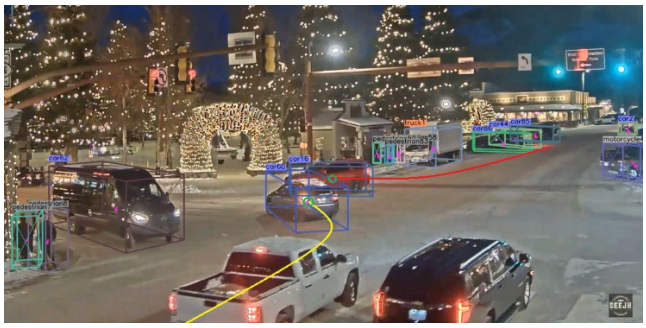
State of Wyoming Department of Transportation

FINAL REPORT

WY-2405F

RAPID SAFETY ASSESSMENT TOOL FOR NON-CONVENTIONAL ROADWAY DESIGNS AND EMERGING TECHNOLOGIES: INNOVATIVE ARTIFICIAL INTELLIGENCE APPLICATION



Principal Investigator
Mohamed M. Ahmed, Ph.D., PE
Professor and Director
Department of Civil and Architectural Engineering and Construction Management,
University of Cincinnati
Authors
Mohamed M. Ahmed, Ph.D., PE
Ahmed Mohamed, M.Sc.
Lizhe Li, Ph.D.
Mandip Sigdel
Vamsi Maddineni

Sponsored by Wyoming Department of Transportation (Grant No: RS04221)

Department of Civil and Architectural Engineering,
University of Wyoming
1000 E. University Ave, Laramie, WY 82071
mahmed@uwyo.edu / Mohamed.ahmed@uc.edu
Phone: 307-766-5550 Fax: 307-766-2221

December 2023

DISCLAIMER

Notice

This document is disseminated under the sponsorship of the Wyoming Department of Transportation (WYDOT) in the interest of information exchange. WYDOT assumes no liability for the use of the information contained in this document. WYDOT does not endorse products or manufacturers. Trademarks or manufacturers' names appear in this report only because they are considered essential to the objective of the document.

Quality Assurance Statement

WYDOT provides high-quality information to serve Government, industry, and the public in a manner that promotes public understanding. Standards and policies are used to ensure and maximize the quality, objectivity, utility, and integrity of its information. WYDOT periodically reviews quality issues and adjusts its programs and processes to ensure continuous quality improvement.

Copyright

No copyrighted material, except that which falls under the "fair use" clause, may be incorporated into a report without permission from the copyright owner, if the copyright owner requires such. Prior use of the material in a WYDOT or governmental publication does not necessarily constitute permission to use it in a later publication.

- Courtesy — Acknowledgment or credit will be given by footnote, bibliographic reference, or a statement in the text for use of material contributed or assistance provided, even when a copyright notice is not applicable.
- Caveat for Unpublished Work —Some material may be protected under common law or equity even though no copyright notice is displayed on the material. Credit will be given, and permission will be obtained as appropriate.
- Proprietary Information — To avoid restrictions on the availability of reports, proprietary information will not be included in reports, unless it is critical to the understanding of a report and prior approval is received from WYDOT. Reports containing such proprietary information will contain a statement on the Technical Report Documentation Page restricting availability of the report.

Creative Commons:

The report is covered under a Creative Commons, CC-BY-SA license. When drafting an adaptive report or when using information from this report, ensure you adhere to the following:

Attribution — You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

Share Alike — If you remix, transform, or build upon the material, you must distribute your contributions under the same license as the original.

No additional restrictions — You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.

You do not have to comply with the license for elements of the material in the public domain or where your use is permitted by an applicable exception or limitation.

No warranties are given. The license may not give you all of the permissions necessary for your intended use. For example, other rights such as publicity, privacy, or moral rights may limit how you use the material.

No Artificial Training

Any use of this publication to train generative artificial intelligence (AI) technologies to generate text is expressly prohibited. The authors, contractor, and WYDOT reserve all rights to license uses of this work for generative AI training and development of machine learning language models.

Declaration of Generative Artificial Intelligence and Artificial Intelligence Assisted Technologies

During the preparation of this work, author(s) used ChatGPT in and edited the content as needed and takes full responsibility for the content of this report. Further, to the best of the author(s) knowledge, the work generated from any artificial intelligence (AI) assisted technologies used in this report does not infringe upon any copyright or plagiarize anyone else's' work. Neither the author nor WYDOT can guarantee the accuracy or completeness of the information generated in this report by AI.

Technical Report Documentation Page

1. Report No. WY- 2405F	**2. Government Accession No.**	**3. Recipient's Catalog No.**	
4. Title and Subtitle Rapid Safety Assessment Tool for Non-Conventional Roadway Designs and Emerging Technologies: Innovative Artificial Intelligence Application		**5. Report Date** December 2023 **6. Performing Organization Code**	
7. Author(s) Mohamed Ahmed, Ph.D., P.E.- orcid.org/0000-0002-1921-0724 Ahmed Mohamed, M.Sc.- orcid.org/0000-0003-1829-4409 Lizhe Li, Ph.D. - orcid.org/0000-0001-8542-0677 Mandip Sigdel - orcid.org/0009-0007-4082-0213 Vamsi Maddineni- orcid.org/0009-0000-7346-5480		**8. Performing Organization Report No.**	
9. Performing Organization Name and Address Department of Civil and Architectural Engineering University of Wyoming 1000 E. University Avenue, Dept. 3295 Laramie, Wyoming 82071		**10. Work Unit No. (TRAIS)** **11. Contract or Grant No.** RS04221	
12. Sponsoring Organization Name and Address Wyoming Department of Transportation 5300 Bishop Blvd, Bldg. 6100 Cheyenne, WY 82009-3340		**13. Type of Report and Period Covered** Final Report December 2023 **14. Sponsoring Agency Code**	
15. Supplementary Notes Matt Carlson, P.E., State Highway Safety Engineer			
16. Abstract Traditional safety assessment methodologies are profoundly dependent on reactive crash data. The Highway Safety Manual (HSM) approach recommends 3-5 worth of crash data before and after the implementation of safety countermeasures. Waiting for such a long period of time might not be feasible to address safety issues at various roadway facilities. Furthermore, with the advent of emerging transportation technologies, rapid safety assessment tools will be required. This research provided a proof-of-concept for the development of a proactive road safety assessment framework which could be utilized for a rapid evaluation of problematic intersections, non-conventional designs, newly adopted countermeasures, as well as emerging transportation technologies. The framework is based on leveraging advanced Artificial Intelligence (AI) and machine vision to identify Surrogate Measures of Safety (SMoS) in near real-time from video cameras installed at intersections. The study established a relationship between the types of crashes and their contributing factors utilizing SMoS. Video analytics used machine vision and object detection algorithms to identify motion paths and trajectories for different road users. From estimated users' trajectories for vehicles and pedestrians, near crashes, known as traffic conflicts, were extracted by identifying critical thresholds for the SMoS such as Time-To-Collision (TTC), Post-encroachment Time (PET), and Deceleration Rate to Avoid a Crash (DRAC). Based on the SMoS identified, and the dominating conflict patterns, safety countermeasures could be recommended. The developed methodology of this study is a first step to cost-effectively assist transportation agencies evaluating hazardous locations and safety countermeasures without the need to wait for traditional crash data.			
17. Key Words Conflict Analysis, Artificial Intelligence, Computer Vision, Vehicle Trajectory, Vehicle Detection, Pedestrian Detection, Monocular 3-D Algorithms, Pose Estimation, Computer Vision, CNN, Conflict Analysis, Signalized Intersections, Rapid Safety Assessment, Surrogate Measures of Safety		**18. Distribution Statement** This document is available through the National Transportation Library and the Wyoming State Library.	
19. Security Classification (of this report) Non-Classified	**20. Security Classification (of this page)** Non-Classified	**21. No. of Pages** **144**	**22. Price**

METRIC CONVERSION FACTORS

SI* (MODERN METRIC) CONVERSION FACTORS

APPROXIMATE CONVERSIONS TO SI UNITS

Symbol	When You Know	Multiply By	To Find	Symbol
		LENGTH		
in	inches	25.4	millimeters	mm
ft	feet	0.305	meters	m
yd	yards	0.914	meters	m
mi	miles	1.61	kilometers	km
		AREA		
in^2	square inches	645.2	square millimeters	mm^2
ft^2	square feet	0.093	square meters	m^2
yd^2	square yard	0.836	square meters	m^2
ac	acres	0.405	hectares	ha
mi^2	square miles	2.59	square kilometers	km^2
		VOLUME		
fl oz	fluid ounces	29.57	milliliters	mL
gal	gallons	3.785	liters	L
ft^3	cubic feet	0.028	cubic meters	m^3
yd^3	cubic yards	0.765	cubic meters	m^3
	NOTE: volumes greater than 1000 L shall be shown in m^3			
		MASS		
oz	ounces	28.35	grams	g
lb	pounds	0.454	kilograms	kg
T	short tons (2000 lb)	0.907	megagrams (or "metric ton")	Mg (or "t")
		TEMPERATURE (exact degrees)		
°F	Fahrenheit	5 (F-32)/9 or (F-32)/1.8	Celsius	°C
		ILLUMINATION		
fc	foot-candles	10.76	lux	lx
fl	foot-Lamberts	3.426	candela/m^2	cd/m^2
		FORCE and PRESSURE or STRESS		
lbf	poundforce	4.45	newtons	N
lbf/in^2	poundforce per square inch	6.89	kilopascals	kPa

APPROXIMATE CONVERSIONS FROM SI UNITS

Symbol	When You Know	Multiply By	To Find	Symbol
		LENGTH		
mm	millimeters	0.039	inches	in
m	meters	3.28	feet	ft
m	meters	1.09	yards	yd
km	kilometers	0.621	miles	mi
		AREA		
mm^2	square millimeters	0.0016	square inches	in^2
m^2	square meters	10.764	square feet	ft^2
m^2	square meters	1.195	square yards	yd^2
ha	hectares	2.47	acres	ac
km^2	square kilometers	0.386	square miles	mi^2
		VOLUME		
mL	milliliters	0.034	fluid ounces	fl oz
L	liters	0.264	gallons	gal
m^3	cubic meters	35.314	cubic feet	ft^3
m^3	cubic meters	1.307	cubic yards	yd^3
		MASS		
g	grams	0.035	ounces	oz
kg	kilograms	2.202	pounds	lb
Mg (or "t")	megagrams (or "metric ton")	1.103	short tons (2000 lb)	T
		TEMPERATURE (exact degrees)		
°C	Celsius	1.8C+32	Fahrenheit	°F
		ILLUMINATION		
lx	lux	0.0929	foot-candles	fc
cd/m^2	candela/m^2	0.2919	foot-Lamberts	fl
		FORCE and PRESSURE or STRESS		
N	newtons	0.225	poundforce	lbf
kPa	kilopascals	0.145	poundforce per square inch	lbf/in^2

*SI is the symbol for the International System of Units. Appropriate rounding should be made to comply with Section 4 of ASTM E380.
(Revised March 2003)

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

Acronym	Description
AAMVA	American Association of Motor Vehicle Administrators
ADAS	Advanced Driver Assistance Systems
ADS	Automated Driving Systems
ARCIS	Automated Roadway Conflict Identification System
AV	Autonomous/Automated Vehicle
CAT	Cooperative Autonomous/Automated Transportation
CAV	Connected and Autonomous/Automated Vehicle
CCTV	Closed-Circuit Television
CNN	Convolutional Neural Network
CSRT	Channel and Spatial Reliability Tracking
DOT	Department of Transportation
DRAC	Deceleration Rate to Avoid a Crash
EVT	Extreme Value Theory
FHWA	Federal Highway Administration
FMVSS	Federal Motor Vehicle Safety Standards
IICS	Intelligent Intersection Control System
IOU	Intersect of Union
mAP	mean Average Precision
MaxD	Maximum Deceleration rate
MLE	Maximum Likelihood Estimate
MO	Moving Object
MTC	Markings Technical Committee
MTTC	Modified Time-to-Collision
mTTC	Minimum Time-to-Collision
MUTCD	Manual on Uniform Traffic Control Devices
NHTSA	National Highway Traffic Safety Administration
NCUTCD	National Committee on Uniform Traffic Control Devices
OCR	Optical Character Recognition
OpenPifPaf	Open Pose in Full Pose Articulation Framework
PET	Post Encroachment Time
POT	Peak Over Threshold
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PWM	Probability-Weighted Moments
R-CNN	Region-based Convolutional Neural Network
SMoS	Surrogate Measures of Safety
SSD	Single Shot Detector
TCD	Traffic Control Devices
TSR	Traffic Signs Recognition
TTC	Time-to-Collision
UAV	Unmanned Aerial Vehicle
VOC	Visual Object Classes
VMT	Vehicle Miles Traveled
WYDOT	Wyoming Department of Transportation
YOLO	You Only Look Once

CHAPTER 1. EXECUTIVE SUMMARY

Motor vehicle crashes remain a leading cause of death in the US and worldwide. In 2023, the US saw 39,508 crashes involving 61,332 motor vehicles, resulting in around 40,000 deaths and over 3 million injuries. The National Highway Traffic Safety Administration (NHTSA) reported a slight decrease in traffic fatalities, with about 9,330 deaths in the first three months, down 3.3 percent from the previous year. Similarly, the National Safety Council (NSC) noted a 3 percent decrease in motor-vehicle deaths in the first half of 2023, totaling 21,130. Despite these improvements, motor vehicle crashes still pose a significant threat to public safety. Advanced road designs, traffic control measures, and intelligent transportation systems continue to be developed to enhance safety and reduce the high frequency of crashes, particularly at intersections. The complexity of intersections with their mixed traffic movements and users, contribute to a significant portion of crashes. However, the effectiveness of new safety countermeasures, especially in rapidly evolving traffic conditions, remains a topic of debate among transportation professionals.

While traditional safety assessment methods have been helpful to better understand crash trends, predict crashes, and evaluate the effectiveness of countermeasures, these techniques depend on aggregate crash data. The primary variable in Safety Performance Functions is average daily traffic, an annual aggregate exposure measure. For intersection safety performance, total entering volumes are key, but this poses a limitation in developing intersection-specific safety functions. Intersections can have 15 distinct crash patterns, and not all are directly linked to traffic volumes of various movements. For instance, left-turn maneuvers are notably hazardous, yet the complexity of gathering detailed turning movement data often leads to reliance on total entering volumes for safety assessments. Furthermore, crash reports typically lack detail, omitting crucial data on crash patterns, weather conditions, road surface status, and real-time traffic. Essential information like speeds, specific turning movements, and exact pre-crash conditions is often missing from these reports, further complicating accurate safety evaluations. The increasing complexity at non-conventional intersections, such as roundabouts and Diverging Diamond Interchanges (DDIs), is often exacerbated by road users' unfamiliarity with these new traffic patterns. This unfamiliarity can lead to unexpected driving behaviors, raising the likelihood of near-crashes or actual crashes. Traditional safety assessment methods primarily rely on

historical data, including past crashes and traffic volumes. However, these methods typically require three to five years of data for a comprehensive traffic safety assessment, indicating a reactive rather than proactive approach. One significant limitation of this approach is its dependency on a substantial sample size of crash observations before safety evaluations can be conducted.

Proactive safety measures, on the other hand, offer several advantages. They aim to identify and mitigate potential crash scenarios before they occur, utilizing traffic conflict analysis, which serves as an indicator of potential crashes. This approach is particularly important in understanding driver's behavior, a critical factor in all traffic incidents. Furthermore, traditional safety studies, which depend on aggregate crash data, often fail to provide real-time or detailed pre-crash information, limiting their ability to convey driver behavior prior to a crash.

With the advent of emerging technologies like Connected and Automated Vehicles (CAVs), traditional safety methodologies are becoming less adequate. These technologies are expected to significantly alter the dynamics of road safety and traffic operations. Consequently, there is a growing need for innovative assessment techniques that can evaluate the safety effectiveness of these innovative technologies. CAVs, equipped with Advanced Driver Assistance Systems (ADAS) such as forward collision warnings, automatic emergency braking, and pedestrian detection systems, have the potential to drastically reduce crash rates.

Moreover, proactive safety techniques are versatile enough to evaluate both conventional and non-conventional intersections, including the interactions between driver behavior and various environmental, vehicular, and roadway factors. They can also assess the impact of CAVs on collision risks. These evolving techniques can complement historical crash data by proactively identifying locations with high crash or near-crash probabilities and recommending effective mitigation strategies in a relatively shorter time frame. This proactive approach challenges the ethical dilemma faced by transportation agencies of waiting for a significant number of crashes, including fatalities and injuries, to occur before identifying and addressing hazardous locations.

Finally, in states like Wyoming, where crashes can be sporadic and unpredictable due to the remote nature of the region, traditional safety techniques may not be feasible. Innovative approaches, such as utilizing video recordings and conflict analysis, can help in developing long-

term safety countermeasures and immediate interventions, enhancing road safety in both conventional settings and with the integration of modern technologies like CAVs.

Expanding on this motivation, this study establishes a proof of concept for employing an alternative proactive technique, conflict analysis, which utilizes Artificial Intelligence (AI) to assess traffic safety at signalized intersections. The report follows a specific organizational sequence. The first chapter provides a general introduction to traffic conflict proactive techniques, and the utilized surrogate safety measures that identify traffic conflict occurrence. This is followed by the problem statement, and the study motivation and objectives. In the second chapter, the manual and automated procedures for identifying traffic conflicts will be presented. Then, the methodologies employed in detecting and tracking road users using surveillance cameras are detailed, encompassing both traditional and advanced techniques. Chapter three introduces post-processing techniques, covering output reduction through outlier removal, trajectory smoothing, and extraction of traffic features. Chapter four is broken into three main sections. The first one is an extensive background on diverse types of Surrogate Measures of Safety (SMoS) accompanied by a synthetic study exploring different conflict indicators and their applicability in identifying conflicts with varying thresholds. The second section provides a detailed explanation of the data collection process at the site of interest combined with heat maps and descriptive analyses. Moreover, different patterns of traffic conflicts that are commonly found at signalized intersections will be illustrated. While in the third section, a subset of the collected data is selected, then, the traffic conflicts identifications frameworks are applied to extract traffic conflicts by employing Extreme Value Theory (EVT) and utilizing different types of SMoS. Finally, the report concludes with chapter five, presenting overall conclusions and recommendations.

In this report, the main goal of utilizing AI technology to identify traffic conflicts has been achieved. The research additionally resulted in benefits like the enhancement of machine vision algorithms, specifically by assessing their performance in relation to the varying heights at which cameras are mounted. Crucially, it was concluded that detecting traffic conflicts relies significantly on incorporating various indicators due to the complex nature of conflicts influenced

by factors like driver behavior, intersection design, sight distances, signal phasing and timing, and weather and surface conditions.

The effectiveness of a single measurement for specific conflict types under certain conditions is limited. A more comprehensive framework necessitates the inclusion of multiple conflict measurements, influenced by road user trajectories. This approach demands varied case studies across different intersections and traffic conflict scenarios to overcome inherent limitations. Our study demonstrates this by analyzing two intersections with varying camera mounting heights, employing four traffic conflict indicators to develop the detection and analysis framework. While this provides a proof of concept, further enhancements are essential to refine and address the limitations of the proposed methodology.

CHAPTER 2. INTRODUCTION AND BACKGROUND STUDY

While conventional methods for assessing safety have been valuable in understanding crash trends, predicting incidents, and assessing countermeasure effectiveness, these approaches primarily rely on aggregated crash data. Safety Performance Functions, for instance, hinge on Average Daily Traffic (ADT) as the key variable, representing an aggregate exposure measured annually. However, when appraising intersection safety performance, Total Entering Volumes (TEV) serve as the primary exposure, presenting a limitation in developing Safety Performance Functions for intersections. Intersections involve 15 distinct crash patterns, with TEV not consistently indicative of all 15 types, as some depend solely on through traffic volumes while others are influenced by turning movements, where left turns are deemed the most hazardous. Despite the complexity of collecting precise turning movement data, researchers often resort to using TEV. Other constraints include the lack of detailed information in crash reports, with data on crash patterns, weather, road conditions, and real-time traffic often missing. Moreover, crash reports rarely provide insights into speeds, turning movements, and specific pre-crash conditions.

Traditional safety techniques are hindered by their reactive nature and face numerous critical limitations. These limitations comprise the subjective and inconsistent nature of crash police reports. The evolving crash reporting system struggles to keep pace with Intelligent Transportation Systems and advancements in the automotive industry. Additionally, there is underreporting of minor crashes. These techniques lack information on driver behavior during normal driving and in pre-, during-, and post-crash events. Furthermore, there is a deficiency of data on violations and near crashes. An incomplete understanding of interactions between different road users is another challenge. Finally, the necessity for crashes to occur is a requirement for meaningful safety analysis, as stipulated by the Highway Safety Manual, which recommends 3-5 years of crash data.

Conversely, proactive safety studies are gaining traction due to their capacity to swiftly assess new safety measures. These studies aim to explore dynamic factors influencing crash and near-crash probability. Modeling techniques in proactive safety studies focus on individual Critical Safety Events (CSE) or conflicts, along with their precursors, to predict real-time crash risks and

comprehend the causal effects. These techniques rely on high-resolution data such as speed, vehicle trajectory, conflicts, traffic, and weather. Traffic conflicts refer to incidents that could escalate to crashes if the involved road users persist in their movements. Proactive safety analyses based on traffic conflicts offer distinct advantages over crash-centric approaches. Traffic conflicts are more frequent and rapidly obtainable, detailed information is extractable from video recordings, direct collection of driver behavior occurs, associated social and economic costs are relatively lower, consistent data can be obtained through well-trained observers and machine learning algorithms, and prompt recommendations for non-conventional countermeasures and emerging technologies can be provided to decision-makers. This study provides proof of concept for employing an alternative proactive technique, conflict analysis, utilizing Artificial Intelligence (AI) to assess traffic safety at signalized intersections. Conflict analysis, employing machine vision approaches, will be employed to swiftly evaluate the operational and safety performance of identified intersection. Video recording data are gathered from existing cameras at case study intersection. A comprehensive 40 hours of video recording data formed the conflict dataset. Moreover, several detection and tracking algorithms are employed to investigate their strengths and limitations. Conflict analysis frameworks are proposed based on the employed detection algorithm and the number of utilized cameras. Finally, conflict sets were identified through the applications of these frameworks. The following sections provide background on proactive safety assessment methodologies, the problem statement, objectives of the study, selected study areas, and the performed tasks to achieve the proposed objectives.

2.1 Background

In recent years with Vision Zero[1], new innovative safety evaluation procedures have been evolved to evaluate non-conventional roadway and intersection designs, as well as new safety and operation countermeasures. Nontraditional approaches, such as Conflict Analysis (CA), have gained popularity to evaluate the safety and operations of roadways and intersections. Transportation researchers have extensively used machine learning and computer vision techniques in recent years, especially with the presence of new emerging technologies such as

[1] Vision Zero is a multi-national set of strategies to prevent fatalities and severe injuries related to road traffic.

Connected and Automated Vehicles. These techniques have been used in multiple aspects to enable vehicles to identify the surrounding features to operate safely and efficiently, to recognize license plates of vehicles for means of traffic surveillance and toll collection [1–3], trip routing estimation [4, 5], and parking management systems [6]. Moreover, machine vision techniques were used to estimate Surrogate Measures of Safety (SMoS) by identifying vehicle trajectories where traffic conflicts could be extracted and further analyzed. The rapid advancement in machine vision technologies, computer processing power, and high-speed connectivity has opened new outlooks of opportunities in the field of transportation engineering to improve traffic safety and operations. The potential of machine learning techniques in identifying unexpected events on the roadways, as well as real-time weather and surface conditions have been thoroughly explored in the literature.

2.1.1 Conflict Analysis for Safety Performance Assessment

Due to the stochastic and scarcity nature of crashes, traffic conflict analysis has gained momentum in recent years. The Federal Highway Administration (FHWA) has developed guidelines for conflict analysis, detailing procedures for trained observers to record traffic conflicts using diagrams and narrations within specific timeframes [7]. Significant contributions in the field of traffic conflict analysis include Sayed and Zein's investigation into traffic conflict standards for intersections [8], the introduction of computer vision techniques for automated conflict data collection [9], and research focusing on pedestrian involvement in collisions [10]. Other researchers conducted a large-scale automated analysis of vehicle interactions and collisions by developing a system to identify traffic conflicts/collisions into 4 categories: head-on, rear-end, side, and parallel near crashes. The results showed the usefulness of the system to study driver behaviors that might lead to a collision [11]. St-Aubin et al. presented a practical framework for implementing an automated, high-resolution, video-based traffic-analysis system by utilizing Closed-Circuit Television CCTV and regular video cameras to collect microscopic traffic flow data [12].

Various studies have employed conflict detection algorithms to expedite traffic safety analysis, given the higher frequency and clearer capture of conflicts compared to crashes. Safety performance functions for roadway facilities have been developed using traffic conflicts [13].

Zhang et al. presented a conflict prediction model for left-turn conflicts in China [14], while Sacchi and Sayed established a relationship between conflicts and crashes using a two-phased process [15]. Bayesian analysis was applied by Sacchi and Sayed (2016) to develop safety performance functions for conflicts at signalized intersections [16].

Several studies utilized Extreme Value Theory (EVT) to estimate crashes from traffic conflicts [17–21], affirming the effectiveness of this approach. Notable studies include Zheng and Ismail's generalized link function for conflict severity [17], Tarko's comparison of crash prediction approaches [18], and a study comparing estimated crashes from traffic conflicts with observed crashes [19]. Wang et al. proposed a bivariate EVT framework for crash prediction using conflict indicators [20], and another study applied bivariate Bayesian hierarchical models for non-stationary conflict extremes to estimate crashes [21]. These studies collectively contribute to the evolving field of traffic conflict analysis for improved road safety.

2.1.2 Surrogate Measures of Safety

The literature on SMoS in traffic conflict analysis has evolved since the 1960s [22]. SMoS offer a means to characterize the interactions between road users, estimating both the probability and severity of potential collisions [23–25]. Various indicators, including Time to Collision (TTC), Post-Encroachment Time (PET), and Deceleration rate (DR), have been widely used to assess the nearness between road users and predict potential collisions [23–25]. The analysis of SMoS reveals research gaps, such as the lack of uniform methods in their analysis, impacting factors like roadway facility, observation method, surrogate measure type, and threshold selection [24].

Recent studies focused on time-series data and explored clusters of conflicts and non-conflicts, considering aspects like TTC, speed, acceleration, and vehicle trajectory [26]. However, systematic reviews highlighted weaknesses in validation, surrogate measure selection, and threshold determination [23, 24, 27]. Some studies have concluded that the choice of measure and its threshold depends on various conditions and application contexts, so measures should be selected depending on their context [27, 28, 29]. Prominent SMoS encompass TTC, with enhancements such as Modified Time-to-Collision (MTTC) demonstrating stronger correlations with crashes [30, 31]. Another indicator, T2, considers the expected time for the second user to

reach a potential collision point, providing a more informative safety measure [32]. Challenges persist in selecting appropriate thresholds, with a common window of 1 to 5 seconds for TTC [25, 26, 33–35]. Other variations of TTC such as time-extended time-to-collision and time-integrated time-to-collision are less commonly used in conflict studies despite their potential [36].

PET is extensively used alongside TTC, offering complementary information on conflicts after one user passes the collision point [37, 38]. Encroachment Time (ET) is a derivative of PET, measuring the duration of time offending vehicles occupy a conflict area [39–41]. Researchers have explored additional SMoS, such as Aggregate Severe Crash Metric (ASCM), proposed by Chen Wang et al., showing promise in predicting fatal or injury crashes [37]. The Aggregate Crash Propensity Metric, developed by Wang and Stamatiadis, incorporates time-to-collision, reaction time, and deceleration rate, but faces challenges related to subjective judgment of evasive maneuvers [42]. Microsimulation studies, like Amir Sobhani et al.'s, include kinematic energy difference in predicting traffic conflicts and crash severity [43].

Despite advancements, challenges persist in establishing uniform methods, validating SMoS, and defining ideal thresholds. The literature suggests the need for a systematic review and Meta-Analysis to determine suitable threshold selection criteria for identifying near-crash events [24].

2.2 Problem Statement and Study Motivation

The evaluation of traffic safety performance at intersections holds significant importance for transportation agencies, particularly considering the shared geographical space where road users interact, leading to complex conflicts. Traditional safety assessment methods relying on historical crashes and traffic volumes face limitations in data requirements through the need of three to five years of data for assessment. This reactive approach, dependent on a decent sample size of crash observations, may be hindered by low exposures. While reactive safety is potent for addressing existing issues, proactive safety analysis offers advantages in diminishing the potential for crashes by investigating traffic conflicts before they occur, serving as crash indicators.

Driver behavior, a crucial factor influencing traffic crashes, is inadequately addressed by traditional safety studies relying on reactive aggregate crash data, lacking direct insights into behavior preceding a crash or detailed pre-crash information. Despite crashes being rare events, their complexity, resulting from multiple factors and failures, is not effectively captured by crash reports. Exposure information, including the frequency of behaviors in normal driving and broader contextual factors contributing to crashes and near-crashes, remains absent.

The advent of emerging technologies like Connected and Automated Vehicles (CAV) renders traditional safety approaches unsuitable for promptly assessing their impacts on roadway safety and operations. These approaches also prove inadequate for assessing newly introduced non-conventional intersections due to data limitations and an unfamiliar driver population. In contrast, proactive innovative safety techniques possess the capability to assess conventional and non-conventional intersections, driver behavior, and performance interactions with various factors, including roadway, environmental, vehicular, and CAV impacts on collision risks. These evolving techniques complement historical crash data, proactively identifying hazardous locations and recommending mitigation strategies more efficiently. This introduces an ethical concern, questioning the necessity for transportation agencies to wait for a sufficient number of crashes, including fatalities and injuries, before identifying and mitigating hazardous locations. The proposed proactive safety approach provides insights into human errors leading to near crashes, enabling the observation of different human actions causing incidents.

Critical for evaluating the safety effectiveness of emerging technologies like CAV, innovative assessment techniques are required. CAV, equipped with multiple safety technologies, including Advanced Driver Assistant Systems (ADAS), are anticipated to significantly reduce crashes. Given the remote nature of Wyoming and the stochastic rarity of crashes, traditional safety techniques become impractical. Utilizing video recordings, conflict analysis, implementing long-term safety countermeasures, and deploying instantaneous interventions emerge as more feasible strategies.

2.3 Research Objectives

This research proposes a proactive road safety assessment for signalized intersections and emerging technologies in Wyoming utilizing traffic conflict analysis. Furthermore, the

innovative assessment methodologies provide detailed observation and analysis of trajectory interpretation, and conflict measures. These methodologies will be based on computer vision applications utilizing recorded video data. The proposed methodologies were applied on Town Square, Jackson Hole signalized intersection within Wyoming. The case study intersection was selected based on the availability of a live feed video channels that connected with three surveillance cameras. Also, the nature of this intersection as a recreational area offers additional complexity in the traffic mixture with various types of road users and higher volume. Moreover, the selection of the case study intersection met the recommendations from WYDOT Safety Office. The objectives of this research include the following:

- Investigate several types of road users' detection and tracking methodologies.
- Perform a conflict analysis to provide insights about the causes that lead to Critical Safety Events (CSE) (i.e., crashes and near crashes).
- Develop an automated conflict analysis system to evaluate the performance of roadway facilities with a focus on intersections/ interchanges.
- Provide WYDOT with a new framework of safety assessment techniques which is more appropriate for evaluating emerging technologies.

2.4 Report Organization

The structure of this report is outlined as follows:

Chapter 3: This section will provide an overview of both manual and emerging technologies on conflict detection. Various tracking methodologies for road users will be applied to the case study intersection, evaluating their strengths and weaknesses. The detection and tracking techniques encompass traditional methods involving feature tracking and background subtraction, along with advanced deep learning algorithms such as those used in autonomous vehicles, bounding boxes, and key points detection.

Chapter 4: Post-processing techniques will be implemented in this chapter to eliminate data outliers, validate output data, demonstrate transformation homography, extract road users' trajectories, smooth them, and calculate speeds and accelerations.

Chapter 5: The initial part will present a background on different traffic conflict indicators. Subsequently, a synthetic study will be conducted to assess severity levels based on various

thresholds. The following section will extensively explore various conflict types present at signalized intersections. This will be followed by the presentation of video data, manual extraction, and annotation of the conflict set. The distribution of extracted conflicts within the intersection area will be analyzed to identify the most prominent conflict types. In the concluding section, two conflict subsets will be scrutinized using the proposed computer vision algorithms and different frameworks.

Chapter 6: Conclusions and recommendations will be provided, offering insights for future applications of various detection and tracking techniques. Strengths and limitations of these techniques will be discussed based on intersection features and camera installation positions. Specific recommendations will be presented, considering the monitoring of the case study intersection and the extracted conflict set.

CHAPTER 3. ROAD USERS' DETECTION METHODOLOGIES

In this chapter, a background on the manual and automated detection methodologies of analyzing traffic conflicts is provided. This is followed by the employment of different computer vision algorithms in road users' detection.

3.1 Background

3.1.1 Manual Techniques in Conflicts Detection

An integral first step in conflict analysis is to identify near crashes. Early studies utilized manual techniques by training human observers on different conflict types and selected Surrogate Measures of Safety (SMoS) thresholds [48, 49]. Manual observation procedures were classified into two categories: subjective, and objective [50]. Subjective methods are performed in two steps: conflicts are investigated by trained observers to classify the severity level, then criticized by researchers since the observers' judgments varied from one observer to another. The objective methods of observation are based on the utilization of time measurements to determine the severity level. Therefore, a SMoS such as TTC was utilized as an essential measure in objective methods. In addition to TTC, several indicators were developed and employed to judge the severity level of traffic conflicts at different roadways conditions. These indicators involved; MTTC, PET, deceleration rate to avoid collision (DRAC), proportion of stopping distance (PSD), among other indicators [51–53]. Zheng et al. classified conflicts indicators into temporal and spatial proximity types according to the nature of each indicator [28]. While Johnsson et al. broke them into three main categories: TTC, PET, and deceleration rate [53]. The selection between the indicators and their specific area of application has been extensively studied through several research works, and it seems to differ based on various contexts [51, 52, 55]. Hence, recommendations for the relevant indicator types and thresholds have been set to detect and assess the traffic conflicts that have occurred at different roadway locations.

In another study, the manual observations of numerous studies were conducted to extract the most fitted statistical models that could describe the relation between conflicts, traffic volumes, and collisions under several cases and different conditions [56–58]. While this

approach achieved decent results on specific sites, it had some deficiencies in covering all the research gaps (e.g., conflicts detection at hazardous locations, and adverse weather conditions). In addition to the excessive cost of hiring and training the observers, another limitation for this technique is the variation in judgements on traffic encounters between the observation teams [59, 60].

3.1.2 Emerging Computer Vision Techniques in Conflict Detection

Based on the rapid and incessant development in the computer vision field, researchers managed to develop tracking techniques for vehicles along roadway sections and intersections utilizing distinct types of cameras. These cameras are either installed for general traffic surveillance or specifically mounted for conflict analysis [61]. Upon the development of tracking methodologies, traffic conflicts were detected through computational algorithms [62, 63]. Five types of objects tracking techniques were developed: 3D model-based tracking, region-based tracking, contour-based tracking, feature-based tracking, and hybrid methods. [64]. Each approach has its pros and cons.

In 3D model-based tracking, vehicle 3D parameters were inserted into the tracking algorithm. Then, vehicle recognition was achieved by matching a projected model to the image data [65]. Although that approach provided high accuracy in orientation detection and tracking for vehicles, it faced two critical drawbacks. Accuracy was achieved for small numbers of vehicles, and it required detailed parameters insertion for all vehicles extracted from cameras, which may not be available in most cases. For region-based tracking, the region for each vehicle was obtained through background subtraction. Then, the vehicle was tracked utilizing the extracted parameters from the region (size, color, shape, etc.). Many researchers [66–69] utilized Kalman filter to perform region-based tracking. That technique showed efficiency in computational processes regarding free-traffic flow. Whereas in congested traffic areas, it faced difficulties in vehicles recognition and separation. There is a similarity between the region and contour-based tracking in which the contour-based tracking technique is creating a boundary surrounding the vehicle, then, a trace line is extracted from the vehicle to describe its path and dynamically update its location. This also applied to other road users. However, contour-based tracking approach was found to be more accurate compared to the region-based tracking, it was

challenged in isolating objects at occlusions. Overcoming this issue depends on creating a contour for each vehicle in the system [70] while the feature-based technique determines specific points or lines (features) on the tracked object. Since each road user is modeled in this approach with a set of features belonging to it, the advantage of this technique is the ability to detect and separate the road users when partial occlusions take place at different times during the day and night. This has been achieved by calculating the geometric properties of road users at clear views, then, predicting them at occluded scenes. Consequently, the integration between Kalman filter and features tracking algorithms could be employed as a hybrid method in achieving more decent results for both detection and tracking accuracy [71], where Kalman filter was utilized efficiently in free-flow conditions to set regions for road users.

3.1.2.1 Computer Vision Applications in Conflicts Detection

Kalman filtering and the Kanade-Lucas-Tomasi feature tracker are two well-known developed techniques applied as hybrid method. Features-based tracking algorithm was utilized to detect vehicles' motion on highway portions and the traffic parameters (flow rate, average speed, and average spatial headway) were computed [72]. A transformation homograph was performed to create a mapping system between perspective image and the 2D top view. The utilization of such homographs serves in easing the calculation of traffic stream parameters. The entrance and exit regions were determined through the shown homograph in **Figure 1.** The homograph illustrates the transformation between a 3D image and a 2D highway section plan.

