LUND UNIVERSITY

Microscopic behaviour analysis using video recordings

A perspective on vulnerable road users

Yastremska-Kravchenko, Oksana

2025

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Microscopic behaviour analysis using video recordings

A perspective on vulnerable road users

OKSANA YASTREMSKA-KRAVCHENKO
FACULTY OF ENGINEERING | LUND UNIVERSITY | 2025

Printed by Media-Tryck, Lund 2025 NORDIC SWAN ECOLABEL 3041 0903

"A Smart City is not measured by the number of shiny gadgets, data gathering novelties, or mobility innovations. Your city is smart when children can cycle there autonomously, joyfully and freely.»
– Lab of Thought (LinkedIn)

Yet, how can we achieve this without accurate data and, crucially, without reliable methods to collect and analyse such data to generate actionable insights into the specific needs required to develop people-centred solutions?

This thesis explores the application of video-derived microscopic traffic data to analyse road user behaviour in urban environments. By capturing real-time, micro-level interactions, it demonstrates the potential of video data to enhance traffic safety analysis, optimise operational performance, and inform infrastructure planning. The research underscores the importance of adopting proactive, evidence-based approaches to tackle traffic safety challenges and improve transport systems.

LUND UNIVERSITY

Faculty of Engineering
Department of Technology and Society

ISBN 978-91-8104-439-3
ISSN 1653-1930

Microscopic behaviour analysis using video recordings

Microscopic behaviour analysis using video recordings

A perspective on vulnerable road users

Oksana Yastremska-Kravchenko

DOCTORAL DISSERTATION

for the degree of Doctor of Philosophy (PhD) at the Faculty of Engineering at Lund University, to be publicly defended on April 25th, 2025, at 9.00 in auditorium V:A, Department of Technology and Society, Klas Anshelms väg 14, Lund.

Faculty opponent
Azra Habibovic, Scania, Sweden

Organization: Lund University

Document name: Doctoral dissertation

Date of issue: 25 April 2025

Author: Oksana Yastremska-Kravchenko

Sponsoring organizations: Vinnova;

The Swedish Energy Agency; Trafikverket

Title and subtitle: Microscopic behaviour analysis using video recordings: A perspective on vulnerable road users

Abstract: This thesis investigates the application of microscopic traffic data derived from video recordings to analyse the behaviour of road users, specifically focusing on pedestrians and cyclists in urban traffic environments. This research addresses the critical need for more accurate and proactive methods in the fields of traffic safety, transport system operational performance, and level of service assessment. The overall aim is to explore the potential of video technology to capture dynamic, real-time behaviours that are otherwise difficult to observe, facilitating a deeper understanding of road users' micro-level interaction with and responses to various traffic conditions. A significant portion of the thesis is dedicated to traffic safety, using conflict-based safety analysis and exploring the integration of Surrogate Measures of Safety (SMoSs) with the Safe System approach to shift the focus from reducing 'collisions' to eliminating 'severe injuries'.

The thesis includes four research papers, each contributing to the study's overall aims and objectives. The first two papers explore methodologies for investigating the impact of physical environmental features, such as facility surfacing and outdoor lighting, on cyclist and pedestrian behaviour. The results show that the proposed method effectively captures behavioural responses to infrastructure changes, revealing that cyclists and pedestrians exhibit distinct reactions. This finding highlights the importance of treating these groups separately to address their unique needs during urban planning and infrastructure improvement processes.

The final two papers address gaps in existing safety frameworks. One proposes a refined, proactive approach to traffic safety assessment and improvement, while the other examines the severity of non-collision traffic events. By analysing video-recorded traffic situations, the study—also forming the final paper included in this thesis—identifies key objective indicators that align with human perceptions of traffic danger, emphasising the importance of two variables—proximity and potential collision consequences—in severity assessments and suggesting their integration into safety evaluations.

The studies included in this thesis demonstrate that video recordings can not only capture simple variables (e.g. the number of road users) but also provide detailed insights into road user behaviours at a micro level. By offering a clear and continuous view of traffic conditions, video data can significantly enhance the 'intelligence input' into traffic analysis, transforming traffic videos into actionable insights.

Key words: Behaviour, conflict-based analysis, cyclist, interactions, microscopic traffic data, outdoor lighting, pedestrian, Safe System, Surrogate Measure of Safety (SMoS), surfacing, traffic event severity, traffic safety, video analysis

Language: English

ISSN and key title: 1653-1930
Bulletin – Lund University, Faculty of Engineering,
Department of Technology and Society, 337

ISBN: 978-91-8104-439-3 (print)
978-91-8104-440-9 (electronic)

Number of pages: 130

Signature

Date 2025-03-05

Microscopic behaviour analysis using video recordings

A perspective on vulnerable road users

Oksana Yastremska-Kravchenko

LUND UNIVERSITY

Faculty of Engineering

Department of Technolodgy and Society

ISBN 978-91-8104-439-3 (print), 978-91-8104-440-9 (electronic)

ISSN 1653-1930

Printed in Sweden by Media-Tryck, Lund University

Lund 2025

To my husband and our wonderful son,

who stood by me through even the toughest moments of this challenging journey

Acknowledgements

After years of research, countless hours of writing, and more than a few moments of self-doubt, I have finally arrived at this milestone—and what a journey it has been! My PhD experience, filled with challenges and rewards, has pushed me to expand my understanding of the world. Looking back, I can see just how much I have grown, not only as a researcher but also as a person. This journey would not have been possible without the unwavering support, encouragement, and kindness of many people to whom I am deeply grateful.

First and foremost, my deepest gratitude goes to my supervisors, Aliaksei Laureshyn and Carmelo D'Agostino, who have guided me from the very beginning and stood by me through every step of this journey. Thank you for believing in me! Your guidance, expertise, and encouragement have been invaluable, and I feel incredibly fortunate to have had the opportunity to learn from you. Beyond your professional mentorship, your support during difficult personal times has meant the world to me. During the challenging period of the russian-Ukrainian war, you are making sure I am never feeling alone, and for that, I am forever grateful. I also want to extend my thanks to András Várhelyi, who supervised me in the early years of my PhD, providing invaluable guidance and shaping my approach to research. My gratitude goes to Carl Johnsson and Kevin Gildea, who stepped in as my supervisors after András and broadened my perspective on the topic while equipping me with practical skills that made this process smoother. My work with you, Carl, began during your PhD journey, when you first introduced me to the fascinating world of road users' trajectories and shared invaluable tips on how to approach them. Kevin, you played a crucial role in speeding up my drone data processing while I was on maternity leave, making my return to work smooth and productive. I am incredibly lucky to have had such a fantastic team of supervisors, and I look forward to maintaining our connection and friendship.

I am also deeply thankful to my colleagues from the '3rd Eye' collaborative team, with whom I had the pleasure of working at the beginning of my PhD. Those were exciting times of experimenting with new camera prototypes, and I am especially grateful to Mikael Nilsson for teaching me how to run Binary Logistic Regression (and survive the process) and Martin Ahrnbom for enlightening me on the mysterious world of computer vision. The latter part of my PhD was spent working on the 'Belysa' project, which brought a new set of challenges and unforgettable experiences, like winter drone flights after dark—because who does not love flying drones in the freezing cold, right? Johan Rahm, thank you for joining during data collection, those long hours under daylight and electrical lighting, for organising and transporting participants to and from the site, and for the insightful discussions on data analysis. Many thanks as well to Anna Niska and Maria Johansson for your deep and intelligent discussions, and for your invaluable contributions as co-authors of my research papers.

I would also like to express my heartfelt appreciation to my colleagues at Transport and Roads, where I have spent a significant portion of my life. I have had the pleasure of working with such amazing people, and I have never felt alone in this journey, thanks to all of you! Whether through in-depth discussions or small but meaningful conversations in the fika room (yes, the fika room holds special powers), you have made this experience all the more special. If I am ever asked to name the best team I have worked with, without hesitation, I will say Transport and Roads. This is a team where support is offered unconditionally—whether in work organisation, administration, or simply through much-needed encouragement during those tough days. Thank you, Astrid, Petra, Helena, Thomas, and Anders, for your immense contributions to making everything run smoothly.

To my fellow PhD student team—Matilda, Zhankun, Michelle, Frida, Grace, Maya, Jakob, Russell, Gustav, Vendela, Annika, and many others who have already earned their degrees before me—Azreena, Ruben, Pajtim, Carl-William, Joel, Maya, Emma, Kah Yong, Tim, Alfred, Ulrik, Erik (I hope I did not forget to mention anyone else—if I did, please consider yourself included with a big thank you!)—you have all made this journey so much brighter. I have learned from each of you and appreciate your friendship and support. To the 'seniors'—Lena, Helena, Lia, Chunli, Fredrik, Vanessa, Hanna, Lars, Sadegh, Joacim, Till, Ebrahim, Andreas, Karin, Hampus, Filbert, and many more—your knowledge and passion have been truly inspiring. I am pretty sure each one of you could teach a PhD course on how to be absolutely amazing!

To my dear friends and fellow colleagues—Janet, Khashayar, and Jean—I am so grateful for our discussions, collaborations, and moments of laughter, which continue even after we are no longer on the same team. And Daria, you were my friend long before you became my colleague. We have supported each other through every challenge, and despite the geographical distance now, it never feels like we are truly apart. Our friendship remains as strong as ever.

I would also like to acknowledge my many colleagues with whom I have had the pleasure of working in the past. While I cannot name you all individually—there are simply too many—you have each played a vital role in this journey. Thank you for your invaluable contributions.

I must also extend my endless gratitude to the entire Technology and Society department. Your support during one of the darkest periods in Ukraine's history, starting in 2022, has been deeply moving. I will never forget your kindness and solidarity.

I would also like to extend an additional thanks to Aliaksei, not only for your excellent supervision but also for introducing me to your wonderful wife, Helene, who has become a dear friend. Helene, your kindness and generosity will always be remembered.

A thank you also goes to Rasa Ušpalytė, my supervisor during my master's studies at Vilnius Gediminas Technical University. You guided me as I shifted my interdisciplinary background towards the field of transportation. You were the first person to believe in my potential to earn a PhD, even before I had made that decision myself. Your support and encouragement, especially in pushing me towards an Erasmus+ internship at Lund University, were turning points in my academic journey.

To my incredible family—my loving mum and dad, my amazing brother and his family, my wonderful mother-in-law, and my cherished grandparents (those who are with us and those who live on in my memories)—your support has been the foundation of my strength. I could not have done this without you.

To my dearest friends from Ukraine—your encouragement and presence in my life have meant so much to me, filling my journey with warmth, joy, and just the right amount of sparkle to keep things exciting. You have truly made this entire adventure much brighter!

Finally, the most special thanks go to the most special people in my life. Vadym, thank you for being my anchor, my strength, and my guiding light. Your unwavering support, your belief in me, and your ability to celebrate even the smallest victories have filled my life with meaning and purpose. And to you, Mark—my heart, my soul, my greatest gift—you arrived in the middle of my PhD journey, bringing with you a love so profound that it reshaped my world. You gave me the space to pause, to reflect, and to see life through new eyes. You have shown me that curiosity is the strongest driving force and that anything is possible if you desire it enough. But above all, you have given me the most precious role of my life—the privilege of being your mother!

Thank you to everyone who has been a part of this remarkable journey—I am forever grateful!

And last but not least, I thank myself—for believing in this dream, for pushing beyond my comfort zone, and for embracing growth, even when it was painful. For every challenge I faced, every moment of doubt I overcame, and every step forward I took despite the obstacles, I acknowledge my own strength and resilience. This journey has not been easy, but I never gave up on myself, and for that, I am truly proud.

'Nothing in life is to be feared, it is only to be understood. Now is the time to understand more, so that we may fear less.'

Marie Curie

Popular science summary

Have you ever wondered why some streets feel more dangerous than others, even if accidents rarely happen there? Is it possible to address hidden risks before they lead to real problems? What if minor details like road surfacing, lighting, or design could significantly affect how people move and interact with traffic? This thesis investigates how we can better understand road users' needs and improve road safety by closely examining small, often unnoticed behavioural responses to traffic environment elements—those invisible to the human eye—at the micro-level in traffic environments using video recording. Analysing these data uncovers patterns that provide deeper insights that can help to make roads safer, inform data-driven, people-centred solutions, and ultimately promote sustainable mobility.

Behavioural studies are commonly used in traffic research. Most of these studies are conducted in real traffic environments by human observers who focus on recording whether specific behaviours occur (e.g. approaching, distractions, eye contact, yielding, and signalling) or how frequently they happen. Video recordings offer an alternative method for conducting behavioural studies, providing several advantages over direct human observation. In addition to tracking the frequency of these behaviours, video recordings enable the collection of highly precise information about how, where, and when these actions take place in traffic. They are also less susceptible to human observer fatigue; they can be saved and reviewed as many times as needed, ensuring more reliable and thorough analysis.

The first part of this thesis provides a comprehensive overview of the entire process of studying traffic behaviour using video recordings, from video data collection to the extraction of accurate microscopic traffic data, which can be used to derive actionable insights.

The second part of the thesis explores new methods for assessing transportation systems by testing changes to infrastructure in real time. For example, such methods could enable municipalities to quickly evaluate whether a new design or safety feature works as expected and make necessary adjustments. Data extracted from video recordings were used to study how physical environmental factors—such as outdoor lighting and path surfacing—affect cyclist and pedestrian behaviour. It was found that these two groups behave in different ways, and small infrastructure interventions, such as improved lighting or smoother surfaces with added segregation, can benefit one group while having a minimal impact on the other. This finding suggests that city planners should treat cyclists and pedestrians as distinct user groups, each with unique needs.

The key focus of the final part of the thesis is conflict-based safety analysis. Rather than waiting for traffic accidents to occur, the thesis examines 'near-misses' and other non-collision interactions between road users. These micro-scale behaviours provide valuable insights into the potential risk of crashes and serve as alternative (surrogate) data that can be used proactively to identify issues before they result in accidents. Additionally, the thesis aims to expand the theoretical application of Surrogate Measures of Safety (SMoSs), focusing on preventing traffic crashes that could lead to serious injuries or fatalities. Research was conducted to develop a measure of the danger (severity) of traffic events. It was identified that, beyond assessing proximity (how close road users come to each other during an interaction), it is crucial to estimate the possible consequences of a collision (the energy of the impact). By incorporating both these variables, traffic engineers and city officials can quickly assess the safety situation at various locations and introduce measures to prevent 'potential risks' leading to 'real outcomes'. This approach enables proactive interventions, preventing unwelcome traffic events (accidents) instead of waiting months or years for them to occur and then analysing limited data afterward.

Overall, this thesis demonstrates that the use of video technology offers far more to traffic research than simply counting how many road users pass through an intersection. It can uncover hidden dynamics in the movements of pedestrians and cyclists—often the most vulnerable road users—ultimately leading to safer, more comfortable, and more efficient urban spaces. By transforming raw footage into actionable, data-driven insights, city planners can make informed adjustments to infrastructure, fine-tuning road systems to better meet the real needs of the people who use them daily.

Populärvetenskaplig sammanfattning

Har du någonsin undrat varför vissa gator känns farligare än andra, även om olyckor sällan inträffar där? Är det möjligt att ta itu med dolda risker innan de leder till verkliga problem? Små detaljer som vägunderlag, belysning eller design kan ha en betydande inverkan på hur människor rör sig och interagerar i trafiken? Denna avhandling undersöker hur vi kan bättre förstå trafikanters behov och förbättra trafiksäkerheten genom att noggrant undersöka små, ofta omärkbara beteendereaktioner på mikronivå i trafikmiljöer med hjälp av videoinspelning. Genom att analysera dessa data avslöjas mönster som ger djupare insikter som kan bidra till att göra vägar säkrare, informera datadrivna, människocentrerade lösningar och slutligen främja hållbar mobilitet.

Beteendestudier är vanliga inom trafikforskning. De flesta av dessa studier genomförs i verkliga trafikmiljöer av mänskliga observatörer som fokuserar på att registrera specifika beteenden (t.ex. närmande, distraktioner, ögonkontakt, väjningsbeteende och signalering) eller hur ofta de inträffar. Videoinspelningar erbjuder en alternativ metod för att genomföra beteendestudier och har flera fördelar jämfört med direkt mänsklig observation. Förutom att registrera hur ofta dessa beteenden förekommer, gör videoinspelningar det möjligt att samla in mycket exakt information om hur, var och när dessa handlingar äger rum i trafiken. De är också mindre känsliga för observatörströtthet, eftersom de kan sparas och granskas så många gånger som behövs, vilket säkerställer en mer tillförlitlig och grundlig analys.

Den första delen av denna avhandling ger en omfattande översikt över hela processen för att studera trafikbeteende med hjälp av videoinspelningar, från insamling av videodata till extrahering av precisa mikroskopiska trafikdata, som kan användas för att generera värdefulla insikter.

Den andra delen av avhandlingen utforskar nya metoder för att utvärdera transportsystem genom att testa förändringar i infrastrukturen i realtid. Till exempel kan sådana metoder göra det möjligt för kommuner att snabbt bedöma om en ny design eller säkerhetsfunktion fungerar som förväntat och göra nödvändiga justeringar. Data som extraherats från videoinspelningar användes för att studera hur fysiska miljöfaktorer—såsom utomhusbelysning och vägunderlag—påverkar cyklisters och fotgängares beteende. Det visades att dessa två grupper beter sig på olika sätt, och små infrastrukturförändringar, som förbättrad belysning eller jämnare ytor med tillagd separation, kan gynna en grupp samtidigt som de har minimal effekt på den andra.

Denna upptäckt antyder att stadsplanerare bör behandla cyklister och fotgängare som distinkta användargrupper, var och en med unika behov.

Huvudfokus i den sista delen av avhandlingen är konfliktbaserad säkerhetsanalys. Istället för att vänta på att trafikolyckor ska inträffa, undersöker avhandlingen "nästan-olyckor" och andra riskabla interaktioner mellan trafikanter. Dessa mikroskopiska beteenden ger värdefulla insikter om potentiell risk för kollisioner och fungerar som alternativa (surrogat) data som proaktivt kan användas för att identifiera problem innan de leder till olyckor. Dessutom syftar avhandlingen till att utveckla den teoretiska tillämpningen av Surrogate Measures of Safety (SMoSs), med fokus på att förhindra trafikolyckor som kan leda till allvarliga skador eller dödsfall. Forskning genomfördes för att utveckla ett mått på faran (allvaret) vid tillbud. Det identifierades att, bortom att bedöma närhet (hur nära trafikanterna kommer varandra vid en interaktion), är det avgörande att uppskatta de möjliga konsekvenserna av en kollision (kollisionens energi). Genom att kombinera dessa två variabler kan trafikingenjörer och stadsplanerare snabbt bedöma säkerhetssituationen på olika platser och vidta åtgärder för att förhindra att 'potentiella risker' leder till 'verkliga utfall'. Detta tillvägagångssätt möjliggör proaktiva insatser och förebygger oönskade trafikincidenter (olyckor), istället för att vänta månader eller år på att de inträffar och sedan analysera begränsad data i efterhand.

Sammanfattningsvis visar denna avhandling att användningen av videoteknik erbjuder mycket mer för trafikforskning än att bara räkna hur många trafikanter som passerar en vägkorsning. Den kan avslöja dolda dynamiker i fotgängares och cyklisters rörelser—ofta de mest sårbara trafikanterna—vilket i slutändan leder till säkrare, mer bekväma och effektivare stadsområden. Genom att omvandla rå videomaterial till handlingsbara, datadrivna insikter kan stadsplanerare göra informerade justeringar i infrastrukturen och finjustera vägsystemen för att bättre möta de verkliga behoven hos de människor som använder dem varje dag

Abstract

This thesis investigates the application of microscopic traffic data derived from video recordings to analyse the behaviour of road users, specifically focusing on pedestrians and cyclists in urban traffic environments. This research addresses the critical need for more accurate and proactive methods in the fields of traffic safety, transport system operational performance, and level of service assessment. The overall aim is to explore the potential of video technology to capture dynamic, real-time behaviours that are otherwise difficult to observe, facilitating a deeper understanding of road users' micro-level interaction with and responses to various traffic conditions. A significant portion of the thesis is dedicated to traffic safety, using conflict-based safety analysis and exploring the integration of Surrogate Measures of Safety (SMoSs) with the Safe System approach to shift the focus from reducing 'collisions' to eliminating 'severe injuries'.

The thesis includes four research papers, each contributing to the study's overall aims and objectives. The first two papers explore methodologies for investigating the impact of physical environmental features, such as facility surfacing and outdoor lighting, on cyclist and pedestrian behaviour. The results show that the proposed method effectively captures behavioural responses to infrastructure changes, revealing that cyclists and pedestrians exhibit distinct reactions. This finding highlights the importance of treating these groups separately to address their unique needs during urban planning and infrastructure improvement processes.

The final two papers address gaps in existing safety frameworks. One proposes a refined, proactive approach to traffic safety assessment and improvement, while the other examines the severity of non-collision traffic events. By analysing video-recorded traffic situations, the study—also forming the final paper included in this thesis—identifies key objective indicators that align with human perceptions of traffic danger, emphasising the importance of two variables—proximity and potential collision consequences—in severity assessments and suggesting their integration into safety evaluations.

The studies included in this thesis demonstrate that video recordings can not only capture simple variables (e.g. the number of road users) but also provide detailed insights into road user behaviours at a micro level. By offering a clear and continuous view of traffic conditions, video data can significantly enhance the 'intelligence input' into traffic analysis, transforming traffic videos into actionable insights.

Glossary of terms

Accident / crash	The definitions of ‘accident’ and ‘crash’ are used interchangeably in this thesis, both referring to traffic incidents that involve a collision.
Accident (crash) severity	The level of impact (seriousness) of a traffic accident (crash), assessed using specific criteria or measures.
Collision course	In this thesis it refers to a situation in which road users are on a path that will result in a collision if they maintain their current speeds and intended paths.
Critical events / conflicts / near-misses	These definitions are used interchangeably in this thesis, all referring to a traffic encounter with a non-zero probability of resulting in a crash.
Encounter / meeting	The definitions of ‘encounter’ and ‘meeting’ are used interchangeably in this thesis, both referring to the simultaneous presence of two road users within a defined area.
Meeting point	In this thesis, this refers to the same point (longitudinally) on the traffic facility (shared-use path) during bi-directional interactions between cyclists.
Microscopic traffic data	Detailed traffic data that focus on the behaviour of individual road users, capturing specific, small-scale movements and interactions within the traffic flow, refers in this thesis to data derived from road users' trajectories.
Safe System approach	A holistic approach to road safety that seeks to eliminate the risk of serious injury or fatality, emphasizing system-wide safety measures, including road design, vehicle safety, road user behaviour, and post-crash care.
Surrogate Measures of Safety (SMoSs)	Objective and measurable parameters used in traffic safety analysis, derived from observable non-collision events, serving as an alternative to traditional crash-based data.
Traffic event severity	In this thesis, it refers to an operational parameter that defines the 'seriousness' of a traffic encounter.
Vision Zero	A road safety strategy aimed at eliminating all traffic fatalities and serious injuries by taking a proactive, system-wide approach that prioritises safe road design, responsible road use, and effective policies.
Vulnerable road users (VRUs)	Road users who are at higher risk of injury in traffic (e.g. pedestrians, cyclists, motorcyclists) due to their lack of physical protection compared to vehicle occupants.

Table of Contents

1 Introduction

In the realm of traffic dynamics, the minute behavioural patterns and interactions of road users at a microscopic level hold a wealth of information. Even the most diligent on-site traffic observers, watching real traffic and trying to catch fleeting moments of these precise details, can miss critical moments in the blink of an eye. Additionally, capturing the precise positions of road users, such as their world coordinates (X, Y, Z), is almost impossible without knowing exactly when the behavioural response occurs. There may be moments (before, during, or after specific developments in a traffic event) that provide important insights. This is where video recordings—representing visual data captured through continuous, time-stamped footage—become invaluable.

Video recordings offer several key advantages for studying microscopic behaviours, which are understood as the tiny, often rapid movements/actions and interactions of road users that occur at a scale too subtle to be easily observed without specialised equipment. Modern high-resolution equipment captures traffic scenes with sufficient pixel density to extract precise motion details while enabling the study of dynamic processes over time. This facilitates a deeper understanding of how behaviours evolve and interactions occur. Video recordings also enable the analysis of external factors such as infrastructure changes or social issues like pandemics. They also create a permanent record that can be reviewed multiple times, ensuring accurate, reproducible results, which is crucial for verification. Advanced video analysis software further strengthens this process, quantifying aspects of microscopic behaviour, such as movement patterns, speed, interaction frequency, and structural changes. Data derived from video recordings can encompass a wide range of information, from monitoring the frequency of specific behaviours to collecting continuous positioning data that are linked to both time and space. Furthermore, video recordings allow for naturalistic observations, minimising interference and ensuring the accurate capture of natural behaviours.

Transforming raw video recordings into actionable traffic insights represents a significant technological advancement, which involves two critical domains: video processing and video analysis. Video processing encompasses the technical aspects of capturing, storing, and preparing video data for analysis. Video analysis focuses on interpreting these data to extract meaningful insights into traffic behaviour. Despite the growing popularity of advanced traffic video analysis techniques, these two domains remain largely separate (Abdel-Aty et al., 2023).

Beyond enhancing the integration of video processing and video analysis, a notable gap exists in the effective translation of research findings into practical applications for practitioners. Specifically, there is a lack of user-friendly tools that transport specialists can easily apply in the field are lacking, particularly those that provide real-time feedback. Such real-time feedback mechanisms could significantly enhance transport system efficiency and service levels by offering municipalities immediate insights into the effectiveness of infrastructure interventions (e.g. modifications to infrastructure). By addressing the critical question, 'Does it work as intended?', these tools would allow cities to quickly assess the impact of their interventions, make necessary adjustments, and optimise transportation systems in a timely manner to better meet user needs while maintaining high levels of service. The development of real-time feedback tools for transport specialists could also reduce the financial costs and human risks associated with traditional trial-and-error methods.

These challenges highlight the need for better integration of research and practice to advance the understanding of subtle and dynamic road user behaviours. Video analysis plays a crucial role in bridging this gap by providing actionable insights that enhance transportation performance and, ultimately, contribute to sustainability. Bridging this gap is essential for advancing the field and forms the foundation of this thesis, which focuses on vulnerable road users (VRUs), particularly cyclists and pedestrians. These road users are more susceptible to injury or fatality in traffic accidents due to their lack of physical protection compared to motor vehicle occupants. The increasing emphasis on VRUs is driven by this greater risk, especially in crashes involving motor vehicles. Furthermore, as cities encourage walking and cycling for sustainability and health, there is a growing need to integrate safe and functional infrastructure for these modes of transport. Behavioural observations of VRUs—especially their interactions with different road features and transport modes—can enhance our understanding of their unique needs, identify specific risks, and inform targeted safety measures.

From a traffic safety perspective, there is also a significant need for improved methods and tools for quickly assessing the effectiveness of various traffic safety measures before implementation. Such a proactive approach would enable the early identification and mitigation of potential safety risks before road users are impacted. Traditional methods for assessing traffic safety focus on analysing past accident data and have significant limitations, including underreporting and a lack of detailed data, which can distort safety analyses. Furthermore, at specific intersections or road segments, the low frequency of accidents makes it difficult to conduct reliable statistical analyses. Video recordings, however, offer valuable microscopic data, thereby enabling a shift toward proactive approaches, such as the use of Surrogate Measures of Safety (SMoSs). SMoSs serve as an alternative to traditional accident-based data and allow traffic safety professionals to gain data-driven insights into traffic behaviour. SMoSs focus on observable non-crash events, such as traffic conflicts, near-misses, or almost-accidents, to identify potential safety risks before they result in actual collisions.

Despite their potential and conception more than six decades ago, SMoSs have not yet become a practical tool for traffic safety professionals. This is largely due to inconsistencies in the underlying theory and the quality of data, which have limited their overall effectiveness. To enhance the applicability of SMoSs, it is crucial to refine the underlying theory and ensure that the data used are accurate and reliable. In this context, the Safe System vision has emerged as a comprehensive, proactive traffic safety framework that prioritises reducing serious injuries—defined as severe physical harm from traffic incidents, often requiring medical intervention and potentially causing long-term impairment—and fatalities on the roads. This approach acknowledges that humans are prone to making errors and that the human body has biomechanical limits, highlighting the importance of designing transportation systems that account for these factors to reduce the risk of serious injuries.

While the Safe System approach is ethically sound, scientifically robust, and has proven successful in reducing accidents resulting in serious injuries and fatalities, it has not yet been fully integrated into the SMoS framework. This presents a significant opportunity for improvement in the field of traffic safety research and practice.

Addressing these challenges is critical for advancing transportation systems that are both safer and more efficient, helping to meet the demands of modern urban environments. This thesis will explore these issues in greater detail and propose potential solutions.

1.1 Scope and objectives

This thesis examines the critical role of video recordings in the study of microscopic behaviour, highlighting their numerous advantages and diverse applications while also addressing some challenges and limitations that could inform future studies. The overarching goal of this research is to explore ***the application of microscopic data extracted from video recordings to study and analyse VRUs' behaviour***, specifically focusing on pedestrians and cyclists in urban traffic environments.

The research focuses on two key areas:

1. Gaining deeper insights into the subtle and dynamic behaviours of road users in everyday traffic situations.

2. Pushing the frontiers of the SMoSs to promote proactive traffic safety analysis in alignment with the Safe System vision.

This thesis' scope encompasses a comprehensive analysis of video recordings, including the following aspects:

- Investigating the processes and methodologies for extracting meaningful information from video recordings.
- Exploring advanced methods to gain detailed insights into the subtle and dynamic behaviours of road users.
- Emphasising conflict-based methods and the application of SMoSs to better understand and improve traffic safety.

This research will support a diverse range of interests in implementing proactive and effective measures that optimise road user comfort and overall traffic safety.

1.2 Structure of the thesis

The thesis comprises this comprehensive summary (referred to as 'Kappa' in Swedish) and four articles, which are included in the appendix and have either been published or submitted for publication in scientific journals. Following this introductory section, the chapters proceed as follows:

The chapter *'Technology-driven journey: Unpacking the process of extracting microscopic traffic data from video footage'* provides a concise overview of the process of converting raw video recordings into detailed microscopic traffic data, discussing the challenges and potential benefits of using video analysis for traffic behavioural studies.

The chapter *'Understanding behavioural observation studies in traffic research: A perspective on VRUs'* delves into the use of microscopic data extracted from video recordings for the observation and analysis of traffic behaviour. It presents findings from a large-scale behavioural study conducted as part of the research project 'Belysa—miljöfaktorers inverkan på fotgängares och cyklisters beteende i skymning och mörker' (grant number: 2019-002011). This chapter provides the foundation for *Papers I* and *II*, which are included in the appendix.

The chapter *'Microscopic behavioural data for enhancing traffic safety: Application of Surrogate Measures of Safety (SMoSs)'* emphasises conflict-based safety analysis and the application of SMoSs to evaluate and improve traffic safety based on observed behaviours. This chapter underpins *Papers III* and *IV*, which are also included in the appendix.

The penultimate chapter, *'Discussion'*, offers reflections on the research, summarises its contributions, and highlights directions for future studies.

Finally, the thesis concludes with the chapter *'Closing remarks'*, offering a compact summary of the main points discussed.

1.3 List of papers and author contributions

Paper I

Yastremska-Kravchenko, O., A. Laureshyn, J. Rahm, M. Johansson, A. Niska, C. Johnsson, C. D'Agostino (2024) Video analysis of bicyclist and pedestrian movement on shared-use paths under daylight and electric lighting conditions—Method exploration. *Journal of Cycling and Micromobility Research*, 2, 100032. DOI: 10.1016/j.jcmr.2024.100032

Authors' contribution:

The majority of the work was carried out by O. Yastremska-Kravchenko, who led the study design, data collection and analysis, visualization, and the initial manuscript drafting. M. Johansson, A. Laureshyn, J. Rahm, and A. Niska contributed to funding acquisition. A. Laureshyn, C. Johnsson, and C. D'agostino provided supervision throughout the project. C. Johnsson and J. Rahm assisted with data curation. A. Laureshyn, M. Johansson, J. Rahm, and A. Niska also reviewed and edited parts of the manuscript.

