A Real-Time Computer Vision Based Framework For Urban Traffic Safety Assessment and Driver Behavior Modeling Using Virtual Traffic Lanes

Awad T. Abdelhalim

Dissertation submitted to the Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Civil Engineering

Montasir M. Abbas, Chair

Antonio A. Trani

Bert Huang

Robert Hildebrand

September 17, 2021

Blacksburg, Virginia

Keywords: Traffic safety, Driver behavior modeling, Microscopic simulation.

A Real-Time Computer Vision Based Framework For Urban Traffic Safety Assessment and Driver Behavior Modeling Using Virtual Traffic Lanes

Awad T. Abdelhalim

(ABSTRACT)

Vehicle recognition and trajectory tracking plays an integral role in many aspects of Intelligent Transportation Systems (ITS) applications; from behavioral modeling and car-following analyses to congestion prevention, crash prediction, dynamic signal timing, and active traffic management. This dissertation aims to improve the tasks of multi-object detection and tracking (MOT) as it pertains to urban traffic by utilizing the domain knowledge of traffic flow then utilize this improvement for applications in real-time traffic performance assessment, safety evaluation, and driver behavior modeling. First, the author proposes an ad-hoc framework for real-time turn count and trajectory reconstruction for vehicles passing through urban intersections. This framework introduces the concept of virtual traffic lanes representing the eight standard National Electrical Manufacturers Association (NEMA) movements within an intersection as spatio-temporal clusters utilized for movement classification and vehicle re-identification. The proposed framework runs as an additional layer to any multi-object tracker with minimal additional computation. The results obtained for a case study and on the AI City benchmark dataset indicate the high ability of the proposed framework in obtaining reliable turn count, speed estimates, and efficiently resolving the vehicle identity switches which occur within the intersection due to detection errors and occlusion. The author then proposes the utilization of the high accuracy and granularity trajectories obtained from video inference to develop a real-time safety-based driver behavior model, which

managed to effectively capture the observed driving behavior in the site of study. Finally, the developed model was implemented as an external driver model in VISSIM and managed to reproduce the observed behavior and safety conflicts in simulation, providing an effective decision-support tool to identify appropriate safety interventions that would mitigate those conflicts. The work presented in this dissertation provides an efficient end-to-end framework and blueprint for trajectory extraction from road-side traffic video data, driver behavior modeling, and their applications for real-time traffic performance and safety assessment, as well as improved modeling of safety interventions via microscopic simulation.

A Real-Time Computer Vision Based Framework For Urban Traffic Safety Assessment and Driver Behavior Modeling Using Virtual Traffic Lanes

Awad T. Abdelhalim

(GENERAL AUDIENCE ABSTRACT)

Traffic crashes are one of the leading causes of death in the world, averaging over 3,000 deaths per day according to the World Health Organization. In the United States alone, there are around 40,000 traffic fatalities annually. Approximately, 21.5% of all traffic fatalities occur due to intersection-related crashes. Intelligent Transportation Systems (ITS) is a field of traffic engineering that aims to transform traffic systems to make safer, more coordinated, and 'smarter' use of transport networks. Vehicle recognition and trajectory tracking, the process of identifying a specific vehicle's movement through time and space, plays an integral role in many aspects of ITS applications; from understanding how people drive and modeling that behavior, to congestion prevention, on-board crash avoidance systems, adaptive signal timing, and active traffic management. This dissertation aims to bridge the gaps in the application of ITS, computer vision, and traffic flow theory and create tools that will aid in evaluating and proactively addressing traffic safety concerns at urban intersections. The author presents an efficient, real-time framework for extracting reliable vehicle trajectories from roadside cameras, then proposes a safety-based driving behavior model that succeeds in capturing the observed driving behavior. This work is concluded by implementing this model in simulation software to replicate the existing safety concerns for an area of study, allowing practitioners to accurately model the existing safety conflicts and evaluate the

different operation and safety interventions that would best mitigate them to proactively prevent crashes.

Dedication

This dissertation is dedicated to my parents, my sisters, my family, and my friends. Your unconditional love, support, and belief in what I can achieve makes me a better person every day.

Acknowledgments

I have started working with Professor Montasir Abbas in the Fall of 2016. One thing that was extremely easy to notice from the get-go is his enthusiasm for and unrelenting pursuit of knowledge. Seeing first-hand that someone so accomplished still seeks, finds joy, and has the humility of admitting that he should still gain more knowledge has definitely been a guiding light for me over the past five years and will continue to be so for the rest of my life. To myself, and all his other advisees, Professor Abbas has always put himself as a mentor and a friend before being a supervisor or employer. I'm extremely thankful and grateful for having had the opportunity to be mentored by him, academically and personally.

I would like to thank my committee members; Dr. Antonio Trani, Dr. Bert Huang, and Dr. Robert Hildebrand for generously setting aside the time to be a part of my doctoral committee. Their inputs, especially at the conceptualization phase of this work, and through their intriguing questions and feedback during my preliminary exam and proposal presentations have added immense value to this work.

I would like to thank Virginia Tech and the Town of Blacksburg for being my home away from home. The friends, colleagues, and the community here at Virginia Tech as a whole, who all truly live by Ut Prosim. I would like to especially thank the Graduate School, the Center for Enhancement of Engineering Diversity, the College of Engineering's Office of Enrollment and Academic Assessment, and the Residential College at West Ambler Johnston for supporting my graduate education, and providing me with fantastic environments to grow academically, professionally, and personally. Being surrounded by such a community of extraordinary people and citizen-scholars can bring out the best in anyone, and this work is theirs as much as it is mine.

Contents

List of Figures

List of Tables

List of Abbreviations

ADAS Advanced Driver-Assistance Systems

CNN Convolutional Neural Network

DLL Dynamic Link Library

GHR Gazis-Herman-Rothery (Car-following Model)

GPS Global Positioning System

GUI Graphical User Interface

IMU Inertial Measurement Unit

ITS Intelligent Transportation Systems

K-NN K-Nearest Neighbors (Supervised Machine Learning Algorithm)

MAE Mean Absolute Error

MPP Meters-Per-Pixel

MS-COCO Microsoft Common Objects in Context (Object Detection Dataset)

MSE Mean Squared Error

NEMA National Electrical Manufacturers Association

OVM Optimal Velocity Model(Car-following Model)

PET Post Encroachment Time

ROI Region of Interest

SSAM Surrogate Safety Assessment Model

TTC Time-to-Collision

VISSIM Verkehr In Städten SIMulationsmodel (Simulation Software)

VLVs Virtual Lane Vehicles

VT-Lane Virtual Traffic Lanes

YOLO You Only Look Once (Object Detection Framework)

Chapter 1

Introduction

1.1 Background and Problem Statement

Traffic crashes are one of the leading causes of death in the world, averaging over 3,000 deaths per day according to the World Health Organization [1]. In the United States alone, there are around 40,000 traffic fatalities annually, an equivalent of a mid-range jet airliner falling off the sky every day! Approximately, 21.5% of all traffic fatalities occur due to intersection-related crashes, which represent 40% of all reported crashes in the United States [2].

Intelligent Transportation Systems (ITS) aim to address transportation safety issues through transforming the typical passive and unaware transport systems into dynamic, proactive systems. Vehicle recognition and trajectory tracking play an integral role in many aspects of ITS applications; from behavioral modeling and car-following analyses to congestion prevention, crash warning, dynamic signal timing, and active traffic management. Vehicle trajectory analysis also plays a crucial role in the applications of path planning and cooperative driving.

Collecting vehicle trajectories from video data was a momentous task back in the day, with one of the major breakthroughs coming through the Next Generation Simulation (NGSIM) project [3]. However, the recent advancements in the areas of computer vision and deep learning have tremendously accelerated this area of research, allowing researchers to relent-

lessly develop new, faster, and more dependable object recognition and tracking models and algorithms that continuously outperform the previous ones. The ever-growing use of cameras to monitor traffic and for overall surveillance has also resulted in video data for traffic being more available, inexpensive, and far less disruptive to obtain compared to loop detector data, which inherently leads to more focus on developing better and more reliable video detection and tracking techniques. The detection and tracking of trajectories of urban traffic, especially through busy intersections, remains a very challenging problem due to a multitude of complications, including privacy concerns regarding the tracked vehicles, and technical barriers related to the quality and area of view provided by on-site cameras, occlusion, and computational requirements.

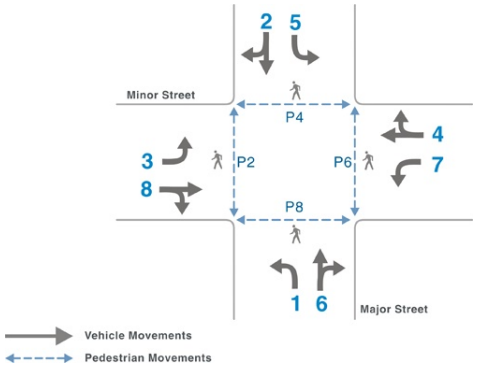


Figure 1.1: NEMA movements and phase labels.

The typical traffic movement assignment for signal controllers in the United States follows the National Electrical Manufacturers Association (NEMA) phases shown in Fig. 2.1 [4]. The author proposes to incorporate this prior knowledge of NEMA traffic movement patterns, alongside other laws of traffic flow, in the video-based trajectory tracking framework to obtain reliable turn counts, assist in addressing the vehicle identity switch problem due to occlusion, and aid in reconstructing missing vehicle trajectories. The proposed framework is evaluated and validated for the tasks of movement tracking and classification, and speed estimation. The produced trajectories are then utilized for automated large-scale driving

behavior calibration, and its application to improve safety assessment via simulation.

1.2 Dissertation Organization

The following chapters of this dissertation are organized as follows:

- Chapter 1 provides an introduction to the background of this work and a review of the relevant literature in the fields of object detection and tracking, trajectory-based safety, and driver behavior modeling.
- Chapter 2 introduces the Virtual Traffic Lanes (VT-Lane) framework, details of its implementation, and case studies for obtaining real-time vehicle trajectories and turn movement classifications at urban intersections.
- Chapter 3 provides an assessment of utilizing image processing and reference object scaling for obtaining reliable speed estimates for the trajectories extracted via VT-Lane.
- Chapter 4 is a value proposition for incorporating the high-granularity of the trajectories and speed estimates obtained via VT-Lane to develop and calibrate a novel real-time safety-based driver behavior model.
- Chapter 5 provides an implementation of the developed driver behavior model in simulation modeling, highlighting the substantial improvements that can be achieved for assessing safety interventions via microsimulation.
- Chapter 6 summarises this dissertation and provides the author's thoughts on future work directions.

1.3 Review of Literature

The work proposed and evaluated in this dissertation covers multiple topics that include object detection and tracking, trajectory-based safety, and driver behavior modeling. This chapter provides a review of the relevant literature in each of those topics, identifying the gaps that exist, and how the work presented in the following chapters of this dissertation contributes to bridging those gaps.

1.3.1 Object Detection and Tracking in Traffic Analysis and Safety Applications

The computer vision task of detecting objects from images and video frames can be subdivided into three major categories:

- Image Classification: The classical problem of identifying the main object in an image. Images might contain more than one object but the task is to identify a single main object within the image.
- Object Recognition: The task of identifying and localizing an object within an image.
- Multi-Object Recognition: Identifying and localizing multiple objects in the same image frame.

The three aforementioned categories of object recognition problems have been classically tackled by a myriad of computer-vision techniques and manipulations. As the field evolved, knowledge about the geometry, scenery, and the deformable part models [5], [6], [7] were utilized. But the massive recent strides have been accomplished due to the evolution of Convolutional Neural Networks (CNNs) which now form the basis of all state-of-the-art

object detection and tracking algorithms. As CNNs proved extremely dependable in object recognition tasks, a major breakthough was achieved by Ross Girshic proposing the Fast Region-based Convolutional Neural Network (Fast R-CNN) [8], followed by the faster, closer to real-time version Faster R-CNN in 2015 [9] which then had a frame rate of 5 fps. Low frame rates limited the number of real-time applications of such models, until the J Redmon et. al. and W. Liu et. al. respectively introduced their You Only Look Once (YOLO) and Single-Shot Detector (SSD) models. [10], [11], achieving performance suitable for real-time detection and tracking applications. Object detection models exponentially improved over the past few years, with newer versions of the aforementioned works massively outperforming their predecessors [12], [13], [14].

In the context of traffic engineering, the research focus in terms of object detection and tracking has classically been focused only on automobiles. It is only recently that researchers began focusing on tracking pedestrians and other roadway users. As early as the early 90s, a multitude of studies were conducted assessing various techniques for vehicle tracking and counting from video data, utilizing belief networks, adaptive background memory, and parametrized 3-D models [15], [16], [17]. The majority of the earlier methods were inapplicable during congestion due to the occlusion problem. In 1996, Beymer and Malik [18] proposed a feature-based model that tracks sub-features of vehicles instead of the entire automobile, making it less sensitive to partial occlusion. Beymer et. al. then extended their work and tested on several hours of video [19]. Feature-based tracking, however, was extremely computationally expensive, and Beymer's proposed feature-based would achieve maximum performance of 2 Hz under congested traffic conditions. Kamijo et. al. [20] proposed a spatio-temporal Markov random field algorithm for traffic monitoring at intersections. Their proposed model demonstrated very high accuracy in tracking vehicles, and they also introduced an incident detection method, albeit their data being insufficient to comprehensively evaluate. Several

studies in the following years provided massive contributions to the applications of vehicle tracking and trajectory analysis [3], [21], [22].

The aforementioned advancements in deep learning, the rapid development of more dependable models, and the rise of GPU computing have managed to tackle the majority of issues related to traffic detection and monitoring in terms of real-time performance and deployability. Recent research has been more focused on reliable means to extracting and predicting vehicle trajectories in urban environments. Dai et. al. proposed a framework to track vehicles and obtain traffic counts based on intersection segmentation on the basis of the entry approach [23]. Wu et. al [24] proposed a Hidden Markov Model to identify the vehicle trajectories within an urban intersection. The basis of their proposed method is that the trajectories of vehicles follow specific routes, hence vehicle trajectories can be classified based on preset motion patterns without the necessity to track vehicles through the entire intersection, making it computationally efficient. Shirazi et. al. [25] developed a video-based trajectory prediction identify the turning movements of vehicles at urban intersections using deep neural networks trained on labeled data for turning movements and achieved a prediction accuracy of 92%. License-plate recognition has been utilized as the main feature of tracking instead of the overall vehicular features [26], [27]. While others utilize the imagery from Unmanned Aerial Vehicles (UAVs) to avoid the problem of occlusion in congested traffic conditions [28]. Multiple benchmark datasets have recently become available to aid in assessing the different models and techniques being developed under this unprecedented growth in this field of study [29], [30], [31]. Issues regarding speed and accuracy of tracking still arise, especially in heavy urban traffic conditions where occlusion is a noteworthy problem. This has led to many researchers exploring applications of obtaining trajectory data from GPS data from mobile phones [32], aerial video data from unmanned aerial vehicles [28], and estimation from traffic simulation that incorporates volume and signal timing knowledge

[33].

1.3.2 The Use of Object Detection and Tracking in Vehicle Speed Estimation

Researchers have utilized different techniques using onboard equipment, optical flow, image scaling, and deep learning methods, among others. Sreedivi and Gupta [34] used optical flow and feature point tracking in an image to estimate the velocity of the vehicle using MATLAB and Simulink. A method utilizing optical flow was also proposed and assessed by Dogan et. al. [35]. Osamura et. al. [36] proposed a method for feature point tracking from video data utilizing an onboard drive recorder. Costa et. al. [37] proposed a method utilizing image scale factoring to calculate the velocities of the vehicles traveling in the longitudinal direction compared to the axis of the camera. While the proposed approach results in accurate speed estimates, it is difficult to automate since it requires knowledge of the actual width of the vehicles being tracked. Garg and Goel [38] proposed a method utilizing license plate recognition. Their method resulted in speed estimates within ±8 km/hr of the actual vehicle speeds. Luvizon et. al. [39] proposed a method based on license plate tracking and were able to estimate speeds with an average error of -0.5 km/hr.

The recent improvements in vehicle and tracking resulted in major strides in vehicle speed estimation. Dong et. al. [40] utilized a method based on 3-dimensional convolutional networks to estimate the average vehicle speed based on information from video footage. The proposed method utilized non-local blocks, optical flow, and a multi-scale convolutional network. The results show that the proposed method results in reliable speed estimation with a mean absolute error of 2.71 km/h. Huang [41] used perspective transformation to estimate vehicle speeds. While using assumptions for the pixels to field measurement conversion, the study by Huang shows that perspective transformation can result in accurate speed

estimation with reduced computational complications. The study also highlights that speed estimation in urban settings is far more challenging than uninterrupted freeway flow.

Accurate and efficient speed estimation remains a challenging task, especially from low-altitude roadside cameras where methods utilizing license plate tracking and deep neural nets outperform those based on vehicle tracking due to detection instability and frequent identity switches. Hence, there is a high potential for speed estimation with reliable accuracy should the critical issue of vehicle identity switches due to occlusion be resolved, eliminating the need for additional complex deep neural nets or the privacy concerns associated with license plate tracking.

1.3.3 Trajectory-Based Safety Assessment

Recent studies and applications utilizing combinations of deep learning and computer vision techniques for vehicle trajectory tracking and obtaining real-time turn counts show very promising results and the ability of at-scale deployability [23], [42]. This growing ability to extract and analyze high-quality trajectory data enabled researchers to move from data mining car crash records to develop predictive crash likelihood models based on historical data to a more proactive approach of real-time assessment of traffic safety surrogate measures, leading to a large number of efforts that have studied and assessed traffic safety from video data under different roadway and infrastructure types [43], [44]. While similar safety-surrogate assessments were made using loop detector data [45], the relatively lower cost and non-disruptive nature of obtaining video data proved more viable option. Those studies assessed different safety surrogate measures, but the time to collision is the most commonly used measure. St-Aubin et. al. have proposed a video inference based safety surrogate framework and thoroughly assessed it in multiple studies [46], [47], [48]. Xie et. al. [49],

[50] developed a framework that incorporates foreground and background separation, non-parametric clustering, and Hidden Markov Models to comprehensively assess traffic safety from video data by analyzing conflict risk and TTC on an hourly basis, and found a significant correlation between the actual number of crashes and the traffic conflicts inferred through their proposed framework. Das and Maurya [51] utilized trajectory data extracted from an urban traffic environment to assess traffic safety based on the interactions of TTC, roadway center line separation, and leader-follower vehicle type pair. Estimating TTC from video data is an ongoing research problem, recent studies have assessed TTC estimation from on-board [52] and UAV [53] cameras.

Recent studies both in the fields of traffic safety and driving behavior modeling show the promising prospects of utilizing trajectory data obtained from video inference and conducting TTC-based safety assessments. A recent review study by Li et. al. [54] concluded that there is a need and immense potential in trajectory-based driver behavior modeling. Therefore, there is an immense need for improved vehicle tracking to accurately and efficiently extract high-resolution vehicle trajectories that would aid in these tasks.

1.3.4 Trajectory-Based Driver Behavior Modeling

Car-following and driver behavior modeling has long been one of the most well-studied topics in traffic engineering. Brackstone and McDonald [55] provided a systemic reexamination of the then-existing car-following models, including the Gazis-Herman-Rothery model, collision avoidance, linear, fuzzy logic-based and psycho-physical models. They concluded that albeit the extensive study of car-following models, and the strong support of those models in terms of conceptual bases and empirical data, the lack of time-series following behavior is a significant limitation to those models. The emergence of naturalistic driving and trajectory

datasets and algorithms [56] alongside advancements in traffic simulation modeling provided a boost to the state-of-the-research allowing researchers to assess driving behavior and car-following under different circumstances and in various roadway geometry settings [57], [58], [59], [60]. This availability of data has even enabled researchers to move from the long-standing models and develop probabilistic driver behavior models utilizing modern methods including neural and Bayesian networks [61], [62], [63].

One of the most widely adopted driver behavior models is the Optimal Velocity Model proposed by Bando et. al. [64]. The base model developed by Bando et. al. was very successful in describing traffic and capturing behavior during congestion, inspiring the development of OVM-derived models to help capture further aspects of traffic flow more realistically. Wang et. al. [65] proposed an OVM variation with a modified optimal velocity function. Mammar et. al. [66] proposed a modified OVM by utilizing the inverse of TTC as a weighing factor at given distance gap increments to better capture the breaking state of following vehicles. The proposed model was only tested it in simulation, finding that the simulated drivers' risk perception of rear-end collision can be significantly improved. Lazar et. al. expanded on this work [67] and also provided an exhaustive review of some of those models [68] concluding that due to the high variability in driving behavior, each model performs well under certain circumstances and has evident weaknesses in other circumstances. There is, therefore, a need and an immense value in developing data-driven models that would take into account the observed driver behavior heterogeneity [69], and in the case of urban intersections, factors such as turn movement classification and signal timing, which all impact the drivers' decision-making processes during car-following episodes. The recent advancements in the fields of vehicle tracking and the growing availability of trajectory data and trajectory-extraction methods make the creation of such data-driven models achievable and necessary.

1.3.5 Driver Behavior Modeling in Microsimulation

Over the past two decades, the growing adoption of microscopic simulation software has provided researchers with powerful means for evaluating different components of transportation networks and their interactions. The simulated vehicle interactions, which form the basis of microscopic driving behavior models, allowed for the assessment of traffic safety and evaluation of countermeasures as opposed to solely relying on analyzing crash data from police reports. The Surrogate Safety Assessment Model (SSAM) [70] is the prime example of such efforts. Using vehicle trajectories exported from microsimulation as an input, SSAM's approach utilizes conflict analysis for near misses, in terms of TTC and post encroachment time (PET), inferred from as a surrogate measure of actual crash data in terms of frequency, types, and severity of crashes. SSAM, however, is dependent on the outputs generated by microsimulation. Hence, adequate parameter calibration for the underlying driver behavior model used to generate the vehicle trajectories is crucial step that the lack thereof was found to result in a significant deviation of simulated safety conflicts from the observed behavior [71], [72], [73].

Given the increasing popularity of VISSIM over the years, numerous studies have been conducted to assess the sensitivity of calibrating its underlying Wiedemann model to better mimic the existing traffic flow and behavior and result in more realistic simulated safety conflicts. Fellendorf and Vortisch [74] conducted one of the earliest studies, where they assessed the impact of the driver model calibration on both a micro and macroscopic level. The results of their study showcased that a well-calibrated model can replicate traffic flow with high accuracy, with the caveat that driver behavior models should at least take national and regional regulations and driving styles into account to produce reliable representations in simulation. Lownes and Machemehl [75, 76] conducted a sensitivity analysis to assess the impact of modifying VISSIM's driver behavior parameters on simulated traffic flow capacity.

Their work has highlighted the impact of changes on driver behavior models at a microscopic level on traffic flow at a macroscopic level, further making the case for the necessity of parameter calibration to ensure adequate representation of traffic modeled in simulation. Other studies have also highlighted the necessity of taking into account the varying traffic patterns and considerations that need to be taken for calibrating VISSIM's driver behavior parameters accordingly, such cases include modeling traffic at urban intersections [77], [78] and freeway merge areas [79]. The work of Fan et. al. [79] particularly stands out in this category, as VISSIM's default driver behavior parameters were calibrated for freeway driving, yet parameter calibration for merging sections specifically was necessary, with the results highlighting the significant improvements obtained. This was attributed to the default VISSIM model parameters being calibrated from driving data from the German autobahn, while the study was conducted in China, hence the driving behavior parameters were not expected to be transferable.

The transferability of the calibrated VISSIM driver behavior model parameters was assessed by Essa and Sayed [80], who calibrated the VISSIM model parameters to maximize the correlation between observed and simulated traffic conflicts in one intersection and tested the calibrated model on a nearby intersection. Their study concluded that while not all of the calibrated model parameters are fully-transferable, the calibrated still significantly outperforms the default model in terms of providing better correlation between the observed and simulated traffic conflicts. Hence, any safety assessment without proper model calibration should be avoided. Essa and Tarek further highlighted this need for rigorous calibration and the limitations of simulation-based safety assessment in other studies [81], [82].

The extensive literature in VISSIM, and microscopic model parameter calibration in general, lacks a clear and transferable framework to overcome the shortcomings of site-by-site data collection and calibration of driver model parameters to reduce errors versus observed data.

In recent years, researchers have opted to utilize neural network models to overpass this tedious process [83], [84]. And while such methods succeed in producing improved results, they lack interpretability, unlike the traditional driving behavior models which are based on traffic flow theory fundamentals. There exists, therefore, a need for a generalizable driving behavior modeling approach for simulation to better mimic the vehicle trajectories being simulated and their resulting safety conflicts, allowing for a more accurate assessment of proposed safety interventions.

1.4 Dissertation Objectives and Contributions

The aim of this dissertation is to aid in advancing the state-of-the-art of traffic safety assessment through video-based vehicle trajectory inference, trajectory-based safety assessment, and driver behavior modeling. To address the gaps identified in the literature review, the following contributions have been made:

- Developing an ad-hoc framework to improve traffic detection and tracking accuracy and efficiency by utilizing the domain knowledge of traffic flow theory and assigned traffic movement patterns in urban intersections.
- Utilizing the ability to extract highly reliable vehicle trajectories and speed estimates for automated, large-scale calibration of safety-based driver behavior modeling. Speed estimation was validated using high-granularity data from an instrumented vehicle with state-of-the-industry sensors.
- Incorporating safety-based, high-fidelity, real-time driving behavior modeling in VISSIM modeling to provide substantial improvements in assessment of safety interventions via microsimulation.

Chapter 2

VT-Lane: An Exploratory Study of an Ad-hoc Framework for Real-time Intersection Turn Count and Trajectory Reconstruction Using NEMA Phases-Based Virtual Traffic Lanes

Abstract[1]**:** In this study, we propose an ad-hoc framework for real-time turn count and trajectory reconstruction for vehicles passing through urban intersections. Our proposed framework utilizes virtual lanes representing the 8 standard National Electrical Manufacturers Association (NEMA) movements within an intersection. A Python Graphical User Interface (GUI) was developed is utilized to identify entry planes for each NEMA movement, obtaining the trajectories and counts for the vehicles that are detected and identified at those planes which are then used as identifiers for other vehicles detected inside the intersection using a nearest neighbors search algorithm. Our proposed framework runs as an additional layer to any multi-object tracker with minimal additional computation, and the results of this preliminary assessment indicate the high potential that this proposed framework has in obtaining reliable turn count and in mitigating identity switches by resolving the vehicle

[1]Abdelhalim, Awad, and Montasir Abbas. "Vt-lane: An exploratory study of an ad-hoc framework for real-time intersection turn count and trajectory reconstruction using nema phases-based virtual traffic lanes." IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC). IEEE, 2020.

re-identification that occurs within the intersection due to detection errors and occlusion, resulting in more accurate vehicle trajectories from video data which will aid developing more reliable intersection safety surrogate measures.

Keywords: Vehicle tracking, intersection turn count, intersection safety.

2.1 INTRODUCTION

Traffic crashes are one of the leading causes of death in the world, averaging over 3,000 deaths per day according to the World Health Organization [1]. In the United States alone, there are around 40,000 traffic fatalities annually, an equivalent of a mid-range jet airliner falling off the sky every day! Approximately, 21.5% of all traffic fatalities occur due to intersection-related crashes, which represent 40% of all reported crashes in the United States [2].

Intelligent Transportation Systems (ITS) aim to address transportation safety issues through transforming the typical passive and unaware transport systems to dynamic, proactive systems. Vehicle recognition and trajectory tracking plays an integral role in many aspects of ITS applications; from behavioral modeling and car-following analyses to congestion prevention, crash warning, dynamic signal timing, and active traffic management.Vehicle trajectory analysis also play a crucial role in the applications of path planning and co-operative driving.

The ever-growing use of cameras to monitor traffic and for overall surveillance has also resulted in video data for traffic being more available, inexpensive and far less disruptive to obtain compared to loop detector data, which inherently leads to more focus on developing better and more reliable video detection and tracking techniques. The detection and tracking of trajectories of urban traffic, especially through busy intersections, remains a very challenging problem due to multitude of complications, including privacy concerns regarding

the tracked vehicles, and technical barriers related to the quality and area of view provided by on-site cameras, occlusion, and computational requirements.

The typical traffic movement assignment for signal controllers in the United States follows the National Electrical Manufacturers Association (NEMA) phases shown in Fig. 2.1 [3]. In this study, we propose to incorporate this prior knowledge of NEMA traffic movement patterns in the video-based trajectory tracking framework to obtain reliable turn counts, assist in addressing the vehicle re-identification problem due to occlusion, and aid in reconstructing missing vehicle trajectories.

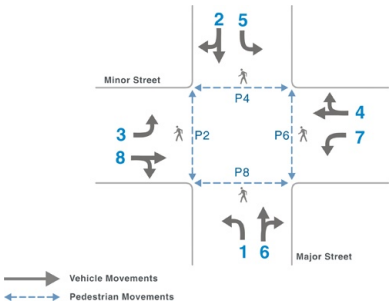


Figure 2.1: Typical 8-phase traffic controller operation.

