TRAFFIC CONFLICT ANALYSIS BY RISK ESTIMATION AND PRIVACY-PRESERVING COLLABORATION FOR MOBILITY SAFETY

Fernando Molano Ortiz

Tese de Doutorado apresentada ao Programa de Pós-graduação em Engenharia Elétrica, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Doutor em Engenharia Elétrica.

Orientador: Luís Henrique Maciel Kosmalski Costa

Rio de Janeiro
Setembro de 2022

TRAFFIC CONFLICT ANALYSIS BY RISK ESTIMATION AND PRIVACY-PRESERVING COLLABORATION FOR MOBILITY SAFETY

Fernando Molano Ortiz

TESE SUBMETIDA AO CORPO DOCENTE DO INSTITUTO ALBERTO LUIZ COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE ENGENHARIA DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE DOUTOR EM CIÊNCIAS EM ENGENHARIA ELÉTRICA.

Orientador: Luís Henrique Maciel Kosmalski Costa

Aprovada por: Prof. Luís Henrique Maciel Kosmalski Costa
Prof. Divanilson Rodrigo de Sousa Campelo
Prof. Marcelo Gonçalves Rubinstein
Prof. Miguel Elias Mitre Campista
Prof. Rodrigo de Souza Couto

RIO DE JANEIRO, RJ – BRASIL
SETEMBRO DE 2022

Ortiz, Fernando Molano

Traffic Conflict Analysis By Risk Estimation And Privacy-Preserving Collaboration For Mobility Safety/Fernando Molano Ortiz. – Rio de Janeiro: UFRJ/COPPE, 2022.

XIX, 107 p.: il.; 29, 7cm.

Orientador: Luís Henrique Maciel Kosmalski Costa

Tese (doutorado) – UFRJ/COPPE/Programa de Engenharia Elétrica, 2022.

Referências Bibliográficas: p. 97 – 107.

1. Exteroceptive Sensors. 2. Smart Mobility. 3. Self-Driving Vehicles. 4. Intelligent Vehicles. I. Costa, Luís Henrique Maciel Kosmalski. II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia Elétrica. III. Título.

To my family.

Acknowledgments

I want to especially thank my advisor Luís for accepting me as student, especially for his patience, support, respect, trust and his friendship during this period, and for all the care he took with me during the pandemic. I have certainly learned many things from him, both professionally and in life. I also thank Professor Miguel, who together with prof. Luís gave me the opportunity to work on the AXA project, from which this thesis emerges. This job demanded a lot of effort from me, and everything I expected to evolve in my learning. I hope I have contributed appropriately to the challenge.

I also thanks to Matteo Sammarco immensely for your patience, the exchange of ideas and all the meetings and suggestions throughout these two years. For me, this has been a learning stage where your participation has been essential for my professional and personal evolution. Each meeting was a challenge and I was happy to work under his guidance.

I also thank Professors Rodrigo de Souza, Pedro Cruz and Pedro Velloso for their patience and help in my training, for their patience in answering my questions. To prof. Otto (*in memoriam*) I thank him for all his advices for my adaptation, because despite his ways, he also had words of praise for me, because he was a great example of dedication. Thank you!

I thank Professors Divanilson Rodrigo de Sousa Campelo, Marcelo Gonçalves Rubinstein, Miguel Elias Mitre Campista and Rodrigo de Souza Couto for their participation in the examination jury.

To my colleagues in GTA lab, you cannot imagine how important it has been to share this time with you, Hugo Sadok, JB (*in memoriam*), Lucas Gomes, Gabriel Rebello, Lucas Airam, Gustavo, Roberto, Luana, who kindly helped me and supported me in the development of my work; in addition to contributing to my training and motivation, I am grateful for the good times shared with you in the lab.

To my colleagues in room 23, I was happy in your company. We laughed, supported and helped each other somehow overcome our challenges. Ana Elisa, much of my success is attributed to you. Your support and effort were vital for our work to progress; I never stop thinking about our “company inc”. Thales, “jovem jovem”, all my admiration for your dedication, your effort and your sacrifice. I just have to

thank you for giving me your friendship and trust during this time. We all make it through this stage. I hope to meet you in new experiences.

To the friends of "Wakanda Pitanga", thank you for your friendship, for the meetings, for all the conversations, the diverse opinions and all the occurrences typical of the group diversity. Without doubt I will continue burning meat and pineapple, I will continue with the RPGs and mocking and laughing at how many crazy things you send.

Dianne and Matheus, I especially thank you for giving me your friendship and sharing pleasant moments with me. Your friendship is unique and I treasure it dearly. I had never tried so many forms of churrasco, enjoyed a few beers so much, played so many board games, in short, so many things that I cherish.

I thank my mother Edilsa, my grandma Ace, and my aunt Nuria, who constantly supported me with their words, advice, and experience, which make me feel very close to them. Without importing the distance, they always have given me the impulse not to give up on my projects and ambitions.

Nori, all of this would not have happened without your help. You covered my back so that I could dedicate myself exclusively to my work, you have accompanied me through the good times and the less good ones. I want to thank you for all your unconditional support. I love you!

I thank the staff of the Electrical Engineering Department (PEE) of COPPE/UFRJ, Daniele, Mauricio, Roberto and Marcos for prompt service in the department secretariat.

Finally, I thank CAPES - Finance Code 001, CNPq, FAPERJ, FAPESP Grant 15/24494-8, COPPETEC and AXA GO Advanced AI/ML & Research Team, for the funding of this work and my support in pandemic.

When you experience some change, there is a lot of new stimuli that would come into your life, and I can only to be graceful for the opportunities received here in Brazil. It was really great to have what I had for 6 years in GTA laboratory, and then to see my evolution, my partners' successes, and that was the part that was gonna challenge me the most. How could I go into a situation where I had no familiarity and try to work as hard as I could to put the best of me everyday? I have grown and learned in ways that I could never have learned if I had not left my comfort zone. That was the greatest challenge I had in my life. Welcome the next ones...

Resumo da Tese apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Doutor em Ciências (D.Sc.)

ANÁLISE DE CONFLITOS NO TRÂNSITO BASEADA EM ESTIMAÇÕES DE RISCO E COLABORAÇÃO COM PRESERVAÇÃO DA PRIVACIDADE PARA A SEGURANÇA DA MOBILIDADE

Fernando Molano Ortiz

Setembro/2022

Orientador: Luís Henrique Maciel Kosmalski Costa

Programa: Engenharia Elétrica

A auditoria e o monitoramento veicular são áreas de importância crescente nos sistemas inteligentes de transportes. Nesse sentido, as seguradoras têm desenvolvido diferentes estratégias para descrever perfis de condução dos motoristas, através de sensores embarcados (p. ex., black-boxes, OBD-II dongles, smartfones, entre outros) que transmitem dados em tempo real. No entanto, com o avanço na implementação de sistemas inteligentes nos veículos, a tendência é do condutor humano ficar em segundo plano. Portanto, é necessário estabelecer parâmetros que descrevam de forma eficiente perfis de condução de veículos inteligentes. Esta tese investiga soluções para propor um modelo de avaliação de risco para veículos inteligentes. Três áreas são estudadas: sensoriamento, infraestrutura e interações com o veículo de interesse, e ambientes colaborativos com preservação da privacidade dos dados. Em princípio, analisam-se sensores exteroceptivos e os dados obtidos destes sensores em veículos autônomos (AVs) experimentais. Em continuação, são analisadas métricas de segurança substitutas baseadas nas leituras dos sensores, contidas em conjuntos de dados de AVs. Baseado nos dados de detecção de objetos, esta tese propõe a métrica de Tempo para Colisão com orientação de movimento, uma medida de estimação de segurança que permite quantificar o risco baseado na orientação na qual se deslocam os objetos que interagem com o veículo de interesse. Dado que os fluxos de dados no ambiente veicular mudam dinamicamente, o objetivo final do trabalho é desenvolver um modelo para avaliar o risco baseado no compartilhamento de dados entre o fabricante de veículos e a seguradora. Para isso, este trabalho implementa a técnica de aprendizado federado vertical para garantir a preservação de segurança e integridade dos dados no ambiente colaborativo.

Abstract of Thesis presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Doctor of Science (D.Sc.)

TRAFFIC CONFLICT ANALYSIS BY RISK ESTIMATION AND PRIVACY-PRESERVING COLLABORATION FOR MOBILITY SAFETY

Fernando Molano Ortiz

September/2022

Advisor: Luís Henrique Maciel Kosmalski Costa

Department: Electrical Engineering

Vehicle's monitoring and audit are areas of growing importance in intelligent transportation systems. In this sense, insurers have developed different strategies to describe driving profiles through embedded sensors (e.g., black-boxes, OBD-II dongles, smartphones, among others) transmitting data in real-time. However, with the advancement in the implementation of self-driving vehicles, the driver goes to the background. Therefore, it is necessary to establish parameters that efficiently describe self-driving vehicle's driving profiles. Thus, this thesis investigates solutions to propose a risk assessment model for self-driving vehicles. Three areas are studied: sensing, infrastructure and interactions with the ego-vehicle, and collaborative environments with privacy-preserving data. In principle, exteroceptive sensors and the data obtained from these sensors in experimental Autonomous Vehicles (AVs) are analyzed. Next, Surrogate Safety Measures (SSMs) based on data from sensor readings contained in AV datasets are analyzed. Based on object detection data, this thesis proposes the Time-to-Collision with motion-orientation ($\mathrm{TTC_{mo}}$) metric, a safety estimation measure that allows risk quantification based on the yaw orientation of detected objects interacting with the ego-vehicle. Given that data flows in the vehicular environment change dynamically, the last objective of this thesis is to develop a risk assessment model based on data sharing between the car manufacturer and the insurer. Thus, this work implements a vertical federated learning framework to ensure privacy protection and data integrity in a two-party collaborative environment.

Contents

List of Figures

List of Tables

List of Abbreviations

ABS	Anti-lock Braking Systems, p. 3
ACC	Adaptive Cruise Control, p. 3
ADAS	Advanced Driving Assistance System, p. 4
AI	Artificial Intelligence, p. vi, 8
AMCW	Amplitude-Modulated Continuous Wave, p. 18
AV	Autonomous Vehicle, p. viii
CAN	Controller Area Network, p. 11
CI	Crash Index, p. 47
CSV	Comma-Separated Values, p. 33
CW	Continuous Wave, p. 16
DCM	Digital Code Modulation, p. 17
DFoV	Diagonal Field of View, p. 20
DNN	Deep Neural Network, p. 85
DRAC	Deceleration Rate to Avoid a Crash, p. 44
DTLS	Datagram Transport Layer Security, p. 80
D	Dimensionality, p. 18
DoA	Direction of Arrival, p. 15
ECG	Electrocardiogram, p. 14
ECU	Electronic Control Unit, p. 12
EM	Electromagnetic, p. 16

ERSO	European Road Safety Observatory, p. 2
ESC	Electronic Stability Control, p. 3
ETTC	Enhanced Time-to-Collision, p. 47
FL	Federated Learning, p. 6
FMCW	Frequency-Modulated Continuous Wave, p. 16
FoV	Field of View, p. 17
GLONASS	Global Navigation Satellite System, p. 15
GNSS	Global Navigation Satellite System, p. 11
GPS	Global Positioning System, p. 15
HCR	High-Contrast Resolution, p. 17
HFL	Horizontal Federated Learning, p. 6
HFoV	Horizontal Field of View, p. 19
HMT	Human-Machine Transition, p. 3
HSV	Hue, Saturation, Value, p. 19
IMU	Inertial Measurement Unit, p. 26
JPEG	Joint Photographic Experts Group, p. 28
JSON	JavaScript Object Notation, p. 28
LBS	Location-Based Services, p. 15
LO	Local Oscillator, p. 18
LRR	Long-Range Radar, p. 17
LiDAR	Light Detection And Ranging, p. 11
LoS	Line-of-Sight, p. 20
MEMS	Microelectromechanical Systems, p. 15
MEO	Medium Earth Orbit, p. 15
MIMO	Multiple Input Multiple Output, p. 17

ML	Machine Learning, p. vi
MTTC	Modified Time-to-Collision, p. 47
NHTSA	National Highway Traffic Safety Administration, p. 2
NIR	Near-Infrared Region, p. 19
NLoS	Non-Line-of-Sight, p. 15
OBD	On-Board Diagnostic, p. 11
OEM	Original Equipment Manufacturer, p. 7
OPA	Optical Phased Array, p. 19
OTS	Off-the-Shelf, p. 10
PAYD	Pay As You Drive, p. 5
PA	Palo Alto, p. 65
PC	Point Cloud, p. 7
PET	Post Encroachment Time, p. 44
PHYD	Pay HowYou Drive, p. 5
PSD	Proportion of Stopping Distance, p. 44
RGB	Red, Green, Blue, p. 19
RTD	Resistance Temperature Detector, p. 11
Radar	Radio Detection And Ranging, p. 11
SAE	Society of Automotive Engineers, p. 3
SG	Singapore, p. 26
SMOTE	Synthetic Minority Oversampling TEchnique, p. 83
SPMD	Safety Pilot Model Deployment, p. 50
SRR	Short-Range Radar, p. 17
SSM	Surrogate Safety Measure, p. viii
TA	Time-to-Accident, p. 65

TET	Time Exposed Time to Collision, p. 48
TIT	Time Integrated Time to Collision, p. 48
TTCD	Time-to-Collision with Disturbance, p. 47
TTC_{mo}	Time-to-Collision with motion orientation, p. viii
TTC	Time-to-Collision, p. 44
ToF	Time-of-Flight, p. 16
UBI	Usage-Based Insurance, p. 5
UDP	User Datagram Protocol, p. 80
US	United States, p. 26
VFL	Vertical Federated Learning, p. 6
VFoV	Vertical Field of View, p. 19
VIS	Visible Imaging Sensor, p. 19

Chapter 1

Introduction

On-board decisions in the vehicle generally have been associated with the driver. However, due to high accidents and crashes involving human errors, car manufacturers have been implementing several driver support services in order to ensure the safety of passengers in vehicles. The evolution of these services has allowed the early adoption of autonomous driving systems. For example, Tesla offers automated driving system features without any expectation that the user will respond to a request to intervene [9]. Nonetheless, some fatal crash events involving AVs have been reported [10]. These incidents make it necessary to establish parameters to monitor the safety of passengers in self-driving vehicles. Nevertheless, self-driving vehicles generate divergences in the analysis of traffic risk events and their severity. This discussion involves numerous questions of technical, ethical, and social nature since driving will correspond to the artificial intelligence system that manages the vehicle [11]. In this scenario, since the driver is no longer responsible, it is necessary to monitor systems in vehicles associated with human perception, which are crucial in the proper functioning of autonomous driving systems. Thus, this work evaluates three aspects that we consider fundamental to assess the risk in self-driving vehicles: sensing, which is directly associated with how the environment around the vehicle is perceived; the infrastructure and interactions with other road users, associated with the object detection and the recognition of properties inherent to the road users and infrastructure; and lastly, data sharing, since the risk assessment will require data from the autonomous driving system in addition to those generated by the insurer, and therefore, it is necessary to evaluate specific moments that can be analyzed more simply in a collaborative environment between manufacturers and insurers, where they can share data without compromising data integrity and privacy.

1.1 Motivation

Mobility safety aims to improve the security of the road users – may they be drivers, passengers, pedestrians – by providing strategies to reduce the negative societal impacts from traffic risk events. Nonetheless, improving road safety is a challenging scenario since it requires covering diverse topics that converge in traffic risk events, for example, driving profiling, vehicle's mechanical state, interactions with other road users, among others. As a matter of fact, traffic accidents are one of the major causes of loss of human lives. In 2018, road traffic crashes represented the eighth leading cause of death globally, causing up to 1.35 million people died, and 50 million non-fatal injuries that year [12]. The European Road Safety Observatory (ERSO) states that traffic injuries produce socio-economic consequences estimated at €120 billion annually [13]. The National Highway Traffic Safety Administration (NHTSA) reports, in 2016, there were 38,824 people killed in motor vehicle crashes on the U.S. roadways, and that human errors were the leading cause in 94% to 96% of all motor vehicle crashes [14].

As a logical consequence, governments and researchers around the globe pursue innovative services to improve road safety policies through autonomous support and assistance systems, as an immediate response to risk events. The key idea of under development technologies consists of monitoring several aspects of driving through the analysis of traffic risk events based on data collected from the driver and the vehicle's surrounding environment. Application examples are driver detection, driver distraction, object detection, among others. These applications are relevant since driving behavior and vehicle interactions depends on multiple sub-areas to describe properly the context involving road safety. In fact, each application has different information requirements from vehicle in-cabin and surroundings. In general, safety monitoring depends on in-cabin and external sensing capabilities available in the vehicle. Therefore, a myriad of sensors come at play. In addition to the sensors needed for any vehicle operation, other sensors are needed to allow the vehicle to capture information about the current situation and informing the driver or, ultimately, taking actions autonomously.

1.2 Sensors in Intelligent Vehicles

Sensing plays a vital role for vehicle monitoring. Sensing can be described as the process of perception of different variables involved with the vehicle operation. Through this process, sensor data readings are organized, identified and interpreted, in order to represent and understand the situation around the vehicle. In-vehicle perception methods are closely related to sensors. Nowadays, vehicles are equipped

with embedded systems that digitally control different vehicle subsystems, resulting from the stimulation of on-board sensors, providing different functionalities [15], e.g., turbo boost, fuel injection, active suspension, vehicle stability, among others. Thus, vehicle sensors generate massive amounts of data, which are used to analyze behavior patterns, environmental conditions and to optimize vehicle performance and security. Therefore, perception systems have to be precise, robust, and frequently processed in real-time.

In-vehicle sensing enables the integration of different systems to improve, adapt and automate vehicle safety and the driving experience. These systems assist drivers by offering precautions to reduce risk exposure or by cooperatively automating driving tasks (e.g., Anti-lock Braking Systems (ABS), Adaptive Cruise Control (ACC), Electronic Stability Control (ESC), among others) with the aim of minimizing human errors, or the effects of human errors at least [16]. In this scenario, it is not only possible to obtain internal measurements of the vehicle associated with the engine and its components. Now the vehicle can acquire information from the surrounding environment, recognizing other factors and objects which coexist with its environment, which makes each variable a measurement to be considered in any scenario. Therefore, the vehicle sensor becomes multi-mode, with the capacity to perceive both internal and external signals. It also means that the vehicle must operate with heterogeneous data from multiple sources. Moreover, this is relevant if we consider that self-driving vehicles will reduce the dependency on the human driver since an intelligent control system will control the vehicle. In this regard, in-vehicle and external environment sensing take on greater relevance, since the decisions in the vehicle will be made from the sensor readings.

Driver assistance and support services are constantly evolved, with functionalities that have some autonomy level, converging towards Autonomous Vehicles (AVs). Currently, the Society of Automotive Engineers (SAE) classifies the vehicles' autonomy level [17]. This classification is linked to specific roles involved in the dynamic driving task: human, vehicle systems, and autonomous systems. A driver fully responsible for driving represents less complexity related to assistance systems; in contrast, vehicle sensing is limited. Meanwhile, it is stated that vehicles with higher complexity in their assistance systems require more advanced sensing devices. Therefore, full automation has complex computational methods, since exhaustive monitoring of the driving environment and full situational awareness are of paramount importance, once its ultimate goal is being capable of replacing the human driver [17]. It is worth mentioning that since vehicles with full autonomy are not yet available, it means the driver must take control and react to situations in which the autonomous support system does not respond adequately. These Human-Machine Transitions (HMT) are explored in [11, 18].

Driving automation requires both "proprioceptive" and "exteroceptive" sensing. Proprioceptive corresponds to data readings from the internal vehicle sensors. On the other hand, exteroceptive sensing corresponds to data readings from the external vehicle sensors. It is worth noting that proprioceptive sensors are not enough to provide safety applications that involve the monitoring of other vehicles, since they do not acquire data from the external environment. Withal, based on readings from proprioceptive sensors, it is possible to obtain intrinsic variables from the vehicle, such as speed, fuel consumption, among others. With exteroceptive sensors instead, vehicles can acquire information on the surrounding environment, recognizing other factors and objects that coexist in the same space. Vehicle external sensing is gaining importance, especially with the proliferation of cameras which, combined with improved image processing and analysis, enables a wide range of applications [19].

Data generated by exteroceptive sensors is of primary interest to drivers and passengers. In fact, environmental sensing is fundamental for Advanced Driving Assistance System (ADAS). ADAS relies on multiple data sources available in the vehicle (e.g., external sensing through cameras, radar, LiDAR, in-car networking, vehicular communications, among others) enabling the implementation of diverse safety applications in the vehicle, such as ACC, collision avoidance, anti-lock braking system, among others. Interestingly enough, the same data has value in the context of smart cities (e.g., sensing the road conditions), and for insurance companies to establish driving profiles and calculate insurance premiums. Ubiquitous sensing methods have allowed monitoring vehicles and driving environments. It is inferred by the proliferation of wireless communications, as well as devices with processing and sensing abilities. As a result, data volume generated by both proprioceptive and exteroceptive sensors can be high depending on the type of sensor, reaching up between 11 TB and 152 TB per day, just for one vehicle [20]. In this way, vehicle data volume shows an exponential growth over the recent years. According to an estimate by IBM, 2.5 quintillion bytes of data are created each day, whilst modern cars have close to 100 sensors that monitor items related to the vehicle [21], generating data from terabytes (TB) to petabytes (PB) level. It is relevant since vehicular safety applications require immediacy in data analysis and real-time operations. Therefore, strategies are required to optimize data analysis to identify risk events, automatic incident detection, monitoring traffic behavior, simulation models, among others. For that, analysis of real vehicular mobility traces is fundamental to extract data concerns to the driving behavior. Moreover, processing data is crucial to safety applications in vehicular telematics. Since traffic data change rapidly, it is necessary to use historical data to identify patterns, and therefore, it must be compared and processed in a short time [22].

1.3 Risk Assessment

In addition to sensing and data analysis, another aspect that must be addressed is the implementation of methods for understanding and predicting safety performance to make an etiological diagnosis that validates those events that really represent a risk for vehicle and passengers. This is a complex task if we consider that the randomness of the vehicular environment requires monitoring in different segments. For instance, driving requires continuous monitoring since driver and vehicle can interact with multiple road users and infrastructure that can become risky at any time. Therefore, it is also necessary to evaluate the interactions with the environment outside the vehicle. Currently, driving style assessment models are based on driving profiles inherent to human drivers [23]. For this, different driver monitoring strategies have been developed to study driving habits in practice. Driver behavior is classified into macro-areas that consider diverse events in driving practice [24], e.g., safety, driving behavior, road monitoring, navigation, among others. In the insurance market, the strategies most used to determine the calculation of insurance policies are based on Usage-Based Insurance (UBI) metrics. Pay As You Drive (PAYD) and Pay How You Drive (PHYD) [23] are techniques that analyze the vehicle displacements, time of use, and driver behavior respectively. Thus, it is possible to consider the driver behavior in defining the value of the insurance policy.

Nowadays, with the advancement in the implementation of self-driving vehicles, the driver goes to the background. Then, a new question arises, which is *how to establish a risk assessment plan to ensure a self-driving vehicle?* To formulate these parameters, it is essential to analyze the data collected from sensors, to describe different profiles of self-driving vehicles in the presence of risk factors for the vehicle and its occupants. Nonetheless, the amount of data can introduce noise for the characterization of these risk factors to establish specific insurance services. One solution can be to filter this large volume of big data to identify critical characteristics of the self-driving vehicle's behavior. Indeed, self-driving vehicles generate divergences in the analysis of traffic risk events and their severity. This discussion involves numerous questions of technical, ethical, and social nature since driving will correspond to the artificial intelligence system that manages the vehicle [11]. Besides that, HMT involves various questions associated with privacy preservation and data security [25], processes are required to guarantee the data integrity, both in handling and anonymization. Moreover, HMT has to deal with the complexity of decisions and how these are approached. It constitutes a challenging scenario since, in addition to technicalities, it is necessary to establish a trade-off between responsibilities and ethical standards, to define policies associated with the AVs functionalities and legislation [18]. It is relevant for insurance companies, given that they must guar-

antee fairness and explainability to the risk events that may occur with the AV [26] to the customers. On the other hand, this is also subject to legislative regulations, which are still lagging when comparing to the AVs' technological advances.

The interpretation of traffic risk events is correlated with the driver/vehicle and his/its response to the variability of interactions with various road users. Nevertheless, this depends on observational estimates that allow the risk to be assessed. Furthermore, these estimates are limited by the absence of crash data or tracking of sequences of events leading to a high-risk event since crashes are the main focus. Nonetheless, the use of non-crash traffic risk events allows us to assume a relationship between the severity and the frequency of different events involving the AV. By understanding these relationships, it is possible to study road safety risks without relying on accident data and improve the prediction of other risk events.

It is important to note that any potential risk event occurs between at least two road infrastructure objects which interact on the road and can be explained by several factors involved in any event. Therefore, any event in traffic is associated with a severity level that can describe or anticipate risk events and their consequences. A strategy to quantify traffic risk events with non-crash traffic data are often evaluated through Surrogate Safety Measures (SSMs). SSMs allow describing the probability of a crash, as the frequency of interactions with different objects on the road (e.g., time difference between vehicles, distance to an object, speed compared with other road users around, among others). Although SSMs are not designed to carry out accident prevention actions, these allow observing aspects of the vehicle's behavior, containing useful information to follow various processes associated with road safety.

1.4 Federated Learning

In addition to SSMs, risk assessment can be done collaboratively. Thus, a goal is to take advantage of the data generated by the self-driving vehicle and the monitoring carried out on it by third parties to properly capture the causality of traffic risk events. Currently, data collection from both car manufacturers and auto insurers is done individually. These data open a window of opportunity for the use of Machine Learning (ML) techniques that enable data sharing to create reliable collaborative ecosystems between manufacturers and policymakers. However, there are privacy and data security requirements to operate with them [27].

One ML approach that implements data privacy-preservation techniques is Federated Learning (FL) [28]. This technique allows data to be manipulated in a distributed way among different users, overlapping datasets with different characteristics. This overlapping can occur Horizontally (HFL) or Vertically (VFL). HFL is a sample-based methodology where data from diverse sources (customers) contain

the same feature space in a distributed manner, but different in samples. On the other hand, VFL analyzes data and labels with different features, i.e., labels are not contained in the data, but their significance is relevant in the model analysis.

We aim to use VFL techniques to establish a partner ecosystem to improve the self-driving model through risk assessment analysis. VFL is applicable in cases where two organizations share a group of customers, but each organization owns different data/features. To implement VFL in this ecosystem, we use AV datasets that contain raw and semantic data from the vehicle while it is in motion. Data exchange is advantageous for both partners: the car manufacturer benefits because a third party can help improve its self-driving model; meanwhile, on the insurer's side it can help reduce claims, which is beneficial for customers. For instance, exploring the partner's ecosystem allows the effective participation of different entities, from Original Equipment Manufacturer (OEMs), communication providers to government entities. Thus, it is possible to evolve towards new learning models that allow coexistence and evolution between current driving profile monitoring models, ensuring data privacy and integrity.

1.5 Objectives

The goal of this thesis manuscript is to present the research work and the obtained results achieved so far. Since self-driving vehicles require special attention to describe risk assessments is scarcely explored up to now, it is the major issue of our investigation. Under the umbrella of this issue, we organized our work in four specific objectives: sensors in intelligent vehicles, datasets and feature selection, risk assessment based on SSMs, and risk assessment based on data sharing between trusted organizations using VFL. Next, we briefly describe these research topics.

- **Sensors in intelligent vehicles:** A state-of-art about sensors in intelligent vehicles review is carried out. This review aims to identify the devices currently used in vehicular telematics and the sensors they use. Moreover, this work emphasizes exteroceptive sensors since these allow to detect variables related to the external environment around the self-driving vehicle.

- **Data analysis:** We analyzed two datasets, nuScenes AV dataset [1], and Lyft5 [2]. nuScenes and Lyft5 are public large-scale datasets for autonomous driving, based on images from camera, point clouds (PC) from LiDAR, and radar signals detected by the sensors installed in the vehicle, besides categorized data. The goal is to analyze the dynamics of road users and the ego-vehicle to evaluate metrics inherent to the objects' motion.

- **Risk assessment based on SSMs:** This work uses SSM techniques for road safety analysis. Through these techniques, one goal of this work is to identify what type of data from the sensors are most appropriate for the calculation of road safety metrics. We propose and implement the Time-to-Collision with motion orientation (TTC_{mo}) technique and introduce the analysis of yaw orientation of each object detected, and position of these w.r.t. the ego-vehicle, in order to analyze only the detected objects that are on a collision course with the self-driving vehicle. Using the 3D object detection data annotations available from the publicly available AV datasets nuScenes and Lyft5 and the TTC_{mo} metric, we find that at least 8% of the interactions with objects detected around the AV present some risk level. This is meaningful, since it is possible to reduce the proportion of data analyzed by up to 60% when replacing regular TTC by our improved TTC computation.

- **Risk assessment based on data sharing between trusted organizations using ML techniques:** The main goal of risk assessment based on data sharing is to create a distributed environment where car manufacturers and insurers can improve the self-driving model and reduce claims, respectively. Thus, it is possible that several partners can share samples in similar time intervals. This messages exchange allows partners to access a part of a complete model at the same time that they can run a training process for the segments they want to analyze. For that, we aim to emulate VFL techniques at the edge using data from self-driving vehicles as data owner, and we define labels from the insurer as data scientist to learn and detect risk from data owner with privacy-preserving. Once the VFL environment is established, we compare its convergence compared to a local model. The results show that the convergence of the models is close, with the VFL model converging faster than the local model (fewer epochs to converge to maximum accuracy and minimum loss), however, with a longer learning time compared to the local model. It is also possible to observe that the classification of the models identify risk events.

The subject of this thesis is part of the research collaboration between AXA, a European insurance company, and Universidade Federal do Rio de Janeiro. Eight students from UFRJ are engaged in the topic "Safety aspects of transports, connected and autonomous vehicles". The project is supported by industrial partner AXA GO Advanced AI/ML & Research Team.

1.6 Organization

This thesis manuscript is organized as follows. Chapter 2 reviews the sensors that can be used in vehicular telematics involved in driving behavior analysis, environmental perception, and object detection in the context of autonomous driving. Also, various sensors for environment detection will be presented with particular attention in the exteroceptive sensors. Chapter 3 describes the data collection, preparation and analysis used to calculate motion properties and dynamics of both AV and detected objects. Chapter 4 proposes a TTC with motion orientation analysis for each object detected by the self-driving vehicles described in Chapter 3. Chapter 5 presents a practical implementation emulating a real scenario of collaborative training of a model. Chapter 6 concludes this work and presents future research directions.

Chapter 2

Sensors in Intelligent Vehicles

This chapter provides the required knowledge related to the diverse sensor types associated with vehicle driving, driving patterns, and safety services, mainly contained in self-driving vehicles. First, basic concepts about exteroceptive and proprioceptive sensors used in telematics are provided. Secondly, Off-the-Shelf (OTS) devices used for insurance telematics are described. Finally, this section emphasizes exteroceptive sensors, their advantages and limitations.

2.1 Background

Sensors in vehicle telematics enable monitoring a broad range of functions inherent to the management of diverse driving activities. Electronic sensing systems and data processing capacity reduce driver's workload and provide innovative services, e.g., ABS, ACC, and ESC systems. This section presents a classification of sensors for vehicle telematics purposes, according to the environment in which they operate (e.g., in-vehicle, cabin, outdoor). Figure 2.1 illustrates a bipartite graph showing the relationship between sensors and OTS devices, where sensors are placed in the right column; exteroceptive sensors are nodes on the top, while proprioceptive on the bottom. Active sensors are represented as gray-colored nodes, passive sensors as white-colored nodes. On the left-hand side, each sensor is connected to OTS telematics devices where it is embedded.

2.2 Proprioceptive vs. Exteroceptive Sensors

A wide variety of sensors is used in regular vehicles, the majority of them to gather information on internal engine mechanisms. Self-driving vehicles on the other hand incorporate sensors with the ability to measure extrinsic variables, whose function is critical to analyze the surrounding environment. Therefore, vehicular telem-

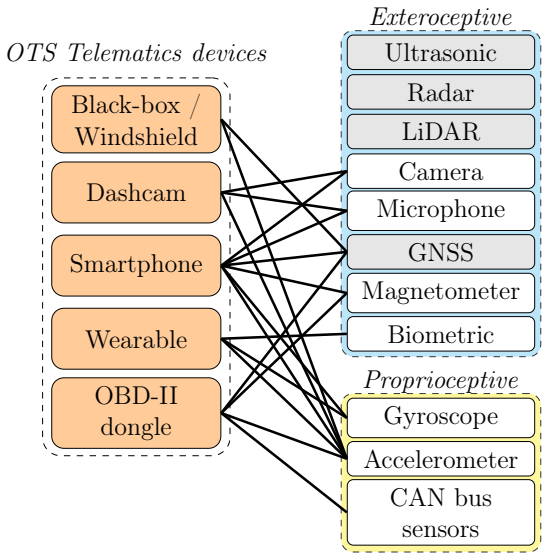


Figure 2.1: Relationship between the most widely used sensors and the most common OTS devices used in vehicular telematics. In the middle, exteroceptive sensors are nodes on the top, while proprioceptive on the bottom. Active sensors are represented as gray colored boxes, passive sensors as white colored boxes.

atics is no longer merely mechanical, leading to the analysis of internal and external variables. As such, a basic classification of the sensors is according to the sensed variables, as *proprioceptive* or *exteroceptive* [29].

Proprioceptive sensors measure variations in signals generated by the vehicle's internal systems (engine speed, battery level, etc.). Those measurements allow estimating different metrics that are specific to the vehicle, such as speed, fluid levels, acceleration, among other topics of interest for vehicle telematics. Tachometers, Resistance Temperature Detector (RTD), encoders, and accelerometers are examples of proprioceptive sensors.

Exteroceptive sensors allow vehicles to be in contact with stimuli coming from the environment surrounding the vehicle. Examples of such external variables are measurements of distance to obstacles, light intensity, sound amplitude, detection of pedestrians, and surrounding vehicles. Therefore, measurements from exteroceptive sensors are interpreted by the vehicle to produce meaningful environmental features. Together, exteroceptive sensors give the AVs a sense of the surrounding environment, which is imperative for autonomous driving.

Proprioceptive sensors, inseparable from vehicle powertrain and chassis, are widely present in production vehicles. In contrast, exteroceptive sensors are mostly available in luxury vehicles, vehicles with some level of autonomy, or experimental vehicles. Conventionally, proprioceptive sensors are designed to measure single-process systems and are therefore limited in capacity. They are unexposed, protected from the external environment. On the other hand, exteroceptive sensors are designed

to analyze and monitor internal (vehicle cabin) and external environments. Thus, they may be designed for operation in different conditions. As such, in some cases they may be subject to harsh environmental conditions, like rain, snow, nighttime, etc. [30].