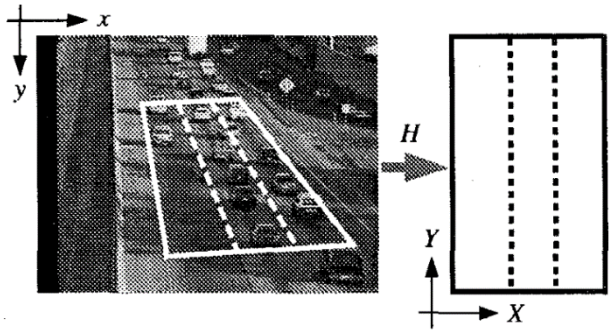


Figure 1 A Projective Transform H, is Used to Map from Image Coordinates (*x, y*) To World Coordinates (*X, Y*) [72]

A simple transformation matrix was put to compute the real-time coordinates of the selected features from the 3D image, **Eq.**1.

$$\begin{Bmatrix} X \\ Y \\ 1 \end{Bmatrix} \propto H \begin{Bmatrix} x \\ y \\ 1 \end{Bmatrix} \quad (1)$$

Corner points were selected, detected, and tracked as sub-features through the dynamic motion of vehicles. It was assumed that the road surface is flat, and the vehicle's movement parallels the road plane. This assumption was put to eliminate the z-coordinate from the 3D image; hence, *H (3, 3)* was chosen to be 1. The tracking module tracked the vehicle sub-features through the highway portion regions (entrance and exit) to determine the sub-features coordinates regarding time. Kalman filter was employed to maintain the noisy measurements utilizing normalized correlation. Then, positions and velocities were calculated $(X, Y, \dot{X}, \dot{Y})$. Real-time computations for traffic parameters (traffic flow, velocity, density, headway) were extracted from the algorithm for each lane separately using the proposed methodology in [73], then, compared to ground truth data. The approach showed a decent accuracy in the correlation between the computation and real traffic parameters. Additionally, the vehicle tracking algorithm and velocity determination model achieved good accuracy in validation with the ground truth data. Table 1 illustrates the error distribution for the computed traffic parameters. The velocity measurements were very accurate. Whereas for other parameters, the errors were interpreted as missing in or overcounting vehicle frequencies. Since the computation time duration is relatively small (5 minutes), it could affect the parameters calculations such that one error or two could be very significant.

Table 1 Error Distribution for Velocity, Flow, Density, and Headway [69]

Error percent less than	Traffic Parameters			
	Velocity Samples percent	Flow Samples percent	Density Samples percent	Headway Samples percent
2.5	86	18	19	19
5	95	31	33	34
10	100	60	59	60
15	100	79	79	81
20	100	91	90	89
25	100	96	96	94

That approach was extended by Saunier, N., and T. Sayed, to include tracking vehicles at intersections [71]. The algorithm was adopted from [72] with a specific modification in the transformation homography matrix. This change is based on the geometric features of the intersections. Four different intersections were selected for experiment conduction. The utilized video sequences for the proposed intersections were gathered such that the main sequence was a set of collected examples used to train conflict observers. Two sequences for two different intersections came from the repository of the University of Karlsruhe. The last video sequence was collected from Vera Kettnaker's former research webpage. Table 2 shows the length of each video sequence in terms of the number of frames.

Table 2 Video Sequences for Evaluation, With Their Length (Number of Frames) [71]

Sequences	Length (Number of frames)
1: Conflicts set	5793
2: Karlsruhe sets	1050
3: Cambridge set	1517

The overall results were satisfying with a vehicle detection percentile equal to 88.4 percent. The pedestrians and two-wheeled vehicles were tracked efficiently, while the errors were summarized as below:

- Computation inaccuracies in features grouping at far distances related to the camera's position.
- Errors caused by camera jitter.
- Problems caused by trucks and buses which are often over segmented.
- Over grouping when two or more vehicles moved together, or one feature is detected moving with two groups.

A hybrid strategy was applied to resolve appearance and disappearance, splitting, and partial occlusion problems [64]. This strategy was based on the interaction between objects and regions' characteristics. The technique was evaluated on rigid and deformable objects, outdoors and indoors, and it has proven good accuracy. All algorithm components could be run in real-time using a standard PC except for the region segmentation stage. It could be beneficial for visual surveillance, video-based hyperlinking, and video manipulation. Moreover, the mechanism could be modified by inserting priori information about objects included in the scenes.

Another study that addressed vehicle tracking at intersections was conducted by [69] to investigate the traffic conflicts at intersections through video monitoring. Whereas the Kalman filter was used in tracking vehicles and pedestrians at intersections. The technique was based on combining low-level image-based tracking with a high-level Kalman filter for position and shape estimation. First, grey scaled images were provided to the system from a stationary camera fixed at the intersection. Then, a mixture of the Gaussian models' method was utilized for the image segmentation process, where each moving object was represented by a blob [74]. The features of each blob (centroid, area, elongation, and first and second moments) were computed in the real-time extraction process during motion from one frame to another. Upon the calculated features, a bounding box was put for each blob to approximately estimate its dimensions and rotations. Then, the tracking process was performed by determining the bounding box for each moving object at each frame regarding timestamp. Labels, timestamps, velocities, and other features were assigned to the blobs. The positions for blobs were estimated, after reassigning blobs as moving objects (MO), using an extended Kalman filter. While shape estimation was done using a standard discrete Kalman filter. Occlusions were investigated from the results of the shape estimator. Finally, a visualization module was created based on the extracted results from the tracking module. The extracted data from the visual module which included positions, velocities, and shapes (lengths and widths), were utilized in the conflicts detection module which was responsible for making predictions of possible collisions between vehicles. The possible collision was estimated by comparing the measured

distances between bounding boxes to a minimum threshold value. The developed visual conflicts detection module provided an easy user interface for online possible collision detection. Figure 2 depicted the conflicts detection interface.

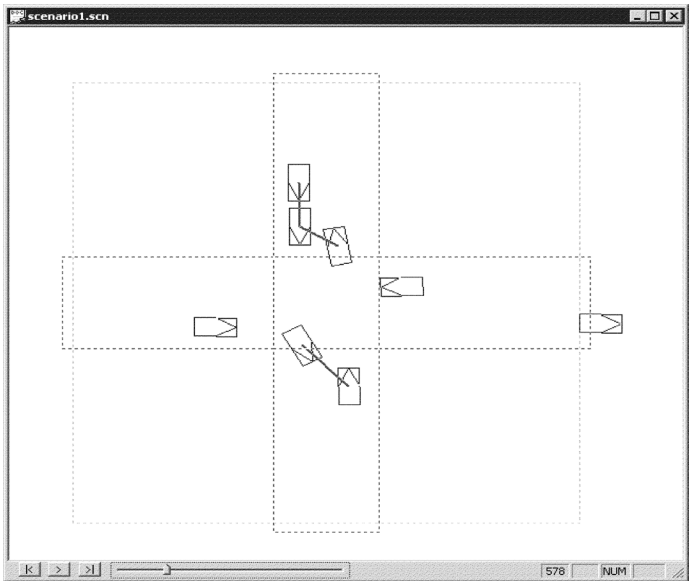


Figure 2 Incident (Conflict) Detection Interface [69]

A recent study using CCTV was presented by Mohamed Abdel-Aty in 2022. The study utilized the videos outputs from CCTV, then applied a Mask-RCNN bounding box and occlusion-Net algorithm to extract the 3D - coordinates of the vehicle's key points (e.g., right-front headlight, left-front headlight, right-back taillight, and left-back taillight) [75]. The 3D-coordinates are transformed to 2D-coodrinates using a transformation algorithm. The method efficiency was assessed by installing a drone video camera at the center of the intersection to provide a plan view (bird eye perspective) which worked as the ground truth data. The study employed PET at pixel level to identify traffic conflicts instances between vehicles. While the intersection was broken into three regions: intersection entrance area (before the stop line), within-intersection area (the region within the intersection that bounded by the stop lines), and intersection exit area (the region located after vehicle pass the within-intersection area) which was recommended in [76]. The utilized framework was an integrated system of algorithms where the first utilized framework is "Automated Roadway Conflict Identification System (ARCIS)" that is employed in detection and tracking vehicles from Unmanned Aerial Vehicle (UAV) images. Then, the undetected vehicles were located using Mask-RCNN from video images. Next, the

vehicles trajectories were obtained by utilizing Channel and Spatial Reliability Tracking algorithm (CSRT). Finally, to assure the detection efficiency, a comparison was held between the detection results and the tracking results using the Intersect of Union (IOU). The study was enriched by including a pre-trained deep neural network object detection model utilizing "COCO 2017" dataset that includes 12,786 cars, 61,377 trucks, and 4,141 buses to find road users from video images. Figure 3 illustrates the utilized vehicle detection framework and 2D image reconstruction. An experiment was performed to assess the effectiveness of the proposed framework. The results have been validated that the presented framework could considerably enhance the precision of localization from the 2D plan view. While PET values were successfully utilized in identifying traffic conflicts between vehicles. The limitation of this study was full dependence on the CCTV to provide usable images under adverse weather conditions.

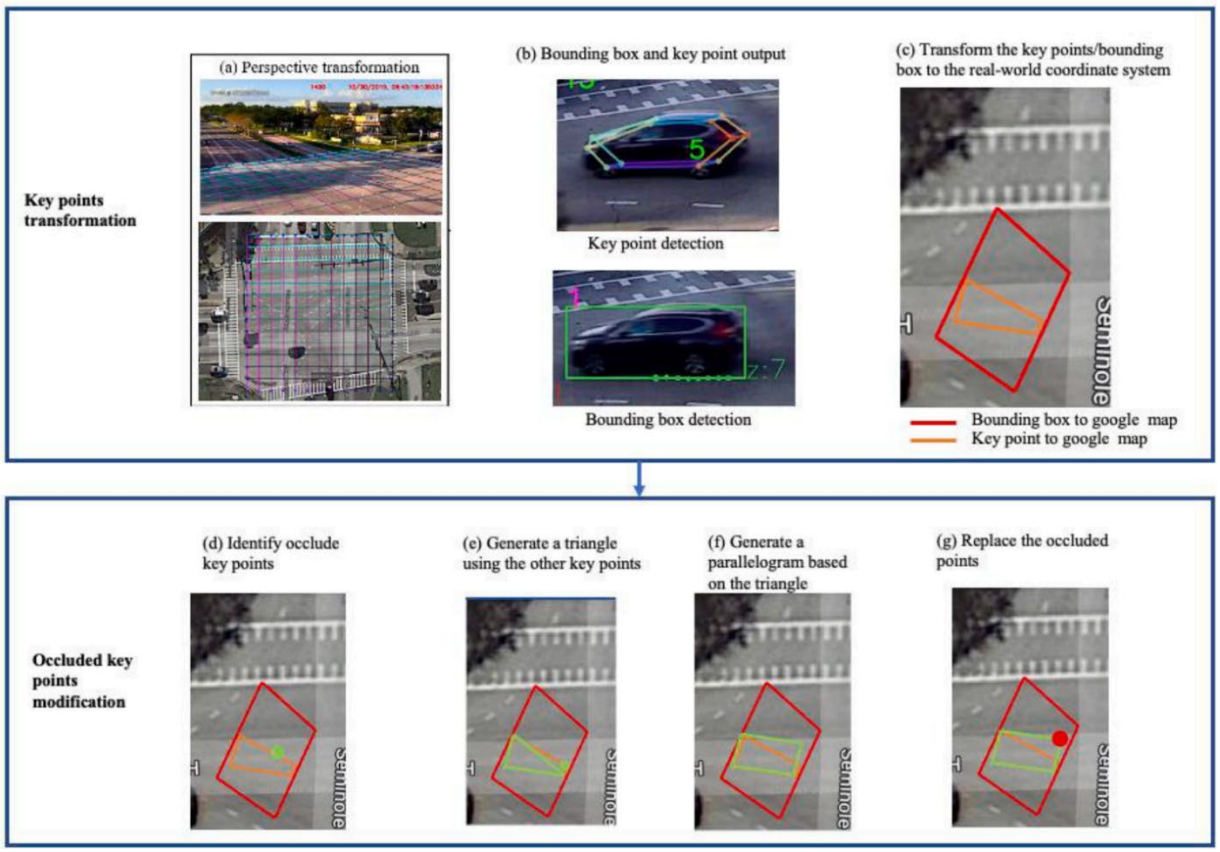


Figure 3 Occluded Key Point Modification [75]

The study was followed by another research work that presented an integrating detection system to overcome the addressed limitations by including CCTV cameras, LiDAR, sensor fusion for trajectory extraction, and safety evaluations [*76*]. The workflow was organized as follows: the real-time detection was performed through CCTV cameras and LiDAR simultaneously, then, the extracted detections were processed by the sensor fusion module to generate vehicle trajectories. The proposed methodology has proved its capability in real-time detection and tracking techniques. The provided precision rate of Camera and LiDAR are 90.32, and 97 percent, respectively, whereas the fusion technique had correctly detected 97.38 percent of the vehicles. Also, the fusion had a recall of 95.32 percent while Camera and LiDAR had recalls of 94.27, and 74.02 percent, respectively. The limitation of this study could be considered as the low recall percentage for the LiDAR, and the un existence of pedestrians, and bicyclists which exclude conflicts that relate to vulnerable road users.

3.2 Detection Methodologies Employment

Initially, the first step of setting a framework for the automated detection of the traffic conflicts was by determining various conflict types that required to be detected. Through an extensively prepared literature study, twenty types were determined including sixteen vehicle-to-vehicle conflicts, and four vehicle-to-pedestrian conflicts. It was found that various methodologies and techniques were performed by utilizing computer vision to detect moving objects (MOs) using several programming languages. The most commonly used language in recent years is Python due to its feasibility, and applicability within computer programmers' community. Hence, Python has been utilized for vehicle detection from the extracted video footage using various approaches. The following procedures are the commonly used techniques in the detection process:

- Background subtraction: This method involves subtracting the background image from each frame of the video and detecting the moving objects, which could be vehicles. Various Python libraries like OpenCV and scikit-image provide functions for background subtraction.
- Feature-based detection: This method involves detecting specific features or patterns that are indicative of vehicles, such as edges, corners, or blob-like structures. These features can be extracted using various techniques such as the Hough Transform, SIFT, or SURF. OpenCV and scikit-image provide functions for feature detection and extraction.
- Deep learning-based detection: This method involves training a deep neural network to recognize vehicles in videos. Popular deep learning frameworks in Python like TensorFlow and PyTorch can be used for this purpose. While other popular object detection algorithms like YOLO (You Only Look Once) and Faster R-CNN (Region-based Convolutional Neural Networks) can be used for vehicle detection.
- Optical flow-based detection: This method involves analyzing the motion of objects in consecutive frames of a video and detecting moving objects, which could be vehicles. Various Python libraries like OpenCV provide functions for optical flow-based detection.

3.2.1 Traditional Techniques in Detection of Road Users

Based on the stated procedures, the analysis of video footage was performed. Initially, the execution of the detection and tracking algorithm was performed using Python 3.11 as the programming language, and PyCharm as the Integrated Development Environment (IDE) for Python. The basic concept for the detection was adopted by integrating features-based, and background subtraction algorithms together. Accordingly, features clustering for each moving vehicle (frame-by-frame) was performed, then, applying a background subtraction technique to determine the surrounding boundary for each moving object (MO). Finally, creating a bounding box with a unique identification (ID) for each vehicle to attain its trajectory as well. Figure 4 illustrates features sorting process. A set of features including object size, shape, color saturation, texture, and motion intensity were selected within a set of limits to achieve a decent detection accuracy. However, the features clustering and matching process showed a promising result, it faced a number of limitations. These limitations could be interpreted as a result of over segmentation of the vehicles that move in close proximity to each other's with association to the camera's angle of view and with almost the same speed.

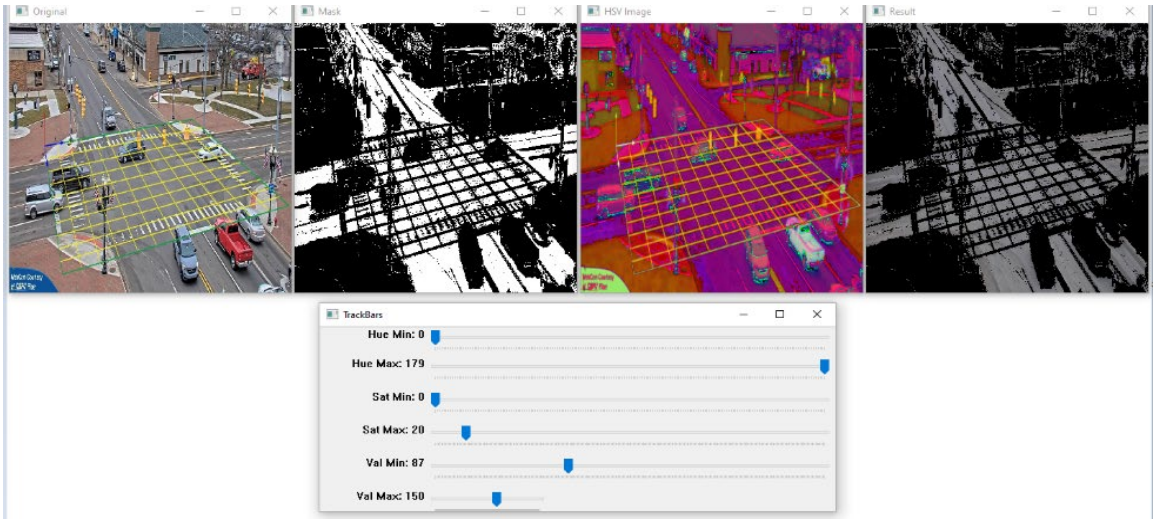


Figure 4 Features Clustering of The Moving Vehicles

The code was performed to detect, track, and count the vehicles that cross an imaginary set of lines which are selected at the stop lines of each approach. The determination of these imaginary lines is associated with the classification and counting of each turning movement volume. Although the algorithm succeeded in detecting vehicles and pedestrians, it faced two critical problems. These problems are the multiple detection of the same vehicle as well as over

segmenting the vehicles that move in the same boundary with almost the same speed. It should be noted as well that the code was examined at Four Corners Camera downtown intersection, city of Cold water, Michigan, to assess the detection and tracking process. Furthermore, a 2D-tranformation matrix was utilized to project the perspective view into a top view plan.

Figure 5 shows the utilized transformation matrix on the provided video footage of Michigan intersection. It is worth mentioning that the transformation to 2D top view aimed to ease the calculation of traffic stream parameters, as well as the traffic conflicts analysis. Additionally, by analyzing the traffic movements from the top view, the occurred distortion resulting from the perspective view is partially eliminated. Furthermore, the mapping results will provide a general aspect that could be utilized widely and followed without taking into consideration the camera's fixation location within the intersection.

a) Image Perspective b) Projected view

Figure 5 3D-2D Image Transformation

The limitations are summarized in the following bullet points:

- Repeated vehicles IDs.
- Overcounts for the crossing vehicles, Figure 6 .
- Over segmentation for the nearby vehicles associated with the provided angle of view, Figure 7.
- Miss detection of vehicles of specific color codes**,** Figure 8.

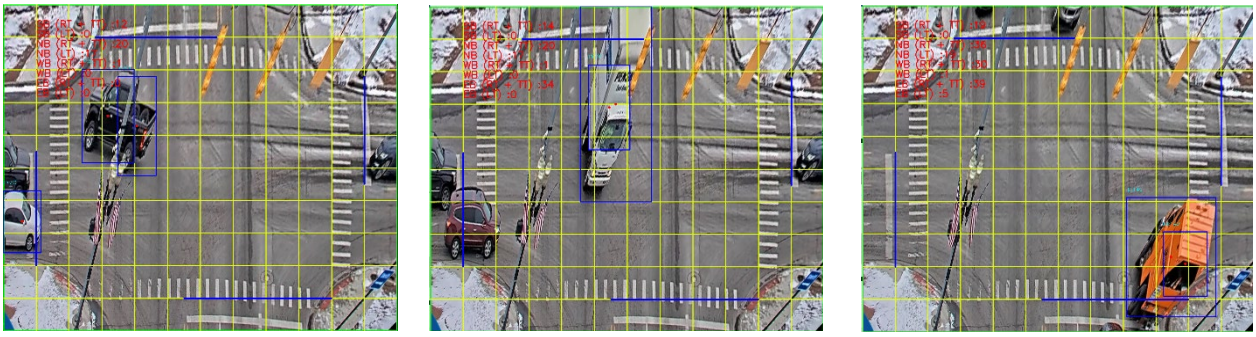


Figure 6 Over Count of Detected Vehicles

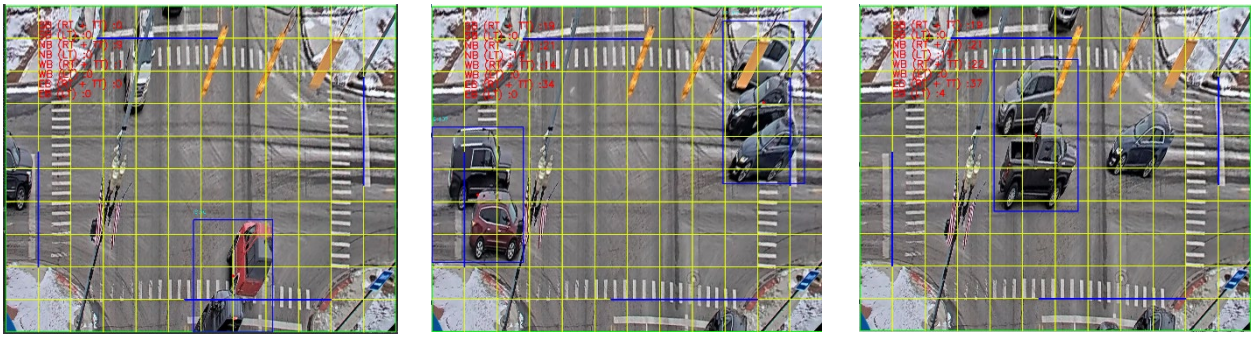


Figure 7 Over Segmentation of Detected Vehicles

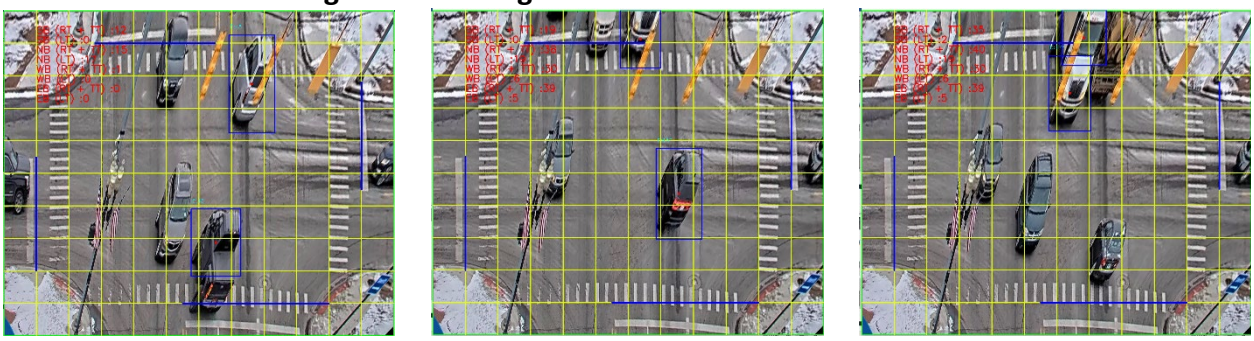


Figure 8 Miss Detection of Vehicles of Specific Color Codes

The extracted gaps could be addressed by fine tuning the inserted detection features options. However, the enhanced overall detection accuracy will not meet the requirements of the proposed traffic conflicts detection framework.

Based on the extracted conclusions from the utilizations of features clustering in the detection process, it was required to follow another approach to acquire higher levels of detection accuracy to serve in the traffic conflicts analysis efficiently. Consequently, a literature review was conducted to follow the recent research work in the object's detection, categorization, and labeling algorithms by utilizing deep learning approaches. Then, a comparison was set to select

the most appropriate algorithm to achieve a decent accuracy accompanied with a reasonable time frame for near real-time video analysis process.

3.2.2 Deep Learning Techniques in Detection

Vehicle detection using deep learning techniques was found to be an active area of research in recent years. Hence, various algorithms and architectures have been proposed for this task. The commonly used deep learning approaches for vehicle detection are YOLO (You Only Look Once), SSD (Single Shot Detector), and Faster R-CNN (Region-based Convolutional Neural Networks).

- YOLO: YOLO is a real-time object detection algorithm that can detect multiple objects, including vehicles, in a single pass through the network. It uses a single neural network to predict the bounding boxes and class probabilities for the detected objects. YOLO is known for its high accuracy and fast inference speed, making it a popular choice for real-time vehicle detection applications [78].
- SSD: SSD is another real-time object detection algorithm that can detect multiple objects, including vehicles, in a single pass through the network. Like YOLO, it uses a single neural network to predict the bounding boxes and class probabilities for the detected objects. SSD is known for its simplicity, speed, and high accuracy, especially for small objects like vehicles [79].
- Faster R-CNN (Region-based Convolutional Neural Networks): Faster R-CNN is a two-stage object detection algorithm that involves first generating a set of object proposals using a separate region proposal network (RPN), and then classifying and refining the proposals using a second-stage network. Faster R-CNN is known for its high accuracy, especially for complex objects and scenes, but it is also slower than YOLO and SSD [80, 81].

Generally, the choice of deep learning approach for vehicle detection depends on the specific requirements and constraints of the application. YOLO and SSD are suitable for real-time applications that prioritize speed and efficiency, while Faster R-CNN is suitable for applications that prioritize accuracy. Further research and development in deep learning-based vehicle

detection algorithms are expected to improve their accuracy, efficiency, and suitability for a wide range of applications [82, 83].

However, YOLO and SSD were found to be the most common algorithms in vehicle detection that will achieve the required overall detection accuracy within the time frame, it was found that YOLO is outperforming SSD in popularity and applicability within the programming community. Consequently, YOLO was chosen as the proposed detection technique at this level of analysis.

At the time of analysis, YOLOv7 was published and was pre-trained on various imagery datasets. However, a summarized overview will be provided for the development of different YOLO algorithms as follows, Table 3:

Table 3 Comparison Between YOLO Algorithm Versions

YOLO version	Utilized Neural Network and Architect formation	Training Procedure	Training Data set	Overall Accuracy
YOLOv1 [75]	Darknet-19, which is a 19-layer convolutional neural network with max pooling, batch normalization, and leaky ReLU activation.	The model is trained on a dataset of labeled images using backpropagation and stochastic gradient descent (SGD). The loss function used is a combination of localization loss (which penalizes errors in bounding box location and size) and confidence loss (which penalizes errors in objectless prediction). During training, the input images are divided into a grid of cells, and each cell predicts a fixed number of bounding boxes and	Combination of PASCAL Visual Object Class (VOC) 2007 and PASCAL VOC 2012	63.4 percent mean Average Precision (mAP) on the VOC 2012 test set.

YOLO version	Utilized Neural Network and Architect formation	Training Procedure	Training Data set	Overall Accuracy
		corresponding class probabilities.		
YOLOv2 [84]	Darknet-19, similar to YOLOv1, but with some additional layers and a skip connection.	The training procedure is similar to YOLOv1, but with some improvements. The loss function now includes a term for classification loss, which penalizes errors in class prediction. Additionally, the model is trained on multiple scales of the input image, which helps improve detection accuracy for objects at different sizes.	Combination of PASCAL VOC 2007, PASCAL VOC 2012, and MS COCO for the testing, which consists of over 330,000 labeled images of 80 object categories.	78.6 percent mAP on the MS COCO test-dev set with a Darknet-19 network.
YOLOv3 [85]	Darknet-53, which is a 53-layer convolutional neural network that includes residual connections and feature pyramid networks.	The model is trained on a large-scale dataset (such as COCO) using the Darknet framework. The training procedure includes data augmentation techniques (such as random scaling, cropping, and flipping) to increase the diversity of training examples. The loss function includes terms for localization loss, confidence loss, and classification loss, and is	MS COCO dataset for the testing.	57.9 percent mAP on the MS COCO test-dev set with a YOLOv3-608 model.

YOLO version	Utilized Neural Network and Architect formation	Training Procedure	Training Data set	Overall Accuracy
		optimized using the Adam optimizer. During training, the model uses feature pyramid networks to perform object detection at multiple scales.		
YOLOv4 [86]	CSPDarknet-53, which is a variant of Darknet-53 that uses cross-stage partial connections to improve information flow and reduce computation.	The training procedure is similar to YOLOv3, but with some improvements. The model is trained on a larger dataset and with more data augmentation techniques (such as mosaic data augmentation). Additionally, the loss function is modified to include a term for focal loss, which helps the model focus on hard examples during training. The training procedure also includes techniques such as self-adversarial training and label smoothing.	Combination of MS COCO, WIDER FACE, and Open Images datasets.	43.5 percent mAP on the MS COCO test-dev set with a YOLOv4-416 model.
YOLOv5 [87, 88]	A variant of Cross-Stage Partial Network (CSPNet) that includes a modified backbone network and detection head.	The model is trained on a dataset of labeled images using the same training procedure as YOLOv4. However, the architecture of YOLOv5 is designed to be more	A variety of datasets were used including MS COCO, PASCAL VOC, and Crowd Human.	51.3 percent mAP on the MS COCO test-dev set with a

YOLO version	Utilized Neural Network and Architect formation	Training Procedure	Training Data set	Overall Accuracy
		lightweight and efficient, which allows for faster training and inference times.		YOLOv5s model, and up to 57.1 percent with larger models.
YOLOv7 [89]	YOLOv7 Backbone, which is a novel backbone network that combines the features of EfficientNet and Squeeze-and-Excitation modules, along with a detection head like YOLOv5.	The model is trained on a dataset of labeled images using backpropagation and a similar loss function as YOLOv5, with the addition of some novel techniques for data augmentation and regularization.	A large-scale dataset with over 20 million images for pre-training and fine-tuned on COCO and other datasets.	51.2 percent mAP on the MS COCO test-dev set with a YOLOv7 model.

The overall conclusions of the comparison between YOLO versions could be stated as follows:

- Training Procedures: The training procedures for YOLO have evolved with each version, with YOLOv4 and YOLOv7 using similar training procedures that involve data augmentation and multi-scale training. YOLOv5 introduced a novel approach called "Mosaic Data Augmentation" that improved the accuracy of small object detection.
- Number of Convolutional Layers and Networks: The number of convolutional layers and networks in YOLO have also evolved with each version, with YOLOv4 and YOLOv7 having more convolutional layers and advanced backbone networks that improve accuracy. YOLOv5 introduced a new network architecture called "CSPNet" that improved accuracy while reducing computational requirements.

- Accuracy: Overall, the accuracy of YOLO in vehicle detection has improved with each version, with YOLOv7 achieving the highest accuracy among all versions evaluated. However, it should be noted that accuracy may vary depending on the dataset and specific application.
- Speed and Computational Requirements: YOLO has consistently maintained real-time speed and low computational requirements throughout its evolution, with each version introducing new optimizations and network architectures that further improve performance.

Based on the comparison, it was concluded that YOLOv7 was the most convenient algorithm to be utilized in the study of traffic conflicts analysis. **Figure 9** shows the comparison between YOLOv7 performance and other YOLO versions used in real-time detection algorithms.

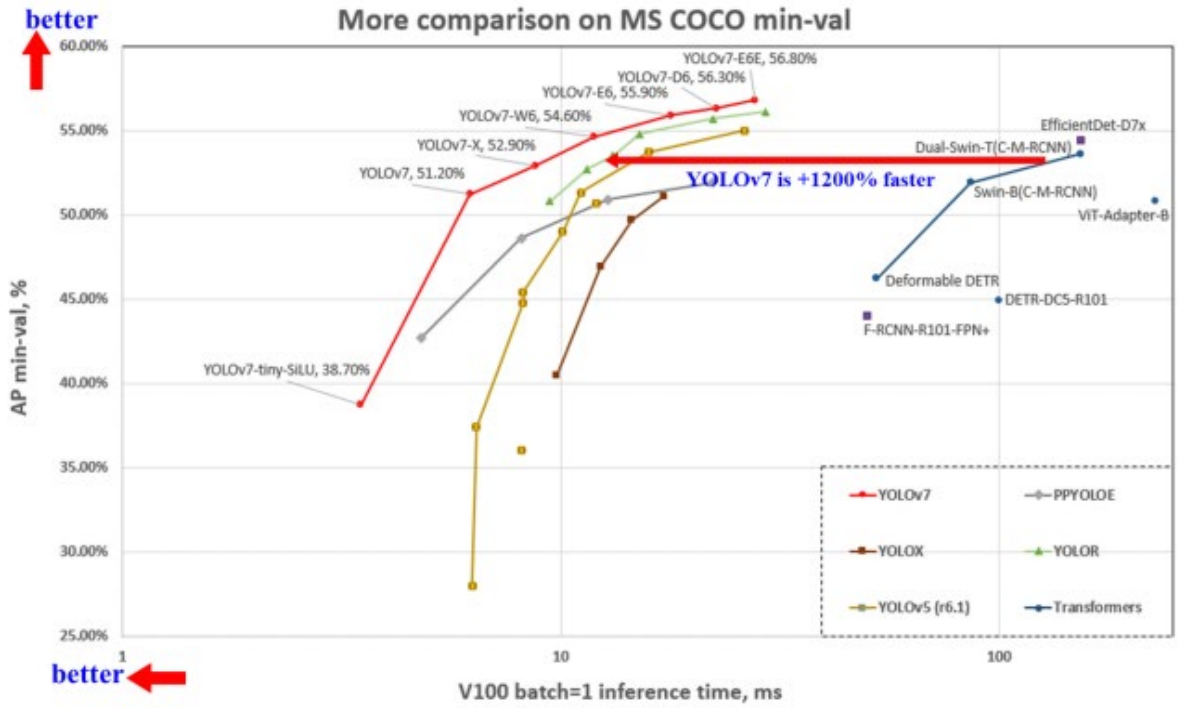


Figure 9 Comparison Yolov7 to Other Real-Time Object Detectors [89]

3.2.2.1 YOLOv7 Employment in Road Users' Detection

YOLOv7 detection and tracking algorithm was utilized in running the analysis of the extracted set of video footages for two of the cases of study intersections (e.g., Jackson Hole Town Square in Wyoming, and Four Corner Camera intersection in Cold Water city, Downtown of Michigan).

For the initial trials, 20 second. video was analyzed to estimate the time frame and the computational power of the utilized device (Intel(R) Xeon(R) CPU E3-1240 v3 @ 3.40GHz 3.40 GHz, 16.0 GB installed RAM, 1 GB NVIDIA Quadro K620). It was found that the analysis took 205 second., and 650 second. by in charging the GPU, and CPU to perform the video analytic process, respectively. Based on this conclusion, it was required to either enhance the analysis computational power of the utilized personal computer (PC) or utilize an external graphic processing unit (GPU) for performing the analysis. Hence, Google external GPU was used through Google Collaboratory system to upload and run the video analysis. It should be mentioned that Google Collaboratory, also known as Google Collab, is a cloud-based development environment provided by Google for machine learning education and research. It allows users to write and execute Python codes in a web browser without requiring any setup or installation. For the 20 second. video, the elapsed time of the analysis performed on Google Collaboratory was 41 second. Consequently, Google Collaboratory environment was selected in running the analysis process. Figure 10 and Figure 11 show samples of video analysis for Jackson Hole Town Square, and Four Corner Camera intersection in Cold-Water Downtown in Michigan intersections using YOLOv7 detect and track algorithm.

Figure 10 Video Analysis of Jackson Hole Intersection Using YOLOv7

However, the extracted results for the detection showed a decent accuracy, the algorithm faced one drawback by inaccurately classifying some of the 4-wheel drive passenger cars as trucks. While for the tracking and labeling objectives, the vehicles that are located on the far side of the camera position fail to be tracked in some frames within the video footage which resulted in taking different label associated with different trajectory. These limitations could be addressed by fine tuning the detection options and narrowing the scope of the detection to include the near side approaches to the camera's view and excluding the far side approaches.

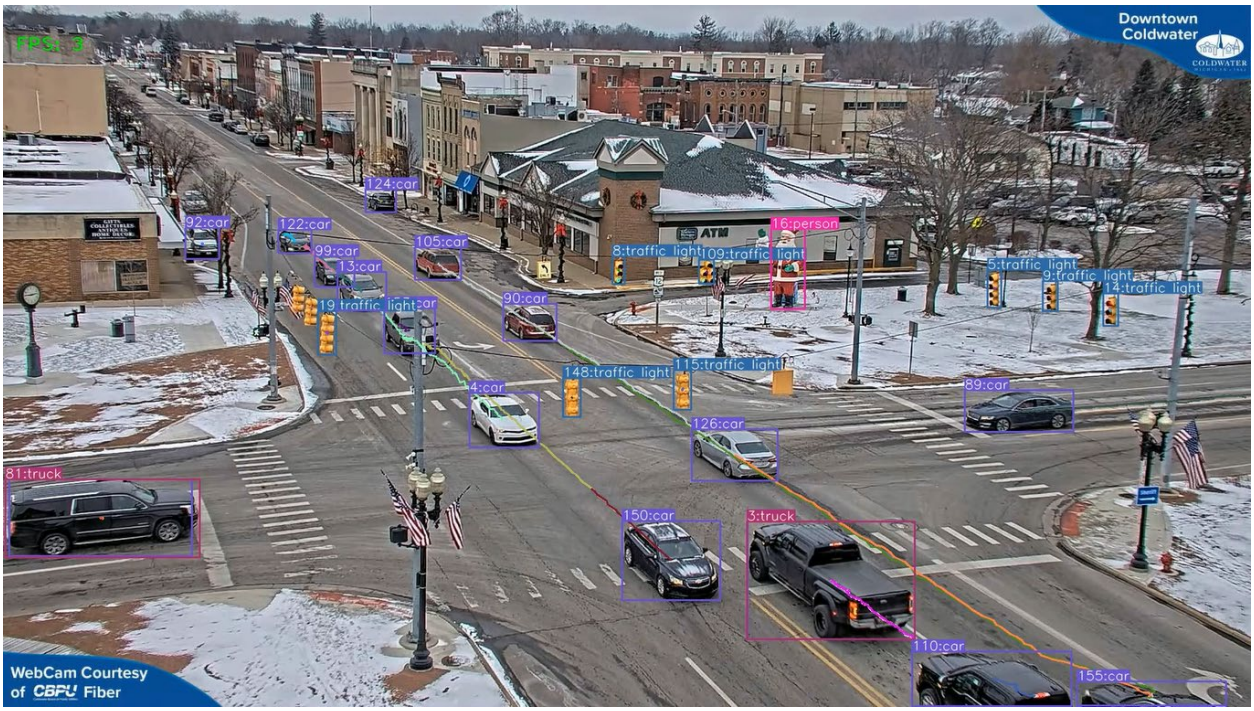


Figure 11 Video Analysis of 4 Corner Camera, Cold-Water City, Michigan Using YOLOv7

Another critical limitation faced the transformation of the centroid and bounding boxes corners coordinates from the 3D view to 2D plan view. This limitation is restricted to the camera's mounting elevation. Since the followed procedure in performing the 3D-2D transformation homograph depends on a set of selected points on the ground level, the output grid is to be constructed on ground level. Hence, any point that does not belong to the ground level will be distorted after projection. The lower the camera was fixed; the more distortion will be resulted after the transformation is conducted. However, this distortion could be neglected for the high-level fixation points of cameras. Assuming that the centroids of the Moving Objects (MOs) are located on the same height from the ground, the amount of distortion will be similar for both

projected objects. Consequently, the utilized measurements will not be critically affected by the distorted projection for the higher elevations of camera's fixation points. To clarify this limitation the transformation of two different frames were performed; the first frame was associated with Jackson Hole intersection (low elevation of camera fixation point), and the second frame was cropped from Cold Water intersection (high elevation of camera fixation point). Figure 12, and Figure 13 illustrate the initiated limitation for the transformation of the videos that are taken from lower cameras mounting elevations.