Paper II

Yastremska-Kravchenko, O., J. Rahm, M. Johansson, A. Niska, C. Johnsson, A. Laureshyn, C. D'Agostino, K. Gildea. The role of path surfacing and lighting conditions in shaping pedestrian and cyclist behaviour: A comprehensive video analysis (*article under review*)

Authors' contribution:

O. Yastremska-Kravchenko led the majority of the work, including study design, data collection and analysis, visualization, and drafting the original manuscript. M. Johansson, A. Laureshyn, J. Rahm, and A. Niska contributed to funding acquisition. C. Johnsson, A. Laureshyn, C. D'agostino, and K. Gildea provided supervision. C. Johnsson and J. Rahm assisted with data curation. J. Rahm, M. Johansson, and A. Niska reviewed and edited parts of the manuscript.

Paper III

Laureshyn, A., O. Yastremska-Kravchenko, C. Johnsson, Z. Chen, C. D'Agostino, F. Patterson. Re-defining the framework for safety analysis using surrogate measures (*article under review*)

Authors' contribution:
The idea for the article was conceptualized and discussed by all co-authors. Writing and investigation were primarily carried out by A. Laureshyn and O. Yastremska-Kravchenko. Funding acquisition was secured by C. D'Agostino.

Paper IV

Yastremska-Kravchenko, O., A. Laureshyn, J, C. D'Agostino, A. Varhelyi (2022) What constitutes traffic event severity in terms of human danger perception? *Transportation Research Part F: Traffic Psychology and Behaviour*, 90, 22-34. DOI: 10.1016/j.trf.2022.08.001

Authors' contribution:
The majority of the work was done by O. Yastremska-Kravchenko, including conceptualization, methodology, data curation, investigation, formal analysis, visualization, and drafting the original manuscript. The other authors contributed to the supervision, particularly in discussions on conceptualization and methodology. A. Laureshyn and C. D'Agostino also assisted with reviewing and editing parts of the manuscript. Funding acquisition was secured by A. Laureshyn.

2 Technology-driven journey: Unpacking the process of extracting microscopic traffic data from video footage

New sources of traffic data have become available thanks to the extensive development and implementation of various technologies, such as video recording tools (e.g. stationary cameras, drone technology, stereo camera setups) and LiDAR (Light Detection and Ranging). The ability of these technologies to capture rich microscopic data has been attracting increasing interest in the field of traffic behavioural research. This thesis primarily focuses on the analysis of microscopic traffic data—data that focus on the behaviour of individual road users, capturing specific, small-scale movements and interactions within the traffic flow—using trajectories extracted from video recordings with computer vision techniques.

A raw video data stream comprises a sequence of individual still images captured at a specific rate (fps—frames/images per second). Each image consists of numerous pixels—tiny units containing specific colour and brightness information that, when combined across the frame, create the video's visual content. Resolution—determined by the number of pixels that compose each image—is another important characteristic of a video, as it significantly impacts the clarity and detail of the recorded content. Common resolutions include 480p (720 x 480 pixels), 720p (1280 x 720 pixels), 1080p (1920 x 1080 pixels), and 4K (3840 x 2160 pixels), each of which serves distinct purposes and has its own advantages, making it important to select the appropriate one based on the application objective.

After completing raw video processing steps (e.g. calibration, stabilisation, stitching), computer vision techniques can be used to analyse pixel data to detect objects, classify the objects according to predefined categories or characteristics, and link detected objects across consecutive frames to create a movement track—trajectory—for each road user in the scene. However, the workload for completing these tasks is complex, spanning from the initial stages of video collection and processing to advanced traffic analysis (see Figure 1), and is compounded by various practical challenges that may arise throughout the process.

Figure 1. Schematic diagram illustrating the sequential stages of transforming video data into advanced traffic analysis

2.1 Video recording methods

2.1.1 Stationary cameras

Stationary cameras are widely used to collect video data across various contexts and provide reliable and consistent monitoring and recording of activities. There are various types of cameras (sensors) available, including RGB and thermal sensors. When choosing a camera type for data collection, it is especially important to consider factors like compliance with local data protection regulations, weather conditions, and desired resolution (Laureshyn & Várhelyi, 2018).

An example of a stationary camera setup is shown in Figure 2. Because these cameras are fixed in place, they offer several advantages, such as uninterrupted surveillance and a consistent viewpoint, which facilitates tracking changes and patterns over time. However, ensuring the optimal functionality and reliability of stationary cameras involves careful consideration of mounting and power supply issues. Methodical planning and implementation are needed to ensure that these cameras operate efficiently, reliably, and safely in various environments.

Figure 2. Example of stationary cameras setup

The first critical aspect to consider is camera placement. Cameras must be positioned strategically to cover the area of interest while avoiding areas prone to frequent obstructions, such as moving objects or vegetation. Cameras should also be placed in locations that protect them from environmental hazards like rain, snow, wind, and excessive sunlight; this may require the use of weatherproof housing and various coverings that provide shade. Additionally, the mounts must be robust enough to withstand wind and various vibrations to maintain image stability and clarity; this might involve using poles or wall mounts. In some locations, particularly in urban or historic areas, mounting options might be restricted by aesthetic guidelines or local regulations.

Power supply is another critical consideration—cameras need to be installed within a reasonable distance from power sources. This might necessitate running power cables through walls, ceilings, or underground conduits, which can be labour-intensive and costly. In remote areas without readily available power sources, alternative solutions such as battery packs or solar panels may be required. Proper cable management is important to prevent damage and ensure safety, and includes using protective conduits, avoiding exposure to harsh environmental conditions, and preventing interference with other electrical systems.

Once installed, stationary cameras require minimal maintenance compared to movable or mobile cameras. However, their deployment must be carefully managed to balance the benefits with ethical considerations, which are discussed in more detail in *Section 2.2*, ensuring the responsible and respectful use of surveillance technology.

2.1.2 Drone technology

In the realm of traffic observation, unmanned aerial vehicles (UAVs)—commonly known as drone technology—have emerged as highly effective tools for capturing video data (Gohari et al., 2023; Bisio et al., 2022; Outay et al., 2020). This technology presents several advantages over traditional ground-based cameras by providing a unique aerial perspective that reduces the need for a 'fisheye' lens and allows more comprehensive coverage of large areas. An example of a drone camera setup is shown in Figure 3. However, the use of video recording tools, such as drones, introduces several challenges. For example, traditional drone footage typically has a duration of only 20–25 minutes, depending on the battery capacity. In the context of long-term data collection, this limitation results in the capture of large volumes of footage, necessitating individual calibration for each drone clip and further complicating the analytical process. A potential solution to this issue is the use of tethered drones, which are typically connected to a high-voltage DC power source through a thin, lightweight cable, thereby enabling extended flight times and prolonged data collection without the constraints imposed by battery life.

Figure 3. Example of a drone camera setup

The commercial UAV market is rapidly growing, leading to significant technological advancements. Modern drones are equipped with high-resolution cameras that capture clear, detailed video footage suitable for thorough analysis while mitigating privacy concerns by operating at high altitudes to prevent the identification of individuals. In the Swedish context, the national aviation authority—the Swedish Transport Agency (Transportstyrelsen)—permits drone flights. However, operators must ensure they adhere to all applicable rules and regulations to operate drones safely (Transportstyrelsen, 2024).

In terms of performance, current UAVs are designed to be stable, manoeuvrable, and waterproof. Despite these advancements, UAVs still require an offline stabilisation step before calibration and subsequent processing. Stability is necessary for maintaining steady footage, particularly in challenging weather conditions or turbulent environments.

The diversity of drone technologies currently available includes lightweight models weighing under 250 grams. These drones are designed to navigate safely in crowded areas, providing flexibility and accessibility for various applications, including naturalistic traffic behavioural studies in densely populated urban settings.

2.1.3 Mobile cameras

Other technologies for collecting video data for traffic analysis include mobile cameras (e.g. dashcams, vehicle-mounted cameras), which offer a mobile perspective of traffic and serve various purposes (Mehrish et al., 2020; Cao et al., 2011). They are valuable for capturing dynamic aspects of traffic behaviour and refining situational awareness, particularly in areas where fixed surveillance is impractical or insufficient. However, they come with several disadvantages that should be considered.

First, their effectiveness is directly tied to the movement of the vehicle on which they are mounted. Consequently, the quality of video footage from mobile cameras can vary significantly depending on factors such as vehicle speed, road conditions, weather, and lighting. Issues like motion blur and shaking, especially during vehicle acceleration, deceleration, or on rough roads, further impact the clarity of the footage. Additionally, captured footage may unintentionally include private properties, individuals, or sensitive information, which creates privacy concerns. One of the most critical drawbacks of mobile cameras, however, is the difficulty in calibrating and synchronising them with other data collection methods, which is necessary for advanced traffic analysis.

2.1.4 Integrated approach

Multiple cameras can be employed to effectively capture larger areas for detailed analysis. This approach may include utilising various video recording methods, such as integrating fixed cameras with drone technologies. However, achieving optimal results requires thorough planning and careful consideration of the specific requirements and capabilities of each method.

Integrating multiple cameras requires the consideration of several critical aspects, including ensuring comprehensive coverage of the area of interest, capturing different perspectives, aligning each camera's technical specifications (resolution, frame rate, etc.), accounting for environmental conditions, and addressing power supply issues.

Furthermore, additional video processing steps are necessary before advanced video analysis techniques can be used to extract meaningful insights from collected video data. First, synchronising the timestamps or frames from multiple cameras is crucial for accurately correlating events across all of the captured footage, ensuring that they are correctly sequenced and analysed together. Second, it is necessary to use the video stitching process to merge footage from overlapping views of multiple cameras into a cohesive video stream to achieve continuous coverage of the area of interest.

2.1.5 Other methods

Advancements in stereo vision technology are ongoing, with various techniques in development. One notable innovation is the Trinocular Linear Camera Array (TLCA), which consists of three cameras aligned linearly (see Figure 4). This setup offers several advantages over traditional stereo cameras, including more accurate 3D reconstruction and improved occlusion handling—where occlusion refers to the temporary blocking of one road user's trajectory by another object or road user—through the simultaneous rectification of all three images (Ahrnbom et al., 2021a). However, further improvements are necessary before it can be widely used for traffic analysis.

Figure 4. A photo of a Trinocular Linear Camera Array (TLCA) prototype

2.2 Ethical considerations and GDPR compliance in video recording

When they contain information about identifiable individuals, video recordings raise significant concerns under the EU General Data Protection Regulation (GDPR, 2016). In Sweden, this issue is further addressed by the Camera Surveillance Act, which establishes additional regulations governing video surveillance. The Swedish Authority for Privacy Protection (IMY, 2024) plays a key role in enforcing these regulations by ensuring that personal data are managed correctly and do not fall into the wrong hands.

Respecting the rights of individuals is critical for ethical data handling. This begins with obtaining informed consent from individuals identifiable in video recordings, ensuring that they are fully aware of how their data will be used. Privacy measures, including anonymisation and secure storage, safeguard personal data in the recordings. Data collection should be minimised, gathering only the information necessary for the project's purpose, and data retention should be limited to meet regulatory requirements.

In some cases, video recordings may fall outside the GDPR's scope, particularly when the footage is sufficiently anonymised or does not capture identifiable personal data. For example, if faces, unique clothing, or distinctive markings are not visible, or if the footage is too distant or its resolution is too low to discern personal characteristics, the recording may not be considered personal data under the GDPR. Additionally, some parts of a recording can be masked. This is especially relevant in naturalistic observational studies, where the primary aim is to capture behaviour without researcher interference, which informed consent procedures necessitate. However, even in cases where anonymisation or masking is applied, secure data storage remains a critical concern to ensure that privacy is maintained.

Other existing GDPR-compliant methods include real-time processing techniques, such as those provided by services like Viscando. This approach involves not storing image frames during data collection, helping to ensure that personal data are not retained or processed beyond what is necessary for the immediate task at hand. By only processing data in real time and not storing identifiable information, these methods reduce the risk of violating privacy rights while still enabling the accurate extraction of microscopic data.

Thermal cameras offer another valuable technology for capturing traffic activities without revealing identifiable personal characteristics, as they detect heat signatures rather than visual details. However, thermal imaging technology tends to be more costly than traditional video recording methods.

Drone technology presents yet another alternative for observational studies. When at high enough altitudes, drones can capture footage that may include people or objects but from which it is impossible to recognise individuals. This ability largely depends on the type of drone and camera used, and it is also necessary to comply with the official flight height limit of 120 m. By utilising drone technology, researchers can obtain aerial perspectives that facilitate the study of large-scale phenomena and natural environments while minimising direct interference with subjects that could alter their behaviour. However, the effectiveness of this approach relies heavily on factors such as weather conditions, observation timing, and adherence to regulations governing drone use, particularly those concerning proximity to sensitive locations like hospitals, airports, military installations, and other critical infrastructure.

2.3 Video processing

2.3.1 Stabilisation

To ensure the accuracy of the microscopic data extracted from a video, the entire video stream must be in optimal condition, free of camera shakes, vibrations or instability. Stability is also a fundamental factor in object detection and tracking, which are the next steps in extracting road users' trajectories. By maintaining stability, the movements of objects or subjects in the video are faithfully represented without introducing distortions caused by camera motion.

Nowadays, many cameras are equipped with hardware stabilisers, which typically feature advanced sensors and lens systems designed to eliminate camera shake during recording. Additionally, there exist various techniques that use software algorithms to process recorded video and correct any residual slight movements that may occur during the stream. These digital image stabilisers use sensors to detect camera motion and have built-in software that adjusts the captured images by cropping them appropriately, compensating for any shaking motion (Ko et al., 1999). However, these solutions are often not sufficient to address sequences that involve more complex motions. In such cases, offline video stabilisation algorithms are still necessary. Numerous methods have been proposed in the computer vision literature to address this challenge. Most existing video stabilisation techniques involve three primary stages: motion estimation, motion compensation, and image composition (Dong et al., 2014;

Wang et al., 2011; Morimoto & Chellappa, 1998), and various approaches have been proposed to address instability problems (Battiato et al., 2007; Matsushita et al., 2006; Chang et al., 2004; Litvin et al., 2003). However, many of these methods rely on assumptions about motion models during motion estimation. To overcome these limitations, Lee et al. (2009) introduced an approach that uses robust feature trajectories from the video sequence to stabilise the video directly, without explicitly estimating camera motion. Additionally, Pant et al. (2021) study evaluated various video stabilisation algorithms aimed at reducing undesired motion in UAV footage.

2.3.2 Calibration

Camera calibration is necessary for translating 2D (two-dimensional) image points into meaningful 3D information. This process involves determining both intrinsic (the camera's focal length and distortion coefficients) and extrinsic (the camera's position and orientation) parameters to accurately transform the point/object location in the image into the local or world coordinate system. Accurate calibration is vital for video-based traffic analysis because incorrect camera setup can lead to imprecise distance and speed measurements, resulting in trajectories that do not accurately reflect the actual movement of road users.

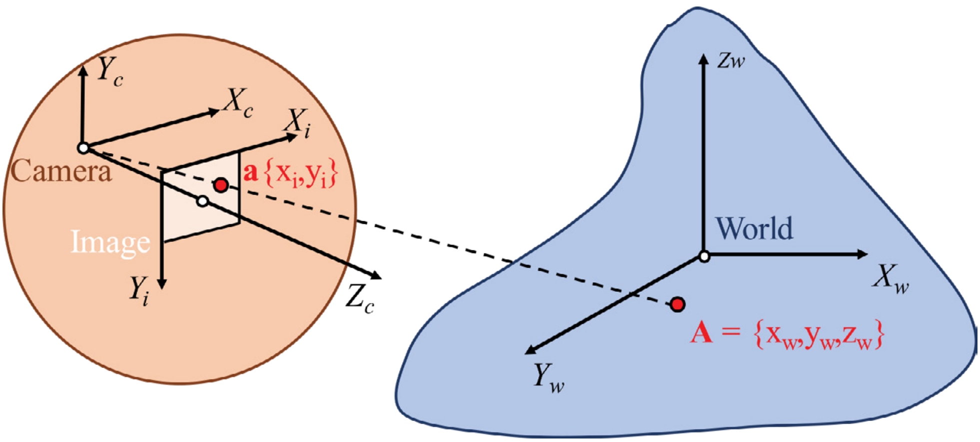

Figure 5. Principle of calibration model (adopted from Laureshyn & Nilsson (2018))

Many camera calibration methods have been proposed in the computer vision and photogrammetry literature (Barone et al., 2020; Hillemann et al., 2019; Abdel-Aziz & Karara, 2015; Gruen & Huang, 2001; Zhang, 2000; Brown, 1996; Tsai, 1986). Each camera calibration model has its strengths and weaknesses, making them suitable for different types of applications. The choice of model depends on the specific task requirements, such as accuracy, computational complexity, and the types of distortions that need to be corrected.

One of the most widely used methods is the Tsai camera calibration model, which has gained recognition for its robustness and high accuracy (Tsai, 1986). It builds upon and significantly refines the basic pinhole camera model to account for real-world imperfections such as lens distortion. The model identifies and extracts the 2D image coordinates of known 3D (three-dimensional) world points from calibration images and then uses these data to determine the camera's intrinsic and extrinsic parameters. This method is highly accurate, making it suitable for traffic study applications, which usually require precise camera parameter estimation. However, its computational intensity and reliance on good initial estimates may present challenges for practical implementation.

Building upon Tsai's model, Heikkila & Silvén (1997), introduced a refined approach that incorporates a more comprehensive distortion model, including both radial and tangential distortions. While this extension enables the method to handle more complex distortions with greater accuracy, it is even more computationally demanding and involves a more complex implementation than Tsai's model.

In contrast to these more computationally intensive methods, Zhang's camera calibration method is widely adopted for its simplicity, robustness, and flexibility (Zhang, 2000). It employs a planar calibration pattern with known dimensions, such as a checkerboard, which is observed from various angles and positions. This method is valued for its straightforward application in camera calibration. However, it is recognised as less resilient to noise and errors when compared to more sophisticated calibration techniques.

Recent research has also introduced new calibration frameworks tailored to specific needs. For instance, Dantsuji et al. (2022) propose a new framework for calibrating dynamic car origin–destination matrices in large networks. Similarly, Zhang & Jin (2019) proposed a mathematical model for transforming coordinate systems that yields accurate results for UAV data; however, the accuracy of vehicle detection from roadside cameras can be further expanded by incorporating 3D ground plane estimation. In the paper, Zhang & Jin (2019) present a new framework for calibrating dynamic car origin–destination matrices in large networks. Additionally, Rangesh & Trivedi (2020) introduced a method to accurately determine the size and position of 3D objects using single 2D images, integrating visual cues and ground plane information to produce robust 3D detection boxes.

In contemporary computer vision field, it is common to distinguish between intrinsic and extrinsic camera calibration. Intrinsic calibration focuses on determining a camera’s internal parameters, such as focal length, optical centre, and distortion coefficients. Typically, a checkerboard or other calibration pattern is used for this purpose, with the pattern placed at various positions and orientations within the camera’s field of view. By capturing multiple images of the pattern, 2D corner points are matched to their corresponding 3D object points, and tools such as the OpenCV Camera Calibration Toolkit (‘*cv2.calibrateCamera()*’ function) are employed to estimate the intrinsic parameters.

Extrinsic calibration is concerned with determining a camera's position and orientation relative to a reference object or coordinate system in 3D space. After completing intrinsic calibration, extrinsic parameters can be computed using techniques such as Perspective-n-Point (PnP). PnP solves for the camera's pose by relating known 3D object points to their 2D projections in the image. OpenCV’s ‘*cv2.solvePnP()*’ function is frequently used to estimate the camera’s position and orientation, leveraging the intrinsic parameters and the 2D–3D point correspondences.

A variety of other software tools are also available for calibrating video cameras using sophisticated calibration models, such as the T-Calibration implementation in the T-analyst software package (T-Analyst, 2019), which is based on Tsai’s model description. The software adjusts parameters such as lens distortion, field of view, and perspective to match real-world coordinates.

Figure 6. Example of camera accuracy output following calibration using T-Calibration software (T-Analyst, 2019)

2.3.3 Optional video processing steps

2.3.3.1 Brightness adjustment

Brightness adjustment is a common technique used to optimise the ability to discern road users in video data recorded under electric lighting conditions, such as streetlights during dark hours. This involves either improving visibility in darker regions or reducing the glare caused by overexposed lights. The adjustment process should be selective, applied only when necessary to improve the visibility of road users (e.g. pedestrians, cyclists, motor vehicles) under varying lighting conditions. In many cases, brightness and contrast can be manually adjusted based on frame analysis, allowing precise control over the final result. More advanced approaches like EnlightenGAN (Jiang et al., 2021) and zero-reference models (Guo et al., 2020) have been designed for low-light optimisation. However, generative models may introduce inaccuracies, such as adding or omitting road users based on inferred patterns, leading to discrepancies in the enhanced image.

Figure 7. An example of brightness adjustment for the same camera view (before adjustment—left; after adjustment—right)

2.3.3.2 Stitching multi-source video feeds

In multi-camera setups, video stitching is necessary to merge footage from multiple cameras into a single, continuous view to create a smooth final video. While video stitching builds on principles similar to those of image stitching, it is a more complex process due to the additional challenges of synchronising motion, managing dynamic scenes, and ensuring consistent transitions across frames (Abdel-Aty et al., 2023). Essentially, it extends and generalises the techniques used in multi-image stitching to accommodate the temporal and spatial consistency required for video (Wei et al., 2019).

Successful video stitching requires overlapping areas in images captured from different video sources. For example, in Figure 8, the green car is the same vehicle viewed by two different cameras. The process involves several key steps, including frame synchronisation, camera calibration, detecting key points (e.g., corners, edges) in the overlapping areas of each video feed and matching them across frames from different cameras.

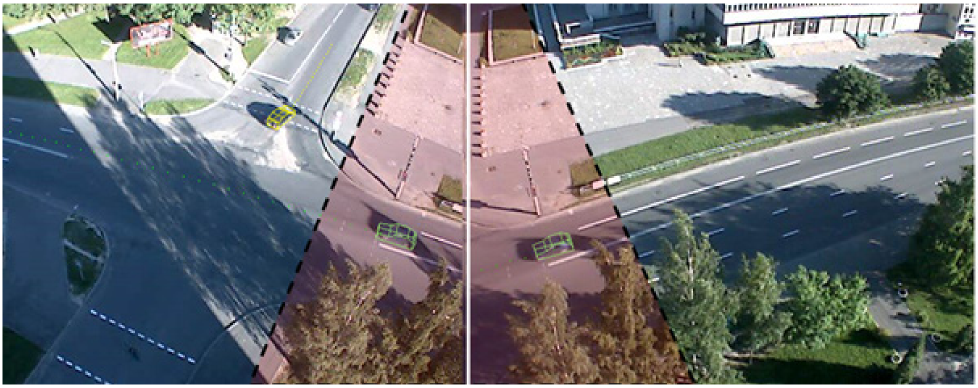

Figure 8. An example of stitching video feeds captured by two different cameras (red highlighting the overlapping areas in the images from both sources)

The literature presents a variety of techniques to manage this process, ranging from traditional (manual) methods to more advanced approaches (Wei et al., 2019; Brown & Lowe, 2007; Szeliski, 2007; Duffin & Barrett, 2001). The effectiveness of video stitching techniques often depends on whether the cameras are fixed or moving. Moving cameras typically face greater challenges, such as shakiness and large parallax, which require more sophisticated solutions (Nie et al., 2017; Lin et al., 2016; Su et al., 2016).

2.4 Object detection algorithms

Object detection is the process of identifying and locating objects that belong to a predefined set of object classes within an image or video. In the context of traffic studies, this typically involves detecting and classifying road users, as shown in Figure 9. Object detection methods are generally categorised into three main approaches: motion-based techniques, which use temporal information to detect objects; appearance-based techniques, which extract relevant features from individual images; and anchor-free approaches, which directly predict object bounding boxes or key points without relying on anchor boxes. Each approach offers distinct advantages and limitations (Ghahremannezhad et al., 2023). The choice of method depends on the specific task requirements, such as the need for quick processing or high detection precision.

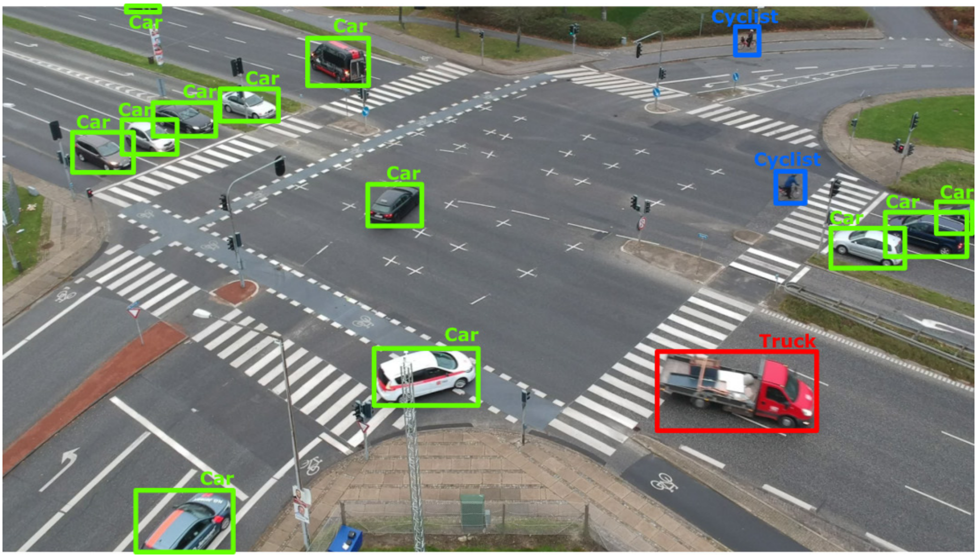

Figure 9. An example of object detection in a traffic surveillance image (Jensen et al., 2019)

2.4.1 Motion-based methods

Motion-based approaches use the motion of objects of interest to distinguish them from the background, allowing their locations to be detected. Common techniques for detecting motion in traffic video analysis include frame differencing, optical flow, and statistical background modelling.

Frame differencing is one of the simplest motion estimation techniques. It calculates the absolute difference in pixel intensities between consecutive frames and applies a threshold to highlight significant changes, thus identifying the locations of moving objects (Al-Smadi et al., 2016). The optical flow technique detects moving objects by analysing the correlation between adjacent frames to find matching key points. This process enables the calculation of optical flow vectors, which represent the instantaneous velocities of specific points in the image (Rauf et al., 2016). Background modelling is the most commonly used motion-based approach in traffic video surveillance. It distinguishes the static background from the moving objects in the foreground and then identifies these objects based on their movement or appearance within the scene (St-Charles et al., 2015).

While motion-based object detection offers advantages in terms of generalisation and computational efficiency, it also presents several challenges, which arise from general issues, such as non-stationary cameras and stationary objects, as well as traffic-specific difficulties, including moving shadows, variations in lighting, and occlusions (Ghahremannezhad et al., 2023; Garcia-Garcia et al., 2020).

2.4.2 Appearance-based methods

Appearance-based object detection methods do not rely on motion or background information to extract and analyse visual features from images or video frames to detect and classify objects of interest. Instead, they use patterns learned from labelled data to identify objects based on their appearance. Common techniques include feature-based approaches, convolutional neural networks (CNNs), and Transformers. These methods are often combined, leveraging the strengths of each approach to improve detection accuracy and robustness.

2.4.2.1 Feature based techniques

Feature-based methods identify specific patterns, shapes, or features in an image to determine whether they belong to a particular category. These traditional techniques were widely used before deep learning took over. For example, Histogram of Oriented Gradients (HOG) extracts gradient features from images and is commonly used for detecting people (Wang et al., 2009). Similarly, techniques like Scale-Invariant Feature Transform (SIFT; Lowe, 2004) and Speeded-Up Robust Features (SURF; Bay et al., 2008) are used to extract key points and descriptors for object recognition. While these methods rely on relatively simple architectures, they tend to be less accurate and slower compared to modern deep learning techniques, especially in complex scenarios (Abdel-Aty et al., 2023).

2.4.2.2 Deep learning-based techniques

The advancement of machine learning and deep learning techniques has introduced new approaches to object detection and classification, offering improved accuracy during both model training and testing (Rani et al., 2022). Deep learning methods for object detection, classification and tracking—particularly those utilizing CNNs—are among the most advanced techniques used in traffic research.

CNNs are specialised deep learning models designed to automatically learn and identify patterns in images. They comprise a computer algorithm divided into several nodes, each performing straightforward operations. The performance of these operations relies on certain parameters, such as weights, which are computed based on training data (Ahrnbom, 2022). CNNs are particularly well-suited for multi-class object detection tasks as they use multiple binary classifiers in the final layer to distinguish between different object categories. Object detection methods based on CNN-driven features are typically categorised as either two-stage or single-stage approaches (Zhang et al., 2022).

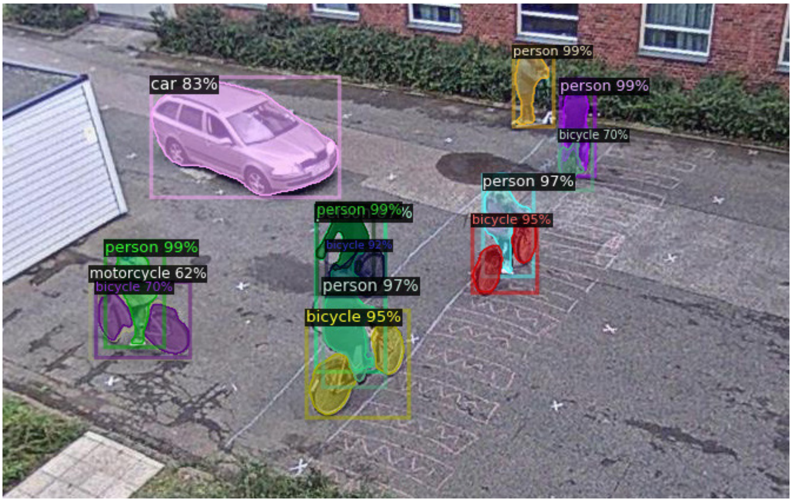

Figure 10. Example output from the Mask R-CNN object detector, a type of convolutional neural network (CNN), adapted from Ahrnbom (2022))

Several factors can influence the performance and evaluation of deep learning models in traffic analysis. First and foremost, training these models requires large, annotated datasets containing images or video frames that clearly identify objects of interest, such as road users. The quality and diversity of these datasets are critical; a well-annotated and varied dataset improves the model's ability to generalise across different scenarios. Additionally, traffic environments are highly variable, ranging from urban streets to rural highways, with changing conditions in terms of lighting, weather, and time of day. To account for this variability, it is important to train the model on a diverse dataset that captures these differences. For instance, training with images showing vehicles from different angles, speeds, and sizes improves the model's ability to generalise and perform accurately in real-world scenarios. While extensive datasets contribute to improved accuracy, the process is resource-intensive and time-consuming, raising concerns about the computational demands and resources required for effective training.

– *Two-stage object detection*

Two-stage methods, while generally more accurate, are often slower than single-stage methods. First, the detector scans the image and suggests regions of interest where objects might be located. Then, it carefully checks these regions to classify the objects and determine their exact locations. Region-based convolutional neural networks (R-CNNs) are an early example of a two-stage method, introduced by Girshick et al. (2014). Over time, R-CNNs evolved into faster models, including Fast R-CNN (Girshick, 2015) and Faster R-CNN (Ren et al., 2015). These methods are often preferred for applications where high accuracy is necessary.

– *Single-stage object detection*

One of the most popular single-stage deep learning object detection algorithms is You Only Look Once (YOLO), introduced by (Redmon et al., 2016). Instead of dividing the process into two steps, YOLO improves speed by scanning the entire image in one pass. It splits the image into a grid and simultaneously predicts both the presence and location of objects. The YOLO framework has evolved over several versions, from YOLOv1 to YOLOv9, with each iteration improving accuracy and speed through advanced architecture and optimisation techniques (Yaseen, 2024; Terven et al., 2023). While YOLO offers significant speed advantages, it may not be as precise as some two-stage methods.

The Single Shot MultiBox Detector (SSD; Liu et al., 2016) and RetinaNet (Lin et al., 2017) are other widely used deep learning object detection algorithms. These methods offer distinct advantages based on specific application needs. SSD is commonly chosen for tasks requiring speed and simplicity, such as real-time traffic monitoring. Meanwhile, RetinaNet excels in applications where accuracy is critical, particularly in detecting small objects within complex scenes.

