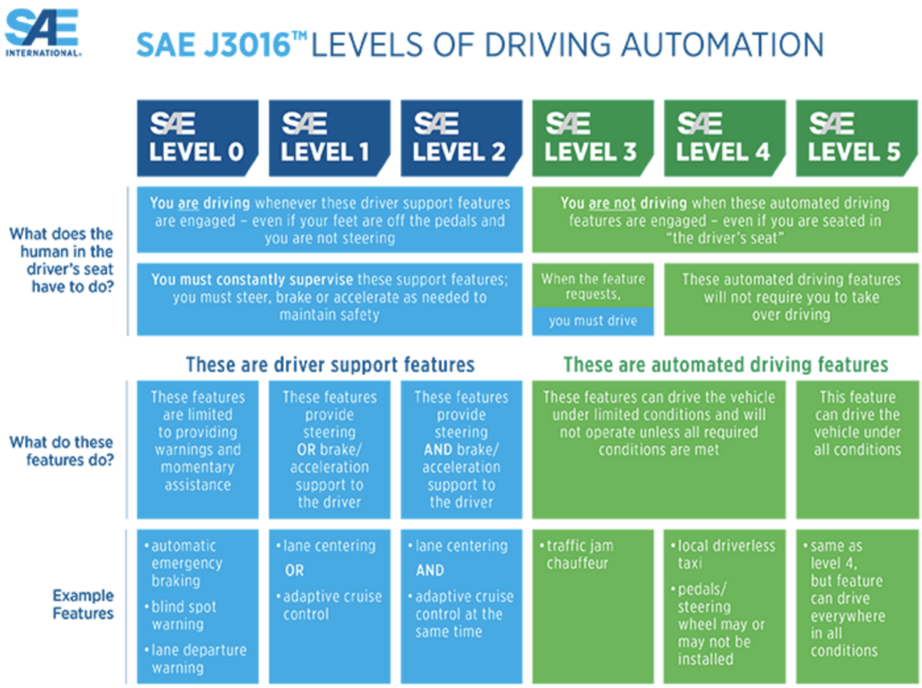


Figure 2.2: $J3016^{(TM)}$ Levels of driving automation. Figure from [122].

2.3 Sensors

Guerrero-Ibáñez et al. [46] classified traffic sensors into two major categories, namely, in-road sensors and in-vehicle sensors.

- **In-road sensors**

In-road sensors, based on their placement, are classified into intrusive sensors (e.g., ILDs) and non-intrusive sensors (e.g., video cameras, radar sensors, and RFIDs) [12]. For instance, an ILD is a type of intrusive sensor that is used to detect a vehicle's presence, movement, number, and occupancy [46]. Intrusive sensors are expensive and may be affected by environmental conditions [46]. Non-intrusive sensors are another type of in-road sensor. For instance, video cameras can be used to perform vehicle detection in multiple lanes, to classify vehicles, and to identify vehicles' presence, flow rate, occupancy, and speed. Non-intrusive sensors provide many of the intrusive sensors' functions with fewer difficulties. However, climate conditions highly affect non-intrusive sensors, such as rain, snow, and fog [46].

- **In-vehicle sensors**

  Mounted sensors on MVs create both a perceptive and locational view of the environment to make decisions in real time [18]. Campbell et al. [18] classified the most commonly used sensors in MVs (with the main focus on AVs) into two categories, as follows:

  1. Exteroceptive sensors: These sensors are utilized to perceive the environment and to calculate the distance to objects [18]. Some of the most popular sensors in this category are described below.

  – LiDAR: LiDAR is a technology that uses a laser of light, which is mostly used to measure distances based on the principle of Time of Flight (ToF). It sends out a pulsed laser of light and measures the time it takes for the pulse to be reflected back. These measurements enable an MV to generate 3D representations of the surrounding environment based on the point cloud data. LiDARs can be classified as a long-range sensor, as they have a range of over 250 m. LiDARs provide high-precision and accurate localization. However, they are costly [18].

  – Radar: Radar uses radio waves to mainly measure the distance, velocity, and angle of objects. Also, they are capable of determining the relative motion of detected objects. Radar sensors are mostly considered as short- to medium-range sensors (50 m - 100 m) [18]. Radars are significantly cheaper than LiDAR. Radar sensors are commonly used in MVs with Advanced Driver Assistance Systems (ADASs) to provide cruise control functions and collision detection [18].

  – Camera: A camera creates a digital image of a covered region using the principle of passive light sensors. The ability to see colors and textures is the key benefit of cameras, which enables MVs to identify road signs, traffic lights, lane markings, etc. Camera systems have a sensing range from 4 m to 80 m [51][109]. Measuring the distance to an object by using a complex processing algorithm is possible as well. The main advantage of a camera is that it is cheaper and better available than LiDAR. However, the weather and light conditions will affect its performance. Cameras are a key technology for fully autonomous navigation, according to most AV manufacturers; however, they tend to fuse their data by radar or LiDAR data [18].

2. Proprioceptive sensors: These sensors are utilized to measure values from within the system (e.g., motor speed) [18]. Some of the most popular sensors in this category are described below.

- GPS: A GPS receiver is a satellite-based radio-navigation system for navigation and localization in AVs. The biggest drawback of using GPS technology for autonomous navigation is that a variety of factors can harm the positioning accuracy. GPS receivers need a direct line of sight with the satellites [18]. The accuracy of the data collected via a GPS receiver depends on several parameters, such as hardware accuracy, satellite geometry, signal blockage, and atmospheric conditions [73].
- Inertial Measurement Unit (IMU): An IMU is an electronic device that is commonly used for the controlling and guiding of MVs. The key drawback of IMU devices is that they can provide only information about the vehicle's motion, not its actual location, which must be calculated by using other sensors, such as a GPS receiver [18].

2.4 Vehicle Communications

Cooperative vehicular networks are widely applied in intelligent transportation-related applications. Two major types of communications are V2V and V2I (e.g., [146][9][7][105][53][63][27][31][45]). The concept behind the V2V communication model is to provide a virtual bridge among nearby vehicles to transfer data. V2V communication enables vehicles to exchange data such as speed, position, and direction with other nearby vehicles; the receiving vehicles will then be able to make smart decisions by aggregating these messages. This connection has a limited lifetime due to the high speed of the vehicles [20]. One potential solution is to extend the transmission range of the sensors used in the vehicles. However, this depends on transmission power, which is an important parameter in this regard. Furthermore, privacy concerns should not be overlooked since vehicles may also transmit personal data [20]. Vehicles in the V2I connectivity model communicate with each other through an intermediate infrastructure that is typically fixed and mounted along the roadside [20]. This approach is ideal when data must be broadcast to all network nodes, as in road hazard detection. Many issues, such as privacy and lifetime connectivity, can be mitigated in V2I communication [20].

2.5 Object Detection Using Deep Learning Algorithms

"Object detection is defined as a process using an image including several objects as input to locate and classify as many target objects as possible in the image" [147]. The proposed deep learning-based object detection approaches can be classified into two-stage and one-stage methods [10]. The two-stage methods are regional proposal-based methods [10] and involve two steps: (1) extracting regions of interest from an image and (2) analyzing the candidate regions for final detection [76]. Two-stage methods include Region-Based Convolutional Neural Networks (R-CNNs) [43], Fast R-CNNs [42], and Faster R-CNNs [119]. One-stage methods are regression/classification-based methods [10] and directly predict the location of an object in an image and classify the object accordingly, without having to go through the regional proposal stage [76]. One-stage methods include You Only Look Once (YOLO) [118][145], Grid Convolutional Neural Networks (G-CNNs) [100], Single-Shot Multibox Detectors (SSDs) [83], Deconvolutional Single-Shot Multibox Detectors (DSSDs) [37], and Reverse connection with Objectness prior Networks (RONs) [70][131].

The main methods of deep convolution neural network-based object detection are shown in Figure 2.3. The object detector used in this thesis is YOLO, and it is briefly introduced below.

- YOLO: Prior to the introduction of the YOLO-V1 algorithm, the R-CNN series algorithm was the main algorithm with high object detection accuracy. However, because of its two-stage network structure, it was unable to satisfy the real-time requirements. As a result, an object detector with a quicker speed was needed [134]. In 2016, Joseph Redmon et al. [118] proposed a single-stage target detection network with a high detection speed and the ability to run in real time (i.e., 45 Frames per Second [FPS]). It attracted wide attention, and the YOLO series have five basic versions so far (V1 - V5) [134]. The basic concept of YOLO is to turn object detection into a regression problem. The entire image is utilized as the input of the network, and it is divided into a uniform grid. The same neural network is also used to determine the positions and categories of the bounding boxes. Anchor boxes and K-means clustering methods are introduced in

YOLO-V2 to enhance the positioning problem on the basis of regression and the use of single neural networks in YOLO-V1, while Darknet-19 is used as the basic classification model to reduce the number of training parameters and to improve the training speed [134]. To improve the accuracy and speed of YOLO-V2, YOLO-V3 was introduced. YOLO-V3 modified the softmax classifier of YOLO-V2 into an independent logistic regression classifier, and a residual structure was included in the backbone to expand the network depth [134]. However, applications were limited due to low accuracy and missed detection of objects with multi-scale features [134]. To reduce the requirements of the experimental equipment, YOLO-V4 was introduced. YOLO-V4 is a lightweight version of YOLO-V3, which can be trained by a single traditional GPU. However, it could not significantly improve the existing problems of YOLO-V3 [134]. YOLO-V5 was proposed in June 2020 to overcome the problem of missed and mis-check in multi-scale feature target detection by YOLO-V3 and YOLO-V4 by adding unique Focus and BottleneckCSP modules [134].

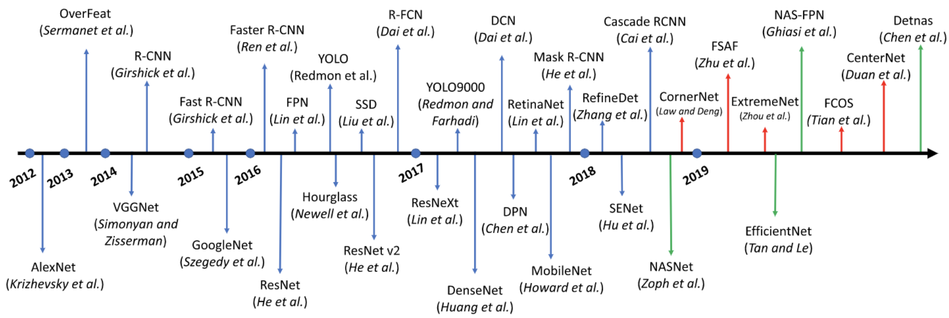


Figure 2.3: The major milestones in deep convolution neural network-based object detection research since 2012. Anchor-free (in red) and AutoML (in green) techniques. Figure from [142].

Chapter 3

Related Work

This chapter gives a brief overview of the related work in the research field, in line with the RQs addressed in this thesis. The chapter includes four sections and begins with reviewing the existing literature, with a focus on ITMSs and MVs. Section 3.2 is concerned with the proposed approaches to estimating the required traffic data (i.e., target vehicle detection and tracking, lane detection, and the target vehicle's distance, speed, and localization). Section 3.3 presents an overview of the MV self-localization approaches. Section 3.4 briefly looks at multiple sensor fusion techniques.

3.1 Literature Review of Intelligent Traffic Management Systems with Modern Vehicles - RQ1

Although there is a growing body of literature on ITMSs and MVs, there are relatively few surveys/reviews on managing signalized and unsignalized intersections intelligently by considering pure traffic of MVs and mixed traffic. Chen et al. [21] surveyed major techniques and solutions regarding Cooperative Intersection Management (CIM) by considering both signalized and unsignalized intersections. The studied cooperative methods in this survey mainly focused on resource (i.e., time slots and space) reservation, trajectory planning, virtual traffic lights, intersection collision avoidance, and vulnerable road users [21]. In addition, several projects concerning CIM were introduced, such as Compass4D [21]. Another survey by Rios-Torres et al. [121] focused on coordinating MVs (i.e., connected MVs and AVs) at intersections and merging at highway on-ramps. In [121], the existing approaches to coordinating vehicles crossing an intersection or merging at the merging zone without rear-end and lateral collisions were classified into (1) centralized approaches, in which a single central controller decides at least one task globally for all vehicles, and (2) decentralized approaches, in which each vehicle is responsible for establishing its own control policy based on the information received from the other vehicles or some coordinator [121]. Guo et al. [48] surveyed urban traffic control approaches and models by considering only signal control systems.

To the best of our knowledge, at the time that we worked on RQ1, there was no SLR available to provide a comprehensive and structured review of our defined study context (i.e., managing a 4-way intersection intelligently in the context of both pure traffic of AVs and mixed traffic) by exploring the research factors (goals), proposed methodologies, and considered traffic data types in the literature.

3.2 Estimating the Target Vehicle's Traffic Data - RQ2

This section briefly presents the related work regarding RQ2. The section includes target vehicle detection and tracking, lane detection, target vehicle's distance estimation, target vehicle's speed estimation, and target vehicle's localization.

3.2.1 Target Vehicle Detection and Tracking

Once the footage is collected via a mounted camera on the mobile MV, the first step is to detect objects. This study mainly focused on vehicles as traffic objects. After vehicle detection, we need to track the vehicles, enabling us to estimate other types of traffic data related to the target vehicle, such as speed. Both tasks (i.e., target vehicle detection and tracking) can be addressed using a variety of approaches [84].

In [84], the approaches to detecting vehicles are classified into four main categories. (1) The static background approach uses mainly stationary cameras to detect vehicles. This approach is based on a thresholding method and background subtraction (e.g., [113][47][23]). (2) The feature-based approach focuses on using different types of features grouped in the region around the vehicle, for instance edges (e.g., [24]), gray-level features (e.g., [98]), binary features or patterns (e.g., [24]), and corners (e.g., [97]). (3) License plate detection can also be used to detect target vehicles (e.g., [39][41][126]). (4) Learning-based methods for detecting target vehicles were introduced in 2007 [4], and they have become increasingly popular in recent years. For instance, Faster R-CNNs (e.g., [14]), extended versions of Mask R-CNNs (e.g., [72]), and different versions of the YOLO detector (e.g., [15]) have been used.

As mentioned, vehicle tracking plays an important role in estimating some of the required traffic data related to the target vehicle (e.g., speed). In [84], the proposed approaches to tracking target vehicles are classified into four main categories: (1)

feature-based tracking, which considers the target vehicle tracking problem as a feature tracking problem and uses features such as gray-level features, binary features, edges, and corners (e.g., [97]), (2) tracking of the centroid of the target vehicle's blob/bounding box (e.g., [49]), (3) tracking of the entire region of the vehicle (known as contour- or bounding box-based tracking) (e.g., [80]), and (4) license plate tracking, which exploits that the movement of the plate and that of the vehicle are linked (e.g., [136]).

3.2.2 Lane Detection

A lane detection system detects lane marks, which can be used to estimate the position of vehicles and their trajectory relative to the lane [101]. Lane detection can be beneficial for measuring the traffic density per lane and decision-making.

The proposed lane detection approaches in the literature are classified into two major groups [99]: (1) non-deep learning approaches, for example, the Hough Transform (HT) (e.g., [137]), the Sobel-Canny hybrid algorithm (e.g., [88]), and a combination of the Improved HT and the Sobel edge detector (e.g., [89]), and (2) deep learning approaches, such as Convolutional Neural Networks (CNNs) (e.g., [74]), Recurrent Neural Networks (RNNs) (e.g., [77]), and Faster R-CNNs (e.g., [133]).

3.2.3 Target Vehicle's Distance Estimation

To date, various studies have investigated the target vehicle's distance estimation, mainly for safety purposes, via monocular cameras. In [84], monocular-based distance estimation approaches are classified into three groups: (1) approaches that use intrusion or augmented lines or regions to measure the real distance between two or multiple virtual lines on the road or to measure the actual size of a road region (e.g., [25][62][149]), (2) homography as a linear transformation of a 3D projective space (the road) to a 2D projective space (the camera image plane) (this approach is also called m/px) (e.g., [35][115][58]), and (3) employing prior knowledge of the real dimensions of objects such as license plates (e.g., [41][39]) or target vehicles (e.g., [52]). In general, estimating distances in the lateral or cross-track directions is simple for a camera system. However, estimating distances in the longitudinal or along-track directions is challenging and less effective with a camera system [109].

3.2.4 Target Vehicle's Speed Estimation

The proposed approaches to estimating the target vehicle's speed are clustered into four main groups [84]. (1) Traffic speed approaches mainly focus on obtaining the average road speed (e.g., [113]) and an individual target vehicle's speed. It is important to know that measuring an individual target vehicle's speed is more popular than measuring the traffic speed. (2) Other approaches employ the time and distance between measurements (also called consecutive/non-consecutive approaches). As presented in Eq. 3.1, image-based speed estimation related to an individual target vehicle in a simple scenario (without considering the distance estimation error) is linked with the distance traveled and the time passed. To compute the target vehicle's speed v, different locations L_1 and L_2 of a target vehicle at times t_1 and t_2 are needed.

$$v = \frac{|L_2 - L_1|}{|t_2 - t_1|} = \frac{\Delta L}{\Delta t} \tag{3.1}$$

Based on Eq. 3.1, to calculate the vehicle's speed, either precise timestamps between measurements (or images) or prior knowledge of the camera's frame rate is required. It is important to know that considering specific timestamps per recorded measurement (or image) based on the recording system's clock is the more accurate approach. Two approaches can measure speed. The first approach is consecutive, which is based on frames t and $t + 1$ (e.g., [64][94]). Although this approach is popular, the obtained value is noisy as it is affected by distance errors. The second approach is non-consecutive and uses several techniques, such as a fixed distance or region between measurements (e.g., [4]), a pre-defined number of frames between measurements (e.g., [124]), the maximum possible distance between the first and the last detection of the tracked vehicle (e.g., [97]), and vehicle detection from two different cameras (e.g., [85]). (3) Measurement integration (also called instantaneous/mean approaches) aims to integrate all the N speed values of a vehicle. This approach is referred to as average speed detection, and the mean value is commonly used in this approach. However, the instantaneous technique calculates the speed based on Eq. 3.1 regardless of whether it is consecutive or non-consecutive. (4) Other uncommon approaches use aspects such as motion blur (e.g., [78]) or regression to estimate the vehicle's speed.

3.2.5 Target Vehicle's Localization Estimation

De Ponte Müller [109] surveyed different techniques for a vehicle's relative positioning. They classified the strategies for the relative positioning of vehicles into two groups [109]: (1) non-cooperative positioning, which is based on using mounted ranging sensors inside the vehicle, such as radio ranging (e.g., radar), laser scanners (e.g., LiDAR), vision (e.g., monocular and stereo cameras), and ToF cameras (also known as 3D cameras), but which has main limitations due to the nature of the sensors, such as a limited sensing range, a limited Field of View (FoV), and sight blockage, and (2) cooperative positioning, which uses other road participants to actively support the estimation of the relative position. The cooperative positioning techniques can be divided into two groups [109]: (1) transponder-based ranging systems, including Time of Arrival (ToA), angle of arrival, round trip delay, and time difference of arrival, and (2) Global Navigation Satellite System (GNSS)-based relative localization, which is based on exchanging GNSS-related information between vehicles.

As we presented, a considerable amount of literature has been devoted to vehicle detection and tracking, lane detection, target vehicle's distance estimation, speed estimation, and localization. Most of the existing studies focused on stationary sensors with the main purpose of enhancing ITMS performance or on mounted sensors on MVs with the purpose of boosting self-awareness and automated driving. However, there is much less empirical knowledge about linking these two topics of interest and using the MV's vision to estimate and provide the traffic data of the target vehicles dynamically based on the ITMSs' needs. In addition, to the best of our knowledge, there are no similar open-source systems available that can estimate all those mentioned traffic data of the target vehicles via the vision of a mobile MV.

3.3 Modern Vehicle Self-localization - RQ3

Several techniques have been proposed to self-localize MVs. In [20], the MV self-localization techniques are classified into five groups. (1) The first technique is based on GPS receivers. In this case, as briefly presented in Section 2.3, the navigation relies on a constellation of 24 satellites circling in orbit about 20,200 km above the Earth's surface. Satellites are positioned in orbit in such a way that four to ten satellites reach each part of the Earth at any specified instant. However, only four satellites are needed to make a location estimation. The device measures the distance to the reached satellites using the receiver's antenna and determines their location by finding the intersection point coordinates [20]. (2) The second technique is based on map-matching [20]. Map-matching aligns locations of vehicles with a previously known map. This technique aims to merge existing positioning systems (e.g., GPS) with newly created geographic information systems (IS) that have access to more precise mapping data [20]. (3) The statistical map-matching algorithm is another technique for localizing vehicles. This technique employs probability computation to determine a vehicle's most likely trajectory. To do this, the algorithm computes the likelihood of previous locations and then compares this with various paths in order to determine the most likely trajectory [20]. (4) Another technique is named cellular localization. This technique is centered on the existing infrastructure used for mobile communication. It estimates the location of a moving device based on signals received from base stations [20]. The position can be estimated by "ranging methods" based on the estimated distance between two points (i.e., base stations and mobile vehicles). To estimate the distance to the base stations, Received Signal Strength Indicators (RSSIs) or ToA can be used. (5) Lastly, the fingerprinting method is a technique that is based on analyzing Received Signal Strengths (RSSs) from multiple transmitters. This approach includes three phases: DB creation, identification, and position estimation [20].

However, making a trade-off between the localization cost and the localization accuracy is vital. In addition, running empirical investigations in urban areas, where tall buildings may affect the location estimation accuracy, is vital for increasing the generalizability and practicality of the proposed approach.

3.4 Multiple Sensor Fusion - RQ4

With the knowledge that sensor estimation may include errors, multiple sensor fusion aims to integrate data from multiple sensors in order to reduce the estimation uncertainty. Also, it addresses the limitations of individual sensors operating independently [144]. Furthermore, sensor fusion aids in developing a consistent model that can perceive the surroundings in a variety of environmental situations [144]. In comparison with using a single sensor, sensor fusion provides several advantages, such as enhanced robustness and reliability, higher resolution, better spatial and temporal coverage, increased confidence, and reduced ambiguity and uncertainty [32].

There are three main approaches to fusing data, namely, High-Level Fusion (HLF), Low-Level Fusion (LLF), and Mid-Level Fusion (MLF) [144][11]. In the HLF approach, each sensor runs its own detection or tracking algorithm before combining the results into a single global decision. Since HLF methods have lower complexity and need less computational load and communication resources than the LLF and MLF approaches, they are often used. However, HLF gives inadequate information as classifications with a lower confidence value are eliminated if, for instance, there are multiple overlapping obstacles. The LLF solution integrates data from each sensor at the lowest level (i.e., raw data). As a result, all data is considered, which can increase the precision of obstacle detection. However, with this approach to accurately fusing sensors, precise extrinsic calibration is needed. Moreover, it produces a huge volume of data, which can cause a memory or communication bandwidth problem. MLF, also known as feature-level fusion, is between LLF and HLF. It fuses multiple target features derived from the corresponding sensor data (raw measurements), such as the color information from images or the radar and LiDAR location features, and then performs identification and classification on the fused multi-sensor features [144].

The existing body of research on multiple sensor fusion mainly focused on three perspectives. (1) Once approach is to fuse multiple stationary sensors that are mounted along the road (e.g., [26]) with the main purpose of enhancing ITMS performance. (2) Another approach is to fuse multiple sensors that are mounted on the same MV (e.g., [67]) with the purpose of using the strengths of one type of sensor to mitigate the weaknesses of another type. For instance, a camera is able to perform lane detection or color perception, while radar and LiDAR are not able to do so. On the other hand, radar and LiDAR are able to estimate distance more accurately than a camera [144]. Also, some researchers fused data from the same sensor type on the same vehicle with the main purpose of enhancing the sensing coverage (e.g., [71]). (3) The third group of studies utilized vehicle sensors and communication (e.g., V2V or V2I communications) to share and fuse the collected information via several MVs (also known as inter-vehicle sensor fusion) (e.g., [81][19][120][16]). Low-cost sensors mounted on MVs may not always provide accurate data, and enhancing the estimation accuracy plays a vital role in ITMS safety and performance. The main purpose of this type of fusion is to enhance the perception area and estimation accuracy by sharing and integrating information. For instance, integrating the GPS data and the vision data by using the V2V communication can improve vehicle tracking whenever the GPS is unavailable or has a poor quality [19]. Also, integrating GNSS data and camera-based measurements of road-boundary locations via V2V communication can enhance positioning accuracy without requiring a stationary reference receiver [120].

Although there have been a few studies on the fusing of multiple low-cost sensors mounted on multiple mobile MVs (e.g., [81]), more empirical investigations are needed to explore practical strategies for fusing uncertain data provided by multiple mobile MVs with different views with the main purpose of mitigating the estimation error, boosting the accurate perception of the traffic scene, and widening the sensing range dynamically.

Chapter 4

Research Methodology

This chapter is composed of five sections. Section 4.1 presents the research overview. Section 4.2 lays out the research methodology adopted during the work. Section 4.3 describes the data generation approaches taken in this study. Section 4.4 gives an overview of the research activities. The research quality is presented in Section 4.5.

4.1 Research Overview

Given the related work in this research field, there is a need for empirical investigations based on proposing new methodologies and developing new algorithms by considering low-cost sensors (i.e., a monocular camera with a built-in GPS receiver) mounted on a mobile MV to estimate the required traffic data for ITMSs.

This thesis addressed four RQs, RQ1 to RQ4, and their Sub-Research Questions (SRQs).

RQ1: What is the state of the art in managing an intersection intelligently in the context of both pure traffic of MVs (i.e., AVs) and mixed traffic at four-way signalized and unsignalized intersections?

- SRQ1.1: What are the factors (goals) that intelligent intersection management studies focused on in terms of utilizing MVs?
- SRQ1.2: What are the proposed methodologies for addressing the potential problems and the considered traffic data types in the pre-defined research context?
- SRQ1.3: What are the challenges and potential research gaps?

RQ2: How can traffic data, such as the number, type, relative position, distance, speed, lane, and geolocation of multiple and mobile target vehicles be estimated via a single mobile MV equipped with a front-facing monocular camera with a built-in GPS receiver in mixed traffic?

- SRQ2.1: How can the target vehicle's number, type, and relative position be estimated?
- SRQ2.2: How can the distance to the target vehicle be estimated?
- SRQ2.3: How can the speed of the target vehicle be estimated?
- SRQ2.4: Besides enhancing the estimation accuracy of a target vehicle's type, how can the lane that the target vehicle is in be estimated?
- SRQ2.5: How can the geolocation of the target vehicle be estimated?

RQ3: How can the self-localization accuracy of an MV be enhanced via two mounted low-cost built-in GPS receivers?

RQ4: How can the accuracy of the estimated geolocation of the target vehicle be increased based on multiple sensor fusion techniques?

- SRQ4.1: How can it be determined that MVs are observing the same target vehicle by considering the estimation uncertainty caused by a low-cost monocular camera mounted on mobile MVs with different views?
- SRQ4.2: How can the estimated geolocations of the re-identified target vehicle be dynamically fused by considering the uncertainty in the geolocation estimation?

4.2 Research Methodology

To conduct this study, we have followed the Design Science Research (DSR) introduced by Hevner et al. [55]. The core principle of DSR is to understand an application domain and problem and to obtain knowledge by building the designed artifact and evaluating it [55].

Hevner et al. [55] described DSR as three tightly connected cycles of activities, namely, relevance, design, and rigor. Understanding these cycles provides important insights into how to perform DSR. DSR cycles [55] are presented in Figure 4.1.

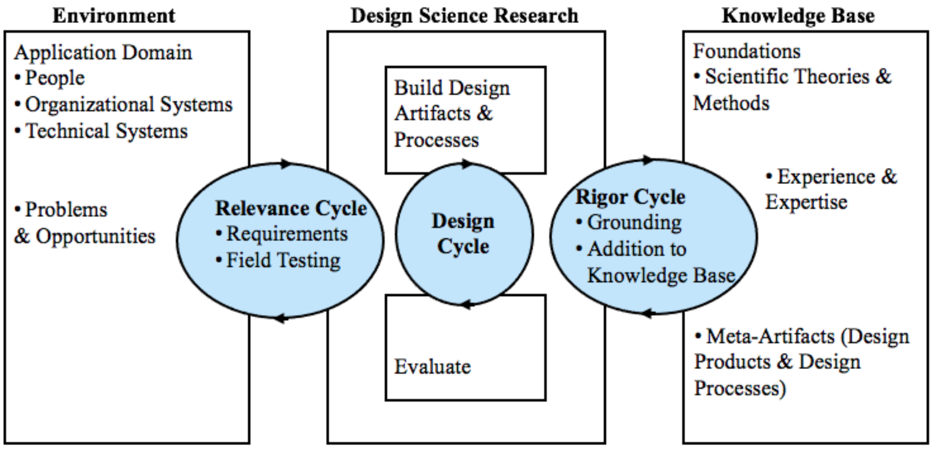


Figure 4.1: DSR cycles. Figure from [55].

- Relevance cycle: The relevance cycle initiates DSR with an application domain to provide the requirements for the research as inputs (e.g., the problems and opportunities to be addressed) and to define the acceptance criteria for the ultimate evaluation of the research results. The iteration allows the requirements to be improved and revised, as well as the outcomes to be studied and evaluated, by utilizing field testing to determine whether additional iterations of the relevance cycle are needed [55].

- Design cycle: The design cycle is the heart of the DSR project and aims to build design artifacts and processes and to evaluate them against the requirements until it achieves a suitable design. The iteration of this cycle helps to generate feedback to refine the design further [55].

- Rigor cycle: The rigor cycle is created by comprehensive knowledge based on scientific theories and methods, experience and expertise, and meta-artifacts. The rigor cycle aims to ensure that the research project is grounded in the relevant literature and that the innovations contribute to the state of the art in the application domain of the research and the existing artifacts [55].

The DSR meets the aim of this study for two main reasons. (1) It grounds the research project in the relevant literature and ensures its novelty. This is done in this study by answering RQ1 (Paper A). (2) Designing and developing new methodologies and algorithms for estimating the pre-defined traffic data by MV(s) based on the ITMSs' needs will answer RQ2 - RQ4 (Paper B - Paper F). It also iteratively evaluates and improves the developed new methodologies and algorithms based on real traffic data. This aspect is considered in RQ2 - RQ4 (Paper B - Paper F).

In this thesis, we have used two major types of research strategies: SLR (Paper A) and a combination of design and creation and case studies (Paper B - Paper F). These research strategies and their connections with our papers are described in the following section.

A Systematic Literature Review (SLR)

Reviewing the relevant published papers provides the foundation of an academic project. It helps researchers understand the field and obtain knowledge about the existing proposed solutions by other researchers in a specific research area, which helps position the work in the context of existing research by exploring the potential research gaps. An SLR is one type of review. The most common reasons for undertaking an SLR are (1) summarizing the existing evidence related to technology or treatment, (2) identifying the current research gaps, and (3) providing a framework/background to position new research activities appropriately [68]. In addition, an SLR explains the procedure and scope of the review regarding the relevant papers that are included. This feature makes the review reproducible [104].

To address RQ1 and as part of the relevance and rigor cycles of the DSR, we performed an SLR of the proposed methodologies regarding ITMSs (Paper A). We considered both pure traffic of MVs (i.e., AVs) and mixed traffic at four-way signalized and unsignalized intersections. The gained knowledge about the factors (goals), methodologies, and traffic data considered by other researchers led us to explore the potential research gaps in the study context, which were used to drive the following research steps and define RQ2 - RQ4. To conduct this SLR, we employed the guidelines proposed by Kitchenham [69][68]. This review used keyword-based searches and included different combinations of keywords and their synonyms, such as autonomous vehicle(s)/car(s), automated vehicle(s)/car(s), intelligent vehicle(s)/car(s), smart vehicle(s)/car(s), driverless vehicle(s)/car(s), unmanned vehicle(s)/car(s), cooperative vehicle(s)/car(s), connected vehicle(s)/car(s), smart intersection(s), intelligent intersection(s), autonomous intersection(s), automated intersection(s), and cooperative intersection(s). For our keyword-based searches, we used seven digital libraries, including Scopus, IEEE, Compendex, Inspec, Transport-Ovid, ACM, and Web of Science. We limited our search to papers published in English between January 2008 and May 10, 2019. This search process produced 2952 primary papers, and we selected 105 of them for review by applying the inclusion and exclusion criteria in six steps, as shown in Figure 4.2. To answer RQ1, data analyses were performed both quantitatively and qualitatively. In addition, a thematic synthesis was performed to classify and analyze the extracted qualitative data.

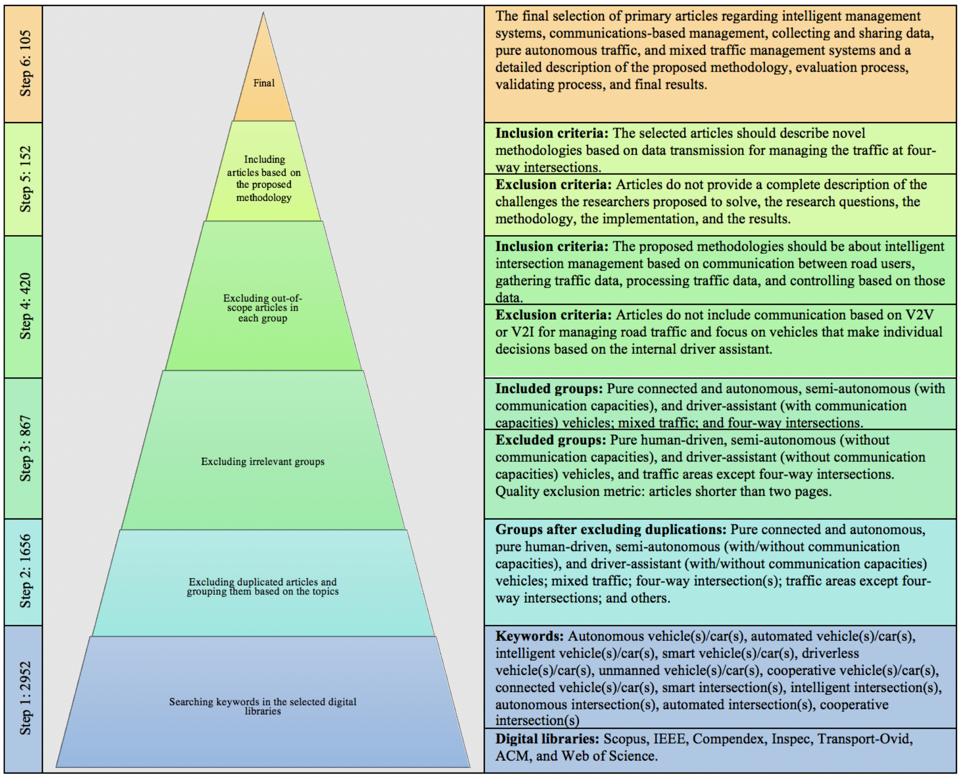


Figure 4.2: The process of selecting primary papers.

B Design and Creation

Design and creation require a problem-solving approach. It is useful in research that is mainly focused on developing new Information Technology IT products (artifacts) [102].

The central core of this research, addressing RQ2 - RQ4, concentrated on proposing new methodologies, developing new algorithms, and evaluating them as a design cycle of the DSR (Paper B - Paper F). This step aims to develop a new system for estimating the most important traffic data identified by the SLR (vehicles' number, type, relative position, distance, speed, lane, and geolocation) mainly based on object detection, lane detection, image processing, geometric computation, map-matching, and sensor fusion techniques. We conducted case studies to validate our proposed methodologies and to evaluate our developed algorithms in practice by considering real-world conditions. We used the data (footage and GPS data, which will be explained in detail in Section 4.3) collected from real traffic. It helped us to iteratively improve and revise our developed algorithms until we achieved our objectives, as the evaluation part of the design cycle related to the DSR.

4.3 Data Generation

In this study, data collection from real traffic is done in two rounds.

- In the first round, to answer four parts of RQ2 (SRQ2.1 - SRQ2.4), data were collected by one front-facing GoPro Hero 7 camera with a built-in GPS receiver on the front window glass of a vehicle that looked forward. The equipped vehicle (i.e., MV) followed a path through Trondheim, Norway, in the metropolitan area. The recording took place from 9 AM on a typical workday in May 2019. The video resolution was 1920×1080, and the frame rate was set to 30 FPS. The recorded data (video and GPS data) were split into manageable sequences. The goal of this round of data collection was to investigate the feasibility of a mobile MV equipped with one GoPro Hero 7 camera with a built-in GPS receiver to estimate the required traffic data based on our proposed methodologies and developed algorithms. These data were used in the evaluation and experimental steps related to Paper B and Paper C.

- In the second round, to answer one part of RQ2 (i.e., SRQ2.5), RQ3, and RQ4, we used three vehicles. Each vehicle was equipped with two GoPro Hero 7 cameras. One of the cameras was mounted on the front window glass and looked forward. Another camera was mounted on the rear window glass and looked backward. We defined several scenarios to drive vehicles in a straight street and an intersection with considering most of the possible real-world combinations of trajectories. The data collection (video and GPS) took place between 1 PM and 2 PM on a Sunday in December 2019 in Chengdu, China, in the metropolitan area. The video resolution and the frame rate were 1920×1440 and 60 FPS, respectively. We resized the video resolution and frame rate based on our needs. This round of data collection has four goals: (1) to investigate the feasibility of a mobile MV equipped with a GoPro Hero 7 camera with a built-in GPS receiver on each camera to estimate the required traffic data based on our proposed methodologies and developed algorithms; (2) to collect the required traffic data, as well as the ground-truth data, for assessing the accuracy of our proposed methodology; (3) to investigate the performance of multiple sensor fusion to deal with the sensor estimation uncertainty and to boost traffic awareness; (4) to collect data with structured and pre-defined scenarios based on our knowledge gained through the data collection in Round 1; and (5) to perform a second round of data collection from a different country, which enabled us to study the generalizability of our proposed methodologies. These data were used in the evaluation and experimental steps related to Paper D - Paper F.

4.4 Research Activities

During this study, three research strategies (i.e., SLR, design and creation, and case studies) and two rounds of data collection contributed to the cycles of the DSR methodology and addressed RQ1 - RQ4. A timeline of the activities is provided in Figure 4.3.

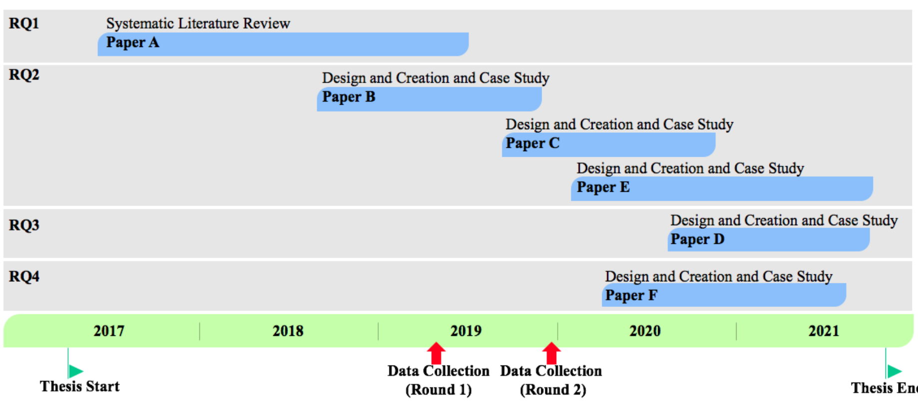


Figure 4.3: Timeline of the research activities.

4.5 Research Quality

Ensuring the research quality and evaluating the research results are critical points. Thus, considering the quality criteria and employing the correct measures are vital.

The philosophical paradigm of this study is the positivist paradigm, which underlies the scientific method [102]. The criteria for judging the quality of positivist research are classified into four headings: objectivity, reliability, internal validity, and external validity [102].

- Objectivity: Objectivity means “ensuring that there are no biases or distortions in the research.” Additionally, “ensuring that researchers have no influence on the results or a vested interest in a specific outcome” [102].
- Reliability: Reliability means “ensuring the neutrality, accuracy, and dependability of research instruments.” Additionally, “ensuring that repeated use of research instruments produces the same outcomes (i.e., repeatability)” [102].
- Internal validity: Internal validity means “ensuring that the research was well-designed and researchers examined the right things or collected the right data from the right sources.” Additionally, “ensuring that the research was coherent and accurate and the data generated to support the researchers’ claims and findings.” Also, “ensuring that the researchers are justified in saying that A causes B and a coincident link exists in reality” [102].
- External validity: External validity means “ensuring that research findings are generalizable to different people, settings, or time” [102].

We consider the related quality criteria to address RQ1 to RQ4 and discuss them in the discussion chapter.

Chapter 5

Results

This chapter summarizes the papers that contain the results of the conducted research. The chapter is composed of four sections and organized based on the RQs defined in this thesis and the corresponding SRQs that were considered in each published/submitted paper included in the thesis.

5.1 Literature Review of Intelligent Traffic Management Systems with Modern Vehicles - Results of RQ1

The nature of RQ1 required performing an SLR to identify the considered factors (goals), proposed methodologies, and considered traffic data types in terms of utilizing MVs (i.e., AVs) in ITMSs by focusing on both pure MV traffic and mixed traffic at four-way unsignalized and signalized intersections. In addition, addressing RQ1 helped us to explore the potential research gaps in the studied scope, which led us to formulate our RQs to conduct this thesis. Therefore, this SLR provided the grounding and identified the gaps in the literature, thus supporting the methodologies designed and developed in the subsequent papers included in this thesis.

As presented in Section 4.2.A, to address RQ1, we searched for papers published between January 2008 and May 10, 2019, and selected 105 primary studies to conduct an SLR. We applied the thematic analysis method to analyze the extracted data and answer the following SRQs.

5.1.1 Intelligent Transportation Systems' Factors (Goals) - Results of SRQ1.1

The results indicated that the considered factors (goals) of the studied papers can be classified into efficiency (e.g., delay, throughput, and congestion mitigation), safety (e.g., collision avoidance and resolving conflict), ecology (e.g., fuel/energy consumption and emission), passenger comfort, and others (e.g., data-sharing features). In addition, some of the papers considered a combination of several factors (goals), such as enhancing both efficiency and safety (e.g., [1]).

5.1.2 Proposed Intelligent Transportation Systems' Methodologies and Considered Traffic Data Types - Results of SRQ1.2

Various methodologies have been proposed in the studied papers. We applied the thematic analysis method to classify the proposed methodologies into four major groups: rule-based, optimization-based, hybrid (i.e., a combination of both rule-based and optimization-based methods), and machine learning methods. The results showed that 40% of the selected papers used rule-based methodologies, 44.76% of them applied optimization methodologies, 11.43% of them were based on hybrid methodologies, and 3.8% of the selected papers used machine learning techniques in the research context.

Moreover, it is important to note that the considered traffic data types in the studied papers are a key component of ITMSs. The most important data types considered in the selected papers include vehicle origin and destination, speed, arrival and existing time, acceleration/deceleration, direction and path, size, Identification (ID), and headway distance.

5.1.3 Identified Research Gaps - Results of SRQ1.3

The most critical potential research gaps that we explored in this SLR helped us to formulate our other RQs in this thesis.

The first potential research gap concerns the appropriate methodology for collecting data of traffic that is a mixture of both HDVs and MVs. In reviewing the papers in the SLR, we found that most of the studied papers focused on pure MV traffic. However, the future traffic would be a combination of both HDVs and MVs, as changing most of the vehicles to the modern version is a time-consuming process. Also, after this transition period, some people might still enjoy driving with traditional vehicles. In addition, traffic might still include pedestrians and cyclists without the ability to collect data and make a connection to transfer data. Therefore, further studies are needed to focus on the potential problems in managing mixed traffic. One important question in this regard is how to collect the required traffic data in mixed traffic. Based on the literature, if the traffic is pure MV traffic, then each MV is responsible for collecting data about itself and share them via V2V or V2I communications, collectively called V2X communication. However, this approach is not effective in mixed traffic as HDVs are not able to collect data and make connections. In this case, stationary sensors can be used, which have a high installation and maintenance cost in order to provide an acceptable coverage range. Therefore, one potential solution would be to employ mounted sensors on MVs as a mobile sensor and use them to collect data of the surrounding HDVs based on their perception features. This SLR confirmed that only a few studies paid attention to collecting the required traffic data in mixed traffic.

Another potential gap concerns the approach to estimating the required traffic data via mounted sensors on vehicles in mixed traffic. Although AI and machine learning techniques are mostly applied to estimate the required data of MVs with the purpose of controlling and driving the vehicles automatically, very little attention has been paid to AI and machine learning techniques for collecting the traffic data of the surrounding vehicles via mounted sensors on MVs by considering the ITMSs' needs with the purpose of improving an ITMS's awareness and performance in mixed traffic.

Furthermore, as another potential gap in dealing with mixed traffic, we need to consider vehicles with various sensing capabilities to enhance the generalizability and practicality of the proposed approach. Therefore, considering a sensor that can easily be mounted on most developed vehicles is needed (e.g., a monocular camera). However, most of the existing literature utilized sensors such as LiDAR on MVs to collect data, which would not be compatible with some vehicles in mixed traffic, thus limiting the generalizability.

Another key potential gap in the literature comes from the evaluation approaches of the proposed methodologies. Ensuring the performance of the proposed approaches is vital to validate them and assure their quality. However, this SLR confirmed that most of the experiments were evaluated and validated by simulators that look at a limited number of scenarios and that do not take into account unpredictable real traffic situations. Therefore, additional empirical studies are needed to validate and evaluate the proposed approaches on real traffic data to confirm the proposed approaches' practicality in the real world.

In addition to the aforementioned potential research gaps, we explored extra gaps, which are not considered in this thesis in order to keep the research scope manageable in view of the limited Ph.D. duration. For instance, to make the proposed methodology practical in the real world, considering all traffic objects, including pedestrians and cyclists, is needed. However, we found that most of the existing papers are mainly focused on the traffic of vehicles. Another important challenge comes from the need to standardize the collected traffic data. To have a compatible and effective ITMS, standardizing the required data types is needed to be able to share and analyse data. In addition, improving communication and data quality are other potential gaps. To provide an efficient V2X connection, possible problems related to data transfer and networking, such as communication delays and failures, privacy, package loss and duplication, and bandwidth limitations, should be considered. Last but not least, one of the most interesting future study directions could be utilizing the data collected via MVs about the surrounding vehicles to develop a digital twin of the traffic to model the traffic scene dynamically. This model can generate a global view of the traffic, which would be helpful in enhancing awareness and managing the traffic effectively.

More detailed results of RQ1 are presented in Paper A, which is included in Part II of this thesis.

5.2 Estimating the Target Vehicle's Traffic Data - Results of RQ2

The preliminary phase of identifying RQ2 originated from the research gaps identified by the SLR, which showed that the target vehicles' number, type, relative position, distance, speed, lane, and geolocation are among the most significant traffic data types that should be considered by ITMSs. Therefore, RQ2 is defined to explore the feasibility of using an MV equipped with a low-cost front-facing GoPro Hero 7 camera with a built-in GPS receiver as a mobile sensor to estimate the required traffic data of the surrounding target vehicles. The results could provide inputs to ITMSs from a variety of perspectives. For instance, this approach could reduce the cost caused by installing and maintaining stationary sensors on the road in order to provide the required sensing coverage range. Additionally, the MVs may help collect data from a wider area than stationary sensors due to their mobility features. In this regard, we proposed new methodologies and developed new algorithms. Following the system development, we conducted empirical experiments on the collected data from real traffic to evaluate our proposed methodologies and developed algorithms to answer RQ2.

5.2.1 Estimating the Target Vehicle's Number, Type, and Relative Position - Results of SRQ2.1

The proposed methodologies to answer this RQ included four main steps.

- Pre-processing: First, we pre-processed the data collected in Round 1 explained in Section 4.3 to decrease computational time. This step included converting frames into grayscale images and blurring them to mitigate noise.
- Lane detection: As the second step, we focused on lane detection to identify lines nearby the MV. To detect lanes, we employed Canny edge detection [28] as our empirical studies showed that it was able to generate enough edges without much noise in comparison with Sobel edge detection [50] and Prewitt edge detection [123]. Then we cropped the image in order to

remove unwanted areas (e.g., the sky) as the Region of Interest (RoI) was defined as a trapezoid at the bottom half of the image that includes a street area. Also, the Progressive Probabilistic Hough Transform (PPHT) [38][91] was used to find edges and draw continuous lines by merging them on the road, as we mainly focused on straight trajectories in this study.

- Vehicle detection: The third step focused on target vehicle detection. In this step, we employed YOLO-V3 [118][117] trained on a Common Objects in Context (COCO) dataset [79] as a foundation for our system. We selected YOLO-V3 as Wang et al. [139] listed YOLO-V3 as the second most popular object detector model. Also, it is open-source software and well documented, and it is easy to employ. In addition, the creators of the YOLO algorithm, Redmon et al., stated in their paper that it is fast and accurate [118][117]. Therefore, we decided to use YOLO-V3, which was the latest released version at the time of our study, to detect target vehicles and identify their types via the MVs' vision and to further develop it to adapt it to our objectives and answer our RQs. We limited our study to detecting vehicles (without considering other traffic objects, such as pedestrians and cyclists), as vehicle detection is an initial step in estimating the pre-defined required traffic data.

- Target vehicles' relative position: Image processing techniques were applied to estimate the relative position of the target vehicles. In this regard, we identified the centroid of the bounding box around the target vehicle generated by YOLO-V3. Then we estimated the position of the centroid based on the detected lines nearby the MV on the road. We classified the detected vehicles into three groups: middle vehicles (when the detected vehicle was located on the same lane as the MV), right vehicles (when the detected vehicle was located on the right side of the MV), and left vehicles (when the detected vehicle was located on the left side of the MV). Then, we counted the number of detected vehicles per group, which can be used to gain knowledge about traffic density.

We developed algorithms and ran experiments on three scenarios of the collected data in Round 1, as presented in Section 4.3, to evaluate the proposed approach to estimating the target vehicles' number, type, and relative position. To evaluate this proposed methodology, the estimated values should be compared with their true values. In this regard, we compared the algorithms' output with the manually counted results as ground truth. Our evaluation was done based on two measurements. Measure 1 considered the overall ability to detect and count the number of target vehicles per lane without respecting the estimated target vehicle's type.

Measure 2 considered the overall ability to detect and count the number of target vehicles per lane by respecting the estimated target vehicle's type.

Our findings are summarized in Table 5.1. As this table shows, the total error of Measure 1, counting target vehicles (including the three lanes named left, middle, and right) without considering the vehicle's type, was between 1.0% and 10.6%. The total error on average (i.e., counting too many or too few target vehicles on the left, middle, and right lanes on average) for Measure 1 was between 12.7% and 29.2%. The total error on average for Measure 2 was between 34.4% and 46.3%. These findings indicated that our proposed algorithms detected and counted vehicles more accurately without considering the target vehicle's type than with considering the target vehicle's type. Together, these results provided important insights into the effect of identifying the detected vehicle's type on the accuracy of our estimations.

Table 5.1: Our findings related to SRQ2.1.

S#	Measure 1		Measure 2
	Total error (%)	**Total error on average (%)**	**Total error on average (%)**
S1	1.0	12.7	34.4
S2	10.6	18.9	46.3
S3	6.1	29.9	34.7

5.2.2 Estimating the Target Vehicle's Distance from the Modern Vehicle - Results of SRQ2.2

The general idea for estimating the distance to the target vehicle was to employ the pinhole camera geometry model (also called the pinhole camera model) [17] and the known pre-defined and standard size of a target vehicle based on its type. Our main contribution was to consider both target vehicle width and height in estimating the distance to the target vehicle. In detail, in order to estimate this distance, we developed a new algorithm on top of the algorithms presented in SRQ2.1, which estimated the vehicle's types (i.e., bus, car, motorbike, truck, and van), and together with the height and width of the bounding box in pixels around the target vehicle. Then we used the identified type of the target vehicle to find the real size of it based on the pre-defined and standard values (e.g., bus width = 2.4 m, bus height = 4.0 m). In the follow-up step, we estimated the distance by considering both the width and the height, as presented in Eq. 5.1 and Eq. 5.2 [17]. In Eq. 5.1 and Eq. 5.2, F_c is the camera focal length and h_v and w_v are the target vehicle's height and width on the image, respectively. H_v and W_v are the target vehicle's height and width in reality, respectively. Finally, as presented in Eq. 5.3, the target vehicle's distance d_v was calculated based on the average of d_h and d_w by using a weight factor to consider the ratio between the height and the width in the distance estimation. The γ factor was used to control which of the height or width values should be prioritized. The appropriate weight factor was obtained with our extra empirical tests. The experiment for this part required its own videos, as the necessary ground-truth data were not available in the collected data presented in Section 4.3; hence these recordings were captured separately from the previously mentioned data collection in Section 4.3. Data for these experiments were collected by recording video clips of stationary vehicles of different lengths and measuring the distance with a basic laser measuring tool to generate ground-truth data. The algorithm then ran multiple video files with different height and width ratios (e.g., 100% of height and 0% of width, 85% of height and 15% of width, 75% of height and 25% of width, 60% of height and 40% of width, 50% of height and 50% of width, 40% of height and 60% of width, 25% of height and 75% of width, 15% of height and 85% of width, and 0% of height and 100% of width). The output was compared with the laser's truth to determine which ratio was more accurate.

$$d_h = F_c \cdot \frac{H_v}{h_v} \tag{5.1}$$

$$d_w = F_c \cdot \frac{W_v}{w_v} \tag{5.2}$$

$$d_v = \frac{(1+\gamma) \cdot d_h + (1-\gamma) \cdot d_w}{2} \tag{5.3}$$

To estimate the target vehicle's distance, we developed algorithms and ran experiments. In this regard, the estimated distance should be compared with its true distance. As we stated already, in Round 1 of the data collection explained in Section 4.3, we did not include the ground-truth data related to the distance as we did not have the required sensor to obtain the true distance to the target vehicle. Therefore, we captured more recordings than previously mentioned in the data collection process. Data for these steps were collected by recording video clips of stationary vehicles with different lengths and at different angles. The ground-truth data were acquired with a simple laser measuring tool. Our experimental results showed that the best ratio for combining the estimated distances is 85% of the height and 15% of the width, which yielded the most stable estimation of the target vehicle's distance.

5.2.3 Estimating the Target Vehicle's Speed - Results of SRQ2.3

The intuition behind the speed estimation of the target vehicle is presented in Eq. 5.4 [127].

$$v = \frac{\Delta d}{\Delta t} \tag{5.4}$$

- Tracking the target vehicle and assigning it an ID: First, we need to track the target vehicle between frames. To do so, we used the centroid of the bounding box generated around the target vehicle by YOLO-V3 and estimated its positional difference between frames based on the Euclidean distance [33]. On the next frame, the calculated centroid was compared with the previous ones, and the closest centroid according to the Euclidean distance [33] was considered to be the same target vehicle.
- Estimating the target vehicle's speed: After we estimated the distance of the tracked target vehicle, for every frame, we calculated its change in distance and stored it. To remove spikes or other sudden changes in distance estimation, for every frame, we calculated the average change in distance for the last 30 frames. Finally, to estimate speed v of the tracked target vehicle, we used the estimated distance traveled Δd over time Δt (we knew that 30 frames equal 1 sec, as the video recording was done with 30 FPS).

To evaluate the proposed methodology for estimating the target vehicle's speed, the estimated speed should be compared with the target vehicle's true speed. However, in Round 1 of the data collection explained in Section 4.3, we did not include the ground-truth data as we did not have the required sensor to obtain the true speed of the target vehicle. Therefore, we assumed that the estimated relative distance in SRQ2.2 was correct. Therefore, we calculated the "true" speed manually. We selected video sequences in which the target vehicle was seen for an extended period of time (i.e., the target vehicle was mostly driving in front of the MV). Then for the first frame of that video sequence, we recorded the values of the MV's speed, the estimated distance to the target vehicle, and the estimated speed of the target vehicle. Then we proceeded to manually step forward an arbitrary number of frames, mostly in the range of 30 - 60 frames, and noted the values of the same parameters. In addition, the number of skipped frames was noted as well. This process continued until a sufficient amount of data were obtained. Then we manually calculated the true speed based on the noted data. The true speed was computed by calculating the time since the last measurement, determining the traveled distance of the target vehicle, and then dividing this distance by the time, which yielded the speed. The distance traveled by the target vehicle was estimated by taking the distance the MV traveled, which was estimated by multiplying the speed with time, adding this to the distance to the target vehicle, and then subtracting the previously recorded distance to the target vehicle; we regarded these manually calculated speeds as ground truth. Our experiments on the 75 manual readings showed that the average difference between our proposed algorithm and the ground truth was between 2.09 m/s and 10.64 m/s.

