Proactive Collision Risk Quantification in Multi-directional Traffic Interactions

Jiao, Y.

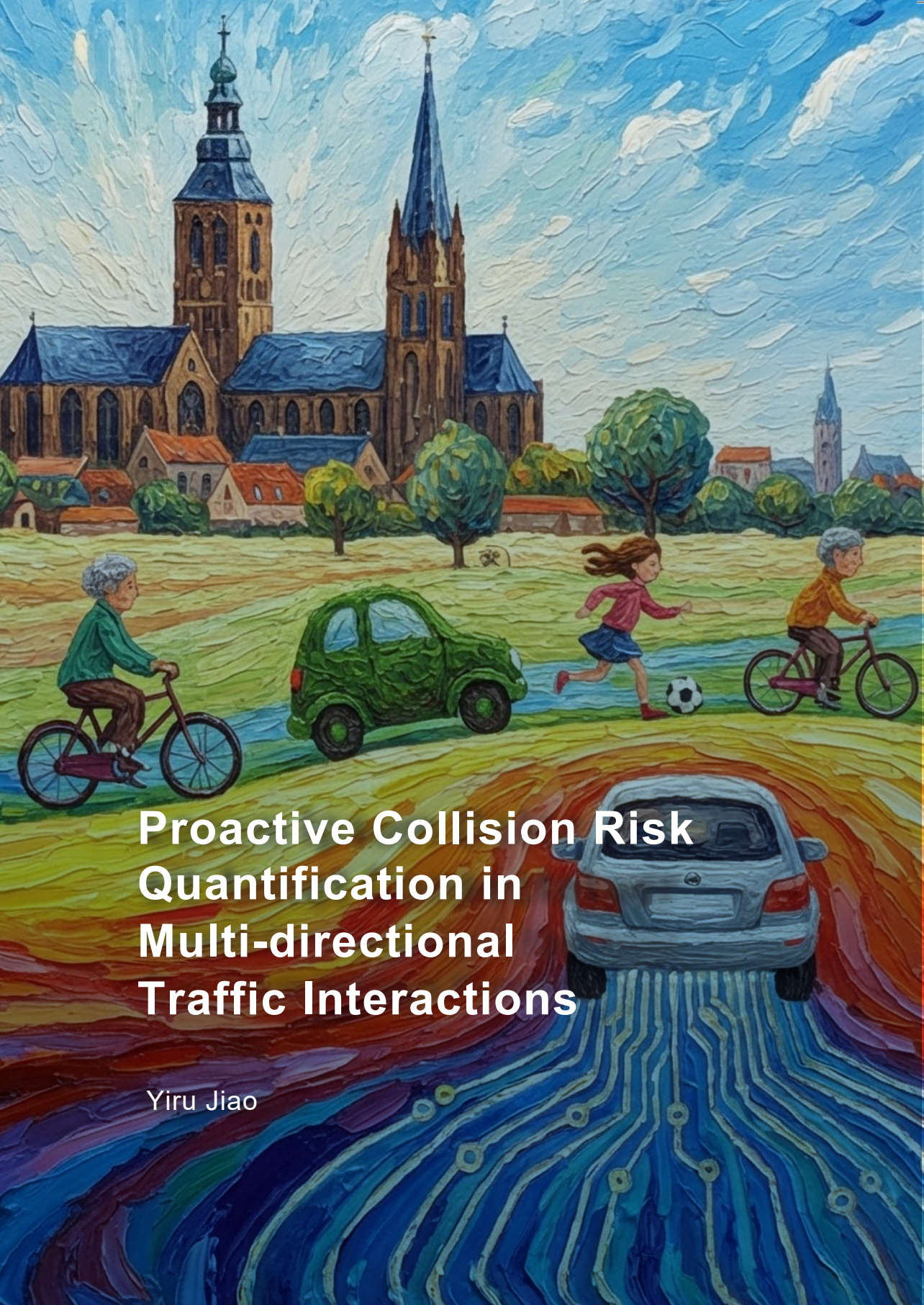
DOI
10.4233/uuid:ba7621cb-babf-42d0-85d6-a3ef97d3e0af

Publication date
2026

Document Version
Final published version

Citation (APA)
Jiao, Y. (2026). *Proactive Collision Risk Quantification in Multi-directional Traffic Interactions* (1 ed.). [Dissertation (TU Delft), Delft University of Technology]. TRAIL Research School. https://doi.org/10.4233/uuid:ba7621cb-babf-42d0-85d6-a3ef97d3e0af

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.



Proactive Collision Risk Quantification in Multi-directional Traffic Interactions

Yiru Jiao

Propositions

accompanying the dissertation

Proactive Collision Risk Quantification in Multi-directional Traffic Interactions

by

Yiru Jiao

1. A continuous measure of collision risk is required because there is no hard boundary between traffic conflicts and safe interactions.
   This proposition pertains to this dissertation (Chapter 3).
2. Traffic interactions that do not end in collisions provide more information than those that do for learning collision risk.
   This proposition pertains to this dissertation (Chapters 4&5).
3. Learning collision risk from unlabelled data overcomes the curse of rarity to achieve safe autonomous driving.
   This proposition pertains to this dissertation (Chapters 5&6).
4. To proactively avoid potential collisions, estimating their likelihood should be prioritised over estimating their severity.
   This proposition pertains to this dissertation.
5. Observable intelligence can be seen as a product of computation.
6. In research, being true to yourself is more helpful than conforming to the mainstream.
7. Research should be motivated by genuine care for society and the environment rather than mere curiosity.
8. Limited resources deter researchers from asking high-risk questions, while abundant resources tempt them to answer the questions less efficiently.
9. Most scientific journals should be replaced by platforms that mandate openly shared studies for reproducibility and transparent discussion.
10. Diversity in people's mental models, rather than in demographic identities such as gender, nationality, or age, brings better performance of a team.

These propositions are regarded as opposable and defendable, and have been approved as such by the promotors prof. dr. ir. J.W.C. van Lint, dr. ir. S. van Cranenburgh, and dr. ir. S.C. Calvert.

Proactive Collision Risk Quantification in Multi-directional Traffic Interactions

Proactive Collision Risk Quantification in Multi-directional Traffic Interactions

Dissertation

for the purpose of obtaining the degree of doctor
at Delft University of Technology
by the authority of the Rector Magnificus Prof. dr. ir. H. Bijl

chair of the Board for Doctorates
to be defended publicly on
Tuesday 20, January 2026 at 15:00 o'clock

by

Yiru JIAO

Master of Science in Management Science and Engineering,
Harbin Institute of Technology, China,
born in Gansu, China.

This dissertation has been approved by the promotors.

Composition of the doctoral committee:

Rector Magnificus	chairperson
Prof. dr. ir. J.W.C. van Lint	Delft University of Technology, promotor
Dr. ir. S. van Cranenburgh	Delft University of Technology, promotor
Dr. ir. S.C. Calvert	Delft University of Technology, copromotor

Independent members:

Prof. dr. M.P. Hagenzieker	Delft University of Technology
Prof. dr. ir. S. Ahn	University of Wisconsin–Madison, USA
Prof. dr. G. Markkula	University of Leeds, UK
Prof. dr. D. Work	Vanderbilt University, USA
Prof. dr. ir. S.P. Hoogendoorn	Delft University of Technology, reserve member

The research was supported by the TU Delft AI Labs programme. In CityAI Lab, data, AI, and behavioural theory come together for the development of more attractive and liveable cities.

An electronic copy of this dissertation is available at TU Delft Repository
https://doi.org/10.4233/uuid:ba7621cb-babf-42d0-85d6-a3ef97d3e0af.

Front & Back: designed by the author and generated with Gemini 3 & Nano Banana Pro.

TRAIL Thesis Series no. T2026/1, The Netherlands Research School TRAIL

TRAIL
P.O. BOX 5017
2600 GA Delft
The Netherlands
E-mail: info@rsTRAIL.nl

ISBN: 978-90-5584-377-0

Printed in the Netherlands

To all the stubborn souls true to themselves.

Contents

Foreword

The research in this thesis has been funded by the TU Delft AI Labs Programme, which offers the author a 5-year employment contract. This is one year longer than the standard Dutch PhD trajectory of 4 years to accommodate an additional commitment to teaching in artificial intelligence (AI), data science, and digitalisation.

The work has been facilitated by four interdisciplinary research environments:

- **CityAI Lab** – where data, machine learning and behavioural theory are combined to understand how cities shape, and are shaped by, human activities.
- **Delft Integrated Traffic & Travel Laboratory (DiTTLab)** – modelling large-scale traffic flow and fusing diverse data streams to spot network vulnerabilities before they manifest on the road.
- **Automated Driving & Simulation (ADaS) Lab** – advancing the safe integration of automated vehicles through high-fidelity simulation and interdisciplinary traffic-behaviour studies.
- **Digitisation & AI for Mobility Network Dynamics (DAIMoND) Lab** – pioneering data-driven, AI-empowered solutions for smarter, more resilient and sustainable transport systems.

The shared ethos of open science, data-centric experimentation, and societally relevant engineering in these labs has profoundly influenced both the ambition and the methodological choices that underpin the thesis.

While the support has been indispensable, all analyses, interpretations and conclusions presented in this thesis remain solely those of the author and do not necessarily reflect the official positions of TU Delft, the AI Labs Programme, or the individual labs.

Summary

Background

Road traffic accidents lead to over a million fatalities and tens of millions more injuries each year worldwide. Fewer than 6% of these accidents are on high-speed motorways, while more than 67% occur in urban environments. Looking at more detailed records, over half of the collisions between road users take place at intersections, and most involve lateral interactions. Collisions rarely result from a single sudden mistake, but emerge from a cascading sequence of increasingly intense interactions that can last seconds or even minutes. Therefore, a time window exists from the detection of a conflict to the occurrence of a collision, providing a valuable opportunity to intervene. To secure the precious time needed for prevention, this thesis is focused on proactively detecting potential collisions and quantifying their risks in multi-directional urban traffic.

Research questions

The overarching research objective is to systematically provide a context-aware, generalisable, and scalable methodology to quantify the collision risk in multi-directional traffic interactions. Beyond car following, these interactions include lane changing, merging, turning, crossing at intersections, etc., which are typical of city streets and intersections. To achieve the objective, this thesis answers four interconnected research questions: *How can multi-directional traffic interactions be effectively characterised*? *How can traffic conflict detection accommodate varying contextual factors*? *How can collision risk in diverse multi-directional traffic interactions be quantified in a unified way*? *How can collision risk quantification be scaled up without annotated data of crashes or near-crashes*? In a nutshell, these questions trace a logical progression from foundational measurement to large-scale deployment.

Methods

The methodology developed in this thesis spans from measuring multi-directional traffic interactions to self-supervised learning of collision risk. It first introduces a two-dimensional (2D, longitudinal and lateral) coordinate transformation to normalise the spacing between road users. This enables microscopic measurement of multi-directional spacing, as well as macroscopic analyses of the required road space for traffic interactions through the interaction Fundamental Diagram (iFD). On this foundation, a conditional formulation and a unified probabilistic framework for conflict detection are proposed. The formulation conditions collision risk on interaction context, i.e., any information that is useful to characterise the interaction, such as relative speed, road user behaviour, and environmental factors. A statistical learning pipeline is then developed to drive

the framework with data. This enables the training of risk metrics that are adaptive to interaction context, thus generalising across traffic scenarios. These data-driven metrics are continuous and can naturally capture a long-tailed risk spectrum ranging from mild conflicts to near-crashes. In order to more effectively exploit the interaction patterns in data, this thesis proposes the Generalised Surrogate Safety Measure (GSSM), an end-to-end assembly of the previously developed methods. GSSM is a self-supervised approach that learns exclusively from abundant naturalistic driving to infer risky extremes without the need for labelled crashes or near-crashes. Further, for more efficient self-supervised learning as increasing data becomes available, contrastive learning is explored to preserve fine-grained similarities in latent interaction patterns.

Results and findings

Extensive experiments with real-world data have validated the proposed methods and revealed important findings. A collection of the data underlying this thesis is prepared [1], where the open-source code repositories and datasets for each chapter are referenced.

The 2D coordinate transformation enables consistent analyses of multi-directional interactions in urban traffic, and finds that lateral interactions are more efficient in the use of road space than longitudinal interactions. There is no hard boundary between safe and unsafe interactions. As a result, context-conditioned conflict detection provides more reliable collision warnings with fewer false detections than traditional heuristic methods. Based on multi-directional spacing and context-conditioned detection, the unified collision risk metric is adaptive in various interaction scenarios and outperforms existing methods in both detection accuracy and timeliness. Further results with GSSM show that environmental information such as weather, lighting, and road surface conditions enhances detection performance, and adverse conditions such as rain and wet roads are found to be associated with increased collision risk. Without requiring collision or conflict annotations, the performance of GSSM improves with increasing training inputs. In addition, preserving fine-grained similarities facilitates learning useful interaction patterns that boost performance.

These results demonstrate the systematic methodology developed in this thesis for proactive collision risk quantification. Each component, including consistent interaction measurement, context-aware conflict detection, generalised risk from normal interactions, and scalable self-supervised learning, contributes to a better understanding and identification of collision risk in multi-directional traffic interactions. Three core conclusions are drawn from the findings. First, there is no universal boundary that separates safe and unsafe traffic interactions; collision risk instead exists as a continuum that must be evaluated dynamically from time to time and location to location. Second, the collision risk in multi-directional interactions can be quantified with a unified probabilistic framework if an appropriate proxy and reference system (specifically, 2D spacing in this thesis) are chosen. Third, extremely conflicting interactions are statistically related to

[1] At https://doi.org/10.4121/8636d409-2f19-4ba3-8093-dff79843537f

everyday interactions; therefore, a learning system can extrapolate collision risk from normal interactions rather than waiting for crashes to occur.

Possible applications and implications

This thesis shows that data-driven proactive collision risk quantification is not only feasible but can be deployed at scale. The methodological advances can be translated into practical applications for traffic management and autonomous driving. Traffic management agencies can utilise the iFD to evaluate urban intersection efficiency and inform design improvement in signal timing or road layout. Automotive engineers can embed the unified conflict detection and adaptive thresholds into omnidirectional collision warning in Advanced Driver Assistance Systems (ADAS), reducing both under-reaction and over-reaction complaints that currently dominate consumer feedback. Developers of autonomous driving systems can leverage the GSSM as a safety reward or validation signal, improving end-to-end training and enabling targeted curation of rare yet safety-critical interaction scenarios. Finally, powered by intelligent transportation systems equipped with connected vehicles and roadside infrastructure, city authorities can monitor a rolling map of road traffic risk in real time and proactively intervene in potential accidents.

Beyond these tangible use cases, this thesis contributes to the ongoing shift from reactive to proactive management of road safety. The research in this thesis promises fewer fatalities and injuries, smoother traffic thanks to less incident-induced congestion, and greater public confidence in automated vehicles. It can also empower policymakers with rapid feedback on traffic safety policies. By learning collision risk from everyday traffic interactions rather than waiting for crashes, the feedback loop between safety policies and their observable effects can be shortened from years to months, accelerating iterative improvements to the road system. These implications align neatly with the principles of Vision Zero to eliminate traffic fatalities and severe injuries, as well as the United Nations Sustainable Development Goal to create inclusive, safe, resilient and sustainable cities.

Samenvatting

Achtergrond

Verkeersongevallen op de weg veroorzaken wereldwijd jaarlijks meer dan een miljoen dodelijke slachtoffers en tientallen miljoenen gewonden. Minder dan 6% van deze ongevallen vindt plaats op autosnelwegen, terwijl meer dan 67% zich in stedelijke omgevingen voordoet. Uit gedetailleerde registraties blijkt dat meer dan de helft van de botsingen tussen weggebruikers plaatsvindt op kruispunten en dat het merendeel laterale interacties betreft. Botsingen zijn zelden het gevolg van één plotselinge fout; ze ontstaan uit een opeenvolging van steeds intensere interacties – een aaneenschakeling van handelingen die zich over seconden tot minuten kunnen uitstrekken. Er bestaat dus een tijdsvenster tussen het waarnemen van een conflict en het moment waarop een botsing daadwerkelijk plaatsvindt, wat ruimte biedt voor preventief ingrijpen. Om de kostbare tijd die nodig is voor preventie effectief te benutten, richt dit proefschrift zich op het proactief detecteren van potentiële botsingen en het kwantificeren van de bijbehorende risico's in multidirectioneel stedelijk verkeer.

Onderzoeksvragen

Het overkoepelende onderzoeksdoel is om een contextbewuste, generaliseerbare en schaalbare methodologie te ontwikkelen om het botsingsrisico bij multidirectionele verkeersinteracties te kwantificeren. Naast het volgen van voorliggers (car following) omvatten deze interacties rijstrookwisselingen, invoegen, afslaan en oversteken, die typerend zijn voor stedelijke straten en kruispunten. Om dit doel te bereiken, beantwoordt dit proefschrift vier onderling verbonden onderzoeksvragen: *Hoe kunnen multidirectionele verkeersinteracties effectief worden gekarakteriseerd*? *Hoe kan conflictdetectie rekening houden met wisselende contextuele factoren*? *Hoe kan het botsingsrisico bij diverse multidirectionele interacties op een uniforme manier worden gekwantificeerd*? *Hoe kan de kwantificatie van botsingsrisico's worden opgeschaald zonder geannoteerde data van ongevallen of bijna-ongevallen*? In essentie volgen deze vragen een logische progressie van fundamentele meting tot grootschalige toepassing.

Methoden

De in dit proefschrift ontwikkelde methodologie strekt zich uit van het meten van multidirectionele verkeersinteracties tot self-supervised leren van botsingsrisico. Allereerst wordt een tweedimensionale (2D) coördinatentransformatie (longitudinaal en lateraal) geïntroduceerd om de onderlinge afstand tussen weggebruikers te normaliseren. Dit maakt microscopische metingen van multidirectionele tussenruimte mogelijk, evenals

macroscopische analyses van de benodigde wegruimte voor verkeersinteracties via het interactie-Fundamenteel Diagram (iFD). Op deze basis worden een conditionele formulering en een uniform probabilistisch raamwerk voor conflictdetectie voorgesteld. De formulering conditioneert het botsingsrisico op de interactiecontext — d.w.z. alle informatie die nuttig is om de interactie te karakteriseren, zoals relatieve snelheid, gedrag van weggebruikers en omgevingsfactoren. Een statistisch leerproces voedt dit raamwerk, waardoor risicomaten ontstaan die adaptief zijn aan de context en dus generaliseren over verschillende scenario's. Om interactiepatronen in data optimaal te benutten, introduceert dit proefschrift de Generalised Surrogate Safety Measure (GSSM), een end-to-end generalisatie van de eerder ontwikkelde methoden. GSSM is een self-supervised aanpak die uitsluitend leert van overvloedige naturalistische rijgegevens om risicovolle extremen af te leiden zonder de noodzaak van gelabelde ongevallen of bijna-ongevallen. Voor efficiënter leren naarmate er meer data beschikbaar komt , wordt bovendien contrastive learning onderzocht om fijnmazige overeenkomsten in latente interactiepatronen te behouden.

Resultaten en bevindingen

Uitgebreide experimenten met praktijkdata hebben de voorgestelde methoden gevalideerd en belangrijke bevindingen opgeleverd. De onderliggende data en open-source code zijn beschikbaar [2].

De 2D-transformatie maakt consistente analyses van multidirectionele interacties mogelijk en toont aan dat laterale interacties de wegruimte efficiënter benutten dan longitudinale. Omdat er geen harde grens bestaat tussen veilige en onveilige interacties, levert context-geconditioneerde conflictdetectie betrouwbaardere botsingswaarschuwingen met minder valse meldingen dan traditionele heuristische methoden. De uniforme metriek voor botsingsrisico is adaptief in verschillende scenario's en presteert beter dan bestaande methoden wat betreft zowel nauwkeurigheid als tijdigheid. Resultaten met GSSM laten zien dat omgevingsinformatie (weer, licht, wegdek) de detectieprestaties verbetert en dat ongunstige omstandigheden, zoals regen en een nat wegdek, geassocieerd zijn met een verhoogd botsingsrisico. Zonder annotaties van ongevallen of conflicten verbeteren de prestaties van GSSM met toenemende trainingsinput, en het behouden van fijnmazige overeenkomsten verbetert de prestaties door het leren van nuttige interactiepatronen te faciliteren.

Deze resultaten demonstreren de systematische methodologie die in dit proefschrift is ontwikkeld voor proactieve kwantificering van botsingsrisico's. Elk onderdeel — waaronder consistente meting van interacties, contextbewuste conflictdetectie, het generaliseren van risico op basis van normale interacties, en schaalbaar self-supervised leren — draagt bij aan een beter begrip en een betere identificatie van het botsingsrisico bij multidirectionele verkeersinteracties. Deze bevindingen leiden tot drie kernconclusies. Ten eerste bestaat er geen universele grens die veilige en onveilige verkeersinteracties scheidt; botsingsrisico bestaat daarentegen als een continuüm dat dynamisch moet worden

[2] Via https://doi.org/10.4121/8636d409-2f19-4ba3-8093-dff79843537f

geëvalueerd per tijd en locatie. Ten tweede kan het botsingsrisico bij multidirectionele interacties worden gekwantificeerd met een uniform probabilistisch raamwerk, mits een geschikte proxy en referentiesysteem (specifiek, 2D-tussenruimte in dit proefschrift) worden gekozen. Ten derde zijn extreem conflicterende interacties statistisch gerelateerd aan alledaagse interacties, waardoor een leersysteem het botsingsrisico kan extrapoleren uit normaal verkeer zonder te wachten tot er ongevallen plaatsvinden.

Mogelijke toepassingen en implicaties

Dit onderzoek toont aan dat datagedreven, proactieve risicokwantificatie niet alleen haalbaar is, maar ook schaalbaar kan worden geïmplementeerd. Verkeersmanagementinstanties kunnen het iFD benutten om de efficiëntie van stedelijke kruispunten te beoordelen en aanpassingen in verkeerslichtregelingen of weginrichting te onderbouwen. Automotive-ingenieurs kunnen de uniforme conflictdetectie met adaptieve drempels integreren in omnidirectionele botsingswaarschuwingen voor Advanced Driver Assistance Systems (ADAS), waardoor klachten over zowel onder- als overreacties verminderen. Ontwikkelaars van autonome rijsystemen kunnen GSSM gebruiken als beloningssignaal voor veilig rijgedrag of als validatietool. Dit verbetert de end-to-end training en maakt het gericht selecteren van zeldzame, maar veiligheidskritieke scenario's mogelijk. Met ondersteuning van intelligente transportsystemen, uitgerust met connected vehicles en wegkantinfrastructuur, kunnen stadsbesturen bovendien een realtime risicokaart monitoren en proactief ingrijpen bij potentiële ongevallen.

Deze bijdrage ondersteunt de overgang van reactief naar proactief verkeersveiligheidsbeheer. Het onderzoek belooft minder slachtoffers, vlotter verkeer dankzij minder incidentgerelateerde congestie en een groter publieksvertrouwen in geautomatiseerd rijden. Omdat botsingsrisico wordt geleerd uit alledaagse interacties in plaats van te wachten op ongevallen, kan de feedbacklus tussen veiligheidsbeleid en de waarneembare effecten worden verkort van jaren naar maanden, wat iteratieve verbeteringen aan het wegennet versnelt. Dit werk sluit naadloos aan bij de principes van Vision Zero om verkeersdoden en ernstige verwondingen uit te bannen en bij de Duurzame Ontwikkelingsdoelstelling van de Verenigde Naties voor inclusieve, veilige, veerkrachtige, en duurzame steden.

总览

【研究背景】

每年，道路交通事故仍在全球范围导致超过一百万人死亡、数千万人受伤。这些事故仅有不到6%发生在高速公路上，却有超过三分之二在城市环境中。更详细的数据显示，交通参与者之间的碰撞半数以上发生在交叉口，且大都涉及横向交互。碰撞很少由某个突发错误直接导致，而是在数秒乃至数分钟内，经历交互冲突强度的不断升级。因此，从发现冲突到发生碰撞之间存在一段时间窗口，留出了实施干预的契机。本论文的研究聚焦于预判城区驾驶多向交互中的潜在碰撞并量化其风险，以赢得宝贵的预防时间。

【研究问题】

本研究的总目标是为预判与量化多向交互的碰撞风险系统性地提供一套可感知情境、可泛化、并可规模化的方法。除单向（即纵向）跟车行为外，多向交互还包括变道、并道、转弯、交叉口穿行等城市交通典型场景。为此，本论文依次回答四个递进的研究问题：如何有效表征多向驾驶交互？交通冲突检测如何适应多变的情境？如何以统一形式量化不同多向交互中的碰撞风险？如何摆脱对碰撞或冲突数据标注的依赖、规模化对碰撞风险的量化？这些问题构成了一条从基础测量到大规模部署的逻辑链。

【研究方法】

本论文开发的方法从多向交互的度量循序渐进至碰撞风险的自监督学习。首先提出的是一种二维（纵向和横向）坐标变换，以规范化道路使用者之间在二维空间的间隔。这使得多向间距能够在微观层面上测量，并通过交互基本图（interaction Fundamental Diagram, iFD）在宏观层面分析多向交互所需的道路空间。在此基础上，本研究提出了用于冲突检测的条件公式和统一概率框架。该公式令碰撞风险依赖于交互情境——即任何有助于描述交互特征的信息，例如相对速度、道路使用者行为、和环境因素。随之开发的统计学习方案利用数据驱动该框架，使可适应交互情境的风险指标得以训练。这样从数据中习得的指标是连续的，能够捕捉从轻微冲突到濒临碰撞的长尾风险谱。为了更充分地挖掘数据中的交互模式，本论文接着提出广义替代安全指标（Generalised Surrogate Safety Measure, GSSM），利用神经网络对前述方法端到端整合。GSSM是自监督的，无需碰撞或濒临碰撞的标注数据，从丰富的自然驾驶场景中学习并外推高风险的极端情况。随着可用数据不断增加，为提升自监督学习效率，本文还探索了对比学习、以在表征空间中保留交互模式间的细粒度相似性。

【研究结果】

本论文在真实世界数据上进行了充分的实验来验证所提出的方法、并揭示有意义的发现。实验数据已公开[3]，可索引至各章的开源代码与复现指南。

通过二维坐标变换度量间距，城市交通中的多向交互得以类比分析；结果表明横向交互比纵向交互更能高效利用道路空间。安全与不安全的交互之间没有绝对界限，因此依赖于情境的冲突检测能提供比传统启发式方法更可靠的碰撞预警、减少误判。基于多向间距与依赖情境的检测，统一的概率性碰撞风险指标可适应各种交互场景，在冲突检测的准确性和及时性方面都优于现有方法。使用GSSM的进一步结果表明，天气、光照、路面状况等环境信息能增强检测效果，而且降雨及湿滑路面等不利条件相关于碰撞风险的增加。无需碰撞或冲突的标注数据，GSSM用于风险预判的准确性会随着自然驾驶的训练数据量增加而提升。同时，保留细粒度相似性有助于神经网络有效学习交互模式，从而进一步提高模型性能。

这些结果论证了本文所构建的一整套方法用于预判和量化碰撞风险的能力。一致的交互测量、适应情境的冲突检测、从正常交互中泛化的风险、以及可规模化的自监督学习，这些共同增进了对多向交互中碰撞风险的理解和识别。从上述研究结果中可以得出三个核心结论。第一，安全与不安全的交互之间不存在通用的边界；碰撞风险是需要随时间与地点动态评估的连续体。第二，只要选取合适的代理量和参照系（本论文使用二维间距），多向交互中的碰撞风险就可以通过统一概率框架进行量化。第三，极端冲突交互在统计上与常态交互相关联；因此，数据驱动的学习方法可以从正常交互中外推碰撞风险，而无需等待碰撞发生。

【研究意义】

本研究表明，数据驱动的碰撞风险预判不仅可行，而且能够大规模训练与部署。本文在研究方法上的进展可以转化为交通管理和自动驾驶中的实际应用。交通规划机构可使用iFD评估城市交叉口的运行效率，据此优化信号配时或道路布局；车辆工程师可将统一冲突检测与自适应阈值嵌入高级驾驶辅助系统（Advanced Driver Assistance Systems，ADAS）全向碰撞预警，缓解目前用户反馈中普遍存在的反应不足与反应过度问题；自动驾驶系统开发者可利用GSSM提供动态、稠密的安全奖励或验证信号，改进端到端算法的训练，并有针对性地筛选罕见但安全关键的交互场景；最后，借助由车联网与路侧设施支持的智能交通系统，城市管理部门可实时监控道路交通风险、及时干预潜在事故。

在这些具体用例之外，本论文贡献于推动道路安全管理从被动响应向主动预防的转变。本文的研究有望直接减少伤亡事故、减少事故引发的拥堵而使交通更加顺畅、以及增强公众对自动驾驶车辆的信任使之反哺交通安全。通过从日常交互中学习碰撞风险、而不是等待事故发生，本研究有潜力将交通安全措施与其可观测效果之间的反馈循环由数年缩短至数月，从而加速道路系统的迭代改进。这些积极影响与“零伤亡愿景”（Vision Zero）消除交通事故引发的死亡与重伤的理念相合，同时回应了联合国可持续发展“建设包容、安全、有韧性、可持续的城市”的目标。

[3]地址https://doi.org/10.4121/8636d409-2f19-4ba3-8093-dff79843537f

Chapter 1

Introduction

1.1 Motivation

Reducing traffic accidents remains a significant challenge. Every year, approximately 1.2 million people lose their lives in road crashes, along with tens of millions of injuries [1]. Substantial societal and economic costs are incurred by even more non-fatal and non-injury accidents through, e.g., congestion and emergency response [2–4]. Despite the common perception that highways pose greater crash risk due to high speeds, as shown in Figure 1.1(a)[1], fewer than 6% of accidents happen there, while more than 67% occur in urban areas. In particular, over half of all crashes between road users take place at urban intersections, as shown in Figure 1.1(b) and 1.1(c)[2]. Of these accidents at intersections, the majority involve lateral interactions.

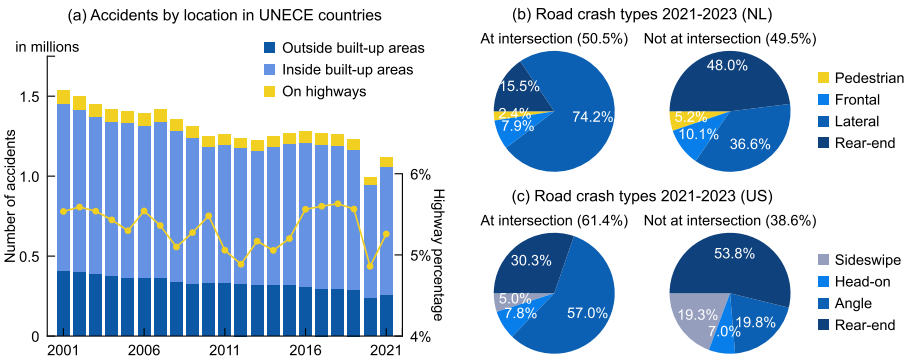


Figure 1.1 Statistics of traffic accidents and road crash types. **(a)** Number of traffic accidents at different locations from 2001 to 2021 in 27 countries for which complete data are accessible. **(b)** Distribution of road crash types in the Netherlands from 2021 to 2023. **(c)** Distribution of motor vehicle crash types in the United States from 2021 to 2023.

[1]These countries are members of the United Nations Economic Commission for Europe (UNECE). The data are sourced from [5].

[2]Only the crashes between road users are included. This means an exclusion of single-vehicle crashes and crashes with parked vehicles or animals. The use of data in the Netherlands (sourced from [6]) and the United States (sourced from [7]) is due to easier data accessibility. These two countries might not be representative of the global situation.

1

Traffic accidents seldom stem from a single catastrophic mistake, but often from the accumulation of multiple events. Prior to a crash, mild and severe conflicts progressively increase the risk until prevention is impossible or fails [8, 9]. From a systematic perspective, a collision is the culmination of an escalating series of crash precursors. These precursors typically unfold for seconds to minutes before a crash, thus leaving a valuable window during which a human or automated system can intervene to prevent a potential collision.

To secure time and space for prevention, proactively detecting the crash precursors and quantifying their risk is essential. This has been extensively researched in longitudinal traffic dynamics (i.e., car-following, see literature reviews, e.g., [10, 11]), but remains challenging in urban traffic where lateral interactions such as lane-changing, overtaking, turning, and crossing abound. Particularly at intersections, road users approach from different traffic directions, negotiate their rights of way, and manoeuvre both longitudinally and laterally. Such interactions are multi-directional and highly variable, making the quantification of collision risk in urban traffic complicated.

The research focus of this thesis is on proactively quantifying the risk of collisions before they occur in multi-directional urban traffic – the very situation in which the majority of road crashes take place. This can inform the prevention of collisions and eventually contributes to reducing traffic accidents.

1.2 Background

Table 1.1 Terminology of traffic safety events

Traffic event	**Definition**
Accident	Event that involves at least one vehicle in motion and results in injury, fatality, or property damage.
Crash	Interchangeable use with accident, covering all harmful road traffic events.
Collision	Contact that a vehicle has with other road users, including vehicles, cyclists, pedestrians, etc.
Non-collision crash	Accidents/Crashes that involve one single vehicle.
Near-crash	Event that requires a rapid evasive manoeuvre to avoid a collision [12].
Conflict	Situation where at least two road users approach each other in time and space to such an extent that there is risk of collision if their movements remain unchanged [13]. Interchangeable use with potential collision.
Interaction	Situation where the behaviour of at least two road users can be interpreted as being influenced by the possibility that they are both intending to occupy the same region of space at the same time in the near future [14].

To avoid confusion in the use of terminology, Table 1.1 provides definitions of the traffic safety events discussed throughout this thesis, and Figure 1.2 shows their relationships. Accidents and crashes are used interchangeably, which include collisions between road users and non-collision crashes with other obstacles. In this thesis, every collision is considered to evolve from a conflict. Conflicts are thus also called potential collisions. Different conflicts vary in intensity, among which near-crashes are those requiring rapid evasive operations. In a broader view, both collisions and conflicts are interactions during which the behaviour of road users influences each other.

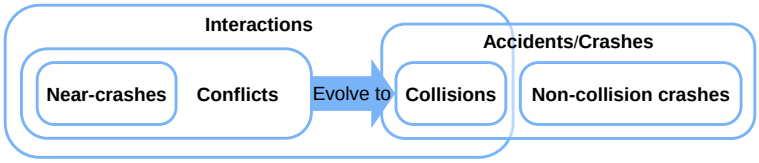


Figure 1.2 Relationships between traffic safety events.

In this thesis, proactive collision risk quantification refers to estimating the time-varying likelihood that a conflict may evolve into a collision in a certain interaction context. The quantified collision risk is therefore not an absolute probability that a collision will occur. Instead, it provides a risk score that is useful for ranking and comparing potential collision situations within and across interaction contexts. For example, two situations with equal risk scores do not necessarily have identical real-world collision probabilities; differences in driver behaviour, perception, and vehicle capabilities may lead to different chances that a conflict is resolved safely. As such, proactively quantifying collision risk first requires consistent characterisation of multi-directional traffic interactions, and then adaptive identification and evaluation of traffic conflicts. This stands on the shoulders of various fields such as traffic dynamics, traffic safety, probability theory, and statistics. The remainder of this section offers an overview of the three most relevant directions and underlines the research need in existing studies.

1.2.1 Traffic interaction characterisation

Modelling traffic interactions has traditionally targeted unidirectional (i.e., longitudinal) vehicle interactions on highways [15, 16]. In this context, interactive behaviours such as car-following and lane-changing are characterised by one-dimensional (1D) speed, headway, spacing, etc. However, urban traffic interactions, especially those at intersections, involve multi-directional dynamics of different road users that existing 1D dynamics struggle to completely characterise. Road users in urban environments move in a two-dimensional (2D) plane with varied orientations, as emphasised in many studies focused on urban intersections (e.g., [17, 18]). In this plane, the notions of safe distance, relative speed, or time headway become unclear when a road user approaches another laterally.

There is not yet an effective measure or consistent quantity to characterise diverse lateral interactions in urban environments, where road users do not simply follow one another, but share a 2D road space. The lack of generic measurements of multi-directional interactions results in a limited ability to model and understand urban traffic. This limitation has also hindered the development of proactive collision risk quantification in multi-directional interactions.

1.2.2 Traffic conflict detection

In the field of traffic safety, traffic conflicts have been widely used as a proxy to quantify collision risk [19, 20]. Different traffic conflicts are detected using separate indicators in existing studies. For example, Time to Collision (TTC) [8, 21] is designed for rear-end conflicts, whereas Post-Encroachment Time (PET) [22, 23] is specialised for path-crossing conflicts. The severity of collision risk quantified by these indicators can vary. Empirical experiments show that driver perception of critical TTC differs among individuals [24] and across environments [25, 26]. More specifically, although a TTC of 3 seconds may imply high risk on highways, it could entail no danger during a cooperative lane-change or deceleration towards an intersection stop line. Similarly, 2-second PET may indicate a near-crash between two cars crossing at an intersection, but it is not unusual for cyclists [27]. These variations necessitate more comprehensive methods for traffic conflict detection.

In response to the variability, commercial collision warning and prevention systems typically implement much more complex decision logic than a single TTC threshold. They fuse multiple sensors, algorithms, heuristics, and often include scenario-specific modules and safeguards. In academic research, enhanced methods for conflict detection are explored. To involve lateral interactions, 2D kinematics-informed indicators such as Anticipated Collision Time [18] and Emergency Index [28] are proposed. To cover a more comprehensive range of conflicts, multiple indicators are combined into a composite model [29–31] and machine learning is utilised [32, 33]. An important line of research focuses on determining more reliable thresholds to distinguish unsafe interactions. When crash or near-crash records are available, the threshold can be optimised for best detection accuracy [34]. In the absence of ground truth labels, heuristic rules [35–37] or statistical hypotheses [38–40] are applied to set thresholds. These enhancements are ultimately limited by the indicators on which they are based, and of which the specific assumptions about conflict emergence are not relaxed.

The limitation in dealing with variations highlights that traffic conflict detection needs to be context-aware, taking into account factors such as road layout, moving speed, interactive behaviour, and road user type. This becomes increasingly important in light of emerging technologies. For instance, automated vehicles and Advanced Driving Assistance Systems (ADAS) require reliable real-time evaluation of collision risk as they navigate in various traffic environments. Failing to account for interaction context can cause underreaction (e.g., overlooking critical threats) or overreaction (e.g., unnecessary emergency braking). According to Ayoub et al. [41], more than 75% of consumer

complaints about proactive safety failures in ADAS are exactly about underreaction and overreaction. In a near future of mixed traffic where human drivers and different levels of automated systems share roads, conflict detection without context-awareness may miss the window to prevent crashes, and even introduce new risks.

1.2.3 Data-driven collision risk quantification

Collision risk has also been quantified through two other approaches in addition to detecting traffic conflicts. These two approaches are typically data-driven and struggle less with context-awareness, but face more challenges in generalisability and scalability. In this thesis, *generalisability* refers to the ability of a model to maintain its effectiveness when applied to road environments, interactive behaviours, or road users that differ from those present in calibration or training data. *Scalability* refers to the ability of a model to learn diverse patterns from massive amounts of data, so that the model can be reliably deployed at scale.

One approach that quantifies collision risk through motion prediction under uncertainty is commonly adopted in robotics and vehicle control. In this approach, probabilistic trajectories or reachable sets of road users are predicted and used to compute the probability of safety constraint violations [42–46]. It is assumed that the uncertainty in road user behaviour can be consistently modelled and that the resulting stochastic model remains valid even when collision risk is high. However, this assumption often fails to generalise to abnormal, aggressive, or highly interactive motions that precede crashes [47], leading to increased misjudgment in safety-critical situations.

Another approach is video-based traffic accident anticipation (TAA), which aims to predict collisions before they occur [48, 49]. TAA trains deep neural networks with dashcam or surveillance videos of crashes to identify early visual cues. Training reliable TAA models requires a large amount of labelled video data of crashes. Yet, compared to daily normal interactions, traffic collisions are rare and variable long-tail events, making such data both scarce and unrepresentative of all risky situations [50]. This data constraint not only limits training TAA models at scale, but also impedes their generalisability to unseen situations that differ from historical collision patterns.

1.3 Knowledge gaps

Despite substantial progress in different directions, existing approaches for proactive collision risk quantification are limited within their specifically designed range. Methods built on predefined rules or assumptions consider restricted information and fail to recognise collision risk in other contexts, while those built on machine learning demand extensive and diverse crash data that are often unavailable. As a result, the development becomes increasingly confined within disconnected comfort zones: each method is useful in specific conditions, but none addresses different traffic interactions in a unified way.

This thesis identifies three progressive knowledge gaps in quantifying collision risk in multi-directional traffic interactions. The foundational gap lies in characterising

multi-directional interactions. Based on effective characterisation, the second gap requires a comprehensive consideration of varying contextual factors of risk. With a context-aware framework, the third gap seeks generalisability across interaction contexts and scalability during data-driven learning process.

Gap1. Effective characterisation of the use of shared road space and conflict dynamics in urban environments, where road users interact across multiple directions.

Gap2. Context-aware framework of traffic conflict detection that accounts for varying factors such as road layout, road user behaviour, and environmental change.

Gap3. Data-driven approach that adapts to various interaction scenarios in a unified way and is feasible for large-scale deployment in reality, to learning collision risk.

1.4 Research objective and questions

Addressing these knowledge gaps is vital for improving proactive traffic safety in urban environments. Accordingly, the research objective of this thesis is to

> systematically provide a context-aware, generalisable, and scalable methodology to quantify the collision risk in multi-directional traffic interactions.

To achieve this objective, the following research questions (RQs) are investigated.

RQ1. *How can multi-directional traffic interactions be effectively characterised?*
What new metrics or methods can capture the characteristics of multi-directional traffic interactions (such as merging, turning, crossing) in a 2D plane? The answer to this question contributes to addressing the first knowledge gap and lays a foundation for later risk quantification. [Chapter 2]

RQ2. *How can traffic conflict detection accommodate varying contextual factors?*
Traffic conflicts are at different levels of collision risk, which is influenced by a variety of contextual factors. What methods can relate interactions, conflicts, and collisions, taking into account the varying factors? Answering this question aims to address the second knowledge gap of context-awareness. [Chapters 3 and 4]

RQ3. *How can collision risk in diverse multi-directional traffic interactions be quantified in a unified way?*
Building upon the answers to RQ1 and RQ2, what methods can enable unified quantification of collision risk across different scenarios of multi-directional traffic interactions? In other words, what methods can consistently detect different types of traffic conflicts and evaluate their intensity? Answering this question aims to address the third knowledge gap regarding generalisability. [Chapter 4]

RQ4. *How can collision risk quantification be scaled up without annotated data of crashes or near-crashes?*
Finally, how can the increasingly abundant data (e.g., naturalistic driving records) be exploited to quantify collision risk at scale, without the need for laborious manual labelling of crashes or near-crashes? Answering this question targets a

scalable and self-supervised approach that addresses the third knowledge gap regarding scalability. [Chapters 5 and 6]

Each RQ aligns with one or two of the main chapters 2∼6. Consequently, the research in this thesis progresses from foundational characterisation of traffic interactions(RQ1), to context-aware and generalisable detection of traffic conflicts (RQ2 and RQ3), and eventually to large-scale implementation of collision risk quantification (RQ4).

1.5 Contributions

By answering the research questions, this thesis addresses the knowledge gaps through providing new theoretical insights, methodological advances, and pathways to practice. Figure 1.3 depicts the connections between these contributions. The scientific and methodological contributions will be detailed in this section, while the practical applications and societal relevance will be summarised in the conclusion chapter. Altogether, these contributions provide the groundwork for traffic safety analysis and automated driving systems to account for the multi-directional and context-dependent nature of road user interactions.

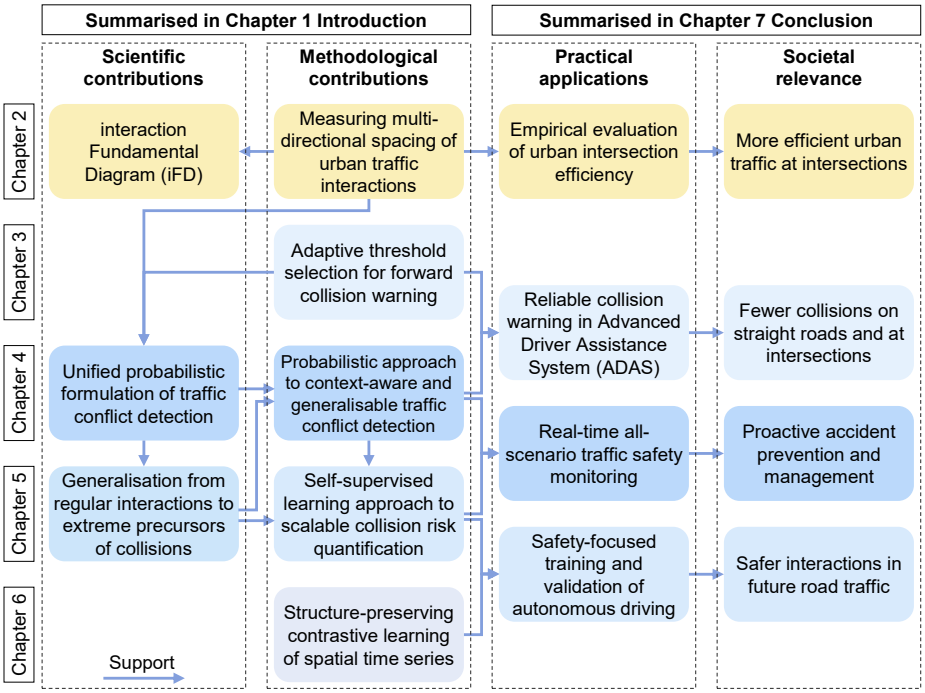


Figure 1.3 Contributions and their connections in the thesis.

1

1.5.1 Scientific contributions

The scientific contributions in this thesis deepen the understanding of how road users interact in complex urban environments and how these interactions are associated with potential collisions. They provide the theoretical basis for subsequent methodological developments.

- **Interaction Fundamental Diagram of average road space required for vehicle interactions**.
  Chapter 2 introduces the interaction Fundamental Diagram (iFD) that empirically relates the efficiency of vehicle interactions at urban intersections, the average amount of road space required for the interactions, and their relative speeds. The iFD reveals that lateral interactions use road space more efficiently than longitudinal interactions at urban intersections. In line with classical longitudinal traffic flow theory, for multi-directional interactions in urban traffic, there also exist optimal states where interaction efficiency is maximised.
- **Unified probabilistic formulation of traffic conflict detection**.
  Chapter 4 formulates the detection of traffic conflicts as quantifying the conditional risk of potential collisions. This formulation unifies numerous surrogate metrics of traffic conflicts under one probabilistic interpretation. It is a conceptual breakthrough that resolves the inconsistencies among different metrics, and thus allows for comparing and evaluating collision risk across different interaction contexts.
- **Generalisation from normal traffic interactions to potential collisions**.
  Every collision evolves from a previously safe interaction. Chapters 4 and 5 theoretically and empirically derive that collisions and conflicts are extreme events deviating from normal interactions to different degrees. Based on this derivation, the patterns of everyday naturalistic interactions can be generalised to safety-critical conflicts and collisions. This generalisation justifies the scalable development of collision risk quantification as increasingly abundant interaction data are collected.

1.5.2 Methodological contributions

Building on the scientific insights, this thesis develops a series of methods and techniques for empirical analysis and collision risk quantification of multi-directional traffic interactions.

- **A new method to measure multi-directional spacing of urban traffic interactions**.
  Chapter 2 introduces a relative coordinate system where the transformed coordinates can represent how two road users approach and leave each other. This provides a method that enables empirical measurement of multi-directional interactions, and thus analysis of the safety and efficiency in real urban traffic.
- **A new method to minimise false warnings for collision avoidance**.
  Chapter 3 proposes to optimise the threshold for issuing collision avoidance warnings by minimising the estimated probabilities of false negatives and false positives. This provides a statistical procedure for setting adaptive collision warning triggers as opposed to heuristic thresholds.

- **A new approach to probabilistic traffic conflict detection**.
  Chapter 4 designs sequential tasks to apply the unified probabilistic formulation of traffic conflict detection. These tasks are intended to extract interaction patterns from real-world data and form a pipeline where each module can be continuously improved for more effective conflict detection.
- **A new self-supervised learning approach for collision risk quantification**.
  As one of the core contributions, Chapter 5 develops a context-aware, generalisable, and scalable approach for proactive collision risk quantification. It enables self-supervised learning and extrapolates the patterns of normal interactions to safety-critical situations without requiring historical crashes or manual annotations of near-crashes.
- **A new technique for structure-preserving contrastive learning of spatial time series**.
  Chapter 6 looks further into self-supervised learning of traffic interactions. It proposes a dynamic mechanism to adaptively balance contrastive learning and preserving fine-grained similarity relations between training samples. This is designed for more effective learning of the patterns in spatio-temporal data to enhance downstream task performance.

1.5.3 Contributions to open science

The source code of the methodological contributions in this thesis has been made publicly available. This is in the spirit of open science and to encourage further innovation.

- Chapter 2, `https://github.com/Yiru-Jiao/DriverSpaceInference`
- Chapter 3, `https://github.com/Yiru-Jiao/Conflict-detection-MFaM`
- Chapter 4, `https://github.com/Yiru-Jiao/UnifiedConflictDetection`
- Chapter 5, `https://github.com/Yiru-Jiao/GSSM`
- Chapter 6, `https://github.com/Yiru-Jiao/SPCLT`

In addition, supporting tools and data resources developed during research are open-sourced for public use. This includes two code repositories for efficient computation of existing surrogate safety measures (SSMs) in the 2D plane, as well as two datasets of real-world crashes and near-crashes reconstructed from naturalistic driving studies.

- Efficient computation of 2D SSMs
  - `https://github.com/Yiru-Jiao/Two-Dimensional-Time-To-Collision`
  - `https://github.com/Yiru-Jiao/SSMsOnPlane`
- Datasets of real-world crashes and near-crashes
  - `https://github.com/Yiru-Jiao/Reconstruct100CarNDSData`
  - `https://doi.org/10.15787/VTT1/EFYEJR`

By sharing both code and data, I hope to foster research transparency, enable broader validation of our findings, and accelerate future research on proactive traffic safety.

1

1.6 Thesis outline

The remainder of this thesis is organised into six chapters as outlined in Figure 1.4. These chapters are based on articles that have been published after peer review or are under peer review, constituting the scientific and methodological contributions in this thesis. They progress from fundamental concepts to generalisable and scalable solutions, while maintaining the focus on quantifying collision risk in multi-directional traffic interactions.

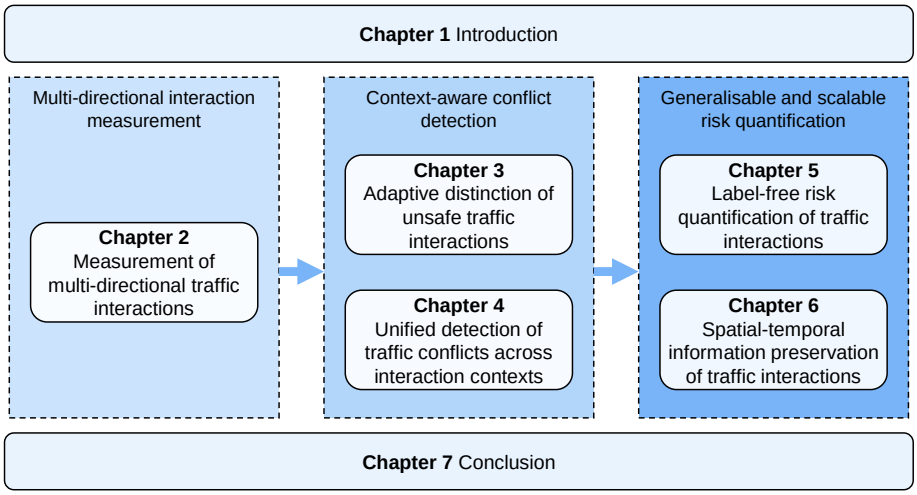


Figure 1.4 Outline of the thesis.

In broad terms, Chapter 2 establishes a foundational methodology for quantifying multi-directional traffic interactions (addressing RQ1). Chapters 3 and 4 then investigate methods for identifying and evaluating traffic conflicts in various interaction contexts (addressing RQ2 and RQ3), introducing both improved metrics and theoretical frameworks for probabilistic traffic conflict detection. Building upon them, Chapter 5 develops a self-supervised framework for context-aware, generalisable, and scalable collision risk quantification, and Chapter 6 further explores self-supervised learning of traffic interactions (addressing RQ4). Finally, Chapter 7 gives the conclusions of this thesis. This concluding chapter synthesises the major findings, evaluates the extent to which the research objective is met, discusses the limitations of the current work, suggests directions for future research, and reflects on potential practical and societal implications of adopting these new approaches to improve traffic safety.

Chapter 2

Measurement of multi-directional traffic interactions

Highlights

- A method to infer average multi-directional spacing from trajectory data is proposed.
- A perspective on the relative movement between interacting vehicles is taken.
- Empirical relations between multi-directional spacing and relative speeds are identified.
- The empirical relations are termed as interaction Fundamental Diagrams (iFDs).
- iFDs describe the variation in required road space for general traffic interactions.

Keywords

Urban traffic, vehicle interaction, two-dimensional spacing, Fundamental Diagram

This chapter is based on the journal article: Yiru Jiao, Simeon C. Calvert, Sander van Cranenburgh, and Hans van Lint. (2023). Inferring vehicle spacing in urban traffic from trajectory data. Transportation Research Part C: Emerging Technologies, 155, 104289. doi: 10.1016/j.trc.2023.104289

Abstract

This study presents a new method to infer the average two-dimensional (2D) spacing between interacting vehicles in urban traffic from trajectory data. In this context, 2D spacing reflects the amount of road space consumed by pairs of interacting vehicles, and is related to 2D density at the macroscopic level. Due to complex interaction and conflicts in urban traffic, the inherent assumptions in traditional traffic flow models, such as unidirectional flow and homogeneity, are often violated. Such violation challenges direct measurement of urban vehicle spacing. The proposed method addresses this challenge by focusing on the relative movement between interacting vehicles and aggregating the accumulated presence of vehicles in similar scenarios. We apply the method to a large-scale urban trajectory dataset called pNEUMA, and validate the consistency of the method through bootstrapping. By applying the method we obtain a new empirical relation between the average 2D spacing and the *relative* speeds between interacting vehicles. There are similarities between this empirical relation with the classical Fundamental Diagram of traffic flow in terms of shape and interpretation, and so we term it the "interaction Fundamental Diagram" (iFD). However, there are also key differences. The iFD does not represent steady-state (homogeneous and stationary) longitudinal behaviour; it describes the average amount of road space needed for vehicle interactions at different relative speeds. We believe these iFD relations contribute to understanding vehicle interaction in urban traffic, and can offer new insights for designing safer and more efficient urban intersections.

Code availability

`https://github.com/Yiru-Jiao/DriverSpaceInference`

Data availability

Raw data sources:

- pNEUMA `https://open-traffic.epfl.ch`
- INTERACTION `https://interaction-dataset.com`

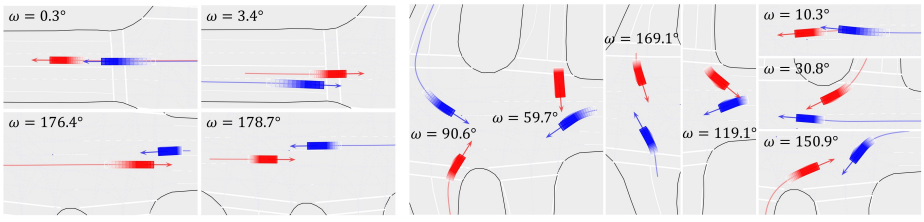
Resulting data: `https://doi.org/10.4121/8cadc255-5fd8-46ab-893a-64b76ca7b7f9`

2.1 Introduction

Whereas on uninterrupted high-speed and high-volume traffic facilities, vehicle interactions are dominated by one-dimensional (1D, longitudinal) dynamics (e.g., car-following); in urban traffic, and particularly at intersections, vehicle interactions are both one and two-dimensional (2D), involving crossing and conflict negotiation aside from car-following and lane-changing. Vehicle interactions collectively shape traffic flow, and conversely, traffic flow influences the interactions between vehicles. Therefore, investigating vehicle interactions in urban traffic and their collective phenomena is essential to manage traffic and promote efficient, safe, and sustainable use of the shared road space.

Traffic flow modelling has predominantly focused on 1D driving, with microscopic and macroscopic models respectively examine individual vehicle interaction and aggregated traffic patterns [15]. At the microscopic level, the characteristics of individual vehicles are described, such as vehicle speed, time headway, and distance headway. These characteristics are aggregated into traffic states at the macroscopic level, i.e., mean speed, flow, and density [51, 52]. The relations between these macroscopic traffic variables are extensively studied as Fundamental Diagrams (FDs), which describe how steady-state traffic states evolve [e.g., 53–56]. The aggregation per se and the relations between aggregated variables rely on assumptions rooted in fluid dynamics. For instance, vehicles are assumed to move in a single direction and follow one another without disruption. The assumptions, however, are often violated in urban traffic. Accordingly, traffic flow models that account for disruptions such as lane-changes [57, 58] and traffic signals [59–61] have been developed, and their corresponding FDs offer valuable information on complex real-world traffic patterns.

A gap persists concerning the assumption of unidirectional interaction, given that vehicle interaction in urban environments is multi-directional. As depicted in Figure 2.1, vehicles can have various orientations on a 2D plane during their interactions in urban traffic [17]. Although pedestrian traffic is also considered 2D, it has been studied within corridors or rings where the flows remain unidirectional or bidirectional [62–65]. Consequently, there is a long-standing need to identify governing quantities that can describe multi-directional vehicle interaction in urban environments and the resulted traffic phenomena.



(a) Longitudinal **(b)** Lateral

Figure 2.1 Real-world examples of vehicle interactions in urban environments. ω is the angle between moving directions of the interacting vehicles. **(a)** Vehicle interactions only in the longitudinal direction. **(b)** Vehicle interactions involving both longitudinal and lateral directions.

In this study, we take the perspective of relative movement between interacting vehicles. This enables mapping the state of each specific interaction onto a consistent system to quantify the 2D spacing between the involved vehicles. By accumulating all these mapped 2D spacings, we reconstruct a spatial distribution of the vehicle interaction states, which allows us to estimate the average spacing that vehicles maintain with one another at various relative speeds. We first present the method to infer average 2D spacing between vehicles in urban traffic, and then test the method using the large-scale urban trajectory dataset pNEUMA [66]. Our results reveal the characteristics of 2D vehicle spacing in different scenarios, with which we derive urban traffic states and discern their fundamental relations. We refer to the 2D spacing relations as interaction Fundamental Diagrams (iFDs), which describe how the necessary space for vehicle interactions change with their relative speeds. We believe these relations contribute towards a deeper understanding of vehicle interaction and road space use in urban traffic, which may, for example, serve traffic engineers in better and safer design of intersection layouts.

The remainder of this paper is organised as follows. Section 2.2 presents the methods to infer 2D vehicle spacing. Section 2.3 explains our experiments on the urban trajectory dataset. Section 2.4 presents the experiment results and findings, and Section 2.5 discusses further on the results. Finally, Section 2.6 recaps the study.

2.2 Methods

Spacing in urban traffic is 2D as the distances drivers maintain from one another can vary depending on the direction. Two-dimensional spacing thus is equivalent to "driver space"[1], which is extended for drivers from the personal space of pedestrians [71, 72]. Here the driver space of a vehicle is not a determined area. Instead, we consider it as a spectrum of spacings, on which the critical spacing is where the driver can feel a rapid change in discomfort caused by proximity. Numerous studies have observed this transition from comfort to discomfort when drivers are approached by other vehicles [73–76]. Thereby, the driver space of a vehicle can be characterised by the less frequent presence of other vehicles in the surrounding, which are expected to be present more frequently if they do not cause higher levels of discomfort. Based on such violation against the expectation, we infer average 2D spacing by estimating the density of the accumulated presence of surrounding vehicles.

Figure 2.2 illustrates the process proposed in this study to infer average 2D vehicle spacing. The following paragraphs will briefly explain the process. Then the specific methods will be introduced in the subsections.

In Subsection 2.2.1, we firstly introduce a coordinate transformation for each pair of interacting vehicles consisting of an ego vehicle and a surrounding vehicle. The transformation emphasises the multi-directionality of the relative movement between

[1] Based on the research of Gibson and Crooks [67], Näätänen and Summala [68], Summala [69], and Bärgman et al. [70] define a similar concept, "comfort zone". The comfort zone is an area within which a driver feels comfortable. Other drivers are unwilling to cross the area without extra motives and will arouse discomfort when they do cross it. For consistency, we uniformly use "driver space" to refer to this area.

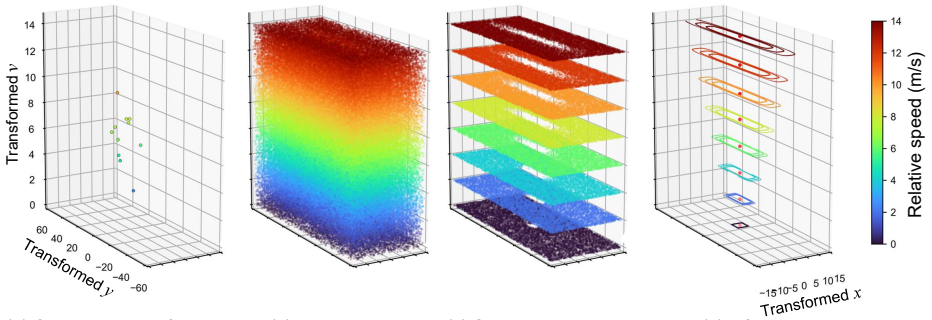


Figure 2.2 Proposed framework to infer average 2D vehicle spacing from trajectory data. **(a)** Coordinate transformation of vehicle pairs in Figure 2.1. **(b)** Accumulated data samples of transformed vehicle pairs. **(c)** Data samples that are conditioned by relative speeds, where the hollows are driver space. **(d)** Spacing inference from the conditioned data samples.

interacting vehicles and establishes a normalised reference frame for analysing vehicle interaction. Consequently, this enables the accumulation of vehicle pairs and consistent analysis across various scenarios.

With the accumulated samples of vehicle pairs in transformed coordinates, Subsection 2.2.2 utilises driver space as a proxy to measure vehicle spacing. The intrusion into driver space can cause varying levels of discomfort, and a driver's response to such discomfort is manifested by maintaining a naturally comfortable distance from other vehicles. We interpret this discomfort response as resistance to the proximity of other vehicles, and parameterise such proximity resistance by adapting the density function of the generalised Gaussian distribution. Proximity resistance describes the comfort-discomfort transition, based on which we can quantify the spacing between vehicles in a probabilistic manner.

Subsection 2.2.3 presents an algorithm to approximate proximity resistance by estimating the density of the accumulated presence of vehicles in the surrounding of an ego vehicle. The change in proximity resistance with spacing is described as a function, which is not fixed, but varies in different scenarios. For this reason, we design our method to be scenario-conditioned with respect to variables such as speed, acceleration, or/and other traffic situations. As an illustration of this method, we use the relative speed between interacting vehicles as the basic condition in this study. Ultimately, a series of functions with different parameters in different scenarios depict the average 2D spacing between interacting vehicles.

2.2.1 Coordinate transformation

To analyse vehicle spacing more consistently, we establish a new coordinate system referring to the relative movement of interacting vehicles. For an ego vehicle i and another vehicle j in the surrounding of i, we transform their global coordinates into a local coordinate system $C(O_i, \boldsymbol{v}_{ij})$. The origin O_i is located at the position of i, and the y-axis

points in the direction of $\boldsymbol{v}_{ij} = \boldsymbol{v}_i - \boldsymbol{v}_j$, the relative velocity of i to j. By doing this, all pairs of interacting vehicles share a common reference point (i.e., the ego vehicle's position) and an aligned orientation (i.e., the direction of the relative velocity of the ego vehicle to the surrounding vehicle).

2

Figure 2.3 visually illustrates the transformation, wherein we rotate and translate the original global coordinate system. We use $(x^{(g)}, y^{(g)})$ to denote the global position of a vehicle and $(x^{(ij)}, y^{(ij)})$ to denote the local coordinates of the vehicle in $C(O_i, \boldsymbol{v}_{ij})$. The transformation equation is formulated as

$$\begin{bmatrix} x^{(ij)} \\ y^{(ij)} \end{bmatrix} = \begin{bmatrix} \cos\rho & -\sin\rho \\ \sin\rho & \cos\rho \end{bmatrix} \left(\begin{bmatrix} x^{(g)} \\ y^{(g)} \end{bmatrix} + \begin{bmatrix} a \\ b \end{bmatrix} \right), \tag{2.1}$$

where ρ represents the counterclockwise rotation angle and $(a, b)^\top$ is the translation vector of the transformation.

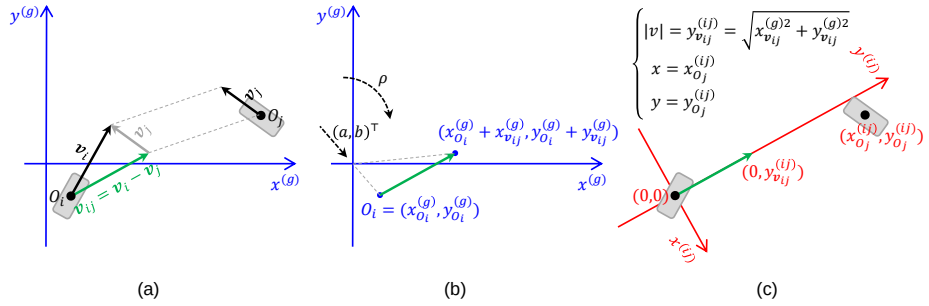


Figure 2.3 Coordinate transformation. **(a)** An ego vehicle i and another nearby vehicle j in the global coordinate system denoted by superscript (g), where O_i and O_j represent their positions, $\boldsymbol{v}_i$ and $\boldsymbol{v}_j$ for their velocities, and $\boldsymbol{v}_{ij}$ for the relative velocity of i to j. **(b)** Rotate the global coordinate system by ρ and translate $(a, b)^\top$ according to the reference point O_i and the reference orientation $\boldsymbol{v}_{ij}$. **(c)** Transformed coordinates in the local coordinate system denoted by superscript (ij).