2.4.3 Anchor-free methods

Traditional object detection methods use anchor boxes—predefined bounding boxes with specified parameters such as height and width—to guide the detection process. In contrast, anchor-free methods simplify this process by directly predicting the centre points or corners of objects, eliminating the need for anchor boxes and their associated computational complexities.

Anchor-free approaches are categorised as keypoint-based or centre-based approaches (Zhang et al., 2023). Keypoint-based methods detect objects by identifying key points and then generating bounding boxes around them. A notable example from this category is the Fully Convolutional One-Stage Object Detector (FCOS; Tian et al., 2022). Centre-based methods define positive samples based on the centre point or area of the object and then predict the distance from the centre to the object's edges. Popular centre-based models include CenterNet (Law & Deng, 2020; Zhou et al., 2019) and Grid R-CNN (Girshick et al., 2014).

While these approaches offer greater flexibility in detecting objects of varying sizes and shapes—a key advantage for identifying road users, as sizes can differ significantly even within the same category (e.g. cars, buses)—they also face several challenges. Issues with accuracy, scalability, and robustness can impact performance, particularly when encountering complex real-world scenarios with factors like occlusions, lighting changes, and diverse environments.

2.5 Object tracking algorithms

Object tracking involves monitoring and connecting the movement of objects (in traffic studies, these are road users) across frames in a video, each represented by a trajectory. In other words, each detected object is assigned an identification, as shown in Figure 11, that remains consistent over time. However, several challenges can arise, such as multiple tracks being assigned to a single object, incorrect splitting or merging of tracks, and track confusion, where different objects are mistakenly followed over time, among others.

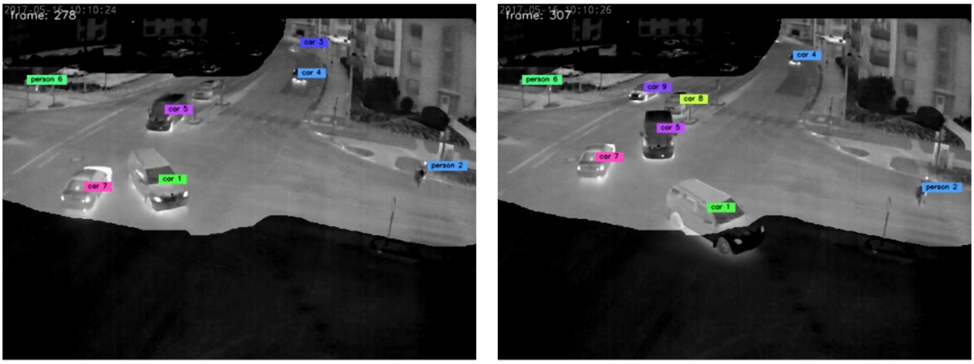

Figure 11. Example of tracking results using the STRUDL tool (adapted from Jensen et al. (2019)).

There are various existing methods for object tracking, each with its strengths and suitability depending on factors such as environmental complexity, the nature of the objects being tracked, and available computational resources. These methods ensure the consistent identification of an object's position as it moves, which is crucial for accurate analysis and interpretation.

2.5.1 Online tracking methods

Online tracking methods are techniques in which the tracker processes frames sequentially and cannot revisit previous frames. In this approach, the tracking algorithm makes real-time decisions and updates as new frames arrive, without relying on past information. A widely used method in this category is Simple Online and Realtime Tracking (SORT; Bewley et al., 2016). This is a straightforward and efficient approach for tracking objects that matches detections (obtained using a detection algorithm) to previously tracked objects based on their position, size and object class and makes assumptions about their movement patterns. In each frame, SORT uses a Kalman Filter (Kalman, 1960) to predict where the object will be in the next frame. It then compares these predicted positions with new detections to update the tracking. The Hungarian algorithm (Kuhn, 1955) is employed to pair detected objects with existing tracks based on proximity and movement. While SORT is efficient and effective in straightforward scenarios, its simplicity and speed may lead to difficulties in very crowded scenes or when dealing with rapid object movements. DeepSORT (Wojke et al., 2017) is an expanded version of SORT that integrates appearance information to handle occlusions and similar-looking objects more effectively. Building on the SORT tracker, Ahrnbom et al. (2021b) introduced two additional methods: SORTS and SORTS+ReID. SORTS improves tracking by incorporating segmentation while maintaining real-time processing speeds, while SORTS+ReID is designed to improve occlusion handling.

2.5.2 Pixel-based methods

Pixel-based methods track objects by analysing changes in their appearance across frames using various techniques. The optical flow technique assesses pixel movement between frames to determine an object's motion (Alfarano et al., 2024). Template matching involves comparing a reference image of an object with the current frame to identify its location. Background subtraction separates the object from the background to track its movement. These methods are frequently employed in real-time vehicle tracking applications and for monitoring vehicles in situations where occlusion occurs (Zheng et al., 2024; Abdel-Aty et al., 2023). While effective, they can be sensitive to changes in lighting and to situations in which objects blend into the background.

2.6 Microscopic traffic data from video

Video recordings offer a wealth of detailed microscopic traffic data, which provide deep insights into both road user behaviour and vehicle dynamics. These data refer to the detailed observation and analysis of individual elements in traffic, such as individual road users, in contrast to aggregated data. They encompass a range of data types, from the frequency of particular behaviours to continuous, highly precise information linked to time and space. These detailed data are fundamental to understanding the nuanced interactions and behaviours that impact traffic safety and the overall comfort of road users.

By applying calibration, detection, classification and tracking models to video footage from specific areas of interest within a transportation network, a series of points or bounding boxes depicting the positions of each road user can be identified, forming trajectories (see Figure 12). These trajectories are represented by data points that describe positions at specific time intervals over a defined period. Various indicators—objective and measurable parameters—such as speed, lateral position, and acceleration/deceleration, and statistical descriptors, like average, median, and standard deviation, can be derived from these continuous data.

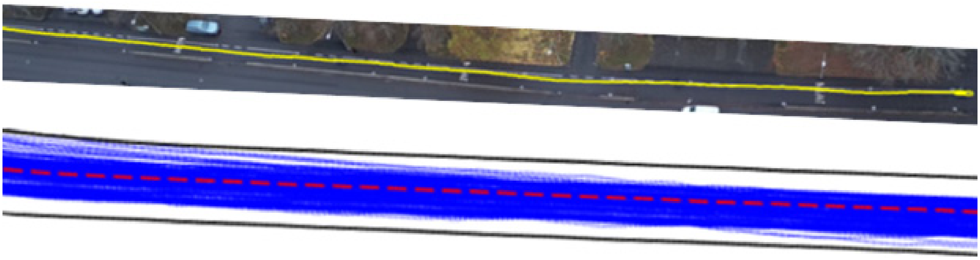

Figure 12. Visual representation of generated trajectories along a shared-use path. The upper section presents an aerial view of the area with one annotated trajectory, while the lower section displays several plotted trajectories in relation to the path

Trajectories can be captured and analysed in 2D or 3D formats, depending on the level of detail and accuracy required. Two-dimensional trajectory analysis tracks object movement within a flat plane using x (horizontal) and y (vertical) coordinates. This analysis typically relies on video recordings or sensor data captured from bird's-eye view or fixed-angle cameras. However, in complex scenarios, such as dense traffic, 2D trajectories often struggle with occlusion issues, making it difficult to accurately distinguish overlapping vehicles. Additionally, 2D trajectories are less reliable when changes in elevation are involved, for which 3D trajectory analysis is necessary. This approach tracks objects in three dimensions, incorporating x, y, and z coordinates.

In addition to choosing between 2D and 3D formats, trajectories can be represented in different ways based on the study's requirements. Typically, road user trajectories are depicted either as a single point within a bounding box (often the central position of the road user on the ground in 2D) or with more detailed 3D poses. Representing trajectories as a single point provides sufficient detail for many traffic analysis tasks. However, for studies focusing on active behaviours, such as the avoidance actions of vulnerable road users or single falls, 3D pose trajectories offer important insights (Gildea et al., 2024; Li et al., 2021). Capturing these detailed 3D pose estimates requires advanced techniques and technologies that offer a more comprehensive view of movement and interactions (Gildea et al., 2022; Reddy et al., 2021; Iskakov et al., 2019).

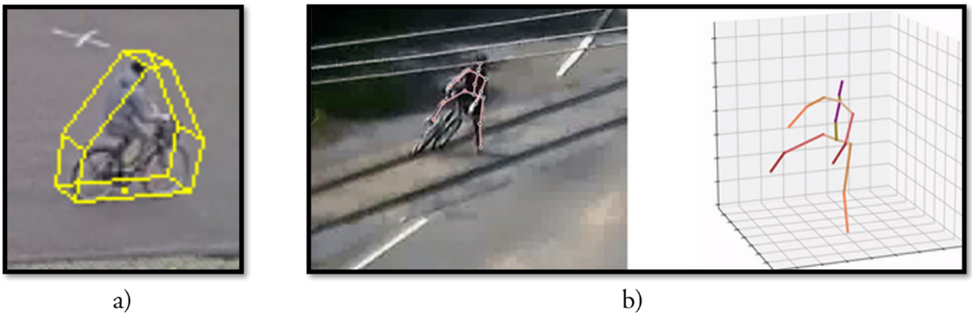

Figure 13. Examples of trajectory representation: a) The central position of the road within the 2D projection on the ground plane, as defined by the bounding box (T-Analyst, 2019), b) sequence of reconstructed 3D poses derived from corresponding 2D keypoints

2.6.1 Tools for data extraction

Extracting road users' trajectories from video recordings is a critical process that can either be fully automated or involve some level of human input (semi-automated) depending on factors like the complexity of the environment, accuracy requirements, and the available technology.

Semi-automated approaches allow users to manually refine detections, adjust tracking parameters, or provide input to improve accuracy in challenging scenarios. This approach is often used when the environment is complex or when high precision is required. The commonly used tool for generating 3D trajectories in traffic studies is T-Analyst (T-Analyst, 2019), a video analysis tool designed specifically for generating 3D trajectories and calculating various measures for further analysis in the field traffic of studies. This tool provides very accurate results, which can also be used as ground truth for training advanced models. However, its detection, classification and tracking functions require extensive manual work before the final output is achieved, making the process labour-intensive and time-consuming. Some manual steps can be streamlined using additional tools. For example, software RUBA (Agerholm et al., 2017; Madsen et al., 2016) can filter hours of video by detecting movements within user-defined regions, allowing for the automatic identification of situations of interest for the human observer to review. While this step still necessitates the observer's manual input, it is significantly faster than rewatching all the recorded videos. However, there are clear limitations: the output depends on the accuracy of the defined regions or detectors in the video. For instance, if a road user moves slightly outside a defined detection area, they will not be registered. Conversely, if the detection area is expanded beyond expected movement zones, the number of false positives may rise. Other factors, such as birds flying near the camera, raindrops, or falling leaves, can also contribute to increased false positives.

Other semi-automated tools commonly used for extracting microscopic traffic data include TVATIC (Shen, 2016), TDE (Munigety et al., 2014), TRAZER (Mallikarjuna et al., 2009), and Kinovea (Kathuria & Vedagiri, 2020). Despite the intensive manual effort required, recent research frequently employs such semi-automated tools (Diwakar et al., 2024; Yastremska-Kravchenko et al., 2024; Johnsson et al., 2021) due to their effective data extraction capabilities. However, these tools are primarily used for evaluating the usability of various indicators or for analysing small datasets that are manageable within a feasible timeframe rather than for automatically computing metrics for every road user interaction in lengthy videos.

In contrast to semi-automated tools, automated tools rely on advanced algorithms and machine learning models to detect, classify, track, and extract trajectories from video footage without requiring manual input. These tools are efficient at processing large datasets and are capable of operating in real time. A common approach in traffic studies involves combining the YOLO object detector (Terven et al., 2023) with the SORT algorithm (Bewley et al., 2016) to infer trajectories automatically, especially in relatively straightforward environments. Other deep learning models, such as SSD and Faster R-CNN, have also proven effective for handling large datasets with real-time processing capabilities. While these automated tools provide consistent results that are free from human biases or errors, their accuracy can be affected by poor lighting, occlusions, or crowded scenes. Additionally, these systems often demand significant computational resources, especially when processing high-resolution video footage.

One of the first cross-disciplinary frameworks for automated traffic analysis is the STRUDL package (Surveillance Tracking Using Deep Learning) which encompasses all processes from video acquisition to traffic analysis output. This system employs deep learning algorithms for the detection and classification of road users, utilising the Hungarian tracking algorithm to maintain tracking continuity. Although STRUDL automates many video processing steps, it still requires some human intervention, particularly in annotating image data to train the object detector. Additionally, the tracking algorithm, while straightforward, can make errors when trajectories become too close to one another or if the detector fails to detect an object over several consecutive frames. However, the system continuously improves it processes and incorporates more data into its learning.

Traffic Intelligence is another example of a tool for automated road safety analysis utilising video recordings (Saunier, 2024; Saunier et al., 2010). This software has been created for multiple projects and applications and continues to be actively developed. It features two main tools for detecting, tracking, and classifying road users, along with a Python library designed for analysing the resulting data, especially for performing road safety evaluations.

Additionally, several organisations, such as Citylog and Viscando, provide various tools and systems for extracting microscopic traffic data. Some solutions offer the complete workflow outlined in Figure 1, starting from data collection to the extraction of microscopic data, while others work with existing videos and focus only on subsequent steps. Additionally, numerous traffic monitoring systems, such as Traficon, are used for tasks like speed violation detection, statistical data collection, and potential enforcement of traffic regulations. These solutions are widely adopted for their advanced technology and reliability in capturing a wide range of traffic data, from basic vehicle counts or speed detection (without spatial context) to detailed trajectory analysis. The selection of the appropriate system should be guided by the study's specific objectives and the required level of detail in the data.

2.6.2 Trajectory accuracy

The accurate representation of road user trajectories is crucial for traffic studies, as they provide insights into movement patterns and interactions between road users and infrastructure. However, extracting high-accuracy trajectories from video recordings is no simple task. As well as video processing techniques (stabilisation, calibration, etc.) and the chosen detection and tracking methods (which often involve a trade-off between accuracy and efficiency), several other factors, ranging from environmental conditions to technical setup challenges, can significantly influence trajectory accuracy.

Among these factors, environmental conditions often present the most unpredictable challenges. For example, when it rains, droplets on the camera lens can blur the image and reflections from wet surfaces can confuse tracking systems, making it difficult to distinguish between actual objects and visual obstructions. In dense fog, visibility is reduced to the point where objects are not fully discernible and sometimes even disappear altogether, leading to fragmented or incomplete trajectory data. Each of these weather conditions adds layers of complexity to tracking systems, which are built to rely on clear and consistent environmental conditions.

Lighting is another critical factor. At night, limited illumination means that objects can be hard to see, leading to poor detection and tracking performance. Even with electric lighting, deep shadows or unevenly lit areas can cause significant inconsistencies in the captured data. Conversely, daylight can introduce its own set of challenges, including glare from reflective surfaces and shadows that change throughout the day, potentially obscuring parts of the scene and leading to mistakes in road user detection and tracking.

The camera's technical setup also plays a key role in determining trajectory accuracy (Laureshyn & Nilsson, 2018). Camera height (the distance between the camera and the studied object) is one key factor; higher positions can make objects appear smaller, which is particularly challenging in low-light conditions and when detecting smaller road users like pedestrians. Additionally, side or angled views can introduce distortions, as objects farther from the camera seem smaller and move more slowly, complicating accurate trajectory estimation. Occlusion is another common issue that can affect trajectory reliability. However, overhead views, increasingly available due to advancements in drone technology, generally provide a clearer and more direct perspective, helping to mitigate these issues and improve trajectory accuracy.

Despite these challenges, it is currently possible to extract meaningful and accurate trajectory data. However, achieving this requires more than just advanced technology. Careful planning and study design, thorough consideration of the study or application's specific purposes and requirements, and thoughtful implementation of modern technology are crucial for effectively addressing most of the challenges associated with trajectory data reliability.

3 Understanding behavioural observation studies in traffic research: A perspective on VRUs

3.1 The evolution of behavioural observation in traffic studies

Understanding the context in which behaviours occur is necessary for accurately interpreting road user actions. Factors such as traffic density, infrastructure characteristics, lighting, weather conditions, and other elements can significantly influence road user behaviour. Traffic observation studies take these contextual elements into account by examining the behaviour of road users in their natural settings to offer a more comprehensive understanding of how frequently specific characteristics of their behaviours occur in various situations.

Behavioural observation is one of the earliest research methods used in traffic research, dating back nearly a century to the 1930s (van Haperen et al., 2019; Eby, 2011). Despite its long history, this method remains effective and widely used. It is commonly employed for monitoring purposes and is particularly valuable in situations in which few data are available or existing data lack detail. Additionally, behavioural observation is effective for assessing whether certain measures have achieved their intended effects and whether they have caused any unintended side effects at an early stage (Polders & Brijs, 2018).

According to literature review by van Haperen et al. (2019), starting in the 1930s, behavioural observation studies primarily focused on motor vehicle road users. This perspective only expanded in the 1970s to include vulnerable road users (VRUs: those without a protective shell, such as pedestrians, cyclists, and motorcyclists), who then became a significant area of interest (see Figure 14).

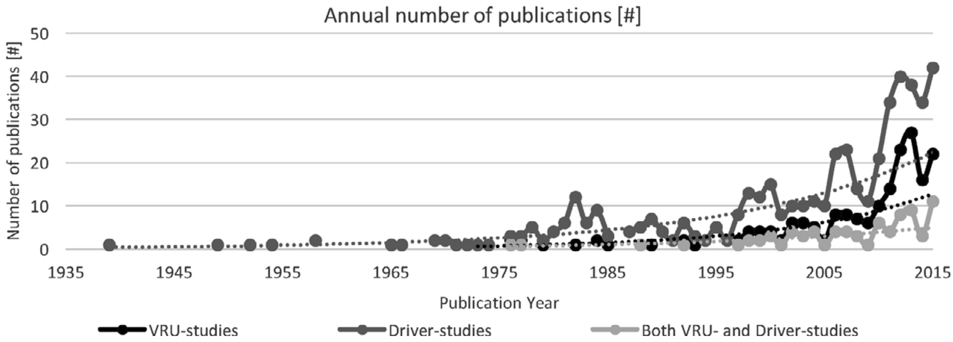

Figure 14. The annual number of publications per study type (adopted from van Haperen et al. (2019))

Several factors caused this increased interest in VRUs. First, VRUs are particularly vulnerable to serious injuries or fatalities in traffic accidents because of their lack of physical protection (Kim et al., 2007; Rowe et al., 1995). Second, they are more likely to be involved in crashes, especially those involving motor vehicles (Schieber & Sacks, 2001; Rowe et al., 1995). In many motor vehicle–cyclist incidents, drivers fail to detect cyclists until it is too late to prevent a collision (Wood et al., 2009). By observing VRUs' behaviour and their interactions with the environment, researchers can identify specific risks and develop targeted safety measures to address these vulnerabilities.

Additionally, the recent emphasis on promoting walking and cycling for sustainability and health benefits has underscored the need for transportation systems to include adequate infrastructure to support these activities (Delaney et al., 2017).As cities work to integrate these modes of transport into their infrastructure, it is important to design road environments with VRUs in mind. Behavioural observation studies are invaluable in this context, providing insights into how VRUs interact with various road features that can inform policies to ensure that these elements are both safe and functional.

As transportation systems increasingly support a mix of travel modes, understanding how VRUs navigate and interact with different forms of transport is key to ensuring smooth and safe intermodal connectivity. Furthermore, VRUs represent diverse groups of individuals with varying needs and behaviours. Behavioural observations can provide additional knowledge about how various transportation modes can work together and where potential conflicts might arise, guiding the development of solutions that strengthen overall system efficiency and user safety. Additionally, this method can shed light on how different demographics (e.g., children, the elderly, or individuals with disabilities) experience and navigate the road environment. This understanding is fundamental for developing inclusive safety measures that protect all VRUs.

The methods used to collect behavioural observation data have evolved over time. Historically, the predominant method used in traffic studies involved on-site human observers. However, advancements in video technology have shifted the preference toward using video recordings, either as the primary source or in combination with trained on-site observers (Laureshyn & Várhelyi, 2018). By the mid-2000s, video cameras had largely surpassed direct human observations as the primary method for this type of research, coinciding with a rapid increase in the number of behavioural observation studies published in the scientific literature (van Haperen et al., 2019).

This shift can be attributed to the several advantages that video recordings offer. They provide a more objective and accurate means of behavioural observation, enabling continuous monitoring of road users and the ability to review footage to verify results. Video recordings also facilitate the collection of detailed, continuous microscopic data, such as speed and positioning. The literature review conducted by van Haperen et al. (2019) revealed that earlier behavioural observation studies of VRUs often relied on binary variables (yes/no measurements). In contrast, more recent studies have shifted toward using detailed continuous variables, replacing binary measures (or adding to them) for a more in-depth data analysis process.

Additionally, collecting video recordings is less labour-intensive and is not subject to on-site observer fatigue, which can affect data quality in traditional on-site observations (Laureshyn, 2010). However, as discussed in Part 1 of this thesis, the process can still be demanding with respect to the subsequent steps of transforming raw video data into advanced analysis.

3.2 Behavioural insights from video: What can be measured?

Traffic behavioural studies differ significantly from other psychological approaches. While traditional psychology often focuses on exploring internal processes (e.g. intention, attitude, individual beliefs) traffic behavioural studies primarily aim to collect and analyse data on observable actions—actions resulting from these internal processes.

Both qualitative and quantitative variables can be captured from video recordings in the form of single or binary (yes/no measurements) values, and quantitative variables can also be recorded as continuous sequences (a set of values over time) (Laureshyn, 2010). Qualitative variables encompass demographic factors such as age and gender, road user categories like motor vehicle drivers and VRUs (e.g. cyclists, motorcyclists, and pedestrians from various demographics, such as children and the elderly), as well as types of interactions, approaching behaviour, distractions, and other observable actions (e.g. looking, eye contact, yielding, crossing, directional indications). These variables reveal how different road users respond to the same environmental stimuli. For example, the crossing behaviour of road users, such as whether they choose to slow down or accelerate while crossing, can highlight their different decision-making strategies based on their transport mode, age, and other factors. Additionally, qualitative data can include infrastructural characteristics of the location, such as traffic rules and how actual behaviour aligns with those rules (e.g. cases of red-light violations).

In contrast, quantitative variables, such as positioning, speed, and distances (both longitudinal and lateral), as well as their variations, are measured to provide objective data on how road users navigate their environment. For example, tracking speed variations allows researchers to evaluate how consistently drivers adhere to speed limits under varying conditions or how speed changes in response to traffic density or infrastructure design. Both qualitative and quantitative data offer valuable insights into the underlying patterns of road user behaviour.

Several studies have demonstrated that the placement and speed of road users can reflect behavioural changes in response to their environment (Johansson et al., 2020; Greibe & Skallebæk Buch, 2016; Franěk, 2012; Fisher & Nasar, 1992). This data can be represented in multiple ways, including as extreme values (e.g. maximum and minimum), measures of central tendency (such as average, median, and mode), percentiles, and others. Each method provides distinct insights into the data, allowing researchers to emphasise specific aspects for analysis in accordance with their study's objectives.

The speed profile, which illustrates how road users adjust their speed over time or distance, is a frequently employed measure in behavioural research. This metric offers valuable insights into behavioural dynamics, revealing patterns of acceleration, deceleration, or consistent speed in response to various environmental or task-related factors. Variations in speed can often reflect internal decision-making processes. For instance, when cyclists slow down, it may signal uncertainty or a more deliberate cognitive process when navigating a complex or novel environment. Conversely, an increase in speed might suggest that cyclists perceive the traffic conditions to be safe, prompting them to move faster, which can be interpreted as a positive response to their surroundings. In contrast, for pedestrians, moving faster might mean they want to reach their destination quickly because they feel less secure or simply do not like the surrounding environment.

Placement offers a range of interpretations and is another important variable in behavioural studies. It can refer to a road user's lateral position or how this position changes over time, as well as the distance between road users (longitudinally and/or laterally). Placement data are often used to assess interactions with infrastructure, revealing whether it is being used as intended. For example, if pedestrians are walking in the road instead of on a sidewalk due to poor conditions, or if cars are parked in a way that disrupts traffic flow, these patterns can indicate potential design issues or safety concerns in the road environment.

Angle-based indicators can also be derived from trajectory placement data. For example, swerving intensity can be assessed by measuring the road user's angular velocity (yaw rate) (Guo et al., 2018; Tageldin et al., 2015).

Another relevant metric is the slalom measure, which evaluates directional changes along a trajectory based on position data. This measure can be expressed in different ways, such as by describing the curvature to demonstrate how sharply the trajectory bends or the oscillations—referring to the frequency with which a trajectory crosses a reference line or axis—to represent how frequently the trajectory crosses a reference line or axis. Furthermore, angular changes derived from the tangent vector help to assess turning behaviour along the trajectory.

In one of my studies (*Paper II*), was employed the slalom measure to determine how much the road users' path deviated from a baseline, a straight line connecting the entrance and exit points of the path (see Figure 15). These points were identified from trajectories sampled every 20 metres to capture the overall pattern and minimise the influence of small, speed-related fluctuations and noise in road users' trajectories. The measure was then calculated as the cumulative sum of integrals between the actual path and the baseline, providing a comprehensive indicator of deviation.

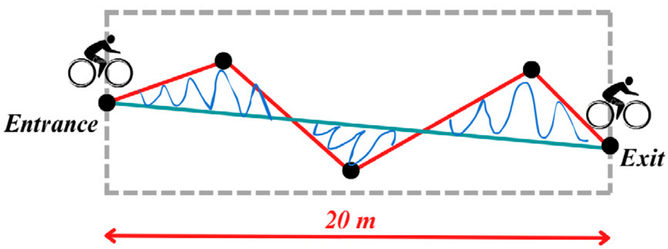

Figure 15. Illustration of the slalom measure: black circles represent the actual trajectory, while the straight line between entrance and exit represents the baseline (adopted from Paper II)

The interpretation of the slalom measure varies depending on the type of path user. Commuters and individuals engaging in work-related tasks (e.g. couriers) often prioritise efficiency, aiming to maintain a straight trajectory to reach their destination swiftly. In contrast, tourists or recreational cyclists may prefer a less direct route, deliberately deviating from a straight path to explore their surroundings or enjoy the journey.

Placement data also provide insights into the interactions between several road users, including various manoeuvres (e.g. overtaking, following, bi-directional), as well as detecting potential critical events in traffic. These insights are crucial for assessing their impact on road safety. Additionally, trajectory placement data offer several insightful metrics, for example, the lateral distribution of road users and the gaps they maintain (both laterally and longitudinally) under different traffic conditions (Mallikarjuna et al., 2009).

Building on these insights, *Paper I* introduces a new measure called '*Safe Lateral Passing Distance'* specifically for cyclists in bi-directional interactions. Using naturalistic data, I found that cyclists approaching each other often swerve to the right, a pattern similar to that noted by Yuan et al. (2018). The measure reflects the lateral space cyclists maintain before meeting and until fully passing another cyclist. In *Paper I*, this *Safe Lateral Passing Distance* was defined as the minimum distance preserved in 95% of the observed interactions. In my case study, this distance was specifically determined to be 1.5 meters (see Figure 16).

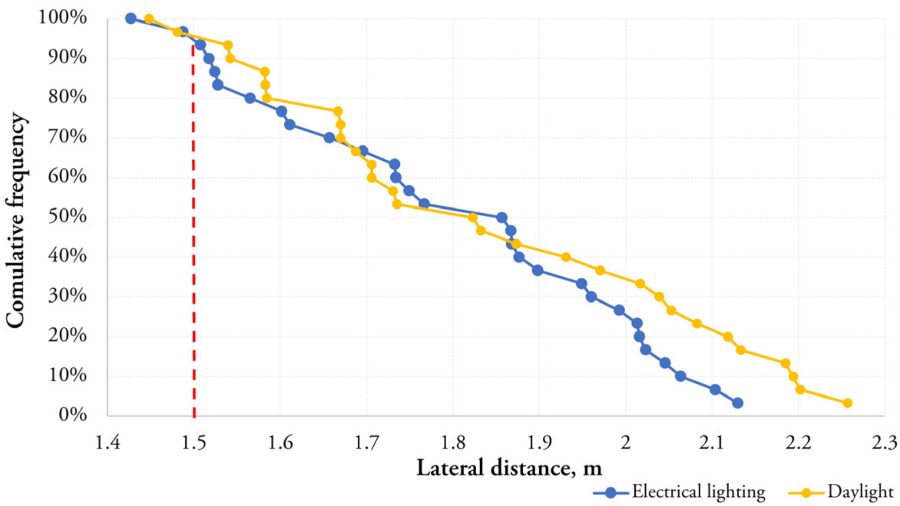

Figure 16. Lateral space between cyclists at the moment they meet each other (adapted from Paper I)

In Figure 17, phases 'p1' and 'p2' illustrate the instance at which the first and second cyclists, respectively, reach the *Safe Lateral Passing Distance*. Two additional variables may also be considered: the longitudinal distance between cyclists (p1–p2) and the time they take to arrive at their meeting point after establishing this distance.

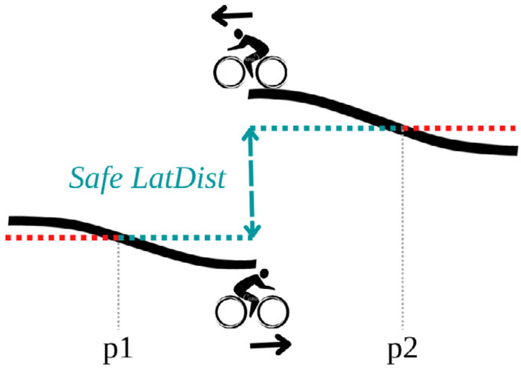

Figure 17. Illustration of the Safe Lateral Passing Distance (adopted from Paper I)

3.3 Practical case study

Despite efforts to promote walking and cycling, little is known about how specific design features impact pedestrian and cyclist behaviour, leaving a gap in our understanding of how to optimise infrastructure for these groups. This knowledge gap is becoming even more relevant as municipalities and urban planners face growing demands for effective methods and tools that provide rapid feedback on the effectiveness of interventions in physical environments.

One of the primary objectives of this thesis was to explore the practical application of microscopic behavioural data in traffic studies, with a particular emphasis on optimising the experience for vulnerable road users (VRUs) and improving their safety. This section delves into the use of microscopic data extracted from video recordings to observe and analyse traffic behaviour and presents findings from a large-scale behavioural study conducted as part of the research project 'Belysa—miljöfaktorers inverkan på fotgängares och cyklisters beteende i skymning och mörker'.

The ability to analyse and respond to behavioural data is necessary for making informed decisions that improve safety and promote sustainable transport options. To address this need, the thesis developed a methodology specifically for examining the travel behaviour of cyclists and pedestrians under varying lighting conditions, including both daylight and electric lighting. This new approach builds on the foundational work of the Video Analysis of Pedestrian Movement (VAPM; Johansson et al., 2020) study by expanding the analysis to include the microscopic movements of both pedestrians and cyclists in outdoor environments under different lighting conditions. Additionally, the studies included in this thesis introduced a progressive data collection method that uses drone technology. This method was first implemented in the study from *Paper I*, where initial observations demonstrated its effectiveness in capturing valuable insights into how cyclists and pedestrians navigate their environments, using data from both individual passages and interactions. In this context, individual passage refers to the trajectory followed by a single road user, capturing their movement and path through the traffic environment. Interactions describe the dynamic interplays between at least two road users, in which each adjusts their intended movement in response to the other, such as when they need to yield or adjust their speed and/or position to ensure smoother travel.

The method was further applied in the large-scale, comprehensive analysis presented in *Paper II.* This study expanded the scope of the initial research, incorporating a broader range of data and examining more diverse settings and environmental conditions. This large-scale analysis offered critical insights into how different interventions on shared-use paths (e.g. road surfacing, facility layout) impact safety and comfort across different environments and lighting scenarios.

3.3.1 Study design

This study used drones as the primary tool for collecting the data used for further analysis and investigation. These data were gathered from two shared-use paths that accommodate both cyclists and pedestrians, located in the cities of Linköping and Lund in Sweden (see Figure 18). These paths were situated between the outskirts and the city centres, commonly serving as walking and cycling commuting routes. The observed sections, captured by drone cameras positioned at a height of 120 metres, spanned approximately 150 metres in length.