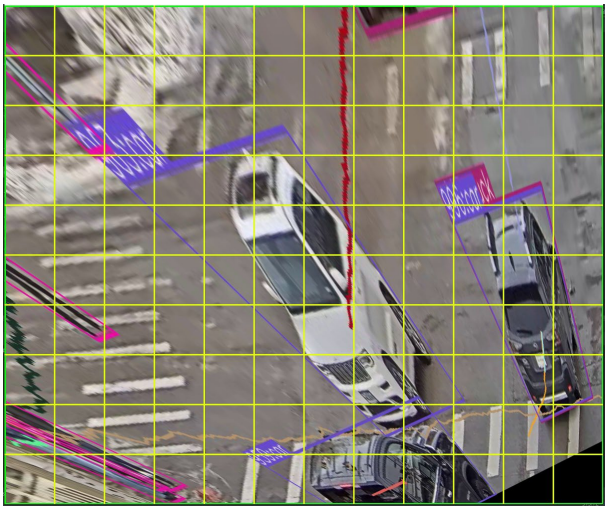


Figure 12 Detection and Tracking for Low Elevations of Camera Mounting Point Using YOLOv7

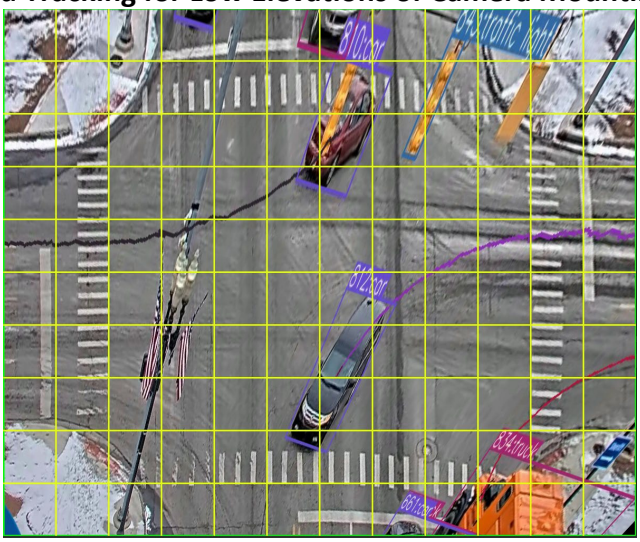


Figure 13 Detection and Tracking for High Elevation of Camera Fixation Point Using YOLOv7

Based on the extracted results from the provided transformations, it was found that for the lower-level fixation cameras points (10-15 ft), YOLOv7 will not be able to serve in the conflict analysis framework unless the selected points for the transformation grid fall on the same

height of the majority of road users' centroid. This required further data acquirement from the field of study itself for the purpose of initiating a unique grid at a special level. However, this could not be applicable for all similar intersections while looking for generalizing the procedure. Hence, YOLOv7 detection and tracking algorithm is to be applied on the intersections of high-elevation fixation points (40-45 ft) for webcams by assuming that the centroids of all road users are located at the same height. While for the lower-elevation fixation points, another approach should be followed.

3.2.2.2 CenterTrack Algorithm Employment in Road Users' Detection

At this stage, the research was directed to study the trajectory level algorithms that were designed for autonomous vehicles. Since these algorithms were developed originally for in-vehicle cameras application, they may accommodate the required vehicle detection and tracking using low-level fixed cameras similar to the one at Jackson Hole Town Square.

The research on trajectory level detection algorithms led to the most recent dataset named "nuScenes" that is used for the training process for convolutional neural networks in the field of vehicle detection algorithms related to autonomous vehicles. nuScenes is a large-scale dataset for autonomous driving developed by nuTonomy, which contains high-resolution sensor data collected from a variety of sensors including LiDAR, radar, and cameras. The dataset includes over 1,000 scenes captured in urban environments, with each scene consisting of 20 seconds of data captured at 20 frames per second.

The nuScenes dataset provides a benchmark for evaluating algorithms for object detection, tracking, and prediction in complex urban environments. It includes annotations for a variety of road users including cars, trucks, pedestrians, cyclists, and scooters. The annotations provide information about the object's class, bounding box, and orientation, as well as additional information such as the object's velocity, acceleration, and heading angle.

In addition to the dataset, nuScenes also provides a set of baseline models for object detection, tracking, and prediction. These models are based on state-of-the-art deep learning algorithms such as CenterNet and PointPillars and have been trained on the nuScenes dataset to achieve high accuracy in object detection and tracking. The baseline models can be used as a starting point for developing more advanced algorithms for autonomous driving.

The research result led to the utilization of CenterTrack algorithm. This algorithm was presented through the original research paper "Objects as Points" by Zhou et al. that provided a comprehensive explanation of the CenterTrack algorithm and its implementation [90]. Additionally, the official GitHub repository for the CenterTrack algorithm contains the source code and a detailed description of the algorithm and its components. The paper proposed a new methodology by representing each object as a single point, then, used it to predict the object's bounding box and its class. This approach differs from traditional object detection methods that use bounding boxes to localize objects.

The method is based on a key observation that most bounding boxes of objects have a similar aspect ratio, which makes it possible to use a single point to represent the center of the object and its size. The paper showed that this approach is not only more accurate but also faster than existing state-of-the-art object detection methods such as Faster R-CNN and RetinaNet. Additionally, it introduced a new training strategy that simplifies the process of training object detectors. Instead of using traditional object-level annotations that require drawing a bounding box around the object, the paper proposed using point-level annotations that only require marking the center point of the object. This approach reduced the annotation effort required and enabled the use of larger datasets for training.

CenterTrack proposed an end-to-end solution for monocular 3D object detection and tracking by exploiting the complementary nature of 2D object detection and 3D object localization. Instead of using bounding boxes to localize 3D objects, it formulates 3D object detection as a keypoint estimation problem, where the location of the object is represented by a set of center points.

The authors of "Objects as Points" argued that this keypoint-based representation was more efficient and effective than traditional bounding boxes, as it allowed for better localization and orientation estimation of 3D objects, especially under challenging occlusion scenarios.

The proposed method consisted of two stages. The first stage was a modified version of the YOLOv3 network, prediction of the center point, height, width, and orientation of the object. The second stage used a simple 2D-to-3D lifting module to obtain the 3D location of the object from its 2D center point. Moreover, the authors introduced a new concept called "associating

the unassociated " which refers to the ability of the model to detect and track objects that were not previously associated with any existing tracks. This is achieved by introducing a new association metric that combines the appearance similarity and motion similarity between object detections. Experiments on the KITTI and nuScenes datasets showed that CenterTrack achieved state-of-the-art performance in both 3D object detection and tracking, surpassing previous methods that use either 2D or 3D information alone. Furthermore, the proposed method is computationally efficient, with a processing speed of over 25 frames per second, making it suitable for real-time applications.

Overall, CenterTrack is a state-of-the-art algorithm for multi-object tracking in complex scenes. Its real-time performance and ability to track multiple object types make it well-suited for applications such as autonomous driving, surveillance, and robotics.

Based on these results, CenterTrack was selected for the analysis of intersections of low-level webcams fixation points, whereas YOLOv7 was utilized for webcams with higher fixation points. Figure 14 shows the video analysis for Jackson Hole intersection using CenterTrack algorithm. It should be noted that the video analysis was run on Google Collab to provide a reasonable time frame for the computational analysis.

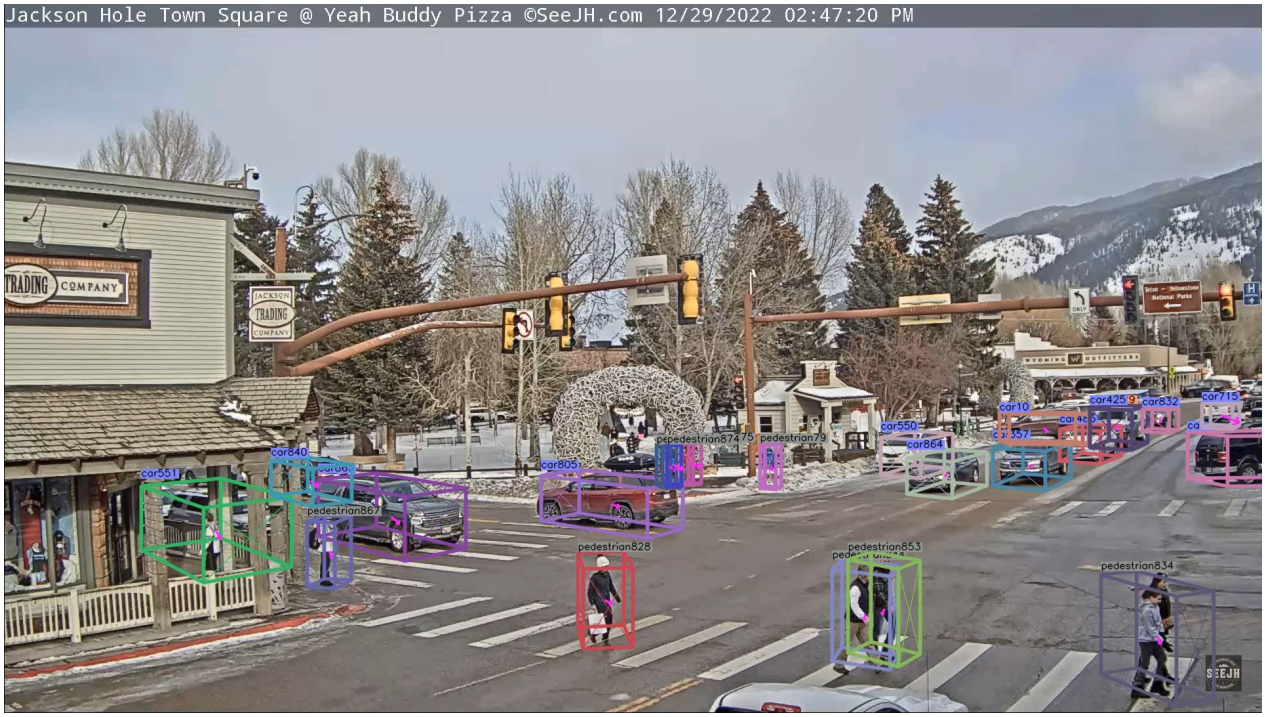


Figure 14 Detection and Tracking for Low Elevation Camera Fixation Point Using CenterTrack

According to the video analysis outputs, CenterTrack has achieved decent accuracy at low-level fixation points for webcams (10-15 ft). Additionally, it provided a well fitted occupation cuboids for road users even at complexes situation and at occlusions. The next step was the identification of the coordinates of the 4 points located on the ground level for each road user and track them frame-by-frame.

The extraction of the coordinates of the four points that were located on the ground level required some modifications to the original code. Consequently, a deep assessment of the code was performed to acquire the coordinates of the cuboid associated with the detected MOs in each frame of the video footage. The IDs of these points were located as 4, 5, 6, and 7. Additionally, the study of the main code resulted in the main concept behind the design of the algorithm. It should be noted that CenterTrack algorithm is setting constraints about the rotation of each (MO) around its center point. It was assumed that the (MO) is rotating about its vertical axis (y-axis) only, while no rotation was permitted around other axes (x, and z). For future studies, several modifications could be performed for generalizing the code for various mounting heights of the cameras.

The base points coordinates were extracted from the code in pixels units, then, the analysis was performed based on this acquirement. Figure 27 illustrates the bounding cuboid corners IDs.

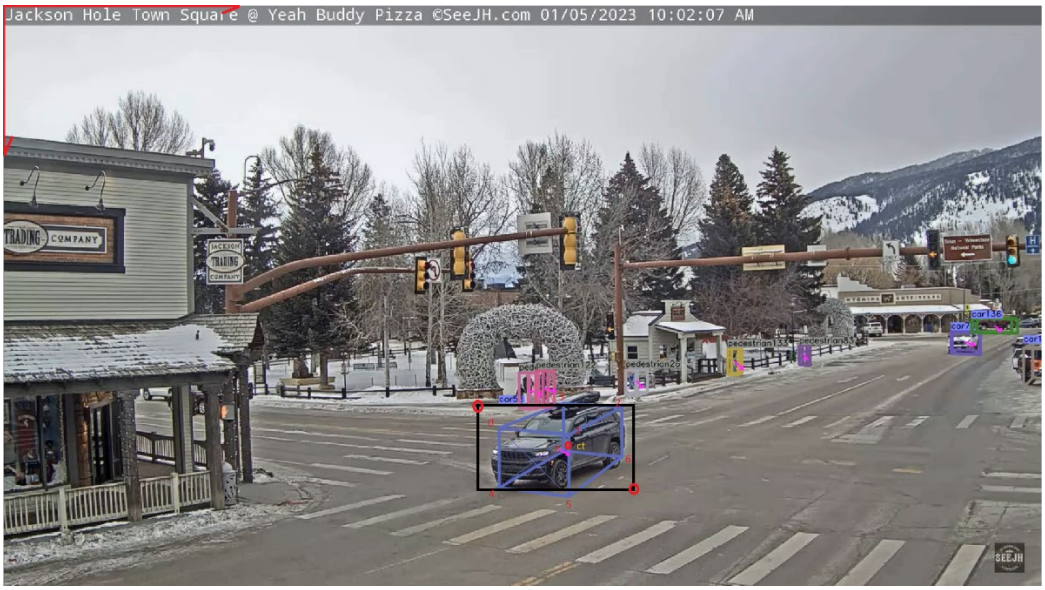


Figure 15 Monocular 3D Bounding Cuboid Corners IDs

3.2.2.2.1 Limitations of CenterTrack Utilization

The limitation of utilizing CenterTrack in the conflict analysis could be summarized as follows:

- Large deformation for the cuboids including road users when approaching near to the camera's location (empirical condition embedded in the original code).
- Suitable for video footage from low-height webcams only.
- Some misclassifications for the far objects from the webcam lens (e.g., pedestrian with trash cans).
- Misclassification between 4WD, and SUV vehicles with trucks.

For the first and second limitations, it is concluded that an essential modification should be applied to the algorithm for addressing these limitations. While the last two limitations could be addressed in the analysis process. For the misclassifications between the far objects, it could be set as a condition if the object is moving in different frames, it could be selected as pedestrian, elsewise, it may be excluded from the analysis. For the misclassification between SUVs and trucks, since the bounding cuboid has the same coordinates and dimensions, the vehicle could be re-classified based on its extracted dimensions. In the following chapter, further analysis of the extracted outputs will be performed to remove the biased data, extract trajectories, and smooth them.

3.2.2.3 OpenPifPaf Algorithm for Vehicles Key Points Detection

The CenterTrack algorithm provided accurate detection for vehicles and pedestrians that are located within the intersection areas. However, it faced a limitation to provide accurate dimensions for each road user since the cuboids are based on average vehicles and human dimensions and their depth in the image perspective. Hence, another two algorithms were examined for extraction of more accurate detection results. The first one is vehicle key points detection algorithm, "Open Pose in Full Pose Articulation Framework" (OpenPifPaf). While the other one is for pedestrian key points detection, YOLOv7_human_pose_estimation.

The original algorithm by Sven Kreiss et. al. presented a new bottom-up methodology for multiple detections of human poses [91]. Accordingly, the study was followed by several plugins' implementation for the algorithm on different classes (e.g., vehicles, and animals) [92].

For the vehicles key points detection, *ApolloCar3D* Dataset was utilized for the training process. The Dataset included 5,277 driving images involving more than 60,000 cars in their captions with a resolution of 4K. The prediction of vehicles key points utilized the pretrained model "Shuffle Net".

Through this study, the vehicles key points detection was performed by applying OpenPifPaf standard command "predict" and a flag "--checkpoint=shufflenetv2k16-apollo-24" that represent the vehicle with 24 key points. Since this research work is investigating the integration of multiple cameras by utilizing the top-down view as the method of representation, the key points of interest will be that they are located at/near to the ground level. Figure 16 demonstrates vehicle's structure and the 24 key point distributions utilizing the employed flag. The IDs of the 8 key points that are located at/near to the ground level are illustrated in the figure.

However, the algorithm achieved decent accuracy in the detection of vehicles key points, it faced one drawback in the detection of the occluded points. Figure 17 presents an analyzed video frame using OpenPifPaf.

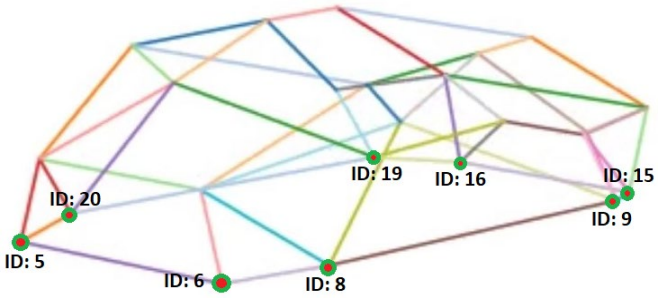


Figure 16 Vehicle Key Points Identification At/Near to Ground Level

Figure 17 Detected Key Points Using a Single Camera

Based on this drawback, it was decided to integrate the utilized surveillance cameras at the intersection to employ OpenPifPaf to depict the detected vehicles by accumulating the detected key points from various cameras. Consequently, employing OpenPifPaf in a conflict analysis framework. The integration was done by utilizing YOLOv7 and OpenPifPaf and leveraging the multiple cameras that are located at the intersection. Figure 18 illustrates the detection framework for a single vehicle.

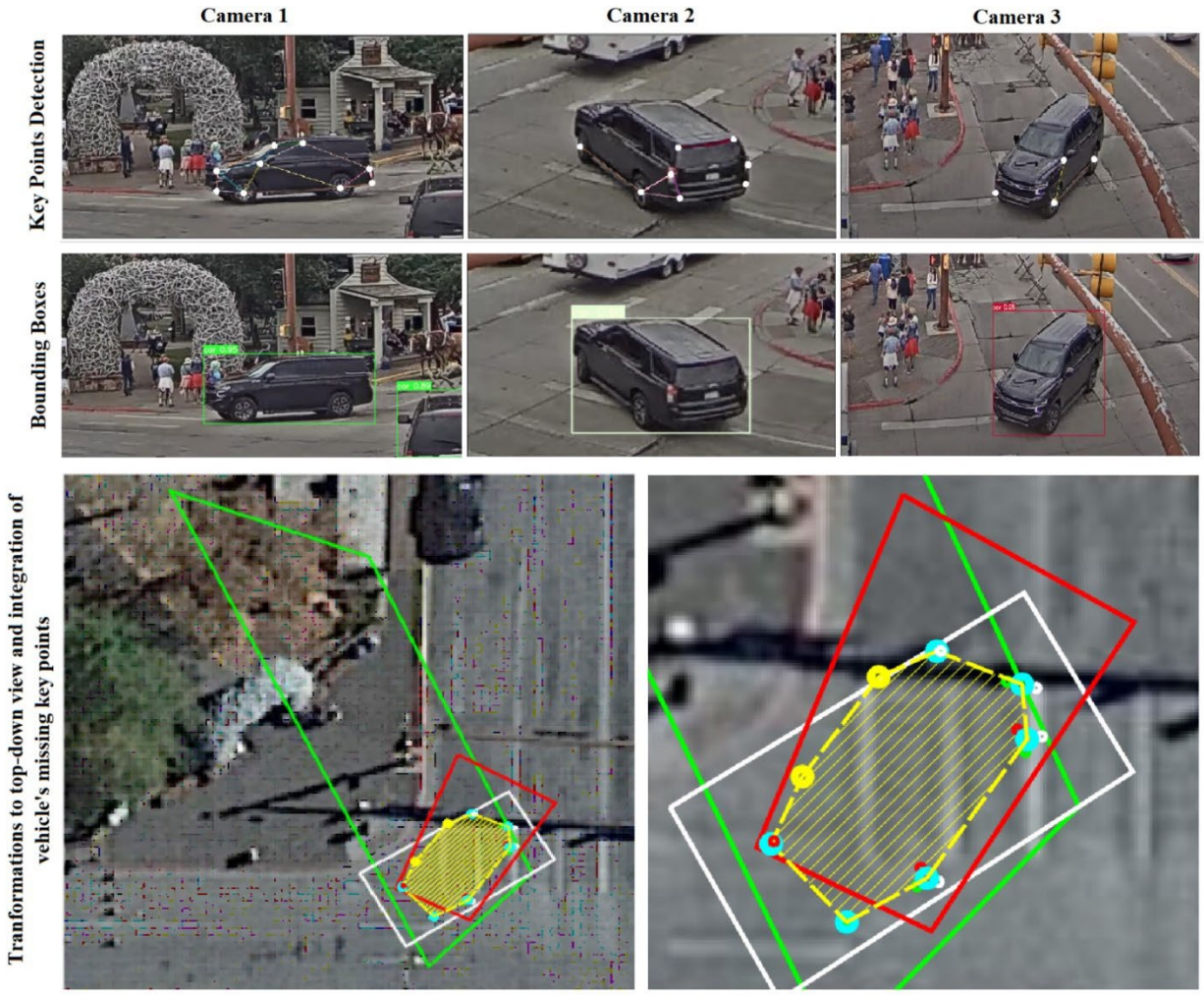


Figure 18 Detection Framework Utilizing OpenPifPaf

In the Figure, the average coordinates of the detected key points with more than one camera were calculated while the undetected points were assumed based on the average distances between the key points. However, the proposed method was put to improve the detection accuracy and provide more accurate projections for vehicle outlines by matching the projected key points on plan view, it faced a transformation problem which was resulted from the distorted images of cameras. Figure 19 shows the projected key points after transformation. Based on the projections output, it was decided to assume the average vehicle dimensions, Figure 21, on plan view by clustering the projected points, Figure 20. An extensive study was performed to restore accurate positions from the linearly and curvilinearly distorted images which will be explained with the application of pedestrian key points detection.

Figure 19 Extracted Key Points from Multiple Cameras.

Figure 20 Key Points Clustering

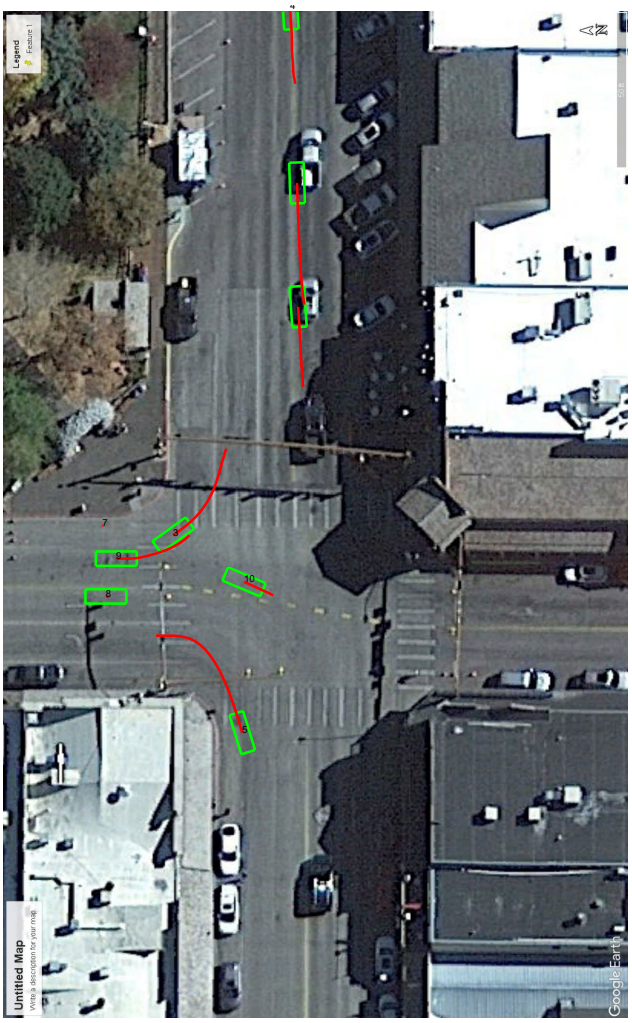


Figure 21 Projected Vehicles Shapes Based on Average Vehicle Dimensions.

3.2.2.4 YOLOv7 Human-Pose-Estimation Algorithm

The human pose estimation algorithm is published on the GitHub repository of Rizwan Munawar, YOLOv7-pose-estimation. The algorithm provides a bounding box for each detected person, and 17 key points distributed between human joints, and face structure. Figure 22 illustrates the key points IDs using the pretrained model “yolov7-w6-pose.pth”.

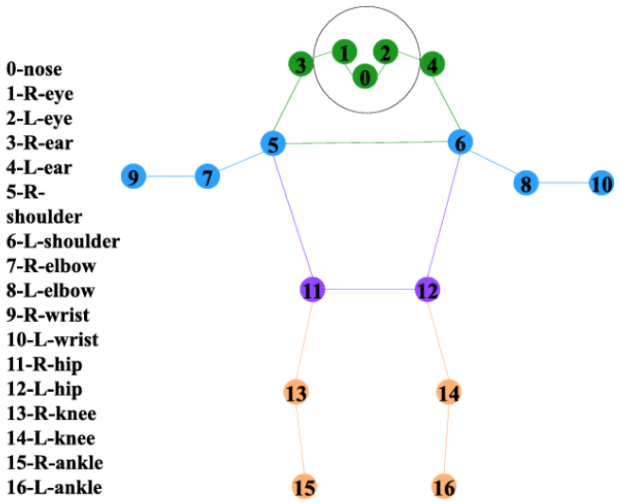


Figure 22 Human Key Points Distribution [93]

The main goal of this section is to extract pedestrian trajectories at crosswalks by accurately detecting their feet key points and tracking them through video recordings. Furthermore, the proposed method utilizes the three surveillance cameras mounted at Town Square intersection, Jackson Hole to improve the accuracy of trajectory extraction at the bounds of each camera. Additionally, a correction strategy was adopted from linear and curvilinear perspective principles to relieve the distortions in cameras outputs resulting in maximizing the feature points and providing more accurate positioning.

The initial steps were organized as follows:

- Examine the field of view for each camera and set the intersected areas between them, Figure 23.
- Set the areas of interest, crosswalks, and set the coverage area of each camera beneficially, Figure 24.

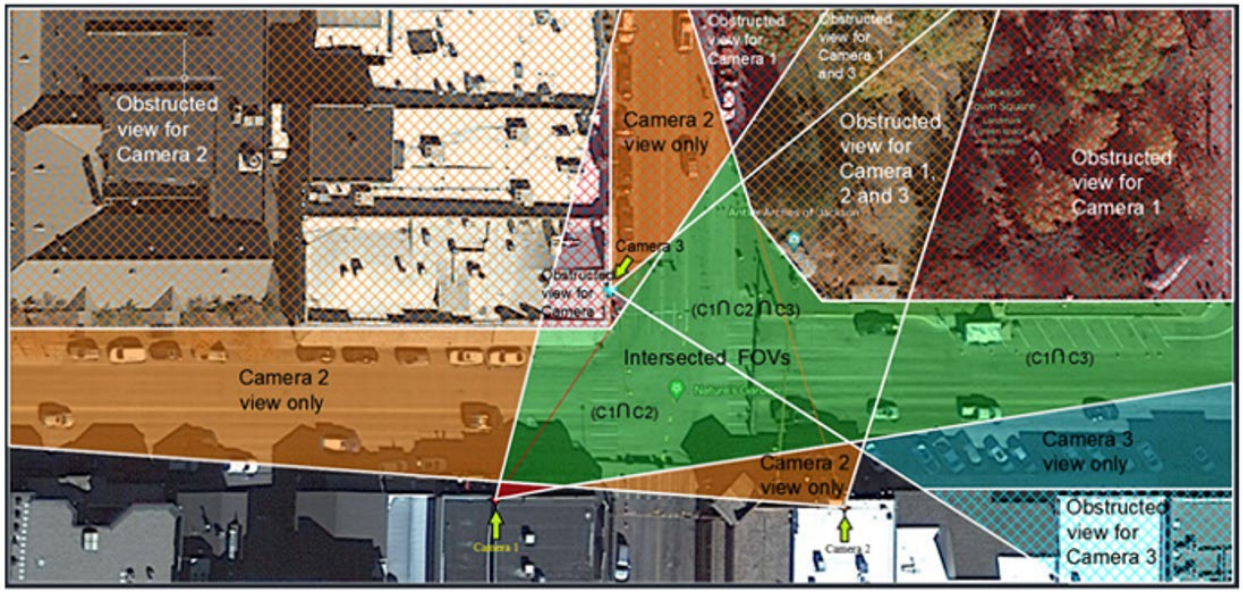


Figure 23 Cameras Coverage and Intersected Fields of Views

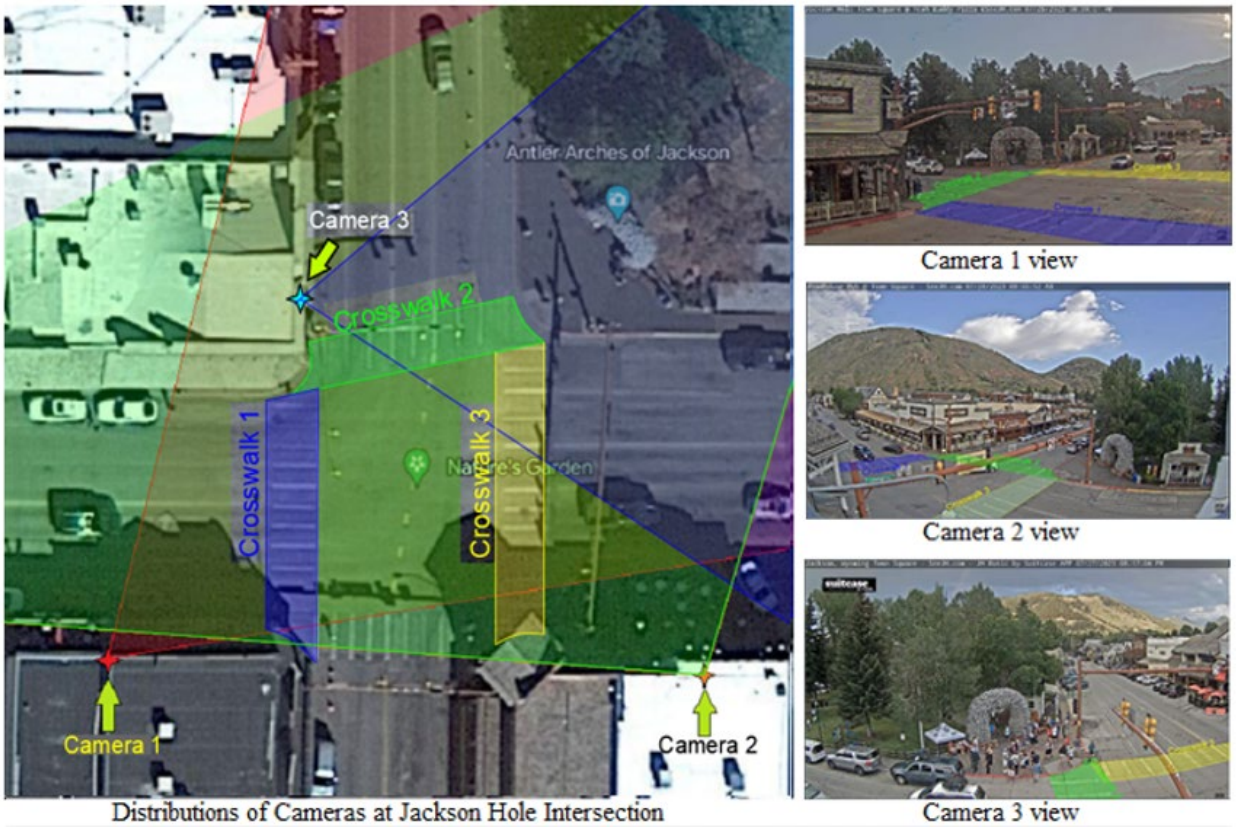


Figure 24 Cameras Positions and Crosswalks determination at Site of Interest

The next step is to improve the transformation accuracy by performing image processing correction methodology based on the linear and curvilinear perspective properties and set a unified grid for each crosswalk between multiple cameras.

3.2.2.5 Linear perspective

In linear perspective each couple of horizontal parallel lines meet at a single point called vanishing point. The group of vanishing points are located at a horizontal line called the horizon line. This line is positioned at the same height of fixation of the camera that provides the linearly distorted image. Figure 25 illustrates the followed technique in creating the grid points for the crosswalks from camera 1-output images.

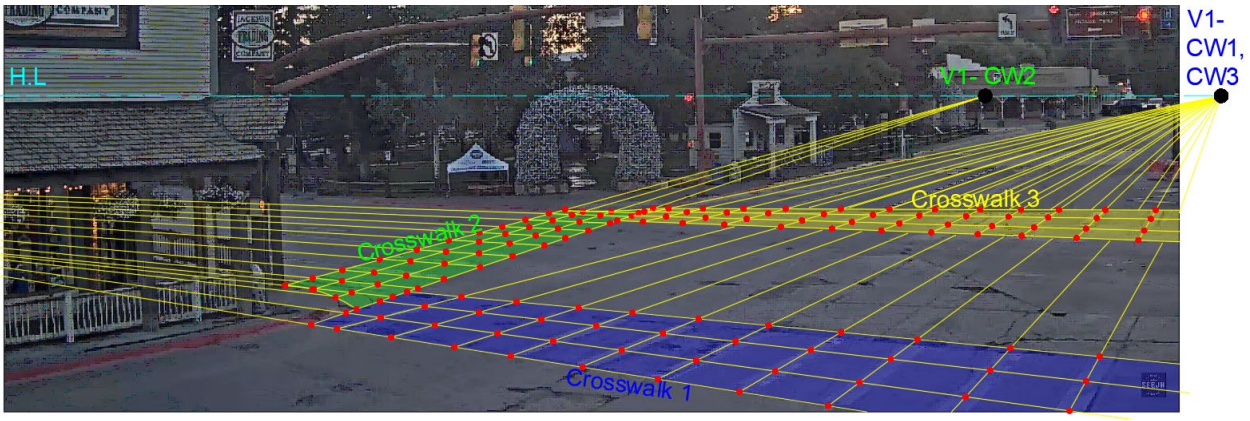


Figure 25 Gridlines Formation on Camera 1 Image Perspective

Where, H.L refers to horizon. V1-CW1, CW3: is the first vanishing point for crosswalks 1, and 3, while V1-CW2: is the first vanishing point for crosswalk 3.

3.2.2.6 Curvilinear perspective

Both horizonal and vertical lines are represented in curvilinear perspective as arc sections. The arc type, whether it is elliptical or circular arc depends on the shape of cameras lens. If the lens is spherical or semi-sphere, the arcs will be circular. If the camera has paraboloidal lens the arc will be elliptical. For practical use, most curvilinear surface cameras are manufactured as semi-sphere.

In the semi-spherical cameras' outputs, each couple of horizontal lines curves will meet at two vanishing points at each end. The line joining the vanishing points of horizontal lines arcs is a horizontal line. Similarly, each couple of vertical lines arcs will meet at two points that are locate on a vertical line. Figure 26 illustrates the extraction of the bounding lines of crosswalks and their grid points. The extracted grid points from the multiple cameras are matched through the plan view to unify the output coordinates of pedestrians. Hence, the integration between cameras outputs is performed to extract accurate pedestrians' trajectories. Figure 27 shows the

grid points on plan view. Consequently, a code was developed to utilize the outputs from multiple cameras and calculate the corrected trajectories by combining the coordinates and remove the biased data, Figure 28.