The transformation parameters are solved as follows by substituting the coordinates in the global and local coordinate systems at the same position into Equation 2.1. We use two positions here. One is the position of i, which is $(x_{O_i}^{(g)}, y_{O_i}^{(g)})$ in the global coordinate system and $(0, 0)$ in $C(O_i, \boldsymbol{v}_{ij})$. The other is the position of the head of the relative velocity vector of i to j when its tail is located at O_i. This position is $(x_{O_i}^{(g)} + x_{\boldsymbol{v}_{ij}}^{(g)}, y_{O_i}^{(g)} + y_{\boldsymbol{v}_{ij}}^{(g)})$ in the global coordinate system and $(0, y_{\boldsymbol{v}_{ij}}^{(ij)})$ in $C(O_i, \boldsymbol{v}_{ij})$, where $y_{\boldsymbol{v}_{ij}}^{(ij)} = \sqrt{x_{\boldsymbol{v}_{ij}}^{(g)2} + y_{\boldsymbol{v}_{ij}}^{(g)2}}$. As a result, the transformation parameters between the global coordinate system and $C(O_i, \boldsymbol{v}_{ij})$ are

$$\begin{cases} a = -x_{O_i}^{(g)}, \\ b = -y_{O_i}^{(g)}, \\ \cos\rho = y_{\boldsymbol{v}_{ij}}^{(g)} / \sqrt{x_{\boldsymbol{v}_{ij}}^{(g)2} + y_{\boldsymbol{v}_{ij}}^{(g)2}}, \\ \sin\rho = x_{\boldsymbol{v}_{ij}}^{(g)} / \sqrt{x_{\boldsymbol{v}_{ij}}^{(g)2} + y_{\boldsymbol{v}_{ij}}^{(g)2}}. \end{cases} \tag{2.2}$$

Given that division by zero is undefined, Equations 2.2 require that $(x_{\boldsymbol{v}_{ij}}^{(g)}, y_{\boldsymbol{v}_{ij}}^{(g)}) \neq (0, 0)$, i.e., $\boldsymbol{v}_i \neq \boldsymbol{v}_j$. This implies that if vehicle i and vehicle j are moving at the same velocity simultaneously, their relative movement cannot be transformed into a local coordinate system defined here. Then we disregard such cases as i and j are temporarily stationary relative to each other.

Finally, we transform the global position of j into $C(O_i, \boldsymbol{v}_{ij})$ as $(x_{O_j}^{(ij)}, y_{O_j}^{(ij)})$. Denoting $x = x_{O_j}^{(ij)}$ and $y = y_{O_j}^{(ij)}$, and together with the mode of the relative velocity $|v| = y_{\boldsymbol{v}_{ij}}^{(ij)}$, this coordinate transformation compresses the information about the relative position and velocity between vehicle i and vehicle j. For easier notation, we will use v in place of $|v|$ in the rest of this paper. These three variables (x, y, v) encode two pieces of information. The first is the position of vehicle j relative to vehicle i. The second is the moving direction of vehicle i and j, where they move along the y-axis either towards ($y > 0$) or away from ($y < 0$) each other. More specifically, vehicles above the transformed x-axis are approaching the vehicle at the origin rather than being ahead of it; similarly, vehicles below the transformed x-axis are leaving it rather than being behind.

In the following sections of this paper, we will use (x, y, v) to refer to the transformed relative position and relative speed of a pair of vehicles i and j. It must be noted that this transformation makes the study specifically focus on the relative movement between vehicles. Readers are reminded that, if not specified, v refers to the relative speed between interacting vehicles rather than their absolute speeds in this study.

2.2.2 Probabilistic spacing

Spacing is considered probabilistic to account for its variation, and we characterise it utilising the concept of driver space. Intrusion into driver space causes discomfort, which does not preclude the spatial intrusion of other vehicles, but motivates a driver to maintain a distance from others at varying levels [77]. For an ego vehicle i, if another vehicle j is far outside the driver space of i, little discomfort is caused and the movement of i remains uninfluenced. In contrast, if vehicle j enters the driver space of i and keeps approaching, the escalating discomfort will compel the driver of i to move away [69, 78, 79]. We therefore introduce proximity resistance to interpret the discomfort response and specify the level of intrusion-caused discomfort. With a value $p \in (0, 1)$ denoting proximity resistance, $p \to 0$ indicates that no discomfort is caused and $p \to 1$ indicates extreme discomfort. In this way, any spacing in a specific scenario corresponds to a level of proximity resistance, which indicates the potential for causing discomfort.

Proximity resistance in one-dimensional spacing

We first consider proximity resistance in 1D spacing. The level of intrusion-caused discomfort has a transition from lower to higher across the the critical spacing where proximity resistance changes the most rapidly. Therefore, the representation of proximity resistance should encode how far the critical spacing is and how rapid the discomfort transition is. To this end, we adapt the density function of the generalised Gaussian distribution (GGD).

The GGD is a family of symmetric probability densities that generalises the Gaussian density [80]. Its function is given as $\alpha \exp(-|\frac{x-\mu}{r}|^{\beta})$, where α is the coefficient making the integral of the function 1, μ is the centre of symmetry, r is the scale parameter, and β is the shape parameter [81]. The family includes the density of Laplace distribution when $\beta = 1$ and of Gaussian distribution when $\beta = 2$. When $\beta \to \infty$, GGD converges to the density of a uniform distribution on $(\mu - r, \mu + r)$. Thereby, with $\beta \geq 2$, GGD covers a continuum of densities from the Gaussian distribution to the uniform distribution.

We use s to denote 1D spacing, which is the relative positions of surrounding vehicles to an ego vehicle. Thereby, the ego vehicle is at the origin and $\mu = 0$. We then present proximity resistance in 1D spacing as Equation (2.3). Proximity resistance is not a probability density, so we remove the coefficient α to ensure $p \in (0, 1)$.

$$p(s|r, \beta) = \exp\left(-\left|\frac{s}{r}\right|^{\beta}\right), \tag{2.3}$$

where $r > 0$ and $\beta \geq 2$. As shown in Figure 2.4, the scale parameter r suggests how far the critical spacing is, and the shape parameter β suggests how fast the discomfort transition is. This parameterisation captures the variation in proximity resistance at different relative positions, which characterises probabilistic spacing.

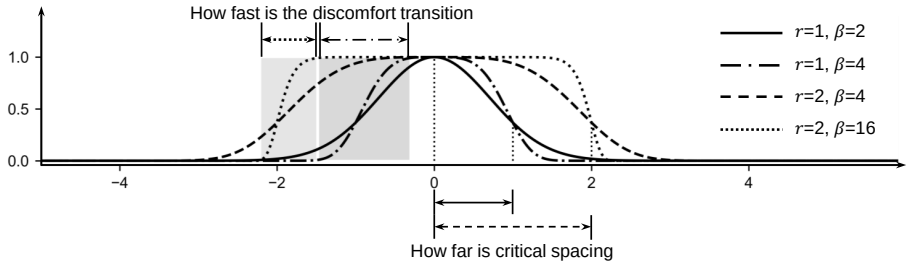


Figure 2.4 Parameters r and β determine the change of proximity resistance in 1D spacing.

Vehicles oriented in different directions may cause different levels of discomfort even being equally distant. We thus further consider two components for r and β, respectively, to allow Equation (2.3) to be asymmetric.

$$\begin{cases} r = \dfrac{1+\mathrm{sgn}(s)}{2} r^{+} + \dfrac{1-\mathrm{sgn}(s)}{2} r^{-}, \\ \beta = \dfrac{1+\mathrm{sgn}(s)}{2} \beta^{+} + \dfrac{1-\mathrm{sgn}(s)}{2} \beta^{-}. \end{cases} \tag{2.4}$$

Equation (2.4) means that $r = r^{+}$ and $\beta = \beta^{+}$ when $s > 0$, and $r = r^{-}$ and $\beta = \beta^{-}$ when $s < 0$.

Proximity resistance in two-dimensional spacing

Coordinate transformation in Subsection 2.2.1 places the ego vehicle at $(0, 0)$ and the surrounding vehicle at (x, y), where x and y are independent[2] in most cases because vehicles can take any place relative to the ego vehicle. Then the proximity resistance in 2D spacing, as shown in Equation (2.5), is an extension of Equation (2.3).

$$p(x, y|\boldsymbol{\theta}) = \exp\left(- \left| \frac{x}{r_x} \right|^{\beta_x} - \left| \frac{y}{r_y} \right|^{\beta_y} \right), \tag{2.5}$$

where each of $\boldsymbol{\theta} = (r_x, r_y, \beta_x, \beta_y)^{\top}$ has two components:

$$\begin{cases} \theta = \dfrac{1+\mathrm{sgn}(x)}{2} \theta^{+} + \dfrac{1-\mathrm{sgn}(x)}{2} \theta^{-} \text{ for } \theta = r_x, \beta_x, \\ \theta = \dfrac{1+\mathrm{sgn}(y)}{2} \theta^{+} + \dfrac{1-\mathrm{sgn}(y)}{2} \theta^{-} \text{ for } \theta = r_y, \beta_y. \end{cases} \tag{2.6}$$

Parameters in Equation (2.5) have realistic explanations. $\boldsymbol{r} = \{r_x^{+}, r_x^{-}, r_y^{+}, r_y^{-}\}$ determine the positions where $p = e^{-1}$ in different directions, and $\boldsymbol{\beta} = \{\beta_x^{+}, \beta_x^{-}, \beta_y^{+}, \beta_y^{-}\}$ determine the rate at which p varies with $x > 0$, $x < 0$, $y > 0$, and $y < 0$. Hereby, $\boldsymbol{r}$ control the spacing resulting in higher proximity resistance, and $\boldsymbol{\beta}$ control the increase rates of proximity resistance when vehicles are moving in different directions. In this way, $\boldsymbol{r}$ and $\boldsymbol{\beta}$ together characterise 2D probabilistic spacing.

Vehicle spacing can be influenced by the movement states of the interacting vehicles, the traffic situation they are in, as well as the driving preference of their drivers. These influences are diverse and may be interrelated, making it difficult to integrate them into an equation as independent variables. In this study, we consider such influences as the conditions for filtering vehicle pair samples. Therefore, $\boldsymbol{\theta}$ vary across conditioning situations.

2.2.3 Parameter inference

We infer the parameters in Equation (2.5) by estimating the density of accumulated presence of surrounding vehicles. The infrequent presence of surrounding vehicles reflects

[2]At some special scenarios such as roundabouts, x and y are correlated but can be transformed into the Frenet-Serret coordinate system to ensure their independence. In order to focus on the main idea, we do not delve into those special scenarios in this paper.

how resistant drivers are to approaching one another. Conversely, the frequent presence of surrounding vehicles reflects how acceptable this approach is. We refer to the former as proximity resistance and the latter as proximity tolerance, and they complement to each other. The density of the accumulated vehicle presence around a vehicle directly corresponds to proximity tolerance, so we can indirectly infer proximity resistance as the opposite to it.

In the transformed coordinate system, a pair of vehicles consists of an ego vehicle i at $(x_i, y_i) = (0, 0)$ and a surrounding vehicle j at (x_j, y_j). Denote the proximity resistance between the drivers of i and j as p_{ij}, then the corresponding proximity tolerance is $1 - p_{ij}$. $1 - p_{ij}$ can be seen as the relative likelihood of a vehicle passing (x_j, y_j), which is in proportion to the density of vehicles accumulatively passing (x_j, y_j). Given n pairs of vehicles i and j, the likelihood of the presence of the surrounding vehicles is

$$L = \prod_{j=1}^{n} \left[1 - p_{ij}(x_j, y_j|\boldsymbol{\theta})\right]. \tag{2.7}$$

We then infer p_{ij} by iteratively estimating $\boldsymbol{\theta}$ until they all converge, where $\boldsymbol{\beta}$ are estimated given $\boldsymbol{r}$ and $\boldsymbol{r}$ given $\boldsymbol{\beta}$.

The estimation of $\boldsymbol{\beta}$ aims to maximise the likelihood L, which allows for the sparsity of distant vehicles in data and prevents too slow discomfort transition. As the area covered by available data is fixed, the more distant the vehicles are, the less they are recorded and the lower their density. In this case, a purely density-based inference would be biased towards a slower discomfort transition. Maximising L avoids this bias as vehicles farther away from the ego vehicle are at high proximity tolerance close to 1, which results in little effect on L.

The estimation of $\boldsymbol{r}$ should distinguish the critical spacing as the boundaries of driver space, within which other vehicles infrequently access. Denoting the area where p is close to 1 as $A_{p\to1}$, the increase of a r (e.g., r_y^+) corresponds to the gradual expansion of $A_{p\to1}$ along the direction indicated by the r (i.e., the relative velocity direction). Holding other parameters constant, L undergoes three phases as each r increases. In the first phase, L increases slightly when almost all the surrounding vehicles are outside $A_{p\to1}$. In the second phase, L decreases slowly when some surrounding vehicles are present within $A_{p\to1}$. In the third phase, L decreases rapidly when more surrounding vehicles are erroneously covered by $A_{p\to1}$. As a result, maximising L leads to smaller $A_{p\to1}$ and closer critical spacing to the ego vehicle. Alternatively to maximising L, we estimate each r by the position at which L decreases from slowly to rapidly. This position is where the second-order derivative of L with respect to r is smallest (negative), signifying the fastest change in the decreasing rate of L.

Further practical details need to be considered in the estimation. For computational convenience, we take the logarithm of L, i.e., the sum of the log-likelihood of vehicle presence. This is equivalent to L for estimating $\boldsymbol{\theta}$ since logarithms are strictly increasing functions. Equation (2.5) assumes no vehicles passing through the central area of driver space. However, some vehicles indeed do so. These cases make the estimation

biased towards overly small $\boldsymbol{r}$ and $\boldsymbol{\beta}$ to keep their relative likelihood not too close to zero. To mitigate this bias, we add $\epsilon = 10^{-4}$ to the proximity tolerance, which is thus $1 + \epsilon - p_{ij} = 1 - (p_{ij} - \epsilon)$. This decreases the proximity resistance between vehicles that are extremely close to each other, and increases their relative likelihood. The adjusted log-likelihood of the presence of n surrounding vehicles is then

$$\ln(L) = \sum_{j=1}^{n} \ln\left[1 + \epsilon - p_{ij}(x_j, y_j|\boldsymbol{\theta})\right]. \tag{2.8}$$

Algorithm 1 summarises the iterative inference of probabilistic spacing. Because the derivative of $\ln(L)$ to one parameter depends on the other parameters, the algorithm is difficult to solve by hand. We thus use numerical methods for the differentiation and optimisation in the algorithm. This requires initial values for the parameters (i.e., $\boldsymbol{r}_0$ and $\boldsymbol{\beta}_0$). We set $\boldsymbol{r}_0$ to the 0.1th percentiles of the surrounding vehicle positions from near to far from the ego vehicle. This heuristic setting follows the belief that few vehicles will intrude into others' driver space. Then we set $\boldsymbol{\beta}_0$ to 2 to initiate a moderate discomfort transition.

The inference is achieved when the estimates of $\boldsymbol{r}$ and $\boldsymbol{\beta}$, i.e., $\hat{\boldsymbol{r}}$ and $\hat{\boldsymbol{\beta}}$, converge to either a set of values or several recurring sets of values. $\hat{\boldsymbol{\beta}}$ are estimated based on maximising $\ln(L)$, by which we can calculate the estimation confidence according to the numerical Hessian matrix of $\ln(L)$. When the inference converges to recurring sets of values, the set with the highest confidence for $\hat{\boldsymbol{\beta}}$ is taken. In addition, a maximum number of iterations is set to force a stop.

Algorithm 1: Parameter inference of 2D probabilistic spacing.

Data: $\boldsymbol{x}$ and $\boldsymbol{y}$ under particular conditions
Result: $\hat{\boldsymbol{r}}$ and $\hat{\boldsymbol{\beta}}$ under the particular conditions
begin
  Initialise $r \leftarrow r_0$ for $r = r_x^+, r_x^-, r_y^+, r_y^-$ and $\beta \leftarrow 2$ for $\beta = \beta_x^+, \beta_x^-, \beta_y^+, \beta_y^-$
  repeat
    $\hat{r} \leftarrow \arg\min_r \frac{\partial}{\partial r}\left(\frac{\partial ln(L)}{\partial r}\right)$ for $r = r_x^+, r_x^-, r_y^+, r_y^-$
    $\hat{\beta} \leftarrow \arg\max_\beta \ln(L)$ subjected to $\beta \geq 2$ for $\beta = \beta_x^+, \beta_x^-, \beta_y^+, \beta_y^-$
  until $\hat{\boldsymbol{r}}$ and $\hat{\boldsymbol{\beta}}$ converge **or** maximum number of iterations reaches

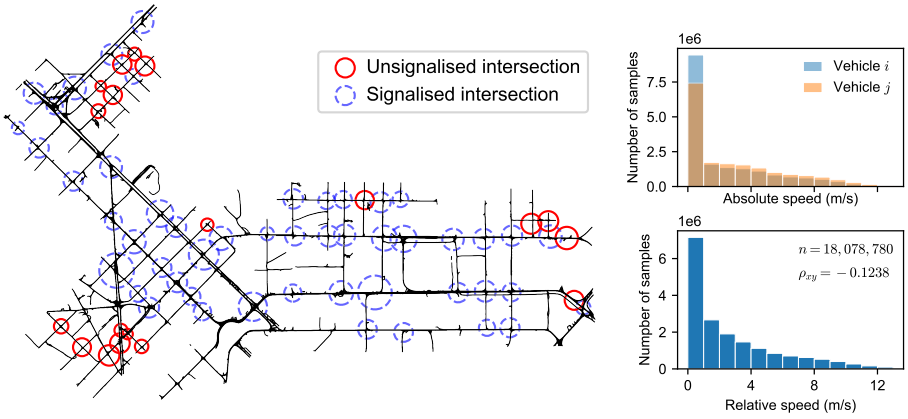
2.3 Experiments

To demonstrate the proposed method, we applied it on pNEUMA [66], a large-scale urban trajectory dataset. With a swarm of drones, pNEUMA recorded vehicle traces in a central business district of Athens, covering an area of 1.3 km^2 with over 100 km-lanes of road

network and around 100 busy intersections. This area was divided into 10 subareas and each was covered by a drone. Due to limited flight time, the drones recorded 9–15 minutes per half hour between 8:00 and 11:00 a.m. on October 24th, 29th, 30th, and November 1st, 2018. To compare the results consistently, we use data during the latter three consecutive days where the trajectories were collected from 8:00 to 10:30 a.m. each day.

2.3.1 Preprocessing

This study pays particular attention to 2D vehicle interaction, therefore, we focused on trajectories at road intersections in the dataset. As shown in Figure 2.5(a), we extracted major T-junctions and crossing intersections following the methods proposed in J. Wang et al. [82] and Cao and Krumm [83]. In total 50 signalised intersections and 19 unsignalised intersections were identified. As 2D vehicle interaction is significantly reduced by signal control at signalised intersections, our experiments were focused on the unsignalised intersections. Most of these unsignalised intersections are single-lane and the recorded trajectories are insufficient to infer driver spaces for each of them, so we combined all the unsignalised intersections as an aggregated study area.



(a) Identified intersections in the pNEUMA dataset **(b)** Speed distribution

Figure 2.5 Map of intersections in the dataset and vehicle sample statistics.

The data preprocessing exists of four steps. First, pedestrians, bicycles, and motorcycles (all are labelled "Mortorcycle" in the dataset) were excluded as they have very different movement characteristics from motor vehicles regarding speed and occupied space. Second, vehicles that never moved were also excluded as they did not interact with any others. Third, vehicles that had overlapping trajectories with another vehicle (position distance $<$0.5 m at every moment) were removed as these are assumed to be data errors or outliers. Such errors could be caused by the vision algorithm used to track the vehicles. Fourth, vehicle pairs were sampled at a series of moments 0.2 s apart. Seeing one in the pair as an ego vehicle and the other as a surrounding vehicle, their position coordinates

were transformed according to Equations (2.1) and (2.2). The histograms in Figure 2.5(b) show the distributions of absolute speeds and relative speeds of extracted samples. The number of samples (n) and the correlation coefficients of transformed sample positions (ρ_{xy}) are also annotated.

2.3.2 Experiment setting

In this study, we use two conditions to specify scenarios. The first is whether the ego vehicle and the surrounding vehicle are interacting in the longitudinal direction (the ego vehicle's heading direction) only or also in the lateral direction (perpendicular to the ego vehicle's heading direction). We calculate the angle between the moving directions of the two vehicles and denote it as ω. If $\omega < 5^\circ$ or $\omega > 175^\circ$, they are considered to interact solely in the longitudinal direction (e.g., car-following and head-on conflict); otherwise, their interaction is considered to also involve the lateral direction (e.g., lane-changing and turning). We will refer to the two cases respectively as *Longitudinal* and *Lateral* in the following. Figure 2.1 in Section 2.1 is referred to for some specific examples. The second condition is the relative speed between the ego vehicle and the surrounding vehicle. Samples separated by the first condition are sorted according to their relative speeds and further grouped. The grouping ensures at least 50,000 samples per group while maintaining the difference in average relative speed between groups at least 0.1 m/s apart.

We obtained 80 *Longitudinal* scenarios and 95 *Lateral* scenarios, where in each scenario an average ego vehicle was abstracted from all vehicle pairs. The inference of average spacing was performed for each scenario, which estimated series of $\hat{\boldsymbol{r}}$ and $\hat{\boldsymbol{\beta}}$. Among the estimated parameters, $\hat{\boldsymbol{r}} = \{\hat{r}_x^+, \hat{r}_x^-, \hat{r}_y^+, \hat{r}_y^-\}$ indicate the critical spacing where the transition of intrusion-caused discomfort is the most significant; $\hat{\boldsymbol{\beta}} = \{\hat{\beta}_x^+, \hat{\beta}_x^-, \hat{\beta}_y^+, \hat{\beta}_y^-\}$ depict the transition buffer from comfort to discomfort.

The values of $\hat{r}_y^+$ and $\hat{r}_y^-$ represent critical spacing along the direction of relative velocity when vehicles are approaching and leaving each other in a given scenario. In contrast, the values of $\hat{r}_x^+$ and $\hat{r}_x^-$ represent spacing along the direction perpendicular to the relative velocity, which is associated with the shape of the road section that is analysed. In *Longitudinal* scenarios, $\hat{r}_x^+$ and $\hat{r}_x^-$ correspond to the road widths, whereas in *Lateral* scenarios, they can be referred to the space within the intersection. $\hat{r}_x^+$ and $\hat{r}_x^-$ can also indicate whether the samples are evenly distributed on both sides of the y-axis. For example, if $\hat{r}_x^- > \hat{r}_x^+$, the sample density is significantly lower on the left side of the y-axis than on the right. To prevent this bias influencing the inference of $\hat{\boldsymbol{r}}_y$, we set an upper limit of 7 m to $\hat{\boldsymbol{r}}_x$ in the experiments. In addition, the inference results were rejected when at least one of the p-values of $\hat{\boldsymbol{\beta}}$ is larger than 0.05, as they are statistically unreliable.

2.4 Results and findings

2.4.1 Critical spacing

Figure 2.6 shows several inference results as examples. Scenarios where the average relative speed is closest to 0, 2, 4, 6, and 8 m/s are shown as five pairs of coloured scatter

plots. The left side of each pair shows *Longitudinal* case and the right side shows *Lateral* case. Scatters in these plots refer to vehicle positions in the transformed coordinate system. They are coloured to indicate the inferred proximity resistance between the surrounding vehicles and the ego vehicles.

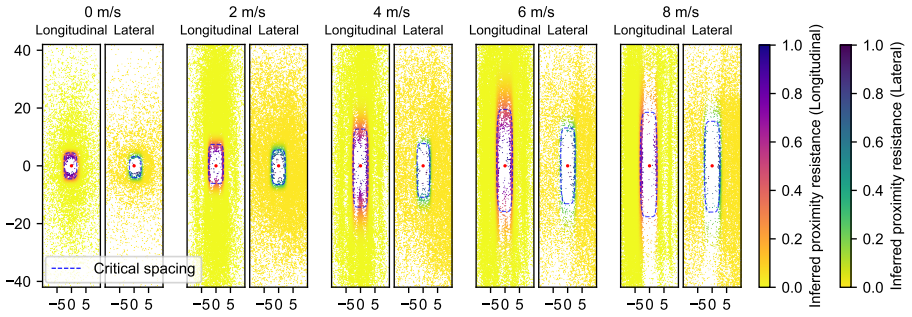


Figure 2.6 Inference results in various scenarios. The averaged ego vehicles are marked with red circles at $(0, 0)$. Vehicles surrounding their ego vehicles are shown as scatters and coloured to indicate the inferred proximity resistance. The contours indicate critical spacing where proximity resistance is e^{-1}.

Driver spaces are the nearly empty areas with significantly fewer vehicles. These spaces are readily visible in the scatter plots, and our method properly depicts them. Their boundaries presented as the contours are the inferred critical spacing. As is seen in Figure 2.6, the critical spacing increases with greater relative speeds in both *Longitudinal* and *Lateral* scenarios. However, it appears to be longer in *Longitudinal* than in *Lateral* scenarios at the same relative speed.

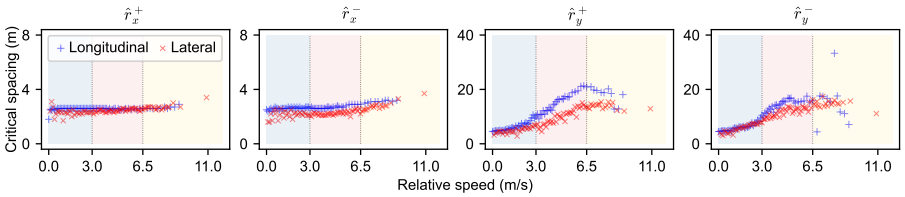


Figure 2.7 Inferred critical spacing in different scenarios. $\hat{r}_x^+$ and $\hat{r}_x^-$ are critical spacing perpendicular to the relative velocity direction between interacting vehicles; $\hat{r}_y^+$ and $\hat{r}_y^-$ are that along the relative velocity direction between approaching and leaving vehicles, respectively. Two dotted lines where the respective relative speed is 3 m/s and 6.5 m/s divide each subplot into 3 cases under lower, medium, and higher relative speeds.

To further see how critical spacing changes in different scenarios, Figure 2.7 displays the inferred critical spacing $\hat{\boldsymbol{r}}$ in various relative speeds. Overall, critical spacing expands along the relative velocity between interacting vehicles. This expansion is not limitless and eventually reaches a plateau after the relative velocity exceeds a certain threshold.

At relative speeds below around 3 m/s, the critical spacing along the relative velocity direction is approximately the same in *Longitudinal* and *Lateral* scenarios. However, when the relative speed is larger than 3 m/s, the critical spacing in *Longitudinal* scenarios exceeds that in *Lateral* scenarios. This exceeding is particularly evident between approaching vehicles, as opposed to vehicles that are leaving away from each other. A possible explanation is different interaction strategies available in different scenarios. Specifically, vehicle interaction in *Longitudinal* scenarios is limited to adjusting spacing along the moving direction; while in interaction scenarios involving other directions, vehicles can make turns and thus require less space.

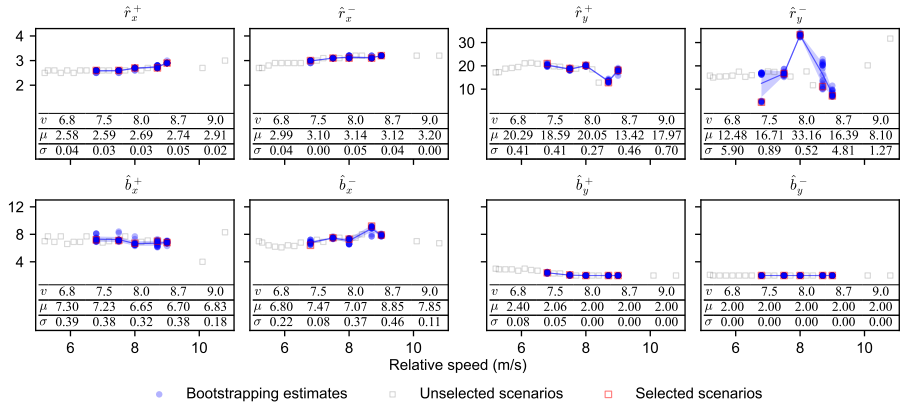
2.4.2 Consistency evaluation

As is seen in Figure 2.7, the inferred critical spacing do not always expand smoothly. Particularly at higher relative speeds larger than 6.5 m/s, some r_y^+ and r_y^- diverge from the general curve and appear to be outliers. This entails the possibility that the proposed methods may produce inconsistent or unreliable inferences.

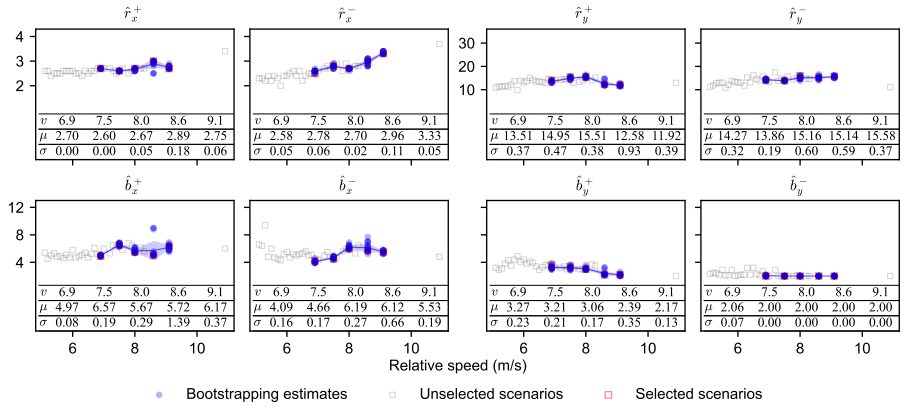
As the ground truth of critical spacing between vehicles is unknown, evaluating the reliability of the inference is not feasible. However, we can assess its consistency through bootstrapping. In this study, $\boldsymbol{r}$ and $\boldsymbol{\beta}$ parameterise proximity resistance change in different spacing and are considered to depend only on the specific scenario where the analysed samples lie. Therefore, the distributions of $\hat{\boldsymbol{r}}$ and $\hat{\boldsymbol{\beta}}$ should be identical for any sampling in the same scenario, which means they are pivotal. In this case, according to Davidson and MacKinnon [84], the minimum number of bootstrapping iterations is 19 when considering a significance level of 0.05. We select 5 scenarios with relative speeds over 6.5 m/s respectively for *Longitudinal* and *Lateral* scenarios, and perform bootstrapping with 20 iterations on them. In every iteration for each scenario, we randomly select 85% of the vehicle samples with replacement, and then apply our method to infer the parameters.

Figure 2.8 displays the bootstrapping results, where two sub-figures correspond to *Longitudinal* and *Lateral* scenarios. In this figure, the estimates of parameters are plotted as dots. The means of the estimates over different iterations in the same scenario are plotted as lines, and the standard deviations are plotted as shaded regions. In addition, the tables at the bottom also numerically show the statistics of bootstrapping results. To read the figure, the consistency of our inference is higher when the dots are more tightly distributed, when the shaded regions are narrower, and when the standard deviations are lower.

These plots and tables show that our method delivers consistent inferences overall. Notably in cases where the estimates appear to be outliers, such as $\hat{r}_y^-$ in *Longitudinal* scenarios, our method still gives consistent results. Referring back to Figure 2.6, one can see that the lower half of the *Longitudinal* scatter plot in the relative speed of 8 m/s does not present a clear boundary between areas of higher and lower sample densities. This unclear boundary violates the assumption that vehicles appear significantly more frequently outside of driver space than inside, and thus challenges the inference of $\hat{r}_y^-$ as critical leaving spacing. Therefore, in contrast with potential inconsistencies in the



(a) *Longitudinal* scenarios



(b) *Lateral* scenarios

Figure 2.8 Consistency evaluation of the proposed method through bootstrapping. In the bottom table of each sub-plot, v indicate the relative speeds of the selected scenarios; μ and σ indicate the mean values and standard deviations of bootstrapping estimates in the selected scenarios, respectively.

inference method, the success in inferring driver space may depend more on the quality of data and sampling that establish clear boundaries as critical spacing.

2.4.3 Interaction Fundamental Diagram

In this study, we quantify average 2D spacing in urban traffic and its relationship with relative speeds of interacting vehicles. This allows for deriving urban traffic states and their relationship, which we term as interaction Fundamental Diagram (iFD).

The accumulated samples of interacting vehicles are transformed to be in the same reference system and are conditioned at the same relative speed. As a result, we can define quasi-density k and quasi-flow q for vehicle interaction, as shown in Equations (2.9). We calculate k as the inverse of D, where D is the amount of necessary space for interaction under a certain relative speed v. As a function of v, k represents for *interaction density*. It describes space occupancy of vehicle interaction at different relative speeds. A higher interaction density means more interaction can occur within given road space, while a lower density means fewer. We then calculate q as the product of k and v, which represents for *interaction rate*. This multiplication is feasible due to coordinate transformation, after which the ego vehicles move uniformly along the y-axis. As also a function of v, q describes time occupancy of vehicle interaction at different relative speeds. A higher interaction rate means faster interaction and is related to interaction efficiency.

$$\begin{cases} k = \dfrac{1}{D}, \\ q = kv. \end{cases} \tag{2.9}$$

The area of necessary space for vehicle interaction can vary given different levels of proximity resistance. With a proximity resistance of p^*, the dimensions of necessary space under a certain scenario can be computed by solving the inverse of Equation (2.5). Equation (2.10) shows the solution from inferred parameters $\hat{\boldsymbol{r}}$ and $\hat{\boldsymbol{\beta}}$. As noted in Subsection 2.3.2, $\hat{\boldsymbol{r}}$ quantify the critical spacing and $\hat{\boldsymbol{\beta}}$ quantify the buffer across them.

$$\boldsymbol{d}(p^*) = \hat{\boldsymbol{r}}(-\ln(p^*))^{1/\hat{\boldsymbol{\beta}}}, \tag{2.10}$$

where $\boldsymbol{d} = \{d_x^+, d_x^-, d_y^+, d_y^-\}$ and correspond to $\{\hat{r}_x^+, \hat{r}_x^-, \hat{r}_y^+, \hat{r}_y^-\}$ and $\{\hat{\beta}_x^+, \hat{\beta}_x^-, \hat{\beta}_y^+, \hat{\beta}_y^-\}$. Then the area of necessary interaction space where the proximity resistance is smaller than p^* is computed as

$$D(p^*) = d_y^+(p^*)[(d_x^-(p^*) + d_x^+(p^*)]. \tag{2.11}$$

Here we do not calculate the space area for vehicles leaving each other (i.e., we ignored $d_y^-(p^*)$) as vehicles moving away are less restricted by one another.

Figure 2.9 presents the relationships between k, q, and v given different levels of proximity resistance. The relationships resemble FDs and so we refer to them iFDs. The blue dashed lines are iFDs at critical spacing (where proximity resistance is e^{-1}), which mark the boundaries of frequent (less transparent) and infrequent (more transparent) occurrence of interaction states. Compared to *Longitudinal* iFDs, *Lateral* iFDs have higher interaction densities and interaction rates in various relative speeds, which suggests that lateral interactions use less road space and are more accommodated than longitudinal interactions at these unsignalised intersections.

Overall, as relative speeds increase, interaction density decreases; meanwhile, interaction rate first rises and then falls. In the initial phase of increasing interaction density, the necessary space for vehicle interaction is reduced. As a result, the limited road

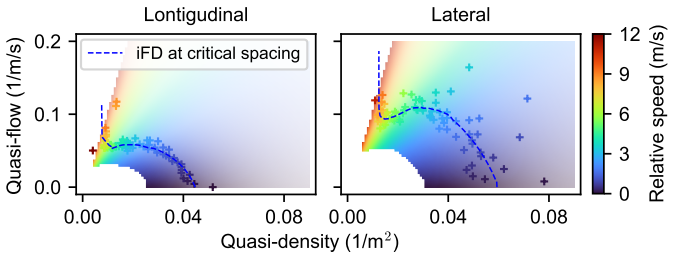


Figure 2.9 Relationships between quasi-density and quasi-flow defined for urban vehicle interaction (grey scale may confuse speed and probability). These diagrams are probabilistic at different levels of proximity resistance, where the surface with varying colours and transparencies was created using smoothed curves to show the transition more clearly. The more transparent the colour, the higher the proximity resistance and the less probable vehicles are in that state.

space can accommodate more interaction, leading to an increased interaction rate. Then in the second phase of rising interaction density, the available road space decreases faster than the reduction in necessary interaction space. Consequently, the interaction rate declines due to insufficient road space for vehicle interaction. The highest level of interaction rate represents the optimal point in the relationship, signifying where vehicle interaction is most often.

We can observe other interesting phenomena in the diagrams. To different extent in both *Longitudinal* and *Lateral* scenarios, there are noticeable drops in interaction rate for high relative speeds, with very limited increase in interaction density. This is comparable to capacity drops in longitudinal traffic flow theory, implying higher variation in driving behaviour at those particular states [85, 86]. Moreover, the diffusion of interaction states at different levels of proximity resistance suggests that stochastic spacing is a contributing factor to the scatters in empirical FDs. This is also consistent with previous studies [e.g., 87–89].

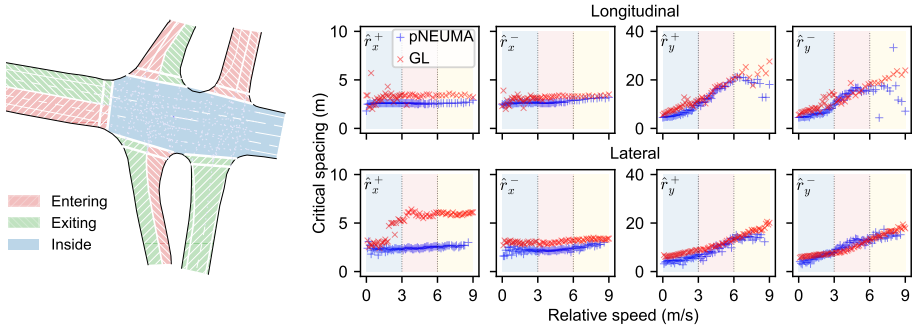
2.5 Discussion

2.5.1 Impact of intersection layout

In Figures 2.6 and 2.8, $\hat{r}_x^+$ and $\hat{r}_x^-$ show very limited variation as relative speeds increase, and approximately correspond to road widths (around 3 m). As stated in Section 2.3.2, the values of $\hat{\boldsymbol{r}}_x$ are correlated with the infrastructure layout where the trajectory data are collected. Therefore, the seemingly constant $\hat{\boldsymbol{r}}_x$ are potentially related to the fact that almost all of the unsignalised intersections in the pNEUMA dataset are junctions of single-lane roads.

To investigate this conjecture further, we applied the same inference to a two-lane unsignalised intersection called *GL* in the INTERACTION dataset [90], which was collected in the U.S. Figure 2.10(a) shows a map of this intersection with the entering lane,

the exiting lane, and the interior section marked. In addition, when sampling vehicle pairs, we excluded instances where one vehicle in a pair is in an entering lane while the other vehicle in the pair is in an exiting lane.



(a) Map of intersection *GL* **(b)** Critical spacing

Figure 2.10 Comparison of critical spacing in different scenarios between the intersection *GL* in INTERACTION and the unsignalised intersections in pNEUMA.

The inference results are displayed in Figure 2.10(b) and are compared with the results from the unsignalised intersections in pNEUMA. As expected, in both *Longitudinal* and *Lateral* scenarios, $\hat{r}_x^+$ and $\hat{r}_x^-$ are larger at the intersection *GL* than at the pNEUMA intersections. This is particularly evident for $\hat{r}_x^+$ in *Lateral* scenarios, which could be a consequence of two factors. One is the existence of yield lines at the entrances of the intersection legs, and the other is the busy entering lanes on the right side of vehicles entering the intersection on the main road.

It is necessary to underline that within the context of our methods, $\hat{\boldsymbol{r}}_x$ do not equate to "lateral spacing" in the sense of the original coordinate system. The coordinate transformation in Section 2.2.1 alters the view to observe the relative movement between vehicles, by simply rotating and translating the coordinate reference frame. This makes $\hat{\boldsymbol{r}}_y$ carry 2D information and be updated as vehicle interaction evolves. In Figure 2.10(b), $\hat{r}_y^+$ and $\hat{r}_y^-$ exhibit comparable patterns at the intersection *GL* and the pNEUMA intersections. This comparability further supports and validates our methods.

2.5.2 Impact of absolute speed

In this study, the importance of velocity differences between vehicles was emphasised for understanding 2D vehicle interaction. In light of such emphasis, our coordinate transformation omits the absolute speed, while it is often considered a crucial factor in determining vehicle spacing and interaction [91, 92].

To investigate the impact of vehicles' absolute speeds, we examined how the spacing between vehicles varies across different absolute speeds of the ego vehicle. Figure 2.11 shows the examination for both *Longitudinal* and *Lateral* scenarios under four different intervals of relative speed.

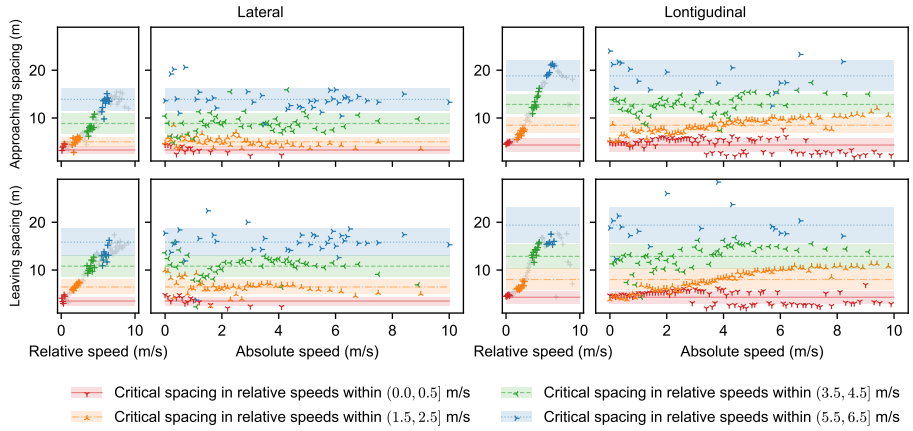


Figure 2.11 Comparison of critical spacing across varying absolute speeds of the ego vehicle. For convenience of comparison, each sub-figure has a plot on the left showing the inferred critical spacing conditioned by relative speeds. The lines and shadows are mean values and standard deviations of the cases with varying absolute speeds but at similar relative speed in the main plot on the right.

In *Lateral* scenarios, the critical spacing under the same interval of relative speed remains fairly consistent across absolute speeds ranging from 0 m/s to 10 m/s. This indicates that the impact of absolute speed is limited for *Lateral* interaction. While in *Longitudinal* scenarios there is a slight trend of increasing critical spacing with higher absolute speeds, it mainly hovers around the spacing determined by the relative speed. This observation is reasonable as we are measuring critical spacing, which, when conditioned by relative speed, is closely related to minimum Time-to-Collision (TTC). In Figure 2.12, we show two plots adapted from other researchers' empirical statistics of minimum TTC during 1D longitudinal interaction. As is clearly seen, minimum TTC is hardly influenced by absolute speeds under 20 m/s. Indeed, in urban trajectory datasets such as pNEUMA

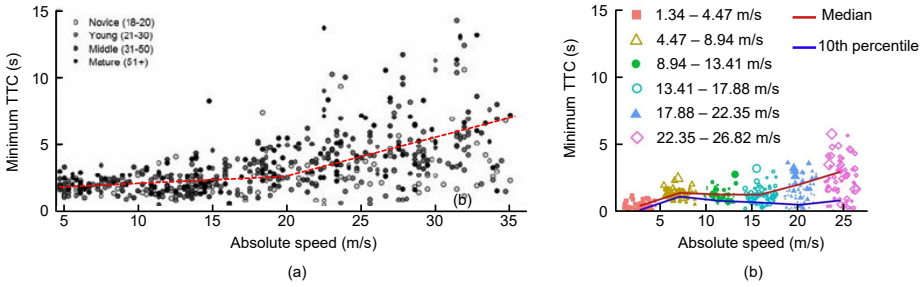


Figure 2.12 The distribution of minimum Time-to-Collision with respect to absolute driving speed in the literature. The unit of speed were miles per hour in the original studies, and we converted them to meters per second. **(a)** Adapted from Figure 4 in Montgomery et al. [93]. **(b)** Adapted from Figure 8 in Kusano et al. [24].

and INTERACTION, the absolute speeds of vehicles at intersections generally do not exceed 15 m/s.

We hereby suggest that the relative velocity direction between vehicles serves as a sufficient reference in the context of 2D spacing in urban traffic. The minimal impact of absolute speed on 2D critical spacing can be potentially explained from two aspects. First, interaction between vehicles occurs more in congested urban traffic and their absolute speeds are significantly slower than on highways. During interaction, the relative velocity between vehicles plays a more critical role [94, 95]. Second, urban disruptions such as districts and intersections make the traffic discontinuous. Drivers may therefore be less sensitive to the changes in absolute speed in urban traffic.

2.6 Conclusion

In this study, we propose a method to infer average 2D spacing between vehicles from urban trajectory data. Our method is built upon interactions of vehicle pairs. We first transform the global coordinates of vehicle pairs into local coordinate systems based on their relative movement during interaction. The coordinate transformation establishes a normalised reference system for consistent analysis. We then accumulate all vehicle pairs together, and by estimating their spatial distribution, we infer the average 2D spacing between interacting vehicles at various relative speeds. Experiments on real-world urban trajectory data demonstrate that the inference results are consistent and behaviourally intuitive. Our inference enables the derivation of urban traffic states from the perspective of relative movement between interacting vehicles. Further, the relations between the derived traffic states are examined as interaction Fundamental Diagrams (iFDs), which describe the average amount of necessary road space for vehicle interactions at different relative speeds. Thereby, this study provides new methods and findings about vehicle interaction and traffic state estimation in cities.

Measuring 2D vehicle spacing through a driver space lens has two methodological limitations. First, the effectiveness of spacing inference is restricted by data. Driver space is delineated from the accumulated presence of vehicles. When the vehicle samples are inadequate to form an intuitively proper driver space, the spacing between vehicles may not be inferred. Second, the inferred spacing is a result of response mixture. Particularly in *Lateral* scenarios, the inference mixes ego vehicle driver's response and that of surrounding vehicle drivers. By setting the condition of static surrounding vehicles, spacing can be inferred as merely a result of discomfort response by the ego vehicle driver. Nevertheless, further research is needed to investigate this response mixture.

This study lays a methodological foundation for future research in the field of urban traffic analysis. Various conditions can be used to specify scenarios of interest in order to investigate urban traffic variations more thoroughly. For example, traffic states and iFDs may be compared across interconnected urban intersections to evaluate their service levels. Additionally, comparisons between different cities and in different time periods are also valuable for urban traffic management. Given adequate data, it would be possible to analyse average spacing between vehicles during distinct interactions, such as

car-following, heading-on encounters, lane-changing, and turning. Such analysis will enhance our understanding about vehicle interaction in urban environments, and can contribute to improving traffic efficiency and reducing congestion at intersections.

Chapter 3

Adaptive boundary between safe and unsafe traffic interactions

Highlights

- Collision warning involves a trade-off between missed and false alarms.
- Probabilities of missed and false alarms are estimated from spacing distributions.
- Critical spacing is optimised to minimise missed and false alarms.
- Validation on synthetic and real-world conflicts confirms superior performance.
- Collision warning can be adaptive in varying traffic contexts and driver preferences.

Keywords

Advanced Driving Assistance System, Forward Collision Warning, collision avoidance, conflict detection, vehicle spacing patterns

This chapter is based on the conference article: Yiru Jiao, Simeon C. Calvert, Hans van Lint. (2024, June 2–5). Minimising Missed and False Alarms: A Vehicle Spacing based Approach to Conflict Detection. 35th IEEE Intelligent Vehicles Symposium, Jeju Island, South Korea. doi:10.1109/IV55156.2024.10588396

Abstract

Safety is the cornerstone of L2+ autonomous driving and one of the fundamental tasks is forward collision warning that detects potential rear-end collisions. Potential collisions are also known as conflicts, which have long been indicated using Time-to-Collision (TTC) with a critical threshold to distinguish safe and unsafe situations. Such indication, however, focuses on a single scenario and cannot cope with dynamic traffic environments. For example, TTC-based crash warning frequently misses potential collisions in congested traffic, and issues false alarms during lane-changing or parking. Aiming to minimise missed and false alarms in conflict detection, this study proposes a more reliable approach based on vehicle spacing patterns. To test this approach, we use both synthetic and real-world conflict data. Our experiments show that the proposed approach outperforms single-threshold TTC unless conflicts happened in the exact way that TTC is defined, which is rarely true. When conflicts are heterogeneous and when the information of conflict situation is incompletely known, as is the case with real-world conflicts, our approach can achieve less missed and false detection. This study offers a new perspective for conflict detection, and also a general framework allowing for further elaboration to minimise missed and false alarms. Less missed alarms will contribute to fewer accidents, meanwhile, fewer false alarms will promote people's trust in collision avoidance systems. We thus expect this study to contribute to safer and more trustworthy autonomous driving.

Code availability

`https://github.com/Yiru-Jiao/Conflict-detection-MFaM`

Data availability

Raw data sources:

- CitySim `https://github.com/UCF-SST-Lab/UCF-SST-CitySim1-Dataset`
- 100-Car Data `https://doi.org/10.15787/VTT1/CEU6RB`

Resulting data: `https://doi.org/10.4121/252a79e7-d9ff-4181-a9e4-842ea7845a77`

3.1 Introduction

The common concern of driving safety is one of the imperative aspects in the development of Advanced Driving Assistance Systems (ADAS). To prevent accidents and mitigate crash severity, collision avoidance systems (CAS) play a critical role. In general, CAS encompasses two proactive components: forward collision warning (FCW) which alerts drivers to imminent collisions, and automatic emergency braking (AEB) which initiates corrective actions such as braking or steering when the driver fails to respond timely. The effectiveness of both FCW and AEB hinges on the accurate detection of conflicting vehicular interactions, which potentially entail collisions.

Conflict detection leverages data collected from various sensors such as radar, lidar, and cameras. In the past two decades, real-time road user detection and tracking have been the predominant challenge in CAS [96], and substantial research efforts have been devoted to this task [97–99]. Along with the rapidly evolving advances in computer vision, constant improvements have been made on object detection and tracking [100, 101]. Nowadays, these techniques are extensively employed in current intelligent cars.

With increasingly accurate localisation of other road users, forward conflicts can be indicated using the surrogate safety measure, Time-to-Collision (TTC). TTC is one of the most effective and broadly used indicators for rear-end collisions [102–105]. It estimates how much time remains until a collision between two vehicles following each other [8, 21]. As an continuous variable, TTC has been used to assess pedestrian-vehicle interaction risk [106, 107], vehicular collision risk [39, 108, 109], and safe autonomous driving [110, 111]. When applied to collision warning, TTC needs to be discretised with critical thresholds in order to distinguish un(safe) situations. More specifically, a TTC value shorter than the critical threshold indicates high enough risk of a collision. Such a threshold determines when CAS should issue emergency warnings and intervene if the driver does not take action [112].

While threshold-based detection is straightforward and computationally efficient, it often falls short in dealing with dynamic traffic environments and more complex driving interactions [113]. For example, when vehicles maintain similar speeds in relatively dense traffic, the TTC between them is very large and suggests a low risk of collision. There then may be missed alarms, as traffic fluctuations can propagate and pose unexpected hazards. For another example during lane-changes, vehicles may exhibit short TTC values that indicate a high risk of crash. However, such alarms can be false as the drivers of the interacting vehicles often anticipate each others' actions and would not perceive an imminent threat. As such, threshold-based conflict detection can yield inconsistent reliability across different driving conditions.

Reliable conflict detection requires minimising missed and false alarms, which remains a challenge. According to an analysis of the consumer complaints about safety-relevant ADAS failures [41], more than 30% of complaints are about AEB and FCW, and over 75% of these complaints are about missed and false alarms. Missed alarms overlook dangerous scenarios and can preclude intervention opportunities, while false alarms might trigger distracting or even disruptive driver responses [114, 115]. In addition, many studies have

found that false alarms may diminish drivers' trust and compliance with the assistance systems [116–118]. Frequent incorrect alarms violate drivers' expectations about system warnings, and hence undermine their behavioral adaptation to ADAS.

Therefore, there is an increasing need for a more reliable method to detect conflicts and prevent collisions. In this paper, we provide a new approach to minimising missed and false detection of conflicts. With preliminary experiments, we will demonstrate that this method is characteristic of

- data-driven detection based on vehicle spacing patterns;
- adjustable balance between missed and false alarms;
- adaptive-tuning to varying traffic scenarios and driver preferences.

3.2 Methods

3.2.1 Indication of conflicts

An interaction between two vehicles can be indicated as safe or unsafe based on the information gathered about this scenario. An unsafe interaction is a conflict[1]. For a scenario at time t with a vehicle i and another vehicle j, a set of variables can describe this scenario and be denoted by $X_{ij}^t = \{\boldsymbol{x}_i^t, \boldsymbol{x}_j^t, \boldsymbol{x}_E^t\}$. Here $\boldsymbol{x}_i^t$ and $\boldsymbol{x}_j^t$ respectively encapsulate the motion of vehicles i and j, and $\boldsymbol{x}_E^t$ is about the physical environment such as weather and road conditions. We can process the information X_{ij}^t as represented in Equation (3.1):

$$X_{ij}^t = \{s_{ij}^t, \theta_{ij}^t\}, \tag{3.1}$$

where s_{ij}^t is the spacing between vehicles i and j, and θ_{ij}^t encodes the conflict situation where this scenario occur. Generally, a smaller value of s_{ij}^t suggests a higher likelihood that vehicles i and j are in a conflict and a collision could happen. A critical spacing s^* is then required to determine whether vehicles i and j are close enough to be considered as a conflict. As formulated in Equation (3.2), $C(X_{ij}^t)$ indicates the scenario X_{ij}^t as a conflict (abbreviated as c) or a non-conflict (nc) by comparing s_{ij}^t and s^*, where s^* depends on the specific situation captured by θ_{ij}^t.

$$C(X_{ij}^t) = C(s_{ij}^t, s^* | \theta_{ij}^t) = \begin{cases} \text{c}, & \text{if } s_{ij}^t \leq s^*(\theta_{ij}^t), \\ \text{nc}, & \text{otherwise.} \end{cases} \tag{3.2}$$

Existing conflict indicators (of which most are surrogate safety measures) can all fit in this expression. For example, TTC is typically calculated by assuming no change in movement of the interacting vehicles, i.e., the drivers are unable to react in time. This assumption considers the relative speed between vehicles as the only condition so that $\theta_{ij}^t = \Delta v_{ij}^t$. In this case, given a threshold TTC*, the critical spacing $s^* = \Delta v_{ij}^t \text{TTC}^*$. Time headway (THW) is another widely used conflict indicator [102]. Let vehicle j follow vehicle i, $\text{THW}_{ij}^t = s_{ij}^t / v_j^t$ thus $\theta_{ij}^t = v_j^t$. Then the speed of the following vehicle j

[1] Here we consider car-following and thus one-to-one interaction. Multiple vehicles can be considered in other interaction scenarios

becomes the condition considered. Given a threshold THW*, equivalently, the critical spacing is v_j^tTHW*.

3.2.2 Probability of missed and false alarms

Under a specific interaction situation θ, conflict detection is a binary classification based on vehicle spacing s and a critical threshold s^*. The probability distributions of spacing s respectively of conflicts and non-conflicts may overlap, as illustrated in Fig. 3.1. Therefore, determining the critical spacing s^* involves a trade-off between missed and false alarms. *Missed alarms* misclassify unsafe scenarios as safe (false negatives) and *false alarms* misclassify safe scenarios as unsafe (false positives). Generally, smaller s^* leads to fewer false alarms and more missed alarms. In contrast, larger s^* reduces missed alarms but increases false alarms.

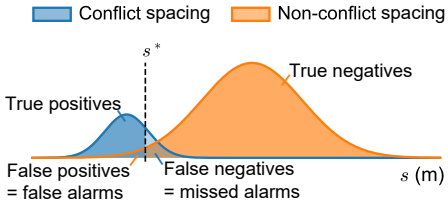


Figure 3.1 Illustration of the trade-off between missed and false alarms.

Considering the spacing between two vehicles as a random variable S, we can estimate the probability of missed alarms and false alarms when considering a spacing s as the critical threshold. As presented in Equations (3.3), PMA(s) denotes the conditional probability of missed alarms (false negatives) and PFA(s) denotes the conditional probability of false alarms (false positives). In this context of conflict indication, a positive means $S \leq s$ and a negative means $S > s$; a true event is a conflict (c) and a false event is a non-conflict (nc).

$$\begin{cases} \text{PMA}(s) = P(S > s|\text{c}) \\ \text{PFA}(s) = P(S \leq s|\text{nc}) \end{cases} \tag{3.3}$$

Given $p(A|B) = p(AB)/p(B)$ in Bayes' theorem and $p(A\bar{B}) = p(A) - p(AB)$ in set theory, we transform PMA(s) into Equation (3.4) and PFA(s) into Equation (3.5), where $s_{\max}$ is a large enough value of spacing and will be specified in Section 3.2.4.

$$\text{PMA}(s) = P(S \leq s_{\max}|\text{c}) - P(S \leq s|\text{c}) \tag{3.4}$$

$$\begin{aligned} \text{PFA}(s) &= \frac{P(S \leq s, \text{nc})}{P(S \leq s_{\max}, \text{nc})} \\ &= \frac{P(S \leq s) - P(S \leq s|\text{c})p(\text{c})}{P(S \leq s_{\max}) - P(S \leq s_{\max}|\text{c})p(\text{c})} \end{aligned} \tag{3.5}$$

Equations (3.4) and (3.5) include two cumulative probabilities, which can be estimated from data. One is of S and we denote its probability density function as $f(x)$ in Equation 3.6. The other is of S in conflict, and we denote its probability density function as $g(x)$ in Equation 3.6.

$$\begin{cases} f(x) = \dfrac{\mathrm{d}}{\mathrm{d}x} P(S \leq x) \\ g(x) = \dfrac{\mathrm{d}}{\mathrm{d}x} P(S \leq x | \mathrm{c}) \end{cases} \tag{3.6}$$

Summarising these derivations, when using s as the critical spacing to distinguish safe and unsafe scenarios, the probability of missed and false alarms are computed according to Equation (3.7).

$$\begin{cases} \mathrm{PMA}(s) = \displaystyle\int_{s}^{s_{\max}} g(x)\mathrm{d}x, \\ \mathrm{PFA}(s) = \dfrac{\int_0^s f(x)\mathrm{d}x - k\int_0^s g(x)\mathrm{d}x}{\int_0^{s^{\max}} f(x)\mathrm{d}x - k\int_0^{s_{\max}} g(x)\mathrm{d}x}, \end{cases} \tag{3.7}$$

where $k = p(\mathrm{c})$ can be counted as the conflict frequency.

3.2.3 Spacing patterns between vehicles

Computing PMA(s) and PFA(s) needs the two spacing distributions $f(x)$ and $g(x)$. In this study, we use Gaussian kernel density estimation (KDE)[2] to give a preliminary demonstration. For a certain interaction situation $\boldsymbol{\theta}$, let $\boldsymbol{s}$ be the set of s_{ij}^t, of which the corresponding $\theta_{ij}^t \in \boldsymbol{\theta}$; we then denote the subset of $\boldsymbol{s}$ in conflict by $\boldsymbol{s}_{\mathrm{c}}$. To estimate $f(x)$, we apply Gaussian KDE to the samples $s_1, s_2, \ldots, s_n$ in $\boldsymbol{s}$. Similarly, to estimate $g(x)$, we apply Gaussian KDE to $\boldsymbol{s}_{\mathrm{c}}$.

3.2.4 Minimising missed and false alarms

Based on the estimated probabilities of missed alarms and false alarms, we can optimise a critical spacing as in Equation (3.8). The parameter α is the weight on minimising missed alarms and $1-\alpha$ on false alarms. This makes the optimisation weigh between less missed alarms or less false alarms. As a result, a reliable critical spacing s^* should minimise the balanced probability of false negatives and false positives. We call this method missed and false alarm minimisation, which can be abbreviated as MFaM.

$$s^* = \underset{0 \leq s \leq s_{\max}}{\arg\min}\ \alpha \mathrm{PMA}(s) + (1-\alpha)\mathrm{PFA}(s) \tag{3.8}$$

Fig. 3.2 gives an example of applying MFaM, where PMA(s) and PFA(s) are estimated based on real spacing samples. As s increases, the probability of missed alarms decreases and the probability of false alarms increases. Then various $s^*(\theta)$ can be obtained by minimising the weighted sum of PMA(s) and PFA(s) given different α.

[2]We applied the function "gaussian_kde" from the python library "scipy" with default arguments.

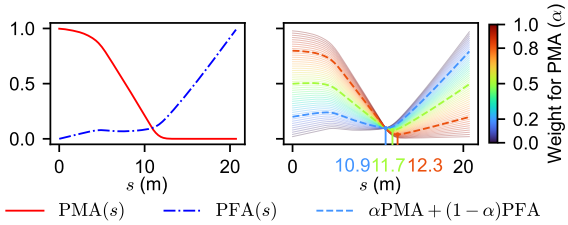


Figure 3.2 An example of minimising missed and false alarm probability.

A proper value for $s_{\max}$ is necessary in the Equations (3.7) and (3.8). In Equation (3.7), $s_{\max}$ determines the range within which PMA(s) and PFA(s) are normalised to the interval [0,1]. In Equation (3.8), $s_{\max}$ sets the searching range for s^*. Although $s_{\max}$ approaches ∞ in theory, a range is important to facilitate computation in practice. In this study, we take the maximum between two values, as shown in Equation (3.9).

$$s_{\max} = \max\{\max \boldsymbol{s}_c, \arg\max_s f(s)\} \tag{3.9}$$

The first value in Equation (3.9) is the maximum of conflict spacing $\boldsymbol{s}_c$. This ensures that all the occurred conflicts are considered. The second value is the most probable spacing in $\boldsymbol{s}$. The spacing maintained between vehicles is based on drivers' perception and preferences. Therefore, we assume that the most frequently maintained spacing is safe enough for most drivers.

3.3 Experiments

This study introduces a new approach, MFaM, aiming for more reliable conflict detection. To demonstrate this approach, we conduct experiments with both synthetic conflicts and real-world conflicts.

3.3.1 Synthetic conflicts

We used a subset called Freeway-B of the CitySim dataset [119] to generate synthetic conflicts. Freeway-B comprises trajectories collected on a 725-m segment of a 6-lane road (three lanes per direction). The movements of 6,555 vehicles in a duration of 0.57 hours were recorded at a frequency of 30 Hz. The average flow was approximately 1,917 veh/hour/lane. This is indicative of congested traffic that is more likely to yield conflicts than fee-following traffic. In total 3,082 car-following pairs were extracted from the dataset.

Existing studies in conflict detection often assume conflicts as when TTC values falling below a critical threshold, to name a few, see C. Yuan et al. [120], Hu et al. [121] and Ding et al. [122]. However, relying solely on relative speed (as assumed by TTC) to determine whether a conflict occurs is inadequate, especially when the absolute speed is slow. In this paper, we define three types of conflicts for a more comprehensive comparison to

demonstrate our approach. As outlined in Table 3.1, we let type I and type II conflicts be conditioned by relative speed only, but additionally consider the absolute speed of following vehicles for type III conflicts. For type I conflicts, we set a uniform threshold to distinguish unsafe scenarios homogeneously; while we use various thresholds for defining type II and type III conflicts heterogeneously.

Table 3.1 Conflict determination for test experiments

Conflict type	Conditions (θ)		Threshold (m)
	Relative speed (m/s)	Follower speed (m/s)	
I	$\Delta v > 0$		$s \leq 3\Delta v$
II	$\Delta v > 5$		$s \leq 2.5\Delta v$
	$2 < \Delta v \leq 5$		$s \leq 3\Delta v$
	$0 < \Delta v \leq 2$		$s \leq 3.5\Delta v$
III	$\Delta v > 5$		$s < 2.5\Delta v$
	$2 < \Delta v \leq 5$	$v > 25$	$s \leq 3.5\Delta v$
		$10 < v \leq 25$	$s \leq 3\Delta v$
		$v \leq 10$	$s \leq 2.5\Delta v$
	$0 < \Delta v \leq 2$	$v > 5$	$s \leq 0.5v$
		$2 < v \leq 5$	$s \leq 0.3v$
		$1 < v \leq 2$	$s \leq 0.6$

3.3.2 Real-world conflicts

For real-world conflict data, we reconstructed trajectories from the 100-Car Naturalistic Driving Study's time-series data [123]. The data was collected during an instrumented-vehicle study conducted in the Northern Virginia / Washington, D.C. area in early 2000s [124]. The instrumentation was designed to be unobtrusive, study participants were given no special instructions, and experimenters were not present.

From the data collection, an event database was compiled consisting of 68 crashes and 760 near-crashes which were manually reviewed and annotated. With the time-series profile for each event, containing radar and accelerometer data spanning 30s before the event and 10s after the event, we reconstructed bird's eye view trajectories for the vehicles involved in these events. Not all of the events can be reconstructed due to the missing values, inaccuracy of sensing, and the lack of a ground truth; matching the conflicting vehicle among the detected vehicles in each event is neither trivial. Eventually, we obtained 219 car-following near-crashes of which vehicle trajectories are properly reconstructed and conflicting vehicles are matched.

With the two conflict datasets, we apply MFaM under varying weights for missed and false alarms, and then compare the detection results with those obtained using TTC with a range of critical thresholds. The next section will present and discuss the results.