2.1.1 Evolution of Object Detection and Tracking

The computer vision task of detecting objects from images and video frames can be subdivided into three major categories:

- Image Classification: The classical problem of identifying the main object in an image.
- Object Recognition: The task of identifying and localizing an object within an image.
- Multi-Object Recognition: Identifying and localizing multiple objects in the same image frame.

The three aforementioned categories of object recognition problems have been classically tackled by a myriad of computer-vision techniques and manipulations. As the field evolved, knowledge about the geometry, scenery, and the deformable part models [4], [5], [6] were utilized. But the massive recent strides have been accomplished due to the evolution of Convolutional Neural Networks (CNNs) which now form the basis of all state-of-the-art object detection and tracking algorithms. As CNNs proved extremely dependable in object recognition tasks, a major breakthough was achieved by Ross Girshic proposing the Fast Region-based Convolutional Neural Network (Fast R-CNN) [7], followed by the faster, closer to real-time version Faster R-CNN in 2015 [8] which then had a frame rate of 5 fps. Low frame rates limited the number of real-time applications of such models, until the J Redmon et. al. and W. Liu et. al. respectively introduced their You Only Look Once (YOLO) and Single-Shot Detector (SSD) models. [9], [10], achieving performance suitable for real-time detection and tracking applications. Object detection models exponentially improved over the past few years, with newer versions of the aforementioned works massively outperforming their predecessors [11], [12], [13].

2.1.2 Object Detection and Tracking in Traffic Analysis and Safety Applications

In the context of traffic engineering, the research focus in terms of object detection and tracking has classically been focused only on automobiles. It is only recently that researchers began focusing on tracking pedestrians and other roadway users. As early as the early 90s, a multitude of studies were conducted assessing various techniques for vehicle tracking and counting from video data, utlilizing belief networks, adaptive background memory and parametrized 3-D models [14], [15], [16]. The majority of the earlier methods were inapplicable during congestion due to the occlusion problem. In 1996, Beymer and Malik [17] proposed a feature-

based model that tracks sub-features of vehicles instead of the entire automobile, making it less sensitive to partial occlusion. Beymer et. al. then extended their work and tested on several hours of video [18]. Feature-based tracking, however, was extremely computationally expensive, and Beymer's proposed feature-based would achieve a maximum performance of 2 Hz under congested traffic conditions. Kamijo et. al. [19] proposed a spatio-temporal Markov random field algorithm for traffic monitoring at intersections. Their proposed model demonstrated very high accuracy in tracking vehicles, and they also introduced an incident detection method, albeit their data being insufficient to comprehensively evaluate. Several studies in the following years provided massive contributions to the applications of vehicle tracking and trajectory analysis [20], [21], [22].

The aforementioned advancements in deep learning, rapid development of more dependable models, and rise of GPU computing have managed to tackle the majority of issues related to traffic detection and monitoring in terms of real-time performance and deployability. Recent research has been more focused on reliable means to extracting and predicting vehicle trajectories in urban environments. Dai et. al. proposed a framework to track vehicles and obtain traffic counts based on intersection segmentation on the basis of entry approach [23]. Wu et. al [24] proposed a Hidden Markov Model to identify the vehicle trajectories within an urban intersection. The basis of their proposed method is that the trajectories of vehicles follow specific routes, hence vehicle trajectories can be classified based on preset motion patterns without the necessity to track vehicles through the entire intersection, making it computationally efficient. Shirazi et. al. [25] developed a video-based trajectory prediction identify the turning movements of vehicles at urban intersections using deep neural networks trained on labeled data for turning movements and achieved a prediction accuracy of 92%. License-plate recognition has been utilized as the main feature of tracking instead of the overall vehicular features [26], [27]. While others utilize the imagery from Unmanned Aerial Vehi-

cles (UAVs) to avoid the problem of occlusion in congested traffic conditions [28]. Multiple benchmark datasets have recently become available to aid in assessing the different models and techniques being developed under this unprecedented growth in this field of study [29], [30].

2.2 METHODOLOGY

2.2.1 Base Object Tracker

For our base object detection and tracking task, we implemented a combination of YOLO v3 [13, 31] and Wojke et. al. Simple Online and Real-time Tracking (SORT) algorithm, which utilizes a combination of correlation matrices, Kalman Filer, and data association based on the Hungarian method [32]. Detailed information about those algorithms are available in the respective original works. The object detector and tracker are pre-trained on the Microsoft Common Objects in Context (COCO) dataset. This combination for vehicle tracking runs at ∼13 fps on NVIDIA GTX 1060 GPU. As most traffic infrastructure refreshes at 10 HZ, this output frame is acceptable, but significant improvements can be achieved with better GPUs. While better baseline tracking performance is expected when training the detector-tracker combination on datasets that are tailored for traffic detection instead of the general MS COCO, we wanted to assess the improvement that can be achieved by solely incorporating our knowledge of predefined traffic routes through intersections and narrowing the search space in that context rather than training a network specifically for the task.

2.2.2 NEMA Phase-Based Virtual Traffic Lanes

A Graphical User Interface (GUI) was developed in Python to allow for the selection of points that represent the edges of the entry planes to each NEMA movement as vehicle enter the intersection. Those planes were defined past the stop bars for each movement, with an additional margin of 4.5 meters (average length of a passenger car), creating a an entry polygon for each NEMA movement. This can also be done in the transformed perspective space for a subset of the trajectories, where the entry polygons defined through the GUI essentially add labels to all vehicles that clearly follow a specific NEMA movement before entering the region of interest (ROI). The trajectories of vehicles identified at entry planes follow specific patterns across the intersection creating virtual lanes within the intersection. Hence, those vehicles are referenced to as *Virtual Lane Vehicles (VLV)*. To account for occlusion, the vehicle detection and tracking is carried out in the original camera space, while trajectories are processed in the transformed perspective space. The perspective transformation was obtained by carrying out a standard checkerboard camera calibration, defining a ROI, calculating the homography transformation matrix (M) while also accounting for lens distortion. The destination x and y coordinates $destination(x, y)$ for all the phase entry polygons and the trajectory points of each vehicle are transformed using the following equation:

$$source \left(\frac{M_{11}x + M_{12}y + M_{13}}{M_{31}x + M_{32}y + M_{33}}, \frac{M_{21}x + M_{22}y + M_{23}}{M_{31}x + M_{32}y + M_{33}} \right) \tag{2.1}$$

Where:

x & y = are the coordinates for each point in the source space from the camera video frame.

$M = 3 \times 3$ homography matrix.

$$NEMA_i = \begin{cases} Phase\ j, & \text{if } x_{1i}, y_{1i} \in p_j\ \forall_{i,j} \\ unknown, & \text{otherwise} \end{cases} \tag{2.2}$$

Where:

$NEMA_i$ = NEMA phase number for vehicle i.

x_{1i} & y_{1i} = are the coordinates for the first point in the trajectory of each detected vehicle i.

$Phase\ j$ = Standard number convention for NEMA phase movement $\{2, 4, 6, 8, 1, 3, 5, 7\}$.

p_j= Polygon defining the entry plane for NEMA phase j.

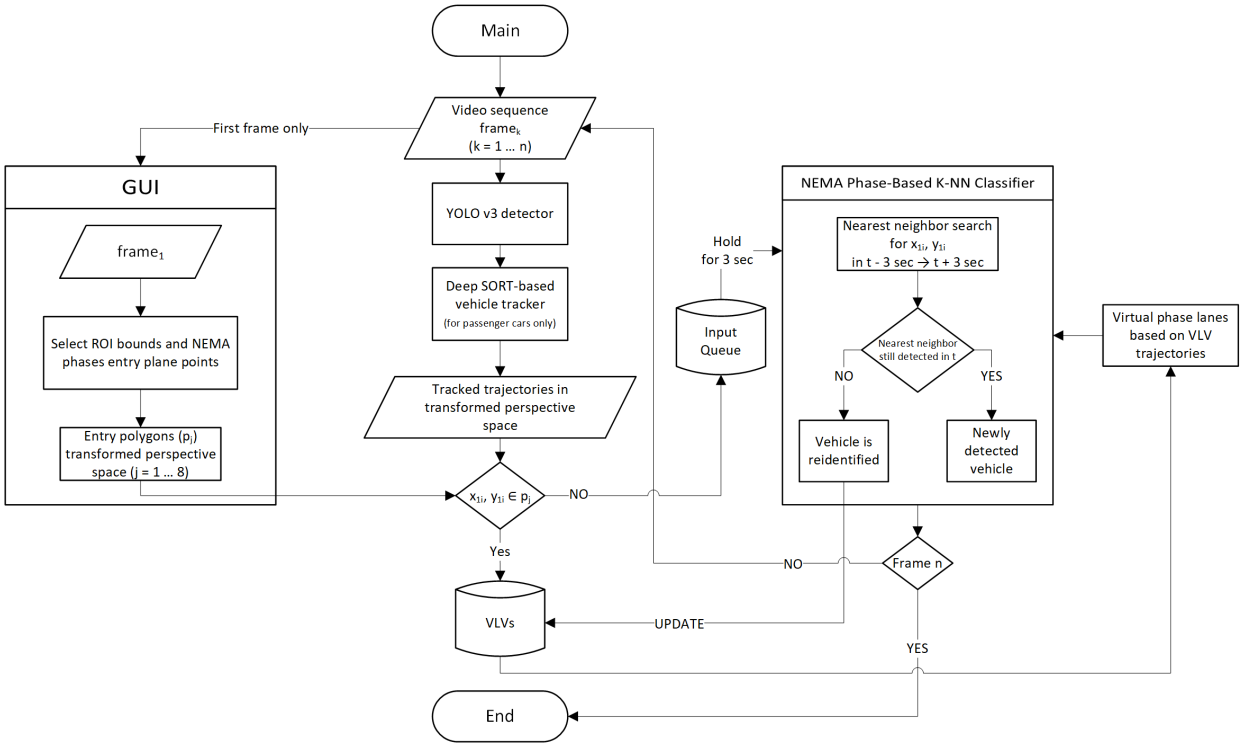


Figure 2.2: Flow chart of the proposed framework's data flow.

Fig. 2.2 shows and overview of the flow of data in the framework. The vehicle trajectories written to the external databases are only those within the ROI. This is to avoid storing trajectories for vehicles before they enter the intersection, hence reducing the amount of

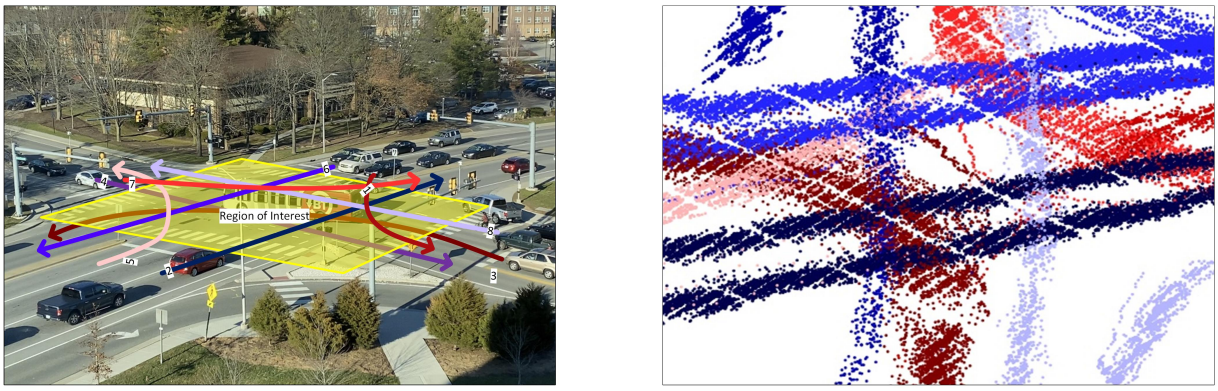


Figure 2.3: Intersection of study (left) and the perspective-transformed virtual lanes from trajectories of sample VLV vehicles (right).

data stored for each video frame and making the K-NN search more efficient. While using entry planes to identify the NEMA movements for the vehicles that were detected entering the intersection from those planes provides a very efficient way of obtaining the expected trajectory and movement patterns for vehicles in an urban intersection, it is insensitive of vehicles that conduct illegal movements within the intersection (e.g. a vehicle that goes straight through the intersection albeit entering from a left-turn pocket lane). While such movements are not very common, however, the use of the aforementioned entry planes to identify vehicle movements does not take those anomalies into account.

Vehicles detected at the entry planes are assigned to their respective NEMA phases and are then used to classify the vehicles that are identified within the intersection. The late detection within the intersection can occur simply due to the late identification of a car by the detector, but may also occur due to identity switches after occlusion, which may happen due to heavy traffic, shadows, or the infrastructure (all can be observed in Fig. 2.3). The NEMA movements and region of interest for our intersection of study are also shown.

The K-Nearest Neighbor (K-NN) search for the vehicles whose movements were not identified at the entry planes is described in Algorithm 1. A 3-second delay is allowed before classifying

unidentified vehicles, this is to account for the vehicles that enter the ROI at the beginning of a phase where the vehicles enter the intersection at saturation flow rate, which a study at the same intersection found to be $\sim$ 2,100 vehicles per hour (1.7 veh/sec) [33]. Such that, the K-NN search space is the VLV trajectories between $frame_{k-90} \rightarrow frame_{k+90}$, where $frame_k$ is the first detection frame of a vehicle with unidentified movement. All video sequences used for this study were recorded from the same location, using the same camera, with a resolution of 1280 $\times$ 720 pixels at 30 fps, and a search back time $t-1$ second allows to identify occlusion (which we observed to average 0.6s in our videos). The majority vote of the minimum of 3 VLVs within the given temporal range is used to identify the NEMA phase. An identity switch is identified if the nearest neighbor of vehicle$_i$ was present in $frame_{k-30}$ and is no longer detected on or after $frame_k$.

Algorithm 1: NEMA Phase-Based K-NN Classifier

Result: Identifying the NEMA phase of vehicles that were unidentified at the entry planes of the intersection.

initialization;
for $frame_k$ do
  for $x_{1i}, y_{1i} \notin P$ do
    neighbors = VLVs $\in \{frame_{k-90} ... frame_{k+90}\}$;
    $NEMA_i = majority\ vote$;
    if nearest $\in frame_{k-30}$ & $\notin \{frame_k ... frame_{k+30}\}$ then
      $ID_i = ID_{nearest} \iff NEMA_i = NEMA_{nearest}$;
    else
      $ID_i = ID_{detection}$;
    end
  end
end

2.2.3 Area of Study and Data Collection

Video data was collected for the intersection of Prices Fork and Tom's Creek Roads in Blacksburg, Virginia, USA shown in Fig. 2.3. Over an hour of video data was recorded

in 10 6-minute video sequences and turning counts through the intersection were collected for those video sequences. The video sequences and traffic counts were taken in different days during the AM and PM peak hours to account for imbalanced directional volumes and changing lighting conditions. The turning counts taken serve as the ground truth for the proposed framework. The metrics used to evaluate the results are the NEMA movement count accuracy and the precision. A 2^{nd} degree polynomial defining the trajectory of each of the 8 nema movements was calibrated from the VLV trajectories and then used to identify the false positives when a vehicle's trajectory doesn't follow the expected pattern for its respective NEMA movement.

$$Count\ Accuracy_j = \frac{TP_j}{GT_j} \tag{2.3}$$

$$Precision_j = \frac{TP_j}{TP_j + FP_j} \tag{2.4}$$

Where TP, FP, GT are respectively the true positive, false positive, and ground truth counts of each NEMA movement.

2.3 RESULTS

Table 2.1 shows the manual movement counts obtained for the video sequences. The video sequences that make a total of 15 off-peak minutes, 15 and 30 minutes respectively during AM and PM hours. A Total of 2,142 vehicles were counted and their corresponding NEMA phase movements were recorded.

Table 2.2 shows the results obtained through our video tracking framework, showing the TP

detected vehicle counts, the numbers of detected identity switches (IDS; difference between ground truth and number of detected vehicles), resolved identity switches (Re_IDs), and total number of vehicles detected whose movement remained unidentified (UIV). The table also shows the weighted average accuracy and precision for each movement.

Table 2.1: Total turn counts for video sequences taken at Off peak, AM and PM peak hours

Phase	Off Peak	AM Peak	PM Peak	Total
2	99	151	438	671
5	7	29	69	105
4	55	45	93	193
7	12	30	53	95
6	123	194	421	738
1	18	51	31	100
8	7	11	52	71
3	8	24	137	169
Total	329	535	1294	2142

Video sequences taken at off-peak time had low traffic, and the sun position during the mid-day period results in less disruptive shadows and no lens glare, unlike AM and PM peaks, resulting in very high detection accuracy and very few identity switch cases. Occlusion due to high traffic at peak hours, alongside infrastructure (signal poles & traffic signs) results in a high number of ID switches, a good percentage of which were resolved by our framework. Isolating the vehicle detections by movement is very beneficial in identifying the movements where the detector is under-performing, which was evident in some turning movements and would otherwise be masked by the overall accuracy due to the higher volumes of through vehicles. This proposed framework shows the substantial improvement on a simple base-line YOLO model by narrowing down the search space based on our knowledge of traffic movement through an intersection, the overall accuracy and accuracy of identifying movements and resolving ID switches is very high.

Table 2.2: Detected vehicle movements and evaluation

Phase	Off Peak	AM Peak	PM Peak	$\overline{Acc}$	$\overline{Prc}$
2	98	141	415	95.71	99.67
5	7	26	62	93.17	93.54
4	52	43	71	88.82	98.02
7	11	30	49	94.71	99.33
6	122	168	385	92.41	99.27
1	16	23	30	76.92	95.06
8	7	11	40	92.31	99.19
3	7	21	129	89.72	97.92
$\overline{Acc}$	95.1	87.22	89.09	90.21	
IDS	7	44	90		
Re_IDs	2	33	62		
Re_IDs %	28%	75%	69%		
UIV	7	80	122		
UIV %	2.02%	13.82%	8.82%		

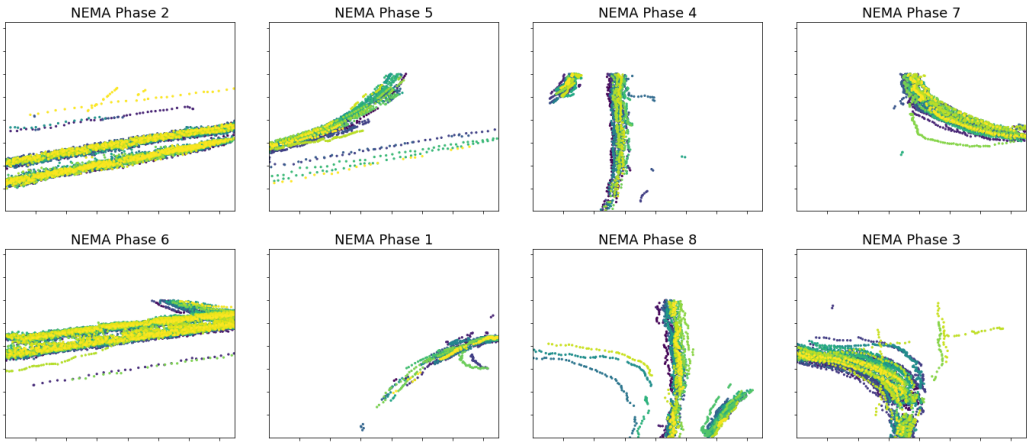


Figure 2.4: Tracked NEMA movements for PM peak video sequences.

Figure 2.4 shows the detected movements for each NEMA phase during the PM peak video sequences. The low number of misclassifications illustrates the very high precision results achieved. Overall, 9.1% of all vehicles remained with unidentified NEMA movement at the end of the process. The percentage of unidentified vehicles at the AM peak is significantly

higher due to the aforementioned traffic and lens glare problems. There is room for improvement on this aspect by relatively compromising precision. Fig. 2.5 shows the trajectories of the unidentified vehicles in PM peak hour video sequences, where the majority of trajectories seem to have an initial detection inside the intersection. This may be due to identity switches or late detection due to occlusion in heavy traffic.

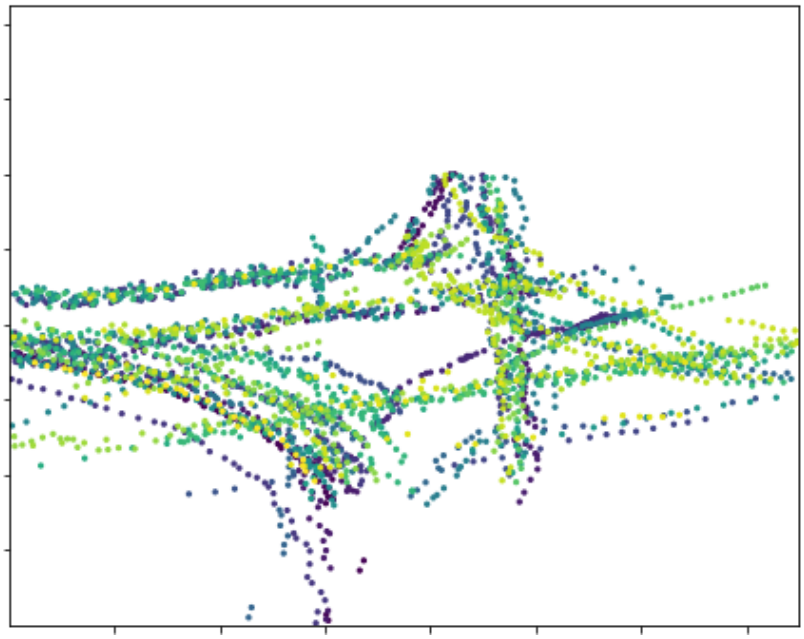


Figure 2.5: Unidentified vehicles during PM peak video sequences.

Fig. 2.6 shows an example of a resolved identity switch, where vehicle 2123 was re-identified as 2137 after occlusion by a transit vehicle. The output trajectory of vehicle 2123 via our framework includes that of 2137, and the missing trajectory during occluded frames can be easily reconstructed using the 2^{nd} degree polynomial representing the expected movement within the virtual lane for the NEMA movement. A total of 97 out of 141 identity switches that occurred in the 1 hour video sequences were identified and corrected.

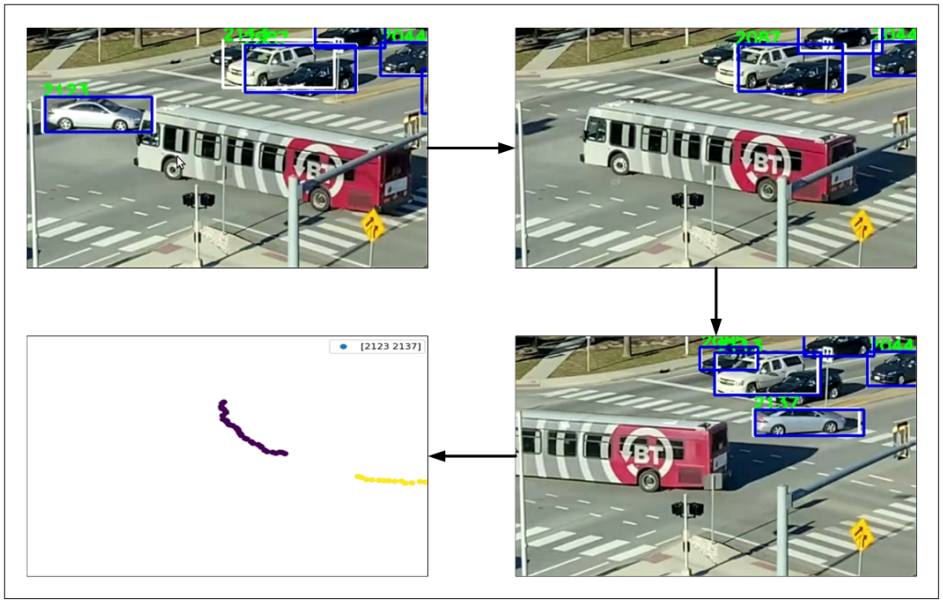


Figure 2.6: Example of resolved ID switch.

2.4 CVPR AI CITY CHALLENGE

Our framework was also assessed and validated by our participation in the 4th AI City Callenge in CVPR 2020 [2]

The dataset provided for this challenge contained 31 video sequences totaling over 9 hours (4 of which were reserved for testing) from 20 different camera locations. Specific movements and ROIs were predefined for each camera location. Figure 2.7 shows a sample video and the processing pipeline in our framework, where the given ROIs are used to define the detection and tracking area, then the user manually defines polygons corresponding to each movement of interest, after which virtual movement lanes obtained via perspective transformation. In the post-processing of detected trajectories for intersection scenes, we only include the trajectories within the intersection (past the stop lines) to avoid storing data for stationary

[2]Abdelhalim, Awad, and Montasir Abbas. "Towards real-time traffic movement count and trajectory reconstruction using virtual traffic lanes." Proceedings of the IEEE CVF Conference on Computer Vision and Pattern Recognition Workshops. 2020.

vehicles awaiting green light, hence all trajectory data starts past the stop lines where the entry polygons are defined. We initially evaluated our the accuracy of our framework on this given subset by evaluating the count accuracy for each of the 12 movement labels in that given intersection before submitting results of the entire dataset to the evaluation server.

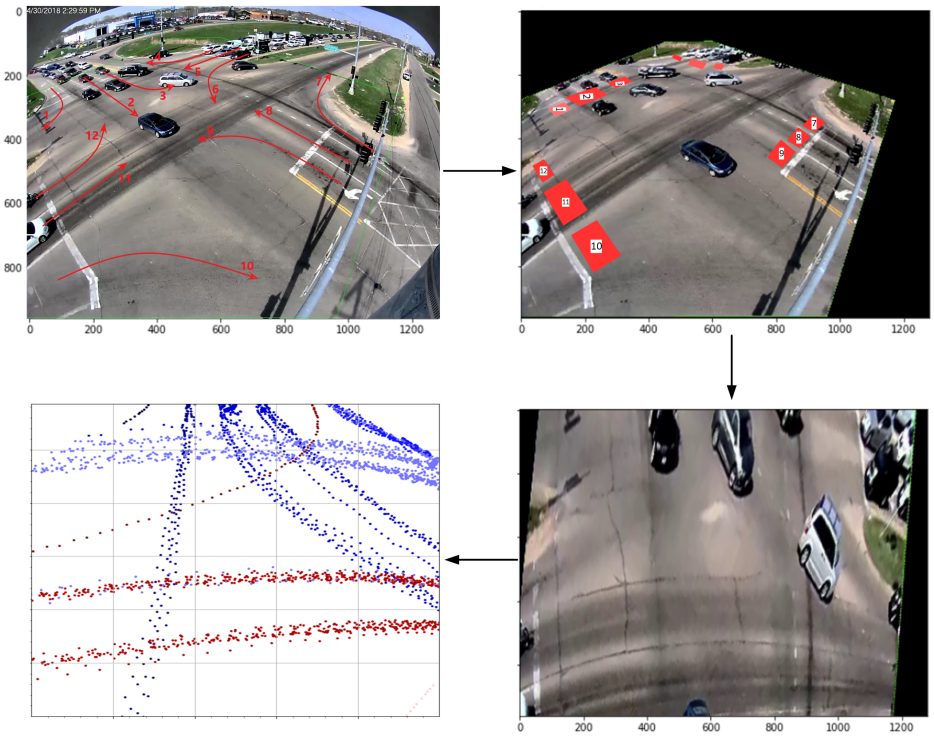


Figure 2.7: Sample of video data in AI City Challenge Dataset.

The conglomerated results from all 31 video sequences were then uploaded to the challenge's evaluation server, which evaluates the overall results on effectiveness and the computational efficiency of the algorithm. The final $S1$ score is a combination of both.

$$S1 = \alpha S1_{efficiency} + \beta S1_{effectiveness} \tag{2.5}$$

Where: $\alpha = 0.3$ $\beta = 0.7$

$$S1_{efficiency} = \max\left(0, 1 - \frac{time \times base\ factor}{5 \times video\ total\ time}\right) \tag{2.6}$$

Where:

time and *video total time* = respectively the execution time and total video time for the given dataset in *sec*.

base factor = score obtained through a Python script provided with the dataset.

$$wRMSE = \sqrt{\sum_{i=1}^{k} w_i(\hat{x}_i - x_i)^2} \tag{2.7}$$

Where:

$$w_i = \frac{i}{\sum_{j=1}^{k}} = \frac{2i}{k(k+1)} \tag{2.8}$$

The $S1_{effectiveness}$ score is calculated based on the weighted average of normalized weighted root mean square error across all video sequences, movements, and vehicle classes.

Table 2.3 shows the movement counts given for the provided sample video sequence and the total and TP counts obtained for each movement through our framework. Table 2.4 shows the results from the evaluation system. We selected two runs where we used slightly different thresholds for re-identification of vehicles within a virtual lane. Run 1 assumes that an identity switch occurs *only* when the nearest neighbor of vehicle i is no longer detected in the following frames, where as in run 2 it could be anyone of the k-nearest neighbors.

Although our code currently distinguishes between car and truck classes, the version of the code that was used before the closure of the evaluation server did not. Hence, these evaluations were affected by the misclassification of all trucks in the dataset. Out of 97 teams that requested the dataset, only 23 teams were able to complete the track challenge

and submit results by the competition's deadline and our submission ranked 8th in the leader board at the time of submission.