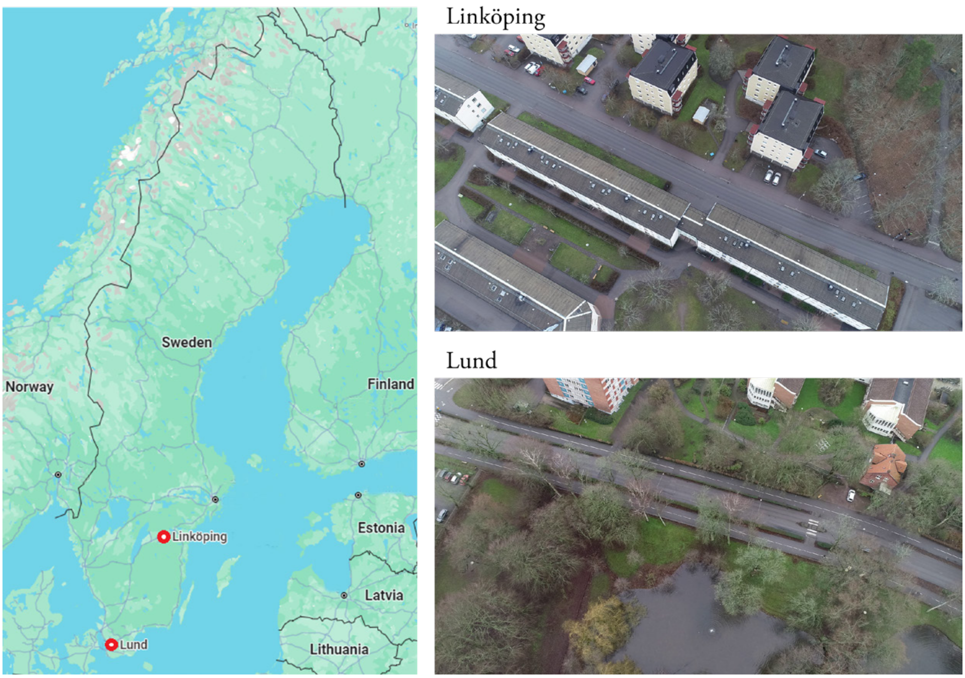

Figure 18. Study sites

In Linköping, cycling was allowed in both directions along the path, unlike in Lund, where the path was designed with a one-way cycling lane alongside a separate footpath for pedestrians. Data were collected over five distinct periods during the months with reduced daylight in Sweden (October–March), the details of which are provided in Table 1. Video data were captured under two lighting conditions: daylight and electric lighting.

Table 1. Overview of the data collection process (adapted from Paper II)

Location	Interventions in the physical environment of the path	Data collection period	Lighting condition	Time
Linköping	**Baseline**	November 2019; March 2020	Daylight	07.00–10.00
			Electric lighting	17.00–20.00
	Intervention 1: *wider path, separation line, new surfacing*	October 2021	Daylight	14.00–16.00
			Electric lighting	17.30–20.00
	Intervention 2: *electric lighting application*	November 2022	Daylight	14.00–17.00
			Electric lighting	18.45–21.00
Lund	**Baseline**	January–February 2020	Daylight	14.00–16.00
			Electric lighting	18.30–21.00
	Intervention: *trimmed greenery, new surfacing, electric lighting application*	February–March 2022	Daylight	13.00–16.00
			Electric lighting	17.30–20.00

Between data collection periods, several changes were made to the paths' physical environments and electric lighting applications. Initially, the path in Linköping (see Figure 19), with its asphalt surface and lighting from high-pressure sodium lamps, did not separate cyclists and pedestrians. The first intervention involved resurfacing and widening the path, as well as adding a white line to separate the two groups. In the second intervention, the path's surface stayed the same, but the electric lighting was changed. The original lampposts were fitted with new LED light sources.

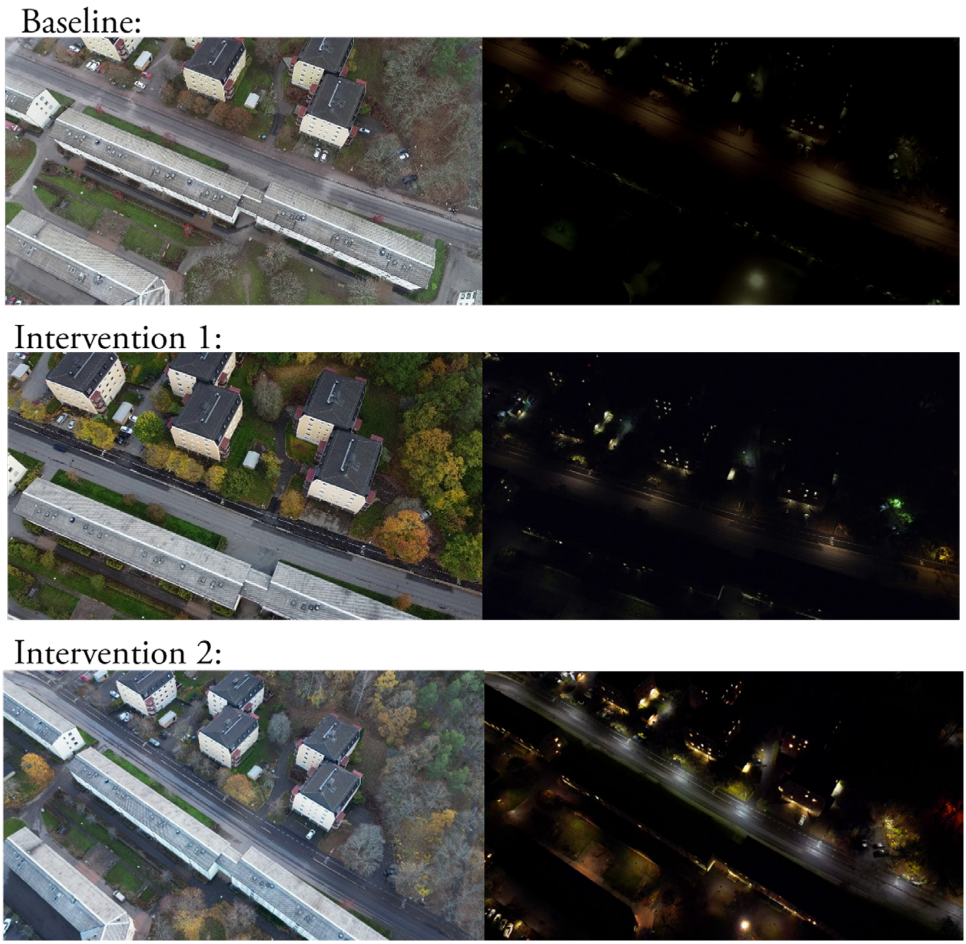

Figure 19. Drone images of the study site captured during each data collection period under both daylight (left) and electric lighting conditions (right) in Linköping (adopted from Paper II)

In Lund (see Figure 20), the path was initially segregated with a white line dividing the cycling and pedestrian areas, and it was one-directional for cyclists. The surface was asphalt, and the lighting was provided by high-pressure sodium lamps. The intervention involved updating both the surface and lighting. The path was resurfaced with new asphalt, keeping the separation line between cyclists and pedestrians, and the greenery was trimmed to improve visibility. New LED light sources were installed, and a luminaire was added to bring the lighting closer to the edge of the path.

Figure 20. Drone images of the study site captured during each data collection period under both daylight (left) and electric lighting conditions (right) in Lund (adopted from Paper II)

Two types of empirical data were collected via video recording: 1) the movements of invited participants—who were aware of the study and either cycling or walking along the paths—aware sample, and 2) the movements of undisturbed traffic—random cyclists and pedestrians in real-life conditions—unaware sample. For the aware sample, students aged 18–35 from Linköping and Lund universities travelled individually along a designated path and completed a questionnaire on their perception of the environment before and after their journey. All participants provided written consent and were briefed on the study's objectives and ethical guidelines.

3.3.1.1 Video processing and microscopic data extraction

Before generating the trajectories of cyclists and pedestrians, each video was stabilised and adjusted for brightness when necessary. Calibration was then performed using T-Calibration software (T-Analyst, 2019) with known reference points measured using a Leica Builder R200M total station.

After processing, the trajectories of path users were generated using T-Analyst (T-Analyst, 2019). Due to the impact of individual road users' speed choices on the number of data points in each unique trajectory, all extracted trajectories were standardised by cropping them at the same start and finish points and normalised using equally spaced reference points. The lateral placement of each road user was recalculated relative to the path's middle line, with positive values indicating the right-hand side and negative values indicating the left-hand side in the direction of travel.

For a more comprehensive analysis of individual cyclist behaviour, key microscopic variables were used, including speed (average and median), average lateral position, and slalom manoeuvres. Since the study sites (shared-use paths) were selected based on a specific criterion—that they were frequently used by commuters traveling between the city outskirts and the city centre, either by foot or by bike—it was assumed that a decrease in the slalom measure indicated a perceived improvement in the pathway for these commuters, suggesting that the design changes had improved their experience. Additionally, the *Safe Lateral Passing Distance* (see Figure 17) was applied to examine cyclist interactions.

3.3.2 Analysis

The study aimed to investigate several key aspects:

1. Differences in movements between aware and unaware samples of path users

To compare the movements of invited participants with those of unaware pedestrians and cyclists, aggregated trajectories were analysed, focusing on average speed and lateral position. Data were collected during baseline periods at both sites under daylight and electric lighting. Interactions with passers-by were excluded, resulting in 146 trajectories per location, which were evenly split between invited and unaware users. Differences between the groups were assessed using a two-tailed t-test.

2. Impact of outdoor lighting conditions on cyclists' path use and positioning during interactions with other cyclists

The analysis focused on 60 bi-directional interactions of the unaware sample at the Linköping site (half in daylight, half under electric lighting). Using the *Safe Lateral Passing Distance* (see Figure 17), it was assessed the point at which cyclists achieved and maintained this space, as well as the longitudinal distance and time to reach the meeting point after establishing the safe distance.

3. Impact of surfacing and lighting modifications on cyclist and pedestrian behaviour

I analysed 353 individual trajectories from the aware sample across all data collection periods to assess how path surfacing and lighting modifications impacted behaviour. The analysis focused on speed and slalom patterns in response to changes in environmental features, with data divided by user mode (cyclists or pedestrians) and location (Linköping or Lund). I examined continuous variables (median speed and cumulative slalom) alongside categorical variables, including surfacing, layout, and lighting conditions (daylight, electric lighting). A two-way ANOVA (Bevans, 2023; Wilcox, 2003) was used to assess the impact of environmental features and lighting on speed and slalom, with a post hoc Tukey's HSD test (Tukey, 1949) to identify significant pairwise differences in speed and slalom among participant groups.

3.3.3 Results

3.3.3.1 Differences between the groups of aware and unaware samples

The results revealed significant differences in movement patterns, especially among cyclists. At both study locations, cyclists in the unaware sample from the baseline data collection period travelled faster and showed greater speed variation than those in the aware group from the same period (see Table 2), with these differences persisting across all lighting conditions. However, differences in cyclists' lateral positions between the two cyclists' samples were only statistically significant during daylight in Linköping, where the path was an unsegregated shared-use path.

Table 2. Cyclists' speed and lateral position profiles during baseline data collection period (adapted from Paper II)

Location	Microscopic behavioural variable	Lighting conditions	Sample	Mean (m/s)	SD	*t*	*P*
Linköping	Speed	Daylight	Aware	4.36	0.54	-2.65	0.01
			Unaware	4.99	0.83		
		Electric lighting	Aware	3.72	0.34	-4.08	<0.01
			Unaware	4.44	0.66		
	Lateral position	Daylight	Aware	0.96	0.36	3.27	<0.01
			Unaware	0.44	0.54		
		Electric lighting	Aware	0.57	0.29	1.67	0.10
			Unaware	0.42	0.26		
Lund	Speed	Daylight	Aware	4.90	0.70	2.04	0.04
			Unaware	5.48	0.94		
		Electric lighting	Aware	4.32	0.93	-2.63	0.01
			Unaware	5.15	1.06		
	Lateral position	Daylight	Aware	0.02	0.26	0.76	0.46
			Unaware	-0.07	0.31		
		Electric lighting	Aware	-0.83	0.36	1.91	0.07
			Unaware	-0.63	0.24		

For pedestrians (see Table 3), speed trends were not statistically significant, but a notable difference in their lateral position on the path was observed in Linköping during daylight hours, similar to the cyclists.

Table 3. Pedestrians' speed and lateral position profiles during baseline data collection period (adapted from Paper II)

Location	Microscopic behavioural variable	Lighting conditions	Sample	Mean (m/s)	SD	*t*	*P*
Linköping	Speed	Daylight	Aware	1.47	0.10	-1.39	0.17
			Unaware	1.53	0.16		
		Electric lighting	Aware	1.72	0.14	1.32	0.20
			Unaware	1.65	0.17		
	Lateral position	Daylight	Aware	0.96	0.30	2.82	0.01
			Unaware	0.58	0.48		
		Electric lighting	Aware	1.24	0.17	1.89	0.07
			Unaware	1.05	0.39		
Lund	Speed	Daylight	Aware	1.54	0.18	0.29	0.77
			Unaware	1.52	0.17		
		Electric lighting	Aware	1.62	0.13	0.25	0.81
			Unaware	1.61	0.19		
	Lateral position	Daylight	Aware	0.29	0.27	0.39	0.70
			Unaware	0.32	0.22		
		Electric lighting	Aware	0.59	0.41	1.02	0.31
			Unaware	0.74	0.47		

3.3.3.2 Bi-directional cyclists' interactions from unaware sample

The results show that bi-directional interactions among the unaware cyclists differed between daylight and after dark. Cyclists initiated interactions earlier in daylight (see Figure 21). For example, at a longitudinal distance of 20 meters, nearly 90% of pairs during daylight had achieved the *Safe Lateral Passing Distance* of 1.5 meters, while only 75% had done so under electric lighting.

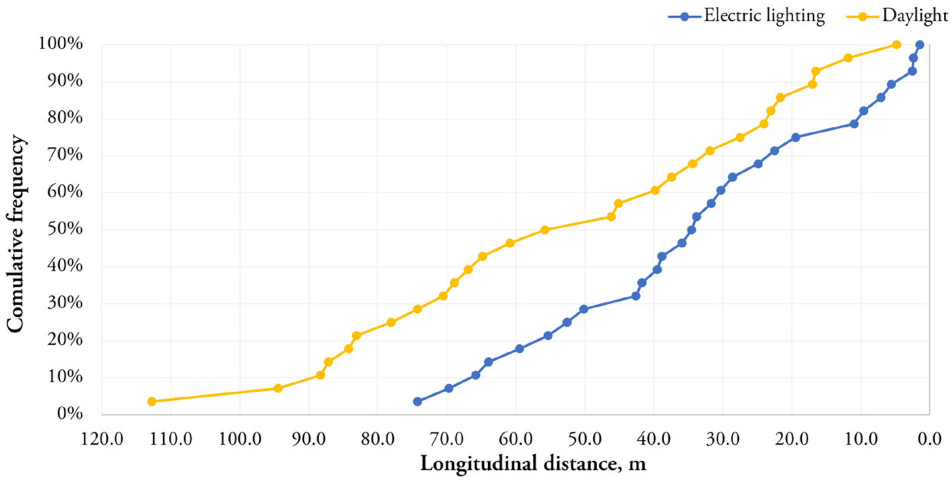

Figure 21. Longitudinal distance between cyclists involved in an interaction after they have established the Safe Lateral Passing Distance (1.5 meters) (adapted from Paper I)

Time perspective—time taken to reach the meeting point—provides additional insight into the dynamics of bi-directional interactions among cyclists, as it accounts for the fact that cyclists typically move at different speeds. The analysis examined the time remaining until cyclists passed each other after reaching the *Safe Lateral Passing Distance* (see Figure 22). The results show that, in daylight, cyclists reached this distance approximately 1.4 seconds earlier than those under electric lighting.

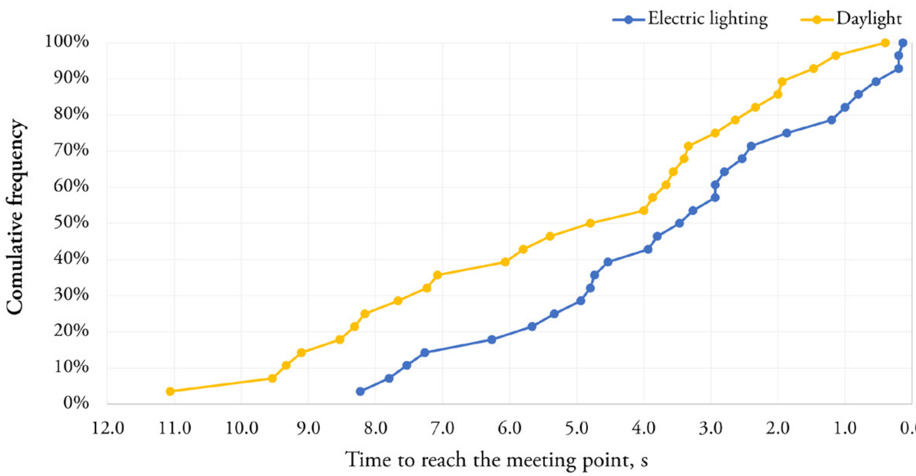

Figure 21. Time taken to reach the meeting point after cyclists have established the Safe Lateral Passing Distance (1.5 meters) (adapted from Paper I)

To ensure that these results are not influenced by differences in the cyclists' free-flow positioning before interactions began, I analysed their lateral positions prior to the onset of interaction. The analysis revealed that daytime cyclists tended to ride closer to the middle of the path during free flow conditions (ranging from -0.7 m to 1 m, where positive values correspond to the roadside and 0 m represents the middle of the path). In contrast, nighttime cyclists were shifted towards the roadside, where the path was better lit, with a narrower distribution of positions (ranging from 0.4 m to 0.8 m). Although the initial positioning of cyclists differed between daylight and nighttime conditions, the pattern of lateral distances between interacting cyclists was similar under both lighting conditions (see Figure 16). This indicates that the observed differences in the maintenance of *Safe Lateral Passing Distances* during bi-directional interactions between the two groups of cyclists (those experiencing the path in daylight and those under electric lighting) are not influenced by variations in their free-flow positioning.

3.3.3.3 Impact of surfacing and lighting modifications on cyclist and pedestrian behaviour

Study findings suggest that changes in path surfacing and electric lighting lead to notable behavioural changes for both pedestrians and cyclists. At the Linköping site, new surfacing, path widening, and segregation (in the form of a painted white line) increased bicycle speed and reduced slalom during the daytime (see Table 4).

Table 4. Tukey HSD test results for cyclists' movements in Linköping (adapted from Paper II)

Group1	Group2	Physical changes	*DV*	Mean diff (G2-G1)	95% CI	*p*	Change (%)
Baseline (Daylight)	Baseline (EL)	Daylight to EL	Speed	-0.24	(-0.34, -0.17)	.01	-9.03
			Slalom	-2.74	(-4.78, -0.74)	.05	-18.77
Baseline (Daylight)	Intervention1 (Daylight)	New surfacing + separation	Speed	0.54	(0.20, 0.88)	.02	+13.57
			Slalom	-3.56	(-7.00, -0.14)	.03	-24.38
Baseline (EL)	Intervention1 (EL)	New surfacing + separation	Speed	-0.02	(-0.15, 0.12)	.90	-
			Slalom	-0.73	(-5.83, 4.37)	.76	-
Intervention1 (Daylight)	Intervention1 (EL)	Daylight to EL	Speed	-0.80	(-1.22, -0.38)	<.01	-17.69
			Slalom	0.09	(-3.44, 3.62)	.74	-
Intervention1 (Daylight)	Intervention2 (Daylight)	No changes	Speed	-0.10	(-0.39, 0.19)	.90	-
			Slalom	0.03	(-2.87, 2.94)	.90	-
Intervention1 (EL)	Intervention2 (EL)	EL application	Speed	0.61	(0.21, 1.04)	.01	+16.40
			Slalom	-2.79	(-4.53, -1.08)	.05	-25.06
Intervention2 (Daylight)	Intervention2 (EL)	Daylight to EL	Speed	-0.09	(-0.27, 0.09)	.90	-
			Slalom	-2.73	(-3.73, -1.79)	.05	-24.66
Baseline (Daylight)	Intervention2 (Daylight)	New surfacing + separation	Speed	0.44	(0.07, 0.82)	.05	+11.06
			Slalom	-3.53	(-5.66, -1.40)	.03	-24.18
Baseline (EL)	Intervention2 (EL)	New surfacing, separation & EL application	Speed	0.59	(0.17, 1.02)	.01	+15.77
			Slalom	-3.52	(-7.69, -0.39)	.01	-29.67

Changes to the path surface alone did not significantly impact the behaviour of cyclists (see Table 4) or pedestrians (see Table 5) after dark. Despite the modifications, both groups maintained their typical movement patterns. The darker surfacing introduced after the first intervention, with poorer reflective properties, may have reduced visibility, potentially influencing cyclists to maintain a lower speed compared to daytime conditions. This suggests that surface alterations alone, in the absence of improved lighting, may not substantially affect behaviour, highlighting consistency or adaptability in how road users navigate their environment.

Table 5. Tukey HSD test results for pedestrians' movements in Linköping (adapted from Paper II)

Group1	Group2	Physical changes	*DV*	Mean diff (G2-G1)	95% CI	*p*	Change (%)
Baseline (Daylight)	Baseline (EL)	Daylight to EL	Speed	0.24	(0.07, 0.42)	<.01	+16.67
			Slalom	-1.87	(-3.52, -0.22)	.05	-22.50
Baseline (Daylight)	Intervention1 (Daylight)	New surfacing + separation	Speed	0.03	(-0.14, 0.09)	.90	-
			Slalom	-0.40	(-12.89, 11.49)	.90	-
Baseline (EL)	Intervention1 (EL)	New surfacing + separation	Speed	0.02	(-0.07, 0.11)	.90	-
			Slalom	0.38	(-9.61, 11.86)	.06	-
Intervention1 (Daylight)	Intervention1 (EL)	Daylight to EL	Speed	0.23	(0.07, 0.39)	< 01	+15.65
			Slalom	-1.09	(-1.59, -0.60)	.05	-13.78
Intervention1 (Daylight)	Intervention2 (Daylight)	No changes	Speed	0.01	(-0.12, 0.10)	.90	-
			Slalom	-0.12	(-11.87, 10.23)	.90	-
Intervention1 (EL)	Intervention2 (EL)	EL application	Speed	-0.36	(-0.61, -0.11)	.04	-21.18
			Slalom	-1.45	(-1.92, -1.00)	<.01	-21.26
Intervention2 (Daylight)	Intervention2 (EL)	Daylight to EL	Speed	-0.09	(-0.21, 0.22)	.10	-
			Slalom	-0.42	(-2.50, 1.50)	.90	-
Baseline (Daylight)	Intervention2 (Daylight)	New surfacing + separation	Speed	0.04	(-0.13, 0.09)	.90	-
			Slalom	-0.52	(-10.45, 10.69)	.90	-
Baseline (EL)	Intervention2 (EL)	New surfacing, separation & EL application	Speed	-0.36	(-0.53, -0.21)	.03	-21.43
			Slalom	-1.07	(-1.23, -0.90)	.05	-16.61

The new electric lighting in Linköping (implemented in Intervention 2) had a notable impact on road users' dynamics. Cyclists' speed increased, and pedestrian speed decreased under the new lighting, both aligning more closely with trends observed in daylight. This shift highlights the significant role of lighting in influencing movement and the perception of environmental conditions.

In Lund, the new surfacing positively influenced cyclists' behaviour (see Table 6), leading to increased speed and reduced deviations during daylight, likely due to cyclists perceiving the improved surface as safer. However, no statistically significant difference was observed in pedestrian behaviour (see Table 7) under the different interventions during daylight.

Table 6. Tukey HSD test results for cyclists' movements in Lund (adapted from Paper II)

Group1	Group2	Physical changes	*DV*	Mean diff (G2–G1)	95% CI	*p*	Change (%)
Baseline (Daylight)	Baseline (EL)	Daylight to EL	Speed	-0.42	(-0.78, -0.07)	.05	-10.05
			Slalom	0.53	(-2.06, 3.32)	.39	-
Baseline (Daylight)	Intervention (Daylight)	Trimmed greenery, new surfacing	Speed	0.54	(0.07, 1.02)	.04	+12.92
			Slalom	-3.43	(-6.35, -0.49)	<.01	-21.68
Baseline (EL)	Intervention (EL)	Trimmed greenery, new surfacing, EL application	Speed	0.83	(0.07, 1.58)	.02	+22.07
			Slalom	-3.71	(-6.65, -0.77)	.03	-22.69
Intervention (Daylight)	Intervention (EL)	Daylight to EL	Speed	-0.13	(-0.59, 0.90)	.90	-
			Slalom	0.25	(-4.16, 5.34)	.90	-

It was not possible to independently analyse the effects of surfacing or lighting under electric lighting conditions in Lund, as both factors were altered simultaneously. The results indicate that the combination of surfacing changes and new lighting led to an increase in cyclists' speed and a decrease in slalom, while pedestrians experienced a reduction in speed and slalom. These findings align with the results observed during the latest intervention in Linköping.

When comparing movements under daylight and electric lighting conditions in Lund, the analysis revealed that, similar to Linköping, the differences in cyclists' speed and slalom (see Table 6) under these two lighting conditions were not statistically significant. However, for pedestrians (see Table 7), there was an increase in speed under electric lighting conditions, while deviations followed trends similar to those observed under daylight.

Table 7. Tukey HSD test results for pedestrians' movements in Lund (adapted from Paper II)

Group1	Group2	Physical changes	*DV*	Mean diff (G2–G1)	95% CI	*p*	Change (%)
Baseline (Daylight)	Baseline (EL)	Daylight to EL	Speed	0.18	(0.06, 0.31)	<.01	+22.07
			Slalom	-0.88	(-3.82, 3.68)	.06	-
Baseline (Daylight)	Intervention (Daylight)	Trimmed greenery, new surfacing	Speed	-0.07	(-0.12, 0.09)	.90	-
			Slalom	-0.58	(-3.41, 2.56)	.17	-
Baseline (EL)	Intervention (EL)	Trimmed greenery, new surfacing, EL application	Speed	-0.15	(-0.27, -0.04)	.05	-9.89
			Slalom	-2.15	(-4.18, -0.15)	<.01	-30.06
Intervention (Daylight)	Intervention (EL)	Daylight to EL	Speed	0.19	(0.04, 0.33)	.01	+13.77
			Slalom	0.15	(-0.70, 1.02)	.08	-

3.3.4 Limitations and additional considerations for future research

The results from these practical case studies highlight the potential efficacy of using drones for data collection, significantly streamlining the process in comparison to previously proposed methods (e.g. the VAPM method proposed by Johansson et al. (2020)). In the Swedish context, drones offer a particular advantage by navigating regulatory requirements related to video data collection in public spaces, ensuring compliance with personal data protection laws. However, to maintain individuals' anonymity, drones must be elevated to a sufficient height, which depends on camera specifications (in our case, approximately 115–118 meters). At this altitude, controlling for confounding factors becomes more challenging, especially during low-light conditions. These factors include identifying the type of road user (e.g. commuters versus recreational cyclists), distinguishing between e-bikes and regular bicycles (as they have similar size and shape), and capturing socioeconomic characteristics like the age and gender of the road users.

Despite the opportunities presented by drone technology for observing events in real-world conditions, several challenges arose during the data collection process. First, adverse weather conditions, such as rain and strong winds, disrupted some of the planned recording activities. Second, mandatory coordination with relevant authorities was required to obtain flight permissions at study sites near strategic locations, such as hospitals or airports (as was the case for the Linköping site). On several occasions, these permissions were denied. Since many of the planned recordings were time-sensitive (with invited participants), rescheduling was not feasible. As a result, these factors significantly reduced the size of the dataset of recorded participant movements (the aware group sample size) available for analysis.

Another critical issue was the study sample size and its implications. Sample size directly affects a study's statistical power, generalisability, variability, and precision. In the Lund and Linköping studies, the sample size was somewhat limited. However, the primary aim of these studies was to develop a method that could be effectively applied for traffic behavioural studies and test its practicality in real-world scenarios rather than to draw broad conclusions representative of the larger population. Increasing and diversifying the sample size would provide significant benefits, refining the ability to generalise the findings. This improvement could be explored in future work as a continuation of these studies.

Additionally, the potential influence of informing experimental participants about the study's aims prior to testing warrants consideration. While investigating the impact of lighting and surfacing on cyclists' and pedestrians' behaviour, we used trajectories obtained from an aware sample (experimental study participants). Comparisons with an unaware sample (undisturbed traffic) revealed differences in speed and lateral positioning despite the aware participants having been instructed to behave naturally. However, since our focus was on behavioural trends rather than precise absolute values, this dataset was deemed suitable for analysis as it reflects movements within a consistent experimental setup. Nevertheless, conducting a similar analysis using naturalistic movements would be beneficial to determine whether comparable behavioural trends emerge under the same physical interventions and to gain additional insights through absolute values.

By applying the developed method, we captured behavioural changes in pedestrians and cyclists, especially among the cyclists. However, pedestrian behaviour often exhibits a higher degree of variability compared to cyclist behaviour, which poses challenges for the accurate measurement of their actions using only quantitative data extracted from video recordings. Unlike cyclists, who typically adhere to designated paths or lanes and follow more structured movement patterns, pedestrians demonstrate greater spontaneity and adaptability in their navigation. This variability stems from factors such as personal preferences, social interactions, and immediate environmental stimuli, making pedestrian movements appear more unpredictable (Schneekloth & Shibley, 1995). Advancing measurement methods by incorporating qualitative data could improve the accuracy and comprehensiveness of pedestrian behavioural analysis in future studies.

In line with this, we have already begun working on such integration, particularly through the research project '*Belysa—miljöfaktorers inverkan på fotgängares och cyklisters beteende i skymning och mörker*' (grant number: 2019-002011), which provided the foundation for the studies mentioned above. In this project, we collaborated with researchers from the Department of Architecture and Built Environment at Lund University, who collected questionnaires during all data collection periods from the same invited participants (aware sample). These questionnaires focused on the participants' experiences of the physical environmental interventions along the shared walking and cycling paths. The questionnaire aimed to explore associations between perceived environmental stimuli and changes in upkeep, surface, and lighting qualities, as well as, to a smaller extent, conceptual environmental appraisals. These assessments were gathered from the participants using the conceptual model introduced by Rahm et al. (2024). The next step in this research is to integrate these qualitative insights with the microscopic behavioural data to form a key focus for subsequent studies.

4 Microscopic behavioural data for enhancing traffic safety: Application of Surrogate Measures of Safety (SMoSs)

4.1 The evolution of approaches to traffic safety management

Approaches to traffic safety management differ in scope, objectives, and methodologies. Traditionally, traffic safety management has been reactive, relying mostly on accident statistics. It has thus primarily focused on mitigating risks identified from past accidents rather than on proactively addressing potential safety issues before they occur (Hauer, 1997).

The limitations of using crash data alone are well-documented in the literature. Key issues include under-reporting, as many non-fatal crashes and near-misses never make it into official statistics (especially with VRUs involved), thus underestimating risk levels and distorting our understanding of accident causation (Kamaluddin et al., 2019; Watson et al., 2015; Aarts & van Schagen, 2006). Additionally, on a smaller scale, such as at specific intersections or road segments, the low frequency of accidents makes it difficult to conduct reliable statistical analyses. This issue is compounded by the rare nature of severe accidents, which poses a challenge to the development of predictive models (Saunier & Laureshyn, 2021; Laureshyn et al., 2010; Lord & Mannering, 2010). Additionally, accident reports often lack the necessary detail for thorough analysis. Key variables such as precise vehicle manoeuvres, driver behaviour prior to the accident, and road conditions are frequently not adequately documented, limiting the depth of safety analyses (Elvik, 2010; Elvik et al., 2009).

With the growing recognition of the limitations of this reactive approach, the Safe System framework has emerged as the global best practice in traffic safety management.

This framework advocates for a more holistic, proactive approach, focusing not only on responding to accidents but also on designing systems that prevent crashes and minimise their severity. This model has been successfully implemented in several countries, including through Vision Zero in Sweden (Edvardsson Björnberg et al., 2022; Johansson, 2009) and Sustainable Safety in the Netherlands (Wegman et al., 2022; Wegman et al., 2005). It stands out because of its ethical focus on protecting human life, its strong scientific basis, and its proven success in reducing road accidents that lead to serious injuries (with remaining or permanent impairment) and fatalities (Elvik, 2023; Elvik & Nævestad, 2023; Edvardsson Björnberg et al., 2022).