3.4 Results and discussion

Our experiments assume that the only known information of conflict situation $\boldsymbol{\theta}$ is relative speed. This aligns with the assumption of TTC-based detection. By doing so, different synthetic conflict types allows for comparisons under different levels of information completeness and conflict heterogeneity. The detection of type I conflicts represents the detection of homogeneous conflicts with complete information of conflict situation; the detection of type II conflicts then represents that of heterogeneous conflicts with complete information of conflict situation; and the detection of type III conflicts represents the cases where incomplete information is known and the conflicts are heterogeneous. From type I to type III, these synthetic conflicts were designed to simulate more realistic and complex conflicts. At the end, we will demonstrate the detection of real-world conflicts.

3.4.1 Detection of synthetic conflicts

Fig. 3.3 shows the detection of type I conflicts using TTC and MFaM. A total number of 21,885 type I conflict moments are defined utilising a uniform TTC* (critical threshold of TTC) of 3 seconds. This criterion makes the conflict situation $\boldsymbol{\theta}$ include solely the relative speed between a vehicle and its preceding vehicle, which is completely considered in conflict detection. With increasing values of TTC* and the weight assigned for missed alarms (α), there are fewer missed alarms and more false alarms. Remarkably, both missed alarms and false alarms reach 0% when TTC* is precisely 3 seconds. In contrast, MFaM does not have such an optimal point. When the weight for missed alarms is larger than 0.2, there are very few missed alarms and the rate of false alarms is also low.

The detection of type II conflicts, as shown in Fig. 3.4, shows similar patterns as of detecting type I conflicts. The 26,912 conflict moments are also conditioned by relative speed only, however, are defined using varied critical thresholds. For these heterogeneous conflicts, implementing MFaM leads to a very similar tendency of increasing false alarms while reducing missed alarms as observed in type I conflict detection. In contrast, when using TTC, both trends of more false alarms and fewer missed alarms with increasing TTC* are slower than in type I conflict detection. Furthermore, there is no longer an optimal TTC* where both missed alarms and false alarms reach 0%.

The challenge of reliable conflict detection is intensified with the 34,203 type III conflict moments, as displayed in Fig. 3.5. The detection is characterised by conflict heterogeneity and incomplete information on conflict situation. Regardless of the magnitude of TTC*, more than 69.36% conflict moments are missed if using TTC. Conversely, MFaM manages to detect nearly all (99.69%) conflict moments, although this is at the expense of a heightened rate of false alarms.

3.4.2 Trade-off between missed and false alarms

Observing the detection results of these three types of conflicts, there exists an inherent trade-off between missed alarms and false alarms. This trade-off is particularly significant when the conflict situation are incompletely known. Fig. 3.6 illustrates the trade-off curves

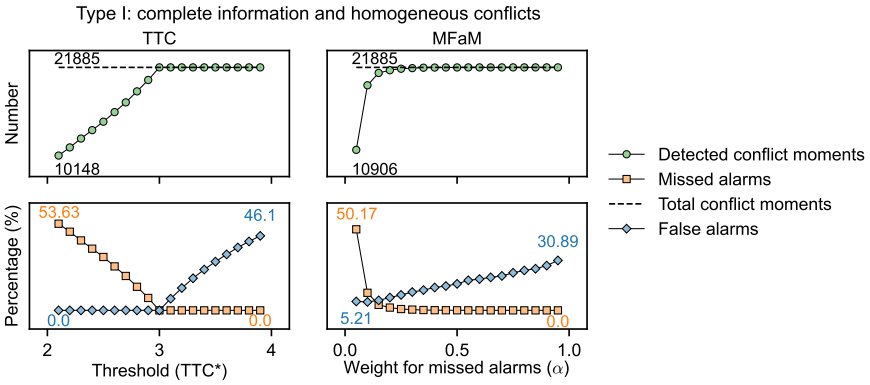


Figure 3.3 Type I conflict detection using TTC and MFaM.

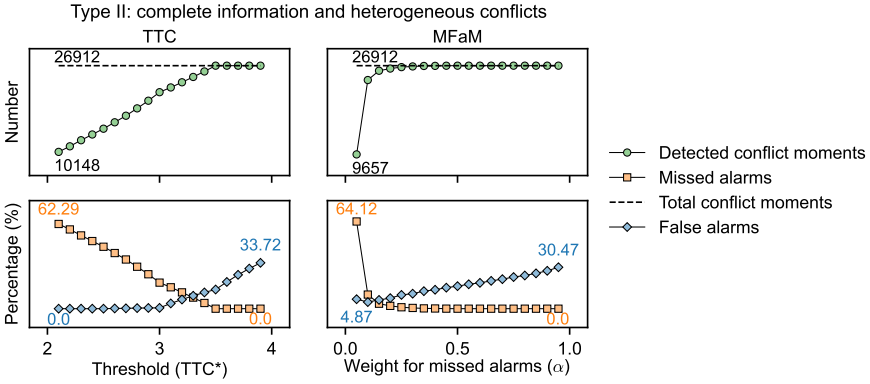


Figure 3.4 Type II conflict detection using TTC and MFaM.

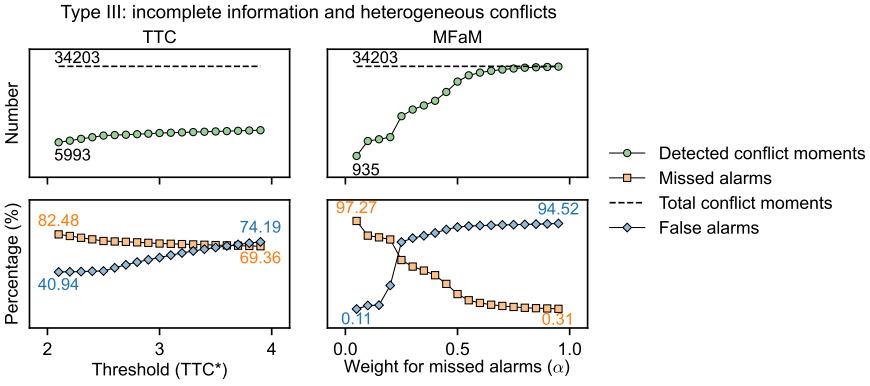


Figure 3.5 Type III conflict detection using TTC and MFaM.

for detecting the three types of synthetic conflicts, where optimal performance is indicated by the proximity to the origin $(0\%, 0\%)$. Next, we will first analyse the effectiveness of MFaM across homogeneous (type I) and heterogeneous (type II) conflicts. Then we will compare its performance when the information of conflict situation is completely known (type II) or only partially known (type III).

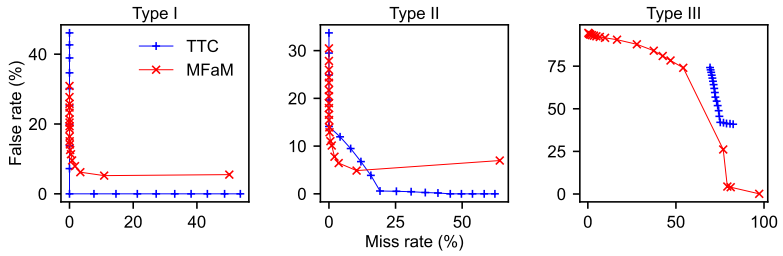


Figure 3.6 Trade-off between miss rate and false rate in conflict detection.

Homogeneous vs. heterogeneous conflicts

Both the detection of type I and type II conflicts have complete information on conflict situation available. Comparing the sub-figures of type I and type II in Fig. 3.6, it is evident that MFaM demonstrates robust effectiveness across homogeneous and heterogeneous conflicts, while TTC does not. For homogeneous conflicts (type I), TTC can ascertain the precise critical threshold, thereby achieving $(0, 0)$ rates of missed and false alarms. In contrast, heterogeneous conflicts (type II) preclude the identification of a single critical threshold applicable to TTC. Conflicts, in reality, are heterogeneous due to factors such as dynamic traffic environments and diverse human driving styles. For this reason, robust detection of conflicts requires managing the heterogeneity of conflicts.

Complete vs. incomplete information

While both type II and III conflicts are heterogeneous, the detection of type II conflicts considers complete information and the detection of type III conflicts operates with only partial information of conflict situation. In the sub-figure of detecting type II conflicts in Fig. 3.6, MFaM can reach a commendable balance of low missed and false alarms. This outperforms TTC, which has comparable performance to MFaM only when the miss rate is around 15%. In the sub-figure of type III conflict detection where the information is incompletely known, MFaM's curve consistently outperforms TTC. Nevertheless, neither TTC nor MFaM attains low rates of missed and false alarms due to information insufficiency.

3.4.3 Detection of real-world conflicts

As presented in Fig. 3.7, the detection results of the real-world conflicts in 100-Car data resembles the detection of type III synthetic conflicts. Around 41.28% conflict moments

cannot be detected if using TTC. In contrast, MFaM can detect around 98.80% of them, but still, along with a high rate of false alarms. The trade-off curves of this real-world conflict detection, at the right of Fig. 3.7, also show similar trends as the detection of type III synthetic conflicts: MFaM consistently outperforms TTC.

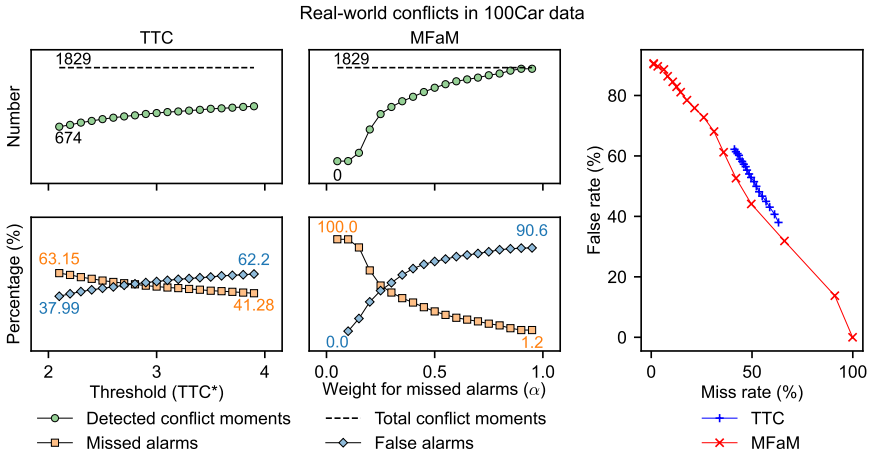


Figure 3.7 Real-world conflict detection using TTC and MFaM.

Real-world conflicts are heterogeneous and the information on conflict situation in detection is always imperfect, which, however, significantly influences the detection effectiveness. Despite the lack of information, MFaM can detect the highest possible number of actual conflicts at the expense of an increased rate of false alarms. To reduce false alarms while preserving minimised missed alarms, it is important to include multi-source information for conflict detection in future studies.

3.5 Conclusion

This study presents a new approach to more reliable conflict detection, which minimises the estimated probabilities of missed and false detection based on vehicle spacing patterns. We abbreviate this method as MFaM representing Missed and False alarm Minimisation. Through comparative experiments of applying MFaM and TTC on both synthetic and real-world conflicts, hereby we summarise the main features:

- MFaM secures a better balance between missed and false alarms compared to TTC in detecting heterogeneous conflicts, both the synthetic and real-world ones;
- MFaM surpasses TTC in accurately identifying true conflicts, especially when the information of conflict situation is incomplete;
- MFaM is flexible to be extended given various vehicle spacing patterns. For example, it can be used to develop user-adaptive collision warning given that drivers perceive different levels of collision risk and react differently to automatic warnings.

Beyond the approach itself, the importance of the information considered in conflict detection is particularly notable. If given limited information of conflict situation, we argue that there is a trade-off curve between missed and false alarms that constrain any algorithms for conflict detection. Nevertheless, this requires further exploration. Our future research will include utilising more effective information of conflict situation and developing adaptive algorithms that can account for varying response patterns of drivers. These developments will enhance the reliability of ADAS collision warning, contributing to safer and more trustworthy autonomous driving.

Chapter 4

Unified detection of potential collisions across interaction contexts

Highlights

- Conflicts are modelled as context-dependent extreme events in normal interactions.
- Unified framework enables consistent conflict detection across traffic environments.
- Statistical learning enables data-driven and comprehensive analysis of conflicts.
- The unified approach detects conflicts at least as well as any existing metrics.
- This approach supports scalable and reliable traffic safety research in the future.

Keywords

Traffic safety, autonomous driving safety, conflict detection, collision risk, Surrogate Measures of Safety

This chapter is based on the journal article: Yiru Jiao, Simeon C. Calvert, Sander van Cranenburgh, and Hans van Lint. (2025). A unified probabilistic approach to traffic conflict detection. Analytic Methods in Accident Research, 45, 100369. doi: 10.1016/j.amar.2024.100369

In the original publication, the term "Bayesian" was incorrectly used referring to conditional probability in the proposed framework. The mistake is corrected in this chapter. The corrections are made to improve theoretical precision. The validity of the methods, experimental results, and conclusions remain unaffected.

Abstract

Traffic conflict detection is essential for proactive road safety by identifying potential collisions before they occur. Existing methods rely on surrogate safety measures tailored to specific interactions (e.g., car-following, side-swiping, or path-crossing) and require varying thresholds in different traffic conditions. This variation leads to inconsistencies and limited adaptability of conflict detection in evolving traffic environments, particularly as the integration of autonomous driving systems adds complexity. Consequently, there is an increasing need for consistent detection of traffic conflicts across interaction contexts. To address this need, we propose a unified probabilistic approach in this study. The proposed approach establishes a unified framework of traffic conflict detection, where traffic conflicts are formulated as context-dependent extreme events of road user interactions. The detection of conflicts is then decomposed into a series of statistical learning tasks: representing interaction contexts, inferring proximity distributions, and assessing extreme collision risk. The unified formulation accommodates diverse hypotheses of traffic conflicts and the learning tasks enable data-driven analysis of factors such as motion states of road users, environment conditions, and participant characteristics. Jointly, this approach supports consistent and comprehensive evaluation of the collision risk emerging in road user interactions. We demonstrate the proposed approach by experiments using real-world trajectory data. A unified metric for indicating conflicts is first trained with lane-change interactions on German highways, and then compared with existing metrics using near-crash events from the U.S. 100-Car Naturalistic Driving Study. Our results show that the unified metric provides effective collision warnings, generalises across distinct datasets and traffic environments, covers a broad range of conflict types, and captures a long-tailed distribution of conflict intensity. In summary, this study provides an explainable and generalisable approach that enables traffic conflict detection across varying interaction contexts. The findings highlight its potential to enhance the safety assessment of traffic infrastructures and policies, improve collision warning systems for autonomous driving, and deepen the understanding of road user behaviour in safety-critical interactions.

Code availability

`https://github.com/Yiru-Jiao/UnifiedConflictDetection` Dynamic visualisations of 100-Car near-crashes are in ./Data/DynamicFigures/ProbabilityEstimation/gifs/; of highD lane-changing conflicts are in ./Data/DynamicFigures/IntensityEvaluation/gifs/.

Data availability

Raw data sources:

- highD `https://levelxdata.com/highd-dataset`
- 100-Car Data `https://doi.org/10.15787/VTT1/CEU6RB`

Resulting data: `https://doi.org/10.4121/06415947-2b9b-4435-833e-e513ae71a6ed`

4.1 Introduction

Collision avoidance and accident prevention are key elements in efforts to improve traffic safety [125, 126]. Motivated by proactive prevention of accidents and not waiting for collisions to happen, Surrogate Measures of Safety were proposed as surrogates of real collisions for safety evaluation and improvement. Over the past decades, traffic conflicts have become one of the most comprehensive and prominent surrogates [19, 20]. A traffic conflict is defined as "an observable situation in which two or more road users approach each other in space and time to such an extent that there is a risk of collision if their movements remain unchanged" [13]. Under this definition, every conflict is a potential collision; and every collision is a conflict until the moment when it becomes unavoidable. Despite the relative rareness when compared with safe daily interactions, successfully resolved conflicts that do not end in collisions offer opportunities to explore the emergence and resolution of collision risk [20]. This has been widely recognised in the traffic safety community, and is being adopted in emerging technologies such as driving assistance and autonomous driving.

Conflicts cannot always be directly measured due to the unclear boundary between safe and unsafe interactions, thus conflict detection often relies on surrogate metrics. These metrics are called "conflict/safety indicators" in the field of traffic safety [102] or "criticality metrics" in the field of autonomous driving [127]. This difference in terminology implies different focuses on the scale of their respective research objective. Traffic safety studies concentrate more on reasoning about the causes of collisions and reducing accidents in general; while autonomous driving studies aim to ensure safe interactions between automated vehicles and other road users [128]. As a result, conflict indicators are predominantly used for identifying contributing factors of collisions [129, 130] and evaluating the safety of transportation infrastructure, a traffic signal system, or a specific type of interaction [23, 131, 132]. In contrast, criticality metrics are more used for predicting trajectories of other road users [133], developing autonomous driving strategies [34], and assessing individual collision risk [44, 134, 135]. In this paper, we use *surrogate metrics of conflicts* to uniformly refer to both conflict/safety indicators and criticality metrics.

A variety of surrogate metrics of conflicts have been developed for different types of interactions. For instance, Deceleration Rate to Avoid Collision [DRAC, 136] and its variants are primarily targeted at rear-end conflicts. Time-To-Collision [TTC, 8, 21] along with its variants can cover both rear-end and side-swipe conflicts. Time advantage [23], also known as predicted Post-Encroachment-Time (PET, [22]), is specifically tailored for path-crossing conflicts. In addition, composite indices are designed by integrating multiple metrics to deal with more complicated conflicts such as those during lane changes [137, 138]. Many summaries [to name a few, 11, 105, 113, 139, 140] are available for an overview of these metrics, and new metrics are being actively proposed for interactions in a two-dimensional plane [18, 141] and in traffic oscillations [29, 142].

The diversity of these metrics entails inconsistency as collision risk is heterogeneous in different traffic conditions, between different vehicles, and is subject to changes in road

user behaviour [126, 143]. Some experiments show that the perception of critical TTC can vary among drivers [24] and in different traffic environments [25, 26, 140]. For example, a TTC of 3 seconds could be dangerous for vehicles rushing on highways, but not necessarily for vehicles making a cooperative lane-change, nor for vehicles decelerating to approach an urban intersection. Similarly, a 2-second PET could be accident-prone for cars crossing their paths at an intersection, but is not uncommon for cyclists [27]. In addition, with the increasing prevalence of driving assistance systems and electric vehicles, behavioural changes in interaction may gradually influence people's perception of collision risk [144, 145].

Increasing efforts are devoted to overcoming these inconsistencies. For example, there is a growing trend towards combining multiple metrics [e.g., 29–31], particularly by applying deep learning methods [32, 33]. Such a combination builds understanding from existing metrics, and thus remains constrained by their underlying assumptions of conflicts. Many studies have also looked at determining robust thresholds to distinguish unsafe conflicts from safe interactions. If there is ground truth, the threshold should optimise the accuracy of conflict identification for, e.g., issuing crash warnings [34]. When no label is available, the threshold selection often follows heuristic rules [35–37], or is guided by the extreme value theory modelling to satisfy certain hypotheses [38–40].

However, as the development of autonomous driving rapidly advances, the inconsistency in conflict detection poses new challenges for traffic safety. Automated vehicles require a unified approach to estimating collision risk as they navigate in various road environments. Without consistent estimation, these vehicles may struggle to reliably interpret and respond to traffic conflicts, thereby jeopardising road safety. As different levels of vehicle automation increasingly share the road, traffic conditions will become more complex. Road user interactions will no longer be solely between humans, but will also involve driving algorithms developed by different manufacturers, creating new layers of complexity. Consequently, in the long run, the assessment and improvement of traffic safety will become more challenging.

To address the challenges, this study introduces a new approach for consistent and comprehensive conflict detection. First, we propose a unified framework of traffic conflict detection by formulating conditional collision risk. Second, we present a series of statistical learning tasks that apply the theoretical framework to practical conflict detection. Our approach considers a traffic conflict as an extreme event of normal interactions, and quantifies its context-dependent and proximity-characterised collision risk. Then the learning tasks break down conflict detection into interaction context representation, proximity distribution inference, and extreme event assessment. Theoretically, any existing surrogate metric of conflicts is a special case under the framework. With this approach, conflict detection can be consistent across different traffic environments, and traffic safety evaluation can involve more comprehensive considerations.

The rest of this paper is organised as follows. Section 4.2 first explains the unified framework of traffic conflict detection. Then Section 4.3 presents the statistical learning tasks to apply conflict detection, including probability estimation and conflict intensity evaluation. In Section 4.4, demonstration experiments are designed, performed, and

analysed to show the performance of the approach. Finally, Section 4.5 concludes this paper and envisions future research.

4.2 A unified framework of traffic conflict detection

This section introduces our unified framework of traffic conflict detection. We frame conflict detection as quantifying the risk of a potential collision c between two or more road users, based on their proximity s and other contextual observables X, i.e., $p(c|s, X)$. This risk quantification uses conditional probabilities to consider the interaction situation of an event, the typical proximity behaviour of road users in the interaction context, and the probability variation with conflict intensity. These considerations are summarised in Equation (4.1) for a preliminary overview, where the symbols are defined in Table 4.1.

$$p(c|s, X) = \iint p(c|s, \phi)p(\phi|\theta)p(\theta|X)\mathrm{d}\phi\mathrm{d}\theta, \text{ where } p(c|s, \phi) := C(n; s, \phi) \quad (4.1)$$

Table 4.1 Symbols and their definitions in the unified framework of traffic conflict detection.

Symbol	Definition
c	A potential collision, i.e., conflict
s	Proximity, the spatial-temporal closeness between road users
X	Observables that can be measured to describe interaction situations
θ	Representation for the interaction context involving key information selected from X
ϕ	Parameters used to characterise the proximity distribution of road users in a given θ
n	Conflict intensity, a conflict at intensity n occurs once per n times in the same interactions
C	Probability of a conflict with intensity n at proximity s in the context characterised by ϕ

The components of the integral in Equation (4.1) represent a probabilistic breakdown of collision risk. Starting from the outermost component, the first term $p(\theta|X)$ denotes the probability of a specific interaction context θ given observable measurements X, such as road user motion states and environmental conditions. The second term $p(\phi|\theta)$ describes the probability of the parameters ϕ that characterise proximity patterns conditioned on the interaction context θ. Finally, the innermost term $p(c|s, \phi)$ expresses the probability of a potential collision occurring at a proximity s, based on the proximity distribution parameterised by ϕ in the specific context. This probability also accounts for conflict intensity n, which reflects the extremeness extent of the conflict within a spectrum of interactions in the same context.

The following subsections provide more detailed explanations and derivations to establish the framework. We begin with defining the probabilistic collision risk that depends on proximity and other situational variables. Then we assume a conflict hierarchy of risk perceptions and reactions in an interaction context, and describe the conflict hierarchy quantitatively by measuring how road user behaviour varies with proximity. Lastly, we consider conflicts as extreme events within a spectrum of interactions, and use extreme value theory to relate conflict probability with conflict intensity.

4.2.1 Context-dependent collision risk

In this subsection, we first present a general probabilistic description of conflict detection which depends on the context of interaction. A major branch of surrogate metrics of conflicts is based on proximity. Proximity means physically less space and time for reactions to prevent a potential collision; and due to the physical constraints, people perceive increased risk when the distance to an approaching object is closer [146, 147]. Therefore, a proximity-based conflict metric can cover both the objective aspect of collision risk and people's subjective perception of the collision risk.

Proximity-based conflict detection can be generally formulated as $p(c|s, X)$, which estimates the probability of a potential collision c based on proximity s ($s \geq 0$) and other observables X of the interaction context. The probabilistic formulation enables consistent evaluation across contexts, as probability has a normalised range between 0 and 1. This consideration has also been taken by some studies, for example, Saunier and Sayed [148] and de Gelder et al. [149]. For an interaction involving two or more road users, s is a measure of the spatial-temporal closeness between the road users; and X can include, but is not limited to, motion states of the road users, traffic states, road layouts, and weather conditions.

Not all of the information in X is used, nor might it all be useful. Every existing surrogate metric of conflicts selects some key information, i.e., variables, and makes a hypothesis based on the selected information. For example, TTC uses the relative velocity between two approaching road users and assumes no movement change (constant acceleration) at the moment of conflict; DRAC also uses the relative speed but assumes an immediate brake. Here we use $\theta = \{\theta_1, \theta_2, \ldots, \theta_k\}$ to denote k variables extracted from X, and θ are assumed to adequately represent the interaction context compressed out of the situation X. Considering all possible selections of θ, Equation (4.2) holds according to the chain rule of conditional probability.

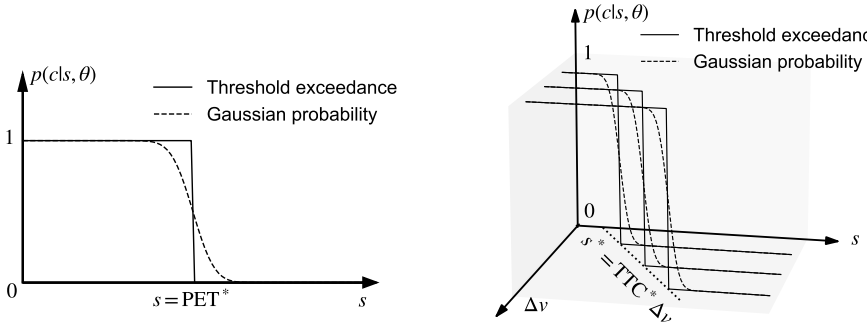
$$p(c|s, X) = \int p(c|s, \theta)p(\theta|X)\mathrm{d}\theta \tag{4.2}$$

If the selection of key information is deterministic as $\theta = \mathbb{R}(X)$, where $\mathbb{R}$ refers to representation learning of the interaction context, $p(\theta|X)$ in Equation (4.2) becomes a Dirac delta function of $\delta(\theta - \mathbb{R}(X))$. A Dirac delta function $\delta(x)$ has a value of 1 when $x = 0$ and 0 for any other values of x. Therefore, Equation (4.2) can be approximated to Equation (4.3).

$$p(c|s, X) = p(c|s, \theta) = p(c|s, \mathbb{R}(X)), \text{ if } p(\theta|X) = \delta(\theta - \mathbb{R}(X)) \tag{4.3}$$

To better explain how a surrogate metric of conflicts fits into our formulation, we provide two examples. The first is Post-Encroachment Time (PET), where s is the time interval between one vehicle leaving a conflict area and another vehicle arriving in the same conflict area. The θ for PET is the existence of a conflict area. As shown in Figure 4.1(a), $p(c|s, \theta)$ can be a step function with probability 1 when PET exceeds a threshold PET* and 0 otherwise; or a continuous function based on a cumulative

Gaussian probability that increases gradually from 0 to 1 and is 0.5 at the threshold PET*. Then another example uses TTC $= s/\Delta v$ as shown in Figure 4.1(b). For two vehicles approaching each other, s is the net distance between them; while Δv is their relative speed, and serves as a univariate context θ. The probability $p(c|s,\theta)$ varies at different relative speeds. If there is a critical threshold TTC* that differentiates safe and unsafe interactions, the threshold of proximity then follows $s^* = \text{TTC}^*\Delta v$.



(a) PET: Post-Encroachment Time. **(b)** TTC: Time-to-Collision.

Figure 4.1 Illustration examples of context-dependent proximity-based conflict probability.

4.2.2 Proximity-characterised conflict hierarchy

Proximity patterns in an interaction context reflect the aggregated behaviour of road users in this context, which is shaped by people's perception of collision risk. Within the same context of interaction, a shorter spatial or temporal gap between road users consistently implies more risk of collision. This increase in collision risk with decreasing proximity motivates road users to maintain an acceptable distance from the others at varying levels [77]. Therefore, the proximity behaviour of road users embodies a conflict hierarchy perceived by the road users. On the one hand, proximity pushes or prevents the road users from approaching each other; on the other hand, the road users adjust their behaviours to maintain comfortable proximity.

We characterise the proximity behaviour of road users with a conditional probability distribution. Seeing proximity as a random variable S, in a certain interaction context θ, the conditional probability distribution of S given a context θ is $p(s|\theta)$. We use $f_S(s|\theta;\phi)$ to describe the density function of $p(s|\theta)$, where ϕ is a set of parameters. To incorporate context-dependent proximity behaviour into conflict detection, we consider all possible sets of ϕ in a context θ, thus integrate $p(\phi|\theta)$ in Equation (4.4).

$$p(c|s,\theta) = \int p(c|s,\phi)p(\phi|\theta)\mathrm{d}\phi \tag{4.4}$$

As we parameterise the conditional probability $p(s|\theta)$ with $f_S(s|\theta;\phi)$, inferring $p(s|\theta)$ is to obtain ϕ, and we denote this inference as $\phi = \mathbb{I}(\theta)$. When the inference is deterministic,

$p(\phi|\theta)$ in Equation (4.4) becomes a Dirac delta function $\delta(\phi - \mathbb{I}(\theta))$. In the same way as explained when deriving Equation (4.3), we can derive Equation (4.5).

$$p(c|s,\theta) = p(c|s,\phi) = p(c|s,\mathbb{I}(\theta)), \text{ if } p(\phi|\theta) = \delta(\phi - \mathbb{I}(\theta)) \tag{4.5}$$

Numerous empirical studies have observed the transition from comfort to discomfort when a road user is approached by other vehicles [73–76]. Therefore, we can generally assume that $p(c|s,\phi)$ monotonically increases while s is decreasing. In a conflict hierarchy characterised by ϕ, the closer the proximity between approaching road users, the less safe they are and the higher the probability of a potential collision. This monotonicity of $p(c|s,\phi)$ implies that Equation (4.6) holds for any ϕ.

$$\begin{cases} \lim\limits_{s\to\infty} p(c|s,\phi) = 0 \\ \lim\limits_{s\to 0} p(c|s,\phi) = 1 \end{cases} \tag{4.6}$$

We use Figure 4.2 to help explain the proximity-characterised conflict hierarchy. On the left of Figure 4.2, we present an adapted pyramid of conflict hierarchy [8], which conceptually relates conflict intensity and frequency, and reflects road user behaviour in balancing safety and efficiency [23]. When the approaching between road users entails a potential collision, the closer they are, the higher the conflict intensity and the lower the frequency of such conflicts. When the proximity is too small for safe interaction, a conflict emerges entailing a potential collision. Furthermore, an accident will happen if there is no (successful) evasion or the proximity is too small to prevent a collision. This hierarchy of conflicts is not fixed, but has varying shapes in different interaction contexts. As illustrated in the right of Figure 4.2, $f_S(s|\theta;\phi)$ is designed to capture such context-dependent proximity patterns. Consequently, the estimated probability $p(c|s,\phi = \mathbb{I}(\theta))$ may vary for the same proximity in different contexts, and evolve as the interaction context changes over time.

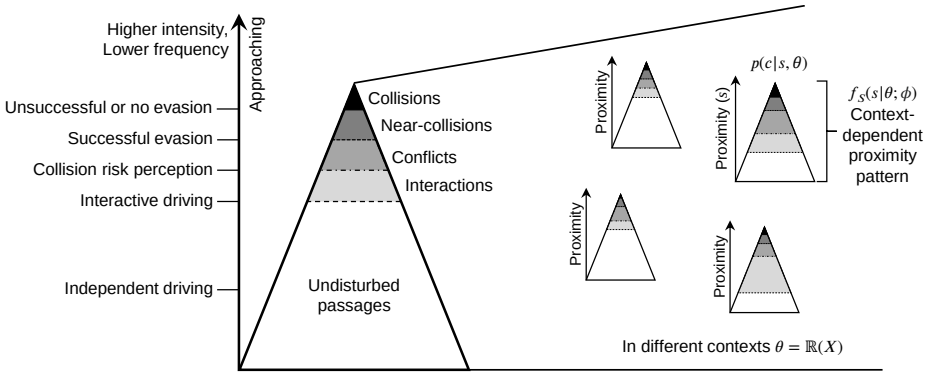


Figure 4.2 Proximity-characterised conflict hierarchy varies in different interaction contexts.

4.2.3 Extreme value theory-based interaction spectrum

Traffic interactions exist on a spectrum varying in conflict intensity and potential consequences. Within this spectrum, we assume that conflicts are extreme events at different levels of collision risk and collisions are the most extreme cases. Then we can establish the relation between conflict intensity and conflict probability utilising extreme value theory. In traffic safety research, the extreme value modelling has been traditionally used to treat crashes as the extreme events of conflicts and extrapolate average crash risk from traffic conflicts [38, 40, 150–153]. Here we zoom out the view and provide a derivation that collisions and conflicts are both extreme events of ordinary interactions, at varying levels of intensity.

Our derivation uses a similar logic to lifetime survival analysis, which is a branch of statistics to model the expected duration of time until one event, such as deaths or failures of medical treatment, occurs [154]. The essential idea of lifetime survival analysis is to estimate the probability of events from the records of human lives or running systems, where the frequency of these events reduces when living time increases. Therefore, the survival function is $S(t) = \Pr(T > t)$, representing the cumulative probability that the duration of survival T is longer than some specified time t. The longer the t, the closer to death or failure.

For estimating the probability of conflicts, collisions are system failures, and our evaluation is based on the records of daily interactions. Contrary to lifetime survival analysis which measures survival duration and *longer* duration is closer to deaths, here conflict analysis measures proximity and *shorter* proximity is closer to collisions. The frequency of conflicts increases when proximity reduces. We thus define a "conflict function" $F(s) = \Pr(S < s)$, which represents the cumulative probability that the interaction proximity S is less than a certain s. Recall that in Section 4.2.2 we use $f_S(s|\theta;\phi)$ to describe the proximity distribution of S in an interaction context θ. Rewriting the function as $f_S(s;\phi)$ for convenience, given that $\phi = \mathbb{I}(\theta)$ and $\theta = \mathbb{I}^{-1}(\phi)$, then Equation (4.7) shows the conflict function in an interaction context θ where the conflict hierarchy is characterised by ϕ.

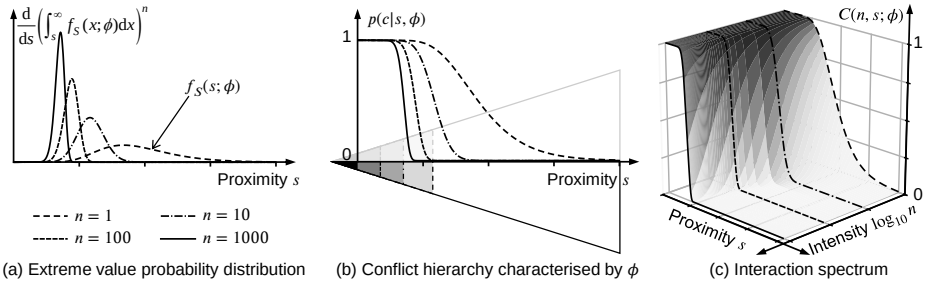
$$F(s;\phi) = \Pr(S < s;\phi) = \int_0^s f_S(x;\phi)\mathrm{d}x \tag{4.7}$$

The shorter the s, the more likely a collision is to occur, but the likelihood varies with conflict intensity. Aligned with the conflict hierarchy in Figure 4.2, we define conflict intensity as the inverse of conflict frequency, i.e., a conflict at intensity n occurs once per n times in the same interaction context. Specifically, assuming an interaction where the proximity between road users is s; if s is always smaller than the observed proximities for n times interactions in the same situation, this interaction is a conflict with a frequency of $1/n$, and we consider it an extreme event of intensity n. According to extreme value theory, we can use $(\Pr(S \geq s;\phi))^n$ to calculate the cumulative probability that a proximity s is the minima in n times of observations in the same interaction context. Therefore, for an interaction at proximity s and in the context θ that is characterised by ϕ,

its probability of conflict is not a value, but a function regarding conflict intensity n. We denote this function by C as shown in Equation (4.8).

$$p(c|s,\phi) := C(n;s,\phi) = (1 - F(s;\phi))^n = \left(\int_s^{\infty} f_S(x;\phi)\mathrm{d}x\right)^n \tag{4.8}$$

Figure 4.3(a) shows the probability distributions of Equation (4.8) under different intensity n. As displayed in Figure 4.3(b), this probability fulfils the requirements in Equations (4.6), monotonically increasing while s decreases. In Figure 4.3(c), we present an extreme value theory-based interaction spectrum described by Equation (4.8). It corresponds to the practical understanding of a conflict: the smaller the proximity and the lower the conflict intensity, the greater the conflict probability. Note that conflict probability does not equate to collision probability in our derivation. Collision probability specifically refers to the case when n is very large and approaches infinity, representing the most intense and least frequent interaction. In that case, Equation (4.8) converges to either Gumbel, Frechet, or Weibull distribution. Here we do not particularly consider this convergence at an infinite n, as traffic and driving safety focuses not just on collisions, but on conflicts at varying levels of intensity.



(a) Extreme value probability distribution (b) Conflict hierarchy characterised by ϕ (c) Interaction spectrum

Figure 4.3 Interaction spectrum described with context-dependent proximity distribution and extreme value theory.

4.3 Statistical learning tasks for conflict detection

The previous section explains the unified theoretical foundation of conflict detection. Now we continue by framing a series of statistical learning tasks for application in practice. Based on the derivation in Section 4.2, conflict detection estimates the probability of context-dependent and proximity-characterised extreme events, and we can decompose it into three tasks as shown in Figure 4.4. The first task $\theta = \mathbb{R}(X)$ maps situational observables X to a compressed information space as θ, which represents interaction context. In a certain interaction context θ, the second task $\phi = \mathbb{I}(\theta)$ infers the conditional probability distribution of typical proximity behaviour of road users. Then the third task uses extreme value theory to determine the relations between proximity, conflict intensity,

and conflict probability. When $\theta = \mathbb{R}(X)$ and $\phi = \mathbb{I}(\theta)$ are deterministic, we can rewrite Equation (4.1) into Equation (4.9).

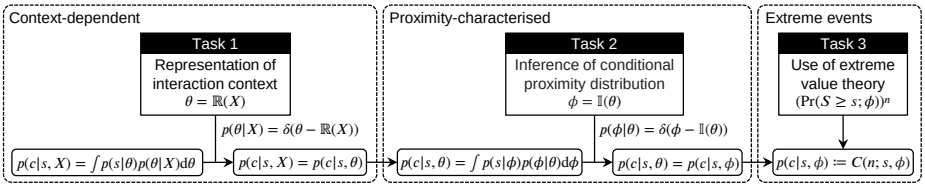


Figure 4.4 The unified framework and statistical learning tasks for conflict detection.

$$p(c|s,X) = \iint p(c|s,\phi)p(\phi|\theta)p(\theta|X)\mathrm{d}\phi\mathrm{d}\theta = p(c|s,\phi) := C(n;s,\phi), \quad \text{if } p(\theta|X) = \delta(\theta - \mathbb{R}(X)) \text{ and } p(\phi|\theta) = \delta(\phi - \mathbb{I}(\theta)) \tag{4.9}$$

These three tasks can be integrated within a pipeline for end-to-end learning, and can also be performed separately. In the following subsections, we will explain the tasks further and present preliminary methods for application. We emphasise here that there are many useful learning methods, and the ones we use in this study are not necessarily the best. Future investigation on optimising suitable statistical learning techniques is required.

4.3.1 Representation of interaction context

The first task $\theta = \mathbb{R}(X)$ is to create an informative representation for interaction contexts by selecting and transforming relevant observables of interaction situations. Traditionally, key variables such as absolute speed, relative speed, and deceleration rate have been widely validated and used in existing surrogate measures of safety. However, there may be more factors at play than those hypothesised in the existing measures. As mentioned in Section 4.2.1, X may cover various aspects such as vehicle motion states, environment factors, and participant characteristics. Incorporating different factors can be particularly helpful for a comprehensive assessment of collision risk and for user-customised collision warnings.

If the observables of an interaction situation are numerous or use time-series and/or image formats, data-driven representation learning can be effectively employed. Such techniques can compress selected observables into a lower-dimensional space, where different interaction contexts are adequately encapsulated. One approach can be auto-encoders, of which the learning objective is to minimise the reconstruction error between compressed representation and original information. Another approach is contrastive learning, which aims to minimise the difference between the representations of similar samples while maximising the difference between dissimilar samples. Representation learning is an active and evolving field in machine learning, where many other methods are worth investigating. It is important to note that purely data-driven methods may not always capture the nuances of complex interaction situations. Therefore,

integrating domain knowledge in traffic safety is necessary and can improve the robustness and reliability of context representation.

To efficiently present the theoretical contributions of this research, we do not perform data-driven representation learning in this study. However, we have some notes for readers on learning representations from variables relevant to motion states. Firstly, the output in the next task of inference is about the proximity between interacting road users. To avoid label leakage, we recommend not including a complete series of position or speed vectors for all involved road users. Secondly, we suggest transforming the variables into local coordinate systems. This transformation can be centred on every road user involved in an interaction situation, as the same situation may be perceived differently by different road users. Such view transformation can ensure the consistency and enhance the comparability of interaction context representations. Once transformed, an interaction context can be represented as instantaneous frames or over a continuous period.

4.3.2 Inference of conflict hierarchy

The second task $\phi = \mathbb{I}(\theta)$ needs to learn the conditional probability distribution of proximity in a certain interaction context, i.e., $p(s|\theta)$. Our previous study [155] used conditional sampling to infer this probability. Such sampling requires a large amount of data to ensure enough samples for each condition range, while the condition ranges are finite samples of an infinite space. In this study, we use Gaussian Process Regression (GPR) to avoid discretising the condition space, and the rest of this subsection will explain more details. Note that there are other useful methods to learn the conditional probability, but we choose GPR in order to obtain the equation of $p(s|\theta)$ and thus its analytical cumulative probability.

Many studies have found that lognormal distribution best fits the distribution of spatial and temporal gaps between road users [156–158]. Therefore, we assume that the proximity s in a certain interaction context θ follows a lognormal distribution. Recall that in Section 4.2.2 we use $f_S(s|\theta;\phi)$ to describe $p(s|\theta)$, and now we have $\phi = \{\mu, \sigma\}$ where μ and σ parameterise a lognormal distribution as shown in Equation (4.10). This implies that the logarithm of proximity, $\ln(s)$, follows a Gaussian distribution that is parameterised by the same μ and σ.

$$f_S(s|\theta;\mu,\sigma) = \frac{1}{s\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln(s)-\mu)^2}{2\sigma^2}\right) \tag{4.10}$$

The Gaussian distribution of $\ln(s)$ can be utilised to learn $p(s|\theta)$ with GPR. First, we consider a mapping $g : \theta \to \ln(s)$ between proximity s and the interaction context θ that s is in. Assuming a normally distributed noise in this mapping, we denote that $\ln(s) = g(\theta) + \epsilon$, where $\epsilon \sim \mathcal{N}(0, \sigma_\epsilon^2)$. Second, given that $\ln(s)$ is derived to be normally distributed, the distribution of function $g(\theta)$ is Gaussian. Then we can consider that $g(\theta)$ is drawn from a Gaussian Process, as shown in Equation (4.11). This suggests that a $g(\theta)$ is one sample from the multivariate Gaussian distribution of all possible mappings.

$$g(\theta) \sim \mathcal{GP}\left(m(\theta), K(\theta, \theta')\right), \tag{4.11}$$

where $\mu = m(\theta)$ and $\sigma = K(\theta, \theta')$ are the mean function and covariance function that specify a Gaussian Process. For two similar interaction contexts θ and θ', $g(\theta)$ is expected to be close to $g(\theta')$.

With GPR, the mean and covariance functions are learned from data. Considering that we may include many variables in the representation of interaction context and thus a fairly large amount of data is needed to serve training, we use the Scalable Variational Gaussian Process (SVGP) model with the python library GPyTorch [159]. As the cost to increase scalability, SVGP does not ensure an exact solution and may underestimate variance. To train this model effectively, we maximise a predictive log-likelihood in Equation (4.12), which is proposed in [160] and we adapt it here. In addition to the symbols that we have consistently used, N is the number of samples; u denotes inducing variables that are introduced to perform sparse training and reduce computational load; $q(u)$ is the variational distribution of u; $p(u)$ is the prior distribution of u; $D_{\mathrm{KL}}\left[q(u)||p(u)\right]$ computes the Kullback–Leibler divergence of $q(u)$ and $p(u)$; and β controls the regularisation effect of KL divergence.

$$\begin{aligned} L &= \mathbb{E}_{p_{\text{data}}(\ln(s),\theta)}\left[\ln\left(p(\ln(s)|\theta)\right)\right] - \beta D_{\mathrm{KL}}\left[q(u)||p(u)\right] \\ &\approx \sum_{i=1}^{N} \ln\left\{\mathbb{E}_{q(u)}\left[\int p(\ln(s_i)|g_i)p(g_i|u,\theta_i)\mathrm{d}g_i\right]\right\} - \beta D_{\mathrm{KL}}\left[q(u)||p(u)\right] \end{aligned} \tag{4.12}$$

Theoretically, learning the mean and covariance functions can approximate any form of mapping g between interaction context θ and proximity s. As explained in Sections 4.2.1 and 4.2.2, a surrogate metric for traffic conflicts essentially assumes such a mapping, based on which a threshold is then determined to distinguish safe interactions and conflicts. Therefore, in theory, any metric based on spatial-temporal proximity is a particular case under our unified framework. Without assuming a specific format of g, the mapping embedded in data can be statistically derived. This suggests that, given the same interaction context, a data-driven metric using the unified probabilistic approach proposed in this study should be no less conflict-indicative than a pre-assumed metric.

4.3.3 Conflict probability estimation and intensity evaluation

In the third task, we use extreme value theory to estimate conflict probability and evaluate conflict intensity. This task plays a similar role to the selection of thresholds when using traditional metrics of conflicts. After the previous two tasks, we can learn parameters $\phi = \{\mu, \sigma\}$ that characterise the conflict hierarchy in different interaction contexts represented with θ. Based on the derivation in Section 4.2.3, then we can write conflict function $F(s;\phi)$ and obtain $C(n;s,\phi)$ that relates conflict intensity n and conflict probability.

Given that we infer $f_S(s|\theta;\phi)$ as the probability density function of the lognormal distribution, the conflict function $F(s;\phi)$ in Equation (4.7) can be further derived as

Equation (4.13). Here $\mathrm{erf}(z) = 2\int_0^z e^{-x^2}\mathrm{d}x/\sqrt{\pi}$ and is the Gaussian error function. Then conflict probability function is as shown in Equation (4.14).

$$\begin{aligned} F(s;\mu,\sigma) &= \int_0^s \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln(x)-\mu)^2}{2\sigma^2}\right)\mathrm{d}x \\ &= \frac{1}{2} + \frac{1}{2}\,\mathrm{erf}\left(\frac{\ln(s)-\mu}{\sigma\sqrt{2}}\right) \end{aligned} \tag{4.13}$$

$$C(n;s,\mu,\sigma) = (1-F(s;\mu,\sigma))^n = \left(\frac{1}{2} - \frac{1}{2}\,\mathrm{erf}\left(\frac{\ln(s)-\mu}{\sigma\sqrt{2}}\right)\right)^n \tag{4.14}$$

4

There are two general purposes for conflict detection. The first is conflict probability estimation. For an interaction, given current proximity s and parameters ϕ that characterise the conflict hierarchy in the interaction context, we can use Equation (4.15a) to estimate the probability of a potential collision at different levels of conflict intensity n. The second purpose is safety/conflict evaluation. Assuming that the probability of a potential collision is larger than the probability of no collision, i.e., conflict probability is larger than 0.5, we can use Equation (4.15b) to evaluate the maximum possible conflict intensity.

$$\hat{p} = C(n;s,\mu,\sigma) \quad = \left(\frac{1}{2} - \frac{1}{2}\,\mathrm{erf}\left(\frac{\ln(s)-\mu}{\sigma\sqrt{2}}\right)\right)^n, \qquad n \geq 1 \tag{4.15a}$$

$$\hat{n} = C^{-1}(p;s,\mu,\sigma) = \frac{\ln p}{\ln\left(\frac{1}{2} - \frac{1}{2}\,\mathrm{erf}\left(\frac{\ln(s)-\mu}{\sigma\sqrt{2}}\right)\right)}, \qquad 0.5 < p < 1 \tag{4.15b}$$

Equations (4.15a) and (4.15b) are useful in different practices. Conflict probability estimation can be used to issue collision warnings, which alert human drivers or driving assistance systems to prevent potential collisions. Conflict intensity evaluation can be used to identify varying levels of conflict cases in daily traffic. This helps to assess the impact of an infrastructure or traffic policies on traffic safety, and then make according improvements.

4.4 Demonstration

This section applies the proposed unified framework and statistical learning tasks on real-world trajectory data, to demonstrate the characteristics of this new conflict detection approach. Table 4.2 presents an overview of our experiment design. In Section 4.4.1, we introduce the datasets used and experiment details. Then respectively in Sections 4.4.2 and 4.4.3, we show the experiment results in conflict probability estimation and conflict intensity evaluation. For convenience, we use *approach* to collectively refer to the unified probabilistic approach proposed in this study; and use the term *unified metric* and its abbreviation *Unified* to refer to the surrogate metric trained with our approach. It is also important to emphasise that these experiments are designed primarily for demonstration. Further exploration is expected in future research.

Table 4.2 An overview of experiment design for performance demonstration.

Purpose	**Training data**	**Application data**	**Characteristics:** this approach
Conflict probability estimation $\hat{p} = C(n; s, \mu, \sigma)$	Trajectories involving lane-changes in highD	Near-crashes in 100-Car NDS	1) is no less conflict-indicative than any of PSD, DRAC, and TTC; 2) is generalisable across datasets.
Conflict intensity evaluation $\hat{n} = C^{-1}(p; s, \mu, \sigma)$	Trajectories involving lane-changes in highD	Lane-changes in highD	1) covers more diverse conflicts during lane-changes than TTC; 2) detects conflicts in a long-tailed distribution of intensity.

4.4.1 Data and experiment details

We use two naturalistic trajectory datasets in this study. First is the highD dataset that was collected at 6 different locations on German highways using drones [161]. It includes detailed information about vehicle types, sizes, and movements, with a positioning error typically less than 10 cm. Considering that a significant part of driving on highways is independent without interactions with other vehicles, we select trajectories that involve lane-changes in highD. In addition to the high-quality drone-collected data for model training, a dataset of real-world conflicts is necessary for demonstration. The second dataset we use is from the 100-Car Naturalistic Driving Study (NDS), which is an instrumented-vehicle study conducted in the U.S. over 2 years in the early 2000s [162]. An event database [123] resulted from the study compiles information on 68 crashes and 760 near-crashes. From the time-series sensor data of radars and accelerometers, we reconstruct bird's eye view trajectories for these events. Due to missing values, inaccuracy of sensing, and the lack of ground truth, not all of the events can be reconstructed. We end up matching 180 events based on the constraint of insufficient space (distance less than 4.5 m) for undetected vehicles, including 11 crashes and 169 near-crashes as summarised in Table 4.3.

Table 4.3 Summary of matched and selected events in 100-Car NDS data.

Event happened with	**Matched crashes**	**Matched near-crashes**	**Selected near-crashes**
leading vehicle	5	119	47
following vehicle	6	30	17
vehicle in adjacent lane	0	13	2
vehicle turning across oncoming traffic	0	4	0
vehicle crossing through an intersection	0	2	0
vehicle in oncoming traffic	0	1	0
In total	11	169	66

The first experiment estimates conflict probability at each time moment for the events in 100-Car NDS data. We utilise the estimation to issue collision warnings, and compare warning effectiveness with using 3 other commonly used metrics. Effective collision warning maximises true positives while minimising false positives. To reserve safe interactions within an event, we select events with a duration of at least 6 seconds, no hard braking (acceleration larger than -1.5 m/s^2) in the first 3 seconds, and speeds larger than 3 m/s at the first time moment for both vehicles involved. After the selection, 4 crashes and 66 near-crashes remained. Considering the relatively low reliability of data in crashes and for a consistent comparison, we use only the selected near-crashes, which are also shown in Table 4.3. The 3 broadly used metrics for comparison are Proportion of Stopping Distance (PSD), Deceleration Rate to Avoid a Crash (DRAC), and Time-to-Collision (TTC). Table 4.4 provides an overview of them, where Δs is the distance between vehicle bounding boxes and Δv is relative velocity. For PSD we use a braking rate of 5.5 m/s^2, as emergency braking is typically between 1.6 and 5.5 m/s^2 [163, 164].

Table 4.4 Existing surrogate metrics of conflicts that are used for warning comparison.

Metric	Calculation	Note	Reference(s)	Proximity s	Context θ
PSD	$\frac{\Delta s}{v_{\text{follower}}^2/2/\text{dec}}$	dec =5.5 m/s^2	[22]	Δs	v_{follower}^2
DRAC	$\frac{\|\Delta v\|^2}{2\Delta s}$	approaching only	[136]	Δs	$\|\Delta v\|^2$
TTC	$\frac{\Delta s}{\|\Delta v\|}$	approaching only	[8, 21]	Δs	$\|\Delta v\|$

The second experiment evaluates conflict intensity at each time moment for lane-change interactions in the highD dataset. We use lane-changes to demonstrate the applicability of our unified metric to two-dimensional conflicts, which remain challenging to be integratively indicated by traditional metrics. We identify lane changes using the lane references provided in highD data. Then we determine the start and end moments of a lane-change based on the time when the vehicle deviates 1/3 of its vehicle width from the centerline of its current lane. In this way, overtaking is considered as two consecutive lane-changes. Seeing the vehicle making a lane-change as an ego vehicle, the lane-change is potentially interactive if there is a vehicle in front or at rear of the ego vehicle in either the original lane or the target lane. Then we use both the unified metric and two-dimensional TTC (2D-TTC) to evaluate conflict intensity (if any) at each time moment. Here 2D-TTC follows the typical definition of TTC (Fig. 2 in [20]), assuming constant velocities for two approaching vehicles at the moment of evaluation. The smaller the TTC value, the higher the conflict intensity; when TTC is infinite, no potential collision is expected to occur.

In both experiments, we use a single unified metric trained with the trajectories involving lane-changes in highD data. In this way, the first experiment can additionally demonstrate the generalisability of our approach; and the metric thresholds calibrated with

real-world near-crashes can be used in the second experiment to distinguish conflicts. For a fair comparison with existing metrics, we avoid complex representations for interaction context θ. Firstly, we include the information used in PSD, DRAC, and TTC, i.e., v^2_{follower}, $\|\Delta v\|^2$, and $\|\Delta v\|$ as analysed in Table 4.4. Since the lead-follow relationship evolves during lane-changing, we use the squared speeds of both ego vehicles (which change lanes) and target vehicles (in the surrounding). Further, we consider the accelerations of ego vehicles [1], heading directions of target vehicles relative to their ego vehicles, and lengths of both ego and target vehicles. In order to indicate two-dimensional conflicts, we define proximity s utilising the two-dimensional spacing proposed in [155]. This spacing is denoted by (x, y) in the real-time transformed relative coordinate systems, and we convert (x, y) into (ρ, s) into polar coordinate systems, where ρ is the angle between (x, y) and $(1, 0)$ within a range of $[-\pi, \pi]$. Then we add ρ as one more variable in θ, and use s as the proximity measure to reflect conflict hierarchy.

We train two sets of SVGP models under different settings of the hyper-parameter β in Equation (4.12), with $\beta = 5$ and $\beta = 10$. Our training is on 60% of the selected interaction trajectories involving lane-changes; and the validation set and test set account for 20% each. During the training, we reduce the learning rate dynamically to avoid overfitting of the models, and stop training early when validation loss converges. Figure 4.5 shows the training progress, where there are 271 batches per epoch and 2,048 samples per batch. Training in both settings converges well, and stops earlier as well as reaches lower loss values when $\beta = 5$. We then evaluate model performance by both loss and negative log-likelihood (NLL) on validation and test sets, as presented in Figure 4.5. As a result, we select the model achieving the minimum test loss and the second minimum test NLL, i.e., the model with $\beta = 5$ after 52 epochs of training, to apply in the subsequent demonstration experiments.

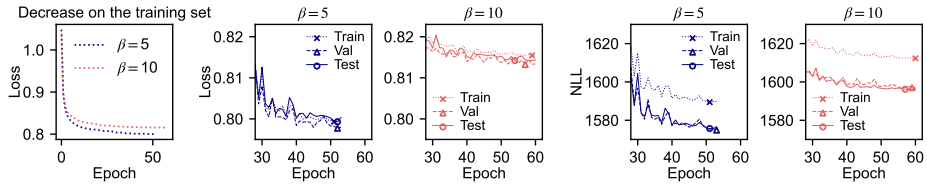


Figure 4.5 Evaluation of SVGP training and selection of the model to apply.

Since we consider two-dimensional interactions, maintaining a consistent coordinate system is necessary. The highD dataset uses a coordinate system where the x-axis points from left to right and the y-axis points downwards. This is a mirrored system from the traditional engineering coordinate system where the y-axis points upwards. When applying a model that is trained on highD data to other datasets, such as the 100-Car dataset, it is essential to adjust for this inconsistency. We make adjustments by exchanging the x

[1] We do not include target vehicle accelerations because this information is lacking in 100-Car NDS data and cannot be reliably derived due to speed fluctuations prior to potential collisions. This lack also prevents comparison with the metric Modified Time to Collision (MTTC).

and y coordinates of positions, velocities, and heading directions in 100-Car NDS data before coordinate transformation. We hereby remind the readers to correct potential inconsistencies in the coordinate systems between training data and application data.

4.4.2 Collision warning compared with existing metrics

The first experiment based on conflict probability estimation is designed to demonstrate two characteristics of our approach. First, the unified metric performs at least as well as any among PSD, DRAC, and TTC in collision warning. This is determined in theory, but experimental results with real-world near-crashes can provide additional evidence. Second, we trained the unified metric with trajectories in the highD dataset while applying it to the events in 100-Car NDS data. This cross-validation involves not only the differences between German and American driving, but also the evolution of driving over more than 15 years. If the highD-trained metric performs well for 100-Car near-crashes, we can argue for the generalisability of this approach, which implies common principles in human road use interactions.

Table 4.5 Definitions of the performance indicators for collision warning effectiveness comparison.

Performance indicator	Definition
True positive	Seeing the moment when conflicting vehicles reach the minimum distance as a critical moment, the 3 seconds prior to the moment are supposed to be dangerous; any warning in this period marks a true positive.
False positive	The first 3 seconds in each of the selected events are supposed to be safe, so any warning in this period marks a false positive.
True positive rate and false positive rate (%)	The rate of true positives and false positives, respectively, among all events. True positive rate ideally approximates 100% and false positive rate 0%.
ROC curve	Receiver Operating Characteristic curve that plots true positive rate against false positive rate at various thresholds.
Warning period (%)	For each event with warnings issued under the optimal threshold, the percentage of warned time moments within the annotated conflict period.
Warning timeliness (s)	For each event with warnings issued under the optimal threshold, the time interval from the last safe-unsafe shift of warning until the critical moment.

In order to systematically compare warning effectiveness, we define performance indicators as explained in Table 4.5. Collision warning classifies safe and unsafe interactions based on the values and specific thresholds of metrics. Seeing the metrics as

different classification models, we can compare their ROC curves, and a metric is better if the area under its curve is larger. For each metric, we then select an optimal threshold such that the corresponding point on their ROC curve is closest to the ideal point with zero false positive and all true positives (0%, 100%). Based on the optimal threshold, we further calculate warning periods and warning timeliness. The warning period of an event should ideally be close to 100%, but good warning timeliness does not necessarily mean early warning. On average, people need 1 to 1.3 seconds in response to an obstacle by braking, and in emergencies this can be less than 1 second [165, 166]. Therefore, good warning timeliness should not be too large, as it may distract people earlier than they need to be warned; neither should it be too small, as the best timing to prevent a potential collision may be missed.

Figure 4.6 presents the comparison of collision warning effectiveness. In the plot of ROC curves, we use circles centred at the ideal point (0%, 100%) and crossing the optimal points of different metrics to facilitate comparing their closeness to the ideal point. The metrics of DRAC, TTC, and Unified have comparable areas under their ROC curves, while PSD is less effective. Zooming in to see more details at the top left corner, TTC and Unified are consistently better than DRAC, and are very close to the ideal point. More specifically, the optimal threshold of Unified is n^* =17, reaching a true positive rate of 95.45% and a false positive rate of 4.55%; TTC* =4.2 s and achieves 93.94% true positive rate and 0.00% false positive rate; DRAC* =0.45, with a true positive rate of 92.42% and a false positive rate of 7.58%. Much less comparably, the optimal threshold of PSD is 0.52, with a true positive rate of 57.58% only but a false positive rate of 25.76%.

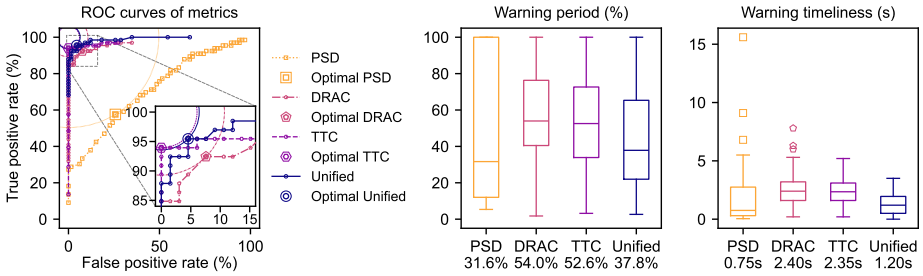


Figure 4.6 Collision warning effectiveness comparison between PSD, DRAC, TTC, and Unified. In the boxplots of warning period and timeliness, median values are marked below the labels of metrics.

These metrics are further compared at their optimal thresholds by the box plots of warning period and warning timeliness in Figure 4.6. Given the weak effectiveness of PSD, its warning period and timeliness are not comparable with other metrics. Both DRAC and TTC have higher median warning periods than Unified. This suggests that the unified metric may not issue a warning at the beginning of the annotated conflicts. This is also verified in the plot of warning timeliness, where DRAC and TTC have similar distributions and their median timeliness is both earlier than that of Unified. In addition, the warning timeliness of Unified is less varied. From a positive perspective, this means that using the unified metric gives consistent warnings and does not distract drivers too much; whereas

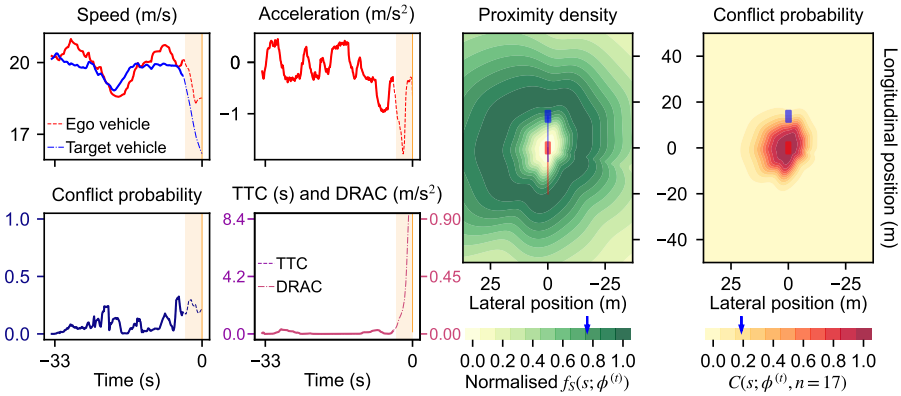
from a negative perspective, this means that the unified metric may miss the chance to prevent a potential collision.

For a more detailed analysis, Figure 4.7 shows two frames in the trip indexed by 8332. This event is warned by DRAC only, while TTC is larger than 10.59 s throughout and the conflict probability estimated by Unified remains lower than 0.5. In this figure, we present the profiles of speed, acceleration, and metric values, with the annotated conflict period shaded. We also plot heatmaps of proximity density distribution $f(s;\phi)$ and conflict probability function $C(s;\phi,n)$, where the y-axis points to the heading direction of ego vehicle and the x-axis points from right to left to align with highD's coordinate system. For each location around the ego vehicle, we assume the target vehicle is at that location and use actual interaction context θ for estimation. Figure 4.7(a) is the frame at 4.4 seconds before the critical moment; and Figure 4.7(b) is 1.4 seconds before. Notably, in Figure 4.7(b), the proximity density distribution, i.e., probable positions of the target vehicle given the interaction context, does not point to the ego vehicle's heading direction. This is due to the lateral interaction between these two vehicles, which is in line with the narrative data of this event: "the target vehicle was merging into the right lane ahead of the ego vehicle, causing the ego vehicle to brake to avoid a collision". Comparing the two frames, the unified metric takes into account the ego vehicle's deceleration as successful prevention of an immediate conflict, and thus the estimated conflict probability stays low during the annotated conflict period.

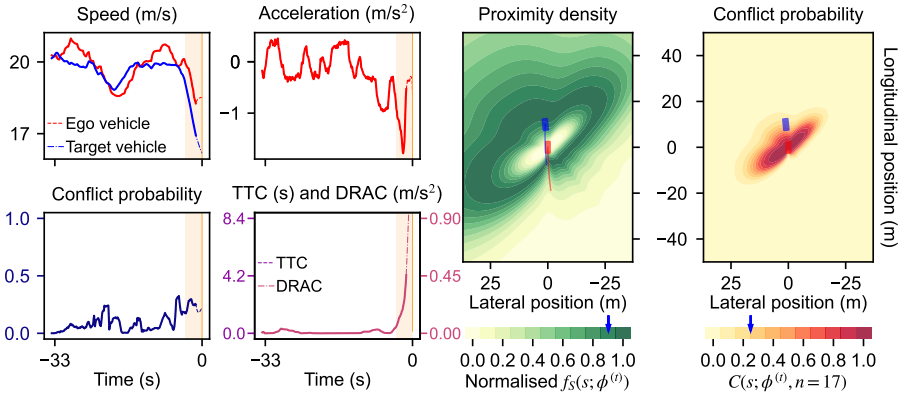
We summarise the success and failure cases of conflict warnings in Table 4.6. PSD is not included due to its lower effectiveness. For the events where at least one of Unified, TTC, and DRAC fails, as well as for the false warnings of Unified, we provide dynamic visualisations along with our open-sourced code. A link can be found in this chapter's Abstract page. Impressively, 58 out of the 61 near-crashes are correctly warned by all the metrics of Unified, TTC, and DRAC. This is because most events in the 100-Car dataset are rear-end conflicts during car following, for which TTC and DRAC have been proven to be useful. In the next experiment, TTC and our unified metric are challenged to handle two-dimensional traffic conflicts.