More detailed results regarding SRQ2.1 - SRQ2.3 are presented in Paper B, included in Part II of this thesis.

5.2.4 Improving the Target Vehicle's Type Estimation and Estimating the Target Vehicle's Lane - Results of SRQ2.4

Based on SRQ1.2, we found that estimating the target vehicle's lane plays an important role in an ITMS, as it is vital for estimating the traffic density and helps to provide an overall view of the traffic scene. In addition, from SRQ2.1, we found that the developed algorithms for detecting the target vehicle's type were not accurate, which might negatively affect the accuracy of estimating the target vehicle's distance and speed. Therefore, SRQ2.4 aimed to deal with the aforementioned issues.

First, in order to improve the accuracy of the vehicle's type estimation, we trained YOLO-V3 by using pre-trained weights on the Karlsruhe Institute of Technology and Toyota Technological Institute (KITTI) dataset [116] instead of the COCO dataset. The COCO dataset includes 80 object categories, such as car, cat, umbrella, and cell phone, which are not mainly focused on traffic objects. However, the KITTI dataset focuses on traffic objects and contains eight categories, namely, car, van, truck, pedestrian, person_sitting, cyclist, tram, and misc (e.g., trailers, segways) [116], which has the potential to boost the target vehicle's type estimation.

Second, in order to estimate the target vehicle's lane in relation to the MV, we developed new algorithms by proposing the two approaches shown in Figure 5.1. As Figure 5.1 shows, it contained three major steps.

- Step 1: Pre-processing and lane detection. This step was developed already to answer SRQ2.1 - SRQ2.3, which included data pre-processing (including the conversion of frames into grayscale images and the removal of noise by blurring) and lane detection (including canny edge detection [28], cropping the RoI, the PPHT [38][91], and merging and drawing lines).

- Step 2: Object detection and estimating the target vehicle's relative position.

This step was already developed to answer SRQ2.1, which included vehicle detection by YOLO-V3 and generating a bounding box around the target vehicle. The only change is that we defined a point on the bottom edge of the bounding box and called it the Central Point (CP). As it is located on the street surface, it may provide more accurate relative positional estimation than the central point of the bounding box. The CP was used to determine the position of the target vehicle in relation to the detected nearby lanes on both sides of the MV. In this step, the target vehicle's relative position was classified into the left, middle, and right groups.

- Step 3: Estimating the target vehicle's lane. This step estimated the exact lane the target vehicle was in by using our proposed Approach 1 and Approach 2 below, by assuming that the vehicle width is less than the lane width.

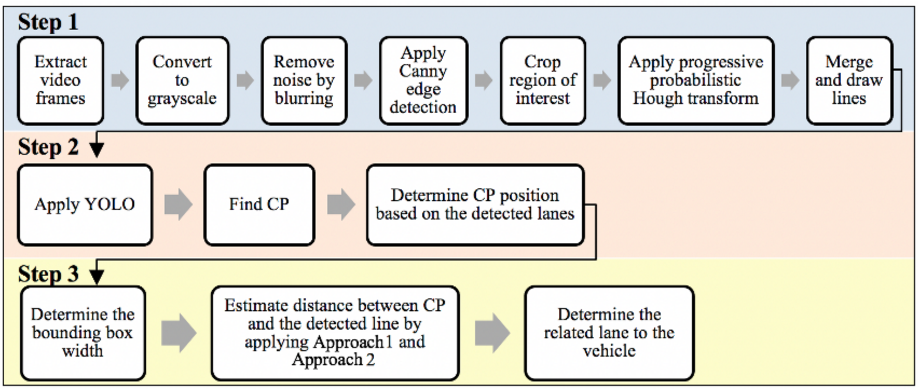


Figure 5.1: Our proposed steps to determine the target vehicle's lane.

- Approach 1: Our first proposed approach to estimating the target vehicle's lane is presented in Figure 5.2. We explain Approach 1 by showing two target vehicles in Figure 5.2. W_{v1} and W_{v2} are the bounding box's width around the target vehicles. LL and RL represent the detected left line and right line nearby the MV, respectively. This approach was based on the shortest distance (D_i, $i := 1, 2$) between central point $CP = (x_{vi,0}, y_{vi,0})$ on the bottom edge of the bounding box around the target vehicle (vi, $i := 1, 2$) and a nearby detected line (which passes through two points $P_1 = (x_1, y_1)$ and $P_2 = (x_2, y_2)$ in Figure 5.2) by using Eq. 5.5 [108].

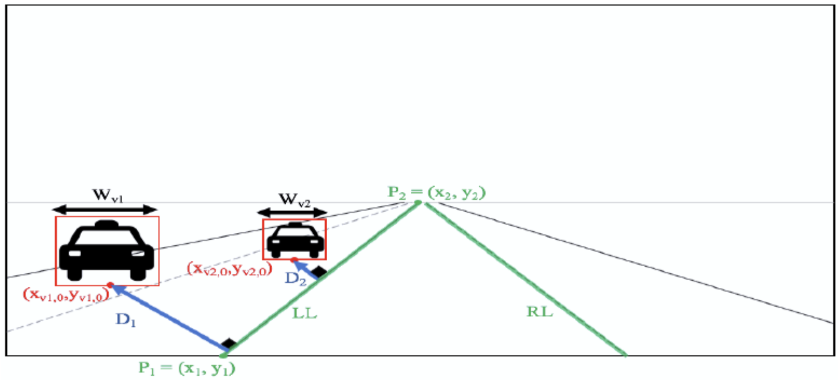


Figure 5.2: The first proposed approach to determining the target vehicle's lane.

$$D_i = \frac{|(x_2 - x_1)(y_1 - y_{vi,0}) - (x_1 - x_{vi,0})(y_2 - y_1)|}{\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}} \tag{5.5}$$

- Approach 2: Our second proposed approach to estimating the target vehicle's lane is presented in Figure 5.3. This approach was based on the horizontal distance d_i between CP and the nearby detected line (which passes through two points P_1 and P_2 in Figure 5.3), by using Eq. 5.6 [135]. In this formula, D_i is the shortest distance between CP and the nearby lane, and θ is the angle between d_i and D_i, which was estimated based on the slope of the nearby lane that passed through P_1 and P_2 [125][130][106], as $\beta = \gamma = \arctan\left(\frac{y_2 - y_1}{x_2 - x_1}\right)$ and $\theta = 90 - \gamma$.

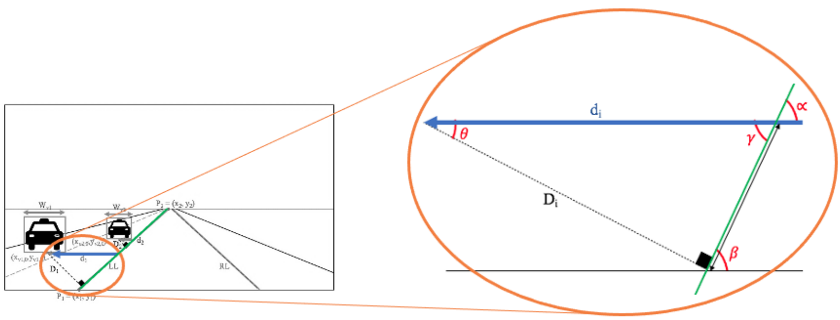


Figure 5.3: The second proposed approach to determining the target vehicle's lane.

$$d_i = \frac{D_i}{cos(\theta)} \tag{5.6}$$

Finally, to estimate the target vehicle's lane, we compared the estimated distance obtained by both proposed approaches with the estimated target vehicle's width based on a bounding box generated by YOLO-V3 (W_{vi}, $i := 1, 2$). We assumed that the vehicle width is less than the lane width; therefore, we used this measure to estimate the lane (e.g., if $0 < Distance_i < W_{vi}$, where $Distance_i$ is the estimated distance by Approach 1 or Approach 2 [i.e., d_i or D_i], then the vehicle is in the first lane on the left or right of the MV depending on the related nearby line). The considered situations to estimate the target vehicle's lane are presented in Table 5.2.

Table 5.2: Conditions for finding the target vehicle's lane in a street with multiple lanes.

Condition	Output
$0 < Distance_i < W_{vi}$	1^{st} lane on the left/right
$W_{vi} < Distance_i < 2 \cdot W_{vi}$	2^{nd} lane on the left/right
$2 \cdot W_{vi} < Distance_i < 3 \cdot W_{vi}$	3^{rd} lane on the left/right
$(n-1) \cdot W_{vi} < Distance_i < n \cdot W_{vi}$	n^{th} lane on the left/right

To evaluate the proposed methodology for enhancing the target vehicle's type estimation and estimating the target vehicle's lane, we developed algorithms and ran experiments by using real traffic data collected in Round 1 presented in Section 4.3. In this regard, the estimated type and lane of the target vehicle should be compared with the ground truth. The required ground-truth data were collected manually by watching videos and documenting the observed target vehicle's type and lane data. First, to enhance the accuracy of the vehicle's type estimation, we trained YOLO-V3 on pre-trained weights with the KITTI dataset [116]. Consistent with the literature, this research found that the accuracy of estimating the target vehicle's type in the studied scenarios was higher than 90.74% for all lanes. For estimating the target vehicle's relative lane in multiple-lane streets, experiments on the studied scenarios showed that the accuracy of the target vehicle's lane estimation was between 71.43% and 90.54% with the first approach and between 71.43% and 94.59% with the second approach, for all lanes.

In summary, these results indicated that an MV equipped with a front-facing GoPro Hero 7 camera was able to effectively identify the target vehicle's lane in a multiple-lane street. By doing extra analyses to explore the reason for the decrease in accuracy to 71.43% with both approaches, we found that the lane marks partially faded in that specific scenario. Thus, as we expected, we can conclude that the performance of these approaches was highly dependent on the lane detection accuracy.

More detailed results regarding SRQ2.4 are presented in Paper C, included in Part II of this thesis.

5.2.5 Estimating the Target Vehicle's Geolocation - Results of SRQ2.5

Based on our findings obtained from SRQ1.2 and keeping RQ4 in our mind, we identified the importance of estimating the target vehicle's geolocations in managing the traffic intelligently, as these data can be used to generate the traffic model dynamically in the GPS coordinate system. SRQ2.5 focuses specifically on image-based target vehicle's geolocation estimation, which depends on MV's self-localization addressed in RQ3.

In this paper, we proposed two new approaches by integrating deep learning, image processing, and geometric computation to use the vision sensing and self-localization capabilities of a mobile MV to estimate the geolocations of a target vehicle. Figure 5.4 illustrates our proposed research strategy.

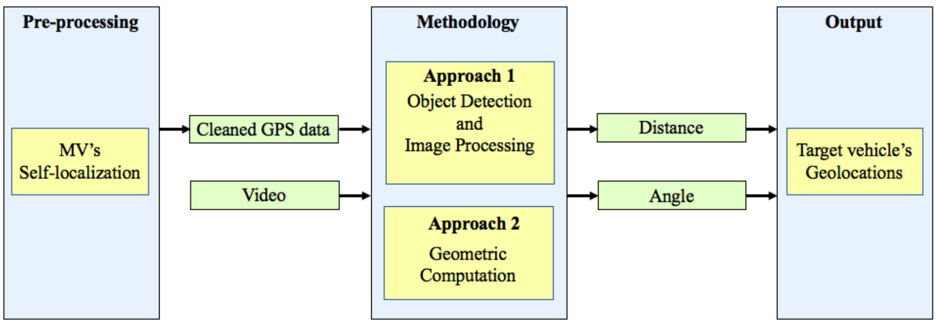


Figure 5.4: The components of our proposed research strategy to address SRQ2.5.

As Figure 5.4 shows, the proposed research strategy includes the following three main steps:

- Pre-processing: For estimating the target vehicle's geolocations, the geolocations of the MV are required. In this study, the geolocations of the MV were collected by a built-in GPS receiver of a GoPro Hero 7 camera mounted on the MV, which might be noisy. To enhance the accuracy of the MV's geolocations, we applied the algorithms that were proposed to address RQ3 (based on cross-GPS validation, interpolation, best-fit, and map-matching techniques).

- Methodology: As shown in Figure 5.5, to estimate the target vehicle's geolocations in the GPS coordinate system, in addition to the MV's geolocations, the distance d between the MV v2 and the target vehicle v1 and the clockwise angle α between the north (N) and d are required [95].

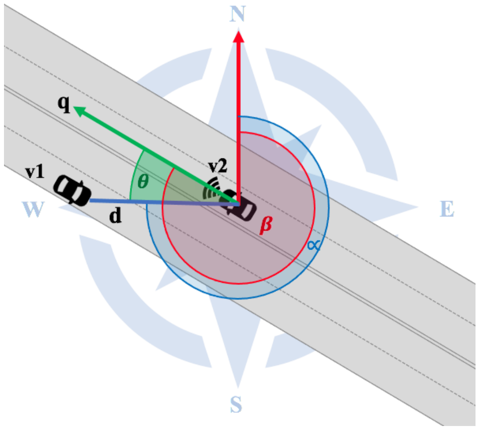


Figure 5.5: The required parameters for estimating the target vehicle's geolocation.

In this regard, as Figure 5.4 shows, we proposed two approaches: (1) Approach 1, which was based on object detection and image processing, and (2) Approach 2, which was based on geometric computations.

- Approach 1: Object detection and image processing

  - Estimating the distance d

  This step was done based on the proposed methodology and developed algorithms to address SRQ2.2, by using the pinhole camera model [17] and by considering both target vehicle width and height. The only change was training YOLO-V3 on the KITTI dataset [116].

  - Estimating the angle α

  As presented in Figure 5.5, in order to estimate the clockwise angle α between the north N and d, we need to estimate angle β, which is the angle between the north N and the MV v2's movement direction q, as well as the angle θ, which is the angle between d and q.

  - Estimating the angle β

  To estimate the angle β, we needed to identify the movement direction of the MV v2 based on its collected GPS coordinates in sequential frames as a start point (ϕ_1, λ_1) and an end-point (ϕ_2, λ_2) for all frames. We used Eq. 5.7 - Eq. 5.9 [95] to estimate the angle β along the whole trajectory dynamically.

$$M = \sin(\lambda_2 - \lambda_1) \cdot \cos\phi_2 \tag{5.7}$$

$$N = \cos\phi_1 \cdot \sin\phi_2 - \sin\phi_1 \cdot \cos\phi_2 \cdot \cos(\lambda_2 - \lambda_1) \tag{5.8}$$

$$\beta = \operatorname{atan2}(M, N) \tag{5.9}$$

- Estimating the angle θ

The idea of estimating the angle θ in Approach 1 is presented in Figure 5.6. In Figure 5.6, the blue bounding box shows the target vehicle v1. P is the central point on the button edge of the bounding box around the target vehicle v1, and H is the central point of the image. The angle θ is estimated based on the horizontal angle per pixel (γ) in degrees and on the number of horizontal pixels between P and the vertical line passing through H, shown by a red line T. γ is estimated based on the camera's horizontal FoV and the video's resolution. In this study, the video's resolution was adjusted to 960 $\times$ 720 pixels, and the camera's horizontal FoV was 86.7 degrees [54]. Therefore, γ was equal to 0.09 degrees. The angle θ in degrees was estimated as follows:

$$\theta = T \cdot \gamma \tag{5.10}$$

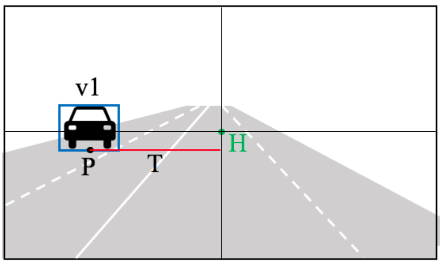


Figure 5.6: The parameters used to estimate the angle θ between the MV and the target vehicle used in Approach 1.

- Estimating the angle α

To estimate the angle α, we considered three different conditions, as follows:

- If the target vehicle drives in the same lane as the MV, then $\alpha = \beta$ and $\theta = 0$.
- If the target vehicle drives on the left side of the MV, then, as Figure 5.7 (a) shows, $\alpha = \beta - \theta$.
- If the target vehicle drives on the right side of the MV, then, as Figure 5.7 (b) shows, $\alpha = \beta + \theta$.

Approach 2: Geometric computation

The main idea of this approach is to transform 2D pixel coordinates of point P into 3D world coordinates. By assessing the 3D world coordinates of point

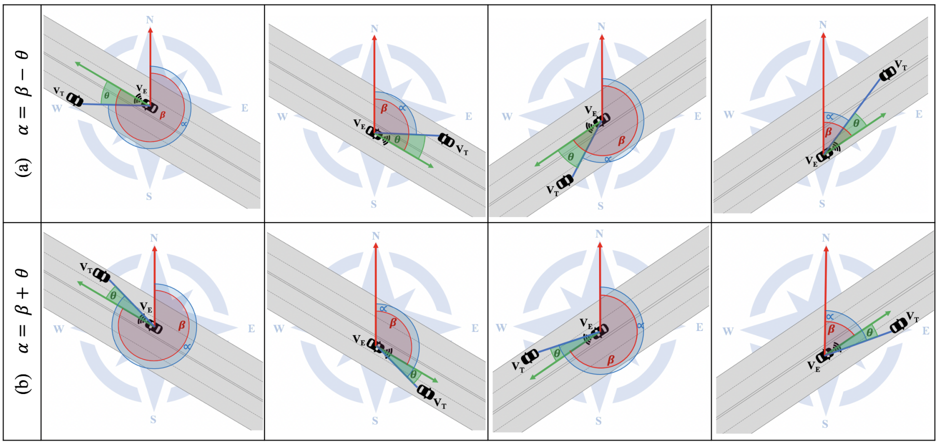


Figure 5.7: The mathematical relations between α, β, and θ.

P, we would be able to estimate the distance d and the angle α, which are needed to estimate the target vehicle's geolocation.

In this regard, we utilized a pinhole camera model as shown in Figure 5.8. In this figure, C is the perspective center of the camera and the origin of the Camera Coordinate Frame (CCF). Three unit vectors of the CCF are represented by y_1, y_2, and y_3. The Image Coordinate Frame (ICF) is centered at principal point H with unit vectors r_1 and r_2. The principal axis passes through C and H and is perpendicular to the image plane. The distance from C to the image plane is f, which is the camera's focal length. The image plane carries a 2D Pixel Coordinate Frame (PCF) with unit vectors z_1 and z_2. The image plane is subdivided into n_h pixels horizontally and n_v pixels vertically. To project the detected vehicle on the image onto the real world, we need to transform point P with pixel coordinates (p_1, p_2) (which is the central point on the bottom edge of the bounding box around the target vehicle) into a 3D world coordinate representation (w_1, w_2, w_3).

In this regard, we need to identify the 3D coordinates of point P in 3D world coordinates, as presented in Eq. 5.11. We assumed that the world coordinates were located on the road surface. In Eq. 5.11, h is the height of the camera mounted on the MV from the road surface. ξ defines any point that lies on the viewing ray from camera center C through point P in world coordinates. The camera's pitch angle is applied with a rotation matrix by the angle σ.

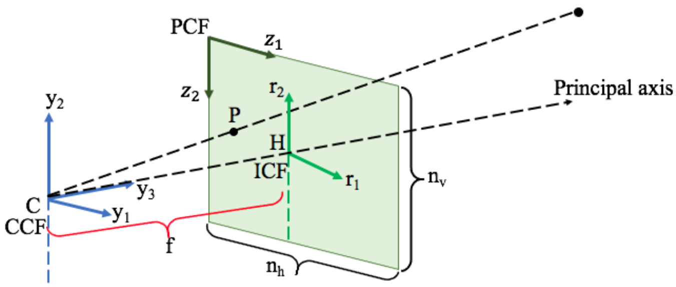


Figure 5.8: The pinhole camera model used in Approach 2 to address SRQ2.5.

$$\begin{pmatrix} w_1 \\ 0 \\ w_3 \end{pmatrix} = \begin{pmatrix} 0 \\ h \\ 0 \end{pmatrix} + \xi \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(-\sigma) & -\sin(-\sigma) \\ 0 & \sin(-\sigma) & \cos(-\sigma) \end{pmatrix} \cdot \begin{pmatrix} p_1 - h_1 \\ -(p_2 - h_2) \\ f \end{pmatrix} \tag{5.11}$$

Therefore, ξ, w_1, and w_3 are calculated as follows:

$$\xi = \frac{-h}{(\cos(-\sigma) \cdot (-p_2 + h_2) - \sin(-\sigma) \cdot f)} \tag{5.12}$$

$$w_1 = \xi \cdot (p_1 - h_1) \tag{5.13}$$

$$w_3 = \xi \cdot (\sin(-\sigma) \cdot (-p_2 + h_2) + \cos(-\sigma) \cdot f) \tag{5.14}$$

To calculate ξ, w_1, and w_3, we need to estimate the camera's focal length f, the camera's pitch angle σ, and the camera's height from the road surface h.

- Estimating the camera's focal length f

In Approach 2, the camera's focal length f in pixels was calculated based on the trigonometric relation presented in Eq. 5.15. We used the horizontal number of pixels n_h from the video's resolution and the camera's horizontal FoV (ρ) in degrees [54].

$$f = \frac{n_h}{2 \cdot \tan(\frac{\rho}{2})} \tag{5.15}$$

- Estimating the camera's pitch angle σ

To estimate the camera's pitch angle σ, we used the camera's focal length f and the vertical differences between the principal point $H = (h_1, h_2)$ and the vanishing point $J = (j_1, j_2)$ based on the lane detection, as shown in Figure 5.9. In this figure, the blue lines represent the detected parallel lines on the road nearby the MV in a perspective view. Based on this figure, the camera's pitch angle σ can be calculated by Eq. 5.16.

$$\sigma = \operatorname{atan2}(j_2 - h_2, f) \tag{5.16}$$

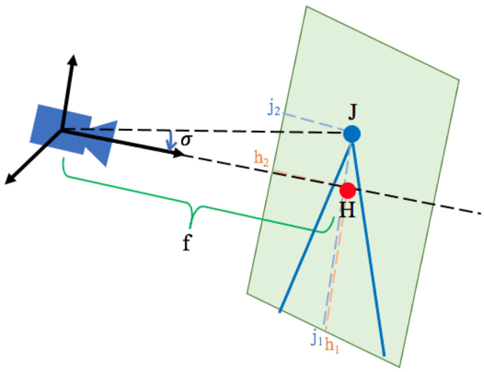


Figure 5.9: The camera's pitch angle σ and vanishing point used in Approach 2.

- Estimating the camera's height h from the road surface

To estimate the height h of the camera mounted on the MV from the road surface by considering the camera's pitch angle σ, we applied Thales's theorem [132]. Thales's theorem in this context is presented in Figure 5.10. The parameters in this figure are defined as follows:

$$A = \frac{f}{\cos\sigma} \tag{5.17}$$

$$B = h \cdot \tan\sigma \tag{5.18}$$

$$E = \frac{h}{\cos\sigma} \tag{5.19}$$

$$G = f \cdot \tan\sigma \tag{5.20}$$

$$N = p_2 - h_2 \tag{5.21}$$

$$K = G + N \tag{5.22}$$

To estimate the camera's height h based on Thales's theorem and the estimated distance d by Approach 1, we have Eq. 5.23.

$$\frac{K}{E} = \frac{A}{d + B} \tag{5.23}$$

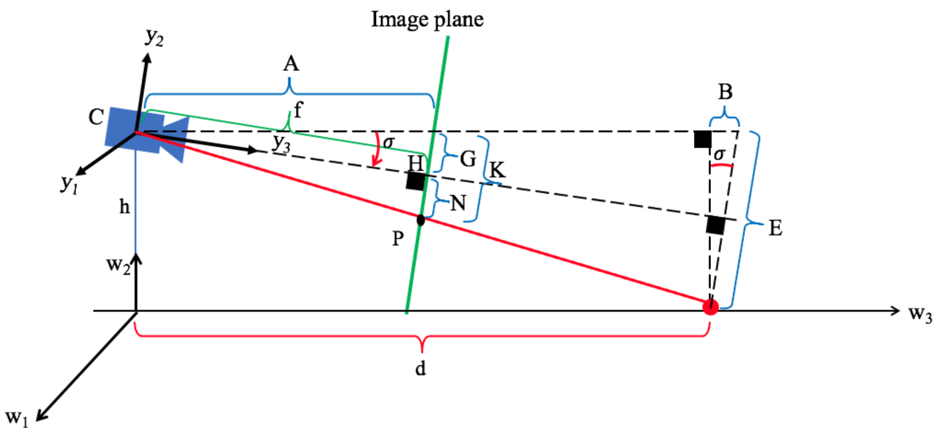


Figure 5.10: The camera's height from the road surface by considering the camera's pitch angle σ used in Approach 2.

By simplifying Eq. 5.23, h in meters was calculated as follows:

$$h = \frac{(N + f \cdot \tan\sigma) \cdot d \cdot (\cos\sigma)^2}{f - (N + f \cdot \tan\sigma) \cdot \tan\sigma \cdot (\cos\sigma)^2} \tag{5.24}$$

Finally, by estimating the camera's height from the road surface h, the camera's pitch angle σ, and the camera's focal length f, we can calculate ξ, w_1, and w_3. Because w_1 and w_3 represent point P in 3D world coordinates, where w_2 = 0, we can estimate the distance d by Approach 2 based on the Euclidean distance [33] between w_1 and w_3, as presented in Eq. 5.25, and estimate the angle θ based on the trigonometry [135] presented in Eq. 5.26. Finally, the angle α can be estimated based on the proposed conditions in Approach 1, as shown in Figure 5.7.

$$d = \sqrt{w_1^2 + w_3^2} \tag{5.25}$$

$$\theta = \operatorname{atan2}(w_3, w_1) \tag{5.26}$$

- Output: To estimate the target vehicle's geolocation with both approaches, we used Eq. 5.27 - Eq. 5.30 [95]. In these formulas, the parameters are as below:

  (ℓ_1, g_1) represent the geolocation of the MV

  (ℓ_2, g_2) represent the geolocation of the target vehicle

R represents the Earth's radius

d represents the estimated distance between the MV and the target vehicle by both approaches

α represents the estimated angle between the north N and d by both approaches

$$\ell_2 = (\sin(\ell_1) \cdot \cos(d/R) + \cos(\ell_1) \cdot \sin(d/R) \cdot \cos(\alpha)) \tag{5.27}$$

$$g_2 = g_1 + \text{atan2}(U, V) \tag{5.28}$$

where,

$$U = \sin(\alpha) \cdot \sin(d/R) \cdot \cos(\ell_1) \tag{5.29}$$

$$V = \cos(d/R) - \sin(\ell_1) \cdot \sin(\ell_2) \tag{5.30}$$

To evaluate the proposed methodology for estimating the target vehicle's geolocation, we developed algorithms and ran experiments by using real traffic data collected in Round 2 presented in Section 4.3. We adjusted the video's resolution to 960 $\times$ 720 and the frame rate to 1 FPS to apply pre-processing and vehicle detection. To run experiments, we focused on two scenarios, called Scenario S1 and Scenario S2, as presented in Figure 5.11. In Scenario S1, MV v3 and target vehicles v1 and v2 drive in the same direction on a straight trajectory. The purpose of this scenario was to evaluate our proposed approaches with one of the target vehicles driving on the same lane as the MV and the other target vehicle driving on the next lane. In Scenario S2, MV v3 and target vehicles v1 and v2 drive in opposite directions on a straight road.

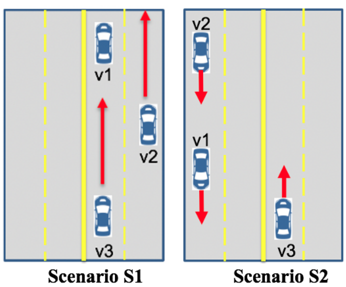


Figure 5.11: The studied scenarios to address RQ2.5 by using both approaches.

To evaluate our proposed approaches, the estimated target vehicle's geolocations should be compared with the ground truth. The required ground-truth data were collected by a built-in GPS receiver of a GoPro Hero 7 camera mounted on the target vehicle. As we stated already, the GPS receiver might be noisy. Thus, to enhance the accuracy of the ground-truth data, we applied our proposed methodology to address RQ3.

We performed the evaluation in two steps: (1) plotting the estimated geolocations (i.e., latitude and longitude) of the target vehicles on Google Maps and (2) analyzing the distance vector between the estimated target vehicle's geolocations and the ground-truth data for both approaches.

As an example, Figure 5.12 represents the outputs of the experiments related to Scenario S1. In this figure, the white polylines in (a) and (d) show the ground-truth trajectory of the observed target vehicle. The red polylines in (b) and (e) show the trajectory of the target vehicle estimated with Approach 1. The blue polylines in (c) and (f) show the trajectory of the target vehicle estimated with Approach 2. This figure shows that the trajectories of the target vehicle estimated with both approaches are plotted on the correct lane of the road and that they almost overlap with the ground-truth trajectory. This means that both approaches enabled us to estimate the trajectory of the target vehicle accurately on the right lane of the road.

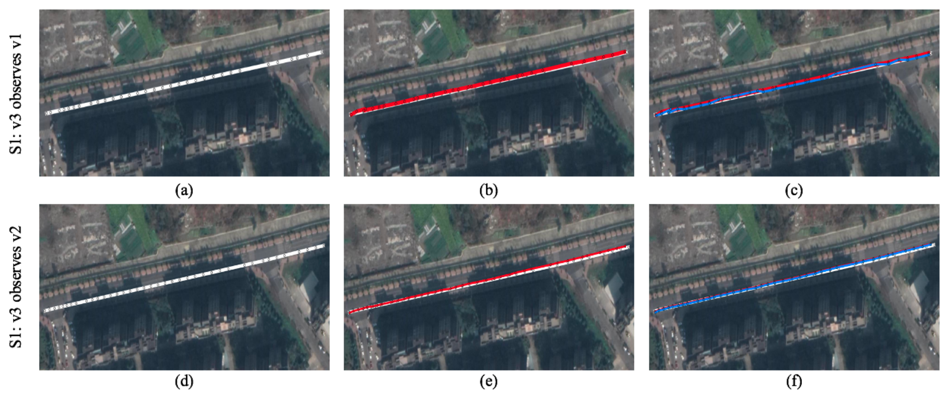


Figure 5.12: The estimated trajectory with Scenario S1 on the map. Two cases are considered: (1) MV v3 observes target vehicle v1 (a-c) and (2) MV v3 observes target vehicle v2 (d-f).

To analyze the estimated geolocations of the target vehicle numerically, we used the distance vector between the ground truth and the geolocations estimated by both approaches per timestamp. Figure 5.13 visualizes our findings related to Scenario S1, as an example. Our numerical findings related to both Scenario S1 and Scenario S2 are summarized in Table 5.3. As Figure 5.13 and Table 5.3 show, the deviation of the geolocation estimation (based on the absolute values) with Approach 1 is on average between 1.38 m and 3.54 m. With Approach 2, this deviation is on average between 1.4 m and 3.51 m. Figure 5.13 (a), (b) shows a slightly upward trend between the plotted points. This result may be explained by the fact that the collected and pre-processed data (see RQ3) by a GPS receiver to provide the MV's location and the ground-truth data of the target vehicle's position were not noise free. In addition, as Table 5.3 shows, the highest on average geolocation estimation deviation with both approaches is obtained in Scenario S2, when v3 observes v1. A possible explanation for this might be that only a few geolocations (2 to 3) were estimated as this scenario focused on vehicle movement in the opposite directions and the sensing lifetime was limited; therefore, the high estimation deviation of one point has a big effect on the average error.

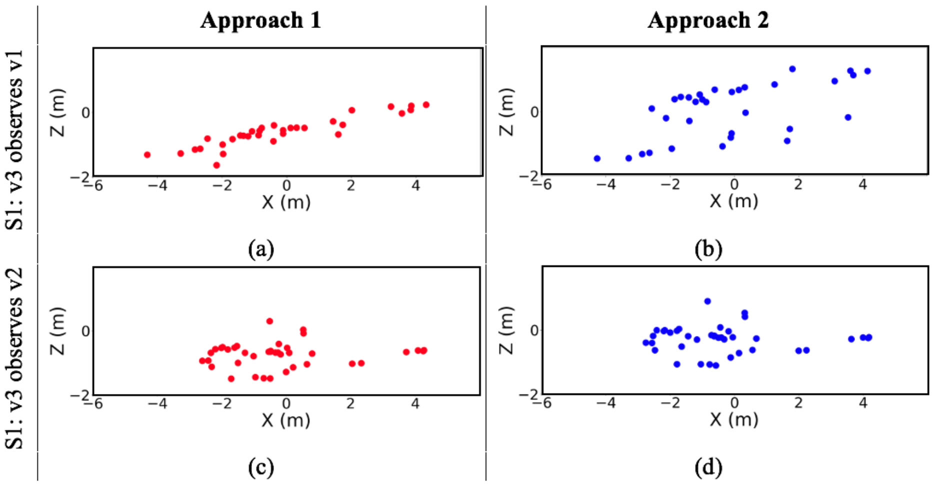


Figure 5.13: The distance vectors between the ground truth and the estimated geolocations with Approach 1 (a, c, e, and g) and Approach 2 (b, d, f, and h) for Scenario S1. Two cases are considered: (1) MV v3 observes target vehicle v1 and (2) MV v3 observes target vehicle v2. X is the longitudinal direction, and Z is the lateral direction.

Table 5.3: Evaluation results.

S#	V#	Estimation deviation of Approach 1 (m)				Estimation deviation of Approach 2 (m)			
		Min	Avg	Max	RMSE	Min	Avg	Max	RMSE
S1	v3 observes v1	0.50	2.03	4.51	2.35	0.35	2.02	4.51	2.34
	v3 observes v2	0.47	1.74	4.31	2.03	0.19	1.63	4.18	1.96
S2	v3 observes v1	2.04	3.54	4.33	3.70	2.18	3.51	4.20	3.63
	v3 observes v2	0.79	1.38	1.97	1.50	0.72	1.4	2.07	1.55

To analyze our proposed approaches further, we applied the Root Mean Square

Error (RMSE) to the distance vector between the estimated geolocations and the ground truth to show the estimation deviation. The calculated RMSE related to Approach 1 was between 1.5 m and 3.7 m (2.39 m on average). The calculated RMSE related to Approach 2 was between 1.55 m and 3.63 m (2.37 m on average). Overall, these results indicate that, in the studied scenarios, Approach 2 is slightly (about 0.02 m on average) better than Approach 1.

In summary, these results indicated that an MV equipped with a front-facing GoPro Hero 7 camera was able to identify the geolocations of the target vehicle dynamically.

More detailed results regarding SRQ2.5 are presented in Paper E, included in Part II of this thesis.

Taken together, the results obtained by RQ2 indicated that an MV equipped with a GoPro Hero 7 camera with a built-in GPS receiver can be used as a mobile sensor to estimate the required traffic data of the target vehicle (e.g., number, type, relative position, distance, speed, lane, and geolocation) in mixed traffic.

5.3 Improving the Self-localization Performance - Results of RQ3

The preliminary phase of identifying RQ3 originated from the research gaps identified in the SLR conducted to answer SRQ1.2, which showed that estimating the MV's location is one of the most significant traffic data types that should be considered by ITMSs. In addition, keeping SRQ2.5 in mind, the target vehicle's geolocation estimation is tightly connected to the MV's geolocation estimation. Therefore, RQ3 is defined to explore the feasibility of employing an MV equipped with two low-cost built-in GPS receivers of a GoPro Hero 7 camera mounted on an MV to improve the MV's self-localization accuracy.

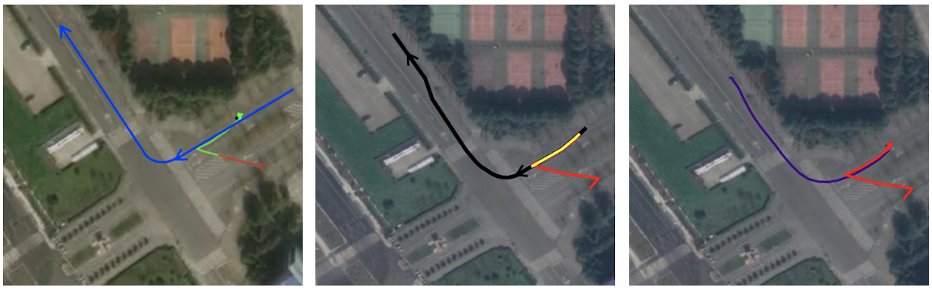
To state the problem, we need to mention that a GPS receiver is commonly used to estimate the MV's location in a GPS coordinate system, as most MVs are equipped with it. However, the accuracy of the data collected via a GPS receiver depends on several parameters, such as hardware accuracy, satellite geometry, signal blockage, and atmospheric conditions [73]. To satisfy an MV's self-localization requirements and mitigate the estimated location error by noisy sensors, the map-matching technique is widely used. Map-matching is a technique that integrates map information and recorded geolocation data from the vehicle in order to increase the accuracy of the MV's location [82]. Map-matching can be classified into online (i.e., during the measurement of the trajectory, e.g., in vehicle navigation) and offline (i.e., after the measurement of the trajectory) techniques [103]. Although map-matching techniques are widely applied to minimize the localization error, in this study, we found that map-matching techniques (e.g., the QGIS-Plug-in Offline-MapMatching [103][110]) do not work well if the GPS data is collected via a low-cost and noisy GPS receiver. By reviewing the literature, we discovered that there is a lack of scientific and empirical information about how to keep the sensors' costs low and the MV localization performance high.

To study the aforementioned problem in practice, we conducted empirical studies on real traffic data collected in Round 2 presented in Section 4.3. As expected, we observed that the positional data collected by using built-in GPS receivers of GoPro Hero 7 cameras were noisy. As a first attempt to mitigate the observed noise, we decided to use map-matching software. We investigated the effectiveness of several existing map-matching software tools (e.g., the map-matching tool developed by Waves [90], Mapbox Directions [29], and the QGIS-Plug-in Offline-MapMatching [103][110]). We found that the QGIS-Plug-in Offline-MapMatching [103][110] was compatible with our data and was a suitable and effective plugin in our research context. In addition, QGIS is one of the most widely used software tools (e.g., NPRA), because it is an open-access and well-documented tool. Moreover, it helps to understand the needs and requirements of future road transport since it does not require expensive licenses. The QGIS-Plug-in Offline-MapMatching [103][110] is a statistical approach based on the principles of Hidden Markov Models (HMM) and the Viterbi algorithm [103]. However, we found that the QGIS-Plug-in Offline-MapMatching [103][110] was not able to identify and map-match the entire trajectory if the vehicle localization error is too high (i.e., an MV location that was far from the road area). Figure 5.14-A and Figure 5.14-B present this problem. From these figures, it is clear that the QGIS-Plug-in Offline-MapMatching [103][110] was not able to identify and map-match the entire trajectory accurately if the vehicle localization error was too high. We applied more experiments by analyzing the data from another GPS receiver on the same vehicle in the studied scenarios. We found varying degrees of positional error between the two GPS receivers. These findings are presented in Figure 5.14-C.

Based on these experiments, we proposed a new approach with the main purpose of identifying the more accurate GPS receiver on the same MV and feeding the data collected by the identified accurate GPS receiver into the QGIS-Plug-in Offline-MapMatching [103][110] to enhance the MV's localization performance.

Figure 5.15 illustrates our proposed research strategy, which comprises data collection, data pre-processing, and data processing.

As Figure 5.15 shows, as part of our proposed approach, the collected data were pre-processed. In this step, first, we need to convert the data from a spherical



A B C

Figure 5.14: Problem formulation related to RQ3. (A) The MV's locations collected via a front-mounted GPS receiver on the vehicle (green-red polyline), compared with the vehicle's movement scenario (blue polyline). (B) Map-matching output (yellow polyline) related to the noisy front GPS receiver (red polyline) by considering the true trajectory of the vehicle (black polyline). (C) The vehicle's locations obtained via two GPS receivers on the same vehicle (front GPS receiver: red polyline; rear GPS receiver: purple polyline).

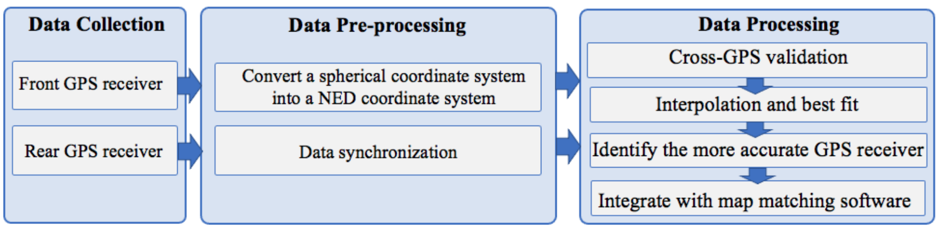


Figure 5.15: The components of our proposed research strategy to address RQ3.

coordinate system [128] into a local North-East-Down (NED) coordinate system [86] on the Earth's surface. The conversion is both practical and justified since we are studying a small, demarcated area on the Earth's surface. Second, since the two mounted GPS receivers on the vehicle are independent and the data collection was not started concurrently, we need to synchronize the receivers in the time domain.

In the follow-up step to process the data, we developed a new algorithm based on cross-validation, interpolation [60], and best-fit [56] techniques, as presented in Figure 5.16. Cross-validation aimed to identify the positions on the trajectory where both GPS receivers were almost in agreement on the vehicle's position (i.e., the position difference obtained by the two GPS receivers was between D_g - e and D_g + e, where D_g shows the fixed and known distance between two GPSs on the same MV and e shows the error threshold). It did so based on the Euclidean distance (Ed) [33] between each pair of pre-processed positions obtained by the front and rear GPS receivers per timestamp. As the number of validated positions obtained by the two GPS receivers can be limited, we applied interpolation [60]. In addition, for the straight vehicle movements, which were determined based on the vehicle's movement slope (we regarded a movement with a slope of less than 20 degrees as a straight movement; otherwise as a turn), the best-fit technique [56] was used to generate more positions along the whole trajectory based on the validated and interpolated positions. To identify the more accurate GPS receiver, we then calculated the average Euclidean distance [33] between the positions calculated through interpolation and best-fit techniques and the positions collected by each GPS receiver. The GPS receiver with a smaller average Euclidean distance was identified as the more accurate one.

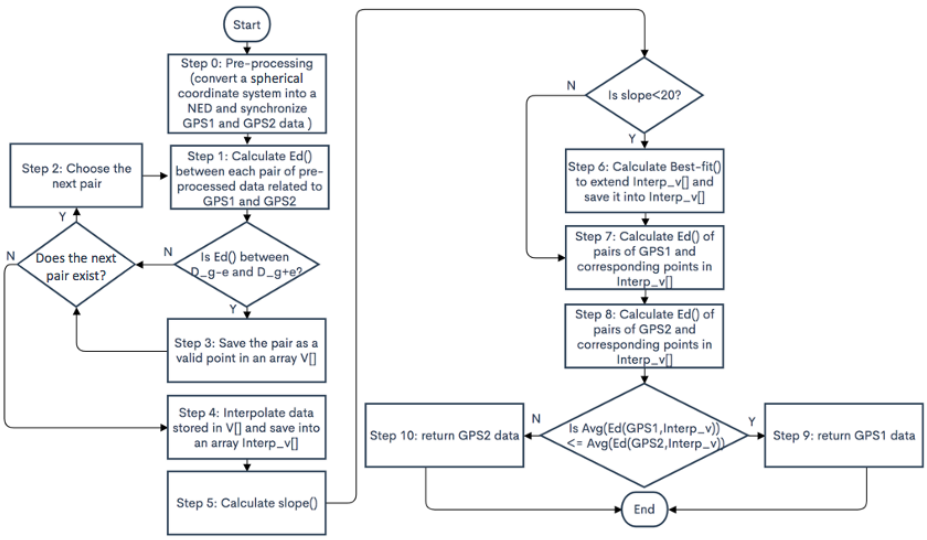


Figure 5.16: Flowchart of our proposed algorithm to address RQ3.

Although we can identify that one GPS receiver is more accurate than the other, the more accurate one may also be noisy. Finally, we inserted the data from the identified more accurate GPS receiver into the QGIS-Plug-in Offline-MapMatching [103][110] to further amend the noisy GPS signal.

To evaluate our proposed methodology for improving the MV's geolocation, we developed algorithms and ran experiments by using real traffic data collected in Round 2 presented in Section 4.3. We adjusted the frame rate to 1 FPS to apply map-matching. To run experiments, in order to provide good data coverage and generalizability, eight different scenarios were defined, comprising both straight-street and intersection movements. In total, 24 trajectories were considered. These scenarios are presented in Figure 5.17.

To assess our proposed approach, the MV's geolocation estimations should be compared with the ground truth. Because we did not use a specific accurate sensor to collect the ground truth, the ground-truth data related to the MV's movements were not available; therefore, we extracted them manually by visually observing the forward-facing video footage and identifying the ground-truth vehicle movements using Google Earth Pro [44].

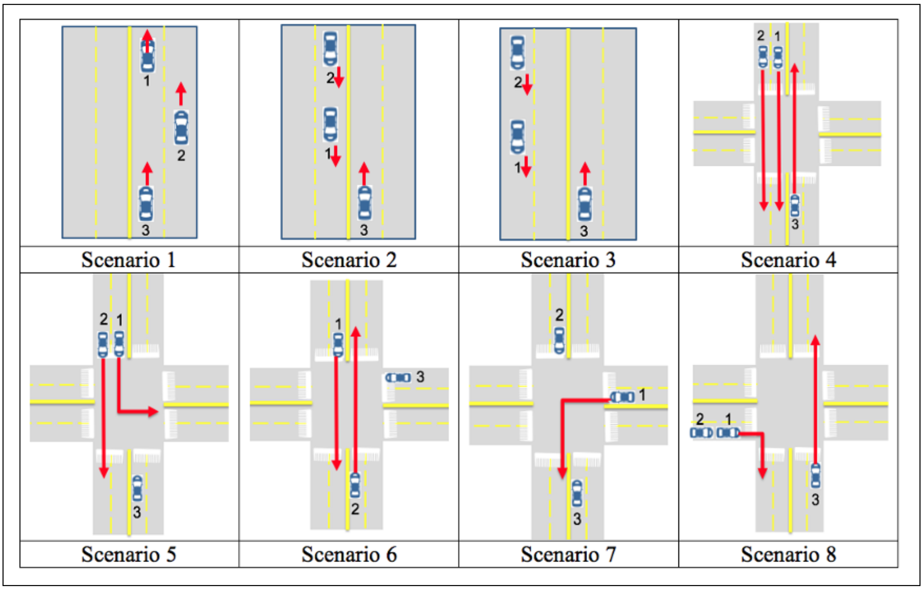


Figure 5.17: The studied scenarios by focusing on a straight street and an intersection related to RQ3.

Table 5.4 summarizes our findings. This table included eight scenarios (S1–S8) and three equipped vehicles (V1–V3). The "Ground-truth" column shows the Cartesian length of a vehicle's movement, and the "Avg. Dis." columns present the average distance between the vehicle's positions collected via each GPS receiver and the ground truth. The GPS receiver with the smaller average distance was labeled as the more accurate GPS receiver. To assess our proposed methodology, we first calculated the Cartesian length of the vehicle's trajectory by using only map-matching on data of both front and rear GPS receivers. Our findings are presented in the "Front GPS" and "Rear GPS" sub-columns of the "Map-matching-based Cartesian length" column. We then calculated the Cartesian length of the vehicle's trajectory, after applying our proposed methodology and identifying the accurate GPS receiver. The results are presented in the "Accurate GPS" (results of Steps 9 and 10 in Figure 5.16) and "Cartesian length" sub-columns of the "Our proposed approach" column. In addition, we compared the deviation from the ground truth by using only map-matching on the collected data and using our proposed approach. The results are shown in the sub-columns of the “Deviation comparison” column.

Table 5.4: Case study evaluation related to RQ3.

S#	V#	Ground truth (m)	Avg. Dis. (m)		Map-matching-based Cartesian length (m)		Our proposed approach		Deviation comparison (m)		
			Front GPS	Rear GPS	Front GPS	Rear GPS	Accurate GPS	Cartesian length (m)	Front GPS	Rear GPS	Our proposed approach
S1	V1	532	12.009	4.935	490	532	Rear	532	-42	0	0
	V2	514	2.044	10.746	513	514	Front	513	-1	0	-1
	V3	441	2.324	4.415	441	437	Front	441	0	-4	0
S2	V1	179	1.457	6.058	179	178	Rear	178	0	-1	-1
	V2	191	4.358	3.385	191	177	Front	191	0	-14	0
	V3	147	1.669	2.19	147	145	Front	147	0	-2	0
S3	V1	191	1.955	1.774	191	155	Rear	155	0	-36	-36
	V2	189	1.552	4.612	189	184	Front	189	0	-5	0
	V3	159	3.608	13.860	159	150	Front	159	0	-9	0
S4	V1	159	4.241	0.665	156	159	Rear	159	-3	0	0
	V2	163	6.044	1.84	163	162	Front	163	0	-1	0
	V3	188	1.388	2.264	188	188	Front	188	0	0	0
S5	V1	174	5.170	1.798	174	162	Rear	162	0	-12	-12
	V2	188	3.126	4.900	188	188	Front	188	0	0	0
	V3	118	1.450	2.385	118	118	Front	118	0	0	0
S6	V1	124	3.752	6.913	124	117	Rear	117	0	-7	-7
	V2	194	1.333	7.131	186	194	Front	186	-8	0	-8
	V3	–	–	–	–	–	–	–	–	–	–
S7	V1	106	1.834	4.460	106	106	Rear	106	0	0	0
	V2	142	7.660	4.803	141	142	Front	141	-1	0	-1
	V3	–	–	–	–	–	–	–	–	–	–
S8	**V1**	**109**	**3.515**	**1.402**	**27**	**109**	**Rear**	**109**	**-82**	**0**	**0**
	V2	107	1.493	2.983	103	107	Front	103	-4	0	-4
	V3	150	1.627	2.939	150	148	Front	150	0	-2	0

To explain the information presented in Table 5.4 in depth, we use scenario S8 and vehicle V3 as an example. In this scenario, the Cartesian lengths of the map-matched positions obtained via both GPS receivers are almost the same (front GPS: 150 m; rear GPS: 148 m). This shows that applying map-matching software would be enough to correct such small errors satisfactorily. However, this table shows that when the GPS error is high, applying only map-matching may not be effective, which is the main focus of this study. For instance, in scenario S8 with vehicle V1, the Cartesian length obtained by applying map-matching associated with the front GPS receiver is 27 m, while it is 109 m for the rear GPS receiver. This means that by using only one GPS receiver (i.e., the front GPS receiver), map-matching is effective only for a small segment of the trajectory (i.e., 27 m). The performance could be improved if we consider another GPS receiver. This confirms that identifying the more accurate GPS receiver is vital, which is the rear GPS receiver in this case. After identifying the more accurate GPS receiver and using its collected data to feed them into the map-matching software, our proposed approach improved the self-localization performance, which is measured using the Cartesian length of the output to 109 m. Therefore, the S8 and V1 case showed that our approach is effective in the presence of extreme GPS signal noise. As can be seen in Table 5.4, using our approach to choose a more accurate GPS receiver first and then apply map-matching does not always give less deviation than using the front or rear GPS receiver randomly. The reason is that we chose to use the GPS2 data in Step 10 in Figure 5.16 when the data from both GPS receivers were acceptable. The GPS2 data may not be better than the GPS1 data in some cases, although the data from both GPS receivers are acceptable. In this table, the information for vehicle V3 in scenarios S6 and S7 is not provided, as the rear GPS receiver did not record during the whole scenario. The reason for this could be that the battery died or that the memory card became full.

Taken together, our proposed approach was able to minimize the measurement error of the low-cost GPS receiver and was able to enhance the vehicle localization performance.

More detailed results regarding RQ3 are presented in Paper D, included in Part II of this thesis.

5.4 Multiple Sensor Fusion - Results of RQ4

The preliminary phase of identifying RQ4 originated from the SLR, which revealed that estimating the target vehicle's location is one of the most significant traffic data types that should be considered by ITMSs. In addition, based on our findings obtained from SRQ2.5, we found that image-based target vehicle localization can be noisy, especially for mobile and low-cost sensors. Therefore, the geolocations of a target vehicle estimated by an MV could be inaccurate. As we stated already in RQ2, this data could be advantageous for ITMSs from a variety of perspectives. Hence the need emerged to identify how the target vehicle localization performance could be improved by fusing the estimated geolocations of the target vehicle via two MVs each equipped with a low-cost GoPro Hero 7 camera with a built-in GPS receiver, which would be beneficial for enhancing the traffic scene awareness and thus for answering RQ4. In this regard, we proposed new methodologies and developed new algorithms.

The aforementioned problem is visualized in Figure 5.18. In Figure 5.18, point P' represents the true position of target vehicle v1 based on the view of MV v2. P represents the true position of target vehicle v1 based on the view of MV v3. P_1 shows the uncertain position of target vehicle v1 estimated by MV v2, and P_2 shows the uncertain position of target vehicle v1 estimated by MV v3. This means that both positions of target vehicle v1 estimated by MV v2 and MV v3 could be inaccurate and vary in a range shown by the blue circles. In Figure 5.18, all the vectors are on a 2D plane. Vector $\vec{x}$ shows the estimated distance between MV v2 and P_1 in coordinate system CCS_1. $\vec{z_t}$ is the true (but unknown) distance to target vehicle v1 in CCS_1, and $\vec{v}$ is the 2D estimation error of $\vec{x}$ caused by MV v2. Similarly, vector $\vec{y}$ shows the estimated distance between MV v3 and P_2 in coordinate system CCS_2. $\vec{g_t}$ is the true (but unknown) distance to target vehicle v1 in CCS_2, and $\vec{w}$ is the 2D estimation error of $\vec{y}$ caused by MV v3. Moreover, $\vec{t}$ shows the transformation vector of $\vec{y}$ between the positions of MV v3 and MV v2.

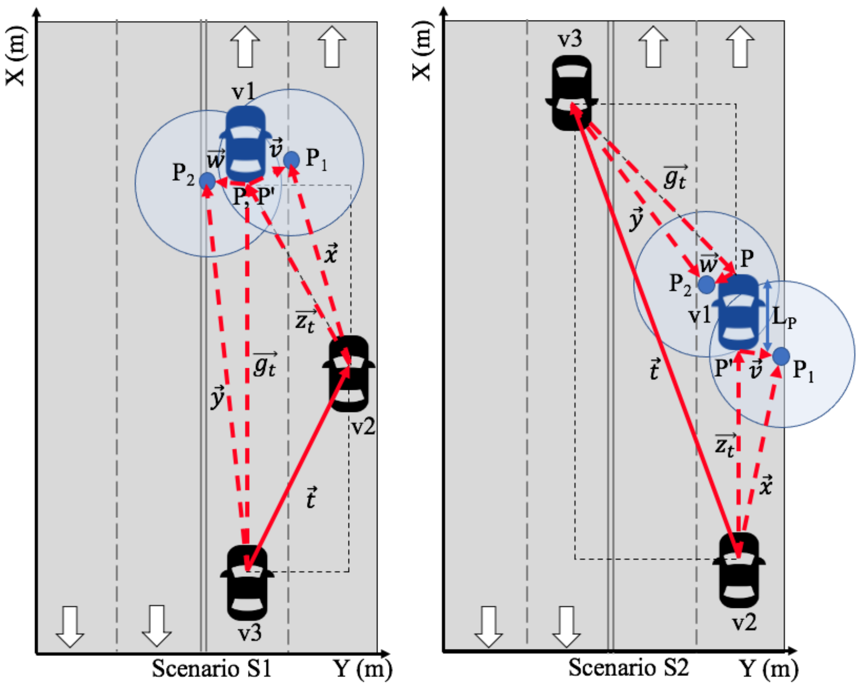


Figure 5.18: Problem statement and illustration of the cases we focused on related to RQ4.