A core principle of the Safe System is its acknowledgement that people will always make mistakes, sometimes intentionally and sometimes without realising it. The Safe System approach therefore designs transportation systems to account for human error, preventing mistakes from leading to serious injuries or fatalities. Instead of focusing solely on the individual level, the Safe System approach views road safety as a responsibility that is shared by not only road users but also road designers, vehicle manufacturers, and policymakers.

Another fundamental concept of the Safe System (see Figure 23) is that the human body has physical limits to the forces it can withstand during a crash. Road infrastructure, vehicle design, and speed limits are planned with these biomechanical tolerances in mind, ensuring they minimise risks to people.

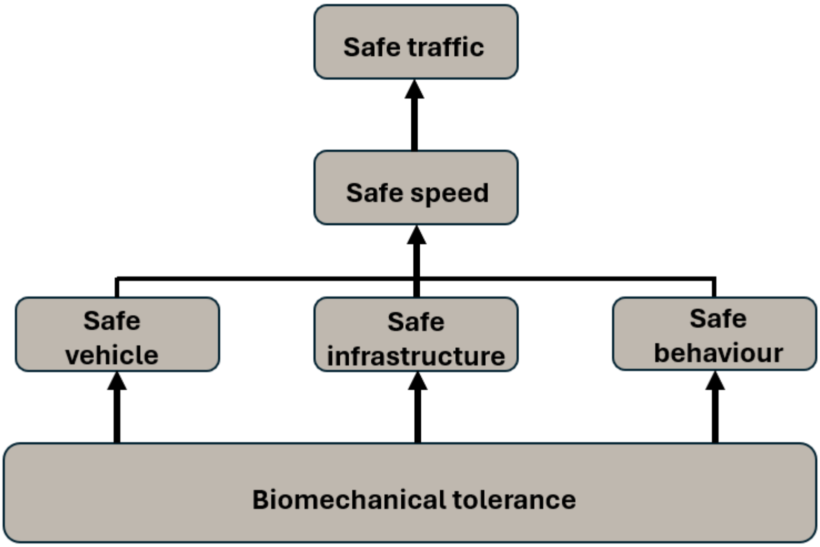

Figure 23. Diagram illustrating the Safe System approach to road safety (Ekman, 2023)

As shown in Figure 23, the Safe System approach is built around four interconnected components: safe speeds, safe vehicles, safe infrastructure, and safe road user behaviours. For a traffic system to be safe, each of these areas must be addressed holistically.

Safe speed is a core component, and speed limits need to be set appropriately for each traffic environment. In areas with high pedestrian traffic, for example, lower limits are necessary to reduce the risks and impacts of collisions.

Equally important is the design of safe infrastructure; road features are carefully selected to mitigate risks while ensuring cost-effectiveness. A key factor in reducing collision severity is the compatibility of road users' mass and speed, as their differences in kinetic energy significantly influence the outcome of accidents (Harms, 1993). Smaller differences in mass and speed indicate higher compatibility. To reduce incompatibility, one effective strategy is to separate road users with differing masses and speeds in time or space (Elvik, 2004). Measures such as protective barriers, dedicated spaces for VRUs, roundabouts, and traffic-calming features all play a crucial role in creating safer road environments for everyone.

Vehicles themselves also play a critical role in road safety. Crash risks are significantly reduced with the promotion of vehicles equipped with advanced safety features, such as passive safety protection. While many other 'safety' features, such as anti-lock braking systems (ABS), lane-keeping assistance, and Intelligent Speed Assistance (ISA), among others, are commonly used, their effectiveness in reducing accidents is often questioned due to drivers' behavioural adaptations. For example, vehicles with ABS technology are not in fewer accidents than vehicles without it (Elvik, 2004), but ABS provides greater steering control, potentially mitigating the consequences of collisions (Evans, 1999). Additional in-vehicle technologies, like alcohol interlock and speed-limiting devices, offer even greater safety potential, though public acceptance may initially vary.

Safe road user behaviour is necessary for reducing accidents and injuries. Encouraging safe behaviours relies on promoting positive social norms and strictly enforcing traffic laws (for example, those mandating seatbelt use, child safety seats, and adherence to speed limits). Combined, these measures create an environment in which safer driving practices become the norm, reinforcing the Safe System's goal of a well-rounded, effective approach to road safety.

As traffic systems grow increasingly complex, a comprehensive understanding of safety risks becomes essential for improving safety outcomes. With the rise of new technologies and data analytics, traffic safety management is shifting towards more data-driven safety analysis approaches. In this context, video recordings, and more specifically, the microscopic behavioural data extracted from video footage, play an important role. By capturing detailed, real-time interactions between road users and infrastructure, video data allow for a deeper analysis of microscopic traffic behaviour.

In this context, Surrogate Measures of Safety (SMoSs) have been proposed as a valuable alternative to traditional accident-based data, providing valuable data sources beyond traditional accident reports. SMoSs focus on observable non-crash traffic events, particularly near-misses, which encompass traffic conflicts of varying severity, to identify potential safety risks before they result in actual collisions (Laureshyn et al., 2010; Tarko et al., 2009). This approach also offers deeper insights into crash mechanisms and contributing factors to inform future accident prevention strategies, thus supporting a more proactive approach to safety (Johnsson, 2020; Svensson, 1992; Hydén, 1987). However, very little work has been done on integrating aspects of the Safe System approach into SMoS. In the following sections (and in *Paper III*), I will outline how such an integration could be implemented.

4.2 Practical framework for safety analysis using SMoSs

This section is largely inspired by Patterson (2021) doctoral thesis, which delves into how traffic safety research often becomes entrenched in outdated paradigms, hindering progress in the field. The author argues that sticking to one paradigm results in fundamental assumptions becoming accepted as unquestionable truths that are rarely challenged or reconsidered, even when they become outdated and are no longer optimal for the present circumstances. Efforts to refine or correct this paradigm often introduce unnecessary complexity and frequently create new challenges.

According to Patterson (2021), the field of SMoS may represent one such case of an outdated paradigm. Despite being a promising concept for over 60 years and generating a significant amount of research, it has not yet been established as a practical tool for use by traffic safety professionals. Validity—the ability of an indicator to accurately reflect the quality it is designed to measure—is a critical factor in this context; however, assessing the extent of SMoS validity is challenging due to the diverse methods employed (Johnsson et al., 2018). While numerous validation studies have attempted to establish the absolute validity of SMoS, to date, none can be regarded as truly convincing. Consequently, efforts to improve the field tend to follow two main paths: to capture conflict severity by proposing new (often more complex) indicators and to develop improved methods for linking observed critical events with predicted accident frequencies.

4.2.1 SMoS: Key concepts

The conceptual foundation of SMoSs rests on the idea that traffic incidents can be categorised into a hierarchy based on the severity of each event. This model, introduced by Hydén (1987) and known as the 'safety pyramid'(see Figure 24), illustrates that in the hierarchy of traffic events, the severity increases from the bottom to the top, while the frequency decreases. The bottom part of the pyramid represents normal interactions between road users, which occur most frequently. In contrast, the most severe traffic events, such as injurious or fatal accidents, are positioned at the pyramid's top, reflecting their rarity in proportion to the total number of events. Viewed as a whole, these events form a continuum that objectively describes the relationship between their frequency and severity (Svensson & Hydén, 2006; Svensson, 1998).

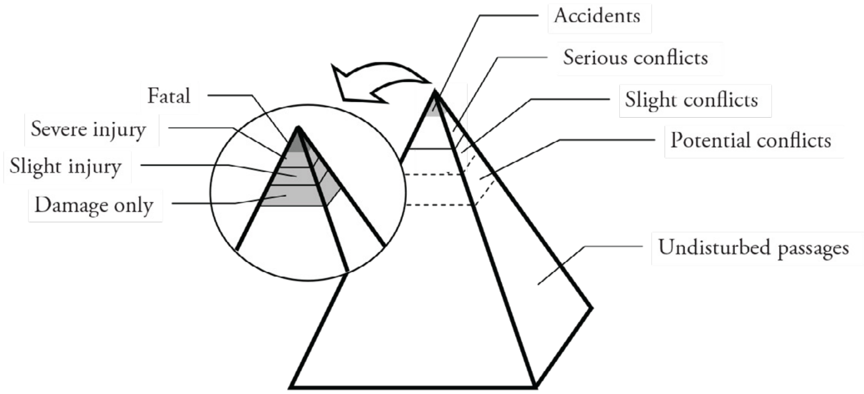

Figure 24. 'Safety pyramid' model (adopted from Laureshyn (2010) based on Hydén (1987))

The frequency distribution across different severity levels offers valuable insights for further exploration. Hydén (1987) research found that serious conflicts occur immediately after accidents in the hierarchy of traffic events, meaning they are observed more frequently than accidents. Additionally, there is a strong statistical correlation between the number of serious conflicts and the number of police-reported injuries. Consequently, the frequency of rare but serious incidents (such as accidents and personal injuries) can be inferred from the frequency of less severe but more commonly observed situations, known as traffic conflict situations (El-Basyouny & Sayed, 2013; Brown, 1994; Hydén, 1987).

The distribution of traffic event frequency at the lower end of the hierarchy, where events have a smaller severity scale, is not fully understood (Laureshyn, 2010). One hypothesis suggests that a high frequency of such events indicates a probability of failure that could lead to unsafe situations; however, the low severity scale implies that this probability is relatively low. Svensson (1998) further posits that these events might make a positive contribution to the overall traffic system's performance by providing road users with the experience and skills to handle more serious traffic situations, thus helping them avoid severe accidents.

4.2.2 Relation between observed critical events and crashes

Although the concept of 'conflict' in traffic safety research was first introduced by Perkins & Harris (1967), with all subsequent studies in the field deriving from their foundational work, the place of 'conflicts' in the traffic process has long been debated. Güttinger (1982) proposed two possible models. In the first model, a conflict differs from normal, undisturbed traffic because of the presence of evasive action—an action taken by a road user (e.g. altering speed or direction) to avoid a potential collision. If the evasive action is unsuccessful or does not occur at all, the conflict results in a collision (see Figure 25 (a)). In this model, a conflict is seen as a potential collision. To be able to detect and measure the severity of a conflict, the indicators used should be related to the conflict's initial phase.

In contrast, the second model distinguishes conflicts from normal traffic only in terms of the 'successes' of evasive action (see Figure 25 (b)); if it was a successful evasive action, then the situation falls within the 'normal traffic' category. In this view, a conflict is the same as a near-miss situation and cannot lead to a collision. Instead, it is considered an event parallel to (separate from) a collision. A significant theoretical drawback of this model is its categorisation of conflicts and accidents into two mutually exclusive causal pathways. This raises concerns about whether accident risk can be accurately represented by conflicts defined in this way.

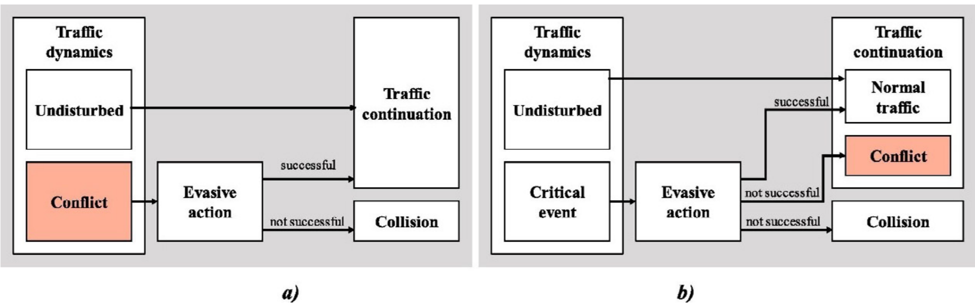

Figure 25. Conflict's place in the traffic process (adopted from Johnsson (2020) based on Güttinger (1982))

The first model has clear theoretical advantages, as it positions conflicts and accidents within the same causal chain, thereby reinforcing their mutual connection and supporting the use of conflicts as proxies for accidents (for example, the Swedish Traffic Conflict Technique follows this model's mechanism (Laureshyn & Várhelyi, 2018; Hydén, 1987)). However, despite these advantages, the second model is more frequently applied in practice. For example, indicators such as the minimum value of Time-to-Collision (TTC_{min}; Hayward, 1972) and Post-encroachment Time (PET; Allen et al., 1978) are more commonly used than Time-to-Accident (TA; Hyden, 1977), which measures TTC at the onset of evasive action. This preference is largely due to practical considerations, as these indicators are easier to operationally define and measure.

Various approaches have been employed to convert conflicts into accident probabilities or scenarios. Most traffic conflict techniques—methods for estimating traffic safety based on observed traffic conflicts—focus on predicting accidents based on conflict data, often expressed as an accident/conflict ratio. However, complications arise from the lack of universal conditions, such as the types of road users involved and their specific manoeuvres. Non-linear statistical models have been proposed in response (Tarko, 2018). Although these advanced methods offer a better fit for the data, they struggle to explain the mechanisms linking conflicts to crashes effectively.

Rather than seeking such a ratio, Davis et al. (2011) proposed a causal model of traffic conflicts and crashes in which a traffic conflict is defined as a situation with a non-zero probability of a collision (traffic conflict and a crash are two alternative outcomes of the same process), based on a detailed description of the traffic environment through corresponding distributions. According to the model, outcomes depend on both the initial conditions (measured according to the proximity of the involved road users in time or space) and the potential evasive actions (e.g. braking, running, swerving). While this approach is grounded in robust theoretical principles and would be ideal, its practical implementation is challenging due to the unavailability of some necessary data.

The application of the Extreme Value Theory (EVT) is another promising approach that appears to be well-suited to the SMoS context. EVT helps estimate the likelihood of rare, extreme events by analysing a large number of normal and slightly unusual occurrences. Initially, the use of EVT for crash estimation was primarily limited to univariate models, in which crashes were viewed as extremes of a single safety measure (single indicators). However, since individual conflict indicators capture only one dimension of traffic event severity (either the proximity of vehicles or potential consequences), there has been a recent shift toward developing multivariate EVT models. This method offers a solid theoretical foundation for the application of SMoS concepts and is worth further exploration.

While EVT is particularly effective for predicting natural events (e.g. extreme water levels or large-magnitude earthquakes, which follow predictable physical laws that are relatively unchanged over time), its application to traffic accidents is more challenging. This difficulty arises from the unpredictable nature of human behaviour, which varies based on a range of factors, including cultural norms, infrastructure, economic conditions, and public policies.

As also previously noted by Hydén (1987), road users behave differently in dangerous situations compared to less severe traffic events. In these cases, subjective perceptions of danger play a key role. For example, humans' ability to perceive hazards improves with age and experience (Horswill, 2016), meaning the same traffic situation can lead to different reactions from different road users (and different consequences). Furthermore, factors such as size, weight and other mobility-related attributes can significantly influence the perception of risk (Geoerg et al., 2023). For VRUs, approaching a vehicle (especially a large one) feels much riskier than meeting another VRU.

Existing publications that apply EVT often focus on highly specific technical details (e.g. the review of analytical methods for modelling traffic conflicts by Zheng et al. (2021)), which can be challenging for many researchers and practitioners to comprehend. In many instances, authors rely on the 'most common' SMoS indicators and frequently combine incompatible measures without adequately exploring the implications of their results or providing clear theoretical frameworks for interpretation. This trend increases the problem of growing complexity, which obscures the underlying essence of the findings.

4.2.3 SMoS indicators and practical use

The definition of SMoS adheres to at least three key criteria for ensuring practical effectiveness rather than being merely theoretical: (i) the indicator should accurately reflect the effects of changes within the traffic system; (ii) there should be a statistical correlation between the indicator and the types of crashes influenced by these changes; and (iii) it should be feasible and useful to implement the measure in real-world settings (Tarko, 2018; Tarko et al., 2009).

Numerous traffic conflict severity indicators have been developed over the decades, which can generally be classified into several main indicator families, such as the Time-to-Collision and Post-Encroachment Time groups, along with their various modifications, deceleration-based measures, various indexes, and other features, as outlined by Laureshyn et al. (2016). Most of these indicators can be derived from road user trajectories along with additional parameters, such as road users dimensions and weight (Laureshyn, 2010). Literature reviews, such as those conducted by Johnsson et al. (2018) and Mahmud et al. (2017), have compiled extensive lists of indicators used in the SMoS context, which have undoubtedly continued to expand.

The newly proposed indicators rarely provide conceptual innovation and instead typically add layers of complexity. This is often achieved by combining multiple measures into a single index or introducing situation-specific criteria, which significantly limit their practical applicability. Johnsson et al. (2018) overview, which focused primarily on indicators relevant to VRUs and their validation in previous research concluded that only a limited number of indicators specifically address the critical aspects of VRU analysis. Furthermore, although some validation studies exist, their results are difficult to compare due to variations in methodology, observation durations, and the number of study locations involved.

The selection of surrogate indicators largely depends on their applicability. This applicability can vary based on factors such as the types of road users involved in the interaction, the specific scenarios being examined (for instance, whether there is a potential collision course), and the complexity of the interactions, which may range from simple to more complex situations involving secondary interactions.

For example, one of the most widely used indicators in the SMoS context is Time-to-Collision (TTC). However, despite its popularity, TTC suffers from complicated issues that many studies tend not to discuss. First, the calculation of TTC depends on assumptions about how road users will move, which can be subject to various interpretations. Traditionally, studies have assumed that the road user's speed and travel path will remain constant (Van der Horst; Hayward, 1972). However, this does not work for complex manoeuvres, such as when a vehicle starts from a complete stop or makes a turn at an intersection, where both speed and direction change over time. Some researchers, like Mohamed & Saunier (2013), have suggested using historical data from other road users to predict how a current road user might behave. While this can improve predictions, it also complicates the calculations and may still not work well in complex traffic situations where a road user is influenced by the previous interaction outcome. According to Johnsson (2020), safety problems are more likely to arise in complex interactions—particularly in those involving secondary interactions—than in simpler scenarios in which only two road users are present. Another significant limitation of TTC and its various modifications is that its calculation relies on the existence of a collision course. As a result, these indicators cannot be applied to situations in which road users are predicted to pass each other with even a small (no matter how small) safety margin. This restricts the ability to rank less severe situations according to their potential risk, even though these encounters remain important for assessing overall severity. While the probability of less severe interactions transitioning into unsafe events is low, it still exists (especially if some failure or unexpected event occurs). These situations differ from 'normal' encounters and should not be dismissed simply due to computational difficulties.

To address the 'presence of a collision course' limitation, several new indicators have been developed that are flexible enough to describe both collision course and non-collision course situations. For instance, indicators such as T_2 (Laureshyn et al., 2010), and Time-to-Zebra (TTZ; Varhelyi, 1998) have been proposed. The T_2 indicator, introduced by Laureshyn et al. (2010) to be later used in their study (Laureshyn et al., 2017a), measures the expected time at which the latest (second) involved road user will reach the conflict point, assuming their speeds and trajectories remain unchanged. In collision course situations, T_2 is equal to TTC; however, in non-collision course scenarios, T_2 still provides a meaningful value for analysis. This allows for a smooth transition between 'no collision course' and 'collision course' states throughout the interaction's duration. Nevertheless, the challenge of assuming 'unchanged' speeds and travel paths remains the same for T_2 as it is for TTC, particularly in complex scenarios where secondary interactions frequently occur.

Deceleration-based indicators have limited applicability, particularly when interactions involve cyclists and pedestrians. As highlighted by Laureshyn et al. (2017b), cyclists in conflict situations tend to prefer swerving over braking as their primary evasive action. Similarly, pedestrians often rely on sudden changes in direction, walking faster, running or completely stopping to avoid collisions (Hussein & Sayed, 2015; Malkhamah et al., 2005). Consequently, these indicators are primarily suited for motor vehicle analysis and may be constrained by the unavailability of key details, such as the pavement friction coefficient, tire conditions, or driver attentiveness, which are often based on certain assumptions.

4.2.4 What do SMoSs measure?

The ranking of traffic events can vary significantly depending on the calculation methods used to assess severity. It is necessary to ensure that this ranking aligns with theoretical expectations; if not, any further analyses or models become questionable or inaccurate. Without this alignment, the validity of the entire analysis could be compromised.

The concept of event severity in traffic safety requires precise clarification. Traditionally, severity has been measured based on the concept of 'proximity to an accident,' which refers to how close a traffic event comes to causing an accident. Most existing indicators express the severity of a traffic encounter in terms of its proximity to a crash, measured in time or space (Zheng et al., 2014). For instance, in their large-scale validation study, Johnsson et al. (2021) observed that in interactions between cyclists and turning motor vehicles at urban intersections (in Denmark, Norway, Sweden), most conflicts that were perceived as severe by observers were associated with low TTC values, and some also exhibited low PET values. However, identifying a clear threshold at which to distinguish between severe and non-severe conflicts proved challenging. For TTC, while low thresholds captured severe conflicts, they also resulted in 95% false alarms. These false alarms occurred in situations where both parties had seen each other well in advance and adjusted their speeds and paths accordingly, which was indicative of efficiency rather than safety issues. With respect to PET, some situations perceived as severe had relatively large PET values due to intense emergency braking—often to a complete stop. In conclusion, both indicators represent different ways of ranking the situations, with significant variations in their outcomes. However, both rankings were counterintuitive, suggesting that neither TTC nor PET fully captured the 'theoretical' severity of the conflicts.

This challenge highlights the need for a more refined approach to traffic safety, one that goes beyond measuring only the proximity to a crash. The adoption of the Vision Zero (Safe System) strategy, which focuses specifically on preventing accidents that lead to serious injuries or fatalities rather than aiming to prevent all accidents in general (Johansson, 2009), represents a paradigm shift in the traffic safety field. This shift implies a strategic trade-off: accepting accidents that cause only property damage while prioritising the prevention of those that result in serious injuries or fatalities. To align with this new approach, the definition of event severity must evolve to include another dimension—namely, the potential consequences of a crash (see Figure 26). As a result, severity should be redefined not as 'proximity to an accident' but as 'proximity to a serious injury' (Laureshyn et al., 2017a).

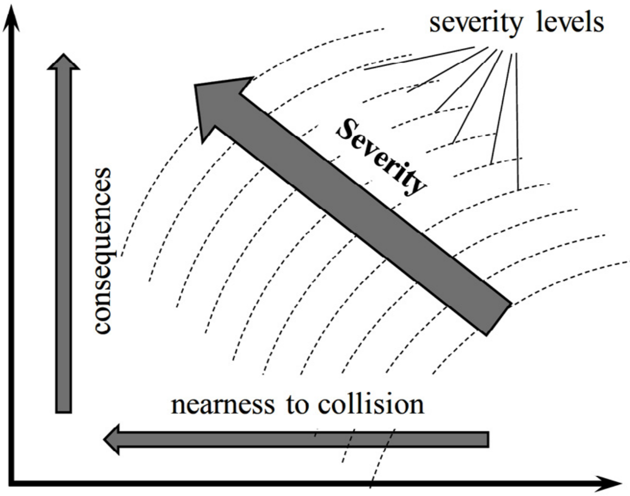

Figure 26. Conceptual illustration of severity (Laureshyn et al., 2017a)

Potential consequences can be represented in various ways, such as by incorporating road users' speed (both absolute and relative, calculated, for example, at the predicted collision point), impulse, kinetic energy, and deceleration rate, among others. Another effective measure is Delta-V, which has been used in numerous studies and has shown promising results (Laureshyn et al., 2017a; Bagdadi, 2013; Shelby, 2011; Gabauer & Gabler, 2008; Evans, 1994). In the context of traffic safety, Delta-V refers to each road user's expected change in speed during a hypothetical collision, assuming they reach the collision point with their current kinetic energy. This measure is particularly useful in situations involving road users of differing masses, such as those with VRUs, as it is sensitive to the mass difference between the colliding objects, reflecting the lighter participant's vulnerability. This aligns with the idea of compatibility in collisions, where differences in the mass and speed of road users exert a decisive influence on accident outcomes (Elvik, 2004; Harms, 1993). However, because Delta-V is continuous and can be computed at any moment during an encounter, it is important to select the appropriate moment based on robust theoretical considerations.

So far, only a few attempts have been made to incorporate both the proximity and the potential consequences of a collision into severity assessments. One example of this is the Swedish Traffic Conflict Technique (Laureshyn & Várhelyi, 2018; Hydén, 1987) which assesses traffic situations based on the TTC at the moment one of the road users first perceives the danger and initiates the first evasive action while also considering the road user's speed at that same moment. The underlying theoretical expectation is that situations with a small TTC (reflecting proximity) become more severe at high speeds (reflecting the potential consequences), as this significantly increases the likelihood of serious outcomes.

The Dutch Objective Conflict Technique for Operation and Research (DOCTOR; Kraay et al., 2013; Van der Horst & Kraay, 1986) is another example of how consequences can be factored into severity assessments. DOCTOR employs a subjective severity score that estimates the likelihood of injuries in the event of a collision. This score is based on various objective criteria, including the types of road users involved, the nature of their manoeuvres, and the presence of specific factors, such as whether evasive actions were taken and whether the situation was controlled or uncontrolled.

Building on similar concerns, Bagdadi (2013) proposed a severity assessment that combines Delta-V, Time-to-Accident (TA)—defined as the Time-to-Collision (TTC) value at the moment the first evasive action is taken (Hydén, 1987)—and the assumed maximum average deceleration. A key limitation of this measure is that it is only applicable in situations with a collision course. Later, Laureshyn et al. (2017a) suggested the Extended Delta-V measure, which combines proximity to a collision, represented by the time remaining in which to take evasive action (T2, which is applicable for scenarios with and without a collision course), with the potential outcome had an accident occurred, reflected by the Delta-V calculated at the moment of T_2^{min}. However, the authors identified several challenges associated with this measure, including the significant issue that a 'true' Delta-V value cannot be observed in traffic conflict studies; only hypothetical values can be generated based on assumptions. Consequently, the resulting Delta-V reflects only the expected outcomes of a crash. This requires an assumption that the involved road users will maintain their current speeds throughout the interaction, which limits the possibility of taking evasive actions that could mitigate the consequences of a potential collision. Additionally, the assumed deceleration is influenced by various factors, including the behaviour of the involved road users, weather conditions, and infrastructure characteristics.

Many other complex indicators have also been introduced; however, their applicability is often limited, as the assumptions underlying these indicators are valid only in specific situations and conditions.

4.2.5 What is the optimal moment in an interaction to measure severity?

Many existing indicators are continuous variables, generating values at every instant during an interaction. Consequently, the identification of an 'important' or 'optimal' moment (or even multiple moments) remains a subject of discussion. The severity of a traffic conflict can vary depending on the phase of the interaction, and selecting the right moment is crucial for accurately capturing the potential risk.

There may be various interpretations regarding what constitutes the starting point of a developing situation. It could be when the road users first become visible in the camera view or when they reach specific positions on the roadway and initiate interaction with one another. Within the framework of SMoS, the conflict's 'start' is often defined as the moment when one of the road users initiates an evasive action (Hyden, 1977). This marks the transition from a state of 'unawareness of danger' to a qualitatively different state of 'avoiding danger'. However, this state of 'avoiding danger' does not guarantee a successful outcome, as it may still result in further consequences, such as a collision.

It is also important to note that not all situations feature a clearly identifiable evasive action. For instance, a study by Niewöhner et al. (2011) found that drivers did not brake in at least 30% of all car-to-pedestrian crashes recorded in the German In-Depth Accident Study (GIDAS). In some cases, road users may anticipate passing each other with a sufficient safety margin and feel no need to take evasive measures. Additionally, it is possible that some road users remain unaware of the impending danger until the moment of collision, leaving them with no opportunity to attempt evasive action.

Detecting the onset of evasive action has become a significant challenge in traffic conflict analysis. Advances in automated video analysis have enabled the precise calculation of various objective features, leading to a growing interest in the investigation of evasive actions during conflict situations (Johnsson & Laureshyn, 2022; Tageldin et al., 2017; Tageldin & Sayed, 2016; Malkhamah et al., 2005). For instance, Johnsson & Laureshyn (2022) proposed an algorithm for identifying such moments. However, while their approach exhibits potential, it remains both computationally demanding and unreliable in complex situations involving secondary interactions, which are more likely to lead to traffic safety problems than simpler, one-on-one interactions.

Other critical moments in the development of a traffic situation ('culmination') can occur when one of the relevant indicators reaches its extreme, for example, when the two road users are closest to one another in time or space (e.g. the instants of TTC_{min} or T_2^{min}). It is important to recognise that different indicators may not reach their extreme values at the same instant. As a result, multiple key instants could be crucial for accurately assessing the severity of the situation. These moments allow for the calculation of additional indicators or features, providing valuable data for further analysis.

Traditional traffic conflict techniques often apply thresholds to indicators like TA, TTC_{min}, and PET (Laureshyn et al., 2016). However, the identification of clear thresholds for distinguishing between safe and unsafe events presents many challenges. For instance, Johnsson et al. (2021) discussed some concerns regarding SMoSs thresholds. First, their study found a strong correlation between encounters and critical events when applying PET threshold values greater than 2s, indicating that such high thresholds may measure exposure rather than actual risk. Second, the authors observed that the results of SMoS analysis using both PET (2s, 1s) and TTC_{min} (1.5s, 2s) thresholds were inconsistent with the ground truth. It was concluded that TTC_{min} thresholds seem to be more meaningful than PET; however, despite effectively capturing severe conflicts at lower threshold values, they also resulted in 95% false alarms.

Following the arguments put forth by Güttinger (1982), which suggest that a conflict may 'occur' before the situation has been resolved thus leaving room for the conflict to escalate into an accident, the final stage that describes severity lasts until the very last moment at which a collision is still (at least hypothetically) possible—the moment at which the first road user leaves the conflict zone. This perspective challenges the common practice of defining a conflict solely by measuring the TTC_{min} (assumed as the maximal proximity to a collision), as a non-zero TTC_{min} value indicates that a collision was avoided. In contrast, the T_2 indicator no longer generates values once the first road user has left the conflict zone.

In addition to this, another key moment is highlighted by Tarko (2021)—the 'moment of failure', when an emergency response is required to prevent a consequential collision (the trigger for evasive action). This moment may serve as a more accurate crash precursor, as it is a direct cause of the crash. The author classifies traffic encounters into two categories: events with 'no failure' and events with 'failure' (see Figure 27). 'No failure' events have zero probability of resulting in a crash and are the most frequently occurring events in everyday traffic, positioned at the very bottom of Hyden's safety pyramid. In contrast, events that pose safety concerns, including traffic conflicts of varying severity and actual crashes, fall into the 'failure' category. There are two potential outcomes within the 'failure' category. The first is 'recovery', where successful evasive actions are taken by one or both road users, preventing a crash. These events are usually easier to observe within a short timeframe. The second outcome, 'no recovery', occurs when evasive actions are either unsuccessful or not taken at all, resulting in a crash. Such events tend to be more random and are generally more challenging to observe in a short timeframe. While Tarko (2021) identifies the moment of failure as a possible key precursor to crashes, developing reliable methods for detecting this critical moment remains challenging.

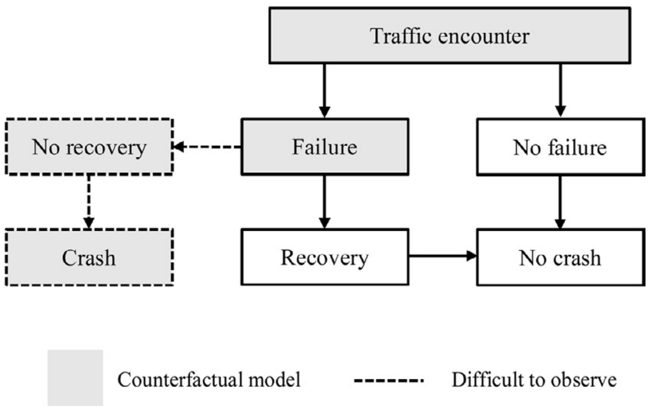

Figure 26. Traffic encounters' outcome mechanism (adopted from Tarko (2021))

4.2.6 Integrating subjective components as a tool in SMoS development

Since the primary purpose of the SMoSs is to describe event severity and to rank situations accordingly, it is important to examine and compare the results to ensure that the outcomes are intuitive and logical. It appears that subjective perception comes closer to the theoretical severity than many objective measures. For example, Johnsson et al. (2021) concluded that while there is no definitive proof that human perception of traffic severity represents an accurate ground truth, it encompasses a broader range of factors, including proximity to collision, potential consequences if a collision occurs, and a level of control during the situation. So, it is at least closer to the 'true' severity.