Table 4.6 Summary of success and failure cases of collision warnings by Unified, TTC, and DRAC.

Unified	TTC	DRAC	Number of events	Trip ID
Succeed	Succeed	Succeed	58	
Succeed	Succeed	Fail	3	8622, 8854, 9101
Succeed	Fail	Succeed	0	
Succeed	Fail	Fail	2	8463, 8810
Fail	Succeed	Succeed	1	8761
Fail	Succeed	Fail	0	
Fail	Fail	Succeed	2	8332, 8702
Fail	Fail	Fail	0	



(a) Frame 4.4 seconds prior to the critical moment of trip 8332.



(b) Frame 1.4 seconds prior to the critical moment of trip 8332.

Figure 4.7 Visualisation example for the collision warning of trip 8332.

4.4.3 Conflict intensity evaluation for lane-changing interactions

The second experiment evaluates the conflict intensity of lane-changing interactions, with which we want to demonstrate two more characteristics of our approach. First, this approach can cover a more diverse range of conflicts during lane-changes than TTC. This stands in theory because TTC assumes constant movements (velocities in this study) of interacting vehicles, which identifies potential lane-change conflicts only if two vehicles have crossing velocity directions and could collide without movement change. In contrast, a unified metric trained with our approach on two-dimensional daily driving data considers conflicts in all directions and varying interaction situations. Second, the intensity of conflicts evaluated by this approach is expected to have a long-tailed distribution. More specifically, the number of detected conflicts should decrease according to a power-law as conflict intensity increases. A long tail in such a distribution indicates that, although high-intensity conflicts are extremely rare, they occur with a non-negligible probability. Not all surrogate metrics of conflicts have this characteristic, but it is crucial to indicate very rare conflicts.

There are 13,364 lane-changes in the highD dataset, and in total 713 (5.34%) of the lane-changes involve at least 1 second consecutively identified as a conflict by either TTC or the unified metric. This identification is under the optimal thresholds found in the first experiment, i.e., of 4.2 s for TTC and 17 for Unified. Among these lane-changes, in 679 the ego vehicle conflicts with one other vehicle, in 31 the ego vehicle conflicts with 2 other vehicles, and in 3 the ego vehicle conflicts with 3 other vehicles. These constitute 750 conflicts, among which 67 are indicated by both TTC and Unified; 18 are indicated by TTC only; and 665 are indicated by Unified only. In Figure 4.8, we plot the scattered relative locations of target vehicles to their ego vehicles aggregated at $(0, 0)$. These locations are where a lane-changing interaction reaches the minimum TTC or the maximum intensity evaluated by Unified.

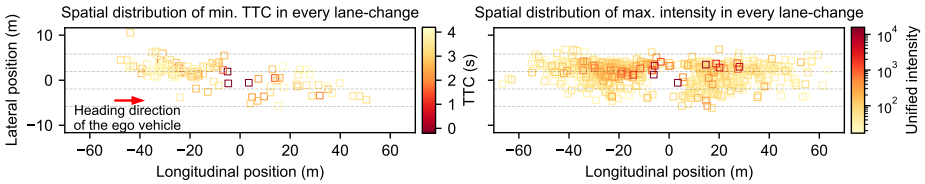


Figure 4.8 Spatial distributions of conflict moments with minimum TTC or maximum intensity in every lane-change. Target vehicle positions are transformed to a coordinate system centred at the ego vehicle position and with the longitudinal axis pointing to the ego vehicle heading direction. The dashed lines mark lanes at the average lane width in highD data.

Figure 4.8 clearly shows that conflict detection by TTC strongly relies on the assumption of constant movements, while training a unified metric can cover all directions around the ego vehicle and accounts for the heterogeneity of proximity behaviour in every different direction. There are significantly more conflicts between the ego vehicles making lane-changes and the target vehicles in the left lane than those in the right lane. This may

be due to more lane-changes from right to left or higher speeds of vehicles on the left lane, but future investigation is needed for more precise reasons. Notably, most of the lane-changing interactions have intensities lower than 100, which means one such conflict on average occurs in every 100 or fewer interactions. In comparison, conflicts with intensities higher than 100 are fewer, and those with intensities higher than 1000 are significantly fewer. However, despite their seeming rarity, they are far from unlikely to occur. In fact, foreseeing and preventing these safety-critical events remain key challenges for safe autonomous driving [50].

We present a closer look at the intensity distribution with Figure 4.9. In the left half of the figure, we plot the histograms of TTC values and Unified intensities respectively, for those at each individual moment and for averaged during each lane-change conflict case. The averaged values consider only lane-changing process, while the individual moments include car-following periods before and after lane-changes. The histograms show a decreasing frequency when intensity increases (TTC decreases), which is aligned with the assumption of conflict hierarchy. To further see if the distributions are long-tailed, in the right half of Figure 4.9, we make log-log scatter plots of the distributions, as well as their dashed trend lines and calibrated functions. All of the logarithm relationships are linear, therefore, all distributions to different extents have the characteristic of power-law. However, the trend line of TTC values averaged during lane-change conflicts has a higher slope than the line of TTC values at individual moments. This suggests that TTC detects fewer lane-changing conflicts than car-following conflicts. In contrast, the unified metric evaluates conflicts in the same theoretical framework, and thus has more consistent trend lines between averaging during lane-changes and at individual moments.

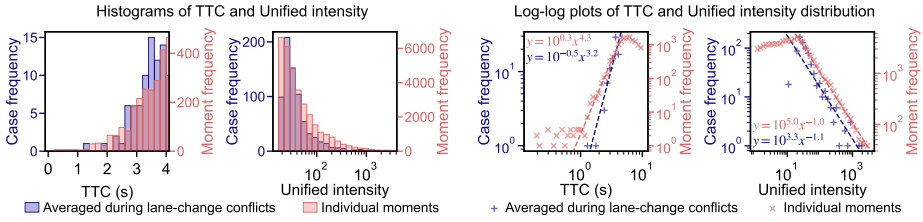


Figure 4.9 Intensity distributions during lane-changing interactions and at individual moments in the highD dataset.

To provide more intuitive information, we make dynamic visualisations for the detected conflicts in different ranges of intensity indicated by TTC and Unified. The readers are referred to this chapter's Abstract page to find links. Here in Figure 4.10 we present an example at location 1 (indexed in the highD dataset) with vehicle track indices of 1860 and 1858. The ego vehicle in red makes two sequential lane-changes, and conflicts with the target vehicle in the intermediate lane. Similar to the visualisation for near-crashes in Section 4.4.2, we plot the profiles of speed, acceleration, evaluated conflict intensity, and TTC values, as well as a real-time heatmap of proximity distribution. In the plots of profiles, lane-change periods are shaded. In addition, we visualise the interaction spectrum

described by $C(n, s; \phi)$, which relates conflict intensity n, current proximity s, and conflict probability p.

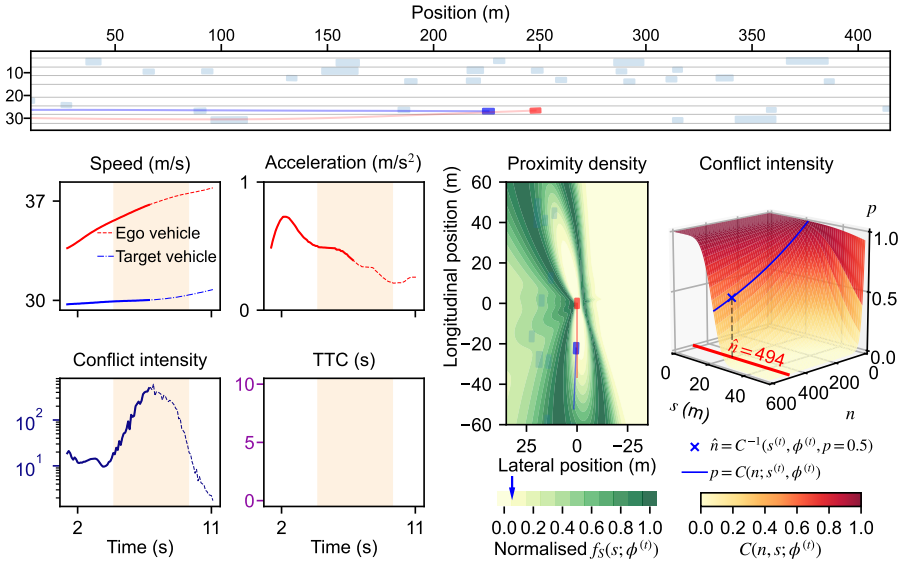
The first frame in Figure 4.10(a) is when the ego vehicle finishes its first lane-change from the right lane to the middle lane. In this frame, the conflict intensity is evaluated to reach the highest level in the whole process of sequential lane-change. On the surface of the interaction spectrum, we plot a line showing the relation between conflict intensity and probability at real-time proximity. This moment is at a very high probability to be considered as a minor conflict, and the maximum intensity to be considered as a conflict, i.e., the maximum intensity if conflict probability is larger than 0.5, reaches 494. This can be considered a serious conflict that occurs once per 494 times in the same interaction situations. Then the evaluated intensity decreases when the ego vehicle leaves the middle lane and continues moving to the left lane. In the frame shown in Figure 4.10(b), at the end of these two sequential lane changes, the evaluated intensity returns to safe levels.

4.5 Conclusion and discussion

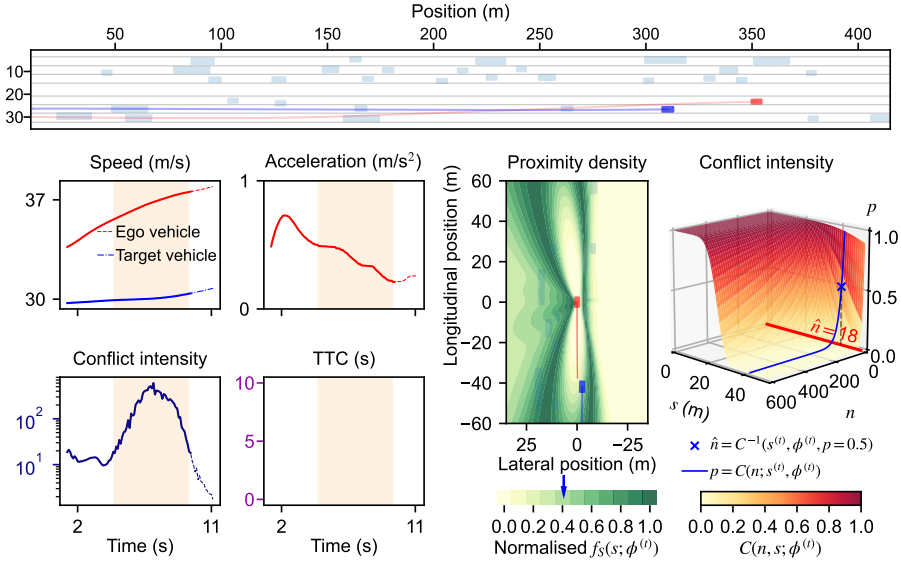
Conflicts do not arise out of nowhere, and every conflict is an extreme continuation of preceding safe interactions. Based on this assumption, this study presents a unified probabilistic approach to detecting traffic conflicts. The unified framework of traffic conflict detection models conflicts as context-dependent extreme events in ordinary interactions. Under this framework, any existing surrogate safety measure that captures certain aspects of conflicts is a special case. Then the statistical learning tasks allow for data-driven hypotheses of traffic conflicts, making the unified detection at least as effective as using a pre-hypothesised surrogate measure. Preliminary experiments with real-world trajectory data demonstrate that the proposed approach provides effective collision warnings, generalises well across datasets, captures various conflicts consistently, and detects conflicts in a long-tailed distribution of intensity.

These features enable consistent and comprehensive conflict detection, which can support scalable and reliable road safety research in the future. For example, more comprehensive surrogate measures of safety can be learned from naturalistic data. In the same way, an integrated estimation of collision risk becomes feasible for automated vehicles. Data-driven conflict detection also allows for the analysis of a broader range of factors underlying unsafe interactions in various situations. When the risk of conflicts is consistently evaluated across interaction contexts, it becomes possible to compare cross-modal road user interactions and provide consistent safety assessment of road infrastructures. Furthermore, safety-critical events that are rarely observed in data may be effectively identified, reinforced, and generated, contributing to training and testing automated vehicles. Pressing societal concerns, such as the impact of vehicle automation on road safety, could also be explored.

The current study has several limitations that should be investigated in future research. First, the demonstration in this study uses spatial proximity and trajectories in car-following and lane-changing scenarios; however, verification in more diverse traffic environments is needed. The framework could also be applied to temporal proximity, and involve other



(a) Frame 9719 at which the ego vehicle finishes first lane-change from the right lane to the middle lane.



(b) Frame 9747 at which the ego vehicle changes lane again to the left lane.

Figure 4.10 Example dynamic visualisation of a lane-changing conflict detected by Unified only.

road users such as pedestrians and cyclists. Future research on cross-modal interaction safety is thus promising. Second, the proposed approach does not directly predict when a collision will occur. This could be addressed by incorporating trajectory prediction models and forming a more complete methodology for safe interaction planning. Third, the theoretical framework is based on proximity and does not indicate the severity of a potential collision. Collision severity depends on the energy released by a collision, which is more physics driven than extreme value theory-based. As severity is a key aspect of collision avoidance, further exploration is necessary. Lastly, the uncertainty in conflict probability estimation, which could inform the reliability of detection, is not quantified in this study. This quantification can be achieved by inferring the distributions of proximity characteristic parameters, given that the parameters are learned using Gaussian models.

An important consideration regarding data should be noted. Although learning the representation of interaction context is identified as a key task, this study does not conduct representation learning. In theory, the use of learning methods can include almost all the information one can collect to describe the interactions between road users. This can go beyond movements and include road layouts, weather, individual characteristics of road users, etc. It is necessary to note, however, that the more information gets involved, the more diverse data is required to train an effective unified metric. Since the framework is well-suited for learning from accumulative evidence, future research could expand into continual learning as more data becomes available, allowing for increasingly robust and adaptive detection of traffic conflicts over time.

Chapter 5

Self-supervised collision risk quantification of traffic interactions

Highlights

- Collision risk is learnt from naturalistic interactions without crash or near-crash labels.
- Context-conditioned distributions of multi-directional spacing characterise interactions.
- Deviations from typical safe spacing towards closer extremes are quantified as risk.
- Outperformance over existing methods is validated on 2,591 real-world (near-)crashes.
- Environmental and historical kinematic features provide performance enhancement.
- Spacing direction, road-surface condition, and past kinematics are main risk factors.

Keywords

Road safety, collision risk, risk quantification, trajectory reconstruction, naturalistic driving data

This chapter is based on the article accepted for journal publication: Yiru Jiao, Simeon C. Calvert, Sander van Cranenburgh, and Hans van Lint. (2026). Learning collision risk proactively from naturalistic driving data at scale. Nature Machine Intelligence (in press). Available at arXiv: 2505.13556.

Abstract

Accurately and proactively alerting drivers or automated systems to emerging collisions is crucial for road safety, particularly in highly interactive and complex urban environments. However, existing approaches to identifying potential collisions either require labour-intensive annotation of sparse risk, struggle to consider varying contextual factors, or are only useful in specific scenarios. To address these limits, this study introduces the Generalised Surrogate Safety Measure (GSSM), a new data-driven approach that learns collision risk exclusively from naturalistic driving without the need for crash or risk labels. GSSM captures the patterns of normal driving and estimates the extent to which a traffic interaction deviates from the norm towards an unsafe state. Diverse data from naturalistic driving, including motion kinematics, weather, lighting, etc., are used to train multiple GSSMs, which are tested with 2,591 reconstructed real-world crashes and near-crashes. These test events are also released here as the largest dataset of its kind to date. A basic GSSM using only instantaneous motion kinematics achieves an area under the precision-recall curve of 0.9 and secures a median time advance of 2.6 seconds to prevent potential collisions. Additional interaction patterns and contextual factors provide further performance gains. Across various types of collision risk scenarios (such as rear-end, merging, and turning interactions), the accuracy and timeliness of GSSM consistently outperform existing baselines. Furthermore, feature attribution analyses reveal the dominant impacts on risk increase of spacing direction, road-surface condition, and historical kinematics in the passed second. These results establish GSSM as a scalable, context-aware, and generalisable foundation to quantify the risk of potential collisions before they occur, supporting proactive safety in autonomous driving systems and traffic incident management.

Code availability

`https://github.com/Yiru-Jiao/GSSM`

Data availability

Raw data sources:

- highD `https://levelxdata.com/highd-dataset`
- Argoverse2 `https://doi.org/10.4121/8d6ee0b0-8ed5-43f3-b1c9-7665cc163e87`
- SHRP2 Naturalistic Driving Study
  - `https://doi.org/10.15787/VTT1/FQLUWZ`
  - `https://doi.org/10.15787/VTT1/DEDACT`

Resulting data: `https://doi.org/10.4121/9caa1e6c-9abd-4e36-ae28-c9ea4542d940`

5.1 Introduction

Road traffic safety remains a critical global concern. Over one million fatalities were recorded on roads due to traffic accidents worldwide every year, as well as ten times as many injuries [1]. With decades of advancements in vehicle safety technologies and policy improvements, a significant number of fatalities have been reduced since 2010. However, this reduction has plateaued in recent years, as evidenced by Figure 5.1(a). Notably, fewer than 6% of all traffic accidents occur on high-speed motorways, as displayed in Figure 5.1(b). Instead, the majority of crashes occur on urban roads, where the traffic situation is more complex due to various types of road users and their multi-directional interactions. Given that 60% of the global population is expected to reside in urban areas by 2030 [167], improving traffic safety in highly interactive urban environments has become both urgent and globally relevant.

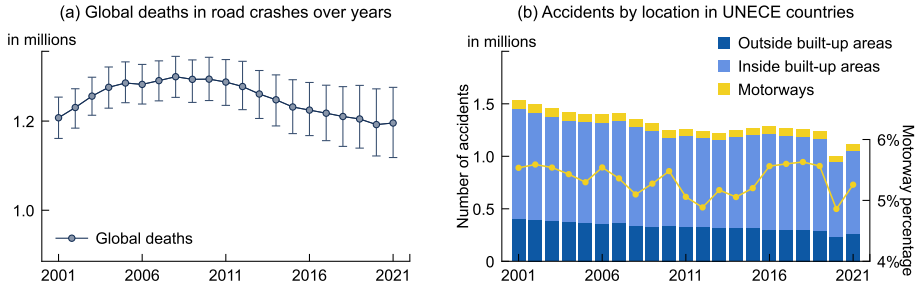


Figure 5.1 Statistics of traffic accidents. **(a)** Estimated numbers of fatalities due to road traffic accidents per year by the Institute for Health Metrics and Evaluation [168]. **(b)** Distribution of traffic accidents that occurred in different locations from 2001 to 2021 in 27 countries where complete data are accessible. The data is sourced from United Nations Economic Commission for Europe (UNECE) Statistical Database [5].

A core challenge lies in accurately quantifying collision risk in real time to allow for prevention before crashes happen. Traffic safety research is traditionally grouped into two categories, reactive and proactive, based on whether the focus is on ex post analysis or ex ante prevention. Reactive research identifies risk factors from real-world crashes to derive improvements in road topology design and infrastructure safety monitoring (see [169, 170] for relevant literature reviews). Relying on historical crash data, such improvements are inevitably delayed after injuries and damage have occurred. In contrast, proactive research anticipates potential crashes to enable timely interventions [171, 172]. This means being alert to the precursors to collisions, such as near misses, hazardous interactions, and evolving traffic conflicts, which potentially lead to crashes if not mitigated.

To address this challenge of proactive collision risk quantification, several methodologies have been explored in different fields. The earliest established methodology is surrogate safety measures (SSM), also referred to as surrogate measures of safety (SMoS) or criticality metrics within autonomous driving research [11, 139, 140]. Developed over decades since the 1970s [13, 21, 136], SSMs estimate the likelihood or severity of potential

collisions through physics-based indicators designed for specific road user behaviour such as car-following and lane-changing. The underlying assumption is that crash risk arises when the current situation will result in a collision unless immediate evasive actions are taken. Well-known examples include Time-to-Collision (TTC), deceleration rate to avoid collision (DRAC), artificial potential fields (e.g., safety field [173] and risk field [174, 175]), and other varying spatiotemporal measures. Although these indicators are intuitive, traffic collisions are not the result of a certain type of road user behaviour, but involve various types of interactions in diverse and dynamic traffic environments [176, 177]. Therefore, the specifically designed SSMs lack adequate ***context-awareness*** [178] to consider heterogeneous interactions, different types of road users, and broader environment factors. This lack also limits the ***generalisability*** of SSMs beyond their designed conditions.

A methodology widely used in robotics and the control field quantifies collision risk based on motion prediction under uncertainty [179, 180]. It predicts the future positions or reachable sets of road users, propagates uncertainty through these predictions, and assesses probabilistic violations of safety constraints (e.g., [42, 44, 46]). Risk quantification in this methodology implicitly assumes that the uncertainties of road user behaviour can be sufficiently captured by a stochastic model, which remains valid in safety-critical situations. In practice, this assumption is often violated by aggressive, inattentive, or other anomalous behaviour that typically precedes crashes. To accommodate such more complex uncertainties, recent studies have increasingly integrated data-driven prediction models [43, 181]. While these models enhance context-awareness as traffic environments are incorporated, the challenge of model ***generalisability*** persists in intensely interactive and safety-critical scenarios that are under-represented in normal training data. When encountering such scenarios, the accuracy and reliability of risk quantification may degrade significantly.

Facilitated by the rapid development of computer vision, a video-data-driven methodology has emerged since 2018. It is known as Traffic Accident Anticipation (TAA, see dedicated literature reviews such as [48, 49]), with the objective to provide early predictions of impending collisions by modelling the visual patterns across sequential video frames. This is commonly achieved through supervised learning with real-world accidents serving as labels. The fundamental assumption is the existence of representative visual cues before accidents happen, so that deep neural networks can identify these cues. For example, vehicles deviating from their lanes may signal loss of control or evasive manoeuvres [182], while motorcycles at high speed can indicate risk in urban areas [183]. Heavily driven by annotated crash data, TAA faces a practical challenge of ***scalability***. The infrequency and variety of traffic accidents, of which the videos are inherently difficult to acquire, impede effective and reliable training [50]. Similar to the previous methodologies, TAA models also struggle with ***generalisability*** to new contexts that differ, even slightly, from training patterns.

In summary, to proactively quantify collision risk, existing methodologies share 3 limitations in

- **scalability**, leveraging large-scale observations instead of relying on factual crashes;

- **context-awareness**, accounting for any information that can be used to characterise interaction [178] in addition to road user behaviour, such as weather, lighting, and road conditions; and
- **generalisability**, covering a wide variety of traffic interactions from on motorways to at urban intersections, and handling new contexts not seen in training data.

In this paper, we propose the generalised surrogate safety measure (GSSM), a novel approach to proactive risk quantification of potential traffic collisions. GSSM evaluates how extreme a given traffic interaction deviates from typical safe behaviour towards an unsafe state in the interaction context, outputting a continuous risk level and the corresponding likelihood of a potential collision. This consideration of extreme interaction makes it possible to naturally generalise from normal interactions to safety-critical situations. GSSM utilises neural networks to enable context-awareness, incorporating various relevant contextual information (e.g., motion states of road users, weather, and road conditions) and thus adapting to diverse scenarios from vehicle-vehicle encounters on highways to pedestrian-vehicle interactions in urban streets. Furthermore, GSSM learns interaction patterns from naturalistic data, i.e., real-world and unconstrained driving behaviour collected under everyday conditions with minimal interference, and does not require any crash records or manually labelled risk. This makes the training of GSSM scalable. In summary, GSSM offers a new paradigm for proactive collision risk quantification. It promises a contribution to advancing proactive safety research, to improve, e.g., autonomous driving systems, infrastructure design and operations, as well as traffic management policies.

In the rest of this paper, Section 5.2 first defines the problem of proactive collision risk quantification and then introduces GSSM. Its learning is explained in Section 5.3. In Section 5.4, we present a real-world dataset of crashes and near-crashes for validating GSSM. Our experiment design is described in Section 5.5, followed by Section 5.6 where the characteristics of GSSM are demonstrated. Finally, Section 5.7 concludes this paper and envisions future research. To improve readability, we place details that do not hinder understanding in appendices.

5.2 Generalised surrogate safety measure

In this section, we introduce the generalised surrogate safety measure (GSSM). We first formulate the general problem of proactive collision risk quantification, and then outline the theoretical basis for GSSM. Our design of GSSM inherits the fundamental knowledge developed over decades of research on surrogate safety measures (SSMs), where potential collisions are also called traffic conflicts. This distinguishes GSSM from existing self-supervised (and unsupervised) approaches that are tailored for computer vision and primarily based on anomaly detection. Despite that, GSSM is adequately flexible to incorporate vision data.

5.2.1 Problem formulation

Although the consequences of traffic accidents vary, this paper considers that every collision should be avoided and thus focuses on the likelihood of potential collisions. Within this scope, we define the proactive risk quantification of potential collisions as a function $Q : \mathcal{S} \mapsto L$, where $\mathcal{S}$ represents a scenario of traffic interaction and $Q(\mathcal{S})$ estimates the likelihood L of an impending collision in this scenario. L is typically expressed as a probability $p \in (0, 1)$, or as a numerical or categorical level M that can be monotonically mapped to p.

The primary objective of Q is to accurately and timely foresee a potential collision. Inaccurate and ill-timed anticipation, e.g., road users acting too quickly or, on the contrary, failing to act, are the most common factors of road crashes (as surveyed in [184], 51% of car drivers, 42% of motorcyclists, 68% of pedestrians, and 46% of cyclists). To alert to a potential collision, L is usually converted into a binary outcome using a threshold. When annotated data of crashes and near-crashes are available, assessing the accuracy of Q then compares these binary classification results with ground truth annotations, while timeliness is assessed by the time between the threshold-based alerts and the occurrence of a crash or near-crash event. In the absence of annotations, surrogates such as close distance and hard braking are used to create near-crash labels (e.g., [173, 185, 186]).

As reviewed in Section 5.1, Q can be predefined in closed-form expressions or based on uncertainty-aware prediction, and can also be learnt in a supervised manner with annotated crash videos. A particular challenge in collision risk quantification is that crash and near-crash data are expensive to collect and inherently rare, despite causing numerous fatalities and injuries. Consequently, there is a growing interest in unsupervised and self-supervised learning approaches that do not rely on annotated crash data, to name a few, see [45, 187–189].

5.2.2 Operational definition

Every collision evolves, often rapidly, from a previously safe interaction. This evolution was first articulated in [8] as the safety pyramid. In Figure 5.2 we depict the conceptual relations in the safety pyramid, which also reflects the interaction behaviours of road users [23]. For two or more approaching road users, a potential collision emerges when they are too close to interact safely. The perceived risk of a potential collision then triggers the road users to take evasive actions. In case of successful evasion, the potential collision will end in a near-crash; otherwise, a crash will occur.

Building on the safety pyramid, GSSM considers the spatial-temporal gap between road users as a proxy that can reflect collision risk across varying interaction contexts. Intuitively, a smaller gap implies less physical space and time available to react and prevent a potential collision. Empirical evidence also shows that people perceive increased risk when other moving objects approach [146, 147]; and motivated by the perceived risk, maintain acceptable separation from each other [77]. In this paper, we let GSSM use the spatial gap, hereafter referred to as spacing. GSSM quantifies collision risk by the extent to

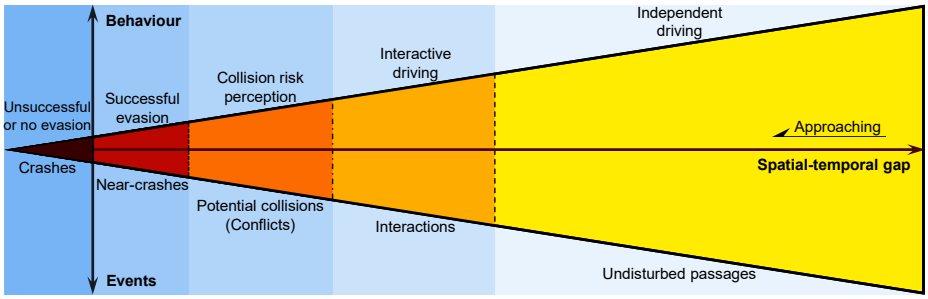


Figure 5.2 The safety pyramid conceptualises the evolution from safe interactions to unsafe interactions up to crashes.

which a spacing is too close for safe interaction in its interaction context. To enable cross-context risk quantification, we establish two key assumptions:

Assumption 1. Potential collisions are precursors to collisions as well as continuation of previously safe interactions.

Assumption 2. In the same interaction context, reduced spacing monotonically indicates increased risk of collision.

Formally, we consider repeated observations in the same interaction context and define the extent of proximity as the number of times a given spacing is the minimum among all observed spacings. Following the problem formulation in Section 5.2.1, we factorise an interaction scenario $\mathcal{S}$ into (s, X), where $s \geq 0$ is the spacing between two or more road users and X denotes all other contextual observables, e.g., motion states, weather, lighting, and road surface conditions. Accordingly, $\mathrm{GSSM} : (s, X) \mapsto M$, where M is a risk level. For $n \in \mathbb{N}_{>0}$ observations with varying spacings in context X, the extreme value theory calculates $p(s, n|X) = (\Pr(S > s|X))^n$ as the probability that s is the minimal spacing in n repeated observations. The smaller s is, over the more observations s remains minimal, and the larger n is.

We interpret the extent of spacing proximity n as conflict intensity. Accordingly, $p(s, n|X)$ can be understood as the probability that an interaction scenario (s, X) is a conflict at intensity n. Then we define the critical intensity of an interaction as the smallest n for which the probability of conflict exceeds the probability of no conflict, i.e., n satisfies $p(s, n|X) > 1 - p(s, n|X)$. Solving the inequality yields $n > \ln 0.5 / \ln(\Pr(S > s|X))$, and we use this critical point to define GSSM.

To summarise, GSSM is operationally defined as the critical conflict intensity $\hat{n}$ of an interaction scenario (s, X). From the view of extreme value theory, the spacing s has a probability of 0.5 to be the minimum in $\hat{n}$ observations in the same interaction context X; from a more intuitive view, if the interaction involves a conflict, its intensity is at least $\hat{n}$. The value of $\hat{n}$ may be very large, e.g., $1,000$ or $10,000$, in high-risk interactions. For convenience, we take the base-10 logarithm of $\hat{n}$, yielding representative scores of 3 and 4, respectively. Equation (5.1) presents the complete definition of GSSM with the following

properties. Let M be the risk level given by GSSM, $M \in \mathbb{R}$. When s is larger than the median spacing in context X, $\Pr(S > s|X) < 0.5$ and $M \leq 0$, which implies safety. In contrast, smaller s corresponds to larger M and a higher risk of potential collision. Particularly, when $s = 0$, M can be infinite and indicate a factual collision. In addition, M can be naturally mapped to a probability $p \in (0, 1)$ by $p(s, X, M) = \Pr(S > s|X)^{10^M}$.

$$\mathrm{GSSM}(s, X) = \log_{10}\left[\frac{\ln 0.5}{\ln(\Pr(S > s|X))}\right] \tag{5.1}$$

In theory, GSSM can be reduced to any SSM that is based on the spatial-temporal gap by specifying appropriate contextual observables. For example, Time-to-Collision (TTC) and deceleration rate to avoid collision (DRAC) account for the relative speed between two vehicles following one another; Proportion of Stopping Distance (PSD) considers the squared speed of a vehicle following another. For more details on such reduction and additional theoretical derivation, please refer to our previous work [190].

5

5.2.3 Multi-directional spacing

Spacing serves as a risk proxy in GSSM and reflects the physical and psychological reaction margin to prevent a potential collision. As mentioned in Section 5.1 and displayed in Figure 5.1(b), traffic collisions in urban traffic significantly outnumber those on highways. Urban traffic involves multi-directional movements beyond longitudinal dynamics. These include not only car-following and lane-changing but also two-dimensional (2D) interactions such as path crossing and conflict negotiation, especially at intersections.

We adapt the method developed in [155] to quantify multi-directional spacing in 2D interactions. Consider two road users, i and j. A relative coordinate system is introduced with its origin at the position of i and y-axis oriented along the direction of the relative velocity $\boldsymbol{v}_{ij} = \boldsymbol{v}_i - \boldsymbol{v}_j$. If i and j have identical velocities, the y-axis is instead aligned with the heading direction of i, denoted by $\boldsymbol{h}_i$. We transform the position of j, given by (x_j, y_j), into the relative coordinate system and get (x_{ij}, y_{ij}). Then the multi-directional spacing between i and j at each time step is represented by $(x_{ij}, y_{ij}, |\boldsymbol{v}_{ij}|)$. This coordinate transformation is explicitly defined in Equation (5.2).

$$\begin{aligned}
\begin{bmatrix} x_{ij} \\ y_{ij} \end{bmatrix} &= \frac{1}{\sqrt{x_{\mathrm{axis}}^2 + y_{\mathrm{axis}}^2}} \begin{bmatrix} y_{\mathrm{axis}} & -x_{\mathrm{axis}} \\ x_{\mathrm{axis}} & y_{\mathrm{axis}} \end{bmatrix} + \left(\begin{bmatrix} x_j \\ y_j \end{bmatrix} - \begin{bmatrix} x_i \\ y_i \end{bmatrix} \right); \\
(x_{\mathrm{axis}}, y_{\mathrm{axis}}) &= \begin{cases} (x_{\boldsymbol{v}_i} - x_{\boldsymbol{v}_j}, y_{\boldsymbol{v}_i} - y_{\boldsymbol{v}_j}), & \text{if } \boldsymbol{v}_i \neq \boldsymbol{v}_j, \\ (x_{\boldsymbol{h}_i}, y_{\boldsymbol{h}_i}), & \text{otherwise.} \end{cases}
\end{aligned} \tag{5.2}$$

To facilitate subsequent model training, we convert the positional components (x_{ij}, y_{ij}) in multi-directional spacing into polar coordinates. We define the polar coordinate system with its pole at the origin of the relative coordinate system and its polar axis aligned with the x-axis of the relative coordinate system. This conversion yields $(\rho_{ij}, s_{ij}, |\boldsymbol{v}_{ij}|)$ as the polar representation of multi-directional spacing, where ρ_{ij} is the angular coordinate and

s_{ij} is the radial coordinate. Respectively, ρ_{ij} represents the direction of spacing and s_{ij} represents the distance of spacing. We then incorporate ρ_{ij} and relative speed $|\boldsymbol{v}_{ij}|$ in the context X, and use s_{ij} as the spacing variable s.

5.2.4 Parameterised GSSM

We now introduce a parametric form of GSSM. Notice that $\Pr(S > s|X)$ in Equation (5.1) equals to $1 - \Pr(S \leq s|X)$. Consider the conditional distribution of spacing s in context X, $\Pr(S \leq s|X)$ is the cumulative probability distribution of $p(s|X)$. We thus parameterise $\Pr(S \leq s|X)$ by $F_S(s; \phi(X))$, where $\phi(X)$ denotes the conditional parameters depending on X. To specify $F_S(s; \phi(X))$, we can learn $\phi(X)$ using standard statistical and machine-learning techniques, thereby making GSSM more tractable for training.

Many studies demonstrate that the spatial gaps between road users are effectively characterised by the lognormal distribution [156–158]. We thus assume the context-conditioned spacing distribution $p(s|X)$ follows a lognormal distribution with two parameters μ and σ. For numerical stability in training, we learn $(\mu, \log(\sigma^2))$ instead of (μ, σ). But considering the convenience of notation, we write the estimated parameters $\hat{\phi}(X) = (\hat{\mu}(X), \hat{\sigma}^2(X))$ in this paper. Then Equation (5.3) presents the corresponding probability density function, and a parameterised GSSM is defined in Equation (5.4), where $\mathrm{erf}(z) = 2\int_0^z e^{-x^2}\mathrm{d}x/\sqrt{\pi}$ is the Gaussian error function within domain $(-1, 1)$.

$$f_S(s; \phi(X)) = \frac{1}{s\sqrt{2\pi\hat{\sigma}^2(X)}} \exp\left[-\frac{(\ln s - \hat{\mu}(X))^2}{2\hat{\sigma}^2(X)}\right] \tag{5.3}$$

$$\begin{aligned} \mathrm{GSSM}(s, X) &= \log_{10}\left[\frac{\ln 0.5}{\ln(1 - F_S(s; \phi(X))}\right] \\ &= \log_{10}\left[\frac{\ln 0.5}{\ln \frac{1}{2}\left(1 - \mathrm{erf}\left(\frac{\ln s - \hat{\mu}(X)}{\sqrt{2\hat{\sigma}^2(X)}}\right)\right)}\right] \end{aligned} \tag{5.4}$$

5.3 GSSM learning

For self-supervised risk quantification using GSSM, the primary learning task aims to estimate the conditional parameters of context-conditioned spacing distributions. Notably, learning a GSSM does not require crash or near-crash data; it leverages normal interactions, of which the available data is far more abundant. In addition to this primary task, we introduce two auxiliary tasks. One is context representation learning to incorporate a wider range of data for training; the other is feature attribution to identify contributing factors to potential collision risk.

5.3.1 Inference of conditional parameters

We use a neural network g_W to estimate the conditional parameters $\phi(X)$, where W denotes the network's learnable weights. Therefore, $g_W(X)$ outputs parameter estimates $\hat{\phi}(X)$ and it needs to make $f_S(s;\hat{\phi}(X))$ approximate the spacing distribution $p(s|X)$ as closely as possible. A closer approximation can be indicated by a smaller Kullback–Leibler divergence $D_{\mathrm{KL}}\left[p(s|X)||f_S(s;g_W(X))\right]$. Minimising this divergence is equivalent to maximising expected log-likelihood $\mathbb{E}_{s\sim p(s|X)}\left[\ln f_S(s;g_W(X))\right]$ given real-world observations of road user interaction. More conveniently, we minimise the negative log-likelihood (NLL) loss defined in Equation (5.5), where (s_i, X_i) represents a sample of interaction scenario described by spacing s_i and contextual observables X_i. This guides $g_W(X)$ to accurately fit the spacing distributions conditioned on various interaction contexts.

$$\begin{aligned}\mathcal{L}_{\mathrm{NLL}} &= -\ln\prod_{i=1}^{N} f(s_i;g_W(X_i)) = \frac{1}{N}\sum_{i=1}^{N}\ell_{\mathrm{NLL}}(s_i,X_i),\\ \ell_{\mathrm{NLL}}(s_i,X_i) &= \frac{1}{2}\left[\ln 2\pi + \ln\hat{\sigma}^2(X_i) + \frac{(\ln s_i - \hat{\mu}(X_i))^2}{\hat{\sigma}^2(X_i)}\right] + \ln s_i\end{aligned} \tag{5.5}$$

To avoid abrupt changes in risk quantification for continuous interactions, we introduce a smoothness regularisation term that penalises sharp differences in the distribution approximations for similar interaction contexts. Specifically, for each X_i, we generate a perturbed X_i' by adding a small Gaussian noise to each continuous variable in the original X_i. As shown in Equation (5.6), then a Jensen-Shannon divergence (D_{JS}) between the estimated distributions in context X and a similar context X' is weighted by β and added to $\mathcal{L}_{\mathrm{NLL}}$. Thereby, $W = \arg\min_W \mathcal{L}_{\mathrm{SmoothNLL}}$.

$$\begin{aligned}\mathcal{L}_{\mathrm{SmoothNLL}} &= \frac{1}{N}\sum_{i=1}^{N}\ell_{\mathrm{SmoothNLL}}(s_i,X_i,X_i',\beta),\\ \ell_{\mathrm{SmoothNLL}}(s_i,X_i,X_i',\beta) &= \ell_{\mathrm{NLL}}(s_i,X_i) + \beta D_{\mathrm{JS}}[f_S(s;g_W(X_i))||f_S(s;g_W(X_i'))]\end{aligned} \tag{5.6}$$

5.3.2 Context representation learning

In this paper, we categorise the contextual observables X describing an interaction scenario into three feature groups. **Current features** (denoted by X_C) represent the instantaneous states of interacting road users, including, e.g., their individual speeds and relative velocity. **Environment features** (X_E) describe external conditions during the interaction, such as weather, lighting, and road surface quality. **Historical kinematic features** (X_T) include time-series data of the interacting road users' speeds and yaw rates within the past 2.5 seconds. To avoid leakage of spacing information as suggested in [190], random values in X_T are dropped out (set as zero). A comprehensive list of the features used in this paper is provided in Appendix Table A1.

Each feature group is processed by a dedicated encoder, yielding encoded representation groups $\theta_C = g^C_{W_C}(X_C)$, $\theta_E = g^E_{W_E}(X_E)$, and $\theta_T = g^T_{W_T}(X_T)$. These groups are then

concatenated and passed to a decoder $g^D_{W_D}$, which estimates conditional parameters as illustrated in Equation (5.7). Details of these modules' architecture are provided in Appendix Section A.3.

$$\phi(X) = g^D_{W_D}([\theta_C; \theta_E; \theta_T]) \tag{5.7}$$

The learning of encoded representation is implicitly achieved when training a $g_W(X)$ as a whole. In practice, the encoders can also be pretrained by separate representation learning. For instance, contrastive learning is well aligned with GSSM training, which essentially aims to capture spacing patterns in similar conditions and distinguish those in dissimilar conditions. To focus on presenting GSSM, we do not perform separate representation learning in the experiments in this paper, but future research is expected for more practical applications.

5.3.3 Attribution of feature importance

In order to improve traffic safety, the features that strongly influence the estimated risk of potential collisions are worth identifying. GSSM learns interaction patterns from real-world data and varies its estimates of collision risk in different contexts. To analyse these learnt patterns and attribute each feature's effect on the estimated risk, we use an explainable artificial intelligence (XAI) method known as Expected Gradients (EG, [191]).

EG extends Integrated Gradients [192] by taking the expectation of gradients over integration path steps $\alpha \sim U(0,1)$ and a set of reference points. For clearer interpretation, we compute attributions with respect to various pieces of latent representation θ (rather than raw inputs). The reference points are typically sampled from a uniform distribution over the entire training set [191]. Here, we apply k-means clustering to Θ, the representations encoded from all samples in the training set, and use cluster centres as representative references θ'. Equation (5.8) summarises our computation of EG. The computed attributions sum to the difference in estimated risk given θ relative to the averaged risk given the references. The positive attribution of a feature means it contributes to increasing the estimated risk, whereas the negative attribution implies a contribution to decreasing the risk. More details about the encoder design to ensure correct dependence of contextual features are referred to Appendix Section A.3.

$$\mathrm{EG}(\theta) = \mathop{\mathbb{E}}_{\theta' \sim \Theta, \alpha \sim U(0,1)} \left\{ (\theta - \theta') \frac{\partial}{\partial \theta} \log_{10} \left[\frac{\ln 0.5}{\ln(1 - F_S(s; g^D_{W_D}(\theta' + \alpha(\theta - \theta'))))} \right] \right\} \tag{5.8}$$

Note that these attributions do not establish causality. For example, a positive attribution of relative speed does not imply that a lower relative speed would have reduced the risk. In a posteriori manner, EG explains the features' computational influence on the inference of a trained neural network. We recommend two strategies for interpretation. First, compare the attributions of all features at a specific time moment to assess their relative importance. Second, track how the attributions evolve over time. For instance, if the estimated collision risk rises along with an increasing attribution for a particular feature, that feature may have a relatively strong correlation with collision risk. Nonetheless, such an explanation alone cannot serve as evidence to justify an intervention, where future research is needed.

5.4 Data of potential collisions

Despite that GSSM does not use real-world traffic accidents for training purposes, a relatively small-scale dataset of crashes and near-crashes is necessary for effectiveness assessment. Such a dataset needs to contain both safe and safety-critical road user interactions in various situations. In Table 5.1, we summarise existing datasets that each have over 500 crashes and/or near-crashes recorded. Among these, we use the database derived from the Second Strategic Highway Research Program's (SHRP2) Naturalistic Driving Study (NDS), where we reconstruct bird's eye view trajectories for easier use. The major shortcoming of the other datasets is the lack of safe baselines to test false positives, except for DAD, which provides 1,130 clips sampled from source videos. Another issue is that most of the datasets were originally designed for accident detection rather than risk quantification, thus the time periods before crashes are not well preserved for assessing risk evolution. In the rest of this section, we first introduce SHRP2 NDS, then describe the trajectory reconstruction, and finally summarise the data we use in the following experiments.

Table 5.1 Existing datasets of crashes (and near-crashes).

Dataset [a]	Year [b]	Annotation [c]	Crashes (and near-crashes)		
			Number	Time range	Period before impact [d]
DAD [193]	2016	T	620	5 s	Designed to be 4.5 s
CADP [194]	2018	T, S	1,416 [e]	Varying	On average 3.7 s
A3D [187]	2019	T, P	1,500	Varying	Unrevealed
CCD [195]	2020	T, P, W	1,500	5 s	At least 3.0 s
DADA [196, 197]	2019, 2022	T, S, C, P, W, L, A	2,000	Varying	On average 5.2 s
DoTA [189]	2022	T, S, C, P	4,677	Varying	On average around 4.0 s
SHRP2 NDS [198, 199]	2010-2013	T, S, C, P, W, L, etc.	8,895 [f]	Varying	At least 20.0 s
Reconstructed SHRP2	This paper	T, S, C, P, W, L, etc.	6,664 [g]	Varying	On average 22.8 s

[a] Data sources: DAD, A3D, CCD, and DoTA were compiled from dashcam recordings posted on YouTube. DADA was compiled from dashcam recordings posted on YouTube, Youku, Bilibili, iQiyi, Tencent, etc. CADP was compiled from surveillance camera recordings posted on YouTube.

[b] For SHRP2 NDS, the listed years are actual data collection time. For the other datasets, the years are publication time, since these sets were sourced from the Internet and have no determinate collection time.

[c] Annotated information includes: event time (T), bounding box positions (S), crash type (C), participants (P), weather (W), lighting (L), driver's attention map (A). The datasets derived from SHRP2 NDS include additional information such as traffic density and road surface quality.

[d] Impact means physical contact in crashes; for near-crashes, the moment of closest proximity.

[e] Among these videos, 205 are annotated while the others are not.

[f] These include 1,942 crashes and 6,953 near-crashes.

[g] These include 1,402 crashes and 5,262 near-crashes.

5.4.1 SHRP2 Naturalistic Driving Study

The SHRP2 NDS is a large-scale research initiative aimed at understanding driver behaviour and performance [200]. Between 2010 to 2013, it collected extensive data using instrumented vehicles in six states in the United States. More than 3,300 participant

vehicles were equipped with data acquisition systems that recorded video footage, vehicle network data (e.g., speed, brake, and accelerator positions), and signals from additional sensors such as forward radar and accelerometers. A significant strength of SHRP2 NDS is its comprehensive set of manually annotated traffic safety events, including crashes and near-crashes that are collectively termed "safety-critical events", along with "safe baselines" selected through stratified random sampling. Table 5.2 outlines the operational definitions of these events as described in [12], where further details of the SHRP2 NDS are also referred to.

Table 5.2 Operational definitions of traffic safety events in the SHRP2 NDS.

Event	Operational definition
Crash	"Any contact that the subject vehicle has with an object, either moving or fixed, at any speed in which kinetic energy is measurably transferred or dissipated is considered a crash. This also includes non-premeditated departures of the roadway where at least one tyre leaves the paved or intended travel surface of the road, as well as instances where the subject vehicle strikes another vehicle, roadside barrier, pedestrian, cyclist, animal, or object on or off the roadway."
Near-crash	"Any circumstance that requires a rapid evasive manoeuvrer by the subject vehicle, or any other vehicle, pedestrian, cyclist, or animal, to avoid a crash is considered a near-crash. A rapid evasive manoeuvrer is defined as steering, braking, accelerating, or any combination of control inputs."
Safe baseline (Non-conflict)	"Normal driving behaviours and scenarios where the driver may react to situational conditions and events, but the reaction is not evasive and the situation does not place the subject or others at elevated risk."

5.4.2 Trajectory reconstruction

The accessible motion data from the SHRP2 NDS do not contain positional information of the subject (participant) vehicles to protect driver privacy. In addition, the other road users are detected by forward radars as surrounding objects in a subject-centric moving coordinate system. To align these data with the mainstream bird's eye view datasets, we reconstruct the trajectories for both subject vehicles and surrounding objects using extended Kalman filters (EKFs). First, we linearly interpolate all time-series signals to a uniform frequency of 0.1 seconds. Next, we reassign the indices of surrounding objects when necessary. Due to detection limitations, an object may be temporarily lost and re-detected later with a new index. If a newly detected object is within a specified distance threshold of a previously tracked object, the new object is assigned the previous index. We define the distance threshold as the object's position displacement relative to the subject over 0.3 seconds, constrained to a minimum of 0.5 m and a maximum of 2.5 m. Then we reconstruct trajectories for each event in two steps.

Step 1: subject vehicle's trajectory reconstruction. We apply an EKF assuming constant yaw rate and acceleration. The motion dynamics are updated according to

Equations (5.9), (5.10a) and (5.10b), where (x_i, y_i) denotes the vehicle's position, ψ_i the heading angle relative to the x-axis, v_i the longitudinal speed, ω_i the yaw rate, a_i the longitudinal acceleration, and $\epsilon = 0.001$ is a threshold to use longitudinal updates when near-zero yaw rates induce numerical instability. The update interval $\Delta t = 0.1$ seconds. We place the subject vehicle initially at $(x_i, y_i) = (0, 0)$ with its heading $\psi_i = 0$ along the x-axis, and set the initial states of speed, yaw rate, and acceleration from original data. The data of yaw rates and accelerations are stably recorded, but speed measurements are not always consistent. We thus consider two orders of time sequence depending on whether the earliest or the latest 0.5-second speed states are missing. If the latest states are missing, we let the EKF propagate forward from the earliest available measurement; if the earliest states are missing, we let the EKF propagate backwards. When both earliest and latest speed states exist, we run two EKFs from both ends and then select the reconstructed trajectory that deviates less in speed and yaw rate from the original sensor data.

$$\begin{bmatrix} x_i \\ y_i \\ \psi_i \\ v_i \\ \omega_i \\ a_i \end{bmatrix}_{t+\Delta t} = \begin{bmatrix} x_i \\ y_i \\ \psi_i \\ v_i \\ \omega_i \\ a_i \end{bmatrix}_{t} + \begin{bmatrix} \Delta x_i \\ \Delta y_i \\ \omega_i \Delta t \\ a_i \Delta t \\ 0 \\ 0 \end{bmatrix}, \text{ where} \tag{5.9}$$

$$\Delta x_i = \begin{cases} \cos(\psi_i)\left(v_i \Delta t + \frac{1}{2} a_i \Delta t^2\right), & \text{if } |\omega_i| \le \epsilon, \\ \frac{v_i \omega_i \left[\sin(\psi_i + \omega_i \Delta t) - \sin(\psi_i)\right]}{\omega_i^2} & \\ + \frac{a_i \left[\cos(\psi_i + \omega_i \Delta t) - \cos(\psi_i)\right]}{\omega_i^2} & \text{and} \\ + \frac{a_i \omega_i \sin(\psi_i + \omega_i \Delta t) \Delta t}{\omega_i^2}, & \text{otherwise;} \end{cases} \tag{5.10a}$$

$$\Delta y_i = \begin{cases} \sin(\psi_i)\left(v_i \Delta t + \frac{1}{2} a_i \Delta t^2\right), & \text{if } |\omega_i| \le \epsilon, \\ \frac{v_i \omega_i \left[\cos(\psi_i) - \cos(\psi_i + \omega_i \Delta t)\right]}{\omega_i^2} & \\ + \frac{a_i \left[\sin(\psi_i + \omega_i \Delta t) - \sin(\psi_i)\right]}{\omega_i^2} & \\ - \frac{a_i \omega_i \cos(\psi_i + \omega_i \Delta t) \Delta t}{\omega_i^2}, & \text{otherwise.} \end{cases} \tag{5.10b}$$

Step 2: surrounding objects' trajectory reconstruction. We then reconstruct the trajectories of the surrounding objects. This reconstruction has the forward field of view only, as only forward radar data are available. During an event, the subject vehicle may detect multiple objects (e.g., vehicles, cyclists, pedestrians, or animals), of which the edge nearest to the subject vehicle is detected. For each object, we first convert its local radar

coordinates into the subject vehicle's reconstructed global coordinate system. Then we determine the object's centroid based on its dimensions and whether the detected edge corresponds to its front or rear, inferred from its heading direction. Since only relative positions and speeds are available for these detected objects, we use an EKF under the assumption of constant heading and speed. The update equations are given in Equation (5.11).

$$\begin{bmatrix} x_j \\ y_j \\ \psi_j \\ v_j \end{bmatrix}_{t+\Delta t} = \begin{bmatrix} x_j \\ y_j \\ \psi_j \\ v_j \end{bmatrix}_{t} + \begin{bmatrix} \cos(\psi_j) v_j \Delta t \\ \sin(\psi_j) v_j \Delta t \\ \psi_j \\ v_j \end{bmatrix} \tag{5.11}$$

Due to limited access to the raw data of cameras and radars, from which the provided data in the SHRP2 NDS were extracted, our reconstruction intends to preserve the first-hand information (i.e., the provided data) as much as possible. Therefore, this study defines reconstruction errors as the deviations of reconstructed trajectories from the provided signals in the SHRP2 NDS. We optimise the EKF parameters of uncertainties and motion ranges by minimising the root mean squared reconstruction error in subject speed, subject yaw rate, subject acceleration, and object speed, as well as the mean displacement error of objects. The distributions of eventual reconstruction errors are presented in Appendix Figure A1. Note that due to the absence of ground truth, the very small errors we achieved do not indicate an accurate capture of reality. Instead, the errors quantify reconstruction deviations from the sensor measurements. On average, the standard deviation is smallest for safe baselines and larger for near-crashes and crashes. But the difference is rather small, with 0.03-0.04 m/s of subject speed, 0.08-0.13 m/s^2 of subject acceleration, and 0.02-0.06 m/s of object speed.

5.4.3 Data used in this paper

We use crashes and near-crashes derived from the SHRP2 NDS as our test set for GSSM demonstration. The safe baselines are used as part of our training data. Therefore, this paper utilises all reconstructed event trajectories, alongside additional information relevant to the safety-critical events.

Not all 8,895 events in the original dataset could be reconstructed due to missing values, sensor inaccuracies, and the absence of ground-truth data. There are 6,664 events where the trajectories of both the subject vehicle and at least one surrounding object are reconstructed. From these, we extract a useful test set by excluding invalid events that meet any of the following criteria:

- no object is detected for more than 5 seconds (too short to observe risk evolution),
- the crash or near-crash is with an object behind the subject vehicle (given the lack of rearward radar data),
- the crash or near-crash involves a "shapeless" obstacle (e.g., roadside pavement).

Eventually, we obtain 4,875 safety-critical events in the test set. Figure 5.3(a) presents the numbers of originally recorded events in the SHRP2 NDS, the subset with both subject and object trajectories reconstructed, and the final filtered subset.

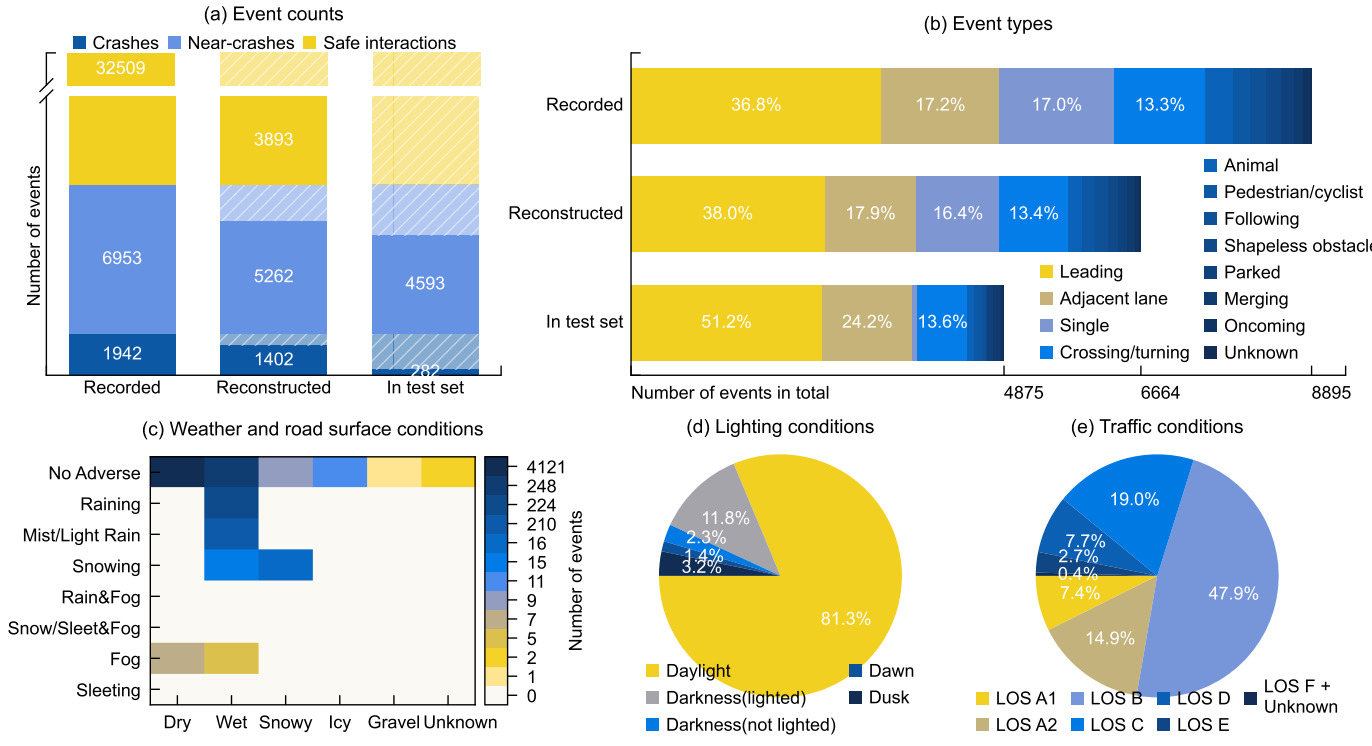


Figure 5.3 Statistics of original and processed events in the SHRP2 NDS. **(a)** Numbers of crashes, near-crashes, and safe baselines that are recorded, reconstructed, and used in the test. **(b)** Distribution of event types that are recorded, reconstructed, and used in the test. **(c)** Distribution of weather and road surface conditions in the test events. **(d)** Distribution of lighting conditions in the test events. **(e)** Distribution of traffic conditions in the test events. LOS stands for level of service, details of which are referred to Appendix Table A1.

The distributions of event types are further displayed in Figure 5.3(b), where the difference is primarily a result of the fact that rear-end events are more likely to be continuously detected by the forward radar. In Figures 5.3(c-e), we also show the numbers and percentages of events in different environmental conditions. More than 80% of the events occurred without adverse conditions, i.e., dry roads, not precipitation, and in daylight. However, considering that adverse conditions are less common by nature, they still play a role in causing traffic accidents.

In addition to the reconstructed trajectories, we attach other information about the safety-critical events. To consider a broader range of factors leading to potential collisions, the environmental conditions, including weather, lighting, road surface, and traffic density, are incorporated. To evaluate whether and when a potential collision is successfully detected, the annotations of time when an event starts and ends, when the driver reacts (if applicable), and when the impact occurs (a physical contact for crashes or the closest proximity in near-crashes) are also included. Further, the specific type (e.g., with a leading object, during crossing or turning) of an event is included to evaluate the detection of potential collisions in different scenarios. Lastly, when necessary, each event's narrative is referred to for further verification. These result in the largest to date trajectory dataset of traffic crashes and near-crashes. This dataset has been made accessible in [201].

5.5 Experiments

We run all experiments with an NVIDIA A100 GPU (80GB RAM) as well as 5 to 50 Intel Xeon CPUs, depending on whether parallel computation is helpful. To ensure fair comparison across settings and methods, we control the following conditions: random seed, maximum training epochs, early stopping criteria, and evaluation sample sets. Hyperparameter choices are listed in Appendix Table A2. We do not conduct an ablation experiment for neural network design, as our core focus by experiments is on validating the proposed generalised surrogate safety measure (GSSM).

5.5.1 Training datasets and experiment design

Our training data consist of 3 datasets collected in different countries, involving different interaction scenarios, and with different equipment. The first dataset is composed of the safe baselines derived from SHRP2 NDS as described in Section 5.4. We term this set as *SafeBaseline*, where the interactions were recorded both on highways and in urban traffic. Considering the detection limitation of forward radar, the majority of SafeBaseline may be on straight roads. To compensate for that, we introduce two more datasets. One is specifically focused on urban intersections. This dataset was collected in the U.S. in 2019 by a fleet of automated vehicles [202]. We use the set of interactions between human road users extracted by [203], and term this set as *ArgoverseHV*. From the other dataset known as *highD*, we extract lane-change interactions using the same methods in our previous work [190]. This dataset was collected by drones on German highways in 2018, with a position error typically less than 10 cm [161]. To make fair use, we draw approximately

equal sample sizes for each dataset, where 80% are used for training and 20% for model validation to monitor training progress.

The experiments are designed as follows to demonstrate the effectiveness of GSSM and its useful characteristics targeted in this research.

- **Effectiveness**. Effective risk quantification accurately identifies dangers and provides timely alerts to prevent potential collisions. Therefore, we compare GSSM with other existing methods on the *accuracy* in alerting crashes and near-crashes as well as the *timeliness* of alerts. For comparability, the GSSM compared in this experiment is trained using only current features (X_C).
- **Scalability**. The training of GSSM and its effectiveness are expected to be scaled with increased interaction patterns. We combine SafeBaseline with varying proportions (random sampling 10%–100%) of ArgoverseHV or highD, train a series of GSSMs on the combined datasets, and then observe the variation in their effectiveness.
- **Context-awareness**. GSSM can consider varying contextual information, and additional contextual features are expected to enhance risk quantification effectiveness. We train multiple GSSMs on SafeBaseline, which uniquely provides environment data. In addition to current features (X_C), these GSSMs progressively incorporate more contextual information of subject acceleration (a_i), environment features (X_E), and historical kinematic features (X_T).
- **Generalisability**. Our training of GSSM does not use any crashes or near-crashes as in the test set. Essentially, the effectiveness of all the GSSMs trained in this study is generalised from normal interactions, which vary in location and equipment during data collection. In addition to that, GSSM is designed to be generally applicable across interaction scenarios. We compare the effectiveness of GSSM against other methods for different types of safety-critical events, including rear-end scenarios, lateral interactions with other road users in adjacent lanes, crossing and turning conflicts, merging situations, and a small number of incidents involving vulnerable road users (e.g., pedestrians, cyclists).
- **Risk attribution**. Lastly, we evaluate the dominant contributing factors to the change in collision risk based on feature attribution.

Table 5.3 summarises the 29 GSSMs that are trained across these experiments. These variations depend on the training datasets (or their combinations) and the selected contextual observables.

Since we do not include video data yet as contextual information, the comparison in this paper's experiments excludes the self-supervised methods based on computer vision or anomaly detection. Instead, we compare with surrogate safety measures (SSMs) tailored for two-dimensional (longitudinal and lateral) interactions as listed in Table 5.4. Their estimations are all based on the instantaneous states of interacting road users, i.e., the current features X_C. For efficient large-scale evaluation, we implement TAdv, ACT, and TTC2D in a vectorised form, enabling parallel computation over thousands of interaction pairs.

Table 5.3 Overview of GSSMs to train in the experiments.

Acronym [a]	Training dataset(s) [b]	Contextual observables	Experiment(s) [c] involved in
S-C	SafeBaseline		S, C
h-C	highD		E, G
A-C	ArgoverseHV		G
SA-C	SafeBaseline ∪ ArgoverseHV$^{\{10\%,20\%,\cdots,100\%\}}$ [d]	$[X_C]$	S
Sh-C	SafeBaseline ∪ highD$^{\{10\%,20\%,\cdots,100\%\}}$		S
SAh-C	SafeBaseline ∪ 10%ArgoverseHV ∪ 100%highD		S
S-Ca		$[X_C, a_i]$	C
S-CE		$[X_C, X_E]$	C
S-CaE	SafeBaseline	$[X_C, a_i, X_E]$	C
S-CET		$[X_C, X_E, X_T]$	C, R, G
S-CaET		$[X_C, a_i, X_E, X_T]$	C

[a] Acronyms follow the format `<Datasets>-<Contextual observables>`. For example, "S-C" is trained on SafeBaseline and uses current features X_C

[b] The dataset names are abbreviated as "S" = SafeBaseline, "A" = ArgoverseHV, "h" = highD. If multiple datasets are used (e.g., "SAh"), the GSSM is trained on their union.

[c] The experiments are abbreviated as "E"=Effectiveness, "S"=Scalability, "C"=Context-awareness, "G"=Generalisability, "R"=Risk attribution. A GSSM may be involved in multiple experiments.

[d] The data combined in addition to SafeBaseline are randomly sampled for 10%-100% from the whole dataset.

Table 5.4 Two-dimensional Surrogate Safety Measures (2D SSMs) used as baseline methods.

Two-dimensional SSM	Year	Brief summary
Time Advantage (TAdv) [23]	2010	Expected time gap between two road users passing a common spatial zone, assuming no behaviour change in the short future. This is also known as predicted Post-Encroachment-Time (PET).
Anticipated Collision Time (ACT) [18]	2022	The shortest distance between two road users divided by their closing-in rate, assuming a constant closing-in rate in the short future.
Two-dimensional Time-to-Collision (TTC2D) [204]	2023	Minimum of the longitudinal and lateral Time-to-Collision values, assuming no behaviour change in the short future.
Emergency Index (EI) [28]	2025	Intensity of evasive action required to prevent a potential collision, based on the change rate in overlapping path.