To overcome the aforementioned problem, RQ4 aims to enhance the geolocation estimation of a target vehicle via two MVs, besides helping to provide a more comprehensive picture of the traffic scene than what can be obtained by using data from only one MV based on re-identification and multiple sensor fusion. Figure 5.19 illustrates the overall view of our proposed research strategy, including three main components. After the data were collected via both MVs in the data collection Round 2 presented in Section 4.3, we performed data pre-processing to mitigate the self-localization noise of the MVs caused by the GPS receivers based on our proposed approach presented in RQ3. Also, we performed pre-estimation to estimate the target vehicle's type, lane, and geolocation based on SRQ2.1 (the only change was training YOLO-V3 with the KITTI dataset) and SRQ2.5 by both MVs. After that, we identified whether both MVs were observing the same target vehicle

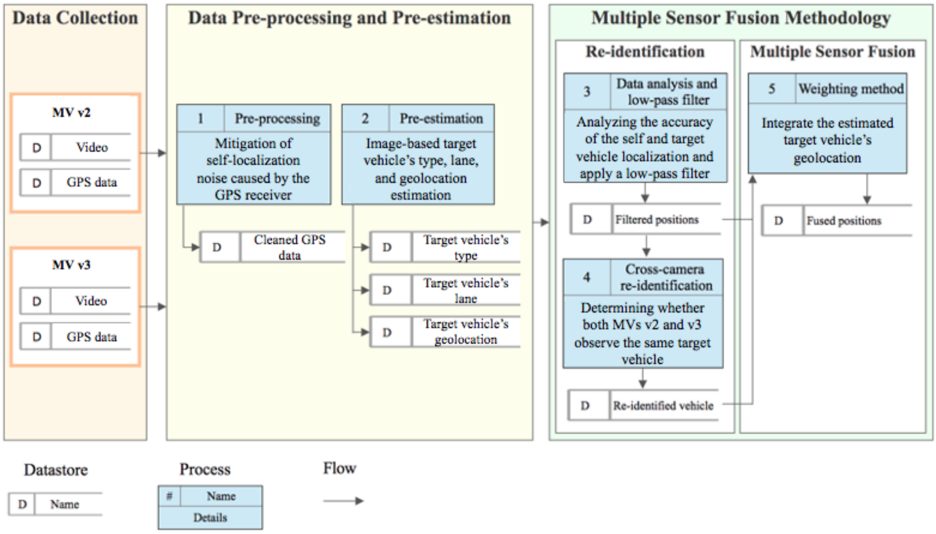


Figure 5.19: The components of our proposed research strategy to address RQ4.

(called re-identification) and dynamically fused their estimated re-identified target vehicle's geolocation based on the weighing method.

5.4.1 Target Vehicle Re-identification - Results of SRQ4.1

SRQ4.1 aims to determine whether the target vehicle observed by the two MVs was the same one by considering the estimation uncertainties. Therefore, as shown in Figure 5.19, we proposed two stages in which to apply target vehicle re-identification. (1) We analyzed the estimated positional data to determine the accuracy of the cleaned GPS data of both MVs based on the proposed methodology in RQ3 and the estimated geolocations of the target vehicle obtained by both MVs based on the proposed methodology in SRQ2.5. After that, to mitigate the uncertainties of the image-based estimated target vehicle's geolocation, we applied a low-pass moving average filter [96] (the term low-pass filter is used interchangeably) to smooth out the noise. (2) We developed a new algorithm based on cross-camera target vehicle re-identification by considering three image-based estimated data types related to the target vehicle: type, lane, and geolocation. If both MV v2 and MV v3 agreed on the value of these three data types, we labeled the observed vehicle as the same target vehicle. Otherwise, we concluded that the MVs observed different target vehicles.

5.4.2 Target Vehicle's Geolocation Estimation by Applying Multiple Sensor Fusion - Results of SRQ4.2

Once the MVs agreed that they were observing the same target vehicle, the estimated positions P_1 and P_2 obtained by both MVs could be fused dynamically. Our proposed approach to fusing P_1 and P_2 was based on the weighing method by exploring appropriate weights for each MV based on its estimation uncertainty level. This process is presented in detail below.

As shown in Figure 5.18, Eq. 5.31, and Eq. 5.32, the distance between MVs and the target vehicle, obtained by considering uncertainties, is as follows.

$$\vec{x} = \vec{z_t} + \vec{v} \tag{5.31}$$

$$\vec{y} = \vec{g_t} + \vec{w} \tag{5.32}$$

In order to fuse the two positions P_1 and P_2, we transformed our measurements to the same coordinate system. We chose CCS_1 to be our reference coordinate system. Eq. 5.33 shows how to calculate the transformed vector, named $\vec{d}$, in CCS_1, as $\vec{x}$ and $\vec{y}$ had different estimation errors. The true measurement part of $\vec{d}$ is equal to the true measurement part of $\vec{x}$, named $\vec{z_t}$, which is the true (but unknown) position of the target vehicle in CCS_1, and $\vec{k}$ is the 2D estimation error of $\vec{d}$ in CCS_1.

$$\vec{d} = \vec{z_t} + \vec{k} \tag{5.33}$$

As shown in Eq. 5.34, we assumed that $\vec{x}$ and $\vec{d}$ were fused in a linear way, and the result is named $\hat{\vec{q}}$. In this equation, $\mathbf{A}$ and $\mathbf{B}$ are two-by-two matrices and represent the corresponding weights for fusing the estimated positions via two MVs, where $\mathbf{B} = 1 - \mathbf{A}$.

$$\begin{aligned} \hat{\vec{q}} &= \mathbf{A} \cdot \vec{x} + \mathbf{B} \cdot \vec{d} \\ &= \mathbf{A} \cdot \vec{x} + (1 - \mathbf{A}) \cdot \vec{d} \end{aligned} \tag{5.34}$$

In the following step, we explored the corresponding weights ($\mathbf{A}$ and $\mathbf{B}$) for each MV to fuse the estimated positions of the re-identified target vehicle.

To decide which MV was more accurate and should receive a higher weight, we utilized the similarity between the low-pass filtered and unfiltered estimated positions of the target vehicle obtained by both MV v2 and MV v3 at each timestamp in both the X and Y directions. The rationale was that the MV with the lower similarity between the low-pass filtered and unfiltered estimated positions of the target vehicle was noisier and less trustworthy than the other one and thus should have less influence on the fusing process. We used a distance vector to estimate the mentioned similarity.

By considering the calculated distance vector at each timestamp, we gained deeper knowledge about the accuracy of the estimated positions. We used two datasets to calculate the accuracy, as follows:

1. The distance vector between the unfiltered and filtered positions of P_1 in both the X and Y directions.
2. The distance vector between the unfiltered and filtered positions of P_2 in both the X and Y directions.

Therefore, Eq. 5.35 and Eq. 5.36 show how to calculate weights $\mathbf{A}$ and $\mathbf{B}$ for each MV, in which $\mathbf{C_{P_1}}$ is the covariance matrix of the dataset related to P_1 and $\mathbf{C_{P_2}}$ is the covariance matrix of the dataset related to P_2, calculated based on the distance vector.

$$\mathbf{A} = \mathbf{C_{P_2}} \cdot (\mathbf{C_{P_1}} + \mathbf{C_{P_2}})^{-1} \tag{5.35}$$

$$\mathbf{B} = \mathbf{C_{P_1}} \cdot (\mathbf{C_{P_1}} + \mathbf{C_{P_2}})^{-1} \tag{5.36}$$

By plugging Eq. 5.35 and Eq. 5.36 into Eq. 5.34, the fused position of the target vehicle can be calculated as follows:

$$\hat{\vec{q}} = \mathbf{C_{P_2}} \cdot (\mathbf{C_{P_1}} + \mathbf{C_{P_2}})^{-1} \cdot \vec{x} + \mathbf{C_{P_1}} \cdot (\mathbf{C_{P_1}} + \mathbf{C_{P_2}})^{-1} \cdot \vec{d} \tag{5.37}$$

We assumed that $\mathbf{C_{P_1}}$ and $\mathbf{C_{P_2}}$ are diagonal matrices. Therefore, the fused position in 2D can be calculated by using Eq. 5.38 and Eq. 5.39, in which the formulas' parameters are as follows:

${\sigma_{P_1}}^2$ is the variance between the filtered and unfiltered positions related to P_1

${\sigma_{P_2}}^2$ is the variance between the filtered and unfiltered positions related to P_2

As already stated, all the vectors are on the 2D plane; therefore, each vector has two components in two dimensions, as follows:

$\hat{\vec{q}} := \begin{pmatrix} \hat{q}_1 \\ \hat{q}_2 \end{pmatrix}$, $\vec{x} := \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, and $\vec{d} := \begin{pmatrix} d_1 \\ d_2 \end{pmatrix}$.

$$\hat{q}_1 = \frac{{\sigma_{P_2}}^2}{{\sigma_{P_2}}^2 + {\sigma_{P_1}}^2} \cdot x_1 + \frac{{\sigma_{P_1}}^2}{{\sigma_{P_2}}^2 + {\sigma_{P_1}}^2} \cdot d_1 \tag{5.38}$$

$$\hat{q}_2 = \frac{{\sigma_{P_2}}^2}{{\sigma_{P_2}}^2 + {\sigma_{P_1}}^2} \cdot x_2 + \frac{{\sigma_{P_1}}^2}{{\sigma_{P_2}}^2 + {\sigma_{P_1}}^2} \cdot d_2 \tag{5.39}$$

To evaluate the proposed methodologies for fusing the estimated geolocations of the target vehicle via two MVs, we developed a new algorithm and ran experiments using real traffic data collected in Round 2 presented in Section 4.3. We found that our proposed methodology and developed algorithms were able to enhance the target vehicle localization accuracy by considering the estimation uncertainty caused by low-cost monocular cameras (i.e., GoPro Hero 7 camera) on mobile MVs while they are following various trajectories with different views. To assess the accuracy of our proposed sensor fusion approach, we compared our findings based on Eq. 5.38 and Eq. 5.39 with the ground-truth positions of the target vehicle, which were collected by the built-in GPS receiver of the GoPro Hero 7 camera (we pre-processed the ground-truth data collected by the GPS receiver based on the proposed approach, in RQ3, to improving the accuracy). Figure 5.20 and Figure 5.21 show our findings related to Scenario 1. The most interesting aspect of these figures is that the fused positions (the red color polyline) are almost between the positions estimated via v2 (the green color polyline) and v3 (the blue color polyline) and fluctuate around 0, which supports our claim that our proposed methodology can improve the localization accuracy.

In addition, Table 5.5 includes our quantitative findings of the comparison of the estimated positions based on the deviation from the ground positions before fusion and the fused positions in both the X and Y directions, based on the distance between the ground-truth position and the estimated position at each timestamp. This table shows that our proposed fusion methodology is able to reduce the estimation noise of the target vehicle's geolocation by MV effectively. For instance, the average error of the estimated position obtained by v2 in Scenario 1 by considering the X direction decreased from 0.38 to 0.14 m, which means a 63.16% improvement.

The experiments confirmed that our proposed approach was able to enhance the accuracy of the image-based target vehicle's geolocation estimation effectively.

More detailed results regarding RQ4 are presented in Paper F, included in Part II of this thesis.

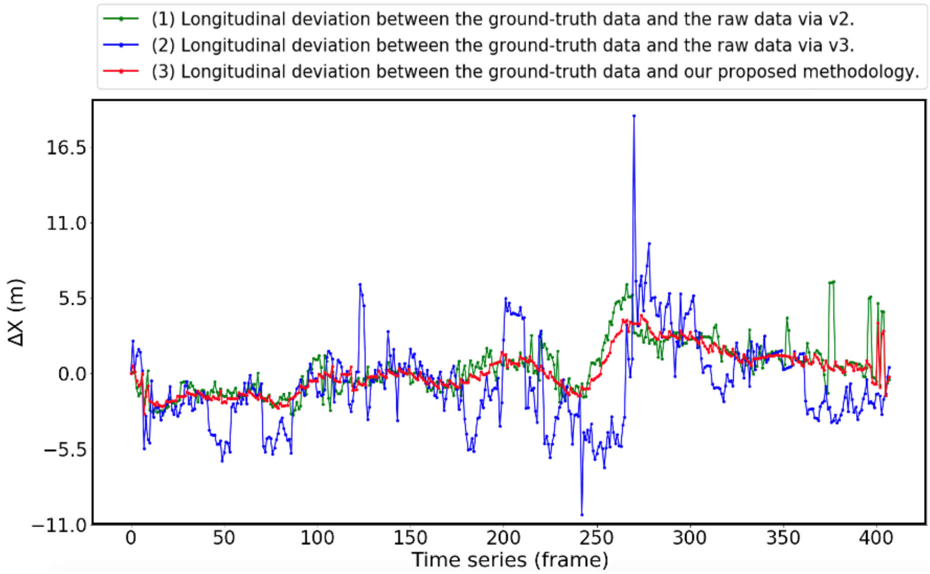


Figure 5.20: Longitudinal deviation between the ground-truth data and three datasets. (1) The results before applying our proposed methodology via v2 are shown by the green polyline. (2) The results before applying our proposed methodology via v3 are shown by the blue color polyline. (3) The results after applying our proposed methodology are shown by the red polyline in the X direction in Scenario 1 related to RQ4.

Table 5.5: Comparison between our proposed approach to localizing the target vehicle and the ground-truth data related to RQ4.

		X (m)			**Y (m)**		
Senario	**Data**	**Min**	**Max**	**Avg**	**Min**	**Max**	**Avg**
S1	**Raw positions via v2**	-3.02	6.68	0.38	-2.20	3.50	-1.00
	Raw positions via v3	-10.28	18.79	-0.84	-4.12	5.28	0.19
	Fused positions	-2.95	4.25	0.14	-1.19	1.03	-0.36
S2	**Raw positions via v2**	-7.38	1.42	-1.01	-1.45	0.34	-0.18
	Raw positions via v3	-0.63	9.41	2.08	-0.99	6.26	0.32
	Fused positions	-2.01	2.27	0.56	-1.2	0.25	-0.14

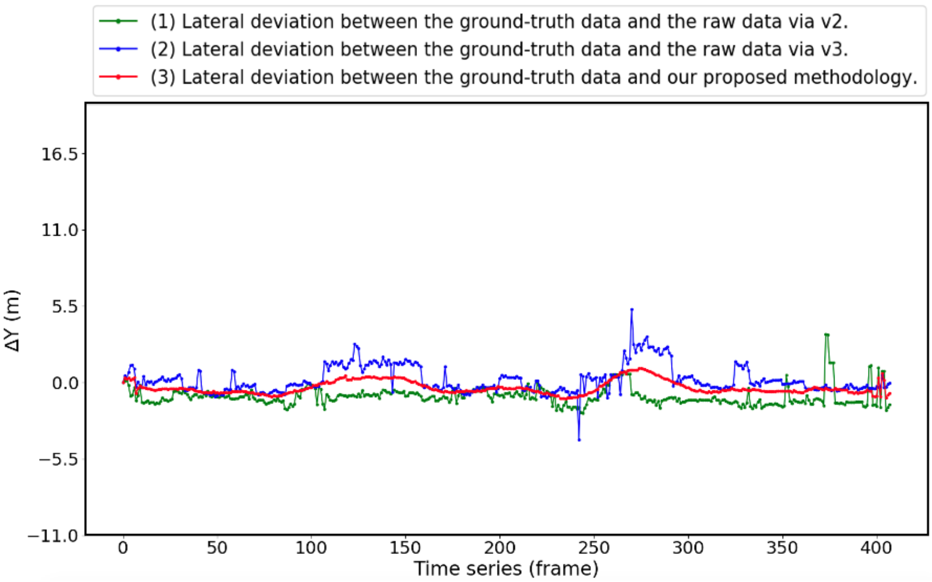


Figure 5.21: Lateral deviation between the ground-truth data and three datasets. (1) The results before applying our proposed methodology via v2 are shown by the green polyline. (2) The results before applying our proposed methodology via v3 are shown by the blue polyline. (3) The results after applying our proposed methodology are shown by the red polyline in the Y direction in Scenario 1 related to RQ4.

Chapter 6

Discussion

This chapter is composed of five sections and organized based on the RQs addressed in this thesis. The chapter synthesizes the research findings of the four RQs by discussing the contributions made in terms of a comparison with related work, implications for academia, implications for practitioners, and related threats to validity. In addition, possible ethical issues are discussed.

6.1 Literature Review of Intelligent Traffic Management Systems with Modern Vehicles - RQ1

The first RQ addressed in this Ph.D. thesis aimed to produce substantive findings regarding the proposed methodologies for ITMSs by focusing on both pure MV (i.e., AV) traffic and mixed traffic at four-way unsignalized and signalized intersections. This contribution was achieved through conducting an SLR. This study presented the current state of the art in the pre-defined scope and explored the potential research gaps.

A Comparison with Related Work

Addressing this RQ contributes to the literature in several ways. (1) We applied an SLR strategy in our study to review the state of the art. One of the advantages of conducting the SLR is that it follows a well-defined methodology (i.e., considers a review protocol or a guideline), which reduces the possibility of bias in presenting the literature results. Also, defining a search strategy to conduct an SLR helps to detect as many relevant papers as possible in the studied area. Moreover, documenting the search strategy and the steps followed to conduct an SLR helps readers to assess the completeness of it and makes the process repeatable [68]. To the best of our knowledge, other available studies related to our study domain at that time did not use an SLR strategy and mostly focused on conducting a survey (e.g., [21][121][48]). (2) We defined the research scope precisely and explicitly to gain deeper knowledge of a specific domain. Our study targeted papers on TMSs at signalized intersections with a combination of both HDV and MV traffic and unsignalized intersections with pure MV traffic. Other available studies (e.g., [21][121][48]) had a different research scope in that they focused on various types of traffic flows and/or different types of scenarios. (3) Available studies mainly summarized the proposed approaches regarding TMSs (e.g., [21][121][48]). In our study, in addition to focusing on the proposed methodologies in TMSs (i.e., optimization and rule-based, hybrid, and machine learning), we extracted extra information and applied a deeper, thematic analysis, which led us to formulate RQ2 - RQ4. For instance, we considered the investigations presented in the studied papers, compared how well the proposed approaches were evaluated, and summarized their results. Moreover, we identified and categorized the presented research goals and sub-goals of the studied papers. In addition, we identified and classified the data types that other researchers considered. This combination of findings provided some support for the conceptual premise that there is a need to estimate the traffic data in mixed traffic by assuming an MV equipped with low-cost sensors as a mobile sensor, as formulated in RQ2 - RQ4.

B Implications for Academia

Our results provide some support for the conceptual premise that ITMSs need to be compatible with mixed traffic and that the benefits of MVs should be considered in this regard. Our findings of SRQ1.1 and SRQ1.2 have important implications for proposing and developing ITMSs that use the data estimated by MVs based on AI and image processing techniques to generate a dynamic model of the traffic scene, which enhances the ITMS's awareness, enables smart decisions to be made, and boosts performance. Also, the presented results in SRQ1.3 related to the explored potential research gaps (e.g., investigating traffic including pedestrians and cyclists, traffic status prediction, communications, and data sharing) revealed many potential new research topics and can guide researchers in further studies and investigations.

C Implications for Practitioners

One of the most interesting findings of this study is that although the applied technologies for managing intersections intelligently and developing MVs are advancing fast, few studies considered mixed traffic in managing an intersection intelligently. Therefore, this finding has significant implications for building a bridge between traffic managers and car producers to allow them to become aware of each other's strengths and needs. For instance, gaining knowledge about the required traffic data types, data processing approaches, and communication types that are employed by traffic managers could be considered by car producers in order to make the sensing and connection capabilities of MVs compatible with ITMSs and to provide the required traffic data for them. Furthermore, traffic managers, by gaining acknowledge about the MVs' needs, would be able to share the required data about the traffic scene and to help MVs make smart decisions and do safe motion planning.

D Threats to Validity

- Objectivity: In order to enhance the objectivity of this study by mitigating the possibility of biases or distortions in the research, we conducted an SLR by following the pre-defined and well-structured guidelines proposed by Kitchenham [68][69]. Moreover, to mitigate our influence on the results or a vested interest in a specific outcome, all steps of the SLR were validated and confirmed by all three authors.

- Reliability: In order to enhance the reliability of this study, we need to improve its accuracy and make the research repeatable. The applied data extraction approach might lead to extracting incomplete information from the selected papers or misunderstanding the provided information in the selected papers. Another bias might occur during the thematic synthesis process of extracting and classifying the required information. To mitigate these threats, the data extraction and analyses were done in several iterations to enhance the accuracy of the extracted information. Also, at least one of the co-authors validated the accuracy of each step (including the search process, extracting data, and analyzing data). To make the research repeatable, we documented all steps (e.g., keywords, DBs, search year, and inclusion and exclusion criteria) that were considered in conducting this SLR.

- Internal validity: In order to enhance the internal validity of this study, we need to consider whether the research is well designed and whether we collected the right data from the right source, which leads us to address the RQs and claimed findings. To increase the internal validity, we generated a conceptual framework of the research in advance to have a full picture of the research process (including RQs and search process). In addition, we continuously kept the RQs in mind and tried to address them in this SLR.

- External validity: In order to enhance the external validity of this study, we need to consider the generalizability of the findings. The possible threats to the external validity of conducting an SLR may come from the search process, which might lead to missing relevant papers. For instance, some papers may not be included because the coverage of the keywords was not complete or the DBs that were used did not contain all the relevant papers. Also, some papers might be discarded during the inclusion/exclusion steps because of mistakes made by the authors. To mitigate these threats, we used a keyword-based search and did a preliminary review to identify the most popular keywords used by other researchers and included most of the synonyms for them. Also, the selected search keywords were evaluated by an expert, Chaoru Lu, in this research area. Then, various combinations of these keywords were used during the search process. Moreover, we searched seven relevant digital libraries to increase the chance that we included the relevant papers. We did not apply the snowballing technique in this process, as we ended up with 2952 primary papers after a keyword-based search. Also, the inclusion and exclusion criteria were applied iteratively to make the process manageable and consistent.

6.2 Estimating the Target Vehicle's Traffic Data - RQ2

The second RQ addressed in this Ph.D. thesis aimed to explore the feasibility of an MV that is equipped with a low-cost monocular camera with a built-in GPS receiver to be used as a mobile sensor that can estimate the required traffic data of the observed target vehicles. The required data to be collected are inspired by the results of RQ1 (i.e., vehicle's number, type, relative position, distance, speed, lane, and geolocation).

A Comparison with Related Work

As stated in Chapter 3, several studies have shown that stationary cameras mounted on the road are able to collect the required traffic data. For instance, stationary cameras are used for vehicle detection, vehicle tracking, and speed estimation (e.g., [15][49]). This makes our approach an important contribution to vision-based traffic data collection, as the camera is one of the popular stationary sensors in ITMSs, which increases the practicality of our proposed approach. In addition, it has previously been observed that MVs would be able to estimate data of the surrounding vehicles (e.g., [109]). Such data is widely used for self-awareness purposes and provides information that enables autopilot/autonomous transportation and mobility (e.g., [144]). Moreover, based on our findings of RQ1, we explored the necessity of considering mixed traffic in data collection and managing the traffic. Also, to enhance the generalizability of the proposed approaches to be compatible with real-world situations, considering low-cost sensors that can easily be mounted on an MV would be helpful.

In this study, we combined the aforementioned findings and considered an MV equipped with a GoPro Hero 7 camera with a built-in GPS receiver as a mobile sensor and proposed new methodologies and developed new algorithms based on the state of the art to dynamically estimate the required traffic data by considering the ITMSs' needs in mixed traffic.

Addressing this RQ contributes to the literature in several ways. As mentioned in Section 3.2, a considerable body of literature has been published on vehicle detection (e.g., [113][47][23][24][98][24][97][39][41][126][14] [72][15]), tracking (e.g., [97][49][80][136]), lane detection (e.g., [137][88][89] [74][77][133]), target vehicle's distance estimation (e.g., [25][62][149][35][115] [58][41][39][52]), and target vehicle's speed estimation (e.g., [64][94][4][124][97] [85][78]). However, we decided to develop our own system based on state-of-the-art algorithms. The main rationale for developing our own system was that, to the best of our knowledge, there were no similar open-source systems available that would be able to estimate all the required traffic data (i.e., number, type, relative position, distance, and speed) of the target vehicle via the vision of a mobile MV. In addition, developing our own algorithms enabled us to have full control over them and to modify them based on the requirements of our RQs. Also, our algorithms had more functionalities than the state-of-the-art algorithms. For instance, to estimate the target vehicle's distance, we used prior knowledge of the real vehicle's size and applied a weighting factor to combine vehicle width and height to obtain a stable distance estimation (presented in SRQ2.2). Our experimental results showed that the best ratio for combining the estimated distances is 85% of the height and 15% of the width. As the target vehicle's speed was estimated based on the traveled distance and time between measurements for individual target vehicles, the accuracy of the target vehicle's distance estimation directly affects the target vehicle's speed estimation (presented in SRQ2.3). Furthermore, realizing that estimating the target vehicle's localization plays an important role in determining the traffic density and modeling the traffic scene dynamically, we focused on dynamically estimating the target vehicle's lane (presented in SRQ2.4) and the target vehicle's geolocations in a GPS coordinate system (presented in SRQ2.5). As presented in Section 3.2.5, a growing body of literature recognizes the importance of localizing the target vehicle by various sensors (e.g., radar, laser scanner, vision, and GNSS). However, as we stated already, using a low-cost sensor to increase the generalizability of the proposed approach is vital. Moreover, de Ponte Müller [109] showed that the accuracy of vision-based relative positioning techniques is between 1 and 5 m. In addition, as surveyed by de Ponte Müller [109], most existing studies are focused mainly on estimating the relative position of the target vehicle (e.g., [129][40][65][3]). Therefore, another contribution of our study was to estimate the target vehicle's lane and to estimate the target vehicle's geolocation in a GPS coordinate system by using a low-cost sensor. Besides enhancing the accuracy of the target vehicle's location estimation, this could be beneficial for modeling the traffic by ITMSs. As presented in SRQ2.4, we proposed two new approaches to estimating the target vehicle's lane dynamically. The main idea of both approaches was to estimate the distance (the horizontal and the shortest distance) between the central point on the bottom edge of the bounding box around the target vehicle and the detected lines nearby the MVs. Additionally, in order to estimate the geolocation of the target vehicle, we proposed two approaches based mainly on object detection, image processing, and geometric computations (presented in SRQ2.5). Our developed algorithms can be regarded as a starting point for generating a dynamic model of the traffic scene (i.e., a digital twin) and enhancing ITMSs' awareness and performance in the future.

B Implications for Academia

Our research revealed many potential new research topics in traffic data estimation of a target vehicle by utilizing MVs as mobile sensors. This study is an initial step in this regard, and more studies are strongly recommended to enhance the accuracy and performance of our developed algorithms. Some examples are presented below.

Vehicle detection is a preliminary step in estimating traffic data. To address RQ2, we used an existing object detection algorithm, assuming that it was reliable to use. This means that studying the accuracy and performance of YOLO-V3 was beyond the scope of this thesis, and further research is needed on these aspects. One of the limitations of YOLO-V3 that we experienced in this study was that the generated bounding boxes around the detected vehicles were not stable between frames, which directly affected the accuracy of our image-based traffic data estimations. Another limitation of YOLO-V3 was that it was not effective in detecting faraway target vehicles (e.g., a vehicle located on the other side of an intersection). Another limitation was caused by the limited vision lifetime of the vehicles driving on the opposite lane of the MV. In addition, the classification of vehicles into five major groups (i.e., car, bus, truck, motorbike, and bicycle) is another limitation of this study. For each category, we assumed a fixed/standard size. However, vehicles of different models/brands, even in the same group, can have different sizes, which can have a negative effect on the accuracy of the traffic data estimation. Lastly, another limitation was that the vehicle tracking approach did not work well in a crowded area (e.g., when two vehicles passed each other, their labels may switch). Further studies that take these limitations into account are therefore needed.

Our research revealed many potential new research topics with respect to lane detection approaches and algorithms. We used canny edge detection [28] and the PPHT [38][91] to identify lanes for estimating the target vehicle's number, relative position, and lane. The employed algorithms for detecting lines are efficient if lane marks are clearly visible and the street is straight, which was the case in the scenarios studied in this thesis. However, these approaches are not sufficiently efficient if lane marks are missing or not visible (due to, for instance, sun reflection or snow coverage) or if there are curvy streets. Also, the detected lanes were not stable between frames, which might have a negative effect on the accuracy of the traffic data estimation. Therefore, an improved lane detection algorithm that is compatible with various traffic scenarios would be useful.

Our results identify new factors that may influence the accuracy of estimating traffic data using a low-cost monocular camera and a GPS receiver. We used the GoPro Hero 7 camera with a built-in GPS receiver to collect traffic data with the purpose of enhancing the generalizability and practicality of our proposed approach. Collecting data with a camera is a common way of generating data for existing ITMSs, and a camera can easily be mounted on most MVs. However, we need to consider its limitations (i.e., required brightness, view angle, observation range, and estimation uncertainty). Another limitation is related to the manual installation of the camera(s) by a suction cup on the MV. Therefore, cameras might not be installed exactly in the middle of the window glass perpendicular to the vehicle's movement direction (i.e., this means that the camera's pitch, yaw, and roll angle might not be zero). Also, this installation is not secure enough, and the vehicle movements may cause some vibrations of the camera, which affects the visual quality and estimation accuracy. Future studies on these topics are therefore recommended to mitigate these limitations.

Moreover, the experiments in this study were carried out offline (collecting data from the real world and data processing were done in the office). However, running the system online (putting a system inside the vehicle and collecting and processing data in real time as the vehicle moves along the trajectory) to assess its accuracy and performance in the real situation is vital. Therefore, more studies are needed to make the system operate in real time and to run experiments in real time.

In addition, dynamically estimating the traffic data by an MV based on the ITMSs' needs can lead to developing a digital twin of the traffic. Therefore, more studies are needed that focus on generating a digital twin of the traffic scene and on making smart decisions, besides enhancing safety and performance.

C Implications for Practitioners

The most interesting finding of RQ2 was that MVs equipped with a low-cost monocular camera (i.e., a GoPro Hero 7 camera with a built-in GPS receiver) could be utilized as mobile sensors to estimate the traffic data required by ITMSs. As monocular cameras are one of the most popular sensors in existing ITMSs, our proposed approach could be beneficial for traffic managers as it could be adapted to their existing systems. As we proposed to utilize cameras mounted on MVs to collect traffic data for managing the traffic intelligently, it would be advantageous to make car producers and traffic managers aware of each other's needs and strengths. Besides exploiting the sensing capability of MVs in intelligent traffic management, this contribution could potentially be used to minimize the costs of installing and maintaining stationary sensors for collecting the required traffic data for the required sensing range.

D Threats to Validity

- Objectivity: In order to enhance the objectivity of RQ2 by mitigating the possibility of biases or distortions in the research, instead of using a simulator or laboratory data (which might omit the complexities of real-life situations), we decided to collect data from real traffic (which reflected some real-life conditions) and use them to assess our developed algorithms.

- Reliability: To enhance the study's reliability, we need to improve its accuracy and make the research repeatable. Therefore, we ran experiments on real traffic data by considering various scenarios from two countries (i.e., Trondheim in Norway and Chengdu in China) to estimate the accuracy of our proposed methodologies and to enhance the reliability and neutrality of the results. In addition, we described our proposed methodologies, developed algorithms, and experimental conditions in detail to enhance the repeatability.

- External validity: In order to enhance the external validity of this study, we need to consider the generalizability of the findings. Collecting data from real traffic helped us to deal with real traffic situations, even in limited scenarios. Also, we used a low-cost camera in order to enhance the generalizability of our proposed approaches as it is a popular sensor for both ITMSs and MVs.

6.3 Modern Vehicle Self-localization - RQ3

The third RQ addressed in this Ph.D. thesis aimed to enhance the self-localization performance of MVs via two low-cost GPS receivers mounted on an MV with a known distance between each other, which has a direct effect on the accuracy of the target vehicle localization addressed in SRQ2.5.

A Comparison with Related Work

Previous studies have noted the importance of identifying and mitigating the measurement error of GPS receivers to enhance the self-localization performance of MVs. As presented in Section 3.3, various techniques for MV self-location have been proposed (e.g., [111][107][22][5][20]). To meet the vehicle localization requirements and mitigate the estimated location error, three major categories of approaches have been proposed in the literature [57][61]. One category of approaches uses a standalone reference station, such as a Wide Area Augmentation System (WAAS) (e.g., [141]). The second category comprises auxiliary hardware-based approaches (e.g., IMU) [92]). Using technologically advanced sensors to determine vehicle's location would boost the estimation accuracy. However, equipping a vehicle with such sensors will also increase the vehicle's cost and decrease the generalizability. The third category uses software, such as map-matching techniques (e.g., [143]). Map-matching is a technique that integrates map information and recorded geolocation data from the vehicle in order to increase the accuracy of the vehicle's location [82]. Although map-matching techniques are widely applied to minimize vehicles' localization error, as presented in RQ3, we found that map-matching techniques (e.g., QGIS-Plug-in Offline-MapMatching [103][110]) do not work well if the GPS data collected via a low-cost GPS receiver are too noisy.

Since making a trade-off between the localization cost and MV localization accuracy is vital, RQ3 aimed to address how to keep the sensors' costs low and the localization performance high. In this regard, we developed a new algorithm (based on cross-validation, interpolation, and best-fit techniques) to identify the more accurate GPS receiver if there are two GPS receivers mounted on the same MV. Compared with the approach relying on expensive or multiple sensors, our approach provides a low-cost solution to precisely identify an MV's location. Compared with the approach that relies solely on map-matching, our strategy of detecting GPS inaccuracy and prioritizing the data from the more accurate GPS receiver helped enhance the performance of the map-matching software.

B Implications for Academia

One of the limitations of this study lies in the fact that the cross-validation step relies on finding overlapping positions collected by both GPS receivers on the same MV. If the localization error of one GPS receiver is too high and there are not enough overlapping points with the other receiver, cross-validation is simply not feasible. This might be the case if one GPS receiver has estimated the vehicle's position totally wrong. More studies are needed to overcome this limitation.

As mentioned, the main contribution of RQ3 is an approach to identifying the most accurate GPS receiver and studying its efficiency to be used as an input for the existing map-matching algorithms. Therefore, the map-matching algorithm is not part of our contribution, and we used the most popular tool (i.e., the QGIS-Plug-in Offline-MapMatching [103][110]). However, an important limitation of this map-matching tool is that it is based on post-processing and offline QGIS, which is not instantaneous. Thus, a method that runs in real time is needed.

Moreover, the MV localization accuracy can be increased by considering a vehicle motion model.

C Implications for Practitioners

This study proposes how car producers and the use of mounted sensors, even low-cost ones, on MVs can help to provide the required traffic data for ITMSs.

D Threats to Validity

- Objectivity: In order to enhance the objectivity of RQ3 by mitigating the possibility of biases or distortions in the research, instead of using a simulator or laboratory data (which might omit the complexity of real-life situations), we decided to collect data from real traffic (which reflect some real-life scenarios) and use them to assess our developed algorithms. For example, one of the scenarios studied represented GPS receiver noise in localizing an MV, which is the case in the presence of tall buildings in a metropolitan area.

- Reliability: We ran experiments on real traffic data collected by three vehicles in several scenarios to estimate the accuracy of our proposed methodologies and to enhance the reliability and neutrality of the results. Furthermore, we evaluated our findings both by visualizing our results on a map and by performing a numerical analysis.

- External validity: Collecting data from real traffic helped us to deal with real traffic situations, by considering several popular scenarios. Also, we used a low-cost GPS receiver in order to enhance the generalizability of our proposed approach as it is a popular sensor on MVs.

6.4 Multiple Sensor Fusion - RQ4

The last RQ addressed in this Ph.D. thesis aimed to use the sensor fusion technique to improve the accuracy of the target vehicle's geolocation estimation than what can be obtained by using data from only one MV.

A Comparison with Related Work

Misestimating the traffic data or blind sensing spots are critical challenges in ensuring the safety and performance of ITMSs. Increasing the number of sensors and advancing their technology may yield improvements; however, it will increase the complexity and costs [81]. Multiple sensor data fusion is often applied to deal with these challenges. As presented in Section 3.4, multiple sensor fusion methodologies in the literature are used mainly for three purposes. The first purpose is to fuse the data collected by stationary sensors (e.g., [26]), a method that is widely used in current ITMSs. The second purpose is to fuse sensors mounted on the same MV (e.g., [67]). This purpose is mostly applied for enhancing self-awareness and automated driving. The third group of studies utilize MV's sensors and communication to share and fuse the estimated data via several MVs in real time (e.g., [81]) [140]. The main purpose of our approach to addressing RQ4 was to enhance the accuracy of the target vehicle's geolocation estimation based on sharing and integrating estimated geolocations between two mobile MVs equipped with low-cost sensors. Moreover, we studied the impact of the uncertainty of low-cost sensors, about which was not much known. Although, some research has been carried out in this regard (e.g., [81]), more studies are needed to understand the impact of sensor uncertainty caused by low-cost and mobile sensors and to understand the impact of fusing the estimated target vehicle's geolocations by multiple and mobile MVs with different views in real traffic.

B Implications for Academia

The main idea of this RQ was to fuse the estimated traffic data by two MVs. The basic requirement of multiple sensor fusion is sharing and transferring traffic data, which was beyond the scope of our study. Sharing and transferring traffic data requires extra studies on privacy, safety, security, networking, communication, etc.

In addition, continued efforts are needed to go beyond V2I communication to fuse the data collected by several road infrastructures based on a combination of Infrastructure-to-Infrastructure (I2I) and V2V communication and a dynamic model of a large area (e.g., city level).

C Implications for Practitioners

Traffic managers and car producers should be aware of each other's needs and utilize each other's strengths. In addition, this study is an initial step to fusing the estimated traffic data via multiple MVs, which can be beneficial for enhancing the estimation accuracy. These data can be utilized to generate a dynamic model of the traffic (e.g., digital twins), which can be used to enhance vehicle motion planning and TMSs in terms of safety and efficiency.

D Threats to Validity

- Objectivity: In order to enhance the objectivity of RQ4 by mitigating the possibility of biases or distortions in the research, instead of using a simulator or laboratory data (which might omit the complexities of real-life situations), we decided to collect data from real traffic (which reflect real-life conditions and sensor estimation uncertainty) and use them to assess our developed algorithms.

- Reliability: We ran experiments on multiple scenarios from real traffic to estimate the accuracy of our proposed methodologies and to enhance the reliability and neutrality of the results. Moreover, we proposed a mathematical model and developed a system to evaluate our proposed approach. In addition, we evaluated our findings both by visualizing our results in charts and by performing a numerical analysis.

- External validity: In order to enhance the external validity of RQ4, we need to consider the generalizability of the findings. Collecting data from real traffic helped us to deal with real traffic situations. Also, we used a low-cost monocular camera with a built-in GPS receiver in order to enhance the generalizability of our proposed approach as it is a popular sensor for both ITMSs and MVs.

6.5 Ethical Issues

As presented in Section 4.3, data collection from real traffic was done in two rounds. The first round of data was collected from Trondheim, Norway. As the traffic data was collected with a camera in a metropolitan area, it might include some people's faces or vehicle plates. Therefore, because of the General Data Protection Regulation (GDPR), we submitted our application to the Norwegian Centre for Research Data (NSD). The NSD's assessment was that "this project will not process data that can directly or indirectly identify individual persons". Therefore, the project did not need approval from the NSD.

Chapter 7

Conclusions and Future Work

The overall research aim of this Ph.D. thesis was to investigate how MVs and the sensing abilities of their mounted low-cost sensors can be employed to estimate the required traffic data of ITMSs in mixed traffic.

This study adopted the DSR approach. The work was grounded in the SLR on utilizing MVs in ITMSs in both pure MV traffic and mixed traffic. The SLR revealed the considered factors (goals) and proposed methodologies and identified the most important traffic data types to be collected according to the state of the art. Then, several algorithms were developed to assess our proposed methodologies for estimating the target vehicle's number, type, relative position, distance, speed, lane, and localization (including self-localization, target vehicle localization, and multiple sensor fusion for enhancing the accuracy of target vehicle localization). Each algorithm was developed through multiple design iterations based mainly on AI, image processing, map-matching, and multiple sensor fusion techniques. Several case studies were performed to evaluate the algorithms developed by considering real traffic data. The scientific work has been published in or submitted to peer-reviewed journals and conference proceedings, and six of these publications were included in this thesis.

The research presented in this thesis is highly interdisciplinary and provides a critical view of how MVs can be employed as mobile sensors to dynamically collect the required traffic data of ITMSs. This is done in four main contributions, as follows:

C1. We generated new knowledge by systematically reviewing, summarizing, and conceptualizing the state of the art in managing an intersection intelligently with a focus on (1) signalized intersections for both pure MVs (i.e., AVs) and mixed traffic and (2) unsignalized intersections if the traffic includes pure MVs.

C2. We proposed new methodologies and algorithms to estimate the required traffic data of the target vehicle via a single MV equipped with a low-cost monocular camera in mixed traffic.

C3. We Proposed a new methodology and algorithms to enhance the self-localization accuracy of an MV by using low-cost GPS receivers.

C4. We proposed a new methodology and algorithms to fuse the estimated geolocations of the observed target vehicle via two MVs equipped with a low-cost monocular camera by considering sensor estimation uncertainty in mixed traffic.

While acknowledging the potential for employing MVs in enhancing an ITMS's traffic awareness, this research has also highlighted the complexity of fully realizing this potential. Transitioning toward an effective ITMS by considering mixed traffic and utilizing low-cost sensors, such as a monocular camera and GPS receivers mounted on MVs, is complex and multifaceted. To develop a more holistic understanding of MVs and ITMSs requires researchers and practitioners to communicate and share their needs and future vision. The main result of this Ph.D. thesis is the feasibility of using an MV equipped with a low-cost monocular camera with a built-in GPS receiver to collect some of the most important traffic data types needed by an ITMS to model the traffic scene and make smart decisions. This result confirms that a digital twin of the traffic scene in mixed traffic might in the future be generated based on the data collected by MVs.

7.1 Avenues for Future Research

The most important challenge related to our system is that we used waterfall-based estimations, which means that the estimated value in a previous step was used as an input for the next estimation. In other words, the vehicle's type was used for estimating the distance, and the distance was used to determine the speed and geolocation. Thus, a mistake in the early stage of the estimations will cascade down the rest of the estimation process. Therefore, future studies should make these data estimation processes more independent from each other or accurate enough to mitigate their negative effect on the estimation accuracy of the next step.

To enhance the generalizability of this study, further investigations and experiments are recommended that consider various potential scenarios (e.g., intersections, roundabouts, highways, freeways, lane changing, giving priority to emergency vehicles, considering pedestrians and cyclists) of various road traffic scenes with different traffic densities and speeds (i.e., low, moderate, and high) in different weather/brightness conditions.

7.2 Final Remarks

The major potential future use of our proposed approach is to make a dynamic digital model of the traffic scene (i.e., a digital twin). This model could provide a global perception of the real-world traffic scene, which would be useful for ITMSs to make smart decisions and to predict the traffic effectively. Therefore, in this study, we have proposed new methodologies and developed new algorithms for estimating the required traffic data by ITMSs. To evaluate them, we ran experiments on footage (including GPS data) collected by mobile MVs from real traffic.

Bibliography

[1] Shawn Adams and Matthew Rutherford. "Towards decentralized waypoint negotiation". In: *Workshops at the Twenty-Sixth AAAI Conference on Artificial Intelligence*. 2012.

[2] M Nadeem Ahangar et al. "A survey of autonomous vehicles: enabling communication technologies and challenges". In: *Sensors* 21.3 (2021), p. 706.

[3] Nima Alam, Asghar Tabatabaei Balaei and Andrew G Dempster. "Positioning enhancement with double differencing and DSRC". In: *Proceedings of the 23rd International Technical Meeting of The Satellite Division of the Institute of Navigation (ION GNSS 2010)*. 2010, pp. 1210–1218.

[4] Bram Alefs and David Schreiber. "Accurate speed measurement from vehicle trajectories using adaboost detection and robust template tracking". In: *2007 IEEE Intelligent Transportation Systems Conference*. IEEE. 2007, pp. 405–412.

[5] Tareq Alhmiedat, Ghassan Samara and Amer O Abu Salem. "An indoor fingerprinting localization approach for ZigBee wireless sensor networks". In: *arXiv preprint arXiv:1308.1809* (2013).

[6] Nawaf O Alsrehin, Ahmad F Klaib and Aws Magableh. "Intelligent transportation and control systems using data mining and machine learning techniques: A comprehensive study". In: *IEEE Access* 7 (2019), pp. 49830–49857.

[7] Hussein Ali Ameen et al. "A review on vehicle to vehicle communication system applications". In: *Indonesian Journal of Electrical Engineering and Computer Science* 18.1 (2020), pp. 188–198.

[8] Ekalux Ua-areemitr, Agachai Sumalee and William HK Lam. “Low-cost road traffic state estimation system using time-spatial image processing”. In: *IEEE Intelligent Transportation Systems Magazine* 11.3 (2019), pp. 69–79.

[9] Fabio Arena and Giovanni Pau. “An overview of vehicular communications”. In: *Future Internet* 11.2 (2019), p. 27.

[10] Lubna Aziz et al. “Exploring deep learning-based architecture, strategies, applications and current trends in generic object detection: A comprehensive review”. In: *IEEE Access* 8 (2020), pp. 170461–170495.

[11] Koyel Banerjee et al. “Online camera lidar fusion and object detection on hybrid data for autonomous driving”. In: *2018 IEEE Intelligent Vehicles Symposium (IV)*. IEEE. 2018, pp. 1632–1638.

[12] Barbara Barbagli et al. “Acoustic sensor network for vehicle traffic monitoring”. In: *Proceedings of the 1st International Conference on Advances in Vehicular Systems, Technologies and Applications*. 2012, pp. 24–29.

[13] *Barents ITS - current and existing ITS status and suggestions for ITS pilots in the Barents Region*. `https://its-norway.no/wp-content/uploads/2021/04/Barents-ITS-report-2020-1.pdf`. 2020.

[14] Vojtěch Bartl and Adam Herout. “OptInOpt: Dual optimization for automatic camera calibration by multi-target observations”. In: *2019 16th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*. IEEE. 2019, pp. 1–8.

[15] D Bell, W Xiao and P James. “Accurate vehicle speed estimation from monocular camera footage.” In: *ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences* 5.2 (2020).

[16] Luis Conde Bento, Ricardo Parafita and Urbano Nunes. “Inter-vehicle sensor fusion for accurate vehicle localization supported by v2v and v2i communications”. In: *2012 15th International IEEE Conference on Intelligent Transportation Systems*. IEEE. 2012, pp. 907–914.

[17] Rein van den Boomgaard. *The Pinhole Camera Matrix*. URL: `https://staff.fnwi.uva.nl/r.vandenboomgaard/IPCV20162017/LectureNotes/CV/PinholeCamera/PinholeCamera.html`.

[18] Sean Campbell et al. “Sensor technology in autonomous vehicles: A review”. In: *2018 29th Irish Signals and Systems Conference (ISSC)*. IEEE. 2018, pp. 1–4.

[19] Georges Challita et al. "An application of V2V communications: Cooperation of vehicles for a better car tracking using GPS and vision systems". In: *2009 IEEE Vehicular Networking Conference (VNC)*. IEEE. 2009, pp. 1–6.

[20] Abdellah Chehri, Nordine Quadar and Rachid Saadane. "Survey on localization methods for autonomous vehicles in smart cities". In: *Proceedings of the 4th International Conference on Smart City Applications*. 2019, pp. 1–6.

[21] Lei Chen and Cristofer Englund. "Cooperative intersection management: A survey". In: *IEEE Transactions on Intelligent Transportation Systems* 17.2 (2015), pp. 570–586.

[22] Mike Y Chen et al. "Practical metropolitan-scale positioning for gsm phones". In: *International Conference on Ubiquitous Computing*. Springer. 2006, pp. 225–242.

[23] Genyuan Cheng et al. "Real-time detection of vehicle speed based on video image". In: *2020 12th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA)*. IEEE. 2020, pp. 313–317.

[24] Zbigniew Czapla. "Vehicle speed estimation with the use of gradient-based image conversion into binary form". In: *2017 Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA)*. IEEE. 2017, pp. 213–216.

[25] Mattias Dahl and Saleh Javadi. "Analytical modeling for a video-based vehicle speed measurement framework". In: *Sensors* 20.1 (2020), p. 160.

[26] Danhui Dan, Liangfu Ge and Xingfei Yan. "Identification of moving loads based on the information fusion of weigh-in-motion system and multiple camera machine vision". In: *Measurement* 144 (2019), pp. 155–166.

[27] Alfred Daniel et al. "Cooperative intelligence of vehicles for intelligent transportation systems (ITS)". In: *Wireless Personal Communications* 87.2 (2016), pp. 461–484.

[28] Lijun Ding and Ardeshir Goshtasby. "On the Canny edge detector". In: *Pattern Recognition* 34.3 (2001), pp. 721–725.

[29] *Directions API*. URL: https://docs.mapbox.com/playground/directions/.

[30] *Directive 2010/40/EU of the European parliament and of the council of 7 July 2010*. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/%5C?uri=CELEX:32010L0040&rid=1.

[31] Soufiene Djahel et al. "Toward V2I communication technology-based solution for reducing road traffic congestion in smart cities". In: *2015 International Symposium on Networks, computers and communications (ISNCC)*. IEEE. 2015, pp. 1–6.

[32] Wilfried Elmenreich. "An introduction to sensor fusion". In: *Vienna University of Technology, Austria* 502 (2002), pp. 1–28.

[33] *Euclidean distance*. URL: `https://en.wikipedia.org/wiki/Euclidean_distance`.

[34] Lola Evans. *Intelligent transport system pilot destined for China*. URL: `https://www.bignewsnetwork.com/news/262123715/intelligent-transport-system-pilot-destined-for-china`.

[35] Mahmoud Famouri, Zohreh Azimifar and Alexander Wong. "A novel motion plane-based approach to vehicle speed estimation". In: *IEEE Transactions on Intelligent Transportation Systems* 20.4 (2018), pp. 1237–1246.

[36] Daniel Firouzimagham, Mohammad Sabouri and Fatemeh Adhami. "An IoT-based system: Big urban traffic data mining through airborne pollutant gases analysis". In: *arXiv preprint arXiv:2002.06374* (2020).

[37] Cheng-Yang Fu et al. "Dssd: Deconvolutional single shot detector". In: *arXiv preprint arXiv:1701.06659* (2017).

[38] Charles Galamhos, Jose Matas and Josef Kittler. "Progressive probabilistic Hough transform for line detection". In: *Proceedings. 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No PR00149)*. Vol. 1. IEEE. 1999, pp. 554–560.

[39] Mayank Garg and Sahil Goel. "Real-time license plate recognition and speed estimation from video sequences". In: *ITSI Transactions on Electrical and Electronics Engineering* 1.5 (2013), pp. 1–4.

[40] Jared L Giesbrecht et al. "A vision-based robotic follower vehicle". In: *Unmanned Systems Technology XI*. Vol. 7332. International Society for Optics and Photonics. 2009, 73321O.

[41] Chaim Ginzburg, Amit Raphael and Daphna Weinshall. "A cheap system for vehicle speed detection". In: *arXiv preprint arXiv:1501.06751* (2015).

[42] Ross Girshick. "Fast r-cnn". In: *Proceedings of the IEEE International Conference on Computer Vision*. 2015, pp. 1440–1448.

[43] Ross Girshick et al. "Rich feature hierarchies for accurate object detection and semantic segmentation". In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2014, pp. 580–587.

[44] *Google earth*. URL: https://www.google.com/earth/download/gep/agree.html.

[45] Juan Antonio Guerrero-Ibanez, Sherali Zeadally and Juan Contreras-Castillo. "Integration challenges of intelligent transportation systems with connected vehicle, cloud computing, and internet of things technologies". In: *IEEE Wireless Communications* 22.6 (2015), pp. 122–128.

[46] Juan Guerrero-Ibáñez, Sherali Zeadally and Juan Contreras-Castillo. "Sensor technologies for intelligent transportation systems". In: *Sensors* 18.4 (2018), p. 1212.

[47] Alexander AS Gunawan, Deasy Aprilia Tanjung and Fergyanto E Gunawan. "Detection of vehicle position and speed using camera calibration and image projection methods". In: *Procedia Computer Science* 157 (2019), pp. 255–265.

[48] Qiangqiang Guo, Li Li and Xuegang Jeff Ban. "Urban traffic signal control with connected and automated vehicles: A survey". In: *Transportation Research Part C: Emerging Technologies* 101 (2019), pp. 313–334.

[49] Pratishtha Gupta, GN Purohit and Manisha Rathore. "Estimating speed of vehicle using centroid method in MATLAB". In: *International Journal of Computer Applications* 102.14 (2014).

[50] Samta Gupta and Susmita Ghosh Mazumdar. "Sobel edge detection algorithm". In: *International Journal of Computer Science and Management Research* 2.2 (2013), pp. 1578–1583.

[51] Mathias Haberjahn. "Multilevel datenfusion konkurrierender sensoren in der fahrzeugumfelderfassung". In: (2013).

[52] J Han et al. "Vehicle distance estimation using a mono-camera for FCW/AEB systems". In: *International Journal of Automotive Technology* 17.3 (2016), pp. 483–491.

[53] Wenxue He et al. "Overview of V2V and V2I wireless communication for cooperative vehicle infrastructure systems". In: *2019 IEEE 4th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*. Vol. 1. IEEE. 2019, pp. 127–134.

[54] *HERO7 Field of View (FOV) Information*. URL: https://gopro.com/help/articles/question_answer/hero7-field-of-view-fov-information?sf96748270=1.

[55] Alan Hevner and Samir Chatterjee. "Design science research in information systems". In: *Design Research in Information Systems*. Springer, 2010, pp. 9–22.

[56] *How do I construct a straight line through data points? Best-fit lines.* URL: https://serc.carleton.edu/mathyouneed/graphing/bestfit.html.

[57] Jiung-yao Huang and Chung-Hsien Tsai. “Improve GPS positioning accuracy with context awareness”. In: *2008 First IEEE International Conference on Ubi-Media Computing*. IEEE. 2008, pp. 94–99.

[58] Tingting Huang. “Traffic speed estimation from surveillance video data”. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*. 2018, pp. 161–165.

[59] *INRIX: Congestion costs each American nearly 100 hours, $1,400 a year*. URL: https://inrix.com/press-releases/2019-traffic-scorecard-us/.

[60] *Interpolation methods*. URL: https://www.mathworks.com/help/curvefit/interpolation-methods.html.

[61] Md Islam, Jong-Myon Kim et al. “An effective approach to improving low-cost GPS positioning accuracy in real-time navigation”. In: *The Scientific World Journal* 2014 (2014).

[62] Saleh Javadi, Mattias Dahl and Mats I Pettersson. “Vehicle speed measurement model for video-based systems”. In: *Computers & Electrical Engineering* 76 (2019), pp. 238–248.

[63] Dongyao Jia and Dong Ngoduy. “Enhanced cooperative car-following traffic model with the combination of V2V and V2I communication”. In: *Transportation Research Part B: Methodological* 90 (2016), pp. 172–191.

[64] Supriya Kamoji et al. “Image processing based vehicle identification and speed measurement”. In: *2020 International Conference on Inventive Computation Technologies (ICICT)*. IEEE. 2020, pp. 523–527.

[65] C Christopher Kellum. “Basic feasibility of gps positioning without carrier-phase measurements as a relative position sensor between two vehicles”. In: *Proceedings of the 2005 National Technical Meeting of the Institute of Navigation*. 2005, pp. 903–910.

[66] Business Analyst Kiisa Wiegand. *Challenges of the day-today operation of a traffic monitoring program*. Georgia Department of Transportation, 2016.

[67] H Kim and I Lee. “Localization of a car based on multi-sensor fusion”. In: *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences* 42 (2018), pp. 247–250.

[68] Barbara Kitchenham and Stuart Charters. "Guidelines for performing systematic literature reviews in software engineering". In: (2007).

[69] Barbara Kitchenham et al. "Systematic literature reviews in software engineering–a systematic literature review". In: *Information and Software Technology* 51.1 (2009), pp. 7–15.

[70] Tao Kong et al. "Ron: Reverse connection with objectness prior networks for object detection". In: *Proceedings of the IEEE Conference on Computer Vision and Pattern recognition*. 2017, pp. 5936–5944.

[71] Geesara Kulathunga, Aleksandr Buyval and Aleksandr Klimchik. "Multi-camera fusion in apollo software distribution". In: *IFAC-PapersOnLine* 52.8 (2019), pp. 49–54.

[72] Amit Kumar et al. "A semi-automatic 2D solution for vehicle speed estimation from monocular videos". In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*. 2018, pp. 137–144.

[73] Jitul Kumar Laphong, Himanta Dihingia and Suren Borah. "Advanced road traffic management system". In: *International Journal of Innovative Science and Research Technology* 4 (2019), pp. 168–169.

[74] Minho Lee et al. "A new lane following method based on deep learning for automated vehicles using surround view images". In: *Journal of Ambient Intelligence and Humanized Computing* (2019), pp. 1–14.

[75] Lone-Eirin Lervåg, Tomas Levin and Ane Dalsnes Storsæter. "Using C-ITS on road status information in winter maintenance operations". In: (2016).

[76] Guofa Li et al. "Detection of road objects with small appearance in images for autonomous driving in various traffic situations using a deep learning based approach". In: *IEEE Access* 8 (2020), pp. 211164–211172.

[77] Jun Li et al. "Deep neural network for structural prediction and lane detection in traffic scene". In: *IEEE Transactions on Neural Networks and Learning Systems* 28.3 (2016), pp. 690–703.