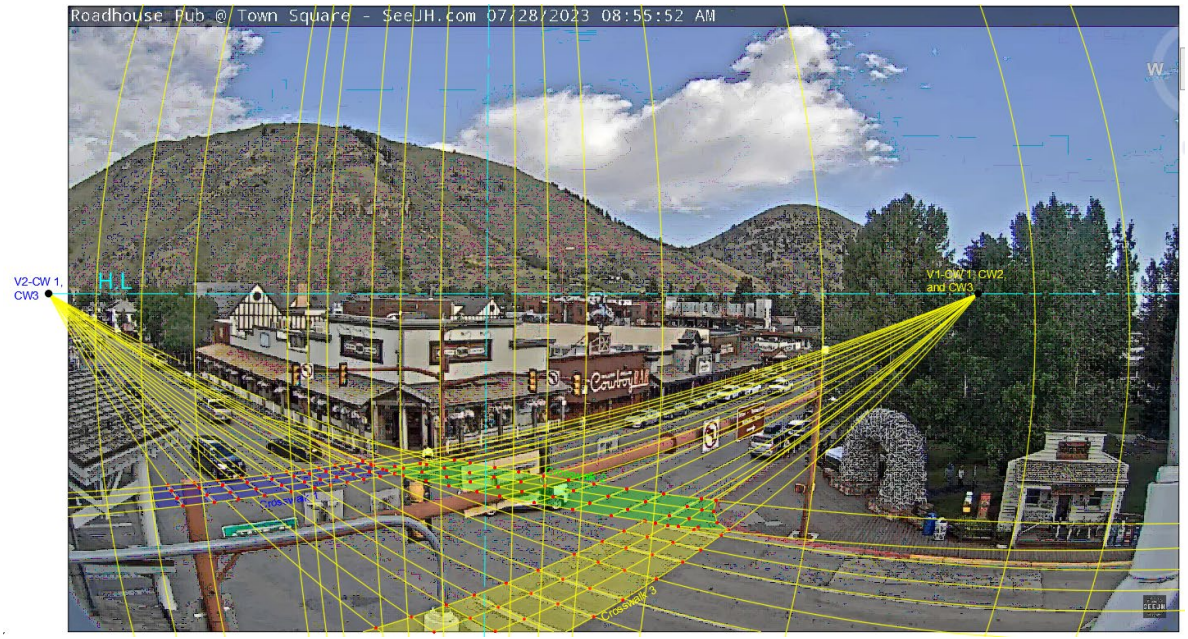


Figure 26 Gridlines Formation on Camera 2 Curvilinear Perspective

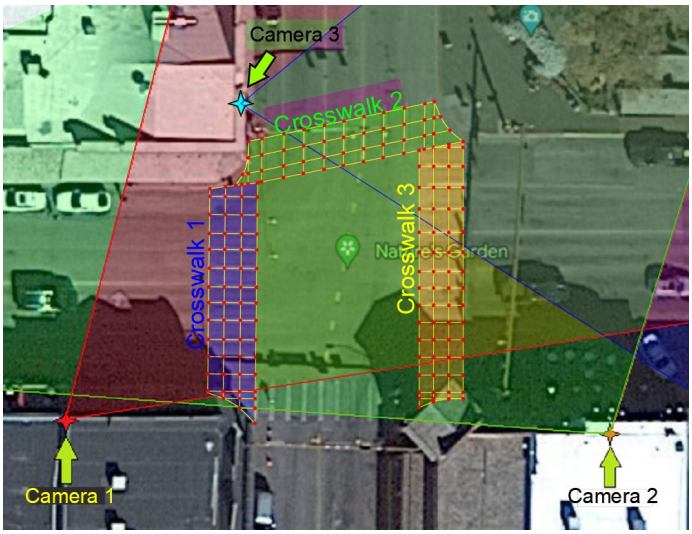
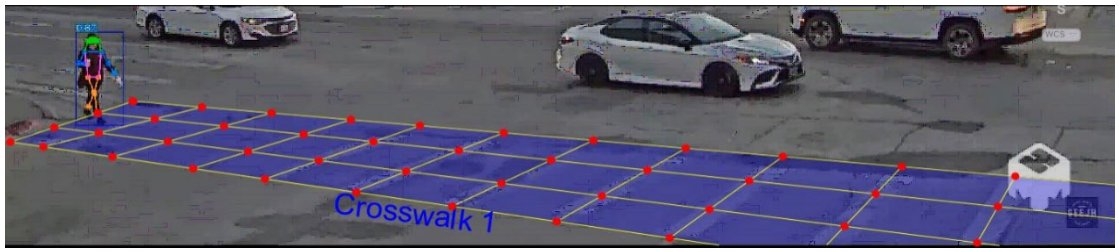
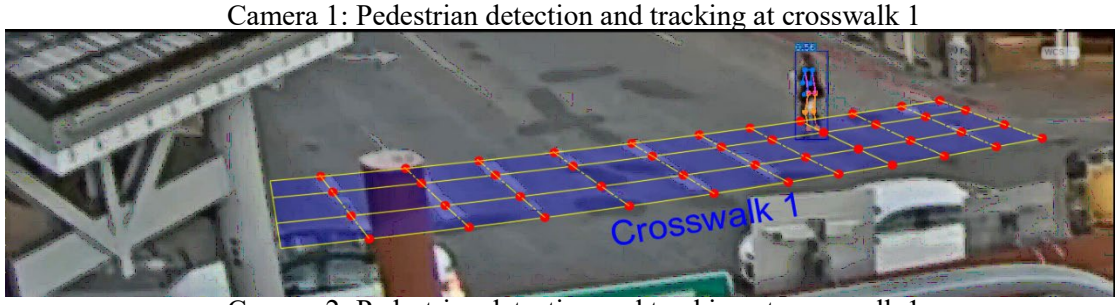


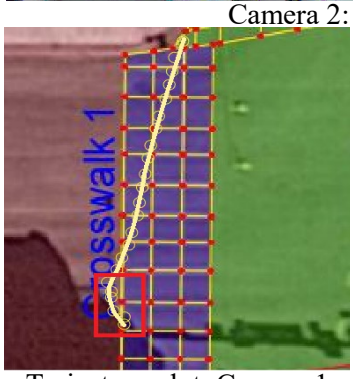
Figure 27 Gridlines of Crosswalks on Plan-View



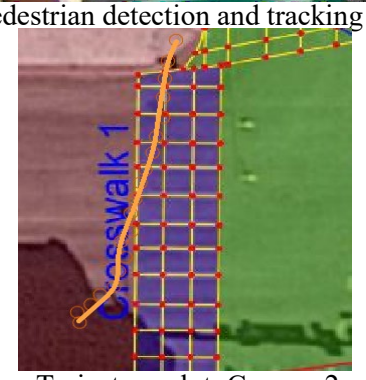
Camera 1: Pedestrian detection and tracking at crosswalk 1



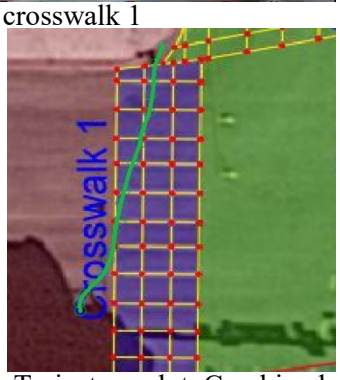
Camera 2: Pedestrian detection and tracking at crosswalk 1



Trajectory plot: Camera 1



Trajectory plot: Camera 2



Trajectory plot: Combined

Figure 28 Detection and Tracking of Pedestrian Using the Proposed Framework

The combined trajectory extraction by employing multiple cameras could correct the mis tracking that occurred by a single camera. This could be illustrated from the plotted trajectory from camera1. The red rectangle refers to false tracking of pedestrian at the exit of the covered area by camera 1. By employing and combining the data from camera 2, the corrected data could be produced to improve the accuracy in the trajectory plotting.

In the next chapter, the utilized post processing techniques to clean the extracted outputs and provide smoothed trajectories from CenterTrack and OpenPifPaf algorithms will be presented.

CHAPTER 4. POST PROCESSING OF ALGORITHMS OUTPUTS

In this chapter, the post processing steps that are applied to extract smoothed trajectories from both CenterTrack and OpenPifPaf algorithms will be presented.

4.1 Post Processing Steps Utilizing CenterTrack as Detection Algorithm

To address the needed post processing steps, a general framework for conflicts detection was proposed, **Figure 29**.

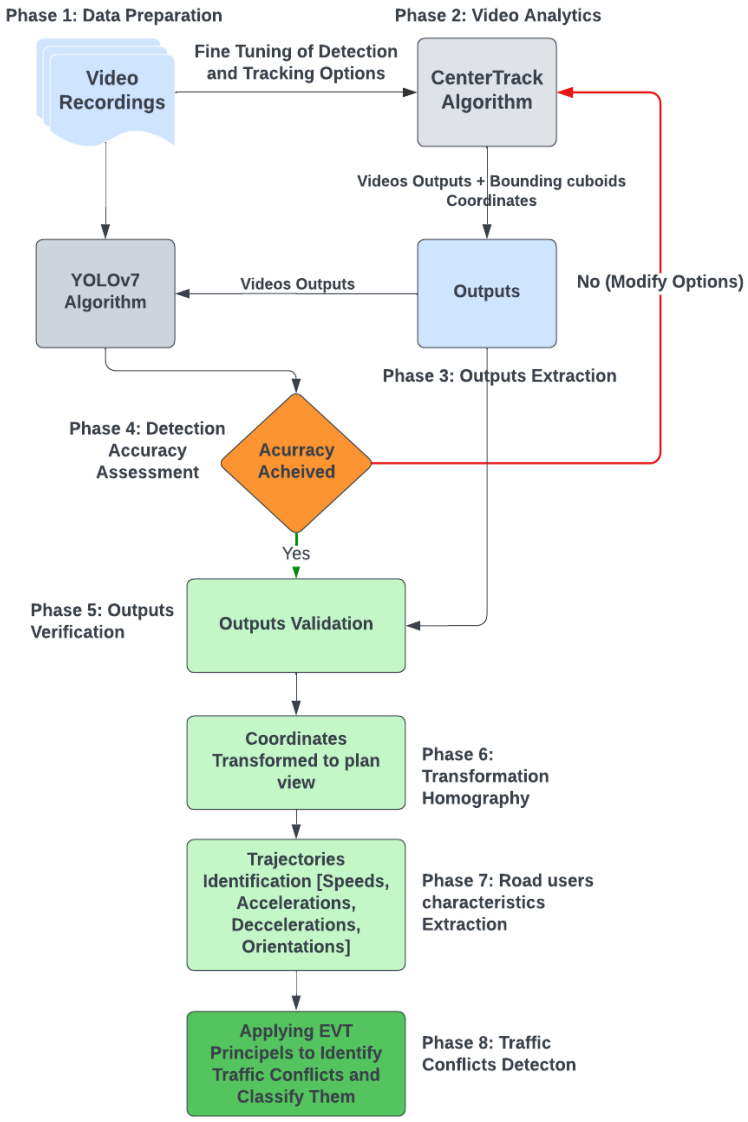


Figure 29 Traffic Conflicts Detection and Analysis Framework

The demonstrated conflict detection framework is utilized in the analysis of a conflict set that will be presented throughout chapter 4.

4.1.1 Output Data Validation Process

According to the nature of the extracted outputs from CenterTrack algorithm, it was required to perform an additional step to assign the extracted coordinates from the TXT file to their

associated road users in JSON file. Also, some road users were double detected (e.g., some vehicles were labeled as cars, and trucks). Accordingly, the resulted redundancy from the multiple detection was handled through this step. MATLAB 2022 was utilized for the alignment process. A MATLAB code was created such that the extracted data from JSON file was organized, then, the number of the detected road users is found. Accordingly, the number of rows from TXT file associated with these road users were selected to be in the same frame (e.g., eight rows for each road user). Then, a polygon was drawn between the lower four points for each cuboid to assess the selection and annotation process visually. This procedure was followed frame-by-frame with a simplification in the MATLAB code such that it was selected from the TXT file the number of detected road users in JSON file in each frame, respectively, without the need of rows calculation. Finally, the doubled coordinates related to the over detected objects were deleted through an enhancement in the code. **Figure 30** shows the visual validation of video output utilizing CenterTrack algorithm.

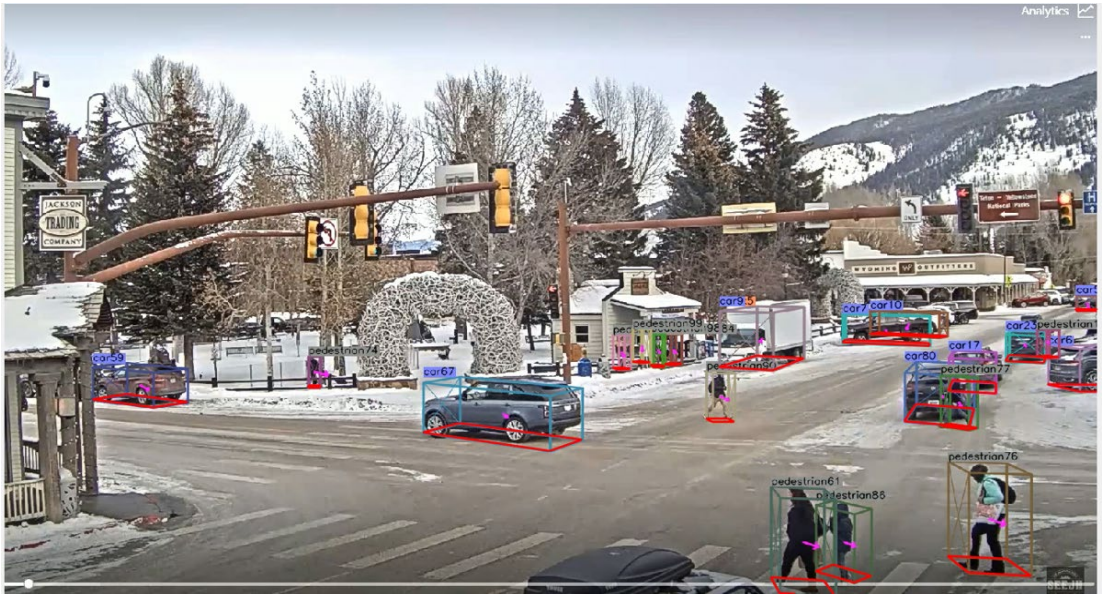


Figure 30 Data Validation Process Using MATLAB Code Applied on CenterTrack Video Output

Referring to Figure 30, the red quadrilateral shapes were drawn using the extracted coordinates from TXT file. A visual revision was set to verify the alignment of these shapes with the existing road users and their location within the image plane at a specific chosen frame. The verification process was done for conflict set 1 by following this procedure frame by frame. It should be noted that data verification process is considered the most time-consuming process throughout the whole analysis duration.

For the YOLOv7 algorithm, since the extracted outputs are decently organized and well separated, data validation process was exclusively included the doubled detection handling. After achieving data validation, it was required to apply the transformation homograph from the image plane to the 2D top view to obtain the traffic stream characteristics and draw road users' trajectories.

4.1.2 Transformation Homography

The transformation homography is the mapping matrix that is required to re-organize the extracted coordinates from the image plane into the 2D top view. As illustrated earlier, this transformation is needed to put all the extracted points on a common grid. Consequently, this will ease the calculation of traffic stream characteristics, as well as the future contributions to the research work without considering the utilized camera's position and properties.

The initial step for homography creation is to select a set of features (e.g., lines, points) that are located in the camera's view (image) and determine their association on the 2D top view. The required number of features should be greater than or equal to four geometrical elements. Obviously, these elements should be found on both images and utilized 2D view. It is favorable for these elements to be long lasting fixed structures since the moving objects are not optimal for locating the feature points.

For the analysis of Town Square intersection's video set, the crosswalk markings were selected as the geometrical features from camera's view for performing the transformation process. Since not all the markings could be viewed thoroughly, it was required to reconstruct an image with neither road users nor snow covering the crosswalks. Therefore, a short period of the video from YouTube channel (https://www.youtube.com/watch?v=1EiC9bvVGnk) was selected for performing such process. By taking the median value of the entire video, all road users and unfavorable surface conditions were removed. Hence, an image showing the clean road surface was reconstructed. Figure 31 shows the reconstructed image that was utilized in transformation homograph.

Figure 31 Jackson Hole Intersection Reconstructed Image for Features Extraction

The feature points were selected from the shown image in Figure 31. Google Earth was used to locate and match the selected feature points on the 2D top view. Figure 32 illustrates the selected points in the image plane (red points) and their associated points from the top view (green points).

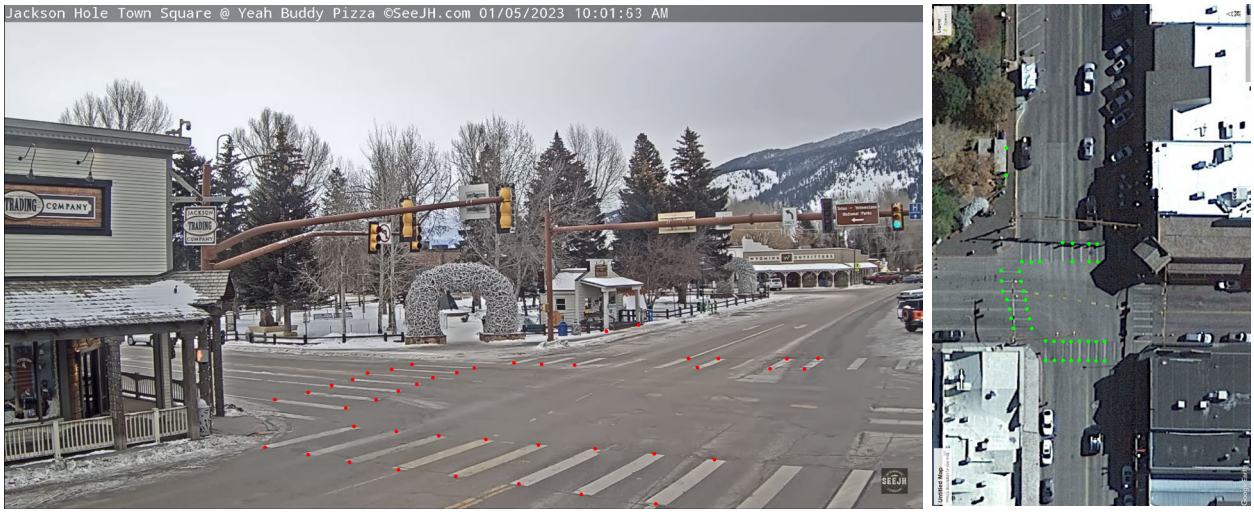


Figure 32 Selected Feature Points and Their Associated Points on Top View

The transformation was performed with respect to the selected feature points and their associated points. It should be mentioned that due to the selection of features points within the intersection area the accurate outcomes for transformation will be associated to this area.

It was challenging to explore more points that fall out of this area due to the poor road infrastructure outside the intersection as well as the image decaying in the far side from the camera's view. The transformation was applied on the reconstructed image to illustrate the occurred distortion outside the intersection area, Figure 33. Accordingly, the resulting distortion will affect the coordinates transforming accuracy for road users that located further from the selected features points (intersection area).

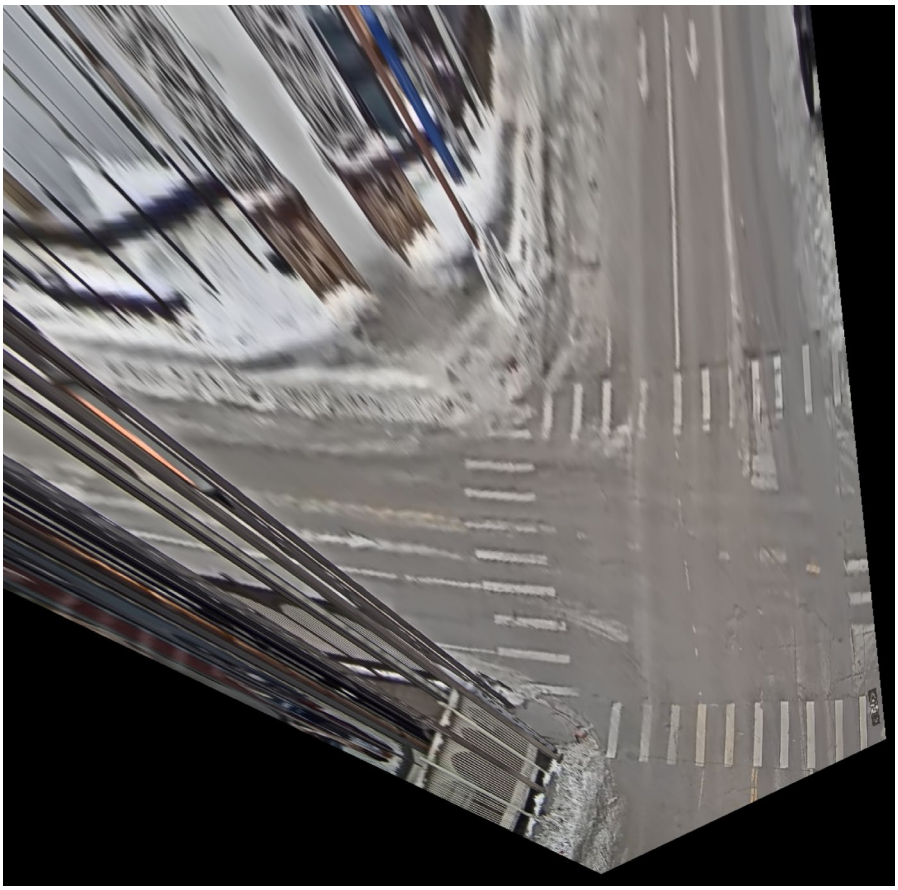


Figure 33 Transformed Image of Jackson Hole Intersection

The transformed road users' coordinates from camera's view were plotted on the Google Earth top view image. Figure 34 shows the projected road users in a single frame. The image shows the accurate plotted rectangles for the road users fall near to the selected features points and the distorted rectangles for the far road users from the intersection area. The following step is to draw the road users' trajectories, and to extract positions, orientations, speeds, and acceleration and deceleration rates. By attaining these characteristics, conflict analysis is to be done.

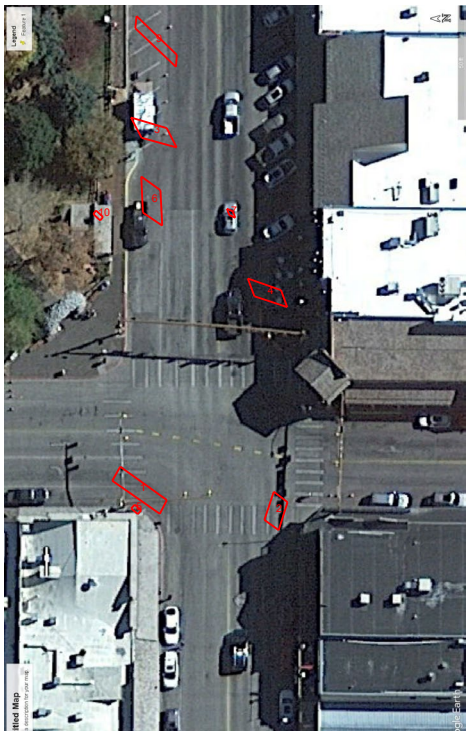


Figure 34 Projected Vehicles Coordinates in Google Earth Image

4.1.3 Road Users' Trajectories Extractions

The tracking process for the road users is organized in the following steps:

1. Start from the first or another specific frame that the tracking process needs to be initiated. Then, select the vehicle ID that is required to be tracked from JSON file.
2. By utilizing MATLAB code for the data validation process, the required vehicle's data is matched from TXT and JSON files.
3. Project the ground coordinates to the top view (red quadrilateral shapes). The background image is from Google Earth.
4. The road user trajectory is drawn by following the center point of the projected quadrilateral frame by frame.

Figure 35 shows the constructed trajectory for one selected vehicle. The red points are the initial coordinates of the vehicle, the green points are the center points, and the blue spline is the detected vehicle trajectory.

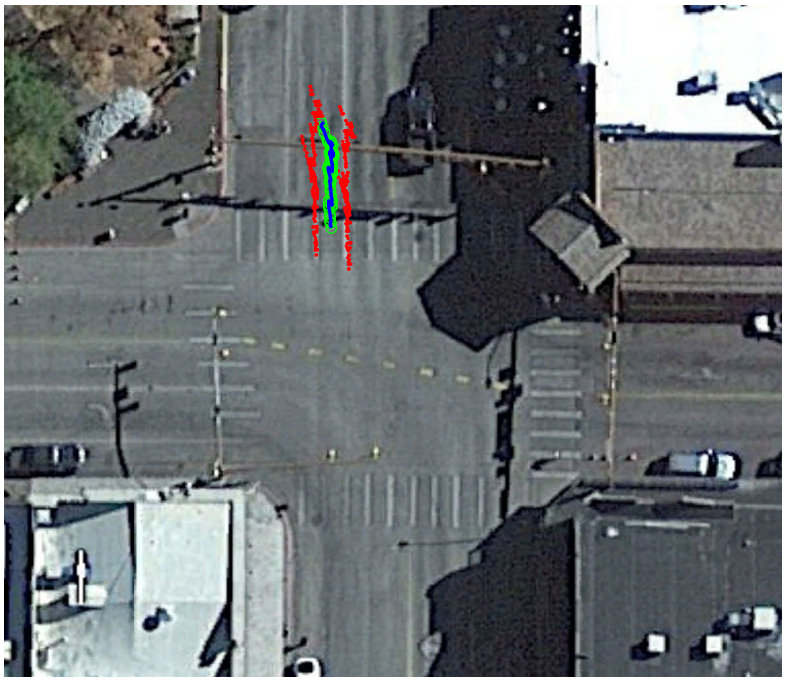


Figure 35 The Detected Coordinates and Trajectories

4.1.3.1 Trajectories reconstruction

However, each tracked vehicle should have four coordinates located on ground level, some duplicates occurred because of the multiple detections of same vehicle, and mis detections (objects categorized as different class in frame or more within the output video). Two errors were defined as the error with previous frame data and error with next frame data. Therefore, an algorithm of coordinates recognition is proposed based on the defined errors. The following two equations illustrate the calculation of these errors such that Eq.2. calculates the error in the previous frame, and Eq.3. calculates the error in the next frame.

$$\boldsymbol{E_1(t) = \sum_{i=1}^{4} (x_i(t) - x_i(t-1))^2} \tag{2}$$

$$\boldsymbol{E_2(t) = \sum_{i=1}^{4} (x_i(t) - x_i(t+1))^2} \tag{3}$$

A correction process was conducted in two steps by setting the following two conditions:

1. Calculate the minimum value from previous, and next frame Eq.2., and Eq.3.

$$\begin{cases} \boldsymbol{Min\ E_1(t,i)} \\ \boldsymbol{Min\ E_2(t,i)} \end{cases} \tag{4}$$

2. Calculate the difference between the errors in the previous and next frame.

$$\boldsymbol{Min\big(E_1(t,i) - E_2(t,i)\big)} \tag{5}$$

The "error difference" refers to the time difference between the current error and the previous error, as well as the time difference between the current error and the next error in a time-series dataset. Minimizing this error difference helps to maintain data consistency and prevent sudden outliers in the trajectory data. If the error difference exceeds a certain threshold, it will be flagged as an outlier or a miss-tracking event.

The road users' coordinates of each frame that matched above equations will be recognized as same vehicles. The outliers are removed from trajectories. Therefore, gaps will exist in some frames. For the gap/missing tracked data, using interpolation to fix the missing data. Figure 36 shows the corrected trajectories for two vehicles traveling through the north bound, and east bound after correction was conducted.

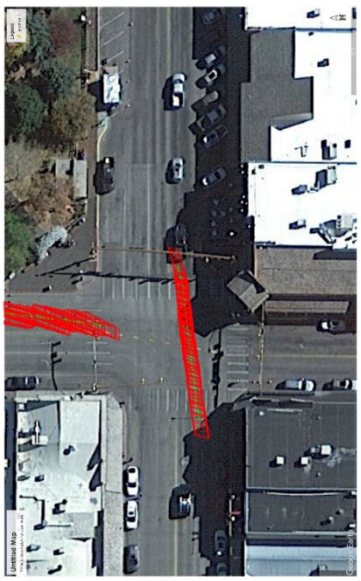


Figure 36 Using Ground Coordinates to Reconstruct Vehicle Trajectories

Finally, the vehicle trajectories were created for the road users travelling through the intersection as shown in Figure 37.

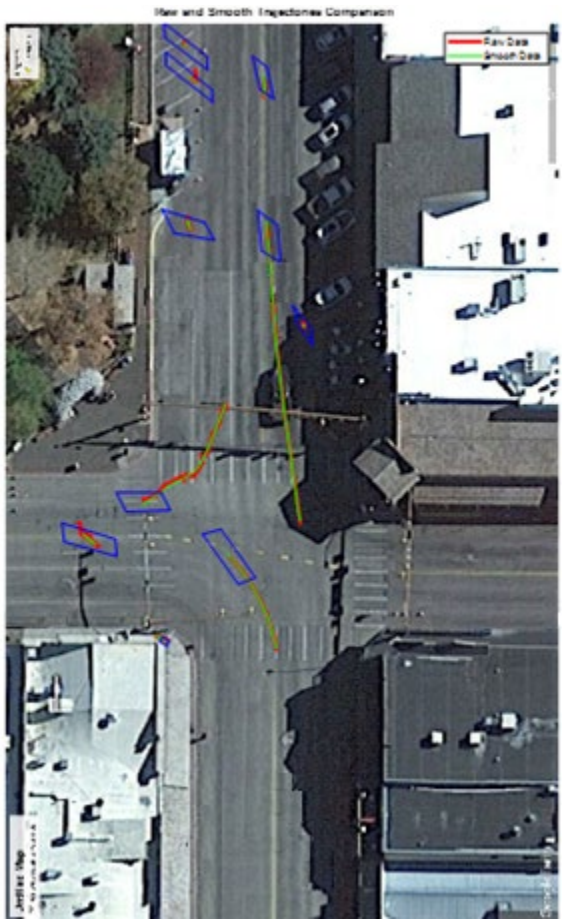


Figure 37 Corrected Vehicle Trajectories

4.1.4 Velocity, Acceleration/Deceleration, and Orientation Calculation

In this subsection the vehicles velocities, their acceleration/deceleration, and orientations are to be calculated. These properties acquisition will serve in the detection and analysis of traffic conflicts.

4.1.4.1 Speed and Acceleration

At current stage, kernel smooth is used to smooth the trajectories, and the speed is calculated by the differential of the position, and acceleration is calculated by the differential of the speed. Both speed and acceleration are separated into two directions, longitudinal and lateral direction. Figure 38 shows the example of vehicle speed and acceleration in both directions. This vehicle a sudden break.

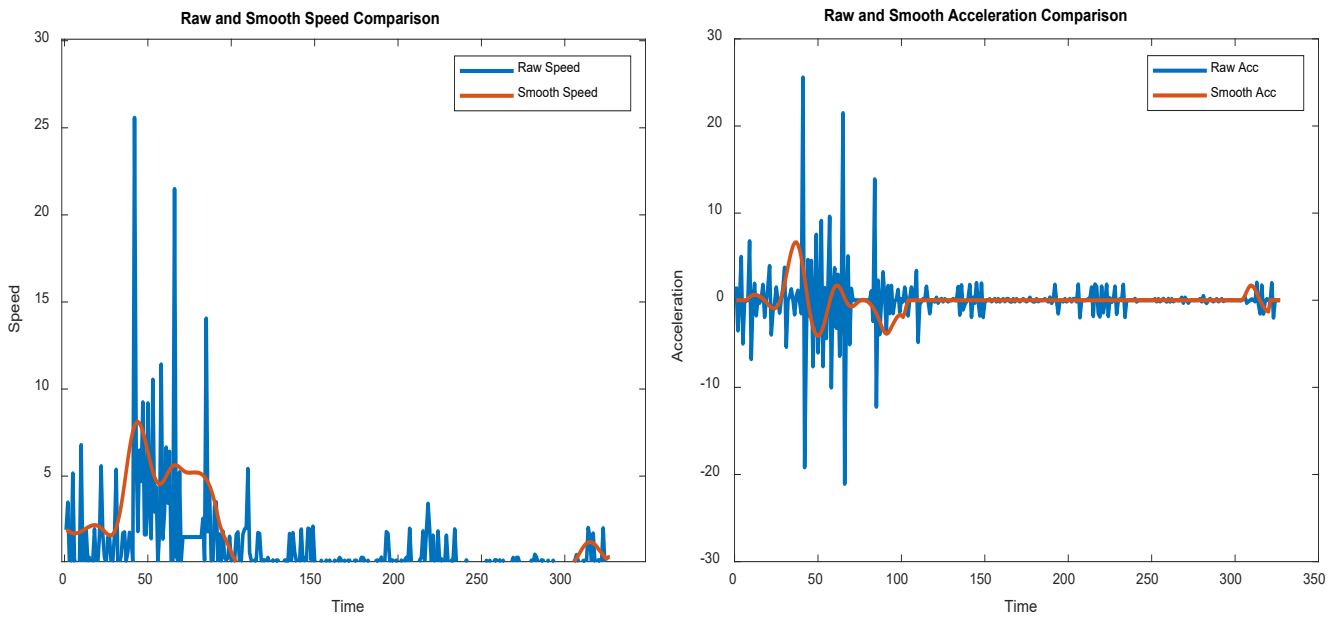


Figure 38 Example of Vehicle Speed and Acceleration That has a Sudden Brake

4.1.4.2 Vehicles' Orientations/Directions

It was chosen to define vehicles orientations by using the polar coordinates system (*r, ϑ*). The following equations represent the magnitude of the traveled distance (*r*) by a specific object between two successive frames in unit time (*t*), and its inclination angle (*ϑ*) with y-axis, Eq.6., and Eq.7. While Figure 39 describes the distance cut by a single vehicle *(r)*, and the orientation *(ϑ)* within a specific period *(t).*

$$\boldsymbol{r(t) = \sqrt[2]{\partial x(t)^2 + \partial y(t)^2} = \sqrt[2]{Vx(t)^2 + Vy(t)^2}} \tag{6}$$

$$\boldsymbol{tan\theta(t) = \frac{\partial y(t)}{\partial x(t)}} \tag{7}$$

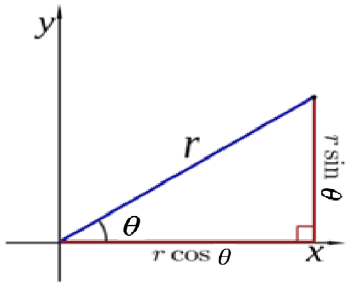


Figure 39 Polar Coordinates System Description

The outputs of traffic conflicts indicators will be presented as a curvilinear relationship between the indicators values versus timeframe. Hence, it was required to remove the outliers from the

output raw data and perform a cleansing process to smooth the curves. Accordingly, a MATLAB code was created to perform this task. The presented code was responsible for smoothening and cleansing the curves by removing the outliers from the raw data.

While processing the raw data, the first step was to add a moving median window to recognize the outliers such that the current step data is always bounded by the median value of a previous window.

$$(1-p)\cdot Median([X_{t-k}, X_{t-k+1}\dots, X_{t-1}])\ X_t \le (1+p)\cdot\ Median([X_{t-k}, X_{t-k+1}\dots, X_{t-1}]) \quad (8)$$

Where k is length of moving window, p is the coefficient for outliers. In our processing, we restrict the coefficient as $0 < p < 0.2$

If the current step data is out of the bound range, it will be marked as outliers, and will not be used for further data smooth. The smooth method is similar to the presented methods in [105, 106]. Figure 40 shows an example for the smoothed curve that extracted from the output raw data.

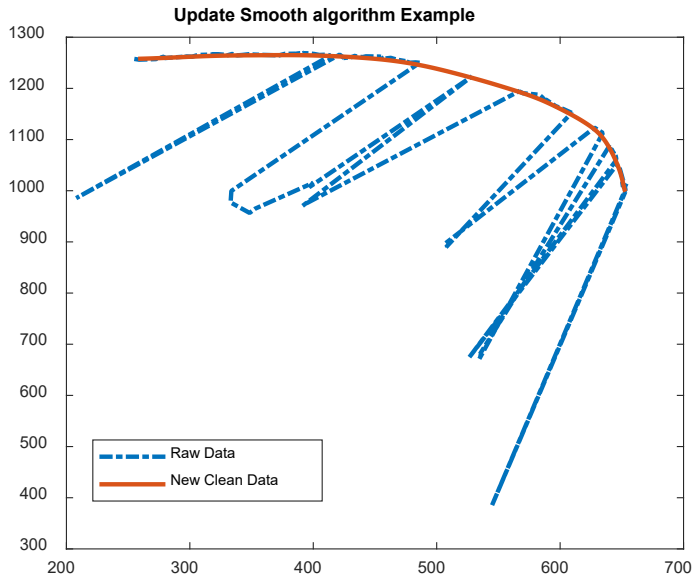


Figure 40 Smoothing and Cleansing Algorithm Output

At this point, the requirements of setting a traffic conflicts detection and analysis framework utilizing CenterTrack algorithm were satisfied. Consequently, in the following chapter the framework will be applied on a conflict set to examine its strengths and weak points.

4.2 Post Processing Steps Utilizing OPenPifPaf as Detection Algorithm

Similar to CenterTrack, a vehicle-vehicle conflict detection framework is proposed utilizing OpenPifPaf to examine the needed post processing steps, Figure 41.

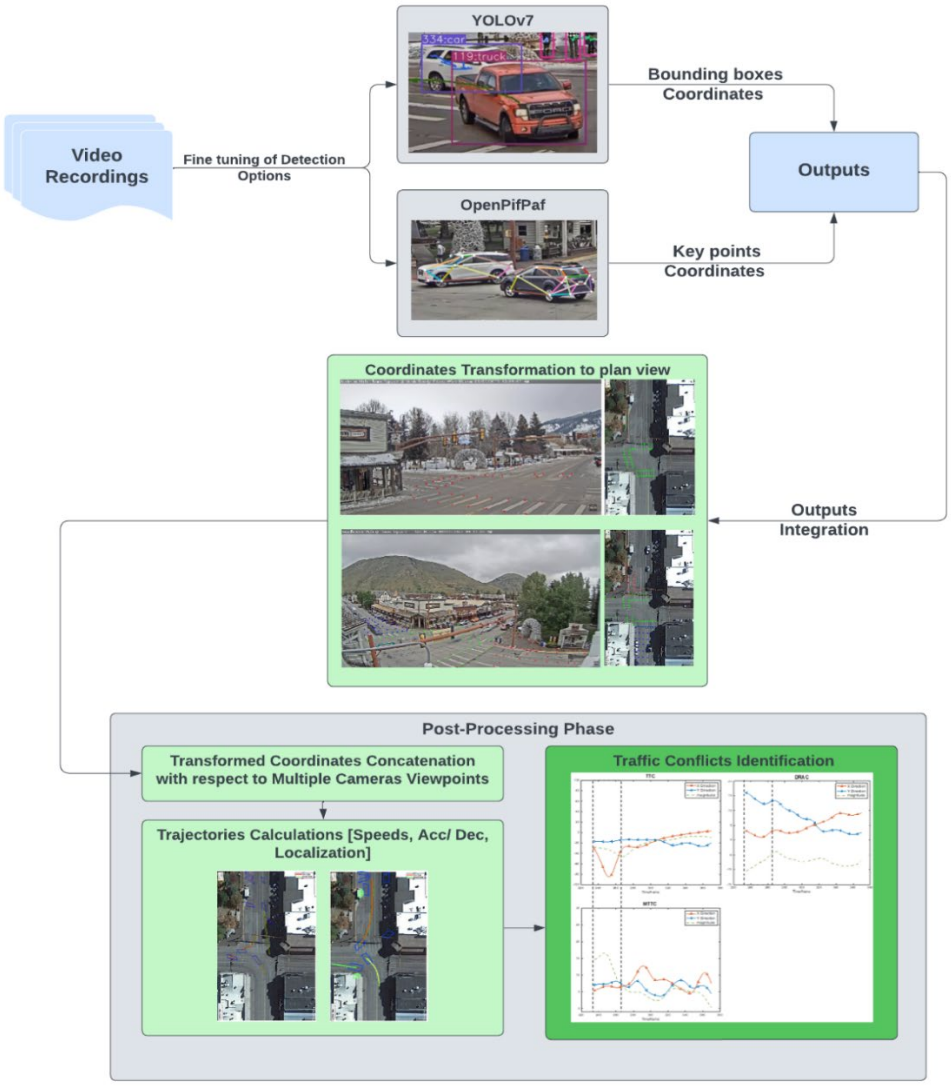


Figure 41 Traffic Conflicts Identification Framework

4.2.1 Detection Error of OpenPifPaf

OpenPifPaf algorithm proves its superiority in tracking vehicles with an unobstructed view, but its performance is notably challenged in urban intersections characterized by heavy traffic. In such congested areas, vehicles may obscure each other, posing a substantial hurdle to tracking accuracy. To assess the algorithm's robustness, we selected traffic videos featuring conflicts,

revealing common tracking errors during post-tracking analysis, as depicted in Figure 42. In this figure, the front detected key points are denoted in red, while the rear points are marked in green.