There also exist extensions that incorporate stochastic behaviour and/or extreme value theory (e.g., survival analysis), such as [149, 205, 206]. We do not compare with these methods because their implementations require specific modelling assumptions about uncertainty (for example, how to represent trajectory noise and behavioural variability), which makes a fair and consistent comparison difficult within our current framework. Nevertheless, we refer interested readers to this line of work as a complementary direction that combines SSMs and probabilistic modelling.

5.5.2 Ground truth of safety-critical events

The original SHRP2 NDS annotates the time moments of the start, impact, and end of a crash or near-crash. For crashes, the impact time means physical contact time; for near-crashes, impact time is when the road users were at the closest proximity. The safety-critical events in the test set have at least 20 seconds before the impact time, which allows for separating the danger period and safe period in each event. Then we can create positive ground truth (i.e., potential collision) between the subject vehicle and a conflicting object in the danger period; and negative ground truth (i.e., safe interactions) between the subject vehicle and other surrounding objects in the safe period. As illustrated in Figure 5.4, we consider the danger period of an event from the start time, or 4.5 seconds before impact time if the annotated start time is less than 4.5 seconds until impact. This danger period is considered to finish at the end time, or in 0.5 seconds after impact time if the annotated end time is delayed. Meanwhile, we consider the safe period for each surrounding object as 2 to 5 seconds after the object is detected for 1.5 seconds and 3 seconds before the event start time. To be considered as in a safe period, the object should not have a hard braking (deceleration greater than 1.5 m/s^2).

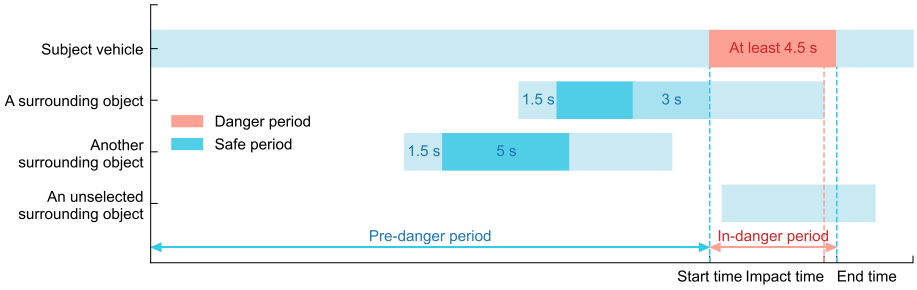


Figure 5.4 Separation of safe and danger periods for each safety-critical event in the test set.

The specific "conflicting object" in an event is not explicitly annotated in the SHRP2 data, and it's possible the conflicting object was not detected or recorded. We design a three-stage evaluation procedure to filter out those invalid cases, create useful ground truth, and then perform a reliable evaluation.

- **In the first stage**, for each event, we use every method among the 2D SSMs and GSSMs (except for the ones used to test scalability only) to evaluate the collision

risk between the subject vehicle and each of all detected surrounding objects, then determine a temporary conflicting object based on the evaluation. An event may have zero or a few candidate objects evaluated by different methods.

- **In the second stage**, we let the methods vote for the most probable conflicting object. Considering that each method votes for a candidate object, and an abstention is when a method determines that no temporary conflicting object exists. We select the object earning the most votes to be the eventual conflicting object. This voted object should earn over 1/3 of the total votes, and have less than 1/3 of the total votes against.
- **In the third stage**, we apply the 2D SSMs and GSSMs to each event that has a voted conflicting object. True and false positives are assessed based on the potential collision between the subject vehicle and the voted conflicting object in the danger period. True and false negatives are assessed based on the safe interactions between the subject vehicle and other surrounding objects in their safe periods.

To determine a temporary conflicting object in the first stage, we assume that the conflicting object in an event is the most risky object in the danger period, and its collision risk is less in the pre-danger period than in-danger period. More specifically, the object with the highest average collision risk during in-danger period is first selected. If the selected object has no data before the danger period, further confirmation is not possible, so we skip the event to be conservative than misleading. If any data of this selected object is recorded before danger, the 25th, 50th, and 75th percentiles of the risk quantification in the pre-danger period should be less than those in the danger period; otherwise, this object is not the real conflicting object. Here we use the percentiles (rather than, e.g., minimum, mean, maximum) for a more robust comparison and to avoid the potential influence of outliers.

5.5.3 Evaluation metrics

As mentioned in Section 5.2.1, the evaluation of proactive risk quantification includes two aspects of detection accuracy and alert timeliness. The following subsections detail the metrics used in this paper.

Detection accuracy

Using a threshold, we convert the risk quantification of a potential collision into binary outcomes of "safe" or "unsafe", then use the unsafe outcomes to issue alerts. Depending on whether the alerts correctly warn of a crash or near-crash and stay silent during safe interactions, each event has 4 possible counts. Targeting the conflicting object in the danger period, a True Positive (TP) issues alerts for at least 0.5 seconds; a False Negative (FN) issues no alerts. For all qualified objects other than the conflicting object in their safe periods, True negatives (TNs) issue no alerts; False positives (FPs) incorrectly issue alerts. Then counting for all test events, Equation (5.12) defines false negative rate R_{FN} and false positive rate R_{FP}; Equation (5.13) defines $\mathrm{Precision}$, Recall, and a combined accuracy score F_1.

$$R_{\mathrm{FN}} = \frac{\mathrm{FN}}{\mathrm{TP} + \mathrm{FN}}, \quad R_{\mathrm{FP}} = \frac{\mathrm{FPs}}{\mathrm{FPs} + \mathrm{TNs}}. \tag{5.12}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FPs}}, \quad \text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \quad F_1 = \frac{2\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}. \tag{5.13}$$

Alert timeliness

We define Time to Impact (TTI) to measure how early a potential collision is recognised relative to impact time (t_{impact}), the moment of physical contact in crashes or the moment of minimum proximity in near-crashes. As a posteriori quantity, TTI counts from the last time moment when risk quantification shifts from safe to unsafe until a (near-)crash in reality. This is formally defined in Equation (5.14), where M_t is the estimated risk level of a potential collision at time t, M^* is a threshold to distinguish safety. We let $M_{-1} \leq M^*$ in case all recorded time moments are estimated as unsafe. Without loss of generality, we assume that a larger M indicates a higher risk; the inequalities can be reversed if a smaller M indicates higher risk.

$$\text{TTI} = t_{\text{impact}} - \max\{t \mid M_t > M^*, M_{t-1} \leq M^*, t \leq t_{\text{impact}}\} \tag{5.14}$$

For a collective evaluation over test events, we first use $P_{\text{TTI}\geq 1.5}$, which indicates the percentage of events with $\text{TTI} \geq 1.5$ s among correctly detected (near-)crashes. The threshold of 1.5 s is motivated by empirical findings that human drivers typically require 1 to 1.3 s to respond to an obstacle [165, 166]. As a second measure of alert timeliness, we use median time-to-impact across all correctly detected (near-)crashes, denoted by $m\text{TTI}$. Based on the definition in Equation (5.14), $\text{TTI} \geq 0$ and its maximal value depends on the time period recorded before impact. We thus restrain $\text{TTI} < 10$ s when calculating $m\text{TTI}$. To reflect the variation of TTI, we report $m\text{TTI}\ [Q1, Q3]; 99\%CI$, where $Q1$ and $Q3$ are the first and third quartiles, respectively, and $99\%CI$ is the 99% confidence interval calculated using the sign test method for $m\text{TTI}$. Similar to the accuracy metrics, $m\text{TTI}$ and $P_{\text{TTI}\geq 1.5}$ vary at different alerting thresholds.

Performance curves

Risk quantification must balance the trade-off between desired yet conflicting characteristics at different thresholds. For example, a threshold resulting in smaller R_{FP} necessarily increases R_{FN}. Such a threshold also entails higher Precision, but lower Recall. Likewise, earlier alerts (i.e., larger TTI) tend to be accompanied by more false positives and less accurate detection. Therefore, our evaluation uses three key performance curves to compare different methods on a common set of events in the test dataset. Each curve is obtained by varying the threshold that distinguishes risk from safety. As summarised in Table 5.5: Receiver operating characteristic curve (ROC) intuitively presents the trade-off between false detection rates; Precision-recall curve (PRC) comprehensively evaluates accuracy performance; and Accuracy-timeliness curve (ATC) evaluates timeliness performance under different accuracy levels.

If a perfect model exists, ROC and PRC have their theoretical optimal points of zero false and all correct detection. ATC does not have a theoretical optimum, but we

Table 5.5 Performance curves of risk quantification at varying thresholds.

Performance curve	Horizontal axis	Vertical axis	Optimal point
ROC	$R_{\mathrm{FP}} \in [0,1]$	$1 - R_{\mathrm{FN}} = R_{\mathrm{TP}} \in [0,1]$	In theory, $(0,1)$
PRC	$\mathrm{Recall} \in [0,1]$	$\mathrm{Precision} \in [0,1]$	In theory, $(1,1)$
ATC	$m\mathrm{TTI} \in [0,10]$	$F_1 \in [0,1]$	$(\max F_1, m\mathrm{TTI}^*)$

consider a practical optimal point to be the $m\mathrm{TTI}$ when F_1 is the highest, defined as $m\mathrm{TTI}^* = \arg\max_{\mathrm{TTI}} F_1$. Accordingly, we denote the $P_{\mathrm{TTI}\geq 1.5}$ at this point $P^*_{\mathrm{TTI}\geq 1.5}$.

Safety-focused evaluation metrics

Based on the performance curves, we first use AUPRC, the area under the PRC, as a comprehensive metric to assess detection accuracy. Here we do not consider the area under ROC, as a method that has a better-performed PRC (i.e., of larger AUPRC) necessarily has a larger area under ROC [207]. For timeliness evaluation, we use $P^*_{\mathrm{TTI}\geq 1.5}$ and $m\mathrm{TTI}^*$ when F_1 is at the highest point on ATC. The larger the values, the earlier the alerts under the optimal detection accuracy.

Risk alerting of potential collisions is different from normal binary classification problems, as false negatives can cause significantly severe consequences than false positives. To emphasise the high cost of false negatives, we define two safety-focused metrics based on the ROC and PRC, respectively. The first is A^{ROC}_R as defined in Equation (5.15), the area under the ROC curve when $1 - R_{\mathrm{FN}}$ is larger than R (i.e., false negative rate is smaller than $1 - R$). For easier comparison, we normalise the area relative to its maximum possible value $1 - R$. The second is $\mathrm{Precision}^{\mathrm{PRC}}_R$ defined in Equation (5.16), the highest $\mathrm{Precision}$ on the PRC when Recall is larger than R. For both metrics, the larger the value, the more accurate a model is, given a necessarily small false negative rate.

$$A^{\mathrm{ROC}}_R = \frac{1}{1-R}\int_R^1 (1 - R_{\mathrm{FP}}(r))\,\mathrm{d}r, \text{ with } r = 1 - R_{\mathrm{FN}} \tag{5.15}$$

$$\mathrm{Precision}^{\mathrm{PRC}}_R = \max(\mathrm{Precision} \mid \mathrm{Recall} \geq R) \tag{5.16}$$

To offer a complete list, we use AUPRC, $A^{\mathrm{ROC}}_{80\%}$, $A^{\mathrm{ROC}}_{90\%}$, $\mathrm{Precision}^{\mathrm{PRC}}_{80\%}$, $\mathrm{Precision}^{\mathrm{PRC}}_{90\%}$, $P^*_{\mathrm{TTI}\geq 1.5}$, and $m\mathrm{TTI}^*$ as evaluation metrics in this paper. We select the restriction rate R to be 80% and 90% to reinforce a focus on safety. Some methods, if they are not safe enough, can have zero area measured by $A^{\mathrm{ROC}}_{80\%}$ or $A^{\mathrm{ROC}}_{90\%}$, or are not applicable (N/A) when measured by $\mathrm{Precision}^{\mathrm{PRC}}_{80\%}$ or $\mathrm{Precision}^{\mathrm{PRC}}_{90\%}$. For all of the metrics, higher values indicate better performance in collision risk quantification.

5.6 Results

The risk quantification of safety-critical interactions by GSSM is generalised from normal traffic interactions. The training of GSSM is on normal interaction data, of which the

collection varies in locations and equipment. Then the test set contains thousands of crashes and near-crashes where at least 20 seconds are recorded before a safety-critical event happens, offering both positive and negative ground truth. In the experiment on Scalability, we use the first-stage evaluation results of the 4,875 safety-critical events in the test set, and determine a proportion of 10% to add ArgoverseHV data and 100% for highD (see Section 5.6.2 for the results), as already noted in Table 5.3. Then the settlement of ground truth excludes GSSMs trained on other proportions of data combinations. Out of the 4,875 test events, 110 crashes and 2,481 near-crashes eventually had their voted conflicting objects. These 2,591 events form a reliable evaluation basis for the other experiments on Effectiveness (Section 5.6.1), Context-awareness (Section 5.6.3), Generalisability (Section 5.6.4), and Risk attribution (Section 5.6.5), where the third-stage evaluation results are presented.

5.6.1 Effective risk quantification

We use Figure 5.5 and Table 5.6 to show the effectiveness of GSSM in detection accuracy and alert timeliness. Compared with other existing two dimensional surrogate safety measures (2D SSMs), GSSM achieves superior accuracy in detecting crashes and near-crashes, while maintaining relatively stable timeliness of alerts. As is shown by the ROC and PRC in Figure 5.5, the curves of GSSM enclose the curves of all the other 2D SSMs. This means an absolutely higher accuracy, with fewer false positives and false negatives. Regarding alert timeliness, as is seen in Table 5.6, the percentage of early alerts issued 1.5 s in advance and the median time to impact at optimal accuracy (i.e., $P^*_{\mathrm{TTI}\geq 1.5}$ and $m\mathrm{TTI}^*$) of GSSM are slightly lower than those of ACT and TTC2D. But GSSM exhibits a narrower interquartile range and 99% confidence interval, indicating more stable performance across events. The accuracy-timeliness curves in Figure 5.5 also shows that GSSM consistently secures a longer time for collision prevention than the other methods at the same accuracy, as indicated by the F_1 score, higher than 0.8.

Table 5.6 Numeric performance evaluation of GSSM and existing methods. The metric $\mathrm{Precision}_R^{\mathrm{PRC}}$ is abbreviated to $\mathrm{P}_R^{\mathrm{PRC}}$ for a more compact table. For all of the metrics, higher values indicate better performance. The best value in each column is underlined and marked bold; the second best value is marked bold.

Method	AUPRC	$A_{80\%}^{\mathrm{ROC}}$	$A_{90\%}^{\mathrm{ROC}}$	$\mathrm{P}_{80\%}^{\mathrm{PRC}}$	$\mathrm{P}_{90\%}^{\mathrm{PRC}}$	$P^*_{\mathrm{TTI}\geq 1.5}$	$m\mathrm{TTI}^*$ $[Q1, Q3]$; $99\%CI$
GSSM	**0.900**	**0.817**	**0.729**	**0.887**	**0.814**	**0.819**	2.60 [1.80, 3.51]; 2.53–2.65
ACT	0.824	**0.721**	**0.561**	0.850	**0.795**	0.769	**2.78** [1.63, 3.91]; 2.70–2.88
TTC2D	0.818	0.493	0.139	**0.865**	0.674	**0.860**	**2.83** [1.97, 3.83]; 2.75–2.90
TAdv	0.700	0.584	0.414	0.738	0.640	0.606	1.73 [0.99, 2.50]; 1.68–1.78
EI	**0.842**	0.426	0.251	0.706	0.539	0.747	2.26 [1.47, 3.20]; 2.20–2.34

5.6.2 Scalable GSSM training

Without requiring crashes or manual labels of near-crashes for training, GSSM leverages a large amount of normal interaction data. This not only means GSSM learns interaction

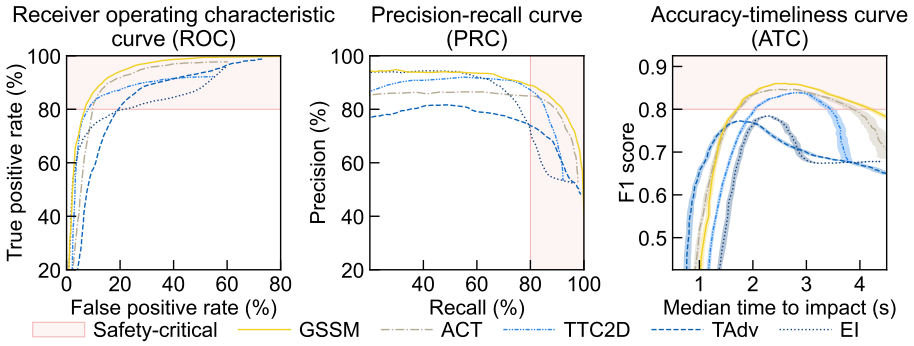


Figure 5.5 Performance curves comparing GSSM and existing methods. In the ATC plot, the shaded bands represent 99% confidence intervals for median time to impact. The GSSM under comparison is trained on the lane changes extracted from the highD dataset and uses instantaneous motion kinematics. For the other methods, TAdv is abbreviated for Time Advantage [23]; ACT for Anticipated Collision Time [18]; TTC2D for Two-dimensional Time-to-Collision [204]; EI for Emergency Index [28].

5

patterns, but also implies effectiveness improvement by feeding more diverse data of traffic interactions. Figure 5.6 illustrates such scalability of GSSM in accurately alerting potential collisions. The GSSMs being compared are all trained with current features, i.e., instantaneous states of road user movement. We use the GSSM trained on SafeBaseline as a benchmark, and increasingly include more data from ArgoverseHV and highD.

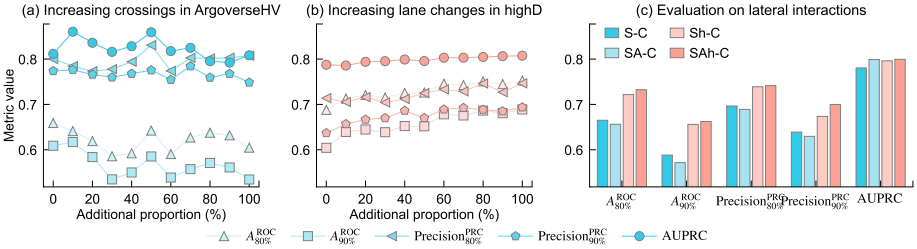


Figure 5.6 Variation in risk alert accuracy of GSSMs trained on different combinations of interaction patterns. **(a)** Accuracy changes when the training data of SafeBaseline are combined with different proportions of ArgoverseHV data, which majorly cover crossing and turning interactions at urban intersections. The accuracy evaluations are performed for event types of Crossing/turning and Pedestrian/cyclist. **(b)** Accuracy changes when the training data of SafeBaseline are combined with different proportions of highD data, which majorly cover lane-change interactions. The accuracy evaluations are performed for event types of Adjacent lane and Merging. **(c)** Accuracy evaluations with events involving lateral interactions. All the GSSMs being compared use current features. The compared models include S-C that is trained on SafeBaseline only, SA-C that combines SafeBaseline and 10% ArgoverseHV, Sh-C that combines SafeBaseline and 100% highD, and SAh-C that combines SafeBaseline, 10% ArgoverseHV, and 100% highD.

Figure 5.6(a) shows the variation in detection accuracy as increasing interactions of crossing and turning in ArgoverseHV are included; while Figure 5.6(b) shows a counterpart for gradually including lane-change interactions in highD. A clear and consistent improvement is seen when more lane changes in highD serve training. As the proportion of ArgoverseHV data increases, however, the accuracy metrics are first enhanced and then drop. This is reasonable because crossing and turning interactions account for around 13.6% in the test set (see Figure 5.3). While these interactions are increasingly included, the training of GSSM gradually shifts its optimisation direction, which may undermine the patterns learnt from the SafeBaseline data.

Based on Figure 5.6(a) and 5.6(b), we consider 10% as the best proportion to include additional data from ArgoverseHV and 100% from highD. Then Figure 5.6(c) compares the enhancement of risk alert accuracy by including more training data in addition to SafeBaseline. It is clearly seen that, to different degrees, incorporating additional lateral interaction data improves the detection accuracy of crashes and near-crashes involving lateral interactions. This implies exciting future work to obtain more powerful GSSMs by training on larger-scale interaction data.

5

5.6.3 Context-aware GSSM

Collision risk is conditioned on interaction context. For example, a usually safe interaction may become unsafe if it is raining. As another example, the same speed could be safe on highways while unsafe in urban traffic. Therefore, being aware of the context where an interaction happens is important, and GSSM allows for that. Figure 5.7 compares all the evaluation metrics of detection accuracy and alert timeliness across GSSMs that consider different contextual information. In each plot specific to a metric, we separate two groups of GSSMs. One group does not consider the acceleration of the subject vehicle as a current feature, while the other group does. Within each group, we progressively include additional features of environmental conditions and historical kinematics.

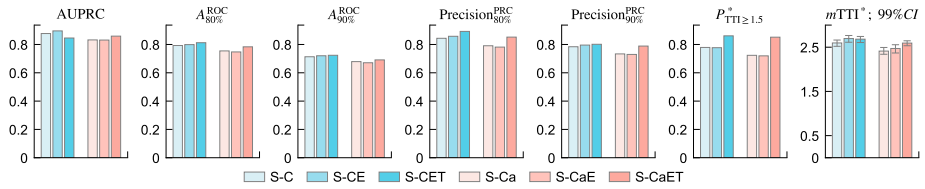


Figure 5.7 Accuracy and timeliness comparison of GSSMs that consider varying contextual information. These GSSMs are all trained on SafeBaseline. S-C uses current features only; S-CE uses current and environment features; S-CET uses current, environment, and historical kinematic features. The other three, i.e., S-Ca, S-CaE, and S-CaET, additionally consider the instantaneous acceleration of the subject vehicle.

The GSSMs using only instantaneous motion kinematics serve as baselines for comparisons within each group. Additionally using environmental conditions in general achieves better performance in terms of both detection accuracy and alert timeliness.

This improvement is further enhanced when historical kinematic features are also used. Interestingly, including historical kinematics improves alert timeliness more significantly than detection accuracy. This suggests that historical kinematics are important to be considered for earlier risk alerts.

Across groups, the GSSMs that exclude and include the subject vehicle's instantaneous acceleration are contrasted. As is seen in Figure 5.7, including acceleration does not make improvements, except for AUPRC when historical kinematics are considered. This implies that the acceleration provides little additional benefit, probably because it is redundant with the historical kinematics, and is also affected by measurement noise and smoothing during trajectory processing.

5.6.4 One GSSM for all interactions

Driven by naturalistic data, GSSM learns interaction patterns across scenarios and effectively alerts to potential collisions in different types of interactions. The crashes and near-crashes in the test set are distributed in various types of events, as shown in Figure 5.3. Accordingly, we compare the performance of GSSM and other existing methods in these different safety-critical events. For better readability, we provide Table 5.7 showing evaluation metrics, while Appendix Figure A2 presents the performance curves. Table 5.7 clearly shows that GSSM is significantly more effective in safety-focused detection accuracy across all interaction scenarios. This is particularly prominent for the conflicts beyond rear-end. Although the timeliness of GSSM is not always the best, it secures at least 1.88 seconds to prevent potential collisions in more than 75% cases.

An interesting observation from Table 5.7 is the superior performance of GSSM in rear-end and merging interactions. These two types of scenarios seem to be less challenging than the other lateral interactions such as crossing and turning. Based on the performance, the most challenging scenarios are interactions with pedestrians, cyclists, and animals. This could be because of the less predictable behaviour of these active and flexible road users, but also could be because the interaction patterns with them are very limitedly covered in the training data used in this paper.

5.6.5 Attribution of collision risk

Figure 5.8 presents the top ranked factors in different situations when GSSM evaluates an interaction to be safe or not, which are respectively considered in safe periods and danger periods as defined in Figure 5.4. At each time moment in a considered period, we calculate the attributions of all features, compare the positive attributions if the considered period is safe or negative attributions otherwise, and record the 3 factors with the largest attributions. Then the ranking of the most contributing factors is obtained after comparing the features at all time moments. Note that because our feature attribution is performed based on the encoded representation rather than raw numbers of variables, the summarised attributions indicate relative importance without implying whether an increasing or decreasing value is associated with the quantified risk.

Table 5.7 Comparison of collision alert performance in different types of safety-critical events. The metric $\mathrm{Precision}_R^{\mathrm{PRC}}$ is abbreviated to $\mathrm{P}_R^{\mathrm{PRC}}$ for a more compact table. The GSSM under comparison is S-CET, which is trained on SafeBaseline and uses contextual information of current features, environment features, and historical kinematic features. For each type of event, the best value in each column is underlined and marked bold; the second-best value is marked bold.

Event	Number of events	Method	AUPRC	$A_{80\%}^{\mathrm{ROC}}$	$A_{90\%}^{\mathrm{ROC}}$	$\mathrm{P}_{80\%}^{\mathrm{PRC}}$	$\mathrm{P}_{90\%}^{\mathrm{PRC}}$	$P^*_{\mathrm{TTI}\geq 1.5}$	$m\mathrm{TTI}^*$ $[Q1, Q3]$; $99\%CI$
Rear-end	1787	GSSM	0.858	**0.926**	**0.887**	**0.952**	**0.930**	**0.877**	**2.53** [1.92, 3.45]; 2.45–2.58
		ACT	**0.869**	**0.852**	**0.786**	0.881	0.877	0.719	2.28 [1.30, 3.09]; 2.19–2.34
		TTC2D	**0.869**	0.702	0.452	**0.932**	**0.906**	**0.856**	**2.49** [1.85, 3.26]; 2.43–2.56
		TAdv	0.726	0.734	0.636	0.798	0.749	0.486	1.44 [0.76, 2.04]; 1.39–1.50
		EI	**0.911**	0.745	0.562	0.930	0.812	0.730	2.10 [1.45, 2.88]; 2.02–2.16
Adjacent lane	611	GSSM	**0.771**	**0.693**	**0.608**	**0.702**	**0.658**	0.826	3.38 [1.88, 5.15]; 3.19–3.61
		ACT	0.685	**0.532**	**0.334**	**0.683**	**0.608**	**0.896**	**4.67** [2.88, 6.14]; 4.40–4.94
		TTC2D	**0.698**	0.317	0.000	0.632	N/A	0.848	**3.68** [2.06, 5.38]; 3.36–3.85
		TAdv	0.588	0.381	0.271	0.512	0.475	0.751	2.38 [1.45, 3.34]; 2.19–2.51
		EI	0.636	0.220	0.042	0.451	0.451	**0.851**	3.48 [1.93, 5.50]; 3.33–3.74
Crossing/turning	93	GSSM	**0.747**	**0.599**	**0.530**	**0.755**	**0.729**	**0.967**	**4.85** [3.63, 6.38]; 4.21–5.46
		ACT	**0.735**	0.350	0.017	**0.792**	**0.677**	0.933	**5.16** [3.48, 6.46]; 4.62–5.40
		TTC2D	0.678	0.000	0.000	N/A	N/A	**0.950**	4.77 [3.21, 6.33]; 4.32–5.31
		TAdv	0.714	**0.366**	**0.182**	0.731	0.649	0.820	2.90 [2.04, 4.07]; 2.39–3.29
		EI	0.694	0.000	0.000	N/A	N/A	0.934	4.13 [2.66, 5.71]; 3.30–4.70
Merging	29	GSSM	**0.865**	**0.795**	**0.719**	**0.828**	**0.771**	0.889	**5.11** [3.23, 7.30]; 3.58–5.60
		ACT	0.705	**0.520**	**0.225**	**0.758**	**0.659**	**1.000**	**5.17** [3.75, 6.80]; 4.52–6.49
		TTC2D	**0.786**	0.000	0.000	N/A	N/A	0.857	3.83 [2.17, 4.85]; 2.66–4.23
		TAdv	0.678	0.251	0.000	0.522	N/A	0.727	2.71 [1.32, 3.87]; 1.52–3.39
		EI	0.777	0.286	0.000	0.553	N/A	**0.897**	3.21 [2.12, 4.85]; 2.52–4.43
With pedestrian/ cyclist/animal	31	GSSM	**0.851**	**0.379**	**0.313**	**0.800**	**0.800**	**1.000**	**5.94** [4.89, 7.16]; 4.89–7.16
		ACT	0.706	0.066	0.000	**0.758**	N/A	**0.966**	**5.38** [3.90, 6.54]; 4.20–6.09
		TTC2D	0.697	0.000	0.000	N/A	N/A	0.897	4.51 [2.31, 5.57]; 2.82–5.14
		TAdv	0.705	**0.090**	**0.005**	0.714	**0.700**	0.806	3.19 [1.76, 5.34]; 2.39–4.62
		EI	**0.733**	0.000	0.000	N/A	N/A	0.867	3.78 [2.48, 5.55]; 2.55–5.36

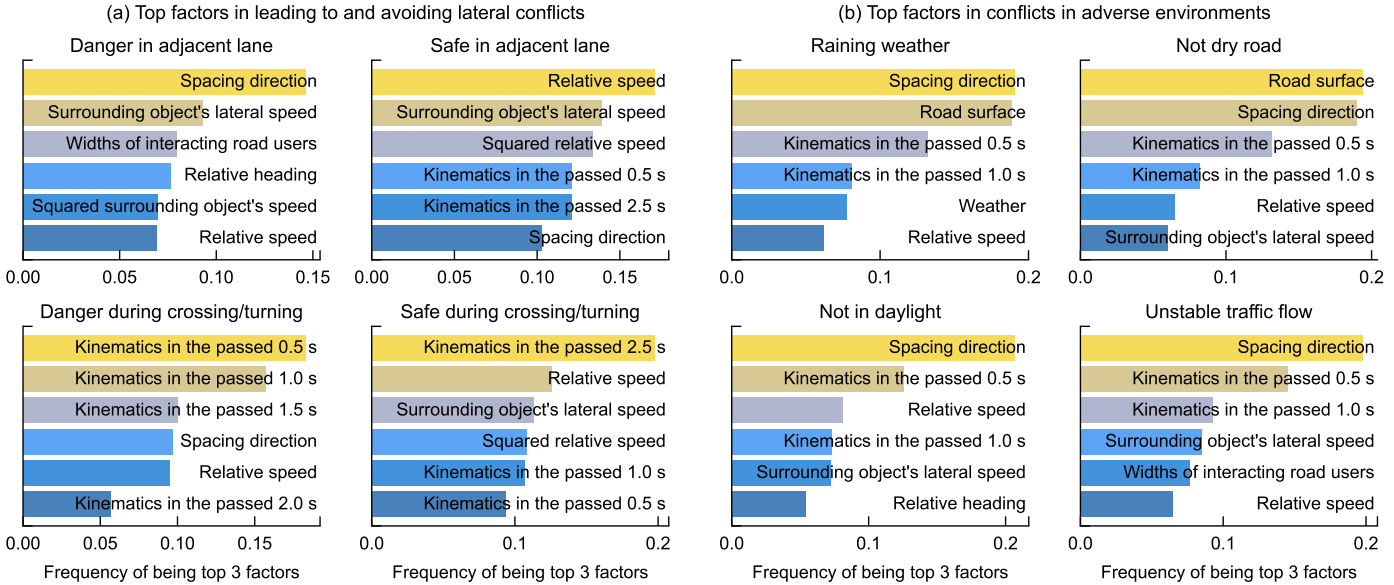


Figure 5.8 Top ranked factors in risk quantification by GSSM. The GSSM used for feature attribution is S-CET, which is trained on SafeBaseline and uses current, environment, and historical kinematic features. The threshold used to distinguish danger and safety is 2.52, which is optimised to achieve the highest $\mathrm{Precision}$ for $\mathrm{Recall} \geq 0.85$. This threshold means an interaction is considered unsafe if its spacing remains minimal for at least $10^{2.52} \approx 331$ times of observations in the same interaction context. **(a)** Top ranked factors in leading to and avoiding lateral crashes and near-crashes. **(b)** Top ranked factors in crashes and near-crashes happened in adverse environments.

Focusing on what leads to potential collisions in lateral interactions, the left half of Figure 5.8(a) shows that spacing direction contributes the most to dangers in the adjacent lane, while the historical kinematics in the past 1.5 seconds contribute the most to dangers during crossing and turning. Spacing direction is related to relative heading and the surrounding object's lateral speed, which are also ranked high. With not much surprise, relative speed also contributes considerably to potential collision. In contrast, the right half of Figure 5.8(a) shows factors contributing to safe interactions. Relative speed and the surrounding object's lateral speed are the most contributing ones. In addition to them, relatively longer historical kinematics are important. These results suggest that relative direction and speed are the main factors for accurate risk quantification. Notably, a short movement history of approximately 1.5 seconds is more influential in detecting danger, whereas a longer history of around 2.5 seconds provides greater context for perceiving safety in lateral interactions.

Adverse environments such as precipitation, wet road surfaces, driving at night, or unstable traffic flow are relatively infrequent conditions, but may significantly increase the risk of collisions. In Figure 5.8(b), we present the factors contributing to dangers in different adverse environments. The top factors harming safety overlap with rain and when the road surface is wet. In both conditions, spacing direction and road surface are ranked as the most important. This implies that the collision risk in such situations is highly influenced by the directional control of the involved road users. Spacing direction remains highly ranked when considering interactions in the dark as well as in unstable traffic flow. The other factors in these two adverse environments resemble those in general interactions in Figure 5.8(a). Interestingly, the widths of interacting road users, i.e., half of the sum of the vehicle widths, play a role in collision risk when the traffic flow is unstable. This could be due to relatively restrained road space for safe interactions.

5.7 Conclusion and discussion

We present the generalised surrogate safety measure (GSSM) to proactively quantify the risk of potential traffic collisions before they happen. GSSM intends to address the challenges of scalability, context-awareness, and generalisability, faced by existing approaches to proactive collision risk quantification. Below, we discuss the main findings.

- Instead of relying on historical records of crashes or near-crashes, GSSM stably learns from the patterns of normal interactions and extrapolates them to safety-critical situations. This allows for ***scalable*** improvement with increasing amounts of data, as are being collected by automated vehicles, and our results show enhanced accuracy in risk quantification by additional lateral interaction patterns.
- The fundamental assumption made by GSSM is that collision risk emerges if interactions become extreme, which is measured by the spacing between road users involved in a specific interaction context. GSSM utilises neural networks to approximate such context-conditioned distributions of multi-directional spacing and is data-driven. Its risk quantification is thus ***context-aware*** by incorporating any

potentially helpful information as inputs, e.g., instantaneous motion states, weather, road surface condition, lighting, and historical kinematics, as we have illustrated.

- Our results also demonstrate that GSSM is highly ***generalisable***. It achieves superior accuracy and timeliness in alerting potential collisions across various interaction scenarios such as rear-end, merging, crossing, turning, etc. Impressively but reasonably, GSSM also generalises well from training data to unseen test data, even when the training data are collected in different countries and by different equipment. This implies that the interaction patterns in urgent conditions, such as crashes and near-crashes, are shared in human behaviour.
- The attribution of collision risk shows that GSSM correctly utilises the information of weather and road surface conditions when it is rainy or the road is not dry. Importantly, ***spacing direction*** has a dominant influence in lateral interactions, where short motion histories of around 1.5 seconds help to recognise risk and longer motion histories of 2.5 seconds are more useful to confirm safety. These findings emphasise the necessity for context-aware collision risk quantification to consider diverse factors.

GSSM is limited by our modelling choices and requires further development and targeted validation in the future. First of all, this study focuses on the probability of potential collisions while omitting to consider collision severity or time of occurrence, which are important aspects of collision risk. Driven by real-world data, GSSM's robustness to sensor noise, missing values, and small perturbations in the input is not guaranteed, yet our training does not provide formal guarantees or certified bounds. The datasets used in this research are carefully processed. To verify GSSM's stability and transferability, especially when extrapolating to extreme behaviours, a larger variety of interaction data with multi-source noise and from more different driving cultures will be essential. In addition, richer modalities such as vision and language can be incorporated to characterise more informative contexts. This motivates representation learning to encode complex context and enable large-scale training in industrial applications. To truly improve traffic safety, it is also important to investigate not only correlations but also causal mechanisms between contextual features and the estimated risk. We regard GSSM as a starting point toward foundation models for proactive risk quantification of potential collisions. Before achieving that goal, more real-world training and verification are required to confirm its stable effectiveness, quantifiable scalability, explainable context-awareness, and controllable generalisability.

The implications of such a methodology are broad for both ex ante prevention and ex post analysis of traffic collision risk. The most direct benefit is for autonomous driving systems (ADS), given that GSSM can compute for a thousand interactions per query within 25 milliseconds (see Section 2.4 in Supplementary Material). Online, GSSM can provide a time-varying risk signal that serves as a safety cost or constraint in motion planning, or as a safety "shield" that warns of potentially unsafe behaviours before they are executed. Offline, it can be used to train and validate ADS in safety-critical situations, e.g., by filtering valuable training materials, generating or ranking validation scenarios at varying levels of collision risk, and supporting closed-loop testing in which high-risk

interactions are selectively evaluated. Beyond ADS, GSSM enables civil engineering and traffic management use cases: it allows for proactive evaluation and thus data-driven improvement in the safety of road designs, operations, and policies. For example, GSSM can help traffic engineers and policy makers detect emerging hazards in road networks and mitigate potential accidents before they occur. Altogether, these applications contribute to a significant step toward the long-term vision of zero traffic fatalities, via a tool that quantifies and manages risk proactively, in any context, at any time.

Appendix A

A.1 SHRP2 trajectory reconstruction error

Appendix Figure A1 shows the error distributions of bird's eye view reconstruction of the trajectories derived in the Second Strategic Highway Research Program's (SHRP2) Naturalistic Driving Study (NDS). These include the root mean squared error (RMSE) in subject speed, subject yaw rate, subject acceleration, and object speed, as well as the mean absolute error (MAE) of object displacement. The errors are evaluated for each event in the categories of crashes, near-crashes, and safe baselines.

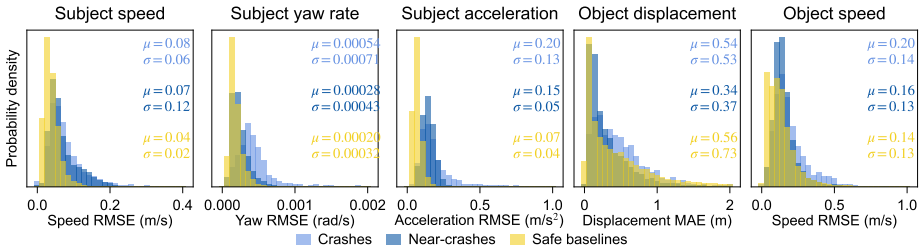


Figure A1 Reconstruction error distributions of crashes, near crashes, and safe baselines in the SHRP2 NDS. The mean values (μ) and standard deviations (σ) are marked in each plot.

A.2 Contextual information details

As explained in Section 5.3.2, this study considers 3 categories of contextual information: current features (X_C) including the instantaneous states of interacting road users, categorical environment features (X_E) of external conditions during the interaction, and historical kinematic features (X_T) composed of time-series speeds and yaw rates within the past 2.5 seconds. Appendix Table A1 provides a detailed list of these contextual features used in this study, as well as their encoding unit.

A.3 Neural network architecture

Out of the aim of validating the generalised surrogate safety measure (GSSM) rather than looking for the most powerful architecture, our neural networks are relatively simple for computational convenience. As mentioned in Section 5.3.2, we have 3 encoders and 1 decoder. The encoders for current features (X_C) and environment features (X_E) are multi-layer perceptrons (MLPs, [208]), while the encoder for historical kinematic features (X_T) uses a single-layer long short-term memory (LSTM, [209]) recurrent neural network. The encoded representation groups are then concatenated and passed to the decoder, which uses the attention mechanism [210] to capture inter-feature relations and convolutional neural networks (CNN, [211]) for intra-feature relations. Throughout the model, we use the activation function of Gaussian error linear units (GELU, [212]). More descriptive details are as follows, while implementation details are referred to in our open-sourced repository.

Table A1 Contextual information considered in this study. Definitions of all variable symbols are given in the table footnote.

Category	Attribution unit	Feature(s)
X_C	Per each feature	$\{l_i, l_j, \frac{w_i+w_j}{2}, \lvert\boldsymbol{v}_i\rvert, x_{\boldsymbol{v}_j}, y_{\boldsymbol{v}_j}, \lvert\boldsymbol{v}_i\rvert^2, \lvert\boldsymbol{v}_j\rvert^2, \lvert\boldsymbol{v}_{ij}\rvert^2, \lvert\boldsymbol{v}_{ij}\rvert\mathrm{sgn}(\Delta v), A_{\boldsymbol{h}_j}, \rho_{ij}\}_{t=0}$, optional $a_{i,t=0}$
X_E	Lighting condition	Darkeness (lighted), darkness (not lighted), dawn, daylight, dusk, unknown.
	Weather condition	No adverse conditions, fog, mist/light rain, rain and fog, raining, sleeting, snow/sleet and fog, snowing, unknown.
	Road surface	Dry, gravel over asphalt, gravel/dirt road, icy, muddy, snowy, wet, unknown.
	Traffic density	LOS A1: free flow, no lead traffic, LOS A2: free flow, leading traffic present, LOS B: flow with some restrictions, LOS C: stable flow, maneuverability and speed are more restricted, LOS D: unstable flow - temporary restrictions substantially slow driver, LOS E: flow is unstable, vehicles are unable to pass, temporary stoppages, etc., LOS F: forced traffic flow with low speeds and traffic volumes that are below capacity, unknown.
X_T	In the passed 0.5 s	$\{\omega_i, \lvert\boldsymbol{v}_i\rvert, x_{\boldsymbol{v}_j}, y_{\boldsymbol{v}_j}\}_{t\in\{-0.5,-0.4,\ldots,-0.1\}}$
	In the passed 1 s	$\{\omega_i, \lvert\boldsymbol{v}_i\rvert, x_{\boldsymbol{v}_j}, y_{\boldsymbol{v}_j}\}_{t\in\{-1.0,-0.9,\ldots,-0.1\}}$
	In the passed 1.5 s	$\{\omega_i, \lvert\boldsymbol{v}_i\rvert, x_{\boldsymbol{v}_j}, y_{\boldsymbol{v}_j}\}_{t\in\{-1.5,-1.4,\ldots,-0.1\}}$
	In the passed 2 s	$\{\omega_i, \lvert\boldsymbol{v}_i\rvert, x_{\boldsymbol{v}_j}, y_{\boldsymbol{v}_j}\}_{t\in\{-2.0,-1.9,\ldots,-0.1\}}$
	In the passed 2.5 s	$\{\omega_i, \lvert\boldsymbol{v}_i\rvert, x_{\boldsymbol{v}_j}, y_{\boldsymbol{v}_j}\}_{t\in\{-2.5,-2.4,\ldots,-0.1\}}$

Note: for two interacting road users i and j, their yaw rates, velocities, lengths, widths, and headings are denoted by $\omega \in [-\pi, \pi]$, $\boldsymbol{v}$, l, w, and $\boldsymbol{h}$; the motion states are in a local coordinate system with the origin at the position of i and y-axis oriented along the velocity direction of i; $\mathrm{sgn}(\Delta v) = \mathrm{sgn}(\lvert\boldsymbol{v}_i\rvert - \lvert\boldsymbol{v}_j\rvert)$; $A_{\boldsymbol{h}} \in [-\pi, \pi]$ denotes the angle between the heading direction and the y-axis; $\rho_{ij} \in [-\pi, \pi]$ is the angle in the polar coordinates of multi-directional spacing.

To ensure that feature attributions are meaningful, we design the encoders to let each attribution unit (a representation vector) carry independent information from the others. As shown in Appendix Table A1, each scalar feature in X_C is mapped to its own representation vector. The categorical features in X_E are first one-hot encoded and grouped into four chunks (weather, lighting, road surface, and traffic density), and each chunk is then mapped to a separate representation vector. For X_C and X_E, we use MLP encoders with 5 and 4 linear layers, respectively. The temporal features X_T are encoded into five representation vectors corresponding to the past 0.5, 1, 1.5, 2, and 2.5 seconds. Concretely, we reverse the time steps in the 2.5-second history and feed the reversed sequence into a single-layer unidirectional LSTM. From the resulting output sequence, we take five representations at 0.5-second intervals, so that each one encodes the kinematic features within 0.5, 1, 1.5, 2, and 2.5 seconds closest to the current moment.

To enrich and regularise the latent representation, we append a set of orthogonal Gaussian random features [213] which are deterministic and unique for a sequence of representation pieces. Seeing the encoded and enriched pieces as tokens of different contextual information, we design the decoder to capture both global and local relations of the tokens. We first use batch normalisation to align the ranges of all feature dimensions, and then stack 6 self-attention blocks, in each of which the feed-forward network has 2 linear layers. Next, we apply a CNN with 2 layers of convolution and a kernel size of 3. Lastly, we use two separate 3-layer MLPs to output the targeted parameters μ and $\log(\sigma^2)$.

A.4 Hyperparameter settings

The performance of GSSM depends on a number of implementation choices such as learning-rate schedules, batch sizes, and regularisation strengths. To ensure transparency and reproducibility, we report a full set of hyperparameters used in the experiments in Appendix Table A2.

Table A2 Hyperparameters set in this paper's experiments.

Hyperparameter	Definition	Value
Random seed	Number that controls generation of pseudo-random values	131
Small value threshold	Threshold to consider a value smaller than it to be near zero	1e-6
Representation dimension	Number of variables in a piece of encoded representation	64
Perturbation noise	Standard deviation of the added Gaussian noise in Section 5.3.1	1% of variable range
β	Weight of JS-divergence in Equation (5.6)	5
Max. epochs	Maximum epochs of training	150
Batch size	Number of samples used per update during model training	512
Initial learning rate	Learning rate at the beginning of model training	1e-4
Dropout	Probability for each value in a tensor to be set 0	0.2

A.5 Computation efficiency

As reported in Appendix Table A3, the GSSMs have 773,286 parameters if the input context has X_C only; when both X_C and X_E are considered, there are 857,486 parameters; if X_T is also included, the number of parameters becomes 885,664. The dominant part of

the parameters is of the decoder, correspondingly accounting for 95.92%, 87.46%, and 85.83% of all parameters in a model. Such a model with fewer than 0.9 million parameters is fairly small for modern deep learning.

To quantify the computational cost, we further report the training time and inference time for the different GSSM variants alongside other existing methods for comparison. Note that we implement TTC2D, ACT and TAdv in a vectorised form for efficient large-scale evaluation. In contrast, EI is used as provided by its authors, as its formulation does not readily admit the same level of vectorisation. Altogether, these results show that the proposed GSSM achieves competitive computational efficiency while providing substantially richer, context-aware risk estimates.

Table A3 Computation efficiency comparison of GSSMs and existing methods.

Acronym [a b]	**Contextual observables**	**Number of parameters**	**Training time per sample (s)**	**Inference time per sample (s)**
TTC2D				6.21E-7
ACT		N/A	N/A	3.93E-6
TAdv				4.78E-6
EI				5.05E-4
S-C	$[X_C]$	773,286	4.43E-5	1.37E-5
h-C		773,286	4.39E-5	1.07E-5
A-C		773,286	4.40E-5	1.34E-5
SAh-C		773,286	4.53E-5	1.19E-5
S-Ca	$[X_C, a_i]$	775,336	5.04E-5	1.33E-5
S-CE	$[X_C, X_E]$	857,486	6.42E-5	1.78E-5
S-CaE	$[X_C, a_i, X_E]$	859,536	6.62E-5	1.88E-5
S-CET	$[X_C, X_E, X_T]$	885,664	7.48E-5	2.21E-5
S-CaET	$[X_C, a_i, X_E, X_T]$	887,714	7.50E-5	2.35E-5

[a]The two-dimensional surrogate safety measures have abbreviations TTC2D for Two-dimensional Time-to-Collision [204], TAdv for Time Advantage [23], ACT for Anticipated Collision Time [18], and EI for Emergency Index [28].

[b]Acronyms for GSSMs follow the format `<Datasets>-<Contextual observables>`. For example, "S-C" is trained on SafeBaseline and uses current features X_C

A.6 Performance curves in different interactions

Corresponding to Appendix Table 5.7, Appendix Figure A2 shows the performance curves to compare GSSM and other two-dimensional surrogate safety measures (2D SSMs) in different types of interactions. The curves are less smooth for interactions of merging and with pedestrian/cyclist/animal because of the small number of test events.

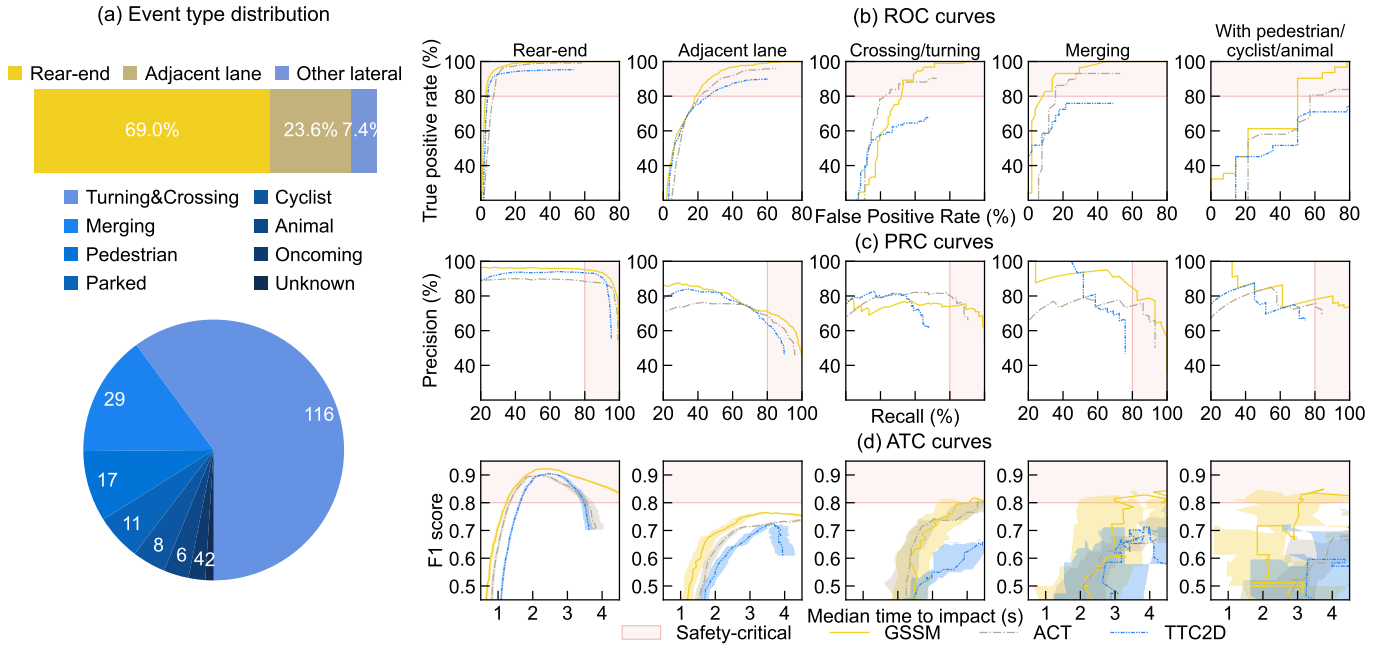


Figure A2 Performance comparison of GSSM and other existing methods in alerting different types of safety-critical events. **(a)** Types of the crashes and near-crashes with determined ground truth. **(b)** Receiver operating characteristic curves for different types of events. **(c)** Precision-recall curves for different types of events. **(d)** Accuracy-timeliness curves for different types of events, where the shaded bands represent 99% confidence intervals for median time to impact. The GSSM under comparison is trained on the SafeBaseline data and uses contextual information of instantaneous motion kinematics, environmental conditions, and historical kinematics in the past 2.5 seconds.

Chapter 6

Spatial-temporal information preservation of traffic interactions

Highlights

- Regularisers at two scales are introduced for contrastive learning of spatial time series.
- The regularisers preserve fine-grained similarity structures across time or instances.
- A dynamic mechanism balances contrastive learning and structure preservation.
- State-of-the-arts in spatial time series classification and multi-scale traffic prediction.
- Better preservation of similarity structures implies more informative representations.

Keywords

Contrastive learning, representation learning, time series, spatio-temporal data, traffic interaction

This chapter is based on the journal article: Yiru Jiao, Sander van Cranenburgh, Simeon C. Calvert, and Hans van Lint. (2025). Structure-preserving contrastive learning for spatial time series. Artificial Intelligence for Transportation, 3–4, 100031. doi: 10.1016/j.ait.2025.100031

Abstract

Neural network models are increasingly applied in transportation research to tasks such as prediction. The effectiveness of these models largely relies on learning meaningful latent patterns from data, where self-supervised learning of informative representations can enhance model performance and generalisability. However, self-supervised representation learning for spatially characterised time series, which are ubiquitous in transportation domain, poses unique challenges due to the necessity of maintaining fine-grained spatio-temporal similarities in the latent space. In this study, we introduce two structure-preserving regularisers for the contrastive learning of spatial time series: one regulariser preserves the topology of similarities between instances, and the other preserves the graph geometry of similarities across spatial and temporal dimensions. To balance the contrastive learning objective and the need for structure preservation, we propose a dynamic weighting mechanism that adaptively manages this trade-off and stabilises training. We validate the proposed method through extensive experiments, including multivariate time series classification to demonstrate its general applicability, as well as macroscopic and microscopic traffic prediction to highlight its particular usefulness in encoding traffic interactions. Across all tasks, our method preserves the similarity structures more effectively and improves state-of-the-art task performances. This method can be integrated with an arbitrary neural network model and is particularly beneficial for time series data with spatial or geographical features. Furthermore, our findings suggest that well-preserved similarity structures in the latent space indicate more informative and useful representations. This provides insights to design and optimise more effective neural networks for data-driven transportation research.

Code availability

`https://github.com/Yiru-Jiao/SPCLT`

Data availability

Raw data source:

- UEA archive `https://www.timeseriesclassification.com/dataset.php`
- Macroscopic traffic `https://github.com/RomainLITUD/uncertainty-aware-traffic-speed-flow-demand-prediction`
- Microscopic traffic `https://interaction-dataset.com`

Resulting data: `https://doi.org/10.4121/3b8cf098-c2ce-49b1-8e36-74b37872aaa6`

6.1 Introduction

Modern transportation systems create massive streams of spatially distributed time series data, such as traffic speeds across road networks and transit ridership through various locations over time. Extracting useful patterns from these data is crucial, especially as neural network models are increasingly used for a range of downstream tasks such as traffic forecasting [214, 215], congestion detection [216, 217], and mobility analysis [218, 219]. In recent years, self-supervised representation learning (SSRL) has emerged as a promising approach to leverage such large-scale datasets [220]. By learning informative latent representations, SSRL can effectively facilitate model performance [221–223] and generalisability [224, 225] in downstream tasks. This advantage is especially valuable in transportation research, where real-world sensor measurements and labels are often noisy or sparse.

In SSRL of time series, contrastive learning is becoming the mainstay technique. This adoption is supported by empirical investigation. In 2022, Lafabregue et al. [226] conducted an extensive experimental comparison over 300 combinations of network architectures and loss functions to evaluate the performance of time series representation learning. One of their key findings is that the reconstruction loss used by traditional autoencoders does not sufficiently fit temporal patterns. Instead, contrastive learning has emerged as a more effective approach, which explicitly pulls similar instances closer and pushes dissimilar instances farther apart in the latent space of representations [227, 228]. This mechanism encourages neural networks to organise the latent space according to the inherent similarities in data, yielding representations that capture meaningful patterns.

Unique challenges arise when learning contrastive representations for spatially characterised time series data. A foremost difficulty is the need to preserve fine-grained similarity structures among data instances in the latent space. The notion of similarity for spatial time series can be subtle and highly domain-specific. For example, financial time series may be considered similar even if some variables show significant divergence, while movement traces with very different spatial features can be anything but similar. Beyond preserving fine-grained similarities, spatially characterised time series such as traffic interactions can involve multiple scales of spatio-temporal patterns. At the macroscopic scale, traffic flow measures collective road usage evolving over the road network; at the microscopic scale, trajectories describe the motion dynamics of individual road users such as car drivers, cyclists, and pedestrians, in local road space. SSRL for spatial time series must accommodate such heterogeneity, capturing patterns at the appropriate level of granularity for the targeted task.

To address these challenges, this study explores contrastive learning regularised by structure preservation to better capture the subtle similarities in spatial time series data. We introduce two regularisers at different scales to preserve the original similarity structure in the latent space. One is a topology-preserving regulariser for the global scale, and the other is a graph-geometry-preserving regulariser for the local scale. This incorporation can be simplified as a weighted loss $\mathcal{L} = \eta_{\text{CLT}} \cdot \ell_{\text{CLT}} + \eta_{\text{SP}} \cdot \ell_{\text{SP}} + r_{\boldsymbol{\eta}}$, where we propose a mechanism to dynamically balance the weights η_{CLT} of contrastive learning

for time series (CLT) and η_{SP} of structure preservation (SP). Within this mechanism, the adaptive trade-off between contrastive learning and structure preservation is based on the uncertainties of their corresponding terms ℓ_{CLT} and ℓ_{SP}; meanwhile, the term $r_{\boldsymbol{\eta}}$ adds regularisation against overfitting of the dynamic weights.

The proposed method is applicable to spatial time series in general, while we highlight its particular usage for traffic interactions in this paper. To thoroughly validate the method, we conduct experiments on tasks of 1) multivariate time series classification, where we benchmark against the current state-of-the-art (SOTA) models, i.e., [229] and [230]; and 2) traffic prediction, where we use [231] for macroscopic benchmark and [232] for microscopic. Along with these experiments, the efficiency of this method is evaluated with multiple network architectures. In addition to performance improvement, we also investigate the impacts of preserving similarity structure during training. Below is a summary of the contributions in this study.

- We introduce a method that incorporates structure-preserving regularisation in contrastive learning of multivariate time series, to maintain finer-grained similarity structures in the latent space of sample representations. We propose a dynamic weighting mechanism to adaptively balance contrastive learning and structure preservation during training. This method can be applied to an arbitrary neural network model for more effective representation learning.
- Preserving similarity structure can enhance SOTA performance on various downstream tasks. The relative improvement on spatial datasets in the UEA archive is 2.96% in average classification accuracy; on macroscopic traffic prediction task is 0.57% in flow speed MAE and 0.55% in the standard deviation of prediction errors; on microscopic trajectory prediction task is 1.87% and 3.40% in missing rates under radii of 0.5 m and 1 m, respectively.
- Considering neural network modelling as learning the conditional probability distribution of outputs over inputs, the similarity structure hidden in the input data implies the distribution of conditions. Our method is therefore important to preserve the original distribution in the latent space for more effective model training. This is particularly beneficial when dealing with spatial time series data in transportation domain, where fine-grained and hierarchical information is required in modelling.

The rest of this paper is organised as follows. In Section 6.2, we briefly review related work in the literature. We use Section 6.3 to systematically introduce the methods. Then we explain the demonstration experiments in Section 6.4 and present according results in Section 6.5. With Section 6.6, we discuss the importance of preserving similarity structure for representation learning. Finally, Section 6.7 concludes this study.

6.2 Related work

6.2.1 Time series contrastive learning

Contrastive learning for time series data is a relatively young niche and is rapidly developing. The development has been dominantly focused on defining positive and

negative samples. Early approaches construct positive and negative samples with subseries within time series [233] and temporal neighbourhoods [234]; and later methods create augmentations by transforming original series [235, 236]. More recently, [229] generates random masks to enable both instance-wise and time-wise contextual representations at flexible hierarchical levels, which exceeds previous state-of-the-art performances (SOTAs). Given that not all negatives may be useful [237, 238], [239] makes hard negatives to boost performance, while [230] utilises soft contrastive learning to weigh sample pairs of varying similarities, both of which reach new SOTAs.

The preceding paragraph outlines a brief summary, and we refer the readers to Section 2 in [230] and Section 5.3 in [240] for a detailed overview of the methods proposed in the past 6 years. These advances have led to increasingly sophisticated methods that mine the contextual information embedded in time series by contrasting similarities. However, the structural details of similarity relations between samples remain to be explored.

6.2.2 Structure-preserving SSRL

Preserving the original structure of data when mapping into a latent space has been widely and actively researched in manifold learning (for a literature review, see [241]) and graph representation learning [242, 243]. In manifold learning, which is also known as nonlinear dimension reduction, the focus is on revealing the geometric shape of data point clouds for visualisation, denoising, and interpretation. In graph representation learning, the focus is on maintaining the connectivity of nodes in the graph while compressing the data space required for large-scale graphs [244]. Structure-preserving has not yet attracted much dedication to time series data. [245] provides a literature review on time series data dimensionality reduction, where none of the methods are specifically tailored for time series. Existing studies that are the most relevant include [246–248], which construct hierarchies of samples or features, while similarity preservation remains under-explored.

Zooming in within structure-preserving SSRL, there are two major branches respectively focusing on topology and geometry. Topology-preserving SSRL aims to maintain global properties such as clusters, loops, and voids in the latent space; representative models include [249] and [250] using autoencoders, as well as [251] and [252] with contrastive learning. The other branch is geometry-preserving and focuses more on local shapes such as relative distances, angles, and areas. Geometry-preserving autoencoders include [253] and [254], while [255] and [256] use contrastive learning. The aforementioned topology and geometry preserving autoencoders are all developed for dimensionality reduction; whereas the combination of contrastive learning and structure-preserving has been explored majorly with graphs.

6.2.3 Traffic interaction SSRL

In line with the literature summary in previous sub-sections, existing exploration of SSRL in the context of traffic interaction data and tasks have been predominantly relied on autoencoders and graph-based contrastive learning. For instance, using a transformer-based multivariate time series autoencoder [257], [258] clusters traffic

scenarios with trajectories of pairwise vehicles. Then a series of studies investigate masking strategies with autoencoders for individual trajectories and road networks, including [259–261]. Combining graph (convolutional) neural networks and contrastive learning, a variety of studies have shown accuracy and stability improvements in traffic flow prediction [262–265].

There are some other use cases. Leveraging data augmentation, [266] utilises graphs and contrastive learning to jointly learn representations for vehicle trajectories and road networks. The authors design road segment positive samples as neighbours in the graph, and trajectory positive samples by replacing a random part with another path having the same origin and destination. In a similar way, [267] learns traffic scene similarity. The authors randomly modify the position and velocity of individual traffic participants in a scene to construct positive samples, with negative samples drawn uniformly from the rest of a training batch. By designing augmentation based on domain-knowledge, [268] focuses on capturing seasonal and holiday information for traffic prediction, while [269] targets abnormal traffic patterns caused by incidents or reduced capacities.

6.3 Methods

6.3.1 Problem definition

We define the problem for general spatial time series, with traffic interaction as a specific case. Learning the representations of a set of samples $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_N\}$ aims to obtain a nonlinear function $f_{\boldsymbol{\theta}} : \boldsymbol{x} \to \boldsymbol{z}$ that encodes each $\boldsymbol{x}$ into $\boldsymbol{z}$ in a latent space. Let T denote the sequence length of a time series and D the feature dimension at each timestamp t. The original space of $\boldsymbol{x}$ can have the form $\mathbb{R}^{T\times D}$, where spatial features are among the D dimensions; or $\mathbb{R}^{T\times S\times D}$, where S represents spatially distributed objects (e.g., sensors or road users). The latent space of $\boldsymbol{z}$ can also be structured in different forms, such as $\mathbb{R}^{P}$, $\mathbb{R}^{T\times P}$, or $\mathbb{R}^{T\times S\times P}$, where P is the dimension of encoded features.

By contrastive learning, (dis)similar samples in the original space should remain close (far) in the latent space. Meanwhile, by structure preservation, the distance/similarity relations between samples should maintain certain features after mapping into the latent space. We use $d(\boldsymbol{x}_i, \boldsymbol{x}_j)$ to denote the distance between two samples i and j, and this also applies to their encoded representations $\boldsymbol{z}_i$ and $\boldsymbol{z}_j$. Various distance measures can be used to define d, such as cosine distance (COS), Euclidean distance (EUC), and dynamic time warping (DTW). The smaller the distance between two samples, the more similar they are. Considering the limitation of storage efficiency, similarity comparison is performed in each mini-batch, where B samples are randomly selected.

6.3.2 SPCLT loss

Equation (6.1) presents an overview of the structure-preserving contrastive learning loss for time series, abbreviated as SPCLT loss, to optimise $f_{\boldsymbol{\theta}}$ for self-supervised representation learning.

$$\mathcal{L} = \frac{1}{2\sigma_{\text{CLT}}^2}\mathcal{L}_{\text{CLT}}\left(1-\exp(-\mathcal{L}_{\text{CLT}})\right) + \frac{1}{2\sigma_{\text{SP}}^2}\mathcal{L}_{\text{SP}}\left(1-\exp(-\mathcal{L}_{\text{SP}})\right) + \log\sigma_{\text{CLT}}\sigma_{\text{SP}} \tag{6.1}$$

Referring to the simplified loss in Section 6.1, i.e., $\mathcal{L} = \eta_{\text{CLT}} \cdot \ell_{\text{CLT}} + \eta_{\text{SP}} \cdot \ell_{\text{SP}} + r_{\boldsymbol{\eta}}$, the contrastive learning loss for time series ($\mathcal{L}_{\text{CLT}}$) and structure-preserving loss ($\mathcal{L}_{\text{SP}}$) are modified using the function $y = x(1-\exp(-x))$ and correspond to ℓ_{CLT} and ℓ_{SP}. This modification is designed to stabilise $\mathcal{L}_{\text{CLT}}$ and $\mathcal{L}_{\text{SP}}$ when they are close to 0. The terms η_{CLT}, η_{SP}, and $r_{\boldsymbol{\eta}}$ controls the trade-off between contrastive learning and structure preservation, depending on two deviations σ_{CLT} and σ_{SP} that dynamically change during training.

In the following sub-sections, we will first introduce the component losses for time series contrastive learning and structure preservation, and then provide more detailed explanations on their stabilisation and dynamic trade-offs.