[78] Huei-Yung Lin, Kun-Jhih Li and Chia-Hong Chang. "Vehicle speed detection from a single motion blurred image". In: *Image and Vision Computing* 26.10 (2008), pp. 1327–1337.

[79] Tsung-Yi Lin et al. "Microsoft coco: Common objects in context". In: *European Conference on Computer Vision*. Springer. 2014, pp. 740–755.

[80] Chenghuan Liu et al. "A vision-based pipeline for vehicle counting, speed estimation, and classification". In: *IEEE Transactions on Intelligent Transportation Systems* (2020).

[81] Hansi Liu et al. "FusionEye: Perception sharing for connected vehicles and its bandwidth-accuracy trade-offs". In: *2019 16th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)*. IEEE. 2019, pp. 1–9.

[82] Lianxia Xi1 Quan Liu and Minghua Li1 Zhong Liu. "Map matching algorithm and its application". In: *Proceedings on Intelligent Systems and Knowledge Engineering (ISKE2007)* (2007).

[83] Wei Liu et al. "Ssd: Single shot multibox detector". In: *European Conference on Computer Vision*. Springer. 2016, pp. 21–37.

[84] David Fernández Llorca, Antonio Hernández Martınez and Iván Garcıa Daza. "Vision-based Vehicle Speed Estimation for ITS: A Survey". In: *arXiv preprint arXiv:2101.06159* (2021).

[85] David Fernández Llorca et al. "Two-camera based accurate vehicle speed measurement using average speed at a fixed point". In: *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE. 2016, pp. 2533–2538.

[86] *Local tangent plane coordinates*. URL: `https://en.wikipedia.org/wiki/Local_tangent_plane_coordinates#Local_north,_east,_down_(NED)_coordinates`.

[87] J Lopes et al. "Traffic and mobility data collection for real-time applications". In: *13th International IEEE Conference on Intelligent Transportation Systems*. IEEE. 2010, pp. 216–223.

[88] Nan Ma et al. "An all-weather lane detection system based on simulation interaction platform". In: *IEEE Access* 8 (2018), pp. 46121–46130.

[89] Kodeeswari Manoharan and Philemon Daniel. "Image processing-based framework for continuous lane recognition in mountainous roads for driver assistance system". In: *Journal of Electronic Imaging* 26.6 (2017), p. 063011.

[90] *Map matching tool*. URL: `https://www.waves.intec.ugent.be/exposure-tool/map-matching-tool`.

[91] Jiri Matas, Charles Galambos and Josef Kittler. "Robust detection of lines using the progressive probabilistic hough transform". In: *Computer Vision and Image Understanding* 78.1 (2000), pp. 119–137.

[92] Hiram McCall and Ching-Fang Lin. *Micro integrated global positioning system/inertial measurement unit system*. US Patent App. 09/911,571. Jan. 2002.

[93] Kristi Miller et al. *Data management life cycle, final report*. Tech. rep. Texas A&M Transportation Institute, 2018.

[94] Md Golam Moazzam, Mohammad Reduanul Haque and Mohammad Shorif Uddin. "Image-based vehicle speed estimation". In: *Journal of Computer and Communications* 7.6 (2019), pp. 1–5.

[95] *Movable Type Scripts, Calculate distance, bearing and more between Latitude/Longitude points*. URL: `https://www.movable-type.co.uk/scripts/latlong.html`.

[96] *Moving average*. URL: `https://en.wikipedia.org/wiki/Moving_average`.

[97] MD Enjat Munajat, Dwi H Widyantoro and Rinaldi Munir. "Vehicle detection and tracking based on corner and lines adjacent detection features". In: *2016 2nd International Conference on Science in Information Technology (ICSITech)*. IEEE. 2016, pp. 244–249.

[98] I Murashov and Y Stroganov. "Method of determining vehicle speed according to video stream data". In: *Journal of Physics: Conference Series*. Vol. 1419. 1. IOP Publishing. 2019, p. 012020.

[99] Muhammad Jefri Muril et al. "A review on deep learning and nondeep learning approach for lane detection system". In: *2020 IEEE 8th Conference on Systems, Process and Control (ICSPC)*. IEEE. 2020, pp. 162–166.

[100] Mahyar Najibi, Mohammad Rastegari and Larry S Davis. "G-cnn: an iterative grid based object detector". In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016, pp. 2369–2377.

[101] Sandipann P Narote et al. "A review of recent advances in lane detection and departure warning system". In: *Pattern Recognition* 73 (2018), pp. 216–234.

[102] Briony J Oates. *Researching information systems and computing*. Sage, 2006.

[103] *Offline-MapMatching*. URL: `https://github.com/jagodki/Offline-MapMatching`.

[104] Chitu Okoli and Kira Schabram. "A guide to conducting a systematic literature review of information systems research". In: (2010).

[105] Panos Papadimitratos et al. "Vehicular communication systems: Enabling technologies, applications, and future outlook on intelligent transportation". In: *IEEE communications magazine* 47.11 (2009), pp. 84–95.

[106] *Parallel (geometry)*. URL: `https://en.wikipedia.org/wiki/Parallel_(geometry)`.

[107] Oliver Pink and Britta Hummel. "A statistical approach to map matching using road network geometry, topology and vehicular motion constraints". In: *2008 11th International IEEE Conference on Intelligent Transportation Systems*. IEEE. 2008, pp. 862–867.

[108] *Point-line distance–2-dimensional*. URL: `https://mathworld.wolfram.com/Point-LineDistance2-Dimensional.html`.

[109] Fabian de Ponte Müller. "Survey on ranging sensors and cooperative techniques for relative positioning of vehicles". In: *Sensors* 17.2 (2017), p. 271.

[110] *QGIS*. URL: `https://qgis.org/en/site/`.

[111] Ahmed El-Rabbany. *Introduction to GPS: the global positioning system*. Artech house, 2002.

[112] Craig B Rafter, Bani Anvari and Simon Box. "Traffic responsive intersection control algorithm using GPS data". In: *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE. 2017, pp. 1–6.

[113] HA Rahim et al. "Vehicle velocity estimation for traffic surveillance system". In: *World Academy of Science, Engineering and Technology (WASET)* (2010), p. 772.

[114] Alok Raj, J Ajith Kumar and Prateek Bansal. "A multicriteria decision making approach to study barriers to the adoption of autonomous vehicles". In: *Transportation Research Part A: Policy and Practice* 133 (2020), pp. 122–137.

[115] YG Anil Rao et al. "Real-time speed estimation of vehicles from uncalibrated view-independent traffic cameras". In: *TENCON 2015-2015 IEEE Region 10 Conference*. IEEE. 2015, pp. 1–6.

[116] Joseph Redmon. *PyTorch-YOLOv3-kitti*. URL: `https://github.com/packyan/PyTorch-YOLOv3-kitti`.

[117] Joseph Redmon and Ali Farhadi. "Yolov3: An incremental improvement". In: *arXiv preprint arXiv:1804.02767* (2018).

[118] Joseph Redmon et al. "You only look once: Unified, real-time object detection". In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016, pp. 779–788.

[119] Shaoqing Ren et al. "Faster r-cnn: Towards real-time object detection with region proposal networks". In: *arXiv preprint arXiv:1506.01497* (2015).

[120] Jason Rife. “Collaborative vision-integrated pseudorange error removal: Team-estimated differential GNSS corrections with no stationary reference receiver”. In: *IEEE Transactions on Intelligent Transportation Systems* 13.1 (2011), pp. 15–24.

[121] Jackeline Rios-Torres and Andreas A Malikopoulos. “A survey on the coordination of connected and automated vehicles at intersections and merging at highway on-ramps”. In: *IEEE Transactions on Intelligent Transportation Systems* 18.5 (2016), pp. 1066–1077.

[122] *SAE International Releases Updated Visual Chart for Its “Levels of Driving Automation” Standard for Self-Driving Vehicles*. URL: https://www.sae.org/news/press-room/2018/12/sae-international-releases-updated-visual-chart-for-its-%5C%E2%5C%80%5C%9Clevels-of-driving-automation%5C%E2%5C%80%5C%9D-standard-for-self-driving-vehicles.

[123] Aramesh Seif, Mohammad Mohammadpour Salut and Muhammad Nadzir Marsono. “A hardware architecture of Prewitt edge detection”. In: *2010 IEEE Conference on Sustainable Utilization and Development in Engineering and Technology*. IEEE. 2010, pp. 99–101.

[124] Dolley Shukla and Ekta Patel. “Speed determination of moving vehicles using Lucas-Kanade Algorithm”. In: *International Journal of Computer Applications Technology and Research* 2.1 (2013), pp. 32–36.

[125] *Slope*. URL: https://en.wikipedia.org/wiki/Slope.

[126] Akash Sonth, Harshavardhan Settibhaktini and Ankush Jahagirdar. “Vehicle speed determination and license plate localization from monocular video streams”. In: *Proceedings of 3rd International Conference on Computer Vision and Image Processing*. Springer. 2020, pp. 267–277.

[127] *Speed*. URL: https://en.wikipedia.org/wiki/Speed.

[128] *Spherical coordinate system*. URL: https://en.wikipedia.org/wiki/Spherical%5C%5C_coordinate%5C%5C_system..

[129] Gideon P Stein, Ofer Mano and Amnon Shashua. “Vision-based ACC with a single camera: bounds on range and range rate accuracy”. In: *IEEE IV2003 Intelligent Vehicles Symposium. Proceedings (Cat. No. 03TH8683)*. IEEE. 2003, pp. 120–125.

[130] *Sum of angles of a triangle*. URL: https://en.wikipedia.org/wiki/Sum_of_angles_of_a_triangle.

[131] Jigang Tang, Songbin Li and Peng Liu. “A review of lane detection methods based on deep learning”. In: *Pattern Recognition* (2020), p. 107623.

[132] *Thales theorem and figure*. URL: https://sites.math.washington.edu/~king/coursedir/m444a02/class/10-21-thales.html.

[133] Yan Tian et al. “Lane marking detection via deep convolutional neural network”. In: *Neurocomputing* 280 (2018), pp. 46–55.

[134] Liu Ting et al. “Ship detection algorithm based on improved YOLO V5”. In: *2021 6th International Conference on Automation, Control and Robotics Engineering (CACRE)*. IEEE. 2021, pp. 483–487.

[135] *Trigonometry*. URL: https://en.wikipedia.org/wiki/Trigonometry.

[136] Elnaz Vakili, Maryam Shoaran and Mohammad R Sarmadi. “Single–camera vehicle speed measurement using the geometry of the imaging system”. In: *Multimedia Tools and Applications* (2020), pp. 1–21.

[137] CH Veerendra and KBR Prasad Reddy. “Lane detection and lane change warning as advanced driver assistance systems using computer vision”. In: *International Journal of Engineering and Advanced Technology* 9.2 (2019), pp. 5540–5543.

[138] V Vijayaraghavan and J Rian Leevinson. “Intelligent traffic management systems for next generation IoV in smart city scenario”. In: *Connected Vehicles in the Internet of Things*. Springer, 2020, pp. 123–141.

[139] Derui Wang et al. “Daedalus: Breaking nonmaximum suppression in object detection via adversarial examples”. In: *IEEE Transactions on Cybernetics* (2021).

[140] Zhangjing Wang, Yu Wu and Qingqing Niu. “Multi-sensor fusion in automated driving: A survey”. In: *IEEE Access* 8 (2019), pp. 2847–2868.

[141] TH Witte and AM Wilson. “Accuracy of WAAS-enabled GPS for the determination of position and speed over ground”. In: *Journal of Biomechanics* 38.8 (2005), pp. 1717–1722.

[142] Xiongwei Wu, Doyen Sahoo and Steven CH Hoi. “Recent advances in deep learning for object detection”. In: *Neurocomputing* 396 (2020), pp. 39–64.

[143] Zhouhao Wu et al. “Map matching based on multi-layer road index”. In: *Transportation Research Part C: Emerging Technologies* 118 (2020), p. 102651.

[144] De Jong Yeong et al. “Sensor and sensor fusion technology in autonomous vehicles: a review”. In: *Sensors* 21.6 (2021), p. 2140.

[145] *YOLO: Real-time object detection*. URL: https://pjreddie.com/darknet/yolo/.

[146] Sherali Zeadally, J Guerrero and J Contreras. "A tutorial survey on vehicle-to-vehicle communications". In: *Telecommunication Systems* 73.3 (2020), pp. 469–489.

[147] Hanbo Zhang et al. "Visual manipulation relationship network for autonomous robotics". In: *2018 IEEE-RAS 18th International Conference on Humanoid Robots (Humanoids)*. IEEE. 2018, pp. 118–125.

[148] Zixuan Zhang, Shengrui Zhang and Shuaiyang Jiao. "A vehicle lane-changing model based on connected vehicles". In: *CICTP 2020*. 2020, pp. 3027–3038.

[149] He Zhiwei, Liu Yuanyuan and Ye Xueyi. "Models of vehicle speeds measurement with a single camera". In: *2007 International Conference on Computational Intelligence and Security Workshops (CISW 2007)*. IEEE. 2007, pp. 283–286.

Part II

Research Papers

Paper A:

Intelligent Intersection Management Systems Considering Autonomous Vehicles: A Systematic Literature Review

Elnaz Namazi, Jingyue Li, and Chaoru Lu

IEEE Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.Doi Number

Intelligent Intersection Management Systems Considering Autonomous Vehicles: A Systematic Literature Review

Elnaz Namazi[1], (Member, IEEE), Jingyue Li[1], (Member, IEEE), and Chaoru Lu[2]

[1]Department of Computer Science, Norwegian University of Science and Technology (NTNU), Trondheim, Norway

[2]Department of Civil and Environmental Engineering, Norwegian University of Science and Technology (NTNU), Trondheim, Norway

Corresponding authors: Jingyue Li (jingyue.li@ntnu.no) and Chaoru Lu (chaoru.lu@ntnu.no)

ABSTRACT Over the past several decades, the development of technologies and production of autonomous vehicles have enhanced the need for intelligent intersection management systems. Subsequently, growing interest in studying the traffic management of autonomous vehicles at intersections has been evident, which indicates a critical need to conduct a systematic literature review on this topic. This article offers a systematic review of the proposed methodologies for intelligent intersection management systems and presents the remaining research gaps and possible future research approaches. We consider both pure autonomous vehicle traffic and mixed traffic at four-way signalized and unsignalized intersection(s). We searched for articles published between January 2008 and May 10, 2019, and identified 105 primary studies. We applied the thematic analysis method to analyze the extracted data, which led to the identification of four main classes of methodologies, namely rule-based, optimization, hybrid, and machine learning methods. We also compared how well the methods satisfy their goals, namely efficiency, safety, ecology, and passenger comfort. This analysis allowed us to determine the primary challenges of the presented methodologies and propose new approaches in this area.

INDEX TERMS Autonomous vehicle, intelligent intersection management system, mixed traffic, vehicle-to-infrastructure (V2I) communication, vehicle-to-vehicle (V2V) communication.

I. INTRODUCTION

The rapid population growth and the attendant increase in vehicle numbers over the last few decades have caused traffic congestion worldwide, with traffic congestion forecast to increase by 60% by 2030 [1]. Because intersections significantly impact the efficiency of traffic management systems in urban areas, this study focuses on intelligent traffic management systems at intersections.

It has previously been observed that traditional traffic lights are inefficient when traffic volumes are high [2]. Moreover, research has shown that intersections play a critical role in collision numbers and traffic delays in urban areas [3]. For instance, Franke et al. mentioned that more than 33% of traffic accidents resulting in injury occur at urban intersections [4]. Likewise, in the United States and Europe, more than 40% of reported traffic accidents occur at intersections [5]. Traffic delays, which affect congestion costs, is another critical matter in traffic management systems. By analyzing the traffic data of 101 urban areas from 1982 to 2014, we found that traffic delays have tended to increase, which has led to rising congestion costs.

In addition, accidents and traffic delays at intersections lead to an enormous waste of human and natural resources [5]. In the United States, accidents at intersections cost $97 billion in 2000 [6], and national congestion costs increased from $42 billion in 1982 to $160 billion in 2014 [7]. Forecasts show that if this trend continues, the national cost of congestion will increase to $192 billion by 2020 [7]. Based on the 2011 Urban Mobility Report, U.S. commuters experienced annual delays of 34 hours—at a cost of more than $100 billion [8].

Data from several studies prove that human error plays a crucial role in traffic congestion and accidents. Recent studies indicate that driver error contributes to up to 75% of all roadway crashes [9]. However, developments in computer science, sensing technology, artificial intelligence (AI), and communication technology have highlighted the possibility of introducing autonomous vehicles (AVs). The major concepts that must be improved by the development of AVs, namely sensing environments, data collection and analysis, planning, decision making, and vehicle control, have the potential to solve current problems with traffic management systems. Additionally, Moody's Investors Service predicts that the vast majority of vehicles will change to autonomous versions after 2045 and that AVs will become close to universal by 2055 [10].

Although several studies (e.g., [11], [12], [13], [14], [15], [16], and [17]) have focused on various aspects of AVs and others (e.g., [18], [19], and [20]) on intersection management related to AVs, our study differs from those in methodology, scope, and research focuses.

- In our study, we applied the systematic literature review (SLR) approach. We began with a keyword-based search and identified 105 primary studies systematically from 2952 search results, whereas other studies mostly used survey or review approaches.
- Our study covers traffic management studies at signalized intersections when AVs and mixed traffic are considered, and at unsignalized intersections when only AVs are considered. Studies [18], [19], and [20] focused on different types of traffic flows and/or different types of intersections.
- Unlike studies [18], [19], and [20], which focus on summarizing the approaches of traffic management systems, our study concentrates on investigating and comparing how well the approaches are evaluated and on the results of the evaluation. We first identified and categorized the goals, for example, improving efficiency, of the approaches. Then, we compared how well different approaches meet a certain goal. In addition, we identified and summarized the data collected from AVs and/or infrastructure for intelligent traffic management at intersections.

The remaining parts of the review have been organized as follows. Section II provides a brief overview of related reviews and surveys, whereas section III defines AVs, intelligent transportation system (ITS), and autonomous intersection management (AIM). Section IV presents the SLR process and our research questions, and illustrates the quantitative analysis of the selected papers and the answers to the research questions. We discuss the findings of our review and potential research directions in section V, and threats to the validity of the study are presented in section VI. The final section contains our conclusions and future work.

II. RELATED WORK

To manage AV-related traffic at an intersection, we need to consider both the traffic flow and the type of intersection. The traffic flow could be pure AV traffic or mixed traffic (i.e., a mixture of human-driven and automated vehicles). The intersection could be signalized or unsignalized and regulated by vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communication. To improve researchers' understanding of these and similar factors, several reviews and surveys have investigated different aspects of AVs, such as adaptive cruise control (ACC) systems [11], cooperative adaptive cruise control (CACC) systems [12], decision-making and control approaches [13], the impact of AVs on traffic [14], techniques related to AV localization [15], communication between AVs and road users [16], and vehicular communication for controlling the traffic [17].

Chen et al. [18] surveyed cooperative intersection management techniques considering V2V and V2I communication at signalized and unsignalized intersections. The cooperative methods were categorized into trajectory planning, time slots and space reservation, and virtual traffic lights. Rios-Torres et al. [19] focused on the coordination of connected and autonomous vehicles (CAVs) at intersection crossings and when merging at highway on-ramps. They covered various proposed approaches based on centralized and decentralized coordination, and they classified the approaches as heuristic rules and optimization. Guo et al. [20] surveyed urban signalized intersection management considering CAVs. The main focus of [20] was to review the proposed methods for estimating traffic flow and for optimizing traffic signal timing.

In addition to studying the approaches to controlling the traffic at intersections, it is also important to summarize and compare how effectively and efficiently the approaches meet their goals to identify gaps and improve the efficiency and effectiveness of the approaches. This insight drives our main research questions. Moreover, it is essential to cover studies related to mixed traffic, which will likely be prevalent in the next 10 to 20 years, and to unsignalized intersections. However, mixed traffic at unsignalized intersections may not be relevant, because human-driven vehicles cannot intelligently communicate and coordinate with other road users. These observations helped us to define the scope of the papers we wanted to review, as shown in Table I.

TABLE I
RESEARCH SCOPE

Intersection type	Traffic type	
	Pure AV	Mixed traffic
Signalized	[18], [19], [20], and our study	[18], [19], [20], and our study
Unsignalized	[18], [19], and our study	--

III. INTRODUCTION TO AV AND INTELLIGENT TRAFFIC MANAGEMENT

In this section, we present a brief description of AVs, intelligent transportation systems, and autonomous intersection management.

A. AUTONOMOUS VEHICLES

The Defense Advanced Research Projects Agency's (DARPA) Grand Challenge was launched in 2004 to demonstrate the technical feasibility of AVs [21]. Since then, numerous companies, such as Tesla, Audi, GM, and Google, have begun to develop and test AV technologies. As shown in Table II, SAE International has classified the automation of vehicles according to six different levels [22].

TABLE II
SAE J3016™ AUTOMATION LEVELS

Level 0	Level 1	Level 2	Level 3	Level 4	Level 5
Users are driving even when driver support features are engaged.			Users are not driving if automated driving features are engaged.		
Drivers supervise the support features.			Drivers must drive if features request them to drive.	Automated driving features do not require users to drive.	
Driver support features			Automated driving features		
Warning Momentary assistance	Steering OR brake/acceleration support	Steering AND brake/acceleration support	Automated driving features can drive the vehicle under limited conditions		Automated driving features can drive the vehicle in all conditions

AVs can gather information about the surrounding environment by using the camera, radar, LiDAR, laser, ultrasonic sensors, and GPS. Therefore, from a transportation engineering perspective, AVs are expected to enhance the safety, efficiency, ecology, and passenger comfort of the transportation system.

B. INTELLIGENT TRANSPORTATION SYSTEMS (ITS) AND AUTONOMOUS INTERSECTION MANAGEMENT (AIM)

Intelligent transportation systems (ITS) manage traffic by using new services for various transport modes [23]. The objective of ITS is to provide an improved system by informing users about traffic situations and by making mobility coordination safer and smarter [24]. In recent years, ITS has been widely applied along with the development of IT technologies such as robotics, signal and image processing, computing, sensing, and communications [25]. By using V2V, V2I, and I2V communication and AV technologies, AIM is expected to improve the efficiency of existing intersections [26]. For instance, Austroads analyzed the potential benefits of C-ITS in Australia and found that V2V communication can reduce serious road collisions by up to 35% [27].

IV. RESEARCH AND IMPLEMENTATION

We followed the Kitchenham et al. SLR process, which was conducted in [28].

A. RESEARCH METHOD AND RESEARCH QUESTIONS

As shown in Table I, in this SLR we focused on pure AV and mixed traffic in signalized and unsignalized four-way intersection(s). We reviewed papers that proposed methodologies to improve intersection performance by considering data collection, data sharing, traffic control, and other aspects.

To achieve our objectives, we formulated three main research questions:

- ***RQ1.*** What factors did intelligent intersection management studies address in terms of utilizing AVs?
- ***RQ2.*** What kinds of methodologies have been proposed to address the potential problems related to intelligent intersection management systems?
    - ***RQ2.1.*** What kinds of ITS methodologies have been proposed for traffic flow consisting of only AVs?
    - ***RQ2.2.*** What kinds of ITS methodologies have been proposed for traffic flow consisting of a mixture of autonomous and human-driven vehicles?
- ***RQ3.*** What challenges and opportunities remain?

B. CONDUCTING THE REVIEW

We focused on articles available online and published in English between January 2008 and May 10, 2019. We included the following digital libraries:

- Scopus
- IEEE
- Compendex
- Inspec
- Transport-Ovid
- ACM
- Web of Science

We used keyword-based searches to identify primary studies and followed six steps to filter relevant articles, as shown in Fig. 1. Table A-1 in Appendix A shows the search strings used in the Scopus digital library as an example.

C. RESULTS OF RESEARCH QUESTIONS

As shown in Fig. 2, the number of papers published on this topic has increased in the last few years. The lower publication number in 2019 is influenced by our search parameters, as our search included articles published only until May 10, 2019.

The top five countries, which generated about 79.6% of the articles, are the United States, China, France, Sweden, and Germany, as shown in Fig. 3.

1) RESULTS OF RQ1

Based on the thematic analysis, we categorized the goals of the primary studies as efficiency, safety, ecology, passenger comfort, and others. The "other" class includes an article about data sharing features. Some goals include several sub-goals to make this analysis more precise. The results are shown in Fig. 4.

2) RESULTS OF RQ2

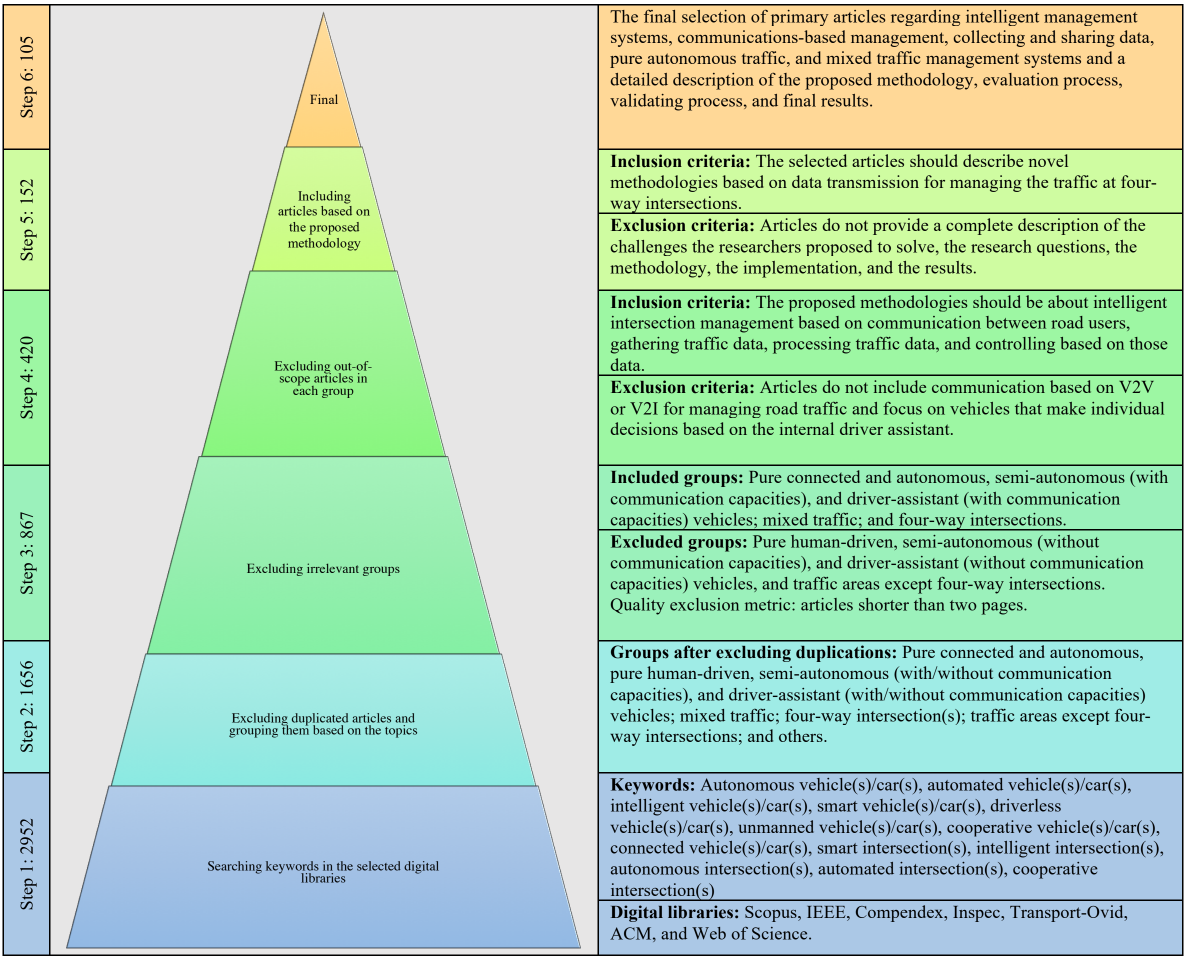


FIGURE 1. **The process of selecting primary articles.**

We divided this question into two sub-questions that yielded the following results:

2.1) RESULTS OF RQ2.1

In this section, we focus on intelligent intersection management methodologies with pure AV traffic. The proposed methodologies have been grouped based on the goals mentioned in RQ1. Some papers proposed new methodologies by focusing on one goal, for example, efficiency, whereas others considered multiple goals, for example, efficiency, safety, and ecology.

Efficiency

Several methods have been proposed to improve the efficiency of AVs in intersections. Various researchers considered different sub-goals, such as decreasing traffic delay, increasing intersection throughput, and mitigating congestion possibility. We reviewed methodologies suggested to improve efficiency at intersections.

To minimize the evacuation time of a set of vehicles, Yan et al., in [29], proposed an approach based on a dynamic programming algorithm to find the optimal vehicle passing sequence according to the arrival and passing time of a vehicle. Likewise, in [30], the authors applied heuristic smallest extra time (SET) and a dynamic programming algorithm. Yan et al. compared the performance of the genetic, dynamic programming, heuristic, and branch-and-bound algorithms to the traditional fixed-cycle-time and adaptive control systems. The results showed that the proposed method can improve evacuation time and reduce average queue length and average vehicle waiting time. Additionally, to improve the performance of the intersection, ShangGuan et al., in [31], proposed a time delay petri net-based (TdPN) control approach to develop a cooperative vehicle–infrastructure system. The results indicated that when the traffic flow rate is higher than 1,200 vehicles per hour, the TdPN method provides better performance than traditional signal control systems in terms of delay, average speed, average queue length, and average stop time.

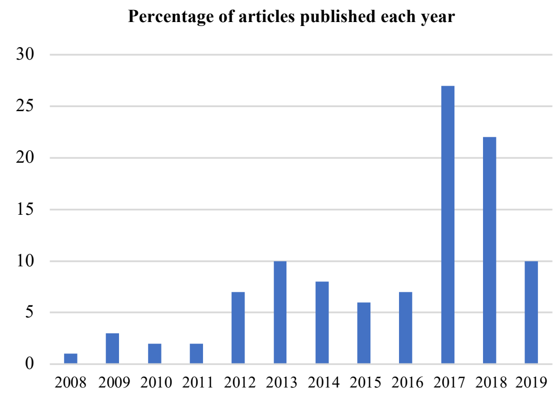


FIGURE 2. Study trends between January 2008 and May 10, 2019.

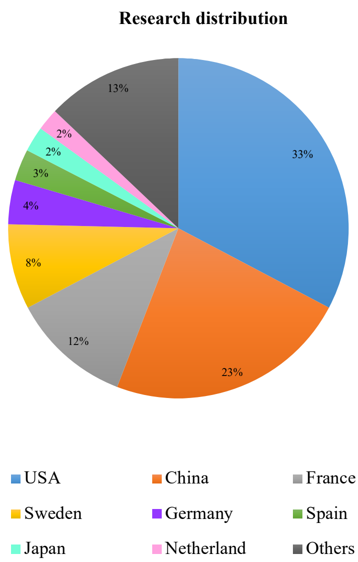


FIGURE 3. Publication distribution based on countries.

Wu et al. [32] proposed an unsignalized intersection control approach considering a new information and communication system for intelligent vehicles based on dynamic programming. They compared the center controller, V2V communication, and global solution based on simulation results and found that the global solution has a greater ability to reduce average queue length than the other two methods. Moreover, to determine the best access order of the intersection, the authors in [33] suggested a new scheduling model by viewing the intersection management problem as a machine scheduling problem, with vehicles treated as jobs and the intersection as a machine. The proposed method is based on dynamic programming. Compared to traditional signal control, the proposed method can reduce average waiting time and queue length, and improve throughput. Furthermore, by considering individual vehicle and real-time intersection control, the authors in [34] presented an AIM strategy based on an ant colony system and discrete optimization algorithm to solve real-time control problems considering a large number of vehicles and lanes. The proposed method outperforms the existing methods in terms of evacuation time, mean vehicle delay, throughput, and mean queue length.

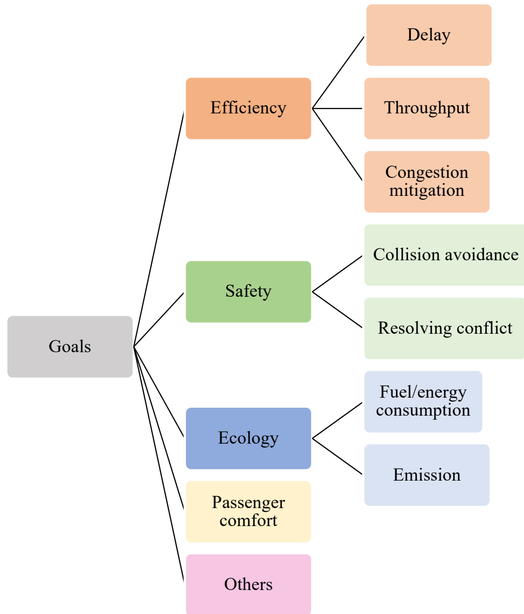


FIGURE 4. Research goals and sub-goals.

To enhance mobility, Vasirani et al. [35] designed a competitive computational market approach for intersection management. In the competitive computational market, the driver agents and the intersection-manager agents trade the use of intersection capacity. The proposed approach outperforms the traffic-light system in terms of average travel time and congestion. Additionally, in [23] the authors presented a novel scheduling model and suggested a hybrid methodology based on the distributed market–inspired approach and reservation-based intersection control model to reduce the delays for drivers who have a higher value of travel time by submitting higher bids. Their idea is to combine the competitive traffic assignment strategy (CTA) with the auction-based (AC) policy, in what is called a CA-CTA mechanism, for traffic control. This model is an extension of the reservation-based intersection control mode, which combines the auction-based policy and reservation concept. The proposed method decreases the probability of deadlock in the reservation concept proposed by Dresner et al. [36]. The results showed that compared to a first-come-first-served (FCFS) policy, the suggested approach decreases average travel time by more than 70%.

Furthermore, a time-sensitive programming method was proposed in [37] to address the round-trip delay (RTD) problem. It performs better than AIM under high input-flow conditions.

Zhang et al. [38] presented a reservation-oriented priority scheduling method, called PriorFIFO, to solve the autonomous passing-through problem. Additionally, novel reservation-based scheduling processing, named csPriorFIFO, was proposed by [39] to model and establish the traffic objects, such as centralized scheduler I-Agent, service-oriented heterogeneous vehicles, and their uniform behavior states. Both of these methods outperform the FCFS method in terms of delay and scheduling performance, respectively.

Moreover, Wei et al. [40] proposed a reservation-based control policy called Batch-Light, which is an adaptive intelligent intersection control policy for AVs. In [40], the authors used a greedy-based conflict matrix decision algorithm to increase the possibility of reservation with fairness. They further applied a k-shift optimization algorithm to help unlucky vehicles pass through the intersection. By simulating the unbalanced and balanced traffic at the intersection, the proposed method outperforms FCFS and traditional traffic-light control policies in terms of average delay and number of vehicles crossed the intersection successfully in one hour.

To optimize arrival time and speed via planning technologies, Au et al. suggested a multi-objective optimization-based method [41]. The authors proposed a planning-based motion controller to prevent stopping before the intersection and to increase throughput. Compared to the optimistic heuristic method described in [42], the proposed method reduces average delay, improves maximum throughput, and improves efficiency. To enforce liveness and prevent deadlock, the authors of [43] proposed a new intersection management policy called the batch policy of reservation in AIM.

Additionally, Carlin et al. [44] proposed an auction-based intersection system that calculates the total bids for all directions to adjust vehicle order in the intersection. Considering increasing fairness, it pays attention to keeping travel time reasonable for drivers with a low budget. When it was simulated on the road networks in four urban cities, the proposed auction-based method outperformed base cases in terms of trip time, except in Baton Rouge.

Wuthishuwong et al. [26] focused on the coordination of traffic information between infrastructures and vehicles. To balance the traffic in the network of intersections, they introduced the coordination method, which considers a network with multiple autonomous intersections. Furthermore, they proposed distributed control for a graph-based intersection network to control traffic at a macroscopic level and implemented a discrete time consensus algorithm to coordinate the traffic density with its neighbors. They used the Greenshields model to define the boundary conditions of various traffic flows to corresponding traffic density and speed. Compared to the traditional traffic signal system, the proposed method can improve the overall traffic flow by up to 20%. In addition, the proposed method outperforms the traffic signal system in terms of flow rate, average traffic speed, and throughput.

To prevent network deadlock and decrease computational delay, the authors in [45] used hierarchical architecture for cooperative intersection management. They proposed a deadlock-free protocol, which is called the advanced cooperative vehicle-actuator system (ACVAS). It can avoid computational overhead, detect and rectify deadlock, and make quick decisions.

Among the methods targeting improved efficiency, we classified methods as rule-based (e.g., [35], [23], and [37]), optimization (e.g., [29], [30], and [32]), and hybrid (e.g., [31]). Most of the proposed methods and base cases were tested in the simulation environment. Overall, the proposed methods outperform the base cases by 14–99.8%, considering different performance indicators. Further, most of the studies used a single intersection with simplified traffic conditions to validate the proposed methods. Details of the efficiency of the surveyed approaches are listed in Table B-1 of Appendix B.

Safety

Improving the safety of a targeted intersection is one of the major goals of AIM. Several methods have been proposed to achieve this goal by focusing on various sub-goals such as avoiding collisions and resolving possible conflicts.

Campos et al. [46] presented a cooperative driving strategy for intersection crossing to decrease the number of accidents and avoid collisions. They proposed a decentralized solution that allows vehicles to sequentially solve local optimization problems to help themselves to cross the intersection safely. Similarly, for considering real-time collision detection, Guangquan et al. [47] proposed a rule-based method to determine proper vehicle order and safe deceleration. The approach is based on the speed control strategy to avoid collisions, clarify the sequence of vehicles, and allow them to pass through the uncontrolled intersection.

In [48], a collaborative method was proposed to minimize collisions between AVs at an unsignalized intersection. The proposed method calculates the optimal action of the vehicle based on cost function when a conflict is detected. Additionally, Riegger et al. [49] proposed a centralized model predictive control (MPC) to control the AVs passing through the intersection and to prevent collisions. They formulated the problem as a convex quadratic program in space coordinates to generate optimal trajectories. They further considered penalized time gaps to increase safety in case of sensor errors. In a similar vein, Altché et al. [50] designed a real-time intersection supervisor based on a mixed-integer quadratic programming (MIQP) approach to monitor the control inputs and improve the safety of vehicles. To guarantee the safe navigation of vehicles, the intersection supervisor can override the vehicle control orders.

Jiang et al. in [51] suggested using a distributed and parallelizable algorithm, named the augmented Lagrangian-based alternating direction inexact Newton (ALADIN) method, to solve the coordination problem at intersections. To achieve collision avoidance at the intersection, each vehicle solves its own optimal control problem and exchanges information (e.g., arrival and departure times) with its neighbors. To provide the optimal control for AVs to safely cross the intersection, Murgovski et al. [52] applied a centralized control strategy with convex modeling steps and transformed the problem from time to space.

Finally, Rahmati et al. [3] developed a game theory–based decision framework for unprotected left-turn maneuvers. It assumes two vehicles as two players who are trying to maximize their awards by deciding to wait or continue. This approach provides the correct result in 80% of test cases.

As shown in Table B-2 in Appendix B, the methods to improve safety can be classified as rule-based (e.g., [47] and [3]), optimization (e.g., [48], [49], and [50]), and hybrid (e.g., [46]) methods to develop collision-free intersection management strategies. Most of the proposed methods and base cases were tested in the simulation environment. Most can guarantee collision avoidance at the intersection (e.g., [46]); other methods minimize conflicts (e.g., [51]). However, collisions can still occur during rush hour.

Efficiency and safety

Creating the ideal balance between several goals plays a key role in increasing the usability of proposed methodologies in real-world settings. Therefore, this section includes articles that simultaneously considered efficiency and safety.

To minimize delays and improve safety, Adams et al. [53] proposed a coordination mechanism that modifies the centralized method proposed by Dresner et al. [54] by turning it into a distributed version. The simulation results showed that the proposed method performs approximately 35–45% better than traffic signal control systems. Fayazi et al. [55] proposed an optimal scheduling strategy considering the arrival time of AVs at the intersection. They applied mixed-integer linear programming (MILP) to solve the scheduling problem, which helps to avoid accidents and reduces the number of stops and delays at intersections. Compared to traditional traffic signal systems, the proposed method reduces average travel time and average stopped delay by 7.5% and 52.4%, respectively. Chen et al. [56] presented a novel reservation management scheme, called win-fit, to reduce average trip delay and increase the average number of vehicles passing through the intersection with guaranteed safety and with starvation avoidance. In comparison to the existing method, the proposed method can reduce the average trip delay by 31–95%.

Moreover, Aoki et al. in [57] presented a safe and practical method called configurable synchronous intersection protocol (CSIP), which is a more general and resilient version of the ballroom intersection protocol (BRIP). Considering the potential for accidents caused by positioning errors in BRIP, CSIP utilizes a specific inter-vehicle distance to overcome this limitation and decreases the number of stops at the intersection, which maximizes intersection throughput. According to the simulation results, CSIP outperforms BRIP in terms of the number of collisions and trip delay. In addition, in [58], Elhenawy et al. proposed a game theory–based algorithm, based on the chicken game, to control the movements of AVs and to reduce average travel time at the intersection. The simulation showed that the proposed method reduces average travel time by 49% and delay by 89% in comparison with the all-way stop-sign intersection.

Savic et al. [59] set out a novel distributed intersection algorithm to avoid collisions and to minimize delays at the intersection in case of communication failure. They found that the proposed method effectively handles unknown and large numbers of communication failures. To minimize total delay and number of accidents, Zohdy et al. [60] presented a method based on game theory decision within a cooperative adaptive cruise control (CACC) system to optimize the movement of AVs at the intersection. In comparison with the stop-sign control intersection, the proposed method reduces total delay by approximately 70%.

Abdelhameed et al. [61] proposed an intelligent intersection control system (ICS) to improve intersection throughput, utility, average and maximum delay, and predicted collision avoidance. ICS uses a hybrid fuzzy-genetic controller to determine proper action for vehicles. In comparison with the existing traffic-light systems and the fuzzy logic controller, the proposed method improves throughput, average delay time, and maximum delay time by 90.7%, 61.6%, and 72.4% respectively. Additionally, considering real-time data processing, Chang et al. [62] suggested a new methodology called autonomous reservation-based intersection control (AReBIC) to decrease conflict and total delay and to improve mobility in an emergency evacuation. The proposed method, which combines reservation methodology and movement priority, outperforms the existing traffic control method in terms of average speed, total delay, and conflicts.

To decrease delays and guarantee safety at intersections, Müller et al. in [63] proposed an optimal arrival time strategy, which determines the optimal arrival time and movement for each vehicle. Compared to fixed-time traffic-signal controls, the proposed method reduces average delays by 97.99–98.88% and average virtual queues by 27.27–98.70%. Additionally, it improves average vehicle speed by 133.35–447.09%.

To improve the performance of the target intersection, Chai et al. [64] proposed a preassigned-slots method using location optimization on sequence evaluation (LOOSE) and the cooperative optimization method for the previous allocation alternatively transforming (COMPACT) for safety and improved efficiency. Applying the proposed method can reduce average delay, and vehicles can cross the intersection without stopping or colliding.

Moreover, Kamal et al. [65] proposed a coordination scheme for AVs to cross an unsignalized intersection safely and efficiently. The evaluations the authors conducted indicated that the proposed coordination scheme outperforms the traditional control method in terms of traffic flow when the turning rate is less than 20%.

To manage AVs at an isolated intersection, Perronnet et al. [66] presented a sequence-based protocol called transparent intersection management. The major advantage of this protocol is that it is robust under conditions of communication latency. Compared to traffic-light systems and existing methods, the proposed method reduces communication latency and evacuation time, with guaranteed safety. Similarly, Lamouik et al. [67] developed a smart multiagent traffic coordinator to provide safe and fast intersection crossing. The proposed method is based on reinforced learning (RL) and deep neural networks designed to learn and estimate the best action for each vehicle. In addition, the authors in [68] proposed an intersection-crossing protocol, which is formulated as a model predictive control problem, to provide a safety-guaranteed trajectory for a vehicle. They further proposed intervehicle coordination rules, a lane-changing protocol, and a yield protocol.

Considering V2I communication, Xie et al. [69] presented a smart in-vehicle decision-support system and used a probabilistic sequential decision-making process to help AVs to make better stop/go decisions and to reduce unnecessary stops. Moreover, to solve the traffic coordination problem, De Campos et al. [70] developed a decentralized coordination approach based on model-based decision heuristics and sequential optimal control. The proposed method is suitable for fast online implementation, and it avoids collisions. Likewise, Katriniok et al. [71] built a distributed MPC for intersection priority management to let AVs pass an unsignalized intersection efficiently.

To avoid collisions, Ze-hua et al. [72] used a discrete control strategy based on a hybrid automata theory to improve the collaboration between AVs at the intersection. They also introduced a market mechanism to improve collaboration efficiency in specific areas. To improve the safety of intersection management systems, Zheng et al. [73] proposed a delay-tolerant protocol that considers communication and network delay. The proposed method outperforms traditional traffic lights in terms of average travel time and performance, and it avoids collisions.

Furthermore, Gregoire et al. [74] developed a hybrid centralized/distributed architecture to coordinate AVs and allow vehicles to safely and efficiently cross intersections. The architecture uses a centralized approach based on a job scheduler to define the crossing time with maximum speed and a decentralized approach to avoid collisions. In the same vein, in [75], Zhang et al. modeled and designed a uniform cooperative mechanism for AVs to help them pass intersections safely, and they created the reserve advance, act later (RAAL) and high-QoS-in-prior policies to achieve these goals.

To avoid collisions and reduce waiting times, Aloufi et al. [76] proposed a model to schedule the AVs at the intersection, which is based on the production line technique. Additionally, they applied the K-Nearest Neighbors (KNN) algorithm to predict the right-turn movement of vehicles. The simulation outputs showed that the proposed model provides higher efficiency than the existing model in the case of average and random-pattern traffic flow.

Considering delay, Chouhan et al. [77] proposed a heuristic approach to avoid space-time conflicts at the intersection. The simulation results show that the proposed approach outperforms the traditional traffic light, FCFS, and CIVIC [78] in terms of average trip delay. Moreover, Creemers et al. [79] designed a centralized supervisory controller based on MPC. The simulation results indicated that the proposed approach achieves a faster transient response and lower average delay than FCFS policy and traditional traffic lights.

To handle external disturbances and model mismatches, Khayatian et al. [80] proposed a time- and space-aware technique for managing intersections with CAV traffic. Experiments on a 1/10 scale intersection with CAVs have shown that the proposed method can improve throughput on average compared to velocity assignment techniques. To navigate CAVs cross the signalized or unsignalized intersection safely and efficiently, Liu et al. [81] proposed a distributed conflict resolution mechanism via V2V communication. The results of their study indicated that the proposed approach can improve intersection efficiency by decreasing the average delay time.

To ensure safe and efficient traffic flow in intersections, Lu et al. [82] proposed a mixed-integer programming-based intersection coordination algorithm (MICA). Based on the simulation outcomes, the proposed approach outperforms the optimized traffic-light mechanism and discrete-time occupancies trajectory-based intersection traffic coordination algorithm [83] in terms of throughput.

To improve traffic throughput, Mo et al. [84] introduced multiple-collision-set strategies by extending the traditional single collision-set (CS) algorithm. Numerical results indicated that the proposed method can provide safe and efficient traffic coordination.

Steinmetz et al. [85] proposed a collision-aware resource allocation (CARA) strategy, based on a self-triggered approach, to coordinate vehicles and to manage the intersection. Moreover, to improve the quality of service (QoS), Wang et al. [86] proposed a dynamic coordination framework based on the queuing theory. Simulation and theoretical analysis results showed that road stability is guaranteed and good QoS can be provided by the proposed method. Wei et al. [87] proposed a game-in-game framework to maximize intersection throughput and mitigate traffic accidents. The simulation outcomes indicated that the proposed framework can decrease accidents and increase throughput.

Cruz-Piris et al. [88] proposed a new method to optimize the throughput of intersections automatically by utilizing the genetic algorithm. A cellular automata simulator was developed to provide a realistic simulation environment. Based on the simulation output, the proposed method can improve throughput by 9.21–36.98% compared to the traditional method.

To deal with the limitation of centralized traffic management systems, Gonzalez et al. [89] suggested a distributed management system to control intersections. The simulation results showed that the proposed method outperforms a conventional traffic control system in terms of throughput. Likewise, to improve the safety and efficiency of an unsignalized intersection, Liu et al. [90] proposed an approach based on trajectory planning for autonomous intersection management (TP-AIM) to assign priority and trajectory to vehicles and determine collision-free trajectory by considering delay. Consequently, the average evacuation time is decreased while the throughput is increased by more than 20%. Moreover, in comparison with the classical traffic light, intersection delay decreases to less than 10%.

Lu et al. [83] proposed an algorithm, named discrete-time occupancies trajectory-based intersection traffic coordination algorithm (DICA), to facilitate safe and efficient intersection crossing. The simulation result showed that DICA improves computational efficiency. Furthermore, enhanced DICA outperforms the optimized traffic light in terms of the standard deviation of trip time and average trip time.

To minimize delays and avoid collisions at the intersection, Wu et al. [91] proposed the decentralized coordination learning of autonomous intersection management (DCL-AIM) to optimize control policy. The sequential movement of vehicles is modeled as multiagent Markov decision processes (MAMDPs) and solved by using reinforcement learning, especially multiagent reinforcement learning. The simulation results showed that the DCL-AIM outperforms existing control methods.

Mirheli at al. [92] proposed a distributed cooperative control to guide connected and autonomous vehicles across an unsignalized intersection without conflict. It is called a distributed coordinated signal-free intersection control logic (DC-SICL). Based on the simulation results, the proposed method outperforms an optimized actuated signal control in terms of travel time, throughput, and safety.

Considering V2I communication, Wuthishuwong et al. [93] proposed a discrete model to manage AVs crossing an intersection without collisions and improve intersection efficiency. The proposed method decreased the waiting time at the intersection compares to the traditional traffic light.

By considering all-direction turn lanes (ADTL), He et al. [94] proposed a conflict-avoidance-based approach for coordinating vehicles at the unsignalized intersection. The simulation results indicated that the proposed approach outperforms traditional traffic lights in terms of throughput and travel time, with guaranteed collision avoidance. Additionally, Xu et al. proposed a scheduling solution to improve the throughput of an unsignalized intersection without collision risk. They developed the individual and platoon-based arrival model, which utilizes the heuristic algorithm and optimal entering time scheduling (OETS) algorithm. The proposed approach decreases traffic delay and improves efficiency compared to traditional traffic lights [95].

As shown in Table B-3 in Appendix B, rule-based (e.g., [53], [56], and [57]), optimization (e.g., [55], [60], and [65]), hybrid (e.g., [61], [63], and [68]), and machine learning (e.g., [67], [76], [84], and [91]) methods have been developed to improve intersection efficiency while considering safety. Researchers claimed that four of the optimization methods are suitable for real-time or online implementation ([70], [71], [84], and [92]). Most of the proposed methods and base cases were tested in the simulation environment. Overall, the proposed methods outperformed base cases with increases of 5–447.09% and decreases of 0–25% when considering different performance indicators. Most of the studies used a single intersection with simplified traffic conditions to validate the proposed methods.

Efficiency and ecology

Some articles considered both efficiency and ecology in managing AV traffic at intersections and proposed various methodologies to achieve this goal.

To reduce travel time, fuel consumption, and pollutant emissions, Jin et al. [96] implement the optimal scheduling of vehicle agents based on departure times in a multiagent system. Compared to the FIFO-based method [97], the proposed method can reduce travel time variability and the number of partial stops by 56–59% and 49–60%, respectively.

By using V2I communications, Saust et al. [98] proposed a cooperative system by considering signal control and vehicles' driving strategies. The idea is based on optimizing longitudinal and lateral control strategies for AVs to reduce delays, emissions, and fuel consumption. The outcomes showed that the total number of required stops decreased by 25%. Likewise, Xu et al. [99] proposed a strategy they named "cooperation between traffic signal and vehicles (CTV)," which calculates the optimal signal timing, vehicle order, and vehicle arrival time. Meanwhile, optimal control is applied to optimize the trajectory, engine power profile, and acceleration/deceleration behavior of AVs. Compared to the actuated signal control method, the proposed method reduces average trip delay and average fuel economy by 19.7% and 23.7%, respectively.

To improve energy consumption, emissions, and traffic throughput, Wang et al. [100] developed an approach called cluster-wise cooperative eco-approach and departure application (coop-EAD), which includes initial vehicle clustering, intra-cluster sequence optimization, and cluster formation control. Compared to the existing ego-EAD method, the proposed coop-EAD improves energy consumption and traffic throughput by 11.01% and 50%, respectively. Additionally, it decreases pollutant emissions by

2.29–19.91%. Tlig et al. [101] created the two-level decentralized multiagent system based on stop-free strategies to optimize network-level traffic flow and make vehicles pass through an intersection without stopping. The results of the simulation confirmed that the proposed method can significantly reduce vehicle-level energy consumption.

As shown in Table B-4 in Appendix B, optimization methods (e.g., [96], [98], [99], and [101]) and the hybrid method (e.g., [100]) have been developed to improve intersection efficiency and environmental impact. Overall, the proposed methods outperform the base cases by 2.29–60%, considering different performance indicators. Moreover, most of the studies used a single intersection with simplified traffic conditions to validate the proposed methods.

Ecology, passenger comfort, and safety

One article paid attention to three goals, namely ecology, passenger comfort, and safety in managing the traffic. Zhang et al. [102] suggested a decentralized optimal control framework to minimize fuel consumption and passenger discomfort during turning at an intersection while guaranteeing safety. The outcomes of the study [102] indicated that the proposed method is suitable for online implementation. The details of the study [102] appear in Table B-5 in Appendix B.

Efficiency, safety, and ecology

This section deals with the articles that simultaneously focused on three goals: efficiency, safety, and ecology.

To optimize energy consumption and collision avoidance, Makarem et al. [103] developed a new decentralized navigation function for AV coordination at intersections. Compared to traffic lights, the mean energy consumption of every vehicle is decreased by 13.29–73.11%. Furthermore, compared to the existing intersection management strategies, the proposed method can improve energy consumption and maximum throughput by 24.34% and 7.33–94.40%, respectively, compared to the central controller. To enhance traffic safety, traffic efficiency, and fuel consumption at an unsignalized intersection, Kamal et al. [104] proposed the vehicle-intersection coordination scheme (VICS) based on the MPC framework. In contrast to a traditional signalized intersection, the proposed method improved intersection performance factors, such as stop delay of vehicles, traffic flows, fuel consumption, and intersection capacity. In addition, Hacıoğlu et al. [105] proposed a new intersection model based on the multiagent reservation approach to decrease total delays and power loss and to improve accident detection by dividing the intersection into three main zones of communication. This strategy decreased the total delay time and total power loss. Moreover, to avoid collisions, improve energy loss, and cross an intersection without stopping, Tlig et al. [106] presented a synchronization-based intersection control to provide proper vehicle speed and arrival time. Considering the worst case, the average vehicle delay of the proposed method does not exceed 6 seconds. However, the average vehicle delay of the signalized intersection exceeds 20 seconds.

Additionally, a multi-objective evolutionary algorithm (MOEA) was proposed [107] to calculate safe routes for AVs in an intersection by routing vehicles in an efficient and safe manner. The method is suitable for low-volume traffic conditions, according to the simulation results. Mirheli et al. [108] further proposed signal-head-free intersection control logic (SICL) to find near-optimal trajectories for CAVs without any conflicts in intersections. The proposed method uses the stochastic lookahead technique to maximize intersection throughput, reduce travel time, decrease the number of stops to zero, and reduce fuel consumption. Considering different traffic situations, the proposed approach can reduce travel time by 59.4–83.7% compared to signal control methods. Malikopoulos et al. [109] proposed a decentralized energy-optimal control framework to minimize travel time, and energy and fuel consumption, and maximize the throughput of an unsignalized intersection with guaranteed safety. Compared to traditional traffic signal control methods, the proposed method can reduce fuel consumption and travel time by 46.6% and 30.9%, respectively.

Based on reservation policy and cost function, Bashiri et al. [110] introduced a centralized platoon-based controller named platoon-based autonomous intersection management (PAIM) to improve delay and its variance at the intersection. The proposed approach outperforms traffic lights in terms of delay and fuel consumption.

Medina et al. [111] introduced a decentralized solution, named cooperative intersection control (CIC) strategy, to decrease the number of accidents and improve the traffic at the intersection. The simulation results showed that the proposed method outperforms the traditional traffic light in terms of throughput and delay.