Table 2.3: The Ground Truth, Obtained Total and True Positive Counts for All Movements in the Provided Sample Video Sequence

Movement	Ground Truth	Total Count	TP
1	1	1	1
2	5	4	4
3	13	14	12
4	10	8	8
5	25	32	25
6	0	0	0
7	1	1	1
8	0	0	0
9	0	0	0
10	2	0	0
11	38	36	36
12	1	1	1

Table 2.4: Select Results From the Evaluation Server

Run	wRMSE	Effectiveness	Efficiency	S1
1	12.2	0.767	0.906	0.809
2	11.87	0.774	0.906	0.814

2.5 CONCLUSIONS

In this study, we explored the potential incorporating our prior knowledge of traffic movement patterns across an urban intersection to improve detection accuracy and obtain real-time turn counts for intersections. The results show very promising potential for obtaining reliable trajectories and speed profiles for each movement, which can aid in improving conflict monitoring, incident avoidance, dynamic signal timing, as well as O-D estimation applications. In this assessment's 1-hour of video sequences, 69% of the identity switches incurred

by the baseline tracker were resolved.

The K-NN search is very time-efficient, running up to 100 HZ in low traffic and between 60-70 HZ in high traffic (on a 2.4 GHz 16-core Intel Xeon e5 2630 v3 processor), a minimal increase to the running time of a multi-object tracking network. As no benchmark datasets for turn counts exist at the time of writing, fully labeling our collected data will allow to comprehensively compare our framework with other existing algorithms, and would allow to optimize the look back and look ahead time frames, number of neighbors to consider, and other framework parameters.

We also attempted the AI City Challenge Track 1: Vehicle Counts by Class at Multiple Intersections. While the movements of interest are predefined for camera locations in this challenge, in general the traffic patterns in a certain road segment or through an intersection do follow similar predefined patterns, and providing practitioners with a simple GUI to identify those movements of interest to obtain reliable traffic counts, vehicle trajectories and speed profiles can significantly aid in developing ITS application of origin-destination estimation, conflict monitoring, incident detection, and dynamic signal timing. The results obtained in the AI City Challenge demonstrated the flexibility, scalability, and efficiency of the proposed framework.

Bibliography

[1] W. H. Organization, "Road traffic deaths," 2015. https://www.who.int/gho/road_safety/mortality/en/.

[2] F. H. Administration, "The national intersection safety problem," tech. rep., 2009.

[3] U. FHWA, "Traffic signal timing manual," 2008.

[4] P. Felzenszwalb, D. McAllester, and D. Ramanan, "A discriminatively trained, multiscale, deformable part model," in *2008 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–8, IEEE, 2008.

[5] M. C. Burl, M. Weber, and P. Perona, "A probabilistic approach to object recognition using local photometry and global geometry," in *European conference on computer vision*, pp. 628–641, Springer, 1998.

[6] D. Crandall, P. Felzenszwalb, and D. Huttenlocher, "Spatial priors for part-based recognition using statistical models," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, vol. 1, pp. 10–17, IEEE, 2005.

[7] R. Girshick, "Fast r-cnn," in *Proceedings of the IEEE international conference on computer vision*, pp. 1440–1448, 2015.

[8] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," in *Advances in neural information processing systems*, pp. 91–99, 2015.

[9] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified,

real-time object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 779–788, 2016.

[10] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "Ssd: Single shot multibox detector," in *European conference on computer vision*, pp. 21–37, Springer, 2016.

[11] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask r-cnn," in *Proceedings of the IEEE international conference on computer vision*, pp. 2961–2969, 2017.

[12] C.-Y. Fu, W. Liu, A. Ranga, A. Tyagi, and A. C. Berg, "Dssd: Deconvolutional single shot detector," *arXiv preprint arXiv:1701.06659*, 2017.

[13] J. Redmon and A. Farhadi, "Yolov3: An incremental improvement," *arXiv preprint arXiv:1804.02767*, 2018.

[14] T. Huang, D. Koller, J. Malik, G. Ogasawara, B. S. Rao, S. J. Russell, and J. Weber, "Automatic symbolic traffic scene analysis using belief networks," in *AAAI*, vol. 94, pp. 966–972, 1994.

[15] K. Karmann and A. Brandt, "Moving object recognition using an adaptive background memory. cappellini v, ed. time-varying image processing and moving object recognition," 1990.

[16] D. Koller, K. Daniilidis, and H.-H. Nagel, "Model-based object tracking in monocular image sequences of road traffic scenes," *International Journal of Computer 11263on*, vol. 10, no. 3, pp. 257–281, 1993.

[17] D. Beymer and J. Malik, "Tracking vehicles in congested traffic," in *Proceedings of Conference on Intelligent Vehicles*, pp. 130–135, IEEE, 1996.

[18] D. Beymer, P. McLauchlan, B. Coifman, and J. Malik, "A real-time computer vision system for measuring traffic parameters," in *Proceedings of IEEE computer society conference on computer vision and pattern recognition*, pp. 495–501, IEEE, 1997.

[19] S. Kamijo, Y. Matsushita, K. Ikeuchi, and M. Sakauchi, "Traffic monitoring and accident detection at intersections," *IEEE transactions on Intelligent transportation systems*, vol. 1, no. 2, pp. 108–118, 2000.

[20] V. G. Kovvali, V. Alexiadis, P. Zhang, *et al.*, "Video-based vehicle trajectory data collection," tech. rep., 2007.

[21] N. Saunier and T. Sayed, "A feature-based tracking algorithm for vehicles in intersections," in *The 3rd Canadian Conference on Computer and Robot Vision (CRV'06)*, pp. 59–59, IEEE, 2006.

[22] A. Sanchez, P. D. Suarez, A. Conci, and E. Nunes, "Video-based distance traffic analysis: Application to vehicle tracking and counting," *Computing in Science & Engineering*, vol. 13, no. 3, pp. 38–45, 2010.

[23] Z. Dai, H. Song, X. Wang, Y. Fang, X. Yun, Z. Zhang, and H. Li, "Video-based vehicle counting framework," *IEEE Access*, vol. 7, pp. 64460–64470, 2019.

[24] C.-E. Wu, W.-Y. Yang, H.-C. Ting, and J.-S. Wang, "Traffic pattern modeling, trajectory classification and vehicle tracking within urban intersections," in *2017 International Smart Cities Conference (ISC2)*, pp. 1–6, IEEE, 2017.

[25] M. S. Shirazi and B. T. Morris, "Trajectory prediction of vehicles turning at intersections using deep neural networks," *Machine Vision and Applications*, vol. 30, no. 6, pp. 1097–1109, 2019.

[26] L. Zhu, S. Wang, C. Li, and Z. Yang, "License plate recognition in urban road based on vehicle tracking and result integration," *Journal of Intelligent Systems*, vol. 29, no. 1, pp. 1587–1597, 2019.

[27] T. K. Cheang, Y. S. Chong, and Y. H. Tay, "Segmentation-free vehicle license plate recognition using convnet-rnn," *arXiv preprint arXiv:1701.06439*, 2017.

[28] H. Zhang, M. Liptrott, N. Bessis, and J. Cheng, "Real-time traffic analysis using deep learning techniques and uav based video," in *2019 16th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, pp. 1–5, IEEE, 2019.

[29] Z. Tang, M. Naphade, M.-Y. Liu, X. Yang, S. Birchfield, S. Wang, R. Kumar, D. Anastasiu, and J.-N. Hwang, "Cityflow: A city-scale benchmark for multi-target multi-camera vehicle tracking and re-identification," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 8797–8806, 2019.

[30] L. Wen, D. Du, Z. Cai, Z. Lei, M.-C. Chang, H. Qi, J. Lim, M.-H. Yang, and S. Lyu, "Ua-detrac: A new benchmark and protocol for multi-object detection and tracking," *Computer Vision and Image Understanding*, p. 102907, 2020.

[31] J. Redmon, "Darknet: Open source neural networks in c." http://pjreddie.com/darknet/, 2013–2016.

[32] N. Wojke, A. Bewley, and D. Paulus, "Simple online and realtime tracking with a deep association metric," in *2017 IEEE international conference on image processing (ICIP)*, pp. 3645–3649, IEEE, 2017.

[33] Q. Wang and M. Abbas, "Optimal urban traffic model predictive control for nema standards," *Transportation Research Record*, vol. 2673, no. 7, pp. 413–424, 2019.

Chapter 3

A Framework for Real-time Traffic Trajectory Tracking, Speed Estimation, and Driver Behavior Calibration at Urban Intersections Using Virtual Traffic Lanes

Abstract[3]**:** In a previous study, we presented VT-Lane, a three-step framework for real-time vehicle detection, tracking, and turn movement classification at urban intersections. In this study, we present a case study incorporating the highly accurate trajectories and movement classification obtained via VT-Lane for the purpose of speed estimation and driver behavior calibration for traffic at urban intersections. First, we use a highly instrumented vehicle to verify the estimated speeds obtained from video inference. The results of the speed validation show that our method can estimate the average travel speed of detected vehicles in real-time with an error of 0.19 m/sec, which is equivalent to 2% of the average observed travel speeds in the intersection of study. Instantaneous speeds (at the resolution of 30 Hz) were found to be estimated with an average error of 0.21 m/sec and 0.86 m/sec respectively for free flowing and congested traffic conditions. We then use the estimated speeds to calibrate the parameters of a driver behavior model for the vehicles in the area of study. The results

[3]Abdelhalim A, Abbas M, Kotha B, and Wicks A. "A Framework for Real-time Traffic Trajectory Tracking, Speed Estimation, and Driver Behavior Calibration at Urban Intersections Using Virtual Traffic Lanes." IEEE 24th International Conference on Intelligent Transportation Systems (ITSC). IEEE, 2021.

show that the calibrated model replicates the driving behavior with an average error of 0.45 m/sec, indicating the high potential for using this framework for automated, large-scale calibration of car-following models from roadside traffic video data, which can lead to substantial improvements in traffic modeling via microscopic simulation.

Keywords: Vehicle tracking, intersection turn count, intersection safety.

3.1 INTRODUCTION

With over 3,000 daily fatalities, traffic crashes remain one of the major causes of death around the globe. In the United States, 40,000 lives are lost annually due to traffic crash-related fatalities, 21.5% of which are due to urban intersection-related crashes. Intersection-related crashes comprise about 40% of all the reported annual crashes in the USA [1].

Intelligent Transportation Systems (ITS) aim to address the safety issues in transportation by revolutionizing the transport systems from being passive to smart, proactive systems. Given the increasing use of traffic cameras for monitoring and enforcement in recent years, offline traffic video data and online traffic video streams are becoming more available to researchers and practitioners. The process of accurate and efficient traffic trajectory tracking remains a challenging task. The quality of data available from video streams of on-site cameras, the full and partial occlusion of vehicles in congested traffic, heavy computational demand, and driver privacy concerns all contribute to additional complications that demand further improvements for the state-of-the-art practices and methodologies.

3.1.1 Object Detection and Tracking for Traffic Safety

Recent years have witnessed gigantic advancements in the fields of object detection and tracking. The development and continuous improvement of Convolutional Neural Networks

(CNNs) played a major role in those advancements. CNNs form the basis of the vast majority of the state-of-the-art detection and tracking methodologies. Achieving real-time performance was made possible with the introduction of Fast R-CNN [2] and shortly after, You Only Look Once (YOLO) and Single-Shot Detector (SSD) [3], [4]. Rapid improvements in those models have followed in recent years [5], [6]. In the transportation field, researchers have tried tackling the problem of vehicle detection and tracking in traffic using a myriad of methodologies including the use of adaptive background memory [7] and scene analysis through belief networks [8]. Tracking vehicles in congested urban conditions has been a challenge for those early methods due to occlusion. Later models utilizing sub-feature tracking were able to better handle this problem [9], [10]. The advancements in GPU-accelerated computing and deep learning methodologies enabled researchers in the field of traffic engineering to develop models that detect, track, and extract vehicle trajectories in real-time [11], [12]. A growing number benchmark datasets is being made available to support the research efforts in the field [13], [14], [15].

3.1.2 The Use of Object Detection and Tracking in Vehicle Speed Estimation

Researchers have utilized different techniques using on-board equipment, optical flow, image scaling, and deep learning methods, among others. Sreedivi and Gupta [16] used optical flow and feature point tracking in an image to estimate the velocity of the vehicle using MATLAB and Simulink. A method utilizing optical flow was also proposed and assessed by Dogan et. al. [17]. Osamura et. al. [18] proposed a method for feature point tracking from video data utilizing an on-board drive recorder. Costa et. al. [19] proposed a method utilizing image scale factoring to calculate the velocities of the vehicles travelling in the longitudinal direction compared to the axis of the camera. While the proposed approach results in accurate speed estimates, it is difficult to automate since it requires knowledge of the actual

width of the vehicles being tracked. Garg and Goel [20] proposed a method utilizing license plate recognition. Their method resulted in speed estimates within ±8 km/hr of the actual vehicle speeds. Luvizon et. al. [21] proposed a method based on licence plate tracking and were able to estimate speeds with an average error of -0.5 km/hr.

The recent improvements in vehicle and tracking resulted in major strides in vehicle speed estimation. Dong et. al. [22] utilized a method based on 3-dimensional convolutional networks to estimate the average vehicle speed based on information from video footage. The proposed method utilized non-local blocks, optical flow, and a multi-scale convolutional network. The results show that the proposed method results in reliable speed estimation with a mean absolute error of 2.71 km/h. Huang [23] used perspective transformation to estimate vehicle speeds. While using assumptions for the pixels to field measurement conversion, the study by Huang shows that perspective transformation can result in accurate speed estimation with reduced computational complications. The study also highlights that speed estimation in urban settings is far more challenging than uninterrupted freeway flow.

Accurate and efficient speed estimation remains a challenging task, especially from low-altitude roadside cameras where methods utilizing license plate tracking and deep neural nets outperform those based on vehicle tracking due to detection instability and frequent identity switches. Given the proven ability of our VT-Lane vehicle tracking framework to provide accurate vehicle trajectories in real-time while resolving vehicle identity switches due to occlusion [24, 25], the authors believe there is a high potential for speed estimation with reliable accuracy from the produced trajectories without the need for additional complex deep neural nets or the privacy concerns associated with license plate tracking.

3.1.3 Objective

In this study, we assess an end-to-end application of our VT-Lane framework to obtain real-time vehicle trajectories and movement classification, followed by image processing and reference object scaling to obtain vehicle speed estimates. We verify the accuracy of the speed estimation using high-granularity data obtained from an instrumented vehicle that is tracked as it is driven through the study site. We conclude with a value proposition of utilizing the high accuracy speed estimates for the calibration of car-following models.

3.2 METHODOLOGY

3.2.1 Vehicle Tracking and Movement Classification Framework

For the base object detection and tracking task in VT-Lane, we implemented a combination of YOLO v4 [6] and Deep-SORT [26] algorithms. Our complete framework is detailed and assessed in [24]. The proposed framework has proven efficient in trajectory tracking, turn movement classification, and trajectory reconstruction after effectively handling identity switches. The efficacy of our work was further verified through the 2020 AI City Challenge [25], [15]. The YOLO v4 object detector was pre-trained on the Microsoft Common Objects in Context (COCO) dataset. For this study, we improve the Deep-SORT tracker by retraining on the UA-Detrac dataset [14] which resulted in significant performance improvements in congested traffic conditions. We also incorporated Hou et. al's work [27] for low confidence track filtering. This pipeline runs at ∼35 fps on NVIDIA Tesla V100 GPU. Given that the majority of traffic control devices run at a 10 Hz rate, the real-time capabilities of this pipeline ensure seamless integration with the existing infrastructure.

3.2.2 Image Processing and Speed Estimation

Detection and tracking of vehicles takes place in the original camera space. We then utilize perspective transformation to flatten the traffic intersection scene of the area of study into a bird's-eye 2D space as shown in Figure 3.1, which also illustrates the NEMA movements for the site of study. The transformed perspective results in a similar view to that of the satellite image of the intersection as shown in Figure 3.2.

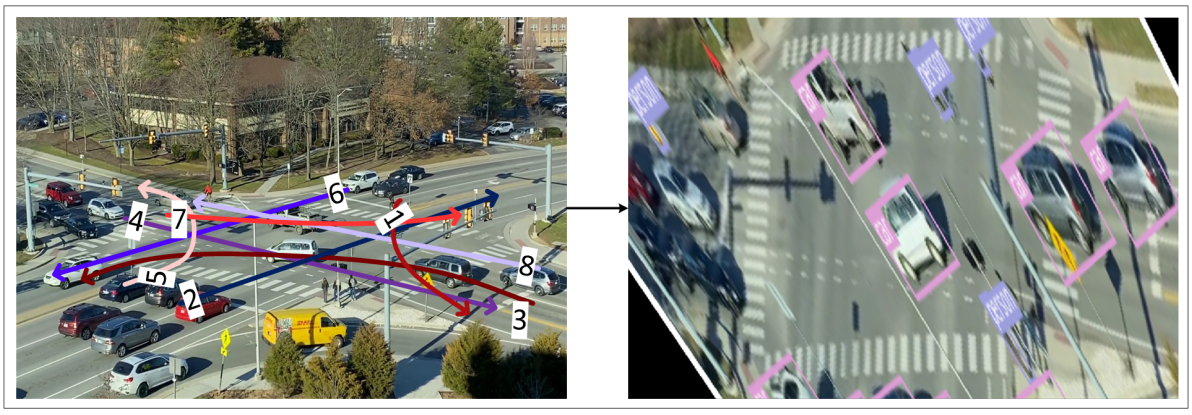


Figure 3.1: Original video view and perspective transformation.

We take advantage of this linearized transformed perspective to calculate the meters per pixel (MPP) ratio using reference objects, which allows the calculation of the actual distance traveled by vehicles. We used the pedestrian crossing markings around our region of interest inside the intersection as a reference, being of standard length on all sides of the intersection. The markings are $\sim$ 2.75 meters in length. The pixel lengths of the markings in the transformed perspective space were found to change from 130 to 90 from the bottom to the top in the Y direction, respectively, and 60 to 100 from the left to the right in the X direction, respectively.

To account for the change in MPP, when calculating distances for speed calculations, we use euclidean distance with applied weights to the changes in MPP for the X and Y coordinates

Figure 3.2: Google Maps image of the study intersection.

as shown in Equation 3.1. A linear interpolation function is applied to calculate the average MPP respectively for X and Y coordinates based on the midpoint of Δ X and Δ Y. For instantaneous speed calculation, Δ X and Δ Y would respectively be the frame-to-frame change in the location of the centroid of the detected vehicle. Knowing the actual distance traversed and the frame rate of the input video, the frame-by-frame speed is calculated following Equation 3.4.

$$Distance_{point_1,\ point_2} = \sqrt{(\Delta X * W_x)^2 + (\Delta Y * W_y)^2} \tag{3.1}$$

Where:

$Distance$ = Weighted Euclidean Distance in m.

X and Y = Pixel coordinates in transformed perspective space of the two points (in this case, the centroids of bounding boxes for the detected vehicles).

W_x and W_y = Average MPP in the X and Y directions, respectively.

$$W_c = \frac{MPP_{c\ max} - MPP_{c\ min}}{c_{max} - c_{min}} * (c_{midpoint} - c_{min}) + MPP_{c\ min} \tag{3.2}$$

$$c = \begin{cases} x, & \text{for } X \textit{ coordinates} \\ y, & \text{for } Y \textit{ coordinates} \end{cases} \tag{3.3}$$

Where:

W_c = Weight of coordinates in meters/pixel.

$MPP_{c\ min}$ and $MPP_{c\ max}$ = The meters/pixel for the coordinates at the minimum (left for X, bottom for Y) and maximum (right for X, top for Y), respectively.

c_{min} and c_{max} = The minimum and maximum pixel location for the coordinate axis, respectively.

$c_{midpoint}$ = The midpoint of the distance being calculated.

The estimated speed profile of all vehicles is calculated by applying Equation 3.4 to all vehicles across all detected frames as shown in Algorithm 2. The speed estimation is skipped for the first appearance of each vehicle as there is no prior knowledge of the location of the vehicle in previous time steps.

$$Speed_{i,k} = \frac{Distance(centroid_{i,\ k}\ ,\ centroid_{i,\ k-1})}{(frame_k - frame_{k-1})/fps} \qquad \forall_{i,\ k} \tag{3.4}$$

Where:

$Speed_{i,k}$ = The estimated speed of vehicle i in frame k in $meter/sec$.

$centroid_{i,k}$ and $centroid_{i,k-1}$ = The centroids of the bounding box for tracked vehicle i in the current and previous frame in which it was detected, respectively.

$frame_k$ and $frame_{k-1}$ = The current and previous frame in which vehicle i was detected, respectively.

fps = The number of frames per second (=30 for this study).

Algorithm 2: Speed Estimation

Result: Estimating the frame-by-frame speed for all identified vehicles.
initialization;
for car_i do
  for $frame_k$ do
    if $frame_{i,k} = frame_{i,1}$ then
      Continue;
    else
      $Speed_{i,k} \leftarrow \frac{Distance(centroid_{i,\ k}\ ,\ centroid_{i,\ k-1})}{(frame_k - frame_{k-1})/fps}$;
    end
  end
end

We utilize a highly instrumented vehicle driving through the study site to validate the speeds. The instantaneous speed of the instrumented vehicle is estimated as it passes through the camera frame, which is then compared to the instantaneous speed recorded by the vehicle's on-board devices.

3.2.3 GHR Model Calibration

Following speed estimation and validation, and identification of car-following episodes based on the vehicles' NEMA movement classification, we assess the value of utilizing the instantaneous speed estimates for calibrating car-following models. For this purpose, we utilize the Gazis-Herman-Rothery (GHR) car-following model. Previous studies show that the GHR model generalizes well [28], [29]. Algorithm 2 shows the steps followed for calibrating the GHR model for each NEMA phase for the intersection of study (Figure 3.1).

The GHR model estimates the instantaneous acceleration of the following vehicle in a car-following episode using the formula below:

Algorithm 3: GHR Model Speed Estimation

Result: Estimating the GHR speed and error for all instances of all car-following episodes within each NEMA movement in the intersection.

initialization;
for $NEMA\_Phase_i$ do
  for $episode_j$ do
    for $instance_k$ do
      if $k \leq Floor(\frac{T}{1/fps})$ then
        Continue;
      else
        $\hat{a}_{j,k} \leftarrow cv_n^m(t)\frac{\Delta v(t-T)}{\Delta x^l(t-T)}$;
        $\hat{v}_{j,k} \leftarrow \hat{v}_{j,k-1} + \frac{\hat{a}_{j,k}+\hat{a}_{j,k-1}}{2*1/fps}$;
        $error_{j,k} \leftarrow |v_{j,k} - \hat{v}_{j,k}|$;
      end
    end
  end
  $MAE_i \leftarrow \frac{1}{n_{phase\_instances}} \sum_{j=1}^{n_{episodes}} \sum_{k=1}^{n_{instances}} error_{j,k}$
end

$$a_n(t) = cv_n^m(t)\frac{\Delta v(t-T)}{\Delta x^l(t-T)} \tag{3.5}$$

Where: $c,\ m,\ l, \&\ T$ are GHR model parameters, $a_n(t)$ and $v_n(t)$ are the instantaneous acceleration and speed of vehicle n at time t, respectively, and Δv & Δx are the differences in speed and distance traveled between the leading and following vehicles in a car-following episode.

For the following vehicle in each episode$_j$ of phase$_i$, the GHR model acceleration $\hat{a}$ during instance$_k$ of the episode is calculated based on the distance and speed difference between the leading and following vehicles in the episode at time $(t-T)$ where T is the following driver's perception-reaction time. An estimated speed $\hat{v}$ is calculated based on the GHR model acceleration. The calculated instantaneous speed for the following vehicle $\hat{v}$ is compared to the speed v estimated by our algorithm (considered to be ground truth speed in this case).

An optimization problem is formulated and solved to find the GHR model parameters that minimize the error in terms of the Mean Absolute Error for each NEMA movement (MAE_i) subject to: $0.2 \leq T \leq 1.8$, $0 \leq m \leq 3$, $0 \leq l \leq 2$, $0 \leq c \leq 3$.

3.3 INSTRUMENTED VEHICLE CONFIGURATION

The instrumented vehicle that has been used to test and validate our framework's speed estimation is shown in Figure 3.3. We use a 2017 Chevrolet Bolt equipped with a sensor suite to provide real-time localization of the car. The localization sensor suite of the car consists of a TORC Robotics Pinpoint Module [30] and a computing platform (AS Rock Beebox). The TORC pinpoint module is internally equipped with an Inertial Measurement Unit (IMU), its own small internal computing platform and has two external Novotel GPS sensors attached to it. The data from the IMU and the GPS sensors is fed into the computing platform of Pinpoint and processed using an extended Kalman filter to provide an update on the position and speed of the car at a rate of 100 Hz.

Figure 3.3: The instrumented vehicle used in this study.

The pose information of the car consists of the location of the car in global coordinates (i.e. latitude and longitude), the acceleration and the velocity of the car. This output from pinpoint is read in at a 100 Hz by our computing platform (Beebox) in the car which is powered by QNX - a real time operating system. QNX is time-synced with the TORC Pinpoint to avoid any time discrepancies. Figure 3.4 illustrates the flow of the information between the sensors in the car.

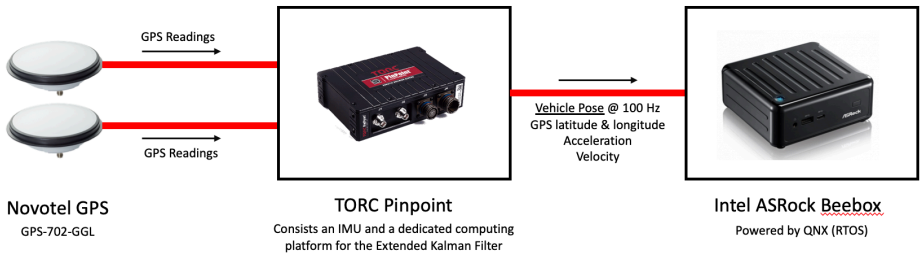


Figure 3.4: Block diagram of the sensor suite in the car.

3.4 RESULTS AND ANALYSIS

3.4.1 Speed Estimation and Validation

A 1-hour video was recorded for the intersection shown in Figure 3.1 during the PM peak hour. The instrumented vehicle was driven to make different movements through the intersection during that time. All vehicles maneuvering the intersection were tracked, and classified based on their NEMA phase movement (also illustrated in Figure 3.1). Speed was estimated for each vehicle on frame-by-frame basis. We used the speed data collected from the instrumented car to compare to the estimated speeds when the car appeared and was tracked in the camera frame within the area of study. The car was also driven during off-peak time to make uninterrupted movements through the intersection. This was made

so as to compare the performance of our speed estimation pipeline when there is no traffic occlusion and stop-start movement at the intersection. Table 3.1 shows the average ground truth speeds during 15 different appearances for the instrumented vehicle at congested and uncongested traffic conditions, and the estimated average speed for those appearances.

Table 3.1: Average Ground Truth vs Estimated Speeds

	Traffic State	NEMA	True Speed ($\frac{m}{s}$)	Estimated Speed ($\frac{m}{s}$)	Error ($\frac{m}{s}$)
1	Light	2	8.81	9.37	+0.56
2	Light	6	10.87	10.39	-0.48
3	Light	2	10.61	11.10	+0.49
4	Light	5	7.04	7.22	+0.18
5	Light	7	5.83	6.21	+0.38
6	Light	1	6.60	7.16	+0.56
7	Light	7	5.84	5.87	+0.03
8	Congested	3	9.74	8.35	-1.39
9	Congested	6	12.17	12.71	+0.54
10	Congested	2	15.20	16.60	+1.40
11	Congested	5	6.89	6.92	+0.03
12	Congested	7	8.20	9.00	+0.80
13	Congested	2	9.16	8.42	-0.74
14	Congested	6	10.81	11.26	+0.45
15	Congested	7	15.09	15.12	+0.03
	Average		9.52	9.71	+0.19

The results indicate a high accuracy in estimating the average speed of the vehicle, with an average error of 0.19 m/sec. The results in Table 3.1 also show that the speed estimation seems to be significantly better in light traffic.