Earlier studies have indicated that human judgments can be reliable and valid (Svensson, 1992; Hydén, 1987; Lightburn & Howarth, 1979). However, some research suggests that initial training may be necessary to improve the accuracy of these judgments.

A large-scale validation study of the Swedish Traffic Conflict Technique tested various TTC and speed thresholds alongside a subjective measure of conflict danger (Svensson, 1992). Conflicts selected based on this danger rating exhibited the strongest correlation with police-reported accidents, and the dataset included several hundred observed intersections. Although the final technique does not use subjective scores, experienced observers often adjust TTC and speed values based on their personal judgment of the situation (Laureshyn & Várhelyi, 2018).

Kruysse (1991) and Kruysse & Wijlhuizen (1992) studied the specific aspects and time instances during traffic conflicts that most influence people's perceptions of danger. They found strong agreement among both expert and non-expert observers regarding their assessments and reported that people typically form their opinions during the early stages of a conflict, particularly when evasive actions begin. However, it was difficult to link these opinions with specific factors, such as the intensity of the evasive action, speed, or distance. People tend to perceive the danger of a situation as a whole rather than by analysing its individual components.

Several studies have demonstrated that people generally agree on the perceived danger of traffic situations. For instance, Madsen (2018) and Jensen (2016) used the Delphi-method, a formalised procedure designed to achieve consensus among experts on a prediction or judgment through structured discussions and iterative adjustments of initial opinions, was used. Both studies reported successful applications of this method, indicating that there is a tendency for individuals' perceptions of danger in traffic situations to align.

4.3 Use case of subjective insights to deepen understanding of severity levels in traffic encounters

4.3.1 Study design

4.3.1.1 Video data and extraction of objective indicators

This study used video recordings from the Horizon 2020 InDeV project (EU Commission, 2020) at urban intersections in Sweden, Denmark, Norway, the Netherlands, and Spain. It focused on cyclist–motor vehicle encounters, in which a cyclist passes through a signalised intersection on green while a turning motor vehicle crosses their path. To ensure representation across severity levels, the dataset was manually composed, including three categories:

- Normal interactions: controllable situations with no or minor speed and path adjustments.
- Slight conflicts: smooth, early interactions with no emergency actions taken.
- Severe conflicts: near-accident situations requiring sudden emergency actions to avoid a collision.

The dataset included 107 traffic encounters. Trajectories were extracted using the semi-automated T-Analyst tool (T-Analyst, 2019), with each situation characterised by objective indicators (see Table 8) derived from trajectories and parameters such as vehicle dimensions and weight (Laureshyn, 2010).

Table 8. Objective indicators (adopted from Paper IV)

Indicator notation	Definition
EA	Presence of evasive action (EA)—a categorical variable indicating whether at least one road user applied an evasive action (1 = yes, 0 = no).
Who takes the first EA?	Categorical variable, 1 for car and 0 for bicycle
Who leaves the conflict area first?	Categorical variable , 1 for car and 0 for bicycle
T_2	The time remaining for the latest road user to arrive at the conflict area—(Laureshyn et al., 2010). This equals TTC in the case of a collision course.
D_1	The Euclidean distance between the two road users.
D_2	The sum of the travel-to-collision-point distances.
PET	The time from the first road user (who reached the conflict area first) leaving the area of potential collision to the moment when the second road user enters it (Allen et al., 1978)
TAdv	Predicted PET value if the road users continue with their intended speeds and trajectories. This value is equal to zero in the case of a collision course (Laureshyn et al., 2010).
DR_{mv}^{min}	The minimum deceleration rate necessary for the motor vehicle to come to a complete stop before reaching the conflict point. This is calculated only when the motor vehicle 'catches up with' the cyclists, otherwise, it is zero.
DR_{mv}	The deceleration rate of the motor vehicle. In the case of no deceleration or positive acceleration, the value is zero—adopted from Gettman et al. (2008).
V_{mv}, V_{bike}	Speeds of the motor vehicle and bicycle, respectively.
V_{rel}	The relative velocity of the involved road users.
KE_{mv}, KE_{bike}	The kinetic energies of the motor vehicle and bicycle, respectively.
$Impulse_{mv}$, $Impulse_{bike}$	The impulse (product of mass and speed) of the motor vehicle and bicycle, respectively.
$DeltaV_{mv}$, $DeltaV_{bike}$	The changes in the velocity of the motor vehicle and bicycle, respectively, due to a hypothetical collision.

The indicators were categorised according to the situation development phase process they most closely corresponded with (see Figure 28).

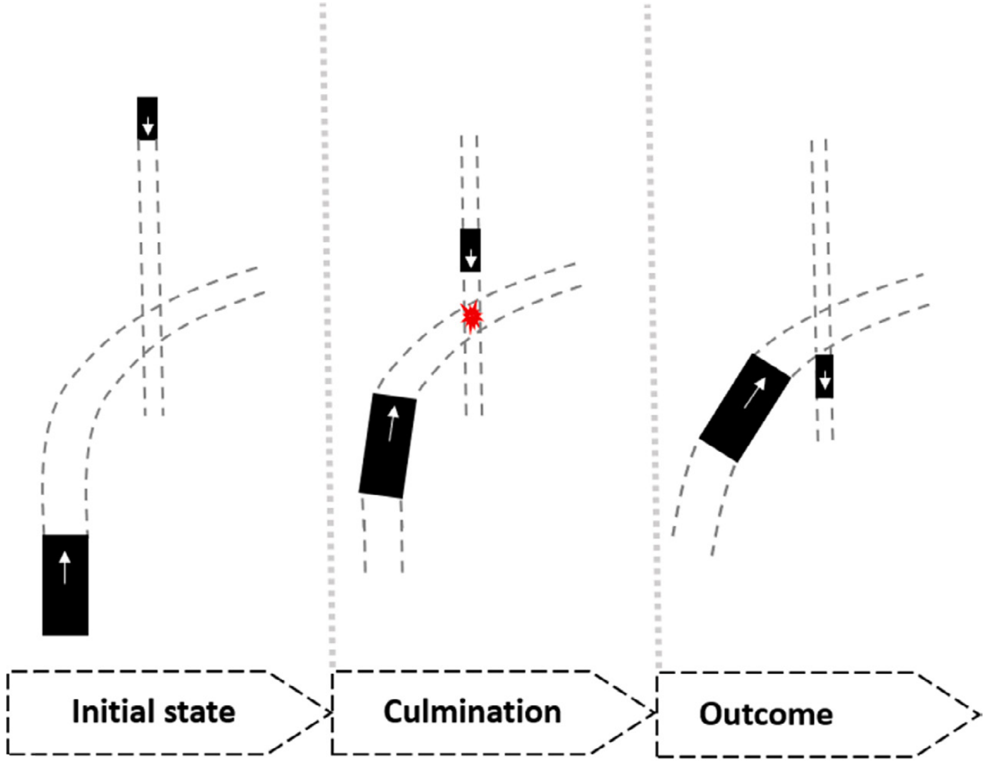

Figure 28. Conceptual model of the key moments during bicycle–motor vehicle interactions considered for the calculation of objective features in this study (adopted from Paper IV)

Table 9 outlines the definitions of these phases. Notably, some of the indicators used in this study are continuous, producing values at various time instances, which can lead to their inclusion in multiple phases.

Table 9. Description of key moments (adopted from Paper IV)

Phase	Definition	Time instance notation	Indicators	Included in model
Initial state	The onset of the first evasive action	EA	*EA, T_2, T_{Adv},* DR_{mv}^{min}, *DR_{MV},* V_{mv}, V_{bike}, V_{rel}, *KE_{mv}, KE_{bike},* $Impulse_{mv}$, $Impulse_{bike}$, *$DeltaV_{mv}$, $DeltaV_{bike}$, D_1, D_2, Who takes EA*	D
Culmination	The moment when T_2 reaches its lowest value during the event.	T_2^{min}	*T_2, T_{Adv},* DR_{mv}^{min}, *DR_{mv},* V_{mv}, V_{bike}, V_{rel}, *KE_{mv}, KE_{bike},* $Impulse_{mv}$, $Impulse_{bike}$, *$DeltaV_{mv}$, $DeltaV_{bike}$, D_1, D_2*	A,B,C,D
	The moment when the accumulated travel-to-collision-point distance of each road user reaches its lowest value during the interaction.	D_2^{min}	*T_2, T_{Adv},* DR_{mv}^{min}, *DR_{mv},* V_{mv}, V_{bike}, V_{rel}, *KE_{mv}, KE_{bike},* $Impulse_{mv}$, $Impulse_{bike}$, *$DeltaV_{mv}$, $DeltaV_{bike}$, D_1, D_2*	A,B,C,D
Outcome	The moment when the first road user (who reached the conflict area first) arrives at the area of a potential collision.	T_1	*T_2, T_{Adv},* DR_{mv}^{min}, *DR_{mv},* V_{mv}, V_{bike}, V_{rel}, *KE_{mv}, KE_{bike},* $Impulse_{mv}$, $Impulse_{bike}$, *$DeltaV_{mv}$, $DeltaV_{bike}$, D_1, D_2, Who leaves the conflict area first*	A,B,C,D
	The moment when the second road user arrives at the area of a potential collision.	T_2	*PET*	A,B,C,D

4.3.1.2 Subjective component

Eighteen observers participated in the study, including eight traffic safety experts (aged 27–65, balanced genders) and 10 non-experts (60% female, aged 24–47). Recruited via social media and personal contacts, participants were shown pairs of traffic situations and were asked to identify which seemed more dangerous. Pairwise comparison was used to assess relative severity, making the task more intuitive than assigning detailed scores. Observers received an Excel file with links to 60 randomly generated YouTube video pairs, enabling rewatching and the option to indicate whether the severity of the situations seemed similar.

4.3.1.3 *Inter-observer reliability and internal consistency tests*

An inter-observer reliability test was used to assess differences between expert and non-expert responses. Fisher's exact test (McDonald, 2009) revealed no significant difference in the proportions of answers between the two groups, consistent with the results of Kruysse & Wijlhuizen (1992), who found strong agreement between the two groups. Internal consistency was evaluated using Cronbach's alpha (Cronbach, 1951), with coefficients ranging from 0.79 to 0.82, indicating high consistency across all responses.

4.3.2 Decision machine development

Binary logistic regression analysis (BLRA) with a forward selection approach was applied. Each observer-processed pair was labelled, with the first event designated '0' and the second '1,' making the responses the dependent categorical variable. Selected explanatory indicators were transformed using a pairwise ranking method based on Melnikov et al. (2017), with each indicator representing the difference between the more and less severe situations. A binary classification algorithm was then used to determine the weight vector (ω) for each indicator, which were treated as explanatory variables.

4.3.2.1 *Models*

The initial dataset used to fit the models included events with and without detected evasive actions, making it impossible to calculate evasive action-related indicators for the entire dataset. Due to this limitation, the analysis was structured with the following objectives:

First, a simple regression model (*Model A*) was developed to identify statistically significant indicators for predicting human responses, assess the accuracy of these indicators in predicting severity judgments, and determine the statistical significance of the presence of evasive actions. Model development used the full dataset (N = 1003), focusing on human judgments of severity in relation to a set of objective indicators that could be calculated for each event. Indicators related to evasive actions were excluded, but a class variable (EA) was retained to indicate whether an evasive action was detected (yes/no). The model can be summarised as follows:

> *MODEL A:* Human responses (entire dataset, N = 1003) ~ Indicators from moments T_2^{min}, D_2^{min}, T_1, T_2 + class variable (EA)

Second, the initial dataset was split into two subsets: one including traffic events for which an evasive action was detected in both situations of a pair (referred to as the EA subset), and another including events where no evasive action was detected in one or both situations of a pair (referred to as the NoEA subset). Two additional models were then developed—*Model B* (EA subset) and *Model C* (NoEA subset)—using the same explanatory variables, excluding those related to evasive action onset. These models aimed to assess whether there was a statistical association between the same indicators and human judgments across both subsets. The models are outlined as follows:

MODEL B: Human responses (EA subset, n = 873) ~ Indicators from moments T_2^{min}, D_2^{min}, T_1, T_2

MODEL C: Human responses (NoEA subset, n = 130) ~ Indicators from moments T_2^{min}, D_2^{min}, T_1, T_2

Finally, a fourth model (*Model D*) was developed using only the EA subset, allowing all selected variables to be included. The objective was to assess whether indicators related to the moment of evasive action added value in estimating severity based on human judgments. The model is schematically described as follows:

MODEL D: Human responses (EA subset, n = 873) ~ Indicators from moments *EA*, T_2^{min}, D_2^{min}, T_1, T_2

4.3.3 Results

The results from *Models B* and *C* were analysed to compare the EA (evasive action) and NoEA subsets (see Table 10). In both models, the lowest T_2 value during the event was statistically significant and contributed most to predictions. However, other significant variables emerged from different phases of the traffic event, indicating differences between situations with and without observable evasive actions. Additionally, the ROC curve for *Model B* showed a slightly lower value, suggesting that the initial conditions phase (when evasive action begins), excluded from this model, may play a key role in evaluating event severity.

Table 10. Regression coefficients (α and β_i), their standard errors and levels of significance, and g.o.f. for the fitted models (AUC, deviance statistics, BIC, and AIC) (adopted from Paper IV)

Parameter	Estimate	Standard error	Wald Chi-square	Pr > ChiSq	Standardised estimate	ROC, % (Individual contribution, %)
Model A (N=1003; NoEA & EA pairs of events)						
Intercept (α)	0.1386	0.1802	0.5913	0.4419		
T_2 (T_2^{min})	-1.459	0.113	166.5692	<.0001	-1.0551	79.01 (90.2)
EA	1.6092	0.3143	26.216	<.0001	0.3158	82.16 (3.6)
D_2 (D_2^{min})	-0.1154	0.0236	23.9194	<.0001	-0.3336	84.87 (3.1)
D_1(D_2^{min})	-0.2363	0.0527	20.1113	<.0001	-0.3636	85.05 (0.2)
D_1 (T_1)	-0.1131	0.0538	4.4131	0.0357	-0.1689	85.91 (1)
V_{mv} (T_2^{min})	0.0535	0.0312	2.9325	0.0868	0.0902	86.22 (0.4)
$DeltaV_{bike}$ (T_2^{min})	0.0893	0.0359	6.1864	0.0129	0.1233	87.01 (0.9)
DR_{mv}^{min} (T_2^{min})	-0.1572	0.0571	7.5756	0.0059	-0.1366	87.55 (0.6)
Deviance Chi-Square: 934.64 (Pr > ChiSq: 0.0387) - BIC = 959.6 - AIC = 910.5						
Model B (N=873; EA pairs of events)						
Intercept (α)	-0.1302	0.0871	2.2336	0.135		
T_2 (T_2^{min})	-1.4552	0.1202	146.6892	<.0001	-1.0196	78.41 (90.6)
D_2 (D_2^{min})	-0.114	0.0258	19.5775	<.0001	-0.3032	81.19 (3.2)
D_1 (D_2^{min})	-0.2428	0.0582	17.4042	<.0001	-0.3594	85.05 (4.5)
D_1 (T_1)	-0.0971	0.0583	2.7786	0.0955	-0.1427	85.91 (1)
DR_{mv}^{min} (T_2^{min})	-0.1522	0.0585	6.7671	0.0093	-0.1379	86.57 (0.8)
Deviance Chi-Square: 744.29 (Pr > ChiSq: 0.5314) - BIC = 845.5 - AIC = 816.9						
Model C (N=130; NoEA pairs of events)						
Intercept (α)	0.3923	0.3459	1.2861	0.2568		
T_2 (T_2^{min})	-1.3889	0.3166	19.2493	<.0001	-1.1976	80.43 (86.4)
D_1 (T_1)	-0.217	0.1096	3.9185	0.0478	-0.3591	86.5 (6.5)
Who leaving the conflict area first	1.3561	0.6707	4.0883	0.0432	0.4832	90.33 (4.1)
$DeltaV_{Bike}$ (D_2^{min})	0.294	0.1359	4.682	0.0305	0.4488	91.86 (1.6)
DR_{mv}^{min} (T_2^{min})	-2.3933	0.8202	8.5146	0.0035	-1.361	93.13 (1.4)
Deviance Chi-Square: 108.54 (Pr > ChiSq: 0.4943) - BIC = 121.3 - AIC = 101.2						
Model D (N=873; EA pairs of events)						
Intercept (α)	-0.3206	0.1146	7.8232	0.0052		
T_2 (EA)	-1.1789	0.1332	78.367	<.0001	-0.826	84.19 (91.8)
$DeltaV_{bike}$ (EA)	0.6275	0.0609	106.2249	<.0001	0.8754	88.12 (4.2)
D_2 (EA)	-0.0594	0.0091	42.5588	<.0001	-0.4491	90.51 (2.6)
$DeltaV_{bike}$ (D_2^{min}))	0.2388	0.0466	26.2291	<.0001	0.3398	91.08 (0.6)
KE_{mv} (EA)	0.000018	4.80E-06	14.0906	0.0002	0.2289	91.51 (0.5)
D_1(T_1)	-0.1406	0.0557	6.3839	0.0115	-0.2065	91.65 (0.2)
DR_{mv}(EA)	-0.0732	0.0442	2.7446	0.0976	-0.0908	91.73 (0.1)
Deviance Chi-Square: 622.48 (Pr > ChiSq: 0.9996) - BIC = 556.7 - AIC = 520.3						

A comparison of *Models B* and *D*, both fitted to the same dataset with detected evasive actions, exhibited differing sets of variables. *Model B* focused on moments with the lowest T_2 and D_2 values, explaining over 85% of human judgments about severity. In contrast, *Model D* emphasised the initial evasive action, with its indicators accounting for over 90% of the data.

The models' performances were assessed using Bayesian information criterion (BIC) and Akaike information criterion (AIC) (Lin et al., 2016), with smaller scores indicating better fit (Cavanaugh & Neath, 2019; Neath & Cavanaugh, 2012). As shown in Table 10, *Model D* had lower BIC and AIC scores, making it the preferred model and also a higher area under the curve (Abdel-Aziz & Karara), indicating better overall performance.

The holdout method (Arlot & Celisse, 2010) was applied to validate *Model D*. The dataset used to fit *Model D* (n = 873) was randomly split into 80% for training and 20% for testing. The results in Figure 29 show that *Model D* can correctly discriminate over 91% of human judgments from both data subsets.

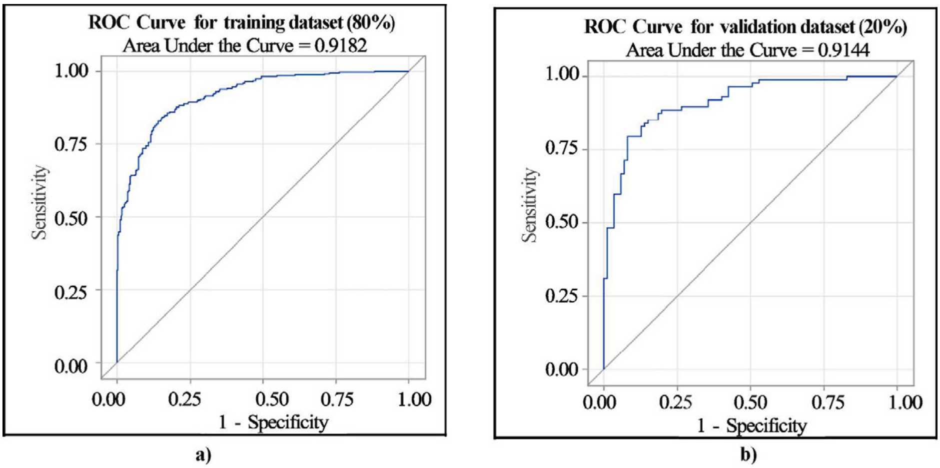

Figure 29. Receiver Operating Characteristic (ROC) curves: a) training set, b) validation set (adopted from Paper IV).

4.3.4 Practical implications and future research direction

The study findings highlight the significance of evasive action as a key factor in human perceptions of traffic event severity. Indicators from an event's initial conditions, defined by the start of an evasive action, are most valuable for explaining its severity.

The next step in understanding the practical implications is to develop a clear definition of what constitutes an evasive action. In situations where drivers naturally adjust their speed or direction due to road design or traffic signals, detecting deviations from expected behaviour can be challenging. Some methods have been proposed, such as Johnsson & Laureshyn (2022) approach, which compares a road user's current path with previous unhindered trajectories at the same location. Evasive action is assumed to start when the current path no longer matches the historical data. While promising, this method becomes less effective in more complex, multi-step interactions, where traffic conflicts are more likely. Additionally, the method is computationally demanding, requiring extensive historical data to ensure accuracy.

This raises an important question: should surrogate safety analysis include all critical events or focus solely on those involving clear evasive actions, which road users perceive as more safety-threatening? What about scenarios in which road users narrowly avoid each other without taking any evasive action until the situation resolves itself or a collision occurs?

Tarko (2021) addresses this by arguing that a failure requiring an emergency response to prevent a collision is a precursor for any non-zero probability of a crash. Thus, identifying this precursor event is necessary for understanding crash causality. The author categorises traffic encounters into two types: preferred (error-free events that cannot result in a crash) and failure-affected (the underlying mechanism we aim to understand) events (see Figure 30). In the latter case, a failure has already occurred, whether or not evasive action is taken, and whether, if taken, it is successful or not. Tarko (2021) defines this failure as '*a temporary lack of response to a traffic encounter, a lapse of awareness, that leads to a short separation (in time or space) between vehicles that necessitate an uncharacteristic evasive maneuver to avoid collision and considerable injury or damage*'.

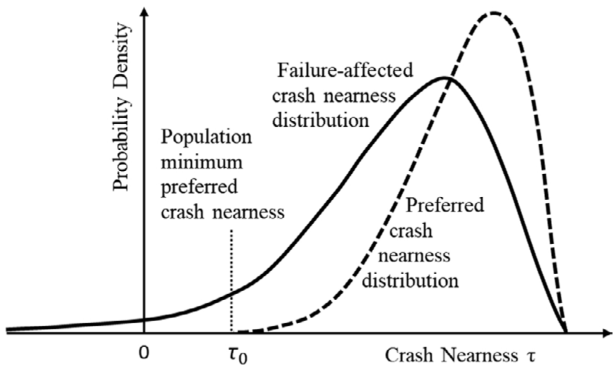

Figure 30. Distributions of preferred (dashed line) and failure-affected (solid line) crash nearness values (Tarko, 2021)

It is also worth noting that in the presence of variables related to evasive actions, those associated with the culmination phase contribute very little to explaining event severity. Additionally, outcome-related variables were not significant in any of the tested models. This aligns with previous studies; for example, Johnsson (2020) observed that the PET indicator does not accurately measure safety but primarily reflects exposure. Similarly, Tarko (2019) concluded that PET values reflect not severity but the outcome of a conflict when all potential paths toward a crash are irrelevant.

The present study also confirms that variables related to both proximity and collision consequences are necessary for severity estimation. Future research should focus on developing an integrated severity measure that incorporates multiple indicators, covering both dimensions: proximity and potential consequences. Advanced tools such as EVT (multivariate models) may increase the predictive power of surrogate measures of safety compared to traditional models that relate single indicators like TTC_{min} or PET to accident risk.

While this study provides important insights, it explores a straightforward case that can serve as a foundation for future research. It would be valuable to investigate potential differences in various types of interactions, road users, and locations. In addition, situations involving secondary interactions were excluded from this dataset to ensure that no other road users influenced the behaviour of the studied pair. However, it has been previously noted that safety issues are more likely to occur in complex interactions (particularly those involving secondary interactions) than in simpler scenarios involving only two road users (Johnsson, 2020).

Finally, the pairwise comparison method used in this study to gather subjective judgments may be inadequate when comparing straightforward interactions to more complex, multi-step interactions. Future research should account for this complexity to further refine the approach to severity estimation in diverse traffic scenarios.

5 Discussion

5.1 Connecting the dots: Bridging computer vision, traffic research, and civil engineering practitioners for smarter collaboration

The development of quick feedback tools for transport specialists, particularly practitioners, could significantly reduce the financial costs and human risks associated with traditional trial-and-error methods. By adopting technologies that facilitate rapid testing and adaptation, we can ensure that various transport measures are not only scientifically grounded but also practically effective. This approach ultimately addresses the needs of road users, improving safety and conserving resources. It marks a groundbreaking shift from reactive to proactive strategies for transportation management.

In this transition, video analysis stands out as a powerful tool among the promising technologies. However, effective and reliable video analysis tools are crucial to unlock its full potential (practically). Historically, the expectations of a usable video analysis system has remained consistent—it should provide detailed descriptions of road users' movements, facilitate the analysis of long observation periods, and, ultimately, automate all video transformation steps (including video processing, accurate detection, classification, and road user tracking) to receive output in the form of high-quality (accurate) microscopic data that will be used for further traffic analysis.

Despite the growing interest in video technologies over the past few decades, a ‘perfect’ tool for traffic video analysis is yet to be developed. One major barrier to progress lies in the communication gap between the following three key components of the successful application of video-based methods for traffic management solutions:

1. Computer vision specialists handle video processing, focusing on the technical aspects of capturing, storing, and preparing video data for analysis.

2. Traffic researchers provide the expertise needed to guide the 'intelligent' analysis of this data and the interpretation of the results, extracting meaningful insights into traffic behaviour and dynamics.
3. Civil engineering practitioners use these findings to implement real-world interventions and infrastructure improvements.

Currently, communication primarily occurs between computer vision professionals and civil engineering practitioners. Practitioners often order traffic analysis for various purposes from various companies that develop video-processing technologies to extract microscopic data for the analyses. While the microscopic data extracted are often of high quality, the insights gained are typically limited to basic metrics, such as speed profiles and traffic counts. While these metrics are important, they represent only a fraction of the valuable insights that can be derived from video data, many of which could be highly beneficial to practitioners.

This highlights a critical missing link—traffic research. Traffic researchers provide in-depth, domain-specific insights, such as behavioural and contextual understanding, which are fundamental for meaningfully interpreting microscopic traffic video data. Without this link, solutions derived from video analysis often lack the depth needed to address complex traffic challenges comprehensively. By incorporating traffic research into the process, a more holistic and effective approach to data-driven traffic management decision-making can be achieved.

While some initiatives have attempted to bridge the gap between computer vision and traffic research (e.g. Ahrnbom, 2022; Jensen et al., 2019; Laureshyn, 2010), these efforts have mostly remained within academic circles. The challenge lies in translating these innovations into practical, everyday tools that can support transportation practitioners. To improve sustainable transportation systems, a tool that is both accessible and actionable for practitioners is essential. Such a tool would provide clear guidelines and deliver results that answer critical questions about new measures: How safe is the location? Are the needs of road users adequately addressed in their interactions with the infrastructure? Has the level of comfort improved? What about perceived safety and how does it align with objective safety metrics?

5.2 From methodology to practical implications: Optimizing decision-making through video analysis

Currently, the development of cycle and pedestrian paths in Sweden (and elsewhere) relies heavily on practice-based knowledge and experience. However, there is still a lack of systematic evaluation of and documentation on which design approaches most effectively meet road users' needs for safety and comfort, leaving limited insight into which approaches work well and which do not. Developing a system to collect, analyse, and use such data for infrastructure design would offer valuable insights and promote informed decision-making.

A promising methodology is introduced in *Papers I* and *II* of this thesis that may assist practitioners in evaluating the impacts of various interventions on infrastructure and road users' behaviour. A common aim of both studies was the creation of an effective, user-friendly approach that practitioners can systematically apply across diverse locations to support sustainable transportation development. As a result, the methodology proved effective in capturing behavioural responses to changes in infrastructure features, such as surface design and lighting conditions. One notable finding was that cyclists and pedestrians (often grouped together as VRUs) exhibit distinct responses to environmental changes at a microscopic level. Based on this observation, a key recommendation is that urban planning and design should avoid treating cyclists and pedestrians as a single homogeneous category and instead recognise and address the unique needs of each group when analysing infrastructure improvements.

However, the existing version of the proposed methodology only represents a partial step toward fully achieving the set goals, as it currently faces several challenges. Addressing these challenges in future work could further refine the methodology and expand its usability, even for large-scale applications.

With respect to the methodology's technical aspects, the initial data collection step, conducted using drone technology, offered several advantages over traditional stationary cameras. Drones helped overcome common issues regarding data privacy (e.g. GDPR compliance), mounting, power supply, and the need for multiple fixed cameras to cover large areas. For example, in the VAPM study by Johansson et al. (2020), covered a similar area and required 12 stationary cameras; the drone approach used for the studies in the present thesis simplified this setup. However, drone-based data collection also had its limitations, requiring the significant effort and skill of the drone operator, especially in terms of maintaining stable flights, and being highly vulnerable to adverse weather conditions like strong winds and heavy rain. Despite carefully planning each session in advance, not all scheduled time slots could be used, resulting in smaller participant sample sizes for the analysis.

After data collection, the footage recorded after dark required stabilisation and brightness adjustments; these tasks had to be performed manually for each drone flight, which was time-consuming. Following these adjustments, the next necessary video processing step was calibration, which directly impacted the accuracy of microscopic data. In the studies presented in this thesis, I employed an approach to calculate the camera's absolute pose by measuring 3D points on the ground or other fixed structures (e.g. building corners or window frames) within the area of interest and matching them to corresponding points on the image. While this method has proven accurate (as demonstrated by Laureshyn & Nilsson (2018)) it is labour-intensive, especially for extended datasets or scenarios in which the camera position frequently changes, as was the case with our drone-based video collection.

The repetitive tasks of video stabilisation, brightness adjustment, and calibration for each 20–25-minute clip should be further automated. One potential improvement, for example, would involve using reflective or illuminated calibration markers at the beginning of each experiment to enable the automatic detection of calibration points, reducing the need for manual identification and significantly streamlining the process.

More advanced, automated calibration techniques, such as the checkerboard method used by Ahrnbom (2022), offer even higher accuracy. However, the key challenge of this method lies in the lack of user-friendly tools to simplify the process. There is therefore a pressing need for intuitive (user-friendly) interfaces that enable users with minimal technical expertise to perform calibration efficiently. The development of such tools would empower practitioners to apply advanced calibration techniques without requiring a deep understanding of the underlying mathematical principles.

The subsequent steps of the analysis, such as detecting, classifying, and tracking road users, were carried out using a semi-automated tool and required a substantial amount of manual input. Although fully automated detection and tracking tools are now available, initial tests showed that their performance on the footage from the studies in this thesis was suboptimal. This is primarily due to the low resolution and varying brightness, especially under low-light conditions, which affects the contrast between individuals and the background. These issues could potentially be solved by using thermal cameras to distinguish people from the environment more effectively while also addressing data privacy concerns. Thermal imaging would allow for closer, lower-altitude drone flights, increasing the visible size of each person in the video without capturing identifiable information. While thermal sensor-equipped drones are available on the market, they are typically expensive. However, as technology advances, it is likely that such tools will become more accessible and affordable in the future.

It is also important to consider the theoretical aspects in addition to the technical challenges, particularly when selecting and interpreting the relevant variables. Any interpretation of the data should be based on a strong theoretical foundation to ensure its validity and relevance.

For example, in *Paper II* of this thesis, the behavioural responses were measured in terms of speed and slalom, which raised a critical question: Is an increase in speed or slalom a positive or negative road user's behavioural response? Initially, it was assumed that if path users (most of whom are likely commuters due to the study sites' locations) perceive the traffic environment as sufficiently safe, they would strive for higher speeds and lower deviations in their movements to reach their destinations as quickly as possible. However, this assumption requires further validation. To achieve this, additional qualitative research in the form of interviews or surveys could be conducted to gain deeper insights into road users' specific needs and to explore any additional factors that may influence their behaviour that may not be captured through video analysis alone.

A key takeaway from this process is that researchers must recognise that if a methodology is to be used by practitioners, it must be easily understandable and applicable. Therefore, it is important to provide suggestions and offer guidance (based on sound theory) on interpreting the results produced by application of the methodology.

5.3 From simplicity to intelligence

For many years, traditional traffic metrics, such as speed, road user counts (traffic flow), and route choice, have been the primary tools used for assessing traffic situations from various perspectives, including safety, accessibility, and level of service. These metrics have been effective in providing a broad overview of traffic dynamics, offering insights into traffic flow, identifying potential safety concerns, and evaluating the efficient use of road infrastructure. However, while these traditional metrics remain valuable, they capture only a portion of the insights that are possible with the use of more advanced methods, particularly video analysis.