2.2.1 Active vs. passive sensors

Proprioceptive and exteroceptive sensors are designed to just capture and read a specific metric, or to interact with the environment by observing and recording changes in it, or reactions from it. This leads to classifying sensors as active or passive. *Passive* sensors are able to perform measurements without interacting with the environment, in other words, the sensor receives energy stimuli from the environment. *Active* sensors emit waves outside the vehicle and measure from the return of the emitted signal. Wave emitters can be lasers or radars, among others.

2.3 OTS Devices

While smartphones include a large number of sensors (e.g., GNSS, camera, microphone, accelerometer) which make them particularly convenient for insurance telematics [31], other sensors require dedicated hardware and installation procedures. Next, we present a background of OTS telematics devices that carry proprioceptive and exteroceptive sensors. Figure 2.2 shows a diagram of OTS telematics devices and their interaction in the vehicle.

2.3.1 OBD-II dongles and CAN bus readers

A modern vehicle can contain more than one hundred sensors, generally related to the mechanics, engine operation, and vehicle systems [32]. Automotive systems are mainly concentrated on three areas: powertrain, chassis, and body; each one contains a set of sensors to measure physical quantities, managed by the ECU of each system and interpreted in a look-up table [33]. Data is stored in profiles used to control the vehicle actuators and their performance, e.g., speed control, vehicle stability, among others. The use of specific sensors may also be associated with other factors such as legislation and safety [32]. Data profiles from the ECUs are used to check the vehicle status information through the On-Board Diagnostics (OBD-II) interface. It provides access to the vehicle sub-systems controlled by the ECUs via the Controller Area Network (CAN) bus. Besides, OBD-II is widely used by automotive manufacturers for diagnosis and data analysis. Nevertheless, the data acquisition through the OBD-II connector is limited to a single port and data are specific to each manufacturer, which defines proprietary message codes.

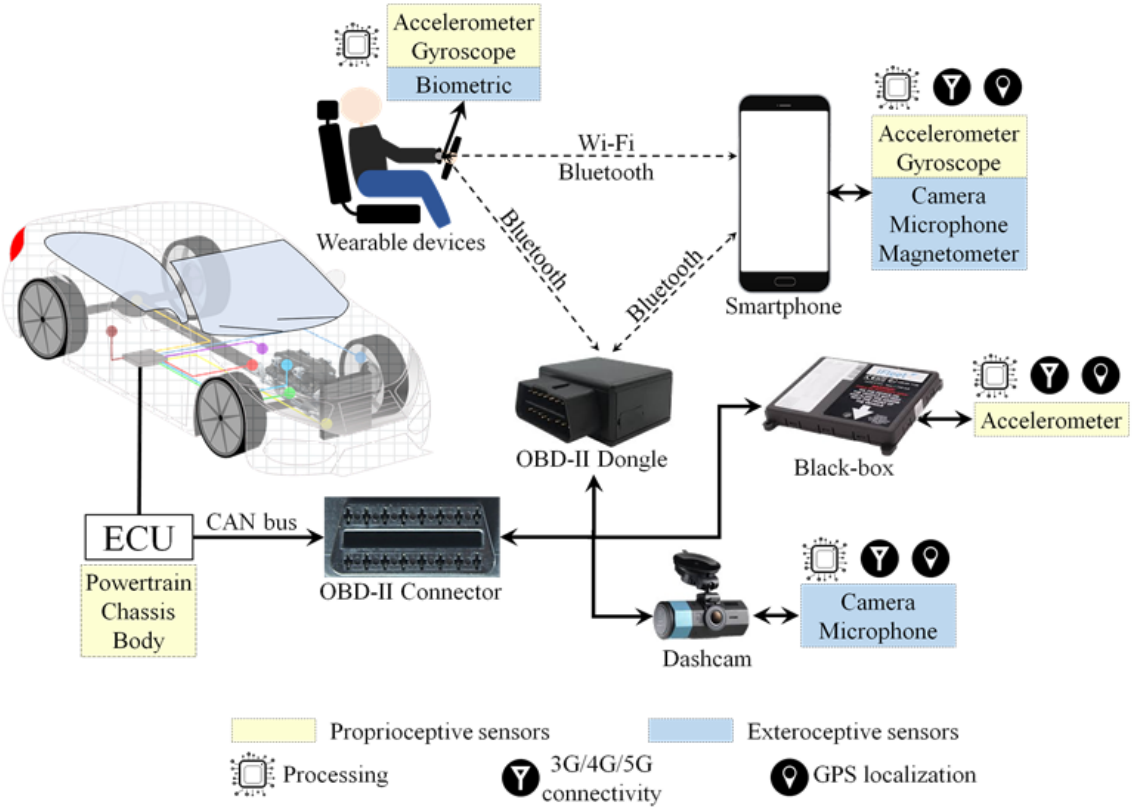


Figure 2.2: Examples of OTS telematics devices, proprioceptive and exteroceptive sensors, and interaction between them in the vehicle.

Meanwhile, commercial OBD-II dongles and CAN bus readers are connected to the vehicle's power source itself, and these can have extra sensors, like a GNSS or an accelerometer.

2.3.2 Black-box and windshield devices

Usually, black-box and windshield devices are installed within the vehicle. They are equipped with self-contained sensor systems, or they can acquire information in a piggyback process via the CAN bus. These devices embed a GNSS and an accelerometer sensor to define driving profiles about harsh acceleration, braking, or impact. In addition, a windshield device may contain a SIM card and a microphone to establish voice communication with remote assistance.

2.3.3 Dashcams

A dashcam is an on-board camera, usually mounted over the dashboard, that records the vehicle's front view. Commonly uses include registering collisions, road hazards, in addition to offering video surveillance services [34]. Since the data volume generated by the video frames is considerable, images are selected beforehand by

the processing system. Additional dashcam functionalities include gesture and voice biometry [35]. Nonetheless, the utilization of dashcams is limited in some countries due to privacy concerns [34].

2.3.4 Smartphones

Smartphones involve diverse technologies that make them sophisticated computers, with the ability to process data and graphics, not to mention communication and sensing capabilities [31]. Smartphones possess a large number of built-in sensors, enabling continuous data collection. Added to mobility, it results in the empowerment of various applications with specific requirements in terms of complexity, granularity, and response time. Moreover, smartphones can acquire data from CAN bus through an OBD-II dongle via Wi-Fi or Bluetooth connection.

2.3.5 Wearable devices

Complementary to smartphones, wearable devices are used to monitor human physiological and biometric signals. In the intelligent vehicles' context, they are used for safety and driving behavior applications. Wearable devices include smartwatches, smart glasses, smart helmets, and electrocardiogram (ECG) sensors [36].

One challenge of built-in sensors in OTS devices is that these have not been designed for vehicular applications and therefore require algorithms to reduce inaccuracies due to their characteristics of manufacturing [31]. Hence, the use of vehicle-fixed sensors is necessary. Exteroceptive sensors used in external and in-vehicle monitoring tasks are described in the next section.

2.4 Exteroceptive sensors in vehicle applications

This section describes in detail exteroceptive sensors, which are used to complement in-vehicle sensing with external information. Therefore, exteroceptive sensing can be used in monitoring systems to audit vehicles during operation, getting accurate information about the vehicle surroundings, relevant in vehicles with some autonomy level, as described in SAE J3016 [17]. Among the main features, exteroceptive sensors can operate in vehicle cabin, and external environments with different extreme conditions (e.g., rain, fog, snow, nighttime, etc.). Moreover, information about the vehicle's surroundings can help to understand better which aspects are involved in traffic conflicts, besides improving other parameters related to the driver/passengers safety.

2.4.1 Global Navigation Satellite System (GNSS)

Some OTS devices implement Location-Based Services (LBS) using an embedded GNSS receiver. GNSS systems enable a quite accurate localization on earth (order of meters), through trilateration signals from dedicated geostationary artificial satellites. GNSS systems are composed of constellations of satellites in the Medium Earth Orbit (MEO) that provide Positioning, Navigation and Timing (PNT) services. The size of the constellations may vary depending on the GNSS system. Among the GNSS systems with global coverage are Global Positioning System (GPS), Global Navigation Satellite System (GLONASS), BeiDou, and Galileo. GNSS systems operate in frequency bands between 1.1 GHz and 1.6 GHz, and varies according to the transmission channels. In fact, different constellations can coexist on the same channel. Depending on the platform on which OEM devices operate, different LBS are offered. In smartphones, some location services merge short and long-range wireless networks such as Wi-Fi, Bluetooth, and cellular networks, in addition to GNSS data [37]. Nowadays, Android and iOS-based devices use messages based on the NMEA 0183 standard [38]. The latest updates to this standard include measurement of the pseudo-range and Doppler shift; this adds simplicity and robustness to the processing of raw GNSS measurements. Nevertheless, GNSS reception exhibits outages due to interference, signal propagation, and measurement accuracy in urban canyons due to multipath effects and Non-Line-of-Sight (NLoS) conditions [37].

2.4.2 Magnetometer

The function of the magnetometer is to read the Earth's magnetic field strength to determine its orientation. Microelectromechanical Systems (MEMS) magnetometers are embedded in commodity devices like smartphones, which inform the magnetic field on the 3-axis (x, y, z) with µT sensibility [39]. Moreover, its miniaturized form factor and low energy consumption favor its availability in multiple devices. Thus, it results as a valuable component for providing navigation and LBS services.

2.4.3 Microphone

A microphone transforms sound waves into electrical energy. These sensors are embedded as MEMS devices or condensed mics that are connected to OTS devices. Microphones are an affordable solution for real-time signal processing. According to ISO 9613-2 standard, their sensing range reaches up to 200 m for high-intensity sounds in an urban scenario [40]. Moreover, microphones consume low energy, have a smaller size, and omnidirectional sensing capability. Devices with an array of microphones are used to estimate the Direction of Arrival (DoA) and localize the

sound source, calculating the time difference of arrival between each microphone pair. On the other hand, their efficiency largely depends on their sensitivity, sound waves amplitude, and environmental noise.

2.4.4 Biometric sensors

Biometric sensors are used to collect measurable biological characteristics (biometric signals) from a human being, which can be used in conjunction with biometric recognition algorithms to perform automated person identification. ECG devices are installed in the steering wheel and in the driver's seat to measure heart activity through touch or photoelectric sensors. To increase driving safety, biometric sensors monitor the driver's stress condition, drowsiness, and fatigue [41].

2.4.5 Ultrasonic sensor

Ultrasonic refers to acoustic waves, where a transmitter sends sound waves, and a receiver captures the bounce off waves from nearby objects. The distance of such object is determined through the Time-of-Flight (ToF). These waves are propagated in a conical shape at the speed of sound (that depends on the density of the propagation medium), and use frequencies higher than those audible by the human ear, between 20 and 180 kHz [29]. The ultrasonic sensor is suitable for low speed, short-range applications (tens or hundreds of cm) like parking assistance, blind spot detection and lateral moving. With a low power consumption (up to 6 W), it is a relatively affordable object detection sensor.

2.4.6 Radar

Radar (Radio Detection and Ranging) detectors use reflected Electromagnetic (EM) waves. The device transmits radio wave pulses that bounce off the objects outside the vehicle. The reflected pulses which arrive some time later at the sensor allow inferring different information. Radar data is collected in a point cloud and provides abstract information about the surrounding objects, such as direction, distance, and estimate the object size [32]. The relative speed of moving targets can be calculated through frequency changes caused by the Doppler shift. Radar implements various techniques to modulate EM waves. Pulse Continuous Wave (CW) uses periodic pulse transmissions and silent periods for object detection [42]. However, pulse CW depends on wave energy and ambient noise, as well as lacks in timing marks, therefore it is unreliable for estimating range to target. To improve detection issues due to noise, Frequency-Modulated Continuous Wave (FMCW) emits a continuous signal, allowing its operating frequency to be changed during measurement.

Thus, instead of relying on time synchronization, the frequency differences between the transmitted and received signals are measured. FMCW considers both range and Doppler information to estimate the target's range and speed [42].

Automotive radars also use antenna diversity techniques such as Multiple Input Multiple Output (MIMO). In MIMO radars, each transmitting antenna can radiate an arbitrary waveform independently of other antennas, and any receiving antenna can receive this signal. Thus, a radar with M_{Rx} and M_{Tx} elements generates a virtual antenna array $M_{Rx} \times M_{Tx}$. Thus, the Field of View (FoV) is improved through their angular resolution (higher azimuth and sensor elevation), in addition to their target detection capabilities [42].

Digital Code Modulation (DCM) radars are also being developed. DCM allows each transmitter to be identified by unique codes. Digital radars aim to improve angular resolution, as well as minimize interference through coding, improving the identification of located targets, detecting their proximity from others through High-Contrast Resolution (HCR) [43].

Radar sensors are used for short and long-range detection at both vehicle front-facing and corners. Short-Range Radar (SRR) systems are employed to monitor environments close to the vehicle that require dealing with complex passive and active safety concerning single or multiple targets. SRR works in the 24 GHz mmWave frequency band. Meanwhile, Long-Range Radars (LRR) are widely used in Adaptive Cruise Control (ACC) systems to monitor the distance to vehicles ahead and control the ego-vehicle speed. LRR operates in the 70, 77 and 79 GHz mmWave frequency band. Unlike SSRs, FoV in LRR reaches longer distances, but azimuth and elevation angles are narrower [42]. Generally, with a fixed number of antennas, a radar with a broad FoV can be obtained at the expense of less angular resolution; on the other hand, a narrow FoV can provide better angular resolution, with opening angles between 9° and 150°, and elevation up to 30° respectively. Radar can operate in distance ranges up to 250 m, with power consumption from 12 W. Typically, radar is used for short, mid, and long-range object detection and adaptive cruise control at high speeds. Figure 2.3 shows the coverage range of an automotive radar sensor and its typical vehicle location. Radars are robust in adverse climatic conditions (e.g., fog or rain) and with scarce or no lighting. Nevertheless, signal processing is harder for classification issues if not combined with other sensor readings.

2.4.7 LiDAR

LiDAR (Light Detection And Ranging) uses laser reflection instead of radio waves. The LiDAR sensor transmits light pulses to identify objects around the vehicle [32]. Typically, LiDARs operating at wavelengths from 850 nm to 940 nm

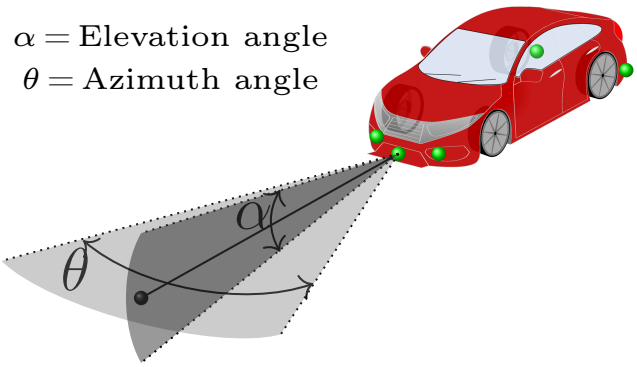


Figure 2.3: Schematic of a typical radar system based on FoV and scanning.

use ToF techniques based on pulsed and Amplitude-Modulated Continuous Wave (AMCW). Pulsed ToF calculates the target distance based on the round-trip time between the transmitted and received photons which bounce off the objects. Different from pulsed ToF, AMCW encodes an intensity pattern in the transmitted light beam, forming a linear radio frequency chirp; the target distance is calculated based on the amplitude of the bounced signal and the phase shift of the chirp sent. Both Pulsed ToF and AMCW are limited by daylight interference and their closeness to the visible light spectrum [44, 45].

On the other hand, LiDARs emitting at 1550 nm wavelength use FMCW. Different from AMCW-based LiDARs, FMCW LiDARs split the laser beam transmitted into a reference signal in a Local Oscillator (LO), and a phase modulated and chirped to the exterior. The light reflected from the target is mixed with the emitted light and compared with the reference signal in the LO. The frequency difference between the emitted and reflected light enables the target distance calculation, besides its speed and distance when it is in motion. FMCW LiDARs reach higher range and resolution in object detection, as well as range of operation in bright environments, with greater depth in dark scenarios [44].

The LiDAR sweeps in a circular and vertical fashion; the direction and distance of the reflected pulses are recorded as a vector of points, where each data point contains a reflectance value and the corresponding 3D coordinate related to the local coordinate system. Moreover, a set of vectors then constitutes a point cloud with spatial representation, enabling 3D model processing with high accuracy.

There are three types of LiDAR in terms of dimensionality (D). 1D LiDARs measure distance to a certain target or direction [29]. Meanwhile, 2D and 3D LiDAR sensors employ electromechanical or MEMS-type scanning methods to go further. 2D LiDAR sensors rotate the light beam in one plane, x or y, and detection occurs sequentially with equal time intervals between samples. Meanwhile, 3D LiDAR sensors operate the planes x, y, z, using axes as pivots to extend the dimensionality of the detected objects [44], i.e., to provide information on the position and distance

along axes, with a Field of View (FoV) extended for both Vertical (VFoV) and Horizontal (HFoV). For instance, LiDAR sensors can cover at 360° the HFoV around the vehicle, and up to 42° in the VFoV. Figure 2.4 shows the coverage range of an automotive LiDAR sensor and its typical vehicle location.

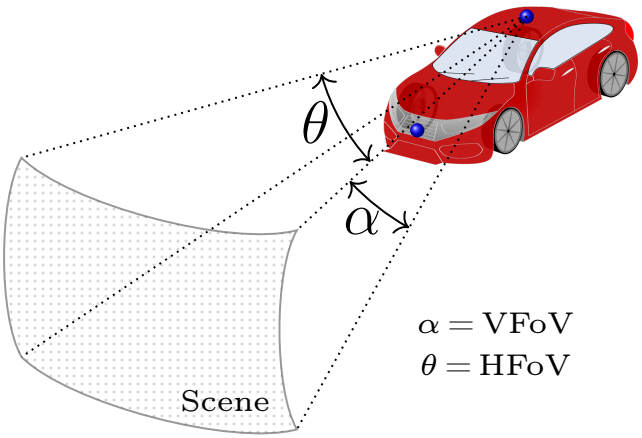


Figure 2.4: Schematic of a typical LiDAR imaging system based on FoV and scanning.

Currently, manufacturers are moving to solid-state LiDARs. Unlike electromechanical LiDARs, these devices use MEMS technology, which uses moving micromirrors to control the laser light beam and focus in a targeted manner. Thus, it is possible to reduce the physical infrastructure and price, but at the expense of reduced detection distance. Meanwhile, Optical Phased Array (OPA) is a MEMS-based LiDAR technique that adjusts the light beam in different directions without requiring mirror movement. Since OPA-based LiDAR does not use moving parts, it becomes more durable and cheaper. Finally, Flash LiDARs emit laser light pulses to detect the entire area around the sensor, similar to the capture imaging process in cameras. These sensors are much faster by not using moving parts or adjustments in the light beam; however, Flash LiDARs are sensitive to brightness by reflection, in addition to demanding a more powerful light beam to reach greater depth when covering an entire scene [44, 45].

2.4.8 Camera

Camera is a vision sensor used to record a visual representation of the surrounding environment. Thus, the camera can be used to detect objects on the road as well as to analyze the driver behavior and his environment inside the vehicle [32]. A frame from the camera is represented in a 2D array, containing the intensity of each pixel encoded in different forms, like HSV (Hue, Saturation, Value), RGB (Red, Green, Blue), or gray levels. Cameras can operate in the Visible (VIS) and Near-Infrared (NIR) spectral region [46]. VIS cameras are largely used because these

reproduce instantaneous images like those perceived by the human eye. Differently, NIR cameras detect objects based on heat radiation. Additionally, the image quality depends on the resolution, the Diagonal FoV (DFoV), HFoV and VFoV. Furthermore, vehicular applications use monocular cameras, stereo cameras, in addition to using so-called fish-eye lenses, which generate optical effects. Besides, Figure 2.5 shows the coverage range of an automotive camera sensor and its typical vehicle location.

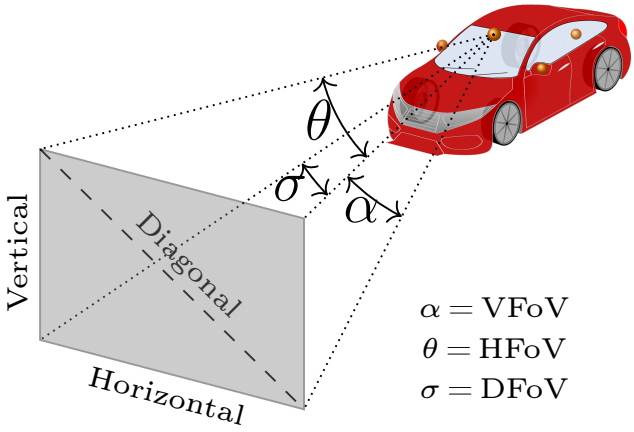


Figure 2.5: Schematic of a typical camera imaging system based on FoV and resolution.

Some active vision sensors operate synchronously while others, asynchronously. ToF cameras use infrared light to give depth to the 2D image, as well as allowing object scanning and measuring the distance to the target. Gated imaging systems use a laser light source synchronized with the camera shutter, and it is activated when the transmitted light beam reaches the object of interest. Meanwhile, event cameras respond to brightness changes in the scene asynchronously and independently for each pixel. Thus, event cameras define a sequence of events from the brightness intensity changes or motion in a scene at random time intervals.

Vehicular applications use CMOS-based cameras, also use monocular cameras, stereo cameras, in addition to using so-called fish-eye lenses [47, 48], which generate optical effects to reach sharpness and large depth of field. Additionally, the quality of the images depends on the resolution and FoV. Some drawbacks exist though: image quality depends on lighting and weather conditions, and scene representation is limited to the pointing direction and Line-of-Sight (LoS).

Given the limitations and advantages inherent to each sensor, single sensory data can be insufficient to make some decisions depending on the task to be executed. Therefore, it is necessary to implement strategies to fuse sensor data in order to process the volume and variety of data generated, in addition to allowing aggregation.

Table 2.1 summarizes the main features of each exteroceptive sensor. It is possi-

ble to observe that all sensors have limitations that can be minimized as the vehicle combines different types of sensors. It is worth noting that GNSS, magnetometer, microphone, and biometric sensors are currently used widely in vehicular telematics since they are present in widespread OTS devices, such as smartphones.

Table 2.1: Multidimensional comparative among exteroceptive sensors used in vehicle telematics.

Sensor	Main usage	Precision	Range	Advantage	Limitation
GNSS	Navigation, positioning	Medium/ High	n/a	High coverage, small form factor	Signal blocking in urban canyons
Magnetometer	Navigation, positioning, orientation	Medium	n/a	Small form factor, low energy consumption	Magnetic interference
Microphone	Surveillance, assistant, environmental sensing	n/a	150 m, omnidirectional	Small form factor, low energy consumption, direction of arrival	Environmental noise
Biometric	Health monitoring	High	n/a	Simple data processing	Uncomfortable
Ultrasonic	Environmental sensing	Low (cm)	150 cm	Small form factor	Low resolution
Radar	Environmental sensing	High	250 m	Robust in adverse climatic conditions and with scarce or absent illumination	Energy consumption, data processing for classification
LiDAR	Environmental sensing	High	200 m, omnidirectional	Low sensitive to light and to weather conditions, 3D representation	Data processing latency
Camera	Environmental sensing	Medium/ High	Line-of-Sight	Multiple techniques for data processing	Sensitive to light and weather conditions

Some exteroceptive sensors have similar characteristics to each other, generating data redundancy. For example, both radar and LiDAR have similar features for object detection, distance estimation, and do not depend on lighting conditions. Nonetheless, LiDAR data volume is higher than radar. These functionalities ensure the sufficiency of security-related data. On the other hand, sensor fusion data can provide solutions in a combined way to reduce computational and operating costs, besides complementing its functionalities to reduce deficiencies between sensors.

2.5 Remarks

It was observed that exteroceptive sensors are used both in-cabin and outdoor environments. Furthermore, these can be used in various safety services and applications that require immediate response times for both perception and reaction to some risk events on the road. An example of this is the driver/controller AI analysis before a vehicle interaction with another road user. It is important to note that multiple sensor readings measuring a specific variable (sensor fusion) can be combined

to improve the precision and accuracy of vehicle perception tasks [49]. Furthermore, sensor fusion reduces the limitations of sensors by complementing each other. Figure 2.6 shows how various exteroceptive sensors built into the vehicle and built into OTS devices can coexist in order to monitor various areas involved in intelligent vehicles safety. Multiple sensor perception systems can be merged to extract features or information from objects detected by exteroceptive sensors. Depending on the application, data from a single sensor may be sufficient to provide both external and in-vehicle monitoring systems.

Vehicular safety applications can be monitored in different ways. For example, as described in Section 1.1, driving behavior is one of the areas of greatest interest since one of the main risk factors on the roads is the human driving. For example, driving profiling can be described through GNSS traces, camera and LiDAR sensors can recognize driver's facial gestures, in addition to the maneuvers he performs with his hands. Moreover, facial analysis allows you to determine fatigue, distractions or drowsiness, and it is also possible to monitor the driver's health with sensors and devices that may be in contact with parts of the body [24]. Another factor that can affect vehicle maneuverability is the road state, so pothole detection is indispensable to understand vehicular safety from the point of view of infrastructure and the reliability of vehicle mechanics [24]. Collision detection applications can be performed through audio recognition by microphones. On the other hand, collision avoidance warnings can include environment sensing for object detection, trajectory analysis, lane departure and impact time estimation [46]. Furthermore, object detection can use data from multiple sensors to identify diverse objects physically, in addition to detect kinematic measurements (e.g., speed, acceleration, distance, among others) related to each specific object [49]. Sensor fusion enables, in addition to observing, perceiving and predicting future interactions [50]. In addition to object detection, vehicle trajectories traceability is important to determine decisions regarding the detection of objects in the vehicle's course, besides to behavior with traffic regulations and other road users [24].

2.6 Challenges

In principle, the analysis of risk events in traffic may benefit from the sensor readings embedded in OTS devices. Nevertheless, there are factors enabling the road safety metrics analysis related to vehicular telematics, in addition to collecting data for services and applications based on navigation, road monitoring, and vehicle safety. On the other hand, the data collection of exteroceptive sensors can improve the risk assessment analysis since it incorporates more characteristics of the objects detected around the self-driving vehicle, for example, speed, distance and location

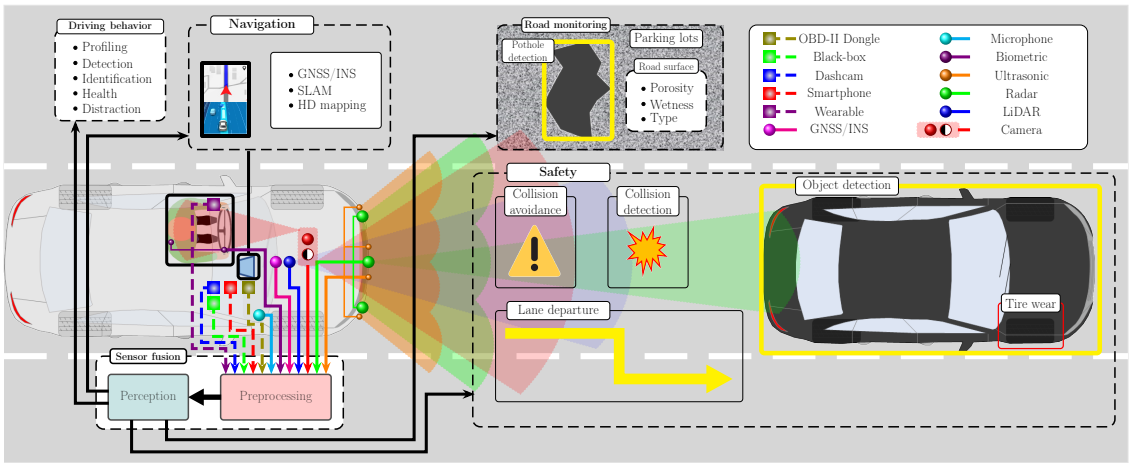


Figure 2.6: Illustration of an intelligent vehicle showing exteroceptive sensors and their applications.

measures from objects detected with respect to itself. Thus, it is possible to define policies to make precise and accurate safety estimations through risk assessment in self-driving vehicles. Some research challenges are open for exteroceptive sensing systems analysis in the context of risk assessment applications:

- Exteroceptive sensor readings analysis can provide detailed descriptions of objects in-motion around the vehicle (e.g., ground truth coordinates obtained from other objects, speed and time intervals related to other objects) with greater data granularity for risk indicators calculation, which determines the severity of traffic events.

- Large-area complex urban environment analysis where the vehicle is traveling can be extensive since the sensing in the vehicle is constant. It is relevant taking into account that each pedestrian, static or moving object, vehicle, road section, among others, can be analyzed as a logical unit of the road infrastructure. Based on this logic, it is possible to analyze the influence of evasive actions (e.g., braking actions of lane departure, harsh acceleration/deceleration, among others) on the AV, considering the actions and reactions related to the road infrastructure and road users.

- The mapping of risky objects, rare traffic events, or accident patterns in specific periods can be analyzed and standardized from the collection of sensing data. It depends on factors associated with the immediateness of the processing, storage, and transfer of data.

- Data volume generated by exteroceptive sensors can be high depending on the type of sensor. For example, the authors of [51] point out that a camera generates up to 40 MB/s, LiDAR up to 70 MB/s, while radar generates up

to 100 kB/s. In this context, it is worth mentioning that the estimated total sensor bandwidth can reach up from 1.4 TB/h to around 19 TB/h just for one AV, reaching up between 11 TB and 152 TB per day [20]. In the same way, the amount of data generated and processed by all types of exteroceptive sensors in the vehicle can be estimated. These data are relevant for the safety analysis in AVs, even more, when the data volume to be analyzed demands a high computational cost. Besides the fact that data redundancy can be noisy in the road safety analysis, sensor fusion data can optimize processes associated with perception and cognition, obtaining results according to the safety services priority. Therefore, it is necessary to assess the feasibility of using sensors with low data rates or to optimize the data analysis process to achieve instantaneous results with an appropriate volume of data for real-time safety services.

The characteristics of the various sensors available for vehicular telematics can facilitate risk assessment for policymakers. It is worth recognizing the functionality and sensor limitations since there are bottlenecks related to the sensor characteristics. As could be seen, there is no ideal sensor, and all are sensitive to environmental variations. Next, we describe datasets containing multi-modal sensors, i.e., proprioceptive and exteroceptive sensor readings. Moreover, these datasets contain information on different climatic conditions and day periods, tracking, location, mapping and semantic data related to the detected objects.

Chapter 3

Self-Driving Vehicle Datasets

In this chapter, we describe AV datasets available on the Internet. These datasets contain information from multi-modal sensor data readings. Moreover, datasets provide information about diversity of locations, weather and day period testing, tracking, localization, mapping and semantic data related to the detected objects. Some of these datasets include data processing to identify objects through bounding boxes, annotations, categories and attributes associated to the objects detected. As described in [24], these datasets contain information that can be analyzed for road safety analysis, e.g., speed, distance, acceleration, among others, besides of semantic data related to perception in the self-driving vehicles.

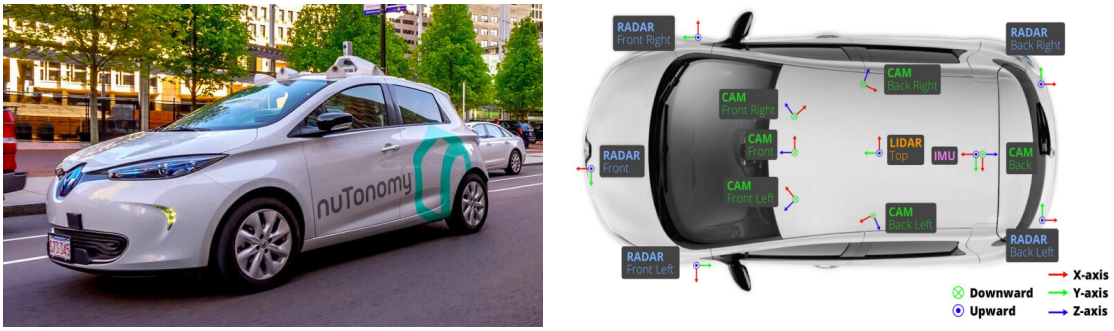
3.1 Datasets overview

Currently, vehicles are equipped with a wide variety of sensors, which integrate different systems to improve, adapt, and automate vehicle safety and the driving experience. Technological advances for driver assistance have paved the way for vehicle autonomy. These vehicles are equipped with exteroceptive sensors, as described in Chapter 2. The AV has the ability to interpret and identify objects, obstacles, traffic signs, among others. To process this data, it is possible to observe that vehicles with full automation have complex computational methods, which result in an exhaustive monitoring of the driving environment. Nowadays, there are a number of public datasets available from experimental autonomous vehicles [24], and these are used to analyze road safety metrics described in Chapter 4. In particular, datasets with semantic data are selected for analysis.

nuScenes: nuScenes [1] is a project by nuTonomy, a MIT start-up focused on developing software for self-driving cars and autonomous robots. Acquired by Delphi Automotive in 2017, the nuScenes project is now part of Motional[1], a company

[1]https://motional.com/

dedicated to the development of driverless technology. nuScenes is a public large-scale dataset of autonomous driving traces which includes images from camera, Point Clouds (PC) from LiDAR, and radar signals detected by the sensors installed on the vehicle. This dataset also provides data from the vehicle internal sensors (e.g., acceleration or speed). As for external sensing, 6 cameras, 1 LiDAR, 5 radars, and a GPS/IMU are deployed on the vehicle. The LiDAR covers 360° around the vehicle. Figures 3.1(a) and 3.1(b) illustrate the vehicle and the sensor setup for nuScenes. In total, the dataset includes almost 6 hours of data gathered by two Renault Zoe supermini electric cars with an identical sensor layout to drive, one in Boston (US), the other one in Singapore (SG). The internal sensing data is acquired from the CAN bus. Data from CAN bus contain navigation data, steering turn, changes in speed, acceleration, braking, rotation and travel records, associated with the location, as well as other vehicle-specific metrics, such as battery level, among others.



(a) nuScenes AV.

(b) nuScenes AV sensors.

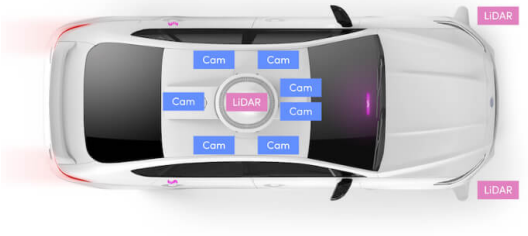
Figure 3.1: AVs and sensor setup for nuScenes AV [1].