Bichiou et al. [112] proposed a new intersection management algorithm considering the nonlinear vehicle dynamic model and weather conditions. Based on the simulation results, the proposed method decreased delay, CO_2 emission, and fuel consumption by up to 80%, 40%, and 42.5%, respectively. However, the proposed algorithm may require a high computational cost to find the optimal solutions.

Philip et al. [113] suggested an approach based on collaboration between AVs and the road-side unit to improve intersection efficiency and decrease fuel consumption. The proposed method outperforms both conventional fixed switching and the state-of-the-art algorithm.

Xu et al. [114] proposed a cooperative method to optimize traffic signal and control the speed of AVs at the intersection. The simulation results indicate that the proposed method yields lower fuel consumption and trip time compared to actuated signal control when the traffic demand is between 800 and 3,200 vehicles per hour.

Bashiri et al. [115] proposed platoon-based approaches to manage the AVs through the intersection. The results showed that the proposed method outperforms stop sign policy in terms of average delay and travel delay variance.

Zhao et al. [116] presented a cooperative speed advice system, named CoDrive, to save vehicular fuel consumption at signalized intersections. Based on the simulation outcomes, fuel consumption is reduced by 7.9–38.2% compared to the GreenDrive.

As shown in Table B-6 in Appendix B, optimization (e.g., [103], [104], and [107]), rule-based (e.g., [105], [106], and [110]), and hybrid (e.g., [112]) methods have been developed to improve intersection efficiency, decrease environmental impact, and maintain traffic safety. Overall, the proposed methods outperform the base cases by 2.7–94.40%, considering different performance indicators. Again, most of the studies used a single intersection with simplified traffic conditions to validate the proposed methods.

Efficiency, safety, and passenger comfort

As efficiency, safety, and passenger comfort play essential roles in managing traffic, in this section we review articles that simultaneously considered these goals.

Considering efficiency, passenger comfort, and collision avoidance, Krajewski et al. [117] proposed a decoupled and decentralized approach, which uses graph-based methods to optimize longitudinal trajectories for multiple vehicles at urban intersections. Compared to the intersection control method for human-driven vehicles and a noncooperative control approach, the proposed method can improve intersection performance.

Dai et al. [118] designed an autonomous intersection control (AIC) to improve the travel experience of passengers, travel time, throughput, system fairness, and safety. The authors proposed a quality-of-experience-oriented autonomous intersection control (QEOIC) algorithm to schedule vehicles and make them cross the intersection efficiently and smoothly. Moreover, by predefining the decision zone and dividing the intersection into multiple collision areas, they created a schedule rule to determine the priority of the vehicles in different collision areas, which linearized the collision constraints. They further claimed that the proposed method can be used for real-time traffic control.

In a similar vein, Mladenović et al. [119] proposed a self-organizing and cooperative framework to guide vehicles across an intersection without conflict. The proposed method outperforms the regular actuated operation in terms of total delay. To decrease the waiting time of the vehicle at the intersection while avoiding collisions, Wuthishuwong et al. [120] introduced the virtual personal traffic signal based on V2I communication protocols and a node reservation algorithm. Compared to the existing traffic-flow model ([121] and [122]), the proposed method improves throughput with guaranteed safety.

In addition, Wang et al. [123] developed a novel intersection driving assistance system (IDAS) designed to deal with multiple objectives and based on V2I communication. IDAS consists of three parts: 1) passing support (PS), which provides a speed recommendation; 2) a traffic-light violation warning to inform the driver in advance about lights changing; and 3) rear-end collision warning. The results of the research indicated that the proposed IDAS can make full use of the capabilities of an infrastructure–vehicle communication system in the way that it not only maintains driving safety but also simultaneously improves passenger comfort and traffic efficiency at the intersection.

As shown in Table B-7 in Appendix B, optimization (e.g., [117] and [118]), rule-based (e.g., [119] and [120]), and hybrid (e.g., [123]) methods have been developed to improve intersection efficiency and environmental impact while considering traffic safety and passenger comfort. Overall, the proposed methods outperform the base cases in terms of total delay and throughput. Additionally, most of the studies used a single intersection with simplified traffic conditions to validate the proposed methods. One method (i.e., [123]) was validated by conducting a field test in a nonpublic intersection.

Efficiency, safety, ecology, and passenger comfort

If the proposed traffic management methodology can consider all four types of goals at the same time, and create an acceptable balance between them, it might be an ideal approach to use in the future.

Ding et al. [124] proposed a centralized cooperative intersection control approach for unsignalized intersections, which is formulated as a nonlinear constrained programming problem. Compared to actuated intersection control, the proposed method can improve traffic flow, reduce traveling time, and improve fuel consumption by 10.49–17.61%, 88.56–95.38%, and 17.18–37.81%, respectively. In addition, it reduces CO_2 emissions by 61.13–67.6%. To improve on-time arrival probability, travel time, driver satisfaction, accident rate, fuel consumption, and emissions, a semi-decentralized multiagent-based vehicle routing approach was developed in [125], considering travel time prediction and computational efficiency. Experimental results confirmed its superior performance over existing methods ([126], [127], and [128]) in areas such as average total travel time, fuel consumption, and air pollution. Qian et al. [129] proposed a decentralized MPC approach for smooth coordination of AVs at intersections to ensure collision-free travel, avoid deadlocks, and improve ecofriendly facets. Compared to MPC, the proposed method reduces fuel consumption by 4%. Furthermore, compared to the bang-bang (BB) law, energy saving is improved by 10%. To avoid collisions and increase traffic throughput, Azimi et al. [5] proposed spatial-temporal intersection protocols (STIP) based on V2V communication and vehicle speed optimization. The proposed method improved the throughput of the intersections up to 87.82% in comparison to traffic lights.

Zhao et al. [130] presented a multi-objective optimization method to coordinate the CAVs at unsignalized intersection to improve fuel consumption, traffic efficiency, and driving comfort. Simulation results showed that the proposed approach improves the efficiency, fuel consumption, and ride comfort of CAVs with low computational cost and guaranteed safety.

To decrease travel time and fuel consumption, Meng et al. [131] proposed a new approach to guide CAVs across an intersection by using traffic-light information and infrastructure-to-vehicle communication. Based on the simulation results, the proposed algorithm outperforms human-driven vehicles in terms of energy consumption and travel time.

As shown in Table B-8 in Appendix B, optimization (e.g., [124], [125], and [129]) and rule-based (e.g., [5]) methods have been developed to improve intersection efficiency and environmental impact while considering traffic safety and passenger comfort. Overall, the proposed methods outperform the base cases in terms of throughput, fuel consumption, and travel time. Most of the studies used a single intersection with simplified traffic conditions to validate the proposed methods.

Other: Data sharing

An extended version of AIM is presented in [132] to decrease the complexity and amount of data sharing in AIM. To avoid redundancy in transmission data, the authors designed an incremental data synchronization policy called ksync for driver agents to optimize the usage of bandwidth and reduce the amount of data transferred. Experimental evaluations indicated that the average data compression rate can improve by more than 80%. The details are shown in Table B-9 in Appendix B.

2.2) RESULTS OF RQ2.2

CAV technologies are likely to be progressively implemented over time, and CAVs and human-driven vehicles are likely to share the same road network. Consequently, intersection management systems with mixed traffic consisting of autonomous and human-driven vehicles have gained increased attention in recent years. Therefore, in this sub-question, we considered articles that proposed new methodologies for managing mixed traffic at intersections.

Dresner et al. [36] proposed a new AIM policy, called FCFS-Light, by using a multiagent approach. It uses a reservation-based system for managing AVs and traffic lights for managing human-driven vehicles to meet the needs of mixed traffic. Based on the simulation results, the proposed method outperforms traditional intersection signal control in terms of delay and safety. By extending the presented model in [36], Sharon et al. [133] proposed a new protocol named hybrid autonomous intersection management (H-AIM) to improve intersection performance under mixed traffic conditions. This protocol used the same FCFS reservation approach for ordering vehicles as FCFS-Light. However, FCFS-Light rejects reservation requests that carry the possibility of conflict on the green trajectory, whereas H-AIM considers conflicts with active green trajectories when rejecting reservation requests. Compared to the existing method, the proposed method can improve congestion and delay once the market penetration of CAVs exceeds 10%.

Li et al. [134] proposed a phase-time-traffic hypernetwork approach, which considers V2I communication, to minimize total control delay. The simulation results showed that the optimal intersection automation policies can serve CAV requests at its maximum potential and maintain acceptable traffic mobility. Similarly, Lin et al. [135] proposed a novel coordination method for CAVs by considering information about human-driven vehicles. Compared to traditional signal control, the proposed method reduces travel delay, the number of stops, and fuel consumption by 24.2–77.1%, 99%, and 22.1–52%, respectively.

Furthermore, based on the model predictive controller and V2I communication, Liu et al. [136] proposed a new intersection management system to manage mixed traffic. Considering the communication between vehicles and the roadside unit, Sayin et al. [137] proposed a novel information-driven intersection control based on payment-based incentive-compatible mechanism and a Vickrey–Clarke–Grove auction. The simulation results showed that the proposed method is universal and able to handle practical situations.

Based on the controller designed by [55], Fayazi et al. [138] proposed a modified MILP-based intersection controller for autonomous and human-driven vehicle traffic. The proposed method outperforms traditional signalized intersections in terms of delay.

As shown in Table B-10 in Appendix B, optimization (e.g., [134], [135], [137], and [138]), rule-based (e.g., [36] and [133]), and hybrid (e.g., [136]) methods have been developed to deal with intersection management problems in the presence of a mixture of autonomous and human-driven vehicles. Overall, the proposed methods outperform the base cases. Most of the studies used a single intersection with simplified traffic conditions to validate the proposed methods.

In summary, several of the primary studies related to RQ2 focused on a single goal (e.g., [29] and [46]). Others worked to achieve multiple goals simultaneously (e.g., [55], [96], and [102]). Fig. 5 shows the number of published articles per goal(s) by considering the categories of the methods.

3) RESULTS OF RQ3

In this section, we discuss the remaining limitations and gaps in the primary studies considering two aspects—methodology and validation environment.

From the methodological aspect, according to the results examined under RQ2, we divided the existing methodologies into four major groups: rule-based, optimization-based, hybrid, and machine learning.

First, most of the existing rule-based methods (e.g., [35], [47], [53], and [36]) have been developed to improve the efficiency and/or safety of intersections with only AV traffic or with mixed traffic. Because of their computational simplicity, rule-based methods can be applied for real-time intersection management systems and vehicle control (e.g., [47]). Moreover, rule-based methods are used to create explainable and interpretable models. Several rule-based methods have been validated by field test or real-world data (e.g., [47]). However, the complexity of the rule-based method significantly increases with the goals and constraints considered in the model. Consequently, if more goals are considered in the rule-based method, the level of improvement of the target factors decreases. Another drawback of the rule-based method is that performance may vary with traffic conditions because the rule-based method involves statistical rules and cannot guarantee the optimality of the results.

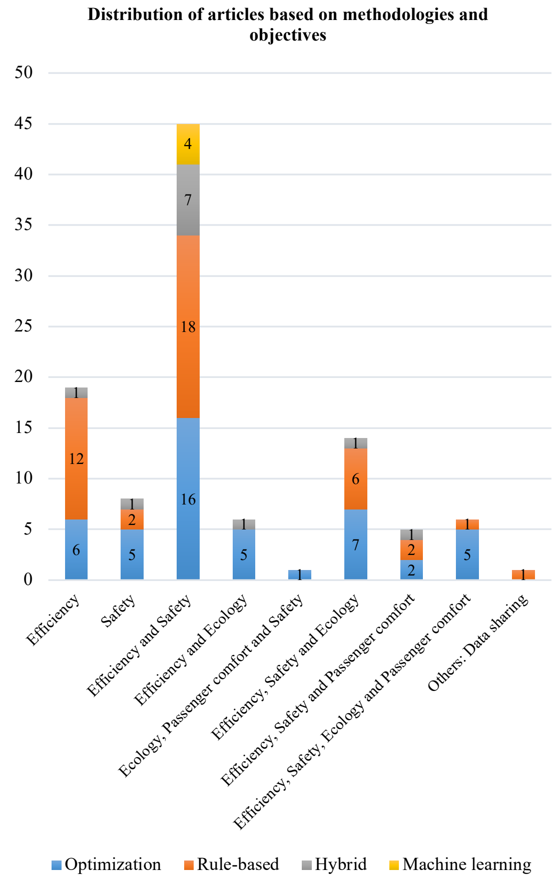


FIGURE 5. Distribution of the published articles based on the proposed methodologies and objectives.

Second, optimization-based methods (e.g., [29], [55], and [102]) have been developed to handle single-goal or multiple-goal problems. Different optimization structures or searching algorithms have been developed or applied to improve computational efficiency and to find optimal solutions. The optimization-based method can easily handle multiple goals and complex conditions by changing objective functions, constraints, and searching algorithms. Optimization-based methods always search for optimal solutions for different traffic conditions. Hence, optimization-based methods guarantee optimum performance under different traffic conditions when optimality is guaranteed. Yet optimization-based methods may not always provide a global optimal solution in the time window required for intersection management. Furthermore, the computational complexity of optimization-based methods significantly increases with the traffic volume and complexity of the situation (e.g., [107]). Therefore, only a few of the existing optimization-based methods were deemed applicable for real-time control (e.g., [50], [71], [92], and [107]). The existing optimization-based methods have been validated based on simulation results.

Third, only a few studies (e.g., [46]) implemented hybrid methods to improve efficiency and safety-related intelligent intersection control problems. Hybrid methods combine both rule-based and optimization-based methods. Since hybrid methods are partially based on rules, their computational complexity is less than optimization-based methods, which leads to lower computational time for producing a solution. Meanwhile, the optimization part of hybrid methods improves their adaptivity compared to rule-based methods. Nevertheless, a different combination of rule-based and optimization-based methods may lead to significantly different performance. Thus, how to combine the rule-based method with the optimization-based method is a challenge. Another common challenge related to the existing methods is effectively balancing multiple goals and ensuring performance.

Furthermore, considering the validation environment of the proposed methodologies revealed several limitations and gaps. First, the traffic conditions considered in the validation process were too simplified to reflect real-world traffic at intersections. Several of the proposed methodologies were tested only under specific traffic conditions, with fixed traffic flow rates. However, the traffic flow rate varies with the time of day, the day of the week, weather conditions, and so on. For example, the approaches presented in [107] are more effective and efficient with low traffic volumes than with high volumes. Few methods (e.g., [51]) were validated by considering different traffic conditions and scenarios. Additionally, only balanced traffic at the intersection was considered in several works, whereas in the real world, traffic types and volumes from different directions of the intersection tend to vary.

Second, most of the vehicle characteristics and car-following behaviors were unrealistic. Deterministic vehicle characteristic (e.g., [37]) and car-following behavior parameters have been applied in existing studies, but driver-behavior parameters (e.g., time, headway, standstill distance, and so on) are stochastic for human-driven vehicles in real-world traffic. Moreover, different car producers are equipping the vehicles they produce with sensors that differ in quality, and they can use various algorithms for automatic movements. Further, the controllers for the different types of vehicles (e.g., truck, passenger car, van, and so on) with variations in size and weight may differ.

Third, most of the methods have been validated in simulation environments (e.g., [31]). Simulation platforms may not be able to present real-world situations accurately, such as geometric limitations, weather conditions, and pedestrian flow. Additionally, developing strategies for considering the limitations of V2X communication technology in simulations remains challenging.

V. DISCUSSION AND POTENTIAL RESEARCH DIRECTIONS

From the survey, we identified several potential research directions to address the limitations of the existing methods.

A. SENSING AND CONSIDERATIONS REGARDING PEDESTRIANS AND CYCLISTS

Pedestrians and cyclists should be considered in the development of intelligent intersection management strategies. AVs can identify pedestrians and cyclists in the sensing range. For signalized intersections, AVs can feed the intersection controller pedestrian and cyclist information. For unsignalized intersections, AVs should avoid conflicts with pedestrians and cyclist and improve intersection performance by exchanging the relevant information between AVs. With the development of the Internet of Things (IoT) and wearable technologies, pedestrians and cyclists are likely to be able to communicate with AVs and intersection controllers. Therefore, an advanced control method must be developed to coordinate AVs, pedestrians, and cyclists in the intersection.

B. LEARNING CONTROL RULES AND PREDICTING TRAFFIC CONDITIONS

The AI method can be applied to improve the smartness of intersection management systems. Additionally, multiple goals should be balanced by the intersection controller under dynamic traffic conditions. Additionally, the controller should be able to control real-time traffic. Hence, based on historical data and supervised learning, we can possibly improve the dynamic rules while considering real-time traffic conditions and balancing different goals. Furthermore, AI has been widely applied [139] to predict traffic conditions based on historical data. Therefore, it can help the controller to generate proper control plans a step ahead of the requirements of the traffic situation to improve traffic management at the intersection.

C. STANDARDIZING DATA COLLECTION

Based on our findings, more studies are required to address the challenges arising from the data aspect. In the extant studies, AVs collected and shared various data, such as vehicle size, position, destination, speed, acceleration/deceleration, and so forth. The summary of the most popular types of data collected is shown in Fig. 6. We suggest that the type of data collected by AVs should be standardized. Likewise, to decrease communication delays, it would be helpful to share only the primary and required data for decision making. For example, by accessing the current speed and location of vehicles, it is possible to calculate their arrival time. This will reduce the data transmission rate and delays, which is critical for real-time management at intersections.

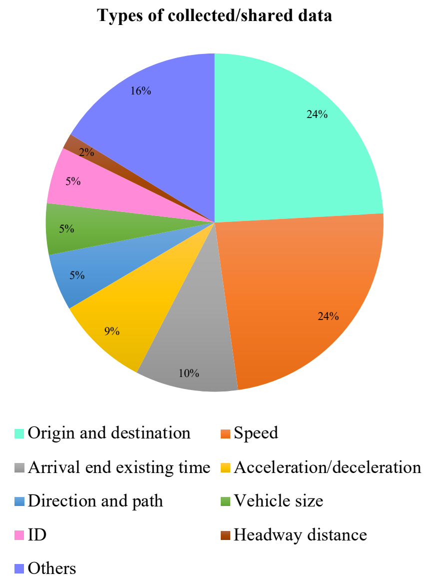


FIGURE 6. Distribution of the data type.

D. IMPROVING COMMUNICATION AND DATA QUALITY

The other matter related to data is caused by communication and data quality problems, for example, communication delays and failures, security, package loss and duplication, bandwidth limitations, low-quality data, and the effect of inclement weather on the data collecting process. Solving these problems is critical for the safety and efficiency of traffic management. For example, the approach presented in [59] will experience a crossing delay in the case of highly correlated failures. The communication network may also cause problems because of a limited communication range. For example, the communication range is set as 500 m in [59], and the experiments showed that by increasing the distance to the intersection, the packet delivery ratio decreases. Similarly, this study [118] shows that by increasing the packet loss, the throughput is decreased and the standard deviation of travel time (SDTT) is increased at the intersection.

E. LOCAL VS. GLOBAL DATA SHARING

The data sharing method is another major factor to consider. Data may be shared locally, for example, only for decision making inside one vehicle or one intersection, or globally between more intersections. This leads to two connected questions: Which approach is more efficient, and what is the effect of the environment in choosing an approach?

Different types of communication exist between vehicles and intersections, which is called V2X. By using V2I communication, data are transferred from vehicles to the infrastructure. Vehicles are responsible for sensing and collecting data and sending this data to the infrastructure. In I2V communication, data are transferred from infrastructures to vehicles. The infrastructure is responsible for sensing and collecting data and processing this data to make a decision for traffic control. V2V communication assumes that there is no central controller, and vehicles are responsible for managing traffic by sensing, collecting, and processing data. The other communication method is a combination of V2I and I2V.

By using all types of communications and accessing the most relevant data, traffic might be managed more precisely. V2V and V2I communication could be continued or discrete. In continued communication, sharing data is possible all the time. In discrete communication, sharing data happens in specifies time slots. To improve efficiency by decreasing data transfer, we suggest sharing data only if some changes occur in the shared data that may improve the performance of data sharing for better traffic management at the intersection.

F. DATA SHARING IN MIXED TRAFFIC

The other research question that could be considered is how can we collect data related to mixed traffic? If the traffic is pure AVs, then AVs are responsible for sharing their data (e.g., [47]). The other idea that is proposed in [98] is that AVs are responsible for providing data about themselves and the surrounding road users. However, these approaches are not considered for mixed traffic, which is a possible condition we might face in the near future. One idea for collecting data in mixed traffic is equipping intersections and streets with roadside sensors, for example, connected vehicle center (CVC) systems and other roadside units responsible for observing vehicle movements (e.g., [135]). However, equipping all intersections with these kinds of devices is costly, and this approach may not be efficient in all weather situations and road conditions, such as the presence of heavy snow on the road or darkness at night.

In [36] and [133], the authors proposed combining light rules with FCFS policy. In those studies, AVs followed the reservation approach, and human-driven vehicles passed through the intersection based on traffic-light rules. Thus, using that approach, AVs could pass through an intersection based on a reservation in the red light, which may be confusing for drivers and other road users such as cyclist and pedestrians. The authors of [134] suggested using data collected from AVs to improve traffic signal timing. Although this is efficient with a low ratio of CAVs (less than 10%), it is not efficient with a higher rate of CAV because “green light ahead” requests are rejected. Although various methodologies have been proposed for managing mixed traffic at intersections, they were not suitable for the real world. A potential solution is using AVs to collect data and sharing the collected data with the intersection manager to control human-driven vehicle traffic by using a dynamic traffic light at the intersection.

G. DISTRIBUTED PROCESSING OF DATA

Where to process the abundant data generated by AVs is a crucial aspect of intelligent intersection management. In existing works, the data is generally processed by either the intersection controller or AV (e.g., [140], [51], and [32]). Considering the computational limitations of intersection controllers and AVs in handling large volumes of data, different computation technologies, such as Cloud, Fog, and distributed computation, should be applied to improve the performance of intersection controllers.

H. ENSURING THE PERFORMANCE OF DATA PROCESSING

It is important to estimate the performance of proposed methods in a realistic validation environment. Ideally, these methods should be applied to control real-world traffic. Due to safety reasons, several studies (e.g., [34] and [37]) have been validated using an isolated intersection with only experimental vehicles. With the development of sensing technology, IoT, big data, digital twin technology, and AI have been gradually introduced to mitigate unpredictable and undesirable emergent behavior in complex systems. In other words, digital twin technology can provide a digital copy of real-world intersections and traffic that can be used to test proposed methods without negative consequences. Additionally, stochastic human-driver behavior should be considered instead of using predetermined parameters in car-following models. Additionally, different vehicle types, such as buses, trucks, and passenger vehicles, should be considered to reflect real-world traffic in the simulation.

VI. THREATS TO VALIDITY

In this section, we discuss the possible threats to validity of our SLR.

A. SEARCH STRATEGY

The search strategy included selecting digital libraries and searching for predefined keywords. This step may face threats from some factors such as missing or excluding relevant articles. To mitigate this risk, we used three strategies. First, to increase the possibility of finding the relevant articles, we searched the seven digital libraries most relevant to our scope. Second, we included synonyms for the search to cover the possible keywords used by various authors. To achieve this, the first author was responsible for performing a primary search to extract and list the synonyms used by different authors for the selected keywords. The second author improved the coverage of the synonyms, and the third author validated this step by considering the predefined research questions and review scope. Third, we searched using different strings by creating various combinations of the selected keywords and synonyms. We did not apply the snowballing process because the first step of our search yielded 2,952 papers, which we believe covered most of the papers relevant to our scope.

B. STUDY SELECTION CRITERIA AND PROCEDURE

Choosing articles to include and discarding others also constitutes a threat to validity, as this can result in omitting relevant articles or including irrelevant articles. To minimize this threat, we predefined the inclusion and exclusion criteria, with all authors contributing to the validation of these criteria. We subsequently strictly adhered to these criteria during the paper selection process. For example, we included papers if the proposed methodology is based on V2I or V2V communication between road users, but we excluded studies involving vehicles that make an individual decision without any communication.

C. DATA EXTRACTION STRATEGY

In this step, threats arise from the potential for incomplete information extraction from the selected articles to answer the SLR questions. To mitigate this threat, after the first author listed the data categories to extract, the second and third authors confirmed the coverage of the data categories in terms of answering the research questions. All authors discussed the categories to finalize the list, and then the first and second authors extracted the data from the selected papers. To decrease bias in the first round, the third author checked and verified the extracted data.

D. DATA SYNTHESIS STRATEGY

To decrease the risk of researcher bias during the interpreting process, we strictly followed the thematic synthesis steps. The first and second authors synthesized the extracted data, and then all the authors discussed the data to validate it.

VII. CONCLUSIONS AND FUTURE WORK

We performed an SLR to study intelligent intersection management systems considering AVs and mixed traffic. We searched seven digital libraries for papers published from January 2008 to May 10, 2019. The initial search yielded 2,952 papers, which we reduced to 105 primary studies by excluding irrelevant candidates. Compared to the surveys published in 2016 [18] [19] and early 2019 [20], in this systematic literature review, we included more articles that were published recently. We included 27, 22, and 10 more articles published in 2017, 2018, and 2019, as shown in Fig. 2. Based on the data we extracted, we observed the following:

1) In the selected articles, 40% used rule-based methodologies, 44.76% optimization methodologies, and 11.43% hybrid methodologies. Only 3.8% of the selected papers used ML approaches. We analyzed and summarized the performance of the proposed methodologies in terms of efficiency, safety, ecology, and passenger comfort. We propose that AI-based traffic management systems may reduce some of the challenges mentioned by improving the data collection process, learning traffic features and human behaviors, predicting traffic features, and making more efficient traffic-management decisions.
2) Researchers used simulators, mathematics, numerical tests, and other tools to validate the concepts they proposed in 92.38% of the selected papers, whereas 7.62% used toy cars, real cars, or field tests. Because vehicle manufacturers install diverse types of sensors with different features and quality to collect data, the proposed methodologies should be evaluated more thoroughly to deal with sensor variation.
3) The data show that 93.33% of studies focused on pure AVs, whereas the reality in the near future will be a mixture of AVs, human-driven vehicles, pedestrians, and cyclists. Therefore, a possible research direction is using the features of AVs to collect environmental data in mixed traffic to improve the performance of traffic management systems.

REFERENCES

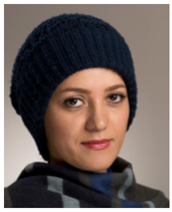
[1] C. B. Rafter, B. Anvari, and S. Box, "Traffic responsive intersection control algorithm using GPS data," in *ITSC*, 2017, pp. 1-6.
[2] M. Shi, H. Jiang, and S. Li, "An intelligent traffic-flow-based real-time vehicles scheduling algorithm at intersection," in *ICARCV*, 2016, pp. 1-5.
[3] Y. Rahmati and A. Talebpour, "Towards a collaborative connected, automated driving environment: A game theory based decision framework for unprotected left turn maneuvers," in *IEEE IV Symp.*, 2017, pp. 1316–1321.
[4] U. Franke, C. Rabe, S. Gehrig, H. Badino, and A. Barth, "Dynamic stereo vision for intersection assistance," in *32nd Congress*, Munich, Germany, 2008, vol. 2, pp. 180–189.
[5] R. Azimi, G. Bhatia, R. R. Rajkumar, and P. Mudalige, "STIP: Spatio-temporal intersection protocols for autonomous vehicles," in *ICCPS*, 2014, pp. 1–12.
[6] S. M. Loos and A. Platzer, "Safe intersections: At the crossing of hybrid systems and verification," in *ITSC*, 2011, pp. 1181–1186.
[7] D. Schrank, B. Eisele, T. Lomax, and J. Bak, (2015). *2015 urban mobility scorecard*. [Online]. Available: https://mobility.tamu.edu/ums/
[8] R. Azimi, G. Bhatia, R. Rajkumar, and P. Mudalige, "Ballroom intersection protocol: Synchronous autonomous driving at intersections," in *IEEE 21st Int. Conf. on Embedded and Real-Time Computing Syst. Applicat.*, 2015, pp. 167–175.
[9] N. A. Stanton and P. M. Salmon, "Human error taxonomies applied to driving: A generic driver error taxonomy and its implications for intelligent transport systems," *Safety Sci*, vol. 47, no. 2, pp. 227–237, 2009.
[10] S. Greenhouse. (March 21, 2017). *Driverless Future.* [Online]. Available: http://prospect.org/article/driverless-future
[11] L. Xiao and F. Gao, "A comprehensive review of the development of adaptive cruise control systems," *Vehicle Syst. Dynamics*, vol. 48, no. 10, pp. 1167–1192, 2010.
[12] K. C. Dey, L. Yan, X. Wang, Y. Wang, H. Shen, M. Chowdhury, L. Yu, C. Qiu, and V. Soundararaj, "A review of communication, driver characteristics, and controls aspects of cooperative adaptive cruise control (CACC)," *IEEE Trans. Intell. Transp. Syst.,* vol. 17, no. 2, pp. 491–509, 2016.
[13] S. M. Veres, L. Molnar, N. K. Lincoln, and C. P. Morice, "Autonomous vehicle control systems—a review of decision making," *Proc. of the Institution of Mech. Engineers, Part I: J Syst. and Control Eng.*, vol. 225, no. 2, pp. 155–195, 2011.
[14] N. Hounsell, B. Shrestha, J. Piao, and M. McDonald, "Review of urban traffic management and the impacts of new vehicle technologies," *IET Intell. Trans. Syst*., vol. 3, no. 4, pp. 419–428, 2009.
[15] S. Kuutti, S. Fallah, K. Katsaros, M. Dianati, F. Mccullough, and A. Mouzakitis, "A survey of the state-of-the-art localization techniques and their potentials for autonomous vehicle applications," *IEEE IoT J*, vol. 5, no. 2, pp. 829–846, 2018.
[16] H. Onishi, "A survey: Why and how automated vehicles should communicate to other road-users," in *VTC*, 2018, pp. 1–6.
[17] L. Li, D. Wen, and D. Yao, "A survey of traffic control with vehicular communications," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 1, pp. 425–432, 2013.
[18] L. Chen and C. Englund, "Cooperative intersection management: A survey," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 2, pp. 570–586, 2016.
[19] J. Rios-Torres and A. A. Malikopoulos, "A survey on the coordination of connected and automated vehicles at intersections and merging at highway on-ramps," *IEEE Trans. Intell. Transp. Syst.,* vol. 18, no. 5, pp. 1066–1077, 2016.
[20] Q. Guo, L. Li, and X. J. Ban, "Urban traffic signal control with connected and automated vehicles: A survey," *Transportation Research Part C: Emerging Technologies*, 2019.
[21] R. Behringer, S. Sundareswaran, B. Gregory, R. Elsley, B. Addison, W. Guthmiller, R. Daily, and D. Bevly, "The DARPA grand challenge-development of an autonomous vehicle," in *IEEE Intell. Vehicles Symp.*, 2004, pp. 226–231.
[22] J. Shuttleworth. (2019). SAE Standards News: J3016 automated-driving graphic update. [Online]. Available: https://www.sae.org/
[23] M. Vasirani and S. Ossowski, "A market-inspired approach for intersection management in urban road traffic networks," *J Artif. Intell. Res.*, vol. 43, pp. 621–659, 2012.
[24] W. Shufeng, Z. Dawei, A. Junhui, and Z. Junyou, "Lane level turning trajectory tracking of intelligent vehicle based on drivers' manipulate habits," in *CCDC*, 2017, pp. 6873–6877.
[25] F. Qu, F.-Y. Wang, and L. Yang, "Intelligent transportation spaces: Vehicles, traffic, communications, and beyond," *IEEE Commun. Mag.*, vol. 48, no. 11, 2010.
[26] C. Wuthishuwong and A. Traechtler, "Consensus-based local information coordination for the networked control of the autonomous intersection management," *Complex Intell. Syst.*, vol. 3, no. 1, pp. 17–32, Mar. 2017.
[27] N. Asselin-Miller, M. Biedka, G. Gibson, F. Kirsch, N. Hill, B. White, and K. Uddin, "Study on the deployment of C-ITS in Europe: Final report," Report for DG MOVE MOVE/C, vol. 3, pp. 2014–794, 2016.
[28] B. Kitchenham and S. Charters, "Guidelines for performing systematic literature reviews in software engineering," 2007.
[29] F. Yan, M. Dridi, and A. E. Moudni, "Autonomous vehicle sequencing algorithm at isolated intersections," in *12th Int. IEEE Conf. on Intell. Transp. Syst.*, 2009, pp. 1–6.
[30] F. Yan, M. Dridi, and A. El Moudni, "an autonomous vehicle sequencing problem at intersections: A genetic algorithm approach," *Int. J. Appl. Math. Comput. Sci.*, vol. 23, no. 1, pp. 183–200, Mar 2013.
[31] W. ShangGuan, J. Yu, B. Cai, and J. Wang, "Research on unsigned intersection control method based on cooperative vehicle infrastructure system," in *CAC*, 2017, pp. 6436–6441.
[32] J. Wu, A. Abbas-Turki, and A. E. Moudni, "Discrete Methods for Urban Intersection Traffic Controlling," in *VTC*, 2009, pp. 1–5.
[33] J. Wu, A. Abbas-Turki, and A. E. Moudni, "Intersection traffic control by a novel scheduling model," in *IEEE/INFORMS Int. Conf. Service Operations, Logistics, Informatics*, 2009, pp. 329–334.
[34] J. Wu, A. Abbas-Turki, and A. El Moudni, "Cooperative driving: An ant colony system for autonomous intersection management," *Appl. Intell.*, vol. 37, no. 2, pp. 207–222, 2012.
[35] M. Vasirani and S. Ossowski, "A computational market for distributed control of urban road traffic systems," *IEEE Trans. Intell. Transp. Syst.,* vol. 12, no. 2, pp. 313–321, 2011.
[36] K. Dresner and P. Stone, "A multiagent approach to autonomous intersection management," *J. Artif. Intell. Res.*, vol. 31, pp. 591–656, 2008.
[37] E. Andert, M. Khayatian, and A. Shrivastava, "Crossroads: Time-sensitive autonomous intersection

management technique," in *Proc. of the 54th Annual DAC*, 2017, p. 50.
[38] K. Zhang, A. D. L. Fortelle, D. Zhang, and X. Wu, "Analysis and modeled design of one state-driven autonomous passing-through algorithm for driverless vehicles at intersections," in *IEEE 16th Int. Conf. Comput. Sci. Eng.*, 2013, pp. 751–757.
[39] K. Zhang, A. Yang, H. Su, A. d. L. Fortelle, K. Miao, and Y. Yao, "Service-oriented cooperation models and mechanisms for heterogeneous driverless vehicles at continuous static critical sections," IEEE *Trans. Intell. Transp. Syst.*, vol. 18, no. 7, pp. 1867–1881, 2017.
[40] X. Wei, G. Z. Tan, and N. Ding, "Batch-Light: An adaptive intelligent intersection control policy for autonomous vehicles," *Proc. IEEE Int. Conf. Progr. Informat. Comput.*, 2014, pp. 98–103.
[41] T.-C. Au and P. Stone, "Motion planning algorithms for autonomous intersection management," in *Bridging the Gap Between Task and Motion Planning*, 2010, vol. WS-10-01, pp. 2–9: AI Access Foundation.
[42] K. M. Dresner, "Autonomous intersection management," University of Texas at Austin, United States, Rep. AD1024611, 2009.
[43] T.-C. Au, N. Shahidi, and P. Stone, "Enforcing liveness in autonomous traffic management," in *AAAI*, San Francisco, CA, United States, 2011, vol. 2, pp. 1317–1322: AI Access Found.
[44] D. Carlin, S. D. Boyles, and P. Stone, "Auction-based autonomous intersection management," in *16th Int. Conf. ITSC*, 2013, pp. 529–534.
[45] F. Perronnet, A. Abbas-Turki, and A. E. Moudni, "Vehicle routing through deadlock-free policy for cooperative traffic control in a network of intersections: Reservation and congestion," in *ITSC*, 2014, pp. 2233–2238.
[46] G. R. d. Campos, P. Falcone, and J. Sjöberg, "Autonomous cooperative driving: A velocity-based negotiation approach for intersection crossing," in *ITSC*, 2013, pp. 1456–1461.
[47] L. Guangquan, L. Lumiao, W. Yunpeng, Z. Ran, B. Zewen, and C. Haichong, "A rule based control algorithm of connected vehicles in uncontrolled intersection," in *ITSC*, 2014, pp. 115–120.
[48] F. Belkhouche, "Control of autonomous vehicles at an unsignalized intersection," in *ACC*, 2017, pp. 1340–1345.
[49] L. Riegger, M. Carlander, N. Lidander, N. Murgovski, and J. Sjoberg, "Centralized MPC for autonomous intersection crossing," in *ITSC*, Piscataway, NJ, USA, 2016, pp. 1372–1377.
[50] F. Altché, X. Qian, and A. d. L. Fortelle, "Least restrictive and minimally deviating supervisor for safe semi-autonomous driving at an intersection: An MIQP approach," in *ITSC*, 2016, pp. 2520–2526.
[51] Y. Jiang, M. Zanon, R. Hult, and B. Houska, "Distributed algorithm for optimal vehicle coordination at traffic intersections,", *IFAC*, Article vol. 50, no. 1, pp. 11577–11582, 2017.
[52] N. Murgovski, G. R. de Campos, and J. Sjöberg, "Convex modeling of conflict resolution at traffic intersections," in *54th IEEE conf. CDC*, 2015, pp. 4708–4713.
[53] S. Adams and M. J. Rutherford, "Towards decentralized waypoint negotiation," in *AAAI Workshop*, Toronto, ON, 2012, vol. WS-12-10, pp. 2–5.
[54] K. Dresner and P. Stone, "Multiagent traffic management: A reservation-based intersection control mechanism," in *Proc. 3rd Int. Joint Conf. Autonomous Agents Multiagent Syst.*, 2004, pp. 530–537.
[55] S. A. Fayazi, A. Vahidi, and A. Luckow, "Optimal scheduling of autonomous vehicle arrivals at intelligent intersections via MILP," in *ACC*, 2017, pp. 4920–4925.
[56] G. Chen and K. Kang, "Win-fit: Efficient intersection management via dynamic vehicle batching and scheduling," in *ICCVE*, 2015, pp. 263–270.
[57] S. Aoki and R. R. Rajkumar, "A configurable synchronous intersection protocol for self-driving vehicles," in *RTCSA*, 2017, pp. 1–11.
[58] M. Elhenawy, A. A. Elbery, A. A. Hassan, and H. A. Rakha, "An intersection game-theory-based traffic control algorithm in a connected vehicle environment," in *IEEE 18th Int. Conf. Intell. Transp. Syst.*, 2015, pp. 343–347.
[59] V. Savic, E. M. Schiller, and M. Papatriantafilou, "Distributed algorithm for collision avoidance at road intersections in the presence of communication failures," in *IEEE IV Symp.*, 2017, pp. 1005–1012.
[60] I. H. Zohdy and H. Rakha, "Game theory algorithm for intersection-based cooperative adaptive cruise control (CACC) systems," in *15th Int. IEEE Conf. Intell. Transp. Syst.*, 2012, pp. 1097–1102.
[61] M. M. Abdelhameed, M. Abdelaziz, S. Hammad, and O. M. Shehata, "A Hybrid Fuzzy-Genetic Controller for a multi-agent intersection control system," in *ICET*, 2014, pp. 1–6.
[62] Y. Chang and P. Edara, "AReBIC: Autonomous reservation-based intersection control for emergency evacuation," in *IEEE IV Symp.*, 2017, pp. 1887–1892.
[63] E. R. Müller, R. C. Carlson, and W. K. Junior, "Intersection control for automated vehicles with MILP," *IFAC*, Article vol. 49, no. 3, pp. 37–42, 2016.
[64] L. Chai, B. Cai, W. ShangGuan, and J. Wang, "Connected and autonomous vehicles coordinating method at intersection utilizing preassigned slots," in *ITSC*, 2017, pp. 1–6.
[65] M. A. S. Kamal, J. Imura, A. Ohata, T. Hayakawa, and K. Aihara, "Coordination of automated vehicles at a traffic-lightless intersection," in *ITSC*, 2013, pp. 922–927.
[66] F. Perronnet, A. Abbas-Turki, and A. E. Moudni, "A sequenced-based protocol to manage autonomous vehicles at isolated intersections," in *ITSC*, 2013, pp. 1811–1816.
[67] I. Lamouik, A. Yahyaouy, and M. A. Sabri, "Smart multi-agent traffic coordinator for autonomous vehicles at intersections," in *ATSIP*, 2017, pp. 1–6.
[68] K. Kim, "Collision free autonomous ground traffic: A model predictive control approach," in *ICCPS*, 2013, pp. 51–60.
[69] X.-F. Xie and Z.-J. Wang, "SIV-DSS: Smart In-Vehicle Decision Support System for driving at signalized intersections with V2I communication," *Transp. Res. Part C: Emerging Technologies*, vol. 90, pp. 181–197, 2018.
[70] G. R. De Campos, P. Falcone, R. Hult, H. Wymeersch, and J. Sjöberg, "Traffic coordination at road intersections: autonomous decision-making algorithms using model-based heuristics," *IEEE Intell. Transp. Syst. Mag.*, vol. 9, no. 1, pp. 8–21, 2017, Art. no. 7823100.
[71] A. Katriniok, P. Kleibaum, C. Ress, and L. Eckstein, "Automation of road vehicles using V2X: An application to intersection automation," in *SAE World Congress Experience*, Southwest Research Institute (SwRI): SAE International.
[72] F. Ze-hua, L. Han-bo, H. Wei, and Y. Tian, "District control strategy for vehicle collision avoidance based on hybrid automaton model," in *Proc. 11th WCICA*, 2014, pp. 5378–5383.
[73] B. Zheng, C. Lin, H. Liang, S. Shiraishi, W. Li, and Q. Zhu, "Delay-aware design, analysis and verification of intelligent intersection management," in *SMARTCOMP*, 2017, pp. 1–8.
[74] J. Gregoire and E. Frazzoli, "Hybrid centralized/distributed autonomous intersection control:

Using a job scheduler as a planner and inheriting its efficiency guarantees," in *CDC*, 2016, pp. 2549–2554.
[75] K. Zhang, A. Yang, H. Su, A. d. L. Fortelle, and X. Wu, "Unified modeling and design of reservation-based cooperation mechanisms for intelligent vehicles," in *ITSC*, 2016, pp. 1192–1199.
[76] N. Aloufi and A. Chatterjee, "Autonomous vehicle scheduling at intersections based on production line technique," in *VTC*, 2018, pp. 1–5.
[77] A. P. Chouhan and G. Banda, "Autonomous intersection management: A heuristic approach," *IEEE Access*, vol. 6, pp. 53287–53295, 2018.
[78] J. Lee and B. Park, "Development and evaluation of a cooperative vehicle intersection control algorithm under the connected vehicles environment," *IEEE Trans. Intell. Transp. Syst.,* vol. 13, no. 1, pp. 81–90, 2012.
[79] F. Creemers, A. I. Morales Medina, E. Lefeber, and N. v. de Wouw, "Design of a supervisory controller for cooperative intersection control using model predictive control," *IFAC*, vol. 51, no. 33, pp. 74–79, 2018.
[80] M. Khayatian, M. Mehrabian, and A. Shrivastava, "RIM: Robust intersection management for connected autonomous vehicles," in *IEEE RTSS*, 2018, pp. 35–44.
[81] C. Liu, C. Lin, S. Shiraishi, and M. Tomizuka, "Distributed conflict resolution for connected autonomous vehicles," *IEEE Trans. Intell. Veh.*, vol. 3, no. 1, pp. 18–29, 2018.
[82] Q. Lu and K. Kim, "A mixed integer programming approach for autonomous and connected intersection crossing traffic control," in *VTC*, 2018, pp. 1–6.
[83] Q. Lu and K.-D. Kim, "Autonomous and connected intersection crossing traffic management using discrete-time occupancies trajectory," *Appl. Intell.*, vol. 49, no. 5, pp. 1621–1635, 2019.
[84] Y. Mo, M. Wang, T. Zhang, and Q. Zhang, "Autonomous cooperative vehicle coordination at road intersections," in *ICCS*, 2018, pp. 192–197.
[85] E. Steinmetz, R. Hult, Z. Zou, R. Emardson, F. Brannstrom, P. Falcone, and H. Wymeersch, "Collision-aware communication for intersection management of automated vehicles," *IEEE Access*, vol. 6, pp. 77359–77371, 2018.
[86] M. Wang, T. Zhang, L. Gao, and Q. Zhang, "High throughput dynamic vehicle coordination for intersection ground traffic," in *VTC*, 2018, pp. 1–6.
[87] H. Wei, L. Mashayekhy, and J. Papineau, "Intersection management for connected autonomous vehicles: A game theoretic framework," in *ITSC*, 2018, pp. 583–588.
[88] L. Cruz-Piris, M. A. Lopez-Carmona, and I. Marsa-Maestre, "Automated optimization of intersections using a genetic algorithm," *IEEE Access*, vol. 7, pp. 15452–15468, 2019.
[89] C. L. Gonzalez, J. L. Zapotecatl, J. M. Alberola, V. Julian, and C. Gershenson, "Distributed management of traffic intersections," in *9th ISAmI*, Toledo, Spain, 2019, vol. 806, pp. 56–64: Springer Verlag.
[90] B. Liu, Q. Shi, Z. Song, and A. El Kamel, "Trajectory planning for autonomous intersection management of connected vehicles," *Simul. Model. Pract. Th.*, vol. 90, pp. 16–30, 2019.
[91] Y. Wu, H. Chen, and F. Zhu, "DCL-AIM: Decentralized coordination learning of autonomous intersection management for connected and automated vehicles," *Transp. Res. Part C: Emerging Technologies*, vol. 103, pp. 246–260, 2019.
[92] A. Mirheli, M. Tajalli, L. Hajibabai, and A. Hajbabaie, "A consensus-based distributed trajectory control in a signal-free intersection," *Transportation research part C: emerging technologies*, 2019, pp.161-176.
[93] C. Wuthishuwong and A. Traechtler, "Vehicle to infrastructure based safe trajectory planning for autonomous intersection management," in *ITST*, 2013, pp. 175–180.
[94] Z. He, L. Zheng, L. Lu, and W. Guan, "Erasing lane changes from roads: A design of future road intersections," *IEEE Trans. Intell. Veh.*, vol. 3, no. 2, pp. 173–184, 2018.
[95] Y. Xu, H. Zhou, B. Qian, H. Wu, Z. Zhang, and X. S. Shen, "When automated vehicles meet non-signalized intersections: A collision-free scheduling solution," in *ICCC*, 2018, pp. 709–713.
[96] Q. Jin, G. Wu, K. Boriboonsomsin, and M. Barth, "Multi-agent intersection management for connected vehicles using an optimal scheduling approach," in *ICCVE*, Beijing, 2012, pp. 185–190.
[97] Q. Jin, G. Wu, K. Boriboonsomsin, and M. Barth, "Advanced intersection management for connected vehicles using a multi-agent systems approach," in *IEEE IV Symp.*, Alcal de Henares, Madrid, 2012, pp. 932–937.
[98] F. Saust, J. M. Wille, and M. Maurer, "Energy-optimized driving with an autonomous vehicle in urban environments," in *VTC*, 2012, pp. 1–5.
[99] B. Xu, X. J. Ban, Y. Bian, J. Wang, and K. Li, "V2I based cooperation between traffic signal and approaching automated vehicles," in *IEEE IV Symp.*, 2017, pp. 1658–1664.
[100] Z. Wang, G. Wu, P. Hao, and M. J. Barth, "Cluster-wise cooperative eco-approach and departure application along signalized arterials," in *ITSC*, 2017, pp. 145–150.
[101] M. Tlig, O. Buffet, and O. Simonin, "Stop-free strategies for traffic networks: Decentralized on-line optimization," in *ECAI*, 2014, pp. 1191–1196.
[102] Y. Zhang, A. A. Malikopoulos, and C. G. Cassandras, "Decentralized optimal control for connected automated vehicles at intersections including left and right turns," in *CDC*, 2017, pp. 4428–4433.
[103] L. Makarem and D. Gillet, "Fluent coordination of autonomous vehicles at intersections," in *SMC*, 2012, pp. 2557–2562.
[104] M. A. S. Kamal, J. Imura, T. Hayakawa, A. Ohata, and K. Aihara, "A vehicle-intersection coordination scheme for smooth flows of traffic without using traffic lights," *IEEE Trans. Intell. Transp. Syst.,* vol. 16, no. 3, pp. 1136–1147, 2015.
[105] F. Hacıoğlu and M. T. Söylemez, "Power consumption based multi agent intersection management method," in *ELECO*, 2017, pp. 954–958.
[106] M. Tlig, O. Buffet, and O. Simonin, "Decentralized traffic management: A synchronization-based intersection control," in *ICALT*, 2014, pp. 109–114.
[107] K. S. N. Ripon, J. Solaas, and H. Dissen, "Multi-objective evolutionary optimization for autonomous intersection management," in *11th Int. Conf. SEAL*, 2017, pp. 297–308.
[108] A. Mirheli, L. Hajibabai, and A. Hajbabaie, "Development of a signal-head-free intersection control logic in a fully connected and autonomous vehicle environment," *Transp. Res. Part C-Emerging Technologies*, vol. 92, pp. 412–425, Jul 2018.
[109] A. A. Malikopoulos, C. G. Cassandras, and Y. J. Zhang, "A decentralized energy-optimal control framework for connected automated vehicles at signal-free intersections," *Automatica*, vol. 93, pp. 244–256, Jul 2018.
[110] M. Bashiri, H. Jafarzadeh, and C. H. Fleming, "PAIM: Platoon-based autonomous intersection management," *ITSC*, 2018, pp. 374–380.
[111] A. I. M. Medina, N. van de Wouw, and H. Nijmeijer, "Cooperative intersection control based on virtual platooning," *IEEE Trans. Intell. Transp. Syst.,* vol. 19, no. 6, pp. 1727–1740, Jun 2018.
[112] Y. Bichiou and H. A. Rakha, "Developing an optimal intersection control system for automated connected

vehicles," *IEEE Trans. Intell. Transp. Syst.,* vol. 20, no. 5, pp. 1908–1916, 2019.
[113] B. V. Philip, T. Alpcan, J. Jin, and M. Palaniswami, "Distributed real-time IoT for autonomous vehicles," *IEEE Trans. Ind. Informat.*, vol. 15, no. 2, pp. 1131–1140, 2019.
[114] B. Xu, X. J. Ban, Y. Bian, W. Li, J. Wang, S. E. Li, and K. Li, "Cooperative method of traffic signal optimization and speed control of connected vehicles at isolated intersections," *IEEE Trans. Intell. Transp. Syst.,* vol. 20, no. 4, pp. 1390–1403, 2019.
[115] M. Bashiri and C. H. Fleming, "A platoon-based intersection management system for autonomous vehicles," in *IEEE IV Symp.*, 2017, pp. 667–672.
[116] Y. Zhao, S. Yao, H. Shao, and T. Abdelzaher, "Codrive: Cooperative driving scheme for vehicles in urban signalized intersections," in *Proc. 9th ACM/IEEE Int. Conf. Cyber-Physical Syst.*, Porto, Portugal, 2018.
[117] R. Krajewski, P. Themann, and L. Eckstein, "Decoupled cooperative trajectory optimization for connected highly automated vehicles at urban intersections," in *IEEE IV Symp.*, 2016, pp. 741–746.
[118] P. Dai, K. Liu, Q. Zhuge, H. M. Sha Edwin, S. Lee Victor Chung, and H. Son Sang, "Quality-of-experience-oriented autonomous intersection control in vehicular networks," *IEEE Trans. Intell. Transp. Syst.,* vol. 17, pp. 1956–1967.
[119] M. N. Mladenović and M. M. Abbas, "Self-organizing control framework for driverless vehicles," *ITSC*, 2013, pp. 2076–2081.
[120] C. Wuthishuwong, A. Traechtler, and T. Bruns, "Safe trajectory planning for autonomous intersection management by using vehicle to infrastructure communication," *EURASIP J. Wirel. Commun. Netw.*, pp. 1–12, 2015.
[121] L. Immers and S. Logghe, "Traffic flow theory," *Faculty of Engineering, Department of Civil Engineering, Section Traffic and Infrastructure, Kasteelpark Arenberg*, 2002.
[122] A. Trani, "Traffic flow models," lecture notes: Introduction to Transportation Engineering, Virginia, USA, Fall 2009.
[123] J. Wang, D. Zhang, J. Liu, M. Lu, and K. Li, "Multi-objective driving assistance system for intersection support," in *13th Int. IEEE Conf. Intell. Transp. Syst.*, 2010, pp. 348–353.
[124] J. Ding, H. Xu, J. Hu, and Y. Zhang, "Centralized cooperative intersection control under automated vehicle environment," in *IEEE IV Symp.*, 2017, pp. 972–977.
[125] Z. Cao, H. Guo, and J. Zhang, "A multiagent-based approach for vehicle routing by considering both arriving on time and total travel time," *ACM Trans. Intell. Syst. Technol.*, vol. 9, no. 3, pp. 1–21, 2017.
[126] Z. Cao, H. Guo, J. Zhang, D. Niyato, and U. Fastenrath, "Improving the efficiency of stochastic vehicle routing: A partial lagrange multiplier method," *IEEE Trans. Veh. Technol.*, vol. 65, no. 6, pp. 3993–4005, 2016.
[127] T. Yamashita, K. Izumi, K. Kurumatani, and H. Nakashima, "Smooth traffic flow with a cooperative car navigation system," in *Proc. 4th Int. joint Conf. Autonomous Agents Multiagent Systems*, 2005, pp. 478–485: ACM.
[128] S. Jiang, J. Zhang, and Y.-S. Ong, "A pheromone-based traffic management model for vehicle re-routing and traffic light control," in *Proc. Int. Conf. Autonomous Agents Multiagent Syst.*, 2014, pp. 1479–1480: Int. Found. Autonomous Agents Multiagent Syst.
[129] X. Qian, J. Gregoire, A. De La Fortelle, and F. Moutarde, "Decentralized model predictive control for smooth coordination of automated vehicles at intersection," in *ECC*, 2015, pp. 3452–3458.
[130] X. Zhao, J. Wang, Y. Chen, and G. Yin, "Multi-objective cooperative scheduling of CAVs at non-signalized intersection," in *ITSC*, 2018, pp. 3314–3319.
[131] X. Meng and C. G. Cassandras, "Optimal control of autonomous vehicles for non-stop signalized intersection crossing," in *Proc. IEEE Conf. Decision Contr.*, 2019, vol. 2018–December, pp. 6988–6993.
[132] C. Yu, G. Tan, and Y. Yu, "Make driver agent more reserved: An AIM-based incremental data synchronization policy," in *IEEE 9th Int. Conf. Mobile Ad-hoc Sens. Netw.*, 2013, pp. 198–205.
[133] G. Sharon and P. Stone, "A protocol for mixed autonomous and human-operated vehicles at intersections," in *AAMAS*, 2017, pp. 151–167.
[134] P. Li and X. Zhou, "Recasting and optimizing intersection automation as a connected-and-automated-vehicle (CAV) scheduling problem: A sequential branch-and-bound search approach in phase-time-traffic hypernetwork," *Transp. Res. Part B: Methodological*, vol. 105, pp. 479–506, 2017.
[135] P. Lin, J. Liu, P. J. Jin, and B. Ran, "Autonomous vehicle-intersection coordination method in a connected vehicle environment," *IEEE Intell. Transp. Syst. Mag.*, vol. 9, no. 4, pp. 37–47, 2017.
[136] X. Liu, P. Hsieh, and P. R. Kumar, "Safe intersection management for mixed transportation systems with human-driven and autonomous vehicles," in *Allerton*, 2018, pp. 834–841.
[137] M. O. Sayin, C. Lin, S. Shiraishi, J. Shen, and T. Başar, "Information-driven autonomous intersection control via incentive compatible mechanisms," *IEEE Trans. Intell. Transp. Syst.,* vol. 20, no. 3, pp. 912–924, 2019.
[138] S. A. Fayazi and A. Vahidi, "Mixed-integer linear programming for optimal scheduling of autonomous vehicle intersection crossing," *IEEE Trans. Intell. Veh.*, vol. 3, no. 3, pp. 287–299, 2018.
[139] H. Nguyen, L.-M. Kieu, T. Wen, and C. Cai, "Deep learning methods in transportation domain: A review," *IET Intell. Transp. Syst.*, vol. 12, no. 9, pp. 998–1004, 2018.
[140] B. Yang and C. Monterola, "A simple distributed algorithm for lightless intersection control based on non-linear interactions between vehicles," in *ITSC*, 2017, pp. 1–6.
[141] F. Webster, "Road research technical paper," *Road Res. Lab.*, London, UK, vol. 39, 1958.
[142] F. C. Fang and L. Elefteriadou, "Development of an optimization methodology for adaptive traffic signal control at diamond interchanges," *J. Transp. Eng.*, vol. 132, no. 8, pp. 629–637, 2006.
[143] S. Cohen, "Ingénierie du trafic routier: Elément de théorie du trafic et applications," Eyrolles, Paris, 1993.
[144] T. Schwerdtfeger, "DYNEMO: A model for the simulation of traffic flow in motorway networks," in *9th Int. Symp. Transp. Traffic Th.*, Netherlands, 11–13 July 1984, 1984.
[145] M. Treiber, A. Hennecke, and D. Helbing, "Congested traffic states in empirical observations and microscopic simulations," *Physical Rev. E*, vol. 62, no. 2, p. 1805, 2000.
[146] K. Nagel and M. Schreckenberg, "A cellular automaton model for freeway traffic," *J. de Physique I*, vol. 2, no. 12, pp. 2221–2229, 1992.
[147] D. Krajzewicz, J. Erdmann, M. Behrisch, and L. Bieker, "Recent development and applications of SUMO-Simulation of Urban MObility," *Int. J. Adv. Syst. Measure.*, vol. 5, no. 3 & 4, 2012.
[148] Gurobi Optimization, "Gurobi optimizer reference manual," (2015). [Online]. Available: http://www.gurobi.com.
[149] "Simulation of Urban Mobility,". [Online]. Available: http://sourceforge.net/projects/sumo/

[150]"AIMSUN User's Manual," Version 7 Transport Simulation Systems, (2011). [Online]. Available: http://www. aimsun.com.
[151]"SUMO,". [Online]. Available: http://www.dlr.de/ts/en/desktopdefault.aspx/tabid-9883/16931 read-41000/
[152]PTV Group, "Ptv vissim 7 user manual," PTV AG, 2015.
[153]"SUMO: Simulation of Urban Mobility,". [Online]. Available: http://sumo.sourceforge.net/. Accessed on July, 2012
[154]Y. J. Zhang, A. A. Malikopoulos, and C. G. Cassandras, "Optimal control and coordination of connected and automated vehicles at urban traffic intersections," in *ACC*, 2016, pp. 6227–6232.
[155]L. Makarem and D. Gillet, "Decentralized coordination of autonomous vehicles at intersections," *IFAC Proc.*, vol. 44, no. 1, pp. 13046–13051, 2011.
[156]"AIMSUN User's Manual" (Barcelona, Spain, 2013). [Online]. Available: http://www.aimsun.com/
[157]D. Zhao, Y. Dai, and Z. Zhang, "Computational intelligence in urban traffic signal control: A survey," *IEEE Trans. Syst.*, vol. 42, no. 4, pp. 485–494, 2012.
[158]PTV America, "PTV VISTRO User Manual," *Ptv Ag*, *2*, 2014.
[159]I. H. Zohdy and H. A. Rakha, "Intersection management via vehicle connectivity: The intersection cooperative adaptive cruise control system concept," *J. Intell. Transp. Syst.: Technol., Plan. Ops*, vol. 20, no. 1, pp. 17–32, 2016.
[160]H. Rakha, P. Pasumarthy, and S. Adjerid, "A simplified behavioral vehicle longitudinal motion model," *Transp. Lett.*, vol. 1, no. 2, pp. 95–110, 2009.
[161] Institute of Transportation Systems German Aerospace Center, "SUMO: Simulation of urban mobility" (2017). [Online]. Available: dlr.de/ts/sumo
[162]S. Yao, Y. Zhao, A. Zhang, L. Su, and T. Abdelzaher, "Deepiot: Compressing deep neural network structures for sensing systems with a compressor-critic framework," in *Proc. 15th ACM Conf. Embedded Netw. Sens. Syst.*, 2017, p. 4: ACM.
[163]M. Behrisch, L. Bieker, J. Erdmann, and D. Krajzewicz, "SUMO–simulation of urban mobility: An overview," in *Proc. SIMUL*, 2011.
[164]M. J. M. Mazo, J. Valencia, and G. O. Tost, "Numerical simulation analysis of a traffic model," in *Analysis, Modelling, Optimization, and Numerical Techniques*: Springer, 2015, pp. 363–371.
[165]H. Schwetman, "CSIM19: A powerful tool for building system models," in *Proc. Winter Simulation Conf.*, 2001, vol. 1, pp. 250–255.
[166]K. Ahn, H. Rakha, A. Trani, and M. Van Aerde, "Estimating vehicle fuel consumption and emissions based on instantaneous speed and acceleration levels," *J. Transp. Eng.*, vol. 128, no. 2, pp. 182–190, 2002.
[167]M. A. S. Kamal, M. Mukai, J. Murata, and T. Kawabe, "Ecological vehicle control on roads with up-down slopes," *IEEE Trans. Intell. Transp. Syst.,* vol. 12, no. 3, pp. 783–794, 2011.
[168]"VISSIM 5.20 COM User Manual, Planung Transport Verkehr (PTV)," Karlsruhe, Germany, 2009.