The ground truth speed profile of the instrumented car for a part of the PM peak hour video is shown in Figure 3.5. The yellow scatter points illustrate the estimated instantaneous speed as the instrumented car passes through the camera frame. It can be clearly noticed that there is plenty of noise in case of the raw, unfiltered estimates. We use a moving average smoothing technique to provide more reliable estimates. Figure 3.5 shows the estimated speeds using

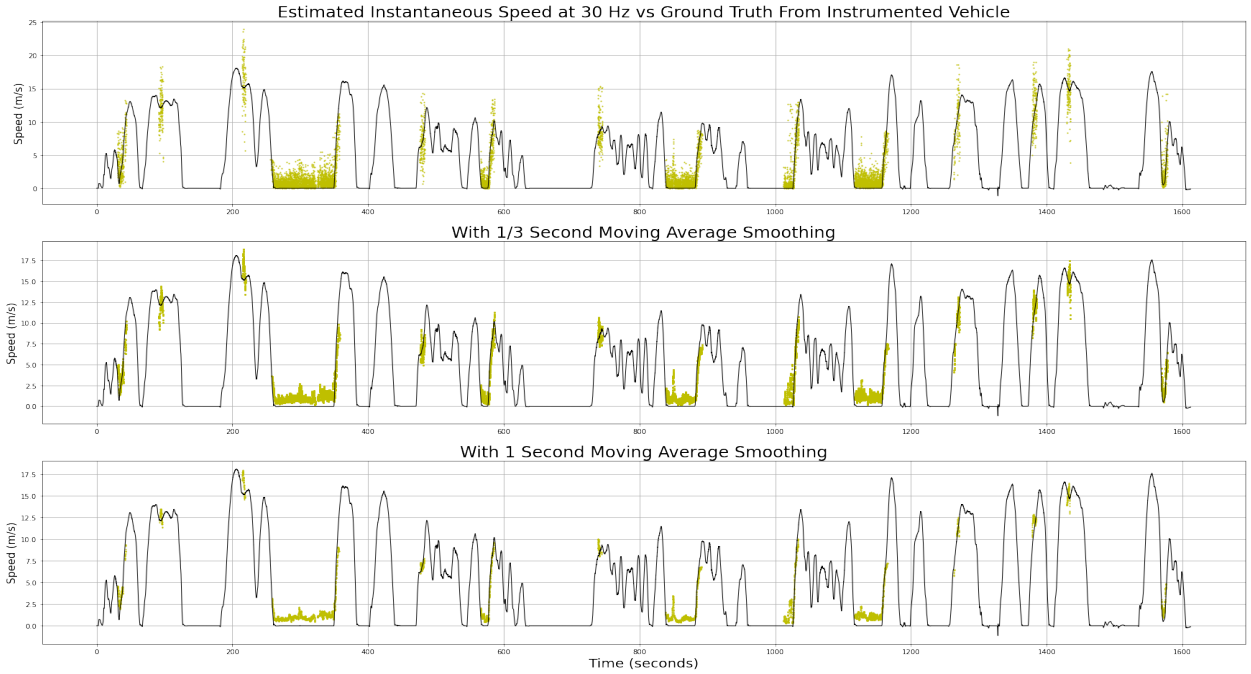


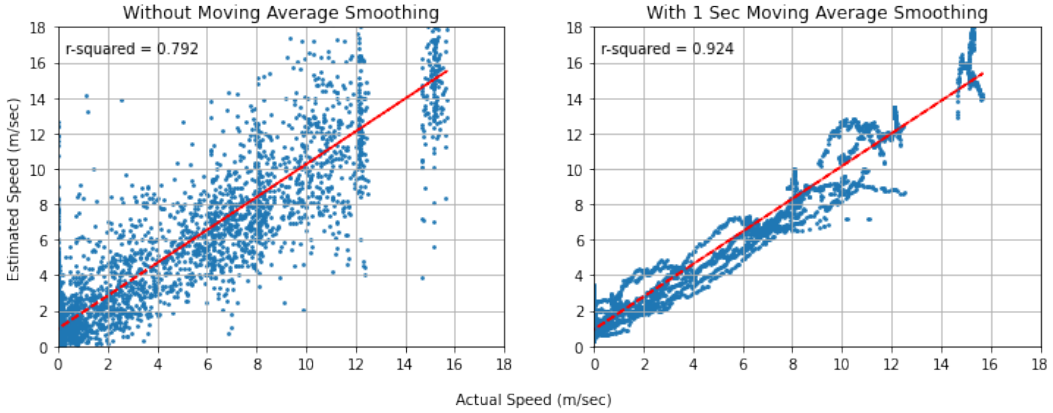
Figure 3.5: Instrumented vehicle's speed profile at 100 Hz (solid line) and estimated speeds (scatter points) with varying moving average window sizes.

1/3-second and 1-second moving average smoothing. While smoothing over longer periods of time would provide improved results, it can be seen that 1-second smoothing is sufficient to provide very reliable speed estimates for the instrumented car. Table 3.2 shows the descriptive statistics of instantaneous speed estimation during the PM peak hour and off-peak times, where N is the number of instances (frames) for which the instantaneous speed of the instrumented car was estimated.

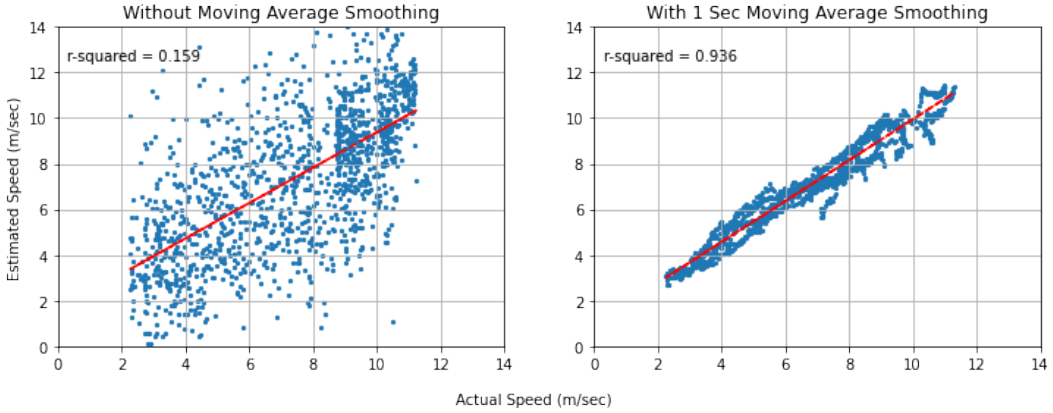
Figure 3.6 shows the actual vs estimated plot for those instantaneous speeds. This illustrates the high accuracy of our method in estimating the speeds when using 1-sec smoothing, as supported by the aforementioned statistics. The speed prediction during light, free-flowing is significantly more accurate than the congested state. There is one instance where a sudden, "v" shaped decrease followed by an increase in the speed of the instrumented car (second-to-last appearance in Figure 3.5) which was not captured by our algorithm, resulting in an

Table 3.2: Descriptive Statistics of Estimated Speeds

Traffic	N	Moving Avg	$\overline{error}$ ($\frac{m}{s}$)	σ ($\frac{m}{s}$)	MAE ($\frac{m}{s}$)	RMSE ($\frac{m}{s}$)
Light	1,594	None	0.61	3.10	2.10	3.16
		$\frac{1}{3}$ sec	0.21	0.80	0.59	0.78
		1 sec	0.21	0.58	0.28	0.45
Congested	8,919	None	0.95	2.07	0.36	1.14
		$\frac{1}{3}$ sec	0.78	0.76	0.16	0.45
		1 sec	0.86	0.46	0.16	0.42



(a) Congested Traffic Condition.



(b) Light Traffic Condition.

Figure 3.6: Actual vs estimated instantaneous speeds.

overestimation of the instantaneous speeds during that time which can be seen in Figure 3.6 (a) for the speeds between 15-16 m/sec. This is attributed to the significantly higher refresh rate of the speed reading from the instrumented car (100 Hz) compared to the input video that provides data to our framework at 30 Hz. Speeds were also found to be over-estimated when vehicles stop at the intersection as a result of the instability of the bounding detection box induced by occlusion as the moving traffic passes in front of idle vehicles.

3.4.2 Driver Behavior Calibration

We conclude with a value proposition of using the speed estimates obtained from video inference to calibrate driver behavior models. Following speed estimation, car-following episodes are identified based on the NEMA phase classification of the vehicle's movement across the intersection. Pairs of vehicles that are simultaneously tracked moving in the same lane of a specific NEMA movement are considered to be in car-following episodes. Table 3.3 shows the GHR model parameters obtained following the optimization process to reduce the MAE of GHR-estimated speeds vs the estimated speeds from video inference for each NEMA phase.

The minimum number of instances indicates the minimum number of frames where both the vehicles in a given car-following episode are tracked simultaneously and have a speed estimate from the video processing (as vehicles are not necessarily detected and tracked through consecutive frames). The average duration ($\overline{D}$) shows the average tracking duration of car-following episodes for a given minimum instances threshold as the vehicles move within the camera frame. Vehicles can be (and most likely are) engaging in car-following episodes well before and after the limited duration they are captured within the camera frame. Blanks indicate that there were not sufficient car-following episodes for the given NEMA movement

Table 3.3: Calibrated GHR Model Parameters For The Intersection of Study

Min Instances	$\overline{D}$ (sec)	NEMA Phase	T	m	l	c	MAE (m/sec)
30	2.63	2	1.74	0.10	2.0	1.39	0.11
		5	1.06	0.10	1.03	1.39	0.25
		4	1.42	2.26	0.98	0.18	0.41
		7	1.17	0.10	0.10	1.19	0.31
		6	0.56	0.30	0.41	1.77	0.71
		1	0.62	0.75	2.00	3.00	0.47
		8	1.08	0.10	2.00	0.10	0.34
		3	-	-	-	-	-
45	3.10	2	1.42	0.10	2.00	2.07	0.32
		5	0.77	0.10	1.27	1.98	0.38
		4	1.29	2.30	1.15	0.23	0.52
		7	0.66	0.41	0.10	0.42	0.51
		6	0.93	0.10	0.10	1.75	0.73
		1	-	-	-	-	-
		8	1.57	0.10	2.00	0.98	0.24
		3	-	-	-	-	-
60	3.60	2	1.02	0.10	1.75	1.34	0.77
		5	1.44	0.14	0.10	0.13	0.24
		4	0.93	2.33	1.57	0.54	0.62
		7	0.52	0.10	0.10	0.69	0.57
		6	1.39	0.10	0.10	2.35	0.91
		1	-	-	-	-	-
		8	1.31	0.10	2.00	0.97	0.37
		3	-	-	-	-	-

to calibrate model parameters. The minimal error obtained via calibration process indicates the high potential of utilizing video-inferred speed estimates for large-scale calibration of driver behavior models.

3.5 DISCUSSION AND CONCLUSIONS

In this study, we presented an end-to-end application of our VT-Lane framework for obtaining real-time vehicle trajectories, movement classification, and speed estimation. The estimated speeds via our framework were verified using high-granularity data obtained from an instrumented vehicle that was tracked as it was driven through the intersection of study. The results of the speed validation show that our framework can estimate speeds in real-time with an error of 0.19 m/sec for estimating the average travel speed of detected vehicles, which is equivalent to 2% of the observed average travel speed (9.52 m/sec) through the intersection of study. Instantaneous vehicle speeds at the resolution of 30 Hz were found to be estimated with an average error of 0.21 m/sec and 0.86 m/sec for free-flowing and congested traffic conditions, respectively, with an overall R^2 of 93%.

We concluded this study with a value proposition of utilizing the high accuracy speed estimates for the calibration of driver behavior models. The parameters of the Gazis-Herman-Rothery model were calibrated for each of the NEMA movements in the intersection of study, with the results showing that the calibrated model replicates the driving behavior with an average error of 0.45 m/sec. The ability of our framework to provide high accuracy, real-time speed estimates, turn movement classification, and identity switch resolution, alongside this demonstrated potential for large-scale driver behavior models calibration from video inference signifies a high practical value for a myriad of applications in traffic safety and simulation modeling.

Our future work will include further expanding the VT-Lane framework and its high accuracy, high granularity outputs for applications in automated real-time performance and safety assessment. The authors will also work to incorporate machine learning for the automated identification and scaling of reference objects from the scenery of a given input traffic video, which would further streamline the speed estimation process and add more value to the framework.

Bibliography

[1] F. H. Administration, "The national intersection safety problem," tech. rep., 2009.

[2] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," in *Advances in neural information processing systems*, pp. 91–99, 2015.

[3] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 779–788, 2016.

[4] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "Ssd: Single shot multibox detector," in *European conference on computer vision*, pp. 21–37, Springer, 2016.

[5] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask r-cnn," in *Proceedings of the IEEE international conference on computer vision*, pp. 2961–2969, 2017.

[6] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "Yolov4: Optimal speed and accuracy of object detection," *arXiv preprint arXiv:2004.10934*, 2020.

[7] K. Karmann and A. Brandt, "Moving object recognition using an adaptive background memory. cappellini v, ed. time-varying image processing and moving object recognition," 1990.

[8] T. Huang, D. Koller, J. Malik, G. Ogasawara, B. S. Rao, S. J. Russell, and J. Weber, "Automatic symbolic traffic scene analysis using belief networks," in *AAAI*, vol. 94, pp. 966–972, 1994.

[9] D. Beymer, P. McLauchlan, B. Coifman, and J. Malik, "A real-time computer vision system for measuring traffic parameters," in *Proceedings of IEEE computer society conference on computer vision and pattern recognition*, pp. 495–501, IEEE, 1997.

[10] S. Kamijo, Y. Matsushita, K. Ikeuchi, and M. Sakauchi, "Traffic monitoring and accident detection at intersections," *IEEE transactions on Intelligent transportation systems*, vol. 1, no. 2, pp. 108–118, 2000.

[11] M. S. Shirazi and B. T. Morris, "Trajectory prediction of vehicles turning at intersections using deep neural networks," *Machine Vision and Applications*, vol. 30, no. 6, pp. 1097–1109, 2019.

[12] C.-E. Wu, W.-Y. Yang, H.-C. Ting, and J.-S. Wang, "Traffic pattern modeling, trajectory classification and vehicle tracking within urban intersections," in *2017 International Smart Cities Conference (ISC2)*, pp. 1–6, IEEE, 2017.

[13] Z. Tang, M. Naphade, M.-Y. Liu, X. Yang, S. Birchfield, S. Wang, R. Kumar, D. Anastasiu, and J.-N. Hwang, "Cityflow: A city-scale benchmark for multi-target multi-camera vehicle tracking and re-identification," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 8797–8806, 2019.

[14] L. Wen, D. Du, Z. Cai, Z. Lei, M.-C. Chang, H. Qi, J. Lim, M.-H. Yang, and S. Lyu, "Ua-detrac: A new benchmark and protocol for multi-object detection and tracking," *Computer Vision and Image Understanding*, p. 102907, 2020.

[15] M. Naphade, S. Wang, D. C. Anastasiu, Z. Tang, M.-C. Chang, X. Yang, L. Zheng, A. Sharma, R. Chellappa, and P. Chakraborty, "The 4th ai city challenge," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pp. 626–627, 2020.

[16] S. Indu, M. Gupta, and A. Bhattacharyya, "Vehicle tracking and speed estimation using optical flow method," *Int. J. Engineering Science and Technology*, vol. 3, no. 1, pp. 429–434, 2011.

[17] S. Doğan, M. S. Temiz, and S. Külür, "Real time speed estimation of moving vehicles from side view images from an uncalibrated video camera," *Sensors*, vol. 10, no. 5, pp. 4805–4824, 2010.

[18] K. Osamura, A. Yumoto, and O. Nakayama, "Vehicle speed estimation using video data and acceleration information of a drive recorder," in *2013 13th International Conference on ITS Telecommunications (ITST)*, pp. 157–162, IEEE, 2013.

[19] L. R. Costa, M. S. Rauen, and A. B. Fronza, "Car speed estimation based on image scale factor," *Forensic science international*, vol. 310, p. 110229, 2020.

[20] M. Garg and S. Goel, "Real-time license plate recognition and speed estimation from video sequences," *ITSI Transactions on Electrical and Electronics Engineering*, vol. 1, no. 5, pp. 1–4, 2013.

[21] D. C. Luvizon, B. T. Nassu, and R. Minetto, "A video-based system for vehicle speed measurement in urban roadways," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 6, pp. 1393–1404, 2016.

[22] H. Dong, M. Wen, and Z. Yang, "Vehicle speed estimation based on 3d convnets and non-local blocks," *Future Internet*, vol. 11, no. 6, p. 123, 2019.

[23] T. Huang, "Traffic speed estimation from surveillance video data," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 161–165, 2018.

[24] A. Abdelhalim and M. Abbas, "Vt-lane: An exploratory study of an ad-hoc framework for real-time intersection turn count and trajectory reconstruction using nema phases-based virtual traffic lanes," in *2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC)*, pp. 1–6, IEEE, 2020.

[25] A. Abdelhalim and M. Abbas, "Towards real-time traffic movement count and trajectory reconstruction using virtual traffic lanes," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pp. 592–593, 2020.

[26] N. Wojke, A. Bewley, and D. Paulus, "Simple online and realtime tracking with a deep association metric," in *2017 IEEE international conference on image processing (ICIP)*, pp. 3645–3649, IEEE, 2017.

[27] X. Hou, Y. Wang, and L.-P. Chau, "Vehicle tracking using deep sort with low confidence track filtering," in *2019 16th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, pp. 1–6, IEEE, 2019.

[28] B. Higgs and M. Abbas, "Segmentation and clustering of car-following behavior: Recognition of driving patterns," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 1, pp. 81–90, 2014.

[29] A. Abdelhalim and M. Abbas, "Vehicle class, speed, and roadway geometry based driver behavior identification and classification," *arXiv preprint arXiv:2009.09066*, 2020.

[30] T. Robotics, "Pinpoint vehicle positioning system," 2014. https://torcrobotics.com/wp-content/uploads/pinpoint_data_sheet1.pdf.

Chapter 4

A Real-Time Safety-Based Optimal Velocity Model

Abstract[4]**:** We propose and evaluate a data-driven Optimal Velocity Model (OVM) based on real-time inference of roadside traffic video data. First, we extract vehicle trajectory data from roadside traffic footage through our advanced‘ video processing algorithm (VT-Lane) for a study site in Blacksburg, VA, USA. Vehicles engaged in car-following episodes are then identified within the extracted vehicle trajectories database, and the real-time time-to-collision (TTC) is calculated for all car-following instances. Then, we analyze the driver behavior to predict the shape of the underlying TTC-based desired velocity function. A clustering approach is used to assess car-following behavior heterogeneity and understand the reasons behind outlying driving behaviors at the intersection to design our model accordingly. Finally, we incorporate this knowledge of on-site driving behavior in an iterative optimization process to calibrate the parameters of a modified OVM in terms of the real-time TTC. The results of this assessment show that the calibrated TTC-based OVM can replicate the observed driving behavior by capturing the acceleration pattern with an error 20% lower than the gap distance-based OVM. This results in an average predicted speed error of 4.74 km/hr, and can provide a better assessment of possible safety interventions modeled via microscopic simulation.

[4]Abdelhalim, Awad, and Montasir Abbas. "A Real-time Safety-Based Optimal Velocity Model." In Review, IEEE Open Journal of Intelligent Transportation Systems IEEE, 2021.

Keywords: Driver behavior calibration, intersection safety, optimal velocity model, vehicle trajectory tracking.

4.1 INTRODUCTION

Vehicle trajectory tracking is one of the major areas of research in Intelligent Transportation Systems (ITS), and is integral to other ITS applications that include driver behavior and car-following analysis, dynamic signal timing, active traffic management, advanced driver-assistance systems (ADAS), among others. The combination of trajectory tracking and driver behavior analysis is key in identifying risks and conflicts that may lead to crashes, allowing practitioners to proactively implement mitigation measures. Both the tasks of accurate and efficient vehicle trajectory tracking and driving behavior modeling remain challenging. Most car-following models (e.g., Gazis-Herman-Rothery (GHR), Wiedemann, Fritzsche) assume that drivers make longitudinal decisions (i.e. acceleration) based on assumed momentary stimuli inputs. For example, the GHR model [1] predicts a driver's acceleration based on the current driver's speed and the sensed difference in speed and distance from the leading vehicle. These models do not account for the driver's planned decision process, where a driver could perceive an impending danger, and plans to avoid that danger, then executes that plan in multiple time steps. This is especially important when modeling safety-critical or close-to-critical situations, when an ego (following) vehicle approaches a vehicle that slows down, forcing the driver of the ego vehicle to start a process of avoidance or adjustment in their speed trajectory.

There is, therefore, a need to estimate the driver intention underlying policy. This policy is expected to potentially vary among drivers (causing heterogeneity in driver behavior). It might also change for the same driver based on the driving context (e.g., driving on

a freeway versus driving on an arterial, executing a turning movement versus continuing straight through an intersection, etc.). Therefore, the policy can take a variable functional form that maps the perceived state into an action plan (a series of maneuvers or a deceleration profile). The closest form of an existing implementation of such an approach can be found in models that attempt to capture a latent driving intention function and then translates that into a predicted desired acceleration. The Optimal Velocity Model (OVM) [2] is an example of such an approach, albeit not explicitly designed with the safety-related input in mind. Instead, the OVM uses the difference in distance between the following and leading vehicle to estimate a desired optimal velocity (using a hyperbolic tangent function form). The OVM then uses the difference between that estimated velocity and current velocity, as well as an intended time-to-execution parameter, to come up with the desired acceleration/deceleration. The issue with the OVM (in this context) is that it assumes that drivers base their "mental intention" on a difference in the distance only. Whereas in safety-critical situations, drivers are more likely to change their intention based on their estimate of the impending danger, as could be measured, for example, by the time-to-collision parameter (TTC).

4.1.1 Objective

This paper examines the hypothesis that using a car-following model based on real-time TTC can better capture the actual driver behavior at urban intersections. We present a case study for a site in Blacksburg, VA, USA. Vehicle trajectories are extracted from roadside video footage. Vehicles engaging in car-following episodes are then identified and an iterative optimization process is formulated to calibrate the parameters of the Optimal Velocity Model (OVM) in terms of the estimated instantaneous time-to-collision. The following sections of this paper are as follows: (a) a survey of related literature in driver behavior modeling, vehicle tracking, and trajectory-based safety, (b) a detailed breakdown of trajectory data

extraction and safety-based model calibration, (c) results and analyses of the case study, and, (d) discussion and conclusions.

4.2 RELATED WORK

4.2.1 Car-Following Behavior Modeling

Car-following and driver behavior modeling has long been one of the most well-studied topics in traffic engineering. Brackstone and McDonald [3] provided a systemic reexamination of the then-existing car-following models, including the GHR model, collision avoidance, linear, fuzzy logic-based and psycho-physical models. They concluded that albeit the extensive study of car-following models, and the strong support of those models in terms of conceptual bases and empirical data, the lack of time-series following behavior is a significant limitation to those models. The emergence of naturalistic driving and trajectory datasets and algorithms [4] alongside advancements in traffic simulation modeling provided a boost to the state-of-the-research allowing researchers to assess driving behavior and car-following under different circumstances and in various roadway geometry settings [5], [6], [7], [8].

One of the most widely adopted driver behavior models is the Optimal Velocity Model proposed by Bando et. al. [2]. The base model developed by Bando et. al. was very successful in describing traffic and capturing behavior during congestion, inspiring the development of OVM-derived models to help capture further aspects of traffic flow more realistically. Wang et. al. [9] proposed an OVM variation with a modified optimal velocity function. Mammar et. al. [10] proposed a modified OVM by utilizing the inverse of TTC as a weighing factor at given distance gap increments to better capture the breaking state of following vehicles. The proposed model was only tested in simulation, finding that the simulated drivers' risk

perception of rear-end collision can be significantly improved. Lazar et. al. expanded on this work [11] and also provided an exhaustive review of some of those models [12] concluding that due to the high variability in driving behavior, each model performs well under certain circumstances and has evident weaknesses in other circumstances. There is, therefore, a need and an immense value in developing data-driven models that would take into account the observed driver behavior heterogeneity [13], and in case of urban intersections, factors such as turn movement classification and signal timing, which all impact the drivers' decision-making processes during car-following episodes. The recent advancements in the fields of vehicle tracking and the growing availability of trajectory data and trajectory-extraction methods make the creation of such data-driven models achievable and necessary.

This newfound availability of trajectory data in recent years has enabled researchers to move from the long-standing models to developing driver behavior models that utilize modern machine learning methods, with a focus on neural and Bayesian networks [14], [15], [16]. Those data-driven models have proven to produce substantial improvements, and overcome the limitations of the conventional, mathematically derived models, especially the tedious process of parameter calibration [17], [18]. Traditional models, however, are based on traffic flow theory fundamentals and hence are highly explainable. The current incomprehensible nature of black-box data-driven models is a major drawback in that regard. Another limitation in the existing literature is the dependence of assessment on long-standing datasets extracted at certain locations from video inference [19] and instrumented vehicles [20] that require further assessment on their transferability to other locations where driver behavior may significantly vary. The evolving automotive technologies over the years may also impact driver behavior in the same location (improved vehicle performance and dynamics, on-board ADAS), etc.), hence there is a need for a more robust and generalizable approach utilizing advances in real-time trajectory extraction.

4.2.2 Vehicle Detection and Tracking for Traffic Safety

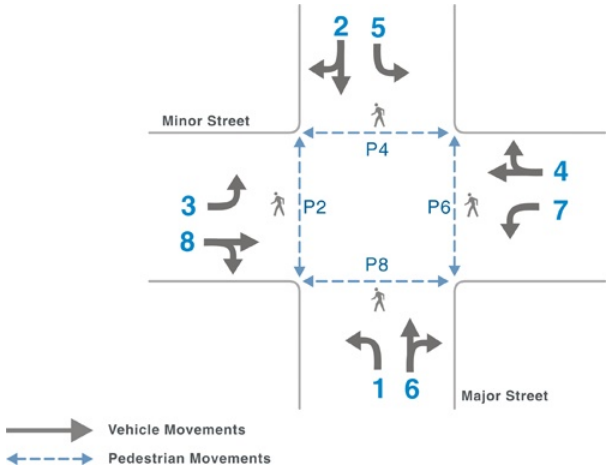
This growing ability to extract and analyze high-quality trajectory data has also enabled researchers to move from developing predictive crash likelihood models based on historical data to a more proactive approach of real-time assessment of traffic safety surrogate measures [21], [22]. Those studies assessed different safety surrogate measures, but the time to collision is the most commonly used measure. St-Aubin et. al. have proposed a video inference-based safety surrogate framework and thoroughly assessed it in multiple studies [23], [24], [25]. Xie et. al. [26], [27] developed a framework to comprehensively assess traffic safety from video data by analyzing conflict risk and TTC on hourly basis, and found a significant correlation between the actual number of crashes and the traffic conflicts inferred through their proposed framework. Das and Maurya [28] utilized trajectory data extracted from an urban traffic environment to assess traffic safety based on the interactions of TTC, roadway center line separation, and leader-follower vehicle type pairing. Estimating TTC from video data is an ongoing research problem, recent studies have assessed TTC estimation from on-board [29] and UAV [30] cameras.

Recent studies both in the fields of traffic safety and driving behavior modeling show the promising prospects of utilizing trajectory data obtained from video inference and conducting TTC-based safety assessments. A recent review study by Li et. al. [31] concluded that there is a need and immense potential in data-driven, trajectory-based driver behavior modeling. Given the proven ability of our VT-Lane vehicle tracking framework to accurately and efficiently extract vehicle trajectories [32] [33] and produce accurate speed estimates [34], this study will investigate utilizing high-resolution video inference-based safety assessment in terms of TTC to propose and calibrate a real-time data-driven car-following behavior model.

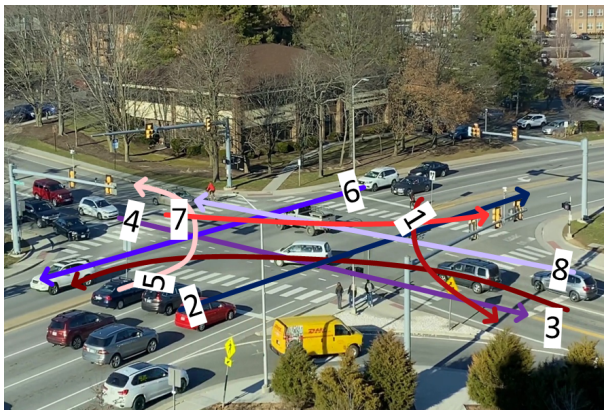
4.3 METHODOLOGY

4.3.1 Area of Study, and the Framework for Vehicle Tracking, Movement Classification, and Speed Estimation

In the United States, the typical traffic pattern assignment for signal controllers follows the National Electrical Manufacturers Association (NEMA) standards shown in Fig. 4.1a [35]. Figure 4.1b shows the site of this study from the perspective of the roadside camera used to obtain the video data for this study. The figure also illustrates the NEMA movement enumeration for the site.



(a) Standard NEMA Phasing.



(b) Site of study and NEMA movements.

Figure 4.1: Site of study and NEMA movement enumeration.

For this study, we utilize an extended VT-Lane framework for the task of vehicle tracking and turn movement classification. The base of the framework is detailed and evaluated in [32] and [33]. The method used for reference object scaling, distance, and speed estimation is discussed and assessed in detail in [34]. A flowchart showing the extended end-to-end framework utilized in this study is shown in Figure 4.2. To improve the accuracy of vehicle

tracking in congested traffic, we retrained the Deep-SORT tracker on the UA-Detrac dataset [36]. Retraining the tracker on different vehicle movements, types, occlusion factors, and glare conditions as illustrated in Figure 4.3 and incorporating low-confidence track filtering [37] resulted in significant performance improvements in congested traffic conditions.