Figure 42 (a) illustrates a tracking discrepancy where the rear bumper of a black sedan (type 1) should have been identified as two rear points. Unfortunately, one of the points is erroneously labeled as being at the front. Figure 42 (b) showcases a prevalent mis-tracking issue, where a vehicle is not tracked at all due to its location being distant from the camera. In Figure 42 (c), an instance of incorrect tracking involves a vehicle with an irregular shape, where key points may be mistakenly recognized as a different part of the vehicle, rather than the front or rear bumper. These tracking errors are visibly evident in the presented results and can significantly impact data extraction from the tracking algorithm. Hence, the development of appropriate solutions is imperative.

Figure 42 Detection Error of OpenPifPaf

4.2.2 Key Points Projection Utilizing a Universal Coordinate System

To rectify the aforementioned errors, the establishment of a universal coordinate system is imperative to represent the detected key points captured by two distinct cameras. Given that the initial perspectives from the rooftop-mounted cameras are inadequate for such a coordinate system, we employ an image of the same intersection area sourced from Google Earth as a reference for the road layout. In projective geometry, a homography transformation serves as a bijection that maps lines to lines. By correlating the feature points from the camera view with those in the reference image, we facilitate the calculation of the homography between the camera view and the reference road. Figure 43 (a) and Figure 43 (b) present the chosen feature points, while Figure 43 (c) underscores their corresponding points on the

reference road. This process allows for the computation of homographies H1 and H2 for the two cameras. Subsequently, by leveraging both homographies, we can project the perspectives of roads captured by Camera 1 and Camera 2 onto the viewpoint of the reference road. Figure 43 (d) and Figure 43 (e) illustrate the roads post-projection, with the central portions closely resembling the reference road when viewed from a similar bird's eye perspective.

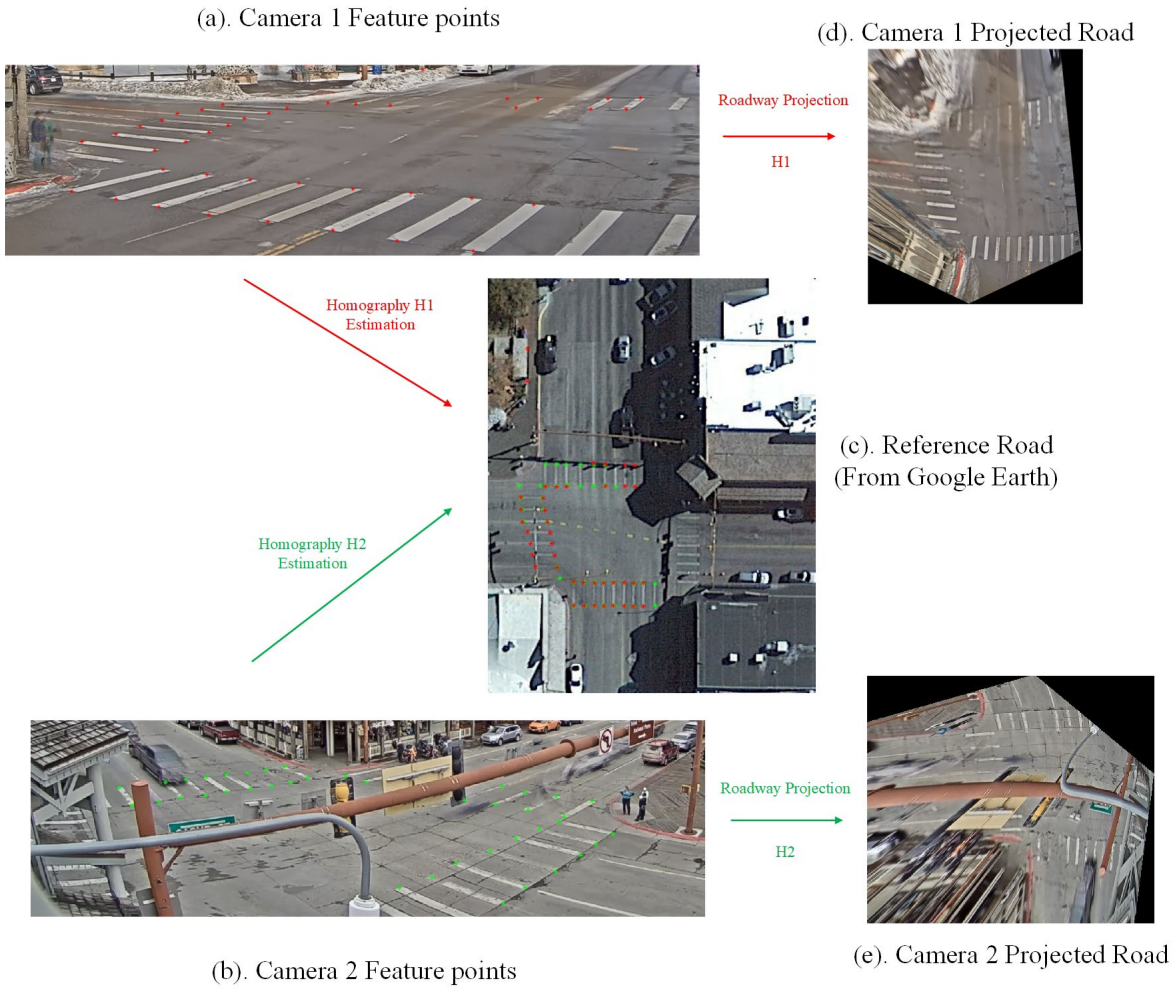


Figure 43 Feature points Matching Between two cameras view into Google Earth Image

4.2.3 Vehicles Key Points Clustering

The default output of the OpenPifPaf algorithm is hampered by a notable limitation. Although it successfully detects the coordinates of all key points, it fails to assign unique vehicle IDs. As a result, in each frame, it becomes challenging to discern which vehicle is being tracked, impeding the association of these points over time to construct comprehensive vehicle trajectories. This

section delves into the methodology for overcoming this limitation by associating key points originating from the same vehicle.

There are two criteria for key point association: (1) Key points originating from the same vehicle must fall within a range that does not exceed the standard dimensions of a typical vehicle. (2) The front key points should exclusively associate with the rear key points. Subsequently, we introduce the error metric as defined in Equation 9. In this equation, the Euclidean distance between front and rear key points is computed. Minimizing this error is crucial for key points originating from the same vehicle. Figure 44 illustrates three examples of key points associations. In this figure, vehicle 1 possesses key points on both sides of its front, while vehicle 3 exhibits key points on both sides of its rear. Consequently, these two vehicles can accurately determine their ground dimensions and locate their central points. In contrast, vehicle 2 features key points on only one side, hindering its ability to ascertain its true position within this frame. For such situations, the central point is approximated as the midpoint of all detected points on the available side.

$$\boldsymbol{Error} = \sum_{i,j} \sqrt{(\boldsymbol{X_{front}} - \boldsymbol{X_{rear}})^2 - (\boldsymbol{Y_{front}} - \boldsymbol{Y_{rear}})^2} \tag{9}$$

4.2.1 Trajectories Extraction and Smoothing

Having established the central points in the previous section, we have successfully identified the central point of each tracked vehicle across all frames. The subsequent step involves extracting the trajectories of the vehicles based on these central points. This process is automated through Eq.10., which defines the error between the current frame data and subsequent frame data. Minimizing this error is crucial to associating a central point in the current frame with a corresponding point in the subsequent frame. Since tracking in the 2D map involves both the X and Y directions in the Google Earth image plane, it is imperative to satisfy this condition for both dimensions. In instances where there is a tracking error in one frame, the minimized error should be smaller than a predetermined threshold to prevent association with the wrong position.

$$\boldsymbol{Min}: \qquad \boldsymbol{Error_{i,j}(t)} = \sqrt{(\boldsymbol{x_i(t)} - \boldsymbol{x_j(t)})^2 - (\boldsymbol{y_i(t)} - \boldsymbol{y_j(t)})^2} \tag{10}$$

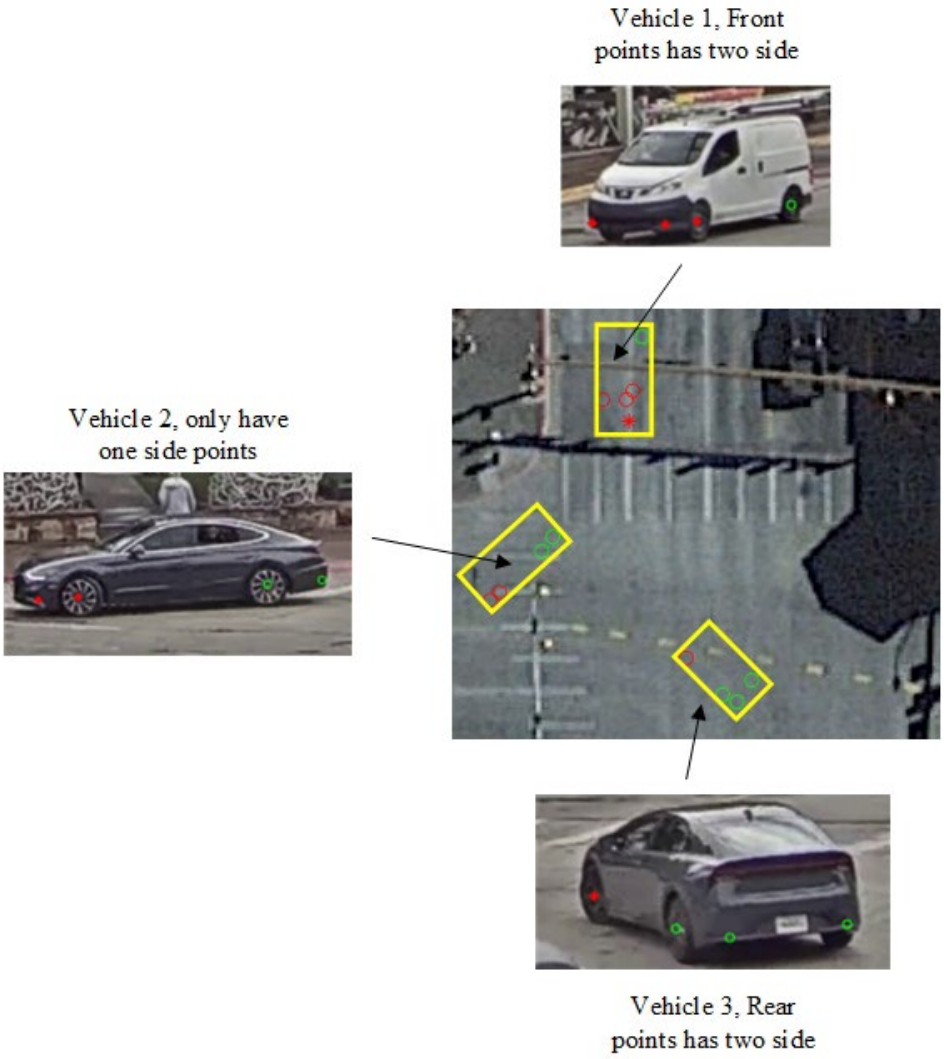


Figure 44 Examples of Key Points Association. Vehicle 1 Has Two Sides of Front Points, Vehicle 3 Has Two Sides of Rear Points, While Vehicle 2 Only Has One Side of Key Points

This approach ensures robust trajectory extraction by considering the continuity and coherence of vehicle movement across frames, thereby enhancing the reliability of the tracking results. Up to this point, we have generated raw trajectories that are prone to errors. To further enhance the accuracy of the trajectories, this research applies a robust locally weighted regression and smoothing procedure. This robust regression technique is designed to protect against aberrant points distorting the smoothed trajectory. The weighting is determined by the bi-square function as per Eq.11**.** In this equation, r_i represents the residual of the i^{th} data point resulting from the regression smoothing procedure. The MAD (median absolute deviation) is the median value of the absolute residuals, defined as MAD = *median*(|r|).

Following the smoothing process, it is advisable to implement a Savitzky-Golay filter to further eliminate any remaining errors. This refined trajectory is essential for accurate speed and acceleration estimation in the next section. Figure 45 provides a comparison between the raw and smooth trajectories, clearly demonstrating a significant improvement in the smooth trajectory compared to the raw one.

$$w_i = \begin{cases} \left[1-\left(\frac{r_i}{6.MAD}\right)^2\right] & |r_i|<6\ .MAD \\ 0 & |r_i|\geq 6\ .MAD \end{cases} \tag{11}$$

Figure 45 Smoothed Trajectories Extraction

4.2.2 Speed and Acceleration Smoothing

The smoothed trajectory derived from the previous section serves as the basis for speed estimation. Speed is estimated by calculating the position differences between successive frames in both the X and Y directions. It is crucial to recognize that, despite the smoothing process, the trajectory may still contain errors. The differential process can potentially amplify these errors, thereby compromising the quality of the final data if not properly address. Therefore, a weighted moving median filter is developed to compute the speed while simultaneously eliminating errors.

The process begins by separately calculating the time series speed for each vehicle, using a range of time steps as defined by Eq.12**.** In this equation, *i* represents the current frame, *n* is the time-step, and *x(i+n)* is the vehicle's position in frame *i+n*. Thus, $\hat{v}_n(i)$ represents the estimated average speed at the current frame with an interval of 2·*n* frames. Subsequently, the

overall average speed is calculated by incrementally increasing the time step (*n)*. Speeds derived from different *n* steps, ranging from 1 to a specified integer, provide a comprehensive set of speed measurements. The vehicle speed is then determined by taking the median value of all these *n* step speeds.

To further mitigate noise, a Savitzky-Golay filter is applied to the median speed. It's important to note that this process is applied to speed measurements in both the X and Y directions. This comprehensive approach ensures robust and accurate speed estimation while addressing potential errors in the trajectory.

$$\widehat{v_n}(i) = \frac{x(i+n)-x(i-n)}{2.nT} \tag{12}$$

Acceleration estimation follows a process similar to that of speed estimation. Building on the cleaned speed data obtained from the previous section, the acceleration for each vehicle is calculated using an equation similar to the one used for speed calculation, Eq.13**.** The key distinction lies in setting *n* to 1, thereby determining acceleration as the median value within a time interval of *2T*. Subsequently, a Savitzky-Golay filter is applied to achieve the final smooth acceleration. Figure 46 and Figure 47 provide comparisons between the raw and smooth speeds and accelerations, demonstrating the effectiveness of the applied estimation process.

$$\widehat{a_n}(i) = \frac{v_{smooth}(i+1)-v_{smooth}(i-1)}{2.T} \tag{13}$$

Based on the extracted results, the conflict detection identification framework employing OpenPifPaf will be examined in the next chapter.

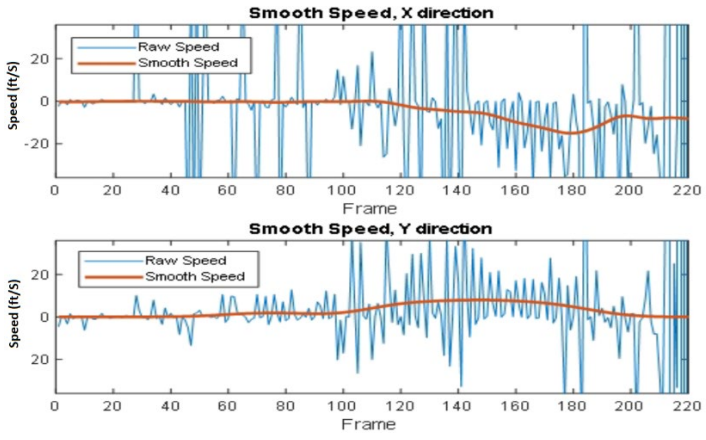


Figure 46 Raw and Smoothed Speeds Extraction

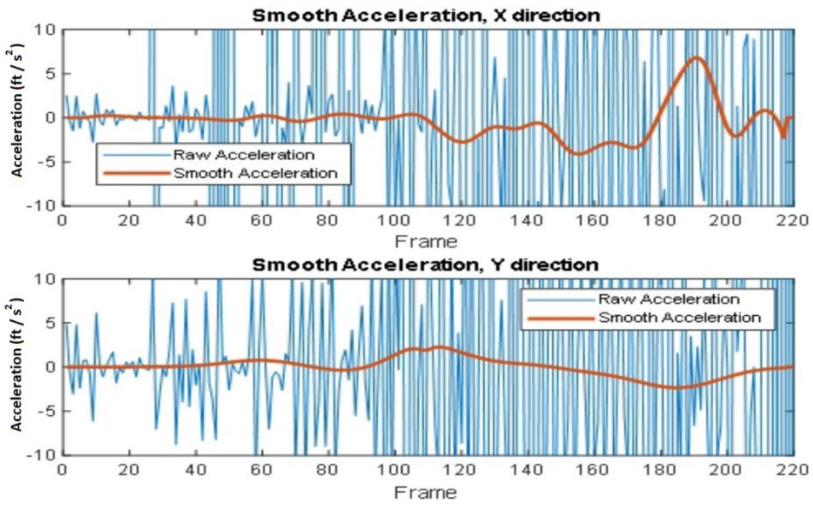


Figure 47 Raw and Smoothed Accelerations Extraction

CHAPTER 5. TRAFFIC CONFLICT ANALYSIS

5.1 Background

The analysis of traffic conflicts has been introduced since 1960 and is ever evolving [22]. By using Surrogate Measures of Safety (SMoS), the interaction between road users can be characterized, and depending on the measure both the probability and severity of a potential collision could be estimated [23–25]. Generally, these SMoS indicate the nearness between road users (or their vehicles) and the proposed collision point.

The analysis in this chapter will firstly provide the preliminary information that will be employed in the proposed detection framework by providing the following subjects:

- Presentation of the extracted conflict types from the literature study.
- The linkage between traffic conflicts and their associated crash types.
- Illustration of the regions diagram and region-conflict relation.
- Presentation of the utilized traffic conflicts indicators, and their preferences in conflicts detection.

Furthermore, the applied procedure in the traffic conflicts detection framework will be presented through the analysis two sets of traffic conflicts that were captured by fixed CCTVs at two signalized intersections (Town Square intersection at Jackson Hole, and Four Corners Camera intersection at Cold Water).

5.2 Literature Review

Traffic conflicts indicators are considered the detection and assessment tools for the traffic conflicts on various roadway types. Many authors have employed various types of traffic conflicts indicators to identify conflicts and their severity levels at different roadway sections [81, 83, 140]. Furthermore, several comparative analyses have been held to extract the outperforming indicator for each specific type of conflict, and road section [97–101]. Zheng et al. classified traffic conflicts measures under temporal and spatial proximity families [28]. While Johnsson et al. broke them into three categories related to time to collision (TTC), post encroachment time (PET), and deceleration rates [54]. A recent systematic review study identified six categories of traffic conflicts measures that are used to assess the safety at intersections including Temporal proximity measures, Spatial proximity measures, Kinematic

measures, Mixed measures, Combination of measures, and crash severity measures [51]. Many of these indicators have a different application and definition and the conflict analysis performed is not found to be uniform amongst researchers. Recent studies have explored Surrogate Measures of Safety by keeping their focus on time-series data of interaction and analyzing it for the clusters of conflicts and non-conflicts by looking at different aspects involved in a conflict, such as time to collision (TTC), speed, acceleration, and vehicle trajectory [26]. They found this method to be more reliable as it uses both proximity and evasive action-based surrogate measures to find traffic conflicts. Some systematic literature reviews have also been performed on traffic conflict-based safety measures, and they indicate the usage of different conflict indicators with varying thresholds for conflict detection [23, 24, 27]. Some studies have concluded that the choice of measure and its threshold depends on various conditions and application contexts, so measures should be selected depending on their context [27, 28].

5.2.1 Time to Collision (TTC)

Time to collision (TTC) is one of the most commonly used SMoS to determine the time to collision, which is defined as the time remaining until a potential collision if the interacting road users remain at the same speed and direction [29]. As a measure of temporal proximity, TTC indicates how close conflicting road users are to one another. Time-to-collision measures the proximity of conflicting road users to a collision point, assuming that both users will arrive at the same time. This surrogate measure of safety can be calculated as long as road users are on a collision course and therefore there is a time series for the TTC values. Thus, the lowest value, which is generally termed as TTC_{min} is the most commonly used indicator.

The following equation provides the calculation of TTC, Eq.14.

$$\boldsymbol{TTC_t = \frac{x_{L,t} - x_{F,t} - D_L}{v_{F,t} - v_{L,t}}; whereas\ (v_{F,t} - v_{L,t}) > 0} \tag{14}$$

Where $x_{L,t}$ and $x_{F,t}$ are the positions of leading and following road users at time t, D_L is the length of the leading road user, and $v_{L,t}$ and $v_{F,t}$ are the speeds of leading and following road users, respectively.

There are multiple variants of TTC that have been developed as an improvement on the traditional time-to-collision. The modified time to collision (MTTC), which incorporates acceleration data to provide more accurate estimation was found to be more strongly correlated with crashes as compared to traditional time to collision measure. [30, 31]. The MTTC was developed by Ozbay et al. to release the speed constraint that was assumed by TTC. The calculation method of TTC proposed that the two colliding road users had constant speed at collision time [102]. While MTTC considered the vehicles acceleration at collision in its calculation, Eq.15.

$$\boldsymbol{MTTC_t} = \frac{\boldsymbol{\Delta v_t} \pm \sqrt{\boldsymbol{\Delta v_t}^2 + 2\boldsymbol{\Delta a_t}(\boldsymbol{x_{L,t}} - \boldsymbol{x_{F,t}} - \boldsymbol{D_L})}}{\boldsymbol{\Delta a_t}} \quad (15)$$

Where Δv_t, and Δa_t are the difference in speeds and accelerations between the two colliding road users, respectively.

Although TTC is the most widely used surrogate measure, little is known about how to select the appropriate threshold. The window is generally between 1 and 5 seconds, with 1.5 seconds being the most common value [25, 26, 33–35]. Although other variations of time-to-collision, such as time-extended time-to-collision and time-integrated time-to-collision, quantify a traffic conflict by considering the duration of time-to-collision values below a certain threshold and the degree to which these values fall below the threshold, respectively, they have also been utilized less frequently in conflict studies [36].

5.2.2 Post-Encroachment Time (PET)

In the SMoS analyses, other methods such as the Post-Encroachment Time (PET) are also found to have been used extensively [37, 38]. PET is a well-known indicator, which is defined as the time difference between the passing out of the offending vehicle from the collision area and the arrival time of the conflicting vehicle to this area or collision point [103]. Additionally, researchers have used Encroachment Time (ET), which is a derivative of PET that measures the duration of time in which offending road vehicles occupy a conflict area, and it can be used when oncoming major road vehicles are traveling at constant speeds.

It should be noted that the analysis of rear-end-crashes using PET required a determination of an imaginary line that proposed as an alternative to the collision area for other (crossing) crash types. The following equation is used to calculate PET, Eq.16.

$$\boldsymbol{PET_t = t_{F,t} - t_{L,t}} \tag{16}$$

Where $t_{L,t}$ and $t_{F,t}$ are the departure time of the offending vehicle, and the arrival time of the conflicting vehicle.

TTC alongside PET are the two most common measures of conflict analysis. The reason for their widespread use is that they provide complementary information and capture different aspects of conflicts. Although both measures are typically used to identify different types of conflicts, time-to-collision is typically used for rear-ending conflicts, while post-encroachment time is used as an additional measure when road users' paths cross [39–41].

5.2.3 Deceleration Rate to Avoid Collision (DRAC)

In addition to TTC and PET, some researchers have attempted to identify near crash events using other SMoS. An indicator called Deceleration Rate to Avoid Collision (DRAC) was developed as a representation of the Kinematic measures for the conflicting vehicles. This preliminary measure could detect the sudden drop in vehicles kinematics (Speeds and accelerations) by detecting the hard brakes which activated before entering the collision area. It is the deceleration rate that should be applied by the following vehicle to avoid crash occurrence with the leading vehicle. The following equation illustrates the calculation of DRAC, **Eq.**17**.** [104]

$$\boldsymbol{DRAC_{F\ t} = \frac{(v_{F,t} - v_{L,t})^2}{2(x_{L,t} - x_{F,t} - D_L)}} \tag{17}$$

As illustrated, usually DRAC is used as a guiding kinematic measurement that could assess the severity of traffic conflicts when the crossing movements are not completing by employing hard brakes. A comparative analysis was performed by Zheng et al. to find the superior traffic conflicts indicator from TTC, MTTC, PET, and DRAC by comparing crashes estimated from conflict measures with the observed crashes [53]. It was concluded that MTTC was the most accurate indicator that managed to estimate the crashes frequencies closer to the observed

one, followed by TTC and then PET. While DRAC achieved the highest number of crash estimations compared to the observed crashes.

5.2.4 Mixed Conflict Measures by Severity

Furthermore, Chen Wang et. al proposed a new surrogate measure "Aggregate Severe Crash metric" that was found to be a better indicator than TTC in determining the crash severity and the paper presents many challenges that come with applying TTC as a final measure of crash severity [37]. This paper develops the new metric ASCM (Aggregate Severe Crash Metric) based on delta-V i.e., change in kinematic energy. This study tested the new metric for twelve different intersections and found it to be particularly good at predicting fatal or injury crashes. Furthermore, Wang and Stamatiadis have developed a mixed conflict measure called "Aggregate Crash Propensity Metric," that comprised time-to-collision, reaction time and deceleration rate etc. in its formulation [42]. Despite showing promising results in comparison to studies that solely rely on temporal proximity measures, this methodology has been underutilized. This may be due to the fact that evaluating evasive maneuvers, such as deceleration and acceleration by conflicting road users, is based on subjective judgment rather than objective criteria, and there is no widely accepted means of quantifying the intensity of such actions. Amir Sobhani et al. conducted micro simulation including driver characteristics [43]. They used two models for identifying traffic conflicts and one of them involves kinematic energy difference. This model predicts the traffic conflicts and crash severity if that conflict turned into a crash.

5.3 Identification of Study Surrogate Measures of Safety

Firstly, a systematic literature review and a meta-analysis were performed on research studies utilizing SMoS for conflict detection in order to aid with the selection of appropriate conflict indicator alongside its optimal threshold. Secondly, a framework for selecting conflict types and their associated collision type was established. Then the regions of probable conflicts were identified and then the conflict detection and analysis were performed.

All these topics have been explored in the subsequent sections below starting with systematic review and meta-analysis followed by framework establishment and conflict region

identification which is later followed by conflict detection and its analysis and then the chapter ends with a discussion section.

5.4 Conflict Indicator and Optimal Threshold Selection

A systematic review of literature was performed using Preferred Reporting Items for Systematic Reviews and Meta-Analyzes (PRISMA) guidelines [44]. In order to formulate relevant and precise eligibility criteria, we followed the Population, Intervention, Comparator, Outcome, Study Design (PICOS) approach. Limited resources and time can be dealt with ease using this method and it is suitable for qualitative research [45]. The population of study considered in this study involved all kinds of road users and included all peer-reviewed journals till 2023 dealing with traffic conflict analysis at intersections, freeways or crossing etc. The aim was to get the most up to date data and to understand the change in trend over time. The study design considered all the studies conducted using Computer Vision on traffic video data. The literature search was conducted to identify relevant papers on several databases including Google Scholar, PubMed, ScienceDirect, Scopus, JSTOR, TRID (Transportation Research Integrated Database), Web of Science and Compendex. A PRISMA flow diagram is shown in Figure 48, illustrating the number of studies included at each stage of the review process. Studies were judged based on title and abstract and those found relevant were included in the full-text screening process. Many studies were out of our context and the results were not sufficiently detailed, these papers whose sufficient information could not be gathered had to be removed from the analysis.

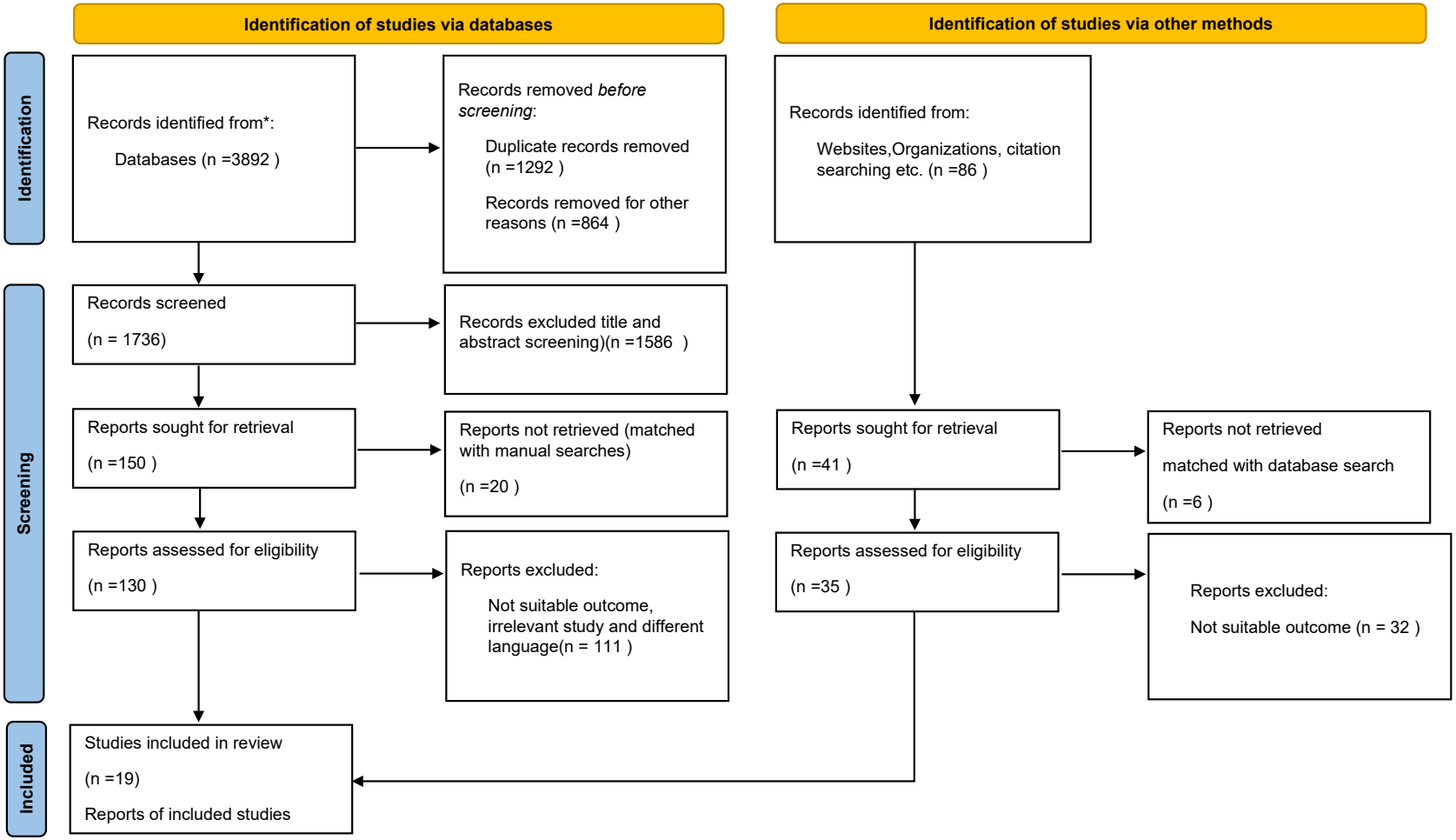


Figure 48 PRISMA Flowchart for Conducting Systematic Review and Meta-Analysis

After full text screening 19 different studies were selected to conduct Meta-Analysis. However, conducting Meta-Analysis of SMoS thresholds has several limitations that are not encountered in Meta-Analyses of studies evaluating road safety. To begin with, for conducting Meta-Analysis, selected studies should follow a basic framework and outcomes need to be conceptually comparable [55]. It should be noted, however, that in SMoS studies, multiple surrogate measures have been used in multiple studies, and there has been a lack of consistency in the methods employed and the reporting of the study outcomes. The absence of comparable outcomes is a key limitation in threshold selection in SMoS studies, which is a vital aspect for conducting meta-analyses. The selected studies didn't follow a comparable framework and the outcomes they provided were diverse in nature. As a result of the different types of outcomes as well as missing values like standard error and sample size, it is impossible to establish a single SMoS threshold based on Meta-Analysis.

While the attempt to conduct a Meta-Analysis was unsuccessful, a systematic review was performed to determine the suitable conflict and their optimal threshold that can be used in this study. The reviewed studies have shown that the major weaknesses in most of the SMoS studies performed have been the lack of validation, selection of SMoS, and problems with threshold selection [24]. In addition, there is a lack of description regarding how the threshold is determined as well as the selection of proper surrogate measures. From review of papers and sample analysis of most prominently used SMoS's (PET and TTC), it can be seen that PET is generally favored and used for pedestrian related conflicts as well as turning vehicle conflicts. Also, TTC is mostly used for rear end conflicts as compared to PET. Both TTC and PET (with their respective variations) are the most used surrogate measures of safety. In general, the PET thresholds mostly used in the studies are 1.5, 3 and 5 s and most of them are based on findings of other studies and out of these, 3 s is found to be used the most to account for severe conflicts. Similar is the case with TTC where most studies are found to be using threshold values in the range of 1.5 to 3 s. In recent studies, in order to account for threshold and determine severity, Extreme Value Theory Analysis method with either Peak Over threshold or Block Maxima Approach have been found to be used extensively. Another method of using CDF plots utilizing wide range of thresholds (generally 1 to 5s) with a certain increment (generally 0.5s)

has been used and 15th, 50th and 85th percentile values have been used to account for various severity levels of conflicts. It was also found that alongside TTC and PET, MTTC and DRAC are also being extensively used for conflict analysis.

Based on the systematic review, this study employs four traffic indicators (TTC, MTTC, PET, and DRAC) whenever possible to detect the various traffic conflict types.

5.5 Traffic Conflict Data Collection and Preliminarily Analysis

In this section the video data collected, and the manual extraction of conflicts will be demonstrated. Consequently, the conflicts distribution within the case study intersection is provided based on spatial and temporal metrics. Moreover, a binary logit model is utilized to examine the relationship between traffic conflict types and other factors that may influence conflicts, in order to investigate the correlation.

As prementioned, the section of the case study intersection is done based on three main advantages on other sites; the intersection includes three surveillance cameras that could provide better field of view for different approaches, each camera is provided with a real-time YouTube channel, and the intersection location in recreational area that high traffic volume with various types of road users.

5.5.1 Conflict Data Collection at Site of Interest

Traffic conflicts were observed in the Jackson Hole intersection through live video data available on YouTube. A total of 39 hours of video footage were reviewed, covering both weekdays and weekends. The details of the video footage used were provided below, Table 4.

Table 4 Video Data Utilized for The Study

Date	Day	Time period (hours)
12/9/2022	Friday	9.5
12/10/2022	Saturday	9.5
12/11/2022	Sunday	9.5
12/20/2022	Tuesday	5.5
12/21/2022	Wednesday	5
	Total	39

The initial step was established by performing extensive literature review to consider the most frequent types of conflicts that occur at signalized intersections. The extraction process of traffic conflict types was mainly based on the "*Traffic Conflict Techniques for Safety and Operations-*

Observers Manual" in addition to the study provided by Ezra Hauer for traffic movements at intersections. Based on these studies, a list consisting of twenty types of traffic conflicts was prepared including four types of pedestrian-involved conflicts [94]. The following two figures illustrate the extracted conflict types, Figure 49 and Figure 50 such that each conflict was given a unique ID. The identification of these types was done in the manual monitoring of the intersection, then, the conflicts data were collected into two sheets, Figure 51 and Figure 52. While for acquiring more details on each collected conflict, each conflict was mapped, and additional details were noted, Figure 53. The first two forms were used to collect a list of conflicts and their timestamps. And the second form was used to collect details for each conflict and includes all relevant information.