Lyft5: Lyft is a carpool company founded in 2012. It started its journey with AVs in 2019, and is currently part of Toyota's Woven Planet[2], a mobility automation development company. Lyft5 [2] is another public large-scale dataset with AV traces, which contains images from cameras and LiDAR PCs. In particular, the Lyft5 vehicle is equipped with 7 cameras, 3 LiDARs, and a GPS/IMU sensor. The LiDAR covers 360° around the vehicle. Figures 3.2(a) and 3.2(b) illustrate the vehicle and the sensor distribution for Lyft5. The perception dataset consists of 2.5 hours of data gathered by twelve vehicles Ford Fusion equipped with autonomous controls in Palo Alto (US) divided into 180 scenes of 25 seconds each. Unlike the nuScenes AV dataset, Lyft5 does not provide CAN bus data from the vehicle.

The raw data from sensors is stored in point cloud data for LiDAR and Radar sensors, and image pixels from the camera. Figure 3.3(a) shows the format of the exteroceptive sensor data available in the dataset. The resolution of the images captured by the image sensors is 1600 × 900 in nuScenes, and 1224 × 1024 in Lyft5,

[2]https://level-5.global/

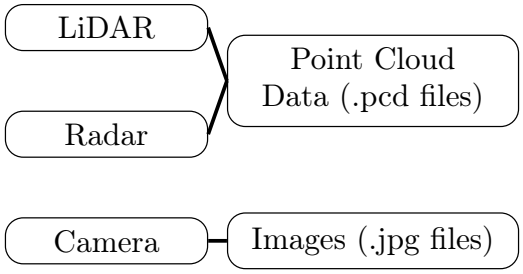
(a) Lyft5 AV.



(b) Lyft5 AV sensors.

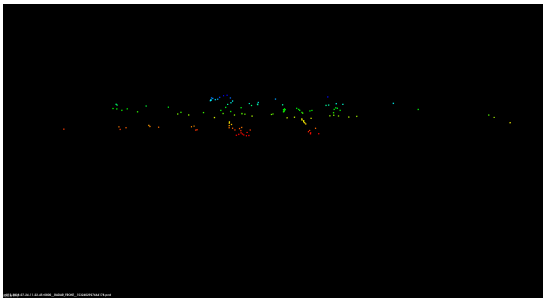
Figure 3.2: AVs and sensor setup for Lyft5 AV [2].

both compressed in JPEG. Figure 3.3(b) shows an example of the images available in the dataset. On the other hand, the distance and speed calculations from the detected objects are performed through the LiDAR sensor. It is worth mentioning that data generated by the radar sensor does not deliver accurate information about the object shapes, and it has no suitable information regarding the 3D location. However, radar PCs data are available for analysis in the dataset. Figures 3.3(c) and 3.3(d) show the point data cloud distributions for the radar sensor and LiDAR, respectively. As expected, the volume of points generated by LiDAR is higher than on radar, as discussed in Section 2.4.

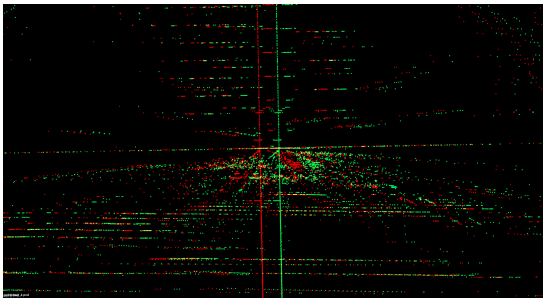


(a) Format data available from exteroceptive sensors readings.

(b) CAM image example.



(c) Point cloud data from radar sensor.



(d) Point cloud data from LiDAR sensor.

Figure 3.3: Representation of images and point cloud data from exteroceptive sensors available in the nuScenes AV Dataset. Lyft5 dataset does not include data from radar sensor.

Both nuScenes and Lyft5 datasets include sweep data from each sensor, based on the sampling frequency of each one. Metadata of all samples are available in JSON nested dictionaries. Moreover, the datasets provide semantic data, that is, data with sample annotations used to describe diverse characteristics of the object itself around the ego-vehicle, based on LiDAR PCs and JPEG images from the cameras. Figure 3.4 shows the data collection features available in nuScenes and Lyft5 datasets[3]. Data is based on images from camera, LiDAR PCs, and radar signals detected by the sensors installed in the vehicles. Table 3.1 describes the meaning of each subset available in the datasets nuScenes and Lyft5.

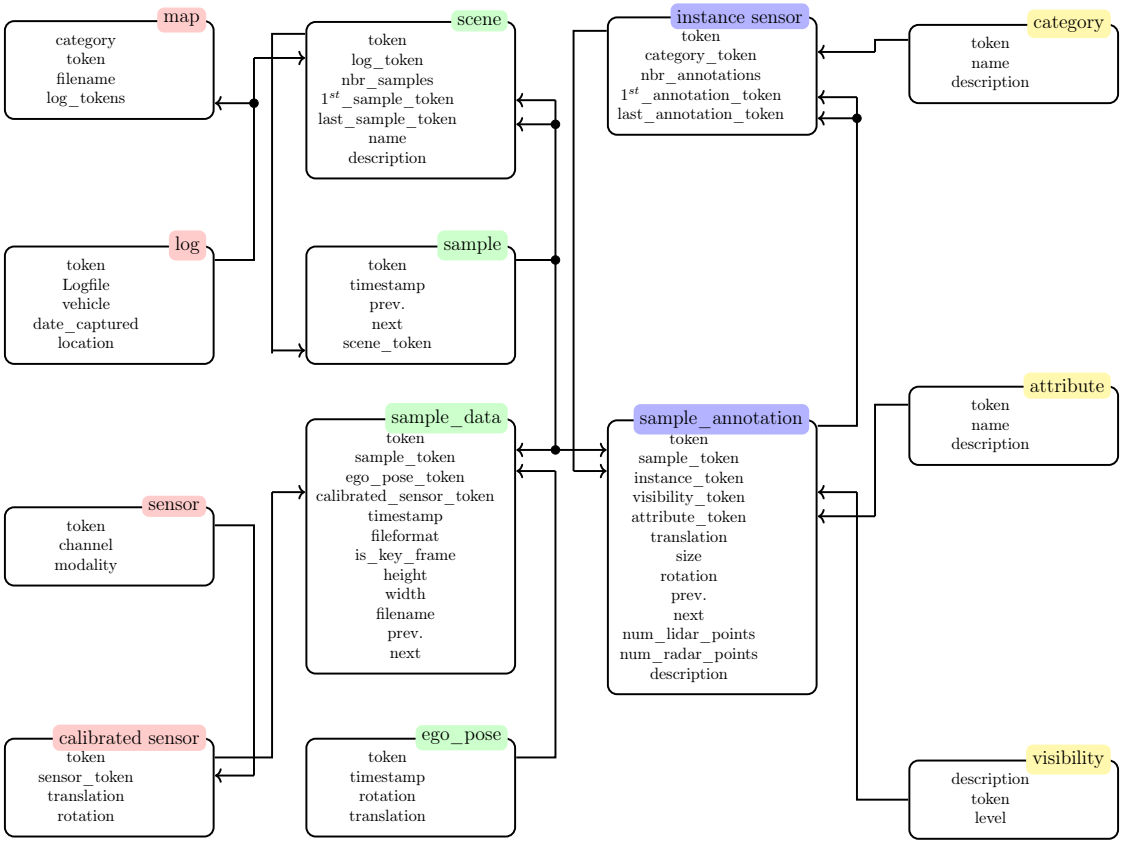


Figure 3.4: Schema for nuScenes and Lyft5 datasets.

Table 3.2 summarizes the main characteristics of nuScenes and Lyft5 datasets. Another advantage of the nuScenes and Lyft5 AV datasets is semantic segmentation, a process used for perception systems to associate LiDAR PCs and pixels from the camera images with predefined classes. Since autonomous systems do not discern the meaning of any detected object by exteroceptive sensors, semantic data analysis enables the recognition of detected objects as they are perceived by human. Semantic data facilitates the explainability processes required for policymakers and stakeholders. Moreover, semantic data is defined by raw data processing perception systems that use sensory systems, expert annotators and software on-board

[3]https://nuscenes.org/nuscenes#data-format

Table 3.1: Characteristics of the nuScenes and Lyft5 AV subsets.

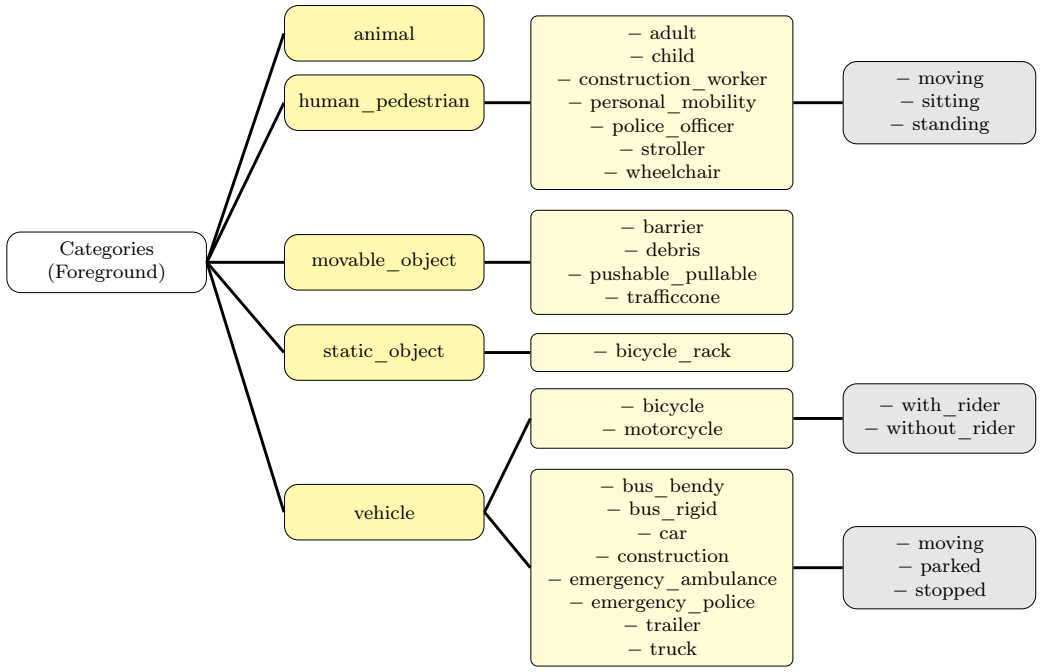
Subset	Description
sensor	A specific sensor type.
calibrated_sensor	Definition of specific sensor and calibration data (extrinsic and intrinsic parameters).
map	Data of binary semantic layers.
log	Information about the log from which the data was extracted.
scene	Description of the scenes (e.g., identifier, number of samples, first and last sample in the scene).
sample	An annotated keyframe at 2 Hz (nuScenes) and 5 Hz (Lyft5).
sample_data	A sensor data registration and its characteristics.
ego_pose	Description of vehicle's pose used for localization analysis.
instance	An object and all its interactions across scenes.
sample_annotation	Description of an instance seen in a sample through bounding boxes specifications (e.g., localization, size, orientation, among others).
category	Taxonomy of object categories and subcategories.
attribute	A property of an instance that can change while the category remains the same.
visibility	The visibility of an instance.

to perform multiple behavioral observations and interactions from different objects around the ego-vehicle, i.e., infrastructure and road users [52]. Each detected object is described as an instance, and each object can have multiple interactions with the AV in a sample. Each instance is marked with a 3D bounding box, category and attribute labels; each interaction of that instance with the AV is recorded as an annotation. Examples of categories are vehicle types, two-wheelers, pedestrians, road infrastructure, among others, and examples of attributes are vehicles or pedestrians stopped, in motion, among others. Figure 3.5 shows the categories available in the dataset. 23 categories and 8 attributes are defined in the nuScenes dataset, as shown in Figure 3.5(a). On the other hand, Lyft5 contains fewer scenes, but the proportion of 3D bounding boxes annotations is similar to that of nuScenes AV dataset. Thus, Lyft5 defines 9 categories and 18 attributes, as shown in Figure 3.5(b).

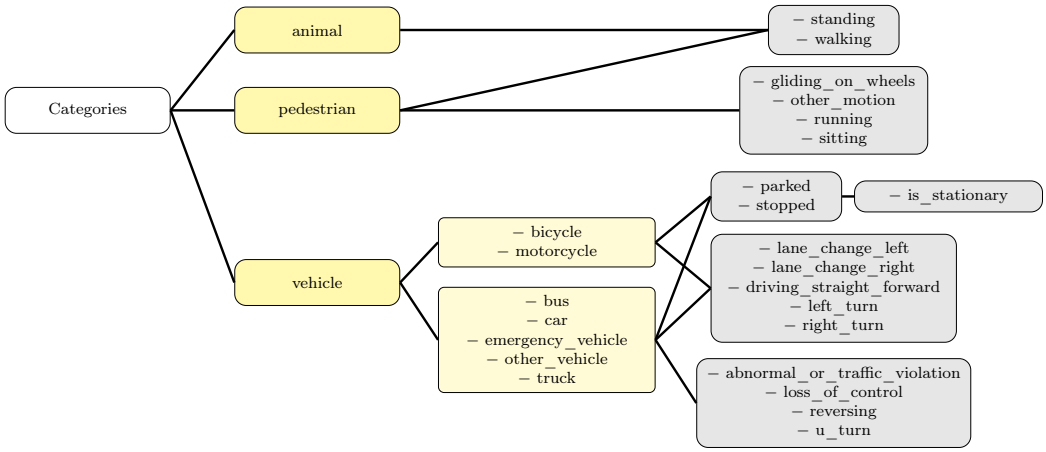
Table 3.2: Characteristics of the nuScenes and Lyft5 AV datasets.

	Scenes	Vehicles	Images	LiDAR PCs	Radar PCs	Bounding Boxes	Day/ Night	Weather	Categories/ Attributes
nuScenes	850	2	1.4 M	400 k	1.3 M	1.4 M	Yes	Yes	23/8
Lyft5	180	12	323 k	46 k	0	1.3 M	No	No	9/18

For both object detection and tracking tasks, nuScenes and Lyft5 datasets provide LiDAR point cloud (PC) data as global reference coordinate system. PC data related to all detected objects was collected using the LiDAR sensor as reference



(a) nuScenes AV dataset.



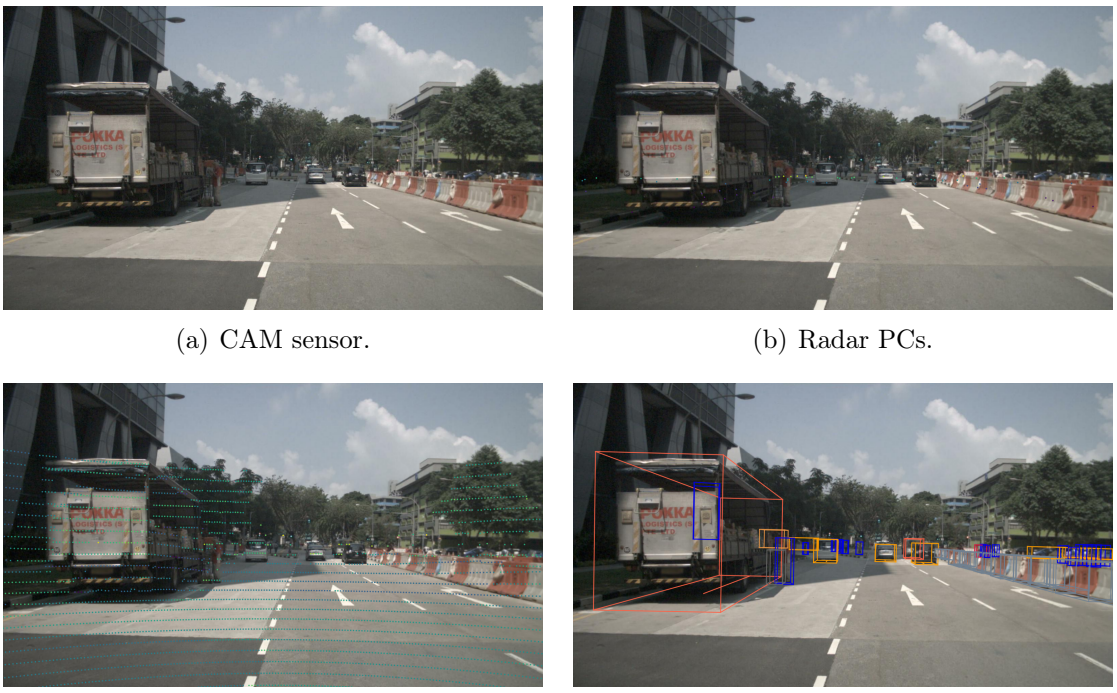
(b) Lyft5 AV dataset.

Figure 3.5: Categories and attributes available in nuScenes and Lyft5 datasets. Categories are represented as yellow boxes, attributes as gray boxes.

system, and therefore, the vehicle's reference coordinate. It is worth noting that individual sensors, calibration and how the orientation and position of the sensors is defined, can greatly affect the motion analysis results. Nonetheless, data labeling depends on sensor fusion of camera images and LiDAR PC data. It provides visual context for labeling, enabling to identify road users and infrastructure, and adjusting 3D point cloud data detection and 2D images projections. Therefore, sensor fusion allows the motion analysis for each detected object, enabling the calculation of vehicle dynamics as distance and speed variables from each sample annotation, among others.

Figure 3.6 illustrates the process of object annotation in the datasets: data an-

notations occur through keyframes based on the sample of image sensor, as shown in Figure 3.6(a), all the annotated objects in the dataset are visible, and they are covered by at least one point cloud returning from LiDAR or Radar sensors, as shown in Figures 3.6(b) and 3.6(c). Thus, image annotation boundaries are defined through 3D bounding boxes annotations technique, and semantic data by expert annotators [1]. Figure 3.6(d) shows how infrastructure and road users are distinguished through bounding boxes with specific colors for each category. AV developers use bounding boxes as a boundary to describe the spatial location of an object, specifically a cuboid. Thus, it is possible to recognize all the objects of interest in the image, but also to estimate their positions, sizes, and kinematic measures.



(a) CAM sensor.

(b) Radar PCs.

(c) LiDAR PCs.

(d) Bounding boxes for each annotation.

Figure 3.6: Exteroceptive sensor detection and bounding boxes.

Besides exteroceptive sensor data readings, the nuScenes AV dataset includes data from proprioceptive sensors acquired through the CAN bus for each scene. CAN bus data includes pedal actions, steering motion, vehicle speed, throttle, braking, acceleration (longitudinal, transversal, linear), IMU coordinates, rotation, translation, and other vehicle-specific metrics like wheels speed, battery level, among others. Figure 3.7 shows in detail the metadata available from CAN bus messages. The proprioceptive sensing data are synchronized with the exteroceptive sensing data, making it possible to analyze the data jointly. Table 3.3 shows the main characteristics for each subset available in nuScenes CAN bus dataset.

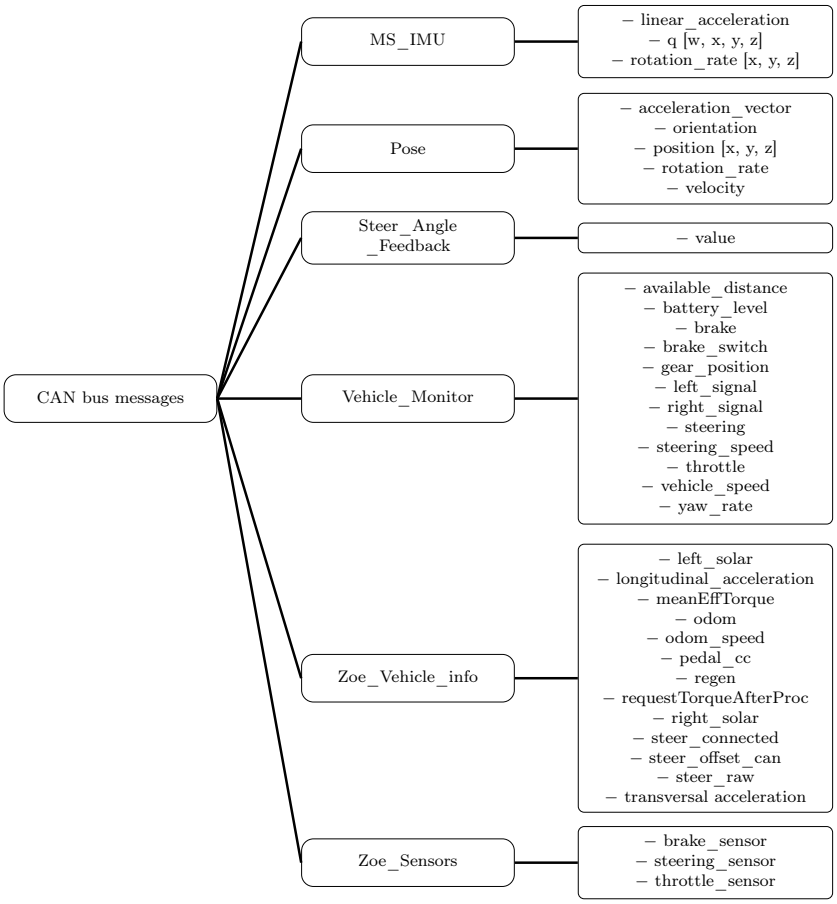


Figure 3.7: CAN bus messages schema in the nuScenes AV dataset.

Table 3.3: Characteristics of the nuScenes CAN bus subset.

Subset	**Description**
IMU	Data from MEMS sensors with linear acceleration and rotation based on IMU coordinate frame, and a quaternion matrix for transformation to fixed reference frame.
pose	Information about acceleration, orientation, rotation and velocity based on ego vehicle frame, position based on global frame.
steer_angle_feedback	Steering turn direction.
vehicle_monitor	Information about diverse sensor into the vehicle.
zoe_sensors	Data from brake, steering and throttle sensors defined by thresholds.
zoe_vehicle_sensors	Data from motion and electric sensors.

3.2 Preliminary Analysis

A preliminary analysis of the data was carried out to understand the scenario in which the vehicle travels, its conditions, and the vehicle's interactions with other road users and infrastructure. For data analysis, Python programming platform is used. The *nuScenes Devkit* [53] tool is used to semantic data analysis, location, and data filters. The *Pandas* and *Multiprocessing* libraries are used for data decom-

pression, parallel data analysis, and data filtering. The *NumPy* library is used for diverse calculations, and the *Matplotlib* library is used to plot the results.

Data analysis requires data from functional areas of the AV related to the autonomous driving system to analyze potential traffic risk events, as shown in Figure 4.7. Based on [3], four functional areas are related to: data *Acquisition* through sensors available in the AV; *Perception* and *Cognition* to processing the data collection into an understanding of the environment around the vehicle, describing vehicle surroundings and tasks scheduled to be executed based on regulations; and finally, the *Action* block defines the decisions to be taken and the devices available to the execution of AV tasks. Figure 3.8 shows the functional blocks for the AV performance analysis.

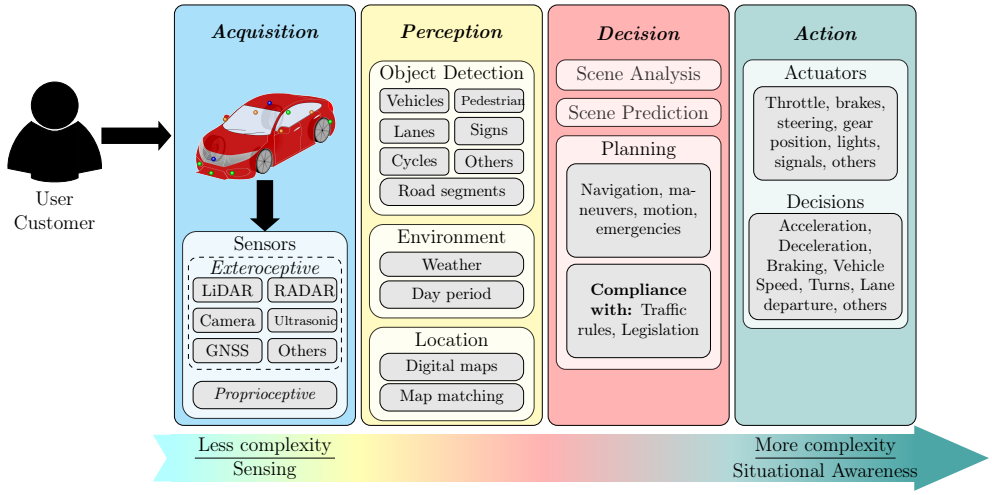


Figure 3.8: Architecture adapted from [3] for the nuScenes and Lyft5 data analysis.

Data analysis uses batch process in offline mode, i.e., analyzed data will be stored in a database to enable historical data analysis. The idea is to improve the classification of mathematical models. The data analysis is divided into three modules: data collection, data extraction and data analysis. The data collection includes data from all four functional areas of the AV. The data extraction module deals with the data collection provided in the datasets. This module handles the raw data contained in JSON files to format, discard outliers, and abnormal samples. The data extraction is heterogeneous, coming from different experimental AV datasets. Therefore, this module uses a json-to-csv converter using *Python*, suitable for the datasets to be analyzed.

Once the data is converted into a format for faster processing, such as CSV, the data analysis module deals with the data coming from the data extraction module. The data analysis module manages and analyzes the data to identify and classify characteristics of the ego-vehicle and the objects detected by the exteroceptive sensors. Furthermore, the data analysis module provides organized and "translated"

data (i.e., data organized according to the data type, its magnitude, when quantitative, or its category and attributes, when qualitative, among other characteristics). On the other hand, the data analysis module implements map matching to show the vehicle location on the map, adding information such as speed limits, road types, layers, addresses, among others, through an external location service server. For data filtering and variable classification, the analysis module uses the *Pandas* and *Multiprocessing* libraries, respectively. Figure 3.9 shows the architecture proposed for the data analysis process.

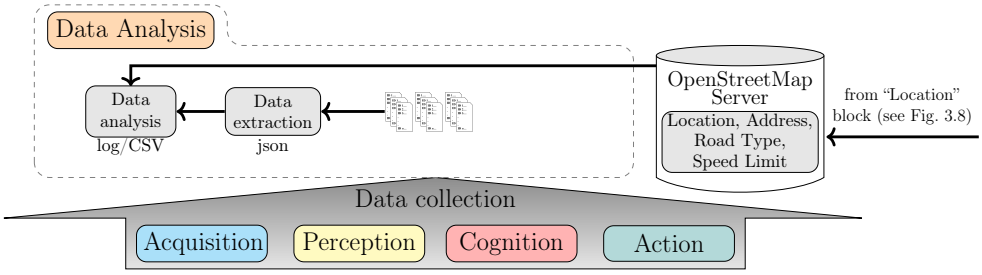


Figure 3.9: Architecture proposed for the data analysis.

An important feature of the data analysis module is the semantic data interpretation. Data interpretation is relevant to understand which objects the AV has interaction with, as well as relate the various events linked to the annotations available in the dataset. For both datasets, the annotations, categories and attributes associated with each instance are identified with syntax tokens. Figure 3.10 shows the logic of semantic data analysis. These data contain information related to the type of objects, attributes, and other classifiers available in the nuScenes and Lyft5 datasets. As can be seen, the perceived objects by the sensors are identified through a syntax token, with color markings related to the bounding box defined for each category defined in the dataset. Therefore, it is not always clear where the matching piece of information can be found, once a triggering token has been located. Therefore, the data analysis module deals with data treatment to execute the syntax token translation process. In the end, the output from the data analysis module will have readable data to provide the reports module and safety estimation analyzed in Chapter 4.

From data extraction and subsequent analysis, conditions that may interfere with the driving profile of AVs are analyzed quantitatively. Table 3.4 shows the observation statistics for the objects detected by the frontal camera. The highest amount of interactions happens with vehicles and traffic objects, while the least amount of interactions happens with animals.

Another factor that can interfere with the correct functioning of the AV is the weather conditions and luminosity. As reported in Section 2.4, sensors may have inherent limitations to climatic factors such as rain, fog, among others, and lu-

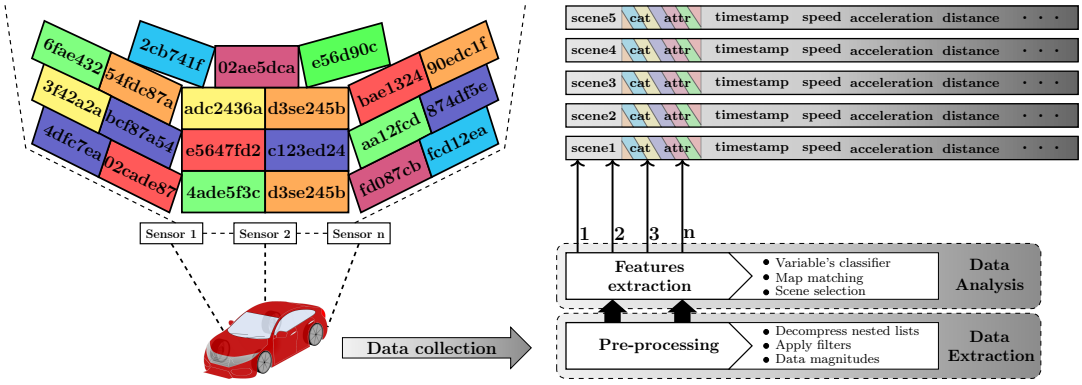


Figure 3.10: Data analysis for semantic data.

Table 3.4: Categories of observations in AV datasets.

Dataset	**nuScenes**		**Lyft5**
	Singapore	**Boston**	**Palo Alto**
Images analyzed	14,106	18,617	21,640
Instance annotations	11,308	21,251	10,525
Sample annotations	107,615	225,957	227,043
Vehicles	46,262	137,927	211,287
Two-wheelers	4,373	2,835	7,039
Pedestrians	25,915	36,221	8,672
Animals	36	121	45
Traffic objects	31,029	49,853	–

minosity such as clear, night, among others. Sequences with different climatic and lighting conditions are described. Table 3.5 shows the environment and the variables on which nuScenes and Lyft5 sensor readings were taken. Descriptions of weather conditions and time of day are provided in nuScenes AV dataset. Meanwhile, data from weather conditions and time of day in Lyft5 dataset was determined by video sequences.

Table 3.5: Weather and day period environment variables in AV datasets.

Environment	**Variable**	**Nr. Scenes**		
		Boston	**Singapore**	**Palo Alto**
day_period	Day	467	284	180
	Night	–	99	–
weather_conditions	Clear	318	367	180
	Rain	149	16	–

To analyze the environment variations, data from climatic conditions and day period is crossed to determine the different variations in which the AV coexist while operating. Table 3.6 shows the number of scenes for each possible crossover.

Another relevant factor in the AV safety analysis is the behavior adjusted to

Table 3.6: Crossed environment variables around the AV datasets.

day_period	weather_conditions	Nr. Scenes		
		Boston	**Singapore**	**Palo Alto**
day	clear	318	284	180
day	rain	149	–	–
night	clear	–	83	–
night	rain	–	16	–

traffic regulations. For that, to determine the road type, the OpenStreetMap Nominatim tool was used to determine the addresses where the vehicle traveled. Table 3.7 shows the interactions of the AVs on different types of roads.

Table 3.7: Vehicle interactions in the traces by road type.

Road type	Nr. Scenes		
	Boston	**Singapore**	**Palo Alto**
Motorway	194	–	–
Primary	2,792	1,037	15,445
Secondary	8,697	1,008	–
Tertiary	6,295	13,163	6,195

3.3 Remarks

The analysis of road safety metrics from data collected by experimental AVs can bring long-term benefits, since the data allows establishing the traceability of different events involved in the road infrastructure. Safety applications are interesting for car manufacturers, insurance companies, smart cities, as well as owners and passengers. Nonetheless, there are uncertainties about which parameters are to be used by the AVs to make complex decisions [11, 18]. Despite the fact that the autonomous system analyzes data through various machine learning techniques to make decisions, explaining the parameters that the AV adopts to decide an evasive action can be complex. Therefore, it is necessary to study the AV driving behavior through the analysis of data available in AV datasets. Nevertheless, there are some pros and cons derived from the needs of road safety analysis to establish AV driving behavior patterns in the presence of traffic conflicts. Open challenges which are crucial for the road safety analysis from AV datasets are:

- **Privacy:** Public datasets from experimental AVs must protect the third-parties privacy.
- **Data fusion:** Experimental AVs diverge on the type of sensing devices used in their vehicles. An example of this is the vehicle manufacturer Tesla, who

renounce to use LiDAR sensors, claiming that they are not reliable [54]. Other AVs do not include the radar in their configurations, and there is not a standard in the quantity of sensors to be used. Therefore, sensor data fusion approach is relevant to achieve better precision and reliability for road safety monitoring.

- **Categorization:** Experimental AV datasets provide detailed information on the objects detected around the vehicle. The availability of this data for safety analysis is relevant for monitoring risk events, in addition to defining which data is important. On the other hand, other strategies are useful to reduce the amount of data to be analyzed, from reducing the number of exteroceptive sensors, to degrading their precision or sampling frequency.

- **Sampling time:** Although experimental AV data is available, the currently available traces include only a few seconds of the vehicle experience, limiting monitoring for other eventualities that may arise while the AV is traveling.

- **Heterogeneity:** Data heterogeneity can lead to divergences in the analysis of diverse experimental AVs. Therefore, road safety analysis through SSMs becomes important to estimate how the vehicle performs while in-motion.

- **Driving limitations:** The trend towards data uniformity limits the analysis of traffic risk events, influenced by the limitations inherent to the AV, such as the top speed below the limit regulations, schedules, or paths available for testing.

In this chapter, experimental AV datasets were studied with the idea of extending their availability to road safety analysis. In particular, exteroceptive sensing enables kinematic measurements' analysis of objects around the vehicle. Furthermore, the vehicle data processing generates semantic data through geometric shapes to generate all the annotations related to each instance throughout the vehicle interaction. Together with this diagram, the specific object categorization and their attributes allow the risk assessment of traffic events and their severity related to road users. On the other hand, through the data processing of the sensor readings, it is possible to assess the SSMs described in Section 4, to determine some AV driving patterns in the presence of any traffic conflict. In the next chapter, nuScenes AV dataset is analyzed through SSM metrics to quantify the severity of all the events to which the AV is exposed while it is traveling.