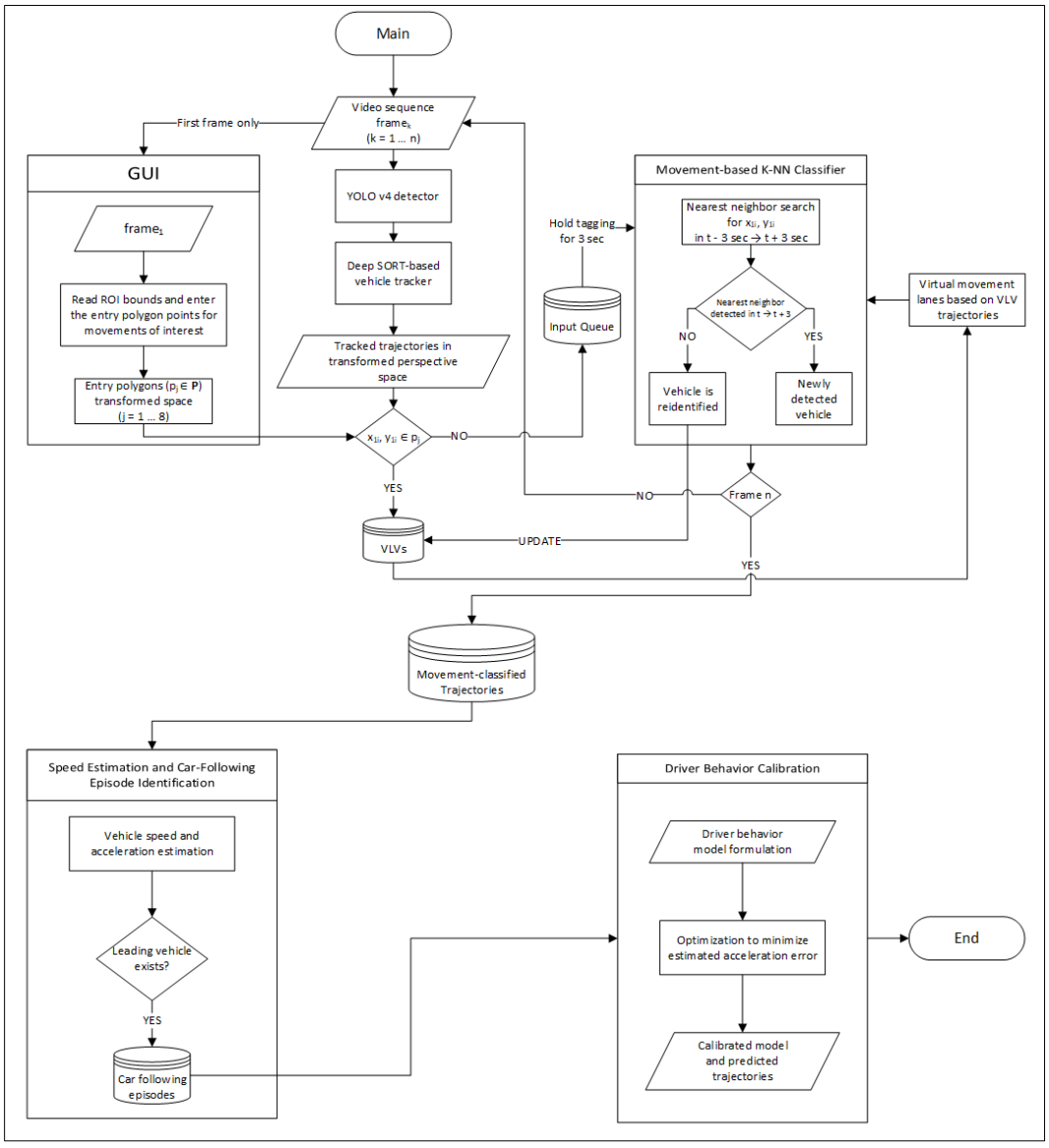


Figure 4.2: Data flow and trajectory data utilization for real-time driving behavior modeling.

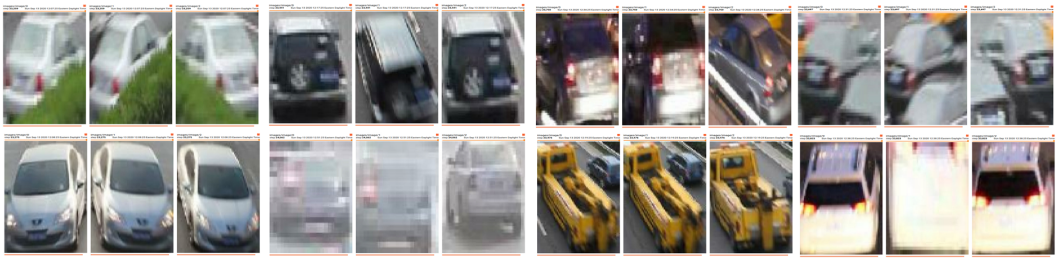


Figure 4.3: Tracker retraining on the UA-Detrac dataset.

4.3.2 Time to Collision Estimation

Following trajectory extraction and speed estimation, the time to collision is calculated for all vehicle pairs broken down by their virtual NEMA movement lane. For the study intersection shown in Figure 4.1b, NEMA movements 2 and 6 have two lanes, resulting in 10 virtual lanes across the intersection (right turners included with through moving vehicles). The classification of vehicle movements via our framework eases the task of calculating the time-to-collision, as it inherently identifies vehicles that are moving in the same lane regardless of their location inside the intersection's geometry. The length of vehicles reported in the output database is used to calculate effective bumper-to-bumper distance as illustrated in Figure 4.4, which is used to calculate the time-to-collision as shown in Equation 4.1 for the following vehicle based on the speed differential for every pair of vehicles detected moving in the same virtual lane within the same frame.

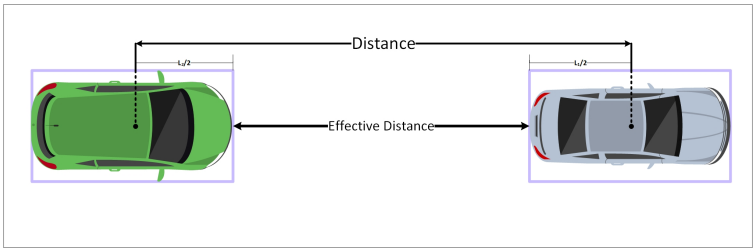


Figure 4.4: Effective bumper-to-bumper distance for time to collision calculation.

Algorithm 4: Time to Collision Calculation

Result: Calculating TTC for vehicle pairs crossing the intersection.

initialization;

for $frame_k$ do

  for $virtual\_lane_n$ do

    if $COUNT(car_n) < 2$ then

      Continue;

    else

      if $car_i \; \exists \; car_{i-1} \;\; \& \;\; \iff \;\; \Delta \; Speed \in \mathbb{R}^+$ then

        Calculate $TTC_{i,n,k}$;

      else

        $TTC_{i,n,k} \leftarrow \infty$

      end

    end

  end

end

$$TTC_{i,n,k} = \frac{\Delta \; Distance - \frac{1}{2}(L_{i,n,k} + L_{i-1,n,k}) * \gamma}{\Delta \; Speed} \; \forall_{i, \; n, \; k} \tag{4.1}$$

Where:

$TTC_{i,n,k}$ = Time to collision in *seconds* for vehicle i to the leading vehicle $i-1$ moving in the same NEMA virtual lane n during frame k.

$\Delta \; Distance$ = The euclidean distance between the centroids of detection boxes of the leading and following vehicle.

$L_{i,n,k}$ = Length of vehicle i during frame k in *pixels* (vehicle length changes slightly as it traverses the intersection due to the vanishing point problem).

γ = Conversion function for vehicle length in *pixels* to *meters* incorporating a pixel-per-meter weight function.

$\Delta \; Speed$ = The speed differential between the following and leading vehicles (i.e. $Speed_{i,n,k} - Speed_{i-1,n,k}$) in *meter/sec*. A negative speed differential occurs when the leading vehicle is moving faster, resulting in a negative TTC meaning that the vehicles will not collide.

4.3.3 TTC-Based Optimal Velocity Model Calibration

Following TTC calculation, vehicles engaging in car-following episodes are identified based on the threshold shown in Equation 4.2. This assumes that a following driver's behavior is not influenced by the leading vehicle if the estimated instantaneous TTC is greater than 20 seconds.

$$\text{Leading Vehicle} = \begin{cases} \text{vehicle}_{\text{ i-1, n}}, & \text{if TTC} \leq 20s \\ \text{None}, & \text{otherwise} \end{cases} \tag{4.2}$$

We then assess the value of utilizing the instantaneous TTC estimates for calibrating car-following models. For the purpose of exploratory assessment in this study, we utilize the base Optimal Velocity Model. The optimal velocity of a following vehicle based on the distance gap is given by Equation 4.3, from which the desired acceleration is calculated using Equation 4.4.

$$v_{opt}(s) = v_o \frac{\tanh\left(\frac{s}{\Delta s} - \beta\right) + \tanh\,\beta}{1 + \tanh\,\beta} \tag{4.3}$$

$$\dot{v} = \frac{v_{opt}(s) - v}{\tau} \tag{4.4}$$

Where:

s = Distance gap between vehicles in a car-following episode (m).

$v_{opt}(s)$ = The theoretical optimal velocity for a given distance gap (km/hr).

v_o = Desired speed (km/hr), Δs = Transition width (m).

β = Form Factor, $\dot{v}$ = OVM acceleration ($km/hr/sec$).

v = Actual speed of a following vehicle (km/hr), τ = Adaptation time (sec).

Alongside calibrating the model in terms of the distance gap between vehicle pairs, we also calibrate in terms of the estimated real-time TTC. We also introduce a data-driven modified acceleration function (MAF) capturing the observed car-following behavior from video inference. Equation 4.4 becomes:

$$\dot{v} = (1-\alpha)\frac{v_{opt}(ttc_{instance}) - v}{\tau} + \alpha f(ttc_{observed}) \tag{4.5}$$

The process of calibrating the parameters of the Optimal Velocity Model from the inferred vehicle trajectories is described in Algorithm 5.

Algorithm 5: OVM Parameter Calibration

Result: Calibrating the parameters of OVM in terms of estimated TTC.
initialization;
for $movement_i \in \{\text{through}, \text{turning}\}$ do
  for $episode_j$ do
    for $instance_k$ do
      $v_{opt_{j,k}} \leftarrow v_o \frac{\tanh\left(\frac{ttc}{\Delta s} - \beta\right) + \tanh\ \beta}{1 + \tanh\ \beta}$;
      $\dot{v}_{j,k} \leftarrow (1-\alpha)\frac{v_{opt_{j,k}} - v_{j,k}}{\tau} + \alpha f(ttc_{observed})$;
      $error_{j,k} \leftarrow \left(a_{j,k} - \dot{v}_{j,k}\right)^2$;
    end
  end
  $MSE_i \leftarrow \frac{1}{n_{instances_i}} \sum_{j=1}^{n_{episodes_i}} \sum_{k=1}^{n_{instances_j}} error_{j,k}$
end

We utilize the turn movement classification from VT-Lane to produce separate calibrated models for through-moving vehicles and turning vehicles for the intersection of study. For the following vehicle in each episode$_j$ of movement type$_i$ (where i is either through or turning movement), the OVM acceleration $\dot{v}$ during instance$_k$ of the episode is calculated based on the estimated TTC between the leading and following vehicles. The true instantaneous

acceleration (a) of vehicles is calculated based on their frame to frame speed changes. An optimization problem is formulated to minimize the Mean Square Error of the true acceleration vs OVM acceleration by calibrating the parameters of the OVM model (v_o, Δs, β, τ, and α).

4.4 RESULTS AND ANALYSIS

4.4.1 Video Inference and Vehicle Trajectory Extraction

A 1-hour footage was recorded and analyzed for the study intersection during the PM peak hour. The trajectory extraction and movement classification obtained via VT-Lane is illustrated in Figure 4.5. A total of 2,656 vehicles were detected and tracked for the full 1-hr footage. Table 4.1 shows a sample of the actual and detected movement counts from a 30-minute segment of the video.

Table 4.1: 30-Minute Actual vs Detected Turn Counts

NEMA Phase	Actual	Detected	Accuracy (%)
2	438	415	95
5	69	62	90
4	93	81	87
7	53	49	92
6	421	398	94
1	31	30	97
8	52	47	90
3	137	129	94
Total	1294	1211	93.5%

Alongside the vehicle trajectories and movement classification, car-following episodes were identified, the instantaneous speed and distance differentials were estimated, and the time-

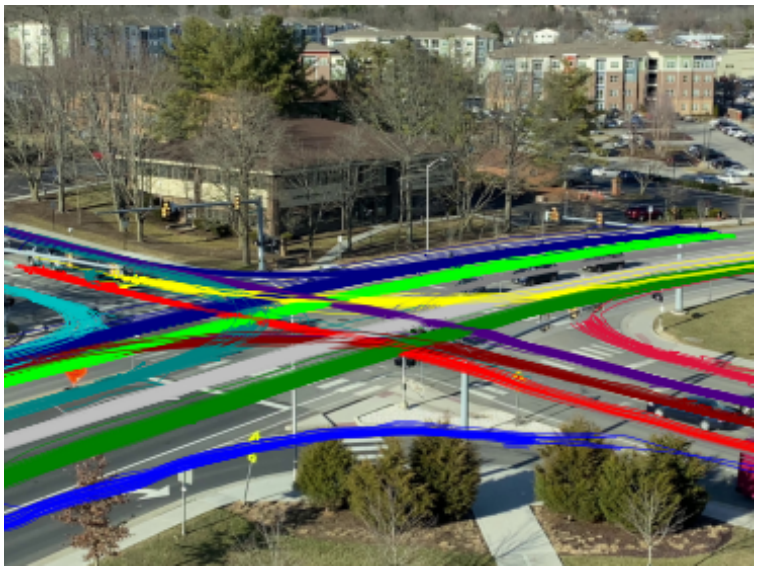


Figure 4.5: Trajectories and turn movement classification.

to-collision was calculated accordingly. A total of 588 car-following episodes were identified for the 1-hr video data. Figure 4.6 shows a sample car-following episode. For both the acceleration and TTC calculations, a Savitzky–Golay filter was applied to obtain smoother estimates.

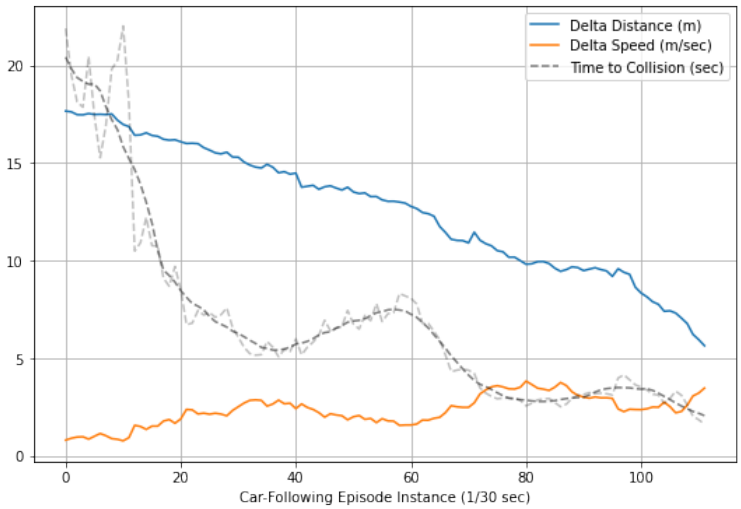
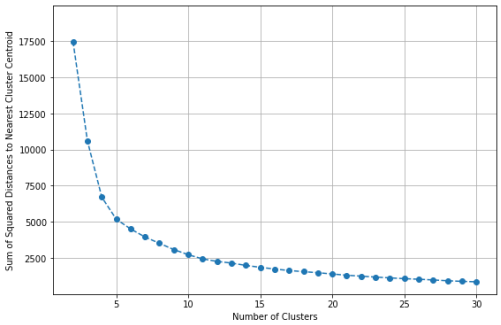


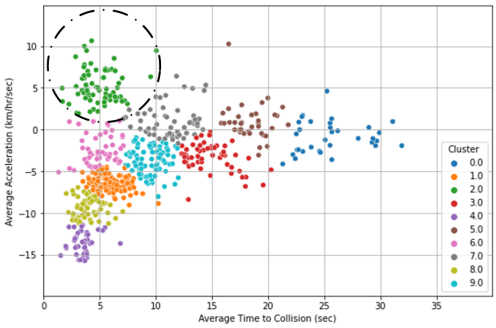
Figure 4.6: Sample car-following episode.

4.4.2 Model Calibration

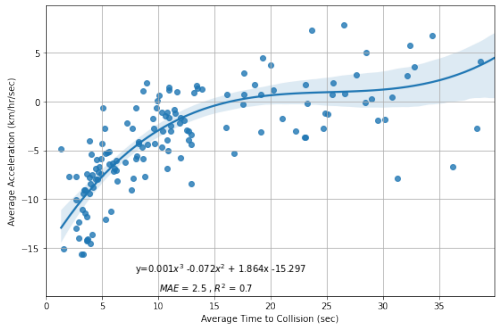
Prior to model calibration, we plot the average acceleration versus time to collision for all 588 car-following episodes. This was carried out to identify outlying behavior that would bias the model. A threshold of ten clusters emerged as one after which no significant reduction in distance to nearest centroid was achieved to justify further clustering as illustrated in Figure 4.7 a. When grouping the episodes into ten clusters, an outlying behavior cluster was identified (shown in Figure 4.7 b) where the following vehicles seemed to be accelerating despite the lower TTC (≤ 5 seconds). All members of this cluster (72 episodes) were found to be episodes belonging to NEMA phase 6, which, alongside phase 2 had the highest volumes. One of the two lanes of phase 6, however, was a shared lane with right turners, resulting in higher queues, and hence the acceleration behavior was significantly higher for episodes in this phase compared to the remaining NEMA movements.



(a) Total error vs number of observed driving behavior clusters.



(b) Identifying outlying driving behavior.



(c) Data-driven deceleration function.

Figure 4.7: Behavior clustering and data-driven function fit.

After identifying the outlying behavior and understanding its causes, we took a 20% subset of the remaining observed episodes (516 episodes) to fit a function that captures the observed deceleration behavior based on TTC. The function is added with a weighing factor (α) to the OVM's acceleration function. A 3rd order polynomial was found enough to capture the observed behavior to a satisfactory extent, with a mean absolute error of 2.5 $km/hr/sec$ and an R^2 of 0.70. Figure 4.8 shows the optimal speed function of the calibrated OVM as a function of the TTC (the dotted line shows the TTC-Based model with modified acceleration function). Table 4.2 shows the optimal model parameters for the base, the TTC-Based OVM, and the modified acceleration function (MAF) model.

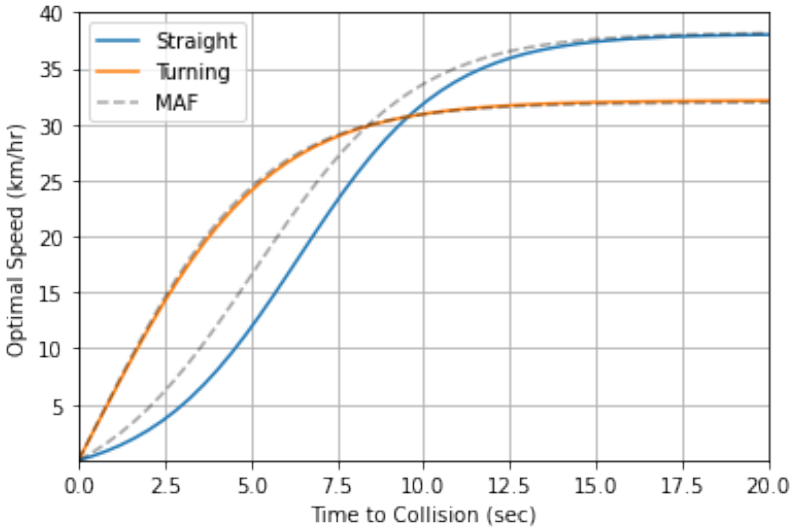


Figure 4.8: Real-time TTC Based Optimal Velocity Model.

Intuitively, both the base and TTC-Based models were optimal with a lower desired speed threshold for turning movements compared to through movements. It should be noted that the Δs parameter is in meters for the base model, and in seconds for the TTC-Based model. The adaptation time τ is in frames (given that the input is in instantaneous, frame-based data), which translates to 0.20-0.25 seconds for the calibrated models. This relatively short adaptation time is attributed to the intersection nature of the area of the study where drivers

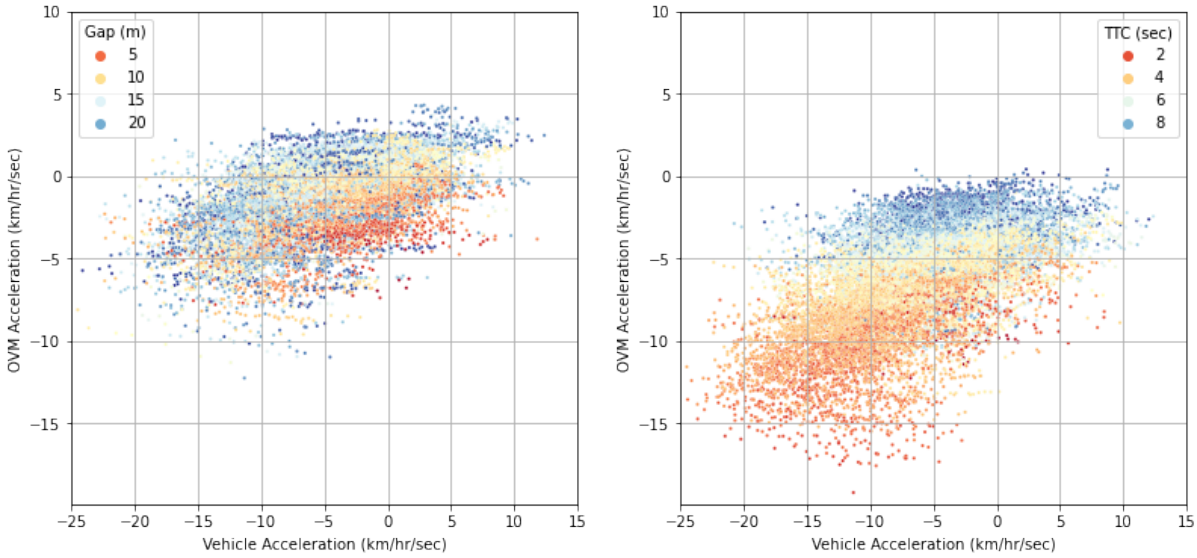
Table 4.2: Optimal Parameters For The Calibrated Models

	Base OVM	Base OVM		TTC-Based		TTC-Based*	
		Thu	Trn	Thu	Trn	Thu	Trn
Δs	9.41	9.98	4.09	4.28	4.93	4.63	4.76
β	0.10	0.10	0.13	1.48	0.11	1.10	0.10
τ	5.12	4.37	4.20	6.49	5.89	6.11	5.83
v_o	36.13	39.63	31.04	38.07	32.13	38.20	31.95
α	-	-	-	-	-	0.18	0.01

make quick decisions. The TTC-Based model with a modified acceleration function (denoted TTC-Based*) provides less conservative optimal velocities as can be seen in Figure 4.8. Given that the OVM doesn't explicitly account for vehicle dynamics, this less conservative optimal velocity function due to the proposed modification based on observed behavior may lead to improved performance in microscopic simulation. The α parameter (weight assigned to the observed acceleration function in the modified TTC-Based model) was close to zero for turning movements. This is attributed to the fact that turning movement car-following episodes only accounted for 7% of all episodes identified, hence the function from the observed 20% subset was more representative of through-moving vehicles.

Figure 4.9 shows the actual vs OVM calculated acceleration for all instances of car-following episodes. The figure clearly illustrates that the TTC-Based model provides a better prediction of the deceleration behavior and can be much better interpreted. Figure 4.9 (a) for the original gap-based OVM is color-coded with the distance gap (0-20+ m) whereas Figure 4.9 (b) for our TTC-based model is color-coded with the TTC (0-10+ seconds). While no clear association can be identified in Figure 4.9 (a), the association between the following vehicles' acceleration and TTC is very clear in the TTC-Based model, where vehicles are accelerating or maintaining speed at higher TTC estimates, and decelerating as TTC decreases. It can be observed that the TTC-Based model provides a significantly better prediction of the

deceleration process.



(a) Gap-Based OVM. (b) TTC-Based OVM.

Figure 4.9: Sample episodes from video inference.

Table 4.3: Instantaneous Acceleration Error (km/hr/sec)

	Base OVM	Base OVM		TTC-Based		TTC-Based*	
		Thu	Trn	Thu	Trn	Thu	Trn
MSE	4.30	4.33	3.80	3.49	3.55	3.45	3.55
mean	-0.01	-0.02	0.09	-0.18	-0.05	-0.10	-0.09
σ	5.46	5.49	4.95	4.42	4.63	4.39	4.63
5%	-8.37	-8.29	-8.15	-7.04	-7.57	-6.92	-7.55
25%	-3.68	-3.79	-2.89	-3.14	-2.98	-3.03	-2.99
50%	-0.33	-0.43	0.07	-0.37	-0.01	-0.31	-0.02
75%	3.44	3.42	2.91	2.56	2.62	2.62	2.61
90%	7.13	7.21	6.38	5.60	5.42	5.61	5.43
95%	9.50	9.63	8.68	7.53	7.75	7.50	7.76

It can also be observed in Table 4.3 that the TTC-Based model results in significant improvements in the MSE and standard deviation of errors compared to the base distance-

based model. Improvements were slightly lower for turning movements. This is attributed to the fact that only 43 of the 588 episodes identified were turning, hence the availability of more data for through movements resulted in a better-calibrated model. The percentage of change in error compared to the base OVM model (distance-based model without accounting for turning movement classification from video inference) is illustrated in Figure 4.10. The figure illustrates that introducing turning movement classification alone leads to significant improvement in the base model's performance for capturing the car-following behavior of vehicles executing turning movements. Substantial improvements are achieved via the TTC-Based model, with the TTC-Based MAF model having an error 19.84% lower than the base model.

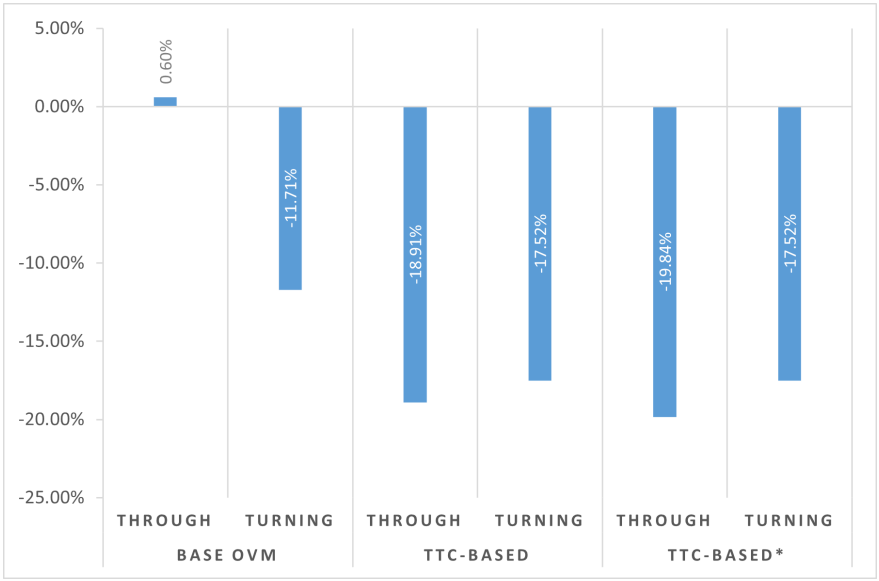


Figure 4.10: Percentage of change in error compared to the base OVM.

Figure 4.11 shows a plot of the actual versus TTC-Based model predicted speeds (color-coded with the TTC) for all instances of car-following inferred from the 1-hr footage. Intuitively, lower TTCs were associated with following vehicles driving at high speeds, and vice-versa. The plot also illustrates that the model performed significantly better at lower following speeds, while the variance increased at speeds higher than 50 km/hr. The average error in

instantaneous speeds estimation was found to be 4.74 km/hr for through movements, and 0.54 km/hr for turning movements.

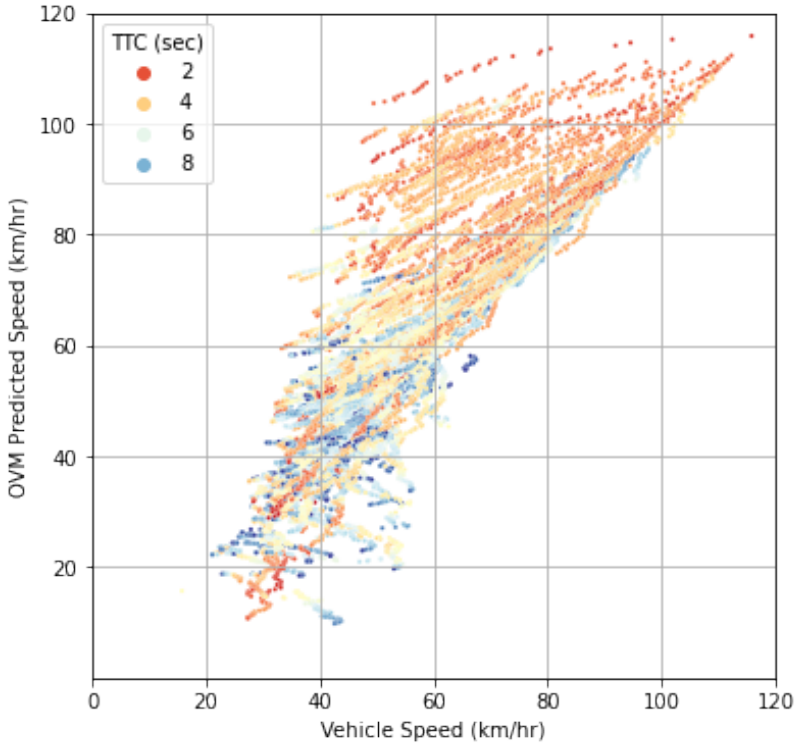


Figure 4.11: Actual vs TTC-Based model predicted speeds.