6.3.3 Contrastive learning for time series

In this study, we use the time series contrastive learning loss introduced in TS2Vec [229] and its succeeder SoftCLT [230] that utilises soft weights for similarity comparison[1]. For each sample $\boldsymbol{x}_i$, two augmentations are created by timestamp masking and random cropping, and then encoded as two representations $\boldsymbol{z}'_i$ and $\boldsymbol{z}''_i$. TS2Vec and SoftCLT losses consider the same sum of similarities for a sample i at a timestamp t, as shown in Equations (6.2) and (6.3). Equation (6.2) is used for instance-wise contrasting, which we denote by the subscript $_{\text{inst}}$; Equation (6.3) is used for time-wise contrasting, denoted by the subscript $_{\text{temp}}$.

$$S_{\text{inst}}(i,t) = \sum_{j=1}^{B}\left(\exp(\boldsymbol{z}'_{i,t}\cdot\boldsymbol{z}''_{j,t}) + \exp(\boldsymbol{z}''_{i,t}\cdot\boldsymbol{z}'_{j,t})\right) + \sum_{\substack{j=1\\j\neq i}}^{B}\left(\exp(\boldsymbol{z}'_{i,t}\cdot\boldsymbol{z}'_{j,t}) + \exp(\boldsymbol{z}''_{i,t}\cdot\boldsymbol{z}''_{j,t})\right) \tag{6.2}$$

$$S_{\text{temp}}(i,t) = \sum_{s=1}^{T}\left(\exp(\boldsymbol{z}'_{i,t}\cdot\boldsymbol{z}''_{i,s}) + \exp(\boldsymbol{z}''_{i,t}\cdot\boldsymbol{z}'_{i,s})\right) + \sum_{\substack{s=1\\s\neq t}}^{T}\left(\exp(\boldsymbol{z}'_{i,t}\cdot\boldsymbol{z}'_{i,s}) + \exp(\boldsymbol{z}''_{i,t}\cdot\boldsymbol{z}''_{i,s})\right) \tag{6.3}$$

[1] We unify the loss function equations in a consistent format following the open-source code provided with the original papers; as such, they are slightly adjusted from the equations in the original papers.

Equation (6.4) then shows the TS2Vec loss. We refer the readers to [229] for more details about the hierarchical contrasting method.

$$\mathcal{L}_{\text{TS2Vec}} = \frac{1}{NT}\sum_i\sum_t\left(\ell^{(i,t)}_{\substack{\text{inst}\\\text{TS2Vec}}} + \ell^{(i,t)}_{\substack{\text{temp}\\\text{TS2Vec}}}\right),$$
$$\text{where}\begin{cases}\ell^{(i,t)}_{\substack{\text{inst}\\\text{TS2Vec}}} = -\log\dfrac{\exp(\boldsymbol{z}'_{i,t}\cdot\boldsymbol{z}''_{i,t})+\exp(\boldsymbol{z}''_{i,t}\cdot\boldsymbol{z}'_{i,t})}{S_{\text{inst}}(i,t)}\\ \ell^{(i,t)}_{\substack{\text{temp}\\\text{TS2Vec}}} = -\log\dfrac{\exp(\boldsymbol{z}'_{i,t}\cdot\boldsymbol{z}''_{i,t})+\exp(\boldsymbol{z}''_{i,t}\cdot\boldsymbol{z}'_{i,t})}{S_{\text{temp}}(i,t)}\end{cases} \tag{6.4}$$

Similarity comparison in TS2Vec is between two different augmentations for the same sample. This is expanded by SoftCLT to also involve other samples in the same mini-batch. Varying instance-wise and time-wise weights are assigned to different comparison pairs as soft assignments, with Equations (6.5) and (6.6). This introduces four hyperparameters, i.e., τ_{inst}, τ_{temp}, α, and m. We use EUC to compute $d(\boldsymbol{x}_i, \boldsymbol{x}_j)$ throughout this paper and set $\alpha = 0.5$, both as recommended in the original paper; the other parameters need to be tuned for different datasets. Specifically, m controls the sharpness of time hierarchical contrasting. TS2Vec uses $m = 1$ (constant) and SoftCLT uses $m(k) = 2^k$ (exponential), where k is the depth of pooling layers when computing temporal loss. In this study, we add one more option $m(k) = k + 1$ (linear), and will tune the best way for different datasets.

$$w_{\text{inst}}(i,j) = \frac{2\alpha}{1+\exp(\tau_{\text{inst}}\cdot d(\boldsymbol{x}_i,\boldsymbol{x}_j)))} + \begin{cases}1-\alpha, & \text{if } i=j\\ 0, & \text{if } i\neq j\end{cases} \tag{6.5}$$

$$w_{\text{temp}}(t,s) = \frac{2}{1+\exp(\tau_{\text{temp}}\cdot m\cdot|t-s|)} \tag{6.6}$$

Then Equation (6.7) shows the SoftCLT loss, where we let λ be 0.5 as recommended in the original paper. For a more detailed explanation and analysis, we refer the readers to [230].

$$\mathcal{L}_{\text{SoftCLT}} = \frac{1}{NT}\sum_i\sum_t\left(\lambda\ell^{(i,t)}_{\substack{\text{inst}\\\text{SoftCLT}}} + (1-\lambda)\ell^{(i,t)}_{\substack{\text{temp}\\\text{SoftCLT}}}\right),$$
$$\text{where}\begin{cases}\ell^{(i,t)}_{\substack{\text{inst}\\\text{SoftCLT}}} = -\displaystyle\sum_{j=1}^{B} w_{\text{inst}}(i,j)\log\frac{\exp(\boldsymbol{z}'_{i,t}\cdot\boldsymbol{z}''_{j,t})+\exp(\boldsymbol{z}''_{i,t}\cdot\boldsymbol{z}'_{j,t})}{S_{\text{inst}}(i,t)}\\ \qquad\qquad -\displaystyle\sum_{\substack{j=1\\j\neq i}}^{B} w_{\text{inst}}(i,j)\log\frac{\exp(\boldsymbol{z}'_{i,t}\cdot\boldsymbol{z}'_{j,t})+\exp(\boldsymbol{z}''_{i,t}\cdot\boldsymbol{z}''_{j,t})}{S_{\text{inst}}(i,t)}\\ \ell^{(i,t)}_{\substack{\text{temp}\\\text{SoftCLT}}} = -\displaystyle\sum_{s=1}^{T} w_{\text{temp}}(t,s)\log\frac{\exp(\boldsymbol{z}'_{i,t}\cdot\boldsymbol{z}''_{i,s})+\exp(\boldsymbol{z}''_{i,t}\cdot\boldsymbol{z}'_{i,s})}{S_{\text{temp}}(i,t)}\\ \qquad\qquad -\displaystyle\sum_{\substack{s=1\\s\neq t}}^{T} w_{\text{temp}}(t,s)\log\frac{\exp(\boldsymbol{z}'_{i,t}\cdot\boldsymbol{z}'_{i,s})+\exp(\boldsymbol{z}''_{i,t}\cdot\boldsymbol{z}''_{i,s})}{S_{\text{temp}}(i,t)}\end{cases} \tag{6.7}$$

6.3.4 Structure-preserving regularisers

We use the topology-preserving loss proposed in [249] and the graph-geometry-preserving loss proposed in [254] as two structure-preserving regularisers, respectively focusing on the global and local structure of similarity relations. The global structure is preserved for instance-wise comparison, and the local structure is preserved for comparison across temporal or spatial features. In the following, we briefly describe the two losses, and the readers are referred to the original papers for more details.

Equation (6.8) presents the topology-preserving loss computed in each mini-batch. Here $\boldsymbol{A}$ is a $B \times B$ EUC distance matrix between the samples in a batch, and is used to construct the Vietoris-Rips complex; $\boldsymbol{\pi}$ represents the persistence pairing indices of simplices that are considered topologically significant. The superscripts X and Z indicate original data space and latent space, respectively.

$$\mathcal{L}_{\text{Topo}} = \frac{1}{2} \left\| \boldsymbol{A}^X \left[\boldsymbol{\pi}^X\right] - \boldsymbol{A}^Z \left[\boldsymbol{\pi}^X\right] \right\|^2 + \frac{1}{2} \left\| \boldsymbol{A}^Z \left[\boldsymbol{\pi}^Z\right] - \boldsymbol{A}^X \left[\boldsymbol{\pi}^Z\right] \right\|^2 \tag{6.8}$$

The graph-geometry-preserving loss is also computed per mini-batch, as is shown in Equation (6.9). $\mathcal{L}_{\text{GGeo}}$ measures geometry distortion, i.e., how much $f_{\boldsymbol{\theta}}$ deviates from being an isometry that preserves distances and angles. The geometry to be preserved of the original data manifold is implied by a similarity graph. To represent temporal and spatial characteristics, instead of using an instance as a node in the graph, we consider the nodes as timestamps or in a spatial dimension such as sensors or road users. Then the edges in the graph are defined by pairwise geodesic distances between nodes.

$$\mathcal{L}_{\text{GGeo}} = \frac{1}{B} \sum_{i=1}^{B} \text{Tr} \left[\tilde{H}_i \left(L, \tilde{f}_\theta(\boldsymbol{x}_i) \right)^2 - 2\tilde{H}_i \left(L, \tilde{f}_\theta(\boldsymbol{x}_i) \right) \right], \tag{6.9}$$

where $\tilde{H}_i$ represents an approximation of the Jacobian matrix of $f_{\boldsymbol{\theta}}$. Note that $\tilde{f}_\theta(\boldsymbol{x}_i)$ as the latent representation of $\boldsymbol{x}_i$ needs to maintain the node dimension. For example, if the nodes are considered as timestamps, $\tilde{f}_\theta(\boldsymbol{x}_i) \in \mathbb{R}^{T \times P}$; if the nodes are spatial objects, $\tilde{f}_\theta(\boldsymbol{x}_i) \in \mathbb{R}^{S \times P}$. With a similarity graph defined, then L is the graph Laplacian that is approximated using a kernel matrix with a width hyperparameter h, which requires tuning for different datasets.

6.3.5 Stabilisation around the theoretical optimal values

This study considers $\mathcal{L}_{\text{TS2Vec}}$ and $\mathcal{L}_{\text{SoftCLT}}$ as $\mathcal{L}_{\text{CLT}}$, and $\mathcal{L}_{\text{SP}}$ can be $\mathcal{L}_{\text{Topo}}$ or $\mathcal{L}_{\text{GGeo}}$. Under this consideration, the optimal values for both $\mathcal{L}_{\text{CLT}}$ and $\mathcal{L}_{\text{SP}}$ are 0. For $\mathcal{L}_{\text{TS2Vec}}$, a value of 0 is reached when $\boldsymbol{z}'_{i,t}$ and $\boldsymbol{z}''_{i,t}$ are identical. Similarly, the optimal case of $\mathcal{L}_{\text{SoftCLT}}$ is when the samples with soft assignments close to 1 are identical, while dissimilar samples have soft assignments close to 0. The topology-preserving loss $\mathcal{L}_{\text{Topo}}$ is 0 when the topologically relevant distances remain the same in the latent space as in the original space, i.e., $\boldsymbol{A}^X \left[\boldsymbol{\pi}^X\right] = \boldsymbol{A}^Z \left[\boldsymbol{\pi}^X\right]$ and $\boldsymbol{A}^X \left[\boldsymbol{\pi}^Z\right] = \boldsymbol{A}^Z \left[\boldsymbol{\pi}^Z\right]$. Finally, $\mathcal{L}_{\text{GGeo}}$ approximates the distortion measure of isometry and is ideally 0, although it may be negative as the approximation of $\tilde{H}_i$ is kernel-based depending on the hyperparameter h.

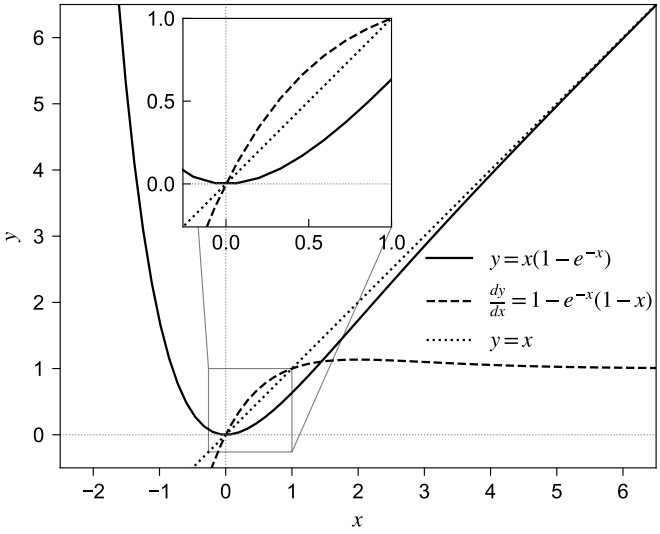


Figure 6.1 Illustration of the modification function.

Section 6.3.2 has briefly described the modification of $\mathcal{L}_{\text{CLT}}$ and $\mathcal{L}_{\text{SP}}$ by $y = x(1 - \exp(-x))$ in Equation (6.1). Now we explain the effect of this function more with Figure 6.1. The first objective of this modification is to penalise negative values of $\mathcal{L}_{\text{CLT}}$ and $\mathcal{L}_{\text{SP}}$. While $x > 0$, y decreases as x decreases; but when $x < 0$, y rapidly increases as x decreases. Once $\mathcal{L} < 0$, the modified term ℓ increases and the direction of gradient descent reverses.

Another objective is to avoid large-step updates when $\mathcal{L}$ is close to 0, so as not to miss its optimum. The modified y approximates x while x is large, but has a slower decreasing rate when $x < 1$. More specifically, the derivative of $x(1 - \exp(-x))$ is $x'(1 - \exp(-x)(1 - x))$, where x' denotes the derivative of x. While $x > 1$, the multiplier in the parentheses is around 1 and x' is less interfered with. While x decreases from 1 to 0, the multiplier decreases from 1 to 0, so the derivative also decreases. As a result, this modification can stabilise training when either $\mathcal{L}_{\text{CLT}}$ or $\mathcal{L}_{\text{SP}}$ approaches its optimal value 0.

6.3.6 Dynamic weighting mechanism to balance contrastive learning and structure preservation

The training needs to balance between contrastive learning and structure preservation to avoid the neural network parameters being biased by either of the two objectives. However, the magnitudes of $\mathcal{L}_{\text{CLT}}$ and $\mathcal{L}_{\text{SP}}$ vary with different datasets and hyperparameter settings. This variation precludes fixed weights for the two modified losses.

Inspired by [270], we weigh ℓ_{CLT} and ℓ_{SP} by considering their uncertainties. We consider the loss values as deviations from their optimal values, and learn adaptive weights according to the deviations. Given the optimal value of 0, we assume a Gaussian distribution of ℓ with standard deviation σ, i.e., $p(\ell) = \mathcal{N}(0, \sigma^2)$. Then we can maximise the log likelihood $\sum \log p(\ell) = \sum(-\log 2\pi - \log \sigma^2 - \ell^2/\sigma^2)/2$ to learn σ. This is equivalent to minimising $\sum \left(\ell^2/2/\sigma^2 + \log \sigma\right)$. When balancing between two losses ℓ_{CLT} and ℓ_{SP} that have deviations σ_{CLT} and σ_{SP}, respectively, we need to use Equation (6.10).

$$\arg\max - \sum \log p(\ell_{\text{CLT}}) p(\ell_{\text{SP}}) \Leftrightarrow \arg\min \sum \left(\frac{1}{2\sigma_{\text{CLT}}^2} \ell_{\text{CLT}} + \frac{1}{2\sigma_{\text{SP}}^2} \ell_{\text{SP}} + \log \sigma_{\text{CLT}} \sigma_{\text{SP}} \right) \tag{6.10}$$

Replacing ℓ_{CLT} in Equation (6.10) with $\mathcal{L}_{\text{CLT}}\left(1 - \exp(-\mathcal{L}_{\text{CLT}})\right)$ and ℓ_{SP} with $\mathcal{L}_{\text{SP}}\left(1 - \exp(-\mathcal{L}_{\text{SP}})\right)$, Equation (6.1) is eventually derived to be the complete loss. The training process trades-off between $\mathcal{L}_{\text{CLT}}$ and $\mathcal{L}_{\text{SP}}$, as well as between the weight regulariser $r_{\boldsymbol{\eta}} = \log \sigma_{\text{CLT}} \sigma_{\text{SP}}$ and the rest of Equation (6.1). When $\mathcal{L}_{\text{CLT}}$ is small and $\mathcal{L}_{\text{SP}}$ is large, σ_{CLT} becomes small and σ_{SP} becomes large, which then increases the weight for $\mathcal{L}_{\text{CLT}}$ while reduces the weight for $\mathcal{L}_{\text{SP}}$. The reverse occurs when $\mathcal{L}_{\text{CLT}}$ is large and $\mathcal{L}_{\text{SP}}$ is small. As the weighted sum of $\mathcal{L}_{\text{CLT}}$ and $\mathcal{L}_{\text{SP}}$ increases by larger weights, $\log \sigma_{\text{CLT}} \sigma_{\text{SP}}$ decreases and discourages the increase from being too much. Similarly, if the weighted sum decreases by smaller weights, $\log \sigma_{\text{CLT}} \sigma_{\text{SP}}$ also regularises the decrease.

6.4 Experiments

We compare 6 losses for self-supervised representation learning (SSRL) of time series: TS2Vec, SoftCLT, Topo-TS2Vec, GGeo-TS2Vec, Topo-SoftCLT, and GGeo-SoftCLT. Among the losses, TS2Vec [229] and SoftCLT [230] are baselines, and the others extend these two with a topology-preserving or a graph-geometry-preserving regulariser. The comparison is then evaluated by downstream task performances using these differently encoded representations. Consequently, the comparison and evaluation serve as an extensive ablation study focusing on the effects of structure-preserving regularisers. Our experiments are conducted with an NVIDIA A100 GPU with 80GB RAM and 5 Intel Xeon CPUs. For fair comparisons, we control the following conditions during experiments: random seed, the space and strategy for hyperparameter search, maximum training epochs, early stopping criteria, and samples used for evaluating local structure preservation.

6.4.1 Baselines and datasets

The evaluation of performance improvement is on 3 downstream tasks: multivariate time series classification, macroscopic traffic prediction, and microscopic traffic prediction. For every downstream task, we split training/(validation)/test sets following the baseline study and make sure the same data are used across models. Each experiment for a task has two stages, of which the first is SSRL and the second uses the encoded representations to perform classification/prediction. Only the split training set is used in the first stage, with 25% separated as an internal validation set to schedule the learning rate for SSRL.

The classification task is on 28 datasets[2] retrieved from the UEA archive [271]. For each dataset, we set the representation dimension to 320 as used in the TS2Vec and SoftCLT studies, train 6 encoders with the 6 losses, and then classify the encoded representations with an RBF-kernel SVM. For traffic prediction, we use the dataset and model in [231] for the macroscopic baseline, and those in [232] for the microscopic baseline. The macroscopic traffic prediction uses 40 minutes (2-minute intervals) of historical data in 193 consecutive road segments to predict for all segments in the next 30 minutes. The microscopic traffic prediction forecasts the trajectory of an ego vehicle in 3 seconds, based on the history of up to 26 surrounding road users in the past 1 second (0.1-second intervals). Both traffic prediction baselines use encoder-decoder structures. We first pretrain the encoder with the 6 different losses for SSRL, and then fine-tune the complete model for prediction. The baseline trained from scratch is also compared.

To facilitate clearer analyses when presenting results, we divide the datasets included in the UEA archive into those with spatial features and those without. According to data descriptions in [271], the UEA datasets are grouped into 6 categories: human activity recognition, motion classification, ECG classification, EEG/MEG classification, audio spectra classification, and other problems. The human activity and motion categories, along with the PEMS-SF and LSST datasets that are categorised as other problems, contain spatial features. We thus consider these as spatial, and the remaining datasets as non-spatial. As a result, each division includes 14 datasets.

6.4.2 Hyperparameters

For each dataset, we perform a grid search to find the parameters that minimise $\mathcal{L}_{\text{CLT}}$ after a certain number of iterations, where we set a constant learning rate of 0.001. Table 6.1 shows the search spaces of various hyperparameters, where bs is abbreviated for batch size and lr_η is a separate learning rate for dynamic weights. When searching for best-suited parameters, we first set them as default values, and then follow the search strategy presented in Table 6.2.

The search spaces and strategy can result in up to 63 runs for one dataset. To save searching time, we adjust the number of iterations to be adequate to reflect the progress of loss reduction but limited to prevent overfitting, as our goal is to identify suitable parameters rather than fully train the models. The number of iterations is scaled according to the number of training samples, with larger datasets receiving more iterations.

6.4.3 Evaluation metrics

Our performance evaluation uses both task-specific metrics and structure-preserving metrics. The former serves to validate performance improvements, while the latter serves to verify the effectiveness of preserving similarity structures. These metrics differ in whether a higher or lower value signifies better performance. To consistently indicate the

[2]The UEA archive collects 30 datasets in total. We omitted the two largest, InsectWingbeat and PenDigits, due to limited computation resources.

Table 6.1 Hyperparameter search space.

	Default	Search space
bs	8	[8, 16, 32] [a]
lr_η	0.05	[0.01, 0.05]
h	1	[0.25, 1, 9, 25, 49]
τ_{temp}	0	[0.5, 1, 1.5, 2, 2.5]
m	constant	[constant, linear, exponential]
τ_{inst}	0	[1, 3, 5, 10, 20]

bs: batch size; lr_η: learning rate for dynamic weights.

[a]Maximum bs does not exceed train size.

Table 6.2 Hyperparameter search strategy.

Stage	bs	lr_η	h	τ_{temp}	m	τ_{inst}
TS2Vec	△					
Topo-TS2Vec	□	△				
GGeo-TS2Vec	□	△	△			
SoftCLT Phase 1	◯			△	△	◯
SofrCLT Phase 2	△			□	□	△
Topo-SoftCLT	□	△		□	□	□
GGeo-SoftCLT	□	△	△	□	□	□

◯: default; □: inherited; △: tuned.

best method, in the tables presented in the following sub-sections, the **best** values are both bold and underlined; the **second-best** values are bold.

For evaluating the classification task, we use accuracy (Acc.) and the area under the precision-recall curve (AUPRC). To evaluate macroscopic traffic prediction, we use mean absolute error (MAE), root mean squared error (RMSE), the standard deviation of prediction errors (SDEP), and the explained variance by prediction (EVar). Dealing with microscopic traffic, we predict vehicle trajectories and assess the minimum final displacement error (min. FDE) as well as missing rates under radius thresholds of 0.5 m, 1 m, and 2 m ($\text{MR}_{0.5}$, MR_1, and MR_2).

As for metrics to evaluate structure preservation, we adopt a combination of those used in [249] and [254]. More specifically, we consider 1) kNN, the proportion of shared k-nearest neighbours according to distance matrices in the latent space and in the original space; 2) continuity (Cont.), one minus the proportion of neighbours in the original space that are no longer neighbours in the latent space; 3) trustworthiness (Trust.), the counterpart of continuity, measuring the proportion of neighbours in the latent space but not in the original space; 4) MRRE, the averaged error in the relative ranks of

sample distances between in the latent and original space; and 5) distance matrix RMSE (dRMSE), the root mean squared difference between sample distance matrices in the latent and original space. We calculate these metrics at two scales to evaluate global and local structure preservation. For global evaluation, our calculation is based on the EUC distances between samples; for local evaluation, it is based on the EUC distances between timestamps in a sample for at most 500 samples in the test set.

6.5 Results

6.5.1 Multivariate time series classification

Table 6.3 displays the classification performance on spatial and non-spatial UEA datasets. Next to the averaged accuracy, we also include the loss values on test sets to offer more information. More detailed results can be found in Tables B1 and B2 in B.1, where we present the classification accuracy with different representation learning losses for each dataset. Then we use Table 6.4 to more specifically compare the relative improvements induced by adding a topology or graph-geometry preserving regulariser. The improvement is measured by the percentage of accuracy difference from the corresponding baseline performance.

Table 6.3 UEA classification evaluation.

Datasets	Method	Acc. (↑)	AUPRC (↑)	$\mathcal{L}_{\text{CLT}}$	$\mathcal{L}_{\text{SP}}$
With spatial features (14)	TS2Vec	0.848	0.872	2.943	
	Topo-TS2Vec	0.851	0.876	2.264	0.085
	GGeo-TS2Vec	0.856	0.881	2.200	186.9
	SoftCLT	0.852	0.876	7.943	
	Topo-SoftCLT	**0.862**	**0.882**	4.900	0.087
	GGeo-SoftCLT	**0.864**	**0.883**	2.316	221.1
Without spatial features (14)	TS2Vec	0.523	0.555	8.417	
	Topo-TS2Vec	**0.553**	**0.561**	11.12	0.122
	GGeo-TS2Vec	0.536	**0.564**	15.58	957.0
	SoftCLT	0.508	0.532	4.714	
	Topo-SoftCLT	0.496	0.534	7.328	0.124
	GGeo-SoftCLT	**0.537**	0.549	10.09	144.7

Note: the **best** values are both bold and underlined; the **second-best** values are bold.

Tables 6.3 and 6.4 clearly show that structure preservation improves classification accuracy, not only when time series data involve spatial features, but also when they do not. The relative improvements in Table 6.4 are higher for non-spatial datasets than for spatial datasets, which is because the datasets without spatial features are more difficult to learn in the UEA archive. As is shown in Table 6.3, the loss of contrastive learning *decreases*

Table 6.4 Classification accuracy improved by Topo/GGeo regulariser. The comparisons are made with corresponding baseline performances.

Datasets	Improvement by method	Persentage in Acc. (%)		
		min.	mean	max.
With spatial features (14)	Topo-TS2Vec	-4.403	0.800	16.54
	GGeo-TS2Vec	-3.783	1.143	10.44
	Topo-SoftCLT	-4.375	2.121	25.94
	GGeo-SoftCLT	-5.674	2.959	28.55
Without spatial features (14)	Topo-TS2Vec	-5.263	8.852	50.00
	GGeo-TS2Vec	-33.33	2.083	44.44
	Topo-SoftCLT	-33.33	-0.815	50.00
	GGeo-SoftCLT	-20.83	18.49	166.7

when a structure-preserving regulariser is added for spatial datasets, while *increases* for non-spatial datasets. This implies that preserving similarity structure is well aligned with contrastive learning for spatial datasets, and can even enhance contrastive learning.

The assessment of similarity preservation is presented in Table 6.5 at both local and global scales. Consistent with the task-specific evaluation, Table 6.5 shows that structure-preserving regularisation preserves more complete information on similarity relations. The improvements are generally more significant on datasets with spatial features, which makes it more evident that our proposed preservation suits spatial time series data better. Although the comparisons in Tables 6.3∼6.5 indicate more notable improvements by preserving graph geometry than preserving topology, we have to note that this does not demonstrate the universal superiority of one over the others. Different datasets have different characteristics that benefit from preserving global or local structure, and domain knowledge is necessary to determine which could be more effective. We will discuss this more in Section 6.6.

6.5.2 Macroscopic and microscopic traffic prediction

In Table 6.6, we present the performance evaluation for both macroscopic and microscopic traffic prediction. This table shows consistent improvements by pretraining encoders with our methods. Notably, single contrastive learning (i.e., TS2Vec and SoftCLT) does not necessarily improve downstream prediction, whereas it does when used together with preserving certain similarity structures. Given that our comparisons are conducted through controlling random conditions, this result effectively shows the necessity of preserving structure when learning traffic interaction representations. In addition, we plot polar heatmaps in Figure 6.2 to visualise the encoded latent representations for the sensors in macroscopic traffic prediction. These sensors are deployed along a ring road, thus adjacent sensors are expected to have similar states and representation patterns. The figure

Table 6.5 Structure preservation evaluation over datasets with and without spatial features in the UEA archive. The standard deviations are computed across datasets.

Datasets	Method	Local mean between timestamps					Global mean between all samples				
		kNN (↑)	Trust. (↑)	Cont. (↑)	MRRE (↓)	dRMSE (↓)	kNN (↑)	Trust. (↑)	Cont. (↑)	MRRE (↓)	dRMSE (↓)
With spatial features (14)	TS2Vec	0.563 ± 0.149	0.875 ± 0.082	0.868 ± 0.073	0.117 ± 0.08	0.346 ± 0.099	0.419 ± 0.149	0.765 ± 0.110	0.784 ± 0.126	0.189 ± 0.124	**0.150** ± 0.065
	Topo-TS2Vec	**0.569** ± 0.151	0.878 ± 0.082	**0.873** ± 0.075	**0.114** ± 0.081	0.344 ± 0.100	0.418 ± 0.150	0.764 ± 0.110	0.783 ± 0.125	0.190 ± 0.126	0.154 ± 0.069
	GGeo-TS2Vec	**0.569** ± 0.154	**0.881** ± 0.08	**0.873** ± 0.076	**0.114** ± 0.079	**0.341** ± 0.096	0.418 ± 0.153	0.762 ± 0.113	0.781 ± 0.129	0.190 ± 0.127	0.157 ± 0.075
	SoftCLT	0.562 ± 0.156	0.875 ± 0.078	0.866 ± 0.074	0.117 ± 0.079	0.348 ± 0.096	0.420 ± 0.153	0.765 ± 0.112	**0.788** ± 0.123	**0.187** ± 0.125	0.171 ± 0.103
	Topo-SoftCLT	0.564 ± 0.156	0.877 ± 0.077	0.869 ± 0.074	0.115 ± 0.077	0.344 ± 0.097	**0.421** ± 0.157	**0.767** ± 0.114	0.784 ± 0.126	0.188 ± 0.128	0.153 ± 0.068
	GGeo-SoftCLT	**0.571** ± 0.150	**0.883** ± 0.073	**0.875** ± 0.067	**0.111** ± 0.076	**0.337** ± 0.091	**0.425** ± 0.149	**0.768** ± 0.110	**0.790** ± 0.121	**0.185** ± 0.125	**0.149** ± 0.065
Without spatial features (14)	TS2Vec	0.423 ± 0.125	**0.835** ± 0.095	**0.820** ± 0.105	0.150 ± 0.105	**0.304** ± 0.162	**0.362** ± 0.175	0.767 ± 0.132	0.767 ± 0.136	**0.252** ± 0.151	0.197 ± 0.136
	Topo-TS2Vec	0.424 ± 0.125	0.831 ± 0.100	**0.820** ± 0.106	0.151 ± 0.105	**0.308** ± 0.164	0.356 ± 0.176	0.767 ± 0.134	0.763 ± 0.139	0.254 ± 0.154	**0.191** ± 0.120
	GGeo-TS2Vec	0.420 ± 0.126	0.832 ± 0.094	**0.820** ± 0.108	0.151 ± 0.105	0.310 ± 0.145	**0.365** ± 0.176	**0.771** ± 0.137	**0.769** ± 0.136	0.253 ± 0.157	**0.189** ± 0.125
	SoftCLT	**0.432** ± 0.127	**0.835** ± 0.099	**0.820** ± 0.106	**0.148** ± 0.105	0.312 ± 0.162	0.354 ± 0.177	0.764 ± 0.134	0.763 ± 0.136	**0.252** ± 0.151	0.197 ± 0.131
	Topo-SoftCLT	0.426 ± 0.119	0.834 ± 0.095	0.818 ± 0.104	**0.148** ± 0.102	0.312 ± 0.163	0.361 ± 0.181	**0.768** ± 0.131	**0.768** ± 0.132	0.254 ± 0.152	0.205 ± 0.123
	GGeo-SoftCLT	**0.430** ± 0.122	**0.835** ± 0.095	**0.822** ± 0.101	**0.147** ± 0.101	0.315 ± 0.151	0.355 ± 0.174	0.762 ± 0.134	0.761 ± 0.136	0.257 ± 0.153	0.203 ± 0.132

Note: the **best** values are both bold and underlined; the **second-best** values are bold.

Table 6.6 Macroscopic and microscopic traffic prediction performance evaluation. Metrics are reported as mean ± standard deviation over 10 stratified folds of the test set.

Method	Macroscopic Traffic				Microscopic Traffic			
	MAE (↓) (km/h)	RMSE (↓) (km/h)	SDEP (↓) (km/h)	EVar (↑) (%)	min. FDE (↓) (m)	$MR_{0.5}$ (↓) (%)	MR_1 (↓) (%)	MR_2 (↓) (%)
No pretraining	**2.850** ± 0.044	5.911 ± 0.100	5.909 ± 0.100	84.784 ± 0.314	0.640 ± 0.013	59.253 ± 1.099	12.161 ± 0.929	0.744 ± 0.256
TS2Vec	2.878 ± 0.046	5.981 ± 0.111	5.981 ± 0.111	84.412 ± 0.337	**0.636** ± 0.008	59.453 ± 0.866	11.899 ± 0.721	**<u>0.558</u>** ± 0.17
Topo-TS2Vec	2.862 ± 0.043	5.914 ± 0.099	5.907 ± 0.099	84.793 ± 0.306	**<u>0.634</u>** ± 0.008	**<u>58.144</u>** ± 1.014	**11.761** ± 0.809	0.737 ± 0.252
GGeo-TS2Vec	2.887 ± 0.045	5.980 ± 0.107	5.977 ± 0.107	84.433 ± 0.326	**0.636** ± 0.010	**58.289** ± 1.007	**<u>11.747</u>** ± 0.797	0.737 ± 0.239
SoftCLT	2.856 ± 0.043	5.937 ± 0.105	5.931 ± 0.104	84.670 ± 0.348	0.641 ± 0.018	59.501 ± 1.742	11.940 ± 0.998	0.785 ± 0.278
Topo-SoftCLT	**2.850** ± 0.045	**5.881** ± 0.111	**5.880** ± 0.111	**84.934** ± 0.327	0.640 ± 0.012	58.626 ± 0.995	12.043 ± 0.735	0.820 ± 0.249
GGeo-SoftCLT	**<u>2.834</u>** ± 0.046	**<u>5.878</u>** ± 0.116	**<u>5.877</u>** ± 0.116	**<u>84.952</u>** ± 0.344	0.652 ± 0.013	60.638 ± 0.596	13.249 ± 0.681	**0.723** ± 0.231
Best improvement	0.568	0.563	0.552	0.198	0.929	1.872	3.401	24.998

Note: the **<u>best</u>** values are both bold and underlined; the **second-best** values are bold.

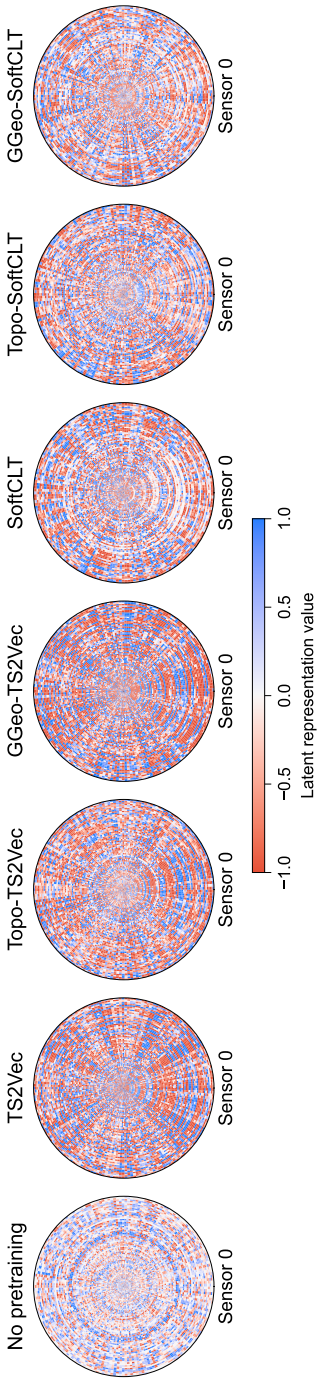


Figure 6.2 Encoded representations after training with different losses on the test set of the Macroscopic traffic prediction task.

intuitively shows better preserved spatial-temporal relations by contrastive learning and structure preservation.

Table 6.7 then displays the corresponding evaluation on similarity structure preservation, which is obtained by assessing the encoders after fine-tuning for traffic prediction. The results show that the better-performing methods in macro-traffic prediction (i.e., GGeo-SoftCLT, Topo-SoftCLT) and micro-traffic prediction (i.e., Topo-TS2Vec, GGeo-TS2Vec) preserve more similarity structures at both global and local scales. In general, the metric values are close for the same task across methods; however, when a method has a significant advantage over the others, it indicates superior performance. Examples for macroscopic traffic prediction are Topo-SoftCLT in local Cont. and GGeo-SoftCLT in global dRMSE; for microscopic traffic prediction are Topo-TS2Vec in local Trust. and GGeo-TS2Vec in global Cont.

Table 6.7 Structure preservation evaluation of encoders after the fine-tuning in traffic prediction tasks.

Method	Macroscopic Traffic					Microscopic Traffic				
	kNN (↑)	Trust. (↑)	Cont. (↑)	MRRE (↓)	dRMSE (↓)	kNN (↑)	Trust. (↑)	Cont. (↑)	MRRE (↓)	dRMSE (↓)
Local mean between timestamps for at most 500 samples										
No pretraining	0.125	0.524	**0.526**	0.496	**0.224**	0.373	0.742	0.552	0.426	**0.478**
TS2Vec	0.125	0.523	0.522	0.501	0.251	**0.398**	**0.761**	**0.592**	**0.393**	0.496
Topo-TS2Vec	**0.128**	**0.533**	0.522	**0.491**	**0.242**	0.397	0.754	**0.590**	0.399	0.506
GGeo-TS2Vec	0.126	0.529	**0.524**	0.496	0.249	0.397	**0.756**	0.589	**0.396**	0.508
SoftCLT	0.127	0.526	0.523	0.500	0.246	0.378	0.746	0.552	0.427	**0.478**
Topo-SoftCLT	**0.129**	**0.536**	**0.524**	**0.492**	0.250	**0.398**	0.755	0.588	0.405	0.480
GGeo-SoftCLT	0.127	0.527	0.523	0.498	0.261	0.397	0.751	0.589	0.405	0.485
Global mean between all samples										
No pretraining	**0.316**	**0.949**	**0.969**	**0.031**	**0.364**	0.218	0.937	0.920	0.049	0.141
TS2Vec	0.264	0.940	0.957	0.039	0.377	0.232	0.953	0.920	**0.044**	**0.139**
Topo-TS2Vec	0.276	0.942	0.963	0.036	0.379	**0.233**	**0.958**	0.917	0.045	**0.138**
GGeo-TS2Vec	0.263	0.940	0.959	0.039	0.400	0.231	**0.959**	**0.923**	**0.041**	0.140
SoftCLT	**0.299**	**0.943**	**0.966**	**0.035**	0.391	0.224	0.924	0.916	0.055	0.148
Topo-SoftCLT	0.288	0.940	0.965	0.036	0.371	0.215	0.909	0.901	0.065	0.150
GGeo-SoftCLT	0.287	0.939	0.964	0.037	**0.359**	**0.240**	0.935	**0.926**	0.046	0.146

Note: the **best** values are both bold and underlined; the **second-best** values are bold.

Notably, in macroscopic traffic prediction, fine-tuning from scratch maintains the greatest global similarities. This implies that the specific model architecture might allow for learning similarity structure without pretraining. This is not crystal clear with the final evaluation only. In the next sub-section, we will add different model architectures for the macro-traffic prediction task, and visualise the fine-tuning progress to further understand the contribution of structure preservation to downstream task performance.

6.5.3 Training efficiency

Incorporating structure-preserving regularisation increases computational complexity, and consequently, training time. The magnitude of this increase depends on the data and model that are applied on. With Table 6.8, we quantify the additional time required for structure preservation and evaluate its impact across diverse model architectures. In prior

experiments, we used Convolutional Neural Network (CNN) encoders for the classification task on UEA datasets, Dynamic Graph Convolution Network (DGCN, [214]) encoder for macroscopic traffic prediction, and Hierarchical Graph Neural Network (HGNN, based on VectorNet [272, 273]) encoder for microscopic traffic prediction. To obtain a more comprehensive evaluation, we include two more Recurrent Neural Network (RNN) models for macroscopic traffic prediction: Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) encoders, paired with simple linear decoders.

Table 6.8 Training time per epoch in the stage of self-supervised representation learning.

Task/data	Encoder	Base (sec/epoch)	TS2Vec	Topo-TS2Vec	GGeo-TS2Vec	SoftCLT	Topo-SoftCLT	GGeo-SoftCLT
Avg. UEA[a]	CNN	11.94	1.00×	1.46×	2.35×	1.00×	1.46×	2.36×
MicroTraffic	HGNN	122.93	1.00×	1.45×	1.16×	1.23×	1.69×	1.42×
	DGCN	74.78	1.00×	1.29×	1.09×	1.02×	1.40×	1.37×
MacroTraffic	LSTM	18.04	1.00×	1.49×	1.12×	1.09×	1.57×	1.22×
	GRU	16.49	1.00×	1.54×	1.14×	1.10×	1.61×	1.24×

[a] Detailed results are referred to Tables B3 and B4 in B.1.

Table 6.8 shows that preserving structure increases training time by less than 50% in most cases, and suits DGCN particularly well with the least additional time. However, when time sequences are very long, the computation of graph-geometry preserving loss becomes intense. For example, the time series length is 1,197 in the Cricket dataset and results in a pretraining time of 2.86 times the base; likewise, the EthanolConcentration dataset has a length of 1,751 and a pretraining time of 4.09 times the base, and the EigenWorms dataset uses 7.59 times of the base pretraining time with a time series length of 17,984.

In more detail, we evaluate the fine-tuning efficiency in macroscopic traffic prediction to further investigate the contribution of structure preservation. Figure 6.3 shows the convergence process of different models with and without pretraining, where RMSE is used to evaluate prediction performance and the other metrics indicate the preservation of global similarity relations.

For all models of DGCN, LSTM, and GRU, structure preservation consistently enhances prediction performance compared to training from scratch (No pretraining). The enhancement is significant when using LSTM and GRU, achieving 6.68% and 10.14% improvement in RMSE, respectively. Meanwhile, the progress of structure preservation is stable when the encoder is pretrained, and maintains the advantage over no pretraining throughout the fine-tuning process. In contrast, for DGCN, which is a more sophisticated model tailored for the task, training from scratch is already very effective and pretraining brings relatively minor improvement. This implies that certain model architectures are better suited to a specific task than others, and the preservation of similarity structure in the latent representations may be a good indicator for model selection.

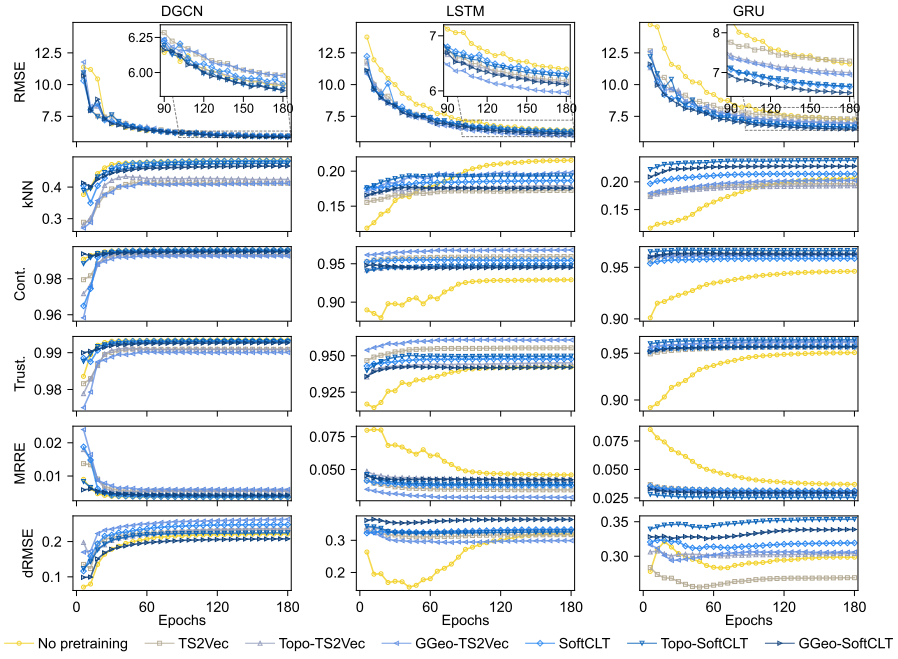


Figure 6.3 Fine-tuning progress of models trained from scratch and pretrained with different losses in macroscopic traffic prediction. Values of the final performance are referred to Table 6.6 for DGCN, to Tables B5 and B6 in B.2 for LSTM and GRU.

6.6 Discussion

The theoretical foundation and experimental results presented in this paper not only demonstrate evident improvements in downstream task performance but also reveal a critical bridge between contrastive learning and similarity structure preservation. In this section, we discuss these findings to guide method selection, interpret the observed performance improvements, and remind of potential failure modes.

6.6.1 Method selection

A key consideration when applying our method lies in selecting an appropriate loss function. This involves two layers of choices: TS2Vec versus SoftCLT, and topology-preserving (Topo-) versus graph-geometry-preserving (GGeo-) regularisation. For the first choice, TS2Vec is generally more suitable for classification tasks with fewer classes as TS2Vec only compares the similarity between two different augmentations of the same sample. In contrast, SoftCLT incorporates all samples in a mini-batch by assigning soft labels based on similarity. This performs a more detailed similarity comparison and thus is advantageous for tasks with a larger number of classes or for regression. As shown in Figure 6.4, in the UEA archive, datasets for which (Topo/GGeo-)TS2Vec achieves the best accuracy tend to have fewer classes than those with the best performance achieved by (Topo/GGeo-)SoftCLT. In essence, SoftCLT implicitly embeds the similarity structure through soft labels.

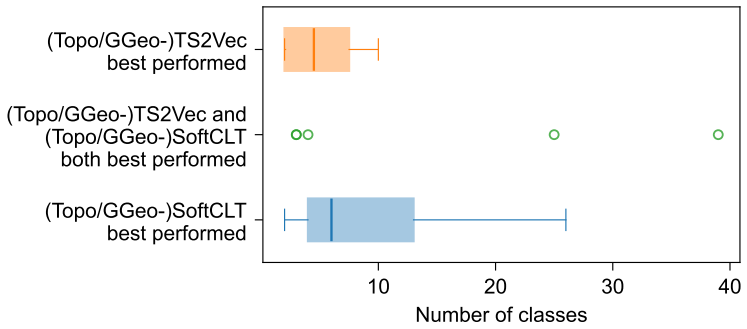


Figure 6.4 Box plots of the number of classes in the UEA datasets for which the best classification accuracy is achieved after contrastive learning based on TS2Vec or SoftCLT.

Then the second layer of choice hinges on the scale of structural relevance to the downstream task. The topology-preserving regulariser is designed to maintain the global similarity structure between data samples, and the graph-geometry-preserving regulariser focuses on locally preserving temporal or spatial similarity structures within samples. Therefore, Topo-regularisation is especially beneficial when the downstream task relies on inter-sample relations; whereas GGeo-regularisation is particularly useful in tasks where subtle intra-sample variations are critical, such as traffic prediction.

Our hypothesis for the performance improvements observed across tasks is that explicit structure-preserving regularisation enforces the encoded latent space to reflect the original distribution of data. This aligns with the theoretical view of neural networks as models that approximate the conditional distributions of outputs on inputs. Contrastive learning typically relies on predefined positive and negative augmentations, which may disrupt the original patterns underlying data and introduce biases. By anchoring the latent space to the original data manifold, structure preservation can effectively mitigate the potential biases by reducing the dependency on augmentations.

6.6.2 Failure modes

While the proposed method demonstrates consistent improvements across a range of tasks, several limitations and potential failure modes need to be acknowledged. These limitations are instructive in identifying the contexts where the method is most effective and where caution is warranted.

One source of difficulty arises from the sensitivity to data scale and complexity. The additional cost by structure-preserving regularisation is acceptable in many cases, but the overhead becomes prohibitive when time series are extremely long. For example, datasets with thousands of time steps result in a large similarity graph that inflates the computation memory and time of the GGeo-loss. Meanwhile, for models that are already well aligned with data, such as the DGCN baseline in macroscopic traffic prediction, pretraining with our method may provide only marginal benefits. This suggests that the utility of structure preservation is most pronounced when the data involve complex similarity relations that base models cannot easily capture.

Another important consideration concerns the benefits of dynamic weighting compared to fixed weights. The dynamic weighting mechanism is introduced to adaptively balance the loss magnitudes of contrastive learning and structure preservation. This works particularly well when the loss components have a large magnitude gap, which biases parameter updates. However, dynamic weighting may not outperform fixed weighting when the magnitudes of loss components are balanced in nature. In such circumstances, the adaptive mechanism can oscillate excessively, delaying convergence or amplifying noise in the latent space. Table 6.9 compares model performance on 5 datasets in the UEA archive. The datasets are selected because of relatively large magnitude gaps between contrastive learning loss and structure preservation loss. Dynamic and fixed weighting yield comparable classification accuracies, but the dynamic mechanism consistently improves structure preservation metrics and achieves lower loss.

Finally, dependency on hyperparameter tuning presents a limitation. The approach proposed in this paper requires careful selection of kernel width for Laplacian approximation, temperature parameters in the contrastive loss, and the learning rate for dynamic weights. Although we provide a scheme for hyperparameter tuning, improper parameters can lead to ineffective training.

Table 6.9 Comparison of dynamic and fixed weighting performance with the UEA archive. By fixed weighting, we let $\sigma_{\text{CLT}} = 1$ and $\sigma_{\text{SP}} = 1$, so that the weights for both contrastive learning and structure preservation are 0.5.

Dataset	Method	Acc. (↑)	kNN (↑)	Trust. (↑)	Cont. (↑)	MRRE (↓)	dRMSE (↓)	$\mathcal{L}_{\text{CLT}}$	$\mathcal{L}_{\text{SP}}$
ERing	GGeo-SoftCLT (dynamic weighting)	0.871 ± 0.011	0.776 ± 0.004	0.954 ± 0.002	0.940 ± 0.003	0.047 ± 0.001	0.209 ± 0.007	0.652 ± 0.023	106.972 ± 15.521
	GGeo-SoftCLT (fixed weighting)	0.876 ± 0.013	0.767 ± 0.003	0.947 ± 0.002	0.933 ± 0.002	0.05 ± 0.001	0.226 ± 0.004	0.602 ± 0.015	149.115 ± 15.738
Libras	GGeo-TS2Vec (dynamic weighting)	0.863 ± 0.010	0.757 ± 0.002	0.945 ± 0.001	0.915 ± 0.001	0.091 ± 0.001	0.281 ± 0.004	4.724 ± 0.038	387.867 ± 47.765
	GGeo-TS2Vec (fixed weighting)	0.864 ± 0.016	0.754 ± 0.002	0.943 ± 0.001	0.915 ± 0.001	0.091 ± 0.001	0.275 ± 0.004	4.818 ± 0.021	519.156 ± 33.440
NATOPS	GGeo-SoftCLT (dynamic weighting)	0.921 ± 0.009	0.640 ± 0.004	0.903 ± 0.004	0.859 ± 0.004	0.108 ± 0.002	0.296 ± 0.004	0.450 ± 0.007	712.646 ± 39.097
	GGeo-SoftCLT (fixed weighting)	0.920 ± 0.011	0.639 ± 0.003	0.902 ± 0.003	0.855 ± 0.003	0.109 ± 0.002	0.295 ± 0.004	0.447 ± 0.006	811.255 ± 55.068
StandWalkJump	GGeo-TS2Vec (dynamic weighting)	0.273 ± 0.049	0.389 ± 0.007	0.947 ± 0.003	0.905 ± 0.004	0.058 ± 0.002	0.226 ± 0.032	12.592 ± 2.747	3826.480 ± 2231.619
	GGeo-TS2Vec (fixed weighting)	0.300 ± 0.057	0.386 ± 0.006	0.945 ± 0.003	0.904 ± 0.004	0.059 ± 0.002	0.239 ± 0.034	12.821 ± 2.847	4173.859 ± 2471.534
UWaveGestureLibrary	GGeo-TS2Vec (dynamic weighting)	0.855 ± 0.011	0.759 ± 0.009	0.950 ± 0.005	0.956 ± 0.006	0.03 ± 0.002	0.308 ± 0.007	1.098 ± 0.170	501.284 ± 47.744
	GGeo-TS2Vec (fixed weighting)	0.853 ± 0.018	0.755 ± 0.009	0.948 ± 0.006	0.954 ± 0.006	0.031 ± 0.002	0.306 ± 0.007	1.102 ± 0.174	604.046 ± 67.679

6.7 Conclusion

This paper presents a method for structure-preserving contrastive learning for spatial time series, where we propose a dynamic mechanism to adaptively balance contrastive learning and structure preservation. The method is generally applicable to time series data with spatial or geographical features, which are particularly abundant in transportation systems. Extensive experiments demonstrate that our methods improve the state of the arts, including for multivariate time series classification in various contexts and for traffic prediction at both macroscopic and microscopic scales.

An important advantage of this method is its ability to enforce the preservation of similarity structures in the latent space, thereby aligning representation learning with the original data manifold. At the same time, the efficiency costs of structure-preserving regularisation are moderate, making the approach practical for many applications. Nonetheless, several limitations deserve mention. The benefits of the method become expensive when time series are extremely long, where computational costs can be prohibitive. Dynamic weighting does not always outperform fixed weighting, but is particularly helpful when the training dynamics are dominated by one loss component. Moreover, the method depends on careful hyperparameter tuning, especially for the graph-geometry-preserving regulariser. These limitations do not undermine the contributions of the work but rather highlight opportunities for future research, including the design of lightweight approximations for long sequences, more stable dynamic weighting schemes, and principled strategies for automatic hyperparameter selection.

Notably, preserving the original similarity structure in the latent space is shown to be well aligned with and beneficial to contrastive learning for spatio-temporal data. This is crucial for informative representational learning from large-scale data, as it impacts a neural network's ability to model the underlying conditional distribution. Our experiments suggest that higher similarity structure preservation is a good indicator of more informative representations, highlighting that the structural information of similarities in spatio-temporal data remains yet to be exploited. In the context of increasingly large neural network models that involve diverse data modalities, we hope this study sheds light on more effective training of large models in transportation systems.

Appendix B

B.1 Detailed results on UEA datasets

This section provides detailed comparisons of evaluation results for the used 28 datasets in the UEA archive. Tables B1 and B2 present the results of classification accuracy. Tables B3 and B4 present the training time for self-supervised representation learning.

Table B1 Detailed evaluation of classification accuracy on spatial datasets in the UEA archive.

Dataset	TS2Vec	Topo-TS2Vec	GGeo-TS2Vec	SoftCLT	Topo-SoftCLT	GGeo-SoftCLT
ArticularyWordRecognition	0.980	**0.987**	0.983	**0.987**	0.977	**0.987**
BasicMotions	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**
CharacterTrajectories	0.971	0.985	0.972	0.980	0.977	**0.986**
Cricket	0.944	0.944	0.972	0.972	0.972	**0.986**
ERing	0.867	0.874	0.881	**0.893**	0.878	0.863
EigenWorms	0.809	0.817	0.863	0.817	**0.901**	0.840
Epilepsy	0.957	0.957	0.949	**0.964**	0.957	0.949
Handwriting	0.498	0.499	0.479	0.487	0.478	**0.580**
LSST	0.485	0.566	0.536	0.452	0.569	**0.581**
Libras	0.883	0.844	0.850	**0.889**	0.850	0.867
NATOPS	0.917	0.917	0.933	0.922	0.917	**0.944**
PEMS-SF	0.792	0.775	**0.815**	0.751	0.803	0.740
RacketSports	0.908	0.914	0.914	**0.928**	0.908	0.875
UWaveGestureLibrary	0.862	0.831	0.834	0.888	0.881	**0.897**
Avg. over spatial datasets	0.848	0.851	0.856	0.852	0.862	**0.864**

Table B2 Detailed evaluation of classification accuracy on non-spatial datasets in the UEA archive.

Dataset	TS2Vec	Topo-TS2Vec	GGeo-TS2Vec	SoftCLT	Topo-SoftCLT	GGeo-SoftCLT
AtrialFibrillation	0.200	**0.267**	0.133	0.133	0.200	**0.267**
DuckDuckGeese	0.360	**0.540**	0.520	0.400	0.420	0.400
EthanolConcentration	0.289	0.274	0.297	0.243	**0.308**	**0.308**
FaceDetection	0.510	0.508	0.505	**0.516**	0.497	0.505
FingerMovements	0.480	0.480	0.480	0.530	0.470	**0.540**
HandMovementDirection	0.324	**0.405**	0.257	0.324	0.230	0.257
Heartbeat	0.751	**0.761**	0.717	0.756	0.737	0.732
JapaneseVowels	0.978	**0.986**	0.978	0.970	0.978	0.978
MotorImagery	0.480	0.500	0.500	**0.520**	0.500	0.500
PhonemeSpectra	0.263	0.258	**0.269**	**0.269**	0.260	0.257
SelfRegulationSCP1	0.778	0.768	**0.788**	0.761	0.730	0.771
SelfRegulationSCP2	0.467	0.550	**0.561**	0.528	0.511	0.511
SpokenArabicDigits	0.973	**0.976**	0.966	0.964	0.968	0.957
StandWalkJump	0.467	0.467	**0.533**	0.200	0.133	**0.533**
Avg. over non-spatial datasets	0.523	**0.553**	0.536	0.508	0.496	0.537

In addition, to visually show the effect of differently regularised contrastive learning losses on representation, we apply t-SNE to compress the encoded representations into 3 dimensions, as plotted in Figure B1 for the dataset Epilepsy, and Figure B2 for

Table B3 Detailed representation training time per epoch (unit: s) on spatial datasets in the UEA archive.

Dataset	TS2Vec	Topo-TS2Vec	GGeo-TS2Vec	SoftCLT	Topo-SoftCLT	GGeo-SoftCLT
ArticularyWordRecognition	3.799 (1.00×)	5.61 (1.48×)	5.863 (1.54×)	3.772 (0.99×)	5.77 (1.52×)	5.983 (1.57×)
BasicMotions	0.475 (1.00×)	0.685 (1.44×)	0.709 (1.49×)	0.457 (0.96×)	0.687 (1.45×)	0.711 (1.50×)
CharacterTrajectories	20.640 (1.00×)	30.863 (1.50×)	33.32 (1.61×)	20.652 (1.00×)	30.948 (1.50×)	33.18 (1.61×)
Cricket	1.903 (1.00×)	2.653 (1.39×)	5.437 (2.86×)	1.904 (1.00×)	2.655 (1.40×)	5.436 (2.86×)
ERing	0.319 (1.00×)	0.482 (1.51×)	0.487 (1.53×)	0.316 (0.99×)	0.483 (1.51×)	0.49 (1.54×)
EigenWorms	19.862 (1.00×)	23.823 (1.20×)	149.05 (7.50×)	20.224 (1.02×)	24.856 (1.25×)	150.7 (7.59×)
Epilepsy	1.737 (1.00×)	2.49 (1.43×)	2.753 (1.58×)	1.686 (0.97×)	2.506 (1.44×)	2.755 (1.59×)
Handwriting	1.875 (1.00×)	2.771 (1.48×)	2.959 (1.58×)	1.88 (1.00×)	2.775 (1.48×)	2.987 (1.59×)
LSST	29.786 (1.00×)	45.273 (1.52×)	45.162 (1.52×)	29.859 (1.00×)	45.216 (1.52×)	45.154 (1.52×)
Libras	2.081 (1.00×)	3.142 (1.51×)	3.142 (1.51×)	2.085 (1.00×)	3.135 (1.51×)	3.141 (1.51×)
NATOPS	1.953 (1.00×)	2.989 (1.53×)	2.949 (1.51×)	2.085 (1.07×)	3.147 (1.61×)	3.159 (1.62×)
PEMS-SF	3.413 (1.00×)	5.069 (1.49×)	5.38 (1.58×)	3.415 (1.00×)	5.064 (1.48×)	5.399 (1.58×)
RacketSports	1.781 (1.00×)	2.685 (1.51×)	2.664 (1.50×)	1.771 (0.99×)	2.711 (1.52×)	2.665 (1.50×)
UWaveGestureLibrary	1.699 (1.00×)	2.395 (1.41×)	2.788 (1.64×)	1.776 (1.05×)	2.595 (1.53×)	2.99 (1.76×)
Avg. over spatial datasets	6.523 sec/epoch	1.46×	2.12×	1.00×	1.48×	2.15×

Table B4 Detailed representation training time per epoch (unit: s) on non-spatial datasets in the UEA archive.

Dataset	TS2Vec	Topo-TS2Vec	GGeo-TS2Vec	SoftCLT	Topo-SoftCLT	GGeo-SoftCLT
AtrialFibrillation	0.182 (1.00×)	0.258 (1.42×)	0.369 (2.03×)	0.177 (0.97×)	0.259 (1.42×)	0.366 (2.01×)
DuckDuckGeese	0.617 (1.00×)	0.973 (1.58×)	1.059 (1.72×)	0.621 (1.01×)	0.968 (1.57×)	1.104 (1.79×)
EthanolConcentration	4.939 (1.00×)	6.655 (1.35×)	20.128 (4.08×)	4.89 (0.99×)	6.664 (1.35×)	20.182 (4.09×)
FaceDetection	70.709 (1.00×)	109.6 (1.55×)	108.83 (1.54×)	71.104 (1.01×)	107.523 (1.52×)	107.092 (1.51×)
FingerMovements	3.826 (1.00×)	5.67 (1.48×)	5.706 (1.49×)	3.779 (0.99×)	5.671 (1.48×)	5.716 (1.49×)
HandMovementDirection	2.221 (1.00×)	3.353 (1.51×)	4.151 (1.87×)	2.226 (1.00×)	3.334 (1.50×)	4.142 (1.86×)
Heartbeat	2.811 (1.00×)	4.218 (1.50×)	5.22 (1.86×)	2.818 (1.00×)	4.216 (1.50×)	5.221 (1.86×)
JapaneseVowels	3.211 (1.00×)	4.871 (1.52×)	4.821 (1.50×)	3.199 (1.00×)	4.846 (1.51×)	4.83 (1.50×)
MotorImagery	7.450 (1.00×)	9.637 (1.29×)	51.0 (6.85×)	7.475 (1.00×)	9.659 (1.30×)	50.881 (6.83×)
PhonemeSpectra	42.956 (1.00×)	63.801 (1.49×)	70.578 (1.64×)	43.015 (1.00×)	63.807 (1.49×)	70.806 (1.65×)
SelfRegulationSCP1	4.178 (1.00×)	6.042 (1.45×)	10.446 (2.50×)	4.237 (1.01×)	6.094 (1.46×)	10.415 (2.49×)
SelfRegulationSCP2	3.295 (1.00×)	4.67 (1.42×)	9.391 (2.85×)	3.269 (0.99×)	4.629 (1.40×)	9.376 (2.85×)
SpokenArabicDigits	96.299 (1.00×)	143.411 (1.49×)	131.577 (1.37×)	86.495 (0.90×)	125.916 (1.31×)	129.073 (1.34×)
StandWalkJump	0.304 (1.00×)	0.404 (1.33×)	1.68 (5.53×)	0.31 (1.02×)	0.400 (1.32×)	1.7 (5.59×)
Avg. over non-spatial datasets	17.357 sec/epoch	1.46×	2.57×	0.99×	1.44×	2.57×

RacketSports. The classes are indicated by colours. We use these two datasets because they are visualisation-friendly, with 4 classes and around 150 test samples.

B.2 Detailed results of macroscopic prediction with LSTM and GRU

This section provides additional tables and figures presenting the evaluation of the final results using LSTM and GRU in macroscopic traffic prediction. Table B5 shows the task-specific metrics and Table B6 shows the metrics for global structure preservation. Figures B3 and B4 show the latent representations encoded by LSTM and GRU models, respectively.

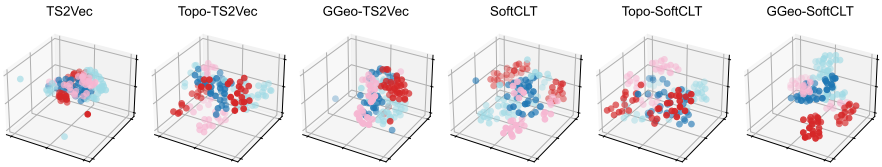


Figure B1 Encoded representations after training with different losses on the test set of Epilepsy.

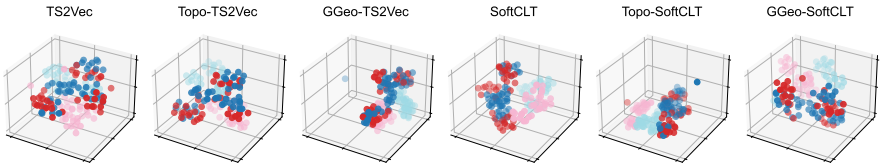


Figure B2 Encoded representations after training with different losses on the test set of RacketSports.

Table B5 Macroscopic traffic prediction evaluation with LSTM and GRU encoders. Metrics are reported as mean ± standard deviation over 10 stratified folds of the test set.

Method	LSTM				GRU			
	MAE (↓) (km/h)	RMSE (↓) (km/h)	SDEP (↓) (km/h)	EVar (↑) (%)	MAE (↓) (km/h)	RMSE (↓) (km/h)	SDEP (↓) (km/h)	EVar (↑) (%)
No pretraining	3.244 ± 0.046	6.401 ± 0.102	6.399 ± 0.102	82.151 ± 0.512	3.552 ± 0.063	7.227 ± 0.119	7.227 ± 0.119	77.241 ± 0.457
TS2Vec	3.158 ± 0.045	6.187 ± 0.089	6.186 ± 0.089	83.322 ± 0.413	3.601 ± 0.062	7.296 ± 0.113	7.296 ± 0.113	76.805 ± 0.427
Topo-TS2Vec	3.139 ± 0.048	6.154 ± 0.107	6.153 ± 0.107	83.499 ± 0.423	3.491 ± 0.057	7.005 ± 0.108	7.005 ± 0.108	78.617 ± 0.442
GGeo-TS2Vec	**3.101** ± 0.046	**5.974** ± 0.089	**5.973** ± 0.089	**84.454** ± 0.354	3.466 ± 0.055	6.947 ± 0.105	6.946 ± 0.105	78.975 ± 0.417
SoftCLT	3.191 ± 0.049	6.319 ± 0.100	6.318 ± 0.100	82.604 ± 0.412	**3.349** ± 0.051	**6.649** ± 0.101	**6.648** ± 0.101	**80.741** ± 0.388
Topo-SoftCLT	3.192 ± 0.049	6.270 ± 0.101	6.270 ± 0.101	82.872 ± 0.317	3.367 ± 0.049	6.660 ± 0.098	6.659 ± 0.098	80.674 ± 0.398
GGeo-SoftCLT	**3.128** ± 0.044	**6.121** ± 0.094	**6.120** ± 0.094	**83.675** ± 0.400	**3.302** ± 0.049	**6.494** ± 0.097	**6.494** ± 0.097	**81.621** ± 0.430
Best improvement	4.415	6.675	6.667	2.803	7.026	10.139	10.143	5.671

Note: the **best** values are both bold and underlined; the **second-best** values are bold.