Chapter 4

Risk Assessment Based on Surrogate Safety Measures

This chapter describes road safety metrics based on direct (i.e., ego-vehicle-specific measurements, e.g., speed variations, acceleration/deceleration rate, among others) and context-aware measures (i.e., estimated ego-vehicle measures relative to other road users or infrastructure, e.g., distance to other objects, time intervals, speed difference with others road users) to define the severity of any traffic risk event or traffic conflict. Table 4.1 lists key terms which are useful to understand road safety metrics analysis in this chapter.

Table 4.1: Glossary of terms related to road safety metrics.

Term	Definition
Ego-vehicle	Vehicle of interest in road safety analysis.
Object ahead	Any road user or infrastructure detected at front of the ego-vehicle.
Road users	Anyone who uses a road, such as a vehicle, pedestrian, cyclist or motorist.
Road infrastructure	Infrastructure related to the road transport, such as movable traffic objects, traffic signals, lanes, among others.
Collision course	A situation where the ego-vehicle and road users/infrastructure can collide in a projected area on the road plane.
Traffic risk event	A potential interactive event that represents a risk to the ego-vehicle and road users, and requires evasive action.
Conflict area	Common spatial area of projected trajectories given momentary measures of speed and distance for the ego-vehicle and one or more road users.
Traffic conflict	A traffic risk event in which the ego-vehicle and one or more road users can be involved, and a crash situation can be imminent if no evasive action is taken [55, 56].
Evasive action	Action taken by the ego-vehicle to avoid a traffic risk event, generally associated with braking, acceleration, deceleration, or swerving.
Severity	A parameter that defines the closeness of a potential traffic risk event in a quantitative way.
Serious event	Severity of an event involving the ego-vehicle, according to the traffic conflict technique threshold that quantifies a traffic risk event as “serious”.
Non-serious event	Severity of an event involving the ego-vehicle, according to the traffic conflict technique threshold that quantifies a traffic risk event as “non-serious”.
Validity	Data validity is a process that analyzes whether a measurement is appropriate to assess a traffic risk event. Validity process shows the precision degree to which an indicator describes a traffic risk event.

The scenario to be analyzed is illustrated in Figure 4.1. An ego-vehicle in motion monitored by an insurer can detect different road users along the way. When the ego-vehicle detects an object with a potential convergence point, estimation of safety based on risk level will be calculated only with those objects that are on a potential collision course. In the illustration, the red vehicle is defined as the ego-vehicle. We use the nuScenes and Lyft5 datasets as data source for the traffic risk analysis. It is important to note that we do not evaluate communication performance in the scenario shown. The illustration shows communication as the medium for transmitting data to the insurer.

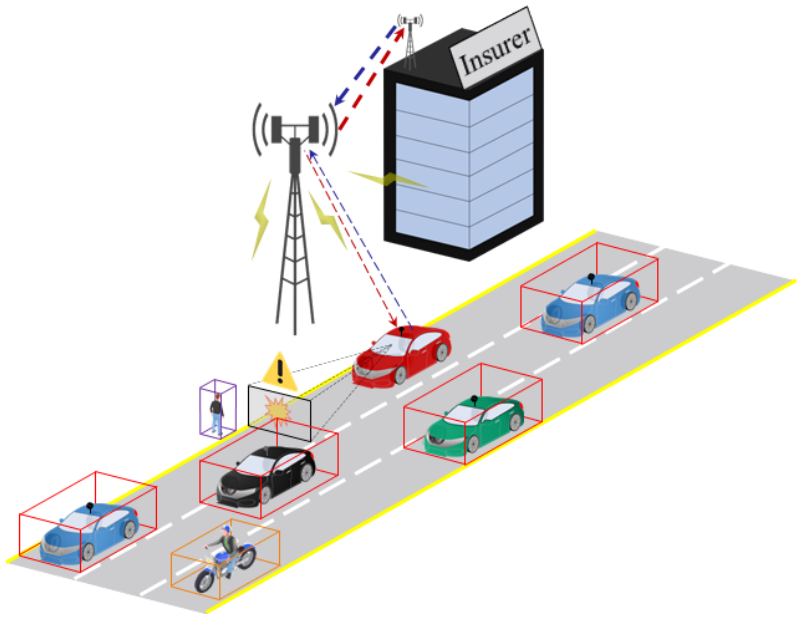


Figure 4.1: Traffic risk events analysis in self-driving vehicles.

4.1 Road Safety Metrics

Road safety metrics can be multifaceted. For example, the traceability of driving patterns, road monitoring, collision detection, among others, can be described through vehicle direct or context-aware measurements. Figure 4.2 illustrates the relationship between direct and context-aware metrics for road safety analysis in this work. The direct metrics are located in the left column; the central column contains the context-aware metrics, and the right column contains object ahead metrics, measured through exteroceptive sensors. Metrics related to the ego-vehicle and objects ahead are represented as gray-colored nodes and context-aware metrics as white-colored nodes. On the left-hand side, each direct metric is connected to context-aware metrics, since context-aware analysis depends on direct metrics; on the right-hand side, metrics acquired via exteroceptive sensors are connected to the

context-aware metrics. Sensor readings in telematics enable monitoring ego-vehicle specific measures, and secondary measures from OTS devices to examine driving patterns, e.g., vehicle tracking, location, driving behavior, as well as other monitoring services involved in the insurance area. However, the complexity of Human-Machine Transition (HMT) requires the characterization of profiles related to the autonomous system managing the self-driving vehicle.

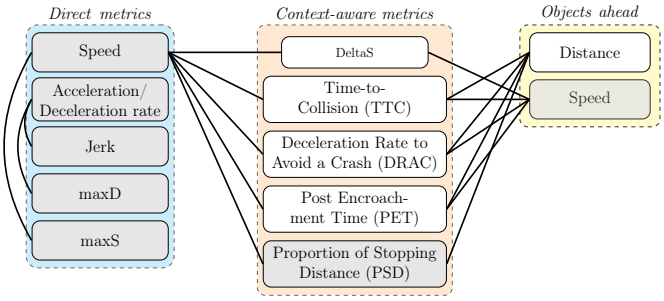


Figure 4.2: Relationship between the most widely used context-aware and direct metrics in SSMs.

Traffic risk events (e.g., jams, accidents, crashes) may take place anywhere at any time, including situations where collisions may or may not occur. These events can be represented as stochastic events, where diverse factors can be analyzed to explain the severity of traffic risk events. Therefore, ego-vehicle requires continuous safety analysis involving all interactions with other road users and infrastructure. From the traffic risk event analysis, it is possible to determine the situation threatening the vehicle safety, establishing a relationship with the severity of the event; subsequently, implement a risk assessment based on safety, guaranteeing a trade-off with the efficiency of driving, comfort, among others.

The severity of a traffic risk event is related to the probability that a vehicle is involved in a traffic conflict, an accident, the potential damages, and the consequences that it entails. The problem is how to determine which indicators reflect that severity, objectively. The relationship between the severity and frequency of conflict events, proposed by Hydén, is shown in Figure 4.3, adapted from [4]. It is important to note that any potential risk event occurs between at least two road infrastructure objects which interact on the road and can be explained by several factors involved in any event. Aiming to determine these indicators, traffic conflict techniques have been developed to show how to estimate a potential traffic risk event through the perception of situations involving road users and their risk probabilities. It is meaningful to note that the risk perception does not reflect the objective risky events, because the behavior is affected by variations corresponding to the individual perception of each vehicle. These behavior variations are more susceptible to occur when the responsibilities are from human drivers and pedestrians.

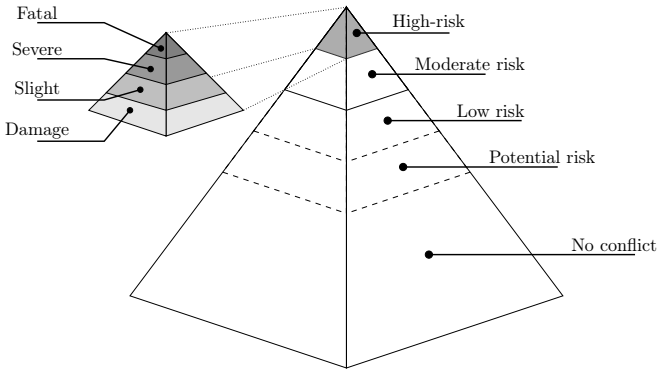


Figure 4.3: Pyramidal representation of traffic events (adapted from [4]).

The definition of a traffic conflict is based on temporal and spatial proximity factors [56]. Through the analysis of these factors, it is possible to estimate and argue the severity of risk events associated with the vehicle. However, traffic conflicts bring a series of limitations that derive from the nature of decisions in the presence of risk events. An example of this is the action/reaction time to minimize accidents, as well as evasive actions and their effectiveness. As a result, it is not possible to estimate the severity of a risk event. On the other hand, since there is no consensus on the reaction pattern in the presence of risk events, then an independent analysis is required for each vehicle to establish the corresponding responsibilities. Therefore, it is necessary to define the conflict measurement according to the nature of traffic risk events investigated. Figure 4.4 shows a distribution of conflict events to define critical events based on [5]. It is crucial to define appropriate risk thresholds to rank traffic conflict events in any scenario.

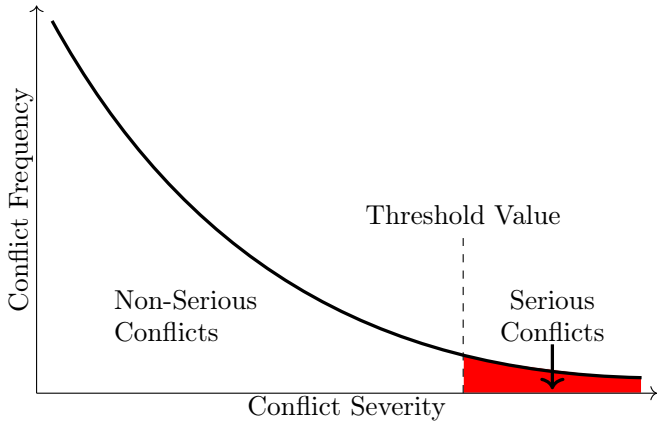


Figure 4.4: Distribution of traffic conflicts defining serious and non-serious conflicts [5].

In addition to direct and context-aware measurements related to the vehicle and the surrounding objects, there are other variables that can be used to model a traffic conflict. For example, we may have to consider geometric attributes (surfaces, align-

ment, lane dimensions), traffic conditions (traffic volume, speed limit, road types), environment (weather conditions, luminosity), and vehicle capacities (maneuverability, braking ability, stability) [57, 58]. In particular, vehicle capacities are improved by exteroceptive sensors enabling recognition of surroundings around the vehicle. Vehicle sensor readings can clarify traffic risk events, providing detailed descriptions of detected objects around. These sensing systems can incorporate features to improve the vehicle safety, besides to make it easier the road safety analysis (considering that sensor readings are more detailed). It should be noted that there are cases in which the vehicle's autonomy level still requires the driver's expertise and attention, and therefore, driver's engagement with the vehicle's situational awareness is crucial.

In the context of self-driving (autonomous) vehicles, through sensor readings, it is possible to determine both internal and external safety metrics. In addition to the analysis of good practices from the car manufacturer and AI developer, new safety requirements may emerge from the analysis and assessment of the driverless vehicle behavior, for example, sensor's liability, on-the-road distance training and experience time of the autonomous system controller, among others. A first instance to evaluate safety metrics in self-driving vehicles is the vehicle's performance related to traffic regulations. Traffic regulation violations can be registered both in data from the sensors and by external agents associated with the vehicular environment. In this sense, historical records can be managed to examine the vehicle's performance related to the surrounding interactions, besides allowing the analysis of the vehicle reliability in terms of autonomous system development. For example, Waymo simulate driving behavior in reconstructed fatal crashes based on real-world fatal collision scenarios that occurred in a specific operational design domain [59].

Another dimension of the safety metrics analysis for self-driving vehicles is related to the ability to interact with the different challenges that arise in a rapidly changing environment such as the urban traffic. In this sense, it is expected that the vehicle will respond appropriately to traffic risk events and that they are minimized. In this aspect, multiple sensing strategies are often used simultaneously so that the AV can be aware of any eventuality. Different traffic risk events revolve around this AV situational awareness:

- **Imminent or near crash situation:** Take precautionary or emergency actions in collision situations.
- **Speeding:** Eventually, a driverless vehicle could exceed the speed limit in unmarked or secondary zones.
- **Harsh acceleration and harsh braking:** These are associated with eventual conditions that require immediate reaction.

- **Lane departure:** It requires the analysis of blind spots, as well as the telemetries' analysis of other vehicles close to it.
- **Time to Collision:** Multiple vehicle sensors can detect a potential object that could cause a crash; however, the driverless vehicle may not make a decision until confirmation from other sensors.

4.2 Surrogate Safety Measures

Surrogate Safety Measures (SSM) are measurements which describe the relationship between two road users in a traffic risk event to quantify the crash probability and/or the potential traffic conflict severity in a meaningful way [60]. Traffic conflicts may differ in the severity level, motivating the analysis of a relationship between severity and risk events frequency; the severity of an event depends on multiple factors. SSMs are useful in vehicular applications, since they satisfy conditions based on crash prediction through analysis of observable non-crash events. Therefore, it is possible to model the frequency of traffic conflict events through practical analysis methods [61]. Traffic conflicts and the severity in which they result can be analyzed as measures of road safety. Furthermore, this area becomes more important considering that the trend of vehicles is increasing levels of autonomy. Thus, the analysis of decision-making and actions/reactions arising from them contributes to the AV audit. In fact, road safety analysis adheres to the explainability and fairness processes concern to policymakers and stakeholders.

A good practice in SSM analysis is to verify the data validity and reliability [62]. SSMs are used for traffic risk event analysis. As such, once SSMs are *estimations* of road safety measures, they can be questioned with regards to ambiguity, unbiasedness, among others. Data validity is a process that analyzes whether a measurement is appropriate to assess a traffic risk event. It shows the precision degree to which an indicator describes a traffic risk event. It means that some metrics may not adequately estimate traffic risk events. Therefore, through data validity, it is possible to understand the chain of events preceding a traffic high-risk event on the road. For example, two vehicles at low speed indicate a high-risk coefficient related to the time-to-collision metric; however, the deceleration coefficient indicates that there is no conflict between the vehicles. In this case, the traffic risk event can be a false positive, since there is no concordance between the metrics used. Data reliability refers to the accuracy and consistency of the measurements. These characteristics are relevant in this work to evaluate the quality of the data collection provided by the AVs related in Chapter 3: they provide objective data through the sensor reading that may decrease their performance by the conditions in which they are operated.

Analysis of traffic conflicts can be studied through direct metrics analysis, performed by exteroceptive sensors, based on images or point cloud data from sensors available in the vehicle. On the other hand, context-aware analysis is conducted through traffic conflict observations, manually. SSMs can be categorized by their objectivity and subjectivity in a traffic conflict. Direct metrics show instantaneous values related to the vehicle motion, such as speed, acceleration, and distance. On the other hand, context-aware metrics depend on direct measures, critical threshold values, and measurements from the objects detected around the ego-vehicle.

4.2.1 Direct metrics

Direct metrics use instantaneous values associated with vehicle motion, such as speed, acceleration, and distance. Therefore, direct metrics are defined as vehicle-specific measures (e.g., speed variations, acceleration/deceleration rate, among others). These measurements are perceived by proprioceptive sensors or OTS devices when available in the ego-vehicle itself, and through exteroceptive sensors, when dealing with other road users.

4.2.2 Context-aware metrics

To identify potential traffic conflicts, there are diverse SSMs as DeltaS [63], Deceleration Rate to Avoid a Crash (DRAC) [64], Post Encroachment Time (PET) [65], and Proportion of Stopping Distance (PSD) [66]. Table 4.2 summarizes the advantages and disadvantages of the SSMs described in this section. We emphasize on Time-to-Collision (TTC) [67]. TTC is one of the most common methods to analyze and describe the severity of traffic risk events.

4.3 Time-to-Collision (TTC)

TTC is defined as the time it would take for the ego-vehicle to collide with an object ahead, if the current relative speed was maintained from the previous advance along the same path [67]. This is a continuous measure of safety that can be calculated at any moment as long as the ego-vehicle and the object are in a conflict area. Thus, TTC enables the collision course analysis for vehicles and predicts how is the vehicle's motion related to other users of the road infrastructure.

Equation 4.1 defines the TTC as the relation of the distance between the ego-vehicle and objects ahead ($d_{(ego,obj)}$) and the speed difference between both ego-vehicle (v_{ego}) and an object ahead (v_{obj}); for simplicity, we assume the object is another vehicle. Typically, the TTC value indicates the minimum time to collide, calculated continuously through the detection process of a potential traffic risk event.

Table 4.2: Advantages and disadvantages of SSM indicators (Adapted from [8]).

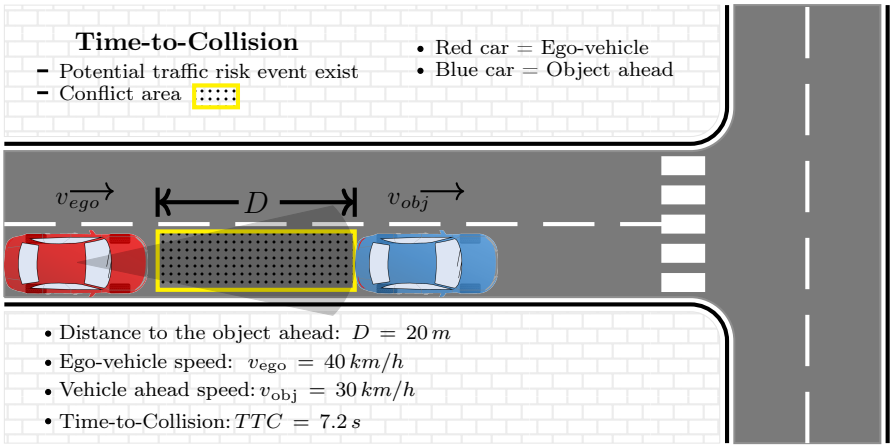
SSM	Advantages	Disadvantages
Direct metrics	Easy to apply: Applicable to objective measures. Variables independent from other values.	Limited metrics for predicting conflicts with other road users.
DeltaS	Reflect the relationship between both ego-vehicle/objects ahead through speed variations. Continuous monitoring of event risks. Applicable to determinate the severity of a crash.	Limited to objects around the vehicle. It is limited to analyze road infrastructure signaling.
TTC	Widely used as road safety metric. Continuous monitoring of event risks. More valid in context-aware measures. Available for dynamic object interactions. Applicable to define safety zones.	Ignores potential conflicts due to acceleration and deceleration actions. Analysis depends on higher speed from the ego-vehicle. It is analyzed just in a course collision. Measures a potential severity conflict, but not the nearness of a crash.
PET	Widely used in intersection conflicts. It is easily estimated. More valid in context-aware measures. Available for dynamic object interactions.	It is not enabled to define a severity conflict. It is not relevant to define the impact of a conflict.
DRAC	Continuous monitoring of acceleration/deceleration events in both ego-vehicle/objects ahead. More valid in context-aware measures. Available for dynamic object interactions.	Limited identification of a traffic conflict situation.
PSD	Available for vehicles interaction and time exposure conflict. It does not consider a response delay from the ego-vehicle.	Focus on specific traffic conflicts limited.

In the situation of imminent collision, TTC values assume finite decreasing values as the severity of the traffic risk event increases. It is worth noting that the TTC value allows inferring the amount of reaction time available for evasive maneuvers as a measurement of the risk level.

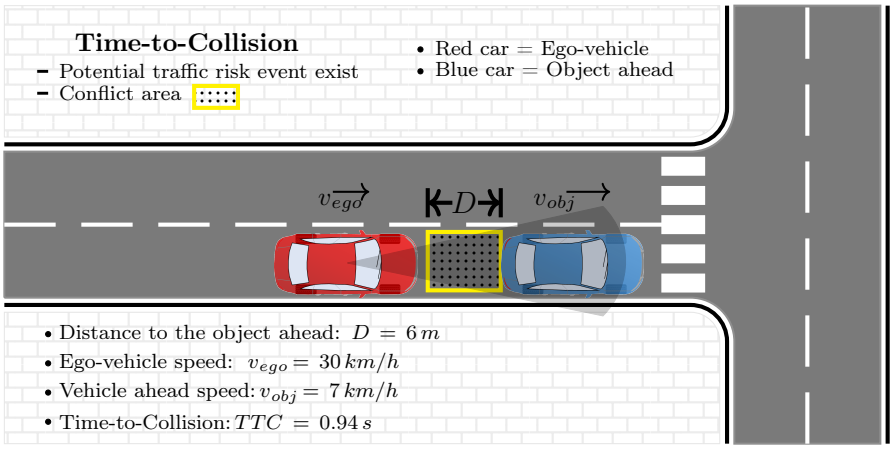
$$TTC = \begin{cases} \frac{d_{(ego,\,obj)}}{v_{ego} - v_{obj}}, & \text{if } v_{ego} > v_{obj} \\ \infty, & \text{otherwise} \end{cases} \tag{4.1}$$

Figure 4.5 shows an example of two vehicles in a conflict area. As soon as the ego-vehicle reaches a higher speed than the vehicle ahead, it is assumed that the pair of users are in a potential traffic event risk, with a higher risk probability. Figure 4.5(a) shows in the first stage that, despite detecting the blue vehicle ahead in-motion at a lower speed, the TTC value does not represent a risk event for the ego-vehicle. In the second stage, the TTC value represents a high-risk event for the ego-vehicle since the blue vehicle reduces its speed almost to stop, as shown in Figure 4.5(b). As a result, an increase in acceleration from the vehicle ahead reduces the risk probability, and the collision course stops existing. Indeed, if the vehicle ahead reaches a speed higher than the ego-vehicle, the TTC value tends to infinity.

The severity of any ego-vehicle risk event through the TTC calculation depends



(a) TTC event in a first stage.



(b) TTC event in a second stage.

Figure 4.5: Example of a TTC event analysis.

on speed and distance measurements. Therefore, at high and low speeds, it is possible to establish the severity of any event. A high TTC value represents a low traffic risk event level, while low TTC values indicate a high-risk event probability. Meanwhile, to quantify the severity of traffic risk events, a risk coefficient must be determined to distinguish the severity of the event.

4.3.1 Related work

Different safety indicators have been designed for risk assessment in traffic conflicts [8]. Indeed, these indicators are characterized by the fact that they allow to quantify the severity of traffic risk events. Additionally, it is possible to estimate the level of risk in scenarios where historical crash data is unavailable. This work focuses on the use of TTC as a technique to assess risk. Related works are briefly discussed in the following.

Approaches of TTC

TTC is a proximal temporal indicator widely used in techniques to analyze traffic conflicts. Due to TTC limitations (it ignores evasive actions, speed restrictions of the ego-vehicle direction related to the object ahead), several modifications have been proposed to improve the accuracy of this metric.

Modified Time-to-Collision (MTTC)

Modified Time-to-Collision (MTTC) [68] uses acceleration as a parameter to analyze the vehicle trajectory and its conflict discrepancies due to acceleration/deceleration. However, MTTC depends on both the acceleration of the following vehicle and the leading vehicle, the latter being difficult to measure or obtain, from the ego-vehicle. Furthermore, MTTC by itself does not allow the severity of potential risk events to be quantified, since various combinations of distance/velocity/acceleration may produce similar MTTC values. For this, the authors propose a Crash Index (CI) that uses kinematic variation factors to estimate the severity of risk events [68].The authors conclude that CI can effectively model the temporal distribution of accidents to the same extent as MTTC.

Enhanced Time-to-Collision (ETTC)

Another TTC variation is Enhanced Time-to-Collision (ETTC) [69]. ETTC assumes that following and leading vehicles do not change their courses until a collision occurs. Moreover, deceleration in the leading vehicle is considered until it stops. On the other hand, following vehicle's deceleration is considered to zero when the brake onset. Thus, ETTC calculation allows defining thresholds for "near" and "far" perception in Forward Collision Warning systems.

Time-to-Collision with Disturbance (TTCD)

Time-to-Collision with Disturbance (TTCD) analyzes collision risks caused by disturbances in the leading vehicles [70]. TTCD also can capture rear-end conflicts in car-following scenarios where the leading vehicle may have higher speed. TTCD considers the deceleration product of the disturbance, and the critical deceleration rate imposed by the leading vehicle deceleration.

Time Exposed TTC (TET) and Time Integrated TTC (TIT)

On the other hand, to determine safety evaluations based on TTC in time intervals, other indicators have been proposed to describe micro-levels of safe and safety-critical events derived from the TTC value analysis. The Time Exposed Time to

Collision (TET) is an indicator proposed in [71] which analyzes the time period that a vehicle remains exposed to high-risk events based on TTC values. These time periods analyze TTC measurements by thresholds defining the risk level. Thus, TET represents the duration of the exposition of safety-critical TTC values over a specified time duration. Thus, all the instants in which the driver is following the leading vehicle, which $0 < TTC < TTC*$ must be summed. Nonetheless, this indicator takes into account a single threshold TTC* (i.e., safety/safety-critical events), and therefore, it does not consider the variation between lower TTC values. To reduce the impact of low TTC values do not affect the TET indicator, Minderhoud *et al.* [71] propose the Time Integrated Time to Collision (TIT) metric which integrates the TTC to define the safety level for each TET interval analyzed in each driver's profile. Thus, TTC values below TTC* are also considered in the calculation process.

In addition to their improvements, TTC's variation metrics (MTTC, ETTC, TTCD, TET and TIT) also have disadvantages. These metrics are limited by the absence of motion analysis of the road users interacting with the ego-vehicle (e.g., evasive maneuvers, motion orientation, among others) when they are in a collision course. Table 4.3 shows a comparison of the approaches to improve TTC calculation.

Table 4.3: Summary of previous approaches using TTC.

Approach	Methodology	Advantages	Disadvantages
TTC [67]	Calculation based on constant speed.	Simple calculation based on distance and speed variations.	Ignores motion characteristics of the ego-vehicle and road users.
MTTC [68]	Calculation uses acceleration in TTC general formulation.	MTTC considers the acceleration of ego-vehicle and other vehicles during collision course.	Ignores motion characteristics of the ego-vehicle and road users.
ETTC [69]	Calculation uses deceleration behavior of the objects ahead.	ETTC considers characteristics of the objects ahead and their behavior when deceleration events occurs.	Ignores motion characteristics of the ego-vehicle and road users.
TTCD [70]	Calculation considers the effects of disturbing events in vehicles ahead of the ego-vehicle.	TTCD Analyzes reactions of the objects ahead that can affect the ego-vehicle.	Ignores motion characteristics of the ego-vehicle and road users. It is not clear how to apply the TTCD in diverse scenarios.
TET|TIT [71]	Calculation considers time duration and extension for the ego-vehicle drives in high-risky situations.	Measures consider time intervals for safety analysis.	Ignores variations occurred in TTC analysis.
TTC_{mo}	Calculation considers motion orientation of the objects ahead with respect to the ego-vehicle.	This metric considers motion orientation on the ego-vehicle's motion axis. Moreover, it considers just the objects ahead in collision course with the ego-vehicle.	Depends on accuracy from semantic segmentation classification and bounding boxes processing in the ego-vehicle.

SSMs based on motion dynamics

Some studies analyze unrestricted road users' motion as part of the dynamics in vehicular environments. Miller *et al.* [72] develop a collision warning system that analyzes traffic risk events and evasive actions, sharing the location and kinematic

measures from both ego-vehicle and vehicles around between peers. The algorithm analyzes the time to collision and the time to avoidance in a parametric way. Laureshyn *et al.* [6] propose a theoretical analysis of SSMs in collision course to determine the severity of traffic risk events. Given that interactions between road users are continuous, the authors suggest some strategies to calculate TTC for conflicts of different angles at constant speed. The authors stated that in potential collisions, a corner of one of the vehicles touches one side of the other vehicle. Thus, a new concept for TTC is developed, which calculates TTC between a moving line section of the ego-vehicle and a point in the other vehicle, in a time instant t. Next, the coordinates of the line section ending after t seconds based on a constant speed motion. Some assumptions about parallel motion are defined, depending on gradient of the line. Jiménez *et al.* [73] make an improved calculation of TTC in [72], assuming the vehicle geometry to be rectangular. In addition to the simplified calculation, the system analyzes the dimensions of the vehicles involved in the interaction, and the areas involved in a potential traffic conflict. Ward *et al.* [74] analyze the interactions between vehicles to define a prediction system and collision avoidance in vehicle-to-vehicle (V2V) communication systems. The method analyzes TTC for vehicles without motion restrictions. The authors calculate TTC in 2D, based on the relative vehicle motion and a looming method (a technique for gating predictions based on the relative motion of the vehicles), which considers the relationship of the vehicle roll angle, linear and angular velocity, and the yaw rate vector. Wachenfeld *et al.* [75] propose a Worst-Time-To-Collision (WTTC) metric to identify risk events related to the mobility dynamics of objects. The authors do a physical analysis of vehicle motion using the Kamm's circle (a theory about the transferable forces from the tire to the road surface) and entering the yaw angle.

Differently from these studies, this thesis analyzes the motion orientation of diverse road users that surround the ego-vehicle, detected through exteroceptive sensors, which enables the analysis not only with vehicles, but also with pedestrians and two-wheelers. Table 4.4 shows a comparison of the approaches involving motion orientation to improve TTC calculation.

SSMs based on naturalistic driving studies

Other works described in the literature analyze SSM metrics in data collection from naturalistic conduction studies, a method that is characterized by the continuous recording of driving information through the sensing of real traffic conditions. These testbeds use diverse exteroceptive sensors such as radars, cameras, GNSS, or V2X communication devices, to detect objects around the vehicle. Data sources, such as 100-Car [76] and SHRP2 [77] have been extensively studied via TTC to formulate safety metrics, analyze risk events, and compare simulated and real envi-

Table 4.4: Summary of works considering motion orientation in the TTC calculation.

Approach	Methodology	Advantages	Disadvantages
Miller *et al.* [72]	The authors propose a collision warning system based on calculation of intersection points.	The system includes an algorithm for intersection collision warning detection and considers communication strategies.	Ignores motion characteristics of the ego-vehicle and road users.
Laureshyn *et al.* [6]	Calculation of TTC based on convergence in different angles at constant speed.	The framework enables to calculate collision probability based on TTC in sideswipe conflicts.	Limited by disregarding motion characteristics of the ego-vehicle and road users.
Jimenez *et al.* [73]	The authors make an improved calculation of TTC based on methodology proposed in [72].	The framework considers vehicle geometry to be rectangular. The tool considers also the dimensions of vehicles involved in the conflict.	The framework is not tested on a real scenario.
Ward *et al.* [74]	The authors propose an indicator that generalizes TTC to the planar case, mapping vehicle trajectories on the road to predict traffic conflicts.	Planar analysis relies heavily on the relative positions of other traffic participants at the moment of predict the risk of a traffic conflict between vehicles. The model considers uncertainties by communication (V2V).	The model ignores other road users in the ego-vehicle vicinity.
Wachenfeld *et al.* [75]	The authors propose a method to reduce the amount of data to estimate the criticality of a conflict.	The method considers the motion orientation through yaw angles.	WTTC can define uncritical events as potential risky, e.g., vehicles travel side by side. WTTC does not consider other road users.
Our proposal (TTC_{mo})	Calculation considers motion orientation of the objects ahead with respect to the ego-vehicle.	This metric considers motion orientation on the ego-vehicle's motion axis. Moreover, it considers just the objects ahead in collision course with the ego-vehicle. TTC_{mo} also discards other objects out the ego-vehicle's path.	Depends on accuracy from semantic segmentation classification and bounding boxes processing in the ego-vehicle.

ronments [78, 79]. In the same way, Safety Pilot Model Deployment (SPMD) used around 3,000 human-driven vehicles, equipped with V2V communication devices and Mobileye sensing devices [80]. He *et al.* [81] evaluate SSMs from SPMD data. The authors implement three metrics: TTC, MTTC, and the Deceleration Rate to Avoid a Crash (DRAC). The authors observed that the MTTC presented the best overall performance. Xie *et al.* [70] propose an analysis of high-risk location identification based on TTC, using hypothetical disturbances in the leading vehicles, which can generate high-risk events in ego-vehicle. The work takes 75 highway segments where the SPMD vehicles circulated, and the information is cross-referenced with crash databases from Michigan City. The authors observed a high correlation coefficient between TTC, DRAC, and TTCD. Meanwhile, TTCD captures risks that were not observed in TTC and DRAC analyzes. Kusano *et al.* [82] develop a methodology to identify situations where the ego-vehicle driver generates an evasive braking action. The authors use radar data and kinematic measures from the ego-vehicle [76] to calculate the TTC as a metric to activate warning actions. Five car-following scenarios are identified to implement the algorithm: scenarios where the leading vehicle or lack of leading vehicle lack is correctly identified by the algorithm; scenarios where the leading vehicle is detected, but it is not in collision course with the following

vehicle; and scenarios where the algorithm failed to identify the leading vehicle or detects other objects different of the visual analysis. The authors conclude that the algorithm can identify 91.8% of the braking events when verified visually.

SSMs based on exteroceptive sensing

Studies on the evaluation of TTC through exteroceptive sensors have been developed to recognize the various entities with which a vehicle can interact. Aycard *et al.* [83] propose a risk assessment system at intersections. The authors use data fusion from camera and LiDAR sensors to detect and establish the dynamics of detected objects. For risk quantification, the TTC is used as a collision risk indicator. The authors conclude that risk assessment through environmental perception can enhance safety applications in the automotive industry. Kilicarslan and Zheng [84] analyze vehicle collisions through TTC using video cameras. The authors analyze the divergence of horizontal and vertical movement in video frames without relying on bounding boxes. To this aim, TTC analysis is based on the size variations of the detected object in the video, divided by the size changes in time intervals. The analysis of the algorithm proposed by the authors is used in videos of naturalistic driving without accidents. Results show 94% accuracy and 93% precision in the relationship between the computed system and the actual video. Meanwhile, compared to the detection of the LiDAR sensor in the KITTI dataset [85], the authors observe that LiDAR-based measurements depend on the depth of detection, discontinued detection, in addition to requiring 3D analysis. In this sense, video frame analysis is robust and can have a higher degree of accuracy.

The analysis of road safety metrics is closely related to the collection of image data from specific areas (mostly intersections), or video analysis in vehicles with embedded devices. Unlike these works, this study explores the potential of using data generated by AVs to develop road safety analysis solutions based on the vehicles' own sensing. Specifically, we focus on the TTC analysis with emphasis on the road users' motion orientation. Depending on the road users' orientation, TTC must be evaluated differently to accurately validate high-risk events involving the AV. This thesis analyzes TTC based on the road users' orientation and position related to the AV. For that, nuScenes AV dataset [1] and Lyft5 dataset [2] are used in this study to analyze the motion orientation and position of the detected objects by the AV while it is moving. The goal is to analyze the TTC based on the yaw angle of the detected object and its position with respect to the AV through data analysis from exteroceptive sensors' data readings in AVs. To the best of our knowledge, this is the first analysis considering orientation for the TTC calculation based on data from AVs.