Finally, Figure 4.12 shows sample episodes, illustrating the actual speed and acceleration profiles of the following vehicles, the calibrated TTC-Based OVM acceleration, and the resulting speed profile. The calibrated model captures the acceleration/deceleration behavior well enough to produce comparable speed profiles to the actual data, albeit the instability of acceleration estimation induced by the high-resolution of video inference (at 28-30 fps). The results not only illustrate the accuracy with which the calibrated model is able to capture the actual deceleration pattern, but also the high stability at which this is accomplished (i.e., no severe abrupt acceleration or deceleration).

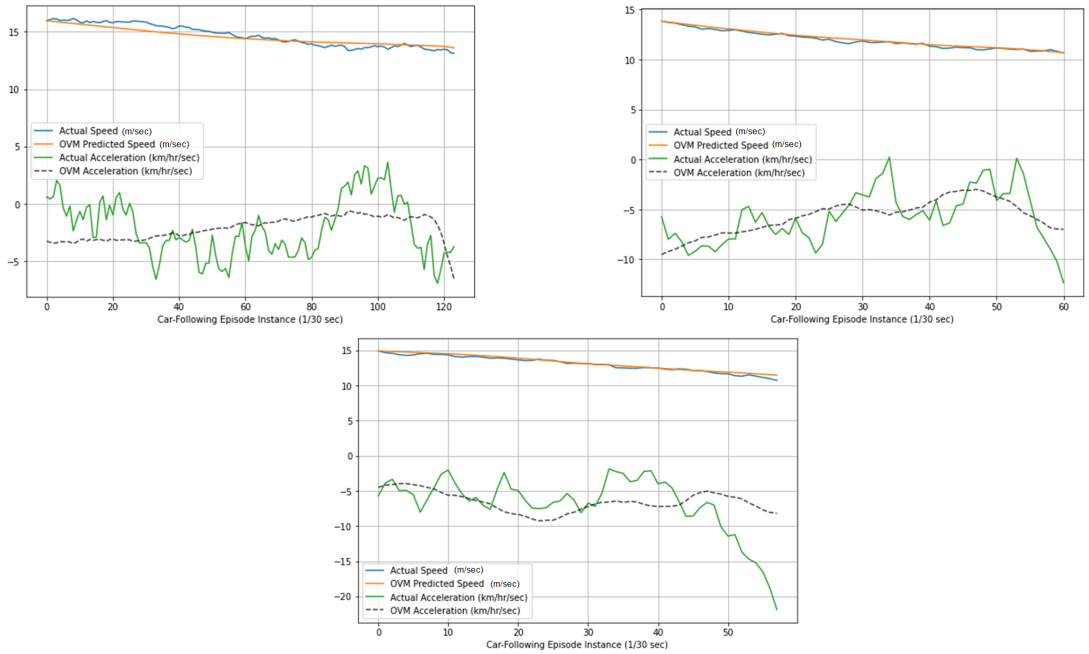


Figure 4.12: Sample episodes and model predicted speed and acceleration profiles.

4.5 DISCUSSION AND CONCLUSIONS

In this study, we examined and assessed the value of driver behavior modeling based on real-time vehicle trajectories and time-to-collision inferred from traffic video data. We utilized the ability of our VT-Lane framework to efficiently extract accurate vehicle trajectories, movement classification, and speed estimates to calibrate the parameters of the Optimal Velocity Model for the area of study both in terms of the distance gap and the estimated TTC from video inference. The TTC-Based model was able to produce improved deceleration estimates for vehicles engaged in car-following episodes, and result in speed profiles with an instantaneous estimated speed error of 4.74 km/hr and an instantaneous acceleration error 19.84% lower than that of the base, distance-based model. The results obtained utilizing real-time data and through the modified acceleration function based on observed behavior show the high value of infusing driving behavior models with data-driven modifications.

The clear differences in model parameters between straight moving vehicles and vehicles executing turning movements at the intersection demonstrates the need for taking into account such characteristics in modeling driving behavior in urban settings. While 3rd order polynomial function was able to capture the observed driving behavior with high reliability (R^2 of 70%), further information that can be obtained through video inference (including, and not limited to signal status time, vehicle type, overall traffic density, and lane occupancy, etc) can all be utilized to develop more sophisticated observed driving behavior functions that would potentially further improve the performance the data-driven model. Given the high-resolution and accuracy of the trajectories obtained through VT-Lane, utilizing a Kalman filter or similar continuous state estimation approach to improve the desired acceleration estimates based on the continuity of the trajectory instead of relying on single measurements is another route worth exploring.

The short length of the car-following episodes, given that vehicles are only tracked as they cross the intersection, was a limiting factor for this study. Although vehicles engaged in car-following within the area of study are expected to be engaging in car-following behavior before and/or after the monitored segment. Future studies will be conducted on longer arterial segments to evaluate leader-follower interactions over longer periods. Future work will also assess and quantify the improvements that can be obtained in microscopic simulation by using the proposed model to calibrate site-specific driving behavior and utilizing that as an external driver model. The end-to-end nature of this study starting with extracting vehicle trajectories from roadside cameras, and utilizing those trajectories for calibrating a data-driven diver behavior model provides a blueprint for practitioners to both accurately assess real-time traffic safety and performance, as well as aid in developing simulation models that better mimic the behavior of existing on-site traffic, allowing for a more accurate assessment of any proposed safety interventions.

Bibliography

[1] D. C. Gazis, R. Herman, and R. B. Potts, "Car-following theory of steady-state traffic flow," *Operations research*, vol. 7, no. 4, pp. 499–505, 1959.

[2] M. Bando, K. Hasebe, A. Nakayama, A. Shibata, and Y. Sugiyama, "Dynamical model of traffic congestion and numerical simulation," *Physical review E*, vol. 51, no. 2, p. 1035, 1995.

[3] M. Brackstone and M. McDonald, "Car-following: a historical review," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 2, no. 4, pp. 181–196, 1999.

[4] V. Punzo, M. T. Borzacchiello, and B. Ciuffo, "On the assessment of vehicle trajectory data accuracy and application to the next generation simulation (ngsim) program data," *Transportation Research Part C: Emerging Technologies*, vol. 19, no. 6, pp. 1243–1262, 2011.

[5] S. H. Hamdar, L. Qin, and A. Talebpour, "Weather and road geometry impact on longitudinal driving behavior: Exploratory analysis using an empirically supported acceleration modeling framework," *Transportation research part C: emerging technologies*, vol. 67, pp. 193–213, 2016.

[6] G. S. Aoude, V. R. Desaraju, L. H. Stephens, and J. P. How, "Driver behavior classification at intersections and validation on large naturalistic data set," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 2, pp. 724–736, 2012.

[7] P. St-Aubin, N. Saunier, L. F. Miranda-Moreno, and K. Ismail, "Use of computer vision

data for detailed driver behavior analysis and trajectory interpretation at roundabouts," *Transportation research record*, vol. 2389, no. 1, pp. 65–77, 2013.

[8] A. Talebpour, H. S. Mahmassani, and F. E. Bustamante, "Modeling driver behavior in a connected environment: Integrated microscopic simulation of traffic and mobile wireless telecommunication systems," *Transportation Research Record*, vol. 2560, no. 1, pp. 75–86, 2016.

[9] H. Wang, Y. Li, W. Wang, M. Fu, and R. Huang, "Optimal velocity model with dual boundary optimal velocity function," *Transportmetrica B: transport dynamics*, vol. 5, no. 2, pp. 211–227, 2017.

[10] S. Mammar, S. Mammar, and H. Haj-Salem, "A modified optimal velocity model for vehicle following," *IFAC Proceedings Volumes*, vol. 38, no. 1, pp. 120–125, 2005.

[11] H. Lazar, K. Rhoulami, and M. D. Rahmani, "A modified full velocity difference model based on time to collision as a safely indicator for braking state," in *2015 2nd World Symposium on Web Applications and Networking (WSWAN)*, pp. 1–6, IEEE, 2015.

[12] H. Lazar, K. Rhoulami, and D. Rahmani, "A review analysis of optimal velocity models," *Periodica Polytechnica Transportation Engineering*, vol. 44, no. 2, pp. 123–131, 2016.

[13] S. Ossen and S. P. Hoogendoorn, "Car-following behavior analysis from microscopic trajectory data," *Transportation Research Record*, vol. 1934, no. 1, pp. 13–21, 2005.

[14] L. Chong, M. M. Abbas, A. M. Flintsch, and B. Higgs, "A rule-based neural network approach to model driver naturalistic behavior in traffic," *Transportation Research Part C: Emerging Technologies*, vol. 32, pp. 207–223, 2013.

[15] J. Morton, T. A. Wheeler, and M. J. Kochenderfer, "Analysis of recurrent neural net-

works for probabilistic modeling of driver behavior," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 5, pp. 1289–1298, 2016.

[16] T. A. Wheeler, P. Robbel, and M. J. Kochenderfer, "Analysis of microscopic behavior models for probabilistic modeling of driver behavior," in *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, pp. 1604–1609, IEEE, 2016.

[17] V. Papathanasopoulou and C. Antoniou, "Towards data-driven car-following models," *Transportation Research Part C: Emerging Technologies*, vol. 55, pp. 496–509, 2015.

[18] Y. Yu, *Revisit of Microscopic Car Following Models: Conventional and Machine Learning Perspectives*. PhD thesis, 2021.

[19] V. G. Kovvali, V. Alexiadis, P. Zhang, *et al.*, "Video-based vehicle trajectory data collection," tech. rep., 2007.

[20] V. Punzo, D. J. Formisano, and V. Torrieri, "Nonstationary kalman filter for estimation of accurate and consistent car-following data," *Transportation research record*, vol. 1934, no. 1, pp. 2–12, 2005.

[21] Y. Wan, Y. Huang, and B. Buckles, "Camera calibration and vehicle tracking: Highway traffic video analytics," *Transportation Research Part C: Emerging Technologies*, vol. 44, pp. 202–213, 2014.

[22] G. Guido, A. Vitale, and V. Gallelli, "Investigating rear-end potential conflicts for roundabout turning movements," *Modern Traffic and Transportation Engineering Research (MTTER)*, vol. 2, no. 4, pp. 20–26, 2013.

[23] P. St-Aubin, L. Miranda-Moreno, and N. Saunier, "An automated surrogate safety analysis at protected highway ramps using cross-sectional and before–after video data," *Transportation Research Part C: Emerging Technologies*, vol. 36, pp. 284–295, 2013.

[24] P. St-Aubin, N. Saunier, and L. F. Miranda-Moreno, "Comparison of various time-to-collision prediction and aggregation methods for surrogate safety analysis," tech. rep., 2015.

[25] P. St-Aubin, N. Saunier, and L. Miranda-Moreno, "Large-scale automated proactive road safety analysis using video data," *Transportation Research Part C: Emerging Technologies*, vol. 58, pp. 363–379, 2015.

[26] K. Xie, C. Li, K. Ozbay, G. Dobler, H. Yang, A.-T. Chiang, and M. Ghandehari, "Development of a comprehensive framework for video-based safety assessment," in *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, pp. 2638–2643, IEEE, 2016.

[27] K. Xie, K. Ozbay, H. Yang, and C. Li, "Mining automatically extracted vehicle trajectory data for proactive safety analytics," *Transportation research part C: emerging technologies*, vol. 106, pp. 61–72, 2019.

[28] S. Das and A. K. Maurya, "Defining time-to-collision thresholds by the type of lead vehicle in non-lane-based traffic environments," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 12, pp. 4972–4982, 2019.

[29] M. Kilicarslan and J. Y. Zheng, "Predict vehicle collision by ttc from motion using a single video camera," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 2, pp. 522–533, 2018.

[30] C. Wang, C. Xu, and Y. Dai, "A crash prediction method based on bivariate extreme value theory and video-based vehicle trajectory data," *Accident Analysis & Prevention*, vol. 123, pp. 365–373, 2019.

[31] L. Li, R. Jiang, Z. He, X. M. Chen, and X. Zhou, "Trajectory data-based traffic flow

studies: A revisit," *Transportation Research Part C: Emerging Technologies*, vol. 114, pp. 225–240, 2020.

[32] A. Abdelhalim and M. Abbas, "Vt-lane: An exploratory study of an ad-hoc framework for real-time intersection turn count and trajectory reconstruction using nema phases-based virtual traffic lanes," in *2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC)*, pp. 1–6, IEEE, 2020.

[33] A. Abdelhalim and M. Abbas, "Towards real-time traffic movement count and trajectory reconstruction using virtual traffic lanes," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pp. 592–593, 2020.

[34] A. Abdelhalim, M. Abbas, B. B. Kotha, and A. Wicks, "A framework for real-time traffic trajectory tracking, speed estimation, and driver behavior calibration at urban intersections using virtual traffic lanes," 2021.

[35] U. FHWA, "Traffic signal timing manual," 2008.

[36] L. Wen, D. Du, Z. Cai, Z. Lei, M.-C. Chang, H. Qi, J. Lim, M.-H. Yang, and S. Lyu, "Ua-detrac: A new benchmark and protocol for multi-object detection and tracking," *Computer Vision and Image Understanding*, p. 102907, 2020.

[37] X. Hou, Y. Wang, and L.-P. Chau, "Vehicle tracking using deep sort with low confidence track filtering," in *2019 16th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, pp. 1–6, IEEE, 2019.

Chapter 5

An Assessment of Intervention Strategies Utilizing Real-Time Safety-Based Driver Behavior Modeling in Microscopic Simulation

Abstract[5]**:** In this study, we showcase an implementation of a data-driven Optimal Velocity Model (OVM) to provide high-fidelity simulation of safety-critical behavior in PTV VISSIM, which is utilized to conduct safety-based assessment of traffic control strategies. The real-time safety-based OVM was proposed and calibrated in a recent study for a site of study for a site of study in Blacksburg, VA, USA. We create a detailed simulation model for the site of study utilizing the traffic information extracted from video inference. The calibrated safety-based OVM is then incorporated as external driver model that overtakes VISSIM's default Wiedemann 74 model during car-following episodes. The results of the preliminary analysis show the significant improvements that are achieved using our model in replicating the existing safety conflicts at the site of study. We then utilize this improved representation of the status-quo to assess the impact different scenarios of signal control and speed limit enforcement in mitigating those existing conflicts. A sensitivity analysis for varying traffic volumes was also conducted, concluding with recommendations of the control strategies that

[5]Abdelhalim, Awad, and Montasir Abbas. "An Assessment of Intervention Strategies Utilizing Real-Time Safety-Based Driver Behavior Modeling in Microscopic Simulation." Manuscript in preparation for submission to Transportation Research Part C: Emerging Technologies.

would minimize the generated safety-critical conflicts for each scenario. The results of this study showcase the considerable improvements that can be achieved by utilizing data-driven driving behavior modeling in microscopic simulation as a decision-support tool to identify appropriate safety interventions.

Keywords: Traffic safety, Microscopic simulation, Driver behavior modeling.

5.1 INTRODUCTION

Traffic modeling and simulation is one of the most powerful tools at the disposal of transportation engineers today. The modeling and simulation of specific site, corridor, or network of interest provides transportation practitioners with an efficient and effective method to assess the current performance and evaluate any proposed performance or safety interventions. Simulation modeling can be sub-categorized into three types:

- Macroscopic Simulation, which provides a high-level representation of traffic streams,
- Microscopic Simulation, where vehicles are modeled at an individual level, and
- Mesoscopic Simulation, which aims to utilize a combination of macro and micro-simulation methods.

While macroscopic and mesoscopic are the more common choices of modeling when simulating larger networks, microscopic simulation is more often used for studying the traffic flow in smaller areas, in greater detail. In microscopic simulation, however, the individual vehicles' behavior is based on mathematically-derived driver behavior models that require a tedious process of data collection and parameter calibration to better replicate the observed behavior at a specific area of interest. This has generally been the challenge for their adoption in

simulation modeling.

PTV's Verkehr In Städten SIMulationsmodel (VISSIM) is a a discrete, stochastic, time step based simulation software which is one of the most popular tools used for traffic simulation due to its extensive capabilities for modeling and evaluating different components of the transportation ecosystem (vehicles, public transit, pedestrians, cyclists, roadside infrastructure, etc.). VISSIM's base car-following model is based on the Wiedemann 74 and 99 psycho-physical models [1]. The car-following behavior of vehicles in a VISSIM simulation is modified through ten driver behavior parameters (labeled CC0 – CC9) that represent different thresholds of four assumed driving states: free-driving, approaching, following, and braking. For each of those modes, the desired instantaneous acceleration is a result of the vehicle's current speed, the speed and distance differences between the vehicle and the lead vehicle, and specific characteristics of the driver and the vehicle. The challenge with the Wiedemann model (and other conventional driver behavior models e.g., Gazis-Herman-Rothery, Fritzsche) is that they assume that drivers make longitudinal decisions (i.e. acceleration) based on assumed momentary stimuli inputs. These models do not account for the driver's planned decision process, where a driver could perceive an impending danger, and plans to avoid that danger, then executes that plan in multiple time steps. This is especially important when modeling safety-critical or close-to-critical situations, when an ego vehicle approaches a vehicle that slows down, forcing the driver of the ego vehicle to start a process of avoidance or adjustment in their speed trajectory. The failure to capture this behavior leads to inaccurate representation of traffic when modeled in simulation, hindering the ability to evaluate safety interventions.

5.1.1 Objective and Contributions

This study aims to showcase the improvements that can be achieved in VISSIM simulation of safety-critical behavior utilizing data-driven driver behavior modeling, and provide a case study for the utilization this improved behavior modeling in identifying appropriate safety interventions. The real-time safety-based Optimal Velocity Model, which was developed and calibrated in a recent study [2], was implemented in VISSIM as an external driver model, and utilized to assess the impact different scenarios of signal control and speed enforcement in mitigating the simulated safety concerns. The following sections of this paper are as follows: (a) a survey of related literature in driver behavior modeling in microscopic simulation and trajectory-based safety, (b) a detailed breakdown the VISSIM model development and external driver model implementation, (c) results and analyses of the case study, and, (d) discussion and conclusions.

5.1.2 Related Work

Driver Behavior and Safety Modeling in Microsimulation

Over the past two decades, the growing adoption of microscopic simulation software has provided researchers with powerful means for evaluating different components of transportation networks and their interactions. The simulated vehicle interactions, which form the basis of microscopic driving behavior models, allowed for the assessment of traffic safety and evaluation of countermeasures as opposed to solely relying on analyzing crash data from police reports. The Surrogate Safety Assessment Model (SSAM) [3] is the prime example of such efforts. Using vehicle trajectories exported from microsimulation as an input, SSAM's approach utilizes conflict analysis for near misses, in terms of time-to-collision (TTC) and post encroachment time (PET), inferred from as a surrogate measure of actual crash data

in terms of frequency, types, and severity of crashes. SSAM, however, is dependent on the outputs generated by microsimulation. Hence, adequate parameter calibration for the underlying driver behavior model used to generate the vehicle trajectories is crucial step that the lack thereof was found to result in a significant deviation of simulated safety conflicts from the observed behavior [4], [5], [6].

Given the increasing popularity of VISSIM over the years, numerous studies have been conducted to assess the sensitivity of calibrating its underlying Wiedemann model to better mimic the existing traffic flow and behavior, and result in more realistic simulated safety conflicts. Fellendorf and Vortisch [7] conducted one of the earliest studies, where they assessed the impact of the driver model calibration on both a micro and macroscopic level. The results of their study showcased that a well-calibrated model can replicate traffic flow with high accuracy, with the caveat that driver behavior models should at least take national and regional regulations and driving styles into account to produce reliable representations in simulation. Lownes and Machemehl [8, 9] conducted a sensitivity analysis to assess the impact of modifying VISSIM's driver behavior parameters on simulated traffic flow capacity. Their work has highlighted the impact of changes on driver behavior models at a microscopic level on traffic flow at a macroscopic level, further making the case for the necessity of parameter calibration to ensure adequate representation of traffic modeled in simulation. Other studies have also highlighted the necessity of taking into account the varying traffic patterns and considerations that need to be taken for calibrating VISSIM's driver behavior parameters accordingly, such cases include modeling traffic at urban intersections [10], [11] and freeway merge areas [12]. The work of Fan et. al. [12] particularly stands out in this category, as VISSIM's default driver behavior parameters were calibrated for freeway driving, yet parameter calibration for merging sections specifically was necessary, with the results highlighting the significant improvements obtained. This was attributed to the default VISSIM model

parameters being calibrated from driving data from the German autobahn, while the study was conducted in China, hence the driving behavior parameters were not expected to be transferable.

The transferability of the calibrated VISSIM driver behavior model parameters was assessed by Essa and Sayed [13], who calibrated the VISSIM model parameters to maximize the correlation between observed and simulated traffic conflicts in one intersection and tested the calibrated model on a nearby intersection. Their study concluded that while not all of the calibrated model parameters are fully-transferable, the calibrated still significantly outperforms the default model in terms of providing better correlation between the observed and simulated traffic conflicts. Hence, any safety assessment without proper model calibration should be avoided. Essa and Tarek further highlighted this need for rigorous calibration and the limitations of simulation-based safety assessment in other studies [14], [15].

The extensive literature in VISSIM, and microscopic model parameter calibration in general, lacks a clear and transferable framework to overcome the shortcomings of site-by-site data collection and calibration of driver model parameters to reduce errors versus observed data. In recent years, researchers have opted to utilizing neural network models to overpass this tedious process [16], [17]. And while such methods succeed in producing improved results, they lack interpretability, unlike the traditional driving behavior models which are based on traffic flow theory fundamentals. There exists, therefore, a need for a generalizable driving behavior modeling approach for simulation to better mimic the vehicle trajectories being simulated and their resulting safety conflicts, allowing for a more accurate assessment of proposed safety interventions. To overcome this gap identified in driver behavior modeling calibration and its application in VISSIM, this study showcases an implementation of a data-driven high-fidelity external driver behavior model in VISSIM and utilize it in a case study to assess appropriate safety interventions. The proposed external driver model is

calibrated from video inference for the site of interest, providing a scalable, automated, and generalizable workflow for replicating the observed behavior in microsimulation.

5.2 METHODOLOGY

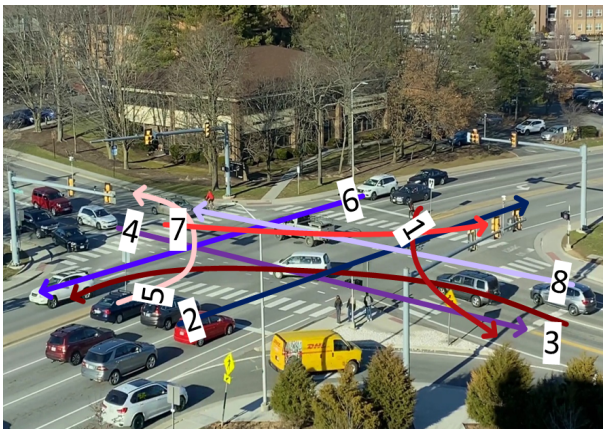
5.2.1 Area of Study, Data Collection, and the Developed VISSIM Model

In a recent study [2], we utilized an extended VT-Lane framework to obtain the trajectories and calibrate a data-driven safety-based driver behavior model for the area of study. The base of the framework is detailed and evaluated in [18] and [19]. The method used for reference object scaling, distance and speed estimation is discussed and assessed in detail in [20]. Figure 5.1 shows the area of this study from the perspective of the roadside camera used to obtain the video data, illustrating the NEMA movement enumeration for the site. The traffic volumes and speed distributions obtained from the inference of 1-hour PM traffic footage were modeled in VISSIM alongside the existing on-site signal timing plan illustrated in Figure 5.2.

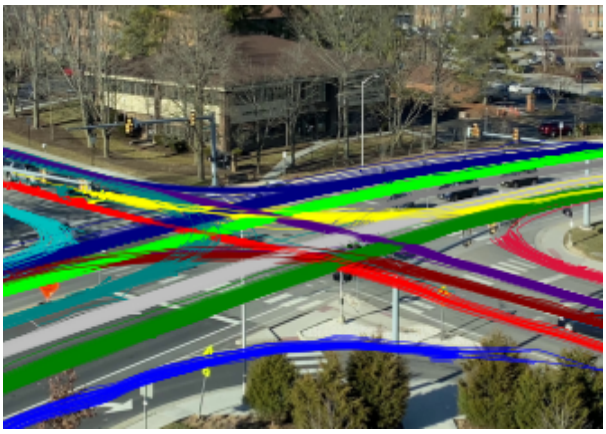
5.2.2 Incorporating the Safety-Based Driver Model

VISSIM's external driver model module is utilized to incorporate our proposed safety-based OVM. The C++ generated Dynamic Link Library (DLL) file is attached to VISSIM's simulation runs once every simulation step for all vehicles in the network during the simulation step. All vehicles in the network are controlled by the default Wiedemann model that VISSIM utilizes, unless they are engaged in car-following episodes. In this study, a vehicle is considered in car-following if it is moving at a speed > 5 m/sec (18 km/hr) and there exists a slower leading vehicle moving in the same lane. This moving speed threshold was set as the

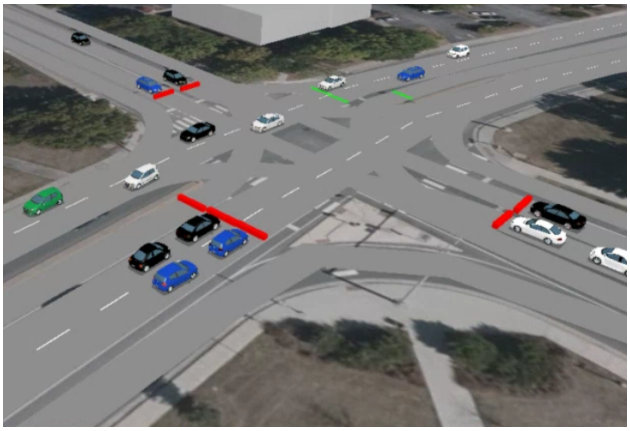
safety-based OVM model utilized was calibrated in [2] for vehicles that are moving across the intersection of study. The threshold ensures that the model is not utilized for vehicles that are stopping at the intersection due to signal control, which could be assumed to be in a car-following model based on estimated TTC. Figure 5.3 shows a flowchart with the code execution logic for driver behavior model selection in VISSIM, which is detailed in Algorithm 6.



(a) Site of study and NEMA movement labels.



(b) Extracted vehicle trajectories.



(c) The VISSIM model developed for this study.

Figure 5.1: Site of study and tracked vehicle trajectories.

For the safety-based OVM, the instantaneous optimal velocity and resulting desired acceleration are calculated using the following equations:

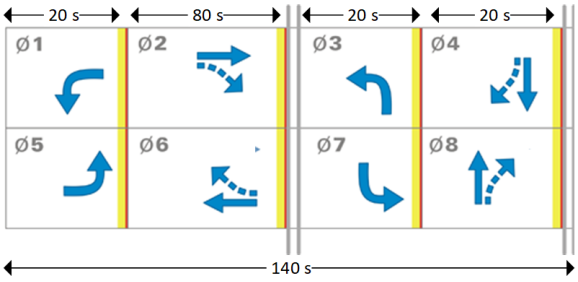


Figure 5.2: NEMA movements and on-site signal timing plan.

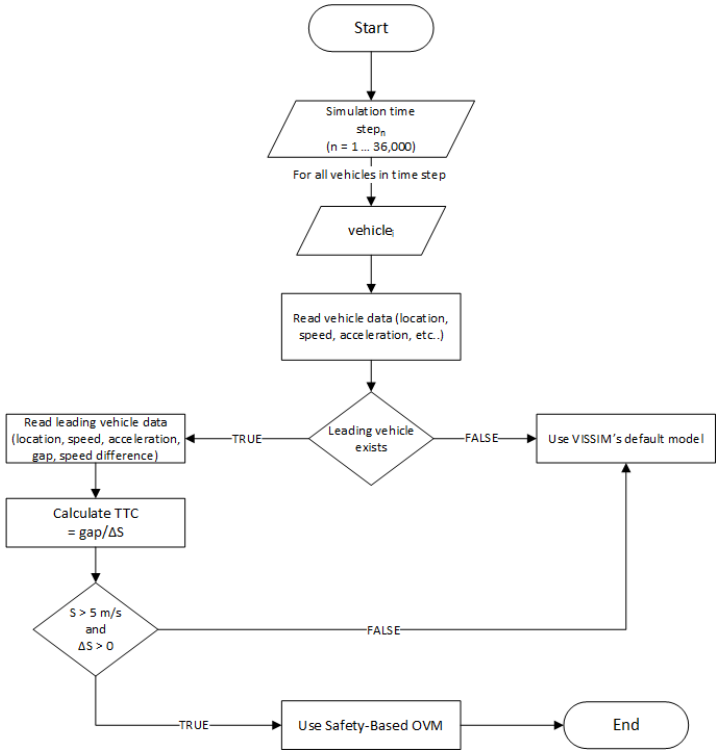


Figure 5.3: VISSIM driver behavior model selection.

$$v_{opt}(ttc)_{i,n} = v_o \frac{\tanh\left(\frac{ttc_{i,n}}{\Delta s} - \beta\right) + \tanh\,\beta}{1 + \tanh\,\beta} \tag{5.1}$$

$$\dot{v} = (1-\alpha)\frac{v_{opt_{i,n}} - v_{i,n}}{\tau} + \alpha f(ttc_{observed}) \tag{5.2}$$

Where:

$ttc_{i,n}$ = Instantaneous time-to-collision between vehicle$_i$ and preceding vehicle in a car-following episode at time step n (sec).