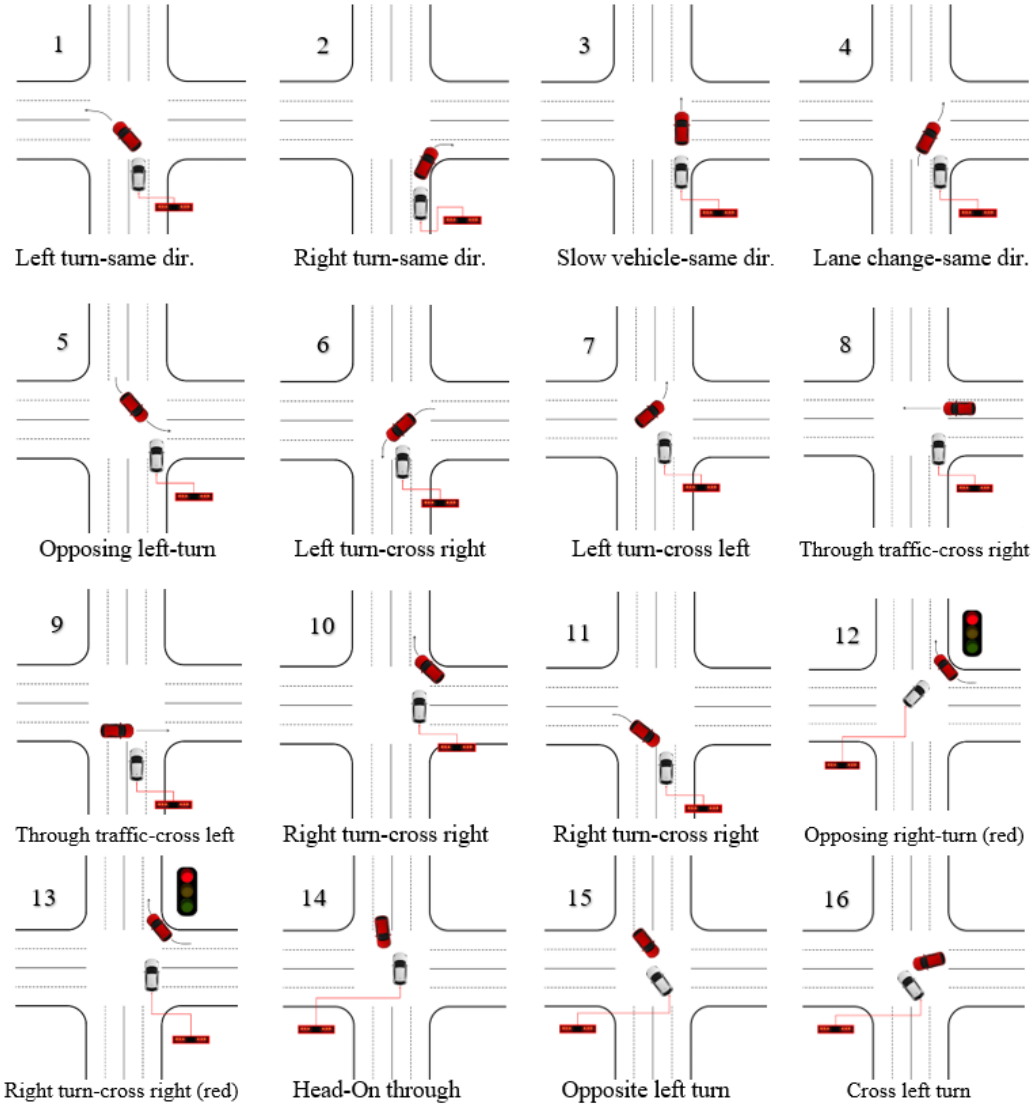


Figure 49 Vehicle-Vehicle Conflict Types

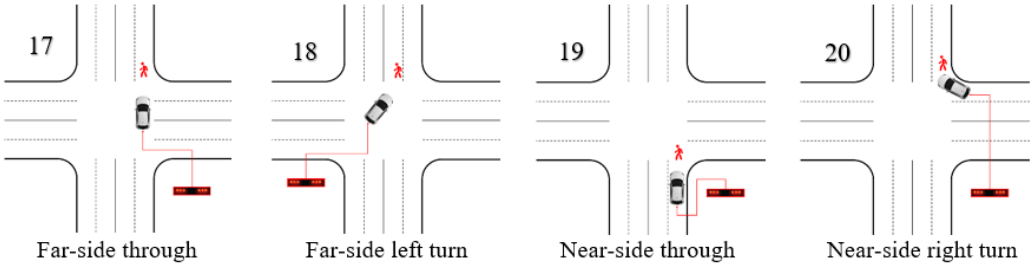


Figure 50 Vehicle-Pedestrian Conflict Types

File name	Time stamp	Serial #	Left Turn Same Direction		Right Turn Same Direction		Slow Vehicle		Lane Change		Opposing Left-Turn		Right Turn From Right		Left Turn From Right		Through From Right		Right Turn From Left		Left Turn From Left		Through From Left		Right Turn On Red		Head On Through		Opposite Left Turn		Cross Left Turn	
			1		2		3		4		5		6		7		8		9		10		11		12		13		14		15	
			C	SC	C	SC	C	SC	C	SC	C	SC	C	SC	C	SC	C	SC	C	SC	C	SC	C	SC	C	SC	C	SC	C	SC	C	SC

Figure 51 Conflict Data Main Sheet (Vehicle – Vehicle)

File name	Time stamp	Serial	Far Side Through		Far Side Left Turn		Near Side Through		Far Side Right Turn	
			16		17		18		19	
			C	SC	C	SC	C	SC	C	SC

Figure 52 Conflict data main sheet (vehicle - pedestrian)

Traffic Conflict Recording Form

Observer and Intersection Information:

Observer Name		Date		Time	
Serial					
Intersection Name		City			

Weather, surface condition, and time period:

Weather	Sunny ☐	Cloudy ☐	Rainy ☐	Snowy ☐
Surface	Dry ☐	Wet ☐		
Time Period	Dawn/Morning ☐	Midday/Noon ☐	Dusk/evening ☐	
	Midnight/Dawn ☐	D/M:6am/11am—M/N:11am/4pm D/E:4pm/12am—M/D:12am/6am		

Traffic Conflicts features and measurements:

Type	Road-User 1	Road-User 2	Road-User 3
Private Car, Truck	☐	☐	☐
Bicycle, Motorbike	☐	☐	☐
Pedestrian	☐	☐	☐
Distance to coll. point	ft	ft	ft
TTC value	s	s	s
Avoiding Action: Braking	☐	☐	☐
Swerving	☐	☐	☐
Acceleration	☐	☐	☐
Possibility to Swerve	Yes No	Yes No	Yes No
At fault	☐	☐	☐

Conflict Pattern	
1	16
2	17
3	18
4	19
5	20
6	21
7	22
8	23
9	24
10	25
11	26
12	27
13	28
14	29
15	30

North

Intersection Type	No. of Vehicles in Queue	Total Entering Volume
Signalized		

Private Car, Truck
Bicycle, Motorbik
Pedestrians

Sketch camera position

Figure 53 Conflict data detail sheet

Additionally, two more conflict types (20, 21) related to parked vehicles were identified. These conflicts were included because parked vehicles were found to cause several rear end conflicts at Jackson Hole intersection. The details for each conflict were manually recorded with TTC value

as the utilized indicator. **Table 5** presents the distribution of the detected conflicts at different TTC values.

Table 5 Conflict Frequency Distribution with TTC Threshold

Conflict type	TTC (Sec)									Grand Total
	1	1.5	2	2.5	3	3.5	4	5	5.5	
1		1	1							2
11			1							1
12		1	4		2	1				8
12,20			1							1
13			1				1			2
16		1	4	3	4			1	1	14
17	1	2	3	1	12	4	4			27
18			1							1
19	1	1	5		5		1	1	1	15
2			1			1				2
2,20			2							2
20	1	3	13	5	18	3	12	1		56
21			2	1	1	1	3			8
3	1	4	51	17	30	3	5			111
3,20				1						1
4	2	1	1	1						5
5	2		9		3		1			15
6			8					1		9
Grand Total	8	14	108	29	75	13	27	4	2	280

From the table, it was observed that the most common conflicts were of Type-3, with 20 instances, followed by Type-17. Type-3 conflicts were rear end collisions that occurred due to slow-moving vehicles, while Type-20 conflicts were also rear end collisions that happened when parked vehicles are exiting inclined parking lots. These two types of conflicts, which involved vehicles colliding with each other, accounted for the majority of the total number of conflicts at the Jackson hole intersection.

Type-17, on the other hand, was considered the most severe among all the conflicts as it involved left-turning vehicles and pedestrians. These types of conflicts had the potential to result in serious injuries or even fatalities, and out of all the pedestrian conflicts, this conflict was the most common.

To gain a better understanding, the conflict locations were transferred from the sheet to geo-coordinates, and heatmaps were developed using Arc-GIS, Figure 54 and Figure 55.

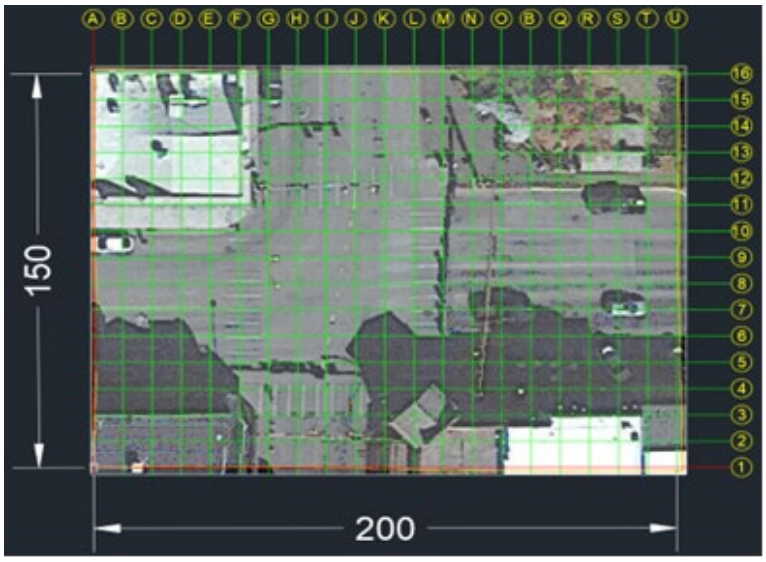


Figure 54 Region Selected for Conflicts

Figure 55 Heatmap of Conflicts

The 85th percentile of the collected conflicts has a TTC alternative value of 3 seconds, Figure 56. While, in order to capture the progression of conflicts at various TTC thresholds, heatmaps were plotted for each threshold ranging from 1.5 seconds to 4 seconds with an increment of 1 second, Figure 57.

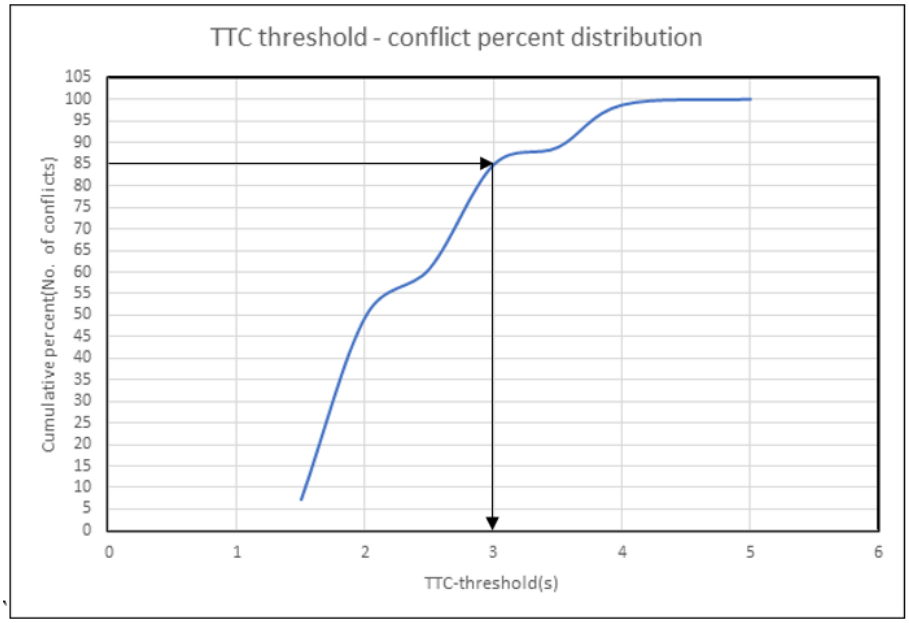


Figure 56 Cumulative Distribution of Conflicts with TTC Threshold

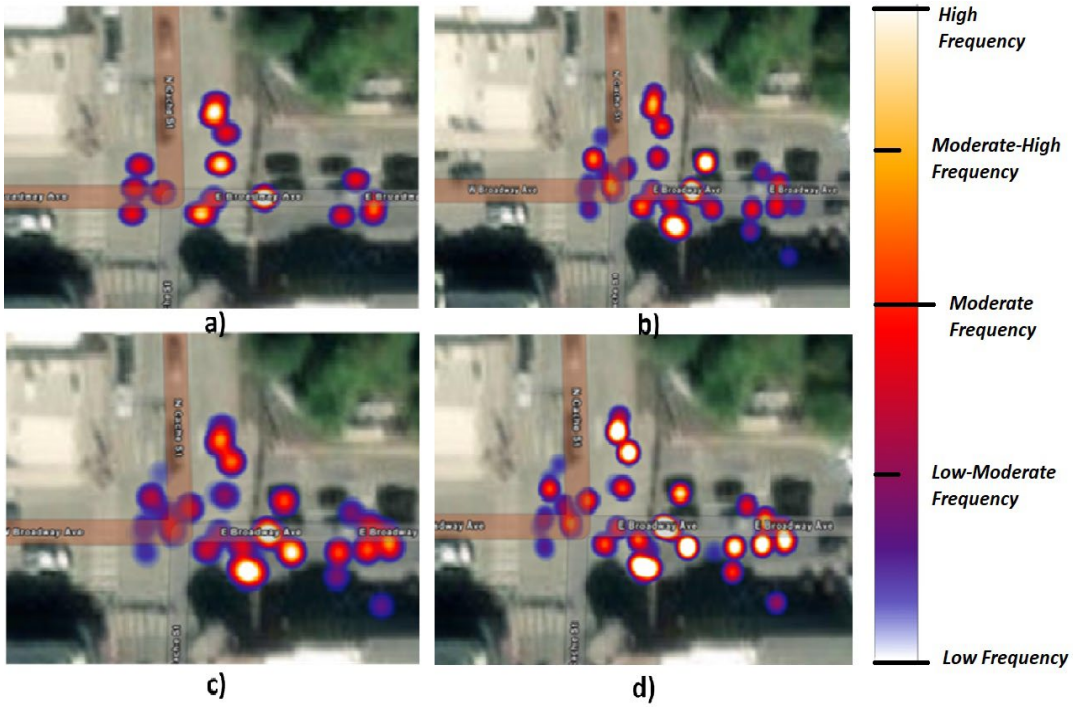


Figure 57 Heat Map Generation with Different TTC Thresholds: a)1.5 seconds, b)2 seconds, c)3 seconds , d)4 seconds

5.5.2 Observations

At Jackson Hole intersection, rear end conflicts were observed at the center of the intersection, as well as most of the occurred crashes were rear end collisions. As the heatmaps were developed

for different TTC values, a significant increase in intensity was observed as the TTC threshold increased from 1.5 seconds to 4 seconds. This increase was due to a greater number of conflicts being included.

It was also observed that as the TTC threshold increases, the conflicts frequencies at eastbound approach became more significant, indicating that the rear end conflicts had higher TTC values.

The traffic conflict analysis identified both rear end and left turning pedestrian conflicts as the most severe conflicts at the Jackson Hole intersection.

At the intersection, the first conflict observed was rear end conflicts, and the majority of them occurred in the eastbound approach. Parked vehicles were involved in a considerable number of these conflicts. During peak hours, due to high traffic, there were large cases of congestion and queuing at this approach, and there was only one lane provided at this approach as other space was occupied by parked vehicles.

The second most common type of conflict observed at this intersection was pedestrian conflicts with left turning vehicles. One reason for this was the high volume of left-turning vehicles and high volume of pedestrians crossing the street at peak hours. Since this was a recreation place, pedestrian violation was observed at night.

The above observations from the video observation led to a further investigation of the relationship between the turning conflicting movements of vehicles and other explanatory variables with the outcomes, i.e., conflicts.

Initially, a count model was proposed, but it did not perform well due to a lack of data. After narrowing the conflicts down to rear end conflicts and pedestrian conflicts, less than 40 conflicts were left for the pedestrian-left turning vehicle interaction. This limitation resulted in changing the model from a count to a case-by-case study model. In the case study model, the explanatory variables that were responsible for the conflict occurrence were considered, and a logit model was used for the analysis.

5.5.2.1 Analysis of Rear End Conflicts

Most conflicts were observed in the east bound approach, so the total east bound through volume was calculated and the case of parked vehicles was studied. A 4-minute interval was chosen for volume calculation after conducting a study with 1-minute, 2-minute, and 8-minute

intervals. It was found that the 4-minute interval was optimal for capturing the conditions of conflict occurrence. Shorter data chunks with 1-minute or 2-minute intervals did not perform well because the cycle length at this intersection varies from 1:15 seconds to 1:50 seconds and these shorter chunks did not capture the entire cycle length in certain situations. On the other hand, the 8-minute interval was able to capture the circumstances of the conflict, but it did not perform well because the cycles vary from each other and the total number of 8-minute data chunks is only half of that for the 4-minute interval. Ultimately, the 4-minute interval was chosen as it was able to show a better relationship between turning volumes and the number of conflicts during the peak hour compared to other intervals. The explanatory variables for this study were rear end conflicts, east bound through volume, and parked vehicle interaction. Other variables, such as time of day and road condition, were also included in the model.

Logit Regression Results

Dep. Variable:	outcome	No. Observations:	90
Model:	Logit	Df Residuals:	85
Method:	MLE	Df Model:	4
Date:	Wed, 08 Mar 2023	Pseudo R-squ.:	0.3803
Time:	23:55:08	Log-Likelihood:	-38.658
converged:	True	LL-Null:	-62.383
Covariance Type:	nonrobust	LLR p-value:	1.229e-09

	coef	std err	z	P>|z|	[0.025	0.975]
const	-7.0554	1.763	-4.001	0.000	-10.511	-3.599
EBT volume	0.2609	0.080	3.267	0.001	0.104	0.417
Parked vehicle	1.8666	0.612	3.051	0.002	0.667	3.066
Snowy	0.9081	0.632	1.436	0.151	-0.331	2.147
Peak time	0.5429	0.737	0.737	0.461	-0.901	1.987

AIC: 87.31685577154327
BIC: 99.8159041231946

Figure 58 Summary of Results for Rear End Conflicts

The odds ratio for eastbound through vehicle volume was 1.32, and the odds ratio for parked vehicle involvement was 2.58. The odds ratio was 1.485 when the road condition was snowy. Additionally, the odds ratio for peak time influence was 1.26.

In this analysis, the ratio of cases to controls was exactly 1:1. Cases represented conflict occurrence, while controls represented 'no conflict' under the same conditions. Cases and controls were chosen from both on-peak and off-peak hours to account for variation. The eastbound through vehicle volume and parked vehicle interaction were both significant with a p-value less than 0.05. The road condition was also significant ($p < 0.2$), which was supported by the crash analysis, indicating a higher chance of rear end conflicts when the road condition was "snowy/wet."

Finally, the peak time (10 – 12 AM, 4-6 PM) showed no significance because there were always cases of congestion in day hours other than those time intervals, where a lot of congestion was observed at the intersection.

5.5.2.2 Analysis of Pedestrian and Left Turning Vehicles Conflict

The Type-17 conflict from the Bayesian conflict types table was identified as more severe than right turning vehicle and pedestrian conflicts due to the higher speed of left-turning vehicles, which turn with more speed because of the higher radius compared to right turns. In some cases, left-turning vehicles were observed to be in a hurry to turn before the opposing vehicles reached the intersection, resulting in conflict.

During the data analysis, a large number of cases were observed where pedestrians were in a rush to cross the crosswalk at night, resulting in conflicts. These incidents were only observed during the night, even on weekdays. In contrast, during the daytime, pedestrians were more cautious, resulting in fewer violations.

Based on these observations, two explanatory variables were added to the model: night peak hour and pedestrian violation, in addition to both left-turning vehicle volumes and pedestrian volumes.

Logit Regression Results

Dep. Variable:	Conflict	No. Observations:	60
Model:	Logit	Df Residuals:	55
Method:	MLE	Df Model:	4
Date:	Thu, 09 Mar 2023	Pseudo R-squ.:	0.5149
Time:	00:22:29	Log-Likelihood:	-20.176
converged:	True	LL-Null:	-41.589
Covariance Type:	nonrobust	LLR p-value:	1.125e-08

	coef	std err	z	P>|z|	[0.025	0.975]
const	-6.4422	2.062	-3.124	0.002	-10.484	-2.401
Ped left	0.2318	0.106	2.189	0.029	0.024	0.439
EBL	0.3388	0.128	2.641	0.008	0.087	0.590
violation	1.9177	1.094	1.753	0.080	-0.226	4.062
Npeak	1.4693	1.021	1.439	0.150	-0.532	3.470

AIC: 50.35293100624543
BIC: 60.82465381735593

Figure 59 Results for Pedestrian Type-17 Conflict Analysis

The odds ratio for conflicting pedestrian volume was 1.32, while the odds ratio for left turning vehicle volume was 1.25. Additionally, the odds ratio for pedestrian violation (running or in a hurry) was 2.97, and the odds ratio for the night peak (7:00 PM – 10:00 PM) was 1.73.

Based on the above results, it can be concluded that pedestrian behavior during night-time hours can increase the conflict risk by three times. Moreover, both pedestrian volume and eastbound left volumes have a significant impact on increasing the conflict risk.

It is also worth noting that there is a higher occurrence of conflicts during the night peak, which is linked to pedestrian behavior during this time. This is further supported by the fact that the total left turning vehicular volume decreases at night.

5.6 Traffic Conflict Framework Establishment

In this section, the analysis of two sets of traffic conflicts that were captured by fixed CCTVs at two signalized intersections (Town Square intersection at Jackson Hole, and Four Corners Camera intersection at Cold Water). The first set consisted of 8 conflicts captured from a low elevation camera (10-15 ft height above ground level). While the second set includes 2 conflicts that captured from high elevation fixation point (40-45 ft height above ground level). The utilization of the second intersection was to illustrate the strength of YOLOv7 algorithm with

higher elevation cameras. The following table illustrates the characteristics of the video footage for the two conflict sets, Table 6.

Table 6 Traffic Conflicts Sets Characteristics

Traffic conflict set	Conflict number	Associated crash type	Number of frames	Video resolution	Video extension	Notes
1	1	Angle-crash	481	1920 x 1080	mp4	
	2	Rear-end crash	434	1920 x 1080	mp4	
	3	Rear-end crash	494	1920 x 1080	mp4	
	4	Angle-crash	327	1920 x 1080	mp4	
	5	Rear-end crash	369	1920 x 1080	mp4	
	6	Right-angle crash	632	854 x 470	mp4	
	7	Head-on crash	418	854 x 470	mp4	
	8	Rear-end crash	710	854 x 470	mp4	Excluded
2	1	Rear-end crash	397	2560 x 1440	mp4	
	2	Rear-end crash	323	1920 x 1080	mp4	

The video analysis was done for the 10 conflicts by utilizing CenterTrack algorithm for the first conflict set. While for the second set, YOLOv7 algorithm was used for their analysis. Accordingly, the bounding cuboids coordinates were extracted for the first 8 conflicts and the bounding boxes coordinates were extracted for the remaining 2 conflicts. For conflict number 8 in conflict set 1, it was excluded from the analysis since the algorithm could not identify one of the conflicting vehicles (Snow removal 3 wheeled truck). The data validation was done, then, the conflict detection and analysis process were performed. Finally, the results were extracted and analyzed at the end of this chapter.

5.6.1 Traffic Conflict Types and Their Associated Types of Collision

Based on the illustrated conflicts IDs in the previous section, a table was created to link the proposed conflict types with their associated type of collision, Table 7.

Table 7 Conflict Types and Their Associated Crash Types

Traffic conflicts ID	Crash type
1, 2, 3 and 4	Rear-end crash
5, 6, 7, 10, 12, 13 and 16	Angle crash
8, and 9	Right-angle crash
11, 14, and 15	Head-on crash
17, 18, 19, and 20	Pedestrian crash

5.6.2 Regions Identification

The following step was to divide the intersection into several regions such that each traffic conflict is supposed to fall within one or more regions. Hence, a proposed regions diagram is drawn to estimate the most common regions in the intersection that are expected to contain specific types of traffic conflicts. By considering N-S approaches as the analysis directions, and with respect to conflicts points chart extracted from the traffic movements at signalized intersection (FHWA), seven proposed expected collision regions were selected. Figure 60 shows the set of conflicting points initiated by different traffic movements at signalized intersections. While Figure 62 shows the proposed regions diagram. The alignment process between the conflict types and their predicted including regions was done by setting approximate areas of occurrence for each conflict within the intersection. Then, these areas were annotated such that each area belongs to one or more regions. The region-conflict relation was summarized in Table 8.

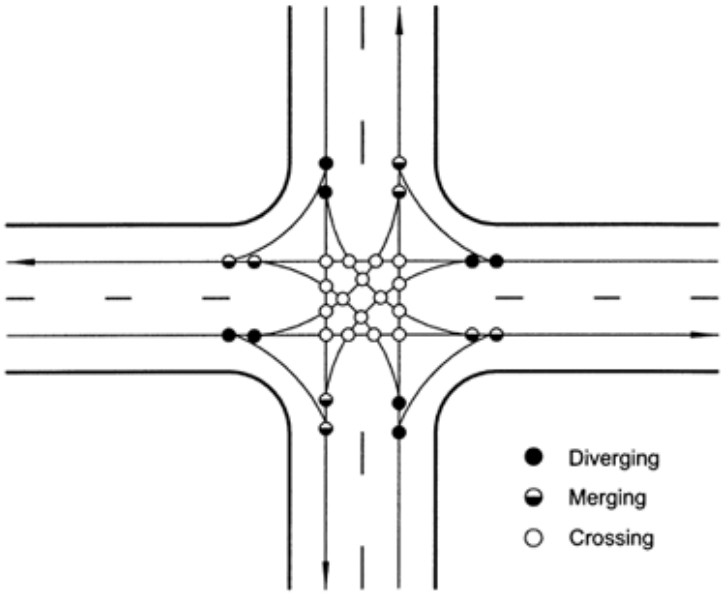


Figure 60 Conflict Points at Four Leg Signalized Intersection [95]

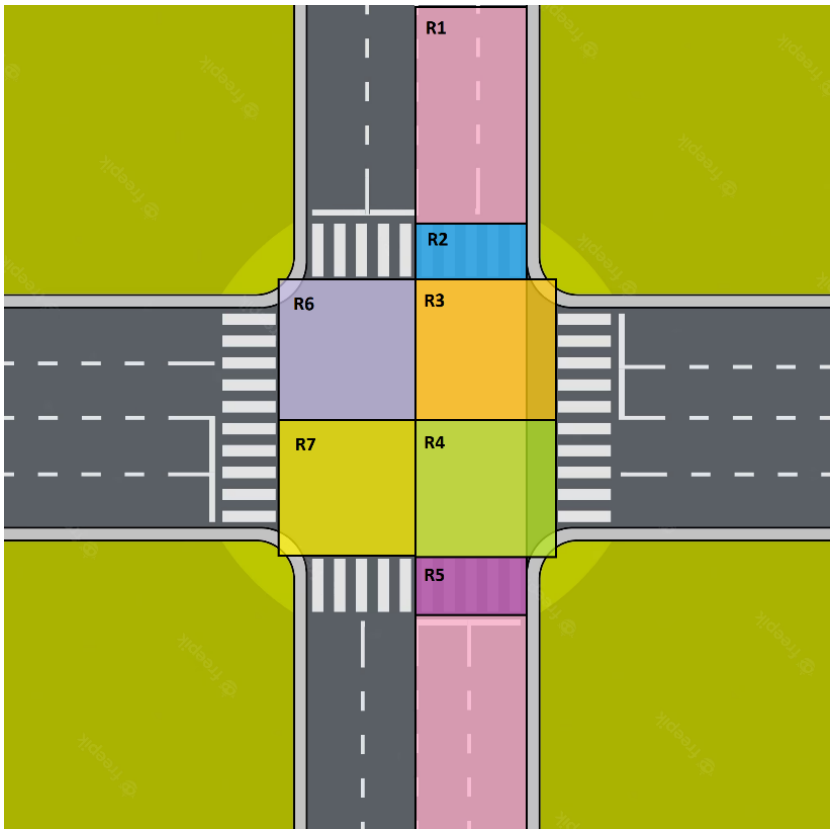


Figure 61 Proposed N-S Approaches Regions Diagram

Table 8 Region-Conflict Relation

Region ID	Predicted conflict type	Notes
1	3, 4 and 12	The regions could be inverted to expect traffic movements from the opposite direction
2	12, 17, and 18	NA
3	7, 8, 10, 12, 13, 14, 15, and 16	NA
4	1, 2, 5, 6, 9, 11, 14, 15, and 16	NA
5	1, 2, and 11	NA
6	15	NA
7	15	NA

Based on Table 7 and Figure 61, it could be noticed that conflict with ID 15 is considered the most critical conflict type since the collision type associated with this conflict is a head-on crash. Additionally, the occurrence of this conflict will be in regions 3, 4, 6, and 7 that increases the probability of the occurrence of secondary conflict in the middle of the intersection. The conflicts 3 and 4 could occupy any section of region 1 and any other sub region of it. That may happen due to the nature of the large number of conflict points at signalized intersection since the vehicles occupying region 1 must pass through 8 conflict points or more. Additionally, with

respect to the analysis approach, the most critical regions are 3 and 4 since each of them could include 8 types out of the 20 conflict types. However, regions 2, and 5 could include severe injuries or fatalities crashes and 3 conflict types are expected with their areas considered to be the minimum within the whole intersection. Regions 2 and 5 could be interpreted as the only regions that include pedestrians' conflicts.

Based on the studied conflict types, and their distribution on the different regions within the signalized intersection, it was required to follow the literature on the assessment tools for the conflict types and their severity level.

5.6.3 Traffic Conflicts Detection and Analysis Framework

This section illustrates the traffic conflicts detection and analysis framework in detail by employing two sets of pre-determined conflicts. The first conflict set consists of 8 videos that were captured using CCTV camera fixed at low elevation point (10-15 ft height above ground level) at Town Square intersection, Jackson Hole. The last 3 videos were taken from YouTube channel with video dating to February 2018. Selection of conflict 6, and 7 is done since they include a right-angle conflict, and head-on conflict, respectively. Conflict 8 was chosen to investigate whether the CenterTrack algorithm can identify unique types of vehicles or not. The remaining 5 conflicts were chosen from a list that was extracted by the research team. This list was created through an extensive study of traffic conflicts where two trained observers were employed to extract traffic conflicts from 39 hours of recorded videos from the real-time streaming YouTube channel for Town Square intersection. In addition to providing the observers with the required set of manuals for observation techniques, observers learnt conflicts extraction process through conducting introductory lectures followed by several group meetings [94]. Additionally, a collection form was created to gather the observed conflicts characteristics (time stamp, conflict type, road user at fault, motion description, weather, and surface condition, ...etc.) and then the data was collected in an ECXEL sheet. Furthermore, a cross-check was performed between the two observers as a revising procedure for achieving more accurate judgements.

The second conflict set included 2 videos that were captured at Four Corners Camera intersection at Cold Water city where the CCTV camera is fixed at a high elevation (40-45 ft

height over ground level). Similar to the first intersection, the conflicts were chosen from a list of manually observed conflicts that were extracted from real-time video feeds. The observation process was held on 12 hours of real-time videos that were captured from YouTube channel of Four Corners Camera intersection. The collected conflict sets were utilized as case studies for verifying the proposed framework for conflict detection and analysis.

5.6.3.1 Video Analytics of Conflicts

Following the videos selection process, the next step involved manually selecting the associated algorithm that is relevant to the video analysis process, refer to chapters 5 and 6. For the first conflict set, CenterTrack algorithm was utilized to analyze the videos, then the extracted data were saved for each conflict individually (video output, JSON file, TXT file). The video output displayed the occupying cuboids for each road user while following them frame by frame. While JSON file presented road user IDs, the coordinates of center points of bounding boxes, and dimensions of bounding cuboids. The TXT file was added to the algorithm code, and it gave the eight-point coordinates of the bounding cuboids of each road user in each frame cumulatively. Thus, cuboids coordinates needed to be integrated into the JSON file and aligned manually to assign cuboids coordinates to their associated road users in each individual frame. The data verification process illustrates the utilized procedure to assign the extracted coordinates for cuboids and road users.

For the second conflict set, the YOLOv7 algorithm was utilized in analyzing the conflicts. Video outputs, and data label TXT files were extracted from the video analytics process using YOLOv7. The data labels TXT files are individual files where each file is titled with the associated frame number and contain the labels for all road users included in the analyzed frame, the tracking id, and bounding box and center point coordinates. Following this is, the outputs were manually checked, road users' detection accuracy was verified, and the redundancy in the double detected objects were removed.

Furthermore, OpenPifPaf was examined on one conflict by employing TTC and vehicle speeds in identifying the conflict occurrence.

5.6.3.1.1 Traffic Conflicts Framework: Cases of Study

In this sub section, the framework of traffic conflicts detection and analysis was applied on the two conflict sets. Seven general steps were proposed and followed throughout the application of the framework. These steps are stated as follows:

- Define the target crash type from the analysis, and the requested approach for the analysis process within the intersection.
- Select the regions that are expected to include the targeted crash type (using region distribution diagram, conflict type-crash type table, and region-conflict type table).
- Exclude any road users' data that falls outside the selected regions.
- Extract road users' characteristics that fall in the selected regions.
- Calculate the traffic conflicts indicators that are associated with the selected crash type for the road users that fall within the selected regions.
- Determine the critical incidents based on the calculated indicators (conflict detection).
- Analyze the trajectories of these incidents to detect the conflict types associated with these incidents.

5.6.3.1.1.1 First Conflict Set

By referring to Table 6, the first conflict set was chosen to start the analysis. From the manual observations the conflict types are determined as illustrated in Table 9.

Table 9 First Conflict Set Characteristics

Traffic conflict set	Conflict number	Associated crash type	Conflict type	Regions expecting the crash	Notes
1	1	Angle-crash	12	2, and 3	SB
	2	Rear-end crash	3	1	
	3	Rear-end crash	2	1	
	4	Angle-crash	12	2, and 3	SB
	5	Rear-end crash	1	1	SB/Veh. Order changed
	6	Right-angle crash	9	3, and 4	
	7	Head-on crash	15	3, 4, 6, and 7	
	8	Rear-end crash	4	4	Excluded

Conflict 1 (Angle-crash/Side-swipe conflict):

The proposed procedure was followed by selecting regions 2, and 3 for the analysis of angle crashes' conflicts with respect to the south bound approach. Data of road users located in these regions were selected and analyzed, then, traffic conflicts indicators were calculated. It is worth mentioning that PET could not be calculated since one of the conflicting vehicles stopped before passing through the point of collision. Figure 62 illustrates the two conflicting vehicles. The grey vehicle (ID: 60) was turning left and had the right of way since the green arrow was turned on while the red vehicle (ID: 16) was about to complete the right turn and stopped after recognition of the coming vehicle.

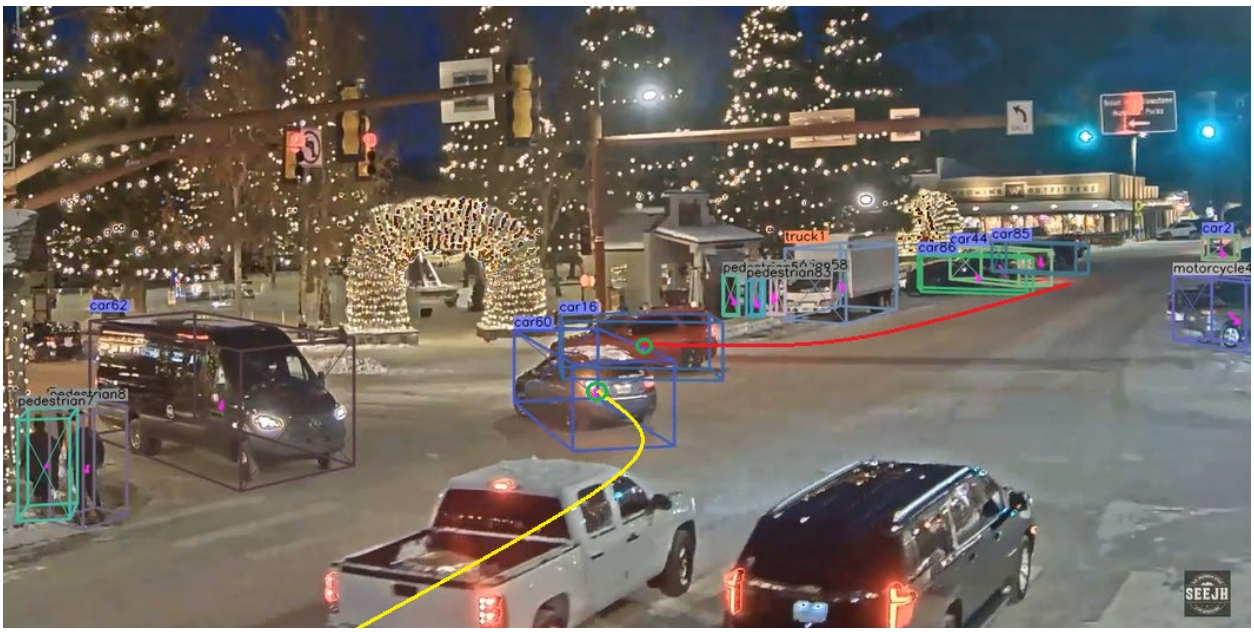


Figure 62 Video Analysis Output for Conflict 1 Utilizing CenterTrack Algorithm

The calculation of the traffic conflicts indicators showed a steep change in the TTC values for the two conflicting vehicles. While for MTTC, it could not detect a noticeable change since the values were fluctuating through different time series. For DRAC curve, it could detect a gradual decrease in the curve within the time of conflict occurrence. Figure 63 illustrates the three traffic conflicts indicators curves for the two conflicting vehicles. Each chart represents 3 curves: the calculation in x-direction, y-direction and the magnitude. That was proposed since the original equation assumed the analysis was in one direction. Finally, the trajectories for the conflicting vehicles were reconstructed to determine the conflict ID.

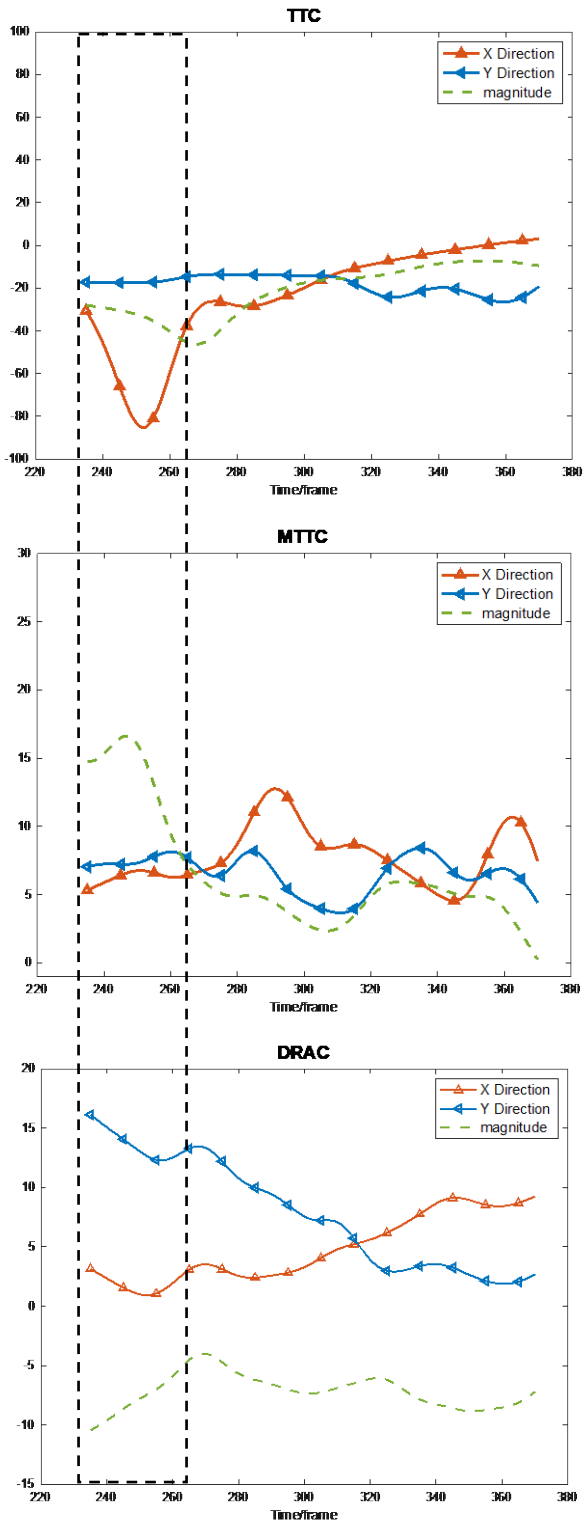


Figure 63 Traffic Conflicts Indicators Extracted Values Versus Time Frame for Conflict 1, Conflict Set 1

Conflict 2 (Rear-end conflict):

Following the proposed framework, rear-end conflicts occupying region 1 with respect to the movement in north bound approach were selected and the road users' data were analyzed within region's boundary. Accordingly, the traffic conflicts indicators were calculated. As for PET, the whole region could be considered as a collision course, thus it was essential to set a specific line as a point of collision to compare PET values for all the vehicles passing this line. The value of PET was 3 seconds between the two conflicting vehicles (e.g., silver vehicle (ID: 17), and black vehicle (ID: 47)) which is not considered as a good indicator for conflict occurrence. The interpretation of occurrence of such action is that the following vehicle, the black vehicle (ID: 47), decelerated before reaching the collision point. Which is another reason for not considering PET as a relevant indicator for such cases. The whole scene could be explained as follows; the black van (ID: 23) was getting out of the inclined parking lot that reserved the right lane. Consequently, the approaching vehicle's driver, silver vehicle (ID: 17), applied the brakes to avoid collision with the reversing vehicle, first conflict. Finally, the following vehicle, black vehicle (ID: 47), decelerated to avoid a rear-end crash with silver vehicle (ID: 17), secondary conflict. Figure 64 illustrates the explained scene. For the traffic conflict indicators curves, only the TTC curves detected the decrease that indicated a potential conflict occurrence, Figure 66.