ELNAZ NAMAZI received a B.S. in Information Technology (IT) in 2009, and an M.S. in Computer Engineering–Software in 2012. She is currently pursuing a Ph.D. in Computer Science at Norwegian University of Science and Technology (NTNU). Her research interest includes the AI-based management of autonomous vehicles in the urban environment.

JINGYUE LI is an Associate Professor at the Computer Science Department of Norwegian University of Science and Technology (NTNU). Jingyue Li obtained his Master in Artificial Intelligence (AI) from the Department of Computer Science of Beijing University of Technology in 2001, and his Ph.D. in software engineering from the Department of Computer Science of NTNU in 2006. After that, he continued as a postdoctoral fellow in the same department. He has been working as a visiting researcher at University College London and the University of Washington. He has published more than 50 scientific reports in peer-reviewed journals and conferences, including the prestigious IEEE Transaction on Software Engineering, International Conference on Software Engineering, Empirical Software Engineering Journal, and IEEE Software. Jingyue Li received the Best Paper Award at the 4th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement.

CHAORU LU received a B.S. in Civil Engineering from Hunan University of Science and Technology in 2011, an M.S. in Civil Engineering from Texas A & M University–Kingsville in 2014, and a Ph.D. in Civil Engineering from Iowa State University in 2017. He is a postdoctoral fellow at Norwegian University of Science and Technology. His research interests include connected and automated vehicles, traffic flow theory, and intelligent transportation systems.

Paper B:

Using Vehicle-Mounted Camera to Collect Information for Managing Mixed Traffic

Elnaz Namazi, Rein Nisja Holthe-Berg, Christoffer Skar Lofsberg, and Jingyue Li

International Conference on Signal-Image Technology & Internet-Based Systems (SITIS)

Using Vehicle-Mounted Camera to Collect Information for Managing Mixed Traffic

Elnaz Namazi
Department of Computer Science
Norwegian University of Science and Technology (NTNU)
Trondheim, Norway
elnaz.namazi@ntnu.no

Rein Nisja Holthe-Berg
Department of Computer Science
Norwegian University of Science and Technology (NTNU)
Trondheim, Norway
reinhb1994@gmail.com

Christoffer Skar Lofsberg
Department of Computer Science
Norwegian University of Science and Technology (NTNU)
Trondheim, Norway
chris.lofsberg@gmail.com

Jingyue Li
Department of Computer Science
Norwegian University of Science and Technology (NTNU)
Trondheim, Norway
jingyue.li@ntnu.no

***Abstract*—With increasingly rapid advances in the field of producing modern and autonomous vehicles, the need for intelligent traffic management systems, which take advantage of the vehicle's abilities to sense and communicate, has increased. A considerable amount of literature has been published on managing traffic that includes only autonomous vehicles. However, changing all vehicles to autonomous versions is a long-term process. In the near future, traffic will be a mixture of human-driven and autonomous vehicles. To date, few studies have investigated mixed traffic in intelligent management systems. The main objective of this research is to study the possibility of using a vehicle-mounted camera to sense and collect the required traffic data of the surrounding vehicles in mixed traffic. To achieve this, a vehicle with a monocular camera is used to collect image information for detecting and counting the vehicles in different lanes and estimating their distance and speed on the defined route. The results indicate that our proposed image processing algorithms can acquire the information needed for intelligent traffic management systems.**

Keywords-intelligent traffic management; autonomous vehicle; image processing; vehicle detection; speed estimation; distance estimation

I. Introduction

Recent developments in autonomous vehicle (AV) technology have heightened the need for intelligent traffic management systems that are suitable for AVs. Managing AV traffic has been studied by many researchers. They used AVs to collect and share information based on vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communications (e.g., [1], [2], [3], [4], [5], [6], and [7]). Despite the importance of managing AVs, changing all vehicles to autonomous versions will take time. Thus, we believe that traffic will be a mixture of human-driven vehicles (HDVs) and AVs for a long period. However, little attention has been paid to the mixed traffic that we will face in the coming years [8].

Intersections play a critical role in enhancing the efficiency and safety of traffic. Intelligent intersection management systems are introduced to manage traffic by using traffic data. There are various approaches to collecting traffic data. If traffic comprises purely AVs, AVs' sensors and V2X communication are used for collecting and sharing traffic data. However, for HDVs, such technologies are not applicable. Therefore, streets and intersections are equipped with various sensors such as inductive loop detectors and stationary cameras for collecting traffic information. Using these sensors can improve a traffic management system through accessing traffic data. However, weather conditions and brightness affect data quality. Furthermore, equipping all intersections and streets with these sensors would be costly.

To deal with these challenges of advanced cars or AVs, the idea of our study is to use vehicles' sensors to collect traffic data of AVs and HDVs. As the main focus of this work is to study the effect of a vehicle sensing and collecting traffic data, and we did not have access to an AV in our experiments, we used a vehicle-mounted camera to achieve our goals. This study seeks to identify how to use sensors, especially the mounted monocular camera, to collect the required data from multiple lanes. We limited our focus to the necessary data for managing traffic, especially at intersections. We developed algorithms to analyze the video data collected from a camera. We collected and analyzed real traffic information from a route in Trondheim to evaluate our algorithms.

Our data collection and analyses focus on answering the following research questions:

- RQ1: How can we identify the number and type of vehicles in front of a vehicle and in the nearby lanes using image data captured by a vehicle-mounted monocular camera?
- RQ2: What is the most accurate combination of width and height when calculating the distance of vehicles in front and in nearby lanes using image data captured by a vehicle-mounted monocular camera?
- RQ3: How can the speed of the detected vehicles be estimated by using image data captured by a vehicle-mounted monocular camera?

The experimental work presented here provides one of the first investigations on how to use a vehicle-mounted camera's ability to collect traffic data to improve traffic management systems at intersections by considering mixed traffic. The results indicate that autonomous vehicles can provide the required mixed traffic data such as number and type of vehicles, their distance, and their speed. However, more studies are needed to improve the accuracy of the outputs.

The remainder of this paper is organized as follows: Section II provides a brief overview of studies related to intersection management methodologies. Section III explains the research objective and approach. Section IV presents the implementation and evaluation of our proposed approach and algorithms. Section V discusses the advantages and remaining challenges of this study. Section VI concludes.

II. Background

Managing intersections plays a critical role in improving the performance and safety of traffic management systems. Developments in software, hardware, networks, and communications and the introduction of AVs have led to the use of intelligent systems to manage traffic. Therefore, more studies recognize the importance of an intelligent traffic management system that includes AV traffic. Many different methodologies have been presented. Most of the current literature has paid particular attention to using rule-based (e.g., [9] and [10]), optimization (e.g., [11], [12], and [13]), and hybrid (e.g., [14] and [15]) methodologies [8]. Moreover, to enhance the smartness of intersection management systems, artificial intelligence (AI) techniques can be applied (e.g., [16], [17], and [18]). Most of the AI-based intersection management research focuses on two issues: One is about decision-making and predicting the traffic situation based on the traffic data, which are collected by stationary sensors at the intersections. Another is to use AVs' AI capabilities to sense, collect, and share information about themselves in purely AV traffic. Our recent literature review [8] showed that in intelligent intersection management systems from 2008 to 2019, only 3.8% of the papers used AI to achieve their goals. However, AI-based traffic management systems have the potential to enhance traffic performance by improving the data collection process and predicting traffic features.

During the last few years, there has been a growing body of literature on using AI in object detection algorithms. Girshick et al. [19] proposed a region-based convolutional network (ConvNet) method (R-CNN) that used a deep ConvNet to classify the object proposals. It is a combination of region proposal and CNN. As the training phase is a multi-stage pipeline, it is costly in terms of space and time. Moreover, it is slow at detecting objects [20]. To mitigate the limitations of R-CNN, the fast region-based convolutional network method (Fast R-CNN) was presented by Girshick [20]. It used a deep ConvNet to classify object proposals efficiently. It tried to improve the training and testing speed and detection accuracy. Fast R-CNN was developed with Python and C++. The experiments indicate that Fast R-CNN trains the deep VGG16 network 19 times faster than R-CNN. Also, test time is 213 times shorter. In addition, it is more accurate.

Faster R-CNN was presented by Ren et al. to detect objects by considering region proposal networks (RPNs) [21]. Mask R-CNN, which is the extended version of Faster R-CNN, was proposed by He et al. [22]. It detects objects in an image efficiently and generates a high-quality segmentation mask simultaneously. Training with Mask R-CNN is simple and adds just a small overhead to Faster R-CNN.

Redmon presented You Only Look Once (YOLO) [23], which is a real-time object detection algorithm. In contrast with previous studies, YOLO is based on a regression problem rather than classification. It uses a single neural network for the detection pipeline. Moreover, classes and bounding boxes are predicted in one run of the algorithm for the whole image.

A number of researchers have considered using various proposed algorithms to detect vehicles, inter-vehicle distance, and vehicle speed. For instance, to detect vehicles, Godha [24] proposed an algorithm using a mounted camera in real-time that could send a warning to the driver. This system was developed in MATLAB as a driver assistant system. Asvadi et al. [25] proposed a real-time and multimodal vehicle detection system. It uses YOLO [23], [26] as a deep ConvNet object detection framework. Moreover, it is based on fusing the data collected by a color camera and 3D-LIDAR. The KITTI object detection dataset is used in the experiments phase. Caltagirone et al. [27] developed a fusion fully convolutional neural network (FCN) for road detection. It uses KITTI as a dataset and LiDAR and camera fusion.

In addition, some studies were done that focused on detection of the inter-vehicle distance. For example, Huang et al. [28] proposed a driver assistant system to detect vehicles and estimate the inter-vehicle distance. This system uses a camera as a sensor and includes image processing, information collection, vanishing point detection, road region segmentation, and estimation of the inter-vehicle distance. Lee [29] presented a method for estimating the inter-vehicle distance using a blackbox camera. The idea is to estimate the distance based on the lane width for the detected vehicle. Chadwick et al. [30] proposed an approach using radar and a camera to estimate the vehicle distance. Moreover, an automatic process was introduced for training and labeling the new dataset from multiple cameras. It used YOLO [23] as an object detector and KITTI as a dataset. Furthermore, several studies consider determining the vehicles' speed. For instance, Gerát et al. [31] used Gaussian mixture models, density-based spatial clustering of applications with noise (DBSCAN), a Kalman filter, and the optical flow method to detect vehicle speed using a stationary camera. Moazzam et al. [32]

proposed a new approach to determine vehicle speed based on video captured by a stationary camera. They used the QMUL dataset [33] for this experiment.

III. Methodology

A. Research objective

An intelligent intersection management system tries to improve traffic flow performance by accessing traffic data. In this study, we have tried to collect the number and type of vehicles, their distance, and their speed in mixed traffic by using a mounted monocular camera installed on a vehicle. We limited our focus to analyzing the monocular camera rather than more expensive sensors, e.g., radar or LiDAR, because we observe that many advanced vehicles have cameras installed by default and not many vehicles will have radar or LiDAR installed in the future. We believe that considering the data from a camera only will make our system more applicable. The objective of this work is to study the possibility of using a mounted monocular camera to collect mixed traffic data from multiple lanes by considering the effect of camera movement.

B. Research approach

This study is exploratory and normative in nature, since vehicle-mounted cameras have not been used to collect vehicle data to manage mixed traffic in existing studies, and new algorithms are developed. In this study, we followed Pfeffers et al.'s Design Science Research Process [34].

IV. Implementation and Evaluation

To answer the research questions, we decided to construct a system based on various state-of-the-art algorithms. The system was developed in Python and used popular frameworks that have well-documented outcomes in various projects. We proposed a system that is a combination of PyTorch [35] for implementation of the pre-trained version of YOLO [23], [36], [37] and OpenCV [38], Canny edge detection [39], and progressive probabilistic Hough transformation [40], [41] for lane detection while driving.

The data were collected from a vehicle equipped with a front-facing camera. We used a GoPro Hero 7 camera [42], since it is able to record GPS data as well. The video resolution was 1920 × 1080, the frame rate was set to 30 frames per second (FPS), and the GoPro had built-in video stabilization. Every 55 ms, the GPS sensor registered information including latitude, longitude, altitude, speed, and a coordinated universal time (UTC) stamp.

To collect data and to evaluate our data analysis algorithms, the route driven was defined by considering the coverage of various road types. For instance, a motorway with multiple lanes, city traffic with traffic lights, buses, and pedestrians, road sections with tunnels or roundabouts, and other mixed traffic were considered. The recording took place between 9 and 10 a.m. on a typical workday. The recorded video was split into manageable sequences. Moreover, the GPS data were extracted to a JSON file by an online tool [43]; then, the GPS file related to the video sequences was split.

In our evaluation, the system ran at around 10–15 FPS on a medium- to a high-end desktop computer with an Intel i7-7700k CPU and NVIDIA GTX 1080ti GPU. This gave a processing time of 60–100 ms per frame. Considering that the videos were captured at 30 FPS, this meant that the system performed at roughly half the speed of the videos themselves.

In the following, we will describe the approaches used to answer each research question and the outcomes.

A. RQ1. Detect number and type of vehicles in nearby lanes

We followed two steps to estimate the vehicles' positions and count them in each lane, namely, vehicle detection and lane detection. The first step was vehicle detection, which was done based on the existing object detection implementation called YOLO [23], [36], [37]. We chose to use YOLO because it is a real-time object detection algorithm. The selected implementation was trained on the COCO dataset [44]. We adapted YOLO to make it fit with our objectives. In the second step, image processing techniques were used to detect the lanes on the road. To achieve this goal, we experimented and compared various edge detection methods: Sobel edge detection [45], Canny edge detection [39], and Prewitt edge detection [46]. As Fig. 1 shows, Sobel has too much noise, and Prewitt is able to recognize only a few edges. Canny showed a good number of lane edges without much noise. Therefore, we decided to use Canny edge detection. In addition, to find the continuous lines, we decided to use progressive probabilistic Hough transform [40], [41], which provided great results for a small computing power cost. To further reduce the computation time, the system uses the grayscale image and regions of interest (ROIs) approach [47].

The steps applied in the lane detection algorithm are shown in Fig. 2. The output of Canny edge detection, cropping of the image, and progressive probabilistic Hough transform for lane detection are displayed in Fig. 3.

To answer RQ1, the findings of the vehicle detection and lane detection were merged and processed to yield information about detected vehicles and their relative positions. The output of the algorithm can be seen in Fig. 4. It shows the lanes and objects by using bounding boxes. It represents the number, speed, and distance of the detected bicycles, buses, vans, motorbikes, trucks, and cars in multiple lanes in front of the equipped vehicle with a monocular camera.

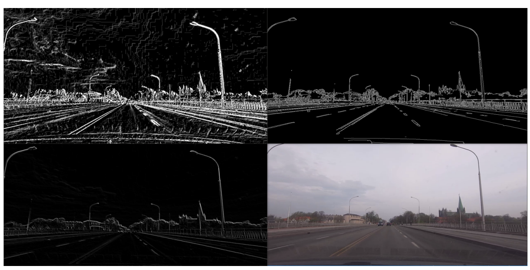


Figure 1. Comparison of different edge detection algorithms. Top left: Sobel edge detection; top right: Canny edge detection; bottom left: Prewitt edge detection; bottom right: the original image

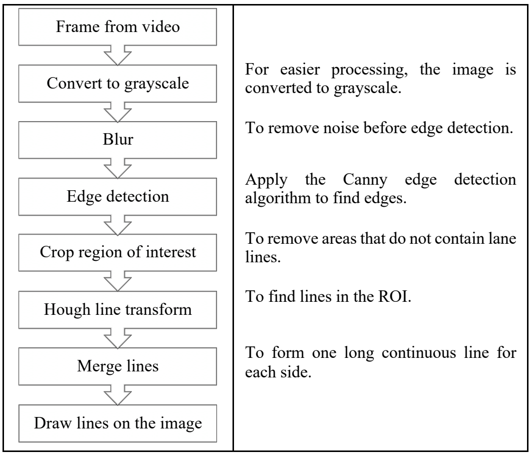
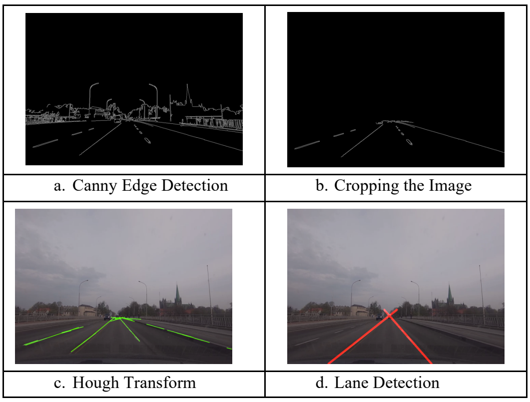


Figure 2. Steps of the lane detection algorithm



a. Canny Edge Detection
b. Cropping the Image
c. Hough Transform
d. Lane Detection

Figure 3. Outcomes of applying Canny edge detection, cropping the image, Hough transform, and lane detection

Figure 4. Lanes and vehicles detection on the road

To evaluate the proposed algorithm, three scenarios are selected with various durations and locations.

- S1. City traffic – Elgeseter Street, Trondheim
  - Includes several traffic light intersections, buses, and pedestrians
  - Video duration is equal to 4 minutes
  - 72 readings, giving a reading approximately every 3.5 seconds
- S2. Mixed traffic – Lade, Trondheim
  - Includes normal to heavy traffic, multiple traffic light intersections, crossing traffic, and surrounding parking lots
  - Video duration is equal to 3.5 minutes
  - 78 readings, giving a reading every 2.7 seconds
- S3. Mixed traffic – Tempe to Lerkendal, Trondheim
  - Includes normal to heavy traffic, and the lanes were separated by a central reservation with a medium-high fence
  - Video duration is equal to 1.5 minutes
  - 27 readings, giving a reading every 3.5 seconds

The evaluation of the algorithm is based on the comparison of the outputs with manually counted results. It is evaluated on two measures:

Measure 1: Overall ability to detect and count objects, not respecting the vehicle type.

Measure 2: Number of times vehicles were counted correctly and incorrectly in different lanes, respecting vehicle type.

The results obtained from the evaluation process of the proposed algorithm are presented in Tables I, II, III, and IV. The outputs show that the total error rate in S1, S2, and S3 is 1.0%–10.6% for measure 1. The total wrong on average for the proposed scenarios in measure 2 is 34.4%–46.3%. These findings show that the proposed algorithm is able to detect and count vehicles with high accuracy without considering their locations and types, but it is still not accurate if it focuses on identifying the type of the vehicle and its position in the lane.

B. *RQ2. Using a camera to estimate the distance*

We proposed a novel approach based on the pinhole camera geometry for calculating the distance of the vehicles in front in the same lane as the camera and in the left, right, and opposite lanes [48]. The pinhole camera is defined as equation 1, where d is the distance to the object, F_c is the focal length of the camera, H_α is the real height of the object, and h_α is the height of the image.

$$d = F_c \times \frac{H_\alpha}{h_\alpha} \quad (1)$$

We used a combination of height and width to estimate the vehicle's size and enhance the accuracy of the estimated distance. The values used for the calculations, based on approximate sizes of vehicles, are presented in Table V. Moreover, the distance estimation algorithm is shown in Fig. 5.

TABLE I. THE OUTPUT OF SCENARIO 1 - BASED ON MEASURE 1

S1	Lane			Total
	Left	***Mid***	***Right***	
Manual	99	49	50	198
System	85	51	60	196
Error	14.1%	4.1%	20.0%	1.0%

TABLE II. THE OUTPUT OF SCENARIO 2 - BASED ON MEASURE 1

S2	Lane			Total
	Left	***Mid***	***Right***	
Manual	265	56	159	480
System	228	70	131	429
Error	14.0%	25.0%	17.6%	10.6%

TABLE III. THE OUTPUT OF SCENARIO 3 - BASED ON MEASURE 1

S3	Lane			Total
	Left	***Mid***	***Right***	
Manual	35	39	24	98
System	40	21	31	92
Error	14.3%	46.2	29.2%	6.1%

TABLE IV. THE OUTPUT BASED ON MEASURE 2

Scenarios	Measure 2			Measure 2 (Total corrects on average)
	Counted too many in average	***Counted too few in average***	***Total wrongs calculation on average***	
S1	19.7%	14.7%	34.4%	65.6%
S2	24.3%	22.0%	46.3%	53.7%
S3	24.7%	10.0%	34.7%	65.3%

TABLE V. THE APPROXIMATE SIZES OF VEHICLES

Vehicle Type	Width	Height
Bus	2.4 m	4.0 m
Car	1.8 m	1.6 m
Motorbike/Bicycle	1.0 m	1.0 m
Truck	2.4 m	4.0 m
Van	1.9 m	2.5 m

As we did not have the ground truth of vehicle distance in our collected videos, we recorded new videos. The goal was to find the most accurate ratio of the vehicles' heights and widths. These videos captured a stationary vehicle at different distances and different angles. Then, we used a laser to measure the ground truth distance to stationary vehicles. We experimented with different ratios of height and width, and the results were compared with the ground truth from the laser. The average error with varying ratios of height and width is shown in Fig. 6. Based on the experiment results, the best ratio is 85% of the height and 15% of the width, which is affected by the detected vehicle angle.

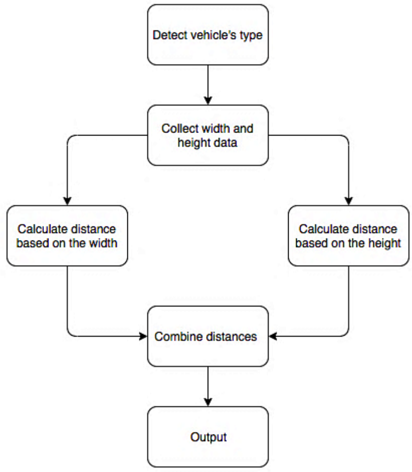


Figure 5. Distance estimation algorithm

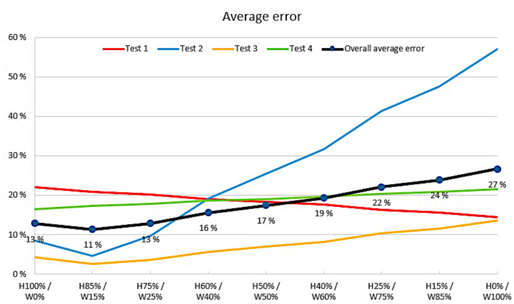


Figure 6. Average error with varying ratios of height and width

C. RQ3. Estimating the speed

Two steps were followed to answer the third research question: tracking the object in different frames and estimating the speed.

- How can we track vehicles between multiple frames? As we used a vehicle-mounted monocular camera, in a given period, the same vehicle could be viewed in the collected video. Then, the centroid of the bounding boxes identified by YOLO object detection and the Euclidean distance between a vehicle's centroids in different frames were used to track the same vehicle.

- How can we estimate the speed (υ)? Based on the physics concepts, distance traveled (Δd) over time (Δt) is needed. The formula is shown in equation 2.

$$\upsilon = \frac{\Delta d}{\Delta t} \tag{2}$$

Referring to the study of Chai and Wong [49], to calculate the speed, we used the known frame rate of the camera. As the FPS in this work is 30, to estimate the speed, the average change in distance over the last 30 frames was used to find the change per second. Moreover, the camera's speed based on GPS data was considered in the estimation process.

The proposed algorithms for object tracking between multiple frames and speed estimation are shown in Fig. 7 and Fig. 8, respectively.

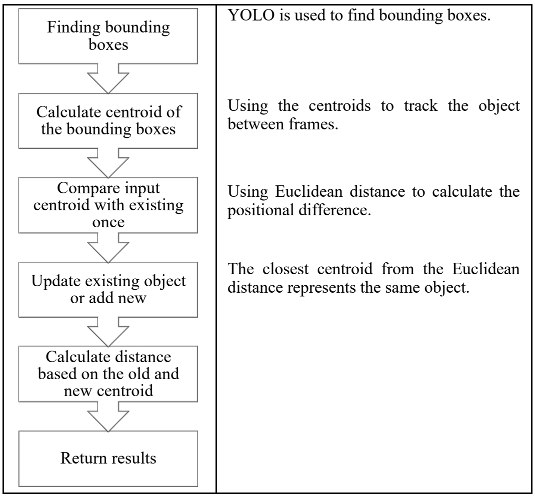


Figure 7. Steps for tracking objects

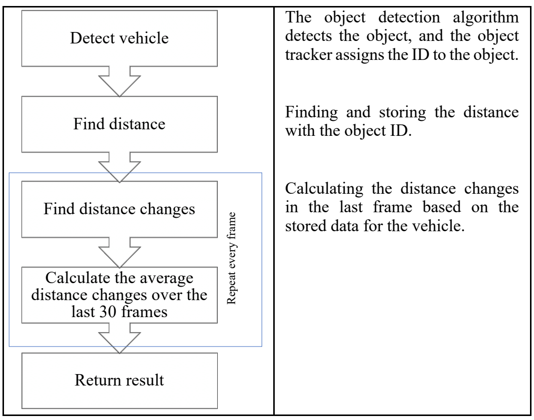


Figure 8. Speed estimation algorithm

To evaluate the algorithms, as we did not have equipment for measuring the true speed of the vehicles, we analyzed the three video sequences manually based on the estimated distance in RQ2 as the traveling distance (Δd) for each selected frame by considering the distance moved by the camera. The results are shown in Fig. 9, which shows the true speed, based on the calculation described in the test procedure, and the speed estimated by the system. The average difference of the 75 total manual readings across all the sequences was 2.09 m/s, and the maximum difference was 10.64 m/s.

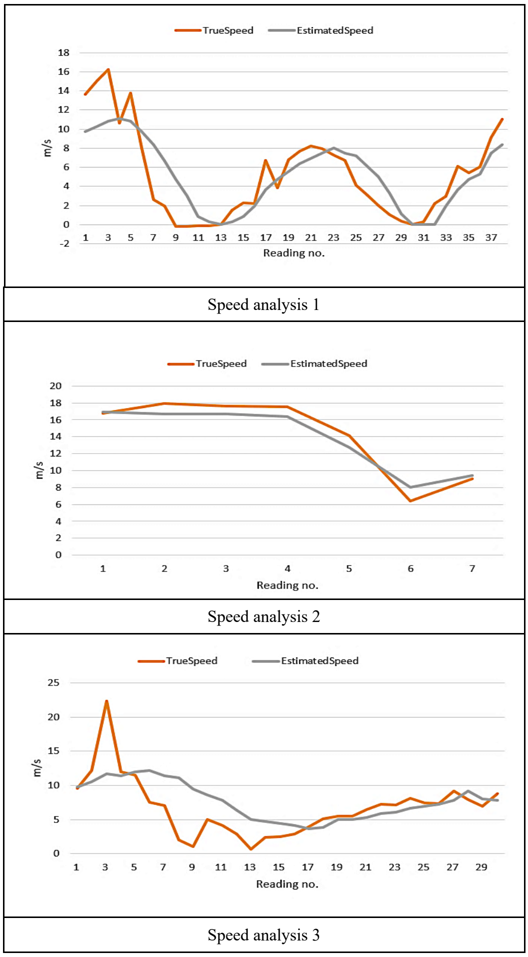


Figure 9. Speed analysis

V. Discussion

This study set out with the aim of assessing the feasibility of using vehicle-mounted sensors to sense the surrounding traffic to collect and share traffic data rather than using stationary sensors on the road. Prior studies focused on collecting traffic data that can be classified into two main groups. One group of papers used stationary sensors [31], [32]. While these methods might be effective in collecting the required data, they count AVs as HDVs without considering AVs' possibilities. Moreover, this methodology is not compatible with using a vehicle-mounted camera, as it does not consider the effect of the sensor's movement on the quality of detection. The other group of studies used vehicle-mounted sensors, which are the main focus of this study. Various autonomous and modern vehicles are equipped with different types of sensors. As there is a strong relationship between the type of sensors and a vehicle's price, some auto manufacturers might use a limited number of sensors to mitigate the vehicle's cost. Therefore, it is vital to find a method that is usable for all types of vehicle-mounted sensors. As the camera is the most common sensor, we tried to extract all the data from video. However, some studies focused on sensors' fusion, which is suitable for vehicles equipped with various types of sensors, such as LiDAR and RADAR (e.g., [25] and [27]). Although it might be effective in the correctness of detection, it is costly as well and might not be practically useful for all vehicles.

Moreover, unlike many other studies (e.g., [27], [25], and [32]), in this research, we equipped a vehicle with a camera and drove it on a defined route that contained many different scenarios to collect real traffic data.

In addition, one of the main goals of collecting traffic data is to improve the performance of intelligent traffic management systems. Therefore, considering the data type required by the traffic management system was a key point in developing our algorithms. To the best of our knowledge, no other paper has determined the number, type, distance, and speed of vehicles at the same time. Most of the papers have tried to extract one data type (e.g., [24], [25], [27], [30], [29], [31], and [32]), which is not enough for managing traffic safely and efficiently.

As a part of our research to determine the number and type of vehicles in each lane, we proposed algorithms for object detection and lane detection. The current study found that overall object detection with YOLO worked quite well. The results show that the average accuracy of vehicle detection is 92.4% in the proposed system. One unanticipated finding was the low accuracy of vehicle classification, which affected the accuracy of the position based on lanes. Almost 60% of the errors were caused by the detector misclassifying vehicles. A possible explanation for this might be that the network was trained on the COCO dataset, which contains 80 different objects, and not only objects related to traffic [44]. Hence, it could conceivably be hypothesized that using a specialized dataset for traffic objects could contribute to improving the classification accuracy, thus lowering the error rate and boosting the accuracy of the system. Moreover, the generated boundary boxed by YOLO may be unstable between frames. This result may be explained by the fact that YOLO's boundary detection approach leads to unstable detections. Non-maximum suppression could be used to fix these multiple detections [23].

Another important finding was the ability of the proposed system to detect lanes and separate vehicles in different lanes. In our evaluation of the system, this worked well on straight roads when the lane markings were clear and easily visible. However, overall lane detection results were not satisfactory with curved lanes. Edge detection with Canny edge worked as anticipated. However, sometimes the edge detector detected curbs as a lane edge. Contrary to expectations, the second-largest source of error was wrongly identified lanes. It would be hard to address some issues with lane detection, such as that faraway lanes are difficult to detect and distinguish, and bad or non-existent lane markings create some difficulties in the detection process. We suspect that, when choosing the lane detection part of the algorithm, progressive probabilistic Hough transform [40], [41] might not be the best choice. The method proposed by Kim [50] of tracking left and right lane markings separately and utilizing an ANN that was trained to detect lines could be used to increase the accuracy. This would also have the added benefit of being able to detect lanes with curves. However, using an ANN might increase the resource demand and would also require training. The use of standardized lane sizes in the algorithm could be another future enhancement. The Norwegian Public Roads Administration [51] handbook includes different standard sizes of lanes and markings. This could ensure that different lane detections were not bigger than a set threshold. The other observed limitation of the system related to lane detection is that the lane detected on the left side was often the lane with opposing traffic, which caused noise in the results.

Another interesting finding of our study concerned estimating the vehicle distance based on the object size. We used the pinhole model and proposed an efficient ratio of object width and length to estimate the vehicle distance. Our evaluations show that combining the height and width of the detected object with a ratio of 85% and 15%, which gave the lowest amount of error with 11% on average. The unexpected finding with this idea was that camera movement and the varying sizes of the generated bounding boxes affected distance estimation. This result may be explained by the fact that images with only 2D information from a camera were used for distance estimation. More advanced equipment such as LiDAR, which generates 3D information, might be more accurate. The other limitation is caused by assuming a fixed true value for vehicle size. Moreover, the accuracy of the distance estimation could be affected by wrongly classified vehicles. Furthermore, in this study, we proved that it is feasible to estimate vehicle speed using a moving monocular camera. Our study found that the speed estimation worked as expected for vehicles in front of the camera that were driving in the same direction as the camera, with a mean difference of 2.09 m/s. However, this experiment was not accurate enough for vehicles going the opposite direction as the camera. A possible explanation for these results may be the lack of adequate time to capture the vehicles driving in the opposite direction, which is needed to calculate the speed accurately. Movement of the camera caused some errors in estimating the speed. Additionally, the estimation was not that accurate for faraway vehicles, since their determined centroid points vanished. On the other hand, as speed was calculated based on distance, any error in distance estimation had a negative effect on the correctness of the estimated speed, and a significant error in estimated speed occurred if the estimated distance suddenly spiked or varied between consecutive frames.

VI. Conclusion and Future Work

This study set out to use vehicle-mounted monocular camera technology to collect the traffic data from multiple lanes required for managing traffic intelligently and efficiently. We tried to achieve this objective by answering three research questions. First, we defined a system based on object detection algorithms and computer vision methods. Experiments on the recorded images from a predefined route in Trondheim confirmed that the proposed approach worked well for object and lane detection in that specific situation. However, more studies are needed to enhance the accuracy of the outputs and generalize the system to various situations. The second aim of this study was to investigate the effects of considering both the length and the width of the detected object in estimating the distance. The results of this investigation show that combination height and width with the ratio of 85% and 15% worked best. The third purpose of the current study was to estimate the speed of the nearby vehicles based on their distance changes over time. In general, this study proved the possibility of collecting traffic data from a camera, which is useful for managing mixed traffic. Our future work will focus on improving the performance of the proposed algorithm to minimize the error rate in real traffic. To achieve this, we will extend the system to be able to work in a broader environment and include more lanes. In addition, we will try to collect more traffic data types, which is required for traffic management systems, considering the state of the art. We will also improve the accuracy of the proposed approach by improving the object detection algorithm and using a training dataset specific to traffic. In addition, we could improve the lane detection approaches by considering the lane width standards.

REFERENCES

[1] F. Yan, M. Dridi, and A. E. Moudni, “Autonomous vehicle sequencing algorithm at isolated intersections,” in 12th Int. IEEE Conf. Intell. Transp. Syst. , 2009, pp. 1–6.

[2] G. R. d. Campos, P. Falcone, and J. Sjöberg, “Autonomous cooperative driving: A velocity-based negotiation approach for intersection crossing,” in 16th Int. IEEE Conf. Intell. Transp. Syst. (ITSC 2013), 2013, pp. 1456–1461.

[3] S. Adams and M. J. Rutherford, “Towards decentralized waypoint negotiation,” in 2012 AAAI Workshop, Toronto, ON, vol. WS-12-10, pp. 2–5.

[4] Y. Zhang, A. A. Malikopoulos, and C. G. Cassandras, “Decentralized optimal control for connected automated vehicles at intersections including left and right turns,” in 2017 IEEE 56th Annu. Conf. Decision and Control (CDC), 2017, pp. 4428–4433.

[5] L. Makarem and D. Gillet, “Fluent coordination of autonomous vehicles at intersections,” in 2012 IEEE Int. Conf. Syst., Man, Cybern. (SMC), 2012, pp. 2557–2562.

[6] R. Krajewski, P. Themann, and L. Eckstein, “Decoupled cooperative trajectory optimization for connected highly automated vehicles at urban intersections,” in 2016 IEEE Intell. Veh. Symp. (IV), pp. 741–746.

[7] J. Ding, H. Xu, J. Hu, and Y. Zhang, “Centralized cooperative intersection control under automated vehicle environment,” in 2017 IEEE Intell. Veh. Symp. (IV), pp. 972–977.

[8] E. Namazi, J. Li, and C. Lu, “Intelligent intersection management systems considering autonomous vehicles: A systematic literature review,” IEEE Access, vol. 7, pp. 91946–91965, 2019.

[9] M. Vasirani and S. Ossowski, “A computational market for distributed control of urban road traffic systems,” IEEE Trans. Intell. Transp. Syst., vol. 12, no. 2, pp. 313–321, 2011.

[10] C. Wuthishuwong and A. Traechtler, “Consensus-based local information coordination for the networked control of the autonomous intersection management,” Complex Intell. Syst., vol. 3, no. 1, pp. 17–32, Mar 2017.

[11] S. A. Fayazi, A. Vahidi, and A. Luckow, “Optimal scheduling of autonomous vehicle arrivals at intelligent intersections via MILP,” in 2017 American Control Conf. (ACC), pp. 4920–4925.

[12] J. Wu, A. Abbas-Turki, and A. E. Moudni, “Intersection traffic control by a novel scheduling model,” in 2009 IEEE/INFORMS Int. Conf. Service Operations, Logistics, Inform., pp. 329–334.

[13] J. Wu, A. Abbas-Turki, and A. E. Moudni, “Discrete methods for urban intersection traffic controlling,” in VTC Spring 2009 – IEEE 69th Veh. Techn. Conf., pp. 1–5.

[14] F. Ze-hua, L. Han-bo, H. Wei, and Y. Tian, “District control strategy for vehicle collision avoidance based on hybrid automaton model,” in Proc. 11th World Congr. Intell. Control Autom., 2014, pp. 5378–5383.

[15] Y. Bichiou and H. A. Rakha, “Developing an optimal intersection control system for automated connected vehicles,” IEEE Trans. Intell. Transp. Syst., vol. 20, no. 5, pp. 1908–1916, 2019.

[16] Y. Wu, H. Chen, and F. Zhu, “DCL-AIM: Decentralized coordination learning of autonomous intersection management for connected and automated vehicles,” Transp. Res. Part C: Emerg. Technol, vol. 103, pp. 246–260, 2019.

[17] Y. Mo, M. Wang, T. Zhang, and Q. Zhang, “Autonomous cooperative vehicle coordination at road intersections,” in 2018 IEEE Int. Conf. Commun. Syst. (ICCS), pp. 192–197.

[18] N. Aloufi and A. Chatterjee, “Autonomous vehicle scheduling at intersections based on production line technique,” in 2018 IEEE 88th Veh. Techn. Conf. (VTC-Fall), pp. 1–5.

[19] R. Girshick, J. Donahue, T. Darrell, and J. Malik, “Rich feature hierarchies for accurate object detection and semantic segmentation,” in Proc. IEEE Conf. Comput. Vision Pattern Recognition, 2014, pp. 580–587.

[20] R. Girshick, “Fast r-cnn,” in Proc. IEEE Int. Conf. Comput. Vision, 2015, pp. 1440–1448.

[21] S. Ren, K. He, R. Girshick, and J. Sun, “Faster r-cnn: Towards real-time object detection with region proposal networks,” Advances Neural Inf. Process. Syst., 2015, pp. 91–99.

[22] K. He, G. Gkioxari, P. Dollár, and R. Girshick, “Mask r-cnn,” in Proc. IEEE Int. Conf. Comput. Vision, 2017, pp. 2961–2969.

[23] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You only look once: Unified, real-time object detection,” in Proc. IEEE Conf. Comput. Vision Pattern Recognition, 2016, pp. 779–788.

[24] S. Godha, “On-road obstacle detection system for driver assistance,” Asia Pacific J. Engineering Science and Technology, vol. 3, no. 1, pp. 16–21, 2017.

[25] A. Asvadi, L. Garrote, C. Premebida, P. Peixoto, and U. J. Nunes, “Multimodal vehicle detection: fusing 3D-LIDAR and color camera data,” Pattern Recognition Lett., vol. 115, pp. 20–29, 2018.

[26] J. Redmon and A. Farhadi, “YOLO9000: Better, faster, stronger,” in Proc. IEEE Conf. Comput. Vision Pattern Recognition, 2017, pp. 7263–7271.

[27] L. Caltagirone, M. Bellone, L. Svensson, and M. Wahde, “LIDAR–camera fusion for road detection using fully convolutional neural networks,” Robot. Auton. Syst., vol. 111, pp. 125–131, 2019.

[28] D.-Y. Huang, C.-H. Chen, T.-Y. Chen, W.-C. Hu, and K.-W. Feng, “Vehicle detection and inter-vehicle distance estimation using single-lens video camera on urban/suburb roads,” J. Visual Commun. Image Representation, vol. 46, pp. 250–259, 2017.

[29] J. Lee, “Intervehicle distance estimation through camera images,” J. Electron. Imag., vol. 27, no. 6, p. 063001, 2018.

[30] S. Chadwick, W. Maddern, and P. Newman, “Distant vehicle detection using radar and vision,” arXiv preprint, arXiv:1901.10951, 2019.

[31] J. Gerát, D. Sopiak, M. Oravec, and J. Pavlovicová, “Vehicle speed detection from camera stream using image processing methods,” in 2017 Int. Symp. ELMAR, pp. 201–204: IEEE.

[32] M. G. Moazzam, M. R. Haque, and M. S. Uddin, “Image-based vehicle speed estimation,” J. Comput. Commun., vol. 7, no. 6, pp. 1–5, 2019.

[33] QMUL Junction Dataset, Chinese University of Hong Kong, Feb. 2018. [Online]. Available: http://personal.ie.cuhk.edu.hk/~ccloy/downloads_qmul_junction.html

[34] K. Pfeffers, T, Tuunanen, C. Gengler, M. Rossi, W. Hui, V. Virtanen, and J. Bragge, “The design science research process: A model for producing and presenting information systems research,” in Proc. 1st Int. Conf. Design Sci. Research Inf. Syst. Techn. (DESRIST 2006), Claremont, CA, USA, pp. 83–106.

[35] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer “Automatic differentiation in pytorch,” 31st Conf. on Neural Information Processing Systems (NIPS), 2017.

[36] J. Redmon and A. Farhadi, “Yolov3: An incremental improvement,” arXiv preprint, arXiv:1804.02767, 2018.

[37] Ayooshkathuria/pytorch-yolo-v3. [Online]. Available: https://github.com/ayooshkathuria/pytorch-yolo-v3

[38] G. Bradski, “The opencv library,” Dr Dobb’s J. Software Tools, vol. 25, pp. 120–125, 2000.

[39] L. Ding and A. Goshtasby, “On the Canny edge detector,” Pattern Recognition, vol. 34, no. 3, pp. 721–725, 2001.

[40] C. Galamhos, J. Matas, and J. Kittler, “Progressive probabilistic Hough transform for line detection,” in Proc. 1999 IEEE Comput. Soc. Conf. Comput. Vision Pattern Recognition (Cat. No PR00149), vol. 1, pp. 554–560: IEEE.

[41] J. Matas, C. Galambos, and J. Kittler, “Robust detection of lines using the progressive probabilistic hough transform,” Comput. Vision Image Understanding, vol. 78, no. 1, pp. 119–137, 2000.

[42] “Gopro hero 7 black.” [Online]. Available: https://shop.gopro.com/EMEA/cameras/hero7-black/CHDHX-701-master.html.

[43] “Gopro telemetry extractor.” [Online]. Available: https://tailorandwayne:com/gopro-telemetry-extractor/#!.

[44] T.-Y. Lin, M. Maire, S. Belongie, L. Bourdev, R. Girshick, J. Hays, P. Perona, D. Ramanan, C. Zitnick, and P. Dollar´, “Microsoft COCO: Common objects in context,” in European Conf. Comput. Vision, 2014, pp. 740–755: Springer.

[45] S. Gupta and S. G. Mazumdar, “Sobel edge detection algorithm,” Int. J. Comput. Sci. Manag. Res., vol. 2, no. 2, pp. 1578–1583, 2013.

[46] A. Seif, M. M. Salut, and M. N. Marsono, “A hardware architecture of Prewitt edge detection,” in 2010 IEEE Conf. Sustain. Utilization and Develop. Eng. Technol., pp. 99–101.

[47] G. Deng and Y. Wu, “Double lane line edge detection method based on constraint conditions Hough transform,” in 2018 17th Int. Symp. Distrib. Comput. Appl. Business Eng. Sci. (DCABES), pp. 107–110: IEEE.

[48] K.-Y. Park and S.-Y. Hwang, “Robust range estimation with a monocular camera for vision-based forward collision warning system,” Scientific World J., vol. 2014, 2014.

[49] C. Chai and Y. D. Wong, “Automatic vehicle classification and tracking method for vehicle movements at signalized intersections,” in 2013 IEEE Intell. Veh. Symp. (IV), pp. 624–629.

[50] Z. Kim, “Robust lane detection and tracking in challenging scenarios,” IEEE Transactions on Intelligent Transportation Systems, vol. 9, pp. 16-26, 2008.

[51] Vegdirektoratet, “Vegoppmerking,” 2015. [Online]. Available: https://www:vegvesen:no/ attachment/69741

Paper C:

Identifying and Counting Vehicles in Multiple Lanes by Using a Low-Cost Vehicle-Mounted Sensor for Intelligent Traffic Management Systems

Elnaz Namazi, Jingyue Li, Rudolf Mester, and Chaoru Lu

International Conference on Hybrid Artificial Intelligence Systems (HAIS)

Identifying and Counting Vehicles in Multiple Lanes by Using a Low-Cost Vehicle-Mounted Sensor for Intelligent Traffic Management Systems

Elnaz Namazi[1[0000-0002-6503-9315]], Jingyue Li[1[0000-0002-7958-391X]],

Rudolf Mester[1[0000-0002-6932-0606]], and Chaoru Lu[2]

[1] Department of Computer Science, Norwegian University of Science and Technology (NTNU), Trondheim, Norway
{elnaz.namazi, jingyue.li, rudolf.mester}@ntnu.no

[2] Department of Civil and Environmental Engineering, Norwegian University of Science and Technology (NTNU), Trondheim, Norway
chaoru.lu@ntnu.no

Abstract. There is evidence that accessing online traffic data is a key factor to facilitate intelligent traffic management, especially at intersections. With the advent of autonomous vehicles (AVs), new options for collecting such data appear. To date, much research has been performed on machine learning to provide safe motion planning and to control modern vehicles such as AVs. However, few studies have considered using the sensing features of these types of vehicles to collect traffic information of the surrounding environment. In this study, we developed new algorithms to improve a traffic management system when the traffic is a mixture of human-driven vehicles (HDVs) and modern vehicles with different levels of autonomy. The goal is to utilize the sensing ability of modern vehicles to collect traffic data. As many modern vehicles are equipped with vehicle-mounted sensors by default, they can use them to collect traffic data. Our algorithms can detect vehicles, identify their type, determine the lane they are in, and count the number of detected vehicles per lane by considering multi-lane scenarios. To evaluate our proposed approach, we used a vehicle-mounted monocular camera. The experimental work presented here provides one of the first investigations to extract real traffic data from multiple lanes using a vehicle-mounted camera. The results indicate that the algorithms can identify the detected vehicle's type in the studied scenarios with an accuracy of 95.21%. The accuracy of identifying the lane the detected vehicle is in is determined by two proposed approaches, which have accuracies of 91.01% and 91.73%.

Keywords: Lane Detection, Multiple Lanes, Vehicle Detection, Intelligent Traffic Management, Vehicle-Mounted Monocular Camera.

1 Introduction

There is a growing body of literature that recognizes the importance of collecting traffic data in intelligent traffic management systems. Developments in machine learning techniques and sensors' capabilities have led to proposing various approaches for collecting different types of traffic data (e.g., [1]). These data can be used to manage traffic safely and efficiently, especially at intersections [2]. When focusing on intersection management systems, detecting vehicles' types [3], identifying the lanes they are in, and counting the number of vehicles per lane are vital to provide a global view of the intersection to manage the traffic with high performance.

Previous research on collecting traffic data has mostly used stationary sensors, which are affected by the brightness and weather condition, besides having high installation and maintenance costs. Moreover, equipping all streets with these types of sensors can be costly. The main contribution of our research is taking advantage of the sensing capabilities of modern vehicles, e.g., AVs, which are equipped with various types of sensors, to collect data of the surrounding vehicles to manage traffic. Moreover, this idea is reachable in pure AVs traffic and mixed traffic (a combination of HDVs and AVs), as managing mixed traffic is one of the most important issues for the near future, since changing all vehicles to autonomous versions will be a time-consuming process. Even after this period, traffic might include HDVs as well, because some people enjoy driving. Another contribution of this research is proposing an approach which is generalizable with various levels of vehicle autonomy. Therefore, we used a vehicle-mounted monocular camera, which is one of the cheapest sensors, so there is a high probability that most modern vehicles will be equipped with one. Moreover, by using the camera vision, we are able to record video from multiple lanes. Therefore, we used the camera data to analyze the surrounding traffic. Our developed algorithms are able to detect and classify vehicles in multiple lanes, detect the lanes next to the equipped vehicle, determine the location of the detected vehicles, and count the number of vehicles in each lane. By accessing this information and sharing it with traffic management systems, these systems would have a better global view of the environment and would be able to make better traffic management decisions, especially at intersections.

Our proposed algorithms attempt to answer two research questions:

- **RQ1.** How can we enhance the accuracy of detected vehicles' types based on existing object detection algorithms?
- **RQ2.** How can we identify the lane the detected vehicle is in on multi-lane streets to estimate the number of vehicles in each lane?

The remainder of the paper proceeds as follows. The next section summarizes related works. Chapter 3 explains the research methodology used in this study. The implementation to answer the proposed RQs is described in chapter 4. Chapter 5 presents the experimental results on real traffic data. The last chapter discusses the findings and concludes.

2 State of the Art

In the past few years, a considerable amount of literature has been published on vehicle detection, lane detection, lane-keeping, and tracking for driver assistant systems (e.g., [4]).

Target detection algorithms can be classified into three categories [5]. The first category is the digital image processing approach, such as the frame difference (FD) approach. The second one is a machine learning approach, which is usually based on an AdaBoost classifier or support vector machine (SVM). The last category is based on deep learning approaches. The proposed algorithms in this group are based on convolution neural networks (CNN), Fast-RCNN, Faster-RCNN, YOLO (You Only Look Once), etc. [5].

To improve the object detection performance, Tian et al. [5] proposed a hybrid method, which combined the FD method and YOLO. The results show that this approach can improve the bounding boxes' precision. Moreover, they introduced a model to estimate the distance and speed of the targets based on video from a stationary monocular camera in real time. To detect and track objects and estimate distance and motion in real time, Chen et al. [4] proposed an approach based on deep learning. First, they compared YOLOv3 with a single shot detector (SSD). Second, their object distance estimation was developed based on the Monodepth algorithm. Third, they proposed a new method to analyze object behavior based on SSD. To validate the proposed methodology, they used real traffic from a city center and a railway.

Moreover, different methodologies have been proposed for lane detection. Hillel et al. classified the purpose of lane understanding into lane departure warning, adaptive cruise control (ACC), lane keeping, lane centering, lane change assist, turn assist, fully autonomous driving for paved roads, and fully autonomous driving for cross-country trips [6]. Lane boundary tracking generally includes three major steps [7]. The first step is lane marking detection. In this step, various types of sensors, such as a camera (e.g., [8]), lidar (e.g., [9]), radar, GPS (e.g., [10]), and a line sensor camera (e.g., [11]), can be used. The second step is lane boundary estimation, which includes position, object type, lane information, and vehicle information. The last step is lane boundary tracking. In this step, different filtering approaches such as a Kalman filter, extended Kalman filter, unscented Kalman filter, and particle filter are used [7].

Jo et al. [12] proposed a new method to build an accurate lane-level road map based on a stereo camera, GPS, and in-vehicle sensors. The lane map generation process includes two main steps. The first step is pre-processing, which includes global optimization, ego-motion estimation, and lane detection. The second step includes coordination conversion, clustering, and polyline fitting. Jia et al. [13] proposed a sequential monocular road detection algorithm. The algorithm is classified into sequential road modeling, probabilistic segmentation, and boundary refinement. The current image, previous image, and previous road maps are the input to this process, and the current road map is its output. The multi-lane detection approach is proposed by Chao et al. based on the deep convolutional neural network. The full connected network (FCN) is applied to the captured image by the monocular camera to extract the lane boundary feature. On the image, perspective transform, Hough transform, and the least square method are applied for the lane fitting [14].

Cao et al. [15] proposed a lane detection algorithm that considered dynamic environments and complex road conditions. It is based on the superposition threshold algorithm and the random sample consensus (RANSAC) algorithm. Another approach proposed color-based segmentation for lane detection; it used global convolution networks (GCN), residual-based boundary refinement, and Adam optimization [16]. Yuan et al. introduced a new approach to segmentation and lane detection [17]. It was based on a normal map, an adaptive threshold segmentation method, denoising operations, Hough transform, and the vanishing point.