Table B6 Global structure preservation of LSTM and GRU encoders in macroscopic traffic prediction task.

Method	LSTM					GRU				
	kNN (↑)	Trust. (↑)	Cont. (↑)	MRRE (↓)	dRMSE (↓)	kNN (↑)	Trust. (↑)	Cont. (↑)	MRRE (↓)	dRMSE (↓)
No pretraining	**0.215**	0.929	0.942	0.046	0.319	0.207	0.946	0.951	0.037	**0.299**
TS2Vec	0.173	**0.959**	**0.955**	**0.035**	**0.318**	0.197	**0.963**	0.957	0.031	**0.269**
Topo-TS2Vec	0.179	0.947	0.945	0.042	0.332	0.193	**0.963**	**0.962**	0.030	0.304
GGeo-TS2Vec	**0.198**	**0.968**	**0.961**	**0.029**	**0.299**	0.203	0.962	**0.962**	**0.029**	0.306
SoftCLT	0.186	0.954	0.948	0.038	0.334	0.214	0.958	0.957	0.031	0.319
Topo-SoftCLT	0.193	0.947	0.950	0.039	0.326	**0.239**	**0.966**	**0.964**	**0.026**	0.354
GGeo-SoftCLT	0.176	0.945	0.942	0.042	0.365	**0.228**	**0.963**	0.957	**0.029**	0.339

Note: the **best** values are both bold and underlined; the **second-best** values are bold.

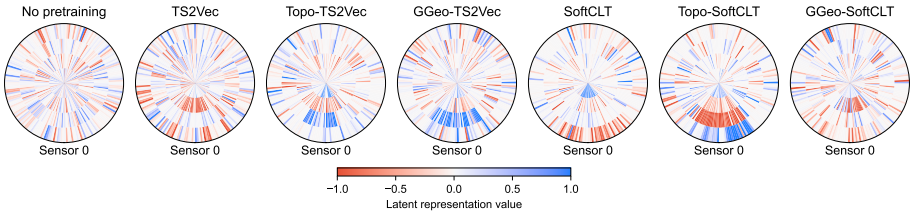


Figure B3 LSTM encoded representations after training with different losses on the test set of the Macroscopic traffic prediction task.

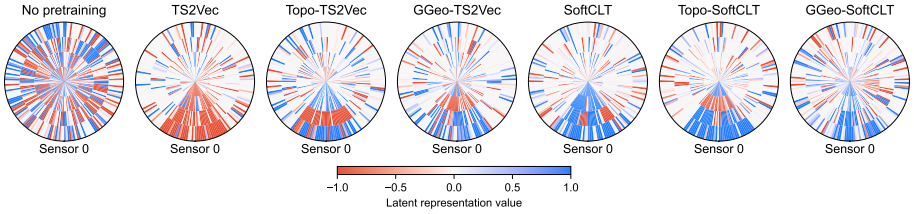


Figure B4 GRU encoded representations after training with different losses on the test set of the Macroscopic traffic prediction task.

Chapter 7

Conclusion

A context-aware, generalisable, and scalable methodology to quantify the collision risk in multi-directional traffic interactions is developed in this thesis, and is demonstrated with state-of-the-art accuracy and all-scenario applicability. This concluding chapter first summarises the main findings by answering the research questions (RQs) that address the identified knowledge gaps. Limitations of the current work are then discussed, shedding light on future research. Finally, this chapter outlines the potential practical implications and societal relevance of this thesis for the proactive improvement of traffic safety.

7.1 Main findings

RQ1. How can multi-directional traffic interactions be effectively characterised?

This question is answered in Chapter 2 by *introducing a coordinate transformation technique from the perspective of relative movement* and *proposing a method to infer the conditionally averaged multi-directional spacing in urban traffic from trajectory data.*

The coordinate transformation technique establishes a normalised reference system to consistently measure the multi-directional spacing between road users in a two-dimensional plane. By accumulating the spacings of numerous road users with varied orientations at different relative speeds, an empirical relation is identified and termed as the interaction Fundamental Diagram (iFD). This relation describes the average road space required for interactions, exhibiting consistency when tested with real-world urban trajectories in different locations. Lateral interactions, such as merging, turning, and crossing at urban intersections, are demonstrated to make more efficient use of road space than longitudinally following interactions. When collectively examining the iFD at a specific intersection, an optimal traffic state exists where the interaction efficiency is maximised. Therefore, the multi-directional spacing and iFD, respectively at the microscopic and macroscopic level, characterise multi-directional traffic interactions.

RQ2. How can traffic conflict detection accommodate varying contextual factors?

This question is answered in Chapters 3 and 4 by *developing a unified framework to quantify collision risk based on conditional spacing*, *in which the posterior probability of a traffic conflict is designed to be conditioned on its interaction context.*

This thesis uses a conditional formulation of traffic conflict detection based on the spacing between road users dependent on their interaction context. This is first proposed

in Chapter 3, where the only contextual factor is relative speed, and the conditional probabilities are used to adaptively optimise the triggering thresholds for forward collision warning. The adaptive method achieves a better balance between missed and false warnings than traditional heuristic threshold selection, for both synthetic and real-world traffic conflicts. The conditional formulation is then further developed in Chapter 4 as quantifying the probabilistic risk of potential collisions conditioned on their interaction contexts, which can be described by various contextual factors that are measured or encoded as quantifiable variables. In the same interaction context, a smaller spacing monotonically indicates a higher collision risk. Built upon the formulation, a unified probabilistic framework for traffic conflict detection is established. This unified framework enables any contextual factors, such as road user behaviour and environmental change, to be incorporated as conditioning variables.

RQ3. How can collision risk in diverse multi-directional interaction scenarios be quantified in a unified way?

This question is answered in Chapter 4 by *proposing and demonstrating a statistical learning pipeline to apply the unified probabilistic framework*, *which integrates the multi-directional spacing measurement* in answering RQ1 and *the conditional spacing-based formulation of traffic conflict detection* in answering RQ2.

The unified probabilistic framework for traffic conflict detection considers potential collisions as unsafe events that deviate from normal interactions to different extents, with contexts varying by interaction scenario. To apply this framework, collision risk quantification is decomposed into a sequence of data-driven statistical learning tasks to represent interaction contexts, infer typical spacing distributions in different contexts, and assess the extent to which a given interaction is in an unsafe state within its typical distribution. A unified metric of collision risk is trained following these tasks. It is adaptive to diverse situations without specifying separate metrics for each case, and is at least as effective as any predefined scenario-specific metrics. In addition, the quantification results present a long-tailed distribution of conflict intensity, suggesting the ability of this approach to comprehensively reflect a continuum of risk levels from minor conflicts to near-crashes. Consequently, such a unified and data-driven approach achieves generalisable collision risk quantification, which remains consistent across interaction scenarios and environments.

RQ4. How can collision risk quantification be scaled up without annotated data of crashes or near-crashes?

This question is answered in Chapters 5 and 6 by *extrapolating potential collisions from normal interactions* and *exploring a self-supervised learning approach to leverage abundant data of everyday road traffic interactions*.

Extrapolating from normal interactions to potential collisions in safety-critical situations, the generalised surrogate safety measure (GSSM) is proposed in Chapter 5. This is theoretically based on the unified probabilistic framework that answers RQ3. GSSM provides a self-supervised approach to quantifying collision risk without requiring

crashes or manual labels of near-crashes for training. It can be trained on large amounts of naturalistic driving data, learn the underlying patterns of traffic interactions, and infer collision risk based on the learnt patterns. Its scalability is demonstrated by the increasing accuracy of collision risk quantification as a larger variety of normal interaction patterns are incorporated. The self-supervised learning of traffic interaction patterns is further explored in Chapter 6. Encoding spatial time series using neural networks is improved by a contrastive learning method to enforce that the latent representations preserve fine-grained structural similarities (e.g., trajectories with similar dynamics remain close in feature space). This structure preservation facilitates learning more effective patterns and thus better performance in downstream tasks. Altogether, by self-supervised learning from normal interactions, GSSM enables scalable collision risk quantification that can be continuously improved based on abundant data without the need for manual annotations.

In summary, this thesis has established a consistent measure for multi-directional traffic interactions, developed a theoretical framework for identifying traffic conflicts across varying contexts, and delivered a self-supervised learning approach for proactive collision risk quantification. The findings note that there is no hard boundary between safe and unsafe interactions. Instead, dynamic and context-aware methods are required to reliably identify potential collisions across varying interaction contexts. By learning multi-directional interaction patterns from large-scale naturalistic data, rather than crash records, traffic safety can be sustainably improved without waiting for accidents.

7.2 Limitations and future research

A thesis inevitably has its limitations due to insufficiently available data, assumptions in model development, or the specific focus of research, etc. Acknowledging these limitations is important to chart a way forward. Each limitation discussed in the following sheds light on a question or challenge that remains to be addressed. Addressing them in future research will advance the contribution of proactive collision risk quantification to real-world traffic safety.

Empirical data availability

- **Limited ground-truth of crash and near-crash data**.
  This thesis uses two reconstructed datasets of crashes and near-crashes for validation. The data collection and reconstruction of these long-tailed safety-critical events could have unknown bias. This results in a limited coverage of risky situations, and the validation of collision risk quantification may be undermined. Future research should use datasets with a more comprehensive coverage of crashes and near-crashes for method validation.
- **Limited interactions with non-vehicle road users**.
  The datasets used for model calibration and training mainly cover the interactions between motor vehicles. As a result, the interactions involving non-vehicle road users, such as pedestrians and cyclists, are not sufficiently modelled or evaluated. This insufficiency is also recognised in [274]. Future research is recommended to use

interaction data between various types of road users, and accordingly adapt model development to account for their distinct behaviours and safety implications.

- **Limited contextual information**.
  The collision risk quantification in this thesis incorporates a limited range of contextual factors, focusing on the movement of road users, with only one dataset containing weather, road surface, and lighting. Not considering important contextual information, such as road layout, may lead to inaccurate evaluation and risk attribution can be misleading [275, 276]. For future research, more comprehensive data such as temperature, humidity, road surface friction, traffic signal timing, etc., should be collected and used. This would provide more information for assessing collision risk in varied real-world conditions.

Modelling assumptions and simplifications

- **Driver space as an observable result of interaction**.
  The driver space and multi-directional spacing in this thesis are visualised as a space around the ego (subject) vehicle. This space is neither maintained by the ego vehicle alone nor by other surrounding road users. Instead, it is a result of their mutual responses during road use interactions. To know more about how individual road users preserve a safe distance from others, future research will require subjective experiments like [277] to distinguish intended behaviour from observed phenomena.
- **Spacing proximity as a proxy for collision risk**.
  The collision risk quantification framework in this thesis uses spacing as a proxy for interaction intensity and assumes that spatial proximity indicates collision risk. This proxy settles a consistent reference across interaction contexts, but omits to consider collision severity or time of occurrence, which are important aspects of collision risk [278, 279]. In future model development, the time gap or speed could serve as a proxy, or a combination of proxies can be considered. These considerations are expected to capture more aspects of collision risk beyond what spacing alone can reveal.
- **Parametric distribution for spacing**.
  This thesis parameterises the conditional spacing distributions by assuming a particular mathematical form, such as the lognormal distribution. However, this assumption is based on empirical experience in the literature, while its validity has not been rigorously verified. Future investigation is needed for a better spacing distribution assumption. Meanwhile, parametric distribution may not capture the true complexity of traffic interaction patterns, especially when real-world spacing data deviates from the assumption. Therefore, future research may also explore non-parametric techniques to learn the conditional spacing patterns.

Underexplored aspects of collision risk

- **Optimisation of neural network models and their training**.
  This thesis uses neural network models to learn conditional spacing patterns. Since the research focus is on proposing a new methodology for data-driven collision risk

quantification and demonstrating its effectiveness, the optimal architecture and training of the neural networks were not pursued. More efficient neural network architectures and training procedures remain to be investigated in future research for performance improvement.

- **Collision severity and occurrence time**.
  This thesis considers collision risk only from a perspective of probability. However, the severity of a potential collision and the time of its occurrence are equally important. In future research, collision risk quantification is expected to incorporate multiple dimensions, including probability, severity, and occurrence time. This direction of future research is also touched on when discussing the limitation of using spacing proximity as a proxy for collision risk.
- **Causal influence of risk factors**.
  The data-driven methodology developed in this thesis correlates collision risk with contextual factors such as spacing direction, speed, weather, etc. Correlation alone does not reveal the causes of traffic collisions [280, 281]. Future research needs to further investigate the causal chain of traffic conflicts and explore how traffic interactions gradually escalate into near-crashes or crashes. This will help not only to estimate the risk of potential collisions, but also to suggest effective interventions to break the chain of events leading to collisions [282].
- **Uncertainty in quantification**.
  The methods presented in this thesis do not indicate the uncertainty or confidence in the collision risk score or probability. This limits insight into the reliability of risk quantification in different conditions, which is important for safety-critical applications such as collision avoidance. Future research needs to incorporate uncertainty quantification in collision risk quantification. This can provide more transparent evaluation of model performance and allow for further improvements, ultimately delivering more reliable and trustworthy applications in reality [283].

Integration into autonomous driving

- **Systematic collision risk in modular autonomous driving**.
  The classical autonomous driving pipeline is composed of three sequential modules: perception, prediction&planning, and control. Although collision risk is not explicitly quantified in this pipeline, the research in this thesis is mainly concerned with the module of prediction and planning. However, collision risk is a system-wide challenge [284]. An error in perception, e.g., failing to detect a child darting into the road, can instantly increase the risk of a collision [285]. Mistakes in the control module, such as misjudging tire grip on wet roads, also lead to increased collision risk [286]. In a complex system like autonomous driving, the failure of any single component can cascade into serious collision risk.
- **Explicit collision risk quantification for end-to-end autonomous driving**.
  The industry is shifting from modular pipelines to end-to-end (e2e) learning that directly maps raw sensor data to motion planning or vehicle control [287]. In this

paradigm, the need for collision risk quantification becomes explicit, and this requires extensive future research.

- **Safety validation**. The most direct e2e approach, imitation learning, encodes all the input information within neural networks [288]. This trades information preservation for reduced transparency and difficulty in monitoring or correcting faults. Therefore, intermediate signals are still vital for interpretation [289]. In addition to common signals such as Bird's-Eye-View (BEV) maps, standards like ISO 21448 on safety of the intended functionality (SOTIF) [290] mandate explicit risk quantification for safety validation.
- **Risk mitigation**. As another rapidly advancing approach, reinforcement learning (RL) uses reward models to define its objectives. Using collisions as a sparse penalty (large negative reward) can lead to unstable training [291]. Collision risk, instead, can serve as a continuous and dense penalty for risk mitigation. This also allows for defining safety-constrained objectives and personalised risk preferences in future development. The article by Wayve [292] similarly emphasises continuous risk assessment and proactive mitigation as part of the core objective function for e2e autonomous driving.
- **Data curation**. World models serve sophisticated training environments for RL [293–295]. Leaving aside the immense challenge in generating realistic and dynamic world representations, a core question is: what kind of virtual world fosters efficient and safe autonomous driving? The consensus is that it must include a significant number of safety-critical scenarios [287, 296]. The generation or identification of these scenarios requires explicit collision risk quantification. In essence, quantifying collision risk is equivalent to quantifying the intensity of driving interactions, enabling better curation of diverse real and synthetic training data and fine-tuning autonomous driving performance.

- **Open benchmark**.
  Collision risk quantification has been an overlooked challenge in autonomous driving. It does not directly contribute to the classical pipelines, suffers from a vague definition with no ground truth, and has been plagued by the scarcity of relevant data. This has prevented the creation of a widely used, systematically maintained open benchmark. In contrast, tasks like perception, prediction, and e2e driving have seen landmark models emerged (e.g., BEVFormer [297], VectorNet [298], UniAD [299]) on open benchmarks such as nuScenes [300], Waymo Open Dataset [301], and nuPlan [302]. As the shift toward e2e autonomous driving continues, explicit, consistent, and unified risk quantification becomes necessary for its training, validation, and regulatory compliance. Therefore, future research needs to clearly define the task, standardise datasets, and establish evaluation metrics, to create open benchmarks and build a sustainable research ecosystem. This is a necessary path to reach a truly safer, more reliable, and more intelligent future for autonomous driving.

7.3 Practical and societal implications

7.3.1 Practical applications

This thesis contributes to addressing the particular challenges in proactively quantifying collision risk in multi-directional urban traffic. The methodological advances and findings can be translated into practical applications. An overview of the translation from theory to practice is referred to Figure 1.3 on Page 7. More specifically, use cases in traffic management, Advanced Driver Assistance System (ADAS), and Autonomous Driving System (ADS) are discussed below.

- **Empirical evaluation of urban intersection efficiency**.
  Many research projects have been invested in to collect and analyse trajectory data for intelligent management of urban traffic, such as the Smart Intersections project [303] (https://sip.umtri.umich.edu) in the US and the SOTERIA project [304] (https://soteriaproject.eu) in the EU, both since 2022. Effectively characterising multi-directional traffic interactions supports trajectory-based evaluation of urban intersection performance. Traffic management agencies can use collected or simulated trajectory data to derive interaction fundamental diagrams (iFDs) and compare them across locations and time. For example, an iFD showing high road space required for the interactions at a specific intersection suggests its suboptimal efficiency. This can inform design improvements such as adjusting signal control, changing traffic rules, or redesigning road layout.
- **Reliable collision warning in ADAS**.
  Context-aware methods improve the accuracy and reliability of traffic conflict detection. In combination with multi-directional spacing, the collision risk in a wide range of driving scenarios can be quantified. These methods are useful for collision warning in ADAS, to alert not only to rear-end conflicts but also to potential collisions in any direction. Such omnidirectional collision avoidance is gaining attention in the industry, and is under development by, e.g., Nissan's Intersection ADAS [305], Ford's Co-Pilot360 [306], and Mobileye's Surround ADAS [307]. The methods proposed in this thesis can reduce false and missed warnings as guaranteed in theory and validated by real-world data. In the long run, more reliable collision warning can increase the trust of human drivers in the system, which in turn more effectively assists driving and contributes to everyday traffic safety.
- **Safety-focused training and validation of autonomous driving**.
  ADS need to cope with safety-critical events. These events are relatively rare in naturalistic road tests [50], making it difficult to accumulate a sufficiently large amount of data for training or validation. The unified conflict detection framework and self-supervised learning of collision risk can help address this difficulty. One way is to build an evolving cycle of data collection and curation, such as Tesla's Data Engine [308]. With collision risk evaluation, various levels of safety-critical interactions can be filtered from continuously collected driving data. Further, the evaluation score can be used to guide the generation of virtual interactions at different collision risks, which is being actively studied [296, 309]. Learning collision risk from

data also enables a context-adaptive objective to reinforce the focus on safety in training autonomous driving algorithms. All these facilitate training and validating safety-focused autonomous driving.

- **Real-time all-scenario traffic safety monitoring**.
  A variety of traffic data are increasingly collected through connected vehicles and sensors in cities. Based on these data, traffic management centres can utilise the scalable and generalisable collision risk quantification to identify potential hazards in a broad range of interaction scenarios. A preliminary example is the Smart Roadways system provided by Drivewyze [310] for truck drivers in multiple states in the US. The quantification methods developed in this thesis do not require historical crash records or manual annotations of collision risk, but learning from everyday interactions in various contexts. It allows real-time estimation of collision risk and provides opportunities for early responses. For instance, if high risks are flagged for multiple interactions at a location, measures such as temporary traffic speed control can be taken to prevent potential accidents. Such proactive monitoring and intervention can help to eliminate traffic fatalities by addressing risks early.

7.3.2 Societal relevance

Road safety remains a pressing societal challenge, and there is a growing consensus on the need to anticipate and mitigate traffic accidents before they occur, rather than reacting afterwards. Leading initiatives such as Vision Zero [311] have been adopted in many countries, underscoring proactive safety and ambitiously aiming at zero road fatalities. Meanwhile, investments in intelligent transportation systems have also been growing [312], such as the 5G C-V2X (Cellular Vehicle-to-Everything) technology in China [313]. Automated vehicles and smart infrastructure can now collect high-quality data on traffic interactions, providing unprecedented opportunities to understand and improve traffic safety.

Responding to the global ambition and trend, this thesis provides data-driven tools to quantify the risk of potential collisions in multi-directional traffic interactions, enabling self-supervised learning of collision risk from extensive naturalistic traffic interactions rather than waiting for crashes to occur. This is particularly relevant to ADAS and autonomous driving by equipping vehicles to anticipate and thus avoid potential collisions in complex and highly interactive urban environments. Preventing collisions directly reduces fatalities, injuries, and associated economic costs. Fewer collisions also mean less congestion due to disruption, contributing to more efficient transport of people and goods. Improved safety assurance will facilitate public acceptance of autonomous driving, which in turn could bring broad societal benefits. In addition to vehicles, this thesis has broader implications for traffic management and policy evaluation in pursuit of safer roads. Policy makers can use proactive risk indicators as performance measures for road safety interventions. For example, the impact of a new policy, such as lowering speed limits or installing a new roundabout, can be evaluated without waiting to see if accidents are reduced, but by learning from everyday observations to see if collision risk declines. This

shorter feedback loop allows the effectiveness of a policy to be assessed more quickly and at lower cost.

Overall, these implications align with the ambitious road safety initiatives by providing practical ways to move toward their goals. The societal relevance of this thesis thus lies in demonstrating how proactive, data-driven, and self-supervised collision risk quantification can be embedded in daily traffic operations, allowing interventions before, rather than after, accidents occur. Certainly, risk quantification is not a stand-alone solution, but an integral component of the ecosystem of intelligent transportation technologies. Parallel developments in, e.g., autonomous driving, V2X communication, and high-fidelity digital twins, are expected to synergistically enhance road safety. Taken together, the developments will accelerate responsible progress that increases public trust in transportation innovations, reduces road traffic injuries and fatalities, and improves the collective efficiency of urban traffic. More broadly, these expected outcomes are in line with the long-term evolution of urban mobility systems to realise the United Nations Sustainable Development Goal of inclusive, safe, resilient, and sustainable cities.

Bibliography

[1] Safety and Mobility (SAM) and Social Determinants of Health (SDH). *Global status report on road safety 2023*. Geneva: World Health Organization, 2023.

[2] W. Wijnen, W. Weijermars, A. Schoeters, W. van den Berghe, R. Bauer, L. Carnis, R. Elvik and H. Martensen. 'An analysis of official road crash cost estimates in European countries'. In: *Safety Science* 113 (2019), pp. 318–327. DOI: 10.1016/j.ssci.2018.12.004.

[3] H. Tan, F. Zhao, H. Hao and Z. Liu. 'Cost analysis of road traffic crashes in China'. In: *International Journal of Injury Control and Safety Promotion* 27.3 (2020), pp. 385–391. DOI: 10.1080/17457300.2020.1785507.

[4] T. Bougna, G. Hundal and P. Taniform. 'Quantitative Analysis of the Social Costs of Road Traffic Crashes Literature'. In: *Accident Analysis & Prevention* 165 (2022), p. 106282. DOI: 10.1016/j.aap.2021.106282.

[5] United Nations Economic Commission for Europe (UNECE). *Road traffic fatalities, injuries and accidents involving injury by location*. URL: https://w3.unece.org/PXWeb2015/pxweb/en/STAT/STAT__40-TRTRANS__01-TRACCIDENTS/02_en_TRAcclocation_r.px/ (visited on 22/03/2025).

[6] Dutch Institute for Road Safety Research (SWOV). *National road crash registration (BRON)*. URL: https://swov.nl/en/data/crashes (visited on 22/03/2025).

[7] National Highway Traffic Safety Administration (NHTSA), U.S. Department of Transportation. *Fatality and Injury Reporting System Tool (FIRST)*. URL: https://cdan.dot.gov/query (visited on 22/03/2025).

[8] C. Hydén. 'The development of a method for traffic safety evaluation: The Swedish Traffic Conflicts Technique'. PhD thesis. Lund University, 1987.

[9] A. P. Tarko. 'Use of crash surrogates and exceedance statistics to estimate road safety'. In: *Accident Analysis & Prevention* 45 (Mar. 2012), pp. 230–240. DOI: 10.1016/j.aap.2011.07.008.

[10] J. Pinnow, M. Masoud, M. Elhenawy and S. Glaser. 'A review of naturalistic driving study surrogates and surrogate indicator viability within the context of different road geometries'. In: *Accident Analysis & Prevention* 157 (2021), p. 106185. DOI: 10.1016/j.aap.2021.106185.

[11] L. Westhofen, C. Neurohr, T. Koopmann, M. Butz, B. Schütt, F. Utesch, B. Neurohr, C. Gutenkunst and E. Böde. 'Criticality metrics for automated driving: A review and suitability analysis of the state of the art'. In: *Archives of Computational Methods in Engineering* 30.1 (2022), pp. 1–35. DOI: 10.1007/s11831-022-09788-7.

[12] J. M. Hankey, M. A. Perez and J. A. McClafferty. *Description of the SHRP 2 Naturalistic Database and the Crash, Near-Crash, and Baseline Data Sets*. Tech. rep. Virginia Tech Transportation Institute, 2016.

[13] P. Cooper. 'Reports from group discussions'. In: *Proceedings of the first workshop on traffic conflicts*. Oslo, Norway, Sept. 1977.

[14] G. Markkula, R. Madigan, D. Nathanael, E. Portouli, Y. M. Lee, A. Dietrich, J. Billington, A. Schieben and N. Merat. 'Defining interactions: a conceptual framework for understanding interactive behaviour in human and automated road traffic'. In: *Theoretical Issues in Ergonomics Science* 21.6 (2020), pp. 728–752. DOI: 10.1080/1463922x.2020.1736686.

[15] S. P. Hoogendoorn and P. H. L. Bovy. 'State-of-the-art of vehicular traffic flow modelling'. In: *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering* 215.4 (2001), pp. 283–303. DOI: 10.1177/095965180121500402.

[16] H. Sadid and C. Antoniou. 'Modelling and simulation of (connected) autonomous vehicles longitudinal driving behavior: A state-of-the-art'. In: *IET Intelligent Transport Systems* 17.6 (2023), pp. 1051–1071. DOI: 10.1049/itr2.12337.

[17] J. Zhao, V. L. Knoop and M. Wang. 'Microscopic Traffic Modeling Inside Intersections: Interactions Between Drivers'. In: *Transportation Science* (2022). DOI: 10.1287/trsc.2022.1163.

[18] S. P. Venthuruthiyil and M. Chunchu. 'Anticipated Collision Time (ACT): A two-dimensional surrogate safety indicator for trajectory-based proactive safety assessment'. In: *Transportation Research Part C: Emerging Technologies* 139 (2022), p. 103655. DOI: 10.1016/j.trc.2022.103655.

[19] L. Zheng, K. Ismail and X. Meng. 'Traffic conflict techniques for road safety analysis: open questions and some insights'. In: *Canadian Journal of Civil Engineering* 41.7 (2014), pp. 633–641. DOI: 10.1139/cjce-2013-0558.

[20] A. P. Tarko. 'Surrogate measures of safety'. In: *Safe mobility: challenges, methodology and solutions*. Ed. by D. Lord and S. Washington. Vol. 11. Leeds: Emerald Publishing Limited, 2018, pp. 383–405. DOI: 10.1108/s2044-994120180000011019.

[21] J. C. Hayward. 'Near miss determination through use of a scale of danger'. In: *51st Annual Meeting of the Highway Research Board* 384 (1972), pp. 24–34.

[22] B. L. Allen, B. T. Shin and P. J. Cooper. 'Analysis of traffic conflicts and collisions'. In: *Transportation Research Record* 667 (1978), pp. 67–74.

[23] A. Laureshyn, Å. Svensson and C. Hydén. 'Evaluation of traffic safety, based on micro-level behavioural data: Theoretical framework and first implementation'. In: *Accident Analysis & Prevention* 42.6 (2010), pp. 1637–1646. DOI: 10.1016/j.aap.2010.03.021.

[24] K. D. Kusano, R. Chen, J. Montgomery and H. C. Gabler. 'Population distributions of time to collision at brake application during car following from naturalistic driving data'. In: *Journal of Safety Research* 54 (2015), 95.e29–104. DOI: 10.1016/j.jsr.2015.06.011.

[25] A. Tageldin and T. Sayed. 'Models to evaluate the severity of pedestrian-vehicle conflicts in five cities'. In: *Transportmetrica A: Transport Science* 15.2 (2019), pp. 354–375. DOI: 10.1080/23249935.2018.1477853.

[26] R. Chauhan, A. Dhamaniya and S. Arkatkar. 'Challenges in rear-end conflict-based safety assessment of highly disordered traffic conditions'. In: *Transportation Research Record* 2677.2 (2022), pp. 624–634. DOI: 10.1177/03611981221108156.

[27] D. Beitel, J. Stipancic, K. Manaugh and L. Miranda-Moreno. 'Assessing safety of shared space using cyclist-pedestrian interactions and automated video conflict analysis'. In: *Transportation Research Part D: Transport and Environment* 65 (2018), pp. 710–724. DOI: 10.1016/j.trd.2018.10.001.

[28] H. Cheng, Y. Jiang, H. Zhang, K. Chen, H. Huang, S. Xu, J. Wang and S. Zheng. 'Emergency Index (EI): A two-dimensional surrogate safety measure considering vehicles' interaction depth'. In: *Transportation Research Part C: Emerging Technologies* 171 (2025), p. 104981. DOI: 10.1016/j.trc.2024.104981.

[29] Y. Kuang, X. Qu and S. Wang. 'A tree-structured crash surrogate measure for freeways'. In: *Accident Analysis & Prevention* 77 (2015), pp. 137–148. DOI: 10.1016/j.aap.2015.02.007.

[30] N. Nadimi, A. M. Amiri and A. Sadri. 'Introducing novel statistical-based method of screening and combining currently well-known surrogate safety measures'. In: *Transportation Letters* 14.4 (2021), pp. 385–395. DOI: 10.1080/19427867.2021.1874184.

[31] A. Mazaheri, M. Saffarzadeh, N. Nadimi and S. S. Naseralavi. 'A revise on using surrogate safety measures for rear-end crashes'. In: *IATSS Research* 47.1 (2023), pp. 105–120. DOI: 10.1016/j.iatssr.2023.02.003.

[32] N. Formosa, M. Quddus, S. Ison, M. Abdel-Aty and J. Yuan. 'Predicting real-time traffic conflicts using deep learning'. In: *Accident Analysis & Prevention* 136 (2020), p. 105429. DOI: 10.1016/j.aap.2019.105429.

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

[34] L. C. Das and M. Won. 'SAINT-ACC: safety-aware intelligent adaptive cruise control for autonomous vehicles using deep reinforcement learning'. In: *Proceedings of the 38th International Conference on Machine Learning* 139 (2021). Ed. by M. Meila and T. Zhang, pp. 2445–2455.

[35] M. C. Panou. 'Intelligent personalized ADAS warnings'. In: *European Transport Research Review* 10.2 (2018), p. 59. DOI: 10.1186/s12544-018-0324-6.

[36] S. Das and A. K. Maurya. 'Defining Time-to-Collision Thresholds by the Type of Lead Vehicle in Non-Lane-Based Traffic Environments'. In: *IEEE Transactions on Intelligent Transportation Systems* 21.12 (2020), pp. 4972–4982. DOI: 10.1109/tits.2019.2946001.

[37] N. Nadimi, S. S. NaserAlavi and M. Asadamraji. 'Calculating dynamic thresholds for critical time to collision as a safety measure'. In: *Proceedings of the Institution of Civil Engineers - Transport* 175.7 (2022), pp. 403–412. DOI: 10.1680/jtran.19.00066.

[38] H. Farah and C. L. Azevedo. 'Safety analysis of passing maneuvers using extreme value theory'. In: *IATSS Research* 41.1 (Apr. 2017), pp. 12–21. DOI: 10.1016/j.iatssr.2016.07.001.

[39] A. Borsos, H. Farah, A. Laureshyn and M. Hagenzieker. 'Are collision and crossing course surrogate safety indicators transferable? A probability based approach using extreme value theory'. In: *Accident Analysis & Prevention* 143 (Aug. 2020), p. 105517. DOI: 10.1016/j.aap.2020.105517.

[40] D. Niu, T. Sayed, C. Fu and F. Mannering. 'A cross-comparison of different extreme value modeling techniques for traffic conflict-based crash risk estimation'. In: *Analytic Methods in Accident Research* 44 (2024), p. 100352. DOI: 10.1016/j.amar.2024.100352.

[41] J. Ayoub, Z. Wang, M. Li, H. Guo, R. Sherony, S. Bao and F. Zhou. 'Cause-and-effect analysis of ADAS: a comparison study between literature review and complaint data'. In: *Proceedings of the 14th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. New York, NY, USA, 2022, pp. 139–149. DOI: 10.1145/3543174.3547117.

[42] M. Althoff and J. M. Dolan. 'Online Verification of Automated Road Vehicles Using Reachability Analysis'. In: *IEEE Transactions on Robotics* 30.4 (2014), pp. 903–918. DOI: 10.1109/TRO.2014.2312453.

[43] R. McAllister, Y. Gal, A. Kendall, M. Van Der Wilk, A. Shah, R. Cipolla and A. Weller. 'Concrete problems for autonomous vehicle safety: advantages of Bayesian deep learning'. In: *Proceedings of the 26th International Joint Conference on Artificial Intelligence*. Melbourne, Australia, 2017, pp. 4745–4753.

[44] J. Kim and D. Kum. 'Collision risk assessment algorithm via lane-based probabilistic motion prediction of surrounding vehicles'. In: *IEEE Transactions on Intelligent Transportation Systems* 19.9 (2018), pp. 2965–2976. DOI: 10.1109/tits.2017.2768318.

[45] C. Li, Y. Qian, C. Sun, W. Yan, C. Wang and M. Yang. 'TTC4MCP: Monocular Collision Prediction Based on Self-Supervised TTC Estimation'. In: *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 2023, pp. 244–250. DOI: 10.1109/IROS55552.2023.10341966.

[46] F. B. Mathiesen, L. Romao, S. C. Calvert, L. Laurenti and A. Abate. 'A data-driven approach for safety quantification of non-linear stochastic systems with unknown additive noise distribution'. In: *arXiv preprint* (2024), arXiv:2410.06662.

[47] Q. Wang, D. Xu, G. Kuang, C. Lv, S. E. Li and B. Nie. 'Risk-Aware Vehicle Trajectory Prediction Under Safety-Critical Scenarios'. In: *IEEE Transactions on Intelligent Transportation Systems* 26.5 (2025), pp. 7156–7171. DOI: 10.1109/TITS.2025.3525744.

[48] H. Kataoka, T. Suzuki, S. Oikawa, Y. Matsui and Y. Satoh. 'Drive Video Analysis for the Detection of Traffic Near-Miss Incidents'. In: *2018 IEEE International Conference on Robotics and Automation (ICRA)*. 2018, pp. 3421–3428. DOI: 10.1109/ICRA.2018.8460812.

[49] J. Fang, J. Qiao, J. Xue and Z. Li. 'Vision-Based Traffic Accident Detection and Anticipation: A Survey'. In: *IEEE Transactions on Circuits and Systems for Video Technology* 34.4 (2024), pp. 1983–1999. DOI: 10.1109/TCSVT.2023.3307655.

[50] H. X. Liu and S. Feng. 'Curse of rarity for autonomous vehicles'. In: *Nature Communications* 15.1 (2024), p. 4808. DOI: 10.1038/s41467-024-49194-0.

[51] L. C. Edie. 'Discussion of traffic stream measurements and definitions'. In: New York: The Port of New York Authority, 1963, pp. 139–154.

[52] Y. Makigami, G. F. Newell and R. Rothery. 'Three-Dimensional Representation of Traffic Flow'. In: *Transportation Science* 5.3 (1971), pp. 302–313. DOI: 10.1287/trsc.5.3.302.

[53] B. D. Greenshields, J. T. Thompson, H. C. Dickinson and R. S. Swinton. 'The Photographic Method of Studying Traffic Behavior'. In: *Highway Research Board Proceedings*. Vol. 13. 1933, pp. 382–399.

[54] F. L. Hall, V. F. Hurdle and J. H. Banks. 'Synthesis of recent work on the nature of speed-flow and flow-occupancy (or density) relationships on freeways'. In: *Transportation Research Record* 1365 (1992), pp. 12–18.

[55] J. Castillo and F. Benítez. 'On the functional form of the speed-density relationship–I: General theory'. In: *Transportation Research Part B: Methodological* 29.5 (1995), pp. 373–389. DOI: 10.1016/0191-2615(95)00008-2.

[56] M. J. Cassidy. 'Bivariate relations in nearly stationary highway traffic'. In: *Transportation Research Part B: Methodological* 32.1 (1998), pp. 49–59. DOI: 10.1016/s0191-2615(97)00012-x.

[57] W.-L. Jin. 'A kinematic wave theory of lane-changing traffic flow'. In: *Transportation Research Part B: Methodological* 44.8-9 (2010), pp. 1001–1021. DOI: 10.1016/j.trb.2009.12.014.

[58] W.-L. Jin. 'A multi-commodity Lighthill-Whitham-Richards model of lane-changing traffic flow'. In: *Transportation Research Part B: Methodological* 57 (2013), pp. 361–377. DOI: 10.1016/j.trb.2013.06.002.

[59] P. Wagner, E. Brockfeld, N. H. Gartner and A. Sohr. 'Fundamental Diagram of Traffic Flows on Urban Roads'. In: *Transportation Research Record* 2124.1 (2009), pp. 213–221. DOI: 10.3141/2124-21.

[60] X. Wu, H. X. Liu and N. Geroliminis. 'An empirical analysis on the arterial fundamental diagram'. In: *Transportation Research Part B: Methodological* 45.1 (2011), pp. 255–266. DOI: 10.1016/j.trb.2010.06.003.

[61] W. Fourati and B. Friedrich. 'A method for using crowd-sourced trajectories to construct control-independent fundamental diagrams at signalized links'. In: *Transportation Research Part C: Emerging Technologies* 130 (2021), p. 103270. DOI: 10.1016/j.trc.2021.103270.

[62] A. Seyfried, B. Steffen, W. Klingsch and M. Boltes. 'The fundamental diagram of pedestrian movement revisited'. In: *Journal of Statistical Mechanics: Theory and Experiment* 2005.10 (2005), P10002. DOI: 10.1088/1742-5468/2005/10/p10002.

[63] J. Zhang, W. Klingsch, A. Schadschneider and A. Seyfried. 'Ordering in bidirectional pedestrian flows and its influence on the fundamental diagram'. In: *Journal of Statistical Mechanics: Theory and Experiment* 2012.02 (2012), P02002. DOI: 10.1088/1742-5468/2012/02/p02002.

[64] G. Flötteröd and G. Lämmel. 'Bidirectional pedestrian fundamental diagram'. In: *Transportation Research Part B: Methodological* 71 (2015), pp. 194–212. DOI: 10.1016/j.trb.2014.11.001.

[65] L. D. Vanumu, K. Ramachandra Rao and G. Tiwari. 'Fundamental diagrams of pedestrian flow characteristics: A review'. In: *European Transport Research Review* 9.4 (2017). DOI: 10.1007/s12544-017-0264-6.

[66] E. Barmpounakis and N. Geroliminis. 'On the new era of urban traffic monitoring with massive drone data: The pNEUMA large-scale field experiment'. In: *Transportation Research Part C: Emerging Technologies* 111 (2020), pp. 50–71. DOI: 10.1016/j.trc.2019.11.023.

[67] J. J. Gibson and L. E. Crooks. 'A theoretical field-analysis of automobile driving'. In: *The American Journal of Psychology* 51 (3 1938), pp. 453–471.

[68] R. Näätänen and H. Summala. 'A model for the role of motivational factors in drivers' decision-making'. In: *Accident Analysis & Prevention* 6.3 (1974), pp. 243–261. DOI: 10.1016/0001-4575(74)90003-7.

[69] H. Summala. 'Towards Understanding Motivational and Emotional Factors in Driver Behaviour: Comfort Through Satisficing'. In: *Modelling Driver Behaviour in Automotive Environments: Critical Issues in Driver Interactions with Intelligent Transport Systems*. Ed. by P. C. Cacciabue. London: Springer, 2007, pp. 189–207. DOI: 10.1007/978-1-84628-618-6_11.

[70] J. Bärgman, K. Smith and J. Werneke. 'Quantifying drivers' comfort-zone and dread-zone boundaries in left turn across path/opposite direction (LTAP/OD) scenarios'. In: *Transportation Research Part F: Traffic Psychology and Behaviour* 35 (2015), pp. 170–184. DOI: j.trf.2015.10.003.

[71] P. Marsh and P. Collett. 'The car as a weapon'. In: *ETC: A Review of General Semantics* 44.2 (1987), pp. 146–151.

[72] D. A. Hennessy, S. Howard and E. Carr. 'Driver space preference: differences across age, gender and traffic conditions'. In: *Traffic Psychology: An International Perspective*. Ed. by D. A. Hennessy. New York: Nova Science Publishers, Jan. 2011, pp. 233–250.

[73] M. Taieb-Maimon and D. Shinar. 'Minimum and comfortable driving headways: reality versus perception'. In: *Human Factors* 43.1 (2001), pp. 159–172. DOI: 10.1518/001872001775992543.

[74] B. Lewis-Evans, D. De Waard and K. A. Brookhuis. 'That's close enough—A threshold effect of time headway on the experience of risk, task difficulty, effort, and comfort'. In: *Accident Analysis & Prevention* 42.6 (2010), pp. 1926–1933. DOI: 10.1016/j.aap.2010.05.014.

[75] F. W. Siebert, M. Oehl and H.-R. Pfister. 'The influence of time headway on subjective driver states in adaptive cruise control'. In: *Transportation Research Part F: Traffic Psychology and Behaviour* 25 (2014), pp. 65–73. DOI: 10.1016/j.trf.2014.05.005.

[76] F. W. Siebert, M. Oehl, F. Bersch and H.-R. Pfister. 'The exact determination of subjective risk and comfort thresholds in car following'. In: *Transportation Research Part F: Traffic Psychology and Behaviour* 46 (2017), pp. 1–13. DOI: 10.1016/j.trf.2017.01.001.

[77] F. Camara and C. Fox. 'Space invaders: pedestrian proxemic utility functions and trust zones for autonomous vehicle interactions'. In: *International Journal of Social Robotics* 13 (2020), pp. 1929–1949. DOI: 10.1007/s12369-020-00717-x.

[78] M. P. Hennessey, C. Shankwitz and M. Donath. 'Sensor-based virtual bumpers for collision avoidance: configuration issues'. In: *Collision Avoidance and Automated Traffic Management Sensors*. Ed. by A. C. Chachich and M. J. de Vries. Vol. 2592. 1995, pp. 48–59. DOI: 10.1117/12.228920.

[79] M. S. Graziano and D. F. Cooke. 'Parieto-frontal interactions, personal space, and defensive behavior'. In: *Neuropsychologia* 44.6 (2006), pp. 845–859. DOI: 10.1016/j.neuropsychologia.2005.09.009.

[80] M. Novey, T. Adali and A. Roy. 'A Complex Generalized Gaussian Distribution–Characterization, Generation, and Estimation'. In: *IEEE Transactions on Signal Processing* 58.3 (2010), pp. 1427–1433. DOI: 10.1109/tsp.2009.2036049.

[81] T. Wang, H. Li, Z. Li and Z. Wang. 'A Fast Parameter Estimation of Generalized Gaussian Distribution'. In: *2006 8th international Conference on Signal Processing*. Vol. 1. 2006. DOI: 10.1109/icosp.2006.345546.

[82] J. Wang, X. Rui, X. Song, X. Tan, C. Wang and V. Raghavan. 'A novel approach for generating routable road maps from vehicle GPS traces'. In: *International Journal of Geographical Information Science* 29.1 (2014), pp. 69–91. DOI: 10.1080/13658816.2014.944527.

[83] L. Cao and J. Krumm. 'From GPS traces to a routable road map'. In: *Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*. ACM, Nov. 2009. DOI: 10.1145/1653771.1653776.

[84] R. Davidson and J. G. MacKinnon. 'Bootstrap tests: how many bootstraps?' In: *Econometric Reviews* 19.1 (2000), pp. 55–68. DOI: 10.1080/07474930008800459.

[85] Y. Yuan, W. Daamen, B. Goni Ros and S. Hoogendoorn. 'Investigating cyclist interaction behavior through a controlled laboratory experiment'. In: *Journal of Transport and Land Use (online)* 11.1 (2018), pp. 833–847. DOI: 10.5198/jtlu.2018.1155.

[86] S. C. Calvert, F. L. van Wageningen-Kessels and S. P. Hoogendoorn. 'Capacity drop through reaction times in heterogeneous traffic'. In: *Journal of Traffic and Transportation Engineering (English Edition)* 5.2 (2018), pp. 96–104. DOI: 10.1016/j.jtte.2017.07.008.

[87] K. Nishinari, M. Treiber and D. Helbing. 'Interpreting the wide scattering of synchronized traffic data by time gap statistics'. In: *Physical Review E* 68.6 (2003). DOI: 10.1103/physreve.68.067101.

[88] N. Geroliminis and J. Sun. 'Properties of a well-defined macroscopic fundamental diagram for urban traffic'. In: *Transportation Research Part B: Methodological* 45.3 (2011), pp. 605–617. DOI: 10.1016/j.trb.2010.11.004.

[89] X. Chen, Z. Li, L. Li and Q. Shi. 'Characterising scattering features in flow–density plots using a stochastic platoon model'. In: *Transportmetrica A: Transport Science* 10.9 (2013), pp. 820–848. DOI: 10.1080/23249935.2013.822941.

[90] W. Zhan, L. Sun, D. Wang, H. Shi, A. Clausse, M. Naumann, J. Kümmerle, H. Königshof, C. Stiller, A. de La Fortelle and M. Tomizuka. 'INTERACTION Dataset: An INTERnational, Adversarial and Cooperative moTION Dataset in Interactive Driving Scenarios with Semantic Maps'. In: *arXiv preprint* (2019), arXiv:1910.03088.

[91] T. Ma and S. Ahn. 'Comparisons of Speed-Spacing Relations under General Car following versus Lane Changing'. In: *Transportation Research Record* 2088.1 (2008), pp. 138–147. DOI: 10.3141/2088-15.

[92] H. Yeo and A. Skabardonis. 'Microscopic fundamental relationships between vehicle speed and spacing in view of asymmetric traffic theory'. In: *14th International IEEE Conference on Intelligent Transportation Systems (ITSC)*. 2011, pp. 1410–1414. DOI: 10.1109/itsc.2011.6082878.

[93] J. Montgomery, K. D. Kusano and H. C. Gabler. 'Age and Gender Differences in Time to Collision at Braking From the 100-Car Naturalistic Driving Study'. In: *Traffic Injury Prevention* 15 (2014), pp. 15–20. DOI: 10.1080/15389588.2014.928703.

[94] E. R. Hoffmann and R. G. Mortimer. 'Scaling of relative velocity between vehicles'. In: *Accident Analysis & Prevention* 28.4 (1996), pp. 415–421. DOI: 10.1016/0001-4575(96)00005-x.

[95] C. G. Wilmot and M. Khanal. 'Effect of Speed limits on speed and safety: A review'. In: *Transport Reviews* 19.4 (1999), pp. 315–329. DOI: 10.1080/014416499295420.

[96] A. Mukhtar, L. Xia and T. B. Tang. 'Vehicle Detection Techniques for Collision Avoidance Systems: A Review'. In: *IEEE Transactions on Intelligent Transportation Systems* 16.5 (2015), pp. 2318–2338. DOI: 10.1109/tits.2015.2409109.

[97] E. Dagan, O. Mano, G. Stein and A. Shashua. 'Forward collision warning with a single camera'. In: *IEEE Intelligent Vehicles Symposium, 2004*. Parma, Italy, 2004, pp. 37–42. DOI: 10.1109/ivs.2004.1336352.

[98] A. Polychronopoulos, M. Tsogas, A. J. Amditis and L. Andreone. 'Sensor fusion for predicting vehicles' path for collision avoidance systems'. In: *IEEE Transactions on Intelligent Transportation Systems* 8.3 (2007), pp. 549–562. DOI: 10.1109/tits.2007.903439.

[99] M. Kilicarslan and J. Y. Zheng. 'Predict Vehicle Collision by TTC From Motion Using a Single Video Camera'. In: *IEEE Transactions on Intelligent Transportation Systems* 20.2 (2019), pp. 522–533. DOI: 10.1109/tits.2018.2819827.

[100] A. Ghasemieh and R. Kashef. '3D object detection for autonomous driving: Methods, models, sensors, data, and challenges'. In: *Transportation Engineering* 8 (2022), p. 100115. DOI: 10.1016/j.treng.2022.100115.

[101] X. Ma, W. Ouyang, A. Simonelli and E. Ricci. '3D Object Detection From Images for Autonomous Driving: A Survey'. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2023), pp. 1–20. DOI: 10.1109/tpami.2023.3346386.

[102] K. Vogel. 'A comparison of headway and time to collision as safety indicators'. In: *Accident Analysis & Prevention* 35.3 (2003), pp. 427–433. DOI: 10.1016/s0001-4575(02)00022-2.

[103] F. Bella and R. Russo. 'A collision warning system for rear-end collision: a driving simulator study'. In: *Procedia Social and Behavioral Sciences* 20 (2011), pp. 676–686. DOI: 10.1016/j.sbspro.2011.08.075.

[104] M. H. Tawfeek and K. El-Basyouny. 'A perceptual forward collision warning model using naturalistic driving data'. In: *Canadian Journal of Civil Engineering* 45.10 (2018), pp. 899–907. DOI: 10.1139/cjce-2017-0592.

[105] C. Lu, X. He, H. van Lint, H. Tu, R. Happee and M. Wang. 'Performance evaluation of surrogate measures of safety with naturalistic driving data'. In: *Accident Analysis & Prevention* 162 (2021), p. 106403. DOI: 10.1016/j.aap.2021.106403.

[106] Z. Zhang, C. Lu, J. Li, Y. Xu, J. Lu and Z. Li. 'Prediction of Pedestrian Risky Level for Intelligent Vehicles'. In: *2020 IEEE Intelligent Vehicles Symposium (IV)*. 2020. DOI: 10.1109/iv47402.2020.9304707.

[107] R. Ezzati Amini, K. Yang and C. Antoniou. 'Development of a conflict risk evaluation model to assess pedestrian safety in interaction with vehicles'. In: *Accident Analysis & Prevention* 175 (2022), p. 106773. DOI: 10.1016/j.aap.2022.106773.

[108] X. Li, O. Oviedo-Trespalacios, A. Rakotonirainy and X. Yan. 'Collision risk management of cognitively distracted drivers in a car-following situation'. In: *Transportation Research Part F: Traffic Psychology and Behaviour* 60 (2019), pp. 288–298. DOI: 10.1016/j.trf.2018.10.011.

[109] R. Happee, C. Gold, J. Radlmayr, S. Hergeth and K. Bengler. 'Take-over performance in evasive manoeuvres'. In: *Accident Analysis & Prevention* 106 (2017), pp. 211–222. DOI: 10.1016/j.aap.2017.04.017.

[110] M. Zhu, Y. Wang, Z. Pu, J. Hu, X. Wang and R. Ke. 'Safe, efficient, and comfortable velocity control based on reinforcement learning for autonomous driving'. In: *Transportation Research Part C: Emerging Technologies* 117 (2020), p. 102662. DOI: 10.1016/j.trc.2020.102662.

[111] Q. Song, P. Runeson and S. Persson. 'A Scenario Distribution Model for Effective and Efficient Testing of Autonomous Driving Systems'. In: *37th IEEE/ACM International Conference on Automated Software Engineering*. Oct. 2022. DOI: 10.1145/3551349.3563239.

[112] M. Sieber and B. Farber. 'Driver perception and reaction in collision avoidance: Implications for ADAS development and testing'. In: *2016 IEEE Intelligent Vehicles Symposium (IV)*. June 2016. DOI: 10.1109/ivs.2016.7535392.

[113] A. Arun, M. M. Haque, A. Bhaskar, S. Washington and T. Sayed. 'A systematic mapping review of surrogate safety assessment using traffic conflict techniques'. In: *Accident Analysis & Prevention* 153 (2021), p. 106016. DOI: 10.1016/j.aap.2021.106016.

[114] D. Ruscio, M. R. Ciceri and F. Biassoni. 'How does a collision warning system shape driver's brake response time? The influence of expectancy and automation complacency on real-life emergency braking'. In: *Accident Analysis & Prevention* 77 (2015), pp. 72–81. DOI: 10.1016/j.aap.2015.01.018.

[115] K. Reinmueller and M. Steinhauser. 'Adaptive forward collision warnings: The impact of imperfect technology on behavioral adaptation, warning effectiveness and acceptance'. In: *Accident Analysis & Prevention* 128 (2019), pp. 217–229. DOI: 10.1016/j.aap.2019.04.012.

[116] J. P. Bliss and S. A. Acton. 'Alarm mistrust in automobiles: how collision alarm reliability affects driving'. In: *Applied Ergonomics* 34.6 (2003), pp. 499–509. DOI: 10.1016/j.apergo.2003.07.003.

[117] M. N. Lees and J. D. Lee. 'The influence of distraction and driving context on driver response to imperfect collision warning systems'. In: *Ergonomics* 50.8 (2007), pp. 1264–1286. DOI: 10.1080/00140130701318749.

[118] F. Naujoks, A. Kiesel and A. Neukum. 'Cooperative warning systems: The impact of false and unnecessary alarms on drivers' compliance'. In: *Accident Analysis & Prevention* 97 (2016), pp. 162–175. DOI: 10.1016/j.aap.2016.09.009.

[119] O. Zheng, M. Abdel-Aty, L. Yue, A. Abdelraouf, Z. Wang and N. Mahmoud. 'CitySim: A drone-based vehicle trajectory dataset for safety-oriented research and digital twins'. In: *Transportation Research Record* 2678.4 (2024), pp. 606–621. DOI: 10.1177/03611981231185768.

[120] C. Yuan, Y. Li, H. Huang, S. Wang, Z. Sun and Y. Li. 'Using traffic flow characteristics to predict real-time conflict risk: A novel method for trajectory data analysis'. In: *Analytic Methods in Accident Research* 35 (2022), p. 100217. DOI: 10.1016/j.amar.2022.100217.

[121] Y. Hu, Y. Li, H. Huang, J. Lee, C. Yuan and G. Zou. 'A high-resolution trajectory data driven method for real-time evaluation of traffic safety'. In: *Accident Analysis & Prevention* 165 (2022), p. 106503. DOI: 10.1016/j.aap.2021.106503.

[122] S. Ding, M. Abdel-Aty, Z. Wang and D. Wang. 'Insights into vehicle conflicts based on traffic flow dynamics'. In: *Scientific Reports* 14.1 (2024). DOI: 10.1038/s41598-023-50017-3.

[123] K. Custer. *100-Car Data*. VTTI, 2018. DOI: 10.15787/vtt1/ceu6rb.

[124] T. A. Dingus, S. G. Klauer, V. L. Neale, A. Petersen, S. E. Lee, J. Sudweeks, M. A. Perez, J. Hankey, D. Ramsey, S. Gupta, C. Bucher, Z. R. Doerzaph, J. Jermeland and R. R. Knipling. *The 100-Car Naturalistic Driving Study, Phase II - Results of the 100-Car Field Experiment*. Tech. rep. DOT-HS-810-593. United State: Department of Transportation. National Highway Traffic Safety Administration., 2006.

[125] A. Vahidi and A. Eskandarian. 'Research advances in intelligent collision avoidance and adaptive cruise control'. In: *IEEE Transactions on Intelligent Transportation Systems* 4.3 (2003), pp. 143–153. DOI: 10.1109/tits.2003.821292.

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

[127] P. Junietz, F. Bonakdar, B. Klamann and H. Winner. 'Criticality metric for the safety validation of automated driving using model predictive trajectory optimization'. In: *IEEE 21st International Conference on Intelligent Transportation Systems (ITSC)*. 2018, pp. 60–65. DOI: 10.1109/itsc.2018.8569326.

[128] A. Razi, X. Chen, H. Li, H. Wang, B. Russo, Y. Chen and H. Yu. 'Deep learning serves traffic safety analysis: A forward-looking review'. In: *IET Intelligent Transport Systems* 17.1 (2022), pp. 22–71. DOI: 10.1049/itr2.12257.

[129] N. Saunier, N. Mourji and B. Agard. 'Mining microscopic data of vehicle conflicts and collisions to investigate collision factors'. In: *Transportation Research Record* 2237 (2011), pp. 41–50. DOI: 10.3141/2237-05.

[130] K. Wu and P. P. Jovanis. 'Crashes and crash-surrogate events: Exploratory modeling with naturalistic driving data'. In: *Accident Analysis & Prevention* 45 (2012), pp. 507–516. DOI: 10.1016/j.aap.2011.09.002.

[131] N. Saunier, T. Sayed and C. Lim. 'Probabilistic collision prediction for vision-based automated road safety analysis'. In: *IEEE Intelligent Transportation Systems Conference*. Bellevue, WA, USA, Sept. 2007. DOI: 10.1109/itsc.2007.4357793.

[132] T. de Ceunynck. 'Defining and applying surrogate safety measures and behavioural indicators through site-based observations'. PhD thesis. Transport and Roads, Hasselt University, 2017.

[133] X. Wang, J. Alonso-Mora and M. Wang. 'Probabilistic risk metric for highway driving leveraging multi-modal trajectory predictions'. In: *IEEE Transactions on Intelligent Transportation Systems* 23.10 (2022), pp. 19399–19412. DOI: 10.1109/tits.2022.3164469.

[134] A. Broadhurst, S. Baker and T. Kanade. 'Monte Carlo road safety reasoning'. In: *Proceedings of IEEE Intelligent Vehicles Symposium*. Las Vegas, NV, USA, 2005. DOI: 10.1109/ivs.2005.1505122.

[135] H. Zhou and Z. Zhong. 'Evasive behavior-based method for threat assessment in different scenarios: A novel framework for intelligent vehicle'. In: *Accident Analysis & Prevention* 148 (2020), p. 105798. DOI: 10.1016/j.aap.2020.105798.

[136] D. F. Cooper and N. Ferguson. 'Traffic studies at T-junctions – A conflict simulation record'. In: *Traffic Engineering and Control* 17 (7 1976), pp. 306–309.

[137] N. Uno, Y. Iida, S. Itsubo and S. Yasuhara. 'A microscopic analysis of traffic conflict caused by lane-changing vehicle at weaving section'. In: *Proceedings of the 13th mini-EURO conference-handling uncertainty in the analysis of traffic and transportation systems*. Bari, Italy, 2002, pp. 10–13.

[138] Y. Li, J. Lu and K. Xu. 'Crash risk prediction model of lane-change behavior on approaching intersections'. In: *Discrete Dynamics in Nature and Society* 2017.1 (2017), p. 7328562. DOI: 10.1155/2017/7328562.

[139] C. Wang, Y. Xie, H. Huang and P. Liu. 'A review of surrogate safety measures and their applications in connected and automated vehicles safety modeling'. In: *Accident Analysis & Prevention* 157 (2021), p. 106157. DOI: 10.1016/j.aap.2021.106157.

[140] A. Arun, M. M. Haque, S. Washington, T. Sayed and F. Mannering. 'A systematic review of traffic conflict-based safety measures with a focus on application context'. In: *Analytic Methods in Accident Research* 32 (2021), p. 100185. DOI: 10.1016/j.amar.2021.100185.

[141] J. R. Ward, G. Agamennoni, S. Worrall, A. Bender and E. Nebot. 'Extending Time to Collision for probabilistic reasoning in general traffic scenarios'. In: *Transportation Research Part C: Emerging Technologies* 51 (2015), pp. 66–82. DOI: 10.1016/j.trc.2014.11.002.

[142] Y. Wang, Z. Li, P. Liu, C. Xu and K. Chen. 'Surrogate safety measures for traffic oscillations based on empirical vehicle trajectories prior to crashes'. In: *Transportation Research Part C: Emerging Technologies* 161 (2024), p. 104543. DOI: 10.1016/j.trc.2024.104543.

[143] F. L. Mannering, V. Shankar and C. R. Bhat. 'Unobserved heterogeneity and the statistical analysis of highway accident data'. In: *Analytic methods in accident research* 11 (2016), pp. 1–16. DOI: 10.1016/j.amar.2016.04.001.

[144] T. Das, M. S. Samandar and N. Rouphail. 'Longitudinal traffic conflict analysis of autonomous and traditional vehicle platoons in field tests via surrogate safety measures'. In: *Accident Analysis & Prevention* 177 (2022), p. 106822. DOI: 10.1016/j.aap.2022.106822.

[145] M. Wessels, S. Kröling and D. Oberfeld. 'Audiovisual time-to-collision estimation for accelerating vehicles: The acoustic signature of electric vehicles impairs pedestrians' judgments'. In: *Transportation research part F: traffic psychology and behaviour* 91 (2022), pp. 191–212. DOI: 10.1016/j.trf.2022.09.023.

[146] W. Schiff and M. L. Detwiler. 'Information used in judging impending collision'. In: *Perception* 8.6 (1979), pp. 647–658. DOI: 10.1068/p080647.

[147] K. H. Teigen. 'The proximity heuristic in judgments of accident probabilities'. In: *British Journal of Psychology* 96.4 (2005), pp. 423–440. DOI: 10.1348/000712605x47431.

[148] N. Saunier and T. Sayed. 'Probabilistic Framework for Automated Analysis of Exposure to Road Collisions'. In: *Transportation Research Record* 2083 (2008), pp. 96–104. DOI: 10.3141/2083-11.

[149] E. de Gelder, K. Adjenughwure, J. Manders, R. Snijders, J.-P. Paardekooper, O. Op den Camp, A. Tejada and B. De Schutter. 'PRISMA: A novel approach for deriving probabilistic surrogate safety measures for risk evaluation'. In: *Accident Analysis & Prevention* 192 (2023), p. 107273. DOI: 10.1016/j.aap.2023.107273.

[150] P. Songchitruksa and A. P. Tarko. 'The extreme value theory approach to safety estimation'. In: *Accident Analysis & Prevention* 38.4 (2006), pp. 811–822. DOI: 10.1016/j.aap.2006.02.003.

[151] L. Zheng, K. Ismail and X. Meng. 'Freeway safety estimation using extreme value theory approaches: A comparative study'. In: *Accident Analysis & Prevention* 62 (2014), pp. 32–41. DOI: 10.1016/j.aap.2013.09.006.

[152] A. P. Tarko. 'A unifying view on traffic conflicts and their connection with crashes'. In: *Accident Analysis & Prevention* 158 (2021), p. 106187. DOI: 10.1016/j.aap.2021.106187.

[153] F. Nazir, Y. Ali and M. M. Haque. 'Effects of sample size on pedestrian crash risk estimation from traffic conflicts using extreme value models'. In: *Analytic Methods in Accident Research* 44 (2024), p. 100353. DOI: 10.1016/j.amar.2024.100353.

[154] R.-D. Reiss and M. Thomas. *Statistical analysis of extreme values*. Basel, Switzerland: Birkhäuser, 1997.

[155] Y. Jiao, S. C. Calvert, S. van Cranenburgh and H. van Lint. 'Inferring vehicle spacing in urban traffic from trajectory data'. In: *Transportation Research Part C: Emerging Technologies* 155 (Oct. 2023), p. 104289. DOI: 10.1016/j.trc.2023.104289.

[156] Q. Meng and X. Qu. 'Estimation of rear-end vehicle crash frequencies in urban road tunnels'. In: *Accident Analysis & Prevention* 48 (2012), pp. 254–263. DOI: 10.1016/j.aap.2012.01.025.

[157] D. S. Pawar and G. R. Patil. 'Critical gap estimation for pedestrians at uncontrolled mid-block crossings on high-speed arterials'. In: *Safety Science* 86 (2016), pp. 295–303. DOI: 10.1016/j.ssci.2016.03.011.

[158] N. Anwari, M. Abdel-Aty, A. Goswamy and O. Zheng. 'Investigating surrogate safety measures at midblock pedestrian crossings using multivariate models with roadside camera data'. In: *Accident Analysis & Prevention* 192 (2023), p. 107233. DOI: 10.1016/j.aap.2023.107233.

[159] J. Gardner, G. Pleiss, K. Q. Weinberger, D. Bindel and A. G. Wilson. 'GPyTorch: blackbox matrix-matrix Gaussian process inference with GPU acceleration'. In: *Advances in Neural Information Processing Systems* 31 (2018). Ed. by S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi and R. Garnett, pp. 587–7597.

[160] M. Jankowiak, G. Pleiss and J. Gardner. 'Parametric Gaussian process regressors'. In: *Proceedings of the 37th International Conference on Machine Learning* 119 (July 2020). Ed. by H. D. III and A. Singh, pp. 4702–4712.

[161] R. Krajewski, J. Bock, L. Kloeker and L. Eckstein. 'The highD Dataset: A drone dataset of naturalistic vehicle trajectories on German highways for validation of highly automated driving systems'. In: *IEEE 21st International Conference on Intelligent Transportation Systems (ITSC)*. Maui, HI, United States, Nov. 2018, pp. 2118–2125. DOI: 10.1109/itsc.2018.8569552.

[162] T. A. Dingus, S. G. Klauer, V. L. Neale, A. Petersen, S. E. Lee, J. Sudweeks, M. A. Perez, J. Hankey, D. Ramsey, S. Gupta and C. Bucher. *The 100-car naturalistic driving study, Phase II-Results of the 100-car field experiment DOT-HS-810-593*. United States. Department of Transportation. National Highway Traffic Safety Administration, 2006.