Figure 64 Video Analysis Output for Conflict 2 Utilizing CenterTrack Algorithm

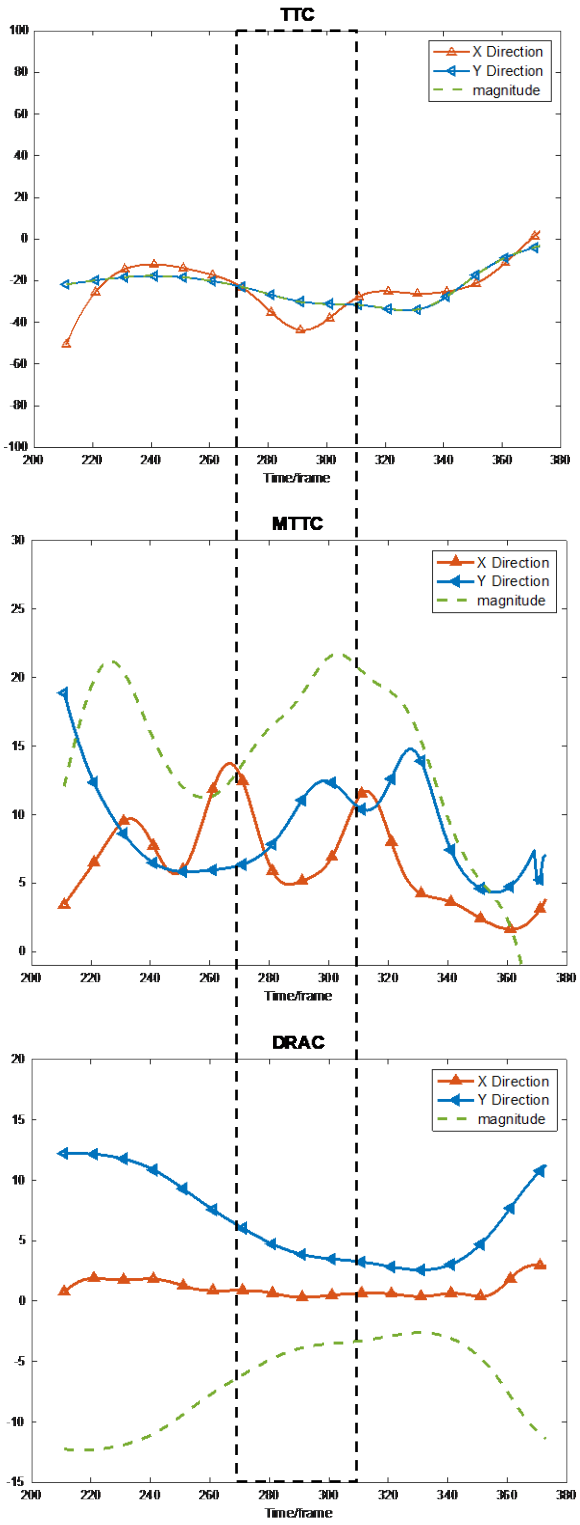


Figure 65 Traffic Conflicts Indicators Extracted Values Versus Time Frame for Conflict 2, Conflict Set 1

Conflict 3 (Rear-end conflict):

Similar to Conflict 2, the inclined parking that reserved the right lane is found to be a common factor for occurrence of rear-end conflicts as well as secondary conflicts and traffic congestions at this specific area of Jackson Hole intersection. In conflict 3, a parked vehicle, silver vehicle (ID: 45), was getting out of the parking lot when a fire truck (ID: 66) was travelling through the intersection. The fire truck had to maneuver to avoid being involved in a rear-end crash with the reversed vehicle. Figure 66 illustrates the vehicles' trajectories.

Figure 66 Video Analysis Output for Conflict 3 Utilizing CenterTrack Algorithm

For the traffic conflicts measurements, it was found that TTC, MTTC, and DRAC could detect the conflict occurrence while PET was not able to identify the conflicts since neither of the conflicting vehicles passed the proposed point of collision. Figure 67 shows the TTC, MTTC, and DRAC curves for silver vehicle (ID: 45), and fire truck (ID: 66). TTC, MTTC, and DRAC curves showed a steep fluctuation at the time of conflict occurrence.

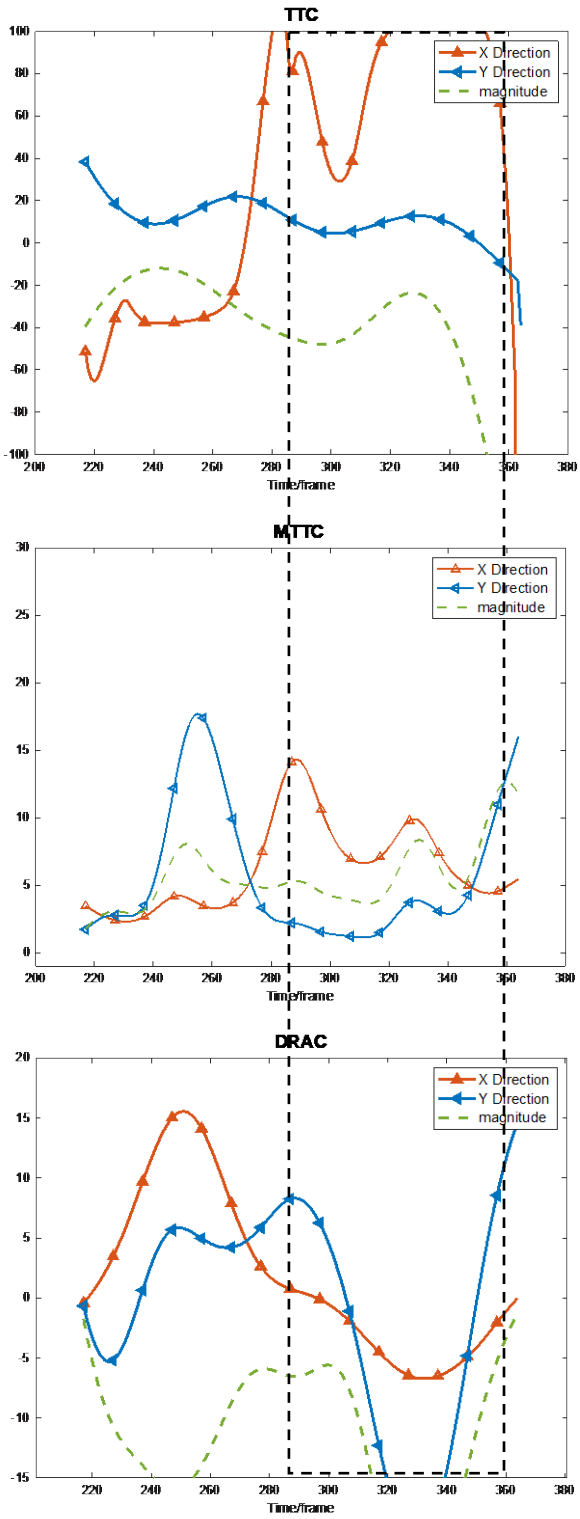


Figure 67 Traffic Conflicts Indicators Extracted Values Versus Time Frame for Conflict 3, Conflict Set 1

Conflict 4 (Angle-crash/Side-swipe conflict):

This conflict is typically similar to conflict 1. The only difference was that conflict 4 occurred in daytime unlike conflict 1 that occurred in nighttime. It should be mentioned that this area within the intersection and with respect to camera's viewpoint included several side swipes conflicts at the snowy weather conditions. This could be interpreted as a result of the usage of the right lane as a parking lot while the remaining part of the lane is covered by snow. This forces the right turning traffic to initiate the right turn from the middle lane. Consequently, side swipes conflicts occurred with the intersecting volumes. Figure 68 describes the scene. The grey SUV (ID: 32) was turning right from the middle lane while the grey Jeep (ID: 2) was taking a protected left turn as illustrated from the traffic signal. The driver of the grey SUV (ID: 32) felt that it will not be possible to end the right turn safely before the arrival of the grey Jeep (ID: 2). Subsequently, he applied brakes before completing the turn. While for the grey Jeep (ID: 2), the driver decelerated while taking the left turn to avoid the occurrence of a potential crash.

Figure 68 Video Analysis Output for Conflict 4 Utilizing CenterTrack Algorithm

The TTC and MTTC traffic conflicts indicators curves identified the conflict occurrence while DRAC could not detect the conflict occurrence since both conflicting vehicles did not use their brakes violently. Figure 69 illustrates the output curves.

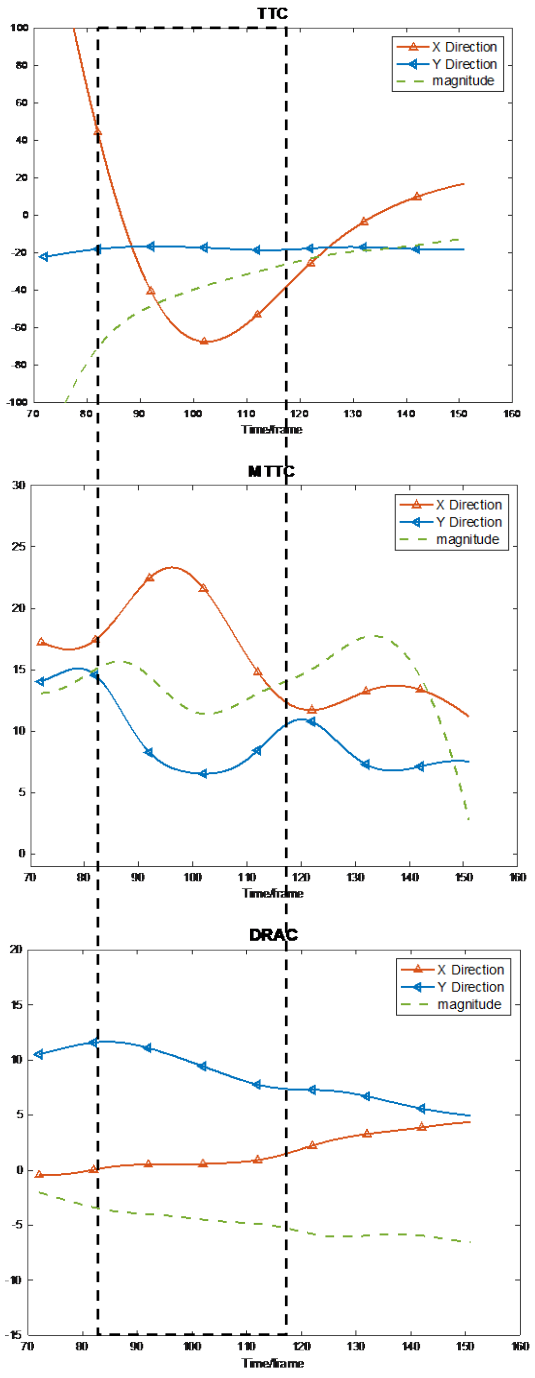


Figure 69 Traffic Conflicts Indicators Extracted Values Versus Time Frame for Conflict 4, Conflict Set 1

Conflict 5 (Rear-end conflict):

This conflict occurred between a left turn vehicle, the grey truck (ID: 51), that was travelling toward the northbound from the eastbound and a reversed vehicle, black van (ID: 15), which

was getting out of the inclined parking lot located on the northbound. The view of grey truck driver (ID: 51) was blocked by the through SB traffic that stopped at the entrance of the intersection at the red-light phase. Hence, grey truck driver (ID: 51) could not recognize the reserved black van (ID: 15) until his vehicle passed the blocking volume. Once, the reversing vehicle was recognized by grey truck driver (ID: 51), he decelerated to avoid the rear end collision. The framework with followed by selecting region 1 and excluding the remaining regions. The included vehicles in region 1 were studied, then, their characteristics were extracted, and conflicts indicators curves were drawn. Figure 71 describes the scene. While the extracted measurements were shown in Figure 72.

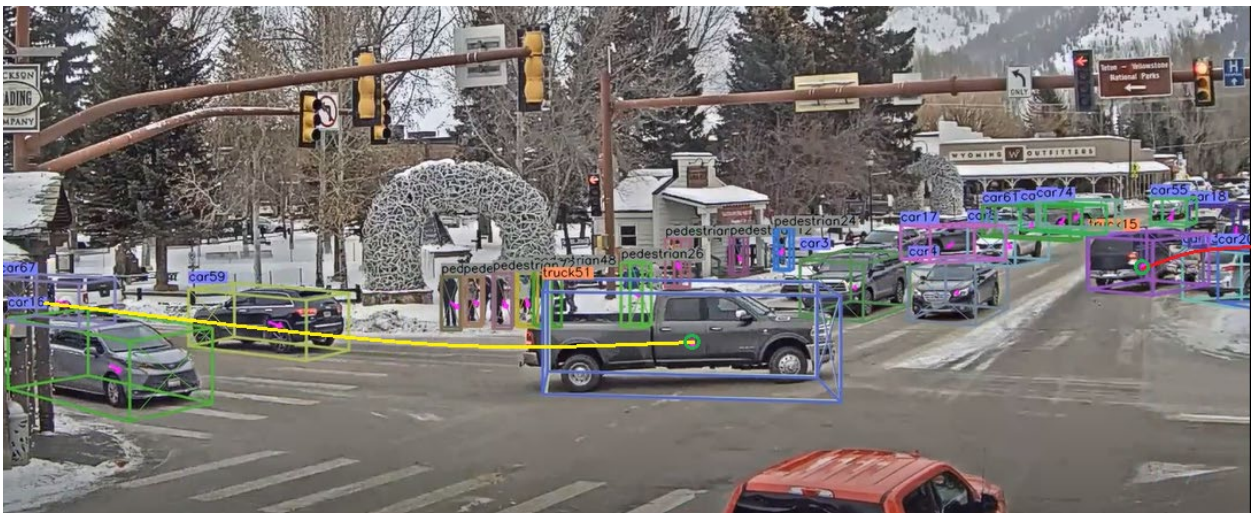


Figure 70 Video Analysis Output for Conflict 5 Utilizing CenterTrack Algorithm

The extracted traffic conflicts indicator curves for TTC, and MTTC identified the conflict occurrence. While, DRAC curve had a gradual change in the curve at the time of conflict. Also, PET could not be utilized in this case since both conflicting vehicles did not pass the potential collision point. The selected traffic conflicts indicators were able to clearly identify the traffic conflict occurrence since the utilization of DRAC is excluded on giving guidance measurements only.

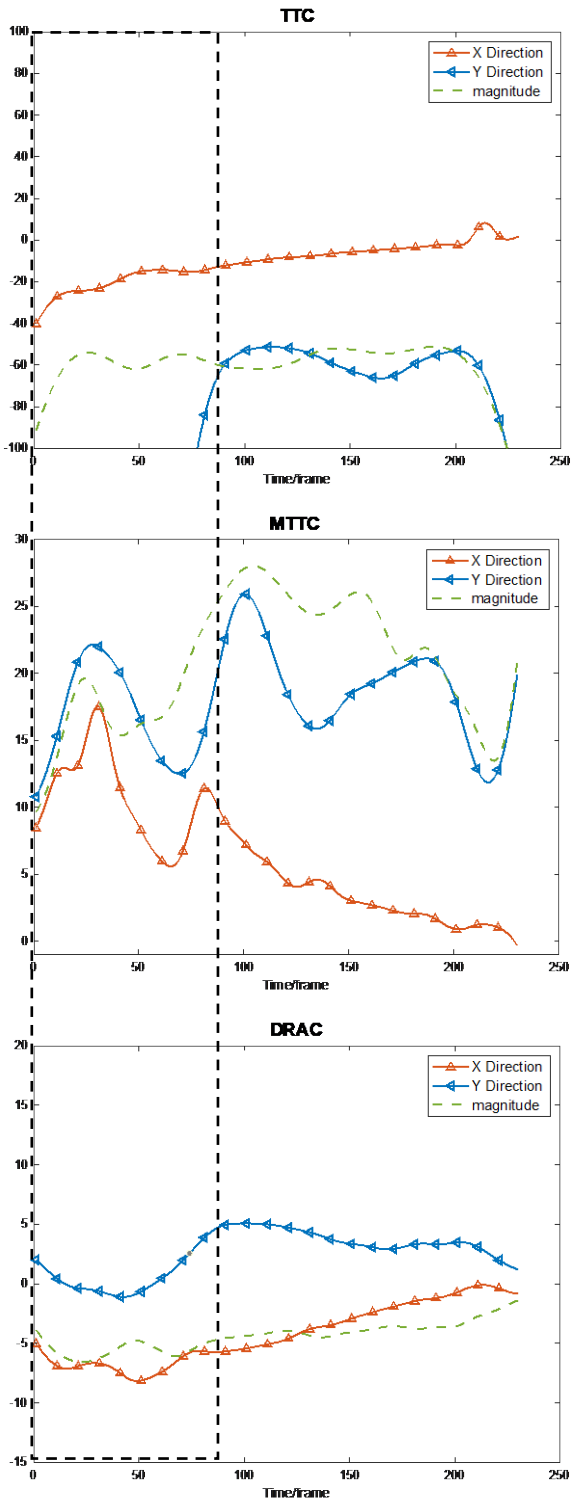


Figure 71 Traffic Conflicts Indicators Extracted Values Versus Time Frame for Conflict 5, Conflict Set 1.

Conflict 6 (Right-angle conflict):

Based on the proposed procedure, right- angle crash was selected. Hence, regions 3, and 4 were chosen while other regions were excluded from the analysis. The following figure shows a captured frame from the output video of conflict 6, Figure 70. The video is captured at nighttime where all flashing red phase was set. This phase is usually applied at low traffic volume intersections whereas all vehicles that approach the intersection are required to fully stop to check the crossing traffic before travelling through the intersection. The figure shows that the red SUV (ID: 5) had the right of the way since the predicted crash will affect the vehicle's driver side.

The vehicle trajectories were drawn. The performed video analytics showed that both vehicles; grey SUV (ID: 1) and red SUV (ID: 5) applied their brakes to avoid collision. Hence, PET could not be applied to detect the traffic conflict since there was no crossing movement. Consequently, other indicators were calculated for the extracted data for vehicles located in regions 3, and 4.

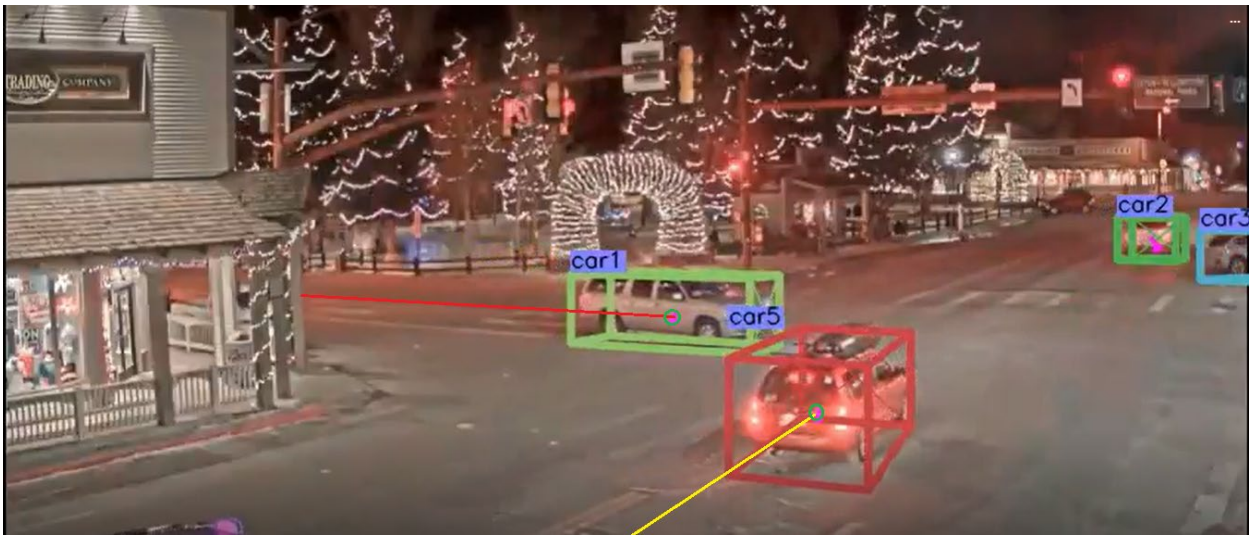


Figure 72 Video Analysis Output for Conflict 6 Utilizing CenterTrack Algorithm

It is concluded that by calculating the TTC, MTTC, and DRAC for the vehicles traversing regions 3, and 4 that the curves have rapid fluctuation for the calculated indicators between grey SUV (ID: 1), and red SUV (ID: 5). Figure 73 illustrates the drops in these curves within the time series including the occurred conflict.

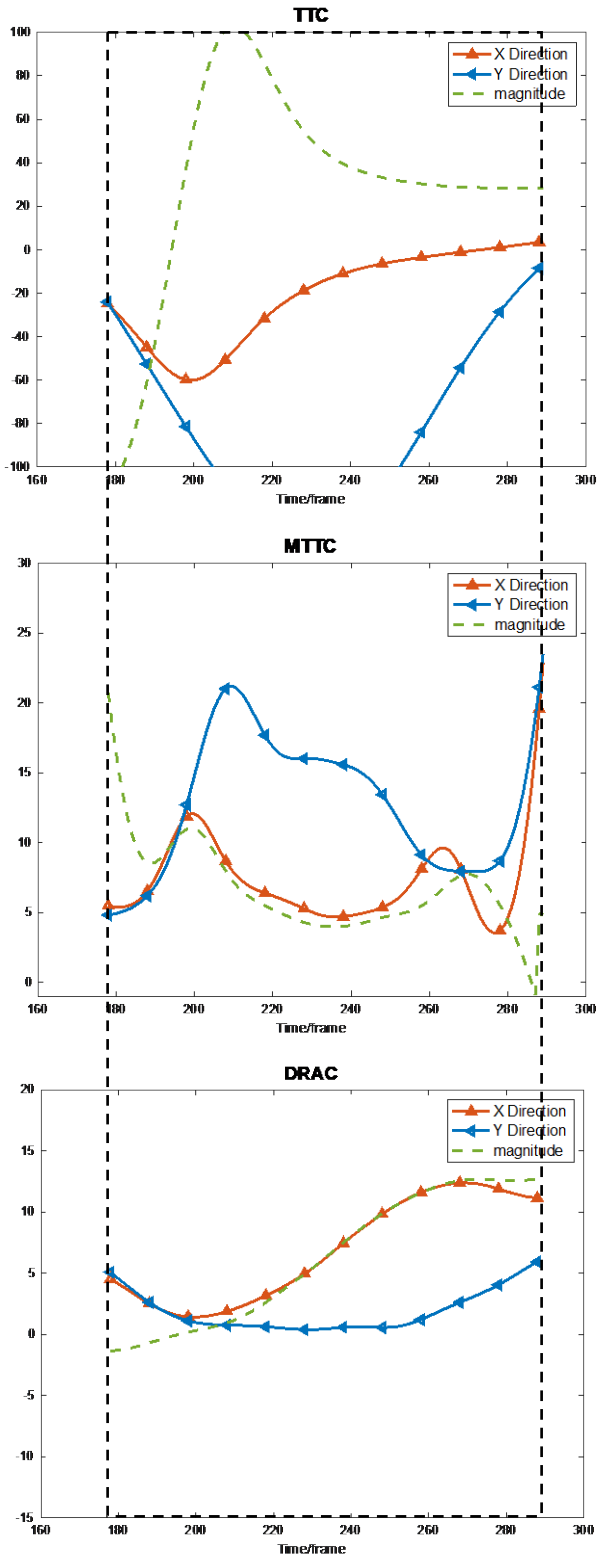


Figure 73 Traffic Conflicts Indicators Extracted Values Versus Time Frame for Conflict 6, Conflict Set 1.

It should be mentioned that DRAC provided a gradual decrease in its value which could be utilized as a guidance for the potential conflict occurrence. While for TTC and MTTC provided a steep change in their values. Additionally, the calculation equations provided a one-dimensional analysis depending on the movement of vehicles occurs in one directional collision course (e.g., rear-end conflict). Consequently, the values were calculated once in x-direction, y-direction, and their magnitude. All of the calculated values for the three curves representing the 3 components could detect the steep change in movement. Finally, returning to the vehicles data, trajectories were drawn, and conflict type was found from the proposed list.

Conflict 7 (Head-on conflict):

For conflict 7, the associated regions with the crash type are 3, 4, 6, and 7. These regions were selected while other data that do not belong to them were excluded from the analysis process. Similar to the applied procedure in conflict 6, the traffic conflicts indicators were calculated for the road users occupying the mentioned regions. Accordingly, the curves were checked for the road users to detect sudden fluctuations. Figure 74 illustrates the video analytics of conflict 7. The black SUV (ID: 15) was taking a left turn at the yellow light after the end of permitted left turn phase. While the silver SUV (ID: 11) had the right of the way since the black SUV (ID: 15) should not make the turn unless a relevant gap is provided. Both vehicles applied sudden brakes to avoid the occurrence of a head on collision. Figure 75 shows the calculated values for TTC, MTTC, and DRAC. For PET, similar to conflict 6, both vehicles applied sudden brakes to avoid collision occurrence. Hence, PET could not be calculated for this conflict.

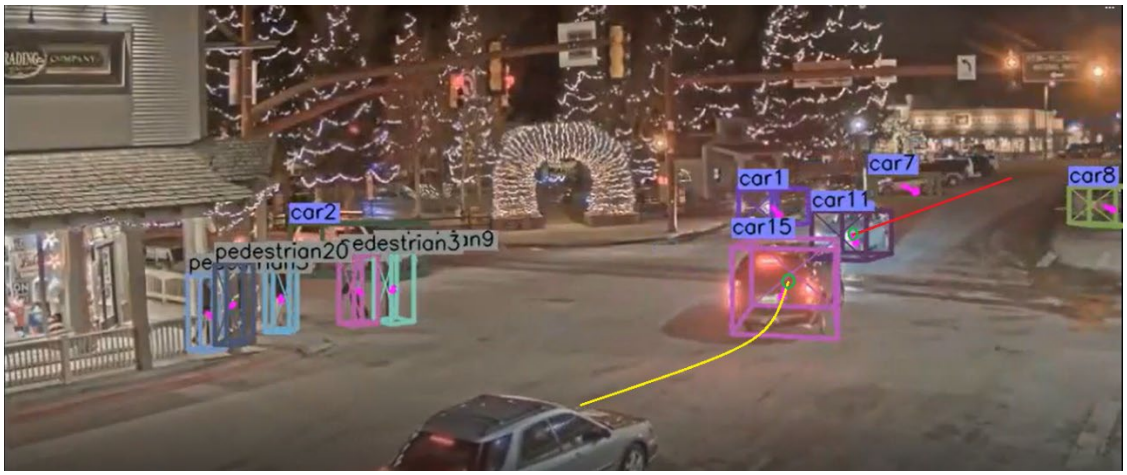


Figure 74 Video Analysis Output for Conflict 7 Utilizing CenterTrack Algorithm

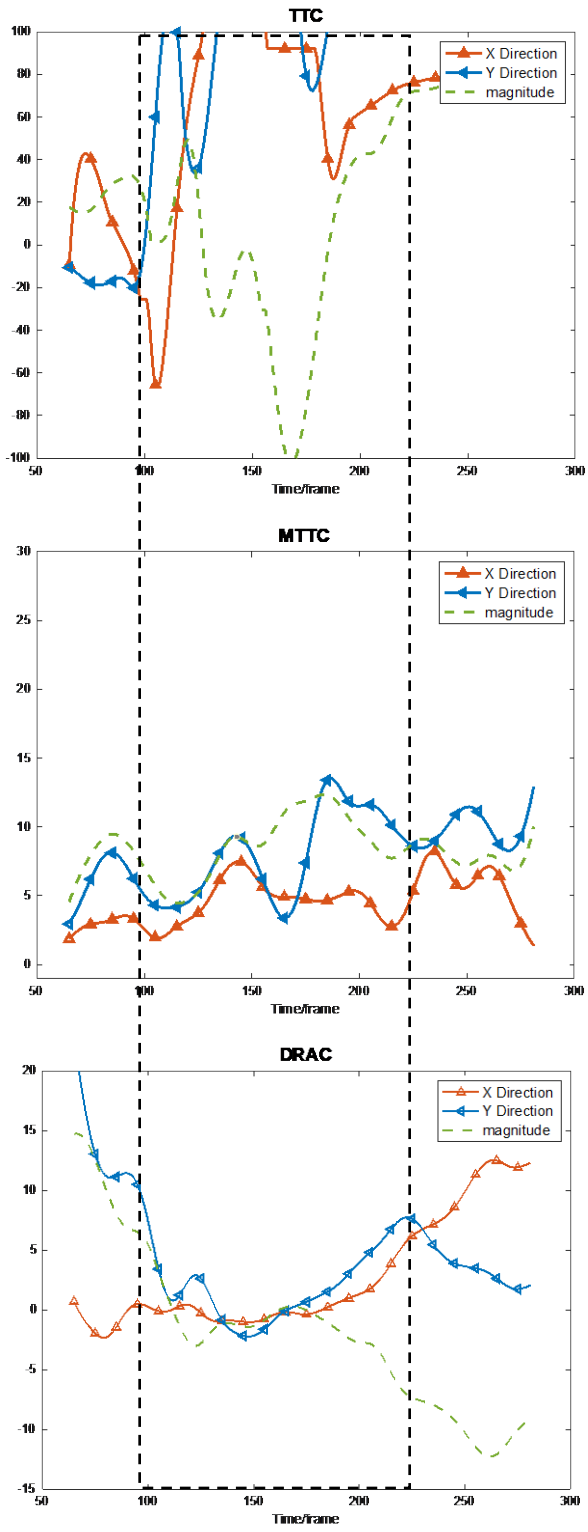


Figure 75 Traffic Conflicts Indicators Extracted Values Versus Time Frame for Conflict 7, Conflict Set 1

However, TTC could detect a steep change in the values for the two conflicting vehicles, MTTC could not detect the occurred conflict. It may be interpreted as the crash type was head-on crash, the different signs for acceleration values that resulted from the different directions could not identify the change in movement. Another issue could occur with the higher volumes traversing the intersection for this type of crash since it associated with four regions, and for the ordinary directional movements that travelling through, false detections could occur. Hence, it is recommended for that type of crash to narrow the selected scope and depend on the integration between two or more measures for more accurate detection.

5.6.3.1.1.2 Second Conflict Set

For the second dataset, YOLOv7 was utilized for the video analytics of traffic conflicts. Table 10 illustrates the characteristics of the second conflict set. The proposed framework for detection and analysis of traffic conflicts was applied to extract the conflicts and identify their types based on the road users' trajectories.

Table 10 Second Conflict Set Characteristics

Traffic conflict set	Conflict number	Associated crash type	Conflict type	Regions expecting the crash	Notes
2	1	Rear-end crash	12	1	SB
	2	Rear-end crash	3	1	

Conflict 1 and 2 (Angle-crash/Side-swipe conflict):

For conflict set 2, the two conflicts were chosen to justify the proposed algorithm selection for the captured videos from the mounted cameras at high elevations. Conflicts were recorded from the real-time channel of 4 corner camera at Cold Water city, Downtown Michigan. By using manual observations, the conflicts were extracted. It should be mentioned that the number of observed hours required to extract these two conflicts exceeded 10 hours. This could be resulted from the near ideal geometric design, satisfying the required drivers sight distances, and relevant signal timings at this intersection. For conflict 1, the silver SUV (ID: 42) was taking a left turn through a permissible phase while at the same instance another vehicle, red SUV (ID: 62), was taking a right turn from the opposite approach. However, the silver SUV is the vehicle at fault, the red SUV while taking the right turn movement deviated from the right

lane in uninterpretable behavior to block the lane on the silver SUV's way. The silver vehicle had to decelerate to almost full stop while entering the opposite direction lane. Figure 76 illustrates the occurred scene.

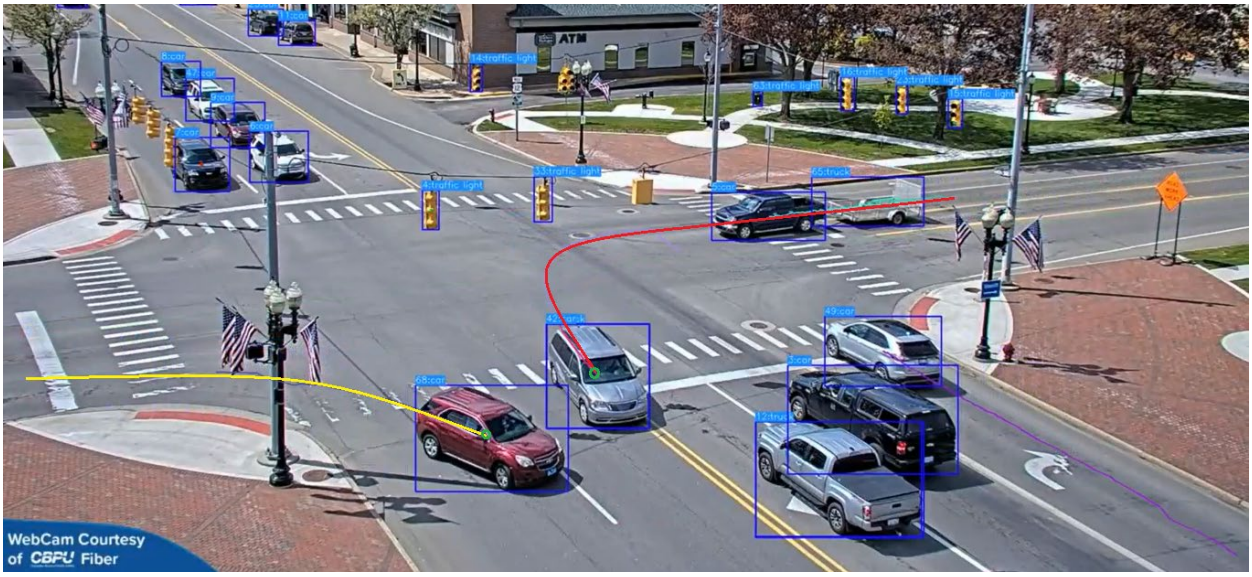


Figure 76 Video Analysis Output for Conflict 1 Utilizing Yolov7 Algorithm

Traffic conflicts indicators curves were extracted as shown in Figure 77. TTC and MTTC curves could detect fluctuations in their curves at the time of conflict occurrence. While the DRAC curve could not identify the conflict occurrence. To calculate PET, an estimated collision point was set then the arrival time of each vehicle is calculated. Consequently, PET was found to be less than one second. The PET value indicates that the occurred conflict is a severe one.

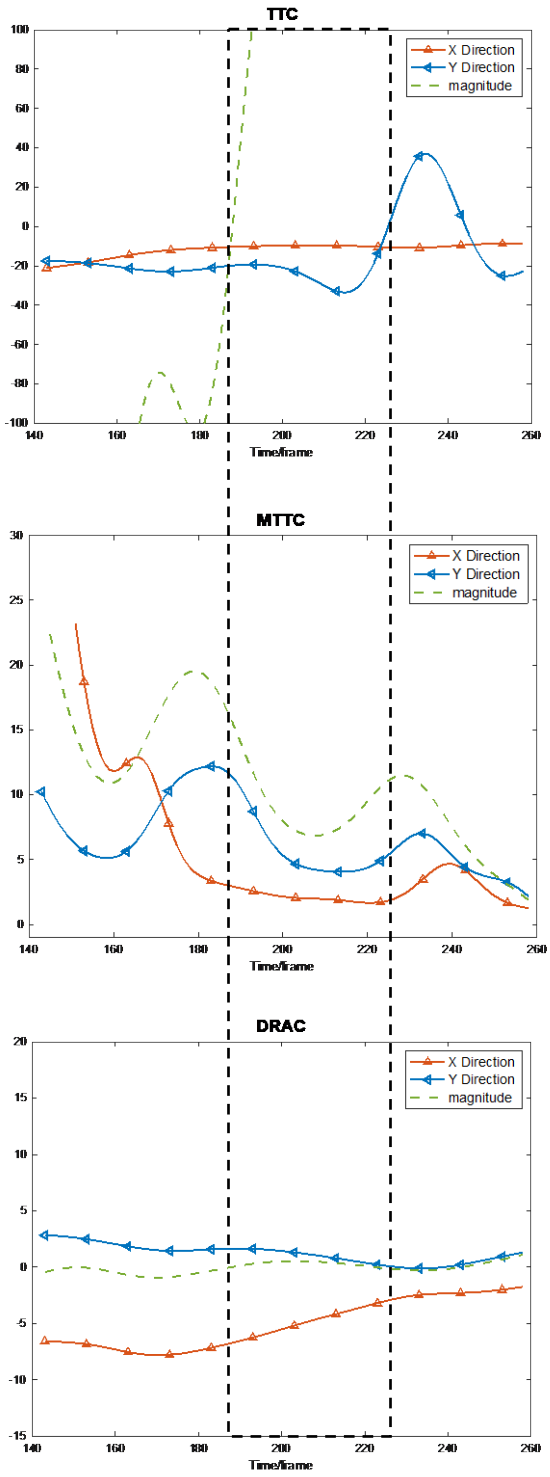


Figure 77 Traffic Conflicts Indicators Extracted Values versus Time Frame for Conflict 1, Conflict Set 2

Another conflict was analyzed by using OpenPifPaf algorithm. The outputs from two cameras were utilized to detect the vehicles positions, then, creating polygons representing vehicles by their average dimensions, Figure 78. The utilized indicator in the conflict detection was TTC. Figure 79 demonstrates the increase in vehicle 3 speed when approaching the intersection to turn right. This action led to a possible collision occurrence that was detected by the TTC when its value falls under 2.0 seconds.