3 Research Methodology

3.1 Research strategy

A case study approach was chosen to evaluate the effectiveness of the proposed algorithms with real traffic in an urban area. A vehicle-mounted monocular camera was driven on a predefined path in Trondheim, Norway. For the purpose of data analysis, the recorded video was divided into smaller scenarios. Five scenarios were selected by considering the situation coverage and the research scope. The studied scenarios are presented in Table 1.

Table 1. Scenarios.

Scenarios	Description	Total frames
S1	Includes streets with 4 lanes and 3 lanes (1 left and 2 right).	994
S2	Includes streets with 4 lanes and 1 reserved lane in the center, 1 four-way intersection with a red traffic light and 2 traffic lights at two-way intersections.	533
S3	Includes a 4-lane street, 1 red traffic light at a four-way intersection, 1 green traffic light at a four-way intersection, and 1 red traffic light at a two-way intersection.	2249
S4	Includes a 4-lane street with a guardrail in the center, 1 green traffic light at a curved four-way intersection, 1 red traffic light at a curved intersection, and 1 red traffic light at a four-way intersection.	1819
S5	Includes 4-lane and 2-lane streets and 1 red traffic light at a three-way intersection.	2278

3.2 Data collection

To test our proposed algorithms with real traffic, we decided to record our own footage. Therefore, we equipped a vehicle with a front-facing GoPro Hero 7 camera [18]. The video resolution and frame rate were 1920 × 1080 and 30 frames per second (FPS), respectively. The GPS information includes latitude, longitude, altitude, speed, and a UTC stamp.

The equipped vehicle was driven along the predefined path in Trondheim, Norway, between 9 and 10 a.m. on a typical workday. In this experiment, we focused on city traffic with various numbers of lanes, intersections, and traffic lights.

The recorded video was split into small scenarios to be analyzable, and one frame was analyzed in every 30. The experiments were run using a desktop computer with an Intel Core i7-4770k CPU 3.40 GHz × 8 and Intel Haswell Desktop graphics.

The data telegram is defined as follows:

- Type of the detected vehicles
- Location of the detected vehicles on the multi-lane streets
- Number of vehicles in each lane

4 Implementation

By extending existing vehicle detection and lane detection algorithms, the proposed method is able to extract the information of the traffic surrounding the camera-mounted vehicle. Several existing algorithms and libraries have been widely applied for vehicle detection and classification, such as YOLO ([24, 25]), PyTorch [19], and OpenCV [20]. Since YOLO is able to run in real-time vehicle detection and classification based on the global context in the image and a single network evaluation [27], it has the potential to provide traffic information to help with real-time traffic management systems [5]. In order to detect lanes, the results of comparing three different edge detection algorithms—Sobel edge detection, Canny edge detection, and Prewitt edge detection—show that Canny edge detection is able to detect the required lanes with less noise than the other two [3]. Therefore, in this paper, we used Canny edge detection [21] and progressive probabilistic Hough transform [22, 23] to deal with lane detection.

The major goal of this paper is to propose a method which can provide lane-based traffic information by extracting data from video via a vehicle-mounted monocular camera. In our last paper [3], we proved that a vehicle-mounted monocular camera can collect traffic data, such as the speed and distance of the detected vehicles. However, traffic management systems need more detailed information on each lane. In this paper, we focused on localizing the detected vehicles in each lane.

4.1 RQ1. Vehicle type detection

As we mentioned before, we used YOLO to do vehicle detection and classification. YOLO was originally trained on the COCO dataset, which includes 80 object categories, such as car, cat, umbrella, cell phone, etc. Therefore, the accuracy of the model is not good enough to extract real-world traffic data [3]. Since the traffic management only requires traffic objects, a pre-trained weight on the KITTI dataset was used to train YOLO to enhance its accuracy in classifying traffic objects. The KITTI dataset focuses on traffic objects and contains eight categories named car, van, truck, pedestrian, person_sitting, cyclist, tram, and misc. [28]. The proposed system architecture is shown in Fig. 1. As shown in Fig. 1, the input of the system is the recorded videos from real-world traffic, as described in section 3.2. The algorithm is based on YOLO trained on the KITTI dataset. Moreover, the output of the system is the processed videos. In these videos, bounding boxes are drawn around the detected vehicles, and the types of detected vehicles are identified. Moreover, lane markers are detected and highlighted. This information is recorded in JSON files for further analysis.

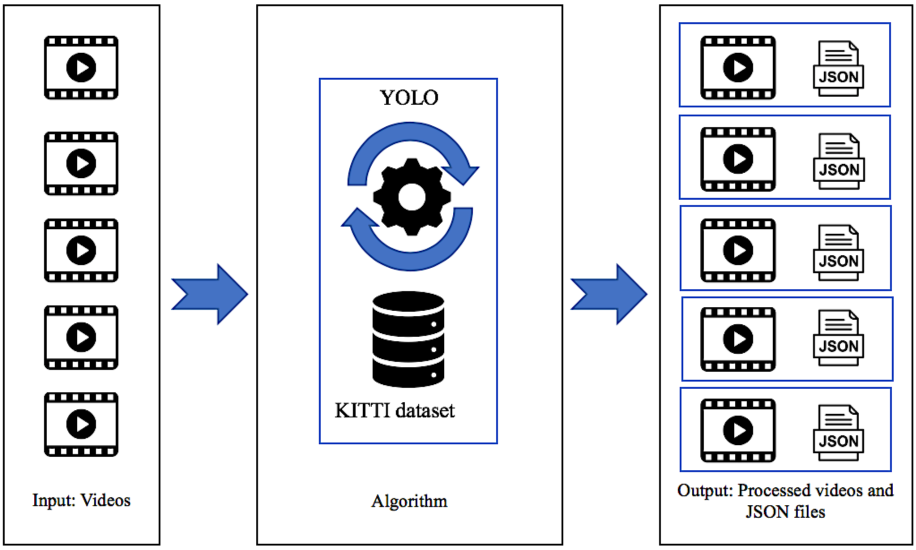


Fig. 1. System architecture.

4.2 RQ2. Extracting traffic data

To answer this research question, we followed three steps, as shown in Fig. 2. The first step was to identify the nearby lanes on both sides of the equipped vehicle. To do this we converted the extracted frames to grayscale to reduce the processing time. To remove the noise, frames were blurred. After that, as we mentioned in section 4, we used Canny edge detection [21], and the regions of interest (RoI) [29] to reduce the computation time. Moreover, progressive probabilistic Hough transform [22, 23] is applied to detect lines. After that, lines were drawn on top of the frames, which are shown in green in Fig. 3 and Fig. 4. The parametrization for the detected lines is based on the starting point (x_1, y_1) and ending point (x_2, y_2) of the line in the defined RoI.

The second step is detecting vehicles and dividing them into three groups. To do this, based on the distance between a central point on the bottom side of the bounding boxes around the detected vehicles and detected lanes, we classified vehicles into three groups, named left, middle, and right. To classify the vehicles, we followed these rules: If the vehicles were driven in the same lane as the equipped vehicle, we classified them as middle; if they were to the left side of that vehicle, we classified them as left; and others were classified as right. The conditions to make these decisions are shown in Table 2. This table includes three figures, in which green lines are the detected lanes on both sides of the equipped vehicle; they are named the left line (LL) and right line (RL). Bounding boxes around the detected vehicle are shown as a red rectangle. The central point on the bottom side of the bounding box is named “central point” (CP). Blue arrows represent the conditions, which are called left of the left line (LoL), left of the right line (LoR), right of the left line (RoL), and right of the right line (RoR).

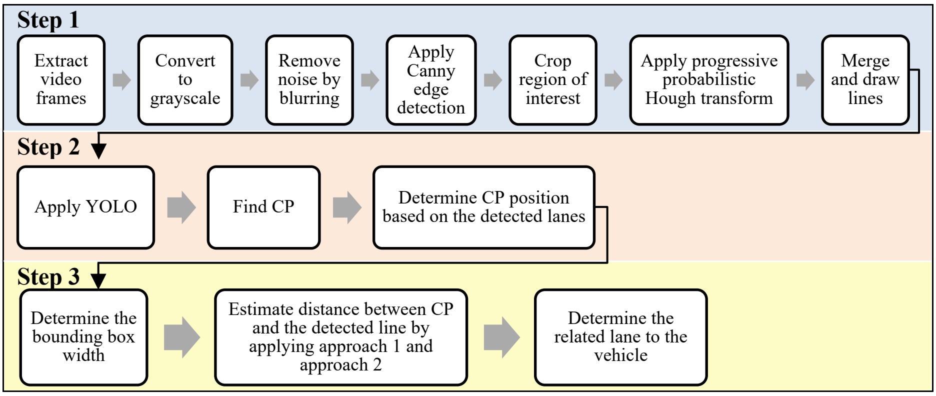


Fig. 2. Lane detection, object detection, and location estimation.

In the third step, we identify the location of each vehicle in multiple lanes. The idea is based on the assumption that vehicle size is less than the lane width. So, the vehicle location is identified based on the distance between the CP in the bounding box around the detected vehicle and the detected lane which that vehicle is in. The distance is measured by two proposed approaches as follows.

Approach 1.

In the first approach, we estimate the location of the detected vehicle based on the shortest path between the CP and the related lane. The shortest distance between a point and a line which is defined by two points, is presented in equation (1) [30]. The distance (D_i) of the point CP on the bounding box around the vehicle i, which is expressed by ($x_{vi,0}$, $y_{vi,0}$) from the line which passes through two points, $P_1:=(x_1,y_1)$ and $P_2:=(x_2,y_2)$, is as follows:

$$\text{distance}\left((P_1,P_2),\ (x_{vi,0}\,,\,y_{vi,0})\right)=\frac{\left|(x_2-x_1)(y_1-y_{vi,0})-(x_1-x_{vi,0})(y_2-y_1)\right|}{\sqrt{(x_2-x_1)^2+(y_2-y_1)^2}} \tag{1}$$

This approach is presented in Fig. 3. In this figure, similar to Table 2, green lines are the detected lanes on both sides of the equipped vehicle, called LL and RL. Red rectangles are bounding boxes around the detected vehicle. CP represents the central point on the bottom side of the bounding box. Blue arrow which is called D_i, shows the shortest distance between a CP on vehicle i and a related line. W_{vi} shows the width of the vehicle i.

Approach 2.

In this approach, we propose a solution to estimate the vehicle distance (d_i) to the related line in the horizontal direction, as shown in Fig. 4. Other variables are named as in Fig. 3.

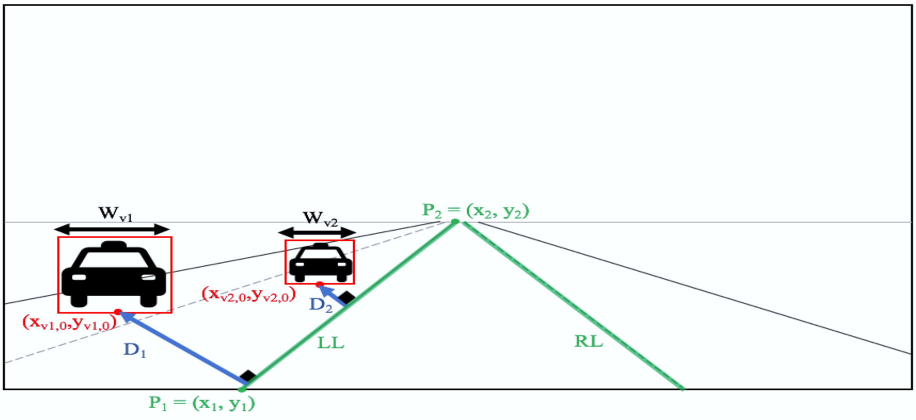


Fig. 3. The first approach to estimating the lane the detected vehicle is in.

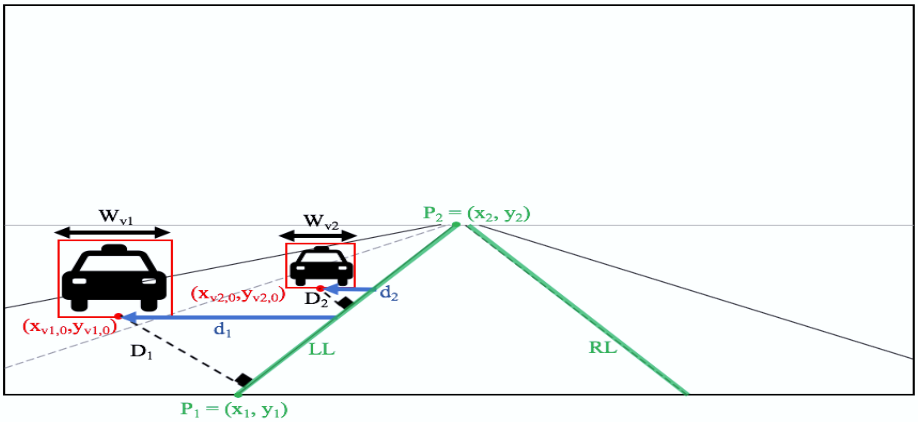
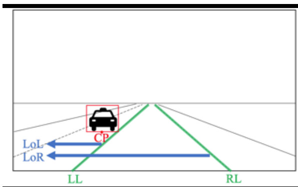
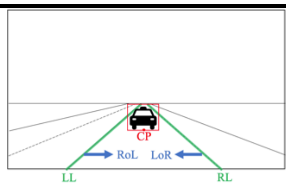
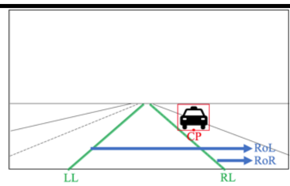


Fig. 4. The second approach to estimating the lane the detected vehicle is in.

Table 2. Dividing vehicles into three main groups, left, middle, and right.

		
If the CP is located on the left side of the left line and on the left side of the right line, then the vehicle is on the left side.	If the CP is located on the right side of the left line and on the left side of the right line, then the vehicle is in the middle.	If the CP is located on the right side of the left line and on the right side of the right line, then the vehicle is on the right side.

To measure d_i, our proposed approach consists of the following steps.

1- Measuring the slope of the related line (jL, j:=L or R), which passes through two points, P_1 and P_2 [31].

$$Slope_{jL} = \frac{(y_2 - y_1)}{(x_2 - x_1)} \tag{2}$$

2- Converting the line's slope to an angle in degrees [31].

$$jL_{degree} = \arctan\left(Slope_{jL}\right) \tag{3}$$

3- Estimating d_i by using triangulation formulas, as shown in Fig. 5.

Based on Euclidean parallelism [26],

$$L \parallel d_i \quad \Rightarrow \quad \beta = \propto = \gamma = N_{degree} \tag{4}$$

$$D_i \perp N \quad \Rightarrow \quad \lambda = 90° \tag{5}$$

By considering the triangle rules [32],

$$\theta + \lambda + \gamma = 180° \;\Rightarrow\; \theta = 180° - 90° - \gamma \quad \Rightarrow \quad \theta = 90° - N_{degree} \tag{6}$$

Based on the trigonometric ratios, the hypotenuse (d_i) is calculated by the following formula [33]:

$$d_i = \frac{D_i}{\cos(\theta)} \quad \Rightarrow \quad d_i = \frac{D_i}{\cos(90° - N_{degree})} \tag{7}$$

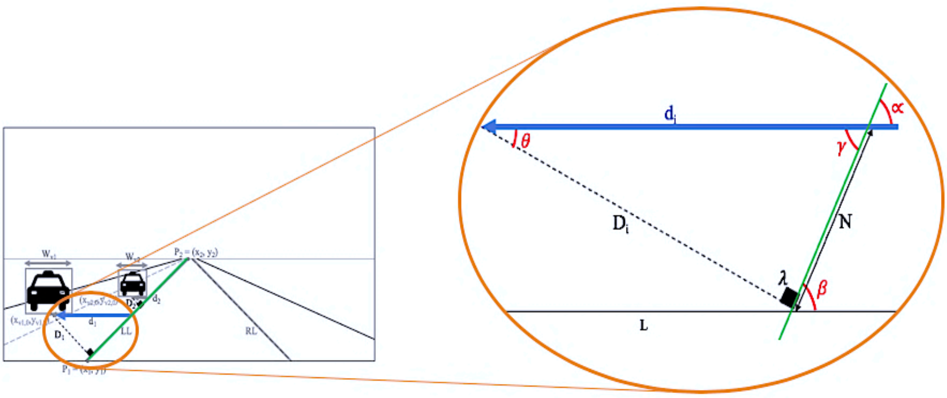


Fig. 5. Identifying the vehicle's lane by the second approach.

Finally, as the last step, the location of the vehicle is estimated by considering the distance and vehicle size, as shown in Table 3, in which $Distance_i$ is the distance calculated for a vehicle i by following approach 1 and approach 2, and S_{vi} is the size of vehicle i.

Table 3. Conditions for finding the detected vehicle's location in multiple lanes.

Condition	Output
$0 < \text{Distance}_i < S_{vi}$	1st lane on the left/right
$S_{vi} < \text{Distance}_i < 2 \times S_{vi}$	2nd lane on the left/right
$2 \times S_{vi} < \text{Distance}_i < 3 \times S_{vi}$	3rd lane on the left/right
$(n-1) \times S_{vi} < \text{Distance}_i < n \times S_{vi}$	nth lane on the left/right

We ran our algorithms on predefined scenarios and extracted the frames. Out of every 30 frames, we analyzed one frame manually as a ground truth. In this study, the scenarios include 7873 frames in total, and we analyzed 262 frames. Then, the outputs of the algorithms were compared with the manually extracted data.

5 Results

The purpose of the first experiment was to determine the accuracy of the improved algorithms in identifying the type of the detected vehicles. Table 4 illustrates our results. It is apparent from this table that the accuracy of identifying the detected vehicles' type is higher than 90.74% for all lanes in the studied scenarios.

Table 4. Vehicle type detection in the predefined scenarios.

		2nd lane on the left	1st lane on the left	Middle	Right	Total
Scenario 1	Manual	5.00	19.00	22.00	2.00	48.00
	System	5.00	19.00	22.00	0.00	46.00
	Correct (%)	100.00	100.00	100.00	0.00	95.83
	Error (%)	0.00	0.00	0.00	100.00	4.17
Scenario 2	Manual	0.00	12.00	9.00	0.00	21.00
	System	0.00	12.00	9.00	0.00	21.00
	Correct (%)	100.00	100.00	100.00	100.00	100.00
	Error (%)	0.00	0.00	0.00	0.00	0.00
Scenario 3	Manual	0.00	70.00	24.00	1.00	95.00
	System	0.00	68.00	24.00	1.00	93.00
	Correct (%)	100.00	97.14	100.00	100.00	97.89
	Error (%)	0.00	2.86	0.00	0.00	2.11
Scenario 4	Manual	1.00	2.00	44.00	7.00	54.00
	System	1.00	1.00	42.00	5.00	49.00
	Correct (%)	100.00	50.00	95.45	71.43	90.74
	Error (%)	0.00	50.00	4.55	28.57	9.26
Scenario 5	Manual	10.00	11.00	5.00	48.00	74.00
	System	9.00	10.00	5.00	45.00	69.00
	Correct (%)	90.00	90.91	100.00	93.75	93.24
	Error (%)	10.00	9.09	0.00	6.25	6.76

In the second experiment, we analyzed the identified location of the detected vehicles in each lane. The results obtained from the selected scenarios are shown in Table 5. App 1 and App 2 indicate approach 1 and approach 2, respectively. The results obtained from the experiments show that the accuracy of vehicle location identification is between 71.43% and 90.54% with the first approach, and between 71.43% and 94.59% with the second approach, for all lanes.

Table 5. Vehicle location detection in the predefined scenarios.

Scenarios	Outputs	2nd lane on the left		1st lane on the left		Middle		Right		Total	
		App 1	App 2	App 1	App 2	App 1	App 2	App 1	App 2	App 1	App 2
Scenario 1	Correct (%)	80.00	100.0	94.74	84.21	86.36	86.36	100.0	100.0	89.58	87.50
	Error (%)	20.00	0.00	5.26	15.79	13.64	13.64	0.00	0.00	10.42	12.50
Scenario 2	Correct (%)	100.0	100.0	75.00	75.00	66.67	66.67	100.0	100.0	71.43	71.43
	Error (%)	0.00	0.00	25.00	25.00	33.33	33.33	0.00	0.00	28.57	28.57
Scenario 3	Correct (%)	100.0	100.0	80.00	80.00	95.83	95.83	100.0	100.0	84.21	84.21
	Error (%)	0.00	0.00	20.00	20.00	4.17	4.17	0.00	0.00	15.79	15.79
Scenario 4	Correct (%)	100.0	100.0	100.0	100.0	90.91	90.91	71.43	71.43	88.89	88.89
	Error (%)	0.00	0.00	0.00	0.00	9.09	9.09	28.57	28.57	11.11	11.11
Scenario 5	Correct (%)	70.00	100.0	72.73	72.73	100.0	100.0	97.92	97.92	90.54	94.59
	Error (%)	30.00	0.00	27.27	27.27	0.00	0.00	2.08	2.08	9.46	5.41

In total, the accuracy of the vehicle type detection and location identification for the vehicles with the correct type detection in all scenarios when considering all lanes is shown in Table 6. As this table shows, the accuracy of the second approach for estimating the lanes the detected vehicles are in is higher than that of the first approach.

Table 6. Total accuracy for all scenarios.

	Type identification	Localization based on App 1	Localization based on App 2
Correct (%)	95.21	91.01	91.73
Error (%)	4.79	8.99	8.27

6 Discussion and Conclusion

The main objective of this research was to study modern vehicles' sensing abilities for collecting traffic data to improve traffic management systems. To achieve this objective, we have developed a system and done experiments with real traffic data. Some of the prior studies that have noted the importance of collecting traffic data used stationary sensors to achieve this goal (e.g., [5]). As using stationary sensors are costly to equip all streets, we have used a vehicle-mounted sensor, as modern vehicles are equipped with various types of sensors, which are powerful and free resources to use.

As we have mentioned, modern vehicles are equipped with various types of sensors, but we decided to use a monocular camera to make our solution more feasible in the real world. Due to lidars are more expensive than cameras, the possibility of equipping all vehicles with a lidar is low, which will limit the generalizability of the proposed approach in reality. Therefore, we decided to use a monocular camera, which is cheap and likely to be mounted on most modern vehicles. Moreover, the camera's field of view gives us the possibility to collect data from multiple lanes to provide a better understanding of the traffic situation.

Our proposed algorithms are a combination of a deep learning algorithm called YOLO, which was trained on the KITTI dataset to detect vehicles and identify their type, and image processing approach to provide robust vehicle location estimation for multiple lanes. Although most of the existing papers in this scope have focused on lane detection (e.g., [15]) or object detection (e.g., [5]), we have combined both methodologies to extract more data types. In reviewing the literature, we found that more recent studies have been limited to lane detection and tracking for driver assistance systems (e.g., [17]). No approaches were found on the dependency between vehicle detection and the related lane, as it is vital for traffic management systems, especially at intersections, to access the traffic volume per lane.

One of the most significant findings from our proposed algorithms is that a vehicle-mounted monocular camera is able to extract traffic data, such as the detected vehicles' type, what lanes they are in, and the number of detected vehicle in each lane. Our experiments on real traffic data with five scenarios confirmed that our algorithms can identify the detected vehicles' type with an accuracy higher than 90.74%. The accuracy of vehicle location identification for all lanes with the first and second approaches is between 71.43% and 90.54%, and between 71.43% and 94.59%, respectively. The observed low accuracy of the second scenario can be explained by the fact that the lane marks on the right side almost vanished, which had a direct effect on the accuracy of the vehicle location detection. Moreover, the accuracy of identifying the lane the detected vehicle with the correct determined type was in by considering the total lanes was 91.01% for the first approach, and 91.73% for the second approach. Although this study was limited by driving an equipped vehicle in the middle lane, the findings prove that this idea would be feasible in reality. However, further experimentation to consider various scenarios is recommended. Moreover, as our proposed algorithms are based on object detection and lane detection algorithms, therefore, by enhancing the accuracy of the object detection and lane detection algorithms, the performance of our proposed algorithm would be enhanced. Our future work will improve the performance and accuracy of our approach further.

References

1. Lamouik, I., Yahyaouy, A., and Sabri, M.A.: Smart multi-agent traffic coordinator for autonomous vehicles at intersections: Book Smart multi-agent traffic coordinator for autonomous vehicles at intersections. IEEE, pp. 1-6 (2017).
2. Namazi, E., Li, J., and Lu, C.: Intelligent intersection management systems considering autonomous vehicles: A systematic literature review. IEEE Access, pp. 91946-91965 (2019).
3. Namazi, E., Holthe-Berg, R.N., Lofsberg, C. S., and Li, J.: Using vehicle-mounted camera to collect information for managing mixed traffic. 15th International Conference on Signal-Image Technology & Internet-Based Systems, pp. 222-230 (2019).
4. Chen, Z., Khemmar, R., Decoux, B., Atahouet, A., and Ertaud, J.-Y.: Real time object detection, tracking, and distance and motion estimation based on deep learning: Application to smart mobility. IEEE, pp. 1-6 (2019).
5. Tian, S., Yu, H., Yang, Z., Jing, X., Zhang, Z., Shi, M., and Wang, Y.: An improved target detection and traffic parameter calculation method based on YOLO with a monocular camera. CICTP, pp. 5696-5708 (2019).
6. Hillel, A.B., Lerner, R., Levi, D., and Raz, G.: Recent progress in road and lane detection: A survey. Machine Vision and Applications, pp. 727-745 (2014).
7. Keatmanee, C., Jakborvornphan, S., Potiwanna, C., San-Uml, W., and Dailey, M.N.: Vision-based lane keeping—A survey. IEEE, pp. 1-6 (2018).
8. Andrade, D.C., Bueno, F., Franco, F.R., Silva, R.A., Neme, J.H.Z., Margraf, E., Omoto, W.T., Farinelli, F.A., Tusset, A.M., and Okida, S.: A novel strategy for road lane detection and tracking based on a vehicle's forward monocular camera. IEEE Transactions on Intelligent Transportation Systems, pp. 1497-1507 (2018).
9. von Reyher, A., Joos, A., and Winner, H.: A lidar-based approach for near range lane detection. IEEE, pp. 147-152 (2005).
10.Goldbeck, J., Hürtgen, B., Ernst, S., and Kelch, L.: Lane following combining vision and DGPS. Image and Vision Computing, pp. 425-433 (2000).
11.Narita, Y., Katahara, S., and Aoki, M.: Lateral position detection using side looking line sensor cameras. IEEE, pp. 271-275 (2003).
12.Jo, Y., Han, S.-J., Lee, D., Min, K., and Choi, J.: An autonomous lane-level road map building using low-cost sensors. In Eleventh International Conference on Machine Vision. (ICMV 2018), vol. 11041 (2019).
13.Jia, B., Chen, J., Zhang, K., and Wang, Q.: Sequential monocular road detection by fusing appearance and geometric information. IEEE/ASME Transactions on Mechatronics, 24 (2), pp. 633-643 (2019).
14.Chao, F., Yu-Pei, S., and Ya-Jie, J.: Multi-lane detection based on deep convolutional neural network. IEEE Access, pp. 150833-150841 (2019).
15.Cao, J., Song, C., Song, S., Xiao, F., and Peng, S.: Lane detection algorithm for intelligent vehicles in complex road conditions and dynamic environments. Sensors, pp. 3166 (2019).

16.Zhang, W., and Mahale, T.: End to end video segmentation for driving: Lane detection for autonomous car. arXiv preprint (2018).
17.Yuan, C., Chen, H., Liu, J., Zhu, D., and Xu, Y.: Robust lane detection for complicated road environment based on normal map. IEEE Access, pp. 49679-49689 (2018).
18. GoPro hero7 black, https://gopro.com/en/us/shop/cameras/hero7-black/CHDHX-701-master.html, last accessed 2020.
19.Paszke, A., Gross, S., Chintala, S., Chanan, G., Yang, E., DeVito, Z., Lin, Z., Desmaison, A., Antiga, L., and Lerer, A.: Automatic differentiation in PyTorch. 31st Conference on Neural Information Processing Systems (NIPS 2017) (2017).
20.Bradski, G.: The opencv library. Software Tools, pp. 120-125 (2000).
21.Ding, L., and Goshtasby, A.: On the Canny edge detector. Pattern Recognition, pp. 721-725 (2001).
22.Galamhos, C., Matas, J., and Kittler, J.: Progressive probabilistic Hough transform for line detection. IEEE, pp. 554-560 (1999).
23.Matas, J., Galambos, C., and Kittler, J.: Robust detection of lines using the progressive probabilistic hough transform. Computer Vision and Image Understanding, pp. 119-137 (2000).
24.Redmon, J., Divvala, S., Girshick, R., and Farhadi, A.: You only look once: Unified, real-time object detection. The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 779-788 (2016).
25.Redmon, J., and Farhadi, A.: YOLOv3: An incremental improvement. arXiv preprint (2018).
26.Geometry, https://en.wikipedia.org/wiki/Parallel, last accessed 2019.
27.YOLO, https://pjreddie.com/darknet/yolo/, last accessed 2019.
28.Use YOLOv3 PyTorch to train KITTI, https://github.com/packyan/PyTorch-YOLOv3-kitti, last accessed 2019.
29.Deng, G., and Wu, Y.: Double lane line edge detection method based on constraint conditions Hough transform. IEEE, pp. 107-110 (2018).
30. Point-Line Distance--2-Dimensional, https://mathworld.wolfram.com/Point-LineDistance2-Dimensional.html, last accessed 2020.
31. Slope, https://en.wikipedia.org/wiki/Slope, last accessed 2019.
32.Sum of angles of a triangle, https://en.wikipedia.org/wiki/Sum_of_angles_of_a_triangle, last accessed 2019.
33.Trigonometry, https://en.wikipedia.org/wiki/Trigonometry, last accessed 2019.

Paper D:

Improving Vehicle Localization with Two Low-cost GPS Receivers

Elnaz Namazi, Rudolf Mester, Chaoru Lu, Markus Metallinos Log, and Jingyue Li

Smart City Applications (SCA) International Conference

Improving vehicle localization with two low-cost GPS receivers

Elnaz Namazi[1], Rudolf Mester[1], Chaoru Lu[2], Markus Metallinos Log[3], and Jingyue Li[1]

[1] Department of Computer Science, Norwegian University of Science and Technology (NTNU), Trondheim, Norway
{elnaz.namazi, rudolf.mester, jingyue.li}@ntnu.no

[2] Department of Civil Engineering and Energy Technology, Oslo Metropolitan University (OsloMet), Oslo, Norway
chaorulu@oslomet.no

[3] Department of Civil and Environmental Engineering, Norwegian University of Science and Technology (NTNU), Trondheim, Norway
markus.log@ntnu.no

Abstract. A primary concern of Intelligent Traffic Management Systems (ITMSs) is to collect the required traffic data. Vehicle position is one of the most important data types to manage traffic effectively. In this regard, Global Positioning System (GPS) receivers are widely used; however, their estimation accuracy is affected by several parameters, such as signal blockage. Map-matching is one of the most popular approaches to dealing with this challenge. In this study, we investigated the performance of map-matching software and found that it cannot locate the vehicle effectively if the positional data are too noisy. This paper aims to propose a new methodology by integrating cross-GPS validation, interpolation, best-fit, and map-matching techniques to enhance the vehicle localization performance in the presence of GPS signal noise and investigate the methodology with real traffic data from a metropolitan area. Our evaluations indicate that the proposed methodology can significantly improve vehicle self-localization performance.

Keywords: Vehicle self-localization, GPS receiver, Map-matching.

1 Introduction

Over the past several years, population growth has led to an increase in vehicle numbers, resulting in increased traffic congestion in many cities. As a result, Intelligent Traffic Management Systems (ITMSs) are introduced to manage traffic based on traffic data and make smart decisions. Such data could originate from stationary sensors, such as inductive loop detectors, or from vehicle-mounted sensors, such as Global Positioning System (GPS) receiver, camera, radar, and Light Detection And Ranging (LiDAR).

Vehicle location is one of the most important kind of traffic data. A GPS receiver is a common solution to estimate vehicle location in a GPS coordinate system (also called vehicle self-localization), as most Modern Vehicles (MVs) are equipped with it. However, the accuracy of the data collected via a GPS receiver depends on several parameters, such as hardware accuracy, satellite geometry, signal blockage, and atmospheric conditions [6].

To satisfy vehicle localization requirements and mitigate the estimated location error, three major categories of approaches have been proposed in the literature [3][5]. One category of approaches uses a standalone reference station, such as Wide Area Augmentation System (WAAS) (e.g., [17]). The second category comprises auxiliary hardware-based approaches (e.g., Inertial Measurement Unit (IMU) [8]). Using technologically advanced sensors to determine vehicle location would boost the estimation accuracy. However, equipping a vehicle with such sensors will also increase the vehicle's cost. The third category uses software, such as map-matching techniques (e.g., [18]). Map-matching is a technique that integrates map information and recorded geolocation data from the vehicle in order to increase the accuracy of the vehicle's location [19]. Although map-matching techniques are widely applied to minimize vehicles' localization error, in this study, we found that map-matching techniques (e.g., QGIS-Plug-in Offline-MapMatching [12][13]) do not work well if the GPS data collected via a low-cost GPS receiver are too noisy.

Therefore, a much-debated question is how to keep the hardware's and sensors' costs low and the localization performance high. This paper proposes a new methodology by integrating cross-GPS validation, interpolation [4], best-fit [2], and map-matching [12][13] techniques to localize a vehicle in the presence of GPS signal noise. Our proposed methodology can identify the more accurate GPS receiver dynamically by considering the fixed and known distance between two GPS receivers. We implemented and evaluated our approach using real traffic data from a metropolitan area in Chengdu, China. The results show that our proposed approach can enhance vehicle self-localization performance.

The paper is organized as follows. Section 2 gives a brief overview of related work. Section 3 explains our proposed research design. Section 4 presents our proposed research approach. Section 5 describes our evaluation of the methodology. Section 6 presents the discussion. The last section concludes and proposes future studies.

2 Related Work

Vehicle localization based on GPS receivers is a key component in managing traffic safely and effectively. However, it can be imprecise, causing operational difficulties. Many approaches have been proposed to process imprecise data from GPS receivers to acquire accurate vehicle localization. For instance, Islam et al. [5] enhanced GPS accuracy by considering a vehicle's movement direction, velocity averaging, and the distance between waypoints using coordinate data. Their experiment used a vehicle-mounted Garmin GPS 19xHVS receiver. In order to examine the accuracy, they plotted the data on Google Maps. The proposed approach achieved improvement of 4–10 m [5]. Acosta et al. [11] proposed an approach based on a Kalman, fuzzy logic, and information selection. In the experimental step, they used three Garmin 18X GPS receivers that were connected to two notebooks. The proposed approach in [11] smoothened the measurement error and mitigated the error that fluctuates in time. Tang et al., in [15], proposed an adaptive map-matching algorithm based on a hierarchical fuzzy system. The experimental results showed that the proposed algorithm in [15] was able to increase the matching accuracy and to outperform the traditional algorithms based on only geometric or topological information of network. Lecce et al. [1] used generalized regression neural networks to increase the GPS position accuracy by correcting the receiver's position. The idea was to use an analytical description of the time series to improve the position accuracy. They proposed an approach based on removing the GPS positioning error by training a neural network to mitigate the periodic components of the GPS positioning error. In the experimental step, they used GPS receiver BU-353. The mean improvement in the accuracy of the GPS position of the proposed approach was 25% [1].

3 Research Design

Our recent studies [9][10] proposed new methodologies that use the ego-vehicle as a mobile sensor, estimating the traffic data for surrounding vehicles in order to share them with an ITMS. Vehicle localization (i.e., ego-vehicle and target vehicle localization) plays an important role in managing the traffic. Our studies revealed that the image-based target vehicle localization accuracy is tightly connected to the localization accuracy of the ego-vehicle.

This paper aims to enhance the ego-vehicle localization performance by using two low-cost GPS receivers. Each vehicle was equipped with two monocular cameras. One camera was mounted on the front window of the vehicle, and another was mounted on the rear window. These two cameras were located at a known distance from each other on the vehicle, helping us to validate the GPS receiver accuracy, as well as collecting footage from both sides of the vehicle, which was needed for further image processing-based studies. All cameras used were of the type GoPro Hero 7. The monocular camera was a low-cost sensor that can be mounted on most ego-vehicles, making our approach generalizable. In addition to collecting video footage, the chosen camera enabled GPS data collection, as it included a built-in GPS receiver (hereafter, we call this camera as a GPS receiver since it embeds a GPS receiver).

To begin this research, we first analyzed the accuracy of the collected GPS data obtained via the GPS receiver mounted on the front window glass by plotting the data on a map (the data collection process is described in detail in Section 5.1). Fig. 1 shows one example of the studied scenarios in which the ego-vehicle turns right at an intersection. In Fig. 1-A, the blue arrow shows the vehicle's movement scenario. The polyline, which is a combination of green and red colors, represents the vehicle's location based on the front GPS receiver mounted on the vehicle. The color of the polyline represents the vehicle's speed.

This polyline and its colors are plotted automatically by using "Telemetry Extractor for GoPro" [16].

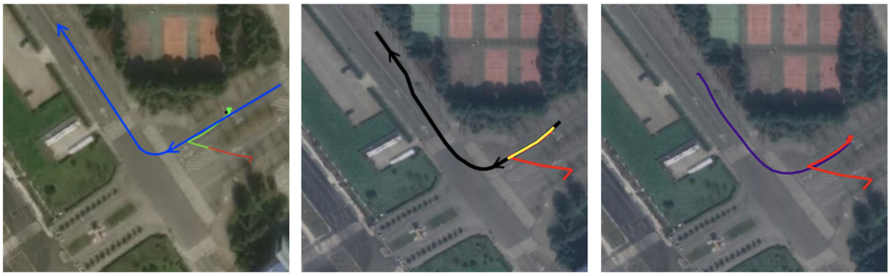


Fig. 1. Problem formulation. A) Vehicle locations collected via a front-mounted GPS receiver on the vehicle (green-red polyline), compared with the vehicle's movement scenario (blue polyline). B) Map-matching output (yellow polyline) related to the noisy front GPS receiver (red polyline) by considering the true trajectory of the vehicle (black polyline). C) Vehicle locations obtained via two GPS receivers on the same vehicle (front GPS receiver: red polyline; rear GPS receiver: purple polyline.)

Our first attempt was to use map-matching software to address the GPS receiver noise issue to obtain the precise vehicle location. We used the QGIS-Plug-in Offline-MapMatching [12][13], which is one of the widely used approaches for minimizing the GPS receiver error. Fig. 1-B shows our findings after applying map-matching to the same scenario shown in Fig. 1-A. In Fig. 1-B, the black polyline is the true trajectory of the vehicle on the road. The red polyline represents the positions collected via the front GPS receiver mounted on the vehicle (part of this red polyline is covered by the yellow polyline), and the yellow polyline represents the map-matched positions of the noisy GPS receiver. From this figure, it is clear that the QGIS-Plug-in Offline-MapMatching [12][13] is not able to identify and map-match the entire trajectory accurately if the vehicle localization error is too high.

We then analyzed data from another GPS receiver on the same vehicle in the scenario shown in Fig. 1-A and Fig. 1-B. Fig. 1-C shows the results. In Fig. 1-C, the red polyline is the vehicle's position based on the GPS receiver mounted on the front window. The purple polyline shows the vehicle's position based on the GPS receiver mounted on the rear window on the same vehicle. As Fig. 1-C shows, the localization error of the front-mounted GPS receiver is much higher than that of the rear-mounted GPS receiver in this scenario.

4 Research Approach

Fig. 2 illustrates our proposed research approach, comprising data collection, data preprocessing, and data processing.

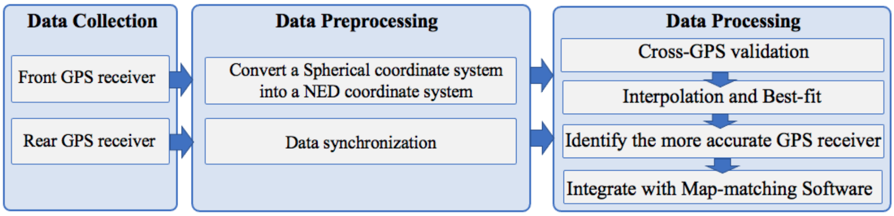


Fig. 2. Our proposed research approach.

4.1 Data Preprocessing

As previously stated, positional data of our study were collected using two GPS receivers mounted on the ego-vehicle. Before applying our approach, the data were preprocessed. In this step, firstly, we need to convert a Spherical coordinate system [14] into a local North-East-Down (NED) coordinate system [7] on the earth's surface. The conversion is both practical and justified, since we are studying a small, demarcated area on the earth's surface. Secondly, since the two mounted GPS receivers on the vehicle are independent and the data collection was not started concurrently, we need to synchronize the receivers in the time domain.

4.2 Data Processing

In this step, first, we need to analyze the accuracy of the two mounted GPS receivers on the same vehicle. To detect whether the GPS signals are accurate, we calculated the vector distance of the estimated positions obtained via the two GPS receivers at equal timestamps, as the two GPS receivers were mounted with a fixed and known distance from each other (in our study, we assumed this fixed distance D_g is 3 m, because we used family vehicles) on the same vehicle. If it is found that the vector distances are different from this fixed distance (with an error threshold e = ± 2 m), we can conclude that at least one of the GPS receivers is inaccurate, which means we need to identify the more accurate GPS receiver.

To identify the more accurate GPS receiver, we developed a new algorithm based on cross-validation, interpolation [4], and best-fit [2] techniques, as presented in Fig. 3. Cross-validation found the positions in the trajectory where both GPS receivers were almost in agreement on the vehicle's position (i.e., the position difference obtained by two GPS receivers was between D_g - e and D_g + e). It did so based on the Euclidean distance (Ed) between each pair of preprocessed positions obtained by the front and rear GPS receivers per each timestamp. In this study, we assumed that the error of the GPS receiver was random error, which means that the GPS receivers can obtain accurate locations in most time (hereafter, the accurate locations are called valid points). Due to possible perturbations [6], GPS receivers can sometimes provide noisy locations.

Our idea is to identify the valid points and then use interpolation [4] technique to calculate the possible locations when the GPS error is identified. In addition, for the straight vehicle movements, which were determined based on the vehicle's movement slope (we assumed the movement with a slope less than 20 degree as a straight movement, otherwise as a turn), the best-fit technique [2] was used to generate more positions in the whole trajectory based on the interpolated position. To identify the more accurate GPS receiver, we then calculated the average Euclidean distance between the positions calculated through interpolation and best-fit and the positions collected by each GPS receiver. The GPS receiver with the smallest average distance was identified as the more accurate one.

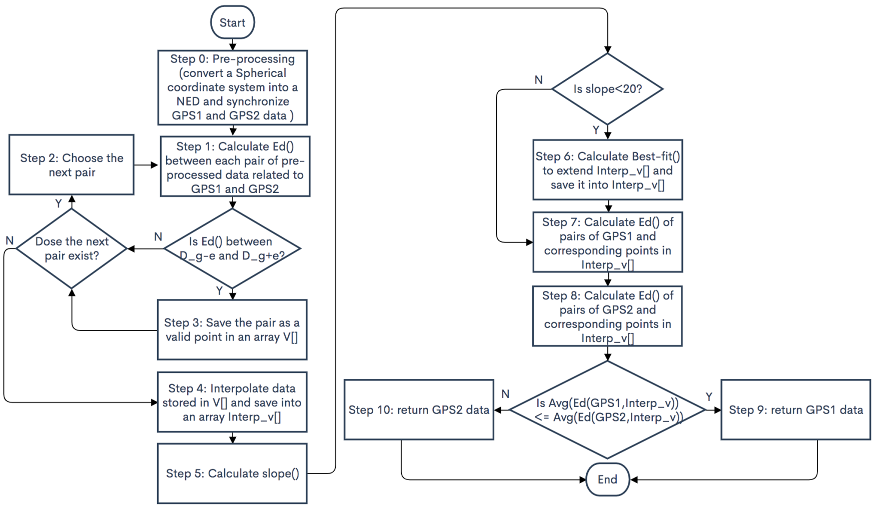


Fig. 3. Flowchart of our proposed algorithm.

Although we can identify that one GPS receiver is more accurate than the other, the more accurate one may also be noisy. Finally, we inserted the data from the identified more accurate GPS receiver into a map-matching algorithm, using it to further amend the noisy GPS signal. We investigated the effectiveness of several existing map-matching software applications and identified the one that was most compatible with our data. We found that the QGIS-Plug-in Offline-MapMatching [12][13] was a suitable and effective tool for map-matching in our research context.

5 Evaluation

5.1 Data Collection

To evaluate our proposed approach, experiments were run using several case studies with real traffic data. We used three equipped vehicles (each vehicle were equipped with two GPS receivers, as described in Section 3) driven in the metropolitan region of Chengdu, China. In order to provide good data coverage and generalizability, eight different scenarios were defined, comprising both straight-street and intersection movements. In total, 24 trajectories were considered. There were many tall buildings around the studied area, which may interfere with GPS signal accuracy and cause GPS data inaccuracies. As the ground truths related to vehicle movements in this study were not available from the GPS receiver data, we extracted them manually by visually observing forward-facing video footage and identifying the ground-truth vehicle movements using Google Earth Pro [20].

5.2 Evaluation of the Results

As we observed in Fig. 1, if the GPS signal was too noisy, the QGIS-Plug-in Offline-MapMatching [12][13] was able to minimize the localization inaccuracy of only a segment of the trajectory. We used the Cartesian length of the trajectory to evaluate the performance of our proposed self-localization approach. Table 1 summarizes our findings. This table included eight scenarios (S1–S8) and three equipped vehicles (V1–V3). The "ground truth" column shows the Cartesian length of a vehicle's movement, and the "avg. dis." columns represents the average distance between the vehicle positions collected via each GPS receiver and the ground truth. The GPS receiver with a smaller average distance was labeled as a more accurate GPS receiver. To assess our proposed methodology, we first calculated the Cartesian length of the vehicle trajectory by using only map-matching on data of both front and rear GPS receivers. Our findings are presented in the "front GPS" and "rear GPS" sub-columns of the "map-matching based Cartesian length" column. We then calculated the Cartesian length of the vehicle trajectory, after applying our proposed methodology and identifying the accurate GPS receiver. The results are presented in the "accurate GPS" (results of steps 9 and 10 in Fig. 3) and "Cartesian length" sub-columns of the "our proposed approach" column. In addition, we compared the deviation from the ground truth by using only map-matching on the collected data and using our proposed approach. The results are shown in the sub-columns of the "deviation comparison" column.

To explain the information presented in Table 1 in depth, we use scenario S8 and vehicle V3 as an example. In this scenario, the Cartesian lengths of the map-matched positions obtained via both GPS receivers are almost the same (front GPS: 150 m; rear GPS: 148 m). This shows that applying map-matching software would be enough to correct such small errors satisfactorily. However, this table shows that when the GPS error is high, applying only map-matching

Table 1. Case study evaluation.

S#	V#	Ground truth (m)	Avg. Dis. (m)		Map-matching-based Cartesian Length (m)		Our proposed approach		Deviation comparison (m)		
			Front GPS	Rear GPS	Front GPS	Rear GPS	Accurate GPS	Cartesian Length (m)	Front GPS	Rear GPS	Our proposed approach
S1	V1	532	12.009	4.935	490	532	Rear	532	-42	0	0
	V2	514	2.044	10.746	513	514	Front	513	-1	0	-1
	V3	441	2.324	4.415	441	437	Front	441	0	-4	0
S2	V1	179	1.457	6.058	179	178	Rear	178	0	-1	-1
	V2	191	4.358	3.385	191	177	Front	191	0	-14	0
	V3	147	1.669	2.19	147	145	Front	147	0	-2	0
S3	V1	191	1.955	1.774	191	155	Rear	155	0	-36	-36
	V2	189	1.552	4.612	189	184	Front	189	0	-5	0
	V3	159	3.608	13.860	159	150	Front	159	0	-9	0
S4	V1	159	4.241	0.665	156	159	Rear	159	-3	0	0
	V2	163	6.044	1.84	163	162	Front	163	0	-1	0
	V3	188	1.388	2.264	188	188	Front	188	0	0	0
S5	V1	174	5.170	1.798	174	162	Rear	162	0	-12	-12
	V2	188	3.126	4.900	188	188	Front	188	0	0	0
	V3	118	1.450	2.385	118	118	Front	118	0	0	0
S6	V1	124	3.752	6.913	124	117	Rear	117	0	-7	-7
	V2	194	1.333	7.131	186	194	Front	186	-8	0	-8
	V3	–	–	–	–	–	–	–	–	–	–
S7	V1	106	1.834	4.460	106	106	Rear	106	0	0	0
	V2	142	7.660	4.803	141	142	Front	141	-1	0	-1
	V3	–	–	–	–	–	–	–	–	–	–
S8	V1	**109**	**3.515**	**1.402**	**27**	**109**	**Rear**	**109**	**-82**	**0**	**0**
	V2	107	1.493	2.983	103	107	Front	103	-4	0	-4
	V3	150	1.627	2.939	150	148	Front	150	0	-2	0

may not be effective, which is the main focus of this study. For instance, in scenario S8 and vehicle V1, the Cartesian length by applying map-matching associated with the front GPS receiver is 27 m, while it is 109 m for the rear GPS receiver. It means by using only one GPS receiver (i.e., front GPS receiver), map-matching is only effective for a small segment of the trajectory (i.e., 27 m). The performance could be increased if we consider another GPS receiver. It confirms that identifying the more accurate GPS receiver is vital, which is the rear GPS receiver in this case. After identifying the more accurate GPS receiver and using its collected data to feed into the map-matching software, our proposed approach increased the self-localization performance which is measured using the Cartesian length of the output to 109 m. Therefore, the S8 and V1 case showed that our approach is effective in the presence of extreme GPS signal noise. In Table 1, using our approach to choose a more accurate GPS receiver first and then use map-matching does not always give less deviation than using front or rear GPS receiver randomly. The reason is that we chose to use GPS2 data in step 10 in Fig. 3 when data from both GPS receivers were acceptable. The GPS2 data may not be better than GPS1 data in some cases, although both data from both GPS receivers are acceptable. In this table, for vehicle V3 in scenarios S6 and S7, information are not provided, as the rear GPS receiver did not record during the whole scenario. The reason for this could be that the battery died or that the memory card became full.

6 Discussion

Previous studies have noted the importance of identifying and mitigating the measurement error of GPS receivers. This paper developed a new algorithm to identify the more accurate GPS receiver if there are multiple, possibly noisy, GPS receivers installed on the same vehicle, based on cross-validation, interpolation [4], and best-fit [2] techniques. Compared to the approach relying on expensive GPS receivers, our approach provides a low-cost solution to identify a vehicle's location precisely. Compared to the approach that relies solely on map-matching, our strategy of detecting GPS inaccuracy and prioritizing the data from the more accurate GPS receiver helped enhance the performance of the map-matching software, when the GPS signal is too noisy. One of the limitations in this study is that the cross-validation step is limited to address random GPS receiver error, which means that our approach relies on the existence of valid points, as explained in Section 4.2, which are collected by both GPS receivers on the same vehicle. If the localization error of one GPS receiver is too high and there are no overlapping points between GPS receivers in the studied trajectories, cross-validation is infeasible. Moreover, the current approach and evaluation are based on postprocessing and the QGIS-Plug-in Offline-MapMatching [12][13]. By analyzing only a small segment of a trajectory each time and using real-time map-matching software, it would be possible to turn the solution to be more real-time.

7 Conclusion and Future Work

In this study, our research context is defined as mounting two low-cost and possibly imprecise GPS receivers on the same vehicle at a fixed and known distance from each other to accurately identify the position of the vehicle based on cross-validation, interpolation, and best-fit techniques while the vehicle is moving. We developed a new algorithm to identify the more accurate GPS receiver in the presence of noise and fed the GPS data from the more accurate GPS receiver into map-matching software. The proposed approach minimized the measurement error of the low-cost GPS receiver and was able to enhance the vehicle localization performance. Since the study was limited to vehicle movements through intersections and along straight streets with limited scenarios, more studies are needed to be able to generalize our approach by considering various vehicle movements, driving speeds, and weather conditions.

References

1. Di Lecce, V., Amato, A., Piuri, V.: Neural technologies for increasing the GPS position accuracy. In: IEEE International Conference on Computational Intelligence for Measurement Systems and Applications 2008, pp. 4–8. IEEE (2008)
2. How do I construct a straight line through data points? Best-fit lines. https://serc.carleton.edu/mathyouneed/graphing/bestfit.html

3. Huang, J.Y., Tsai, C.H.: Improve GPS positioning accuracy with context awareness. In: First IEEE International Conference on Ubi-Media Computing 2008, pp. 94–99. IEEE (2008)
4. Interpolation Methods. https://www.mathworks.com/help/curvefit/interpolation-methods.html
5. Islam, M., Kim, J.M.: An effective approach to improving low-cost GPS positioning accuracy in real-time navigation, pp. 1–8. The Scientific World Journal 2014 (2014)
6. Laphong, J.K., Dihingia, H., Borah, S.: Advanced road traffic management system. IJISSET 4, 168–169 (2019)
7. Local tangent plane coordinates. https://en.wikipedia.org/wiki/Local_tangent_plane_coordinates
8. McCall, H., Lin, C.F.: American GNC Corp, assignee. Micro integrated global positioning system/inertial measurement unit system. United States patent application US (2002)
9. Namazi, E., Li, J., Mester, R., Lu, C.: Identifying and counting vehicles in multiple lanes by using a low-cost vehicle-mounted sensor for intelligent traffic management systems. In: International Conference on Hybrid Artificial Intelligence Systems 2020, pp. 598–611. Springer, Cham (2020)
10. Namazi, E., Holthe-Berg, R.N., Lofsberg, C.S., Li, J.: Using vehicle-mounted camera to collect information for managing mixed traffic. In: 15th International Conference on Signal-Image Technology Internet-Based Systems (SITIS) 2019, pp. 222–230. IEEE (2019)
11. Acosta, N., Toloza, J.: Techniques to improve the GPS precision. International Journal of Advanced Computer Science and Applications 3, pp. 125–130, no. 8 (2012)
12. Offline-MapMatching. https://github.com/jagodki/Offline-MapMatching
13. QGIS. https://qgis.org/en/site/
14. Spherical coordinate system. https://en.wikipedia.org/wiki/Spherical_coordinate_system
15. Tang, J., Zhang, S., Zou, Y., Liu, F.: An adaptive map-matching algorithm based on hierarchical fuzzy system from vehicular GPS data. Plos One 12, no. 12, e0188796 (2017)
16. Telemetry Extractor for GoPro. https://goprotelemetryextractor.com/free/
17. Witte, T.H., Wilson, A.M.: Accuracy of WAAS-enabled GPS for the determination of position and speed over ground. Journal of Biomechanics 38, no. 8, 1717–1722 (2005)
18. Wu, Z., Xie, J., Wang, Y., Nie, Y.M.: Map matching based on multi-layer road index. Transportation Research Part C: Emerging Technologies, 118, 102651 (2020)
19. Liu, L.X.Q., Liu, M.L.Z.: Map matching algorithm and its application. Proceedings on Intelligent Systems and Knowledge Engineering (ISKE2007) (2007)
20. Google Earth. https://www.google.com/earth/download/gep/agree.html?hl=en-GB

Paper E:

Geolocation Estimation of Target Vehicles Using Image Processing and Geometric Computations

Elnaz Namazi, Rudolf Mester, Chaoru Lu, and Jingyue Li

Neurocomputing Journal

Geolocation estimation of target vehicles using image processing and geometric computation

Elnaz Namazi[a,*], Rudolf Mester[a], Chaoru Lu[b], Jingyue Li[a]

[a] *Department of Computer Science, Norwegian University of Science and Technology (NTNU), Trondheim, Norway*
[b] *Department of Civil Engineering and Energy Technology, Oslo Metropolitan University (OsloMet), Oslo, Norway*

Abstract

Estimating vehicles' locations is one of the key components in intelligent traffic management systems (ITMSs) for increasing traffic scene awareness. Traditionally, stationary sensors have been employed in this regard. The development of advanced sensing and communication technologies on modern vehicles (MVs) makes it feasible to use such vehicles as mobile sensors to estimate the traffic data of observed vehicles. This study aims to explore the capabilities of a monocular camera mounted on an MV in order to estimate the geolocation of the observed vehicle in a global positioning system (GPS) coordinate system. We proposed a new methodology by integrating deep learning, image processing, and geometric computation to address the observed-vehicle localization problem. To evaluate our proposed methodology, we developed new algorithms and tested them using real-world traffic data. The results indicated that our proposed methodology and algorithms could effectively estimate the observed vehicle's latitude and longitude dynamically.