4.4 Time-to-Collision with Motion Orientation

Traffic conflict analysis based on Surrogate Safety Measures (SSMs) helps to estimate the risk level of an ego-vehicle interacting with other road users. Nonetheless, risk assessment for driverless vehicles is still incipient, given that most of the AVs are currently prototypes and current SSMs do not directly apply to autonomous driving styles. Therefore, to assess and quantify the potential risk arising from self-driving vehicles interactions with other road users, this study introduces the Time-to-Collision with motion orientation (TTC_{mo}), a metric that considers the yaw angle of conflicting objects. In fact, the yaw angle represents the orientation of the other road users and objects detected by the AV sensors, enabling a better identification of potential risk events from changes in the motion orientation and position through the geometric analysis of the boundaries for each detected object.

An important factor for TTC_{mo} analysis is determining which road users and infrastructure are detected in the ego-vehicle path. For this, we use data from bounding boxes generated data processing into the AVs for objects detected by LiDAR and camera sensors. The goal is to evaluate when the ego-vehicle detects an object with a potential convergence point, an estimation of safety based on risk level will be calculated only with those objects that are on a potential collision course.

The probability of a risk event depends on the way an AV drives as well as on the other objects around interacting with the AV. Both have to be accounted for when considering the dynamics of the urban environment in terms of time and space. Therefore, it is necessary to analyze interactions with other road users. This is important whether we take into account that the AV may be on collision course with another user within an immediate time span. Figure 4.6 shows the logical relationship for determining whether an object detected in front of the AV represents a traffic risk event or not. In principle, any detected object by the vehicle sensors is continuously analyzed by the AV to check its motion orientation and position with respect to it. This is important considering that the mobility scenario changes quickly, and as such the possible interactions may change. Thus, we can detect whether a detected object is on a collision course with the AV. A detected object is not on a collision course when the object's direction course does not converge with the AV, i.e., the detected object moves on an adjacent track, or the trajectory of both the object and the AV diverge. Otherwise, an object in a driving course converging with the AV course can represent a potential risk.

4.4.1 Data Preparation

Analysis of safety assessment requires data from functional areas of the AV described in Section 3.2 to analyze potential traffic risk events, as shown in Figure 4.7.

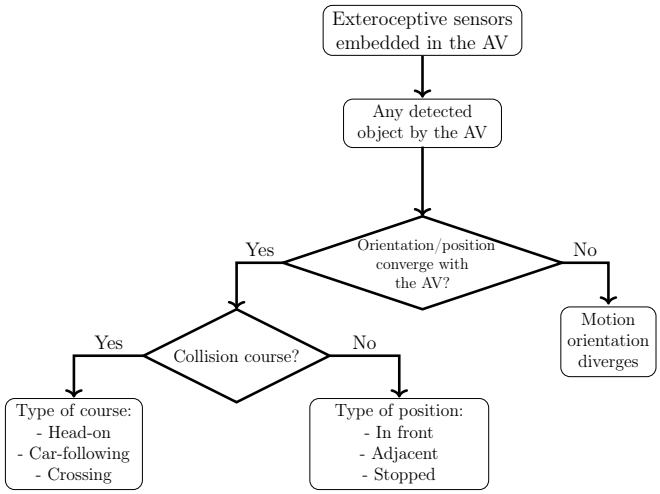


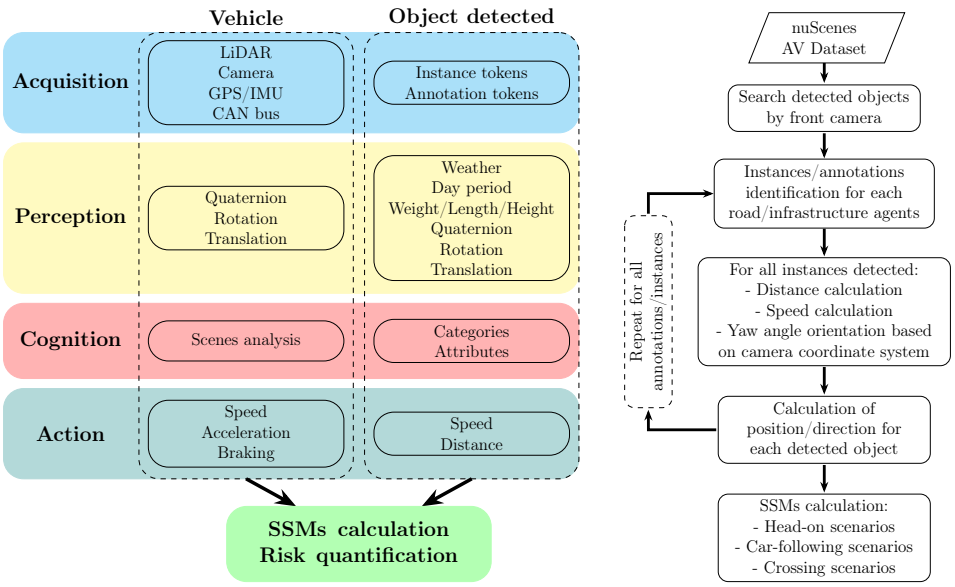
Figure 4.6: Relationship between object detection and motion orientation/position based on [6].

Figure 4.7(a) shows the data used for both AV and detected objects through the exteroceptive sensors, with respect to functional areas for the AV performance analysis related in Section 3.2. Thus, we aim to assess potential traffic risk events based on raw and semantic data from the AVs and their interactions with various road users and infrastructure. Categorized data make possible to assess safety with respect to road users different from vehicles, such as pedestrians and two-wheelers, among others when available. Moreover, data from all functional stages in AVs are used to assess risk events for the categories of detected objects in the dataset, in order to establish a standard of AV driving with respect to the road users' motion. Although traffic accidents are unexpected and rare events that can be associated with multiple causing factors, this analysis can help to explain more clearly potential traffic accidents since any collision describes a convergence approach between the users involved in the collision, as described in Figure 4.7(b).

Motion orientation and position angle

From the analysis of the dynamics of road users and the ego-vehicle, it is possible to evaluate metrics inherent to the objects' motion. For that, we use the nuScenes and Lyft5 devkits [53, 86], which provide a set of libraries to manipulate their datasets. We compute the bounding box orientation, swapping the global coordinate vector [*x, y, z*] of the LiDAR ([*1, 0, 0*]) for the camera ([*0, -1, 0*]), according to the coordinate frames defined for each sensor. In this way, we set the yaw angles for each object detected based on the global coordinates of the frontal camera, as observed in Figures 3.1 and 3.2.

Furthermore, the yaw rate describes the spatial orientation of the vehicle through the angle variation on the z-axis, as shown in Figure 4.8(a). Yaw angles (ψ) indicate



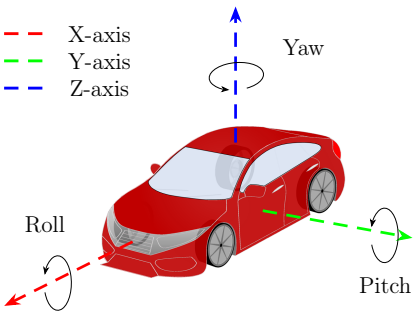
(a) Key variables used in the data analysis. (b) Data analysis procedure.

Figure 4.7: Summary of data used to analyze traffic risk events based on motion orientation.

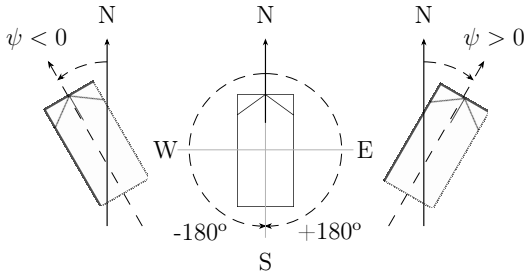
the orientation of each detected object. North is $0\,^{o}$ (ψ_0), east is $90\,^{o}$ (ψ_1), west is $-90\,^{o}$ (ψ_2) and south is $180\,^{o}$ (ψ_3). Objects have positive heading in clockwise direction and negative value in counterclockwise direction. About the ego-vehicle, we assume that yaw angle is $\psi_{ego} = 0\,^{o}$. Thus, detected objects with yaw angle between $\psi_2 < \psi_0 < \psi_1$ indicate that the direction on z-axis is forward to the ego-vehicle; meanwhile, yaw angles between $\psi_1 < \psi_3 < \psi_2$ indicate that the direction is opposite to the ego-vehicle, as shown in Figure 4.8(b). On the other hand, position angles (θ) indicate the location of an object with respect to the ego-vehicle. Position angles $0\,^{o} < \theta < 90\,^{o}$ indicate object locations at right-side with respect to the driving direction, while $-90\,^{o} < \theta < 0\,^{o}$ at left-side, as shown in Figure 4.8(c). Bounding box centroid coordinates (x, z) are used to determine θ. Thus, it is possible to establish when the ego-vehicle path is converging with detected objects. It is important to explainability requirements to assess traffic conflicts based on the road users' motion.

Geometric analysis of objects and ego-vehicle

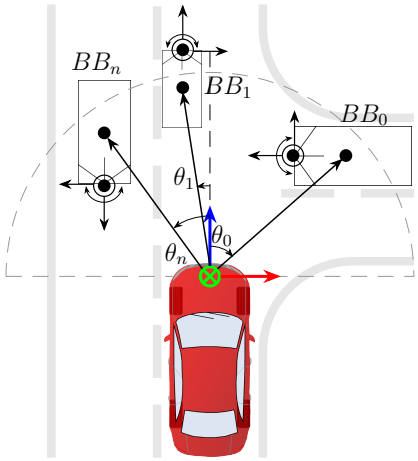
As shown in Figure 4.9, we use the ego-vehicle size specification to obtain a geometric representation and to analyze the interaction with surrounding objects. The width and length of the detected objects, available from the bounding boxes, are considered in the geometric analysis. In this sense, each vertex is labeled to determine its location and orientation when the object moves and rotates. The ego-vehicle is also represented as a bounding box. It is important to note that since $\psi_{ego} = 0$, the position of its vertices will always be the same for the analysis. In



(a) Roll, Pitch and Yaw rotations.



(b) Illustration of the yaw orientation in road users.



(c) Object position with respect to the ego-vehicle.

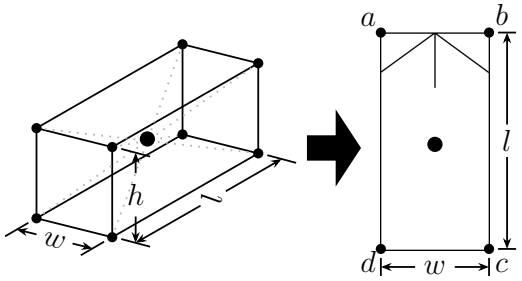
Figure 4.8: Relationship between AV and detected objects via motion orientation and position angle.

addition, the remaining space between the lane width and the ego-vehicle width is used as a safety area (sa_n), to identify objects adjacent to the AV that may represent potential traffic conflicts. The lane width is based on the respective road city regulations. Vehicle size specifications are reported in Table 4.5.

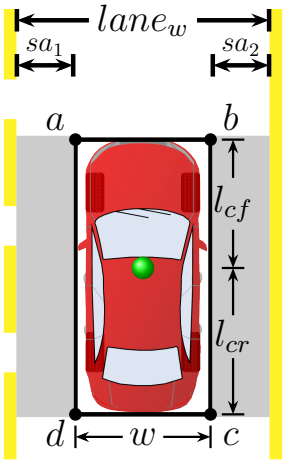
Table 4.5: nuScenes and Lyft5 vehicle overall dimensions. The width (w) includes external mirrors. The length between camera and vehicle front-side (l_{cf}) and the length between camera and vehicle rear-side (l_{cr}) are based on the camera location on the vehicle's rooftop.

Dimensions	**Vehicle**	
	nuScenes (Renault Zoe)	**Lyft5 (Ford Fusion)**
w [m]	1.945	2.121
l [m]	4.087	4.871
h [m]	1.562	1.478
l_{cf} [m]	1.81	2.302
l_{cr} [m]	2.277	2.569

To determine which objects are in collision course with the ego-vehicle, it is important to identify adjacent or overlapping trajectories between the ego-vehicle



(a) Bounding box geometry.



(b) Ego-vehicle geometry.

Figure 4.9: Geometric representation for an object (a) and the ego-vehicle (b).

and other objects. It is possible to determine objects trajectories via the motion analysis. In this way, we identify behavior indicators according to the AV reaction in several possible interactions with the objects around. Thus, we process AVs data to identify these interactions. In this analysis, we consider data annotations of the camera's coordinate system as reference. Information of the bounding box like the yaw rate (ψ), centroid position data in the image ($x,\ y,\ z$), and the size ($w,\ l,\ h$) are extracted from each annotation. Each vertex of a bounding box and the AVs are calculated by the relationship between sizes and the centroid coordinates, as described in Equation 4.2:

$$\begin{aligned} a_x, d_x &= x - \frac{w}{2}, \\ b_x, c_x &= x + \frac{w}{2}, \\ a_z, b_z &= z + \frac{l}{2}, \\ c_z, d_z &= z - \frac{l}{2}. \end{aligned} \tag{4.2}$$

Table 4.6 shows the vertices' calculation for both the ego-vehicle and bounding boxes. We also model the ego-vehicle as a bounding box to analyze the interaction

of each corner of it with the detected objects. Thus, we consider the position of the camera on the vehicle's rooftop as the origin x, z. It is important to note that the camera position does not correspond to the vehicle's centroid, and therefore, it is necessary to calculate l_{cf} and l_{cr}, as shown in Figure 4.9(b). Furthermore, we assume that $\psi = 0$ since we analyze the interactions with objects detected from images captured by the AV front camera.

Table 4.6: Relation between the centroid position in the bounding box and the ψ_{obj} rotation.

Vertices	**Bounding Box**		**nuScenes/Lyft5**	
	x_{obj}	z_{obj}	x_{ego}	z_{ego}
a	$a_{x_{obj}}cos(\psi_{obj}) + a_{z_{obj}}sin(\psi_{obj}) + x_{obj}$	$-a_{x_{obj}}sin(\psi_{obj}) + a_{z_{obj}}cos(\psi_{obj}) + z_{obj}$	$a_{x_{ego}}$	$z_{ego} + l_{cf}$
b	$b_{x_{obj}}cos(\psi_{obj}) + b_{z_{obj}}sin(\psi_{obj}) + x_{obj}$	$-b_{x_{obj}}sin(\psi_{obj}) + b_{z_{obj}}cos(\psi_{obj}) + z_{obj}$	$b_{x_{ego}}$	$z_{ego} + l_{cf}$
c	$c_{x_{obj}}cos(\psi_{obj}) + c_{z_{obj}}sin(\psi_{obj}) + x_{obj}$	$-c_{x_{obj}}sin(\psi_{obj}) + c_{z_{obj}}cos(\psi_{obj}) + z_{obj}$	$c_{x_{ego}}$	$z_{ego} - l_{cr}$
d	$d_{x_{obj}}cos(\psi_{obj}) + d_{z_{obj}}sin(\psi_{obj}) + x_{obj}$	$-d_{x_{obj}}sin(\psi_{obj}) + d_{z_{obj}}cos(\psi_{obj}) + z_{obj}$	$d_{x_{ego}}$	$z_{ego} - l_{cr}$

Next, we analyze when an intersection exists between ego-vehicle vertices and bounding boxes converging to the AV path. For this, data from the detected object vertices and ego-vehicle vertices are analyzed to determine the interactions between them. For this analysis, a line segment is defined as the line connecting the adjacent vertices of the bounding box. We define line's equation for each selected bounding box segment of both object and the ego-vehicle and potential intersections are calculated, as shown in Equation 4.3:

$$\begin{aligned} A_{ego}x + B_{ego}z = C_{ego}, \\ A_{obj}x + B_{obj}z = C_{obj}, \end{aligned} \tag{4.3}$$

where A, B, and C correspond to the line's equation values for each segment of the bounding box (object) interacting with the ego-vehicle. These values are given by a set of conditions that depend on the detected object's orientation.

Once the line equations have been calculated, the resulting values are used to compute the intersection coordinates at x, z, as shown in Equation 4.4:

$$\begin{aligned} x_{ego \cap obj} = \frac{(B_{ego}C_{obj}) - (B_{obj}C_{ego})}{(A_{ego}B_{obj}) - (A_{obj}B_{ego})}, \\ z_{ego \cap obj} = \frac{(A_{obj}C_{ego}) - (A_{ego}C_{obj})}{(A_{ego}B_{obj}) - (A_{obj}B_{ego})}. \end{aligned} \tag{4.4}$$

Thus, the distance d is calculated between the potential conflict vertices and segments between the detected object and the ego-vehicle, as shown in Equation 4.5:

$$d = \sqrt{(x_{seg_{ego}} - x_{ego \cap obj})^2 + (z_{seg_{ego}} - z_{ego \cap obj})^2}. \tag{4.5}$$

Finally, it is possible to identify the location of objects around the AV. Nevertheless, in order to determine not only moving objects, as reported in [82], the goal is to define also when movable/static objects (e.g., vehicles parked, traffic signals, among others) can provoke AV evasive actions that may represent potential risk events immediately. Thus, this work aims to evaluate the interactions between vertices of both AV and movable/static objects. This is important considering that although the proximity of the AV to other objects is inherent in the vehicular environment (e.g., adjacent vehicles, crosswalks, crossing vehicles, among others), and therefore some risk events can result in false positives.

Through this analysis, it is possible to describe various interactions with surrounding objects detected by the AV. Nevertheless, it is necessary to quantify the risk when the AV is on a collision course. For that, this work uses the TTC considering the detected objects' orientation as a metric to improve the analysis of traffic risk events involving the AV. The goal is to propose an improved TTC and test it with real data from AVs.

4.5 $\mathbf{TTC_{mo}}$ Calculation from Camera Images

From the analysis of camera images, it is possible to determine the position of objects. We can derive both the absolute location of the object and its position in the image through projections from 2D camera frames. As shown in Figure 4.10, it is possible to analyze the mapping between the world coordinate system and camera coordinate system that corresponds to the coordinate system used for vehicle navigation. Also, the object's speed related to the ego-vehicle is calculated by measuring the time difference between the sending and rebounding laser pulses from the LiDAR sensor.

To reduce the shortcomings of SSMs proposed in the literature, we include the motion orientation and position of objects detected by the AV as a parameter for the TTC calculation. The goal is to improve the accuracy of TTC to assess risk events for AV. Equation 4.5 summarizes the computation of the proposed $\mathrm{TTC_{mo}}$:

$$TTC_{mo} = \begin{cases} \frac{d_{(obj,ego)}}{v_{ego} - v_{obj}\cos(\psi_{obj})}, \text{ if } (v_{ego} - v_{obj}\cos(\psi_{obj})) > 0, \\ \infty : \begin{cases} \text{if } (b_{ego} + sa_2) < (a,b,c,d)_{obj} < (a_{ego} - sa_1), \\ \text{or } (v_{ego} - v_{obj}\cos(\psi_{obj})) < 0, \\ \text{or } v_{ego} = 0, \end{cases} \end{cases}$$

where d is the distance between the segment/vertex on the ego-vehicle's course and the front-side of the ego-vehicle, v_{ego} is the speed of the ego-vehicle, v_{obj} is the speed

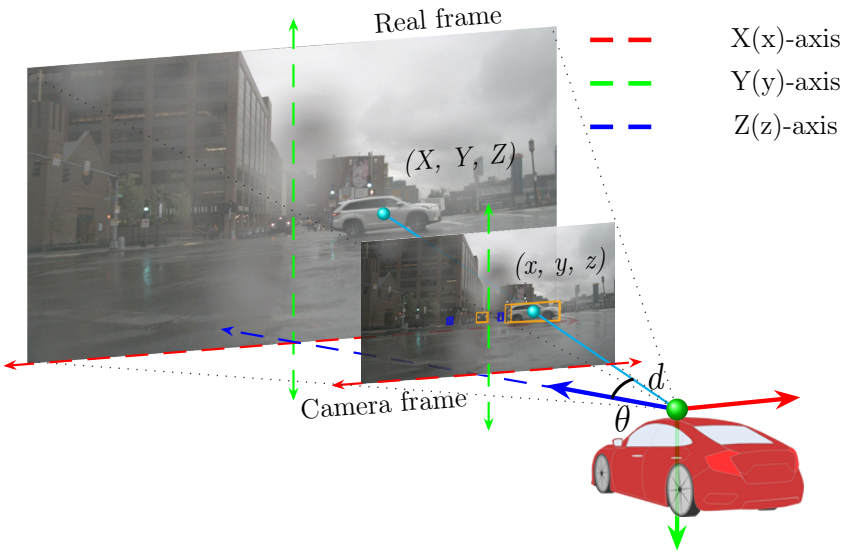


Figure 4.10: Mapping between a real frame and the camera frame.

of the detected object, and ψ_{obj} is the yaw angle of the detected object. The product of v_{obj} and ψ_{obj} captures the influence of the speed component on the same axis of the ego-vehicle shift (z-axis), since the geometric analysis uses the camera's reference system. On the other hand, TTC_{mo} tends to infinity when none of the bounding box vertices of detected objects are in the path of the AV or invading the safety area (sa). Likewise, it is assumed that when detected objects with speed higher than the AV. Finally, when the AV is stopped, it is inferred that there will be no risk event. The speed values of the detected objects and the ego-vehicle in the AV nuScenes are obtained directly from the dataset. On the other hand, the speed data of the ego-vehicle in the Lyft5 dataset is obtained from the analysis of translation data by means of the haversine formula [87].

As a result, it is possible to simultaneously analyze and differentiate traffic risk events in both car following and head-on scenarios. Therefore, TTC is conditioned to the yaw orientation of each road user detected by the ego-vehicle. Positive yaw angles describe road users moving in the same direction as the ego-vehicle, which represents a car following scenario. On the other hand, negative yaw angles describe road users moving in the opposite direction to the ego-vehicle, which defines a head-on scenario.

To quantify the risk level from the TTC analysis, we employ the risk coefficients proposed in [88]. This criterion gathers values which correspond to the reaction time requirements in AVs, based on the parameters described in [89]. Table 4.7 shows the risk coefficient defined according to the TTC values.

Motion orientation has a direct impact on the safety analysis. The road users' is random by nature, therefore it is inferred that traffic risk events require a mapping analysis of the detected objects around the ego-vehicle. Next, TTC_{mo} analysis is

Table 4.7: Risk coefficient as a function of TTC.

Severity grade	TTC [s]	Description	Risk coefficient
0	>4.0	No safety risk	0.0
1	2.5 to 4.0	Accident-to-conflict ratio stable	0.2
2	1.5 to 2.5	Low risk level	0.3
3	1.0 to 1.5	Moderate risk level	0.6
4	≤ 1.0	High risk level	0.8

used on the AV datasets presented in Section 3.

4.6 Performance Evaluation

As stated in Chapter 1, this work proposes a strategy to analyze road safety metrics for AVs, using traffic conflict techniques based on SSMs. To extract the information from sensor datasets and calculate SSMs, a Monitoring subsystem is proposed. The data sequences ordered in the Data Analysis module are distributed in both the SSMs analysis and Reports modules, as shown in Figure 4.11. The SSM analysis module calculates the TTC and TTC_{mo} discussed in Section 4.3 and 4.4. On the other hand, the Reports module shows the data resulting from the data analysis module since it organizes and calculates direct metrics, defined in Section 4.2.1 and 4.2.2, in addition to quantitative characteristics related to data semantic interpretation. Meanwhile, data from the SSM analysis module are delivered to the reports module in an organized way, corresponding to the order delivered from the data analysis module. The SSM analysis module uses the *NumPy* library for the SSMs calculation, while the reports module uses the *Pandas* library for the report format organization and the *Matplotlib* library for graphical analysis of road safety metrics used in the Monitoring subsystem shown in Figure 4.11.

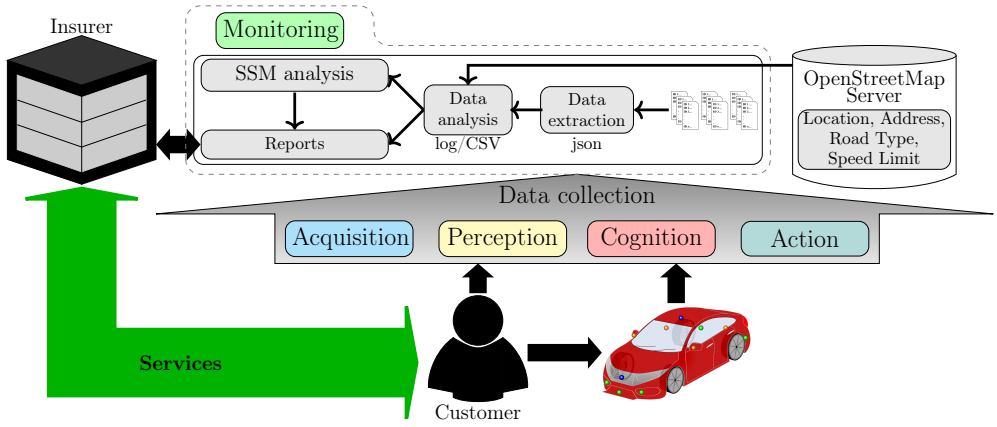


Figure 4.11: Architecture proposed for the monitoring analysis.

4.6.1 Vehicle Tracking

The frequency of each event is influenced by the topology of the cities where the AVs circulate, as shown in Figure 4.12. To analyze the vehicle tracking, we enriched the datasets with data related to road type and speed limit. The ego pose data encoded in translation data are transformed into geodetic coordinates to track the vehicle. Geodetic coordinates are used to make queries in Nominatim[1] and Overpass API[2].

(a) Singapore. (b) Boston. (c) Palo Alto.

Figure 4.12: Trajectories of the AVs in the datasets.

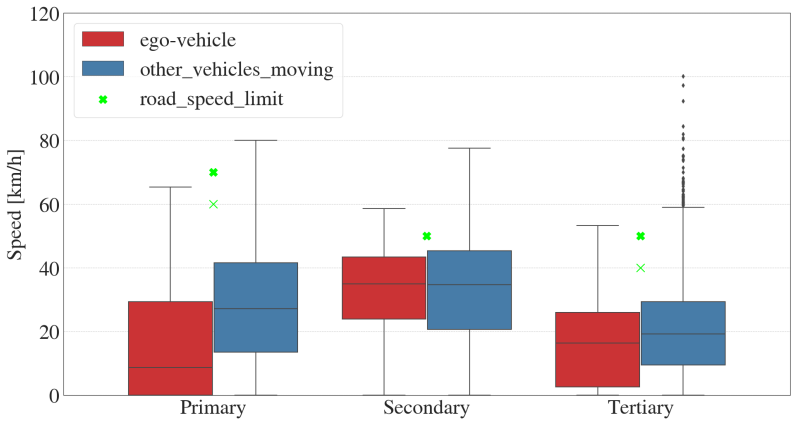
4.6.2 Speed Limit Analysis

From the vehicle tracking analysis in Section 4.6.1, the ego vehicle speed profile is verified. Figure 4.13 shows that the ego vehicle maintains an average speed between 15 km/h and 30 km/h in Boston, 20 km/h and 40 km/h in Singapore, and between 30 km/h and 50 km/h in Palo Alto. Likewise, the speed of vehicles moving in front of the ego vehicle is analyzed. It is possible to observe that some samples exceed the threshold speed limit established by the traffic regulations; obviously relevant information, given that speeding increases the probability of risky events.

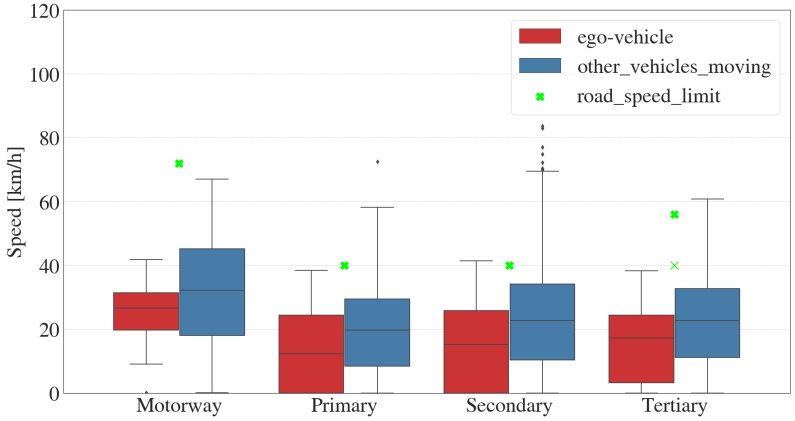
4.6.3 TTC_{mo} Evaluation

Kinematic measures like speed of and distance to the detected objects are used for the TTC_{mo}. Speed and distance are estimated through LiDAR measurements, while the images are used for the recognition of the various objects around the AV. Data is available in the datasets in form of annotations and metadata for each instance (object) detected by the AV. Moreover, annotations are identified by categories, each one associated with each object detected.

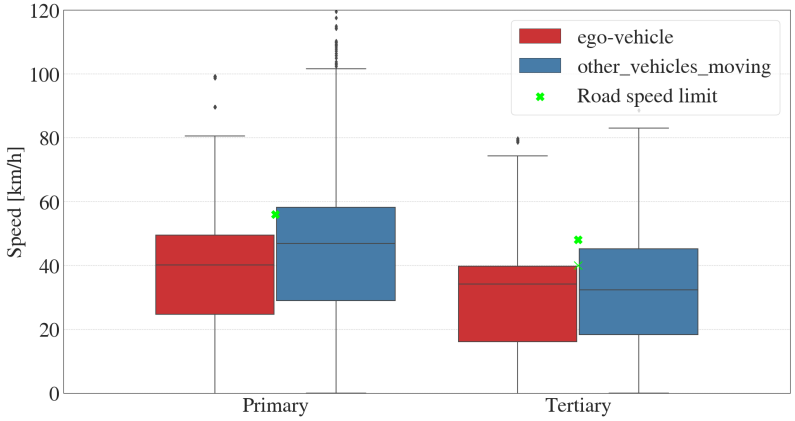
[1]https://nominatim.org/
[2]https://overpass-turbo.eu/



(a) Ego-vehicle speed vs. road speed limit in Singapore.



(b) Ego-vehicle speed vs. road speed limit in Boston.



(c) Ego-vehicle speed vs. road speed limit in Palo Alto.

Figure 4.13: Relationship between the ego-vehicle speed and the road speed limit. x-axis shows the main key used for identifying any kind of road, street or path. y-axis shows the road speed limit for each road type (green marks), ego-vehicle speed (red box plots), and other_vehicles speed (blue box plots). Variations in road speed limits for a same road type are provided with different green marks.

To analyze potential risk events, AV datasets are examined to assess the driving behavior. For that, annotations made to images captured by the front camera are analyzed. Annotations with no speed data are discarded: 5% from the Lyft5 dataset,

4.3% from the nuScenes Boston subset, and 1.1% from the nuScenes Singapore subset. Next, we evaluate the TTC for all remaining valid annotations, in order to observe the proportion of objects that are analyzed through the regular TTC defined in Equation 4.1. In proportion, approximately 70% of the samples represent some risk level w.r.t. valid ones, as shown in Figure 4.14. Different from the analysis with the regular TTC, which only discards events when $v_{obj} > v_{ego}$, the methodology proposed in Section 4.4.1 allows to determine which objects may be in the ego-vehicle's course. Therefore, objects that are not in the course of the ego-vehicle are discarded since they do not represent a potential traffic conflict. Thus, annotations of objects detected in the AV course or the safety zone defined in Section 4.4.1 are used to analyze the TTC. Thus, the proportion of samples representing some risk w.r.t. valid ones corresponds approximately to 8% for the Singapore subset, 5% for the Boston subset, and 4% for the Palo Alto subset, as observed in Figure 4.14.

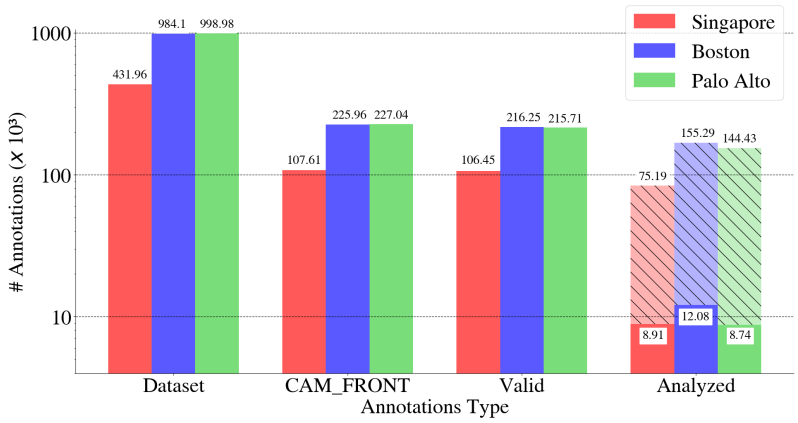
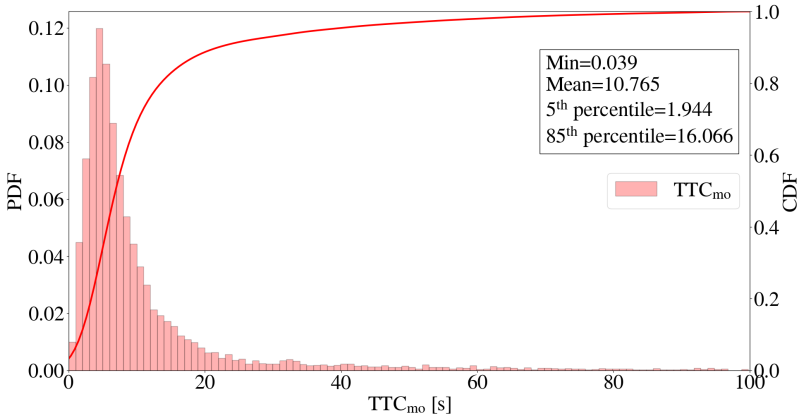


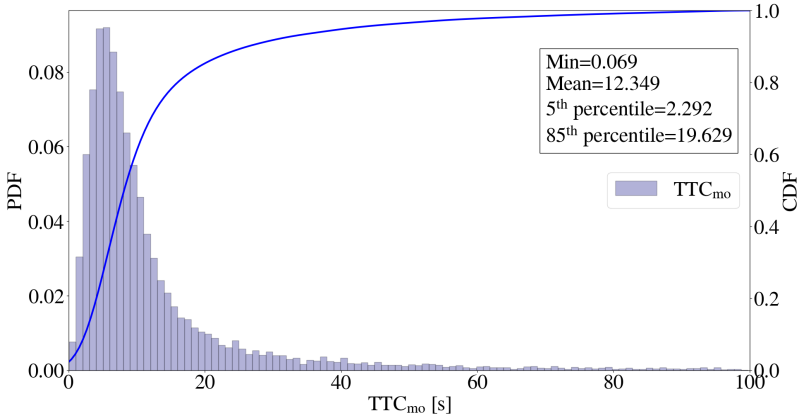
Figure 4.14: Number of annotations ($\times 10^3$) assessed for the analysis of potential risk events in the AV datasets studied. x-axis shows the quantity of annotations available in the dataset, annotations from the CAM_FRONT and the valid and analyzed annotations. Hatch pattern bars in Analyzed label on x-axis correspond to the TTC general formulation analysis; solid color bars correspond to the TTC_{mo} proposed in this work.