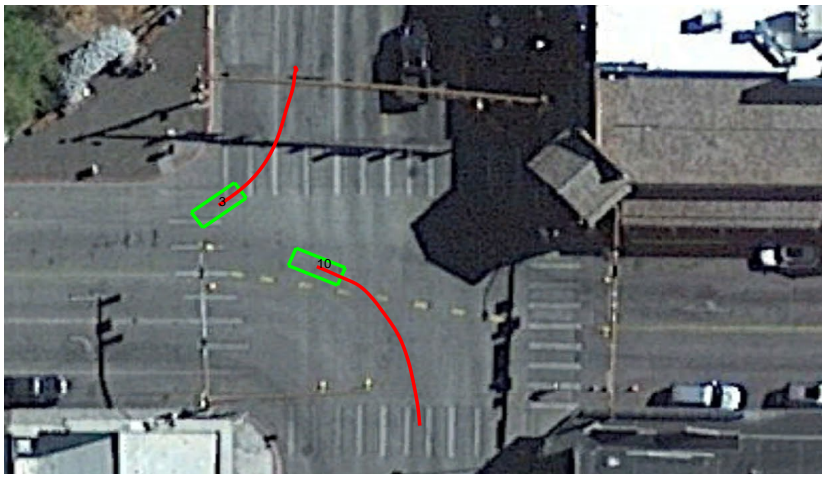


Figure 78 Sideswipe/Angle Traffic Conflict

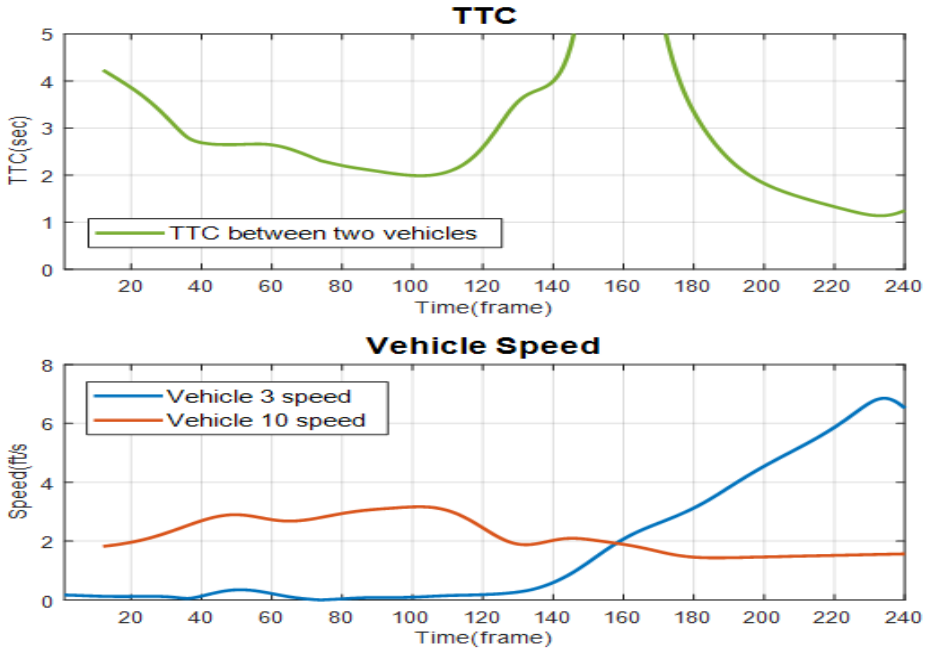


Figure 79 TTC and Conflicting vehicles Plots to Identify Side Swipe Conflict Occurrence

5.7 Integration of YOLOv7_human_pose_estimation Algorithm and OpenPifPaf

The integration between human and vehicles key points detection algorithms is investigated by analyzing the movement of pedestrians using YOLOv7_human_pose_estimation algorithm. Then, analyzing the video output using OpenPifPaf algorithm to extract vehicle key points. The integration between the outputs will serve in the traffic conflicts detection by calculating the road users' trajectories and apply either of the followed procedures in identifying conflicts occurrences. The dependence on key points detection algorithms provides more accurate localization of the road users which leads to minimizing the false identifications for traffic conflicts. Figure 80 shows an analyzed video frame using of both YOLOv7_human_pose_estimation and OpenPifPaf algorithms.

Figure 80 Integration of Human and Vehicles Key Points Detection Algorithms

5.8 Discussions

From the presented work throughout this chapter, the main concept of utilizing AI technology in identifying traffic conflicts occurrence has been proven. Furthermore, other gains were attained such as the manipulation between algorithms to assign the superiority of each one with respect to the camera's mounting elevation. Additionally, it was concluded that based on

the nature of occurrence of traffic conflicts and it dependance on several factors (e.g., drivers' behaviors, geometric design of intersections, sight distances, signal timings, weather, and surface conditions, etc.), the implementation of a detection system relies heavily on the integration between traffic conflicts indicators. The utilization of a single measurement could be superior for a single crash type under a specific condition. While for the sake of constructing a general framework that can provide solid interpretations and descriptions for the whole scene, an implementation of number of conflicts measurements should be done with the guidance of the road users' trajectories. This will require various case studies of intersections and traffic conflicts to cover most of the surroundings and address the limitations. For the presented study through this research work two intersections were included as examples for the low and high elevations of the mounting points of traffic surveillance cameras. Also, four traffic conflicts indicators were employed as a base for the proposed framework of detection and analysis of traffic conflicts. The proof of concept was achieved. While further modifications need to be performed to address the limitations in the proposed procedure.

CHAPTER 6. CONCLUSIONS AND RECOMMENDATIONS

The study addresses the prevalent practice of employing surrogate measures for identifying traffic conflicts with often arbitrarily determined threshold values. Current thresholds, typically chosen without clear justification, lack a standardized approach. While some researchers attempt to establish optimal threshold values through empirical methods, the present study employs systematic review and Meta-Analysis techniques to synthesize findings across various case studies, locations, and data configurations. However, Meta-Analysis proved to be challenging due to the diverse methods and frameworks used in reporting research outputs, compounded by insufficient studies with similar outputs and a lack of essential information, such as standard error and sample size. The study also quantitatively assessed the quality of selected papers, identifying empirical methods as potentially the most effective for determining surrogate measures of safety (SMoS) thresholds based on location characteristics and available data. It suggests that future research should focus on refining existing empirical and mathematical methods to enhance threshold value identification in the context of traffic safety studies.

The investigations of traffic conflicts and crashes at a specific intersection, employing heat maps for both data types and conducting a visual comparison. Notably, crashes were concentrated at the intersection's center, while conflicts were observed predominantly on the right side. This disparity resulted from the camera's westbound focus, skewing conflict data. Manual observation, despite limitations, captured all visible conflicts but displayed inconsistencies in quantitative variables like Time-to-Collision (TTC).

Despite manual observation's qualitative advantage in incident breakdown, computer vision excelled in precisely detecting quantitative factors such as speed, deceleration, and TTC—essential for conflict severity assessment. The study advocates a combined approach leveraging the strengths of both methods for more accurate conflict analysis, with recommendations to address camera limitations through increased height and distance.

Further analysis highlighted rear-end conflicts and those involving left-turning vehicles and pedestrians as most severe. Parking orientation in the eastbound approach, particularly angular

or side-by-side parking, contributed significantly to severe rear-end conflicts. The study recommends adopting parallel parking in this approach to align with established traffic guidelines, emphasizing the need for additional research to solidify the correlation between parking orientation and conflict occurrence. The study underscores the value of integrating manual observation and computer vision to devise effective strategies for reducing crashes and conflicts at intersections. While advocating a combined approach, the study acknowledges challenges in capturing all conflicts when the camera's scope is limited, proposing camera height and distance adjustments as potential solutions.

The study initiates a comprehensive exploration of vehicle tracking methodologies, spanning traditional approaches like GPS and RFID tracking to more advanced techniques incorporating AI and deep learning, such as YOLO and OpenPifPaf. Through an extensive literature review, a common limitation is identified: many existing tracking techniques prioritize tracking itself, often overlooking adaptability and effectiveness in diverse traffic scenarios. This gap leads to notable tracking errors, particularly in complex and varied traffic conditions.

In response to this identified limitation, our research introduces a tailored post-processing procedure aimed at enhancing tracking accuracy and adaptability. Taking OpenPifPaf as an example, common tracking errors are analyzed, resulting in the development of a novel algorithm proficient in extracting accurate feature points from raw data and constructing refined vehicle trajectories. Augmented by a uniquely developed filter for trajectory smoothing, this innovative approach facilitates precise estimations of speed and acceleration, significantly improving the reliability and accuracy of vehicle tracking in challenging scenarios.

In conclusion, this research not only highlights the shortcomings of existing methodologies but also pioneers a multifaceted approach that addresses these challenges. The post-processing improvement enhances tracking algorithm accuracy, opening new avenues for continued investigation and refinement in the dynamic domain of vehicle tracking technology. This contribution is significant in advancing the broader discourse on intelligent transportation systems.

Throughout this report, the primary concept of employing AI technology for identifying traffic conflicts has been substantiated. The study also achieved additional benefits, including

optimizing algorithms to discern their superiority based on the camera's mounting elevation. Importantly, it was deduced that the detection of traffic conflicts relies heavily on integrating various indicators due to the multifaceted nature of conflicts influenced by factors like driver behavior, geometric intersection design, sight distances, signal timings, weather, and surface conditions.

While a single measurement may excel for a specific crash type under particular conditions, constructing a comprehensive framework necessitates the implementation of multiple conflict measurements guided by road users' trajectories. This entails conducting diverse case studies to encompass different intersections and traffic conflict scenarios, addressing inherent limitations. The current study exemplified this with two intersections representing low and high camera mounting elevations, utilizing four traffic conflicts indicators as a foundation for the proposed detection and analysis framework. The proof of concept has been established, but further refinements are required to overcome limitations in the proposed procedure.

Based on the findings and analysis presented in this report, the following recommendations are made:

- Utilize CenterTrack algorithm in low-elevation camera fixation point: The results showed that CenterTrack algorithm performed better in detecting traffic conflicts at Town Square intersection, which had a low-elevation camera mounting point. Therefore, we recommend that this algorithm be used by applying some modifications to the original code, then, to be utilized at locations where the camera is closer to the ground.
- Utilize YOLOv7 algorithm in high-elevation camera fixation point: The analysis revealed that YOLOv7 algorithm outperformed CenterTrack algorithm in detecting traffic conflicts at Four Cameras square intersection, which had a high-elevation camera fixation point. Therefore, we recommend that this algorithm be used in similar settings where the camera is mounted at a higher elevation.
- Utilize OpenPifPaf algorithm at the intersections that have multiple cameras with intersected fields of view. The application of this algorithm will improve the detection accuracy resulting in accurate safety assessments by detecting vehicles key points.

- Cameras positions and orientations: The analysis results at both intersections revealed that the installed cameras should be covering the whole approaches of the intersection for better detection of the road users as well as achieving more accuracy in the transformation of the coordinates from image plane to the top view plan of the intersection.
- It is recommended that future research endeavors focus on the continuous refinement and development of post-processing techniques, extending beyond the presented methodology. Exploring alternative algorithms and filters, as well as conducting comparative analyses with different tracking systems, could provide valuable insights into optimizing the overall performance of vehicle tracking in various dynamic traffic environments. This iterative approach to advancing post-processing methodologies would contribute to the evolution of intelligent transportation systems and ensure their effectiveness across a broad spectrum of scenarios.
- Expand testing to include diverse weather conditions and explore day and night detection capabilities. Given that most featured case studies were conducted in clear weather, validating algorithm robustness requires additional testing in varying conditions like rain, fog, or snow.

Despite limitations encountered during the implementation of several types of algorithms in traffic conflict detection, the research team has proactively addressed these issues. The identified limitations included miss detections, multiple detections, missing coordinates, and an incorrect number of tracked objects. Proposed solutions include refining the training dataset, fine-tuning algorithm parameters, and implementing modifications to rectify output discrepancies. The results from this study leveraging computer vision techniques underscores the viability of effectively tracking road users, analyzing traffic conflicts, and assessing the efficacy of countermeasures in a more rapid and cost-effective manner.

REFERENCES

[1] Du, S., M. Ibrahim, M. Shehata, and S. Member. Automatic License Plate Recognition (ALPR): A State-of-the-Art Review. Vol. 23, No. 2, 2013.

[2] Arth, C., H. Bischof, and C. Leistner. TRICam - An Embedded Platform for Remote Traffic Surveillance. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol. 2006, 2006, pp. 0–8. https://doi.org/10.1109/CVPRW.2006.208.

[3] Silva, J. G., G. C. Marques, P. M. Jorge, A. J. Abrantes, A. L. Osório, J. S. Gomes, and J. C. Braga. Evaluation of an LPR-Based Toll Enforcement System on Portuguese Motorways. 2006.

[4] Rao, W., Y. J. Wu, J. Xia, J. Ou, and R. Kluger. Origin-Destination Pattern Estimation Based on Trajectory Reconstruction Using Automatic License Plate Recognition Data. *Transportation Research Part C: Emerging Technologies*, Vol. 95, No. June 2018, pp. 29–46. https://doi.org/10.1016/j.trc.2018.07.002.

[5] Hu, L., and Q. Ni. IoT-Driven Automated Object Detection Algorithm for Urban Surveillance Systems in Smart Cities. *IEEE Internet of Things Journal*, Vol. 5, No. 2, 2018, pp. 747–754. https://doi.org/10.1109/JIOT.2017.2705560.

[6] Rashid, M. M., A. Musa, M. A. Rahman, N. Farahana, and A. Farhana. Automatic Parking Management System and Parking Fee Collection Based on Number Plate Recognition. *International Journal of Machine Learning and Computing*, Vol. 2, No. 2, 2012, pp. 93–98. https://doi.org/10.7763/ijmlc.2012.v2.95.

[7] Parker, M., and C. Zegeer. *Traffic Conflict Techniques for Safety and Operations (Observers Manual)*. 1989.

[8] Sayed, T., and S. Zein. Traffic Conflict Standards for Intersections. *Transportation Planning and Technology*, Vol. 22, No. 4, 1999, pp. 309–323. https://doi.org/10.1080/03081069908717634.

[9] Saunier, N., and T. Sayed. A Feature-Based Tracking Algorithm for Vehicles in Intersections. *Third Canadian Conference on Computer and Robot Vision, CRV 2006*, Vol. 2006, 2006. https://doi.org/10.1109/CRV.2006.3.

[10] Ismail, K., T. Sayed, N. Saunier, and C. Lim. Automated Analysis of Pedestrian-Vehicle Conflicts Using Video Data. *Transportation Research Record*, No. 2140, 2009, pp. 44–54. https://doi.org/10.3141/2140-05.

[11] Saunier, N., T. Sayed, and K. Ismail. Large-Scale Automated Analysis of Vehicle Interactions and Collisions. *Transportation Research Record*, No. 2147, 2010, pp. 42–50. https://doi.org/10.3141/2147-06.

[12] St-Aubin, P., N. Saunier, and L. Miranda-Moreno. Large-Scale Automated Proactive Road Safety Analysis Using Video Data. *Transportation Research Part C: Emerging Technologies*, Vol. 58, 2015, pp. 363–379. https://doi.org/10.1016/j.trc.2015.04.007.

[13] El-Basyouny, K., and T. Sayed. Safety Performance Functions Using Traffic Conflicts. *Safety Science*, Vol. 51, No. 1, 2013, pp. 160–164. https://doi.org/10.1016/j.ssci.2012.04.015.

[14] Zhang, X., P. Liu, Y. Chen, L. Bai, and W. Wang. Modeling the Frequency of Opposing Left-Turn Conflicts at Signalized Intersections Using Generalized Linear Regression Models. *Traffic Injury Prevention*, Vol. 15, No. 6, 2014, pp. 645–651. https://doi.org/10.1080/15389588.2013.860526.

[15] Sacchi, E., and T. Sayed. Conflict-Based Safety Performance Functions for Predicting Traffic Collisions by Type. *Transportation Research Record*, Vol. 2583, No. 2583, 2016, pp. 50–55. https://doi.org/10.3141/2583-07.

[16] Sacchi, E., and T. Sayed. Bayesian Estimation of Conflict-Based Safety Performance Functions. *Journal of Transportation Safety and Security*, Vol. 8, No. 3, 2016, pp. 266–279. https://doi.org/10.1080/19439962.2015.1030807.

[17] Zheng, L., and K. Ismail. A Generalized Exponential Link Function to Map a Conflict Indicator into Severity Index within Safety Continuum Framework. *Accident Analysis and Prevention*, Vol. 102, 2017, pp. 23–30. https://doi.org/10.1016/j.aap.2017.02.013.

[18] Tarko, A. P. Estimating the Expected Number of Crashes with Traffic Conflicts and the Lomax Distribution – A Theoretical and Numerical Exploration. *Accident Analysis and Prevention*, Vol. 113, No. January 2018, pp. 63–73. https://doi.org/10.1016/j.aap.2018.01.008.

[19] Zheng, L., and T. Sayed. Comparison of Traffic Conflict Indicators for Crash Estimation Using Peak Over Threshold Approach. *Transportation Research Record*, Vol. 2673, No. 5, 2019, pp. 493–502. https://doi.org/10.1177/0361198119841556.

[20] Wang, C., C. Xu, and Y. Dai. A Crash Prediction Method Based on Bivariate Extreme Value Theory and Video-Based Vehicle Trajectory Data. *Accident Analysis and Prevention*, Vol. 123, No. October 2018, 2019, pp. 365–373. https://doi.org/10.1016/j.aap.2018.12.013.

[21] Zheng, L., and T. Sayed. A Bivariate Bayesian Hierarchical Extreme Value Model for Traffic Conflict-Based Crash Estimation. *Analytic Methods in Accident Research*, Vol. 25, 2020, p. 100111. https://doi.org/10.1016/j.amar.2020.100111.

[22] Perkins, S. R., and J. I. Harris. *Traffic Conflict Characteristics-Accident Potential at Intersections*.

[23] Wang, C., Y. Xie, H. Huang, and P. Liu. A Review of Surrogate Safety Measures and Their Applications in Connected and Automated Vehicles Safety Modeling. *Accident Analysis and Prevention*, Vol. 157, 2021. https://doi.org/10.1016/j.aap.2021.106157.

[24] Arun, A., M. M. Haque, S. Washington, T. Sayed, and F. Mannering. A Systematic Review of Traffic Conflict-Based Safety Measures with a Focus on Application Context. *Analytic Methods in Accident Research*, Vol. 32, 2021. https://doi.org/10.1016/j.amar.2021.100185.

[25] Jiang, R., S. Zhu, P. Wang, Q. C. Chen, H. Zou, S. Kuang, and Z. Cheng. In Search of the Consequence Severity of Traffic Conflict. *Journal of Advanced Transportation*, Vol. 2020, 2020. https://doi.org/10.1155/2020/9089817.

[26] Lu, J., O. Grembek, and M. Hansen. Learning the Representation of Surrogate Safety Measures to Identify Traffic Conflict. *Accident Analysis and Prevention*, Vol. 174, 2022. https://doi.org/10.1016/j.aap.2022.106755.

[27] Johnsson, C., A. Laureshyn, and T. De Ceunynck. In Search of Surrogate Safety Indicators for Vulnerable Road Users: A Review of Surrogate Safety Indicators. *Transport Reviews*, Vol. 38, No. 6, 2018, pp. 765–785. https://doi.org/10.1080/01441647.2018.1442888.

[28] Zheng, L., K. Ismail, and X. Meng. Traffic Conflict Techniques for Road Safety Analysis: Open Questions and Some Insights. *Canadian Journal of Civil Engineering*, Vol. 41, No. 7, 2014, pp. 633–641. https://doi.org/10.1139/cjce-2013-0558.

[29] Hayward, J. C. *NEAR-MISS DETERMINATION THROUGH USE OF A SCALE OF DANGER*.

[30] Ozbay, K., H. Yang, B. Bartin, and S. Mudigonda. Derivation and Validation of New Simulation-Based Surrogate Safety Measure. *Transportation Research Record*, No. 2083, 2008, pp. 105–113. https://doi.org/10.3141/2083-12.

[31] Zheng, L., and T. Sayed. Comparison of Traffic Conflict Indicators for Crash Estimation Using Peak Over Threshold Approach. *Transportation Research Record*, Vol. 2673, No. 5, 2019, pp. 493–502. https://doi.org/10.1177/0361198119841556.

[32] Laureshyn, A., T. De Ceunynck, C. Karlsson, Å. Svensson, and S. Daniels. In Search of the Severity Dimension of Traffic Events: Extended Delta-V as a Traffic Conflict Indicator. *Accident Analysis and Prevention*, Vol. 98, 2017, pp. 46–56. https://doi.org/10.1016/j.aap.2016.09.026.

[33] Essa, M., and T. Sayed. Simulated Traffic Conflicts: Do They Accurately Represent Field-Measured Conflicts? *Transportation Research Record*, Vol. 2514, 2015, pp. 48–57. https://doi.org/10.3141/2514-06.

[34] Sayed, T., G. Brown, and F. Navin. *SIMULATION OF TRAFFIC CONFLICTS AT UNSIGNALIZED INTERSECTIONS WITH TSC-Sim*. 1994.

[35] Sacchi, E., and T. Sayed. Conflict-Based Safety Performance Functions for Predicting Traffic Collisions by Type. *Transportation Research Record*, Vol. 2583, 2016, pp. 50–55. https://doi.org/10.3141/2583-07.

[36] Rahman, M. S., M. Abdel-Aty, J. Lee, and M. H. Rahman. Safety Benefits of Arterials' Crash Risk under Connected and Automated Vehicles. *Transportation Research Part C: Emerging Technologies*, Vol. 100, 2019, pp. 354–371. https://doi.org/10.1016/j.trc.2019.01.029.

[37] Wang, C., and N. Stamatiadis. Derivation of a New Surrogate Measure of Crash Severity. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2432, No. 1, 2014, pp. 37–45. https://doi.org/10.3141/2432-05.

[38] Zhang, S., M. Abdel-Aty, Y. Wu, and O. Zheng. Modeling Pedestrians' near-Accident Events at Signalized Intersections Using Gated Recurrent Unit (GRU). *Accident Analysis and Prevention*, Vol. 148, 2020. https://doi.org/10.1016/j.aap.2020.105844.

[39] Ismail, K., T. Sayed, and N. Saunier. Methodologies for Aggregating Indicators of Traffic Conflict. *Transportation Research Record*, No. 2237, 2011, pp. 10–19. https://doi.org/10.3141/2237-02.

[40] Mohamed, M., and N. Saunier. Motion Prediction Methods for Surrogate Safety Analysis. *Transportation Research Record*, No. 2386, 2013, pp. 168–178. https://doi.org/10.3141/2386-19.

[41] Mohamed, M. G., and N. Saunier. Behavior Analysis Using a Multilevel Motion Pattern Learning Framework. *Transportation Research Record*, Vol. 2528, 2015, pp. 116–127. https://doi.org/10.3141/2528-13.

[42] Wang, C., and N. Stamatiadis. Surrogate Safety Measure for Simulation-Based Conflict Study. *Transportation Research Record*, No. 2386, 2013, pp. 72–80. https://doi.org/10.3141/2386-09.

[43] Sobhani, A., W. Young, and M. Sarvi. A Simulation Based Approach to Assess the Safety Performance of Road Locations. *Transportation Research Part C: Emerging Technologies*, Vol. 32, 2013, pp. 144–158. https://doi.org/10.1016/j.trc.2012.10.001.

[44] Moher, D., A. Liberati, J. Tetzlaff, and D. G. Altman. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *International Journal of Surgery*, Vol. 8, No. 5, 2010, pp. 336–341. https://doi.org/10.1016/j.ijsu.2010.02.007.

[45] Methley, A. M., S. Campbell, C. Chew-Graham, R. McNally, and S. Cheraghi-Sohi. PICO, PICOS and SPIDER: A Comparison Study of Specificity and Sensitivity in Three Search Tools for Qualitative Systematic Reviews. *BMC Health Services Research*. 1. Volume 14.

[46] Renard, A., L. Novačko, K. Babojelić, and N. Kožul. Analysis of Child Traffic Safety near Primary School Areas Using UAV Technology. *Sustainability (Switzerland)*, Vol. 14, No. 3, 2022. https://doi.org/10.3390/su14031144.

[47] Marisamynathan, S., and P. Vedagiri. Pedestrian Safety Evaluation of Signalized Intersections Using Surrogate Safety Measures. *Transport*, Vol. 35, No. 1, 2020, pp. 48–56. https://doi.org/10.3846/transport.2020.12157.

[48] *Traffic Conflict Techniques for Safety and Operations-Observers Manual*. 1989.

[49] Salman, N. K., K. J. Al-Maita, and N. K. Salman. *Transportation Research Record No. 1485, Human Performance and Safety in Highway, Traffic, and ITS Systems*.

[50] van der, A. *Manual Conflict Observation Technique DOCTOR Dutch Objective Conflict Technique for Operation and Research Foundation Road Safety for All, The Netherlands*. 2013.

[51] Arun, A., M. M. Haque, S. Washington, T. Sayed, and F. Mannering. A Systematic Review of Traffic Conflict-Based Safety Measures with a Focus on Application Context. *Analytic Methods in Accident Research*, Vol. 32, 2021. https://doi.org/10.1016/j.amar.2021.100185.

[52] Li, S., Q. Xiang, Y. Ma, X. Gu, and H. Li. Crash Risk Prediction Modeling Based on the Traffic Conflict Technique and a Microscopic Simulation for Freeway Interchange Merging Areas. *International Journal of Environmental Research and Public Health*, Vol. 13, No. 11, 2016. https://doi.org/10.3390/ijerph13111157.

[53] Zheng, L., and T. Sayed. Comparison of Traffic Conflict Indicators for Crash Estimation Using Peak Over Threshold Approach. *Transportation Research Record*, Vol. 2673, No. 5, 2019, pp. 493–502. https://doi.org/10.1177/0361198119841556.

[54] Johnsson, C., A. Laureshyn, and T. De Ceunynck. In Search of Surrogate Safety Indicators for Vulnerable Road Users: A Review of Surrogate Safety Indicators. *Transport Reviews*, Vol. 38, No. 6, 2018, pp. 765–785. https://doi.org/10.1080/01441647.2018.1442888.

[55] Guido, G., A. Vitale, V. Astarita, F. Saccomanno, V. P. Giofré, and V. Gallelli. Estimation of Safety Performance Measures from Smartphone Sensors. *Procedia - Social and Behavioral Sciences*, Vol. 54, 2012, pp. 1095–1103. https://doi.org/10.1016/j.sbspro.2012.09.824.

[56] El-Basyouny, K., and T. Sayed. Safety Performance Functions Using Traffic Conflicts. *Safety Science*, Vol. 51, No. 1, 2013, pp. 160–164. https://doi.org/10.1016/j.ssci.2012.04.015.

[57] Tarko, A. P. A Unifying View on Traffic Conflicts and Their Connection with Crashes. *Accident Analysis and Prevention*, Vol. 158, 2021. https://doi.org/10.1016/j.aap.2021.106187.

[58] Zheng, L., and T. Sayed. A Bivariate Bayesian Hierarchical Extreme Value Model for Traffic Conflict-Based Crash Estimation. *Analytic Methods in Accident Research*, Vol. 25, 2020. https://doi.org/10.1016/j.amar.2020.100111.

[59] Glauz, W. D., K. M. Bauer, and D. J. Migletz. *Expected Traffic Conflict Rates and Their Use in Predicting Accidents*.

[60] Ismail, K., T. Sayed, N. Saunier, and C. Lim. Automated Analysis of Pedestrian-Vehicle Conflicts Using Video Data. *Transportation Research Record*, No. 2140, 2009, pp. 44–54. https://doi.org/10.3141/2140-05.

[61] Karim Aldin Ismail, by. *APPLICATION OF COMPUTER VISION TECHNIQUES FOR AUTOMATED ROAD SAFETY ANALYSIS AND TRAFFIC DATA COLLECTION*. 2010.

[62] Essa, M., and T. Sayed. Simulated Traffic Conflicts: Do They Accurately Represent Field-Measured Conflicts? *Transportation Research Record*, Vol. 2514, 2015, pp. 48–57. https://doi.org/10.3141/2514-06.

[63] Hou, J., G. F. List, and X. Guo. New Algorithms for Computing the Time-to-Collision in Freeway Traffic Simulation Models. *Computational Intelligence and Neuroscience*, Vol. 2014, 2014. https://doi.org/10.1155/2014/761047.

[64] Cavallaro, A., O. Steiger, and T. Ebrahimi. Tracking Video Objects in Cluttered Background. *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 15, No. 4, 2005, pp. 575–584. https://doi.org/10.1109/TCSVT.2005.844447.

[65] Dahlkamp, H., A. E. C. Pece, A. Ottlik, and H.-H. Nagel. *Differential Analysis of Two Model-Based Vehicle Tracking Approaches*.

[66] Magee, D. R. Tracking Multiple Vehicles Using Foreground, Background and Motion Models. No. 22, 2004, pp. 143–155.

[67] Maurin, B., O. Masoud, and N. P. Papanikolopoulos. Tracking All Traffic: Computer Vision Algorithms for Monitoring Vehicles Individuals, and Crowds. *IEEE Robotics and Automation Magazine*, Vol. 12, No. 1, 2005, pp. 29–36. https://doi.org/10.1109/MRA.2005.1411416.

[68] Stauffer, C., and W. E. L. Grimson. Learning Patterns of Activity Using Real-Time Tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 8, 2000, pp. 747–757. https://doi.org/10.1109/34.868677.

[69] Veeraraghavan, H., O. Masoud, and N. P. Papanikolopoulos. Computer Vision Algorithms for Intersection Monitoring. *IEEE Transactions on Intelligent Transportation Systems*, Vol. 4, No. 2, 2003, pp. 78–89. https://doi.org/10.1109/TITS.2003.821212.

[70] Koler, D., J. Weber, and J. Malik. *UC Berkeley Working Papers Title Robust Multiple Car Tracking With Occlusion Reasoning Permalink Https://Escholarship.Org/Uc/Item/49c0g7p8 Publication Date*.

[71] Saunier, N., and T. Sayed. A Feature-Based Tracking Algorithm for Vehicles in Intersections. No. 2006, 2006, pp. 59.

[72] Beymer, D., P. Mclauchlan, B. Coifman, and J. Malik. *A Real-Time Computer Vision System for Measuring Traffic Parameters **.

[73] Gazis, D. C., and L. C. Edie. Traffic Flow Theory. *Proceedings of the IEEE*, Vol. 56, No. 4, 1968, pp. 458–471. https://doi.org/10.1109/PROC.1968.6336.

[74] Stauffer, C., and W. E. L. Grimson. Adaptive Background Mixture Models for Real-Time Tracking. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol. 2, 1999, pp. 246–252. https://doi.org/10.1109/cvpr.1999.784637.

[75] Abdel-Aty, M., Y. Wu, O. Zheng, and J. Yuan. Using Closed-Circuit Television Cameras to Analyze Traffic Safety at Intersections Based on Vehicle Key Points Detection. *Accident Analysis and Prevention*, Vol. 176, 2022. https://doi.org/10.1016/j.aap.2022.106794.

[76] Cai, Q., M. Abdel-Aty, J. Lee, L. Wang, and X. Wang. Developing a Grouped Random Parameters Multivariate Spatial Model to Explore Zonal Effects for Segment and Intersection Crash Modeling. *Analytic Methods in Accident Research*, Vol. 19, 2018, pp. 1–15. https://doi.org/10.1016/j.amar.2018.05.001.

[77] Anisha, A. M., M. Abdel-Aty, A. Abdelraouf, Z. Islam, and O. Zheng. Automated Vehicle to Vehicle Conflict Analysis at Signalized Intersections by Camera and LiDAR Sensor Fusion. *Transportation Research Record: Journal of the Transportation Research Board*, 2022, p. 036119812211288. https://doi.org/10.1177/03611981221128806.

[78] Redmon, J., S. Divvala, R. Girshick, and A. Farhadi. *You Only Look Once: Unified, Real-Time Object Detection*.

[79] Liu, W., D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg. SSD: Single Shot MultiBox Detector. 2015. https://doi.org/10.1007/978-3-319-46448-0_2.

[80] Ren, S., K. He, R. Girshick, and J. Sun. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. 2015.

[81] He, K., G. Gkioxari, P. Dollar, and R. Girshick. Mask R-CNN. No. 2017-October 2017, pp. 2980–2988.

[82] Agarwal, S., J. Ogier Du Terrail, F. Jurie, and J. Ogier du Terrail. *Recent Advances in Object Detection in the Age of Deep Convolutional Neural Networks*. 2019.

[83] Jiao, L., F. Zhang, F. Liu, S. Yang, L. Li, Z. Feng, and R. Qu. A Survey of Deep Learning-Based Object Detection. *IEEE Access*, Vol. 7, 2019, pp. 128837–128868. https://doi.org/10.1109/ACCESS.2019.2939201.

[84] Redmon, J., and A. Farhadi. *YOLO9000: Better, Faster, Stronger*.

[85] Redmon, J., and A. Farhadi. YOLOv3: An Incremental Improvement. 2018.

[86] Bochkovskiy, A., C.-Y. Wang, and H.-Y. M. Liao. YOLOv4: Optimal Speed and Accuracy of Object Detection. 2020.

[87] Wang, C.-Y., H.-Y. M. Liao, I.-H. Yeh, Y.-H. Wu, P.-Y. Chen, and J.-W. Hsieh. CSPNet: A New Backbone That Can Enhance Learning Capability of CNN. 2019.

[88] Horvat, M., L. Jelečević, and G. Gledec. *A Comparative Study of YOLOv5 Models Performance for Image Localization and Classification Hascheck-Croatian Academic Spelling Checker View Project A Comparative Study of YOLOv5 Models Performance for Image Localization and Classification*. 2022.

[89] Wang, C.-Y., A. Bochkovskiy, and H.-Y. M. Liao. YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors. 2022.

[90] Zhou, X., D. Wang, and P. Krähenbühl. Objects as Points. 2019.

[91] *Kreiss, S., L. Bertoni, and A. Alahi. PifPaf: Composite Fields for Human Pose Estimation.*

[92] *Kreiss, S., L. Bertoni, and A. Alahi. OpenPifPaf: Composite Fields for Semantic Keypoint Detection and Spatio-Temporal Association. 2021.*

[93] *Fu, H., J. Gao, and H. Liu. Human Pose Estimation and Action Recognition for Fitness Movements. Computers and Graphics, 2023. https://doi.org/10.1016/j.cag.2023.09.008.*

[94] *Traffic Conflict Techniques for Safety and Operations-Observers Manual*. 1989.

[95] Trentacoste Director, M. F. *NOTICE QUALITY ASSURANCE STATEMENT*.

[96] Sayed, T., and S. Zein. Traffic Conflict Standards for Intersections. *Transportation Planning and Technology*, Vol. 22, No. 4, 1999, pp. 309–323. https://doi.org/10.1080/03081069908717634.

[97] Guo, Y., T. Sayed, and M. H. Zaki. Exploring Evasive Action–Based Indicators for PTW Conflicts in Shared Traffic Facility Environments. *Journal of Transportation Engineering, Part A: Systems*, Vol. 144, No. 11, 2018. https://doi.org/10.1061/jtepbs.0000190.

[98] Vogel, K. *A Comparison of Headway and Time to Collision as Safety Indicators*. 2003.

[99] Mullakkal-Babu, F. A., M. Wang, H. Farah, B. van Arem, and R. Happee. Comparative Assessment of Safety Indicators for Vehicle Trajectories on Highways. *Transportation Research Record*, Vol. 2659, No. 1, 2017, pp. 127–136. https://doi.org/10.3141/2659-14.

[100] Reza Mamdoohi, A., A. Professor, and M. Fallah Zavareh. *A. R. Mamdoohi et al.: Comparative Analysis of Safety Performance Indicators Based on Inductive Loop Detector Data COMPARATIVE ANALYSIS OF SAFETY PERFORMANCE INDICATORS BASED ON INDUCTIVE LOOP DETECTOR DATA*. 2014.

[101] Guido, G., F. Saccomanno, A. Vitale, V. Astarita, and D. Festa. Comparing Safety Performance Measures Obtained from Video Capture Data. *Journal of Transportation Engineering*, Vol. 137, No. 7, 2011, pp. 481–491. https://doi.org/10.1061/(ASCE)TE.1943-5436.0000230.

[102] Ozbay, K., H. Yang, B. Bartin, and S. Mudigonda. Derivation and Validation of New Simulation-Based Surrogate Safety Measure. *Transportation Research Record*, No. 2083, 2008, pp. 105–113. https://doi.org/10.3141/2083-12.

[103] Co, P. J. *EXPERIENCE WITH TRAFFIC CONFLICTS IN CANADA WITH EMPHASIS ON "POST ENCROACHMENT TIME" TECHNIQUES*.

[104] Gettman, D., and L. Head. *Surrogate Safety Measures from Traffic Simulation Models*.

[105] Coifman, B., and L. Li. A Critical Evaluation of the Next Generation Simulation (NGSIM) Vehicle Trajectory Dataset. *Transportation Research Part B: Methodological*, Vol. 105, 2017, pp. 362–377. https://doi.org/10.1016/j.trb.2017.09.018.

[106] Coifman, B., L. Li, and W. Xiao. Resurrecting the Lost Vehicle Trajectories of Treiterer and Myers with New Insights into a Controversial Hysteresis. *Transportation Research Record*, Vol. 2672, No. 20, 2018, pp. 25–38. https://doi.org/10.1177/0361198118786473.