Keywords: Target-vehicle localization, vision-based localization, geometric computation, GPS coordinate system, mixed traffic

[*] Corresponding author

1. Introduction

A growing body of literature recognizes the importance of intelligent traffic management systems (ITMSs) to manage traffic safely and efficiently. ITMSs mainly rely on traffic data to enhance traffic scene awareness and make smart decisions [1].

There are two main approaches to collecting traffic data for ITMSs. The first approach is based on stationary sensors placed toward road networks, such as inductive loop detectors (e.g., [2]) and closed-circuit television cameras (e.g., [3]). Although this approach is nowadays widely applied to collect traffic data, installing and maintaining these sensors to provide an acceptable coverage range on all roads might be costly [4]. The second approach is based on using modern vehicles (MVs) equipped with sensors. An MV with sensing and communication abilities can collect traffic data mainly about itself and transfer it based on, in general, vehicle-to-vehicle (V2V) or vehicle-to-infrastructure (V2I) communications. To collect enough traffic data with this approach, most vehicles in the traffic need to be MVs with an advanced sensor mounted. However, converting most vehicles into MVs is time-consuming. Studies predict that only 50% of vehicles in the United States will have autonomy in Level 4 (vehicles in Level 4, based on the Society of Automotive Engineers (SAE), have high automation, with which the automated driving features can drive the vehicle under limited conditions, and the driver holds control only if the automated situation turns unsafe [5][6]) by 2050 [7]. Thus, the near-future traffic would be a mixture of human-driven vehicles (HDVs) and MVs with various levels of sensing capabilities, which is called mixed traffic hereafter. Therefore, it is necessary to explore the possibility of using an MV equipped with a low-cost and popular sensor (e.g., a monocular camera), with the purpose of enhancing generalizability in mixed traffic to collect traffic data of the observed vehicles and feed them into the ITMSs.

In our previous studies [8][9], we have investigated the feasibility of using a vehicle equipped with a low-cost front-facing monocular camera with a built-in global positioning system (GPS) receiver (hereafter, we call this vehicle an ego vehicle) to observe another vehicle (hereafter, we call this vehicle an target vehicle) and estimate its speed, distance, and the lane it is in. After studies [8][9], a follow-up research question has been raised about the use of an ego vehicle in estimating the geolocation of the target vehicle, as accessing the vehicle's geolocation plays a critical role in modeling the traffic scene and making smart decisions by ITMSs.

Therefore, in this paper, we go beyond the lane-level target-vehicle localization presented in [9] and find the latitude and longitude of a target vehicle in a GPS coordinate system dynamically while both the ego vehicle and the target vehicle are moving in a metropolitan area. Although some research has been carried out on utilizing an ego vehicle as a mobile sensor to estimate traffic data of the target vehicle, there is still very little scientific understanding of estimating the geolocation of HDVs based on ego-vehicle self-localization, image-based estimated distance to the target vehicle, and the relative angle between them by using a monocular camera with a built-in GPS receiver mounted on a mobile ego vehicle.

Therefore, the objective of this paper is to investigate the feasibility of using data from low-cost sensors (i.e., a monocular camera with a built-in GPS receiver) mounted on an ego vehicle to estimate the geolocation of a moving target vehicle. Our research question is defined as follows:

- RQ: How can the geolocation of a mobile target vehicle be dynamically estimated in a GPS coordinate system based on the vision of a front-facing low-cost monocular camera with a built-in GPS receiver on a mobile ego vehicle?

To address this research question, we proposed two approaches based on (1) object detection and image processing and (2) geometric computation by considering the camera's pitch angle and height from the road surface. In this regard, we extended the proposed algorithms presented in [9] by including the estimation of the distance and angle between the ego vehicle and the target vehicle.

To evaluate our proposed approaches and develop algorithms, we ran empirical experiments using real traffic data from a metropolitan area in Chengdu, China. We analyzed the findings by plotting the estimated target vehicle's trajectory on Google Maps and compared it with the ground-truth trajectory of the target vehicle. Additionally, the vector distance was used to quantitatively analyze the deviations between the estimated and the ground-truth geolocations of the target vehicle. The evaluation results confirmed that both approaches could estimate the geolocation of the target vehicles accurately.

The rest of the paper is organized as follows. Section 2 gives a brief overview of the recent history related to vehicle localization approaches. Section 3 explains the research strategy and methodology we propose. Section 4 presents the experiments and results of our approaches. The discussion is presented in Section 5. The last section concludes and proposes future studies.

2. Related work

To estimate the target vehicle's geolocation in a GPS coordinate system, we need to know the ego vehicle's geolocation and the target vehicle's location (the distance and angle between the ego vehicle and the target vehicle) [10]. This section presents a related work of these aspects briefly.

2.1. Ego-vehicle geolocation

There has been an increasing amount of literature on estimating the geolocation of ego vehicles, which is usually called self-localization (e.g., [11]). A GPS receiver is one of the most popular sensors for localizing ego vehicles [12]. Standard GPS receivers in the market have an accuracy of about 10-15 meters in 95% of the time [13]. To minimize the GPS receiver's estimation error in ego-vehicle localization, map matching is applied widely [14]. Huang et al, [14] classified the map matching algorithms into four categories: geometric theory, topology, probability statics, and advanced model.

2.2. Target vehicle's location estimation

To date, various studies have investigated target vehicle's location estimation via monocular cameras regarding driving safety measures, assistance, and autonomous navigation. For instance, Ifthekhar et al., [15] introduced an optical camera communications (OCC)-based cooperative vehicle positioning (CVP) technique. They proposed two approaches: (1) a neural network-based approach and (2) a computer vision-based approach to estimate the target vehicle's location. They considered two vehicles, one as an observing vehicle equipped with front-left and front-right cameras. Another vehicle was treated as a target vehicle, and its positioning was estimated based on its rear light-emitting diodes (LEDs). Simulation results showed that the accuracy achieved by the proposed neural network-based method was higher than the computer vision-based method [15]. Hayakawa et al., in [16] proposed a new approach based on integrating three deep neural networks to estimate the ego-motion and the target vehicle's state (e.g., 3D vehicle bounding box, depth, and optical flow). The experimental evaluations demonstrated that the distance error in the lateral and longitudinal directions were 1.19 m and 1.70 m, respectively.

Lee [17] focused on inter-vehicle distance estimate based on lane width. The proposed technique had a distance estimate error of less than 7%. Huang et al. [18] proposed a novel approach to estimate the inter-vehicle distance based on vanishing point detection, road segmentation, and vehicle detection. The ratio of true distance to image pixel was used to calculate the distance. In [18], a single-lens camera was utilized to capture data from urban/suburban roadways. Five image sequences of urban/suburban roads were utilized to verify the performance of the suggested method. The results showed average detection rate (DR), and false alarm rate (FAR) values of the approach are 82.21% and 16.16%, respectively [18]. Giesbrecht et al. [19] proposed a vision-based leader/follower system for an ego vehicle. The system was a combination of three main components: (1) a computer vision system for tracking the target vehicle based on color and the scale-invariant feature transform (SIFT), (2) a control system based on linear quadratic Gaussian control, and (3) a path following system. Their experiments showed that the mean and maximum error in the visual distance estimate were 0.72 m and 2.42 m, respectively, the follower speed was between 7.6 km/h and 10.2 km/h, and the follower separation was between 10.46 m and 23.71 m.

Taken together, although some research has been carried out on vehicle localization, more detailed empirical investigations are needed to dynamically estimate the target vehicle's geolocation in a GPS coordinate system via a monocular camera with the purpose of generating data to model the traffic scene and improve the ITMS performance.

3. Research strategy and methodology

In this paper, we proposed two new approaches by integrating deep learning, image processing, and geometric computation to use the vision sensing and self-localization capabilities of a mobile ego vehicle to estimate the geolocation of target vehicles. Figure 1 illustrates our proposed research strategy. The components included in Figure 1 are as follows:

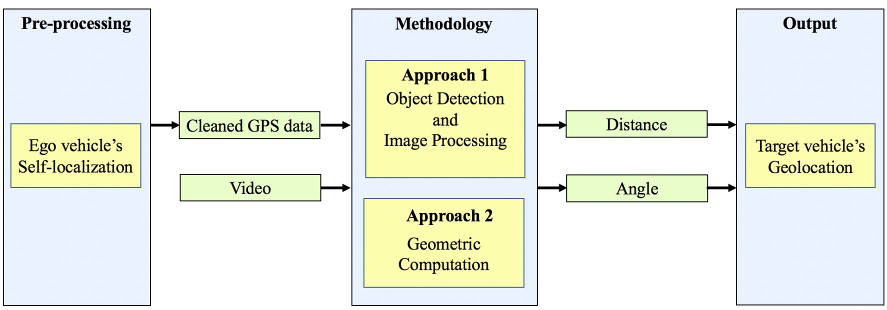


Figure 1: The proposed steps in our research strategy.

3.1. Pre-processing

For the estimation of the target vehicle's geolocation, the geolocation of the ego vehicle is required. The collected latitude and longitude of the ego vehicle, which are usually collected by a GPS receiver, might be noisy. Collecting the geolocations of the ego vehicle accurately plays a vital role in accurately estimating the target vehicle's geolocations. Therefore, to enhance the accuracy of the ego-vehicle localization in the GPS coordinate system, the proposed approach in [20], based on cross-GPS validation, interpolation, best-fit, and map-matching techniques, is used.

3.2. Methodology

As shown in Figure 2, to estimate the target vehicle's geolocation in the GPS coordinate system, in addition to the ego vehicle's geolocation, the distance d between the ego vehicle V_E and the target vehicle V_T and the clockwise angle α between the north (N) and d are required [10].

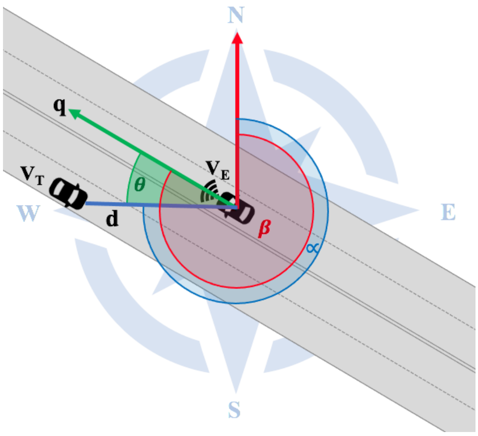


Figure 2: The required parameters for estimating the target vehicle's geolocation.

In this regard, as Figure 1 shows, we proposed two approaches, as follows:

I) Approach 1: Object detection and image processing

- Estimating the distance d

  To begin the distance estimation process, we employed you only look once (YOLO)-v3 [21][22] to detect target vehicles via the ego vehicle's vision. YOLO-v3 is a well-documented open-source one-stage method for detecting objects. “YOLO-v3 is extremely fast and accurate” [23]. For example it is more than 1000x faster than R-CNN and 100x faster than Fast R-CNN [21][22]. Wang et al., [24] listed YOLO-v3 as the second most popular object detector models. Therefore, we chose to use YOLO-v3 in this study.

We trained YOLO-v3 on the KITTI dataset [25], as this study focuses on traffic objects, and KITTI includes eight categories of traffic objects: car, van, truck, pedestrian, person_sitting, cyclist, tram, and misc [25]. To estimate the distance d from the ego vehicle to the target vehicle, we followed the approach proposed by Namazi et al. [8]. This approach [8] was based on the pinhole camera model by considering the real, pre-known size of the target vehicle and the size of the bounding box added by YOLO-v3 around the target vehicle on the image plane. Distance d was calculated based on the average of the computed distances for both vehicle width and vehicle height by using a weight factor (i.e., 85% of the height and 15% of the width).

- Estimating the angle α

  As presented in Figure 2, in order to estimate the clockwise angle α between the north N and d, we need to estimate angle β, which is the angle between the north N and the ego vehicle V_E's movement direction q, as well as angle θ, which is the angle between d and q.

  i. Estimating the angle β

     To estimate the angle β, we need to identify the movement direction of the ego vehicle V_E based on its collected GPS coordinates in sequential frames as a start point (ϕ_1, λ_1) and an end-point (ϕ_2, λ_2) for all frames. We used Eq. 1 - Eq. 3 [10] to estimate angle β along the whole trajectory dynamically.

$$M = \sin(\lambda_2 - \lambda_1) \cdot \cos\phi_2 \tag{1}$$

$$N = \cos\phi_1 \cdot \sin\phi_2 - \sin\phi_1 \cdot \cos\phi_2 \cdot \cos(\lambda_2 - \lambda_1) \tag{2}$$

$$\beta = \text{atan2}(M, N) \tag{3}$$

ii. Estimating the angle θ

The idea of estimating angle θ in our first approach is presented in Figure 3. In Figure 3, the blue bounding box shows the target vehicle V_T. P is the central point on the button edge of the bounding box around the target vehicle V_T, and H is the central point of the image. Angle θ is estimated based on the horizontal angle per pixel (γ) in degrees and on the number of horizontal pixels between P and the vertical line passing through H, shown by a red line T. γ is estimated based on the camera's horizontal field of view (FOV) and the video's resolution. In this study, the video's resolution was 960×720 pixels, and the camera's horizontal FOV was 86.7 degrees [26]. Therefore, γ is equal to 0.90 degrees. Angle θ in degrees is estimated as follows:

$$\theta = T \cdot \gamma \tag{4}$$

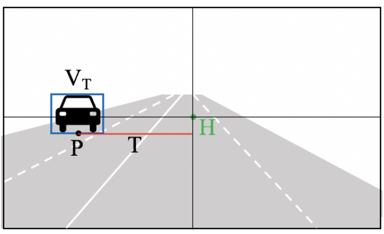


Figure 3: The parameters used to estimate angle θ between the ego vehicle and the target vehicle used in Approach 1.

iii. Estimating the angle α

To estimate the angle α, we considered three different conditions, as follows:

– If the target vehicle drives in the same lane as the ego-vehicle, then $\alpha = \beta$ and $\theta = 0$.

– If the target vehicle drives on the left side of the ego vehicle, then, as Figure 4, (a) shows, $\alpha = \beta - \theta$.

– If the target vehicle drives on the right side of the ego vehicle, then, as Figure 4, (b) shows, $\alpha = \beta + \theta$.

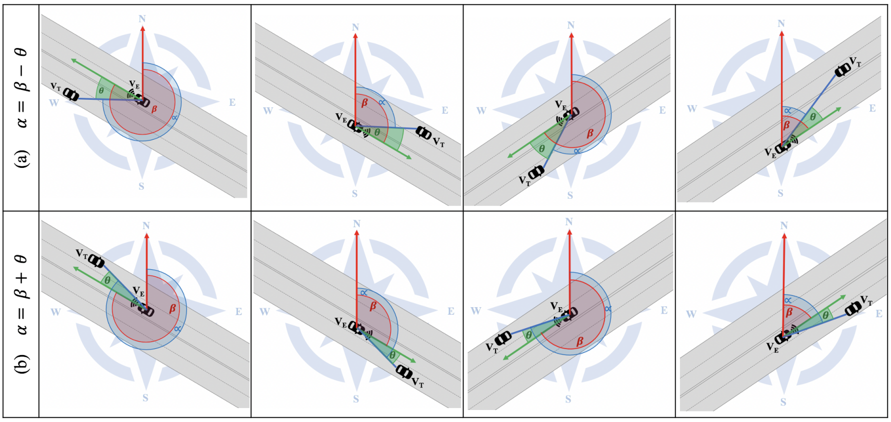


Figure 4: The mathematical relations between α, β, and θ.

II) Approach 2: Geometric computation

The main idea of this approach is to transform 2D pixel coordinates of point P into 3D world coordinates. By assessing the 3D world coordinates of point P, we would be able to estimate distance d and angle α, which are needed to estimate the target vehicle's geolocation.

In this regard, we utilized a pinhole camera model as shown in Figure 5. In this figure, C is the perspective center of the camera and the origin of the camera coordinate frame (CCF). Three unit vectors of the CCF are represented by y_1, y_2, and y_3. The image coordinate frame (ICF) is centered at principal point H with unit vectors r_1 and r_2. The principal axis passes through C and H and is perpendicular to the image plane. The distance from C to the image plane is f, which is the camera's focal length. The image plane carries a 2D pixel coordinate frame (PCF) with unit vectors z_1 and z_2. The image plane is subdivided into n_h pixels horizontally and

n_v pixels vertically. To project the detected vehicle on the image onto the real world, we need to transform point P with pixel coordinates (p_1, p_2) (which is the central point on the button edge of the bounding box around the target vehicle) into a 3D world coordinate representation (w_1, w_2, w_3).

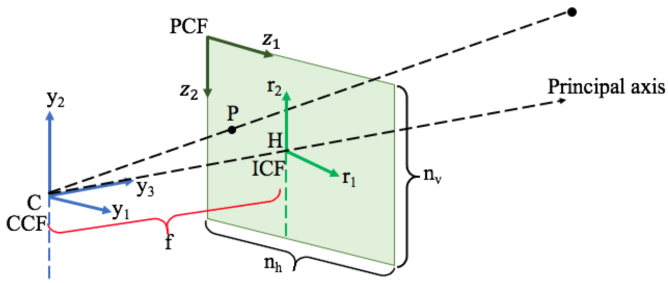


Figure 5: The pinhole camera model used in Approach 2.

In this regard, we first need to identify the 3D coordinates of point P in 3D camera coordinates. Based on Figure 6, the 3D coordinates of point P in 3D camera coordinates are presented in Eq. 5.

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} p_1 - h_1 \\ -(p_2 - h_2) \\ f \end{pmatrix} \tag{5}$$

In the follow-up step, we need to identify the 3D coordinates of point P in 3D world coordinates. In this step, we temporally assumed that the camera's pitch angle, yaw angle, and roll angle were equal to 0. The height of the camera mounted on the ego vehicle from the road surface is named h, and we assumed that the world coordinates are located on the road surface. The 3D coordinates of point P in 3D world coordinates (w_1, w_2, w_3) are given in Eq. 6.

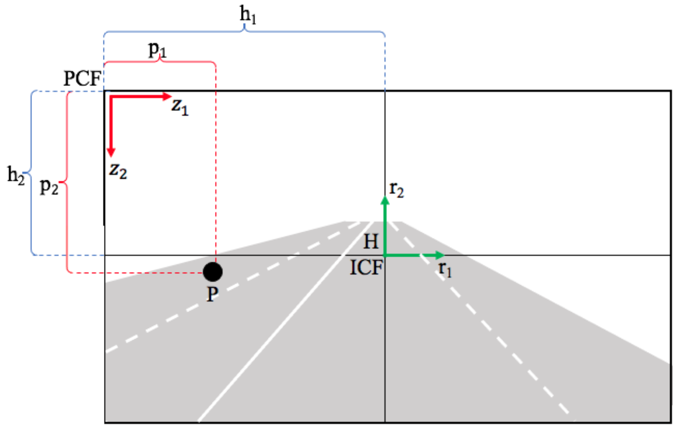


Figure 6: The image plane used in Approach 2.

$$\begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = \begin{pmatrix} p_1 - h_1 \\ -(p_2 - h_2) + h \\ f \end{pmatrix} \tag{6}$$

After that, we need to find a mathematical expression for all points that lie on the viewing ray from camera center C through point P in world coordinates, as presented in Eq. 7, where ξ defines any position along the viewing ray.

$$\begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = \begin{pmatrix} 0 \\ h \\ 0 \end{pmatrix} + \xi \cdot \begin{pmatrix} p_1 - h_1 \\ -(p_2 - h_2) \\ f \end{pmatrix} \tag{7}$$

As we assumed, the world coordinate system is located on the road surface; therefore, the height of any point on the road surface in the world coordinate system is equal to 0. So we can express Eq. 7 as Eq. 8.

$$\begin{pmatrix} w_1 \\ 0 \\ w_3 \end{pmatrix} = \begin{pmatrix} 0 \\ h \\ 0 \end{pmatrix} + \xi \cdot \begin{pmatrix} p_1 - h_1 \\ -(p_2 - h_2) \\ f \end{pmatrix} \tag{8}$$

Finally, to increase the accuracy of this estimation, we need to consider the camera's pitch angle σ. Rotating the camera by angle σ has no effect on the location of point P in the w_1 direction, but the point's location in the w_2 and w_3 directions will be affected by this rotation. Therefore, the following rotation is applied:

$$\begin{pmatrix} w_1 \\ 0 \\ w_3 \end{pmatrix} = \begin{pmatrix} 0 \\ h \\ 0 \end{pmatrix} + \xi \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(-\sigma) & -\sin(-\sigma) \\ 0 & \sin(-\sigma) & \cos(-\sigma) \end{pmatrix} \cdot \begin{pmatrix} p_1 - h_1 \\ -(p_2 - h_2) \\ f \end{pmatrix} \tag{9}$$

Therefore, ξ, w_1, and w_3 are calculated as follows:

$$\xi = \frac{-h}{(\cos(-\sigma) \cdot (-p_2 + h_2) - \sin(-\sigma) \cdot f)} \tag{10}$$

$$w_1 = \xi \cdot (p_1 - h_1) \tag{11}$$

$$w_3 = \xi \cdot (\sin(-\sigma) \cdot (-p_2 + h_2) + \cos(-\sigma) \cdot f) \tag{12}$$

To calculate ξ, w_1, and w_3, we need to estimate the camera's height from the road surface h, the camera's pitch angle σ, and the camera's focal length f.

- Estimating the camera's focal length f

  In Approach 2, the camera's focal length f in pixels is calculated based on the trigonometric relation presented in Eq. (13). We used the horizontal number of pixels n_h from the video's resolution and the camera's horizontal FOV ρ in degrees [26].

  $$f = \frac{n_h}{2 \cdot \tan(\frac{\rho}{2})} \tag{13}$$

- Estimating the camera's pitch angle σ

  To estimate the camera's pitch angle σ, we used a vanishing point estimated based on the lane detection.

  To detect lanes, as we presented in [8] and [9], we used canny edge detection [27] and the progressive probabilistic Hough transform [28][29].

In this study, we go further to identify the vanishing point based on the detected parallel lines on the road nearby the ego vehicle.

To estimate the pitch angle σ, we used the camera's focal length (f) and the vertical differences between the principal point $H = (h_1, h_2)$ and the vanishing point $J = (j_1, j_2)$, as shown in Figure 7. In this figure, the blue lines represent the detected parallel lines on the road nearby the ego vehicle in a perspective view. Based on this figure, the camera's pitch angle σ can be calculated by Eq. 14.

$$\sigma = \text{atan2}(j_2 - h_2, f) \tag{14}$$

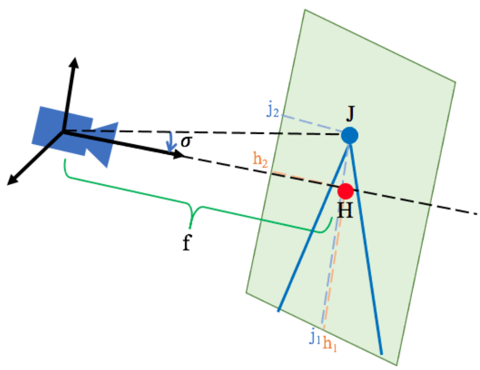


Figure 7: The camera's pitch angle σ and vanishing point used in Approach 2.

- Estimating the camera's height h from the road surface

  To estimate the height h of the camera mounted on the ego vehicle from the road surface by considering the camera's pitch angle σ, we applied Thales's theorem [30]. Thales's theorem in this context is presented in Figure 8. The variables in this figure are defined as follows:

$$A = \frac{f}{\cos \sigma} \tag{15}$$

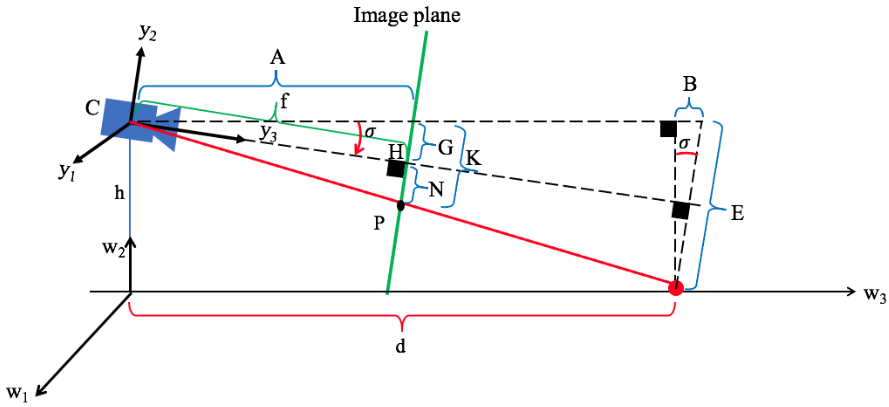


Figure 8: The camera's height from the road surface by considering the camera's pitch angle σ used in Approach 2.

$$B = h \cdot \tan\sigma \tag{16}$$

$$E = \frac{h}{\cos\sigma} \tag{17}$$

$$G = f \cdot \tan\sigma \tag{18}$$

$$K = G + N \tag{19}$$

To estimate the camera's height h based on Thales's theorem and the estimated distance d by Approach 1, we have the following equation.

$$\frac{K}{E} = \frac{A}{d + B} \tag{20}$$

By simplifying Eq. 20, h in meters is calculated as follows:

$$h = \frac{(N + f \cdot \tan\sigma) \cdot d \cdot (\cos\sigma)^2}{f - (N + f \cdot \tan\sigma) \cdot \tan\sigma \cdot (\cos\sigma)^2} \tag{21}$$

Finally, by estimating the camera's height from the road surface h, the camera's pitch angle σ, and the camera's focal length f, we can calculate ξ, w_1, and

w_3. Because w_1 and w_3 represent point P in 3D world coordinates, where w_2=0, we can estimate distance d by Approach 2 based on the Euclidean distance between w_1 and w_3 as presented in Eq. 22 [31], and estimate angle θ based on trigonometry presented in Eq. 23 [32]. Finally, angle α can be estimated based on the proposed conditions in Section 3.2.I.iii.

$$d = \sqrt{w_1^2 + w_3^2} \tag{22}$$

$$\theta = \text{atan2}\left(w_3, w_1\right) \tag{23}$$

3.3. Estimating the geolocation of the target vehicle

To estimate the target vehicle's geolocation with both approaches, we used Eq. 24 - Eq. 27 [10]. In these formulas, the variables are as below:

(ℓ_1, g_1) represent the geolocation of the ego vehicle

(ℓ_2, g_2) represent the geolocation of the target vehicle

R represents the Earth's radius

d represents the estimated distance between the ego vehicle and the target vehicle by both approaches

α represents the estimated angle between the north N and d by both approaches

$$\ell_2 = \text{asin}(\sin(\ell_1) \cdot \cos(d/R) + \cos(\ell_1) \cdot \sin(d/R) \cdot \cos(\alpha)) \tag{24}$$

$$U = \sin(\alpha) \cdot \sin(d/R) \cdot \cos(\ell_1) \tag{25}$$

$$V = \cos(d/R) - \sin(\ell_1) \cdot \sin(\ell_2) \tag{26}$$

$$g_2 = g_1 + \text{atan2}(U, V) \tag{27}$$

4. Experiments and results

In this paper, we carried out experiments using real-world traffic data to demonstrate the effectiveness of both proposed approaches for estimating the target vehicle's geolocation by an ego vehicle's vision.

4.1. Data collection

We used three vehicles and drove them by the following the pre-defined scenarios in Chengdu, China. All vehicles were equipped with two GoPro Hero 7 cameras, and each camera included a built-in GPS receiver. We used the GoPro Hero 7 camera as it provides us with both visual information and GPS data of the vehicles. One of the cameras mounted on the front window glass looked forward through the window, and another mounted on the back window glass looked backward. The purpose of mounting two cameras on the same vehicle was to improve the ego vehicle's self-location in the pre-processing step [20], and to collect more data for future studies. To estimate the target's geolocation, this study used only the footage collected from the camera mounted on the front window glass. We used one of these vehicles as an ego vehicle. The other two vehicles were treated as target vehicles. The GPS data collected by the target vehicles were used as ground truth to assess our proposed approaches.

The settings of the GoPro Hero 7 cameras we used were as follows. The used mode was 4×3 and linear with a zoom $= 0\%$. During the recording of the footage, the video's resolution and the frame rate were 1920×1440 and 60 frames per second (FPS), respectively. We adjusted the video's resolution to 960×720 and the frame rate to 1 FPS to apply pre-processing and vehicle detection.

Figure 9 shows the studied scenarios, called Scenario S1 and Scenario S2. In Scenario S1, ego vehicle v3 and target vehicles v1 and v2 are driven in the same direction on a straight trajectory. The purpose of this scenario was to evaluate our proposed approaches with one of the target vehicles driving on the same lane as the ego vehicle and the other target vehicle driving on the next lane. In

Scenario S2, ego vehicle v3 and target vehicles v1 and v2 are driven in opposite directions on a straight trajectory. In the scenario in which the vehicles are driven in opposite directions, the period between detecting a target vehicle via an ego vehicle until both vehicles pass each other is short, so the number of estimated locations is limited to this short period.

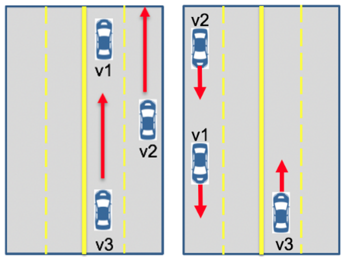


Figure 9: The scenarios studied using both approaches.

4.2. Experiments

The experiments were run using a laptop with a 3.1 GHz Intel Core i5 processor and Intel Iris Plus Graphics 650 1536 MB. As we explained in Section 4.1, we chose to use 1 FPS when analyzing the video, with the purpose of making a trade-off between the amount of generated data and the running time of the system. Also, as the vehicles' speeds were low (between 19.55 km/h and 30.18 km/h on average), there were hardly any informative changes in the vehicle's speed, distance, angle, and location within less than one second. Therefore, analyzing the data with a higher frequency could not provide much extra information. We measured the running time based on the experimental studies using Approaches 1 and 2 after the data pre-processing. We found that the system based on Approach 1 ran 0.680 FPS on average, and that the system based on Approach 2 ran 0.684 FPS on average, to output the geolocation of the target vehicle from the input videos and the pre-processed ego vehicle's geolocations.

4.2.1. Evaluation

The evaluation is done in two steps: (1) plotting the estimated geolocations (i.e., latitude and longitude) of the target vehicles on Google Maps and (2) analyzing the distance vector between the estimated target vehicle's geolocations by both proposed approaches and the ground-truth data.

Figure 10 and Figure 11 present the outputs of the experiments related to Scenario S1 and Scenario S2, respectively. In these figures, the white polyline in (a) and (d) shows the ground-truth trajectory of the observed target vehicle. The red polyline in (b) and (e) shows the trajectory of the target vehicle estimated with Approach 1. The blue polyline in (c) and (f) shows the trajectory of the target vehicle estimated with Approach 2. These figures show that the trajectories of the target vehicle estimated with both approaches are plotted on the correct lane of the road and that they almost overlapped with the ground-truth trajectory. This means that both approaches enable us to estimate the trajectory of the target vehicle accurately on the right lane of the road. As expected, with Scenario 2, the number of plotted positions along the trajectory, presented in Figure 11, are limited (2-3 points) because of the opposite movement directions of the ego vehicle and target vehicle and the short sensing time.

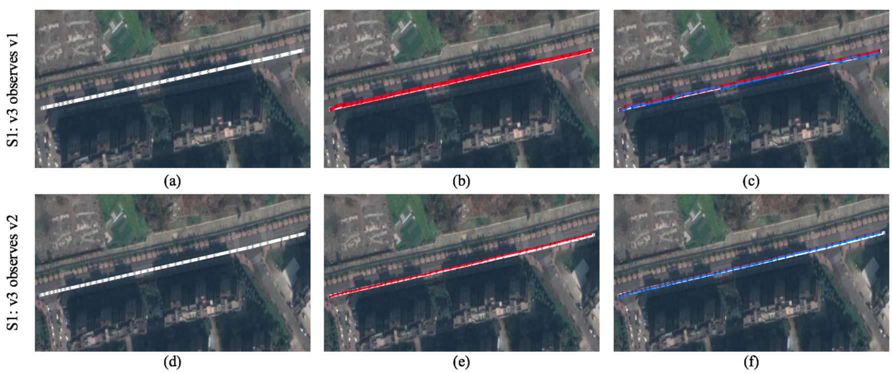


Figure 10: The estimated trajectory with Scenario S1 on the map. Two cases are considered: (1) ego vehicle v3 observes target vehicle v1 (a-c) and (2) ego vehicle v3 observes target vehicle v2 (d-f).

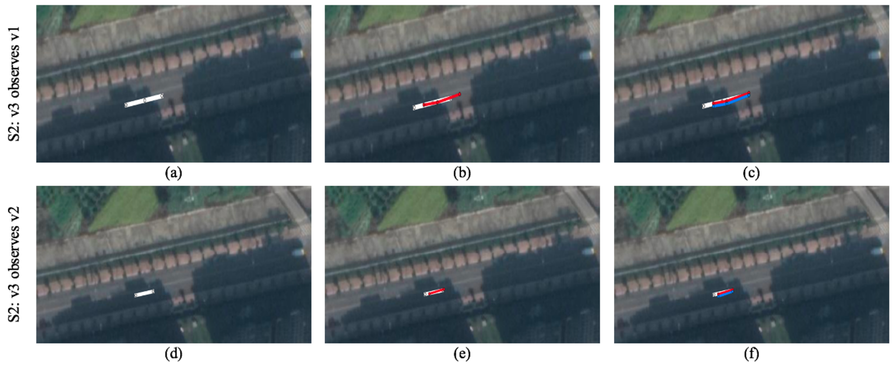


Figure 11: The estimated trajectory with Scenario S2 on the map. Two cases are considered: (1) ego vehicle v3 observes target vehicle v1 (a-c) and (2) ego vehicle v3 observes target vehicle v2 (d-f).

To analyze the estimated geolocations of the target vehicle numerically, we used the distance vector between the ground truth and the geolocations estimated by both approaches. This analysis provides information regarding the estimation deviation in our proposed approaches. Figure 12 visualizes our findings related to Scenario S1 and Scenario S2. Our numerical findings are summarized in Table 1, as well. As Figure 12 and Table 1 show, the geolocation estimation deviation (based on the absolute values) with Approach 1 is on average between 1.38 m and 3.54 m. The geolocation estimation deviation with Approach 2 is on average between 1.4 m and 3.51 m. Figure 12 (a) and (b) shows a slightly upward trend between the plotted points. This result may be explained by the fact that the collected data by a GPS receiver to provide the ego vehicle's location and ground truth data of the target vehicle's position were not noise-free. As we expected, Figure 12 (e) - (h) represent the limited points as Scenario S2 focused on studying the vehicle movement in the opposite directions and the sensing lifetime was limited. In addition, as Table 1 shows, the highest on average geolocation estimation deviation with both approaches is obtained in Scenario S2, when v3 observes v1. A possible explanation for this might be that as in this scenario, only limited geolocations (2-3) were estimated, so the estimation deviation of one point has a big effect on the average error.

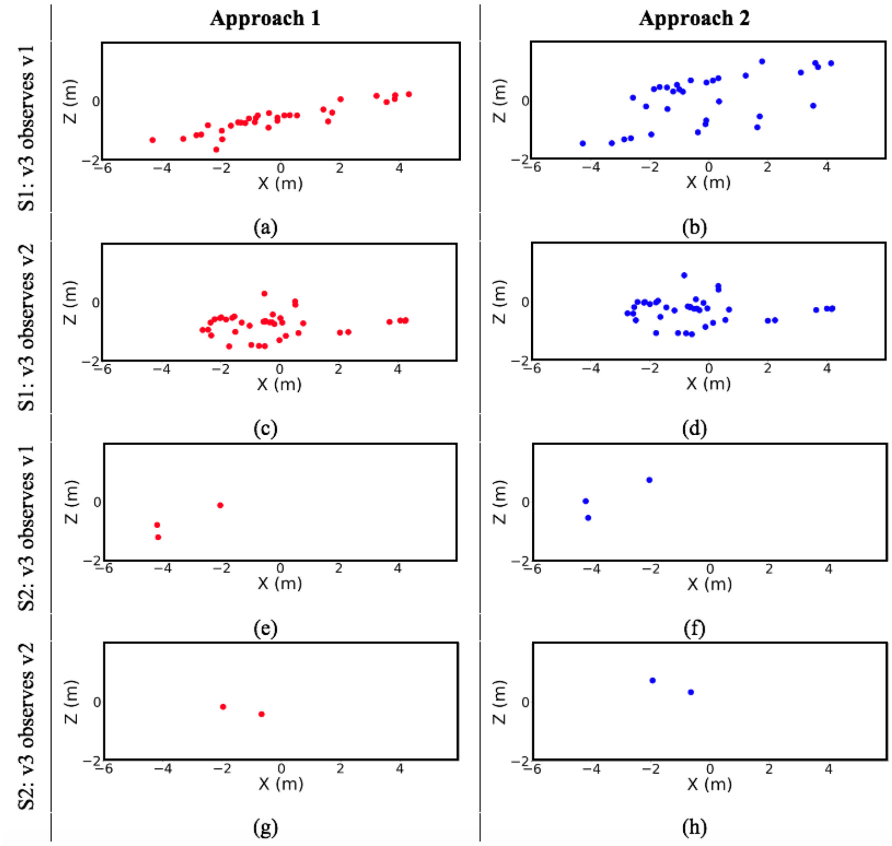


Figure 12: The distance vectors between ground truth and estimated geolocations with Approach 1 (a, c, e, and g) and Approach 2 (b, d, f, and h) for Scenario 1 and Scenario 2. Two cases are considered: (1) ego-vehicle v3 observes target vehicle v1, and (2) ego-vehicle v3 observes target vehicle v2. X shows longitudinal and Z shows lateral directions

To analyze our proposed approaches further, we applied the root mean square error (RMSE) to the distance vector between the estimated geolocations and the ground truth to show the estimation deviation. The calculated RMSE related to Approach 1 was between 1.5 m and 3.7 m (2.39 m on average). The calculated RMSE related to Approach 2 was between 1.55 m and 3.63 m (2.37 m on average). Overall, these results indicate that, in the studied scenarios, Approach 2 is slightly (about 0.02 m on average) better than Approach 1.

Table 1: Evaluation results.

S#	V#	Estimation deviation of Approach 1 (m)				Estimation deviation of Approach 2 (m)			
		Min	**Avg**	**Max**	**RMSE**	**Min**	**Avg**	**Max**	**RMSE**
S1	**v3 observes v1**	0.50	2.03	4.51	2.35	0.35	2.02	4.51	2.34
	v3 observes v2	0.47	1.74	4.31	2.03	0.19	1.63	4.18	1.96
S2	**v3 observes v1**	2.040	3.54	4.33	3.70	2.18	3.51	4.20	3.63
	v3 observes v2	0.79	1.38	1.97	1.50	0.72	1.4	2.07	1.55

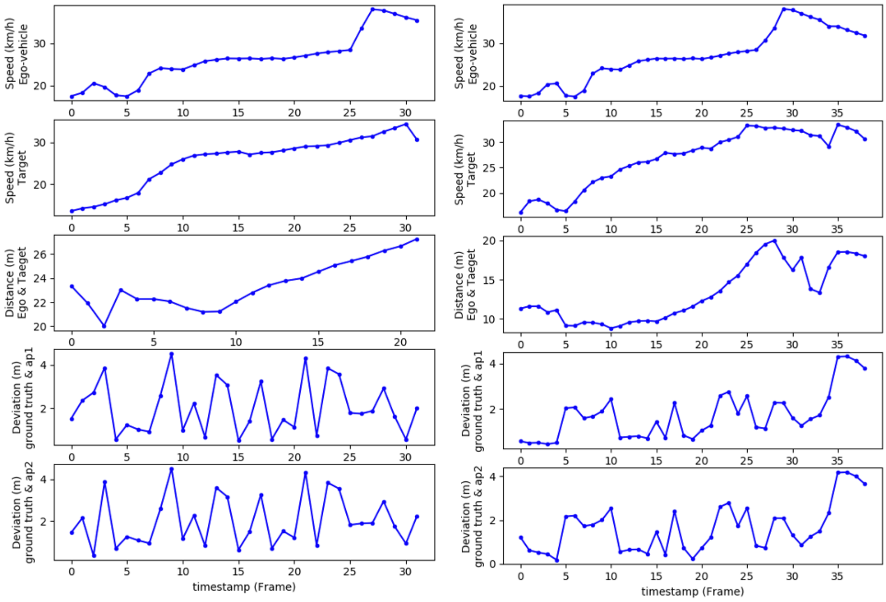
Table 2: Traffic data measurements with Scenario 1 and Scenario 2.

S#	V#	Ego vehicle speed (km/h)			Target vehicle speed (km/h)			Dist. ego and target vehicles (m)		
		Min	**Avg**	**Max**	**Min**	**Avg**	**Max**	**Min**	**Avg**	**Max**
S1	**v3, v1**	17.50	26.38	38.04	13.64	25.65	34.35	20.04	25.02	38.84
	v3 , v2	17.50	27.14	38.04	16.19	26.94	33.4	8.81	13.25	20.01
S2	**v3, v1**	16.98	19.55	22.04	29.12	29.40	29.55	17.35	30.78	44.14
	v3 , v2	22.04	23.2	24.36	29.63	30.18	30.74	20.53	27.95	35.37

As the speed of the vehicles and the distance between them may influence our estimation accuracy, we studied the effect of the speed of the ego vehicle and target vehicle and the distance between them on estimating the target vehicle's geolocation. As an example, we presented our findings related to Scenario 1 in Figure 13. Figure 13 shows that in the case in which v3 observes v1, with both vehicles driving on the same lane, changes in the distance and speed have no significant effect on our estimation accuracy. However, when v3 observes v2, with both vehicles driving on different lanes, increasing the distance between the vehicles, as caused by changes in the vehicles' speed, increases the estimation deviation. However, as the speed of the vehicles and the distance between them were limited in the studied scenarios, more studies are needed to validate these findings in the future. The extracted traffic data (i.e., the ego vehicle's and the target vehicle's speed and the distance between the vehicles) are summarized in Table 2.

4.2.2. Comparison of Approach 1 with Approach 2

Although Approach 2 can estimate the target vehicle's geolocation slightly better than Approach 1 on average, it cannot always be the optimum approach. Therefore, we investigated the deviations between the ground truth and esti-

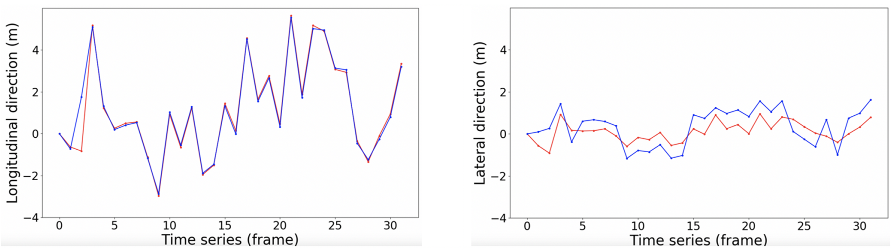


Scenario S1: v3 observes v1

Scenario S1: v3 observes v2

Figure 13: Plotting the deviations with regard to ground truth to study the effect of speed and distance (here shown for Scenario S1.)

mated geolocations in the longitudinal and lateral directions in time series using both approaches. Our findings related to Scenario 1 as an example are presented in Figure 14.



(a) (b)

Figure 14: The deviation between the ground truth and the estimated geolocations with Approach 1 (the red polyline) and Approach 2 (the blue polyline) in the lateral and longitudinal directions in Scenario S1.

As Figure 14 shows, the deviations between the ground truth and the estimated geolocations in the longitudinal direction (a) with Approach 1 and Approach 2 are almost overlapped. However, this parameter in the lateral direction (b) is more different. This result may be explained by the fact that the ego vehicle's geolocation was utilized in estimating the angle between the ego vehicle and the target vehicle. However, the used ego vehicle's geolocation is not noise-free. In addition, the ground truth data to apply the estimation deviation were collected by such GPS receivers, as well. Another possible explanation for this is that the employed methodology to estimate the distance between ego vehicle and target vehicle with both approaches were different. The first approach relied on the accuracy of the bounding box added by YOLO-v3 around the target vehicle, and the second approach used the central point on the bottom edge of the bounding box. In general, by considering the proposed methodology in each approach, we can conclude in case the lane marks and the vanishing point, which are needed by the second approach, are available, we can use Approach 2 as it is identified as the more accurate approach; otherwise, Approach 1 can be used. In both cases, enhancing the accuracy of the GPS receiver and the vehicle detection algorithms are vital.

5. Discussion

To date, most of the existing studies on collecting traffic data focused on two main approaches: (1) stationary sensors and (2) V2V and V2I communications, which require most of the vehicles to have sensing and communication capabilities. However, the much-debated question is how to estimate the geolocation of the target vehicle in mixed traffic. To address this gap, this study proposed two approaches to dynamically estimating the target vehicle's geolocation based on an ego vehicle's vision capability.

5.1. Comparison with related work

As presented in Section 2, most of the existing studies regarding estimating the position of the target vehicle focused on the relative target vehicle's location estimation (e.g., [15][19]) and inter-vehicle distance estimation (e.g., [17][18]). Despite the importance of the target-vehicle localization, there remains a paucity of empirical studies on estimating the target vehicle's geolocation in a GPS coordinate system in order to enhance the ITMS awareness about the traffic scene. In addition, to be able to generalize the proposed approaches in reality, using real traffic data to run experiments is vital; however, most of the studies applied the experiments by using a simulator (e.g., [15]). Moreover, studying the real traffic data from the urban area, which can reflect the possible estimation uncertainties and sources of noise (e.g., tall buildings that affect the GPS receiver accuracy) are important.

Therefore, the main objective of this study is to go beyond the target vehicle localization and estimate the relative angles besides estimating distance and ego vehicle's geolocation to be able to estimate the geolocation of the target vehicle dynamically. To estimate the ego vehicle's geolocation based on a low-cost GPS receiver, we used the proposed approach in [20]. To estimate distance and relative angle, we proposed two new approaches based on object detection and image processing and geometric computation. To assess our proposed approaches, we developed a new system and ran experiments on real traffic data collected from the urban area.

The experiments on estimating the target vehicle's geolocation showed that the estimation deviation with Approach 1 was on average between 1.38 m to 3.54 m. The results with Approach 2 were between 1.4 m and 3.51 m. In our study, the ego vehicle's speed was between 16.98 km/h and 38.04 km/h, the target vehicle's speed was between 13.64 km/h and 34.35 km/h, and the distance between the ego vehicle and the target vehicle varied between 8.81 m and 44.14 m. Comparison of our findings with those of similar studies focused on estimating the distance to the target vehicle (e.g., [19]) confirms that our geolocation estimation deviation is reasonable. For example, the approach proposed by Giesbrecht et al. [19] yielded an estimated distance with a mean and maximum error of 0.72 m and 2.42 m, respectively, with a follower speed between 7.6 km/h and 10.2 km/h and a follower separation between 10.46 m and 23.71

m. However, our experiments were applied to real traffic data collected from the urban area by considering the higher speed and distance.

5.2. Limitations of our proposed approaches

Although the experiments confirmed that both our proposed approaches were able to estimate the geolocation of the target vehicle accurately on the right lane of the road, each approach has some pros and cons. For instance, Approach 1 is tightly connected to the vehicle detection accuracy because we employed the size of the bounding box added by YOLO around the target vehicle to estimate its distance. However, during the experiments, we observed that these bounding boxes were shaking during the trajectory, and the size of them varies between frames, which can affect the accuracy of the distance estimation.

Approach 2 has some limitations as well. For instance, like Approach 1, Approach 2 is tightly dependent on the accuracy of the central point on the bottom edge of the bounding box around the target vehicle, which is used to transform 2D space into a 3D space. Therefore, it has a direct effect on estimating the distance of the target vehicle. In addition, the camera's pitch angle was considered in Approach 2 to enhance the localization accuracy. To estimate the camera's pitch angle, we used a vanishing point based on the parallel lane on the road. Therefore, enhancing the lane detection accuracy would be beneficial for enhancing the accuracy. Furthermore, this pitch angle was caused by our manual installation of the camera with a suction cup on the vehicle; therefore, installing the camera more precisely is highly recommended.

In addition, in the both Approaches, to estimate the angle, we used the ego vehicle's movement direction and the central point on the bottom edge of the bounding box around the target vehicle. Therefore, the estimation error of each of these parameters has a negative effect on the angle estimation accuracy. As the ego vehicle's movement direction was estimated based on the ego vehicle's GPS data, if the GPS data is noisy, the estimation deviation will increase. The same is true for the central point on the bottom edge of the bounding box around the target vehicle, which is estimated by YOLO. The experiments showed that the central point was not stable and was shaking between frames. Therefore, to mitigate the estimation deviation, enhancing the accuracy of the YOLO and GPS data is needed.

Another source of uncertainty is that the noisy low-cost GPS receiver used to collect the ground truth data in a metropolitan area surrounded by many tall buildings. Although we applied data pre-processing to mitigate the GPS receiver noise, it is still not noise-free. Moreover, since the study was limited to vehicle movements along straight streets with limited scenarios, more studies are needed to be able to generalize our proposed approaches by considering various vehicle movements scenarios.

6. Conclusion

The main goal of the current study was to dynamically estimate the geolocation of mobile target vehicles via a low-cost front-facing monocular camera on a mobile ego vehicle. To estimate the target vehicle's geolocation, the distance between the ego vehicle and the target vehicle and the relative angle are needed. In this regard, we proposed two approaches: (1) object detection and image processing and (2) geometric computation.

The results of the evaluation using real traffic data confirmed that our algorithms were able to estimate the geolocation of the target vehicles accurately. Taken together, these findings confirmed the feasibility of a vehicle-mounted monocular camera for estimating the location of target vehicles in mixed traffic. The present study lays the groundwork for future research on using an ego vehicle as a mobile sensor to collect traffic data to reduce the traffic cost and improve ITMS efficiency. Further studies that take these data types into account will be needed to increase the accuracy and enhance the generalizability by considering various scenarios.

References

[1] E. Namazi, J. Li, C. Lu, Intelligent intersection management systems considering autonomous vehicles: A systematic literature review, IEEE Access 7 (2019) 91946–91965.

[2] H. A. Oliveira, F. R. Barbosa, O. M. Almeida, A. P. Braga, A vehicle classification based on inductive loop detectors using artificial neural networks, in: 2010 9th IEEE/IAS International Conference on Industry Applications-INDUSCON 2010, IEEE, 2010, pp. 1–6.

[3] H. A. Kurdi, Review of closed circuit television (cctv) techniques for vehicles traffic management, International Journal of Computer Science & Information Technology 6 (2) (2014) 199.

[4] B. A. Kiisa Wiegand, Challenges of the day-today operation of a traffic monitoring program (2016).

[5] M. N. Ahangar, Q. Z. Ahmed, F. A. Khan, M. Hafeez, A survey of autonomous vehicles: enabling communication technologies and challenges, Sensors 21 (3) (2021) 706.

[6] Sae international releases updated visual chart for its "levels of driving automation" standard for self-driving vehicles.
URL `https://www.sae.org/news/press-room/2018/12/sae-international-releases-updated-visual-chart-for-its-%E2%80%9Clevels-of-driving-automation%E2%80%9D-standard-for-self-driving-vehicles`

[7] A. Raj, J. A. Kumar, P. Bansal, A multicriteria decision making approach to study barriers to the adoption of autonomous vehicles, Transportation research part A: policy and practice 133 (2020) 122–137.

[8] E. Namazi, R. N. Holthe-Berg, C. S. Lofsberg, J. Li, Using vehicle-mounted camera to collect information for managing mixed traffic, in: 2019 15th In-

ternational Conference on Signal-Image Technology & Internet-Based Systems (SITIS), IEEE, 2019, pp. 222–230.

[9] E. Namazi, J. Li, R. Mester, C. Lu, Identifying and counting vehicles in multiple lanes by using a low-cost vehicle-mounted sensor for intelligent traffic management systems, in: International Conference on Hybrid Artificial Intelligence Systems, Springer, 2020, pp. 598–611.

[10] Movable-Type, Calculate distance, bearing and more between latitude/longitude points.
URL `https://www.movable-type.co.uk/scripts/latlong.html`

[11] E. Javanmardi, Y. Gu, M. Javanmardi, S. Kamijo, Autonomous vehicle self-localization based on abstract map and multi-channel lidar in urban area, IATSS research 43 (1) (2019) 1–13.

[12] A. Chehri, N. Quadar, R. Saadane, Survey on localization methods for autonomous vehicles in smart cities, in: Proceedings of the 4th International Conference on Smart City Applications, 2019, pp. 1–6.

[13] N. Acosta, J. Toloza, Techniques to improve the gps precision, International Journal of Advanced Computer Science and Applications 3 (8).

[14] Z. Huang, S. Qiao, N. Han, C.-a. Yuan, X. Song, Y. Xiao, Survey on vehicle map matching techniques, CAAI Transactions on Intelligence Technology 6 (1) (2021) 55–71.

[15] M. S. Ifthekhar, N. Saha, Y. M. Jang, Stereo-vision-based cooperative-vehicle positioning using occ and neural networks, Optics Communications 352 (2015) 166–180.

[16] J. Hayakawa, B. Dariush, Ego-motion and surrounding vehicle state estimation using a monocular camera, in: 2019 IEEE Intelligent Vehicles Symposium (IV), IEEE, 2019, pp. 2550–2556.

[17] J. Lee, Intervehicle distance estimation through camera images, Journal of Electronic Imaging 27 (6) (2018) 063001.

[18] D.-Y. Huang, C.-H. Chen, T.-Y. Chen, W.-C. Hu, K.-W. Feng, Vehicle detection and inter-vehicle distance estimation using single-lens video camera on urban/suburb roads, Journal of Visual Communication and Image Representation 46 (2017) 250–259.

[19] J. L. Giesbrecht, H. K. Goi, T. D. Barfoot, B. A. Francis, A vision-based robotic follower vehicle, in: Unmanned Systems Technology XI, Vol. 7332, International Society for Optics and Photonics, 2009, p. 73321O.

[20] E. Namazi, R. Mester, C. Lu, M. M. Log, J. Li, Improving vehicle localization with two low-cost gps receivers, in: The sixth smart city applications international conference, 2021.

[21] J. Redmon, S. Divvala, R. Girshick, A. Farhadi, You only look once: Unified, real-time object detection, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 779–788.

[22] J. Redmon, A. Farhadi, Yolov3: An incremental improvement, arXiv preprint arXiv:1804.02767.

[23] Yolo: Real-time object detection.
URL `https://pjreddie.com/darknet/yolo/`

[24] D. Wang, C. Li, S. Wen, Q.-L. Han, S. Nepal, X. Zhang, Y. Xiang, Daedalus: Breaking nonmaximum suppression in object detection via adversarial examples, IEEE Transactions on Cybernetics.

[25] J. Redmon, Pytorch-yolov3-kitti.
URL `https://github.com/packyan/PyTorch-YOLOv3-kitti`

[26] Gopro, Hero7 field of view (fov) information.
URL `https://gopro.com/help/articles/question_answer/hero7-field-of-view-fov-information?sf96748270=1`

[27] L. Ding, A. Goshtasby, On the canny edge detector, Pattern Recognition 34 (3) (2001) 721–725.

[28] C. Galamhos, J. Matas, J. Kittler, Progressive probabilistic hough transform for line detection, in: Proceedings. 1999 IEEE computer society conference on computer vision and pattern recognition (Cat. No PR00149), Vol. 1, IEEE, 1999, pp. 554–560.

[29] J. Matas, C. Galambos, J. Kittler, Robust detection of lines using the progressive probabilistic hough transform, Computer vision and image understanding 78 (1) (2000) 119–137.

[30] Thales theorem and figure.
URL `https://sites.math.washington.edu/~king/coursedir/m444a02/class/10-21-thales.html`

[31] Euclidean distance.
URL `https://en.wikipedia.org/wiki/Euclidean_distance`

[32] Trigonometry.
URL `https://en.wikipedia.org/wiki/Trigonometry`

Paper F:

Traffic Awareness Through Multiple Mobile Sensor Fusion

Elnaz Namazi, Rudolf Mester, Jingyue Li, Chaoru Lu, Meng Tang, and Ying Xiong

IEEE Sensors Journal

This paper is not yet published and is therefore not included.

ISBN 978-82-326-5727-8 (printed ver.)
ISBN 978-82-326-5631-8 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (online ver.)

NTNU

Norwegian University of
Science and Technology