Figure 4.15 shows the TTC_{mo} frequency distributions for each analyzed dataset. It is possible to observe that the distribution in all cities is very similar, with distributions skewed to the right. Therefore, the 5^{th} and the 85^{th} percentiles are evaluated, which represent the most pronounced inflection points in the cumulative distribution. Values below the 5^{th} percentile represent TTC_{mo} values $< 2.4\,\text{s}$ in all datasets. We also note that the bulk of representative TTC_{mo} samples are concentrated in up to 33 s, with an average of maximum 18 s.

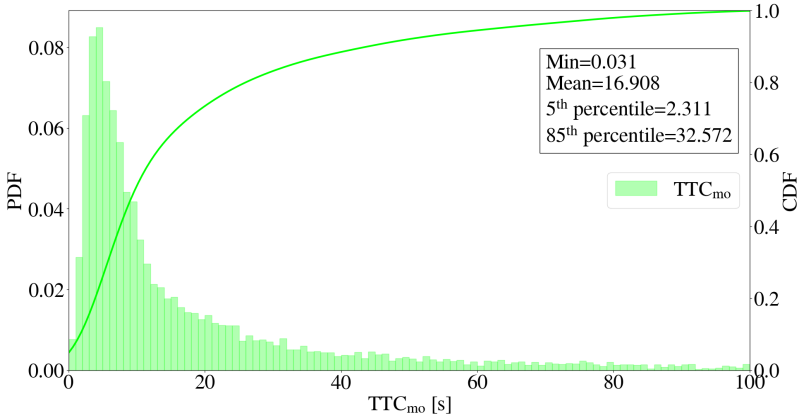
From the annotations analyzed in Figure 4.14, it is possible to observe the frequency and the type of events concerning potential risk events, when both the objects and the ego-vehicle are in collision course. Table 4.8 shows the total frequency of event types based on the course of detected objects, as described in Section 4.4.1,



(a) Singapore.



(b) Boston.



(c) Palo Alto.

Figure 4.15: Cumulative and Probability Density Functions for $TTC_{mo} < 100\,s$ for each dataset.

classified as following, head-on, and crossing events. Course analysis can help to analyze the way in which these objects converge with the AVs. These data are important to consider the severity of the event. For example, a car-following event can have a different effect than a head-on event.

To analyze the risk level of the ego-vehicle interactions with other objects, we use the severity hierarchy based on the level and severity zones proposed by Hydén [4].

Table 4.8: Conflict types defined by position and orientation w.r.t. the ego-vehicle.

Event/City	Singapore	Boston	Palo Alto
Following	3,094	4,452	7,022
Head-on	595	824	147
Lane-change	267	349	104
Crossing	4,947	6,456	1,463
Total events	8,903	12,081	8,736

Severity level defines a threshold for serious and non-serious conflicts. On the other hand, severity zones quantitatively define severity levels. Both severity level and zones are based on a relationship between time and speed. A fixed threshold to define a high-risk event is based on the Time-to-Accident (TA) under a traffic conflict. This value was established at 1.5 s [4], which is consistent with the studies reported in [89], and that corresponds to the response time of the sensors readings, processing, recognition and planning tasks of the AV between the detection of an obstacle and the evasive action. The diagram in Figure 4.16 is used to define the severity level of the conflict.

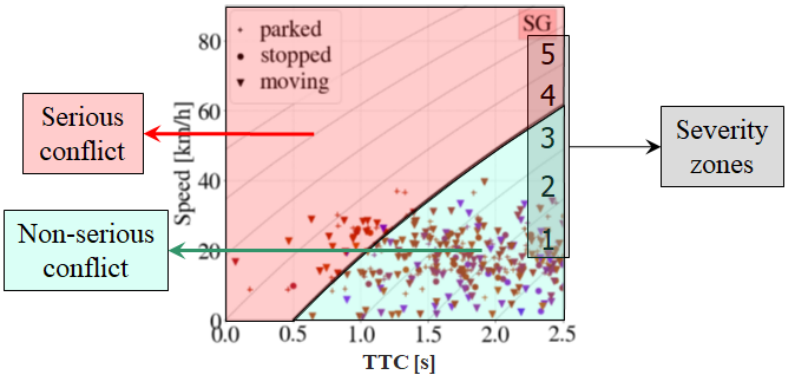
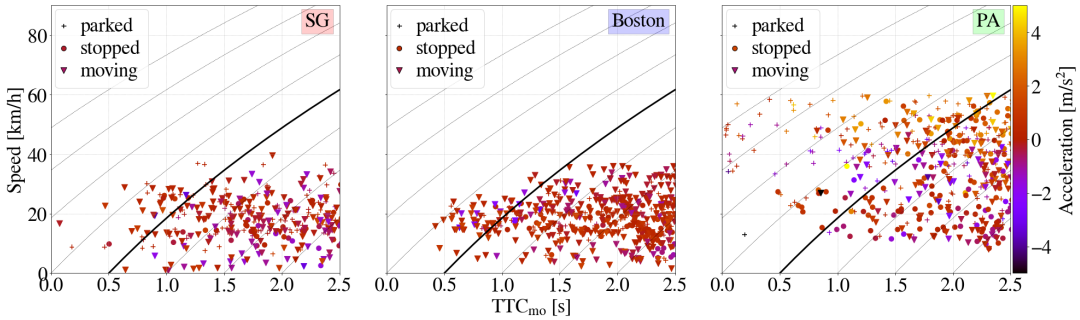


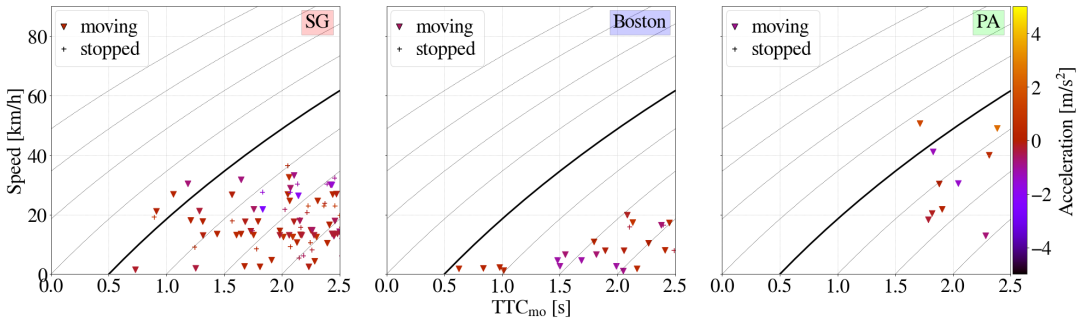
Figure 4.16: Conflict severity diagram.

All interactions that represent some level of risk for the ego-vehicle are presented in Figure 4.17. All interactions within the 5$^{\text{th}}$ percentile are plotted, as observed in the cumulative distributions of Figure 4.15. We note that most of the observed interactions in Singapore (SG) and Boston occur with vehicles and objects. Figure 4.17(a) shows that interactions with $\text{TTC}_{\text{mo}} < 1.5\,\text{s}$ occur with other moving vehicles, with a deceleration pattern as the TTC_{mo} decreases. On the other hand, in Palo Alto (PA) we observe more interactions with parked vehicles: this characteristic is due to the interaction with vehicles that invade the safety area we have defined for the AV. It is interesting to note that interactions with pedestrians show some events that represent lower risk of collision, as shown in Figure 4.17(b). The same behavior is observed for the two-wheelers in Figure 4.17(c). Finally, Figure 4.17(d) shows the interactions with objects of the vehicular infrastructure like barriers, traffic cones,

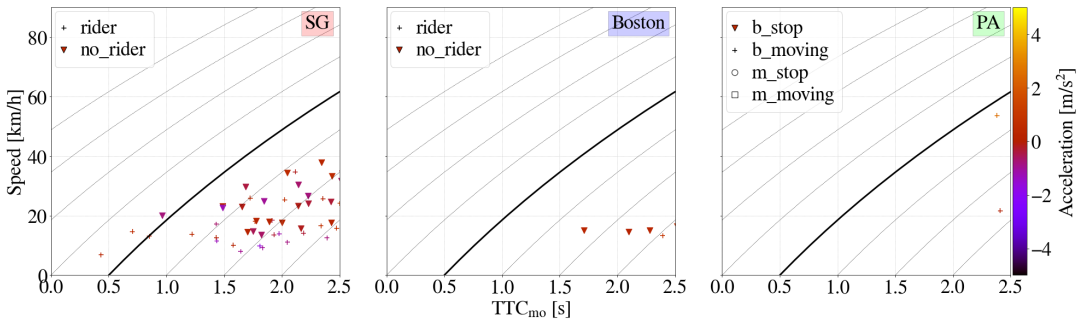
among others.



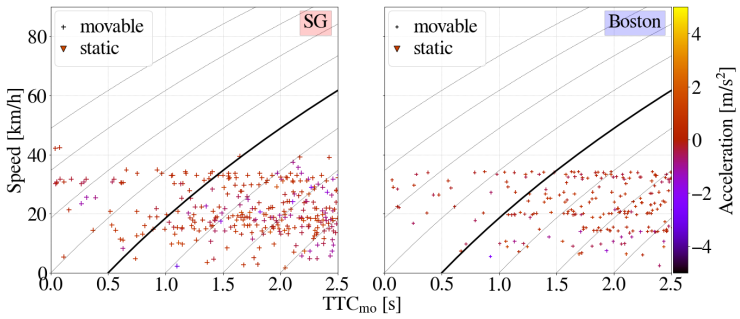
(a) Vehicles.



(b) Pedestrians.



(c) Two-wheelers.



(d) Objects.

Figure 4.17: TTC$_{mo}$ 5th percentile indicators for each scenario. The columns describe the city where the interactions take place: to the left Singapore (SG), to the center Boston, and to the right Palo Alto (PA). Meanwhile, the rows describe the general category of objects interacting with the AV. Conflicts above the black line on the graphs are ranked as serious; below the black line, non-serious.

To summarize, the proportion of interactions for all the annotations analyzed represents less than 1% for high-risk events, whereas events with some risk represent approximately 10%. Events that do not represent any risk represent more than 70%, as shown in Figure 4.18. Compared to valid annotations, the proportion of interactions that represent some risk level is less than 2%.

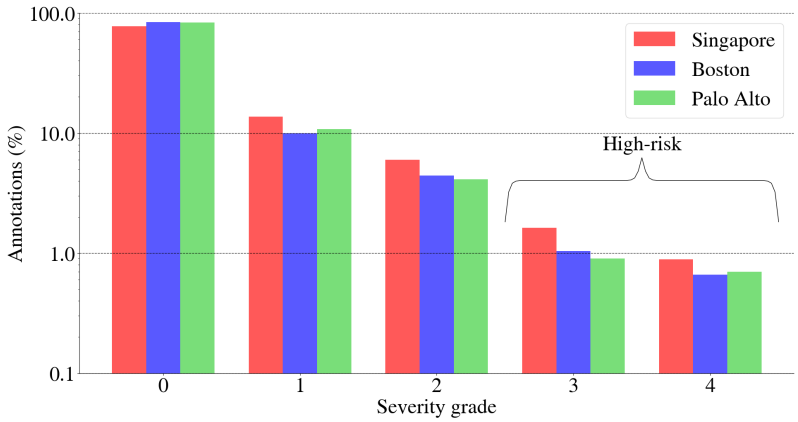


Figure 4.18: Annotation volume based on severity grade ratio.

The present TTC_{mo} analysis allows assessing risk events through the geometric analysis of the boundaries associated with each object detected by the AVs. Thus, it is possible to limit the analysis to objects in a possible collision course. This is relevant for TTC_{mo} analysis since it is possible to identify how interactions occur with various road users and objects. Nevertheless, further investigation is needed to establish a pattern of AV behavior with a longer time sequence in the scenes, mainly to obtain more parameters to describe driving behavior patterns related the AI system that controls the vehicle.

An advantage of data analysis through exteroceptive sensors available in autonomous vehicles is that risk assessment is not limited to claims related to vehicles only; as shown in Figure 4.17, where the TTC_{mo} is assessed for various categories and attributes available in the datasets. Moreover, the distribution of risk events was similar among the three datasets, with 85% of the sampling concentrated in less than 33 s, and the highest risk events below 2.4 s, as shown in Figure 4.15. It is also important to note that the analysis of safety metrics for various road users will depend on the data labeling available. This can be observed for example in Table 3.4, where the Lyft5 dataset does not have data related to traffic infrastructure objects.

It should be noted nonetheless that there are some limitations in the used AV datasets. The sampling time of each scene is limited to a maximum of 25 s (Lyft5), and 20 s (nuScenes), in most cases without sequence, which prevents observing a greater number of events with potential risk. Another limitation is related with the speed of the AVs analyzed, which is much lower than the limit speeds of the road infrastructure. The speed uniformity of the AVs reduces the possibility of observing

the effect of the evasive actions by the AV. Finally, the number of vehicles limits the risk assessment analysis since the age and learning experience of the autonomous system may still be limited.

Finally, the calculation of TTC_{mo} considering the motion orientation of the detected objects reduces the overload generated by the volume of data in the safety analysis. Thus, TTC_{mo} reduces by up to 60% the proportion of data to be analyzed when compared to the regular TTC, as shown in Figure 4.14. Motion orientation and geometry analysis enable to discard all objects that, despite interacting with the AV, they do not converge on a collision course, and therefore, they do not represent a risk for AV. It is relevant if we consider that safety monitoring requires immediate analysis when exists potential high-risk events.

4.7 Remarks

In this chapter, SSMs were studied with the idea of quantifying risk from the interactions of self-driving vehicles with other road users and infrastructure. SSMs enable the analysis of safety metrics based on variations in speed, distance, acceleration/deceleration, and time. However, these measurements require validity. This concept involves various features, such as sufficient evidence, an appropriate description of conflicts, identification of high-risk locations, among others. Therefore, the validity process requires constant analysis. Hence, SSMs analysis can help to improve the risk assessment process describing traffic risk events accurately. On the other hand, validity processes are important for the explainability and fairness paradigms in charge of policymakers. To validate traffic risk events, the method selected in this section was the TTC. TTC has proven to be an effective measure for severity of traffic conflicts. In fact, TTC allows risk quantification for each AV interaction with other road users, discriminating critical from normal behavior. Besides that, semantic data enable to analyze not just interactions with other vehicles, it is also possible to evaluate the behavior of the autonomous controller system with other road users, which also allows understanding the logic and priorities of the autonomous controller system, which are part of the paradigm in HMT. In this context, this chapter proposes TTC_{mo}, an improvement of the regular TTC that allows analyzing interactions with road users and objects in collision course with the ego-vehicle. The metric uses the yaw orientation and the geometric analysis of bounding boxes generated in the autonomous controller system for the recognition and perception of objects detected by the exteroceptive sensors. This is important to understand the risk when detected objects converge with the ego-vehicle. Some highlights about the SSMs are described below:

- All metrics require data validity to define their potential importance in the road safety analysis.
- Road safety analysis through observation of real traffic risk events enables the behavioral and situational assessment through parameters and their importance in potential conflicts between an ego-vehicle and road users.
- Road safety monitoring through SSMs allows defining assumptions about the projected trajectories of the road infrastructure, and planned routes of road users.
- Data collection of AVs allows validating SSMs through sensor readings inherent to the vehicle. Although AV datasets do not have information on fatalities or crashes, predictable event risks analysis makes available the development of AI patterns in road safety monitoring.
- All metrics require real-time analysis. This is essential for the immediateness of risk assessments. On the other hand, real-time analysis depends on processing data from the AV sensor. Data interpretation is crucial to identify detected objects around the vehicle, and consequently, that is reflected in a high computational cost. Therefore, SSMs analysis depend on AV perception and cognition processes to define road safety assessments, diagnostic activities based on automated observations, and risk assessment models adjusted to the real scenario.

Risk assessment in self-driving vehicles is an initial step towards understanding potential AV driving patterns on the road. This is a challenging scenario considering that the insurer currently interacts with drivers through metrics, alerts and suggestions to offer benefits and improve safety on the road; now, considering a self-driving vehicle, the question is how to pass these monitoring results to the autonomous controller system. It is well known that the insurer cannot influence the decision-making of the autonomous system other than the car manufacturer or the autonomous system administrator. In this context, it is necessary to consider that both insurer and car manufacturer have data with different features since car manufacturer has data coming from all the sensors of the AV, while the insurer has more derived data (from OTS devices), but an overlap exists, e.g., speed or position. Therefore, it is necessary to establish a collaborative environment in which organizations can acquire information from one trusted another to improve their corporate processes. Thus, it is possible then that the insurer's monitoring helps the car manufacturer to improve its autonomous driving model; on the other hand, the insurer can reduce claims and the self-driving vehicle owner can be rewarded.

In the next chapter, this thesis evaluates the use of ML techniques to establish a collaborative environment in which it is possible to share data, while ensuring data integrity and privacy-preserving.

Chapter 5

Privacy-Preserving Collaboration for Mobility Safety

Self-driving vehicles have numerous sensing systems to meet the challenges arising from the complexity of the varying vehicular landscape and preserving the safety of the vehicle, passengers, and road users. This is an advantage if we consider the sensors' mobility a source to obtain data that, when interpreted, allow providing accurate, personalized and real-time services. This is a challenging scenario for vehicle risk assessment, considering that the growth of data volume generated by AV is exponential [20], and therefore, data analysis can require high computational cost in terms of complexity and efficiency. Thus, it is necessary to make use of edge computing infrastructure or even cloud access to preprocessing or processing tasks associated with road safety monitoring. These data are processed generally through centralized algorithms based on machine learning techniques [90]. Nevertheless, this data may contain sensitive personal information that could compromise the privacy and data integrity, whether through data leakage or eavesdropping. This issue is very relevant considering that several data protection and regulatory institutions prohibit the upload of user data content without the user's consent [91].

To address these challenges, a potential solution is the use of Federated Learning (FL) techniques. The concept of federated learning is proposed by Google [28, 92–94]. Their main idea is to build ML models based on datasets that are distributed across multiple devices or sites while preventing data leakage [7, 95]. Unlike ML, where the learning process is centralized, FL aims to train data in the location where the data is stored, then, each party communicates the results of its "local" model in order to reach a consensus to obtain a "global" model. Finally, the global model feeds back the local model, and the learning process continues. As a result, information of the model training can be exchanged between parties, but not the data. Depending on the application, FL may contain multiple parties, i.e., individuals (e.g., IoT devices, OTS devices, among others), or organizations (e.g., hospitals,

car manufacturers, financial entities, among others). In addition to the information exchange process, FL implements local algorithms to preserve party privacy and data integrity. Furthermore, the FL architecture can be peer-to-peer, so it does not require a coordinator.

The scenario to be analyzed is illustrated in Figure 5.1. The objective is to use data from car manufacturers for risk assessment based on the insurance monitoring analysis. In this context, car manufacturers learn from data sent by the AVs; in the meantime, the insurer monitors the ego-vehicle and learns from the data sent by OTS devices located on-board the ego-vehicle (red vehicle in the illustration), making a driving profile audit. It should be noted that the car manufacturer learns from its own vehicles, while the insurer can monitor multiple vehicles from multiple car manufacturers. We assume that ego-vehicle sends data from the sensors readings to the car manufacturer. In the meantime, we assume that ego-vehicle sends data via OTS device to the insurer. We use the nuScenes AV dataset as car manufacturer data. On the other hand, we evaluate the scenes individually in order to generate risk assessment labels as part of insurer data. It is important to note that we do not evaluate communication performance in the scenario shown. We assume that the communication efficiency between vehicles, car manufacturer and insurer does not pose a problem for the risk assessment based on Vertical Federated Learning (VFL).

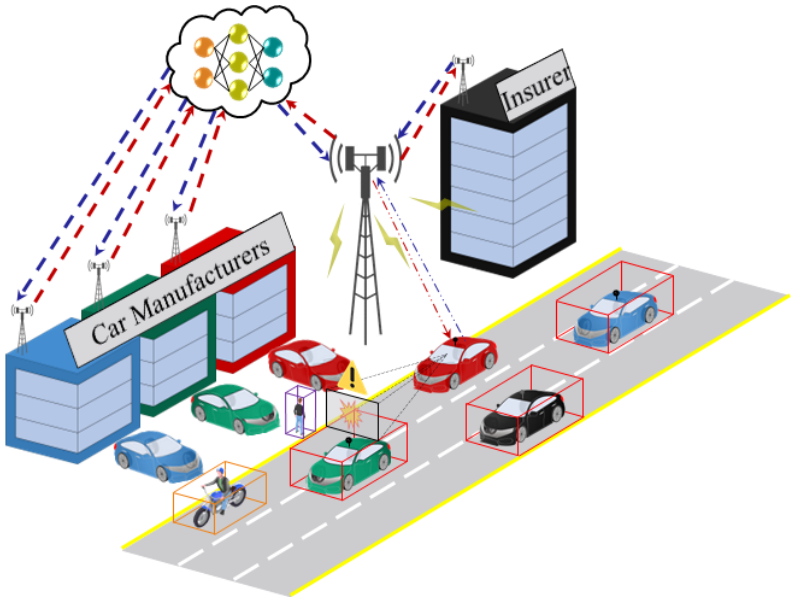


Figure 5.1: Federated learning in the context of car manufacturers and insurance companies sharing data from self-driving vehicles.

The use of VFL for risk assessment can bring several advantages within the context of mobility safety:

- Self-driving vehicles can be individually evaluated to make them smarter and make better decisions.

- Car manufacturers and insurers can have a different way/perspective/metrics to score AV driving software, but they don't share data as they are different companies, and they could work with different competitors.

- Both the car manufacturers and the insurance company would be glad to improve the driving model, though (better driving experience and fewer claims to pay). The driving model is supposed to be a reinforcement learning model, so we need data labels along with new data to train the model with bonus/penalty rewards. We assume that the label corresponds to a level of risk (i.e., 5 for less risky and 0 for quite risky vehicles).

- Car manufacturers and insurers create a partner ecosystem to work together towards self-driving model improvement and claim reduction. Figure 5.2 summarizes the challenges of the collaborative environment.

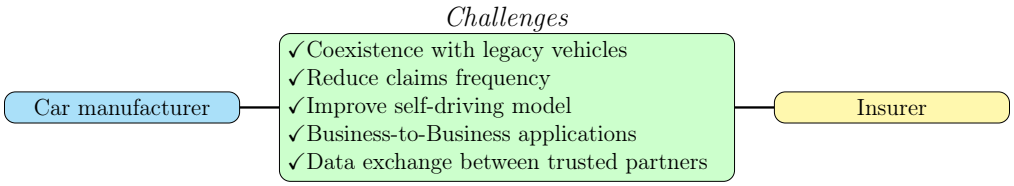


Figure 5.2: Advantages and challenges of data sharing between automakers and insurance companies.

In this chapter, the aim is to emulate VFL to assess risk based on shared data between different automakers partners. We demonstrate the VFL results compared to a traditional centralized ML environment. Thus, we aim to improve the autonomous driving model (whatever it is) putting together data from insurer and car manufacturer via the FL approach (so that they collaborate without sharing data). Note that we consider that both insurer and car manufacturer have data with different sets of features, but some overlap exists, for example both may have speed or positioning data. Moreover, we also assume that both the insurer and the car manufacturer have their own labels. Ideally, the car manufacturer has data which is collected from the vehicles, while the insurer has data which is collected from OTS devices (e.g., an app installed in a smartphone). For this thesis, data from nuScenes AV dataset is used as car manufacturer data. On the other hand, as we do not have nuScenes AV data from the insurer, risk assessment data is provided by visually evaluating the videos of scenes available in the nuScenes AV dataset. The result of this visual analysis is used as data by the insurer in the collaborative environment.

5.1 Federated Learning

One ML approach that implements data privacy-preservation techniques is Federated Learning (FL). FL allows building ML models based on datasets that are distributed across multiple devices or sites, while preventing data leakage [7, 95]. Furthermore, a goal of FL is to take advantage of data silos, where data is collected and located at individuals (e.g., IoT devices, OTS devices, among others), or organizations (e.g., hospitals, car manufacturers, financial entities, among others), in order to improve learning processes to become more accurate at predicting outcomes process. Besides that, federated learning seeks to maximize computational power in order to share computed results, rather than raw data, improving communication efficiency [95].

Federated learning allows data to be manipulated in a distributed way among different users, overlapping datasets with different features. This overlapping can occur Horizontally (HFL) or Vertically (VFL). HFL assumes that samples of parties contain distributed data and labels. On the other hand, VFL analyzes data and labels with different features, i.e., labels are not contained in the data, but their significance is relevant in the model analysis. Figure 5.3 shows the vertical federated learning for a two-party scenario. Two parties A and B collect data samples from a common user. Data from A and B can be overlapped or simply contain partitioned data features or labels. Data user in A and B can be aligned, either by a common ID or by time intervals, however, they contain different features. Party A can learn from party B's data and vice versa. Moreover, parties can learn jointly.

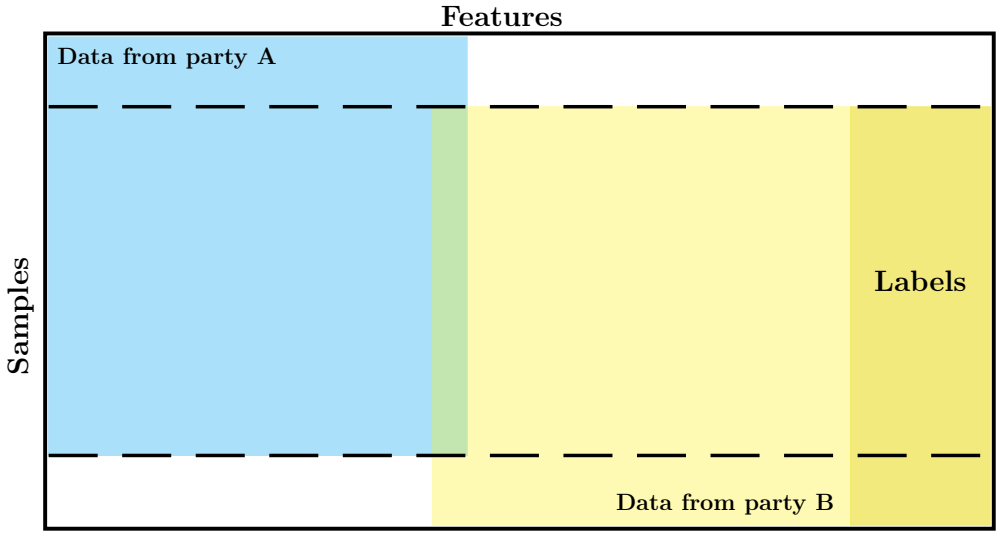


Figure 5.3: Illustration of VFL [7]. Data from party A is highlighted with blue color; data from party B is highlighted with yellow color. Dashed lines delimit the samples in which A and B are taken to train a federated learning process.

In general, VFL system defines the following considerations:

$$\mathcal{X}_i \neq \mathcal{X}_j, \mathcal{Y}_i \neq \mathcal{Y}_j, \mathcal{I}_i = \mathcal{I}_j \ \ \forall \ \mathcal{D}_i, \mathcal{D}_j, \ i \neq j,$$

where $\mathcal{X}$ and $\mathcal{Y}$ denote the feature space and the label space, respectively. $\mathcal{I}$ is the sample ID space, and matrix $\mathcal{D}$ represents data held by different parties [7]. It is important to note that the identity and the status of each party in VFL is the same, allowing each party in the federation being helped by the learning process.

5.1.1 Vertical Federated Learning

The purpose of VFL is to generate a learning environment where different parties with different interests, with different feature spaces, but share a pool of common users, can take advantage of the heterogeneity of their features to improve their local learning models without raw data exchange or expose private data. Therefore, we aim to use VFL to establish a partner ecosystem to improve the self-driving model through risk assessment analysis. To implement VFL in this ecosystem, we use AV datasets that contain raw and semantic data from the vehicle while it is on the road. For instance, exploring the partner's ecosystem allows the effective participation of different entities, from Original Equipment Manufacturer (OEMs), communication providers to government entities. Thus, it is possible to evolve towards new learning models that allow coexistence and evolution between current driving profile monitoring models, preserving privacy and data security.

5.2 Related Work

Vertical federated learning is a machine learning technique in which it is possible to share data that have similar information, without being exactly the same, and that aims to predict predetermined activity patterns [7, 28]. Moreover, VFL enables to build high-performance models shared along multiple parties considering concerns regarding user privacy and data integrity, creating a trusted environment that conforms to the legislative requirements for data protection [91]. VFL has been explored by researchers and industry in different business activities, such as finance, and healthcare.

Webank is a Chinese bank that together with its AI division become developing FL-based solutions. Webank [96] developed *FedRiskCtrl*, a framework for assessing the financial risk in loans for small and medium-sized companies. This model defines two parties: an invoice agency A, which has data associated with each company's invoices; a bank, which has credit-related data and label features of each company. To align company data, the framework uses Private Set Intersection (PSI) or secure

entity alignment. The framework is based on Federated AI Technology Enabler (FATE) [97], a framework that implements secure computation protocols based on homomorphic encryption and Multi-Party Computation (MPC). The authors observed that the model built with FL performs significantly better than the model built only with the centralized dataset from the bank data.

Another tool developed by Webank is *FedVision* [96], an edge computing application that aims to detect objects to train risk event recognition in video surveillance. In FedVision, an initial object detection model is sent from the FL server to surveillance companies that use local stored data to train the object detection model. At the end, each surveillance company sends encrypted parameters of the local model to the FL server to feed a global model. The process is repetitive until a criterion consensus is reached. Despite the VFL having a better performance than the centralized model, the frameworks are susceptible to biased model by imbalanced data samples, and also by unequal distributions.

Cheng *et al.* [98] proposed a system called *SecureBoost*, in which all participants combine user features to obtain an improved accuracy in the prediction model. SecureBoost is implemented over the FATE framework. The authors claim that the system is lossless compared to the centralized model. The model defines two parties: an active party, which is the data provider, who holds both a data matrix and the labels. Data provider is defined as a server; and a passive party, which has only the data matrix. Passive parties are clients in the model. The model is tested with two credit public datasets: the first one, a credit scoring dataset based on loans[1]; the second one contains information on default payments, credit data, among others, of credit card clients in Taiwan[2]. The goal is to analyze financial problems and payments on time. The authors conclude that theoretically proof shows that the proposed framework is as accurate as non-federated gradient tree boosting counterparts.

Another area where VFL has been explored is healthcare. Duverle *et al.* [99] presents a statistical analysis model based on logistic regression to analyze hypertension by smoking status in vertical data partition. The model uses the Paillier cryptosystem for privacy-preserving framework. The authors use data collected from a cohort study on genetic variants linked to chronic kidney disease. The experiment analyze data from 4,257 patients clinically assessed with hypertension. In the clinical data, the authors selected smoking status. The authors conclude that the method is not only as secure and accurate as existing methods, but provides a marked improvement in performance when compare with other models.

Chen *et al.* [100] proposes a Vertical Asynchronous Federated Learning (VALF)

[1]https://www.kaggle.com/competitions/GiveMeSomeCredit/

[2]https://www.kaggle.com/datasets/uciml/default-of-credit-card-clients-dataset

method to perform stochastic gradient algorithms without requiring coordination between clients. For privacy-preserving data, the authors use a local perturbation embedding technique. Numerical tests were carried out to test VAFL for logistic regression using a disease dataset, and VAFL for deep learning using the Medical Information Mart for Intensive Care-III (MIMIC-III) [101] dataset. The result shows that VAFL learns a federated model with accuracy comparable to that of the centralized model, and requires less time relative to the synchronous FL algorithm.

Budrionis *et al.* [102] benchmarks the PySyft framework in medical data [103]. The authors deployed a realistic scenario to compare the performance in terms of execution time. Three scenarios were tested: (i) Data: the number of nodes in the network was fixed, while the data as uniformly increased and distributed across the nodes; (ii) Nodes: the number of nodes was increased from 1 to 128, while the amount of data was fixed to maximum; and (iii) Distribution: the number of nodes constant ($n = 32$) and the amount of data was fixed to the maximum. The benchmark architecture uses a model-centric setup, where a coordinating node host the global model, while computational nodes download the model and train it on local datasets. The experiment setup uses the MIMIC-III dataset. The framework's performance shows that predictions of the federated model are equivalent to those of the centralized model. Nonetheless, federated model training and inference took between 9 and 40 times longer than equivalent tasks in a centralized model.

In the vehicular environment, an advantage with self-driving vehicles is the ubiquity of sensors embedded in AVs, and therefore, an expectation that there will be a greater volume of data and training ML models.

Peng *et al.* [104] propose the Blockchain-based Federated Learning Pool (BFLP) framework solution that uses blockchain for privacy-preserving and VFL to determine complex road conditions and accident-prone areas. Depending on the application, the authors propose a federated learning pool, a module carried in servers with the capacity to select the most suitable FL method according to the data sources. Moreover, the authors propose a lightweight encryption algorithm to be executed in the local model. Specifically, in VFL, the authors consider data from vehicle sensing and transport department data to analyze complex road conditions and accident-prone areas. FL implementation in BFLP uses FATE framework. BFLP was tested with simulated data from TOSSIM simulator. Results about road conditions and avoiding obstacles show that success rates with BFLP was higher accuracy than centralized model based on logistic regression model.

Yuan *et al.* [105] presents Federated Deep Learning based on the Spatial Temporal Long and Short-Term Networks (FedSTN), a traffic flow prediction model to improve mobility in smart cities. FedSTN uses collected vehicle trajectory data to predict future road traffic stored in edge computer servers. The authors use VFL to

design a layer of their model called the Federated Graph Attention (FedGAT), which allows capturing short-term temporal information without loss of spatial information among areas and sharing these parameters based on VFL. Specifically, in addition to traffic flow prediction, FedGAT allows learning traffic flow from meteorological data acquired by external sensors. The authors employed two datasets: TaxiNYC and TaxiBJ. FedSTN is compared with other baselines, and the authors conclude that FedSTN has higher traffic flow prediction accuracy than the centralized model.