[163] D. B. Fambro, R. J. Koppa, D. L. Picha and K. Fitzpatrick. 'Driver braking performance in stopping sight distance situations'. In: *Transportation Research Record* 1701 (2000), pp. 9–16. DOI: 10.3141/1701-02.

[164] S. P. Deligianni, M. Quddus, A. Morris, A. Anvuur and S. Reed. 'Analyzing and modeling drivers' deceleration behavior from normal driving'. In: *Transportation Research Record* 2663 (2017), pp. 134–141. DOI: 10.3141/2663-17.

[165] H. Summala. 'Brake Reaction Times and Driver Behavior Analysis'. In: *Transportation Human Factors* 2.3 (2000), pp. 217–226. DOI: 10.1207/sthf0203_2.

[166] G. Markkula, J. Engström, J. Lodin, J. Bärgman and T. Victor. 'A farewell to brake reaction times? Kinematics-dependent brake response in naturalistic rear-end emergencies'. In: *Accident Analysis & Prevention* 95 (2016), pp. 209–226. DOI: 10.1016/j.aap.2016.07.007.

[167] Department of Economic and Social Affairs. *Policies on spatial distribution and urbanization have broad impacts on sustainable development*. 2020/2. New York: United Nations, 2020.

[168] Institute for Health Metrics and Evaluation (IHME). *Global Burden of Disease Study 2021 (GBD 2021) Results*. 2022. URL: https://vizhub.healthdata.org/gbd-results/ (visited on 23/03/2025).

[169] E. Papadimitriou, A. Filtness, A. Theofilatos, A. Ziakopoulos, C. Quigley and G. Yannis. 'Review and ranking of crash risk factors related to the road infrastructure'. In: *Accident Analysis & Prevention* 125 (2019), pp. 85–97. DOI: 10.1016/j.aap.2019.01.002.

[170] M. D. Jakobsen, K. Glies Vincents Seeberg, M. Møller, P. Kines, P. Jørgensen, L. Malchow-Møller, A. B. Andersen and L. L. Andersen. 'Influence of occupational risk factors for road traffic crashes among professional drivers: systematic review'. In: *Transport Reviews* 43.3 (2022), pp. 533–563. DOI: 10.1080/01441647.2022.2132314.

[171] W. J. Horrey, M. F. Lesch, M. J. Dainoff, M. M. Robertson and Y. I. Noy. 'On-Board Safety Monitoring Systems for Driving: Review, Knowledge Gaps, and Framework'. In: *Journal of Safety Research* 43.1 (2012), pp. 49–58. DOI: 10.1016/j.jsr.2011.11.004.

[172] A. Chang, N. Saunier and A. Laureshyn. *PROACTIVE METHODS FOR ROAD SAFETY ANALYSIS*. White Paper. SAE, 2017. DOI: 10.4271/wp-0005.

[173] J. Wang, J. Wu and Y. Li. 'The Driving Safety Field Based on Driver–Vehicle–Road Interactions'. In: *IEEE Transactions on Intelligent Transportation Systems* 16.4 (2015), pp. 2203–2214. DOI: 10.1109/TITS.2015.2401837.

[174] F. A. Mullakkal-Babu, M. Wang, X. He, B. van Arem and R. Happee. 'Probabilistic field approach for motorway driving risk assessment'. In: *Transportation Research Part C: Emerging Technologies* 118 (2020), p. 102716. DOI: 10.1016/j.trc.2020.102716.

[175] S. Kolekar, J. de Winter and D. Abbink. 'Human-like driving behaviour emerges from a risk-based driver model'. In: *Nature Communications* 11.1 (2020). DOI: 10.1038/s41467-020-18353-4.

[176] World Health Organization. *Road traffic injuries*. Dec. 2023. URL: https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries (visited on 14/03/2025).

[177] E. Papadimitriou. *Road safety thematic report – Main factors causing fatal crashes*. Brussels: Directorate General for Transport, European Commission, 2024.

[178] A. K. Dey. 'Understanding and using context'. In: *Personal and Ubiquitous Computing* 5.1 (2001), pp. 4–7. DOI: 10.1007/s007790170019.

[179] S. Lefèvre, D. Vasquez and C. Laugier. 'A survey on motion prediction and risk assessment for intelligent vehicles'. In: *ROBOMECH Journal* 1.1 (2014). DOI: 10.1186/s40648-014-0001-z.

[180] J. Dahl, G. R. de Campos, C. Olsson and J. Fredriksson. 'Collision Avoidance: A Literature Review on Threat-Assessment Techniques'. In: *IEEE Transactions on Intelligent Vehicles* 4.1 (2019), pp. 101–113. DOI: 10.1109/TIV.2018.2886682.

[181] D. Li, B. Liu, Z. Huang, Q. Hao, D. Zhao and B. Tian. 'Safe Motion Planning for Autonomous Vehicles by Quantifying Uncertainties of Deep Learning-Enabled Environment Perception'. In: *IEEE Transactions on Intelligent Vehicles* 9.1 (2024), pp. 2318–2332. DOI: 10.1109/TIV.2023.3297735.

[182] P. Patera, Y.-T. Chen and W.-H. Fang. 'A Multi-modal Architecture with Spatio-Temporal-Text Adaptation for Video-based Traffic Accident Anticipation'. In: *IEEE Transactions on Circuits and Systems for Video Technology* (2025). DOI: 10.1109/tcsvt.2025.3552895.

[183] Y. Kumamoto, K. Ohtani, D. Suzuki, M. Yamataka and K. Takeda. 'AAT-DA: Accident Anticipation Transformer with Driver Attention'. In: *Proceedings of the Winter Conference on Applications of Computer Vision (WACV) Workshops*. Feb. 2025, pp. 1142–1151.

[184] P. Thomas, A. Morris, R. Talbot and H. Fagerlind. 'Identifying the causes of road crashes in Europe'. In: *AAAM 57th Annual Conference*. Vol. 57. Quebec City, 2013.

[185] R. Nahata, D. Omeiza, R. Howard and L. Kunze. 'Assessing and Explaining Collision Risk in Dynamic Environments for Autonomous Driving Safety'. In: *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*. 2021, pp. 223–230. DOI: 10.1109/ITSC48978.2021.9564966.

[186] J. Chen, Q. Zhang, J. Chen, J. Wang, Z. Fang, Y. Liu and G. Yin. 'A Driving Risk Assessment Framework Considering Driver's Fatigue State and Distraction Behavior'. In: *IEEE Transactions on Intelligent Transportation Systems* 25.12 (2024), pp. 20120–20136. DOI: 10.1109/TITS.2024.3446832.

[187] Y. Yao, M. Xu, Y. Wang, D. J. Crandall and E. M. Atkins. 'Unsupervised Traffic Accident Detection in First-Person Videos'. In: *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 2019.

[188] J. Fang, J. Qiao, J. Bai, H. Yu and J. Xue. 'Traffic Accident Detection via Self-Supervised Consistency Learning in Driving Scenarios'. In: *IEEE Transactions on Intelligent Transportation Systems* 23.7 (2022), pp. 9601–9614. DOI: 10.1109/TITS.2022.3157254.

[189] Y. Yao, X. Wang, M. Xu, Z. Pu, Y. Wang, E. Atkins and D. Crandall. 'DoTA: unsupervised detection of traffic anomaly in driving videos'. In: *IEEE transactions on pattern analysis and machine intelligence* (2022).

[190] Y. Jiao, S. C. Calvert, S. van Cranenburgh and H. van Lint. 'A Unified Probabilistic Approach to Traffic Conflict Detection'. In: *Analytic Methods in Accident Research* 45 (2025), p. 100369.

[191] G. Erion, J. D. Janizek, P. Sturmfels, S. M. Lundberg and S.-I. Lee. 'Improving performance of deep learning models with axiomatic attribution priors and expected gradients'. In: *Nature Machine Intelligence* 3.7 (2021), pp. 620–631.

[192] M. Sundararajan, A. Taly and Q. Yan. 'Axiomatic Attribution for Deep Networks'. In: *Proceedings of the 34th International Conference on Machine Learning (ICML)*. Vol. 70. 2017, pp. 3319–3328.

[193] F.-H. Chan, Y.-T. Chen, Y. Xiang and M. Sun. 'Anticipating accidents in dashcam videos'. In: *Asian Conference on Computer Vision*. 2016, pp. 136–153.

[194] A. P. Shah, J.-B. Lamare, T. Nguyen-Anh and A. Hauptmann. 'CADP: A Novel Dataset for CCTV Traffic Camera based Accident Analysis'. In: *2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*. 2018, pp. 1–9. DOI: 10.1109/AVSS.2018.8639160.

[195] W. Bao, Q. Yu and Y. Kong. 'Uncertainty-based Traffic Accident Anticipation with Spatio-Temporal Relational Learning'. In: *ACM Multimedia Conference*. May 2020.

[196] J. Fang, D. Yan, J. Qiao, J. Xue, H. Wang and S. Li. 'DADA-2000: Can Driving Accident be Predicted by Driver Attention$f$ Analyzed by A Benchmark'. In: *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*. 2019, pp. 4303–4309. DOI: 10.1109/ITSC.2019.8917218.

[197] J. Fang, D. Yan, J. Qiao, J. Xue and H. Yu. 'DADA: Driver Attention Prediction in Driving Accident Scenarios'. In: *IEEE Transactions on Intelligent Transportation Systems* 23.6 (2022), pp. 4959–4971. DOI: 10.1109/TITS.2020.3044678.

[198] E. Sears, M. A. Perez, K. Dan, T. Shimamiya, T. Hashimoto, M. Kimura, S. Yamada and T. Seo. *A Study on the Factors That Affect the Occurrence of Crashes and Near-Crashes*. Version V2. 2019. DOI: 10.15787/VTT1/FQLUWZ. URL: https://doi.org/10.15787/VTT1/FQLUWZ.

[199] C. K. Layman, M. A. Perez, T. Sugino and J. Eggert. *Research of Driver Assistant System*. Version V3. 2019. DOI: 10.15787/VTT1/DEDACT. URL: https://doi.org/10.15787/VTT1/DEDACT.

[200] Federal Highway Administration. *Strategic Highway Research Program (SHRP2)*. 2018. URL: https://highways.dot.gov/safety/data-analysis-tools/rsdp/rsdp-tools/strategic-highway-research-program-shrp2 (visited on 28/03/2025).

[201] Y. Jiao and S. Calvert. *Bird's eye view trajectory reconstruction of naturalistic crashes and near-crashes in the SHRP2 NDS*. Version V1. 2025. DOI: 10.15787/VTT1/EFYEJR.

[202] B. Wilson, W. Qi, T. Agarwal, J. Lambert, J. Singh, S. Khandelwal, B. Pan, R. Kumar, A. Hartnett, J. K. Pontes, D. Ramanan, P. Carr and J. Hays. 'Argoverse 2: Next Generation Datasets for Self-Driving Perception and Forecasting'. In: *Thirty-fifth Conference on Neural Information Processing Systems Datasets and Benchmarks Track (Round 2)*. 2021. URL: https://openreview.net/forum?id=vKQGe36av4k.

[203] G. Li, Y. Jiao, S. C. Calvert and J. (van Lint. 'Lateral conflict resolution data derived from Argoverse-2: Analysing safety and efficiency impacts of autonomous vehicles at intersections'. In: *Transportation Research Part C: Emerging Technologies* 167 (2024), p. 104802. DOI: 10.1016/j.trc.2024.104802.

[204] H. Guo, K. Xie and M. Keyvan-Ekbatani. 'Modeling driver's evasive behavior during safety–critical lane changes: Two-dimensional time-to-collision and deep reinforcement learning'. In: *Accident Analysis & Prevention* 186 (2023), p. 107063. DOI: 10.1016/j.aap.2023.107063.

[205] J. Eggert and T. Puphal. 'Continuous risk measures for ADAS and AD'. In: *Future Active Safety Technology Symposium*. 2017.

[206] T. Puphal, M. Probst and J. Eggert. 'Probabilistic uncertainty-aware risk spot detector for naturalistic driving'. In: *IEEE Transactions on Intelligent Vehicles* 4.3 (2019), pp. 406–415. DOI: 10.1109/tiv.2019.2919465.

[207] J. Davis and M. Goadrich. 'The relationship between Precision-Recall and ROC curves'. In: *Proceedings of the 23rd international conference on Machine learning*. 2006, pp. 233–240.

[208] D. E. Rumelhart, G. E. Hinton and R. J. Williams. 'Learning representations by back-propagating errors'. In: *Nature* 323 (1986), pp. 533–536. DOI: 10.1038/323533a0.

[209] S. Hochreiter and J. Schmidhuber. 'Long Short-Term Memory'. In: *Neural Computation* 9.8 (1997), pp. 1735–1780. DOI: 10.1162/neco.1997.9.8.1735.

[210] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser and I. Polosukhin. 'Attention is All you Need'. In: *Advances in Neural Information Processing Systems*. Ed. by I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan and R. Garnett. Vol. 30. 2017.

[211] Y. Lecun, L. Bottou, Y. Bengio and P. Haffner. 'Gradient-based learning applied to document recognition'. In: *Proceedings of the IEEE* 86.11 (1998), pp. 2278–2324. DOI: 10.1109/5.726791.

[212] D. Hendrycks and K. Gimpel. 'Gaussian error linear units (gelus)'. In: *arXiv preprint* (2016), arXiv:1606.08415.

[213] F. X. X. Yu, A. T. Suresh, K. M. Choromanski, D. N. Holtmann-Rice and S. Kumar. 'Orthogonal Random Features'. In: *Advances in Neural Information Processing Systems*. Ed. by D. Lee, M. Sugiyama, U. Luxburg, I. Guyon and R. Garnett. Vol. 29. 2016.

[214] G. Li, V. L. Knoop and H. van Lint. 'Multistep traffic forecasting by dynamic graph convolution: Interpretations of real-time spatial correlations'. In: *Transportation Research Part C: Emerging Technologies* 128 (July 2021), p. 103185. DOI: 10.1016/j.trc.2021.103185.

[215] W. Jiang and J. Luo. 'Graph neural network for traffic forecasting: A survey'. In: *Expert Systems with Applications* 207 (2022), p. 117921. DOI: 10.1016/j.eswa.2022.117921.

[216] T. T. Nguyen, P. Krishnakumari, S. C. Calvert, H. L. Vu and H. van Lint. 'Feature extraction and clustering analysis of highway congestion'. In: *Transportation Research Part C: Emerging Technologies* 100 (2019), pp. 238–258. DOI: 10.1016/j.trc.2019.01.017.

[217] K. Kumar, M. Kumar and P. Das. 'Traffic congestion forecasting using multilayered deep neural network'. In: *Transportation Letters* 16.6 (2023), pp. 516–526. DOI: 10.1080/19427867.2023.2207278.

[218] P. Krishnakumari, O. Cats and H. van Lint. 'Estimation of metro network passenger delay from individual trajectories'. In: *Transportation Research Part C: Emerging Technologies* 117 (2020), p. 102704. DOI: 10.1016/j.trc.2020.102704.

[219] X. Chen, S. Saidi and L. Sun. 'Understanding bus delay patterns under different temporal and weather conditions: A Bayesian Gaussian mixture model'. In: *Transportation Research Part C: Emerging Technologies* 171 (2025), p. 105000. DOI: 10.1016/j.trc.2025.105000.

[220] X. Liu, F. Zhang, Z. Hou, L. Mian, Z. Wang, J. Zhang and J. Tang. 'Self-supervised learning: Generative or contrastive'. In: *IEEE transactions on knowledge and data engineering* 35.1 (2023), pp. 857–876. DOI: 10.1109/TKDE.2021.3090866.

[221] N. Saunshi, O. Plevrakis, S. Arora, M. Khodak and H. Khandeparkar. 'A theoretical analysis of contrastive unsupervised representation learning'. In: *Proceedings of the 36th International Conference on Machine Learning*. 2019, pp. 5628–5637.

[222] J. Z. HaoChen, C. Wei, A. Gaidon and T. Ma. 'Provable Guarantees for Self-Supervised Deep Learning with Spectral Contrastive Loss'. In: *Advances in Neural Information Processing Systems*. Ed. by M. Ranzato, A. Beygelzimer, Y. Dauphin, P. Liang and J. W. Vaughan. Vol. 34. 2021, pp. 5000–5011.

[223] J. Ge, S. Tang, J. Fan and C. Jin. 'On the Provable Advantage of Unsupervised Pretraining'. In: *The Twelfth International Conference on Learning Representations*. 2024.

[224] A. Tendle and M. R. Hasan. 'A study of the generalizability of self-supervised representations'. In: *Machine Learning with Applications* 6 (2021), p. 100124. DOI: 10.1016/j.mlwa.2021.100124.

[225] K. Zhou, Z. Liu, Y. Qiao, T. Xiang and C. C. Loy. 'Domain generalization: A survey'. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 45.4 (2022), pp. 4396–4415. DOI: 10.1109/TPAMI.2022.3195549.

[226] B. Lafabregue, J. Weber, P. Gançarski and G. Forestier. 'End-to-end deep representation learning for time series clustering: a comparative study'. In: *Data Mining and Knowledge Discovery* 36.1 (2022), pp. 29–81. DOI: 10.1007/s10618-021-00796-y.

[227] J. Wu, J. Chen, J. Wu, W. Shi, X. Wang and X. He. 'Understanding Contrastive Learning via Distributionally Robust Optimization'. In: *Advances in Neural Information Processing Systems*. Ed. by A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt and S. Levine. Vol. 36. 2023, pp. 23297–23320.

[228] Z. Yang, M. Ding, T. Huang, Y. Cen, J. Song, B. Xu, Y. Dong and J. Tang. 'Does Negative Sampling Matter? a Review With Insights Into its Theory and Applications'. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 46.8 (2024), pp. 5692–5711. DOI: 10.1109/TPAMI.2024.3371473.

[229] Z. Yue, Y. Wang, J. Duan, T. Yang, C. Huang, Y. Tong and B. Xu. 'TS2Vec: Towards Universal Representation of Time Series'. In: *Proceedings of the AAAI Conference on Artificial Intelligence* 36.8 (June 2022), pp. 8980–8987. DOI: 10.1609/aaai.v36i8.20881.

[230] S. Lee, T. Park and K. Lee. 'Soft Contrastive Learning for Time Series'. In: *The Twelfth International Conference on Learning Representations*. 2024.

[231] G. Li, V. L. Knoop and H. van Lint. 'How predictable are macroscopic traffic states: a perspective of uncertainty quantification'. In: *Transportmetrica B: Transport Dynamics* 12.1 (2024). DOI: 10.1080/21680566.2024.2314766.

[232] G. Li, Z. Li, V. L. Knoop and H. van Lint. 'Unravelling uncertainty in trajectory prediction using a non-parametric approach'. In: *Transportation Research Part C: Emerging Technologies* 163 (2024), p. 104659. DOI: 10.1016/j.trc.2024.104659.

[233] J.-Y. Franceschi, A. Dieuleveut and M. Jaggi. 'Unsupervised Scalable Representation Learning for Multivariate Time Series'. In: *Advances in Neural Information Processing Systems*. Ed. by H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox and R. Garnett. Vol. 32. 2019.

[234] S. Tonekaboni, D. Eytan and A. Goldenberg. 'Unsupervised Representation Learning for Time Series with Temporal Neighborhood Coding'. In: *The Ninth International Conference on Learning Representations*. 2021.

[235] E. Eldele, M. Ragab, Z. Chen, M. Wu, C. K. Kwoh, X. Li and C. Guan. 'Time-Series Representation Learning via Temporal and Contextual Contrasting'. In: *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, IJCAI-21*. 2021, pp. 2352–2359.

[236] E. Eldele, M. Ragab, Z. Chen, M. Wu, C.-K. Kwoh, X. Li and C. Guan. 'Self-Supervised Contrastive Representation Learning for Semi-Supervised Time-Series Classification'. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 45.12 (2023), pp. 15604–15618. DOI: 10.1109/TPAMI.2023.3308189.

[237] T. T. Cai, J. Frankle, D. J. Schwab and A. S. Morcos. 'Are all negatives created equal in contrastive instance discrimination?' In: *arXiv preprint* (2020), arXiv:2010.06682.

[238] S. Jeon, D. Min, S. Kim and K. Sohn. 'Mining Better Samples for Contrastive Learning of Temporal Correspondence'. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. June 2021, pp. 1034–1044.

[239] J. Liu and S. Chen. 'Timesurl: Self-supervised contrastive learning for universal time series representation learning'. In: *Proceedings of the AAAI Conference on Artificial Intelligence* 38.12 (2024), pp. 13918–13926. DOI: 10.1609/aaai.v38i12.29299.

[240] P. Trirat, Y. Shin, J. Kang, Y. Nam, J. Na, M. Bae, J. Kim, B. Kim and J.-G. Lee. 'Universal time-series representation learning: A survey'. In: *arXiv preprint* (2024), arXiv:2401.03717.

[241] M. Meilă and H. Zhang. 'Manifold Learning: What, How, and Why'. In: *Annual Review of Statistics and Its Application* 11 (2024), pp. 393–417. DOI: 10.1146/annurev-statistics-040522-115238.

[242] W. Ju, Z. Fang, Y. Gu, Z. Liu, Q. Long, Z. Qiao, Y. Qin, J. Shen, F. Sun, Z. Xiao, J. Yang, J. Yuan, Y. Zhao, Y. Wang, X. Luo and M. Zhang. 'A Comprehensive Survey on Deep Graph Representation Learning'. In: *Neural Networks* 173 (2024), p. 106207. DOI: 10.1016/j.neunet.2024.106207.

[243] S. Khoshraftar and A. An. ‘A Survey on Graph Representation Learning Methods’. In: *ACM Transactions on Intelligent Systems and Technology* 15.1 (2024), pp. 1–55. DOI: 10.1145/3633518.

[244] H.-Y. Yao, C.-Y. Zhang, Z.-L. Yao, C. P. Chen and J. Hu. ‘A recurrent graph neural network for inductive representation learning on dynamic graphs’. In: *Pattern Recognition* 154 (2024), p. 110577. DOI: 10.1016/j.patcog.2024.110577.

[245] M. Ashraf, F. Anowar, J. H. Setu, A. I. Chowdhury, E. Ahmed, A. Islam and A. Al-Mamun. ‘A Survey on Dimensionality Reduction Techniques for Time-Series Data’. In: *IEEE Access* 11 (2023), pp. 42909–42923. DOI: 10.1109/ACCESS.2023.3269693.

[246] Q. Meng, H. Qian, Y. Liu, L. Cui, Y. Xu and Z. Shen. ‘MHCCL: Masked Hierarchical Cluster-Wise Contrastive Learning for Multivariate Time Series’. In: *Proceedings of the AAAI Conference on Artificial Intelligence* 37.8 (June 2023), pp. 9153–9161. DOI: 10.1609/aaai.v37i8.26098.

[247] Y. Wang, Y. Han, H. Wang and X. Zhang. ‘Contrast Everything: A Hierarchical Contrastive Framework for Medical Time-Series’. In: *Advances in Neural Information Processing Systems*. Ed. by A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt and S. Levine. Vol. 36. 2023, pp. 55694–55717.

[248] S. Liu, T. Kimura, D. Liu, R. Wang, J. Li, S. Diggavi, M. Srivastava and T. Abdelzaher. ‘FOCAL: Contrastive Learning for Multimodal Time-Series Sensing Signals in Factorized Orthogonal Latent Space’. In: *Advances in Neural Information Processing Systems*. Ed. by A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt and S. Levine. Vol. 36. 2023, pp. 47309–47338.

[249] M. Moor, M. Horn, B. Rieck and K. Borgwardt. ‘Topological autoencoders’. In: *Proceedings of the 37th International Conference on Machine Learning*. 2020, p. 653.

[250] I. Trofimov, D. Cherniavskii, E. Tulchinskii, N. Balabin, E. Burnaev and S. Barannikov. ‘Learning topology-preserving data representations’. In: *The Eleventh International Conference on Learning Representations*. 2023.

[251] H. Madhu and S. P. Chepuri. ‘TopoSRL: Topology preserving self-supervised Simplicial Representation Learning’. In: *Advances in Neural Information Processing Systems*. Ed. by A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt and S. Levine. Vol. 36. 2023, pp. 64306–64317.

[252] Y. Chen, J. Frias and Y. R. Gel. ‘TopoGCL: Topological Graph Contrastive Learning’. In: *Proceedings of the AAAI Conference on Artificial Intelligence* 38.10 (Mar. 2024), pp. 11453–11461. DOI: 10.1609/aaai.v38i10.29026.

[253] P. Nazari, S. Damrich and F. A. Hamprecht. ‘Geometric autoencoders: what you see is what you decode’. In: *Proceedings of the 40th International Conference on Machine Learning*. Hawaii, USA, 2023, p. 1075.

[254] J. Lim, J. Kim, Y. Lee, C. Jang and F. C. Park. 'Graph Geometry-Preserving Autoencoders'. In: *Proceedings of the 41st International Conference on Machine Learning*. 2024.

[255] S. Li, J. Zhou, T. Xu, D. Dou and H. Xiong. 'Geomgcl: Geometric graph contrastive learning for molecular property prediction'. In: *Proceedings of the AAAI conference on artificial intelligence* 36.4 (2022), pp. 4541–4549. DOI: 10.1609/aaai.v36i4.20377.

[256] Y. Koishekenov, S. Vadgama, R. Valperga and E. J. Bekkers. 'Geometric Contrastive Learning'. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops*. Oct. 2023, pp. 206–215.

[257] G. Zerveas, S. Jayaraman, D. Patel, A. Bhamidipaty and C. Eickhoff. 'A transformer-based framework for multivariate time series representation learning'. In: *Proceedings of the 27th ACM SIGKDD conference on knowledge discovery & data mining*. 2021, pp. 2114–2124. DOI: 10.1145/3447548.3467401.

[258] J. Lu, O. Grembek and M. Hansen. 'Learning the representation of surrogate safety measures to identify traffic conflict'. In: *Accident Analysis & Prevention* 174 (2022), p. 106755. DOI: doi.org/10.1016/j.aap.2022.106755.

[259] J. Cheng, X. Mei and M. Liu. 'Forecast-MAE: Self-supervised Pre-training for Motion Forecasting with Masked Autoencoders'. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*. Oct. 2023, pp. 8679–8689.

[260] H. Chen, J. Wang, K. Shao, F. Liu, J. Hao, C. Guan, G. Chen and P.-A. Heng. 'Traj-MAE: Masked Autoencoders for Trajectory Prediction'. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*. Oct. 2023, pp. 8351–8362.

[261] Z. Lan, Y. Jiang, Y. Mu, C. Chen and S. E. Li. 'SEPT: Towards Efficient Scene Representation Learning for Motion Prediction'. In: *The Twelfth International Conference on Learning Representations*. 2024.

[262] A. Prabowo, H. Xue, W. Shao, P. Koniusz and F. D. Salim. 'Traffic forecasting on new roads using spatial contrastive pre-training (SCPT)'. In: *Data Mining and Knowledge Discovery* 38.3 (2023), pp. 913–937. DOI: 10.1007/s10618-023-00982-0.

[263] L. Pan, Q. Ren, Z. Li and X. Lv. 'Rethinking spatial-temporal contrastive learning for Urban traffic flow forecasting: multi-level augmentation framework'. In: *Complex & Intelligent Systems* 11.1 (2024). DOI: 10.1007/s40747-024-01754-z.

[264] W. Zhou, G. Shen, Y. Zhang, Z. Deng, X. Kong and F. Xia. 'Sequence-to-Sequence Traffic Missing Data Imputation via Self-Supervised Contrastive Learning'. In: *IEEE Transactions on Intelligent Transportation Systems* 26.7 (2025), pp. 9948–9961. DOI: 10.1109/tits.2025.3564578.

[265] D. Zhang, P. Wang, L. Ding, X. Wang and J. He. 'Spatio-Temporal Contrastive Learning-Based Adaptive Graph Augmentation for Traffic Flow Prediction'. In: *IEEE Transactions on Intelligent Transportation Systems* 26.1 (2025), pp. 1304–1318. DOI: 10.1109/tits.2024.3487982.

[266] Z. Mao, Z. Li, D. Li, L. Bai and R. Zhao. 'Jointly Contrastive Representation Learning on Road Network and Trajectory'. In: *Proceedings of the 31st ACM International Conference on Information & Knowledge Management*. Atlanta, USA, Oct. 2022, pp. 1501–1510. DOI: 10.1145/3511808.3557370.

[267] M. Zipfl, M. Jarosch and J. M. Zöllner. 'Traffic Scene Similarity: a Graph-based Contrastive Learning Approach'. In: *IEEE Symposium Series on Computational Intelligence*. Mexico City, Mexico, Dec. 2023. DOI: 10.1109/ssci52147.2023.10372060.

[268] X. Zheng, S. A. Bagloee and M. Sarvi. 'TRECK: Long-Term Traffic Forecasting With Contrastive Representation Learning'. In: *IEEE Transactions on Intelligent Transportation Systems* (2024). DOI: 10.1109/tits.2024.3421328.

[269] Y. Huang and J. J. Yang. 'Symmetric contrastive learning for robust fault detection in time-series traffic sensor data'. In: *International Journal of Data Science and Analytics* (2024). DOI: 10.1007/s41060-024-00521-0.

[270] A. Kendall, Y. Gal and R. Cipolla. 'Multi-Task Learning Using Uncertainty to Weigh Losses for Scene Geometry and Semantics'. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. June 2018.

[271] A. Bagnall, H. A. Dau, J. Lines, M. Flynn, J. Large, A. Bostrom, P. Southam and E. Keogh. 'The UEA multivariate time series classification archive, 2018'. In: *arXiv preprint* (2018), arXiv:1811.00075.

[272] J. Gao, C. Sun, H. Zhao, Y. Shen, D. Anguelov, C. Li and C. Schmid. 'VectorNet: Encoding HD Maps and Agent Dynamics From Vectorized Representation'. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. June 2020.

[273] J. Gu, C. Sun and H. Zhao. 'DenseTNT: End-to-End Trajectory Prediction From Dense Goal Sets'. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*. Oct. 2021, pp. 15303–15312.

[274] M. B. Lee, C. T. Lee, M. A. Abas and W. W. Fong Chong. 'Advancing pedestrian safety in the era of autonomous vehicles: A bibliometric analysis and pathway to effective regulations'. In: *Journal of Traffic and Transportation Engineering (English Edition)* (2025). DOI: 10.1016/j.jtte.2024.05.004.

[275] S. Moosavi and R. Ramnath. 'Context-aware driver risk prediction with telematics data'. In: *Accident Analysis & Prevention* 192 (2023), p. 107269. DOI: 10.1016/j.aap.2023.107269.

[276] L. Masello, G. Castignani, B. Sheehan, M. Guillen and F. Murphy. 'Using contextual data to predict risky driving events: A novel methodology from explainable artificial intelligence'. In: *Accident Analysis & Prevention* 184 (2023), p. 106997. DOI: 10.1016/j.aap.2023.106997.

[277] F. Pazhoohi, G. Choi and A. Kingstone. 'Larger distances from larger vehicles: effect of vehicle size, viewing side and their facia on comfort distance in virtual reality'. In: *Australian Journal of Psychology* 73.2 (2021), pp. 179–187. DOI: 10.1080/00049530.2021.1882272.

[278] O. Bagdadi. 'Estimation of the severity of safety critical events'. In: *Accident Analysis & Prevention* 50 (2013), pp. 167–174. DOI: 10.1016/j.aap.2012.04.007.

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

[280] X. Wang, Q. Liu, F. Guo, S. Fang, X. Xu and X. Chen. 'Causation analysis of crashes and near crashes using naturalistic driving data'. In: *Accident Analysis & Prevention* 177 (2022), p. 106821. DOI: 10.1016/j.aap.2022.106821.

[281] L. Sun, S. Cheng, X. Wang, J. Zhang, H. Xiang and H. Zhao. 'Exploring autonomous vehicle crash risk: system coupling effects and key causal factors'. In: *Journal of Safety Research* 94 (2025), pp. 284–293. DOI: 10.1016/j.jsr.2025.06.029.

[282] R. Elvik. 'Risk factors as causes of accidents: Criterion of causality, logical structure of relationship to accidents and completeness of explanations'. In: *Accident Analysis & Prevention* 197 (2024), p. 107469. DOI: 10.1016/j.aap.2024.107469.

[283] K. Wang, C. Shen, X. Li and J. Lu. 'Uncertainty Quantification for Safe and Reliable Autonomous Vehicles: A Review of Methods and Applications'. In: *IEEE Transactions on Intelligent Transportation Systems* 26.3 (2025), pp. 2880–2896. DOI: 10.1109/tits.2025.3532803.

[284] W. M. D. Chia, S. L. Keoh, C. Goh and C. Johnson. 'Risk Assessment Methodologies for Autonomous Driving: A Survey'. In: *IEEE Transactions on Intelligent Transportation Systems* 23.10 (2022), pp. 16923–16939. DOI: 10.1109/tits.2022.3163747.

[285] K. Moller, R. Trauth and J. Betz. 'Overcoming Blind Spots: Occlusion Considerations for Improved Autonomous Driving Safety'. In: *2024 IEEE Intelligent Vehicles Symposium (IV)*. June 2024, pp. 819–826. DOI: 10.1109/iv55156.2024.10588481.

[286] J. Wang, X. Yuan, Z. Liu, W. Tan, X. Zhang and Y. Wang. 'Adaptive Dynamic Path Planning Method for Autonomous Vehicle Under Various Road Friction and Speeds'. In: *IEEE Transactions on Intelligent Transportation Systems* 24.10 (2023), pp. 10977–10987. DOI: 10.1109/tits.2023.3273545.

[287] L. Chen, P. Wu, K. Chitta, B. Jaeger, A. Geiger and H. Li. 'End-to-End Autonomous Driving: Challenges and Frontiers'. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 46.12 (2024), pp. 10164–10183. DOI: 10.1109/tpami.2024.3435937.

[288] L. Le Mero, D. Yi, M. Dianati and A. Mouzakitis. 'A Survey on Imitation Learning Techniques for End-to-End Autonomous Vehicles'. In: *IEEE Transactions on Intelligent Transportation Systems* 23.9 (2022), pp. 14128–14147. DOI: 10.1109/TITS.2022.3144867.

[289] S. Teng, L. Chen, Y. Ai, Y. Zhou, Z. Xuanyuan and X. Hu. 'Hierarchical Interpretable Imitation Learning for End-to-End Autonomous Driving'. In: *IEEE Transactions on Intelligent Vehicles* 8.1 (2023), pp. 673–683. DOI: 10.1109/TIV.2022.3225340.

[290] ISO Central Secretary. *Road vehicles – Safety of the intended functionality*. EN. Standard. Geneva, CH: International Organization for Standardization, June 2022. URL: https://www.iso.org/standard/77490.html.

[291] L. Wang, J. Liu, H. Shao, W. Wang, R. Chen, Y. Liu and S. L. Waslander. 'Efficient Reinforcement Learning for Autonomous Driving with Parameterized Skills and Priors'. In: *Robotics: Science and Systems 2023*. Daegu, Republic of Korea, 2023.

[292] E. Dagan. *Solving the long-tail with e2e AI: "The revolution will not be supervised"*. 2024. URL: https://wayve.ai/thinking/e2e-embodied-ai-solves-the-long-tail (visited on 08/07/2025).

[293] D. Hafner, T. Lillicrap, J. Ba and M. Norouzi. 'Dream to Control: Learning Behaviors by Latent Imagination'. In: *International Conference on Learning Representations*. 2020. URL: https://openreview.net/forum?id=S1lOTC4tDS.

[294] X. Wang, Z. Zhu, G. Huang, X. Chen, J. Zhu and J. Lu. 'DriveDreamer: Towards Real-World-Drive World Models for Autonomous Driving'. In: *European Conference on Computer Vision*. 2024, pp. 55–72.

[295] M. Goff, G. Hogan, G. Hotz, A. du Parc Locmaria, K. Raczy, H. Schäfer, A. Shihadeh, W. Zhang and Y. Yousfi. 'Learning to Drive from a World Model'. In: *Proceedings of the Computer Vision and Pattern Recognition Conference (CVPR) Workshops*. June 2025, pp. 1964–1973.

[296] T. Feng, W. Wang and Y. Yang. 'A survey of world models for autonomous driving'. In: *arXiv preprint* (2025), arXiv:2501.11260.

[297] Z. Li, W. Wang, H. Li, E. Xie, C. Sima, T. Lu, Q. Yu and J. Dai. 'BEVFormer: Learning Bird's-Eye-View Representation From LiDAR-Camera via Spatiotemporal Transformers'. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 47.3 (2025), pp. 2020–2036. DOI: 10.1109/TPAMI.2024.3515454.

[298] J. Gao, C. Sun, H. Zhao, Y. Shen, D. Anguelov, C. Li and C. Schmid. 'VectorNet: Encoding HD Maps and Agent Dynamics From Vectorized Representation'. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 2020.

[299] Y. Hu, J. Yang, L. Chen, K. Li, C. Sima, X. Zhu, S. Chai, S. Du, T. Lin, W. Wang, L. Lu, X. Jia, Q. Liu, J. Dai, Y. Qiao and H. Li. 'Planning-Oriented Autonomous Driving'. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 2023, pp. 17853–17862.

[300] H. Caesar, V. Bankiti, A. H. Lang, S. Vora, V. E. Liong, Q. Xu, A. Krishnan, Y. Pan, G. Baldan and O. Beijbom. 'nuScenes: A Multimodal Dataset for Autonomous Driving'. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 2020.

[301] P. Sun, H. Kretzschmar, X. Dotiwalla, A. Chouard, V. Patnaik, P. Tsui, J. Guo, Y. Zhou, Y. Chai, B. Caine, V. Vasudevan, W. Han, J. Ngiam, H. Zhao, A. Timofeev, S. Ettinger, M. Krivokon, A. Gao, A. Joshi, Y. Zhang, J. Shlens, Z. Chen and D. Anguelov. 'Scalability in Perception for Autonomous Driving: Waymo Open Dataset'. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 2020.

[302] H. Caesar, J. Kabzan, K. S. Tan, W. K. Fong, E. Wolff, A. Lang, L. Fletcher, O. Beijbom and S. Omari. 'nuPlan: A closed-loop ML-based planning benchmark for autonomous vehicles'. In: *arXiv preprint* (2021), arXiv:2106.11810.

[303] N. Singer. *U.S. Department of Transportation Awards $9.95 Million to the University of Michigan for Smart Intersections*. 2020. URL: https://highways.dot.gov/newsroom/us-department-transportation-awards-995-million-university-michigan-smart-intersections (visited on 09/07/2025).

[304] NETCOMPANY S.A. and the other partners. *Systematic and orchestrated deployment of safety solutions in complex urban environments for ageing and vulnerable societies*. Project. Horizon Europe, 2022. DOI: 10.3030/101077433.

[305] Nissan. *Nissan adds intersection collision avoidance to its in-development LIDAR-based driver-assistance technology*. 2023. URL: https://global.nissannews.com/en/releases/nissan-lidar-based-driver-assistance-technology (visited on 09/07/2025).

[306] Ford. *Ford Co-Pilot360 Technology*. 2024. URL: https://www.ford.com/technology/driver-assist-technology/ (visited on 09/07/2025).

[307] Mobileye. *Unveiling Surround ADAS: A new standard of safety and tech*. 2025. URL: https://www.mobileye.com/blog/how-surround-adas-delivers-the-new-standard-of-safety-and-tech/ (visited on 09/07/2025).

[308] A. Karpathy. *Keynote - CVPR 2021 Workship on Autonomous Driving*. 2021. URL: https://youtu.be/g6bOwQdCJrc?si=UgsJwV_TEPl5BtZ2 (visited on 09/07/2025).

[309] M. Gyllenhammar, G. Rodrigues de Campos and M. Törngren. 'The Road to Safe Automated Driving Systems: A Review of Methods Providing Safety Evidence'. In: *IEEE Transactions on Intelligent Transportation Systems* 26.4 (2025), pp. 4315–4345. DOI: 10.1109/TITS.2025.3532684.

[310] Drivewyze. *Smart Roadways*. 2025. URL: https://drivewyze.com/products/smart-roadways/ (visited on 09/07/2025).

[311] K. Edvardsson Björnberg, S. O. Hansson, M. Å. Belin and C. Tingvall. *The Vision Zero Handbook: Theory, Technology and Management for a Zero Casualty Policy*. Springer International Publishing, 2022. DOI: 10.1007/978-3-030-23176-7.

[312] The Insight Partners. *Intelligent Transportation System Market is expected to reach US$ 66,355.13 Million by 2031*. 2024. URL: https://www.theinsightpartners.com/pr/intelligent-transportation-system-market (visited on 09/07/2025).

[313] J. Falkiner. *China Set to Lead the World in V2X-Connected Vehicle Technology Uptake*. 2024. URL: https://www.idtechex.com/en/research-article/china-set-to-lead-the-world-in-v2x-connected-vehicle-technology-uptake/30516 (visited on 10/07/2025).

Acknowledgements

This PhD started in the middle of a global pandemic, continued as a war unfolded in Europe, and evolved against rising living costs, changing attitudes toward international students, and uncertainties in higher education funding and policy. I do not mention this to blame politics or politicians; it is simply the world we lived and worked in. At the same time, there were more personal and painful moments that touched on questions of identity, belonging, fairness, and academic integrity. I will not describe them in detail here, but they were real, they were difficult, and they have shaped who I am as a person and as a researcher. Despite all of this, or perhaps because of it, I believe even more strongly that scientific research brings hope. In an uncertain world, the commitment to understanding and improving things is one way we refuse to be defeated. This thesis is my modest contribution to that shared effort, and it has been carried to the finish line by many hands.

Before thanking specific people, I would like to first thank the broader academic community that made this work possible. To the researchers who insist on open-source, who share datasets, tools, and code, often with little recognition, so that strangers they will never meet can learn and build on their work. To the editors and reviewers who invest their time and energy into improving other people's manuscripts, making our work sharper and more rigorous (even when it hurts a little to read the comments). To the academic systems and institutions that supported my development, including funding, infrastructure, courses, conferences, and administrative help, even as they themselves were under pressure from reforms, budget cuts, and shifting expectations. And, on a light-hearted but honest note, I also want to thank large language models and other digital tools that have accompanied me. They helped me learn quickly, discover unknown papers, debug stubborn code, and explore ideas from different angles. They also occasionally produced confidently misleading answers, which still somehow pushed my thinking forward.

I owe thanks to an entire constellation of people who helped me navigate this adventure. The following is my attempt, necessarily incomplete, to dedicate my gratitude to them.

I am deeply grateful to my supervisory team for their constant support, patience, and guidance. **Hans**, I have always seen you as a role model, not just as a scientist, but even more as a person who sincerely cares about the society we live in. Thank you for your sharp questions, your broad perspective, and your willingness to connect technical detail with societal relevance. I greatly enjoyed our discussions every time I knocked on your door. Sparks flew from our differing opinions, and a quick question often turned into a real conversation. You showed me that being a good researcher is not only about methods and models, but also about values and responsibility. **Simeon**, you gave me the freedom to explore my own ideas, the structure to turn them into research, and the encouragement to continue when I doubted myself. You treated me not only as a PhD student but also as an

independent researcher, so I always felt that my voice mattered in our discussions. This is the greatest support I can imagine from a daily supervisor. **Sander**, thank you for your trust, your kindness, and your willingness to listen to both my scientific arguments and my non-scientific worries. You have always been responsive, and I appreciate it deeply, for reading my drafts (especially those messy early ones), for your careful comments, and for showing me how to transform an intuition into a clear contribution. Your steady support provided the ground on which this thesis was built.

I am also honoured to have had such a thoughtful and supportive thesis committee. Having two inspiring female professors, **Marjan** and **Soyoung**, at the table, and a female Rector Magnificus presiding, meant a great deal to me. Representation is not just symbolic. It quietly shifts what feels possible, and I felt that strongly. **Gustav**, thank you for your detailed feedback on this thesis. Reading your remarks felt like we were discussing the ideas face to face. Your comments pushed me to sharpen the arguments while still being encouraging, and I truly appreciate that balance. **Dan**, thank you for your support and for the warmth you have shown. I owe you many thanks for the research visit that did not happen in the end, and for your considerate hospitality in Nashville. The fact that you remained supportive despite changing circumstances meant a lot.

I would like to extend my gratitude to other faculty members who have helped, supported, and encouraged me. Dear, dear **Haneen**, it is my great pity not to have you in my committee. For so many times, when you asked me how I was doing, you reassured me no matter whether I was struggling or overly excited. I learned a lot from you about being a good researcher, a good friend, a good lecturer, and a good person who is at peace with herself. You set an example of kindness and strength that I will carry with me. **Victor**, thank you for our weekly discussions on traffic dynamics. Your insights were always interesting and inspiring, bringing a different angle that helped me think more deeply. This thankfulness also goes to **Maria**, **Andreas**, and **Marco**. I really enjoyed our discussions and the fresh perspectives you brought. **Bart**, thank you for the random chats in the corridor and for being willing to offer suggestions when I nervously asked about career considerations. Those informal conversations helped me feel less lost. **Irene**, thank you for being around and for sharing your ideas so openly. I always appreciated that you care about PhD students' lives. I would also like to thank **Lóri**, **Shadi**, **Jan Anne**, and **Gonçalo** for their encouragement and for sharing their opinions and experiences. Your feedback, questions, and friendly comments all contributed to an environment in which I felt supported and seen.

I am also grateful to the supporting staff, who quietly keep everything running. **Charlotte** and **Simon**, many thanks for not only your efforts in maintaining and developing the DAI Labs programme, but also your help in dealing with the challenges during my preparation for the (unfortunately cancelled) research visit. In the department, I owe **Dian** and **Maaike** a big thank you for helping me arrange my internship. Without your help, it would not have been possible. Thank you, **Monique**, for truly caring about everyone's life, and for a smile that is always encouraging. **Pris**, thank you for your continued and reliable support. On many early mornings, I saw that you had already arrived and started working long before most of us. The department is calmer and more organised because of you.

Thanks to **Marije** and **Suzanne**, for coordinating meetings for my supervisory team and me, for dealing with complicated calendars, and for being so patient with last-minute changes. Thanks to **Dehlaila** and **Moreen** for all your work for the department, and for being consistent and approachable. **Conchita** and **Esther**, thank you a lot for being the bridge between TRAIL and us students. You made the logistics around courses and events much easier than they might have been. Finally, a heavy thank you to **Serge** and **Oded**. Your work may be behind what we usually see, but your efforts are always appreciated in this department.

To my friends and colleagues, I am truly grateful that our paths have approached, overlapped, and intertwined during these years. Thank you for the discussions, the shared frustrations over experiments and deadlines, the coffee breaks, the jokes, the walks, the memes, and the quiet moments of understanding. You turned an individual journey into a shared experience. I am aware that I am also leaving in some sense, moving on to a next stage, but I carry your influence, your support, and our memories with me.

First, to my cityAI lab mates, our conversations were as valuable as any seminar, and we really should have more table game nights. **Francisco**, you are a true friend. I always feel happy when I see your smile. That is your magic. Thank you, and **Cony** as well, for many treats of food from all over the world. If one day I opened a café helado restaurant, please be my chef and chief happiness officer. **Gabriel**, thank you for being considerate and kind. You are a real gentleman, quietly helping, noticing when someone is stressed, and offering support. **Lucas**, thank you for being serious–just kidding–for being insightful, patient, and helpful. You were always ready to discuss research questions, challenge assumptions, or share a clever idea. I learned a lot from our exchanges. **Lion**, thank you for your humour and for connecting the lab. The way you bring people together is incredibly valuable.

Then I would like to thank my office mates in DiTTlab, who always share snacks and chocolates at our snack corner, which means a great deal to PhD students. **Xue**, our friendship extends far beyond being office mates. Thank you for always being open and supportive. I love listening to your adventures. Talking with you makes me feel more vibrant and alive. **Saeed**, thank you for being a ballast in the lab. You are someone everyone trusts, and with good reason. Your approachability and reliability created a sense of safety for many of us. **Zili**, thank you for being straightforward and stubborn in the best possible way. Many times I feel we are alike, even if we do not say it out loud. **Jonah**, thank you for your interest in discussing history, politics, and animations. I saw a different angle from you, and I hope I also brought one to you. **Lucas**, thank you for being kind, easy-going, and for sharing travel suggestions. **Kexin**, thank you for bringing new things into the lab, new energy, new ideas, and new presence. Change is important, and you contributed to that. **Samir**, thank you for sharing your thoughts and for organising events for ADaS lab. Your efforts did not go unnoticed. **Mariko**, thank you for your kindness and the gifts from your home country. They brightened many days. **Robin**, thank you for being you. From you I see a soft heart and open mind. **Wouter**, thank you for your cold jokes! I genuinely love them. **Srinath**, thank you for always being there. Sometimes support simply means someone's steady presence, and you offered that generously. **Saman** and

Mahsa, although you did not share the office room ;) I want to thank you for organising activities for DiTTlab and keeping the community spirit alive. I also want to thank our doctors who graduated and left DiTTlab: **Panchamy**, **Tin**, **Zahra**, and **Ali**. You helped me in my early years here, answered my beginner questions, and offered help whenever I needed it. I appreciate our interactions and the paths you showed ahead.

Now I go out of the office and into the corridor. **Ziyulong**, we started our PhD journeys at almost the same time. Thank you for your company, your considerate help, your unconditioned encouragement, and for sharing food and snacks along the way. **Elif**, sometimes I regret not having known you earlier. Thank you for being fun and kind, for your honesty, and for sharing your stories. You brought laughter and warmth into many days. **Nirvana**, **Renate**, **Nina**, and **Merve**, you are four golden flowers in my mind. Our interactions, your energy, and the way you take responsibility for PhDs in the department always encourage me. You make the place softer and more vibrant. **Alex**, thank you so much for entertaining discussions and for sharing fun things that brightened regular weekdays. **Ting**, thank you for hanging out and sharing your thoughts without any reservation. I value your honesty and the way you speak your mind. **Nagarjun**, thank you for encouraging me when I was in my first year. To some extent, your encouragement and help made me feel courageous enough to interact more with others. **Qiaochu**, thank you for home-made food and kitten petting. Those soft memories warmed me for a long time. **Yufei**, thank you for organising summer barbecues for Chinese colleagues. Thanks to the visiting PhDs in 2022, **Bing**, **Chao**, **Yunxue**, **Kequan** & **Yuxuan**, **Yumin**, and **Shiqi**. I have great memories from when you were here. Even though your stay was temporary, your presence left a lasting trace. Extended thanks to **Jie**, **Yuqing**, **Jinyang**, **Yanyan**, **Sara**, **Yiyun**, **Yuxia**, and **Dingshan**. Thank you for being around, being supportive, being encouraging, and being inspiring in many small and big ways.

I had great casual time with my Chinese friends in the other faculties and beyond Delft. **Yujie**, **Xiaohuan**, **Ziqing**, and **Jingjing**, the time with you, whether two, three, or all of us together, cooking or hanging out, formed my most relaxed memories. Speaking our language, sharing inside jokes, and cooking familiar dishes together created a little island of home. I am so lucky to have you around. **Zirui** & **Zheyu**, **Sherry** & **Sihao**, **Tianlong** & **Jing**, **Xiaolin**, **Zhaochong**, **Qiuju**, **Yifei**, **Desong**, **Junyi**, and **Zhiyao**, having dinner with you was always refreshing. Sharing Chinese food together reassured a lot of my nostalgia. Thank you for your openness, your invitations, and your company.

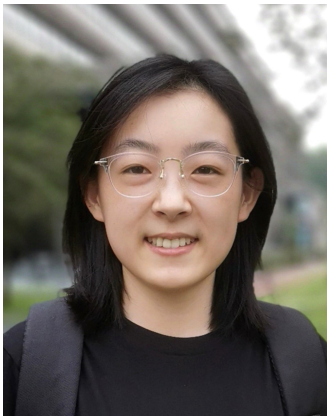
To my family and loved ones, thank you for your absolute support. Your belief in me, your patience with my stress, and your gentle reminders that there is life outside research have been invaluable. 谢谢你们，爸爸妈妈。谢谢你们一直以来的支持，情感上的、经验上的、有时默默的。谢谢弟弟，我很开心看到你成长、分享你的经历。即使远隔重洋，我能感受到你们的关心。(Thank you, **Mom** and **Dad**, for supporting me whenever I needed it, emotionally, practically, and silently. Thank you to my younger brother. I am really happy seeing you grow up and telling me your growth stories. Even from far away, I felt your care.) My closest friends, **Zitong** and **Mengyue**, thank you for always having my back. I deeply cherish our remote communication, trips together, gifts on important days, and unyielding trust across time zones and oceans.

Deeply from my heart, Zitong, I feel so lucky to have you in my life, for such a long time of almost 15 years. Finally, thank you, **Guopeng**. I have too many mixed emotions and feelings when I think of you now. Words cannot summarise the journey we have walked along. Thank you for your continued support and love, for standing by my side through the hardest moments, and for celebrating the small victories on the way.

Interactions have been at the heart of my research and my life during these years. My master's thesis was about opinion interaction, by which people communicate, agree, disagree, and together create complex dynamics. This PhD thesis turns to road-user interactions, when people approach, coexist briefly in the same road space, negotiate their movements, and then depart toward different destinations. Looking back, I see that my own journey followed a similar pattern. Human connection is an eternal dance of approaching and leaving. People come into each other's lives, walk alongside for a while, and then step away onto their own paths, leaving traces behind. This thesis is about interactions, and it has been built through interactions. Some are intense and short, some are gentle and long-lasting, but all of them shape us in ways we often only understand afterwards. To everyone who has been part of that dance, in big ways or small, whether your name appears here or not: thank you.

Yiru
Shanghai, 10 December 2025

About the Author



Yiru Jiao was born and raised in Dingxi (定西), a historic city in Gansu (甘肃) Province, China. The city was once a northwestern frontier in ancient times, and its history of resistance has shaped a local culture of unyielding loyalty.

In 2014, she travelled across half the country to Shenyang to earn a bachelor's degree in Industrial Engineering at Northeastern University, where she became fascinated with operations research and programming. In 2018, she obtained an exam-free admission to Harbin Institute of Technology, where she pursued her master's degree in Management Science and Engineering, specialised in Management Information Systems.

These learning experiences sparked her interest in studying the collective patterns emerging in individual interactions, motivating her to pursue a PhD. In January 2021, she began her doctoral journey at Delft University of Technology (TU Delft). Supported by the TU Delft AI Initiative, her research focuses on the proactive safety quantification of road user interactions. More broadly, she aims to enable safer traffic for all road users by utilising automated technologies, including, but not limited to, automated vehicles, data-empowered traffic monitoring, and data-driven improvements to transportation systems.

List of publications

First-authored journal papers

1. **Yiru Jiao**, Simeon C. Calvert, Sander van Cranenburgh, and Hans van Lint. (2026). *Learning collision risk proactively from naturalistic driving data at scale*. Nature Machine Intelligence, in press. Available at arXiv: 2505.13556. [Chapter 5]

2. **Yiru Jiao**, Sander van Cranenburgh, Simeon C. Calvert, and Hans van Lint. (2025). *Structure-preserving contrastive learning for spatial time series*. Artificial Intelligence for Transportation, 3–4, 100031. doi: 10.1016/j.ait.2025.100031 [Chapter 6]

3. **Yiru Jiao**, Simeon C. Calvert, Sander van Cranenburgh, and Hans van Lint. (2025). *A unified probabilistic approach to traffic conflict detection*. Analytic Methods in Accident Research, 45, 100369. doi: 10.1016/j.amar.2024.100369 [Chapter 4]

4. **Yiru Jiao**, Guopeng Li, Simeon C. Calvert, Sander van Cranenburgh, and Hans van Lint. (2024). *Beyond behaviour change: Investigating alternative explanations for shorter time headways when human drivers follow automated vehicles*. Transportation Research Part C: Emerging Technologies, 164, 104673. doi: 10.1016/j.trc.2024.104673

5. **Yiru Jiao**, Simeon C. Calvert, Sander van Cranenburgh, and Hans van Lint. (2023). *Inferring vehicle spacing in urban traffic from trajectory data*. Transportation Research Part C: Emerging Technologies, 155, 104289. doi: 10.1016/j.trc.2023.104289 [Chapter 2]

First-authored conference proceedings

6. **Yiru Jiao**, Simeon C. Calvert, Hans van Lint. (2024, June 2–5, poster). *Minimising Missed and False Alarms: A vehicle spacing based approach to conflict detection*. 35th IEEE Intelligent Vehicles Symposium, Jeju Island, South Korea. doi: 10.1109/IV55156.2024.10588396 [Chapter 3]

7. **Yiru Jiao**, Simeon C. Calvert, Sander van Cranenburgh, Hans van Lint. (2023, Sep 24–28, oral). *Identifying vehicle interaction at urban intersections: A comparison of Proximity Resistance, Time-to-Collision, and Post-Encroachment-Time*. 26th IEEE International Conference on Intelligent Transportation Systems ITSC 2023, Bilbao, Spain. doi: 10.1109/ITSC57777.2023.10422682

Peer-reviewed conference presentations

8. **Yiru Jiao**, Simeon C. Calvert, Sander van Cranenburgh, Hans van Lint. (2024, Oct 17–18, oral). *A generalised data-driven framework for conflict detection in autonomous driving*. 36th Annual Conference of International Cooperation on Theories and Concepts in Traffic safety (ICTCT), the Heague, Netherlands.

9. **Yiru Jiao**, Simeon C. Calvert, Hans van Lint. (2024, Jan 7–11, poster). *Impacts of automated vehicles on human drivers following them: A focus on distance and time headway*. Transportation Research Board 103rd Annual Meeting, Washington, D.C., United States.
10. **Yiru Jiao**, Simeon C. Calvert, Sander van Cranenburgh, Hans van Lint. (2023, Sep 6–8, oral). *Varying critical time to collision: A perspective of driver space*. hEART 2023: 11st Symposium of the European Association for Research in Transportation, Zürich, Switzerland.
11. **Yiru Jiao**, Sander van Cranenburgh, Simeon C. Calvert, Hans van Lint. (2022, June 1–3, oral). *Probabilistic representation of driver space and its inference from trajectory data*. hEART 2022: 10th Symposium of the European Association for Research in Transportation, Leuven, Belgium.

Co-authored collaboration papers

Journal papers

12. Guopeng Li, **Yiru Jiao**, Simeon C. Calvert, J.W.C. (Hans) van Lint. (2024) *A conflict resolution dataset derived from Argoverse-2: Analysis of the safety and efficiency impacts of autonomous vehicles at intersections*. Transportation Research Part C: Emerging Technologies, 167, 104802. doi: 10.1016/j.trc.2024.104802
13. Kequan Chen, Zhibin Li, Pan Liu, Victor L. Knoop, Yu Han, and **Yiru Jiao**. (2024) *Evaluating the safety and efficiency impacts of forced lane change with negative gaps based on empirical vehicle trajectories*. Accident Analysis & Prevention, 203, 107622. doi: 10.1016/j.aap.2024.107622

Conference proceeding

14. Guopeng Li, **Yiru Jiao**, Victor L. Knoop, Simeon C. Calvert and J.W.C. (Hans) Van Lint. (2023, Sep 24–28, oral) *Large car-following data based on Lyft level-5 open dataset: Following autonomous vehicles vs. human-driven vehicles*. IEEE 26th International Conference on Intelligent Transportation Systems (ITSC), Bilbao, Spain. doi: 10.1109/ITSC57777.2023.10422574

Education activities

Apr. 2025 – Nov. 2025	**Master thesis supervision** Daily supervisor – *Driving style-aware car-following model and simulation: Impact of styles on traffic breakdown*
May 2025	**CIEQ6222 Traffic Safety** Guest lecturer – *Data-driven surrogate safety measures*
Nov. 2024	**TIL6022 TIL Python Programming** Final group assignment grading
Apr. 2024 – Jun. 2024	**CTB3000 Bachelor eindwerk** Advisor – *Human factors in roundabout design in urban areas*
Sep. 2023 – Jun. 2024	**Master thesis supervision** Daily supervisor – *Impacts of micro-scale built environment features on residential location choice: a computer vision-aided assessment*
Sep. 2023 – Nov. 2023	**IFEEMCS520100 Fundamentals of AI Programme** Project supervision and grading
Sep. 2023 – Oct. 2023	**TIL 4020-20 Research Project** Group workshop
Apr. 2023 – Jun. 2023	**CTB3000 Bachelor eindwerk** Supervision and evaluation
Sep. 2022 – Jan. 2023	**CIE5050 Additional Master Thesis** Daily supervisor – *Analysis on the heterogeneity of proximity resistance in car following*
Oct. 2022	**MUDE (Modelling, Uncertainty and Data for Engineers)** Poster and coding session teaching assistance
Apr. 2022 – Jul. 2022	**CIE4845 Emerging Topics for Transport & Planning** Supervision and grading
Sep. 2021 – Dec. 2021	**CS4320TU Applied AI project** Supervision and grading – *Detecting traffic jam with urban trajectory data*

TRAIL Thesis Series

The following list contains the most recent dissertations in the TRAIL Thesis Series. For a complete overview of more than 400 titles, see the TRAIL website: www.rsTRAIL.nl.

The TRAIL Thesis Series is a series of the Netherlands TRAIL Research School on transport, infrastructure and logistics.

Jiao, Y., *Proactive Collision Risk Quantification in Multi-directional Traffic Interactions*, January 2026, TRAIL Thesis Series, the Netherlands

Asadi, M., *Accessibility and Road Safety: Integration of road safety in accessibility evaluation*, November 2025, TRAIL Thesis Series, the Netherlands

Akse, R., *Understanding and untangling the uncertainty knot: How to catalyse decision-making in mobility innovations*, November 2025, TRAIL Thesis Series, the Netherlands

Führer, K., *Participatory Decision-making under Deep Uncertainty: Modelling mobility transitions*, November 2025, TRAIL Thesis Series, the Netherlands

Picco, A., *Monitoring and Feedback in Driving*, T2025/17, October 2025, TRAIL Thesis Series, the Netherlands

Cebeci, M.S., *Behaviour of Prosumers in Last-mile Logistics: The case of crowdshipping*, T2025/16, September 2025, TRAIL Thesis Series, the Netherlands

Kuijpers, A., *Enabling Inter-Organizational Collaboration Through Platforms: The role of trust*, T2025/15, September 2025, TRAIL Thesis Series, the Netherlands

Song, R., *Human-MASS Interaction in Decision-Making for Safety and Efficiency in Mixed Waterborne Transport Systems*, T2025/14, June 2025, TRAIL Thesis Series, the Netherlands

Destyanto, A.R., *A Method for Evaluating Port Resilience in an Archipelago*, T2025/13, June 2025, TRAIL Thesis Series, the Netherlands

Karademir, C., *Synchronized Two-echelon Routing Problems: Exact and approximate methods for multimodal city logistics*, T2025/12, May 2025, TRAIL Thesis Series, the Netherlands

Vial, A., *Eyes in Motion: A new traffic sensing paradigm for pedestrians and cyclists*, T2025/11, May 2025, TRAIL Thesis Series, the Netherlands

Chen, Q., *Towards Mechanical Intelligence in Soft Robotics: Model-based design of mechanically intelligent structures*, T2025/10, April 2025, TRAIL Thesis Series, the Netherlands

Eftekhar, Z., *Exploring the Spatial and Temporal Patterns in Travel Demand: A data-driven approach*, T2025/9, June 2025, TRAIL Thesis Series, the Netherlands

Reddy, N., *Human Driving Behavior when Interacting with Automated Vehicles and the Implications on Traffic Efficiency*, T2025/8, May 2025, TRAIL Thesis Series, the Netherlands

Durand, A., *Lost in Digitalisation? Navigating public transport in the digital era*, T2025/7, May 2025, TRAIL Thesis Series, the Netherlands

Dong, Y., *Safe, Efficient, and Socially Compliant Automated Driving in Mixed Traffic: Sensing, Anomaly Detection, Planning and Control*, T2025/6, May 2025, TRAIL Thesis Series, the Netherlands

Droffelaar, I.S. van, *Simulation-optimization for Fugitive Interception*, T2025/5, May 2025, TRAIL Thesis Series, the Netherlands

Fan, Q., *Fleet Management Optimisation for Ride-hailing Services: from mixed traffic to fully automated environments*, T2025/4, April 2025, TRAIL Thesis Series, the Netherlands

Hagen, L. van der, *Machine Learning for Time Slot Management in Grocery Delivery*, T2025/3, March 2025, TRAIL Thesis Series, the Netherlands

Schilt, I.M. van, *Reconstructing Illicit Supply Chains with Sparse Data: a simulation approach*, T2025/2, January 2025, TRAIL Thesis Series, the Netherlands

Ruijter, A.J.F. de, *Two-Sided Dynamics in Ridesourcing Markets*, T2025/1, January 2025, TRAIL Thesis Series, the Netherlands

Fang, P., *Development of an Effective Modelling Method for the Local Mechanical Analysis of Submarine Power Cables*, T2024/17, December 2024, TRAIL Thesis Series, the Netherlands

Zattoni Scroccaro, P., *Inverse Optimization Theory and Applications to Routing Problems*, T2024/16, October 2024, TRAIL Thesis Series, the Netherlands

Kapousizis, G., *Smart Connected Bicycles: User acceptance and experience, willingness to pay and road safety implications*, T2024/15, November 2024, TRAIL Thesis Series, the Netherlands

Lyu, X., *Collaboration for Resilient and Decarbonized Maritime and Port Operations*, T2024/14, November 2024, TRAIL Thesis Series, the Netherlands

Nicolet, A., *Choice-Driven Methods for Decision-Making in Intermodal Transport: Behavioral heterogeneity and supply-demand interactions*, T2024/13, November 2024, TRAIL Thesis Series, the Netherlands

Kougiatsos, N., *Safe and Resilient Control for Marine Power and Propulsion Plants*, T2024/12, November 2024, TRAIL Thesis Series, the Netherlands

Uijtdewilligen, T., *Road Safey of Cyclists in Dutch Cities*, T2024/11, November 2024, TRAIL Thesis Series, the Netherlands