Traditional metrics often lack the capacity to assess safety potential directly (Johnsson & Laureshyn, 2024; SWOV, 2018). In contrast, video recordings add a new layer of depth to traffic studies, enabling a more nuanced understanding of traffic conditions and road user behaviours. One example of this more advanced approach is conflict-based studies, which are now primarily conducted using video data. Unlike traditional methods that focus on aggregated data, such as speed and traffic counts, conflict-based studies analyse microscopic behavioural data (linked to time and space) from the interactions of individual road users.

A further example, distinct from conflict-based studies but also illustrating intelligent analysis, is a recent study by Johnsson & Laureshyn (2024), which assesses potential safety at specific locations using video recordings, going beyond the traditional analysis of speed and traffic counts. This method identifies the region carrying the highest injury risk within the vicinity of a target area, providing an estimate of the worst-case safety scenario or 'potential safety risk' at that site. This approach considers differences in road users' speed, direction, and type, adjusting for these factors based on an injury model framework originally introduced by Lubbe et al. (2022).

In addition to safety, it is also critical to understand road users' experiences and needs when developing effective and sustainable transportation systems. In *Papers I* and *II*, a video-based method was employed to analyse how specific physical environmental features impact user behaviour, focusing on metrics such as speeds, lateral positions, slalom and passing distances (linked with metrics like time and longitudinal distance). This approach aims to provide a comprehensive understanding of usability.

Similarly, other researchers have used video technology to examine various factors that influence road user comfort. For example, the level of service (LOS) metric is commonly used to quantify the comfort a road facility offers from a user's perspective, and video analysis has become an invaluable tool for capturing detailed data on user behaviours and interactions with which to assess this metric. Studies by Kazemzadeh & Bansal (2021); Kazemzadeh et al. (2020) applied video analysis to investigate e-bike riders' behaviour and comfort when navigating through pedestrian crowds, which are key factors in defining the LOS index for bike lanes. By extracting microscopic data from video recordings, these studies measured key metrics such as pedestrian crowd density, e-bike lateral distance, and e-bike speeds, offering valuable insights into how these factors influence the overall navigation experience.

Another illustrative application of video analysis can be found in Kuipers et al. (2024) study of the speed of passengers alighting and boarding trains under different conditions, such as when travelling with luggage. The video analysis was able to highlight passenger behaviour, making it possible to obtain a deeper understanding of this process.

Overall, these studies underscore the broader potential of video analysis in traffic behavioural studies. By enabling the measurement of a wide range of variables beyond traditional metrics, video analysis offers a level of detail that can deepen our understanding of traffic dynamics. These rich data can inform better decision-making regarding road infrastructure design and proactive safety measures.

5.4 Exploring human behaviour as a social construct

Traffic behavioural studies aim to examine the interplay between road users and the traffic environment they navigate. A key aspect of this research is human behaviour, a complex social phenomenon that is often difficult to quantify. Unlike natural phenomena governed by predictable physical laws human behaviour is influenced by various dynamic factors that can change over time. While numerous parameters can be analysed to capture certain dimensions of human behaviour, many inherent characteristics that are influenced by various external factors may be overlooked.

Within this context, two important dimensions of human behaviour emerge—reflexivity and randomness (Patterson, 2021; Noland, 2013)—both of which significantly influence traffic patterns and safety outcomes. These dimensions are essential to understanding how individuals interact with their environment and with one another on the road.

The first dimension, reflexivity, refers to an individual's ability to reflect on their own behaviour and the behaviours of others, often leading to adaptive decision-making in response to changing traffic conditions. For example, a driver may adjust their speed or route based on observed traffic flow, weather conditions, or the behaviour of nearby road users. This reflective process is based on both past experiences and real-time observations. Many proposed approaches for measuring behaviour work well in simple traffic scenarios but fall short when applied to more complex scenarios. Incorporating reflexivity into traffic behavioural studies can potentially address this limitation, as this dimension is particularly critical in complex scenarios involving secondary (multi-step) interactions, which are more likely to lead to safety problems.

The second dimension, randomness, captures the unpredictable nature of human behaviour. In a traffic context, people may make spontaneous decisions while on the move, driven by a variety of factors, such as emotions, distractions, fatigue or social influences. For instance, a driver might suddenly accelerate to avoid stopping at a red light, or a pedestrian might run at a red signal because they are in a hurry or simply not paying attention. This randomness can lead to unexpected traffic events, increasing the risk of accidents and complicating traffic management efforts.

Introducing reflexivity and randomness into traffic behavioural studies allows for a more nuanced view of human behaviour in traffic contexts, helping to build a more comprehensive framework. They are not only theoretical constructs but also have practical implications for designing safer and more efficient transportation systems. Without considering these dimensions, behavioural models could miss critical aspects of how individuals respond to complex, real-world traffic situations. Therefore, this theoretical perspective may help expand our understanding and provide insights into how interventions might be designed to address the complex and often unpredictable nature of human actions on the road.

5.5 Should a universal measure of road user behaviour be the goal?

Establishing a universal measure for road user behaviour is challenging, if not impossible, due to the highly varied and context-specific nature of traffic environments. Behaviour differs significantly depending on factors such as the type of road user (e.g. motor-vehicles, cyclists, pedestrians) and the specific location (e.g. intersections, roundabouts, highways, shared-use spaces, and pathways). For instance, metrics like acceleration and deceleration are often used to assess motor vehicle behaviour, but applying these straightforwardly to VRUs is less practical.

Speed, on the other hand, is a metric that appears to be more universally applicable. My studies found that speed is an effective metric for measuring cyclist behaviour. However, it is less effective in describing pedestrian movement patterns and preferences. Unlike cyclists, who generally follow designated paths or lanes and exhibit more predictable patterns, pedestrians display much greater spontaneity and adaptability in their movements. This variability results from factors such as individual preferences, social interactions, and responses to immediate environmental stimuli, making pedestrian movements appear more unstructured and often unpredictable.

Therefore, in developing a tool intended for practitioner use, it is important to offer adaptable metrics that can be adjusted according to the specific user groups and traffic settings under study. In addition to contributing a tool that allows for metric customisation, clear and practical guidance should also be provided on how to select and apply the most appropriate metrics for each context (e.g., urban vs. rural settings or areas with high pedestrian/cyclist traffic vs. motor vehicle traffic).

It is equally important to provide guidance that helps practitioners interpret and respond effectively to the insights derived. This support will ensure that the data are meaningful and actionable, helping practitioners to understand their implications and identify appropriate responses.

5.6 Beyond complexity: Developing clear and effective measures for traffic safety

Another important consideration when choosing a measure is the complexity of the metrics involved. *Paper III* attempted to reframe both the problem and its potential solutions in SMoS, drawing on current knowledge and methodologies. This study was inspired by Patterson (2021), who argues for moving beyond outdated paradigms that only add complexity and often hide core issues. Today's discussions within the SMoS context are often focused on very specific technical details while leaving fundamental principles underexplored.

Most current research seeks to advance commonly used indicators (such as TTC, PET, etc.) by adapting them for computational analysis despite inconsistencies in both the underlying theory and input data quality. This results in complicating the interpretation of findings, which is contrary to the original intentions. A timely reassessment of how SMoS is applied is therefore necessary, and *Paper III* proposes a framework for this purpose.

A key concern in this context is the adoption of the Safe System approach to traffic safety management, which is currently considered world best practice and is known as Vision Zero in Sweden. Despite its name, Vision Zero does not aim for zero crashes overall but rather for zero crashes that result in serious consequences. The primary aim is to eliminate severe injuries and fatalities rather than to simply reduce the overall number of accidents. In essence, the policy no longer focuses on collision risks but instead on personal injury risks. However, many existing implementations rely primarily on proximity-based measures, which fail to capture the actual severity of traffic events. To align more closely with Safe System's objectives, it is necessary to incorporate measures of potential consequences (if a crash does occur) alongside proximity indicators in order to more accurately evaluate safety outcomes.

These measures of potential consequences do not necessarily need to rely on overly complex computations, which can often hinder both accurate calculation and interpretation. In some studies, for instance, incompatible SMoS measures are combined into single indexes without clear logic or theoretical grounds. Another example is the various advanced, sophisticated algorithms designed to predict traffic risks or estimate injury numbers. If they generate results that are challenging to interpret, it can lead to ineffective or misguided safety interventions. Policies or countermeasures based on these misrepresented outcomes could then be directed toward the wrong priorities, ultimately failing to tackle the most critical traffic safety concerns.

Excessive complexity can also obscure potential biases or inaccuracies, making it difficult to identify and correct errors that distort severity rankings. If these methods lack transparency or have not been validated against theoretical expectations, policymakers risk misinterpreting data, leading to misguided resource allocation and ineffective traffic safety strategies.

To address these challenges, Patterson (2021), for example, suggests that a straightforward proximity metric, such as minimum physical distance and consequence reflection, in the form of Delta-V could be an effective approach to enable Vision Zero to capture the risk of severe traffic events. Of course, additional work will be needed to perfect the analytical method, but unnecessary complexity should be avoided, ensuring that clarity and practical applicability remain central.

5.7 Subjective component as a tool for measuring severity

The reliability of human judgement—the ability to produce consistent accuracy in measurements, regardless of location, conditions, or the person performing the measurement—has been a topic of debate in both academic and practical contexts, with some arguing that human assessments can lack consistency and may be influenced by cognitive biases, personal experiences, and situational stress, among other factors. However, extensive empirical studies conducted since the early 1980s have aimed to validate the reliability of human assessments in determining traffic event severity, demonstrating that human judgements produce consistent results across various geographical and temporal contexts.

At present, humans may still be the most reliable tool for reflecting the severity of traffic events. Their assessments tend to provide a rating that aligns more closely with the theoretical true values compared to the objective measures currently available. Unlike automated systems that depend solely on quantitative data, humans have the capacity to evaluate qualitative factors such as emotional context and environmental conditions. They possess an intuitive understanding of danger and its intensity, which strengthens their ability to assess the situation comprehensively.

The findings from *Paper IV* of this thesis further support the notion that the subjective component of human judgement can serve as a valuable calibration tool for automated systems, helping us move closer to 'true' severity ratings.

5.8 Contributions of the thesis

The primary aim of this thesis was to explore the use of microscopic traffic data extracted from video recordings to study and analyse road user behaviour. The focus was on improving the experience of road users, particularly VRUs, within traffic environments. This thesis serves several purposes, including evaluating road users' safety and examining the dynamic behaviours of road users in everyday traffic situations, which are closely linked to levels of service. By demonstrating various applications and new methods, the thesis highlights the value of microscopic data in gaining deeper insights into road user behaviour.

One of the main focuses of this work is safety, particularly for VRUs. Maintaining a balance between objective and subjective safety is essential. When subjective safety is lower than objective safety, it can make green modes of transport less attractive, leading to sustainability challenges. Conversely, when subjective safety exceeds objective safety, the situation is even more concerning, as road users may perceive the traffic environment to be safer than it actually is, increasing the risk of serious injuries or fatalities. To address these issues, reliable methods are needed to continuously monitor and improve safety features. This thesis attempted to contribute to meeting this need by advancing the SMoS approach through its integration with the Safe System vision as the underlying theoretical framework.

Beyond safety, there are other critical needs related to traffic studies., such as assessing the level of service, which still includes safety aspects but places greater emphasis on the operational performance of road facilities. This thesis introduces a new methodology to address this requirement that encompasses the entire process from data collection to data analysis. While I acknowledge that there is significant room for improvement, this work represents a promising starting point. If integrated into real-world applications and used by practitioners for large-scale projects, this methodology has the potential to deliver valuable results. It can enable the quick evaluation of various interventions on traffic environment features (e.g. road facility design elements, infrastructure, signage, electrical lighting), identify the need for adjustments, and contribute to a systematic record of 'what works' and 'what does not'. This can help track challenges and highlight necessary considerations that may arise during implementation, ensuring that adequate planning processes are in place to anticipate and mitigate potential obstacles.

Like most research, this thesis is not without its limitations. These include the use of semi-automated tools for trajectory extraction, which was time and resource-consuming; the small sample sizes used for statistical analyses; and the reliance on simple scenarios for road user interaction analysis, among others. Achieving 'ideal conditions' in real-world situations is often impractical, as perfection is an elusive concept, particularly in research. However, while these limitations present challenges, they also offer opportunities for improvement, innovation, and progress.

5.9 Looking to the future

Following current rapid trends in the development of artificial intelligence (AI), we are moving toward a future where advanced technology will be part of every aspect of our lives. In the field of transportation, this shift is very clear. Traditional methods for managing traffic often rely on fixed rules and simple assumptions. However, these methods cannot easily handle the complex and constantly changing conditions of a busy traffic environment. This is where large-scale models (or massive neural networks (NNs)) and reinforcement learning (RL) approaches become valuable.

Looking ahead, RL methods will enable traffic management systems to learn directly from their environments. Rather than relying on fixed rules, these systems will monitor road conditions, test various strategies, and continuously improve their performance. Upcoming research can focus on making these RL systems more flexible and better at predicting and preventing unexpected events, such as accidents. By using real-time data—especially from autonomous vehicles—to provide new and more detailed information, researchers can develop smarter systems that reduce congestion, improve safety, and lower emissions. Ultimately, these advancements can lead to traffic management solutions that are not only more efficient but also more sustainable and centred on human needs, improving urban mobility for everyone.

Collecting high-quality data is a key goal for achieving this, as AI systems' output is ultimately driven by the input data they receive—the more data they receive, the more they can learn. As demonstrated in this thesis, video recordings can capture not only the number of road users but also provide detailed insights into their behaviours at a micro level. By offering a clear and continuous view of traffic conditions, such video inputs can significantly strengthen the 'intelligence' of RL-based systems. Future studies can explore ways to integrate this detailed information into RL models to facilitate their understanding of their environment and effective decision-making.

The approaches introduced in this thesis are quite complex, involving multiple steps, some of which could potentially be streamlined or improved through the use of AI. Reinforcement learning, for instance, could be used to detect behavioural patterns that current indicators fail to capture. As these systems evolve, the ability to automate and optimise certain processes will only increase, providing an opportunity to further refine traffic management strategies.

In my view, even as AI systems improve over time, continuously learning from human input and improving their mimicry of the human cognitive system, humans will still remain at the top of the hierarchy, largely because we think holistically –as confirmed by my studies. We combine context, intuition, abstract reasoning, experience, knowledge, cultural understanding, and even moral judgments when making decisions. While the combination of NNs and RL will undoubtedly accelerate progress, I believe that humans will still oversee quality control and guide the implementation of these systems to ensure they align with human values, ethical considerations, safety, and long-term goals. Ultimately, AI-generated solutions will still be subject to human review before final decisions are made. In this sense, humans serve not simply as 'quality controllers' but also as the decision-makers about the directions in which AI systems should evolve.

That said, the pace of AI advancements is rapid, and the future holds many possibilities. Let's see what happens tomorrow.

6 Closing remarks

While extracting rich microscopic data from video recordings for advanced analysis has become increasingly feasible thanks to recent developments in techniques, several obstacles still hinder the accuracy and effectiveness of the results. To ensure that the extracted data are meaningful and accurate, careful and precise planning is required at every stage of the process, as outlined in Figure 1.

Data derived from video recordings are subject to various external factors that might sometimes be beyond the control of researchers and practitioners. Reflexivity, the inherent randomness of human behaviour, variations in vehicle types (e.g., regular bicycles vs. e-bikes), and unforeseen circumstances all introduce variability that can significantly influence outcomes. These factors cannot always be fully assessed using video recordings alone, highlighting the need for a comprehensive approach that accounts for such complexities when interpreting the data.

Clarity, validation, and theoretical alignment are vital to accurately representing real-world situations for traffic analysis. These principles form the foundation for generating reliable and actionable insights that inform effective interventions. However, the systematic evaluation and documentation of design solutions that address road users' safety and comfort remain underdeveloped. To address this gap, a comprehensive system for evaluating and registering which solutions work and which do not should be established, providing valuable insights and a solid foundation for informed decision-making. This system could facilitate the development of user-centric infrastructure and improve the overall impact of traffic interventions.

At the same time, researchers must design methodologies that balance reliability with practical usability, ensuring outputs are both accurate and accessible to practitioners. Beyond providing results and recommendations, it is necessary to offer theoretical guidance to equip practitioners with the tools they need to interpret and apply findings effectively. This approach not only fosters deeper understanding but also ensures that solutions are both actionable and impactful.

The successful application of video-based methods in traffic management depends on the collaboration of three critical disciplines: computer vision specialists, traffic research experts, and civil engineering practitioners. The effective integration of these domains depends on a well-structured process that ensures a cohesive and efficient synthesis of expertise. Improved collaboration would lead to the development of more accurate, actionable, and practical traffic management solutions, aligning technical precision with the needs of real-world implementation.

Simplification is another critical factor for fostering understanding, particularly in selecting appropriate metrics for traffic analysis. While straightforward and interpretable measures are beneficial, simplicity should not come at the expense of analytical depth and intelligence. Video-based analysis holds potential far beyond that of basic metrics like speed or traffic counts, offering deeper insights into road user behaviour, safety, and infrastructure usability. Through improved collaboration and communication across the key domains, these opportunities can be fully realised, leading to more effective, data-driven, and people-centred decision-making in the field of traffic management.

7 References

Aarts, L., I. van Schagen (2006) 'Driving speed and the risk of road crashes: A review'. *Accident Analysis & Prevention*, 38 (2), 215–224: https://doi.org/10.1016/j.aap.2005.07.004

Abdel-Aty, M., Z. Wang, O. Zheng, A. Abdelraouf (2023) 'Advances and applications of computer vision techniques in vehicle trajectory generation and surrogate traffic safety indicators'. *Accident Analysis & Prevention*, 191, 107191: https://doi.org/10.1016/j.aap.2023.107191

Abdel-Aziz, Y. I., H. M. Karara (2015) 'Direct linear transformation from comparator coordinates into object space coordinates in close-range photogrammetry'. *Photogrammetric Engineering & Remote Sensing*, 81 (2), 103–107: https://doi.org/10.14358/pers.81.2.103

Agerholm, N., C. Tønning, T. K. O. Madsen, C. H. Bahnsen, T. B. Moeslund, H. S. Lahrmann (2017) 'Road user behaviour analyses based on video detections: Status and best practice examples from the RUBA software' *24th ITS World Congress*, Montreal, Canada, 29 October–2 November 2017: https://vbn.aau.dk/ws/portalfiles/portal/273569946/Road_user_behaviour_analyses_based_on_video_detections_Status_and_best_practice_examples_from_the_RUBA_software.pdf

Ahrnbom, M. (2022) 'Computer vision for automated traffic safety assessment: A machine learning approach'. PhD thesis, *Lund University*: https://www.maths.lth.se/matematiklth/personal/ahrnbom/dokt_thesis.pdf

Ahrnbom, M., M. Nilsson, H. Ardö, K. Åström, O. Yastremska-Kravchenko, A. Laureshyn (2021a) 'Calibration and absolute pose estimation of trinocular linear camera array for Smart City applications' *25th International Conference on Pattern Recognition (ICPR)*, Milan, Italy, 10–15 January 2021: https://doi.org/10.1109/icpr48806.2021.9412816

Ahrnbom, M., M. G. Nilsson, H. Ardö (2021b) 'Real-time and online segmentation multi-target tracking with track revival re-identification' *16th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, Online, 8–10 February 2021: https://doi.org/10.5220/0010190907770784

Al-Smadi, M. m., K. Abdulrahim, R. Abdul Salam (2016) 'Cumulative frame differencing for urban vehicle detection' *First International Workshop on Pattern Recognition*, Tokyo, Japan, 11–13 May 2016: https://doi.org/10.1117/12.2242959

Alfarano, A., L. Maiano, L. Papa, I. Amerini (2024) 'Estimating optical flow: A comprehensive review of the state of the art'. *Computer vision and image understanding*, 249, 104160: https://doi.org/10.1016/j.cviu.2024.104160

Allen, B. L., B. T. Shin, P. Cooper (1978) 'Analysis of traffic conflicts and collisions'. *Transportation research record* (667), 67–74: https://onlinepubs.trb.org/Onlinepubs/trr/1978/667/667-009.pdf

Arlot, S., A. Celisse (2010) 'A survey of cross-validation procedures for model selection'. *Statistics Surveys*, 4, 40–79: https://doi.org/10.1214/09-SS054

Bagdadi, O. (2013) 'Estimation of the severity of safety critical events'. *Accident Analysis & Prevention*, 50, 167–174: https://doi.org/10.1016/j.aap.2012.04.007

Barone, F., M. Marrazzo, C. J. Oton (2020) 'Camera calibration with weighted direct linear transformation and anisotropic uncertainties of image control points'. *Sensors*, 20 (4), 1175: https://doi.org/10.3390/s20041175

Battiato, S., G. Gallo, G. Puglisi, S. Scellato (2007) 'SIFT features tracking for video stabilization' *14th International Conference on Image Analysis and Processing (ICIAP)*, Modena, Italy, 10–14 September 2007: https://doi.org/10.1109/iciap.2007.4362878

Bay, H., A. Ess, T. Tuytelaars, L. Van Gool (2008) 'Speeded-up robust features (SURF)'. *Computer vision and image understanding*, 110 (3), 346–359: https://doi.org/10.1016/j.cviu.2007.09.014

Bevans, R. (2023) 'Two-way ANOVA | Examples & when to use it'. *Scribbr*: https://www.scribbr.com/statistics/two-way-anova/ (*Acc.* 2024-12-20)

Bewley, A., Z. Ge, L. Ott, F. Ramos, B. Upcroft (2016) 'Simple online and realtime tracking' *IEEE International Conference on Image Processing (ICIP)*, Phoenix, AZ, USA, 25–28 September 2016: https://doi.org/10.1109/icip.2016.7533003

Bisio, I., C. Garibotto, H. Haleem, F. Lavagetto, A. Sciarrone (2022) 'A systematic review of drone based road traffic monitoring system'. *IEEE*, 10, 101537-101555: https://doi.org/10.1109/access.2022.3207282

Brown, D. (1996) 'Decentering distortion of lenses'. *Photogrammetric engineering*, 32 (3), 444–462: http://close-range.com/docs/Decentering_Distortion_of_Lenses_Brown_1966_may_444-462.pdf

Brown, G. R. (1994) 'Traffic conflicts for road user safety studies'. *Canadian Journal of Civil Engineering*, 21 (1), 1–15: https://www.ictct.net/wp-content/uploads/SMoS_Library/LIB_Brown_1994.pdf

Brown, M., D. G. Lowe (2007) 'Automatic panoramic image stitching using invariant features'. *International Journal of Computer Vision* 74, 59–73: https://doi.org/10.1007/s11263-006-0002-3

Cao, M., W. Zhu, M. Barth (2011) 'Mobile traffic surveillance system for dynamic roadway and vehicle traffic data integration' *14th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, Washington, DC, USA, 5–7 October 2011: https://doi.org/10.1109/itsc.2011.6083096

Cavanaugh, J. E., A. A. Neath (2019) 'The Akaike information criterion: Background, derivation, properties, application, interpretation, and refinements'. *Wiley Interdisciplinary Reviews*, 11 (3), e1460: https://doi.org/10.1002/wics.1460

Chang, H.-C., S.-H. Lai, K.-R. Lu (2004) 'A robust and efficient video stabilization algorithm' *IEEE International Conference on Multimedia and Expo (ICME)*, Taipei, Taiwan, 27–30 June 2004: https://doi.org/10.1109/icme.2004.1394117

Cronbach, L. J. (1951) 'Coefficient alpha and the internal structure of tests'. *Psychometrika*, 16, 297–334: https://doi.org/10.1007/bf02310555

Dantsuji, T., N. H. Hoang, N. Zheng, H. L. Vu (2022) 'A novel metamodel-based framework for large-scale dynamic origin–destination demand calibration'. *Transportation research part C: emerging technologies*, 136, 103545: https://doi.org/10.1016/j.trc.2021.103545

Davis, G. A., J. Hourdos, H. Xiong, I. Chatterjee (2011) 'Outline for a causal model of traffic conflicts and crashes'. *Accident Analysis & Prevention*, 43 (6), 1907–1919: https://doi.org/10.1016/j.aap.2011.05.001

Delaney, H., G. Parkhurst, S. Melia (2017) 'Walking and cycling on shared-use paths: The user perspective'. *Proceedings of the Institution of Civil Engineers: Municipal Engineer*, 170 (3), 175–184: https://doi.org/10.1680/jmuen.16.00033

Diwakar, P., V. S. Landge, U. Jain (2024) 'A pro-active approach for assessment of vehicle-vehicle interaction at unsignalized intersection in mixed traffic conditions'. *Innovative Infrastructure Solutions*, 9, 296: https://doi.org/10.1007/s41062-024-01608-0

Dong, J., Y. Xia, Q. Yu, A. Su, W. Hou (2014) 'Instantaneous video stabilization for unmanned aerial vehicles'. *Journal of Electronic Imaging*, 23 (1), 013002: https://doi.org/10.1117/1.jei.23.1.013002

Duffin, K. L., W. A. Barrett (2001) 'Fast focal length solution in partial panoramic image stitching' *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Kauai, HI, USA, 8–14 December 2001: https://doi.org/10.1109/cvpr.2001.991031

Eby, D. W. (2011) 'Naturalistic observational field techniques for traffic psychology research' *Handbook of Traffic Psychology*: Elsevier

Edvardsson Björnberg, K., S. O. Hansson, M.-Å. Belin, C. Tingvall (2022) 'The vision zero handbook: Theory, technology and management for a zero casualty policy'. Springer Nature

Ekman, L. (2023) 'Good practice in building safer roads, Sweden'. *Trafikverket*: https://etsc.eu/wp-content/uploads/05.-Lars_Ekman_Good_practice_Sweden.pdf

El-Basyouny, K., T. Sayed (2013) 'Safety performance functions using traffic conflicts'. *Safety Science*, 51 (1), 160–164: https://doi.org/10.1016/j.ssci.2012.04.015

Elvik, R. (2004) 'To what extent can theory account for the findings of road safety evaluation studies?'. *Accid Anal Prev*, 36 (5), 841-849: 10.1016/j.aap.2003.08.003

Elvik, R. (2010) 'Why some road safety problems are more difficult to solve than others'. *Accident Analysis & Prevention*, 42 (4), 1089–1096: https://doi.org/10.1016/j.aap.2009.12.020

Elvik, R. (2023) 'What would a road safety policy fully consistent with safe system principles mean for road safety?'. *Accident Analysis & Prevention*, 193, 107336: https://doi.org/10.1016/j.aap.2023.107336

Elvik, R., A. Høye, T. Vaa, M. Sørensen (2009) 'The Handbook of Road Safety Measures'. Bingley, UK: Emerald Group Publishing Limited

Elvik, R., T.-O. Nævestad (2023) 'Does empirical evidence support the effectiveness of the Safe System approach to road safety management?'. *Accident Analysis & Prevention*, 191, 107227: https://doi.org/10.1016/j.aap.2023.107227

EU Commission (2020) 'InDeV: In-Depth understanding of accident causation for Vulnerable road users': https://ec.europa.eu/inea/en/horizon-2020/projects/h2020-transport/safety/indev (*Acc.* 2024-12-20)

Evans, L. (1994) 'Driver injury and fatality risk in two-car crashes versus mass ratio inferred using Newtonian mechanics'. *Accident Analysis & Prevention*, 26 (5), 609–616: https://doi.org/10.1016/0001-4575(94)90022-1

Evans, L. (1999) 'Antilock brake systems and risk of different types of crashes in traffic'. *Journal of Crash Prevention and Injury Control*, 1 (1), 5–23: https://doi.org/10.1080/10286589908915737

Fisher, B. S., J. L. Nasar (1992) 'Fear of crime in relation to three exterior site features: Prospect, refuge, and escape'. *Environment and Behavior*, 24 (1), 35–65: https://doi.org/10.1177/0013916592241002

Franěk, M. (2012) 'Pedestrian walking speed as a response to environmental properties' *5th WSEAS International Conference on Urban Rehabilitation and Sustainability*, Sliema, Malta, 7–9 September 2012: https://doi.org/10.2466/06.50.pms.116.3.992-1019

Gabauer, D. J., H. C. Gabler (2008) 'Comparison of roadside crash injury metrics using event data recorders'. *Accident Analysis & Prevention*, 40 (2), 548–558: https://doi.org/10.1016/j.aap.2007.08.011

Garcia-Garcia, B., T. Bouwmans, A. J. R. Silva (2020) 'Background subtraction in real applications: Challenges, current models and future directions'. *Computer Science Review*, 35, 100204: https://doi.org/10.1016/j.cosrev.2019.100204

GDPR (2016) 'General Data Protection Regulation 2016/679 of 27 April 2016', European Union: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679

Geoerg, P., A. K. Boomers, M. Berthiaume, M. Boltes, M. Kinateder (2023) 'Exploring perceived vulnerability of pedestrians: Insights from a forced-choice experiment'. *arXiv preprint arXiv:2310.19524*: https://arxiv.org/abs/2310.19524

Gettman, D., L. Pu, T. Sayed, S. G. Shelby, S. Energy (2008) 'Surrogate safety assessment model and validation: Final report', Turner-Fairbank Highway Research Center, FHWA-HRt-08-051: https://www.fhwa.dot.gov/publications/research/safety/08051/08051.pdf

Ghahremannezhad, H., H. Shi, C. Liu (2023) 'Object detection in traffic videos: A survey'. *IEEE*, 24 (7), 6780–6799: https://doi.org/10.1109/tits.2023.3258683

Gildea, K., D. Hall, C. R. Cherry, C. Simms (2024) 'Forward dynamics computational modelling of a cyclist fall with the inclusion of protective response using deep learning-based human pose estimation'. *Journal of Biomechanics*, 163, 111959: https://doi.org/10.1016/j.jbiomech.2024.111959

Gildea, K., C. Mercadal-Baudart, R. Blythman, A. Smolic, C. Simms (2022) 'KinePose: A temporally optimized inverse kinematics technique for 6DOF human pose estimation with biomechanical constraints'. *arXiv preprint arXiv:2207.12841*: https://arxiv.org/abs/2207.12841

Girshick, R. (2015) 'Fast r-cnn' *IEEE international conference on computer vision*, Santiago, Chile, 7–13 December 2015: https://doi.org/10.1109/iccv.2015.169

Girshick, R., J. Donahue, T. Darrell, J. Malik (2014) 'Rich feature hierarchies for accurate object detection and semantic segmentation' *IEEE conference on computer vision and pattern recognition*, Columbus, OH, USA, 23–28 June 2014: https://doi.org/10.1109/cvpr.2014.81

Gohari, A., A. B. Ahmad, R. B. A. Rahim, N. I. M. Elamin, M. S. M. Gismalla, O. O. Oluwatosin, R. Hasan, A. S. A. Latip, A. Lawal (2023) 'Drones for road accident management: A systematic review'. *IEEE*, 11, 109247–109256: https://doi.org/10.1109/access.2023.3321142

Greibe, P., T. Skallebæk Buch (2016) 'Capacity and behaviour on one-way cycle tracks of different widths'. *Transportation research procedia*, 15, 122–136: https://doi.org/10.1016/j.trpro.2016.06.011

Gruen, A., T. Huang (2001) 'Calibration and Orientation of cameras in Computer Vision'. Springer Berlin, Heidelberg

Guo, C., C. Li, J. Guo, C. C. Loy, J. Hou, S. Kwong, R. Cong (2020) 'Zero-reference deep curve estimation for low-light image enhancement' *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA, USA, 13–19 June 2020: https://doi.org/10.1109/cvpr42600.2020.00185

Guo, Y., T. Sayed, M. H. Zaki (2018) 'Exploring evasive action based indicators for PTW conflicts in shared traffic facility environments'. *Journal of Transportation Engineering, Part A: Systems*, 144 (11): https://doi.org/10.1061/jtepbs.0000190

Güttinger, V. (1982) 'From accidents to conflicts: alternative safety measurement' *3rd International Workshop on Traffic Conflict Techniques*, Leidschendam, The Netherlands: https://www.ictct.net/wp-content/uploads/XX-Leidschendam-1982/1982-Proceedings.pdf

Harms, P. L. (1993) 'Crash injury investigation and injury mechanisms in road traffic accidents'. London, UK: Her Majesty's Stationary Office