$v_{opt}(ttc)_{i,n}$ = The theoretical optimal velocity for vehicle i during time step n (km/hr).

v_o = Desired speed (km/hr), Δs = Transition width (sec).

β = Form Factor, $\dot{v_{i,n}}$ = Desired acceleration for vehicle i during time step n $(km/hr/sec)$.

$v_{i,n}$ = Actual speed of vehicle i during time step n (km/hr).

τ = Adaptation time $(simulation\ time\ steps)$.

The weighted observed acceleration, $\alpha f(ttc_{observed})$, allows our OVM to implicitly learn location-specific driving behavior characteristics from a subset of the observed driving behavior on-site to supplement the base OVM model's assumptions, while maintaining the benefits of interpretability of the remaining model parameters (e.g., the optimal model had a lower desired speed for vehicles executing turning movements compared to through moving vehicles, which is logical). We utilized VT-Lane's ability to classify the vehicles' movements across the intersection to calibrate different model parameters for through moving vehicles and vehicles executing turning movements at the intersection. The model calibration process and substantial improvements achieved are detailed in [2]. The calibrated model parameters for each vehicle movement type utilized in the external driver model for this study are detailed in Table 5.1. Algorithm 6 details the steps for the external driver model application during each simulation time step, where it is applied to all the vehicles in simulation.

Algorithm 6: VISSIM External Driver Behavior Model

Result: Desired acceleration of vehicles in VISSIM simulation.

initialization;

for $time\_step_n$ do

  for $vehicle_i$ do

    if $speed_{i,n} \leq 5\ (m/s)$ then

      $acceleration_{i,n} \leftarrow Default$;

    else

      if $vehicle_{i,n}\ \exists\ \Delta\ Distance\ \ \&\ \ \Longleftrightarrow\ \ \Delta\ Speed \in \mathbb{R}^+$ then

        $v_{opt_{i,n}} \leftarrow v_o \frac{\tanh\left(\frac{ttc_{i,n}}{\Delta s} - \beta\right)\ +\ \tanh\ \beta}{1 + \tanh\ \beta}$;

        $acceleration_{i,n} \leftarrow (1-\alpha)\frac{v_{opt_{i,n}} -\ speed_{i,n}}{\tau} + \alpha f(ttc_{observed})$;

      else

        $acceleration_{i,n} \leftarrow Default$;

      end

    end

  end

end

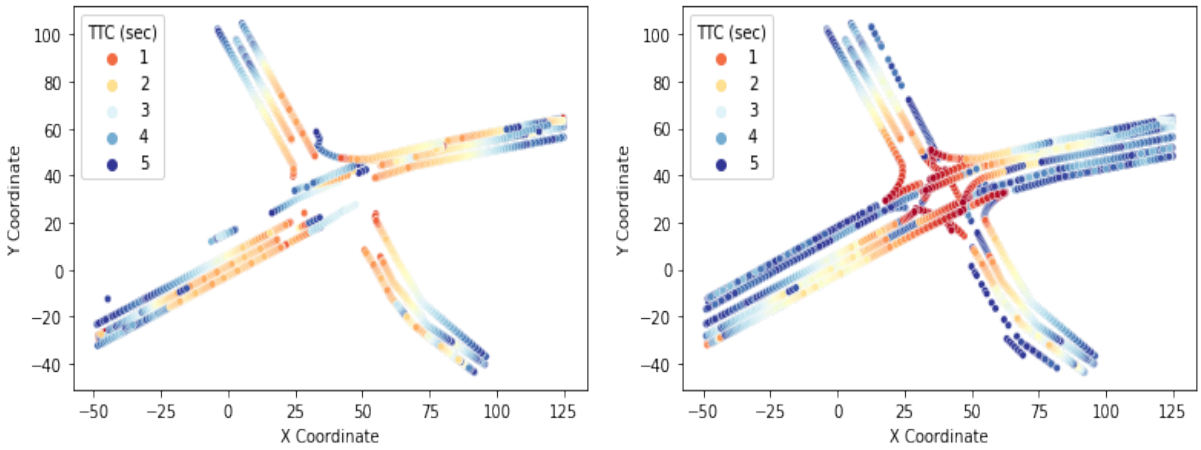
Table 5.1: Optimal Parameters For The Calibrated Safety-Based OVM

	Vehicle Movement	
	Through	Turning
Δs	4.63	4.76
β	1.10	0.10
τ	6.11	5.83
v_o	38.20	31.95
α	0.18	0.01

5.3 RESULTS AND ANALYSIS

5.3.1 Observed Vehicle Behavior and Safety Conflicts

A 1-hr simulation was conducted and the time-to-collision for vehicles involved in car-following episodes was calculated for comparison with the 1-hr data from video inference. Figure 5.4 illustrates the locations of safety conflicts and the associated average TTC for car-following instances where the TTC was ≤ 5 seconds.



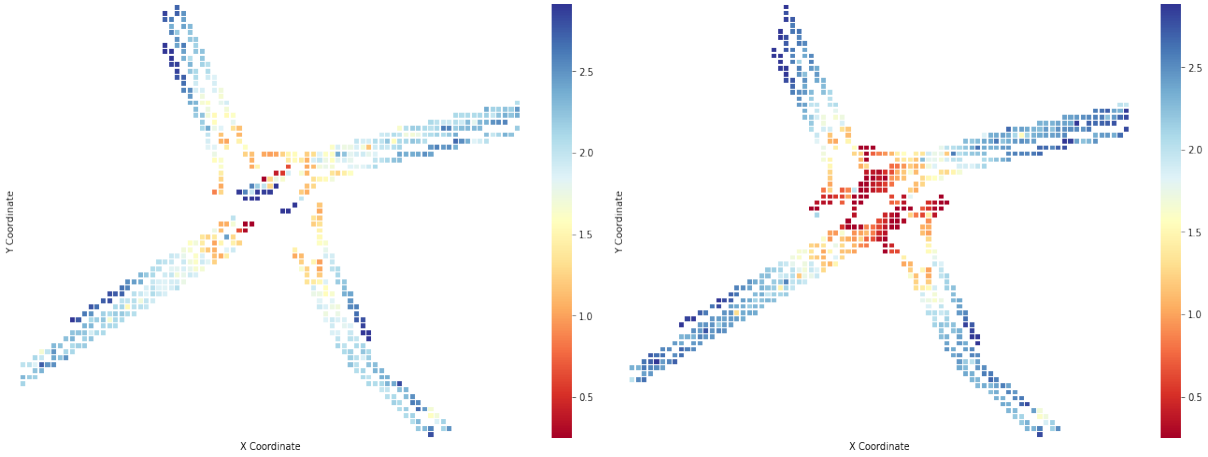
(a) VISSIM's base Wiedemann model. (b) Safety-Based Optimal Velocity Model.

Figure 5.4: Comparison of simulated episodes with low TTC (≤ 5 sec) for the intersection of study.

It can be clearly observed that VISSIM's base Wiedemann model is extremely conservative in generating safety conflicts during car-following episodes, with the majority of low TTCs occurring as vehicles approach the stop bars at the intersections' four approaches. The base VISSIM model generates minimal safety conflicts as vehicles traverse the intersection or downstream, as following vehicles seem to always be accelerating at a rate slower than the leading vehicles, leading to severe underestimation of safety conflicts as found by numerous studies in the literature. The Safety-Based OVM utilized in VISSIM as an external driver

model generates a more realistic driving behavior that would be expected at an urban intersection, especially as vehicles accelerate to clear the intersection and at the merge areas of through lanes with turning lanes as clearly illustrated in Figure 5.4 (b).

This distinction between VISSIM's default behavior and the external driver model is most evident when looking at safety-critical situations where the TTC between vehicles involved in a car-following episode momentarily drops to 1 second or less as can be seen in Figure 5.5, which illustrates the location of those instances and the average TTC associated with each location. Out of the 2667 generated in the 1-hour simulation, 31 vehicles (1.1%) had safety-critical car-following instances using VISSIM's default model, compared to 210 (7.8%) when utilizing the external driver model.



(a) VISSIM's base Wiedemann model. (b) Safety-Based Optimal Velocity Model.

Figure 5.5: Comparison of simulated episodes with safety-critical TTC ($\leq$ 1 sec) for the intersection of study.

5.3.2 Impact Assessment of Mitigation Measures

Based on this observed driving behavior, we assess the impact of different mitigation measures on reducing the number of generated conflicts. The area of study has a posted speed

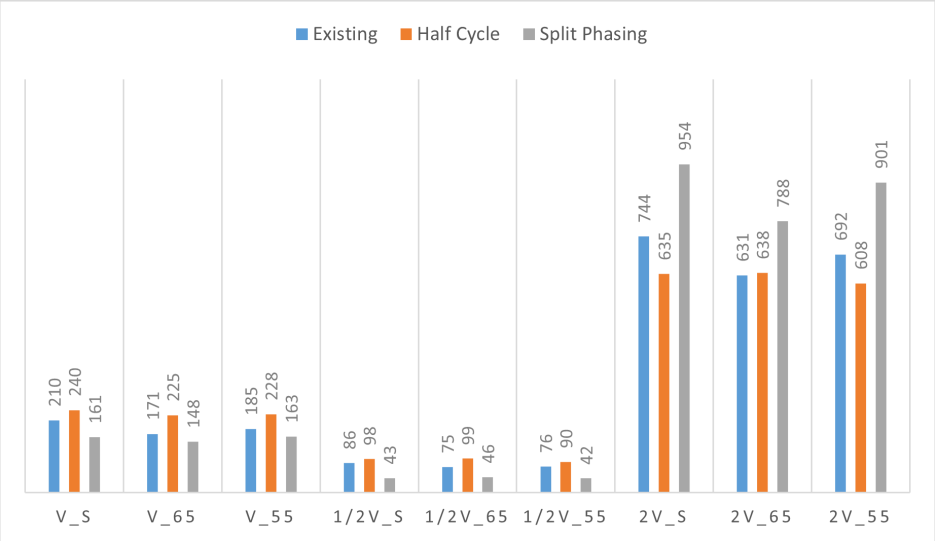
limit of 35 mph (55 km/hr). The observed speeds from video inference that were utilized in creating this simulation model had an 80th percentile speed of 35 mph (55 km/hr) and 95th percentile speed of 47 mph (75 km/hr), which is expected as there is no speed enforcement on-site and drivers typically drive above this posted speed limit. We assess the impact of enforcing speeds at 65 km/hr and 55 km/hr, respectively. Those enforcement measures are assessed with three scenarios of signal control, the current signal control previously illustrated in Figure 5.2, a similar control half-cycle length (70 seconds instead of 140), and a split-phasing signal control scenario. A sensitivity analysis for those enforcement and signal control scenarios is conducted for the observed, as well as half, and double the traffic volume. For brevity, the conventions for the different scenarios used hereafter are shown in Table 5.2. Signal control scenarios are labeled by name.

Table 5.2: Conventions for Volume/Speed States

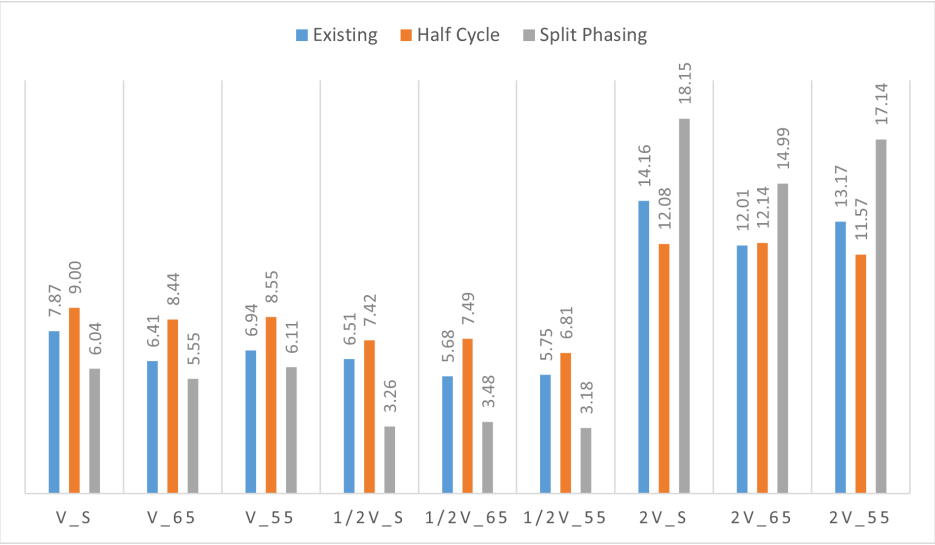
	Observed Speed	Enforce 65 km/hr	Enforce 55 km/hr
Observed Volume	V_S	V_65	V_55
Half Volume	1/2V_S	1/2V_65	1/2V_55
Double Volume	2V_S	2V_65	2V_55

Figure 5.6 shows the counts and percentages of car-following episodes with instances of safety-critical TTC ≤ 1 sec. For the existing volume and given turn movement classifications at the intersection, utilizing a split-phasing control was found to reduce the number of generated safety-critical conflicts by over 23%. Enforcing a 65 km/hr speed provides further but not substantial improvements. Split-phasing was also found to produce substantially better results in terms of safety in the scenario of lower traffic volumes. Conversely, it was the lowest performer in the case of double the observed volume.

Figures 5.7 to 5.9 show the heatmaps of those safety conflicts for the different simulation scenarios. The split-phasing which was found to significantly reduce the exiting safety conflicts can be seen in Figure 5.9 almost eliminating the conflicts generated at the turning



(a) Count of car-following episodes.



(b) Percentage of episodes relative to total volume.

Figure 5.6: Car-following episodes with safety-critical TTC ($\leq$ 1 sec) for the different simulated scenarios.

lanes. The extended queues for phases 2 and 6 due to split-phasing resulted in an extended shockwave as vehicles approach and stop at the intersection during a red light, hence it is to be avoided in case of an increased volume at this intersection.

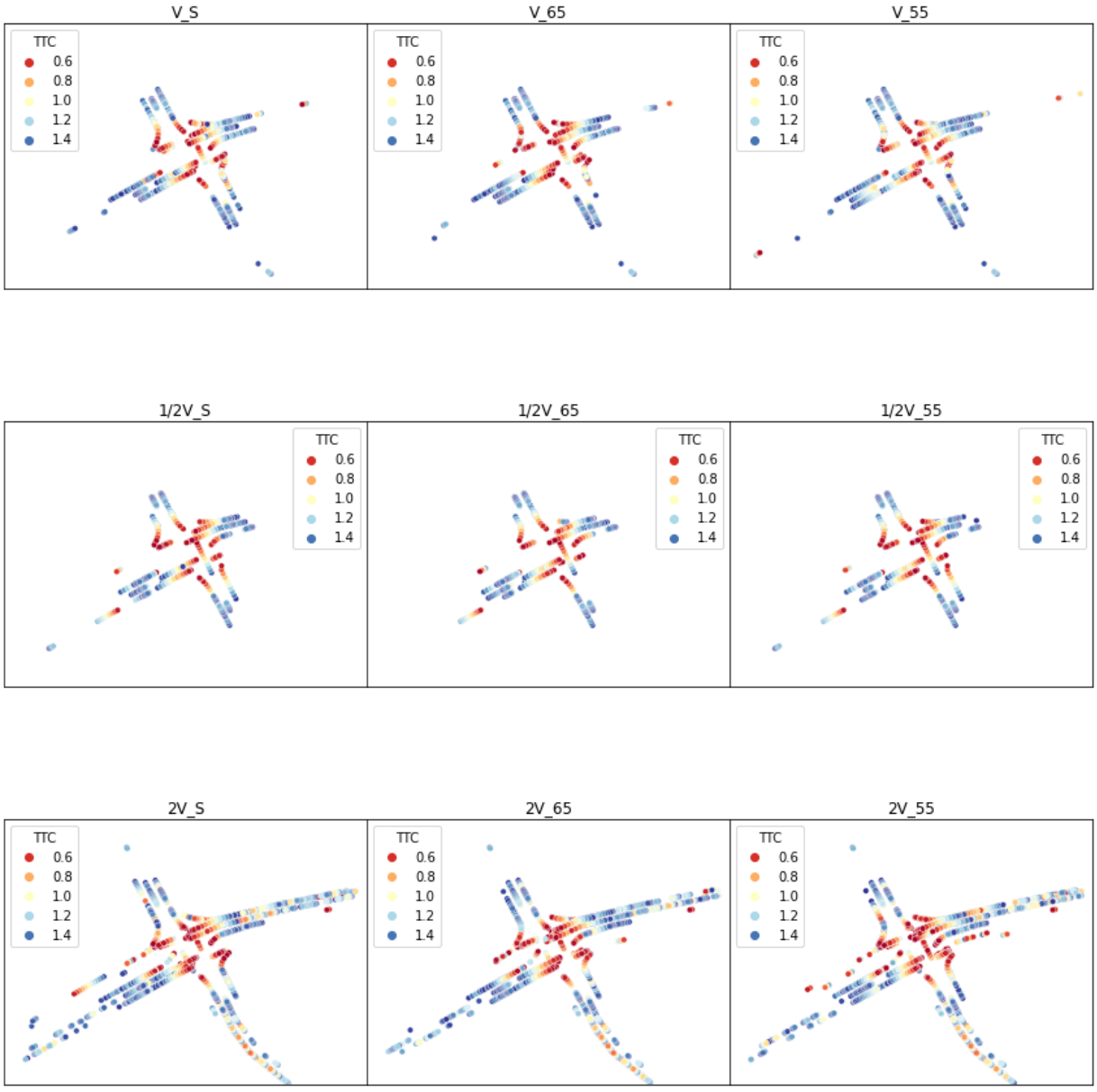


Figure 5.7: Current signal control.

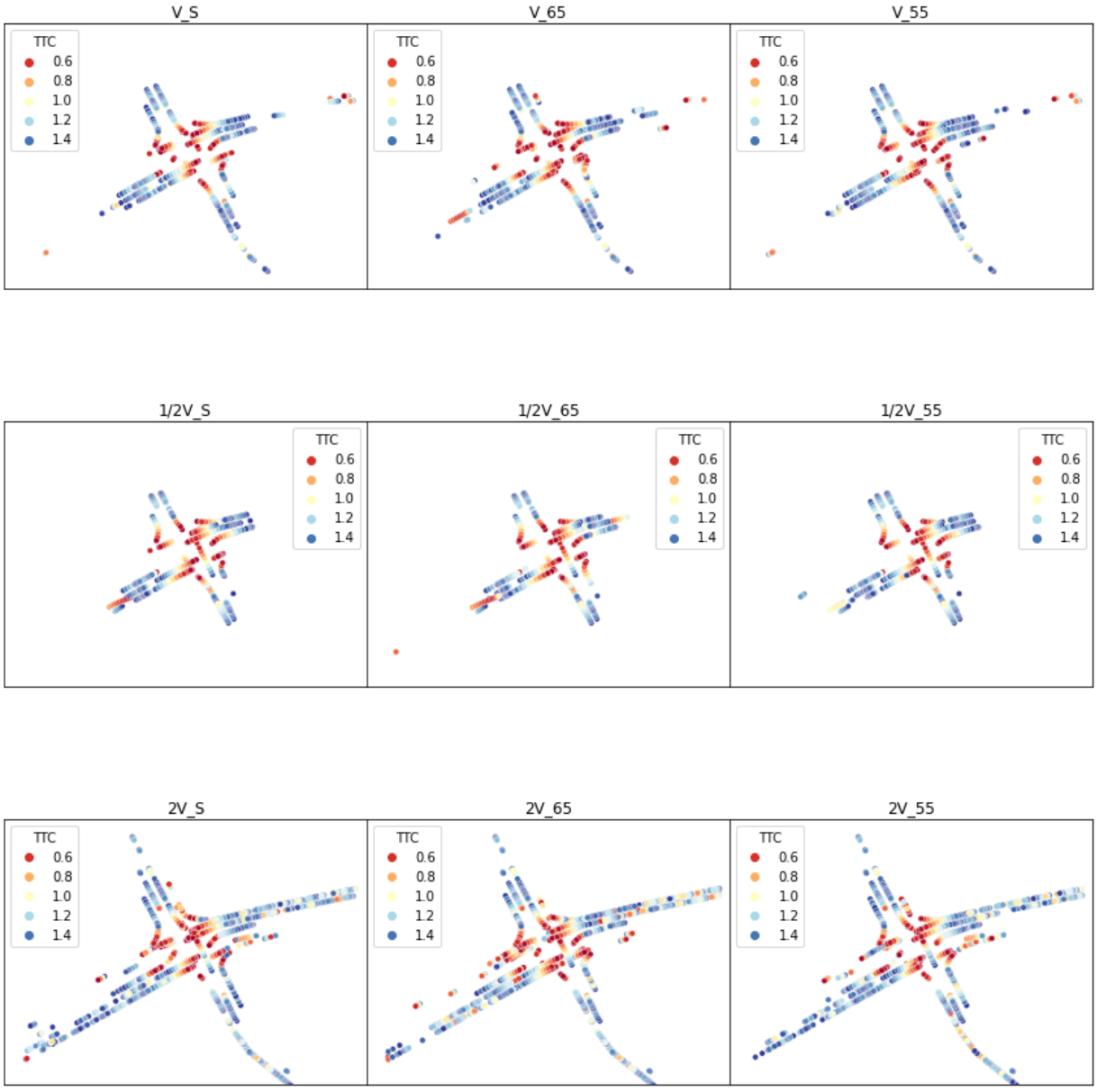


Figure 5.8: Half cycle length.

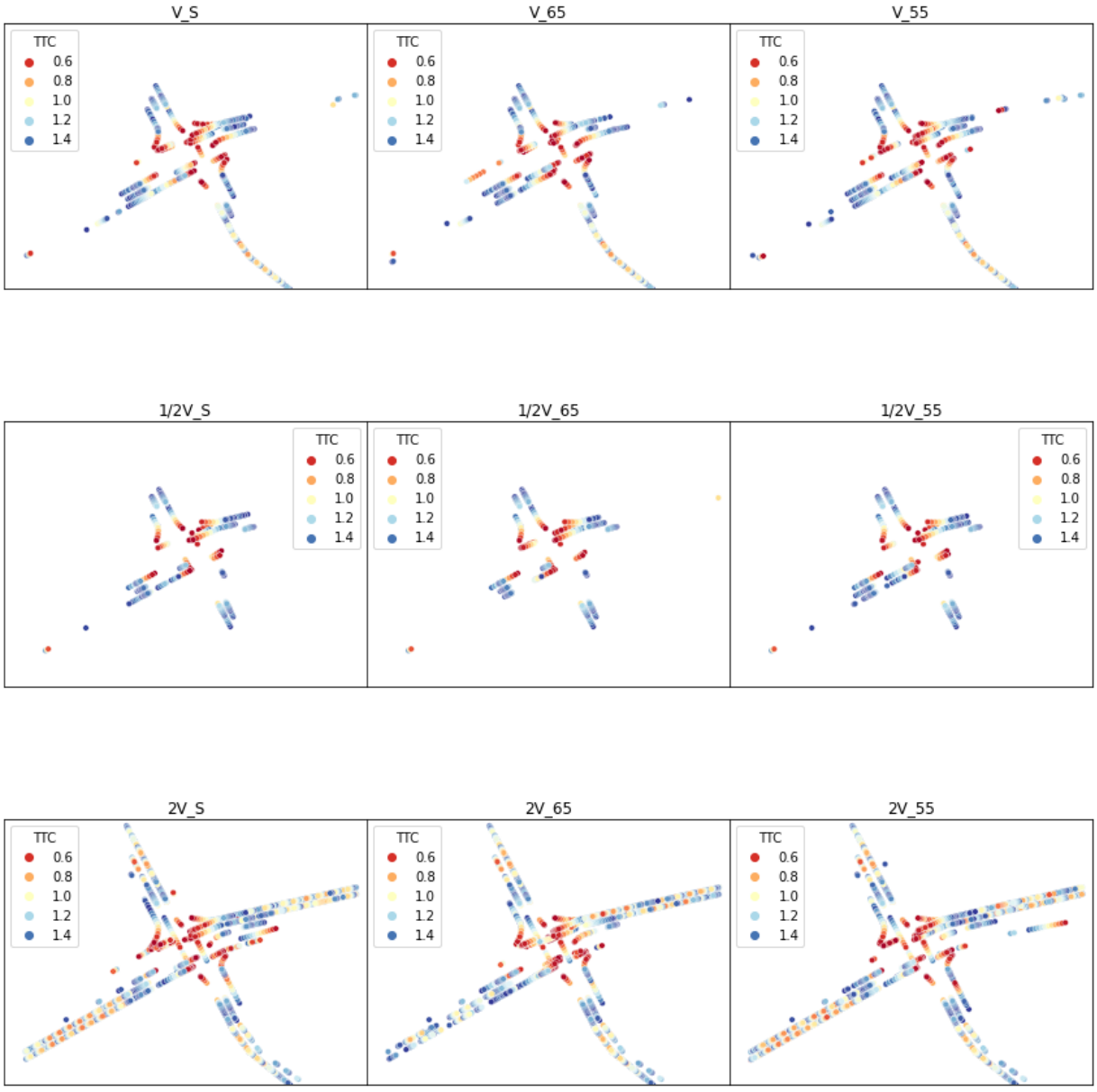


Figure 5.9: Split-phasing.

5.4 CONCLUSIONS

This study provided an assessment of utilizing a data-driven driver behavior model to provide high-fidelity simulation of safety-critical behavior. A safety-based Optimal Velocity Model that was developed and calibrated for the location of study in Blacksburg, VA was implemented in VISSIM as an external driver model. The resulting safety conflicts were compared to that of VISSIM's default Wiedemann model, and the results showcase the substantial improvements that can be achieved in simulation utilizing a data-driven driver behavior model in terms of replicating the existing safety concerns. We utilized this improved behavior modeling to quantify the impact of different signal control and speed enforcement scenarios in mitigating the simulated safety concerns. We concluded this study by providing recommendations for appropriate signal control and speed enforcement measures to reduce the simulated conflicts at the intersection across varying traffic volumes.

Unlike existing studies in the literature, this assessment did not involve a process of calibrating VISSIM's model parameters to replicate the expected safety concerns at the area of study. Instead, the safety-based OVM utilized as an external driver model in this study was calibrated from video inference, and the direct implementation of that calibrated model resulted in a highly accurate representation of the status-quo at the area of study. The results highlight the value of this approach for video inference-based driver behavior modeling which provides a highly scalable, automated, and generalizable workflow and aids in better replicating the observed behavior in microsimulation. The results obtained and discussed in this study do not take into account the trade-off between capacity and operational safety, which could be the subject of future works. Future studies will also assess the scalability and transferability of this approach, which could provide traffic safety practitioners with a decision-support tool for evaluating the safety impact of different traffic demand management, signal control, and enforcement measures.

Bibliography

[1] R. Wiedemann, "Modelling of rti-elements on multi-lane roads," in *Drive Conference (1991: Brussels, Belgium)*, vol. 2, 1991.

[2] A. Abdelhalim and M. Abbas, "A real-time safety-based optimal velocity model," *In Review, IEEE Open Journal of Intelligent Transportation Systems*, 2021.

[3] D. Gettman, L. Pu, T. Sayed, S. G. Shelby, S. Energy, *et al.*, "Surrogate safety assessment model and validation," tech. rep., Turner-Fairbank Highway Research Center, 2008.

[4] L. Vasconcelos, L. Neto, Á. M. Seco, and A. B. Silva, "Validation of the surrogate safety assessment model for assessment of intersection safety," *Transportation Research Record*, vol. 2432, no. 1, pp. 1–9, 2014.

[5] F. Huang, P. Liu, H. Yu, and W. Wang, "Identifying if vissim simulation model and ssam provide reasonable estimates for field measured traffic conflicts at signalized intersections," *Accident Analysis & Prevention*, vol. 50, pp. 1014–1024, 2013.

[6] A. Dijkstra, P. Marchesini, F. Bijleveld, V. Kars, H. Drolenga, and M. Van Maarseveen, "Do calculated conflicts in microsimulation model predict number of crashes?," *Transportation research record*, vol. 2147, no. 1, pp. 105–112, 2010.