Wang *et al.* [106] implements vertical federated factorization machine algorithm by homomorphic encryption to realize EV charging point recommendation. The authors use data from the historical charging point data, and at the same time generate labels from the historical charging orders. The goal is to obtain recommendations for electric vehicle charging services. Data from two parties are used in the VFL environment: charge point data, and vehicle data. The architecture also implements a coordinator node or third-party platform to manage the encrypted features. The authors concluded that the federated model is almost lossless with respect to the centralized model.

This thesis evaluates the use of VFL for risk assessment based on data sharing between car manufacturers and insurers. Thus, it is possible that several partners can share samples in similar time intervals. This message exchange allows partners to access a part of a learning model at the same time that they can run a training process for the segments they want to analyze. For that, we aim to emulate the VFL techniques at the edge using data from self-driving vehicles as data owner, and we define labels from the insurer as data scientist to learn and detect risk from data owner with privacy preservation.

5.3 Architecture for Risk Assessment based on VFL

The architecture used to describe the VFL system used in this thesis is shown in Figure 5.4. The model employed is peer-to-peer, i.e., without a coordinator. Thus, organizations A and B communicate directly without the help of a third party. For purposes of this thesis, we define organization A as the car manufacturer, and organization B as the insurer. We assume the car manufacturer is honest; meanwhile, the insurer is honest-but-curious [107]. Honest-but-curious parties indicate that parties involved in the federated environment try to deduce as much information as possible from the data provided by the partners, without disturbing them while the federated environment is active. Furthermore, it is assumed that the federated environment is secure and reliable, while the communication is lossless and unaltered while exchanging intermediate results. In this scenario, the car manufacturer provides its data in order to allow the insurer to improve its risk assessment model.

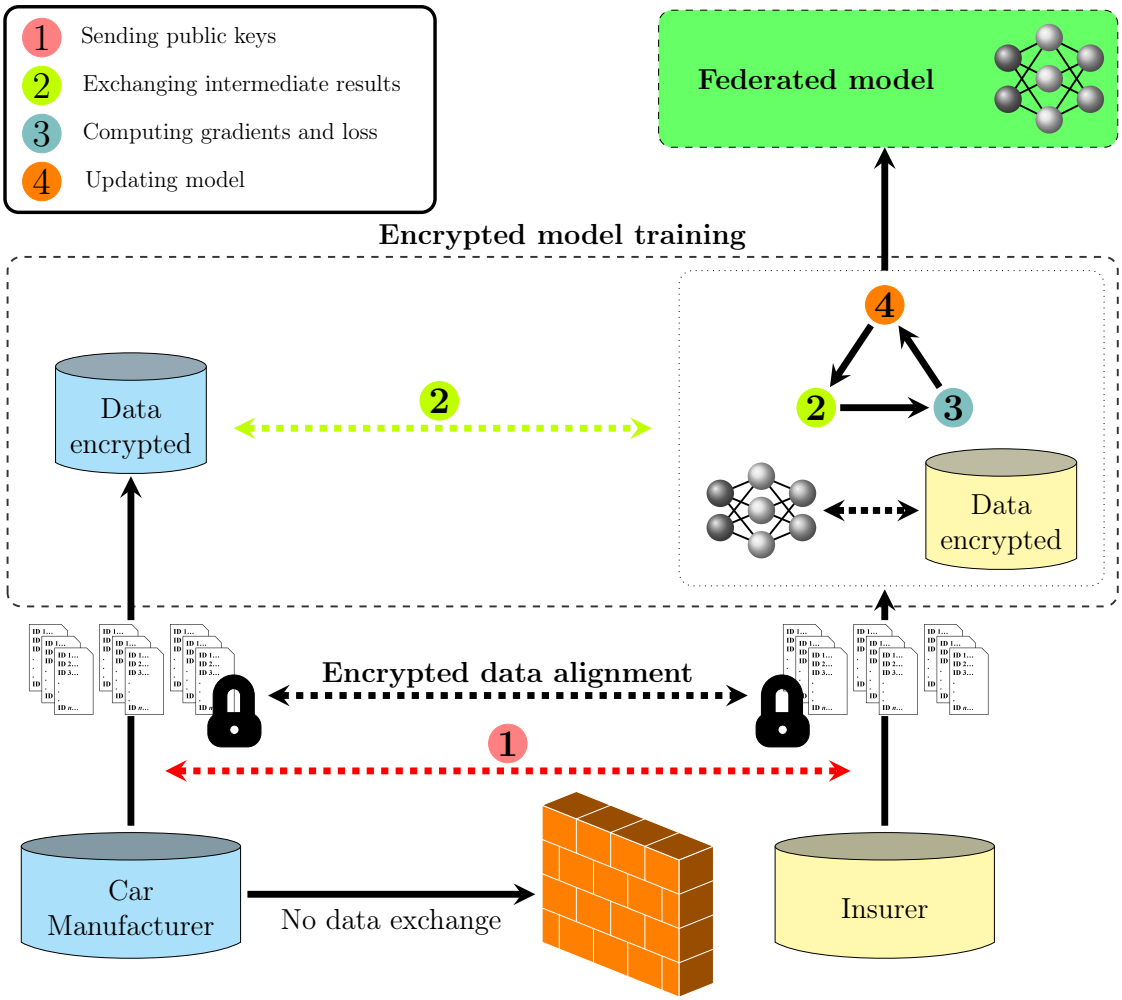


Figure 5.4: Proposed architecture for VFL.

5.3.1 Problem Statement

Consider a set of n automotive industries: $\mathcal{A} := a_1, a_2, ..., a_n$. For example, a_1 may be a car manufacturer and a_2 an insurance company. Both a_1 and a_2 can learn in a collaborative scenario through a distributed machine learning model with a vertical setting. In the problem definition, we assumed that a_1 shares a complete dataset $\mathbf{X} \in \mathbb{R}$ containing m samples, and that a_2 contains a label dataset $\mathbf{Y} \in \mathbb{R}$. Each label is associated with a unique set of features. Thus, we assume that a customer is affiliated with each organization, and these hold different data from the customer's information. Thus, a new dataset $\mathbf{X}_{a_1}|\mathbf{Y}_{a_2}$ is created, being that it does not exist in one place, and each a_n in the network maintains data privacy through encryption techniques. This work considers a passive partner (a_1) that makes available the self-driving vehicle data, and an active partner (a_2) that analyzes data from a_1 to learn about mobility safety.

5.3.2 The Syft Framework

Experiments were performed using Syft [108] version 0.5.0. Syft is an OpenMined[3] open source federated learning framework developed for building secure and scalable ML models [108, 109]. Syft provides diverse methods for privacy preservation like federated learning with differential privacy, encrypted computation through Multi-Party Computation (MPC) and homomorphic encryption. Moreover, Syft is supported on PyTorch [110], an open source machine learning framework, and compatible with other tensor libraries like Tensorflow, scikit-learn, among others [108]. A goal of Syft is to provide an approach where parties can perform models with data encrypted. For that, Syft uses Duet, a framework focused on providing coordination functionalities between parties in a FL environment. Duet enables parties to perform data analysis, in addition to data control, for the data owners [108, 111].

Syft's architecture defines two parties, the Data Owner (DO) and the Data Scientist (DS). The Data Owner is the party who creates a Duet session. Moreover, DO has control over permissions to share and access information. On the other hand, the DS is the party that connects to the DO. Once a DS requests a connection, it requires an input verification key. Signing is done through PyNaCl library, using a 256-bit public keys [108]. Once public key authentication is satisfactory, the socket connection between DO and DS is established.

Once the connection is established, Duet shares data via serial communication. Remote computation in Syft uses Abstract Syntax Tree (AST), which is a tool that allows to map Python modules to support third-party Python libraries, in addition to generating Agents to resolve issues using credentials, attributes, among others. Also, all remote computing operations are performed locally and marked as network pointers. Thus, all operations performed on these network pointers are encapsulated as remote procedure calls performed on the remote machine [108]. Finally, all interactions and metadata are stored in the DO, this includes all intermediate data created on remote execution in DS and results from the learning process. Serial communication on Duet is based on sockets through WebRTC[4] library. Communication connection takes place over UDP, and uses Datagram Transport Layer Security (DTLS) as data integrity mechanism. In addition to integrity, Syft uses the Multi-Party Computation solution (SyMPC)[5] encryption method or homomorphic encryption through TenSEAL [112].

[3]https://github.com/OpenMined/PySyft
[4]https://webrtc.org/
[5]https://github.com/OpenMined/SyMPC

5.4 Methodology

Next, the methodology for data preparation is described, and the establishment of a federated environment to assess risk from the car manufacturer's data is described.

5.4.1 Manual Risk Level Classification

In this section, image sequences from each scene available in the nuScenes AV dataset are manually (visually) classified. The objective is to generate labels to define the risk level of the self-driving vehicle from the perception of different evaluators. These labels will be used as part of the insurer's risk assessment model.

First, all images corresponding to the keyframes of each scene are selected to generate a video sequence of 20 s each scene. In this step, three human evaluators classify the risk level into three levels: 0 (*low_risk*), 1 (*moderate_risk*) and 2 (*high_risk*). Each evaluator was responsible for the manual evaluation of 850 scenes. In this way, approx. 32,965 frames (98.37%) were evaluated as *low_risk* events, approx. 480 frames (1.37%) were evaluated as *moderate_risk* events, and approx. 87 frames (0.3%) were evaluated as *high_risk* events. It is important to note that the evaluators selected for this task have experience in driving, and the evaluation was carried out for all available video sequences. Figure 5.5 shows some examples of how each video/sequence of scenes was evaluated, and the labels generated from the risk perception. "*name*" is defined as the ID; "*driver/no_driver*" defines if the evaluator is driver or not; and "*Time video player*" denotes the second of the video sequence where is detected some risky event.

name	driver / no_driver	scene_id	Time video player																			risk levels
			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	1 - low_risk
																							2 - moderate_risk
																							3 - high_risk
Norida Hoyos	driver	0001	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
Norida Hoyos	driver	0002	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
Norida Hoyos	driver	0003	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
Norida Hoyos	driver	0004	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
Norida Hoyos	driver	0005	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
Norida Hoyos	driver	0006	1	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
Norida Hoyos	driver	0007	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
Norida Hoyos	driver	0008	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
Norida Hoyos	driver	0009	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	

Figure 5.5: Example of the manual labeling classification.

We evaluate the confidence degree of the information manually classified by the judges by means of the *Alpha Krippendorff* coefficient, which is a statistical evaluation strategy between 0 and 1 used to estimate the agreement degree of the evaluators [113]. An agreement degree of 0.73 was observed in our evaluation, which indicates that there is moderate to strong correlation among the labels suggested

by the evaluators.

5.4.2 Imbalanced Data

A classic risk assessment problem is data imbalance, since high risk events are considered unusual cases, as shown in Figures 4.3 and 4.4. This is even more visible considering in AVs, where risky situations are limited. Moreover, AV coexists with other road users, not only vehicles, and therefore, it is inferred that more no risk events are more frequent. This is an expected situation, considering that one of the self-driving vehicles goals is to increase traffic safety. However, it is inferred that even though it is safer, it is susceptible to issues due to the randomness of the vehicular environment.

In a first stage, data resulting from the TTC_{mo} analysis is filtered in a first stage by TTC events with *low_risk* below 20 s. Meanwhile, we kept the records for *moderate_risk* and *high_risk*. Figure 5.6 shows the distribution of data filtered by TTC_{mo}. Therefore, we adopted resampling techniques. These consist of removing samples from the majority class (undersampling) and/or adding more samples from the minority class (oversampling) [114]. To reduce misclassification and biases by unbalancing, we employ data level solutions based on hybrid undersampling and oversampling techniques to balance the class distribution. First, we undersampling the class with the most samples. Since potentially important data is not expected to loss, the majority of data reduction is controlled. Next, after the undersampling process, oversampling of the minority classes is done to increase the sampling for training and validation. To apply these techniques, we use *imblearn*, a Python toolbox used to tackle the curse of imbalanced datasets in machine learning [115].

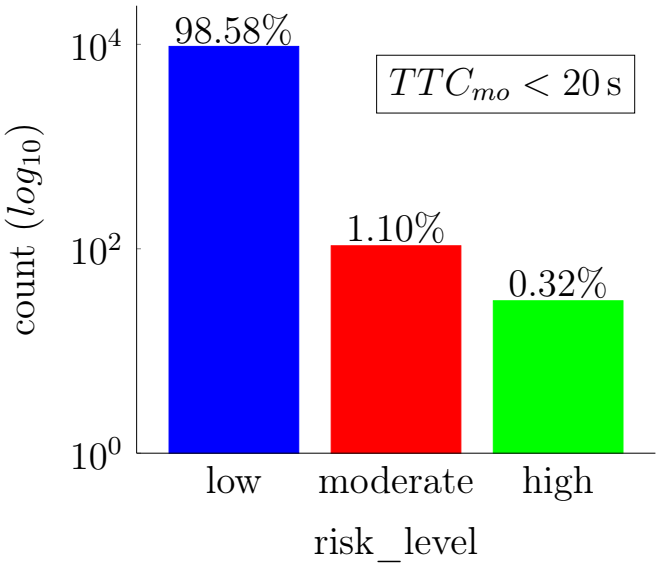
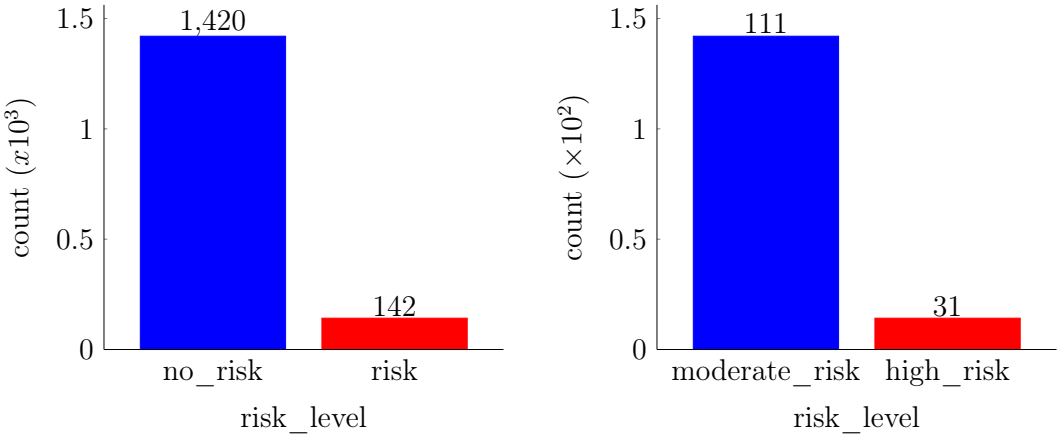


Figure 5.6: Multilabel data distribution.

We apply an undersampling process based on Heinrich's pyramid (see Figure 4.3), a theory of industrial accident prevention which postulates that there is a numerical relationship between low risk events, with minor or major injuries [116]. In vehicular environments, this theory can be used to define the relationship between the severity and frequency of conflict events [4]. A widely used representation is that for every 10 risk events with minor injuries, 1 event with major injuries occurs [117]. We also use this theory to reduce the sampling of low-risk events. Thus, *low_risk* events are defined as *"no_risk"* or 0, while *moderate_risk* and *high_risk* are defined as *"risk"* events or 1. Therefore, we assume that *risk* events correspond to events with imminent risk of crash, while *no_risk* events represent low-risk events that may become *risk* events. We undersample *no_risk* class to 10 times the sum of events detected as *risk* events, as shown in Figure 5.7(a). Figure 5.7(b) shows the data distribution for *moderate_risk* and *high_risk* events.



(a) Binary data distribution for *no_risk* and *risk* events.

(b) Binary data distribution for *moderate_risk* and *high_risk* events.

Figure 5.7: Data distribution for multilabel and binary risk classification.

Since the dimensionality of the labels in the selected data can affect the classification, with an Imbalanced Ratio, $IR > 341$, we use a *binarization* technique to decompose the imbalance problem [118] into binary classification sub-problems [119]. Next, we did an oversampling process using Synthetic Minority Oversampling TEchnique (SMOTE)[120]. The goal is to increase the minority class samples by introducing synthetic samples rather than by over-sampling with replacement.

5.4.3 Preprocessing

The preprocessing stage encompasses the following actions: (i) Data alignment between $\mathbf{X}_{a_1}$ and $\mathbf{Y}_{a_2}$; (ii) partitioning $\mathbf{X}_{a_1}|\mathbf{Y}_{a_2}$ into training, validation and test subsets; (iii) imbalanced data treatment; and (iv) data normalization.

- a_1 and a_2 are aligned based on the *scene_id* to ensure data alignment in the partition process.
- Partitioning data is defined as follows: **validation** subset uses 10% of the dataset's samples and **test** subset uses 20% of the dataset's samples. The remaining 70% is used as **training** data.
- Undersampling process is applied for the majority class in training, validation and test subsets to reduce it to 10 times the quantity of samples in the minority class.
- Oversampling process is applied for the minority class in training and validation subsets to equalize the majority class samples. Test subset is not oversampled. Figure 5.8 shows the data distribution of training, validation and test subsets.

Once the oversampling process is performed, the resulting subsets from $\mathbf{X}_{a_1}$ are normalized to obtain the minimum and maximum value of the features between 0 and 1. Normalized values are transformed to tensor format, a vector n-dimensional that represents all the data contained in the subsets from $\mathbf{X}_{a_1}$. Next, tensor subsets will be used in the next Section 5.4.4. Meanwhile, subsets resulting from partitioning $\mathbf{Y}_{a_2}$ were not modified.

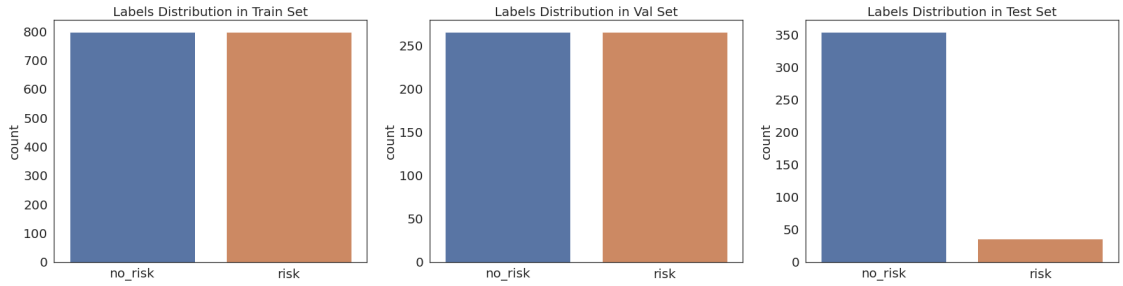


Figure 5.8: Data distribution for training, validation and test subsets.

5.4.4 Framework Setup

We use Syft framework, Python version 3.7.13, and PyTorch version 1.8.1. The experiments were performed on the Google Compute Engine platform (Intel® Xeon® CPU @ 2.30GHz, 12 GB RAM). Syft allows computing data by implementing encryption methods that enable privacy preservation. A Data Owner (DO) and Data Scientist (DS) setup model is used. In this approach, we define DO as the car manufacturer (a_1), and DS as the insurer (a_2). Figure 5.9 shows the details

of the model used to VFL implementation. The Syft-based federated environment happens as follows:

- **DO** launches the Duet server on its own machine and generates a session ID and a public key that needs to be shared with **DS** to join the session. **DO** waits until **DS** sends its public key to start the session.
- **DS** uses the session ID and public key from **DO** to join the Duet session. Next, a **DS** public key is generated and returned to the **DO** that prompt the public key from **DS** to validate the start of session.
- Once the **DS** is authenticated, the connection is established and the parties can start exchanging information.
- **DO** does all the preprocessing processes in $\mathbf{X}_{a_1}$.
- **DO** sends the data to Duet framework, who performs all encryption processes and pointer identification on this data to preserve privacy and data integrity. A Pointer is the main handler when interacting with remote data. Next, the data pointer from **DO** is sent to **DS**.
- **DS** receives the data pointer from **DO**.
- **DS** loads $\mathbf{Y}_{a_2}$ and data pointers from **DO** on its own machine.
- **DS** setups hyperparameters, DNN and manages parallel iteration between encrypted $\mathbf{X}_{a_1}$ and $\mathbf{Y}_{a_2}$.
- **DS** trains the model, computes gradients and loss, and updates the model until obtain a federated model. Intermediate results are shared with **DO**.

The model used for the experiments is a simple Deep Neural Network (DNN) model using Pytorch, with the architecture shown in Figure 5.10. The deep learning network consists of 2 different hidden layers and involves use of ReLU activation function in the first hidden layer in the network. The model also has Dropout set to 0.5. The input layer consists of 3 different perceptrons which corresponds to each of the input feature for the training dataset. The layers of the model are described below:

- **Input layer**: For both classifiers we use the following data from a_1: vehicle speed, in [km/h], $\mathrm{TTC}_{\mathrm{mo}}$, in [s], and quantity of objects detected. The input vector of the model is expressed as $[TTC_{mo}$*, vehicle_ speed, qty_ obj_ detected*$]$.

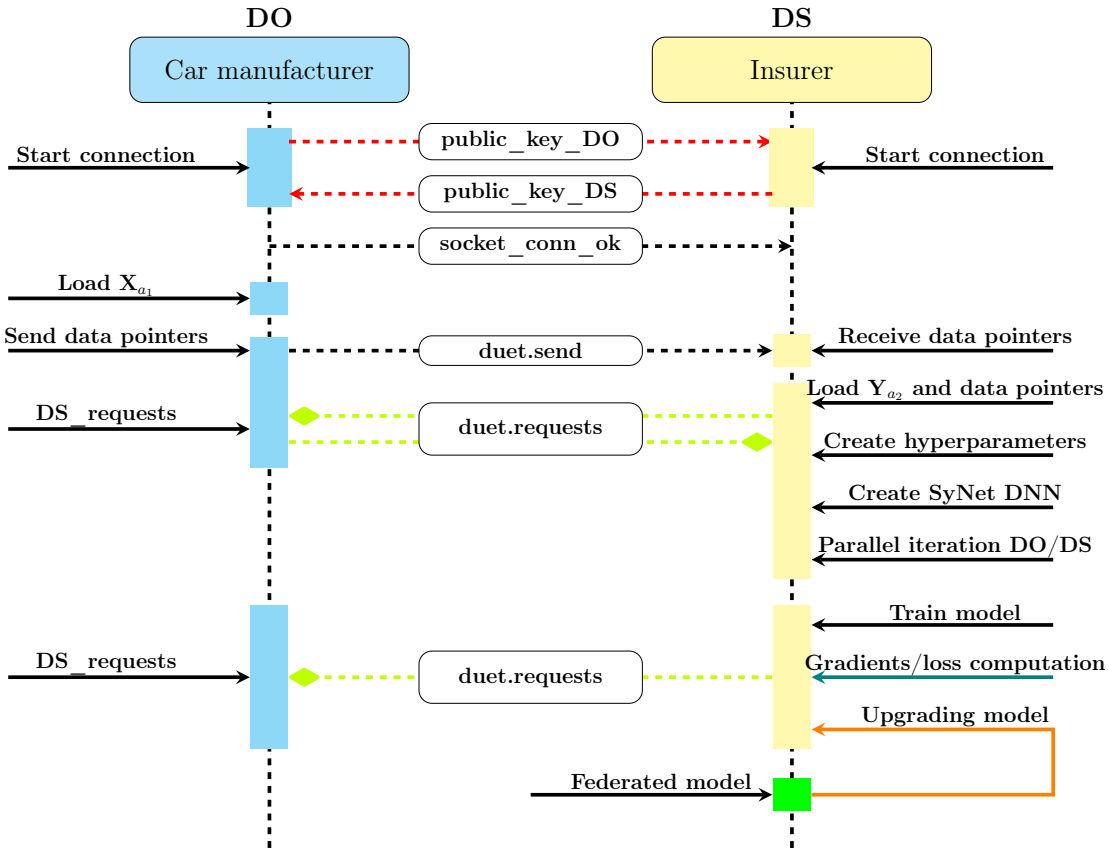


Figure 5.9: Illustration of the model used for VFL evaluation using the Syft framework.

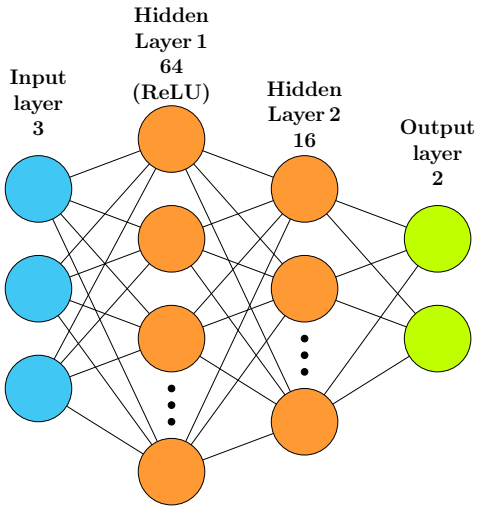


Figure 5.10: Neural network for risk prediction.

- **Hidden layers**: The first hidden layer of the model contains 64 neurons and is a full connection layer, with rectified linear unit (ReLU) as the activation function. The second hidden layer contains 16 neurons fully connected.
- **Output layer**: The output layer corresponds to the prediction of *risk* and *no_risk* events.

Once the processes described in the methodology were elaborated, the simulation experiments were performed. The results are described below.

5.5 Results

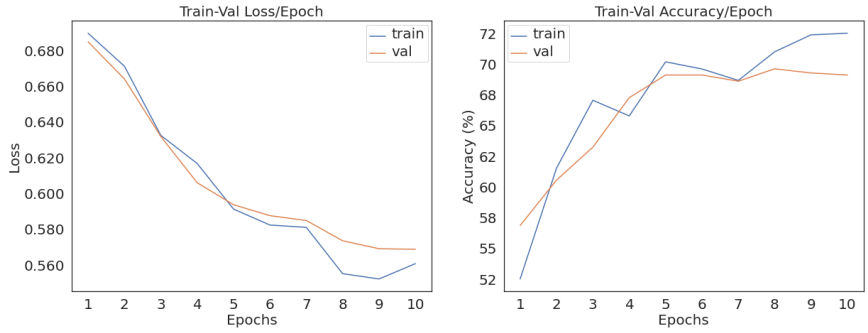
In this section, we provide the details of our experiments and compare a local machine learning against vertical federated learning. To evaluate the results, we consider the same subsets distribution a_1 and a_2. Table 5.1 shows the parameters defined for both models, as shown in presents these parameters. The learning rate was set different until we observed a quick drop in the loss function in the train/validation learning.

Table 5.1: Parameters defined for the learning stage.

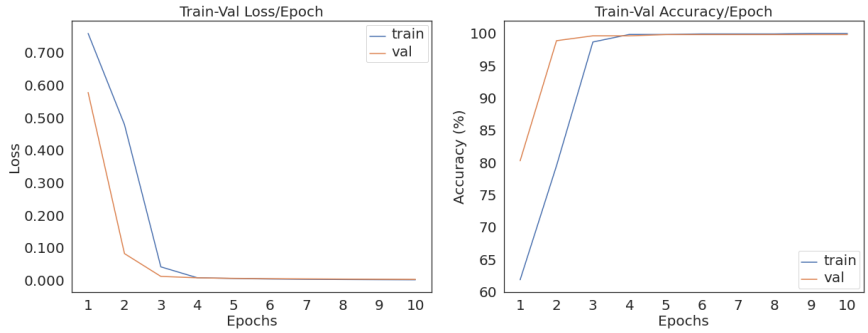
Parameter	Values	
	VFL	Local
Optimizer	Adam	Adam
Loss model	CrossEntropyLoss	CrossEntropyLoss
Learning rate	0.01	0.001
Training epochs	10	10
Batch size	32	32

Figure 5.11 shows that validation subset in both local and VFL models maintain the same progress of the loss function on the training subset. Indeed, it is possible noting high accuracy in each model, with approx. 70% for the local model (see Figure 5.11(a)), and 98% in the VFL model (see Figure 5.11(b)). In fact, the VFL model converges faster than the local model. Moreover, the findings show that VFL accuracy is 30% higher than local model accuracy. This may be associated with the data randomness at the time of preprocessing, as well as in the synthetic sampling inserted by oversampling process. On the other hand, the model training time was 5 s, while VFL training time is approx. 30 min. This difference may be associated with problems in synchronization with the Duet socket due to communication delays in the gradient aggregation process, intermediate exchange and model updates.

The classification results for the test subset in the local model and VFL can be analyzed through the confusion matrices in Figure 5.12. The values are within the range of 0 to 100%. The higher the color tone, the greater the number of classifications belonging to the label. In the local model, it is possible to observe that 100% of the "*risk*" events were detected, while 91.92% of "*no_risk*" events were detected, and 8.08% were classified as false negatives (FN). In the VFL model, 94.74% of the "*risk*" events were detected, and 5.26% were classified as false positives (FP); on the other hand, 100% of the "*no_risk*" events were detected.

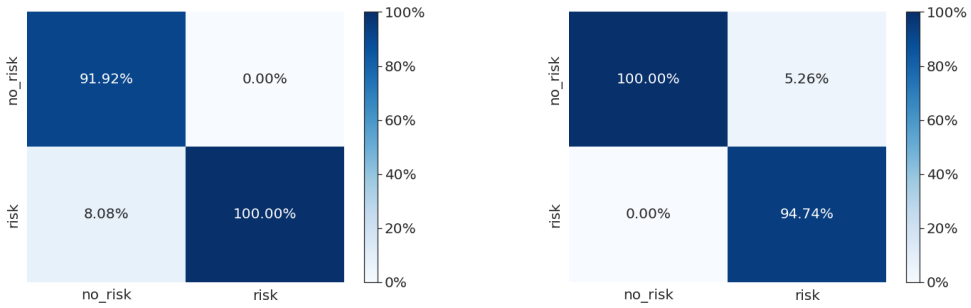


(a) Local model.



(b) VFL model.

Figure 5.11: Loss function vs. accuracy for train-validation analysis.



(a) Local model.

(b) VFL model.

Figure 5.12: Confusion matrix for test analysis.

Besides Accuracy, we also measure other metrics, such as Precision, Recall and F1-Score. The accuracy of a method is the ratio of the total of correctly sorted samples (True Positives (TP) + True Negatives (TN)) divided by the total number of samples. The precision is the ratio between the number of samples correctly classified for the positive class (TP), divided by the total of samples classified for this class (TP + FP). The Recall, also known as sensitivity or true positive rate, is the ratio of the number of correctly classified samples to the positive class (TP), divided by the total of samples belonging to this class (TP+ FN). F1-score is the harmonic mean of

the precision with the sensitivity. Table 5.2 and 5.3 show the classification reports of the local model and VFL. By observing the values of Accuracy and Precision, it is possible to see the good performance of the models in the classification of events. As shown in Figure 5.12, both models showed high accuracy in all classes, with a low rate of false positives. Moreover, both models achieve high precision values, almost 100%. Altogether, the overall accuracy for the local model is 0.927, while the VFL model is 0.995. Nonetheless, recall and F1-score in local model gets low values, seems to underfitting. This behavior can be influenced by the dataset skewed (imbalanced data), as well as by the undersampling/oversampling methods adopted in Section 5.4.2, since data nature do not allow observing a trend to detect risk events, e.g., risk events do not always occur at high speeds. In fact, since there is no clear trend, the randomness inserted in the oversampling process can result in data synthetic sampling that impacts on the learning process.

Table 5.2: Evaluation metrics for local model.

	Precision	Recall	F1-Score	Accuracy
no_risk	1	0.919	1	0.927
risk	1	0.125	0.222	

Table 5.3: Evaluation metrics for VFL model.

	Precision	Recall	F1-Score	Accuracy
no_risk	1	0.994	0.997	0.995
risk	0.947	1	0.973	

The experiments presented in this chapter illustrate the initial results of comparing ML models trained in local and federated manners. Models are evaluated in terms of predictive performance, training and inference duration. The effects of data amount, imbalanced data are studied, giving an indication that the model performance and system scalability depends on data distribution. Both local and VFL models converged appropriately, with similar trends in the loss function and accuracy. However, it was possible to observe that the VFL model converged faster both the loss function and the accuracy, showing a better performance than the local model. Minimal differences are observed in models performance when measured overall accuracy. This is also observed that there are minimal differences in no_risk events with precision, recall, and F1-score metrics; nonetheless, there are differences between local and VFL results for the risk events perception. It can be an issue for the perception of risk events since they are limited. Although the accuracy is maximum (1), the Recall may indicate that there may be False Negatives that can

interfere with the perception of risk events. This trend is not observed in the VFL model, which indicates that VFL can deliver accurate results from the analysis from other parties data.

These results coincide with those observed in Section 5.2, where it was observed that in common, all VFL models converged with a similar performance of the centralized models. On the other hand, experiments with VFL were more extensive with PySyft, as reported in Budrionis *et al.* [102], which indicates that the framework can extend the learning time up to 40 times longer; it was possible to observe this behavior.

The results obtained also have some limitations. Due to the randomness of vehicular environment, in addition to the data uniformity in the nuScenes AV data, it can prevent to correctly learn the risk through VFL. Furthermore, the use of synthetic samples can generate random trends that may not correctly represent the original trend of the data. In fact, oversampled data distribution can insert wrong data or repetitive data in VFL, resulting in False Negative detection. In this sense, a solution could be to filter only the time section where some type of risk was detected, in order to avoid imbalanced data.

5.6 Remarks

Risk assessment from data analysis in federated learning environments promises to be a functional solution that can help designing shared models which require privacy preservation and data integrity. In addition, it is a promising solution to integrate data silos and reduce data fragmentation due to isolated data. As noted in this chapter, collaboration between organizations can allow higher interaction between partners who share user groups in industrial sectors.

- Although there are some frameworks that work for specific tasks, there is still no framework or base toolbox for developing VFL-based solutions. This can be a limiting factor for other participants to be willing to share their data in federated environments, since there is no common framework for all potential participants.

- Since data cannot be inspected, there is a general lack of knowledge about how to manipulate the data. VFL processes can be difficult due to data distribution, specifically how it will be handled. That is, definition of hyperparameters, optimizers, type of neural network and number of neurons, among others, which can take a long time to understand how the data distribution of participants is arranged in the federated environment.

- As observed in the experiments carried out, VFL can be susceptible to communication delays in the gradient aggregation process, intermediate exchange and model updates. This can be reflected in delays, high computational cost, and slow data transfers.