Hauer, E. (1997) 'Observational before/after studies in road safety. Estimating the effect of highway and traffic engineering measures on road safety'. Bingley, UK: Emerald Group Publishing Limited

Hayward (1972) 'Near miss determination through use of a scale of danger'. *Pennsylvania Transportation and Traffic Safety Center*: https://onlinepubs.trb.org/Onlinepubs/hrr/1972/384/384-004.pdf

Heikkila, J., O. Silvén (1997) 'A four-step camera calibration procedure with implicit image correction' *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 17–19 June 1997: https://doi.org/10.1109/cvpr.1997.609468

Hillemann, M., M. Weinmann, M. S. Mueller, B. Jutzi (2019) 'Automatic extrinsic self-calibration of mobile mapping systems based on geometric 3D features'. *Remote sensing*, 11 (16), 1955: https://doi.org/10.3390/rs11161955

Horswill, M. S. (2016) 'Improving fitness to drive: The case for hazard perception training'. *Australian Psychologist*, 51 (3), 173–181: https://doi.org/10.1111/ap.12132

Hussein, M., T. Sayed (2015) 'Microscopic pedestrian interaction behavior analysis using gait parameters'. *Transportation research record*, 2519 (1), 28–38: https://doi.org/10.3141/2519-04

Hyden, C. (1977) 'A traffic conflicts technique for examining urban intersection problems' *Proceedings of the first workshop on traffic conflicts*

Hydén, C. (1987) 'The development of a method for traffic safety evaluation: The Swedish Traffic Conflicts Technique'. *Bulletin Lund Institute of Technology* (70): https://www.scirp.org/reference/referencespapers?referenceid=2384888

IMY (2024) 'The Swedish Authority for Privacy Protection (IMY)': https://www.imy.se/en/ (*Acc.* 10-14 2024)

Iskakov, K., E. Burkov, V. Lempitsky, Y. Malkov (2019) 'Learnable triangulation of human pose' *IEEE/CVF International Conference on Computer Vision (ICCV)*, Seoul, Korea, 27 October–2 November 2019: https://doi.org/10.1109/iccv.2019.00781

Jensen, M. B., M. Ahrnbom, M. Kruithof, K. Åström, M. Nilsson, H. Ardö, A. Laureshyn, C. Johnsson, T. Moeslund (2019) 'A framework for automated traffic safety analysis from video using modern computer vision'. *Transportation Research Board 98th Annual Meeting*, Washington, US, 13–17 January 2019: https://lucris.lub.lu.se/ws/portalfiles/portal/102387560/3_strudl.pdf

Jensen, M. L. (2016) 'Can the Delphi method complement or replace time-based measuring of conflicts between vulnerable road users?'. *ICTCT workshop*, Via Trafik Rådgivning: https://ictct.net/wp-content/uploads/29-Lund-2016/29-Jensen-Presentation.pdf

Jiang, Y., X. Gong, D. Liu, Y. Cheng, C. Fang, X. Shen, J. Yang, P. Zhou, Z. Wang (2021) 'Enlightengan: Deep light enhancement without paired supervision'. *IEEE transactions on image processing*, 30, 2340–2349

Johansson, M., A. Laureshyn, M. Nilsson (2020) 'Video analysis of pedestrian movement (VAPM) under different lighting conditions—Method exploration'. *Energies*, 13 (16), 4141: https://doi.org/10.3390/en13164141

Johansson, R. (2009) 'Vision Zero–Implementing a policy for traffic safety'. *Safety Science*, 47 (6), 826–831: https://doi.org/10.1016/j.ssci.2008.10.023

Johnsson, C. (2020) 'Surrogate Measures of Safety with a focus on vulnerable road users: An exploration of theory, practice, exposure, and validity'. PhD thesis, *Lund University*, Faculty of Engineering, Technology and Society, Transport and Roads

Johnsson, C., A. Laureshyn (2022) 'Identification of evasive manoeuvres in traffic interactions and conflicts'. *Traffic safety research*, 3, 000012: https://doi.org/10.55329/erqd8683

Johnsson, C., A. Laureshyn (2024) 'Worst case analysis: A video analysis approach to injury potential in traffic'. *Journal of Road Safety*, 35 (4), 10–18: https://doi.org/10.33492/JRS-D-24-4-2450246

Johnsson, C., A. Laureshyn, C. Dágostino (2021) 'Validation of surrogate measures of safety with a focus on bicyclist–motor vehicle interactions'. *Accident Analysis & Prevention*, 153, 106037: https://doi.org/10.1016/j.aap.2021.106037

Johnsson, C., A. Laureshyn, T. De Ceunynck (2018) 'In search of surrogate safety indicators for vulnerable road users: a review of surrogate safety indicators AU - Johnsson, Carl'. *Transport Reviews*, 38 (6), 765–785: https://doi.org/10.1080/01441647.2018.1442888

Kalman, R. E. (1960) 'A new approach to linear filtering and prediction problems'. *Journal of Basic Engineering*, 82 (1), 35–45: https://doi.org/10.1115/1.3662552

Kamaluddin, N. A., M. F. Abd Rahman, A. Várhelyi (2019) 'Matching of police and hospital road crash casualty records—A data-linkage study in Malaysia'. *International journal of injury control and safety promotion*, 26 (1), 52–59: https://doi.org/10.1080/17457300.2018.1476385

Kathuria, A., P. Vedagiri (2020) 'Evaluating pedestrian vehicle interaction dynamics at un-signalized intersections: A proactive approach for safety analysis'. *Accident Analysis & Prevention*, 134, 105316: https://doi.org/10.1016/j.aap.2019.105316

Kazemzadeh, K., P. Bansal (2021) 'Electric bike navigation comfort in pedestrian crowds'. *Sustainable Cities and Society*, 69, 102841: https://doi.org/10.1016/j.scs.2021.102841

Kazemzadeh, K., A. Laureshyn, E. Ronchi, C. D'Agostino, L. W. Hiselius (2020) 'Electric bike navigation behaviour in pedestrian crowds'. *Travel behaviour and society*, 20, 114–121: https://doi.org/10.1016/j.tbs.2020.03.006

Kim, J.-K., S. Kim, G. F. Ulfarsson, L. A. Porrello (2007) 'Bicyclist injury severities in bicycle-motor vehicle accidents'. *Accident Analysis & Prevention*, 39 (2), 238–251: https://doi.org/10.1016/j.aap.2006.07.002

Ko, S.-J., S.-H. Lee, S.-W. Jeon, E.-S. Kang (1999) 'Fast digital image stabilizer based on gray-coded bit-plane matching'. *IEEE Transactions on Consumer Electronics (Cat. No.99CH36277)*, Los Angeles, US, 22–24 June 1999: https://doi.org/10.1109/icce.1999.785181

Kraay, J., A. Van Der Horst, S. Oppe (2013) 'Manual conflict observation technique DOCTOR (Dutch Objective Conflict Technique for Operation and Research)'. *SWOV*, Foundation Road safety for all: https://www.ictct.net/wp-content/uploads/SMoS_Library/LIB_Kraay_et_al_2013.pdf

Kruysse, H. W. (1991) 'The subjective evaluation of traffic conflicts based on an internal concept of dangerousness'. *Accident Analysis & Prevention*, 23 (1), 53–65: https://doi.org/10.1016/0001-4575(91)90035-4

Kruysse, H. W., G. Wijlhuizen (1992) 'Why are experts not better in judging the danger of filmed traffic conflicts?'. *Accident Analysis & Prevention*, 24 (3), 227–235: https://doi.org/10.1016/0001-4575(92)90002-Z

Kuhn, H. W. (1955) 'The Hungarian method for the assignment problem'. *Naval Research Logistics*, 52 (1), 7–21: https://doi.org/10.1002/nav.20053

Kuipers, R., A, F. Carlvika, J. Rahm, C.-W. Palmqvist (2024) 'A naturalistic study into the flow of alighting and boarding passengers of commuter trains'. *26th Euro Working Group on Transportation Meeting (EWGT)*, Lund, Sweden, 4–6 September 2024

Laureshyn, A. (2010) 'Application of automated video analysis to road user behaviour'. PhD thesis, *Lund University*, Transport and Roads: https://portal.research.lu.se/files/5337337/1522970.pdf

Laureshyn, A., T. De Ceunynck, C. Karlsson, Å. Svensson, S. Daniels (2017a) 'In search of the severity dimension of traffic events: Extended Delta-V as a traffic conflict indicator'. *Accident Analysis & Prevention*, 98, 46–56: https://doi.org/10.1016/j.aap.2016.09.026

Laureshyn, A., M. D. Goede, N. Saunier, A. Fyhri (2017b) 'Cross-comparison of three surrogate safety methods to diagnose cyclist safety problems at intersections in Norway'. *Accident Analysis & Prevention*, 105, 11–20: https://doi.org/10.1016/j.aap.2016.04.035

Laureshyn, A., C. Johnsson, T. D. Ceunynck, A. Svensson, M. de Goede, N. Saunier, P. Wlodarek, A. R. A. van der Horst, S. Daniels (2016) 'Scoping review: surrogate measures of safety in site-based road traffic observations'. *InDeV, H. p.*, Review of current study methods for VRU safety, part 4 https://lup.lub.lu.se/record/2a154419-d104-4a35-8dc1-1e6bdea244e4

Laureshyn, A., M. Nilsson (2018) 'How accurately can we measure from video? Practical considerations and enhancements of the camera calibration procedure'. *Transportation Research Board*, 2672 (43), 24–33: https://doi.org/10.1177/0361198118774194

Laureshyn, A., Å. Svensson, C. Hydén (2010) 'Evaluation of traffic safety, based on micro-level behavioural data: Theoretical framework and first implementation'. *Accident Analysis & Prevention*, 42 (6), 1637–1646: https://doi.org/10.1016/j.aap.2010.03.021

Laureshyn, A., A. Várhelyi (2018) 'The Swedish traffic conflict technique'. *Observer's manual, Lund*: https://lup.lub.lu.se/search/ws/files/51195704/TCT_Manual_2018.pdf

Law, H., J. Deng (2020) 'CornerNet: Detecting objects as paired keypoints'. *International journal of computer vision*, 128, 642–656: https://doi.org/10.1007/s11263-019-01204-1

Lee, K.-Y., Y.-Y. Chuang, B.-Y. Chen, M. Ouhyoung (2009) 'Video stabilization using robust feature trajectories' *IEEE 12th International Conference on Computer Vision*, Kyoto, Japan, 29 September–2 October 2009: https://doi.org/10.1109/iccv.2009.5459297

Li, Q., S. Shang, X. Pei, Q. Wang, Q. Zhou, B. Nie (2021) 'Kinetic and kinematic features of pedestrian avoidance behavior in motor vehicle conflicts'. *Frontiers in bioengineering and biotechnology*, 9, 783003: https://doi.org/10.3389/fbioe.2021.783003

Lightburn, A., C. Howarth (1979) 'A study of observer variability and reliability in the detection and grading of traffic conflicts' *Second international traffic conflicts technique workshop*, Paris, France, 10–12 May 1979: https://www.ictct.net/wp-content/uploads/XX-Paris-1979/1979-Lightburn-Howarth.pdf

Lin, K., S. Liu, L. F. Cheong, B. Zeng (2016) 'Seamless video stitching from hand-held camera inputs' *Computer Graphics Forum*: https://doi.org/10.1111/cgf.12848

Lin, T.-Y., P. Goyal, R. Girshick, K. He, P. Dollár (2017) 'Focal loss for dense object detection' *IEEE International Conference on Computer Vision (ICCV)*, Venice, Italy, 22–29 October 2017: https://doi.org/10.1109/iccv.2017.324

Litvin, A., J. Konrad, W. C. Karl (2003) 'Probabilistic video stabilization using Kalman filtering and mosaicing'. *Image and video communications and processing 2003*, Santa Clara, US, 21–24 January 2003: https://doi.org/10.1117/12.476436

Liu, W., D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, A. C. Berg (2016) 'Ssd: Single shot multibox detector'. *Computer Vision–ECCV 2016: 14th European Conference*, Amsterdam, The Netherlands, 11–14 October 2016: https://doi.org/10.1007/978-3-319-46448-0_2

Lord, D., F. Mannering (2010) 'The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives'. *Transportation Research Part A: Policy and Practice*, 44 (5), 291–305: https://doi.org/10.1016/j.tra.2010.02.001

Lowe, D. G. (2004) 'Distinctive image features from scale-invariant keypoints'. *International journal of computer vision*, 60, 91–110: https://doi.org/10.1023/b:visi.0000029664.99615.94

Lubbe, N., Y. Wu, H. Jeppsson (2022) 'Safe speeds: fatality and injury risks of pedestrians, cyclists, motorcyclists, and car drivers impacting the front of another passenger car as a function of closing speed and age'. *Traffic safety research*, 2, 000006: https://doi.org/10.55329/vfma7555

Madsen, T. K. O. (2018) 'Video Analysis and Mapping of Vulnerable Road Users' Safety', *Aalborg Universitetsforlag*: https://doi.org/10.5278/vbn.phd.eng.00057

Madsen, T. K. O., C. H. Bahnsen, M. B. Jensen, H. S. Lahrmann, T. B. Moeslund (2016) 'Watchdog System'. *European Commission*, Project No. 635895—InDeV: https://vbn.aau.dk/ws/portalfiles/portal/270264451/Watchdog_System.pdf

Mahmud, S. M. S., L. Ferreira, M. S. Hoque, A. Tavassoli (2017) 'Application of proximal surrogate indicators for safety evaluation: A review of recent developments and research needs'. *IATSS Research*, 41 (4), 153–163: https://doi.org/10.1016/j.iatssr.2017.02.001

Malkhamah, S., M. Tight, F. Montgomery (2005) 'The development of an automatic method of safety monitoring at Pelican crossings'. *Accident Analysis & Prevention*, 37 (5), 938–946: https://doi.org/10.1016/j.aap.2005.04.012

Mallikarjuna, C., A. Phanindra, K. R. Rao (2009) 'Traffic data collection under mixed traffic conditions using video image processing'. *Journal of Transportation Engineering*, 135 (4), 174–182: https://doi.org/10.1061/(asce)0733-947x(2009)135:4(174)

Matsushita, Y., E. Ofek, W. Ge, X. Tang, H.-Y. Shum (2006) 'Full-frame video stabilization with motion inpainting'. *IEEE Transactions on pattern analysis and Machine Intelligence*, 28 (7), 1150–1163: https://doi.org/10.1109/tpami.2006.141

McDonald, J. H. (2009) 'Handbook of biological statistics'. Baltimore, US: Sparky House Publishing

Mehrish, A., P. Singh, P. Jain, A. V. Subramanyam, M. Kankanhalli (2020) 'Egocentric analysis of dash-cam videos for vehicle forensics'. *IEEE Transactions on Circuits and Systems for Video Technology*, 30 (9), 3000–3014: https://doi.org/10.1109/tcsvt.2019.2929561

Melnikov, V., P. Gupta, B. Frick, D. Kaimann, E. Hüllermeier (2017) 'Pairwise versus pointwise ranking: A case study'. *Schedae Informaticae*, 25, 73–83: https://doi.org/10.4467/20838476si.16.006.6187

Mohamed, M. G., N. Saunier (2013) 'Motion prediction methods for surrogate safety analysis'. *Transportation research record*, 2386 (1), 168–178: https://doi.org/10.3141/2386-19

Morimoto, C., R. Chellappa (1998) 'Evaluation of image stabilization algorithms'. *IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP'98*, Seattle, US, 15 May 1998: https://doi.org/10.1109/icassp.1998.678102

Munigety, C. R., V. Vicraman, T. V. Mathew (2014) 'Semiautomated tool for extraction of microlevel traffic data from videographic survey'. *Transportation research record*, 2443 (1), 88–95: https://doi.org/10.3141/2443-10

Neath, A. A., J. E. Cavanaugh (2012) 'The Bayesian information criterion: background, derivation, and applications'. *WIREs Computational Statistics*, 4 (2), 199–203: https://doi.org/10.1002/wics.199

Nie, Y., T. Su, Z. Zhang, H. Sun, G. Li (2017) 'Dynamic video stitching via shakiness removing'. *IEEE transactions on image processing*, 27 (1), 164–178: https://doi.org/10.1109/tip.2017.2736603

Niewöhner, W., F. Roth, J. Gwehenberger, C. Gruber, M. Kuehn, R. Sferco, C.-H. Pastor, U. Nagel, M. Stanzel (2011) 'Proposal for a test procedure of assistance systems regarding preventive pedestrian protection' *22nd Enhanced Safety of Vehicles Conference*, Washington, US, 13–16 June 2011

Noland, R. B. (2013) 'From theory to practice in road safety policy: Understanding risk versus mobility'. *Research in Transportation Economics*, 43 (1), 71–84: https://doi.org/10.1016/j.retrec.2012.12.007

Outay, F., H. A. Mengash, M. Adnan (2020) 'Applications of unmanned aerial vehicle (UAV) in road safety, traffic and highway infrastructure management: Recent advances and challenges'. *Transportation Research Part A: Policy and Practice*, 141, 116–129: https://doi.org/10.1016/j.tra.2020.09.018

Pant, S., P. Nooralishahi, N. P. Avdelidis, C. Ibarra-Castanedo, M. Genest, S. Deane, J. J. Valdes, A. Zolotas, X. P. V. Maldague (2021) 'Evaluation and selection of video stabilization techniques for UAV-based active infrared thermography application'. *Sensors*, 21 (5), 1604: https://doi.org/10.3390/s21051604

Patterson, F. L.-H. (2021) 'The adverse effects of paradigm and pragmatism on road safety with case studies in traffic conflicts technique and cyclist safety at roundabouts'. PhD thesis, *University of Adelaide, School of Social Sciences*: https://hdl.handle.net/2440/132844

Perkins, S. R., J. I. Harris (1967) 'Traffic conflict characteristics: Freeway curve and exit area F1, December, 1966'. *Corporation, G. M.*

Polders, E., T. Brijs (2018) 'How to analyse accident causation? A handbook with focus on vulnarable road users'. Hasselt, Belgium: Hasselt University

Rahm, J., A. Niska, M. Johansson (2024) 'The environmental experience during daylight and dark conditions: a conceptual model for pedestrians and cyclists/La experiencia ambiental en condiciones de luz diurna y oscuridad: Un modelo conceptual para peatones y ciclistas'. *PsyEcology: Bilingual Journal of Environmental Psychology*, 15 (1), 85–113: https://doi.org/10.1177/21711976241232869

Rangesh, A., M. M. Trivedi (2020) 'Ground plane polling for 6DoF pose estimation of objects on the road'. *IEEE Transactions on Intelligent Vehicles*, 5 (3), 449–460: https://doi.org/10.1109/tiv.2020.2966074

Rani, S., D. Ghai, S. Kumar (2022) 'Object detection and recognition using contour based edge detection and fast R-CNN'. *Multimedia Tools and Applications*, 81, 42183–42207: https://doi.org/10.1007/s11042-021-11446-2

Rauf, R., A. R. Shahid, S. Ziauddin, A. A. Safi (2016) 'Pedestrian detection using HOG, LUV and optical flow as features with AdaBoost as classifier'. *6th International Conference on Image Processing Theory, Tools and Applications (IPTA)*, Oulu, Finland, 12–15 December 2016: https://doi.org/10.1109/ipta.2016.7821024

Reddy, N. D., L. Guigues, L. Pishchulin, J. Eledath, S. G. Narasimhan (2021) 'TesseTrack: End-to-end learnable multi-person articulated 3D pose tracking'. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Nashville, US, 20–25 June 2021: https://doi.org/10.1109/cvpr46437.2021.01494

Redmon, J., S. Divvala, R. Girshick, A. Farhadi (2016) 'You only look once: Unified, real-time object detection'. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, US, 27–30 June 2016: https://doi.org/10.1109/cvpr.2016.91

Ren, S., K. He, R. Girshick, J. Sun (2015) 'Faster R-CNN: Towards real-time object detection with region proposal networks'. *IEEE Transactions on pattern analysis and Machine Intelligence*, 39 (6), 1137–1149: https://doi.org/10.1109/tpami.2016.2577031

Rowe, B. H., A. M. Rowe, G. W. Bota (1995) 'Bicyclist and environmental factors associated with fatal bicycle-related trauma in Ontario'. *CMAJ: Canadian Medical Association Journal*, 152 (1), 45–53: https://www.cmaj.ca/content/152/1/45

Saunier, N. (2024) 'Traffic Intelligence': https://pypi.org/project/trafficintelligence/ (*Acc.* 10-24 2024)

Saunier, N., A. Laureshyn (2021) 'Surrogate Measures of Safety'. in Vickerman, R. (Ed.) *International Encyclopedia of Transportation*, Oxford, UK: Elsevier

Saunier, N., T. Sayed, K. Ismail (2010) 'Large-scale automated analysis of vehicle interactions and collisions'. *Transportation research record*, 2147 (1), 42–50: https://doi.org/10.3141/2147-06

Schieber, R. A., J. J. Sacks (2001) 'Measuring community bicycle helmet use among children'. *Public health reports*, 116 (2), 113–121: https://doi.org/10.1093/phr/116.2.113

Schneekloth, L. H., R. G. Shibley (1995) 'Placemaking: The art and practice of building communities'. *Choice Reviews Online*: https://doi.org/10.5860/choice.33-1144

Shelby, S. G. (2011) 'Delta-v as a measure of traffic conflict severity'. *3rd International Conference on Road Safety and Simulation*, Indianapolis, US, 14–16 September 2011: https://onlinepubs.trb.org/onlinepubs/conferences/2011/RSS/1/Shelby,S.pdf

Shen, A. (2016) 'BeaverDam: Video annotation tool for computer vision training labels'. Master thesis, *University of California*: https://www2.eecs.berkeley.edu/Pubs/TechRpts/2016/EECS-2016-193.pdf

St-Charles, P.-L., G.-A. Bilodeau, R. Bergevin (2015) 'A self-adjusting approach to change detection based on background word consensus'. *2015 IEEE Winter Conference on Applications of Computer Vision*, Waikoloa, US, 5–9 January 2015: https://doi.org/10.1109/wacv.2015.137

Su, T., Y. Nie, Z. Zhang, H. Sun, G. Li (2016) 'Video stitching for handheld inputs via combined video stabilization'. *SIGGRAPH ASIA 2016 Technical Briefs*, Macau, China, 5–8 December 2016: https://doi.org/10.1145/3005358.3005383

Svensson, Å. (1992) 'Vidareutveckling och validering av den svenska konflikttekniken'. *University, L.*, Department of Technology and Society

Svensson, Å. (1998) 'A method for analysing the traffic process in a safety perspective'. PhD thesis, *Lund University*, Bulletin 166: https://lucris.lub.lu.se/ws/portalfiles/portal/4456913/1653512.pdf

Svensson, Å., C. Hydén (2006) 'Estimating the severity of safety related behaviour'. *Accident Analysis & Prevention*, 38 (2), 379–385: https://doi.org/10.1016/j.aap.2005.10.009

SWOV (2018) 'Sustainable Safety 3rd edition – The advanced vision for 2018-2030': https://swov.nl/system/files/publication-downloads/dv3_en_kort_rapport.pdf

Szeliski, R. (2007) 'Image alignment and stitching: A tutorial'. *Foundations and Trends® in Computer Graphics and Vision*, 2 (1), 1–104: https://doi.org/10.1561/0600000009

T-Analyst (2019) 'Software for semi-automated video processing ': https://bitbucket.org/TrafficAndRoads/tanalyst/wiki/Home

Tageldin, A., T. Sayed (2016) 'Developing evasive action-based indicators for identifying pedestrian conflicts in less organized traffic environments'. *Journal of advanced transportation*, 50 (6), 1193–1208: https://doi.org/10.1002/atr.1397

Tageldin, A., T. Sayed, X. Wang (2015) 'Can time proximity measures be used as safety indicators in all driving cultures? Case study of motorcycle safety in China'. *Transportation research record*, 2520 (1), 165–174: https://doi.org/10.3141/2520-19

Tageldin, A., M. H. Zaki, T. Sayed (2017) 'Examining pedestrian evasive actions as a potential indicator for traffic conflicts'. *IET Intelligent Transport Systems*, 11 (5), 282–289: https://doi.org/10.1049/iet-its.2016.0066

Tarko, A. (2019) 'Measuring road safety with surrogate events'. Amsterdam, The Netherlands: Elsevier

Tarko, A., G. Davis, N. Saunier, T. Sayed, S. Washington (2009) 'White paper: surrogate measures of safety'. *Committee on Safety Data Evaluation and Analysis (ANB20)*: http://n.saunier.free.fr/saunier/stock/tarko09surrogate.pdf

Tarko, A. P. (2018) 'Surrogate measures of safety' *Safe Mobility: Challenges, Methodology and Solutions*, Leeds, England: Emerald Publishing Limited

Tarko, A. P. (2021) 'A unifying view on traffic conflicts and their connection with crashes'. *Accident Analysis & Prevention*, 158, 106187: https://doi.org/10.1016/j.aap.2021.106187

Terven, J., D.-M. Córdova-Esparza, J.-A. Romero-González (2023) 'A comprehensive review of YOLO architectures in computer vision: From YOLOv1 to YOLOv8 and YOLO-NAS'. *Machine Learning and Knowledge Extraction*, 5 (4), 1680–1716: https://doi.org/10.3390/make5040083

Tian, Z., X. Chu, X. Wang, X. Wei, C. Shen (2022) 'Fully convolutional one-stage 3D object detection on LiDAR range images'. *36th Conference on Neural Information Processing Systems (NeurIPS 2022)*, New Orleans, US, 28 November–9 December 2022: https://proceedings.neurips.cc/paper_files/paper/2022/file/e1f418450107c4a0ddc16d008d131573-Paper-Conference.pdf

Transportstyrelsen (2024) 'Drones–unmanned aircraft': https://www.transportstyrelsen.se/en/aviation/Aircraft/drones--unmanned-aircraft/ (*Acc.* 25 November 2024)

Tsai, R. Y. (1986) 'An efficient and accurate camera calibration technique for 3D machine vision'. *IEEE Conference on Computer Vision and Pattern Recognition*, Miami Beach, US, 22–26 June 1986

Tukey, J. W. (1949) 'Comparing individual means in the analysis of variance'. *Biometrics*, 5 (2), 99–114: https://doi.org/10.2307/3001913

Van der Horst, R., J. Kraay (1986) 'The Dutch Conflict Observation Technique–DOCTOR'. *Workshop "Traffic Conflicts and Other Intermediate Measures in Safety Evaluation"*, Budapest, Hungary, 8–10 September 1986

van Haperen, W., M. S. Riaz, S. Daniels, N. Saunier, T. Brijs, G. Wets (2019) 'Observing the observation of (vulnerable) road user behaviour and traffic safety: A scoping review'. *Accident Analysis & Prevention*, 123, 211–221: https://doi.org/10.1016/j.aap.2018.11.021

Van der Horst, A. (1990) 'A time-based analysis of road user behaviour in normal and critical encounters', TU Delft: https://resolver.tudelft.nl/uuid:8fb40be7-fae1-4481-bc37-12a7411b85c7

Varhelyi, A. (1998) 'Drivers' speed behaviour at a zebra crossing: A case study'. *Accident Analysis & Prevention*, 30 (6), 731–743: https://doi.org/10.1016/s0001-4575(98)00026-8

Wang, X., T. X. Han, S. Yan (2009) 'An HOG-LBP human detector with partial occlusion handling'. *IEEE 12th International Conference on Computer Vision*, Kyoto, Japan, 29 September 2009: https://doi.org/10.1109/iccv.2009.5459207

Wang, Y., Z. Hou, K. Leman, R. Chang (2011) 'Real-time video stabilization for unmanned aerial vehicles'. *MVA2011 IAPR Conference on Machine Vision Application*, Nara, Japan, 13–15 June 2011: https://www.mva-org.jp/Proceedings/2011CD/papers/09-27.pdf

Watson, A., B. Watson, K. Vallmuur (2015) 'Estimating under-reporting of road crash injuries to police using multiple linked data collections'. *Accident Analysis & Prevention*, 83, 18–25: https://doi.org/10.1016/j.aap.2015.06.011

Wegman, F., L. Aarts, P. van der Knaap (2022) 'Sustainable safety: a short history of a Safe System approach in the Netherlands' *The Vision Zero Handbook: Theory, Technology and Management for a Zero Casualty Policy*, Heidelberg, Germany: Springer

Wegman, F., A. Dijkstra, G. Schermers, P. van Vliet (2005) 'Sustainable Safety in the Netherlands: the vision, the implementation and the safety effects'. *SWOV*: https://swov.nl/system/files/publication-downloads/R-2005-05.pdf

Wei, L., Z. Zhong, C. Lang, Z. Yi (2019) 'A survey on image and video stitching'. *Virtual Reality & Intelligent Hardware*, 1 (1), 55–83: https://doi.org/10.3724/sp.j.2096-5796.2018.0008

Wilcox, R. R. (2003) 'Two-way ANOVA' *Applying Contemporary Statistical Techniques*, Burlington, US: Academic Press

Wojke, N., A. Bewley, D. Paulus (2017) 'Simple online and realtime tracking with a deep association metric'. *IEEE International Conference on Image Processing (ICIP)*, Beijing, China, 17–20 September 2017: https://doi.org/10.1109/icip.2017.8296962

Wood, J. M., P. F. Lacherez, R. P. Marszalek, M. J. King (2009) 'Drivers' and cyclists' experiences of sharing the road: Incidents, attitudes and perceptions of visibility'. *Accident Analysis & Prevention*, 41 (4), 772–776: https://doi.org/10.1016/j.aap.2009.03.014

Yaseen, M. (2024) 'What is YOLOv9: An in-depth exploration of the internal features of the next-generation object detector'. *arXiv preprint arXiv:2409.07813*: https://arxiv.org/abs/2409.07813

Yastremska-Kravchenko, O., A. Laureshyn, J. Rahm, M. Johansson, A. Niska, C. Johnsson, C. D'Agostino (2024) 'Video analysis of bicyclist and pedestrian movement on shared-use paths under daylight and electric lighting conditions—Method exploration'. *Journal of Cycling and Micromobility Research*, 2, 100032: https://doi.org/10.1016/j.jcmr.2024.100032

Yuan, Y., W. Daamen, B. Goñi-Ros, S. P. Hoogendoorn (2018) 'Investigating cyclist interaction behavior through a controlled laboratory experiment'. *Journal of Transport and Land Use*, 11 (1), 833–847: https://doi.org/10.5198/jtlu.2018.1155

Zhang, T., P. J. Jin (2019) 'A longitudinal scanline based vehicle trajectory reconstruction method for high-angle traffic video'. *Transportation research part C: emerging technologies*, 103, 104–128: https://doi.org/10.1016/j.trc.2019.03.015

Zhang, X., Y. Feng, P. Angeloudis, Y. Demiris (2022) 'Monocular visual traffic surveillance: A review'. *IEEE Transactions on Intelligent Transportation Systems*, 23 (9), 14148–14165: https://doi.org/10.1109/tits.2022.3147770

Zhang, X., W. Guo, Y. Xing, W. Wang, H. Yin, Y. Zhang (2023) 'AugFCOS: Augmented fully convolutional one-stage object detection network'. *Pattern Recognition*, 134, 109098: https://doi.org/10.1016/j.patcog.2022.109098

Zhang, Z. (2000) 'A flexible new technique for camera calibration'. *IEEE Transactions on pattern analysis and Machine Intelligence*, 22 (11), 1330–1334: https://doi.org/10.1109/34.888718

Zheng, L., K. Ismail, X. Meng (2014) 'Traffic conflict techniques for road safety analysis: Open questions and some insights'. *Canadian Journal of Civil Engineering*, 41 (7), 633–641: https://doi.org/10.1139/cjce-2013-0558

Zheng, L., T. Sayed, F. Mannering (2021) 'Modeling traffic conflicts for use in road safety analysis: A review of analytic methods and future directions'. *Analytic Methods in Accident Research*, 29, 100142: https://doi.org/10.1016/j.amar.2020.100142

Zheng, O., M. Abdel-Aty, L. Yue, A. Abdelraouf, Z. Wang, N. Mahmoud (2024) 'CitySim: a drone-based vehicle trajectory dataset for safety-oriented research and digital twins'. *Transportation research record*, 2678 (4), 606–621: https://doi.org/10.1177/03611981231185768

Zhou, X., J. Zhuo, P. Krahenbuhl (2019) 'Bottom-up object detection by grouping extreme and center points'. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, US, 15–20 June 2019: https://doi.org/10.1109/cvpr.2019.00094