[7] M. Fellendorf and P. Vortisch, "Validation of the microscopic traffic flow model vissim in different real-world situations," in *transportation research board 80th annual meeting*, vol. 11, 2001.

[8] N. E. Lownes and R. B. Machemehl, "Vissim: a multi-parameter sensitivity analysis," in *Proceedings of the 2006 Winter Simulation Conference*, pp. 1406–1413, IEEE, 2006.

[9] N. E. Lownes and R. B. Machemehl, "Sensitivity of simulated capacity to modification of vissim driver behavior parameters," *Transportation Research Record*, vol. 1988, no. 1, pp. 102–110, 2006.

[10] P. Manjunatha, P. Vortisch, and T. V. Mathew, "Methodology for the calibration of vissim in mixed traffic," in *Transportation research board 92nd annual meeting*, vol. 11, Transportation Research Board Washington, DC, United States, 2013.

[11] M. Arafat, S. R. Nafis, E. Sadeghvaziri, and F. Tousif, "A data-driven approach to calibrate microsimulation models based on the degree of saturation at signalized intersections," *Transportation Research Interdisciplinary Perspectives*, vol. 8, p. 100231, 2020.

[12] R. Fan, H. Yu, P. Liu, and W. Wang, "Using vissim simulation model and surrogate safety assessment model for estimating field measured traffic conflicts at freeway merge areas," *IET Intelligent Transport Systems*, vol. 7, no. 1, pp. 68–77, 2013.

[13] M. Essa and T. Sayed, "Transferability of calibrated microsimulation model parameters for safety assessment using simulated conflicts," *Accident Analysis & Prevention*, vol. 84, pp. 41–53, 2015.

[14] M. Essa and T. Sayed, "Simulated traffic conflicts: do they accurately represent field-measured conflicts?," *Transportation research record*, vol. 2514, no. 1, pp. 48–57, 2015.

[15] M. Essa and T. Sayed, "Comparison between surrogate safety assessment model and real-time safety models in predicting field-measured conflicts at signalized intersections," *Transportation research record*, vol. 2674, no. 3, pp. 100–112, 2020.

[16] I. I. Otković, A. Deluka-Tibljaš, and S. Šurdonja, "Validation of the calibration methodology of the micro-simulation traffic model," *Transportation Research Procedia*, vol. 45, pp. 684–691, 2020.

[17] H. Naing, W. Cai, N. Hu, T. Wu, and L. Yu, "Data-driven microscopic traffic modelling and simulation using dynamic lstm," in *Proceedings of the 2021 ACM SIGSIM Conference on Principles of Advanced Discrete Simulation*, pp. 1–12, 2021.

[18] A. Abdelhalim and M. Abbas, "Vt-lane: An exploratory study of an ad-hoc framework for real-time intersection turn count and trajectory reconstruction using nema phases-based virtual traffic lanes," in *2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC)*, pp. 1–6, IEEE, 2020.

[19] A. Abdelhalim and M. Abbas, "Towards real-time traffic movement count and trajectory reconstruction using virtual traffic lanes," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pp. 592–593, 2020.

[20] A. Abdelhalim, M. Abbas, B. B. Kotha, and A. Wicks, "A framework for real-time traffic trajectory tracking, speed estimation, and driver behavior calibration at urban intersections using virtual traffic lanes," 2021.

Chapter 6

Discussion, Conclusions, and Future Work

In this dissertation, the author proposes VT-Lane, a framework incorporating the prior knowledge of NEMA traffic movement patterns, alongside laws of traffic flow and continuity, in a video-based trajectory tracking framework to obtain reliable turn counts, assist in addressing the vehicle identity switch problem due to occlusion, and aid in reconstructing missing vehicle trajectories. The proposed framework is evaluated and validated for the tasks of movement tracking and classification through a case study of manually collected data and utilizing the AI City benchmark dataset. The task of accurate speed estimation is then evaluated using an instrumented vehicle equipped with a state-of-the-industry sensor suite. The validated high-granularity trajectories are then utilized for automated large-scale safety-based driving behavior calibration for the area of study. This calibrated safety-based driver behavior model is then utilized in microsimulation to showcase the improvements that can be achieved using this approach for the assessment of safety interventions.

The results obtained for the validation of VT-Lane show very promising potential for obtaining reliable trajectories and speed profiles for each movement, which can aid in improving conflict monitoring, incident avoidance, dynamic signal timing, as well as O-D estimation applications. In the local case study using the 1-hour of video sequences, 69% of the identity switches incurred by the baseline tracker were resolved. The K-NN search in the VT-Lane framework is very time-efficient, running up to 100 HZ in low traffic and between 60-70 HZ in high traffic (on a 2.4 GHz 16-core Intel Xeon e5 2630 v3 processor), a minimal increase

to the running time of a multi-object tracking network. The results obtained in the AI City Challenge further demonstrated the flexibility, scalability, and efficiency of the proposed framework.

The estimated speeds via our framework were verified using high-granularity data obtained from a vehicle instrumented with a state-of-the-industry sensor suite that was tracked as it was driven through the intersection of the case study. The results of the speed validation show that VT-Lane can estimate speeds in real-time with an error of 0.19 m/sec for estimating the average travel speed of detected vehicles, which is equivalent to 2% of the observed average travel speed (9.52 m/sec) through the intersection of study. Instantaneous vehicle speeds at the resolution of 30 Hz were found to be estimated with an average error of 0.21 m/sec and 0.86 m/sec for free-flowing and congested traffic conditions, respectively, with an overall R^2 of 93%. The high-fidelity trajectories and speed profiles can be utilized to provide a real-time spatio-temporal assessment of crash risk via the estimation and monitoring of the time to collision, providing practitioners with a multitude of high-quality performance and safety-surrogate measures for urban traffic intersections.

Utilizing the ability of the VT-Lane framework to efficiently extract accurate vehicle trajectories, movement classification, and speed estimates, the parameters of a safety-based Optimal Velocity Model in terms of the distance the instantaneous TTC estimated from video inference. The proposed TTC-Based OVM was able to produce improved deceleration estimates for vehicles engaged in car-following episodes and resulted in speed profiles with an instantaneous estimated speed error of 4.74 km/hr and an instantaneous acceleration error 19.84% lower than that of the base, distance-based model. The results obtained utilizing real-time data and through the proposed modified acceleration function based on observed behavior show the high value of infusing driving behavior models with data-driven modifications.

Finally, the safety-based OVM was incorporated in VISSIM, and utilized in a case study

to assess appropriate safety interventions. The resulting safety conflicts were compared to that of VISSIM's default Wiedemann model, showcasing the substantial improvements that can be achieved in simulation utilizing a data-driven driver behavior model in terms of replicating the existing safety concerns. This improved behavior modeling was utilized to quantify the impact of different signal control and speed enforcement scenarios in mitigating the simulated safety concerns, providing traffic safety practitioners with a decision-support tool for evaluating the safety impact of different traffic demand management, signal control, and enforcement measures.

The short length of the car-following episodes, given that vehicles are only tracked as they cross an intersection, was a limiting factor for this work. Although vehicles engaged in car-following within the intersection are expected to be engaging in car-following behavior before and/or after the monitored segment, that is merely the segment where the trajectories are extracted. Future studies will be conducted to assess VT-Lane on longer arterial segments and evaluate leader-follower interactions through longer periods. Studying this behavior within those limits, however, was sufficient to capture the driving behavior via the proposed safety-based model and replicate it in simulation. The clear differences in driving behavior model parameters between straight moving vehicles and vehicles executing turning movements at the intersection demonstrate the need for taking into account such characteristics in modeling driving behavior in urban settings. While 3rd order polynomial function was able to capture the observed driving behavior with high reliability for the presented case study (R^2 of 70%), further information that can be obtained through video inference (including, and not limited to signal status time, vehicle type, overall traffic density, and lane occupancy, etc.) can all be utilized to develop more sophisticated observed driving behavior functions that would potentially further improve the performance the data-driven model. Future work will also include investigating and data-mining the relationship between

the observed historic crashes for an area of study and the safety assessments produced via VT-Lane, both in terms of real-time inferred measures and through simulation.

Bibliography

[1] W. H. Organization, "Road traffic deaths," 2015. https://www.who.int/gho/road_safety/mortality/en/.

[2] F. H. Administration, "The national intersection safety problem," tech. rep., 2009.

[3] V. G. Kovvali, V. Alexiadis, P. Zhang, *et al.*, "Video-based vehicle trajectory data collection," tech. rep., 2007.

[4] U. FHWA, "Traffic signal timing manual," 2008.

[5] P. Felzenszwalb, D. McAllester, and D. Ramanan, "A discriminatively trained, multiscale, deformable part model," in *2008 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–8, IEEE, 2008.

[6] M. C. Burl, M. Weber, and P. Perona, "A probabilistic approach to object recognition using local photometry and global geometry," in *European conference on computer vision*, pp. 628–641, Springer, 1998.

[7] D. Crandall, P. Felzenszwalb, and D. Huttenlocher, "Spatial priors for part-based recognition using statistical models," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, vol. 1, pp. 10–17, IEEE, 2005.

[8] R. Girshick, "Fast r-cnn," in *Proceedings of the IEEE international conference on computer vision*, pp. 1440–1448, 2015.

[9] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," in *Advances in neural information processing systems*, pp. 91–99, 2015.

[10] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 779–788, 2016.

[11] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "Ssd: Single shot multibox detector," in *European conference on computer vision*, pp. 21–37, Springer, 2016.

[12] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask r-cnn," in *Proceedings of the IEEE international conference on computer vision*, pp. 2961–2969, 2017.

[13] C.-Y. Fu, W. Liu, A. Ranga, A. Tyagi, and A. C. Berg, "Dssd: Deconvolutional single shot detector," *arXiv preprint arXiv:1701.06659*, 2017.

[14] J. Redmon and A. Farhadi, "Yolov3: An incremental improvement," *arXiv preprint arXiv:1804.02767*, 2018.

[15] T. Huang, D. Koller, J. Malik, G. Ogasawara, B. S. Rao, S. J. Russell, and J. Weber, "Automatic symbolic traffic scene analysis using belief networks," in *AAAI*, vol. 94, pp. 966–972, 1994.

[16] K. Karmann and A. Brandt, "Moving object recognition using an adaptive background memory. cappellini v, ed. time-varying image processing and moving object recognition," 1990.

[17] D. Koller, K. Daniilidis, and H.-H. Nagel, "Model-based object tracking in monocular image sequences of road traffic scenes," *International Journal of Computer 11263on*, vol. 10, no. 3, pp. 257–281, 1993.

[18] D. Beymer and J. Malik, "Tracking vehicles in congested traffic," in *Proceedings of Conference on Intelligent Vehicles*, pp. 130–135, IEEE, 1996.

[19] D. Beymer, P. McLauchlan, B. Coifman, and J. Malik, "A real-time computer vision system for measuring traffic parameters," in *Proceedings of IEEE computer society conference on computer vision and pattern recognition*, pp. 495–501, IEEE, 1997.

[20] S. Kamijo, Y. Matsushita, K. Ikeuchi, and M. Sakauchi, "Traffic monitoring and accident detection at intersections," *IEEE transactions on Intelligent transportation systems*, vol. 1, no. 2, pp. 108–118, 2000.

[21] N. Saunier and T. Sayed, "A feature-based tracking algorithm for vehicles in intersections," in *The 3rd Canadian Conference on Computer and Robot Vision (CRV'06)*, pp. 59–59, IEEE, 2006.

[22] A. Sanchez, P. D. Suarez, A. Conci, and E. Nunes, "Video-based distance traffic analysis: Application to vehicle tracking and counting," *Computing in Science & Engineering*, vol. 13, no. 3, pp. 38–45, 2010.

[23] Z. Dai, H. Song, X. Wang, Y. Fang, X. Yun, Z. Zhang, and H. Li, "Video-based vehicle counting framework," *IEEE Access*, vol. 7, pp. 64460–64470, 2019.

[24] C.-E. Wu, W.-Y. Yang, H.-C. Ting, and J.-S. Wang, "Traffic pattern modeling, trajectory classification and vehicle tracking within urban intersections," in *2017 International Smart Cities Conference (ISC2)*, pp. 1–6, IEEE, 2017.

[25] M. S. Shirazi and B. T. Morris, "Trajectory prediction of vehicles turning at intersections using deep neural networks," *Machine Vision and Applications*, vol. 30, no. 6, pp. 1097–1109, 2019.

[26] L. Zhu, S. Wang, C. Li, and Z. Yang, "License plate recognition in urban road based on vehicle tracking and result integration," *Journal of Intelligent Systems*, vol. 29, no. 1, pp. 1587–1597, 2019.

[27] T. K. Cheang, Y. S. Chong, and Y. H. Tay, "Segmentation-free vehicle license plate recognition using convnet-rnn," *arXiv preprint arXiv:1701.06439*, 2017.

[28] H. Zhang, M. Liptrott, N. Bessis, and J. Cheng, "Real-time traffic analysis using deep learning techniques and uav based video," in *2019 16th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, pp. 1–5, IEEE, 2019.

[29] Z. Tang, M. Naphade, M.-Y. Liu, X. Yang, S. Birchfield, S. Wang, R. Kumar, D. Anastasiu, and J.-N. Hwang, "Cityflow: A city-scale benchmark for multi-target multi-camera vehicle tracking and re-identification," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 8797–8806, 2019.

[30] L. Wen, D. Du, Z. Cai, Z. Lei, M.-C. Chang, H. Qi, J. Lim, M.-H. Yang, and S. Lyu, "Ua-detrac: A new benchmark and protocol for multi-object detection and tracking," *Computer Vision and Image Understanding*, p. 102907, 2020.

[31] M.-C. Chang, C.-K. Chiang, C.-M. Tsai, Y.-K. Chang, H.-L. Chiang, Y.-A. Wang, S.-Y. Chang, Y.-L. Li, M.-S. Tsai, and H.-Y. Tseng, "Ai city challenge 2020-computer vision for smart transportation applications," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pp. 620–621, 2020.

[32] M. Forghani, F. Karimipour, and C. Claramunt, "From cellular positioning data to trajectories: Steps towards a more accurate mobility exploration," *Transportation Research Part C: Emerging Technologies*, vol. 117, p. 102666, 2020.

[33] A. H. Ghods and L. Fu, "Real-time estimation of turning movement counts at signalized intersections using signal phase information," *Transportation Research Part C: Emerging Technologies*, vol. 47, pp. 128–138, 2014.

[34] S. Indu, M. Gupta, and A. Bhattacharyya, "Vehicle tracking and speed estimation

using optical flow method," *Int. J. Engineering Science and Technology*, vol. 3, no. 1, pp. 429–434, 2011.

[35] S. Doğan, M. S. Temiz, and S. Külür, "Real time speed estimation of moving vehicles from side view images from an uncalibrated video camera," *Sensors*, vol. 10, no. 5, pp. 4805–4824, 2010.

[36] K. Osamura, A. Yumoto, and O. Nakayama, "Vehicle speed estimation using video data and acceleration information of a drive recorder," in *2013 13th International Conference on ITS Telecommunications (ITST)*, pp. 157–162, IEEE, 2013.

[37] L. R. Costa, M. S. Rauen, and A. B. Fronza, "Car speed estimation based on image scale factor," *Forensic science international*, vol. 310, p. 110229, 2020.

[38] M. Garg and S. Goel, "Real-time license plate recognition and speed estimation from video sequences," *ITSI Transactions on Electrical and Electronics Engineering*, vol. 1, no. 5, pp. 1–4, 2013.

[39] D. C. Luvizon, B. T. Nassu, and R. Minetto, "A video-based system for vehicle speed measurement in urban roadways," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 6, pp. 1393–1404, 2016.

[40] H. Dong, M. Wen, and Z. Yang, "Vehicle speed estimation based on 3d convnets and non-local blocks," *Future Internet*, vol. 11, no. 6, p. 123, 2019.

[41] T. Huang, "Traffic speed estimation from surveillance video data," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 161–165, 2018.

[42] Z. Liu, W. Zhang, X. Gao, H. Meng, X. Tan, X. Zhu, Z. Xue, X. Ye, H. Zhang, S. Wen, *et al.*, "Robust movement-specific vehicle counting at crowded intersections," in *Pro-*

ceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, pp. 614–615, 2020.

[43] Y. Wan, Y. Huang, and B. Buckles, "Camera calibration and vehicle tracking: Highway traffic video analytics," *Transportation Research Part C: Emerging Technologies*, vol. 44, pp. 202–213, 2014.

[44] G. Guido, A. Vitale, and V. Gallelli, "Investigating rear-end potential conflicts for roundabout turning movements," *Modern Traffic and Transportation Engineering Research (MTTER)*, vol. 2, no. 4, pp. 20–26, 2013.

[45] C. Xu, A. P. Tarko, W. Wang, and P. Liu, "Predicting crash likelihood and severity on freeways with real-time loop detector data," *Accident Analysis & Prevention*, vol. 57, pp. 30–39, 2013.

[46] P. St-Aubin, L. Miranda-Moreno, and N. Saunier, "An automated surrogate safety analysis at protected highway ramps using cross-sectional and before–after video data," *Transportation Research Part C: Emerging Technologies*, vol. 36, pp. 284–295, 2013.

[47] P. St-Aubin, N. Saunier, and L. F. Miranda-Moreno, "Comparison of various time-to-collision prediction and aggregation methods for surrogate safety analysis," tech. rep., 2015.

[48] P. St-Aubin, N. Saunier, and L. Miranda-Moreno, "Large-scale automated proactive road safety analysis using video data," *Transportation Research Part C: Emerging Technologies*, vol. 58, pp. 363–379, 2015.

[49] K. Xie, C. Li, K. Ozbay, G. Dobler, H. Yang, A.-T. Chiang, and M. Ghandehari, "Development of a comprehensive framework for video-based safety assessment," in *2016*

IEEE 19th International Conference on Intelligent Transportation Systems (ITSC), pp. 2638–2643, IEEE, 2016.

[50] K. Xie, K. Ozbay, H. Yang, and C. Li, "Mining automatically extracted vehicle trajectory data for proactive safety analytics," *Transportation research part C: emerging technologies*, vol. 106, pp. 61–72, 2019.

[51] S. Das and A. K. Maurya, "Defining time-to-collision thresholds by the type of lead vehicle in non-lane-based traffic environments," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 12, pp. 4972–4982, 2019.

[52] M. Kilicarslan and J. Y. Zheng, "Predict vehicle collision by ttc from motion using a single video camera," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 2, pp. 522–533, 2018.

[53] C. Wang, C. Xu, and Y. Dai, "A crash prediction method based on bivariate extreme value theory and video-based vehicle trajectory data," *Accident Analysis & Prevention*, vol. 123, pp. 365–373, 2019.

[54] L. Li, R. Jiang, Z. He, X. M. Chen, and X. Zhou, "Trajectory data-based traffic flow studies: A revisit," *Transportation Research Part C: Emerging Technologies*, vol. 114, pp. 225–240, 2020.

[55] M. Brackstone and M. McDonald, "Car-following: a historical review," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 2, no. 4, pp. 181–196, 1999.

[56] V. Punzo, M. T. Borzacchiello, and B. Ciuffo, "On the assessment of vehicle trajectory data accuracy and application to the next generation simulation (ngsim) program data," *Transportation Research Part C: Emerging Technologies*, vol. 19, no. 6, pp. 1243–1262, 2011.

[57] S. H. Hamdar, L. Qin, and A. Talebpour, "Weather and road geometry impact on longitudinal driving behavior: Exploratory analysis using an empirically supported acceleration modeling framework," *Transportation research part C: emerging technologies*, vol. 67, pp. 193–213, 2016.

[58] G. S. Aoude, V. R. Desaraju, L. H. Stephens, and J. P. How, "Driver behavior classification at intersections and validation on large naturalistic data set," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 2, pp. 724–736, 2012.

[59] P. St-Aubin, N. Saunier, L. F. Miranda-Moreno, and K. Ismail, "Use of computer vision data for detailed driver behavior analysis and trajectory interpretation at roundabouts," *Transportation research record*, vol. 2389, no. 1, pp. 65–77, 2013.

[60] A. Talebpour, H. S. Mahmassani, and F. E. Bustamante, "Modeling driver behavior in a connected environment: Integrated microscopic simulation of traffic and mobile wireless telecommunication systems," *Transportation Research Record*, vol. 2560, no. 1, pp. 75–86, 2016.

[61] L. Chong, M. M. Abbas, A. M. Flintsch, and B. Higgs, "A rule-based neural network approach to model driver naturalistic behavior in traffic," *Transportation Research Part C: Emerging Technologies*, vol. 32, pp. 207–223, 2013.

[62] J. Morton, T. A. Wheeler, and M. J. Kochenderfer, "Analysis of recurrent neural networks for probabilistic modeling of driver behavior," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 5, pp. 1289–1298, 2016.

[63] T. A. Wheeler, P. Robbel, and M. J. Kochenderfer, "Analysis of microscopic behavior models for probabilistic modeling of driver behavior," in *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, pp. 1604–1609, IEEE, 2016.

[64] M. Bando, K. Hasebe, A. Nakayama, A. Shibata, and Y. Sugiyama, "Dynamical model of traffic congestion and numerical simulation," *Physical review E*, vol. 51, no. 2, p. 1035, 1995.

[65] H. Wang, Y. Li, W. Wang, M. Fu, and R. Huang, "Optimal velocity model with dual boundary optimal velocity function," *Transportmetrica B: transport dynamics*, vol. 5, no. 2, pp. 211–227, 2017.

[66] S. Mammar, S. Mammar, and H. Haj-Salem, "A modified optimal velocity model for vehicle following," *IFAC Proceedings Volumes*, vol. 38, no. 1, pp. 120–125, 2005.

[67] H. Lazar, K. Rhoulami, and M. D. Rahmani, "A modified full velocity difference model based on time to collision as a safely indicator for braking state," in *2015 2nd World Symposium on Web Applications and Networking (WSWAN)*, pp. 1–6, IEEE, 2015.

[68] H. Lazar, K. Rhoulami, and D. Rahmani, "A review analysis of optimal velocity models," *Periodica Polytechnica Transportation Engineering*, vol. 44, no. 2, pp. 123–131, 2016.

[69] S. Ossen and S. P. Hoogendoorn, "Car-following behavior analysis from microscopic trajectory data," *Transportation Research Record*, vol. 1934, no. 1, pp. 13–21, 2005.

[70] D. Gettman, L. Pu, T. Sayed, S. G. Shelby, S. Energy, *et al.*, "Surrogate safety assessment model and validation," tech. rep., Turner-Fairbank Highway Research Center, 2008.

[71] L. Vasconcelos, L. Neto, Á. M. Seco, and A. B. Silva, "Validation of the surrogate safety assessment model for assessment of intersection safety," *Transportation Research Record*, vol. 2432, no. 1, pp. 1–9, 2014.

[72] F. Huang, P. Liu, H. Yu, and W. Wang, "Identifying if vissim simulation model and ssam provide reasonable estimates for field measured traffic conflicts at signalized intersections," *Accident Analysis & Prevention*, vol. 50, pp. 1014–1024, 2013.

[73] A. Dijkstra, P. Marchesini, F. Bijleveld, V. Kars, H. Drolenga, and M. Van Maarseveen, "Do calculated conflicts in microsimulation model predict number of crashes?," *Transportation research record*, vol. 2147, no. 1, pp. 105–112, 2010.

[74] M. Fellendorf and P. Vortisch, "Validation of the microscopic traffic flow model vissim in different real-world situations," in *transportation research board 80th annual meeting*, vol. 11, 2001.

[75] N. E. Lownes and R. B. Machemehl, "Vissim: a multi-parameter sensitivity analysis," in *Proceedings of the 2006 Winter Simulation Conference*, pp. 1406–1413, IEEE, 2006.

[76] N. E. Lownes and R. B. Machemehl, "Sensitivity of simulated capacity to modification of vissim driver behavior parameters," *Transportation Research Record*, vol. 1988, no. 1, pp. 102–110, 2006.

[77] P. Manjunatha, P. Vortisch, and T. V. Mathew, "Methodology for the calibration of vissim in mixed traffic," in *Transportation research board 92nd annual meeting*, vol. 11, Transportation Research Board Washington, DC, United States, 2013.

[78] M. Arafat, S. R. Nafis, E. Sadeghvaziri, and F. Tousif, "A data-driven approach to calibrate microsimulation models based on the degree of saturation at signalized intersections," *Transportation Research Interdisciplinary Perspectives*, vol. 8, p. 100231, 2020.

[79] R. Fan, H. Yu, P. Liu, and W. Wang, "Using vissim simulation model and surrogate

safety assessment model for estimating field measured traffic conflicts at freeway merge areas," *IET Intelligent Transport Systems*, vol. 7, no. 1, pp. 68–77, 2013.

[80] M. Essa and T. Sayed, "Transferability of calibrated microsimulation model parameters for safety assessment using simulated conflicts," *Accident Analysis & Prevention*, vol. 84, pp. 41–53, 2015.

[81] M. Essa and T. Sayed, "Simulated traffic conflicts: do they accurately represent field-measured conflicts?," *Transportation research record*, vol. 2514, no. 1, pp. 48–57, 2015.

[82] M. Essa and T. Sayed, "Comparison between surrogate safety assessment model and real-time safety models in predicting field-measured conflicts at signalized intersections," *Transportation research record*, vol. 2674, no. 3, pp. 100–112, 2020.

[83] I. I. Otković, A. Deluka-Tibljaš, and S. Šurdonja, "Validation of the calibration methodology of the micro-simulation traffic model," *Transportation Research Procedia*, vol. 45, pp. 684–691, 2020.

[84] H. Naing, W. Cai, N. Hu, T. Wu, and L. Yu, "Data-driven microscopic traffic modelling and simulation using dynamic lstm," in *Proceedings of the 2021 ACM SIGSIM Conference on Principles of Advanced Discrete Simulation*, pp. 1–12, 2021.

[85] J. Redmon, "Darknet: Open source neural networks in c." http://pjreddie.com/darknet/, 2013–2016.

[86] N. Wojke, A. Bewley, and D. Paulus, "Simple online and realtime tracking with a deep association metric," in *2017 IEEE international conference on image processing (ICIP)*, pp. 3645–3649, IEEE, 2017.

[87] Q. Wang and M. Abbas, "Optimal urban traffic model predictive control for nema standards," *Transportation Research Record*, vol. 2673, no. 7, pp. 413–424, 2019.

[88] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "Yolov4: Optimal speed and accuracy of object detection," *arXiv preprint arXiv:2004.10934*, 2020.

[89] M. Naphade, S. Wang, D. C. Anastasiu, Z. Tang, M.-C. Chang, X. Yang, L. Zheng, A. Sharma, R. Chellappa, and P. Chakraborty, "The 4th ai city challenge," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pp. 626–627, 2020.

[90] A. Abdelhalim and M. Abbas, "Vt-lane: An exploratory study of an ad-hoc framework for real-time intersection turn count and trajectory reconstruction using nema phases-based virtual traffic lanes," in *2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC)*, pp. 1–6, IEEE, 2020.

[91] A. Abdelhalim and M. Abbas, "Towards real-time traffic movement count and trajectory reconstruction using virtual traffic lanes," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pp. 592–593, 2020.

[92] X. Hou, Y. Wang, and L.-P. Chau, "Vehicle tracking using deep sort with low confidence track filtering," in *2019 16th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, pp. 1–6, IEEE, 2019.

[93] B. Higgs and M. Abbas, "Segmentation and clustering of car-following behavior: Recognition of driving patterns," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 1, pp. 81–90, 2014.

[94] A. Abdelhalim and M. Abbas, "Vehicle class, speed, and roadway geometry based driver behavior identification and classification," *arXiv preprint arXiv:2009.09066*, 2020.

[95] T. Robotics, "Pinpoint vehicle positioning system," 2014. https://torcrobotics.com/wp-content/uploads/pinpoint_data_sheet1.pdf.

[96] D. C. Gazis, R. Herman, and R. B. Potts, "Car-following theory of steady-state traffic flow," *Operations research*, vol. 7, no. 4, pp. 499–505, 1959.

[97] V. Papathanasopoulou and C. Antoniou, "Towards data-driven car-following models," *Transportation Research Part C: Emerging Technologies*, vol. 55, pp. 496–509, 2015.

[98] Y. Yu, *Revisit of Microscopic Car Following Models: Conventional and Machine Learning Perspectives*. PhD thesis, 2021.

[99] V. Punzo, D. J. Formisano, and V. Torrieri, "Nonstationary kalman filter for estimation of accurate and consistent car-following data," *Transportation research record*, vol. 1934, no. 1, pp. 2–12, 2005.

[100] A. Abdelhalim, M. Abbas, B. B. Kotha, and A. Wicks, "A framework for real-time traffic trajectory tracking, speed estimation, and driver behavior calibration at urban intersections using virtual traffic lanes," 2021.

[101] R. Wiedemann, "Modelling of rti-elements on multi-lane roads," in *Drive Conference (1991: Brussels, Belgium)*, vol. 2, 1991.

[102] A. Abdelhalim and M. Abbas, "A real-time safety-based optimal velocity model," *In Review, IEEE Open Journal of Intelligent Transportation Systems*, 2021.