Chapter 6

Conclusions

The development of this thesis demanded an interdisciplinary effort which explores the risk assessment in traffic risk events. We focus on self-driving vehicles since it is a challenging field for companies offering driver-based insurance services. We start by giving an overview of sensors in intelligent vehicles, exploring the most used sensors in OTS devices and exteroceptive sensors. It is important to note that sensors play an essential role in self-driving vehicles since they allow monitoring surroundings, detect oncoming obstacles and plan routes safely. Therefore, from the performance of these and other functionalities in the vehicle, it is possible to determine the reliability and liability of each sensor installed in the self-driving vehicle or in the OTS device. Furthermore, we note that depending on the complexity of the vehicle's functionalities, sensor fusion techniques are required to improve the accuracy and precision of vehicle telemetry in order to improve sensing performance. These analyses were reported in [24].

From sensor analysis, we carried out an analysis of road users and infrastructure of the environment around the self-driving vehicle. In a first stage, we analyzed public AV datasets with available exteroceptive sensor readings [1, 2]. These datasets contain raw data from the sensor readings, as well as semantic data from the detection of objects, categorization, attributes and detected object metrics stand out.

Data from the AVs are used to analyze, in offline mode, traffic risk events based on Surrogate Safety Measures (SSMs). We established a monitoring layer to calculate SSMs for all objects detected and categorized in the AV datasets. However, many of the detected objects, despite interacting with the AV, may not represent a risk for it. Thus, we formulated the implementation of TTC with motion orientation ($\mathrm{TTC_{mo}}$). We use data from the bounding boxes generated for each detected object. We calculate which of these objects are in the course of the AV as they interact. Furthermore, the velocity of the ego-vehicle and the object are calculated about the axis of displacement of the ego-vehicle and its yaw orientation. Meanwhile, the distance is calculated to the closest potential impact point between the ego-vehicle and the

object. In this way it is possible to explain interactions in car-following, head-on and intersection scenarios. This analysis is important for the explainability of traffic risk events, in addition to clarifying which of all the detected objects represent a real risk for the ego-vehicle. The results show that motion orientation analysis optimizes the regular TTC calculation by evaluating those objects interacting with the ego-vehicle on the path of motion. Likewise, in our analysis, it is also possible to determine the first point of potential impact of the ego-vehicle. Furthermore, risk assessment by analyzing other road users allows understanding what the decision pattern is like and can help to intuit what are the priorities of the autonomous system controller.

Another strategy to analyze risk assessment is data sharing. It is becoming a relevant topic to analyze various factors in traffic. This is an area that requires continuous development of machine learning and artificial intelligence techniques in order to obtain the maximum amount of information from the available data with a low computational cost. In this sense, we implemented the vertical federated learning technique, which is a method to create collaborative environments between different entities, allowing data sharing between trusted partners, preserving privacy and data integrity. However, due to the learning time, it was possible to observe that there is an overhead that impacts the model performance in terms of time and computational cost, which we associate with the encryption and privacy methods implemented in the framework used. Therefore, we propose a neural network as simple as possible to try to improve the processing overhead of the framework. We observed that a simple neural network becomes practical for the type of data to be analyzed, managing to converge to the detection of risk events from the alignment of shared features in different automakers. Our results show that VFL model classification was successful for *risk* event detection, showing a higher precision and accuracy in the classification.

6.1 Future Work

In the vehicular environment, it is possible to observe that it is a promising solution since the generation of data in self-driving vehicles is exponential, and it is expected the analysis of this data will contribute to the improvement of road safety, since it can be collaborated globally with other vehicles and computing centers. Nonetheless, some challenges remain open:

- **Liability:** Self-driving vehicles have a set of sensors that aim to emulate a sense of perception similar to the human. In this respect, self-driving vehicles trust the set of sensors and will execute its processes from the readings of its sensors. Therefore, sensor's liability requires continued monitoring. Moreover, this is also directly linked to the sensing of the vehicle's mechanics.

- **Perception data:** Misclassifying data can be crucial in self-driving vehicles. This can happen because it is assumed that AV operation may not be perfect due to limitations associated with the performance of learning algorithms executed by the autonomous driving system. In this sense, federated learning can enrich the information of the autonomous driving system from the evaluation of third-party decision-making to improve the vehicles, passengers and other road users safety.

- **Data standardization:** Still under development, self-driving vehicles and manufacturers do not have a standard defined for semantic data analysis. Data heterogeneity can make the autonomous driving system able to live with the randomness of the vehicular environment, however, it requires stable and robust approaches to ensure the safety of drivers in each and every situation. In this sense, there are several AV datasets [24], but it was observed that there is no standard in the data categorization, which can be an issue to assess the safety of both vehicles and passengers as well as other road users. Therefore, it is necessary to standardize the categorization of data from objects detected in AVs.

- **Driving profiling vs. Experience:** Self-driving vehicles are still a young industry, so it is possible to assume that they need to mature in terms of driving/learning time [121]. For example, Tesla has millions of miles driven, which allows it to be one of the few manufacturers to offer autonomy services up to level 4 [9]. Nevertheless, there is no collaborative environment in which other manufacturers can learn from Tesla's experience in order to converge towards a highly autonomous vehicle environment. In this sense, although there is no minimum experience/learning threshold, self-driving vehicles require continuous monitoring of their processes. In this regard, collaborative environments based on federated learning could be a solution, since most of these AVs have similar data acquisition and processing methodologies for their perception and action processes. In addition, it is important noting that the sensing is not unique to the vehicle, it also contains information on the surroundings and also on the infrastructure monitoring, which can enrich experience in new AVs.

- **Learning from human driving:** Coexistence of the self-driving vehicle with vehicles with a human driver is another challenging environment within the risk assessment in self-driving vehicles. These vehicles can be used to monitor human driver behaviors and share this data with partner organizations interested in controlling their vehicles, as well as insurance. Thus, well-informed driving behaviors can make the vehicle environment more intelligent to improve the road safety of human driver vehicles.

In the course of this thesis work, the following publications were produced.
As first author:

- **Ortiz, F.M.**, Ortiz, F. M., Sammarco, M., Costa, L. H. M. K., and Detyniecki, M. - "Applications and Services Using Vehicular Exteroceptive Sensors: a Survey", in IEEE Transactions on Intelligent Vehicles, June 2022.

- **Ortiz, F.M.**, Sammarco, M., Detyniecki, M., and Costa, L. H. M. K. - "Análise de Métricas Telemáticas para Auditoria de Veículos Autônomos", in V Workshop de Computação Urbana (CoUrb SBRC 2021), Uberlândia, MG, Brazil, August 2021.

- **Ortiz, F. M.**, Almeida, T. T., Ferreira, A., E. and Costa, L. H. M. K. "Experimental vs. Simulation Analysis of LoRa for Vehicular Communications", in Computer Communications, vol. 160, pp. 299-310, Elsevier, ISSN 0140-3664, DOI 10.1016/j.comcom.2020.06.006, 2020.

- **Ortiz, F. M.**, Almeida, T. T., Ferreira, A., E. and Costa, L. H. M. K. - "Caracterização de Desempenho de uma Rede LoRa em Ambientes Urbanos: Simulação vs. Prática", in III Workshop de Computação Urbana (CoUrb SBRC 2019), Gramado, RS, Brazil, May 2019. **Best paper award**.

As co-author:

- Pinto Neto, J. B., Gomes, L. C., **Ortiz, F. M.**, Almeida, T. T., Campista, M. E. M., Costa, L. H. M. K., Mitton, N. - "An Accurate Cooperative Positioning System for Vehicular Safety Applications", Computers and Electrical Engineering, Elsevier, 2020.

- Ferreira, A. E., **Ortiz, F. M.**, Almeida, T. T., Costa, L. H. M. K. - "A Visitor Assistance System Based on LoRa for Nature Forest Parks", in Electronics, vol. 9 (4), 696, MDPI, ISSN: 2079-9292, DOI 10.3390/electronics9040696, 2020.

- Ferreira, A., E. **Ortiz, F. M.**, Costa, L. H. M. K., Foubert, B., Amadou, I., and Mitton, N. - "A Study of the LoRa Signal Propagation in Forest, Urban, and Suburban Environments", in Annals of Telecommunications, vol. 75, no. 5, pp. 333-351, Springer, ISSN 1958-9395, DOI 10.1007/s12243-020-00789-w, 2020.

- Almeida, T. T., Gomes, L. C., **Ortiz, F. M.**, Junior, J. G. R. and Costa, L. H. M. K. - "Comparative Analysis of a Vehicular Safety Application in NS-3 and Veins", in IEEE Transactions on Intelligent Transportation Systems, ISSN 1558-0016, DOI 10.1109/TITS.2020.3014840, 2020.

- Ferreira, A. E., **Ortiz, F. M.** and Costa, L. H. M. K. - "Avaliação de Tecnologias de Comunicação Sem-Fio para Monitoramento em Ambientes de Floresta", in XXXVII Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos - SBRC 2019, Gramado, RS, Brazil, May 2019.

- Ferreira, A. E., **Ortiz, F. M.** and Costa, L. H. M. K. - "Caracterização de Disrupções na Faixa ISM de 900 MHz em Ambiente de Floresta", in XXXVII Simpósio Brasileiro de Telecomunicações e Processamento de Sinais - SBrT 2019, Petrópolis, RJ, Brazil, September 2019.

- Almeida, T. T., Gomes, L. C., **Ortiz, F. M.**, Junior, J. G. R. and Costa, L. H. M. K. - "IEEE 802.11p Performance Evaluation: Simulations vs. Real Experiments", in 21st IEEE International Conference on Intelligent Transportation Systems - ITSC 2018, Maui, Hawaii, USA, November 2018.

References

[1] CAESAR, H., BANKITI, V., LANG, A. H., et al. "nuScenes: A Multimodal Dataset for Autonomous Driving". In: *IEEE/CVF Conf. Comput. Vision Pattern Recognit. (CVPR)*, 2020.

[2] KESTEN, R., USMAN, M., HOUSTON, J., et al. "Lyft Level 5 AV Dataset 2019". 2019. Available in: <`https://lft.to/3cGvAwk`>.

[3] SOCIETY, I. E. P. "Automotive". In: *Heterogeneous Integration Roadmap*, IEEE, ch. 5, pp. 1–22, USA, 2019.

[4] HYDÉN, C. *The Development of a Method for Traffic Safety Evaluation: The Swedish Traffic Conflicts Technique*. Doctoral thesis, Lund University, Lund, Sweeden, Jul. 1987.

[5] CHIN, H., QUEK, S. T. "Measurement of traffic conflicts", *Saf. Sci.*, v. 26, pp. 169–185, 1997.

[6] LAURESHYN, A., ÅSE SVENSSON, HYDÉN, C. "Evaluation of traffic safety, based on micro-level behavioural data: Theoretical framework and first implementation", *Accid. Anal. Prev.*, v. 42, n. 6, pp. 1637–1646, 2010.

[7] YANG, Q., LIU, Y., CHEN, T., et al. "Federated Machine Learning: Concept and Applications", *ACM Trans. Intell. Syst. Technol.*, v. 10, n. 2, jan. 2019.

[8] MAHMUD, S. S., FERREIRA, L., HOQUE, M. S., et al. "Application of proximal surrogate indicators for safety evaluation: A review of recent developments and research needs", *IATSS Res.*, v. 41, n. 4, pp. 153–163, 2017.

[9] TESLA. "Full Self-Driving Capability Subscriptions". `https://bit.ly/3KChi2O`, 2020. Accessed June, 2022.

[10] BETZ, J., HEILMEIER, A., WISCHNEWSKI, A., et al. "Autonomous Driving—A Crash Explained in Detail", *Appl. Sci.*, v. 9, n. 23, 2019.

[11] MORDUE, G., YEUNG, A., WU, F. "The looming challenges of regulating high level autonomous vehicles", *Transp. Res. Part A Policy Pract.*, v. 132, pp. 174–187, 2020.

[12] WORLD HEALTH ORGANIZATION (WHO). "Global Status Report on Road Safety". `https://bit.ly/2VXS2KK`, 2018. Accessed January, 2020.

[13] EUROPEAN COMMISSION. "Road Safety in the European Union: Trends, statistics and main challenges". `https://bit.ly/3cKouad`, 2018. Accessed January, 2020.

[14] U.S. DEPARTMENT OF TRANSPORTATION. "Overview of Motor Vehicle Crashes in 2020". `https://bit.ly/3cuNqIn`, 2022. Accessed July, 2022.

[15] KOSCHER, K., CZESKIS, A., ROESNER, F., et al. "Experimental Security Analysis of a Modern Automobile". In: *IEEE Symp. Secur. Privacy*, pp. 447–462, May 2010.

[16] GUO, C., SENTOUH, C., POPIEUL, J.-C., et al. "Cooperation between driver and automated driving system: Implementation and evaluation", *Transp. Res. Part F Psychol. Behav.*, v. 61, pp. 314–325, 2019.

[17] SAE INTERNATIONAL. *SAE Standard J3016: Taxonomy and Definitions for Terms Related to On-Road Motor Vehicles*. Revision, Society of Automotive Engineers (SAE), 2018.

[18] BELLET, T., CUNNEEN, M., MULLINS, M., et al. "From semi to fully autonomous vehicles: New emerging risks and ethico-legal challenges for human-machine interactions", *Transp. Res. Part F: Psychol. Behav.*, v. 63, pp. 153–164, 2019.

[19] IHS MARKIT. "8 in 2018: The top transformative technologies to watch this year". 2018. Available in: <`https://bit.ly/3elVRkm`>.

[20] TIFFINY ROSSI. "Autonomous and ADAS test cars produce over 11 TB of data per day". 2018. Available in: <`http://bit.ly/2CWCuBy`>.

[21] IBM. "The Four V's of Big Data". `http://ibm.co/39KHENq`, 2018. Accessed January, 2020.

[22] ASSUNÇÃO, M. D., CALHEIROS, R. N., BIANCHI, S., et al. "Big Data computing and clouds: Trends and future directions", *J. Parallel Distrib. Comput.*, v. 79-80, pp. 3 – 15, 2015.

[23] BAECKE, P., BOCCA, L. "The value of vehicle telematics data in insurance risk selection processes", *Decis. Support Syst.*, v. 98, pp. 69–79, 2017.

[24] MOLANO ORTIZ, F., SAMMARCO, M., COSTA, L. H. M. K., et al. "Applications and Services Using Vehicular Exteroceptive Sensors: a Survey", *IEEE Trans. Intell. Veh.*, pp. 1–20, 2022.

[25] LI, H., MA, D., MEDJAHED, B., et al. "Secure and Privacy-Preserving Data Collection Mechanisms for Connected Vehicles". In: *SAE Techn. Paper*, Mar. 2017.

[26] BONNEFON, J.-F., CERNY, D., DANAHER, J., et al. *Ethics of Connected and Automated Vehicles: Recommendations on Road Safety, Privacy, Fairness, Explainability and Responsibility.* UE, European Commission, Sep. 2020.

[27] DERIKX, S., DE REUVER, M., KROESEN, M. "Can privacy concerns for insurance of connected cars be compensated?" *Electron. Markets*, v. 26, n. 1, pp. 73–81, 2016.

[28] KAIROUZ, P., MCMAHAN, H. B., AVENT, B., et al. "Advances and Open Problems in Federated Learning", *Found. Trends Mach. Learn.*, v. 14, n. 1-2, pp. 1–210, 2021.

[29] SIEGWART, R., NOURBAKHSH, I. R. *Introduction to Autonomous Mobile Robots.* USA, Bradford Company, 2004.

[30] SJAFRIE, H. *Introduction to Self-Driving Vehicle Technology.* USA, Taylor & Francis Group, 2020.

[31] WAHLSTRÖM, J., SKOG, I., HÄNDEL, P. "Smartphone-Based Vehicle Telematics: A Ten-Year Anniversary", *IEEE Trans. Intell. Transp. Syst.*, v. 18, n. 10, pp. 2802–2825, 2017.

[32] FLEMING, W. J. "New Automotive Sensors – A Review", *IEEE Sensors J.*, v. 8, n. 11, pp. 1900–1921, 2008.

[33] WONG, P. K., TAM, L. M., KE, L. "Automotive engine power performance tuning under numerical and nominal data", *Control Eng. Pract.*, v. 20, n. 3, pp. 300–314, 2012.

[34] KIM, J., PARK, S., LEE, U. "Dashcam Witness: Video Sharing Motives and Privacy Concerns Across Different Nations", *IEEE Access*, v. 8, pp. 110425–110437, 2020.

[35] TYMOSZEK, C., ARORA, S., WAGNER, K., et al. "DashCam Pay: A System for In-vehicle Payments Using Face and Voice", *arXiv:2004.03756*, 2020.

[36] SENEVIRATNE, S., HU, Y., NGUYEN, T., et al. "A Survey of Wearable Devices and Challenges", *IEEE Commun. Surv. Tutorials*, v. 19, n. 4, pp. 2573–2620, 2017.

[37] EUROPEAN GLOBAL NAVIGATION SATELLITE SYSTEMS AGENCY. "GNSS Market Report, Issue 6". 2019. Available in: <`https://bit.ly/3es9XRQ`>.

[38] SIRF TECHNOLOGY, INC. "NMEA Reference Manual". 2007. Available in: <`https://bit.ly/2YBt5Yy`>.

[39] JONES, W. D. "A compass in every smartphone", *IEEE Spectrum*, v. 47, n. 2, pp. 12–13, 2010.

[40] ISO 9613-2. "Acoustics — Attenuation of sound during propagation outdoors — Part 2: General method of calculation". 1996. Available in: <`https://bit.ly/3hjkbWa`>.

[41] DOUDOU, M., BOUABDALLAH, A., BERGE-CHERFAOUI, V. "Driver Drowsiness Measurement Technologies: Current Research, Market Solutions, and Challenges", *Int. J. Intell. Transp. Syst. Res.*, v. 18, pp. 297–319, 2020.

[42] PATOLE, S. M., TORLAK, M., WANG, D., et al. "Automotive radars: A review of signal processing techniques", *IEEE Signal Process. Mag.*, v. 34, n. 2, pp. 22–35, 2017.

[43] UHNDER. "Uhnder Launches Industry's First 4D Digital Imaging Radar". 2022. Available in: <`https://bit.ly/3MWRvSi`>.

[44] LI, Y., IBANEZ-GUZMAN, J. "Lidar for Autonomous Driving: The Principles, Challenges, and Trends for Automotive Lidar and Perception Systems", *IEEE Signal Process. Mag.*, v. 37, n. 4, pp. 50–61, 2020.

[45] RORIZ, R., CABRAL, J., GOMES, T. "Automotive LiDAR Technology: A Survey", *IEEE Trans. Intell. Transp. Syst.*, pp. 1–16, 2021.

[46] GRIMM, M. "Camera-based driver assistance systems", *Adv. Opt. Technol.*, v. 2, n. 2, pp. 131–140, 2013.

[47] BOUTTEAU, R., SAVATIER, X., BONARDI, F., et al. "Road-line detection and 3D reconstruction using fisheye cameras". In: *IEEE Int. Conf. Intell. Transp. Syst. (ITSC)*, pp. 1083–1088, 2013.

[48] WEN, W., BAI, X., KAN, Y. C., et al. "Tightly Coupled GNSS/INS Integration via Factor Graph and Aided by Fish-Eye Camera", *IEEE Trans. Veh. Technol.*, v. 68, n. 11, pp. 10651–10662, 2019.

[49] WANG, Z., WU, Y., NIU, Q. "Multi-Sensor Fusion in Automated Driving: A Survey", *IEEE Access*, v. 8, pp. 2847–2868, 2020.

[50] VARGHESE, J. Z., BOONE, R. G., OTHERS. "Overview of autonomous vehicle sensors and systems". In: *Int. Conf. Oper. Exc. Serv. Eng.*, pp. 178–191, set. 2015.

[51] KATHY WINTER. "For Self-Driving Cars, There's Big Meaning Behind One Big Number: 4 Terabytes". 2018. Available in: <`http://intel.ly/2KvU6YG`>.

[52] HOUSTON, J., ZUIDHOF, G., BERGAMINI, L., et al. "One Thousand and One Hours: Self-driving Motion Prediction Dataset", *arXiv:2006.14480*, 2020.

[53] NUTONOMY. "nuscenes-devkit". `https://github.com/nutonomy/nuscenes-devkit`, 2018. Accessed October, 2020.

[54] ELIOT, LANCE. "LIDAR Put Into A Tesla – And The World Doesn't Come To An End". 2019. Available in: <`https://bit.ly/3g85jte`>.

[55] PARKER JR, M., ZEGEER, C. V. *Traffic conflict techniques for safety and operations: Observers manual.* Technical report, U.S. Federal Highway Administration, 1989.

[56] ZHENG, L., ISMAIL, K., MENG, X. "Traffic conflict techniques for road safety analysis: Open questions and some insights", *Can. J. Civil Eng.*, v. 41, 07 2014.

[57] VITALE, A., GUIDO, G., SACCOMANNO, F., et al. "Comparing Safety Performance Measures Obtained from Video Capture Data", *J. Transp. Eng.*, v. 137, pp. 481–491, 07 2011.

[58] ELVIK, R., HØYE, A., VAA, T., et al. *The Handbook of Road Safety Measures.* UK, Emerald Group, 2009.

[59] SCANLON, J. M., KUSANO, K. D., DANIEL, T., et al. "Waymo simulated driving behavior in reconstructed fatal crashes within an autonomous vehicle operating domain", *Accid. Anal. Prev.*, v. 163, pp. 106454, 2021.

[60] JOHNSSON, C., LAURESHYN, A., CEUNYNCK, T. D. "In search of surrogate safety indicators for vulnerable road users: a review of surrogate safety indicators", *Transp. Rev.*, v. 38, n. 6, pp. 765–785, 2018.

[61] TARKO, A., DAVIS, G., SAUNIER, N., et al. "Surrogate Measures of Safety", *Safe Mobility: Challenges, Methodology and Solutions-(Transport and Sustainability)*, v. 11, pp. 383–405, Jan. 2009.

[62] HAUER, E., GARDER, P. "Research into the validity of the traffic conflicts technique", *Accid. Anal. Prev.*, v. 18, n. 6, pp. 471–481, 1986.

[63] GETTMAN, D., PU, L., SAYED, T., et al. *Surrogate safety assessment model and validation*. Technical report, United States. Federal Highway Administration. Office of Safety Research and ..., 2008.

[64] ALMQVIST, S., HYDÉN, C., RISSER, R. "Use of speed limiters in cars for increased safety and a better environment", *Transp. Res. Rec.*, 1991.

[65] HYDÉN, C. "Traffic conflicts technique: state-of-the-art", *Traffic Saf. Work Video Process.*, v. 37, pp. 3–14, 1996.

[66] ALLEN, B. L., SHIN, B., COOPER, P. J. "Analysis of Traffic Conflicts and Collisions", *Transp. Res. Rec.*, 1978.

[67] HAYWARD, J. C. "Near-Miss Determination Through Use of a Scale of Danger", *Highway Res. Rec.*, 1972.

[68] OZBAY, K., YANG, H., BARTIN, B., et al. "Derivation and Validation of New Simulation-Based Surrogate Safety Measure", *Transp. Res. Rec.*, v. 2083, n. 1, pp. 105–113, 2008.

[69] KIEFER, R. J., LEBLANC, D. J., FLANNAGAN, C. A. "Developing an inverse time-to-collision crash alert timing approach based on drivers' last-second braking and steering judgments", *Accid. Anal. Prev.*, v. 37, n. 2, pp. 295–303, 2005.

[70] XIE, K., YANG, D., OZBAY, K., et al. "Use of real-world connected vehicle data in identifying high-risk locations based on a new surrogate safety measure", *Accid. Anal. Prev.*, v. 125, pp. 311–319, 2019.

[71] MINDERHOUD, M. M., BOVY, P. H. "Extended time-to-collision measures for road traffic safety assessment", *Accid. Anal. Prev.*, v. 33, n. 1, pp. 89–97, 2001.

[72] MILLER, R., HUANG, Q. "An adaptive peer-to-peer collision warning system". In: *IEEE Veh. Technol. Conf. (VTC Spring)*, v. 1, pp. 317–321, 2002.

[73] JIMÉNEZ, F., NARANJO, J. E., GARCÍA, F. "An improved method to calculate the time-to-collision of two vehicles", *Int. J. Intell. Transp. Syst. Res.*, v. 11, n. 1, pp. 34–42, 2013.

[74] WARD, J. R., AGAMENNONI, G., WORRALL, S., et al. "Extending Time to Collision for probabilistic reasoning in general traffic scenarios", *Transp. Res. Part C: Emerging Technol.*, v. 51, pp. 66–82, 2015.

[75] WACHENFELD, W., JUNIETZ, P., WENZEL, R., et al. "The worst-time-to-collision metric for situation identification". In: *IEEE Intell. Veh. Symp.*, pp. 729–734, 2016.

[76] DINGUS, T. A., KLAUER, S. G., NEALE, V. L., et al. *The 100-car naturalistic driving study, Phase II-results of the 100-car field experiment.* Technical report, U.S. NHTSA, 2006.

[77] CAMPBELL, K., (U.S.), S. S. H. R. P. *The SHRP 2 Naturalistic Driving Study: Addressing Driver Performance and Behavior in Traffic Safety.* USA, Transp. Res. Board, 2012.

[78] MONTGOMERY, J., KUSANO, K. D., GABLER, H. C. "Age and Gender Differences in Time to Collision at Braking From the 100-Car Naturalistic Driving Study", *Traffic Inj. Prev.*, v. 15, n. sup1, pp. S15–S20, 2014.

[79] MARKKULA, G., ENGSTRÖM, J., LODIN, J., et al. "A farewell to brake reaction times? Kinematics-dependent brake response in naturalistic rear-end emergencies", *Accid. Anal. Prev.*, v. 95, pp. 209–226, 2016.

[80] NODINE, E., STEVENS, S., LAM, A., et al. *Independent evaluation of light-vehicle safety applications based on vehicle-to-vehicle communications used in the 2012-2013 safety pilot model deployment.* Technical report, U.S. NHTSA, 2015.

[81] HE, Z., QIN, X., LIU, P., et al. "Assessing Surrogate Safety Measures using a Safety Pilot Model Deployment Dataset", *Transp. Res. Rec.*, v. 2672, n. 38, pp. 1–11, 2018.

[82] KUSANO, K. D., MONTGOMERY, J., GABLER, H. C. "Methodology for identifying car following events from naturalistic data". In: *IEEE Intell. Veh. Symp.*, pp. 281–285, 2014.

[83] AYCARD, O., BAIG, Q., BOTA, S., et al. "Intersection safety using lidar and stereo vision sensors". In: *IEEE Intell. Veh. Symp.*, pp. 863–869, 2011.

[84] KILICARSLAN, M., ZHENG, J. Y. "Predict Vehicle Collision by TTC From Motion Using a Single Video Camera", *IEEE Trans. Intell. Transp. Syst.*, v. 20, n. 2, pp. 522–533, 2019.

[85] GEIGER, A., LENZ, P., URTASUN, R. "Are we ready for Autonomous Driving? The KITTI Vision Benchmark Suite". In: *IEEE Conf. Comput. Vision Pattern Recognit. (CVPR)*, pp. 3354–3361, 2012.

[86] LYFT LEVEL 5. "l5kit-devkit". `https://github.com/lyft/l5kit`, 2018. Accessed October, 2020.

[87] IVIS, F. "Calculating geographic distance: Concepts and Methods". `https://bit.ly/3wWn7AJ`, 2006. Accessed July, 2021.

[88] LI, Y., LU, J., XU, K. "Crash risk prediction model of lane-change behavior on approaching intersections", *Discrete Dynamics in Nature and Society*, v. 2017, Aug. 2017.

[89] RYDZEWSKI, A., CZARNUL, P. "Human awareness versus Autonomous Vehicles view: comparison of reaction times during emergencies". In: *IEEE Intell. Veh. Symp. (IV)*, pp. 732–739, 2021.

[90] MURSHED, M. G. S., MURPHY, C., HOU, D., et al. "Machine Learning at the Network Edge: A Survey", *ACM Comput. Surv.*, v. 54, n. 8, out. 2021.

[91] EUROPEAN PARLIAMENT AND COUNCIL OF THE EUROPEAN UNION. "The General Data Protection Regulation (EU) 2016/679 (GDPR)". 2016. Available in: <`https://eur-lex.europa.eu/eli/reg/2016/679/oj`>.

[92] MCMAHAN, H. B., MOORE, E., RAMAGE, D., et al. "Federated Learning of Deep Networks using Model Averaging", *arXiv preprint arXiv:1602.05629*, 2016.

[93] KONEČNÝ, J., MCMAHAN, H. B., YU, F. X., et al. "Federated learning: Strategies for improving communication efficiency", *arXiv preprint arXiv:1610.05492*, 2016.

[94] KONEČNỲ, J., MCMAHAN, H. B., RAMAGE, D., et al. "Federated optimization: Distributed machine learning for on-device intelligence", *arXiv preprint arXiv:1610.02527*, 2016.

[95] WAHAB, O. A., MOURAD, A., OTROK, H., et al. "Federated Machine Learning: Survey, Multi-Level Classification, Desirable Criteria and Future Directions in Communication and Networking Systems", *IEEE Commun. Surv. Tutorials*, v. 23, n. 2, pp. 1342–1397, 2021.

[96] CHENG, Y., LIU, Y., CHEN, T., et al. "Federated Learning for Privacy-Preserving AI", *Commun. ACM*, v. 63, n. 12, pp. 33–36, nov 2020.

[97] LIU, Y., FAN, T., CHEN, T., et al. "FATE: An Industrial Grade Platform for Collaborative Learning With Data Protection", *J. Mach. Learn. Res.*, v. 22, n. 226, pp. 1–6, 2021.

[98] CHENG, K., FAN, T., JIN, Y., et al. "SecureBoost: A Lossless Federated Learning Framework", *IEEE Intell. Syst.*, v. 36, n. 6, pp. 87–98, 2021.

[99] DUVERLE, D. A., KAWASAKI, S., YAMADA, Y., et al. "Privacy-Preserving Statistical Analysis by Exact Logistic Regression". In: *IEEE Secur. Privacy Workshops*, pp. 7–16, 2015.

[100] CHEN, T., JIN, X., SUN, Y., et al. "VAFL: a method of vertical asynchronous federated learning", *arXiv preprint arXiv:2007.06081*, 2020.

[101] JOHNSON, A. E., POLLARD, T. J., SHEN, L., et al. "MIMIC-III, a freely accessible critical care database", *Sci. Data*, v. 3, n. 1, pp. 1–9, 2016.

[102] BUDRIONIS, A., MIARA, M., MIARA, P., et al. "Benchmarking PySyft Federated Learning Framework on MIMIC-III Dataset", *IEEE Access*, v. 9, pp. 116869–116878, 2021.

[103] ZILLER, A., TRASK, A., LOPARDO, A., et al. "PySyft: A Library for Easy Federated Learning". In: *Federated Learning Systems: Towards Next-Generation AI*, pp. 111–139, Cham, Springer International Publishing, 2021.

[104] PENG, Y., CHEN, Z., CHEN, Z., et al. "BFLP: An adaptive federated learning framework for internet of vehicles", *Mobile Inf. Syst.*, v. 2021, 2021.

[105] YUAN, X., CHEN, J., YANG, J., et al. "FedSTN: Graph Representation Driven Federated Learning for Edge Computing Enabled Urban Traffic Flow Prediction", *IEEE Trans. Intell Transport. Syst.*, pp. 1–11, 2022.

[106] WANG, X., ZHENG, X., LIANG, X. "Charging Station Recommendation for Electric Vehicle Based on Federated Learning", *J. Phys. Conf. Ser.*, v. 1792, n. 1, pp. 012055, fev. 2021.

[107] BONAWITZ, K., IVANOV, V., KREUTER, B., et al. "Practical Secure Aggregation for Privacy-Preserving Machine Learning". In: *ACM SIGSAC Conf. Comput. Commun. Secur.*, p. 1175–1191, 2017.

[108] HALL, A. J., JAY, M., CEBERE, T., et al. "Syft 0.5: A platform for universally deployable structured transparency", *arXiv preprint arXiv:2104.12385*, 2021.

[109] RYFFEL, T., TRASK, A., DAHL, M., et al. "A generic framework for privacy preserving deep learning", *arXiv preprint arXiv:1811.04017*, 2018.

[110] PASZKE, A., GROSS, S., MASSA, F., et al. "PyTorch: An Imperative Style, High-Performance Deep Learning Library". In: *33rd Int. Conf. Neural Inf. Proc. Syst.*, p. 8026–8037, 2019.

[111] ZILLER, A., TRASK, A., LOPARDO, A., et al. "PySyft: A Library for Easy Federated Learning". In: *Federated Learning Systems: Towards Next-Generation AI*, pp. 111–139, Cham, Springer International Publishing, 2021.

[112] BENAISSA, A., RETIAT, B., CEBERE, B., et al. "TenSEAL: A library for encrypted tensor operations using homomorphic encryption", *arXiv preprint arXiv:2104.03152*, 2021.

[113] KRIPPENDORFF, K. "Reliability in content analysis: Some common misconceptions and recommendations", *Hum. Commun. Res.*, v. 30, n. 3, pp. 411–433, 2004.

[114] HE, H., YUNQIAN, M. *Imbalanced Learning: Foundations, Algorithms, and Applications*. USA, Wiley-IEEE Press, 2013.

[115] LEMAÎTRE, G., NOGUEIRA, F., ARIDAS, C. K. "Imbalanced-learn: A Python Toolbox to Tackle the Curse of Imbalanced Datasets in Machine Learning", *J. Mach. Learn. Res.*, v. 18, n. 17, pp. 1–5, 2017.

[116] HEINRICH, H. *Industrial Accident Prevention, A Scientific Approach*. USA, McGraw-Hill, 1931.

[117] BIRD, F., GERMAIN, G. *Practical Loss Control Leadership*. USA, International Loss Control Institute, 1985.

[118] HE, H., GARCIA, E. A. "Learning from Imbalanced Data", *IEEE Trans. Knowl. Data Eng.*, v. 21, n. 9, pp. 1263–1284, 2009.

[119] FERNÁNDEZ, A., LÓPEZ, V., GALAR, M., et al. "Analysing the classification of imbalanced data-sets with multiple classes: Binarization techniques and ad-hoc approaches", *Knowl.-Based Syst.*, v. 42, pp. 97–110, 2013.

[120] CHAWLA, N. V., BOWYER, K. W., HALL, L. O., et al. "SMOTE: synthetic minority over-sampling technique", *J. Artif. Intell. Res.*, v. 16, pp. 321–357, 2002.

[121] KALRA, N., PADDOCK, S. M. *Driving to Safety: How Many Miles of Driving Would It Take to Demonstrate Autonomous Vehicle Reliability?* Santa Monica, CA, RAND Corporation, 2016.